EuroComb'25
Booklet of extended abstracts

HUN-REN Alfréd Rényi Institute of Mathematics

Contents

Invited Talks **1**

Matija Bucic: Pass to expander paradigm 1

Vida Dujmović: Clustered Hadwiger Conjecture 1

Peter Frankl: Half a Century With and Without Extremal Set Theory 1

Hong Liu: Chromatic, homomorphism, blowup thresholds and beyond 2

László Lovász: Matroids, submodular functions, and bubbles 2

Wojciech Samotij: Stability of large cuts in random graphs 3

Julia Wolf: Efficient regularity in hypergraphs 3

Contributed Talks **5**

Eyal Ackerman, Balázs Keszegh: The Zarankiewicz Problem for Polygon Visibility Graphs 6

Kristina Ago: Variations on an intersection problem 14

Ron Aharoni, Eli Berger, He Guo, Dani Kotlar: Coloring, list coloring, and fractional coloring in intersections of matroids 19

Sandra Albrechtsen, Matthias Hamann: Asymptotic half-grid and full-grid minors 25

Jack Allsop, Dani Kotlar, Ian M. Wanless: On the Wide Partition Conjecture for free matroids 32

Nóra Almási, Gábor Simonyi: Alternating odd cycles in orientations of topologically χ-chromatic graphs 37

Yuyu An, Binlong Li, Fangfang Wu, Shenggui Zhang: Rainbow Triangles in Edge-colored Tripartite Graphs 42

Michael Anastos, Zhihan Jin, Matthew Kwan, Benny Sudakov: 46

Andrés Aranda, David Bradley-Williams, David Hartman: Homomorphism-homogeneous Graphs from Finite Reflexive Graphs 57

Patrick Arras, Frederik Garbe, Felix Joos: Asymptotically Enumerating Independent Sets in Regular k-Partite k-Uniform Hypergraphs 63

Maria Axenovich, Leon Kießle, Arsenii Sagdeev, Maksim Zhukovskii: Faces of maximal plane graphs without short cycles have bounded degrees 69

Matías Azócar Carvajal, Giovanne Santos, Mathias Schacht: Canonical Ramsey numbers for partite hypergraphs 74

Martin Balko, Klára Grinerová: Estimating multicolor ordered Ramsey numbers 81

Igor Balla, Marek Filakovský, Bartłomiej Kielak, Niklas Schlomberg: Curves on the torus with prescribed intersections 90

József Balogh, Andrea Freschi, Andrew Treglown: Ramsey problems for tilings in graphs 96

János Barát, Zoltán L. Blázsik, Balázs Keszegh, Zeyu Zheng: On the number of edges in saturated partial embeddings of maximal planar graphs 102
Julian Becker, Konstantinos Panagiotou, Matija Pasch: Improved Universal Graphs for Trees 105
Anton Bernshteyn, Jing Yu: Borel Local Lemma: Arbitrary Random Variables and Limited Exponential Growth 113
Gaétan Berthe, Caroline Brosse, Brian Hearn, Jan van den Heuvel, Pierre Hoppenot, Théo Pierron: Determining a graph from its reconfiguration graph . 122
Kerrin Bielser: On a Property of the 0-1 Subword Order 127
Olivier Bodini, Francis Durand: Efficient Random-Bit Algorithms for Uniform Involutions 133
Marthe Bonamy, Mónika Csikós, Anna Gujgiczer, Yelena Yuditsky: On graph classes with constant domination-packing ratio 142
Nicolas Bonichon, Thomas Muller, Adrian Tanasa: Higher dimensional floorplans and Baxter d-permutations 149
Fábio Botler, Cristina G. Fernandes, Carla N. Lintzmayer, Rui A. Lopes, Suchismita Mishra, Bruno L. Netto, Maycon Sambinelli: Immersions of large cliques in graphs with independence number 2 and bounded maximum degree 156
F. Botler, Y. S. Couto, C. G. Fernandes, E. F. de Figueiredo, R. Gómez, V. F. dos Santos, C. M. Sato: Extremal Problems on Forest Cuts and Acyclic Neighborhoods in Sparse Graphs 163
F. Botler, A. Kaique, W. Mendonça, G. O. Mota, C. M. Sato: Rainbow Path Covers of Sparse Random Graphs 168
Candida Bowtell, Richard Montgomery: Decomposing random Latin squares into transversals 174
Simona Boyadzhiyska, Micha Christoph, Tibor Szabó: Almost-perfect colorful matchings in bipartite graphs 180
David Bradley-Williams, Sofia Brenner: EPPA witnesses with two blocks of imprimitivity on edges 186
Christoph Brause, Rafał Kalinowski, Monika Pilśniak, Ingo Schiermeyer: On the distinguishing chromatic number in hereditary graph classes 193
Hector Buffière, Rutger Campbell, Kevin Hendrey, Sang-il Oum: Blind copwidth and balanced minors of graphs 198
Fabian Burghart, Marc Kaufmann, Noela Müller, Matija Pasch: The Hitting Time Of Nice Factors 204
Pjotr Buys, Sven Polak, Jeroen Zuiddam: A limit theorem for Shannon capacity from lattice packings 213
Pjotr Buys, Anna Luchnikov, Jim Wittebol, Jeroen Zuiddam: Universality of asymptotic graph cohomomorphism 222
Bartłomiej Bychawski: Classification of unstable circulant graphs of twice odd order 230
Pablo Candela, Diego González-Sánchez, Balázs Szegedy: Higher-order Fourier analysis via spectral algorithms 237
Daniela Černá, Bartłomiej Kielak, Jan Volec: On Shattering Triples with Six Permutations 245

Debsoumya Chakraborti, Micha Christoph, Zach Hunter, Richard Montgomery, Teo Petrov: Almost Full Transversals in Equi-n-squares 250
Parinya Chalermsook, Ly Orgo, Minoo Zarsav: A Freeable Matrix Characterization of Bipartite Graphs of Ferrers Dimension Three 257
Yulin Chang, Yangyang Cheng, Tianjiao Dai, Qiancheng Ouyang, Guanghui Wang: An Exact Ore-Type condition for Hamilton cycles in oriented graphs 263
Yaobin Chen, Zhicheng Liu, Yihang Xiao, Junchi Zhang: Extreme Points of Base Polytope of Submodular Set Function and Limit of Quotient Convergent Graph Sequence 269
Nina Chiarelli, Vesna Iršič Chenoweth, Marko Jakovac, William B. Kinnersley, Mirjana Mikalački: The burning game on graphs 278
David Chodounský, Monroe Eskew, Thilo Weinert: Colors of the Pseudotree . . 285
Micha Christoph, Anders Martinsson, Aleksa Milojević: Universality for transversal Hamilton cycles in random graphs 291
Maria Chudnovsky, Linda Cook, James Davies, Sang-il Oum, Jane Tan: Colouring t-perfect graphs 297
Quentin Chuet, Tianjiao Dai, Qiancheng Ouyang, François Pirot: New bounds for proper h-conflict-free colourings 304
Quentin Chuet, Hugo Demaret, Hoang La, François Pirot: Fractional domatic number and minimum degree 311
Sylwia Cichacz, Claire Hilaire, Tomáš Masařík, Jana Masaříková, Martin Milanič: Path Eccentricity and Forbidden Induced Subgraphs 318
Simone Costa, Stefano Della Fiore, Eva Engel: Graham's rearrangement for dihedral groups 326
Péter Csikvári, Ivan Damnjanović, Dragan Stevanović, Stephan Wagner: Lazy walk-based bounds on the spectral radius of threshold graphs 332
Bence Csonka: Fractional Haemers' bound and the Mycielski construction . . . 332
Tianjiao Dai, Qiancheng Ouyang, François Pirot, Jean-Sébastien Sereni: Beyond the fractional Reed bound for triangle-free graphs 336
Ivan Damnjanović: Maximum index of connected graphs 344
David de Boer, Pjotr Buys, Jeroen Zuiddam: The asymptotic spectrum distance, graph limits, and the Shannon capacity 350
Francesco Di Braccio, Freddie Illingworth: The Zarankiewicz problem on tripartite graphs 360
Travis Dillon: Sparse volumetric weak ϵ-nets 366
Gabriel D. Do Couto: On Meyniel's conjecture for Random Hypergraphs 377
Baran Düzgün, Ago-Erik Riet, Vladislav Taranchuk: New constructions of unbalanced $\{C_4, \theta_{3,t}\}$-free bipartite graphs 383
Mengyu Duan, Binlong Li, Minghui Yu, Shenggui Zhang: Rainbow Directed C_4's in Arc-colored Digraphs 389
Adrian Dumitrescu, Jeck Lim, János Pach, Ji Zeng: A Purely Geometric Variant of the Gale–Berlekamp Switching Game 394
Zdeněk Dvořák, Ross J. Kang, David Mikšaník: The Borodin–Kostochka conjecture and correspondence coloring 401
Zdeněk Dvořák, Bernard Lidický, Bojan Mohar: On a conjecture concerning 4-coloring of graphs with one crossing 411

Idan Eisner, Shlomo Hoory: Entropy and the growth rate of universal covering trees . . . 417
Joshua Erde, Max Gutkin, Mihyun Kang, Dominik Schmid: Froböse percolation in the hypercube . . . 423
Monroe Eskew, Yair Hayut: The Erdős-Hajnal graphs and reflection of infinite chromatic numbers . . . 429
Alberto Espuny Díaz, Lyuben Lichev, Alexandra Wesolek: Local resilience of random geometric graphs with respect to connectivity . . . 432
Alberto Espuny Díaz, Frederik Garbe, Tássio Naia, Zak Smith: Factors and powers of cycles in the budget-constrained random graph process . . . 438
Filip Filipi: Simple Graph Covers of Dipoles with Loops . . . 445
Nóra Frankl, Attila Jung: Helly-type theorems for monotone properties of boxes 450
V. G. Gorbounov, A.A. Kazakov: Metric properties of electrical networks and the graph reconstruction problems . . . 456
Cristina G. Fernandes, Carlos Hoppen, George Kontogeorgiou, Guilherme Oliveira Mota, Danni Peng: Separating Path Systems for 2-Degenerate Graphs . . . 461
Florian Galliot, Hoang La, Raphaëlle Maistre, Matthieu Petiteau, Dimitri Watel: Graph reconstruction from queries on triples . . . 467
Florian Galliot, Jonas Sénizergues: A unified convention for achievement positional games . . . 473
Guillermo Gamboa Quintero, Ida Kantor: Minimum blocking sets for families of partitions . . . 480
Dániel Gerbner, Hilal Hama Karim, Gaurav Kucheriya: On Turán-type problems and the abstract chromatic number . . . 486
Babak Ghanbari, Robert Šámal: Facial diagrams and cycle double cover . . . 493
Anubhab Ghosal, Peter Keevash: On the largest k-product-free subsets of the Alternating Groups . . . 499
António Girão , Zach Hunter, Yuval Wigderson: Blowups of triangle-free graphs 505
Lior Gishboliner, Stefan Glock, Peleg Michaeli, Amedeo Sgueglia: Defect and transference versions of the Alon–Frankl–Lovász theorem . . . 512
Stefan Glock, Olaf Parczyk, Silas Rathke, Tibor Szabó: On the maximum diameter of d-dimensional simplicial complexes . . . 518
Miriam Goetze, Fabian Klute, Kolja Knauer, Irene Parada, Juan Pablo Peña, Torsten Ueckerdt: Strong odd coloring in minor-closed classes . . . 524
J. Pascal Gollin, Attila Joó: Wild generalised truncation of infinite matroids . . . 531
Manuel González-Regadera, Raúl M. Falcón, Daniel Kotlar: Partial quasigroup digraphs . . . 537
Alexa Gopaulsingh, Zalán Molnár: On Distinguishing Graphs and Cost Number using Automorphism Representations . . . 543
Alexey Gordeev, Klas Markström, Lars-Daniel Öhman: Near Triple Arrays . . . 551
Alexey Gordeev: A Christofides-based approach to the travelling salesman problem in the unit cube . . . 559
Alexandr Grebennikov, Letícia Mattos, Tibor Szabó: On almost Gallai colourings in complete graphs . . . 566

Andrzej Grzesik, Justyna Jaworska, Bartłomiej Kielak, Piotr Kuc, Tomasz Ślusarczyk: Generalized Turán problem for directed cycles 572
Andrzej Grzesik, Justyna Jaworska, Piotr Kuc: Outdegree conditions forcing directed cycles 579
Andrzej Grzesik, Marek Skrzypczyk: Antidirected paths in oriented graphs . . . 586
Krystal Guo, Ross J. Kang, Gabriëlle Zwaneveld : Spectral approaches for d-improper chromatic number 591
Ervin Győri, Binlong Li, Nika Salia, Casey Tompkins, Kitti Varga and Manran Zhu: On graphs without cycles of length 0 modulo 4 597
Benedikt Hahn, Bettina Klinz, Birgit Vogtenhuber: On the geometric k-colored crossing number 603
Anastasia Halfpap, Bernard Lidický, Tomáš Masařík: Proper Rainbow Saturation Numbers for Cycles 612
Aneta Pokorná, David Hartman and Daniel Trlifaj: Reconstructing graphs and their connectivity using graphlets 618
David Hartman, Tomáš Hons and Jaroslav Nešetřil: All Vertices Of Nowhere Dense Modelings Are Traceable 625
Jędrzej Hodor, Hoang La, Piotr Micek, Clément Rambaud: Quickly excluding an apex forest 633
Nadzieja Hodur, Monika Pilśniak, Magdalena Prorok and Paweł Rzążewski: Finding large k-colorable induced subgraphs in (bull, chair)-free and (bull,E)-free graphs 640
Lawrence Hollom, Lyuben Lichev, Adva Mond, Julien Portier, Yiting Wang: Approximate Itai-Zehavi conjecture for random graphs 647
Jan Hubička, Matěj Konečný, Štěpán Vodseďálek, Andy Zucker: Counting big Ramsey degrees of the homogeneous and universal $\mathbf{K}_4$-free graph 653
Jan Hubička, Matěj Konečný, Stevo Todorcevic, Andy Zucker: On Big Ramsey degrees of universal ω-edge-labeled hypergraphs 660
Karolína Hylasová, Tomáš Kaiser: Hypertree shrinking avoiding low degree vertices 666
Daniel Iľkovič, Jun Yan: An improved hypergraph Mantel's Theorem 670
Jared León: Multi-cyclic graphs in the random graph process with restricted budget 676
Dawid Ignasiak, Lyuben Lichev: Hitting times and the power of choice for random geometric graphs 682
Mikhail Isaev, Tamás Makai, Brendan D. McKay: Enumeration of regular multipartite hypergraphs 687
Barnabás Janzer, Oliver Janzer, Abhishek Methuku, Gábor Tardos: Tight bounds for intersection-reverse sequences, edge-ordered graphs and applications 693
Barnabás Janzer, Raphael Steiner, Benny Sudakov: Chromatic number and regular subgraphs 700
Jialu Zhu and Jaroslav Nesetril: The interpolation theorem for asymmetric trees 706
Attila Jung, Dömötör Pálvölgyi: A note on infinite versions of (p, q)-theorems . . 712
Victoria Kaial, Annegret K. Wagler: Minimal non-EPT graphs based on minimal non-interval graphs 718

Ilyas Kalimullin, Eduard Lerner: The α-representation for Tait coloring and sums over spanning trees 724
Thomas Karam: Distance 2 versions and relative versions for difference patterns 731
Thomas Karam: Unions of intervals in codes based on powers of sets 737
Thomas Karam: On small densities defined without pseudorandomness 743
Thomas Karam: On distinguishing subsets of the Boolean cube by mod-p forms 749
Jasmin Katz, Matías Pavez-Signé, Jozef Skokan: Monochromatic trees in 2-coloured dense graphs 755
Derman Keskinkılıç, Lale Özkahya: Coloring Grids Avoiding Bicolored Paths . . 761
Benedek Kovács, Zoltán Lóránt Nagy, Dávid R. Szabó: Blocking Planes by Lines in PG(n, q) 766
Benedek Kovács, Zoltán Lóránt Nagy, Dávid R. Szabó: Resolution of the no-$(k+1)$-in-line problem when k is not too small 774
Jakub Kozik, Grzegorz Ryn, Toullec Streicher Ronan: Two-colorability of random non-uniform hypergraphs 781
Daniel Kráľ, Filip Kučerák, Ander Lamaison and Gábor Tardos: Uniform Turán density—palette classification 789
Péter L. Erdős, Gábor Lippner, Lajos Soukup: Every fully graphic degree sequence region is P-stable 795
Aleksei L. Perezhogin, Vladimir N. Potapov, Anna A. Taranenko, Sergey Yu. Vladimirov: Polystochastic matrices of order 4 with zero permanent 801
Ander Lamaison: Palettes determine uniform Turán density 807
Richard Lang, Nicolás Sanhueza-Matamala: A hypergraph bandwidth theorem . 812
Inês Legatheaux Martins: The rank partition of a matroid and partial symmetries on tensors 818
Shoham Letzter, Arjun Ranganathan: Tight Hamilton Cycles in Hypergraphs with Dichotomous Degree Conditions 826
Yuquan Lin, Patrice Patrice Ossona de Mendez: Path degeneracy and applications 833
Xujun Liu, Yan Wang: Partition subcubic planar graphs into independent sets . 840
Hexuan Liu, Bing-Ze Lu, Taoyang Wu, Guan-Ru Yu: Asymptotic Enumeration of Subclasses of Level-2 Phylogenetic Networks 849
Borut Lužar, Paweł Pękała, Jakub Przybyło, Roman Soták: On degree-balanced edge partitions 855
Yulai Ma, Davide Mattiolo, Eckhard Steffen, Isaak H. Wolf, Junxue Zhang: Perfect matchings in r-graphs 862
Tamás Makai, Matija Pasch, Kalina Petrova, Leon Schiller: Sharp thresholds for higher powers of Hamilton cycles in random graphs 869
Anna Margarethe Limbach, Robert Scheidweiler, Eberhard Triesch: 877
Klas Markström , Eero Räty: First-order convergence law for small permutation classes 882
Anders Martinsson, Raphael Steiner: Local Shearer Bound 888
Haruhide Matsuda: Local Degree Sum Conditions for the Existence of Paths and k-trees in Graphs 899
Luka Milićević: Low-codimensional Subvarieties Inside Dense Multilinear Varieties 904
Aleksa Milojević, Richard Montgomery, Alexey Pokrovskiy, Benny Sudakov: Properly colored Hamilton cycles and the Bollobás-Erdős conjecture 910

Kerri Morgan, Lluís Vena: Using induced subgraphs to express the coefficients of the chromatic polynomial . . . 916
Patrick Morris, Miquel Ortega, Juanjo Rué: Counting subsets of integers with no arithmetic progression of a given length . . . 921
Reza Naserasr, Lan Anh Pham, Cyril Pujol and Huan Zhou: Planar Signed Graphs with Fractional Balanced Chromatic Number Greater than 2 . . . 926
Rajko Nenadov, Lander Verlinde: Multipartite nearly orthogonal sets over finite fields . . . 932
Gyula O.H. Katona: Sharpenings of the Erdős-Ko-Rado theorem . . . 938
Alexander Omelchenko, Igor Labutin: Enumeration of 3-regular one-face maps on orientable or non-orientable surface up to all symmetries . . . 941
Alexander Omelchenko: Lattice Path Interpretations of the Pearson Distribution Systems . . . 947
Péter P. Pálfy: On the isomorphism problem of Cayley digraphs . . . 951
Dömötör Pálvölgyi, Kristóf Zólomy: Stabbing non-piercing sets and face lengths in large girth plane graphs . . . 955
Panna Tímea Fekete, Kitti Varga, Roland Molontay: Complexity of finding maximal color-avoiding connected subgraphs . . . 962
Irena Penev, R. B. Sandeep, D. K. Supraja, S Taruni: Isometric path partition: an upper bound and a characterization of some extremal graphs . . . 970
Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, Mihir Neve: Robustness of the Sauer–Spencer Theorem . . . 977
Alexey Pokrovskiy, Leo Versteegen, Ella Williams: Embedding trees using minimum and maximum degree conditions . . . 984
Clément Rambaud: Excluding a rectangular grid . . . 991
Pedro S. M. Arraes, Carla N. Lintzmayer, Guilherme O. Mota, Maycon Sambinelli: An improvement on the K_2-irregular index of bipartite graphs . . . 997
Igal Sason: On Counting Graph Homomorphisms by Entropy Arguments . . . 1003
Igal Sason: On H-Intersecting Graph Families . . . 1012
Robert Šámal: Loops in random duals . . . 1019
Robert Šámal, Hadi Zamani: Typical properties of countable graphs: flows and bridges . . . 1023
Thomas Selig, Haoyue Zhu: Prime graphical parking functions and strongly recurrent configurations of the Abelian sandpile model . . . 1030
Deping Song, Xuding Zhu: Bound on shortest cycle covers of graphs . . . 1039
Cameron Strachan, Konrad Swanepoel: Edge Isoperimetry of Lattices . . . 1045
He Sun, Danny Vagnozzi: Three hardness results for graph similarity problems . . . 1051
Lilla Tóthmérész: On canonical sandpile actions of embedded graphs . . . 1057
Chaoliang Tang, Hehui Wu, Junchi Zhang: The linear Turán number of the 3-graph P_5 . . . 1063
Matěj Trödler, Jan Volec, Jan Vybíral: Lower bounds on the minimal dispersion of point sets via cover-free families . . . 1069
Hans U. Simon, Jan Arne Telle: The Hierarchy of Saturating Matching Numbers 1074
Michael Wallner: The decompressed tree size of k-ary chains . . . 1081
Guanghui Wang, Yun Wang, Zhiwei Zhang: Arbitrary orientations of cycles in oriented graphs . . . 1087

Invited Talks

Matija Bucic: Pass to expander paradigm

The pass-to-expander paradigm is a very powerful and versatile tool in modern combinatorics. It allows one to reduce solving many questions about completely arbitrary graphs to solving them on (usually weak) expander graphs. We will discuss several variants of this method by means of recent applications, which include a solution to Graham's rearrangement conjecture over $\mathbb{F}_2^n$, an essentially tight bound for rainbow Turán numbers for all cycles (which found surprising applications to coding theory, additive combinatorics, and discrete geometry) and the classical Erdős-Gallai cycle decomposition conjecture.
This is based on joint works with: Bedert, Kravitz, Montgomery, and Müyesser; Alon, Sauermann, Zakharov, and Zamir; and Montgomery.

Vida Dujmović: Clustered Hadwiger Conjecture

Hadwiger's Conjecture asserts that every K_h-minor-free graph is properly $(h-1)$-colourable. Even the weaker bound of $O(h)$ colours remains open. I will discuss a relaxed variant, where monochromatic components of size bigger than one are allowed but must remain of bounded size. I will show that Hadwiger's Conjecture holds in this setting. Specifically, for fixed h, every K_h-minor-free graph is $(h-1)$-colourable with monochromatic components of bounded size. The number of colours is the best possible regardless of the size of monochromatic components. It solves an open problem of Edwards, Kang, Kim, Oum and Seymour [SIAM J. Disc. Math. 2015], and concludes a line of research initiated in 2007. This is joint work with Louis Esperet, Pat Morin and David R. Wood.

Peter Frankl: Half a Century With and Without Extremal Set Theory

My first paper, "The proof of a conjecture of G. O. H. Katona," appeared exactly 50 years ago (in *Journal of Combinatorial Theory, Series A*). In the lecture, I am going to state this result, but in this short abstract, let me describe the generic problem of Extremal Set Theory, instead.

Let $[n] = \{1, 2, \ldots, n\}$ be an n-set and $2^{[n]}$ its power set. Subsets of $2^{[n]}$ are called *families* (or *hypergraphs*). If all members of a family share the same size, say k, then it is called *k-uniform*. The generic problem is as follows: Impose some conditions on the family (e.g., no member contains another, no two are disjoint) and then ask for the maximum size that the family might have.

Many of the basic problems were solved by Paul Erdős and his collaborators. The major part of my research is connected, in one way or another, to papers of Erdős, in particular, the Erdős–Ko–Rado Theorem (proved in 1938). This fundamental result states that for positive integers k and t, and for $n > n_0(k,t)$, the largest k-uniform family in which any two members intersect in at least t elements is the *Full Star*: all k-sets containing a fixed t-subset.

In 1976, I proved that $n_0(k,t) = (k-t+1)(t+1)$ for $t > 14$. I will elaborate on later developments that played a great role in pushing me to return to the mathematical battlefield after a twenty-year hiatus. Along with recent results, I am going to mention several exciting open questions.

Hong Liu: Chromatic, homomorphism, blowup thresholds and beyond

The classical chromatic/homomorphism threshold problems study density conditions that guarantee an H-free graph to have bounded complexity. In this talk, I will survey some recent developments, including an unexpected connection to the theory of VC dimension and also discrete geometry, an asymmetric version that we introduce to interpolate the two problems. If time permits, I will discuss two related problems, blowup and VC thresholds.

László Lovász: Matroids, submodular functions, and bubbles

Limit theories of growing graph sequences are reasonably well developed, at least in the two extreme cases: dense graphs and bounded-degree graphs. It is a natural further development to study limits of various structures associated with graphs: spectra, automorphism groups, flows, etc. One of these objects is the cycle matroid of the graph. To describe its limit, we need to consider "matroids" on infinite sets. Our approach has been to construct the limits of the rank functions of these matroids in the form of submodular functions on the Borel sets of (say) $[0,1]$.

It turns out that such setfunctions have been studied in the analysis literature since a famous paper of Choquet in 1954. In the meanwhile, matroids and other finite submodular functions (like cut capacities) have become a large and important field of combinatorial optimization, with applications to rigidity theory, information theory, and more. The surprising fact is that the two research lines have had virtually no interaction.

In this talk we illustrate how cross-fertilization occurs once this connection is discovered, and sketch a limit theory for submodular setfunctions and its connection with the limit theories of graphs.

Wojciech Samotij: Stability of large cuts in random graphs

A cut in a graph G is the set of edges that cross some partition of the vertices of G into two sets and a maximum cut of G is a cut with the largest size among all cuts. We will prove that the family of largest cuts in the binomial random graph $G_{n,p}$ exhibits the following stability property: If $1/n \ll p \leq 1 - \Omega(1)$, then, with probability $1 - o(1)$, there is a set of $n - o(n)$ vertices that is partitioned in the same manner by all maximum cuts of $G_{n,p}$. Moreover, the analogous statement remains true when one replaces maximum cuts with nearly-maximum cuts. We will then demonstrate how one can use this statement as a tool for showing that certain properties of $G_{n,p}$ that hold in a fixed cut hold simultaneously in all maximum cuts.

This talk is based on a joint work with Ilay Hoshen (Tel Aviv University) and Maksim Zhukovskii (University of Sheffield).

Julia Wolf: Efficient regularity in hypergraphs

Szemerédi's celebrated regularity lemma states, roughly speaking, that the vertex set of any large graph can be partitioned into a bounded number of sets in such a way that all but a small proportion of pairs of parts induce a 'regular' graph.

Gowers proved in 1997 that in general, the number of parts needed in such a regularity partition may be as large as a tower in the regularity parameter. However, Alon, Fischer and Newman showed in 2007 that if the large graph in question has bounded VC-dimension, then the number of parts in the partition is polynomial in the regularity parameter, and in addition the density of each regular pair is either close to zero or close to 1.

The example of the half-graph demonstrates that the existence of irregular pairs cannot be ruled out in general. Recognising the half-graph as an instance of the so-called 'order property' from model theory, Malliaris and Shelah proved in 2014 that if one assumes that the large graph contains no half-graph of a fixed size (as a bi-induced subgraph), then it is possible to obtain a regularity partition with no irregular pairs.

In joint work with Caroline Terry in 2021, we obtained analogous results in 3-uniform hypergraphs. These have prompted further research in both model theory and combinatorics, and given rise to several open questions. The talk will survey these recent developments.

Contributed Talks

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The Zarankiewicz Problem for Polygon Visibility Graphs

(Extended abstract)

Eyal Ackerman* Balázs Keszegh†

Abstract

We prove a quasi-linear upper bound on the size of $K_{t,t}$-free polygon visibility graphs. For visibility graphs of star-shaped and monotone polygons we show a linear bound. In the more general setting of n points on a simple closed curve and visibility pseudo-segments, we provide an $O(n \log n)$ upper bound and an $\Omega(n\alpha(n))$ lower bound.

1 Introduction

Determining the maximum number of edges in a bipartite graph which does not contain the complete bipartite graph $K_{s,t}$ as a subgraph is known as the Zarankiewicz Problem and is a central problem in Extremal Graph Theory. Since every graph has a bipartite subgraph with at least half of the edges, it is essentially equivalent to consider this problem for arbitrary host graphs. It is also customary to study the symmetric case, that is, the maximum number of edges in a $K_{t,t}$-free n-vertex graph. The classical Kővári-Sós-Turán Theorem [21] gives the upper bound $O_t(n^{2-1/t})$ which is known to be asymptotically tight in some cases, namely, for $t = 2, 3$. It is a major open problem in combinatorics to determine whether this bound is tight for greater values of t.

The Kővári-Sós-Turán upper bound can be improved for certain classes of host graphs. For example, Fox, Pach, Sheffer, Suk and Zahl [7] provided an $O_{t,d}(n^{2-1/d})$ upper bound for $K_{t,t}$-free graphs of VC-dimension at most d (see also [8, 17, 18]), whereas Girão and Hunter [13] and Bourneuf, Bucić, Cook, and Davies [2] proved an $O_t(n^{2-\varepsilon(H)})$ bound for $K_{t,t}$-free graphs that avoid a given bipartite graph H as an induced subgraph. For other results concerning

*Department of Mathematics, Physics and Computer Science, University of Haifa at Oranim, Tivon 36006, Israel. ackerman@math.haifa.ac.il

†HUN-REN Alfréd Rényi Institute of Mathematics and ELTE Eötvös Loránd University, Budapest, Hungary. Research supported by the National Research, Development and Innovation Office – NKFIH under the grant K 132696 and by the ERC Advanced Grant "ERMiD". This research has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the ELTE TKP 2021-NKTA-62 funding scheme.

graphs of certain structural properties see [16] and the references within and also the recent survey by Du and McCarty [5].

A different yet related line of research studies $K_{t,t}$-free graphs that arise from geometry, usually incidence or intersection graphs of geometric objects. For example, the incidence graph between n points and n lines[1] is $K_{2,2}$-free and hence has $O(n^{3/2})$ edges by the Kővári-Sós-Turán Theorem. The celebrated Szemerédi-Trotter Theorem [24] improves this bound to $O(n^{4/3})$ which is tight. Moreover, in some geometric settings linear or almost linear upper bounds are known, for instance, the maximum size of a $K_{t,t}$-free intersection graph of n curves in the plane[2] is $O_t(n)$ [6]. For further results on the Zarankiewicz Problem in geometric settings see the recent survey by Smorodinsky [23].

In this work we also study the Zarankiewicz Problem in a geometric setting, however, rather than incidence or intersection graphs we consider *visibility* graphs, and more specifically, *polygon visibility* graphs, their generalizations and some special cases. Given a simple polygon P and two points $p, q \in P$ we say that p and q are *mutually visible* with respect to P (or that they 'see' each other), if the straight-line segment $\overline{pq}$ is disjoint from the exterior of P. The *visibility graph* of P consists of vertices that correspond to the vertices of P and edges that correspond to pairs of mutually visible vertices. We prove the following quasi-linear bound on the size of a $K_{t,t}$-free visibility graph of a simple polygon.

Theorem 1.1. *For every integer $t > 0$ there is a constant c_t such that the following holds. Let P be a simple n-gon and let G be the visibility graph of P. If G is $K_{t,t}$-free, then it has $O(n2^{\alpha(n)^{c_t}})$ edges, where $\alpha(n)$ is the inverse Ackermann function.*

A simple polygon is *star-shaped* if it contains a point that sees every other point in the polygon. If every vertical line intersects the boundary of a polygon at most twice, then this polygon is *x-monotone.* We obtain linear upper bounds on the size of $K_{t,t}$-free visibility graphs of star-shaped polygons and of x-monotone polygons.

Theorem 1.2. *Let $t > 0$ be an integer, let P be a star-shaped polygon and let V be a set of n points on its boundary. If the visibility graph of V with respect to P is $K_{t,t}$-free, then it contains $O_t(n)$ edges.*

Theorem 1.3. *Let $t > 0$ be an integer, let P be an x-monotone polygon and let V be a set of n points on its boundary. If the visibility graph of V with respect to P is $K_{t,t}$-free, then it contains $O_t(n)$ edges.*

Note that the bounds for star-shaped and x-monotone polygons hold for any set of n points on their boundaries, whereas in Theorem 1.1 every vertex of the polygon is a vertex of the visibility graph. It is known that polygon visibility graphs are not hereditary, that is, an induced subgraph of a polygon visibility graph may not be the visibility graph of some other polygon. Induced subgraphs of polygon visibility graphs are equivalent to *curve visibility graphs* which are visibility graphs of points on a Jordan curve.[3] Du and McCarty [5] mention the following open problem which was the motivation for this work.

[1] In this graph the vertices correspond to points and lines and there is an edge between a vertex that represents a point and a vertex that represents a line if that point lies on that line.

[2] In this graph, known as *string graph*, every curve is represented by a vertex and every pair of intersecting curves by an edge.

[3] Two vertices in such a graph are adjacent if the straight-line segment between their corresponding points is disjoint from the exterior of the curve.

Problem 1 ([5]). *Does every n-vertex $K_{t,t}$-free curve visibility graph have $O_t(n)$ edges?*

Although we do not settle Problem 1, we were able to provide a negative answer for the more general class of *curve pseudo-visibility graphs* which are defined as follows (see [4]). Let K be a Jordan curve, let V be a set of points on K and let $\mathcal{L}$ be a set of pseudolines such that every pair of points in V lies on a pseudoline from $\mathcal{L}$. The *curve pseudo-visibility graph* $G_{\mathcal{L}}(K, V)$ consists of the vertex set V and edges between every pair of points such that the segment of the pseudoline connecting them is disjoint from the exterior of K.

Theorem 1.4. *For every positive integer n there is a set of n points V, a Jordan curve K and a set of pseudolines $\mathcal{L}$ such that the curve pseudo-visibility graph $G_{\mathcal{L}}(K, V)$ is $K_{3,3}$-free and has $\Omega(n\alpha(n))$ edges, where $\alpha(n)$ is the inverse Ackermann function.*

As for an upper bound, we have:

Theorem 1.5. *Let $t > 0$ be an integer and let $G_{\mathcal{L}}(K, V)$ be a $K_{t,t}$-free curve pseudo-visibility graph of a set of n points V that lie on a Jordan curve K such that every pair of points in V also lie on a pseudoline from a set of pseudolines $\mathcal{L}$. Then $G_{\mathcal{L}}(K, V)$ has $O_t(n \log n)$ edges.*

Related work. Visibility graphs of polygons (and other geometric creatures) have been studied extensively both from a combinatorial and a computational point of view, see, e.g., [10, 12, 11, 15, 29]. It is worth mentioning a recent result of Davies, Krawczyk, McCarty and Walczak [4] who proved that curve pseudo-visibility graphs are χ-*bounded*, that is, their chromatic number can be bounded from above in terms of their clique number.

If the maximum size of any $K_{t,t}$-free n-vertex graph in a certain family of graphs is at most $c_t n$, where c_t depends only on t, then this family of graphs is called *degree-bounded*.[4] By a remarkable result of Girão and Hunter [13], if a hereditary family of graphs is degree-bounded, then c_t can be bounded by a polynomial function. For a further discussion on degree-boundedness see the recent survey of Du and McCarty [5].

Organization. The bounds for curve pseudo-visibility graphs are discussed in Section 3. In Section 4 we consider the upper bound for polygon visibility graphs. Most of the proofs are omitted due to space limitations and can be found in the full version of this paper [1].

2 Terminology and tools

In our proofs we make use of ordered graphs, cyclically-ordered graphs and pattern-avoiding 0-1 matrices. Due to the space limitation we omit the definitions of these well-known terms.

Lemma 2.1. *Every ordered graph $G = (V, E, <)$ has an ordered bipartite subgraph $G' = (L \cup R, E', <)$ such that $v < N_{G'}(v)$ for every vertex $v \in L$; $v > N_{G'}(v)$ for every vertex $v \in R$ and $|E'| \geq |E|/4$.*

Any curve pseudo-visibility graph $G_{\mathcal{L}}(K, V)$ can naturally be considered as a cyclically-ordered graph. Furthermore, edges cross in the sense of cyclically-ordered graphs if and only if they cross in the geometric sense in these graphs. This fact is trivial for polygon and curve

[4]Since every $K_{t,t}$-free graph in the family has a vertex of a small degree – at most $2c_t$.

visibility graphs, and is not hard to prove for curve pseudo-visibility graphs, see [4, Lemma 2.3].

Let u and v be two vertices of an ordered graph G such that $u < v$. Following [4] we say that a sequence of edges $e_1, e_2, \ldots, e_k$ is a *crossing sequence from u to v* if u is the smaller endpoint of e_1, v is the greater endpoint of e_k, and e_i crosses e_{i+1}, for every $k = 1, 2, \ldots, k-1$. If G is cyclically-ordered, then $e_1, e_2, \ldots, e_k$ is a crossing sequence from u to v, if it is a crossing sequence from u to v in G_u. It is easy to see that if there are crossing sequences from u to v and from v to u in a curve visibility graph, then these vertices must be adjacent. Proving this for curve pseudo-visibility graphs is more tricky [4, Proposition 3.2].

3 Bounds for $K_{t,t}$-free curve pseudo-visibility graphs

Let M be a 0-1 matrix. We say that row i and column j *cross* if $M_{ij} = 1$ or $M_{ij_1} = M_{ij_2} = M_{i_1j} = M_{i_2j} = 1$ for some $i_1 < i < i_2$ and $j_1 < j < j_2$.

Theorem 3.1. *Let $t > 0$ be an integer and let M be an $n \times n$ 0-1 matrix without t rows $r_1, r_2, \ldots, r_t$ and t columns $c_1, c_2, \ldots, c_t$ such that for every $1 \leq i, j \leq t$ row r_i and column c_j cross. Then M has $O_t(n)$ 1-entries.*

Proof. Suppose that M is drawn as an $n \times n$ squares board in which there is a point at the center of every cell that corresponds to a 1-entry. For every row of the board draw a horizontal straight-line segment connecting the leftmost and rightmost points in that row (that might be the same point, in which case we get a degenerate segment). Similarly, draw for every column a vertical straight-line segment connecting the topmost and bottom-most points in that column. Clearly, every point that corresponds to a 1-entry is the intersection point of a horizontal segment and a vertical segment. Therefore, it is enough to bound the number of such intersections, that is, the number of edges in the intersection graph of the at most $2n$ horizontal and vertical segments. Since an intersection point at cell (i, j) implies that row i and column j cross, it follows that the intersection graph of the segments is $K_{t,t}$-free. By a result of Fox and Pach [6] this graph has at most $t(\log t)^{O(1)}n$ edges and the claim follows. □

Let $G = (V, E, <)$ be an (cyclically-) ordered graph. We say that two vertices u and v form a *double cherry* if $(u, v) \in E$ or there are edges $(u, v_1), (u, v_2), (v, u_1), (v, u_2) \in E$ such that $u_1 < u < u_2 < v_1 < v < v_2$. Note that in the latter case a double cherry in a cyclically-ordered graph is a special case of crossing sequences from u to v and from v to u.

Corollary 3.2. *Let $t > 0$ be an integer and let $G = (A \cup B, E, <)$ be a bipartite ordered graph such that $A < B$. If G does not contain two t-subsets $A' \subseteq A$ and $B' \subseteq B$, such that every two vertices from different subsets form a double cherry, then $|E| = O_t(|V|)$.*

Corollary 3.3. *Let $t > 0$ be an integer and let $G = (V, E, <)$ be an ordered graph. If G does not contain two t-subsets of vertices, such that every two vertices from different subsets form a double cherry, then $|E| = O_t(|V| \log |V|)$.*

Corollary 3.4. *Let $t > 0$ be an integer and let $G = (V, E, <)$ be a cyclically-ordered graph. Suppose that G does not contain two t-subsets of vertices, such that every two vertices from different subsets form a double cherry. Then $|E| = O_t(|V| \log |V|)$. If G is bipartite with bipartition $V = A\dot{\cup}B$ such that $A < B$, then $|E| = O_t(|V|)$.*

It follows from [4, Proposition 3.2] that if u and v form a double cherry in an ordered curve pseudo-visibility graph, then they are adjacent. Therefore, we can conclude the upper bound for curve pseudo-visibility graphs from Corollary 3.4.

Corollary 3.5. *Let $t > 0$ be an integer and let $G_{\mathcal{L}}(K, V)$ be a $K_{t,t}$-free curve pseudo-visibility graph of a set of n points V that lie on a Jordan curve K such that every pair of points in V also lie on a pseudoline from a set of pseudolines $\mathcal{L}$. Then $G_{\mathcal{L}}(K, V)$ has $O_t(n \log n)$ edges. Furthermore, for every $A, B \subseteq V$ such that $A < B$, G has $O_t(|A| + |B|)$ edges (a, b) such that $a \in A$ and $b \in B$.*

For the lower bound on the size of $K_{3,3}$-free curve pseudo-visibility graphs we use the following result of Walczak.

Theorem 3.6 (Walczak [28])**.** *For every integer n there is an ordered graph with n vertices and $\Omega(n\alpha(n))$ edges which does not contain H_1 as an ordered subgraph, where H_1 denotes the ordered graph consisting of five vertices $a < b < c < d < e$ and three edges (a, e), (b, d) and (c, e).*

Using Lemma 2.1 it is not hard to show that one can remove a fraction of the edges of such a graph and obtain a $K_{3,3}$-free subgraph. We then show that this subgraph can be realized as a curve pseudo-visibility graph.

4 Bounds for $K_{t,t}$-free polygon visibility graphs

Recall that in a polygon visibility graph every vertex of the given polygon is a vertex of its visibility graph. Clearly, every polygon visibility graph contains a triangulation of the corresponding polygon, and therefore contains a copy of $K_{2,2}$ for every two triangles with a common edge. We can also show that as soon as a polygon visibility graph contains crossing edges, it contains a copy of K_4 (and hence $K_{2,2}$) as a subgraph.

Theorem 4.1. *Let P be an n-gon and let G be the visibility graph of P. If G has more than $2n - 3$ edges, then G contains K_4 as a subgraph.*

Next we prove the bound on the size of a $K_{t,t}$-free polygon visibility graph which is stated in Theorem 1.1. For this we need the following.

Proposition 4.2. *Let P be a simple polygon and let $r_1, r_2, c_1, c_2, c_3, c_4$ be vertices of P that appear in this clockwise order on the boundary of P. If r_1 sees c_1 and c_3 and r_2 sees c_2 and c_4, then P has a vertex v such that $c_2 \leq_P v \leq_P c_3$ and v sees r_1 and r_2.*

Corollary 4.3. *Let P be a simple polygon and let $r_1, r_2, \ldots, r_k, c_1, c_2, \ldots, c_{2k}$ be vertices of P that appear in this clockwise order on the boundary of P. If r_i sees c_i and c_{i+k} for every $i = 1, 2, \ldots, k$, then P has a vertex v such that $c_k \leq_P v \leq_P c_{k+1}$ and v sees r_j for every $j = 1, 2, \ldots, k$.*

Proof. By applying Proposition 4.2 for $r_1, r_k, c_1, c_k, c_{k+1}$ and c_{2k} we conclude that there is a vertex v that sees r_1 and r_k such that $c_k \leq_P v \leq_P c_{k+1}$. Consider r_j for some $1 < j < k$. It follows from the cyclic order of the vertices that the edges (r_1, v) and (r_j, c_{k+j}) cross and so do the edges (r_k, v) and (r_j, c_j). Therefore, v and r_j form a double cherry and thus they are adjacent in the visibility graph of P. $\square$

Proof of Theorem 1.1. Let $G = (V, E)$ be a $K_{t,t}$-free visibility graph of a simple n-gon P and let $v_1, v_2, \ldots, v_n$ be the vertices of P listed in their clockwise order starting from an arbitrary vertex v_1. Denote by $G' = (L \cup R, E')$ the ordered bipartite subgraph guaranteed by Lemma 2.1 and let A' be its adjacency matrix. That is, the rows of A' correspond to vertices in L according to their order and its columns correspond to vertices in R according to their order.

Denote by M_t the $t \times (4t-2)$ 0-1 matrix we get by: taking the identity matrix I_t; concatenating to it $t-1$ copies of the $t \times 2$ matrix with exactly two 1-entries, at cells $(1,1)$ and $(t,2)$; and concatenating to the result the identity matrix I_t.

Given an $m \times n$ 0-1 matrix M, we denote by M^+ the matrix we get by adding to M a new last row and a new first column, setting the $(m+1,1)$ entry to 1 and all the other new entries to 0.

Consider M_t^+ and observe that it has exactly one 1-entry at each column, therefore, it follows from a result of Klazar [19] that the maximum number of 1-entries in an $n \times n$ 0-1 matrix which avoids M_t^+ is $O(n2^{\alpha(n)^{c_t}})$ for some constant c_t. Thus, it is enough to show that A' avoids M_t^+.

Suppose for contradiction that A' contains M_t^+ as a submatrix, let $l_1 < l_2 < \cdots < l_{t+1}$ be the vertices of G' that correspond to the rows of this submatrix and let $r_0 < r_1 < \cdots < r_{4t-2}$ be those vertices that correspond to its columns. Since $(l_{t+1}, r_0) \in E'$, it follows that $l_{t+1} < r_0$. Thus, we have a subgraph of G' whose vertices are $l_1 < l_2 < \cdots < l_t < r_1 < r_2 < \cdots < r_{4t-2}$ and whose edges correspond to M_t. Now for every $s = 0, 2, \ldots, t-1$, by applying Corollary 4.3 on $l_1, l_2, \ldots, l_t$, $r_1, r_2, \ldots, r_{t-1}$, r_{t+2s}, r_{t+2s+1}, $r_{3t}, \ldots, r_{4t-2}$ we conclude that there is a vertex u_s such that $r_{t+2s} \le u_s \le r_{t+2s+1}$ and u_s is a neighbor of each of $l_1, l_2, \ldots, l_t$. Therefore, $\{l_1, \ldots, l_t\} \cup \{u_0, \ldots, u_{t-1}\}$ is a bipartition of a $K_{t,t}$ subgraph in G, a contradiction. $\square$

Acknowledgments. We are grateful to Bartosz Walczak for his permission to include his proof of Theorem 3.6 in this paper. We also thank Gábor Tardos for informing us of this result.

References

[1] Eyal Ackerman and Balázs Keszegh. The Zarankiewicz problem for polygon visibility graphs. *CoRR*, abs/2503.09115, 2025. URL: https://arxiv.org/abs/2503.09115, arXiv:2503.09115.

[2] Romain Bourneuf, Matija Bucić, Linda Cook, and James Davies. On polynomial degree-boundedness. *Advances in Combinatorics*, 2024. URL: https://doi.org/10.19086/aic.2024.5.

[3] Vasilis Capoyleas and János Pach. A Turán-type theorem on chords of a convex polygon. *Journal of Combinatorial Theory, Series B*, 56(1):9–15, 1992. URL: https://www.sciencedirect.com/science/article/pii/009589569290003G, doi:10.1016/0095-8956(92)90003-G.

[4] James Davies, Tomasz Krawczyk, Rose McCarty, and Bartosz Walczak. Coloring polygon visibility graphs and their generalizations. *Journal of Combinatorial Theory, Series B*, 161:268–300, 2023. URL: https://www.sciencedirect.com/science/article/pii/S0095895623000187, doi:10.1016/j.jctb.2023.02.010.

[5] Xiying Du and Rose McCarty. A survey of degree-boundedness. *European Journal of Combinatorics*, 2024. URL: https://www.sciencedirect.com/science/article/pii/S019566982400177X, doi:10.1016/j.ejc.2024.104092.

[6] Jacob Fox and János Pach. Applications of a new separator theorem for string graphs. *Combinatorics, Probability and Computing*, 23(1):66–74, 2014. doi:10.1017/S0963548313000412.

[7] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. A semi-algebraic version of Zarankiewicz's problem. *J. Eur. Math. Soc.*, 19(6):1785–1810, 2017. URL: `https://ems.press/journals/jems/articles/14758`.

[8] Nóra Frankl and Andrey Kupavskii. On the Erdős-Purdy problem and the Zarankiewitz problem for semialgebraic graphs. *CoRR*, abs/2112.10245, 2021. URL: `https://arxiv.org/abs/2112.10245`, `arXiv:2112.10245`.

[9] Zoltán Füredi and Péter Hajnal. Davenport-Schinzel theory of matrices. *Discrete Mathematics*, 103(3):233–251, 1992. URL: `https://www.sciencedirect.com/science/article/pii/0012365X92903168`, `doi:10.1016/0012-365X(92)90316-8`.

[10] Subir Kumar Ghosh. On recognizing and characterizing visibility graphs of simple polygons. *Discret. Comput. Geom.*, 17(2):143–162, 1997. `doi:10.1007/BF02770871`.

[11] Subir Kumar Ghosh and Partha P. Goswami. Unsolved problems in visibility graphs of points, segments, and polygons. *ACM Comput. Surv.*, 46(2):22:1–22:29, 2013. `doi:10.1145/2543581.2543589`.

[12] Subir Kumar Ghosh, Thomas Caton Shermer, Binay Kumar Bhattacharya, and Partha Pratim Goswami. Computing the maximum clique in the visibility graph of a simple polygon. *Journal of Discrete Algorithms*, 5(3):524–532, 2007. Selected papers from Ad Hoc Now 2005. URL: `https://www.sciencedirect.com/science/article/pii/S1570866706000955`, `doi:10.1016/j.jda.2006.09.004`.

[13] António Girão and Zach Hunter. Induced subdivisions in $K_{s,s}$-free graphs with polynomial average degree, 2024. URL: `https://arxiv.org/abs/2310.18452`, `arXiv:2310.18452`.

[14] Sergiu Hart and Micha Sharir. Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6(2):151–177, Jun 1986. `doi:10.1007/BF02579170`.

[15] John Hershberger. An optimal visibility graph algorithm for triangulated simple polygons. *Algorithmica*, 4(1):141–155, 1989. `doi:10.1007/BF01553883`.

[16] Zach Hunter, Aleksa Milojevic, Benny Sudakov, and István Tomon. Kővári-Sós-Turán theorem for hereditary families. *J. Comb. Theory B*, 172:168–197, 2025. URL: `https://doi.org/10.1016/j.jctb.2024.12.009`, `doi:10.1016/J.JCTB.2024.12.009`.

[17] Oliver Janzer and Cosmin Pohoata. On the Zarankiewicz problem for graphs with bounded VC-dimension. *Comb.*, 44(4):839–848, 2024. URL: `https://doi.org/10.1007/s00493-024-00095-2`, `doi:10.1007/S00493-024-00095-2`.

[18] Chaya Keller and Shakhar Smorodinsky. Zarankiewicz's problem via ϵ-t-nets. In *International Symposium on Computational Geometry*, 2024. URL: `https://api.semanticscholar.org/CorpusID:270286589`.

[19] Martin Klazar. A general upper bound in extremal theory of sequences. *Comment. Math. Univ. Carolin.*, 33(4):737–746, 1992.

[20] Martin Klazar. Extremal problems for ordered (hyper)graphs: applications of Davenport–Schinzel sequences. *European Journal of Combinatorics*, 25(1):125–140, 2004. URL: `https://www.sciencedirect.com/science/article/pii/S0195669803001239`, `doi:10.1016/j.ejc.2003.05.001`.

[21] Tamás Kővári, Vera T. Sós, and Paul Turán. On a problem of K. Zarankiewicz. *Colloquium Mathematicum*, 3:50–57, 1954. URL: `https://api.semanticscholar.org/CorpusID:118684929`.

[22] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 107(1):153–160, 2004. URL: `https://www.sciencedirect.com/science/article/pii/S0097316504000512`, `doi:10.1016/j.jcta.2004.04.002`.

[23] Shakhar Smorodinsky. A survey of Zarankiewicz problem in geometry, 2024. URL: `https://arxiv.org/abs/2410.03702`, `arXiv:2410.03702`.

[24] Endre Szemerédi and William T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3:381–392, 1983. URL: `https://api.semanticscholar.org/CorpusID:1750834`.

[25] Gábor Tardos. personal communication.

[26] Gábor Tardos. On 0-1 matrices and small excluded submatrices. *Journal of Combinatorial Theory, Series A*, 111(2):266–288, 2005. URL: `https://www.sciencedirect.com/science/article/pii/S009731650500004X`, `doi:10.1016/j.jcta.2004.11.015`.

[27] Gábor Tardos. *Extremal theory of vertex or edge ordered graphs*, page 221–236. London Mathematical Society Lecture Note Series. Cambridge University Press, 2019.

[28] Bartosz Walczak. personal communication.

[29] Onur Çağırıcı, Petr Hliněný, and Bodhayan Roy. On colourability of polygon visibility graphs. *European Journal of Combinatorics*, 117:103820, 2024. Selected papers of EuroComb 2019. URL: `https://www.sciencedirect.com/science/article/pii/S0195669823001373`, `doi:10.1016/j.ejc.2023.103820`.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Variations on an intersection problem

(Extended abstract)

Kristina Ago*

Abstract

Let $\mathscr{F} \subseteq 2^{[n]}$ be an s-uniform family such that every two distinct sets have a nonempty intersection but intersect in at most $k-1$ elements. By the well-known Ray-Chaudhuri–Wilson theorem, since the intersections can take at most $k-1$ different values, we have $|\mathscr{F}| \leqslant \binom{n}{k-1}$. We prove that the assumption that the family is intersecting, when n is large enough compared to s (and $k < s$), implies that $|\mathscr{F}| \leqslant \frac{\binom{n-1}{k-1}}{\binom{s-1}{k-1}}$, which improves the existing upper bound. Further, we discuss how we can apply the obtained results for some non-uniform families.

This is a joint work with Gyula O. H. Katona.

1 Introduction

For a positive integer n, we write $[n]$ for the set $\{1, 2, ..., n\}$ and $2^{[n]}$ for the power set of $[n]$. A family $\mathscr{F} \subseteq 2^{[n]}$ is called *s-uniform* if $|F| = s$ for all $F \in \mathscr{F}$. We write $\binom{[n]}{s}$ for the family of all s-element subsets of $[n]$. We say that the family $\mathscr{F}$ is *intersecting* if $A \cap B \neq \varnothing$ for every $A, B \in \mathscr{F}$.

Many intersection theorems have appeared in the literature. Let us begin by presenting some relevant results. For a uniform set system $\mathscr{F}$, if the possible intersection sizes are limited to at most k distinct values, the Ray-Chaudhuri–Wilson theorem [7] provides an upper bound on $|\mathscr{F}|$.

Theorem 1.1. *If $\mathscr{F} \subseteq 2^{[n]}$ is a uniform family such that for every two distinct sets $A, B \in \mathscr{F}$ we have $|A \cap B| \in \{l_1, l_2, \ldots, l_k\}$, then*

$$|\mathscr{F}| \leqslant \binom{n}{k}.$$

*Department of Mathematics and Informatics, University of Novi Sad, Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia. E-mail: `kristina.ago@dmi.uns.ac.rs`. Supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (grants no. 451-03-137/2025-03/200125 and 451-03-136/2025-03/200125) and by the Hungarian Academy of Sciences (the *Domus* program).

If the considered family is not uniform, then the Frankl–Wilson theorem [4] provides an upper bound on the cardinality of the system (again, assuming that the intersection sizes are limited to at most k distinct values).

Theorem 1.2. *If $\mathscr{F} \subseteq 2^{[n]}$ is a family such that for every two distinct sets $A, B \in \mathscr{F}$ we have $|A \cap B| \in \{l_1, l_2, \ldots, l_k\}$, then*

$$|\mathscr{F}| \leqslant \sum_{i=0}^{k} \binom{n}{i}.$$

The Erdős–Ko–Rado theorem [3], considered a classic in the field, provides an upper bound on the cardinality of an r-uniform set system in which the intersection of any two sets is nonempty.

Theorem 1.3. *If $\mathscr{F} \subseteq \binom{[n]}{r}$ is an intersecting family, $n \geqslant 2r$, then*

$$|\mathscr{F}| \leqslant \binom{n-1}{r-1}.$$

The original proofs of Theorems 1.1 and 1.2 rely on techniques from linear algebra. Alon, Babai and Suzuki provided a remarkably elegant and concise proof for both theorems using the so-called *polynomial method* [1]. They also gave a generalization of the Ray-Chaudhuri–Wilson theorem. In our work we investigate a variant of the Ray-Chaudhuri–Wilson theorem in which we put some restrictions on the intersection sizes $\{l_1, l_2, \ldots, l_k\}$. Let $\mathscr{F} \subseteq 2^{[n]}$ be an s-uniform family such that every two distinct sets have a nonempty intersection but intersect in at most $k-1$ elements. By the Ray-Chaudhuri–Wilson theorem, since the intersections can take at most $k-1$ different values, we have $|\mathscr{F}| \leqslant \binom{n}{k-1}$. In this paper we prove that the assumption that the family is intersecting, when n is large enough compared to s, implies that $|\mathscr{F}| \leqslant \frac{\binom{n-1}{k-1}}{\binom{s-1}{k-1}}$. This result significantly improves the existing upper bound. Further, we discuss how we can apply the obtained results for some non-uniform families. Füredi and Frankl [5] conjectured that, for a non-uniform family $\mathscr{F}$ in which every two distinct sets have a nonempty intersection but intersect in at most $k-1$ elements, it holds that $|\mathscr{F}| \leqslant \sum_{i=1}^{k} \binom{n-1}{i-1}$. This was proved by Snevily first when n is sufficiently large [8] and later for the general case using the polynomial method [9]. Our approach provides a combinatorial proof for the case when n is sufficiently large.

2 Main results

The following theorem represents the main contribution of this paper.

Theorem 2.1. *Let $\mathscr{F} \subseteq \binom{[n]}{s}$ be a family, $3 \leqslant k < s < n$, such that for every two distinct sets $A, B \in \mathscr{F}$ it holds that $1 \leqslant |A \cap B| \leqslant k-1$. Then, if $n > n_0(s)$, we have*

$$|\mathscr{F}| \leqslant \frac{\binom{n-1}{k-1}}{\binom{s-1}{k-1}}. \tag{1}$$

Proof (Sketch). First, if there is an element, for example 1, such that $1 \in F$ for all $F \in \mathscr{F}$, then for all those $(k-1)$-element subsets that do not contain 1 it holds that they are included in at most one element of $\mathscr{F}$. Since each $F \in \mathscr{F}$ contains $\binom{s-1}{k-1}$ such $(k-1)$-element sets, we have $|\mathscr{F}|\binom{s-1}{k-1} \leqslant \binom{n-1}{k-1}$, thus (1) follows even without any restriction on n.

Next, if there is no element which is contained in all the sets, but there are two elements, for example 1 and 2, such that for all $F \in \mathscr{F}$, either $1 \in F$ or $2 \in F$ (or both), then there exist sets A and B such that $1 \in A$ but $2 \notin A$ and $2 \in B$ but $1 \notin B$. We obtain the desired result by counting those Fs for which it holds that $\{1, b\} \subseteq F$ or $\{2, a\} \subseteq F$, but here we will need the assumption that $n \geqslant 2s^2$.

Finally, if there are no two elements such that at least one of them is contained in every member of $\mathscr{F}$, then we first prove the following lemma.

Lemma 2.2. *Let us assume that there are no $r-1$ points such that at least one of them is contained in each member of $\mathscr{F}$. Let $l_r(s)$ denote the minimal cardinality of a collection of r-element sets such that, for each $F \in \mathscr{F}$, there exists a set R from the collection such that $R \subseteq F$. It holds that $l_3(s) \leqslant s^3$.*

Using this lemma, we count the number of sets in our system and obtain the desired result but now under the assumption that $n \geqslant s^{\frac{5}{2}}$. □

Note. An interesting observation about Lemma 2.2 is that we can prove $l_2(s) \leqslant s^2$, which is an almost optimal bound. Namely, we can consider for a prime power q the finite projective plane on $n = q^2 + q + 1$ points, where $s = q + 1$, thus $s^2 = q^2 + 2q + 1$. On the other hand, because of properties of finite geometries, we have $l_2(q + 1) \geqslant q^2 + q + 1$.

However, regarding Theorem 2.1, one can observe that particular counterexamples arise when n is not sufficiently large compared to s. For instance, consider the complements of the lines in the Fano plane. In this case $n = 7$, $s = 4$ and $k = 3$, so Theorem 2.1 would imply $|\mathscr{F}| \leqslant 5$, whereas we know that $|\mathscr{F}| = 7$. This counterexample belongs to a (somewhat) larger class of counterexamples determined by certain designs. Let X be a finite set of v points and β be a finite family of distinct s-subsets of X, called *blocks.* Then the pair $D = (X, \beta)$ is called a t-(v, s, λ) *design* if every t-subset of X occurs in exactly λ blocks. (We note that in the standard terminology in design theory—see, for example, [2]—we write k for the size of the blocks instead of s, but here we will use s to avoid any confusion with the notation in the rest of the paper.) In a t-(v, s, λ) design the number of the blocks b is $b = \frac{\lambda\binom{v}{t}}{\binom{s}{t}}$. Thus, for the number of the blocks in a 2-(v, s, λ) design we have $b = \frac{\lambda v(v-1)}{s(s-1)}$. A 2-$(v, s, \lambda)$ design is called *symmetric* if $b = v$ (this is one of a few equivalent conditions). From these relations we have that in case of symmetric 2-(v, s, λ) designs it holds that

$$v = \frac{s(s-1)}{\lambda} + 1. \tag{2}$$

In a symmetric 2-(v, s, λ) design every two distinct blocks have λ points in common. Let us consider a family $\mathscr{F}$ that is determined by a symmetric 2-$(v, s, 2)$ design. It is defined on $n = v$ elements, thus, from (2) we have $n = \frac{s(s-1)}{2} + 1$, all sets in this family have s elements, while the size of the intersection of every pair of its sets is 2. It can be calculated that for these parameters Theorem 2.1 would give

$$|\mathscr{F}| \leqslant \frac{s(s+1)}{4},$$

whereas, by the listed properties of symmetric designs, we have

$$|\mathscr{F}| = \frac{s(s-1)}{2} + 1.$$

For a symmetric $D = (X, \beta)$ design its *residual* is a $D' = (X, \beta')$ design in which $\beta' = \{B \setminus B' : B \in \beta, B' \neq B\}$ for some block $B' \in \beta$. A residual design of a symmetric 2-(v, s, λ) design has the following properties. The number of the points is $V = v - s$, the size of its blocks is $S = s - \lambda$, while the number of its blocks is $B = \frac{s(s-1)}{\lambda}$. A design is called *quasi-residual* if its parameters correspond to the parameters of a residual design of a symmetric design. The following theorem holds (see [6]).

Theorem 2.3. *In a quasi-residual design with $\lambda = 2$ we have $V = \frac{S(S+1)}{2}$ points, the size of the blocks is S, the number of the blocks is $B = \frac{(S+2)(S+1)}{2}$, while the size of the intersection of every pair of its sets is 1 or 2.*

Let us consider a family $\mathscr{F}$ which is determined by a quasi-residual 2-$(V, S, 2)$ design. It can be calculated that for these parameters Theorem 2.1 would give

$$|\mathscr{F}| \leqslant \frac{(S+2)(S^2+S-4)}{4(S-2)},$$

whereas, by properties of these designs, we have

$$|\mathscr{F}| = \frac{(S+2)(S+1)}{2}.$$

The following proposition can be proved in a similar manner to Theorem 2.1.

Proposition 2.4. *Let $\mathscr{F} \subseteq \binom{[n]}{k} \cup \binom{[n]}{k+1} \cup \cdots \cup \binom{[n]}{s}$ be a family such that for every two distinct sets $A, B \in \mathscr{F}$ it holds that $1 \leqslant |A \cap B| \leqslant k-1$. Then, if $n > n_0(s)$, we have*

$$|\mathscr{F}| \leqslant \binom{n-1}{k-1}.$$

This proposition enables us to provide a simple combinatorial proof of the following theorem when n is sufficiently large.

Theorem 2.5. *Let $\mathscr{F} \subseteq 2^{[n]}$ be a family such that for every two distinct sets $A, B \in \mathscr{F}$ it holds that $1 \leqslant |A \cap B| \leqslant k-1$. Then, if $n > n_0(s)$, we have*

$$|\mathscr{F}| \leqslant \sum_{i=1}^{k} \binom{n-1}{i-1}.$$

Proof (Sketch). Let $\mathscr{F} = \mathscr{F}_1 \cup \mathscr{F}_2 \cup \cdots \cup \mathscr{F}_s$, where $\mathscr{F}_i$ denotes the collection of the i-element sets in $\mathscr{F}$. Using the EKR theorem separately for the families $\mathscr{F}_i$, $i = 1, 2, \ldots, k-1$, we have $|\mathscr{F}_i| \leqslant \binom{n-1}{k-1}$ (we can use it since in these systems the sizes of the intersections are certainly not greater than $k-1$, so we only have to be sure that they are nonempty). For the rest of the system, if n is large enough, we use the previous proposition, and thus together we have the desired result. $\square$

The bound from this theorem is the best possible, since the system that contains those at most k-element subsets of $[n]$ which all have one element in common satisfies the mentioned conditions, and its size is exactly $\sum_{i=1}^{k} \binom{n-1}{i-1}$.

Finally, we mention that in Proposition 2.4, if n is not sufficiently large, we are able to construct counterexamples.

Claim. *For a positive integer d, let k be a positive integer, such that $(d-1) \mid (k-2)$, and let $n = 2k - 2 + \frac{k-2}{d-1}$. Then there is a family $\mathscr{F} \subseteq \binom{[n]}{\geqslant k}$ that is intersecting, the size of the intersection of every pair of its elements is at most $k-1$, but*

$$|\mathscr{F}| = \binom{n-1}{k-1} + d.$$

Proof (Sketch). Let $\mathscr{F}_k = \{F \in \binom{[n]}{k} : 1 \in F\}$ and let $\mathscr{F}^* = \{\{2, 3, \ldots, k\} \cup \{k+1+i \cdot l, k+2+i \cdot l, \ldots, k+l+i \cdot l\} : i \in \{0, 1, \ldots, d-1\}\}$, where $l = \frac{k-2}{d-1}$. Finally, let $\mathscr{F} = \mathscr{F}_k \cup \mathscr{F}^*$. It can be checked that $\mathscr{F}$ satisfies all the requirements. □

References

[1] Noga Alon, László Babai and Hiroshi Suzuki. Multilinear polynomials and Frankl–Ray-Chaudhuri–Wilson type intersection theorems. *Journal of Combinatorial Theory, Series A*, **58**:165–180, 1991.

[2] Charles J. Colbourn and Jeffrey H. Dinitz. *Handbook of Combinatorial Designs.* Chapman & Hall/CRC, 2007.

[3] Pál Erdős, Chao Ko and Richard Rado. Intersection theorems for systems of finite sets. *The Quarterly Journal of Mathematics. Oxford. Second Series*, **12**:313–320, 1961.

[4] Peter Frankl and Richard M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, **1**:357–368, 1981.

[5] Peter Frankl and Zoltán Füredi. Families of finite sets with missing intersections. *Colloquia Mathematica Societatis János Bolyai*, **37**:305–318, 1984.

[6] Jerald F. Lawless, Ronald C. Mullin and Ralph G. Stanton. Quasi-residual designs. *Aequationes Mathematicae*, **2**:274–281, 1969. 10.

[7] Dijen K. Ray-Chaudhuri and Richard M. Wilson. On t-designs. *Osaka Journal of Mathematics*, **12**:737–744. 1975.

[8] Hunter S. Snevily. On generalizations of the de Bruijn-Erdős theorem. *Journal of Combinatorial Theory, Series A*, **68**:232–238, 1994.

[9] Hunter S. Snevily. A sharp bound for the number of sets that pairwise intersect at k positive values. *Combinatorica*, **23**:527–533, 2003.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Coloring, list coloring, and fractional coloring in intersections of matroids

(Extended abstract)

Ron Aharoni* Eli Berger† He Guo‡ Dani Kotlar§

Abstract

It is known that in matroids the difference between the chromatic number and the fractional chromatic number is smaller than 1, and that the list chromatic number is equal to the chromatic number. We investigate the gap within these pairs of parameters for hypergraphs that are the intersection of a given number k of matroids. We prove that in such hypergraphs the list chromatic number is at most k times the chromatic number and at most $2k-1$ times the maximum chromatic number among the k matroids. We study the relationship between three polytopes associated with k-sets of matroids, and connect them to bounds on the fractional chromatic number of the intersection of the members of the k-set. This also connects to bounds on the matroidal matching and covering number of the intersection of the members of the k-set. The tools used are in part topological.

1 Preliminaries

1.1 Hypergraphs and complexes

A *hypergraph* $\mathcal{H}$ is a collection of subsets, called *edges*, of its *vertex set* (or *ground set*), a finite set $V = V(\mathcal{H})$. Throughout the paper we assume that there is no isolated vertex in $\mathcal{H}$, i.e., every vertex $v \in V(\mathcal{H})$ belongs to some edge of $\mathcal{H}$. If all edges of a hypergraph $\mathcal{H}$ are of the same size k we say that $\mathcal{H}$ is *k-uniform*, or that it is a *k-graph*. A hypergraph $\mathcal{H}$ is *k-partite* if its vertex set can be divided into k parts $V_1, \dots, V_k$ so that for every edge $S \in \mathcal{H}$, $|S \cap V_i| = 1$ for each $1 \le i \le k$. For $U \subseteq V$, $\mathcal{H}[U] = \{S \in \mathcal{H} \mid S \subseteq U\}$ is the *subhypergraph* of $\mathcal{H}$ *induced* on U.

*Faculty of Mathematics, Technion, Haifa 32000, Israel. E-mail: `raharoni@gmail.com`.

†Department of Mathematics, University of Haifa, Haifa 31905, Israel. E-mail: `berger.haifa@gmail.com`.

‡Department of Mathematics and Mathematical Statistics, Umeå University, Umeå 90187, Sweden. E-mail: `he.guo@umu.se`. Supported by the Kempe Foundation grant JCSMK23-0055.

§Computer Science Department, Tel-Hai College, Upper Galilee 12210, Israel. E-mail: `dannykot@telhai.ac.il`.

A *matching* in a hypergraph is a set of disjoint edges, and a *cover* is a set of vertices meeting all its edges. The set of matchings in $\mathcal{H}$ is denoted by $\mathcal{M}(\mathcal{H})$. The maximum size of a matching in $\mathcal{H}$ is denoted by $\nu(\mathcal{H})$, and the minimum size of a cover by $\tau(\mathcal{H})$. Obviously, $\tau(\mathcal{H}) \geq \nu(\mathcal{H})$.

A set of vertices of $\mathcal{H}$ is said to be *independent* if it does not contain an edge. The complex of independent sets in $\mathcal{H}$ is denoted by $\mathcal{I}(\mathcal{H})$. Clearly, both $\mathcal{M}(\mathcal{H})$ and $\mathcal{I}(\mathcal{H})$ are non-empty (both include the empty set) and closed under taking subsets. Hypergraphs satisfying these two conditions are called *(abstract simplicial) complexes*, and their edges are also called *faces*. This terminology is borrowed from topology.

We assume (except for one explicit deviation — Definition 1.1 of the "join" below) that all complexes have the same set, denoted by V, as their ground set.

For a complex $\mathcal{C}$ and $U \subseteq V$, let $rank_{\mathcal{C}}(U) = \max_{S \in \mathcal{C}[U]} |S|$. The *rank* of $\mathcal{C}$, denoted by $rank(\mathcal{C})$, is $rank_{\mathcal{C}}(V)$.

A complex $\mathcal{C}$ is said to be a *flag complex* if it is 2-determined, meaning that $e \in \mathcal{C}$ whenever $\binom{e}{2} \subseteq \mathcal{C}$. (Note that $\binom{S}{m} = \{T \subseteq S \mid |T| = m\}$.)

Definition 1.1. The *join* $\mathcal{C} * \mathcal{D}$ of two complexes on disjoint ground sets is $\{A \cup B \mid A \in \mathcal{C}, B \in \mathcal{D}\}$. If $V(\mathcal{C}) \cap V(\mathcal{D}) \neq \emptyset$, we define $\mathcal{C} * \mathcal{D}$ by first making a copy of $\mathcal{D}$ on a ground set that is disjoint from that of $\mathcal{C}$, and then taking the join of $\mathcal{C}$ and the copy of $\mathcal{D}$.

1.2 Colorings, list colorings, and fractional colorings

Given a complex $\mathcal{C}$, a *coloring* by $\mathcal{C}$ is a set of faces of $\mathcal{C}$ whose union is the ground set V. The *chromatic number* $\chi(\mathcal{C})$ of $\mathcal{C}$ is the minimum size (number of faces) of a coloring.

Let L_v be a set of *permissible colors* at every $v \in V$. A *list coloring* with respect to these is a function $f : V \to \cup_{v \in V} L_v$ satisfying $f(v) \in L_v$ for every $v \in V$. It is said to be $\mathcal{C}$*-respecting* if $f^{-1}(c) \in \mathcal{C}$ for every color $c \in \cup_{v \in V} L_v$. The list chromatic number $\chi_\ell(\mathcal{C})$ is the minimum number p such that any system of lists L_v of size p has a $\mathcal{C}$-respecting list coloring.

Obviously, $\chi_\ell(\mathcal{C}) \geq \chi(\mathcal{C})$.

The second main concept we shall study is that of fractional colorings.

Definition 1.2. Given a complex $\mathcal{C}$ on V and $\vec{w} \in \mathbb{R}^V_{\geq 0}$, a $\vec{w}$*-fractional coloring* of $\mathcal{C}$ is a function $f : \mathcal{C} \to \mathbb{R}_{\geq 0}$ satisfying $\sum_{S \in \mathcal{C}: v \in S} f(S) \geq w(v)$ for every $v \in V$. The *fractional* $\vec{w}$*-chromatic number*, denoted by $\chi^*(\mathcal{C}, \vec{w})$, is the minimum of $\sum_{S \in \mathcal{C}} f(S)$ over all $\vec{w}$-fractional colorings f of $\mathcal{C}$. A $\vec{1}$-fractional coloring is plainly called a *fractional coloring*, and $\chi^*(\mathcal{C}, \vec{1})$ is denoted by $\chi^*(\mathcal{C})$.

1.3 Matroids

A complex $\mathcal{M}$ is called a *matroid* if for all $S, T \in \mathcal{M}$ satisfying $|S| < |T|$ there exists $v \in T \setminus S$ such that $S \cup \{v\} \in \mathcal{M}$. The edges of a matroid are said to be *independent*. A *base* of a matroid is a maximal independent set. A *circuit* is a minimal dependent set (the name comes from graphic matroids, namely matroids consisting of acyclic sets of edges in a graph). This is compatible with the terminology of independent sets in hypergraphs, once we note that $\mathcal{M} = \mathcal{I}(\mathcal{H})$ where $\mathcal{H}$ is the set of circuits.

For $A \subseteq V(\mathcal{M})$, let

$$span_{\mathcal{M}}(A) = A \cup \Big\{x \in V \mid \{x\} \cup S \notin \mathcal{M} \text{ for some } S \in \mathcal{M}[A]\Big\}.$$

$A \subseteq V$ is called *spanning* if $span_{\mathcal{M}}(A) = V$.

Throughout the paper we assume that all matroids are loopless, namely all singletons are independent.

Given a partition $\mathcal{P} = (P_1, P_2, \ldots, P_m)$ of V, let $\mathcal{T}(\mathcal{P})$ be the set of all subsets of V meeting each P_i for $1 \leq i \leq m$ in at most one vertex. $\mathcal{T}(\mathcal{P})$ is called *partition matroid.*

In a seminal paper [4], Edmonds showed how combinatorial duality (min-max results) can sometimes be formulated in terms of the intersection of two matroids. The classical case is the König–Hall theorem, which can be viewed as a statement on the intersection of two partition matroids, defined on the edge set of a bipartite graph. The parts in each are the stars in one of the sides. A matching in this graph is a set of edges belonging to the intersection of the two matroids, and a marriage is a base in one matroid that is independent in the other.

Intersections of more than two matroids are more complex: $\min - \max$ results are no longer available, algorithms for finding maximum size sets in the intersection are no longer polynomial, and necessary and sufficient conditions for the existence of certain objects are replaced by sufficient conditions.

Let $\mathcal{D}^k$ be the collection of sets $\mathcal{L} = \{\mathcal{M}_1, \ldots, \mathcal{M}_k\}$ of matroids. Let $MINT_k = \{\bigcap \mathcal{L} \mid \mathcal{L} \in \mathcal{D}^k\}$, namely the set of complexes that are the intersection of k matroids on the same ground set.

The topic of the paper is properties of complexes belonging to $MINT_k$. Though possibly familiar, it is worthwhile mentioning the primary facts about such objects. The first, that was proved more than once, is that every complex is in some $MINT_k$.

Proposition 1.3. *[9, 5, 11] Every complex is the intersection of matroids.*

Proof. For $e \subseteq V$, let $\mathcal{M}_e = \{S \subseteq V \mid S \not\supseteq e\}$. $\mathcal{M}_e$ is easily seen to be a matroid, and $\mathcal{C} = \bigcap_{e \notin \mathcal{C}} \mathcal{M}_e$. □

In a beautiful M.Sc thesis [8], András Imolay addressed the question of how many matroids are needed in the intersection. Let $\kappa(n)$ be the maximum, over all complexes $\mathcal{C}$ on n vertices, of the minimum number of matroids whose intersection is $\mathcal{C}$.

Theorem 1.4 (Theorem 6.7 in [8])**.** $\binom{n-1}{\lfloor (n-1)/2 \rfloor} \leq \kappa(n) \leq \binom{n}{\lfloor n/2 \rfloor}$.

A special role will be played by systems of partition matroids. Given a k-partite hypergraph $\mathcal{H}$, we construct a collection $\mathcal{L} = \{\mathcal{M}_1, \ldots, \mathcal{M}_k\}$ of k partition matroids, each having $E(\mathcal{H})$ as ground set, where each parts of $\mathcal{M}_i$ is a star S_x, namely the set of edges incident to a vertex x in the ith side of $\mathcal{H}$. This system of matroids is denoted by $\mathcal{L}(\mathcal{H})$.

Conversely, let $\mathcal{L} = \{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k\}$ be a system of partition matroids, defined by the partitions $\mathcal{P}^i = (P^i_1, P^i_2, \ldots, P^i_{m_i})$ of the ground set V for $i \in [k]$. Then the complex $\cap_{i=1}^k \mathcal{M}_i$ is the matching complex of a k-partite k-uniform hypergraph $\mathcal{K} = \mathcal{K}(\mathcal{L})$, whose vertices are the parts P^i_j, and each of its sides S_i is $\{P^i_j \mid 1 \leq j \leq m_i\}$. Every vertex $v \in V$ corresponds to an edge $e(v) = \{P^i_j \mid 1 \leq i \leq k, v \in P^i_j\}$ of $\mathcal{K}$. Then $\cap_{i=1}^k \mathcal{M}_i = \mathcal{M}(\mathcal{K})$, the set of matchings in $\mathcal{K}$. Clearly,

$$\mathcal{L}(\mathcal{K}(\mathcal{L})) = \mathcal{L} \quad \text{and} \quad \mathcal{K}(\mathcal{L}(\mathcal{K})) = \mathcal{K}. \tag{1}$$

These constructions are useful in characterizing intersections of partition matroids.

Proposition 1.5. *The following conditions are equivalent:*

(i) $\mathcal{C} = \mathcal{I}(G)$ *for some graph* G,

(ii) $\mathcal{C}$ is a flag complex,

(iii) $\mathcal{C}$ is the intersection of k partition matroids for some k,

(iv) $\mathcal{C}$ is the matching complex of a k-partite hypergraph for some k.

Moreover, (iii) *and* (iv) *are equivalent for each k separately.*

Proof. (i) $\Rightarrow$ (ii) is true by the definition of an independent set as a set in which each pair of elements is independent. To prove (ii) $\Rightarrow$ (i) note that if $\mathcal{C}$ is a flag complex then $\mathcal{C} = \mathcal{I}(G)$ for the graph whose edge set is $\{xy \mid xy \notin \mathcal{C}\}$. (iv) $\Rightarrow$ (i) is true since the matching complex $M(H)$ of a hypergraph H is $\mathcal{I}(L(H))$, where $L(H)$ is the line graph of the hypergraph H. (iii) $\Leftrightarrow$ (iv) follows from (1). To prove (ii) $\Rightarrow$ (iii) note that if $|e| = 2$ then the matroid $\mathcal{M}_e$ in the proof of Proposition 1.3 is a partition matroid, with parts e and all singletons disjoint from e. $\square$

1.4 Polytopes

Another viewpoint on the intersection of matroids, introdcued by Edmonds, is that of polytopes, which are particularly useful in studying fractional colorings.

For a subset A of V let $\mathbf{1}_A \in \mathbb{R}^V$ be the *characteristic function* of A, namely the function taking value 1 on elements of A and 0 elsewhere. Functions will also be viewed as vectors, so a real-valued function f on a set V is also denoted by $\vec{f} \in \mathbb{R}^V$. We write $f[V]$ for $\sum_{v \in V} f(v)$.

The polytope $P(\mathcal{C})$ of a complex $\mathcal{C}$ on V is the convex hull in $\mathbb{R}^V_{\geq 0}$ of the characteristic vectors of the edges of $\mathcal{C}$.

A polytope $Z \subseteq \mathbb{R}^V_{\geq 0}$ is *closed down* if $z \in Z$ and $0 \leq y \leq z$ imply $y \in Z$.

Observation 1.6. *$P(\mathcal{C})$ is closed down.*

Proof. Let
$$\vec{v} = \sum_{S \in \mathcal{C}} \lambda_S \mathbf{1}_S \in P(\mathcal{C}), \tag{2}$$
where $\lambda_S \geq 0$ and $\sum_{S \in \mathcal{C}} \lambda_S = 1$. Let $\vec{0} \leq \vec{u} \leq \vec{v}$. We claim that $\vec{u} \in P(\mathcal{C})$. Using induction, it suffices to consider the case that $\vec{u}$ and $\vec{v}$ differ just in one coordinate, namely $\vec{u} = \vec{v} - \alpha e_x$ for some $x \in V$. Furthermore, by the convexity of $P(\mathcal{C})$, it is enough to prove the case $\vec{u}(x) = 0$. Since $\mathcal{C}$ is a complex, we can replace all S's containing x in (2) by $S \setminus \{x\}$ to obtain $\vec{u}$. $\square$

1.5 Matroids intersection vs. matroidal cooperative covers

Given a k-set of matroids $\mathcal{L} = \{\mathcal{M}_1, \ldots, \mathcal{M}_k\} \in \mathcal{D}^k$, let $\bar{\nu}(\mathcal{L}) = rank(\bigcap \mathcal{L})$ and $\bar{\tau}(\mathcal{L}) = \min\{\sum_{i=1}^k rank_{\mathcal{M}_i}(V_i) \mid \cup_{i=1}^k V_i = V\}$. A set belonging to $\bigcap \mathcal{L}$ is called a *matoridal matching* of $\mathcal{L}$, and a k-tuple of functions $(\mathbf{1}_{V_1}, \ldots, \mathbf{1}_{V_k})$ for which $\cup_{i=1}^k span_{\mathcal{M}_i}(V_i) = V$ is called a *matroidal cooperative cover*. This is the reason for the notation $\bar{\nu}$ and $\bar{\tau}$.

Given a maximum size set X in $\bigcap \mathcal{L}$ and any representation of V as $\cup_{i=1}^k V_i$ we have
$$|X| = |X \cap (\cup_{i=1}^k) V_i| \leq \sum_{i=1}^k |X \cap V_i| \leq \sum_{i=1}^k rank_{\mathcal{M}_i}(V_i).$$
This shows that
$$\bar{\nu}(\mathcal{L}) \leq \bar{\tau}(\mathcal{L}). \tag{3}$$

Edmonds' famous two-matroids intersection theorem is that for $k = 2$ equality holds.

Theorem 1.7 ([4])**.** *If $k = 2$ then $\bar{\tau}(\mathcal{L}) = \bar{\nu}(\mathcal{L})$.*

We shall later prove that equality holds also for fractional versions of $\bar{\nu}$ and $\bar{\tau}$, for every k. It is not hard to prove that $\bar{\tau}(\mathcal{L}) \leq k\bar{\nu}(\mathcal{L})$ — we shall return to variants of this inequality, and also to the basic conjecture in the field, of which Edmonds' theorem is a special case.

Conjecture 1.8 (Conjecture 7.1 in [1])**.** *$\bar{\tau}(\mathcal{L}) \leq (k-1)\bar{\nu}(\mathcal{L})$.*

[1, Theorem 7.3] yields the conjecture for $k = 3$.

If $\mathcal{L} \in \mathcal{D}^k$ is a collection of partition matroids and $H = \mathcal{K}(\mathcal{L})$, which is a k-partite hypergraph, then $\bar{\nu}(\mathcal{L}) = \nu(H)$ and $\bar{\tau}(\mathcal{L}) = \tau(H)$. Thus a conjecture of Ryser, stating that in k-partite hypergraphs $\tau \leq (k-1)\nu$, is a special case of Conjecture 1.8.

2 Preview

The unifying theme of the paper is that belonging to $MINT_k$ entails quantifiably "tame" behavior of the complex. This is manifested in two ways. One is that the list chromatic number is not far from the ordinary chromatic number. This is proved via an observation, that a topological tool used to bound the chromatic number works just as well for the list chromatic number. This leads to two results. One is that if $\mathcal{C} \in MINT_k$ then

$$\chi_\ell(\mathcal{C}) \leq k\chi(\mathcal{C}).$$

This solves a conjecture posed by Király [10] and also by Bérczi, Schwarcz, and Yamaguchi [2]. The other result is that if $\mathcal{C} = \cap_{i=1}^k \mathcal{M}_i \in MINT_k$, then

$$\chi_\ell(\mathcal{C}) \leq (2k-1) \max_{1 \leq i \leq k} \chi(\mathcal{M}_i).$$

The proof of the latter requires new topological tools. If all matroids $\mathcal{M}_i$ are partition matroids then the $2k-1$ factor can be replaced by k, i.e.,

$$\chi_\ell(\mathcal{C}) \leq k \max_{1 \leq i \leq k} \chi(\mathcal{M}_i).$$

The latter is probably true (and also not sharp) for general matroids.

Then we deal with another aspect of the "tame-ness" of intersections of matroids — the behavior of fractional colorings. The motivation comes from two theorems of Edmonds on the intersection of two matroids. One is Theorem 1.7. The other, closely related, is that for any two matroids, $\mathcal{M}$ and $\mathcal{N}$, the following holds.

Theorem 2.1 ([3])**.** *$P(\mathcal{M} \cap \mathcal{N}) = P(\mathcal{M}) \cap P(\mathcal{N})$.*

An obvious challenge is to extend this theorem to any number of matroids. We shall define two notions: matchings and cooperative covers, as well as their fractional and weighted versions, for k-sets of matroids. We prove that the fractional weighted cooperative covering number equals the fractional weighted matching number for any k-set of matroids.

Following the footsteps of a result of Füredi [6] and its weighted version proved by Füredi, Kahn and Seymour [7], we shall consider a matroidal fractional weighted Ryser-type conjecture. We prove its equivalence to the following conjecture on polytopes.

Conjecture 2.2. *For $\mathcal{C} = \cap_{i=1}^{k} \mathcal{M}_i \in MINT_k$,*

$$(k-1)P(\mathcal{C}) \supseteq \cap_{i=1}^{k} P(\mathcal{M}_i).$$

Theorem 2.1 is the case $k = 2$.

Furthermore, we show an equivalence to yet another conjecture.

Conjecture 2.3. *For $\mathcal{C} = \cap_{i=1}^{k} \mathcal{M}_i \in MINT_k$ and $\vec{w} \in \mathbb{R}^V_{\geq 0}$,*

$$\chi^*(\mathcal{C}, \vec{w}) \leq (k-1) \max_{1 \leq i \leq k} \chi^*(\mathcal{M}_i, \vec{w}).$$

Using known results, we prove these conjectures for $k \leq 3$ and for partition matroids. For general k and general matroids, we prove these conjectures with k replacing $k-1$.

The second to the last section is devoted to the study of the fractional $\vec{w}$-chromatic number $\chi(\mathcal{C}, \vec{w})$ for general complexes $\mathcal{C}$.

The last section is devoted to a brief discussion of the combination of the two main themes — a fractional version of list coloring.

References

[1] R. Aharoni and E. Berger. The intersection of a matroid and a simplicial complex. *Transactions of the American Mathematical Society* **358**, (2006) 4895–4917.

[2] K. Bérczi, T. Schwarcz, and Y. Yamaguchi. List coloring of two matroids through reduction to partition matroids. *SIAM J. Discrete Math.* **35**, (2021) 2192–2209.

[3] J. Edmonds. Matroid intersection. *Ann. Discrete Math.* **4**, (1979) 39–49.

[4] J. Edmonds. Submodular Functions, Matroids, and Certain Polyhedra. *Combinatorial Optimization — Eureka, You Shrink!* Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. **2570** (2003), 11–26.

[5] S. P. Fekete, R. T. Firla, and B. Spille. Characterizing matchings as the intersection of matroids. *Math. Methods Oper. Res.* **58**, (2003) 319–329.

[6] Z. Füredi. Maximum degree and fractional matchings in uniform hypergraphs. *Combinatorica* **1**, (1981) 155–162.

[7] Z. Füredi, J. Kahn, and P. Seymour. On the fractional matching polytope of a hypergraph. *Combinatorica* **13**, (1993) 167–180.

[8] A. Imolay. Systems of representatives and their applications. *Master thesis, Eötvös Loránd University* (2023).

[9] K. Kashiwabara, Y. Okamoto, and T. Uno. Matroid representation of clique complexes. *Discrete Appl. Math.* **155**, (2007) 1910–1929.

[10] T. Király. Egres open: Research forum of the egerváry research group. 2013.

[11] B. Korte and D. Hausmann. An analysis of the greedy heuristic for independence systems. *Ann. Discrete Math.* **2**, (1978) 65–74.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Asymptotic half-grid and full-grid minors

(Extended abstract)

Sandra Albrechtsen* Matthias Hamann†

Abstract

We prove that every locally finite, quasi-transitive graph with a thick end whose cycle space is generated by cycles of bounded length contains the full-grid as an asymptotic minor and as a diverging minor. This in particular includes all locally finite Cayley graphs of finitely presented groups that are not virtually free, and partially solves problems of Georgakopoulos and Papasoglu and of Georgakopoulos and Hamann.

Additionally, we show that every (not necessarily quasi-transitive) graph of finite maximum degree which has a thick end and whose cycle space is generated by cycles of bounded length contains the half-grid as an asymptotic minor and as a diverging minor.

1 Introduction

Fat minors are a coarse or metric variant of graph minors. They first appeared in works of Chepoi, Dragan, Newman, Rabinovich and Vaxes [5] and of Bonamy, Bousquet, Esperet, Groenland, Liu, Pirot and Scott [3]. They play an important role in many (open) problems at the intersection of structural graph theory and coarse geometry – an area which can be described as 'coarse graph theory'.

A *model* of a graph X in a graph G is a collection of connected *branch sets* and *branch paths* in G such that after contracting each branch set to a vertex, and each branch path

*University of Hamburg, Department of Mathematics, Hamburg, Germany. E-mail: sandra.albrechtsen@uni-hamburg.de.

†University of Hamburg, Department of Mathematics, Hamburg, Germany. E-mail: matthias.hamann@math.uni-hamburg.de. Supported by DFG – Project No. 549406527.

to an edge, we obtain a copy of X. A model of X is *K-fat* (in G), for some $K \in \mathbb{N}$, if its branch sets and paths are pairwise at least K apart, except that we do not require this for incident branch set-path pairs. We say that X is a *(K-fat) minor* of G if G contains a (K-fat) model of X. The graph X is an *asymptotic minor* of G if X is a K-fat minor of G for every $K \in \mathbb{N}$. An important advantage of asymptotic minors over the usual minors is that they are preserved under quasi-isometries, and in particular, it does not depend on the choice of a finite generating set whether a Cayley graph of a finitely generated group contains a fixed graph as an asymptotic minor [14].

Recently, Georgakopoulos and Papasoglu [14] gave an overview of the area of 'coarse graph theory', where they presented results and open problems regarding the interplay of geometry and graphs, many of which concern fat minors. These problems have already attracted quite some attention; some (partial) solutions can be found in [1,2,5,6,11,12,17]. Our main contribution is a partial resolution of a problem of Georgakopoulos and Papasoglu about asymptotic grid minors in quasi-transitive graphs [14, Problem 7.3]. To state this problem, we first need some definitions.

An *end* of a graph G is an equivalence class of rays where two rays in G are equivalent if there are infinitely many pairwise disjoint paths between them in G. An end is *thick* if it has infinitely many pairwise disjoint rays. The *full-grid* is the graph on $\mathbb{Z} \times \mathbb{Z}$ in which two vertices (m, n) and (m', n') are adjacent if and only if $|m - m'| + |n - n'| = 1$, and the *half-grid* is its induced subgraph on $\mathbb{N} \times \mathbb{Z}$.

One of the cornerstones of infinite graph theory is *Halin's Grid Theorem* [15, Satz 4′], which asserts that every graph with a thick end contains the half-grid as a minor. Following this approach, Heuer [16] characterised the graphs containing the full-grid as a minor. These graphs form a proper subclass of the graphs with a thick end: while it is clearly true that every graph with a full-grid minor has a thick end, the converse is false in general, as the half-grid itself already witnesses. However, as it turned out, if we only consider graphs which are *quasi-transitive*, i.e. graphs whose vertex set has only finitely many orbits under its automorphism group, then these two graph classes coincide. Indeed, Georgakopoulos and the second author [13] showed that every locally finite, quasi-transitive graph with a thick end contains the full-grid as a minor.

Georgakopoulos and Papasoglu [14] asked whether this result can be generalised to the coarse setting in the following sense.

Problem 1.1. [14, Problem 7.3] *Let G be a locally finite Cayley graph of a one-ended finitely generated group. Must the half-grid be an asymptotic minor of G? Must the full-grid be an asymptotic minor of G?*

Note that every Cayley graph of a group is (quasi-)transitive. Moreover, the unique end of a one-ended, quasi-transitive graph is always thick [4,18].

Our main theorem partially answers both questions in the affirmative, under the additional assumption that G is a locally finite Cayley graph of a finitely presented group. In fact, we show the following result.

Theorem 1. *Let G be a locally finite, quasi-transitive graph whose cycle space is generated by cycles of bounded length. If G has a thick end, then the full-grid is an asymptotic minor of G.*

Note that Theorem 1 includes all locally finite Cayley graphs of finitely presented groups. Examples such as inaccessible graphs and groups [8, 9] or Diestel-Leader graphs [7, 10] indicate that the geometry of arbitrary locally finite, quasi-transitive (or Cayley) graphs may be far more involved. This is why generalising Theorem 1 to locally finite Cayley graphs of arbitrary finitely generated groups or even to all locally finite, quasi-transitive graphs may be much harder, and will require a different approach.

For the proof of Theorem 1 we construct, for every such graph G, a single model of the full-grid, which can be turned into a K-fat model of the full-grid, for every $K \in \mathbb{N}$, by deleting some of its branch sets and paths. Moreover, it can be turned into a model of the full-grid that *diverges*: for any two diverging sequences of vertices and/or edges of the full-grid also their branch sets/paths diverge in G.

Theorem 2. *Let G be a locally finite, quasi-transitive graph whose cycle space is generated by cycles of bounded length. If G has a thick end, then the full-grid is a diverging minor of G.*

This partially solves a question of Georgakopoulos and the second author [13, Problem 4.1].

As a first step in the proof of Theorem 1, we find the half-grid as an asymptotic minor. For this, we do not need the transitivity assumption on G. Indeed, we prove the following theorem.

Theorem 3. *Let G be a graph of finite maximum degree whose cycle space is generated by cycles of bounded length. If G has a thick end, then the half-grid is an asymptotic minor of G.*

Note that every graph satisfying the premise of Theorems 1 and 2 has finite maximum degree as it is locally finite and quasi-transitive.

Similar as in the proof of Theorem 1, we again construct a single model of the half-grid (see Theorem 2.1), which can be turned into a K-fat model of the half-grid, for every $K \in \mathbb{N}$, and into a diverging model of the half-grid.

Theorem 4. *Let G be a graph of finite maximum degree whose cycle space is generated by cycles of bounded length. If G has a thick end, then the half-grid is a diverging minor of G.*

This partially solves a question of Georgakopoulos and the second author [13, Problem 4.2].

2 The unifying results

We say that a model of the countably infinite clique $K_{\aleph_0}$ with branch sets $(V_i)_{i\in\mathbb{N}}$ and branch paths $(E_{ij})_{i\neq j\in\mathbb{N}}$ in a graph G is *ultra fat* if

- $d_G(V_i, V_j) \geq \min\{i, j\}$ for all $i \neq j \in \mathbb{N}$,
- $d_G(E_{ij}, E_{k\ell}) \geq \min\{i, j, k, \ell\}$ for all $i, j, k, \ell \in \mathbb{N}$ with $\{i, j\} \neq \{k, \ell\}$, and
- $d_G(V_i, E_{k\ell}) \geq \min\{i, k, \ell\}$ for all $i, k, \ell \in \mathbb{N}$ with $i \notin \{k, \ell\}$.

Further, we say that $K_{\aleph_0}$ is an *ultra fat minor* of G, and write $K_{\aleph_0} \prec_{UF} G$, if G contains an ultra fat model of $K_{\aleph_0}$. The idea is that an ultra fat model of $K_{\aleph_0}$ in a graph G witnesses that G contains $K_{\aleph_0}$ as an asymptotic minor. Indeed, if we remove all branch sets and branch paths with an index smaller than K from an ultra fat model of $K_{\aleph_0}$ in G, then this results in a K-fat model of $K_{\aleph_0}$ in G. So it contains every countable graph as a K-fat minor. Furthermore, in this situation, the graph contains every countable graph as a diverging minor, and in particular, it contains every countable graph of maximum degree at most 3 as a diverging subdivision.

Let G be a graph and let $H \subseteq G$ be a subdivision of the hexagonal half-grid with double rays S^i and paths P_{ij} connecting S^{i-1} and S^i. We say that H is *escaping* if there are $0 := M_0 < M_1 < \ldots \in \mathbb{N}$ such that $M_i > M_{i-1} + 2i$ for all $i \geq 1$ and

(i) $S^i \subseteq G[S^0, M_i] - B_G(S^0, M_{i-1} + 2i)$ for all $i \in \mathbb{N}_{\geq 1}$, and

(ii) $P_{1j} \subseteq G[S^0, M_1]$ and $P_{ij} \subseteq G[S^0, M_i] - B_G(S^0, M_{i-2} + i)$ for all $i \in \mathbb{N}_{\geq 2}$ and $j \in \mathbb{Z}$.

A subdivision $H \subseteq G$ of the hexagonal full-grid with double rays S^i and paths P_{ij} connecting $S^{|i|-1}$ and $S^{|i|}$ is *escaping* if the S^i with $i \geq 0$ and P_{ij} with $i \geq 1$ form an escaping subdivision of the hexagonal half-grid as well as the S^i with $i \leq 0$ and P_{ij} with $i \leq -1$, and if there is some $M \in \mathbb{N}$ such that the S^i with $i > 0$ are contained in a different component of $G - B_G(S^0, M)$ than the S^i with $i < 0$.

In order to obtain K-fat or diverging hexagonal half- or full-grids from an escaping hexagonal half- or full-grid, it essentially suffices to remove some of the double rays S^i with $1 \leq |i| \leq n$ for some $n \in \mathbb{N}$ and some of the paths P_{ij}.

We will prove Theorems 3 and 4 simultaneously by showing the following stronger result.

Theorem 2.1. *Let G be a graph with finite maximum degree whose cycle space is generated by cycles of bounded length. Then either $K_{\aleph_0} \prec_{UF} G$ or G contains an escaping subdivision of the hexagonal half-grid.*

Additionally, we may prescribe the thick end of G that contains all rays of the minor or subdivision.

Similarly as before, we prove Theorems 1 and 2 simultaneously by showing the following stronger result.

Theorem 2.2. *Let G be a locally finite, quasi-transitive graph whose cycle space is generated by cycles of bounded length. Then either $K_{\aleph_0} \prec_{UF} G$ or G contains an escaping subdivision of the hexagonal full-grid.*

Additionally, we may prescribe the thick end of G that contains all rays of the minor or subdivision.

3 A sketch of the proof

For the proof of Theorem 2.1, we first show that G contains for every thick end ε a diverging double ε-ray S^0. Second, we show that G contains double rays $S^1, S^2, \dots$ such that the S^i are contained in increasingly distant 'thickened cylinders' around S^0 of the form $G[S^0, M_i] - B_G(S^0, M_{i-1} + 2i)$ for some $M_0 < M_1 < \ldots \in \mathbb{N}$, as required by 2 for the double rays of an escaping subdivision of the hexagonal half-grid. Finally, we connect the S^i by infinitely many paths so that infinitely many of them either form the double rays of an escaping subdivision of the hexagonal half-grid or they form the branch sets of an ultra fat model of $K_{\aleph_0}$.

Let us describe the second step in more detail. We will choose the S^i recursively, starting from the diverging double ε-ray $S^0 = \ldots r_{-1} r_0 r_1 \ldots$. For this, we first show that $C \cap G[B_G(S^0, L + \lfloor \frac{\kappa}{2} \rfloor)]$ is connected for every $L \in \mathbb{N}$ and every component C of $G - B_G(S^0, L)$, where $\kappa \in \mathbb{N}$ is such that the cycle space of G is generated by cycles of length $\leq \kappa$. Note that this is the only part in in the proofs of Theorems 1 to 4 where we use the assumption on the cycle space; nevertheless, the assumption is crucial here, and the rest of the proof relies on this.

We then show that for every $L \in \mathbb{N}$ some component C of $G - B_G(S^0, L)$ is 'long', i.e. it has a neighbour in $B_G(Q, L)$ for all tails Q of S^0. Combining that C is long and $C \cap G[B_G(S^0, L + \lfloor \frac{\kappa}{2} \rfloor)]$ is connected then allows us to find a double ray in $C \cap G[B_G(S^0, L + \lfloor \frac{\kappa}{2} \rfloor)]$. Hence, we may proceed recursively by increasing the radius L of the ball around S^0 by a summand of $\lfloor \frac{\kappa}{2} \rfloor + 2i$ in each step.

The proof of Theorem 2.2 builds upon Theorem 2.1. From the proof of Theorem 2.1 it follows that we have control over where the escaping subdivision of the hexagonal half-grid lies. For this, let ε be a thick end of G, and let R be a diverging double ε-ray. Given a 'thick' component C of $G - B_G(R, K)$ for some $K \in \mathbb{N}$, that is one which includes a long component of $G - B_G(R, L)$ for every $L \geq K$, we in fact obtain an escaping subdivision H of the hexagonal half-grid whose first double ray is R and which is 'mostly' contained in C (unless we find an ultra fat model of $K_{\aleph_0}$, in which case we are immediately done). Now suppose that for some large enough $L \in \mathbb{N}$ there is another thick component D of $G - B_G(R, L)$. Then Theorem 2.1 yields another escaping subdivision H' of the hexagonal half-grid whose first double ray is R and which is 'mostly' contained in D (or an ultra fat model of $K_{\aleph_0}$). Gluing H and H' together along their common first double ray R then yields the desired subdivision of the hexagonal full-grid.

It thus suffices to prove that G contains a diverging double ε-ray R such that, for some large enough $K \in \mathbb{N}$, there are two distinct thick components of $G - B_G(R, K)$. This step is divided into two parts. We first show that if R' is a double ε-ray which is not only diverging but even quasi-geodesic, then it is enough that for some large enough $K \in \mathbb{N}$ there are distinct components $C \neq D$ of $G - B_G(R', K)$ such that C is thick but D is only 'half-thick' because then we can use the quasi-transitivity of G to find another quasi-geodesic double ε-ray R such that $G - B_G(R, K)$ has two distinct thick components. Here, a component of $G - B_G(R, K)$ is half-thick if it includes for every $L \geq K$ a component of $G - B_G(R, L)$ which is 'half-long', i.e. which has neighbours in $B_G(Q, L)$ for some tail Q of R.

Next, we show that such a double ray R' exists. For this, we first prove that G contains three ε-rays R_1, R_2, R_3 that intersect pairwise in a single common vertex such that $R_1 \cup R_2 \cup R_3$ is quasi-geodesic. Applying Theorem 2.1 to the quasi-geodesic, and hence diverging, double ray $R_1 \cup R_2$ then yields either an ultra fat model of K_∞, in which case we are immediately done, or an escaping subdivision H of the hexagonal half-grid whose first double ray is $R_1 \cup R_2$. Now for every $K \in \mathbb{N}$, by the definition of escaping, H will lie 'mostly' in one component C_K of $G - B_G(R, K)$, which then needs to be thick. We then analyse where R_3 lies in relation to H. If, for some large enough $L \in \mathbb{N}$, R_3 has a tail in a component $D_L \neq C_L$ of $G - B_G(R_1 \cup R_2, L)$, then we are done since D_L needs to be half-thick as R_3 diverges from $R_1 \cup R_2$ but lies in the same end as R_1 and R_2.

Otherwise, again since R_3 diverges from $R_1 \cup R_2$, it has a tail in C_K for all $K \in \mathbb{N}$. We then distinguish two cases. First assume that R_3 is far away from H. Then, since R_3 has a tail in each C_K, we can connect R_3 and H by infinitely many paths. These paths together with R_3 then yield infinitely many H-paths that 'jump over' H. We then use these paths together with H to find an ultra fat model of $K_{\aleph_0}$. Otherwise, R_3 lies close to H. Then either R_3 separates H into an 'upper half' containing (a tail of) R_1 and a 'lower half' containing (a tail of) R_2, and then $R_1 \cup R_3$ (or symmetrically $R_2 \cup R_3$) is the desired double ray R', or there are infinitely many H-paths that 'jump over' R_3, which then again yield an ultra fat model of $K_{\aleph_0}$.

References

[1] S. Albrechtsen, R. W. Jacobs, P. Knappe, and P. Wollan, *A characterisation of graphs quasi-isometric to K_4-minor-free graphs*, arXiv preprint (2024), arXiv:2408.15335.

[2] S. Albrechtsen, R. Diestel, A.-K. Elm, E. Fluck, R. W. Jacobs, P. Knappe, and P. Wollan, *A structural duality for path-decompositions into parts of small radius*, arXiv preprint (2023), arXiv:2307.08497.

[3] M. Bonamy, N. Bousquet, L. Esperet, C. Groenland, C.-H. Liu, F. Pirot, and A. Scott, *Asymptotic dimension of minor-closed families and Assouad–Nagata dimension of surfaces*, Journal of the European Mathematical Society **26** (2023), no. 10, 3739–3791.

[4] J. Carmesin, F. Lehner, and R. G. Möller, *On tree-decompositions of one-ended graphs*, Mathematische Nachrichten **292** (2019), no. 3, 524–539.

[5] V. Chepoi, F. F. Dragan, I. Newman, Y. Rabinovich, and Y. Vaxes, *Constant approximation algorithms for embedding graph metrics into trees and outerplanar graphs*, Discrete & Computational Geometry **47** (2012), 187–214.

[6] J. Davies, R. Hickingbotham, F. Illingworth, and R. McCarty, *Fat minors cannot be thinned (by quasi-isometries)*, arXiv preprint (2024), arXiv:2405.09383.

[7] R. Diestel and I. Leader, *A conjecture concerning a limit of non-Cayley graphs*, Journal of Algebraic Combinatorics **14** (2001), no. 1, 17–25.

[8] M. J. Dunwoody, *An inaccessible graph*, Random walks, boundaries and spectra, 2011, pp. 1–14.

[9] M. J. Dunwoody, *An inaccessible group*, Geometric group theory, Vol. 1 (Sussex, 1991), 1993, pp. 75–78.

[10] A. Eskin, D. Fisher, and K. Whyte, *Coarse differentiation of quasi-isometries I: Spaces not quasi-isometric to Cayley graphs*, Annals of Mathematics. Second Series **176** (2012), no. 1, 221–260.

[11] L. Esperet and U. Giocanti, *Coarse geometry of quasi-transitive graphs beyond planarity*, Electronic Journal of Combinatorics **31(2)** (2024).

[12] K. Fujiwara and P. Papasoglu, *A coarse-geometry characterization of cacti*, arXiv preprint (2023), arXiv:2305.08512.

[13] A. Georgakopoulos and M. Hamann, *A full Halin grid theorem*, Discrete & Computational Geometry (2024).

[14] A. Georgakopoulos and P. Papasoglu, *Graph minors and metric spaces*, arXiv preprint (2023), arXiv:2305.07456.

[15] R. Halin, *Über die Maximalzahl fremder unendlicher Wege in Graphen*, Mathematische Nachrichten **30** (1965), 63–85.

[16] K. Heuer, *Excluding a full grid minor*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **87** (2017), no. 2, 265–274.

[17] J. P. MacManus, *Fat minors in finitely presented groups*, arXiv preprint (2024), arXiv:2408.10748.

[18] C. Thomassen, *The Hadwiger number of infinite vertex-transitive graphs*, Combinatorica **12** (1992), no. 4, 481–491.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On the Wide Partition Conjecture for free matroids

(Extended abstract)

Jack Allsop* Dani Kotlar† Ian M. Wanless‡

Abstract

A Young diagram Y is called *wide* if every sub-diagram Z formed by a subset of the rows of Y dominates Z', the conjugate of Z. A Young diagram Y is called *Latin* if there exists an assignment of numbers to its squares such that the ith row is assigned the numbers $1, \dots, l_i$, where l_i is the length of the ith row, and the assignment is injective in each column. Chow and Taylor conjectured that every wide Young diagram is Latin. Chow et al. (2003) proved the conjecture for Young diagrams with two distinct row lengths. We prove it for three distinct row lengths.

1 Introduction

In [3] Chow et al. posed the following question:

Suppose the squares of a Young diagram are filled, each with an element of the ground set of a given matroid M, so that the rows are independent sets in M. Can the elements in each row be rearranged so that the columns are independent?

Their conjecture, named the Wide Partition Conjecture, or WPC, is that the answer is positive if and only if the Young diagram is so-called *wide*, which means that any subdiagram formed by a subset of its rows dominates its conjugate. The 'only if' part of the Wide Partition Conjecture conjecture can easily be shown to hold, so the conjecture is about the 'if' part.

Handling this problem for general matroids is hard, as even the 'easy' case in which all rows of the given Young diagram have the same length, is an open problem, known as Rota's bases conjecture [6]. So, as in [3], we focus on the seemingly much simpler case of a free matroid. We define a Young diagram to be *Latin* if there is an assignment of integers to its cells so that each row i of length l_i is populated by the numbers $1, \dots, l_i$, and the assignment is injective in each column. Chow et al. showed that for free matroids, the WPC can be reduced to the following conjecture:

*School of Mathematics, Monash University, Vic 3800, Australia. E-mail: Jack.Allsop@monash.edu.
†Computer Science Department, Tel-Hai College, Israel. E-mail: dannykot@telhai.ac.il.
‡School of Mathematics, Monash University, Vic 3800, Australia. E-mail: ian.wanless@monash.edu.

Conjecture 1.1. *Every wide Young diagram is Latin.*

Chow et al. [3] proved Conjecture 1.1 for Young diagrams with two distinct row lengths. They also proved that Conjecture 1.1 is true for self-conjugate Young diagrams with three distinct row lengths, and that proving the conjecture for all self-conjugate Young diagrams will suffice for establishing the conjecture for all Young diagrams. However, these results do not imply that Conjecture 1.1 holds for all Young diagrams with three distinct row lengths. The latter statement is the main result of our work. The proof is based on filling the diagram starting from the bottom, each time filling all rows of equal length using 1-factorizations of the bipartite graph that assigns a set of legal symbols to each column.

Two recent works apply other approaches to Conjecture 1.1. Chow and Tiefenbruck [4] obtained some partial results regarding a generalization of Conjecture 1.1 known as the Latin Tableau conjecture. This conjecture is about the existence of a Latin filling of a Young diagram, given the frequencies of the symbols. Aharoni et al. [1] took a hypergraph approach and showed that Conjecture 1.1 would follow from a conjectured equality between a matching parameter of a 3-partite 3-hypergraph related to a given Young diagram, and its covering dual.

2 Preliminaries

As a preliminary result, we have the following theorem about the complexity of verifying whether a Young diagram is wide:

Theorem 2.1. *Let Y be a Young diagram with p distinct row lengths $a_1 < a_2 < \cdots < a_p$, such that for $i = 1, \ldots, p$ the number of rows of length i is e_i. To check whether Y is wide, we need to perform no more than the following $O(p^2)$ verifications:*

$$a_k \geq \sum_{i=1}^{k} e_i, \quad k = 1, \ldots, p \tag{1}$$

and for $j = 1, \ldots, p-1$

$$\sum_{t=i}^{p} e_t a_t + (a_j - \mathrm{se}_{i,p}) a_{i-1} \geq \sum_{t=1}^{j-1} a_t e_t + a_j \sum_{t=j}^{p} e_t \quad \textit{if} \quad \mathrm{se}_{i,p} < a_j \leq \mathrm{se}_{i-1,p} \quad \textit{for} \quad i \in [2,p]. \tag{2}$$

where $\mathrm{se}_{i,p} := \sum_{t=i}^{p} e_t$ is the number of rows in the upper $p - i + 1$ row-blocks.

Hilton [5] defined the notion of an *outline rectangle*:

Definition 2.2. *Let C be an $m \times m$ matrix of multisets of symbols from $\{\tau_1, \tau_2, \ldots, \tau_m\}$. For $i \in [m]$ let ρ_i be the sum of the cardinalities of the multisets in row i, let c_i be the sum of the cardinalities of the multisets in column i, and let σ_i be the number of occurrences of symbol τ_i among all the multisets. Then C is an outline rectangle if there is some integer n such that the following is satisfied:*

(i) n divides each ρ_i, c_i, and σ_i,

(ii) cell (i,j) contains $\rho_i c_j / n^2$ symbols,

(iii) the number of occurrences of τ_k in row i is $\rho_i\sigma_k/n^2$, and

(iv) the number of occurrences of τ_k in column j is $c_j\sigma_k/n^2$.

Hilton [5] showed that an outline rectangle can be obtained from any $n \times n$ Latin square L with symbols in $[n]$ by introducing three sequences of positive integers $P = (p_i)_{i=1}^m$, $Q = (q_i)_{i=1}^m$, and $S = (s_i)_{i=1}^m$, so that $\sum_{i=1}^m p_i = \sum_{i=1}^m q_i = \sum_{i=1}^m s_i = n$, and amalgamating rows $p_0 + \cdots + p_{i-1} + 1, \ldots, p_0 + \cdots + p_i$, columns $q_0 + \cdots + q_{i-1} + 1, \ldots, q_0 + \cdots + q_i$, and symbols $s_0 + \cdots + s_{i-1} + 1, \ldots, s_0 + \cdots + s_i$, for $i = 1, \ldots, m$, where we assume that $p_0 = q_0 = s_0 = 0$. The array C thus obtained is called *the reduction modulo* (P, Q, S) *of* L.

Hilton's main result is:

Theorem 2.3 (Hilton [5])**.** *Each outline rectangle is the reduction modulo (P, Q, S) of some Latin square L, for some sequences P, Q, and S.*

We generalize the notion of outline rectangle to Young diagrams, by introducing the notion of *allocation*.

We assume that a Young diagram has p distinct row lengths $a_1 < a_2 < \cdots < a_p$ and for $i = 1, \ldots, p$ the number of rows of length i is e_i. Assuming this, we define a *row block* as a maximal set of rows consisting of all rows of the same length. Similarly, a column block is a maximal set of columns consisting of all columns of the same length. The symbol blocks are $S_1, S_2, \ldots, S_p$, where S_i corresponds to the symbols from $\{1, 2, \ldots, a_1\}$ if $i = 1$ and from $\{a_{i-1} + 1, a_{i-1} + 2, \ldots, a_i\}$ otherwise.

Definition 2.4. *For a Young diagram Y consisting of e_i rows of size a_i as defined above define $a_0 = 0$ and $b_i = a_i - a_{i-1}$ for $1 \le i \le p$. An* allocation *for Y is a choice of nonnegative integers z_{ijk} for $1 \le j, k \le i \le p$, satisfying*

$$\sum_{k=1}^{i} z_{ijk} = e_i b_j \qquad \text{for } 1 \le j \le i \le p, \tag{3}$$

$$\sum_{j=1}^{i} z_{ijk} = e_i b_k \qquad \text{for } 1 \le k \le i \le p, \tag{4}$$

$$\sum_{i=\max(j,k)}^{p} z_{ijk} \le b_j b_k \qquad \text{for } 1 \le j, k \le p. \tag{5}$$

If Y is Latin, then it has an allocation; specifically, given a Latin filling of Y we can let z_{ijk} be the number of symbols from symbol block k that appear in the rectangular region in Y where the rows from row block i and the columns from column block j intersect.

3 Main results

We prove the following:

Theorem 3.1. *If a Young diagram has an allocation, then it is Latin.*

We prove Theorem 3.1 by showing that a Young diagram that has an allocation can be embedded in the upper left corner of a larger outline rectangle. By Hilton's theorem, this outline rectangle is the reduction modulo (P, Q, S) of some Latin square L. So, the Young diagram is embedded in the this Latin square and the Latin filling of the square induces a Latin filling of our Young diagram.

Thus, Conjecture 1.1 reduces to the following conjecture:

Conjecture 3.2. *A wide Young diagram has an allocation.*

As mentioned above, the case of two distinct row lengths was proven in [3]. Two proofs were presented there. Showing that a wide Young diagram with two row lengths has an allocation is a simple matter. So, our work contains yet another proof of Conjecture 1.1 for this case.

For the general case with p distinct row lengths, we assume, by induction, that the sub-diagram consisting of the $p-1$ lower blocks has an allocation, that is, there exist non-negative values z_{ijk} with $1 \le j, k \le i \le p-1$ satisfying the requirements (3), (4), and (5). We define $(p-1)^2$ variables z_{pjk} with $j, k = 1, \dots, p-1$. The values of the remaining variables z_{pjk} where $j = p$ or $k = p$ are calculated from the values of the first $(p-1)^2$ variables by the allocation constraints (3) and (4). Then, we look at the constraints arising from the requirements (5) and $z_{pjk} \ge 0$ for all $j, k = 1, \dots, p$.

These constraints define upper and lower bounds for the individual variables z_{pjk} with $j, k = 1, \dots, p-1$:

$$0 \le z_{pjk} \le b_j b_k - \sum_{i=\max(j,k)}^{p-1} z_{ijk} \quad 1 \le j, k \le p-1 \tag{6}$$

They also define upper and lower bounds for sums of $p-1$ variables:

$$b_j(e_p - b_p) \le \sum_{k=1}^{p-1} z_{pjk} \le b_j e_p \quad j = 1, \dots, p-1, \tag{7}$$

and

$$b_k(e_p - b_p) \le \sum_{j=1}^{p-1} z_{pjk} \le b_k e_p \quad k = 1, \dots, p-1, \tag{8}$$

and finally for the sum of all $(p-1)^2$ variables:

$$(a_{p-1} - b_p)e_p \le \sum_{k=1}^{p-1} \sum_{j=1}^{p-1} z_{pjk} \le b_p^2 - b_p e_p + a_{p-1} e_p. \tag{9}$$

We show, using the wideness properties (1) and (2), that these bounds are valid. This means that each lower bound is smaller or equal to the corresponding upper bound, and that there are no contradictions among constraints in (6)–(9). It remains to show that we can find nonnegative z_{pjk} for $j, k = 1, \dots, p-1$ that satisfy both (7) and (8). If we show that the z_{pjk} satisfying (7) also satisfy the symmetry property $z_{pjk} = z_{pkj}$ for all $j, k = 1, \dots, p-1$, then we are done, since this will imply (8).

So far, we managed to show this only for the case $p = 3$. Thus, conjecture 3.2 is true when $p = 3$ and we have:

Theorem 3.3. *A wide Young diagram with at most three distinct row lengths is Latin.*

4 Alon-Tarsi conjecture for Young tableaux

In this work, we also discuss the question floated by Chow et al. [3] of whether the Alon-Tarsi conjecture [2] could be generalized from Latin squares to Latin Young tableaux. We show that if we generalize the notion of parity of rows and columns to Latin Young tableaux in the most obvious way, then the analogous conjecture fails to hold. However, the parity of columns in Latin Young tableaux may be generalized in different ways. For one possible such generalization, there are no counterexamples among diagrams with up to at least 20 boxes.

References

[1] Ron Aharoni, Eli Berger, He Guo, and Dani Kotlar. 2-covers of wide young diagrams. *arXiv preprint arXiv:2311.17670*, 2023.

[2] Noga Alon and Michael Tarsi. Colorings and orientations of graphs. *Combinatorica*, 12:125–134, 1992.

[3] Timothy Y Chow, C Kenneth Fan, Michel X Goemans, and Jan Vondrak. Wide partitions, latin tableaux, and rota's basis conjecture. *Advances in Applied Mathematics*, 31(2):334–358, 2003.

[4] Timothy Y Chow and Mark G Tiefenbruck. The latin tableau conjecture. *arXiv preprint arXiv:2408.04086*, 2024.

[5] A. J. W. Hilton. The reconstruction of Latin squares with applications to school timetabling and to experimental design. *Math. Programming Stud.*, (13):68–77, 1980.

[6] Rosa Huang and Gian-Carlo Rota. On the relations of various conjectures on latin squares and straightening coefficients. *Discrete Mathematics*, 128(1-3):225–236, 1994.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Alternating odd cycles in orientations of topologically χ-chromatic graphs

(Extended abstract)

Nóra Almási* Gábor Simonyi †

Abstract

An oriented odd cycle is called alternating if it has only one vertex with both its in-degree and out-degree positive. We investigate whether certain graphs can be oriented so that alternating odd cycles are avoided as subgraphs, or in such a way that all their shortest odd cycles become alternating. Our focus is on topologically χ-chromatic graphs, that are graphs for which the topological method gives a sharp lower bound on their chromatic number. Among others, these include Kneser graphs, Schrijver graphs and generalized Mycielski graphs.

The talk is based on the forthcoming paper [1].

1 Introduction

The local chromatic number of graphs was introduced in [4] and the notion was generalized to directed graphs in [6]. The latter parameter, called directed local chromatic number, is the minimum number of colors that must appear in the most colorful (closed) out-neighborhood of a vertex in a proper coloring, i.e., for a directed graph D it is defined as

$$\psi_d(D) := \min_c \max\{|c(N_+[v])| : v \in V(D)\},$$

where $N_+[v]$ is the closed out-neighborhood of vertex v (that includes v itself), c runs over all proper colorings of (the underlying undirected graph of) D and for a set $A \subseteq V(D)$ $c(A)$ denotes the set of colors appearing on vertices in A. (Considering undirected graphs as directed ones containing all edges in both directions, the above definition gives back the local chromatic number of undirected graphs). For undirected graphs it follows from the results in [4] that, somewhat surprisingly, there exist graphs with local chromatic number 3 and arbitrarily high

*Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, E-mail: almasinori@gmail.com

†HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary and Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, E-mail: simonyi@renyi.hu. Supported by the National Research, Development and Innovation Office (NKFIH) grant K–132696 of NKFIH Hungary.

chromatic number. It takes a similar argument to show that a directed graph can have directed local chromatic number 2 even when its chromatic number (meant to be the chromatic number of its underlying undirected graph) is arbitrarily large.

In what follows, we will use the notation $\psi_{d,\min}(G) := \min_{\overrightarrow{G}} \psi_d(\overrightarrow{G})$, where the minimization is over all oriented versions $\overrightarrow{G}$ of graph G.

We have $\psi_d(D) > 2$ if and only if D contains an alternating odd cycle, that is an oriented odd cycle, all but one vertex of which is a source or a sink, as a subgraph, cf. [13]. In this sense alternating odd cycles play a similar role in the theory of the directed local chromatic number as odd cycles do in the chromatic theory of undirected graphs.

The symmetric shift graph S_m can be defined as the line graph of the complete directed graph on m vertices, where complete means that all edges are included in both directions. Recall that for a digraph D its line graph is defined by

$$V(L(D)) = E(D), \quad E(L(D)) = \{\{(a,b),(c,d)\} : (a,b),(c,d) \in E(D), b = c \text{ or } a = d\}.$$

It is not hard to prove the following.

Fact 1. *An undirected graph G has $\psi_{d,\min}(G) = 2$ and it is attained by a coloring with m colors if and only if G admits a homomorphism into the symmetric shift graph S_m.*

We are mainly interested in the orientations of t-chromatic graphs that, using the terminology of [12], are also topologically t-chromatic. This basically means that (a particular implementation of) the topological method introduced by Lovász [7] to bound the chromatic number from below gives t as a sharp lower bound for the chromatic number. These graphs include, among others, Kneser graphs, Schrijver graphs and (generalized) Mycielski graphs. The respective definitions can be found in [7, 11, 5] or in the excellent book [8], but we also summarize them in the next section. Our main concern is whether various versions of these graphs can be oriented so that alternating odd cycles are completely avoided. We also investigate the question of whether we can orient them so that all of the shortest odd cycles become alternating. This latter question also fits into the framework called $\mathcal{F}$-free orientation problem in [2]. This problem asks whether, given a fixed finite set $\mathcal{F}$ of oriented graphs, a finite undirected graph admits an orientation in which no $F \in \mathcal{F}$ appears as an induced oriented subgraph.

2 Topologically χ-chromatic graphs

Definition 1. *The Kneser graph* $\mathrm{KG}(n,k)$ *with parameters $n \geq 2k$ is defined on the vertex set $\binom{[n]}{k}$ consisting of all k-element subsets of the n-element set $[n] = \{1,\dots,n\}$ with edge set*

$$E(\mathrm{KG}(n,k)) = \{\{A,B\} : A, B \in \binom{[n]}{k}, A \cap B = \emptyset\}.$$

Proving Kneser's conjecture, Lovász [7] proved that $\chi(\mathrm{KG}(n,k)) = n - 2k + 2$, and shortly afterwards Schrijver showed that a significantly smaller induced subgraph of $\mathrm{KG}(n,k)$ already has the same chromatic number, furthermore, it is also vertex-color-critical. These graphs are now called Schrijver graphs $\mathrm{SG}(n,k)$. Both proofs are famously based on topological tools, in particular, on the Borsuk-Ulam theorem.

Definition 2. *The Schrijver graph* $\mathrm{SG}(n,k)$ *with parameters* $n \geq 2k$ *is the induced subgraph of* $\mathrm{KG}(n,k)$ *on the following subset of the vertices:*

$$V(\mathrm{SG}(n,k)) = \{A \in \binom{[n]}{k} : \ \forall\, i \in [n-1],\ \{i,i+1\} \not\subseteq A, \{1,n\} \not\subseteq A\}.$$

Definition 3. *The r-level generalized Mycielskian $M_r(G)$ of a graph G is defined on the vertex set*

$$V(M_r(G)) = \{(v,i) : v \in V(G), 0 \leq i \leq r-1\} \cup \{z\}$$

with edges

$$E(M_r(G)) = \{\{(u,i),(v,j)\} : \{u,v\} \in E(G) \text{ and } (|i-j| = 1 \text{ or } i = j = 0)\} \cup$$
$$\{\{z,(v,r-1)\} : v \in V(G)\}.$$

The Mycielskian $M(G)$ of a graph G simply means the 2-level generalized Mycielskian $M_2(G)$.

When we simply say Mycielski graphs, it means the iterated Mycielskians of the single-edge graph K_2. Similarly, by generalized Mycielski graphs we mean the iterated generalized Mycielskians of K_2. Note that this way every odd cycle is a generalized Mycielski graph since $M_r(K_2) \cong C_{2r+1}$. Notice that the definition above also makes sense for $r = 1$.

It is proven by Stiebitz [14] (cf. also [5, 8]) that when starting with K_2, then the generalized Mycielski construction increases the chromatic number by 1 at every iteration. (For the original Mycielski construction this was proven by Mycielski [9], along with the easy fact that the clique number does not increase. These facts made Mycielski graphs one of the basic examples of arbitrarily high chromatic graphs with clique number 2.)

It is already proven in [13] that a topologically t-chromatic graph G has $\psi_{d,\min}(G) \geq \lceil t/4 \rceil + 1$, in particular, if $t \geq 5$ then it cannot have an orientation avoiding alternating odd cycles. If a graph is 3-colorable, then it is easy to orient it so, that we attain $\psi_{d,\min}(G) = 2$ (see [13]). Thus it is in the case of topologically 4-chromatic graphs when the question of orientability without creating alternating odd cycles is particularly interesting. It is also shown in [13] that 4-chromatic Schrijver graphs and generalized Mycielski graphs, whose defining parameters are large enough, admit an orientation which has directed local chromatic number 2 and this can be attained with 4 colors. Below we will see that requiring large parameters is not necessary for such a statement.

3 Avoiding alternating odd cycles

The above-mentioned result of [13], stating that 4-chromatic Schrijver graphs and generalized Mycielski graphs of large enough defining parameters admit an orientation and a 4-coloring attaining directed local chromatic number 2, implies by Fact 1 that these graphs admit a homomorphism to S_4. As noted above, we can show that this does not need the large parameters. At the same time, a similar statement is not true for Kneser graphs. This is summarized in the following two theorems. We use the usual notation $F \to G$ to indicate that a homomorphism from graph F to graph G exists, as well as $F \nrightarrow G$ for its negation.

Theorem 1. *1. We have for every $k \geq 2$*

$$\mathrm{SG}(2k+2,k) \to S_4.$$

2. For any positive integer k we have

$$\mathrm{KG}(n,k) \nrightarrow S_4.$$

Notice that for $k=1$ the graph $\mathrm{SG}(n,k)$ coincides with $\mathrm{KG}(n,k)$. While the proof of the first statement here is constructive, the proof for the second statement is based on a result of Peng-An Chen [3] that is also obtained via topology.

Theorem 2. *Let r_1, r_2 be positive integers. We have*

$$\psi_{d,\min}(M_{r_1}(M_{r_2}(K_2))) = 2$$

if and only if both $r_1, r_2 \geq 2$. In this case, the directed local chromatic number can be attained with 4 colors, that is,

$$M_{r_1}(M_{r_2}(K_2)) \to S_4,$$

while for every positive integer r and m

$$M_1(M_r(K_2)) \nrightarrow S_m \quad \text{and} \quad M_r(M_1(K_2)) \nrightarrow S_m$$

holds.

Note that Theorem 1 does not tell us if $\psi_{d,\min}(\mathrm{KG}(2k+2,k)) = 2$ can be true for large enough k (using more than 4 colors) or it never happens. At least for $k=2$ we can confirm that the latter is the case.

Proposition 1.

$$\psi_{d,\min}(\mathrm{KG}(6,2)) = 3.$$

4 Making odd cycles alternating

One can also turn around the question investigated in the previous section and ask if we can orient a graph so that we make at least all its shortest odd cycles alternating. The answer is obviously affirmative if the odd girth, i.e., the length of the shortest odd cycle is 3: it is enough to take an acyclic orientation. This is so just because the 3-cycle has only two non-isomorphic orientations, a cyclic and an acyclic one. Since this is not the case for longer odd cycles, for those the question becomes more interesting. For the 4-chromatic Mycielski graph $M(M(K_2)) = M(C_5)$, called Grötzsch graph, one can find an orientation where all its 5-cycles are alternating. In fact, even the so-called Clebsch graph, a remarkably symmetric graph of odd girth 5 that contains the Grötzsch graph as an induced subgraph, admits such an orientation where all C_5 subgraphs are alternating. Although these orientations are not immediate to find, they can be constructed using the following simple observation.

Lemma 1. *If the edge set of a graph G with odd girth 5 can be partitioned into a bipartite graph and a matching, then it admits an orientation in which all of its 5-cycles are alternating.*

Notice that every graph satisfying the conditions of Lemma 1 must be 4-colorable as it is the union of two bipartite graphs.

The next theorem about certain Kneser graphs shows, however, that for any positive integer $h \geq 3$ and odd number g we have graphs with odd girth equal to g and chromatic number at least h that admit an orientation where all the shortest odd cycles are alternating. Note the well-known fact [10], that the odd girth of Kneser graph $\mathrm{KG}(n, k)$ equals $1 + 2\left\lceil \frac{k}{n-2k} \right\rceil$. In particular, all the Kneser graphs in the statement below have odd girth equal to $2k+1$.

Theorem 3. *For any positive integers ℓ and k the Kneser graph $\mathrm{KG}(\ell(2k+1), \ell k)$ admits an orientation in which all its subgraphs isomorphic to C_{2k+1} become an alternating odd cycle.*

References

[1] Nóra Almási, Gábor Simonyi. Alternating odd cycles and orientations of Kneser-like graphs, paper in preparation.

[2] Manuel Bodirsky, Santiago Guzmán-Pro, Forbidden tournaments and the orientation completion problem, *SIAM J. Discrete Math.*, 39(1) (2025), 170–205.

[3] Peng-An Chen, A new coloring theorem of Kneser graphs, *J. Combin. Theory, Ser. A*, 118 (2011), 1062–1071.

[4] Paul Erdős, Zoltán Füredi, András Hajnal, Péter Komjáth, Vojtěch Rödl, Ákos Seress, Coloring graphs with locally few colors, *Discrete Math.*, 59 (1986), 21–34.

[5] András Gyárfás, Tommy Jensen, Michael Stiebitz, On graphs with strongly independent colour-classes, *J. Graph Theory*, 46 (2004), 1–14.

[6] János Körner, Concetta Pilotto, Gábor Simonyi, Local chromatic number and Sperner capacity, *J. Combin Theory Ser. B*, 95 (2005), 101-117.

[7] László Lovász, Kneser's conjecture, chromatic number, and homotopy, *J. Combin. Theory Ser. A*, 25 (1978), no. 3, 319–324.

[8] Jiří Matoušek, Using the Borsuk-Ulam theorem, *Springer-Verlag, Berlin*, (2003).

[9] Jan Mycielski, Sur le coloriage des graphs, *Colloq. Math.*, 3 (1955), 161–162.

[10] Svatopluk Poljak and Zsolt Tuza, Maximum bipartite subgraphs of Kneser graphs, *Graphs and Combinatorics*, vol. 3, (1987) no. 1, pp. 191–199.

[11] Alexander Schrijver, Vertex-critical subgraphs of Kneser graphs, *Nieuw Arch. Wisk. (3)*, 26 (1978), no. 3, 454–461.

[12] Gábor Simonyi, Gábor Tardos, Local chromatic number, Ky Fan's theorem, and circular colorings, *Combinatorica*, 26 (2006), 587–626.

[13] Gábor Simonyi, Gábor Tardos, On directed local chromatic number, shift graphs, and Borsuk-like graphs, *J. Graph Theory*, 66 (2011), 65–82.

[14] Michael Stiebitz, Beiträge zur Theorie der färbungskritischen Graphen, *Habilitation, TH Ilmenau, (1985).*

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Rainbow Triangles in Edge-colored Tripartite Graphs

(Extended abstract)

Yuyu An* Binlong Li* Fangfang Wu† Shenggui Zhang*

Abstract

Let G be an edge-colored tripartite graph with parts of sizes n_1, n_2 and n_3 such that $n_1 \geq n_2 \geq n_3 \geq 1$. Denote by $e(G)$ and $c(G)$ the number of edges of G and colors used in G. A triangle in G is called rainbow if all the colors on its edges are distinct. In this paper, we prove that if $e(G) + c(G) \geq \max\{2n_1(n_2 + n_3), 2n_1n_2 + n_3(n_1 + n_2 + 1)\} + 1$, then G contains a rainbow triangle, and show that this result is best possible.

1 Introduction

Let G be a graph. We use $e(G)$ to denote the number of edges in G. An *edge-coloring* of a graph G is a mapping $C : E(G) \to \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers. We call G an *edge-colored graph* if it is assigned such an edge-coloring. We say that an edge-colored graph G is *rainbow* if all the colors on its edges are distinct. We denote the number of colors of edges in G by $c(G)$.

How many edges are needed in a graph on n vertices to ensure the existence of a triangle? In 1907, Mantel provided the maximum edge number of a graph containing no triangles, which is one important starting point of extremal graph theory.

Theorem 1 (Mantel [8]). *Any graph G with n vertices and at least $\lfloor n^2/4 \rfloor$ edges contains a triangle, unless $G \cong K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.*

The systematic study of these problems began with Turán [9], who generalized Mantel's result to arbitrary complete graphs. What happens if we try to forbid other subgraphs and the ground graph is not necessarily complete? There is a general definition. Given graphs

*School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an, 710072, China. E-mail: anyuyu@mail.nwpu.edu.cn, binlongli@nwpu.edu.cn, sgzhang@nwpu.edu.cn. Supported by NSFC (Nos. 12131013, 12171393, 12301456, 12471334) and Shaanxi Fundamental Science Research Project for Mathematics and Physics (No. 22JSZ009).

†School of Mathematics and Statistics, Hunan Normal University, Changsha, 410081, China. E-mail: ffwu@hunnu.edu.cn

G and H, the *Turán number* $\mathrm{ex}(G,H)$ is defined to be the maximum number of edges of a subgraph of G without copies of H. So Mantel's theorem states that $\mathrm{ex}(K_n, C_3) = \lfloor n^2/4 \rfloor$.

Given graphs G and H, the *anti-Ramsey number* $\mathrm{ar}(G,H)$ is defined to be the maximum number of colors of an edge-coloring of G without rainbow copies of H. In 1973, Erdős, Simonovits and Sós [3] first studied the anti-Ramsey number of graphs and showed that these are closely related to Turán numbers. For the rainbow triangle in complete graphs, they got:

Theorem 2 (Erdős, Simonovits and Sós [3])**.** $\mathrm{ar}(K_n, C_3) = n - 1$.

Considering the sum of the number of edges and colors, Li, Ning, Xu and Zhang [6] extended the anti-Ramsey number result due to Erdős, Simonovits and Sós, and showed the following rainbow version of Mantel's theorem.

Theorem 3 (Li, Ning, Xu and Zhang [6])**.** *Let G be an edge-colored graph on n vertices. If $e(G) + c(G) \geq \binom{n}{2} + n$, then G contains a rainbow triangle.*

The bound for $e(G) + c(G)$ in the above theorem is best possible. Fujita, Ning, Xu, and Zhang [5] characterized the extremal graphs for Theorem 3.

It is natural to ask whether G contains a rainbow H if $e(G)+c(G) \geq e(K_n)+\mathrm{ar}(K_n, H)+1$. The main purpose of this article is to investigate the problem on the minimum possible sum of the number of edges and colors in an edge-colored graph containing certain rainbow subgraphs. Given graphs G and H, we define $f(G,H)$ to be the maximum sum of the number of edges and colors of an edge-colored subgraph of G without rainbow copies of H. By the definition, we have $f(G,H) = \max\{e(F) + \mathrm{ar}(F,H) : F \subseteq G\}$. By taking F to be G and the extremal graphs for $\mathrm{ex}(G,H)$, respectively, in the above formula, one may easily obtain the following inequality.

$$f(G,H) \geq \max\{2\mathrm{ex}(G,H), e(G) + \mathrm{ar}(G,H)\}. \tag{1}$$

Therefore, by inequality (1) and Theorem 3, one can obtain

$$f(K_n, C_3) = \binom{n}{2} + n - 1 = e(K_n) + \mathrm{ar}(K_n, C_3).$$

Afterwards, several results on $f(G,H)$ were established for a number of graphs, including, among others, cliques [11], quadrilaterals [10] and multiple cliques [2, 7].

2 Main result

In the present paper, we consider the rainbow triangles in tripartite graphs. The Turán number and the anti-Ramsey number of triangles in complete tripartite graphs K_{n_1,n_2,n_3} were given by Bollobás et al., and by Fang et al., respectively.

Let n_1, n_2, n_3 be three integers with $n_1 \geq n_2 \geq n_3 \geq 1$,

Theorem 4 (Bollobás, Erdős and Szemerédi [1])**.** $\mathrm{ex}(K_{n_1,n_2,n_3}, C_3) = n_1(n_2 + n_3)$.

Theorem 5 (Fang, Győri, Li and Xiao [4])**.** $\mathrm{ar}(K_{n_1,n_2,n_3}, C_3) = n_1 n_2 + n_3$.

We provided a sharp condition on the sum of the edge number and color number for the existence of triangles in general tripartite graphs.

Theorem 6. *Let G be an edge-colored tripartite graph with partite sets of cardinalities n_1, n_2, n_3 such that $n_1 \geq n_2 \geq n_3 \geq 1$. If*

$$\begin{aligned} e(G) + c(G) &> \max\{2n_1(n_2 + n_3),\ 2n_1n_2 + n_3(n_1 + n_2 + 1)\} \\ &= \max\{2\mathrm{ex}(K_{n_1,n_2,n_3}, C_3),\ e(K_{n_1,n_2,n_3}) + \mathrm{ar}(K_{n_1,n_2,n_3}, C_3)\}, \end{aligned}$$

then G contains a rainbow triangle.

Inequality (1) implies that the lower bound on $e(G) + c(G)$ in Theorem 6 is the best possible.

3 Sketch of the proof

Set $f(n_1, n_2, n_3) = \max\{2n_1(n_2 + n_3),\ 2n_1n_2 + n_3(n_1 + n_2 + 1)\}$ as in Theorem 6. Suppose that G is a counterexample to Theorem 6 with order as small as possible. We first show that G is not rainbow and then $e(G) + c(G) = f(n_1, n_2, n_3) + 1$ exactly. Next we distinguish the proof four cases, namely: (1) $n_1 = n_2 = n_3$; (2) $n_1 = n_2 > n_3$; (3) $n_1 > n_2 = n_3$; (4) $n_1 > n_2 > n_3$. For each case, we give bounds on $d(A) + d^s(A)$, the sum of the number of edges incident to A and the number of colors saturated by A, where A is some proper set of vertices we take. Using these bounds we compute a bound of $e(G) + c(G)$, which contradicting that $e(G) + c(G) = f(n_1, n_2, n_3) + 1$.

4 Remarks

Based on existing results, we propose the following conjecture.

Conjecture 1. *Let H be a given graphs. For n large enough,*

$$f(K_n, H) = \max\{2\mathrm{ex}(K_n, H),\ e(G) + \mathrm{ar}(K_n, H)\}.$$

References

[1] B. Bollobás, P. Erdős and E. Szemerédi, On complete subgraphs of r-chromatic graphs, Discrete Math., 13 (1975) 97–107.

[2] S. Ehard and E. Mohr, Rainbow triangles and cliques in edge-colored graphs, European J. Combin., 84 (2020) 103037.

[3] P. Erdős, M. Simonovits and V.T. Sós, Anti-Ramsey theorems, Colloquia Mathematica Societatis János Bolyai, 10. Infinite and finite sets, Keszthely (Hungary), 1973.

[4] C. Fang, E. Győri, B. Li and J. Xiao, The anti-Ramsey numbers of C_3 and C_4 in complete r-partite graphs, Discrete Math., 344 (2021) 112540.

[5] S. Fujita, B. Ning, C. Xu and S. Zhang, On sufficient conditions for rainbow cycles in edge-colored graphs, Discrete Math., 342 (2019) 1956–1965.

[6] B. Li, B. Ning, C. Xu and S. Zhang, Rainbow triangles in edge-colored graphs, European J. Combin., 36 (2014) 453–459.

[7] X.C. Liu, D. Peng and X. Yang, More on rainbow cliques in edge-colored graphs, European J. Combin., 124 (2025) 104088.

[8] W. Mantel, Vraagstuk XXVIII, Wiskundige Opgaven, 10 (1907) 60–61.

[9] P. Turán, On an extremal problem in graph theory, (Hungarian), Mat. Fiz. Lapok, 48 (1941) 436–452.

[10] F. Wu, H.J. Broersma, S. Zhang and B. Li, Properly colored and rainbow C_4's in edge-colored graphs, J. Graph Theory, 105 (2024) 110–135.

[11] C. Xu, X. Hu, W. Wang and S. Zhang, Rainbow cliques in edge-colored graphs, European J. Combin., 54 (2016) 193–200.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

EXTREMAL, ENUMERATIVE AND PROBABILISTIC RESULTS ON ORDERED HYPERGRAPH MATCHINGS

(EXTENDED ABSTRACT)

Michael Anastos* Zhihan Jin† Matthew Kwan‡ Benny Sudakov§

Abstract

An *ordered r-matching* is an r-uniform hypergraph matching equipped with an ordering on its vertices. These objects can be viewed as natural generalisations of *r-dimensional orders*. The theory of ordered 2-matchings is well-developed and has connections and applications to extremal and enumerative combinatorics, probability, and geometry. On the other hand, in the case $r \geq 3$ much less is known, largely due to a lack of powerful bijective tools. Recently, Dudek, Grytczuk and Ruciński made some first steps towards a general theory of ordered r-matchings, and in this paper we substantially improve several of their results and introduce some new directions of study. Many intriguing open questions remain.

1 Introduction

A *matching* is a graph where every vertex is incident to exactly one edge. Given a partition of the vertex set into two equal-size parts V_1, V_2, we say that a matching is *bipartite* if every edge

*Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400, Klosterneuburg, Austria. E-mail: michael.anastos@ist.ac.at. Supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101034413.

†Department of Mathematics, ETH Zürich, Rämistrasse 101, Zürich, 8092, Switzerland. E-mail: zhihan.jin@math.ethz.ch. Supported by SNSF grant 200021-228014.

‡Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400, Klosterneuburg, Austria. E-mail: matthew.kwan@ist.ac.at. Supported by ERC Starting Grant "RANDSTRUCT" No. 101076777 and the European Union.

§Department of Mathematics, ETH Zürich, Rämistrasse 101, Zürich, 8092, Switzerland. E-mail: benjamin.sudakov@math.ethz.ch. Supported by SNSF grant 200021-228014.

is between the two parts. Matchings are basic structures in graph theory. However, when one puts an *ordering* on the set of vertices, matchings will become very interesting objects in their own right. In particular, in an ordered matching there are three different ways that a pair of distinct edges e, e' can "interact with each other", as follows (write $e[1] < e[2]$ for the two vertices of e, write $e'[1] < e'[2]$ for the two vertices of e', and assume without loss of generality that $e[1] < e'[1]$).

- We could have $e[1] < e[2] < e'[1] < e'[2]$ (that is to say, one edge fully comes before the other one). This configuration is called an *alignment*.
- We could have $e[1] < e'[1] < e[2] < e'[2]$ (that is to say, the two edges are interleaved with each other). This configuration is called a *crossing*.
- We could have $e[1] < e'[1] < e'[2] < e[2]$ (that is to say, one of the two edges is "within" the other). This configuration is called a *nesting*.

It is a classical result (first proved by Errera [33]) that the crossing-free matchings on $\{1, \ldots, 2n\}$ are enumerated by the Catalan numbers C_n, and an ingenious bijection (see [79]) shows that there are also exactly C_n nesting-free matchings on $\{1, \ldots, 2n\}$. Note that a matching M on $\{1, \ldots, 2n\}$ is alignment-free if and only if every edge is between $\{1, \ldots, n\}$ and $\{n+1, \ldots, 2n\}$ (i.e., if M is a bipartite matching with these two parts). Such matchings are in correspondence with permutations $\sigma \in \mathcal{S}_n$, and there are therefore $n!$ of them.

Two of the most important parameters of a permutation $\sigma \in \mathcal{S}_n$ are the length $L_{\nearrow}(\sigma)$ of its longest increasing subsequence and the length $L_{\searrow}(\sigma)$ of its longest decreasing subsequence (see for example [72,78] for surveys on the study of these parameters). Two of the highlights in this area are the *Erdős–Szekeres theorem* [32], which says that we always have $L_{\nearrow}(\sigma) \geq \sqrt{n}$ or $L_{\searrow}(\sigma) \geq \sqrt{n}$, and the *Robinson–Schensted–Knuth correspondence* [24, 57, 74] between permutations and Young tableaux, which can be used to enumerate permutations $\sigma \in \mathcal{S}_n$ by their values of $L_{\nearrow}(\sigma)$ and $L_{\searrow}(\sigma)$ (and in particular, to study the behaviour of $L_{\nearrow}(\sigma)$ and $L_{\searrow}(\sigma)$ for a *random* permutation $\sigma \in \mathcal{S}_n$).

It turns out that increasing and decreasing subsequences can be described in the language of configurations in matchings: an increasing subsequence corresponds to a set of edges which are pairwise crossing (which we call a *crossing-clique*), and a decreasing subsequence corresponds to a set of edges which are pairwise nesting (which we call a *nesting-clique*).

Extending the huge body of work on increasing and decreasing subsequences, there has been quite some work (see for example [7, 8, 22, 51, 53, 55, 59, 78]), studying nesting-cliques and crossing-cliques (and alignment-cliques, which have the obvious definition) in general matchings. It turns out that many of the techniques that are effective for permutations have natural analogues for general matchings (in particular, there is a variant of the Robinson–Schensted–Knuth correspondence relating matchings to *oscillating tableaux*; see [78]).

Very recently, Dudek, Grytczuk and Ruciński [27, 28] made some first steps towards extending the theory of ordered matchings to ordered *hypergraph* matchings (which will be defined later). The jump from graphs to hypergraphs seems to introduce a number of serious difficulties, which should perhaps not be surprising: in much the same way that ordered matchings generalise permutations, ordered r-uniform hypergraph matchings generalise $(r-1)$-tuples of permutations (which are sometimes called "r-dimensional orders", as they can be described by sets of points in r-dimensional space). When $r \geq 3$, there is no known analogue of the Robinson–Schensted–Knuth correspondence for r-dimensional orders, and there are a number of longstanding open problems (see for example the survey [17]).

In this paper we substantially extend and improve the results in [27, 28], discovering some surprisingly intricate phenomena and moving towards a more complete theory of ordered hypergraph matchings. A large number of compelling open problems are listed in the full paper.

1.1 Ordered hypergraph matchings: basic notions

We say that a (hyper)graph H is *ordered* if its vertex set $V(H)$ is equipped with a total order (one may think of (hyper)graphs which have vertex set $\{1, \ldots, N\}$ for some N).

Definition 1.1. An ordered hypergraph is said to be an *r-matching* if every edge has exactly r vertices, and every vertex is contained in exactly one edge. An r-matching is said to be *r-partite* if, when we divide the vertex set into r contiguous intervals of equal length, every edge of the matching has exactly one vertex in each interval. The *size* of a matching M is its number of edges.

Note that there are $\frac{(rn)!}{(r!)^n n!}$ different r-matchings on the vertex set $\{1, \ldots, rn\}$. The r-partite r-matchings on $\{1, \ldots, rn\}$ are in correspondence with $(r-1)$-tuples of permutations $(\sigma_1, \ldots, \sigma_{r-1}) \in \mathcal{S}_n^{r-1}$, and there are therefore exactly $(n!)^{r-1}$ of them.

Definition 1.2. An *r-pattern* is an r-matching of size 2 (on the vertex set $\{1, \ldots, 2r\}$, say). We can represent an r-pattern by a string of "A"s and "B"s starting with "A" (where the vertices from one edge are represented with "A", and the vertices of the other edge are represented with "B").

Definition 1.3. For an r-pattern P, an r-matching M is said to be a *P-clique* if every pair of edges of M are order-isomorphic to P. For any r-pattern P and r-matching M, let $L_P(M)$ be the size of the largest P-clique in M. We say that an r-pattern P is *collectable* if there are arbitrarily large P-cliques.

As shown in [28], among the $\frac{1}{2}\binom{2r}{r}$ possible r-patterns, precisely 3^{r-1} of them are collectable, and for the non-collectable r-patterns, the largest possible clique has size 2.

1.2 Ramsey-type questions

Recall that the Erdős–Szekeres theorem says that every permutation $\sigma \in \mathcal{S}_n$ has an increasing or decreasing subsequence of length at least $\sqrt{n}$, and it is tight. One may ask similar questions for r-matchings.

Definition 1.4. Let $L(M) = \max_P L_P(M)$ be the size of the largest clique (of any pattern) in M and let $L_r(n)$ be the minimum value of $L(M)$ among all ordered r-matchings of size n.

When $r = 2$, Huynh, Joos and Wollan [46] showed $L_2(n) = \lceil n^{1/3} \rceil$. For $r \geq 3$, Dudek, Grytczuk and Ruciński [28] proved $c_r n^{1/3^{r-1}} \leq L_r(n) \leq n^{1/(2^{r-2}+2)}$ for some $c_r > 0$ depending on r. Our first result is to narrow down the gap and provide evidences that $L_r(n) = \Theta_r(n^{1/(2^r-1)})$.

Theorem 1.5. *For $r \geq 2$, we have $\frac{1}{r-1} \cdot n^{1/((r+1)2^{r-2})} \leq L_r(n) \leq \lceil n^{1/(2^r-1)} \rceil$. In addition, we have $L_3(n) = \Theta(n^{1/7})$ and $L_4(n) = \Theta(n^{1/15})$.*

After our work, it was shown by Sauermann and Zakharov [73] that $L_r(n) = \Theta_r(n^{1/(2^r-1)})$.

1.3 Random matchings

One of the most notorious problems in the theory of permutations is the *Ulam–Hammersley problem*, to describe the distribution of the longest increasing permutation $L_{\nearrow}(\sigma)$ in a random permutation $\sigma \in \mathcal{S}_n$. This problem was famously resolved by Baik, Deift and Johansson [6], but one of the most important milestones along the way was a theorem of Logan and Shepp [61] and Vershik and Kerov [86] (see also the alternative proofs [2, 44, 48, 75]), establishing that the expected value of $L_{\nearrow}(\sigma)$ is asymptotically $2\sqrt{n}$.

These results were extended to random matchings by Baik and Rains [8] (see also [7, 78]). Specifically, they proved that in a random matching M on $\{1, \ldots, 2n\}$, the expected values of $L_{\text{crossing}}(M)$ and $L_{\text{nesting}}(M)$ are both $(\sqrt{2} + o(1))\sqrt{n}$ (as part of a tour-de-force where they found the asymptotic distribution of these quantities).

In [28], Dudek, Grytczuk and Ruciński proved that there are constants $c'_r, c''_r \geq 0$ such that for any collectable r-pattern P, and a random r-matching M, we have $c'_r n^{1/r} \leq L_P(M) \leq c''_r n^{1/r}$ whp. They also conjectured that $L_P(M)/n^{1/r}$ converges in probability, to a constant that only depends on r.

Our second result is to show that for a collectable pattern P and a random matching M, the random variable $L_P(M)/n^{1/r}$ does converge to a limit. This limit may depend on P, but only through the *type* of P, which is defined in the full paper.

Theorem 1.6. *Fix an r-pattern P, and let M be a random r-matching of size n. Then we have*

$$\frac{L_P(M)}{n^{1/r}} \xrightarrow{p} b_P$$

for some $b_P > 0$ depending only on the type of P.

Actually determining the constants b_P seems to be a hard problem. However, we are able to prove some bounds on the b_P in some special cases. In particular, we show that the conjecture of Dudek, Grytczuk and Ruciński is false for all $r \geq 2$. (Recall that the Γ function is the analytic continuation of the factorial function).

Proposition 1.7. *Let the constants b_P be as in Theorem 1.6.*

1. *If $P = \mathrm{AA}\ldots\mathrm{ABB}\ldots\mathrm{B}$, then $b_P = 1/\Gamma((r+1)/r)$.*
2. *If P is r-partite, then $b_P = c_r(r!)^{1/r}/r > 1/\Gamma((r+1)/r)$, where c_r is the constant from the Bollobás–Winkler theorem [12].*

1.4 Enumeration

For an r-pattern P, let $N_P(n)$ denote the number of ordered r-matchings on the vertex set $\{1,\ldots,rn\}$ which are P-free (i.e., no two edges form P). First notice that if P is not r-partite then every r-partite matching is P-free, which indicates $N_P(n) = e^{O_r(n)}n^{(r-1)n}$.

The case where P is r-partite is more delicate (and the value of $N_P(n)$ is quite different), but we are able to obtain estimates of similar quality, as follows.

Theorem 1.8. *Fix a constant $r \in \mathbb{N}$. If P is r-partite then*

$$N_P(n) = e^{O_r(n)}n^{(r-1-1/(r-1))n}.$$

The lower bound in Theorem 1.8 is a direct consequence of a result due to Brightwell [16] on linear extensions of r-dimensional posets, but the upper bound is new (it is obtained by a general connection to extremal ordered hypergraph theory).

We are also interested in enumerating ordered matchings by the size of their largest P-clique. We write $N_{P,m}(n)$ for the number of ordered r-matchings M on the vertex set $\{1,\ldots,rn\}$ which satisfy $L_P(M) < m$ (i.e., they do not contain a P-clique of size m). Note that if P is not r-partite, then the same considerations above show that $N_{P,m}(n) = e^{O_r(n)}n^{(r-1)n}$ for all m. However, if P is r-partite, we get a significant dependence on m, as per the following generalisation of Theorem 1.8 (note that Theorem 1.8 corresponds to the case $m = 2$).

Theorem 1.9. *Fix a constant $r \in \mathbb{N}$. If P is r-partite then for any $2 \le m \le n^{1/r}$ we have*

$$N_{P,m}(n) = e^{O_r(n)}(m-1)^{(r/(r-1))n}n^{(r-1-1/(r-1))n}.$$

We prove the upper bound in Theorem 1.9 via a general lemma which estimates $N_{P,m}(n)$ in terms of a certain *extremal* parameter (namely, the maximum number of edges in an ordered r-uniform hypergraph on $\lceil n/2 \rceil$ vertices with $L_P(M) < m$). This type of reduction goes back to Alon and Friedgut [3] (see also [34, 54]). We also acquire a few new results on the extremal numbers of certain P-cliques; see the full paper for the details.

We remark that for the lower bound in Theorem 1.9, we obtain a new estimate on the number of $(r-1)$-tuples of length-n permutations which have no common increasing subsequence of length m, by "reverse-engineering" the techniques in the upper bound.

References

[1] Hüseyin Acan. On a uniformly random chord diagram and its intersection graph. *Discrete Math.*, 340(8):1967–1985, 2017.

[2] D. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields*, 103(2):199–213, 1995.

[3] Noga Alon and Ehud Friedgut. On the number of permutations avoiding a given pattern. *J. Combin. Theory Ser. A*, 89(1):133–140, 2000.

[4] Noga Alon and Ehud Friedgut. On the number of permutations avoiding a given pattern. *J. Combin. Theory Ser. A*, 89(1):133–140, 2000.

[5] Boris Aronov, Vida Dujmović, Pat Morin, Aurélien Ooms, and Luís Fernando Schultz Xavier da Silveira. More Turán-type theorems for triangles in convex point sets. *Electron. J. Combin.*, 26(1):Paper No. 1.8, 26, 2019.

[6] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[7] Jinho Baik and Robert Jenkins. Limiting distribution of maximal crossing and nesting of Poissonized random matchings. *Ann. Probab.*, 41(6):4359–4406, 2013.

[8] Jinho Baik and Eric M. Rains. The asymptotics of monotone subsequences of involutions. *Duke Math. J.*, 109(2):205–281, 2001.

[9] József Balogh, Béla Bollobás, and Robert Morris. Hereditary properties of partitions, ordered graphs and ordered hypergraphs. *European J. Combin.*, 27(8):1263–1281, 2006.

[10] Béla Bollobás and Graham Brightwell. The height of a random partial order: concentration of measure. *Ann. Appl. Probab.*, 2(4):1009–1018, 1992.

[11] Béla Bollobás and Graham Brightwell. The height of a random partial order: concentration of measure. *Ann. Appl. Probab.*, 2(4):1009–1018, 1992.

[12] Béla Bollobás and Peter Winkler. The longest chain among random points in Euclidean space. *Proc. Amer. Math. Soc.*, 103(2):347–353, 1988.

[13] P. Braß, G. Rote, and K. J. Swanepoel. Triangles of extremal area or perimeter in a finite planar point set. *Discrete Comput. Geom.*, 26(1):51–58, 2001.

[14] Peter Braß. Turán-type extremal problems for convex geometric hypergraphs. In *Towards a theory of geometric graphs*, volume 342 of *Contemp. Math.*, pages 25–33. Amer. Math. Soc., Providence, RI, 2004.

[15] Peter Braß, Gyula Károlyi, and Pavel Valtr. A Turán-type extremal theory of convex geometric graphs. In *Discrete and computational geometry*, volume 25 of *Algorithms Combin.*, pages 275–300. Springer, Berlin, 2003.

[16] Graham Brightwell. Random k-dimensional orders: width and number of linear extensions. *Order*, 9(4):333–342, 1992.

[17] Graham Brightwell. Models of random partial orders. In *Surveys in combinatorics, 1993 (Keele)*, volume 187 of *London Math. Soc. Lecture Note Ser.*, pages 53–83. Cambridge Univ. Press, Cambridge, 1993.

[18] Matija Bucić, Benny Sudakov, and Tuan Tran. Erdős-Szekeres theorem for multidimensional arrays. *J. Eur. Math. Soc. (JEMS)*, 25:2927–2947, 2023.

[19] Boris Bukh and Jiří Matoušek. Erdős–Szekeres-type statements: Ramsey function and decidability in dimension 1. *Duke Math. J.*, 163(12):2243 – 2270, 2014.

[20] H. Burkill and L. Mirsky. Monotonicity. *J. Math. Anal. Appl.*, 41:391–410, 1973.

[21] Vasilis Capoyleas and János Pach. A Turán-type theorem on chords of a convex polygon. *J. Combin. Theory Ser. B*, 56(1):9–15, 1992.

[22] William Y. C. Chen, Eva Y. P. Deng, Rosena R. X. Du, Richard P. Stanley, and Catherine H. Yan. Crossings and nestings of matchings and partitions. *Trans. Amer. Math. Soc.*, 359(4):1555–1575, 2007.

[23] Josef Cibulka and Jan Kynčl. Better upper bounds on the Füredi-Hajnal limits of permutations. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2280–2293. SIAM, Philadelphia, PA, 2017.

[24] G. de B. Robinson. On the Representations of the Symmetric Group. *Amer. J. Math.*, 60(3):745–760, 1938.

[25] Anna de Mier. k-noncrossing and k-nonnesting graphs and fillings of Ferrers diagrams. *Combinatorica*, 27(6):699–720, 2007.

[26] Jean-Dominique Deuschel and Ofer Zeitouni. On increasing subsequences of I.I.D. samples. *Combin. Probab. Comput.*, 8(3):247–263, 1999.

[27] Andrzej Dudek, Jarosł aw Grytczuk, and Andrzej Ruciński. Ordered unavoidable substructures in matchings and random matchings. *Electron. J. Combin.*, 31(2):Paper No. 2.15, 27, 2024.

[28] Andrzej Dudek, Jarosł aw Grytczuk, and Andrzej Ruciński. Erdős-Szekeres type theorems for ordered uniform matchings. *J. Combin. Theory Ser. B*, 170:225–259, 2025.

[29] Andrzej Dudek, Jarosław Grytczuk, and Andrzej Ruciński. *On weak twins and up-and-down subpermutations*, page 187–202. De Gruyter, April 2022.

[30] Vida Dujmović and David R. Wood. On linear layouts of graphs. *Discrete Math. Theor. Comput. Sci.*, 6(2):339–357, 2004.

[31] Paul Erdős and Daniel J. Kleitman. On coloring graphs to maximize the proportion of multicolored k-edges. *J. Combinatorial Theory*, 5:164–169, 1968.

[32] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[33] Alfred Errera. Analysis situs: un problème d'énumération. *Mémoires de la Classe des sciences. Académie royale de Belgique. Collection 8*, 11(6):3–26, 1931.

[34] Jacob Fox. Stanley-Wilf limits are typically exponential. *arXiv preprint arXiv:1310.8378*, 2013.

[35] Jacob Fox, János Pach, Benny Sudakov, and Andrew Suk. Erdős-Szekeres-type theorems for monotone paths and convex bodies. *Proc. Lond. Math. Soc. (3)*, 105(5):953–982, 2012.

[36] Peter Frankl, Andreas Holmsen, and Andrey Kupavskii. Intersection theorems for triangles. *Discrete & Computational Geometry*, pages 1–10, 2020.

[37] Zoltán Füredi, Tao Jiang, Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Tight paths in convex geometric hypergraphs. *Adv. Comb.*, pages Paper No. 1, 14, 2020.

[38] Zoltán Füredi, Tao Jiang, Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Extremal problems for convex geometric hypergraphs and ordered hypergraphs. *Canad. J. Math.*, 73(6):1648–1666, 2021.

[39] Zoltán Füredi, Tao Jiang, Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Partitioning ordered hypergraphs. *J. Combin. Theory Ser. A*, 177:Paper No. 105300, 18, 2021.

[40] Zoltán Füredi, Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Ordered and convex geometric trees with linear extremal function. *Discrete & Computational Geometry*, 64(2):324–338, 2020.

[41] Zoltán Füredi, Dhruv Mubayi, Jason O'Neill, and Jacques Verstraëte. Extremal problems for pairs of triangles. *J. Combin. Theory Ser. B*, 155:83–110, 2022.

[42] António Girão, Gal Kronenberg, and Alex Scott. A multidimensional Ramsey theorem. *arXiv preprint arXiv:2210.09227*.

[43] W. T. Gowers and J. Long. The length of an s-increasing sequence of r-tuples. *Combin. Probab. Comput.*, 30(5):686–721, 2021.

[44] Piet Groeneboom. Ulam's problem and Hammersley's process. *Ann. Probab.*, 29(2):683–690, 2001.

[45] Benjamin Gunby and Dömötör Pálvölgyi. Asymptotics of pattern avoidance in the klazar set partition and permutation-tuple settings. *European Journal of Combinatorics*, 82:102992, 2019.

[46] Tony Huynh, Felix Joos, and Paul Wollan. A unified Erdős-Pósa theorem for constrained cycles. *Combinatorica*, 39(1):91–133, 2019.

[47] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[48] Kurt Johansson. The longest increasing subsequence in a random permutation and a unitary random matrix model. *Math. Res. Lett.*, 5(1-2):63–82, 1998.

[49] Jakob Jonsson. Generalized triangulations and diagonal-free subsets of stack polyominoes. *Journal of Combinatorial Theory, Series A*, 112(1):117–142, 2005.

[50] Jakob Jonsson and Volkmar Welker. A spherical initial ideal for Pfaffians. *Illinois J. Math.*, 51(4):1397–1407, 2007.

[51] Joyce Justicz, Edward R. Scheinerman, and Peter Winkler. Random intervals. *Amer. Math. Monthly*, 97(10):881–889, 1990.

[52] Kenneth Kalmanson. On a theorem of Erdős and Szekeres. *J. Combinatorial Theory Ser. A*, 15:343–346, 1973.

[53] Anisse Kasraoui and Jiang Zeng. Distribution of crossings, nestings and alignments of two edges in matchings and partitions. *Electron. J. Combin.*, 13(1):Research Paper 33, 12, 2006.

[54] Martin Klazar. The Füredi-Hajnal conjecture implies the Stanley-Wilf conjecture. In *Formal power series and algebraic combinatorics (Moscow, 2000)*, pages 250–255. Springer, Berlin, 2000.

[55] Martin Klazar. On identities concerning the numbers of crossings and nestings of two edges in matchings. *SIAM J. Discrete Math.*, 20(4):960–976, 2006.

[56] Martin Klazar and Adam Marcus. Extensions of the linear bound in the Füredi-Hajnal conjecture. *Adv. in Appl. Math.*, 38(2):258–266, 2007.

[57] Donald E. Knuth. Permutations, matrices, and generalized Young tableaux. *Pacific J. Math.*, 34:709–727, 1970.

[58] Dániel Korándi, Gábor Tardos, István Tomon, and Craig Weidert. On the turán number of ordered forests. *Journal of Combinatorial Theory, Series A*, 165:32–43, 2019.

[59] C. Krattenthaler. Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes. *Adv. in Appl. Math.*, 37(3):404–431, 2006.

[60] Nathan Linial and Michael Simkin. Monotone subsequences in high-dimensional permutations. *Combin. Probab. Comput.*, 27(1):69–83, 2018.

[61] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Advances in Math.*, 26(2):206–222, 1977.

[62] L. Lovász and M. D. Plummer. *Matching theory*, volume 121 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam; North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29.

[63] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Combin. Theory Ser. A*, 107(1):153–160, 2004.

[64] Adam W Marcus. *New combinatorial techniques for nonlinear orders*. Georgia Institute of Technology, 2008.

[65] Jon McCammond. Noncrossing partitions in surprising locations. *Amer. Math. Monthly*, 113(7):598–610, 2006.

[66] Colin McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

[67] L. Mirsky. A dual of Dilworth's decomposition theorem. *Amer. Math. Monthly*, 78:876–877, 1971.

[68] László Ozsvárt. Counting ordered graphs that avoid certain subgraphs. *Discrete Mathematics*, 339(7):1871–1877, 2016.

[69] János Pach and Rom Pinchasi. How many unit equilateral triangles can be generated by N points in convex position? *Amer. Math. Monthly*, 110(5):400–406, 2003.

[70] János Pach and Gábor Tardos. Forbidden paths and cycles in ordered graphs and matrices. *Israel Journal of Mathematics*, 155(1):359–380, 2006.

[71] Sriram Venkata Pemmaraju. *Exploring the powers of stacks and queues via graph layouts*. ProQuest LLC, Ann Arbor, MI, 1992. Thesis (Ph.D.)–Virginia Polytechnic Institute and State University.

[72] Dan Romik. *The surprising mathematics of longest increasing subsequences*, volume 4 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, New York, 2015.

[73] Lisa Sauermann and Dmitrii Zakharov. A sharp ramsey theorem for ordered hypergraph matchings. *arXiv preprint arXiv:2309.04813*, 2023.

[74] C. Schensted. Longest increasing and decreasing subsequences. *Canadian J. Math.*, 13:179–191, 1961.

[75] Timo Seppäläinen. Large deviations for increasing sequences on the plane. *Probab. Theory Related Fields*, 112(2):221–244, 1998.

[76] Timo Seppäläinen. A growth model in multiple dimensions and the height of a random partial order. In *Asymptotics: particles, processes and inverse problems*, volume 55 of *IMS Lecture Notes Monogr. Ser.*, pages 204–233. Inst. Math. Statist., Beachwood, OH, 2007.

[77] Rodica Simion. Noncrossing partitions. *Discrete Math.*, 217(1-3):367–409, 2000.

[78] Richard P. Stanley. Increasing and decreasing subsequences and their variants. In *International Congress of Mathematicians. Vol. I*, pages 545–579. Eur. Math. Soc., Zürich, 2007.

[79] Richard P. Stanley. Catalan addendum (version of 25 may 2013). `http://www-math.mit.edu/\%7Erstan/ec/catadd.pdf.`, 2013.

[80] J. Michael Steele. Limit properties of random variables associated with a partial ordering of R^d. *Ann. Probability*, 5(3):395–403, 1977.

[81] Sheila Sundaram. The Cauchy identity for $\mathrm{Sp}(2n)$. *J. Combin. Theory Ser. A*, 53(2):209–238, 1990.

[82] Tibor Szabó and Gábor Tardos. A multidimensional generalization of the Erdős-Szekeres lemma on monotone subsequences. *Combin. Probab. Comput.*, 10(6):557–565, 2001.

[83] Michel Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, 81(81):73–205, 1995.

[84] Gábor Tardos. Extremal theory of ordered graphs. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 3235–3243. World Scientific, 2018.

[85] Gábor Tardos. Extremal theory of vertex or edge ordered graphs. In *Surveys in combinatorics 2019*, volume 456 of *London Math. Soc. Lecture Note Ser.*, pages 221–236. Cambridge Univ. Press, Cambridge, 2019.

[86] A. M. Veršik and S. V. Kerov. Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Dokl. Akad. Nauk SSSR*, 233(6):1024–1027, 1977.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Homomorphism-homogeneous Graphs from Finite Reflexive Graphs

(Extended abstract)

Andrés Aranda* David Bradley-Williams† David Hartman‡

Abstract

Despite years of study, few examples and constructions of HH-homogeneous graphs are described in detail in the literature. We introduce a new notion of homogeneity for reflexive graphs (multimorphism-homogeneity) and prove that new homomorphism-homogeneous graphs can be produced from finite multimorphism-homogeneous reflexive graphs.

1 Introduction

A countable graph G is *homomorphism-homogeneous*, or HH-homogeneous, if every partial homomorphism with finite domain is a restriction of some endomorphism of G. The concept was introduced in [3] as a variation of homogeneity, where every partial *isomorphism* of G with finite domain is a restriction of some *automorphism* of G. The first examples of HH-homogeneous graphs are those graphs in which every finite subset F has a *cone*, that is, a vertex c such that $c \sim d$ for all $d \in F$. Graphs with that property are homomorphic images of the Rado graph. The first examples of HH-homogeneous graphs which are not images of the Rado graph were produced by M. Rusinov and P. Schweitzer in [5]. Their examples seem to have what we call a "finite soul", meaning that they can be represented as blown-up versions of the intersection graphs of finite set systems. This property pops up again in the proof

*Independent scholar, Prague, Czech Republic. E-mail: andres.aranda@gmail.com.

†Computer Science Institute of Charles University, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, Prague 1, 118 00, Czech Republic and Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 110 00 Prague 1, Czech Republic. E-mail: williams@math.cas.cz. Supported by Project 24-12591M of the Czech Science Foundation (GAČR).

‡Computer Science Institute of Charles University, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, Prague 1, 118 00, Czech Republic and Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 271/2, 182 07 Prague, Czech Republic E-mail: hartman@iuuk.mff.cuni.cz. Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Synergy Grant DYNASNET, grant agreement no. 810115)

that if G is connected, HH-homogeneous, and not an image of the Rado graph, then it has finite independence number: the authors of [2] note that any HH-homogeneous graph in which a great independent set has no cone interprets a graph of Rusinov-Schweitzer type (see the paragraph following claims 13 and 14 in [2]).

In this abstract, we prove that certain finite reflexive graphs produce countably infinite HH-homogeneous graphs that are new to the literature. We produce these by blowing up intersection graphs of finite set systems in a way that is strongly reminiscent of the Rusinov-Schweitzer construction, of [5], yet produces a distinctly different family of homomorphism-homogeneous graphs. Our first candidates were homomorphism-homogeneous reflexive graphs, but we soon discovered that homomorphism-homogeneity is not enough to guarantee the HH-homogeneity of the resulting infinite graph; thus, we introduce the concept of *multimorphism* for reflexive graphs and a stronger form of homogeneity, *multimorphism-homogeneity* which is necessary and sufficient to produce HH-homogeneous graphs.

2 Finite reflexive graphs and HH-homogeneous graphs

Notation 2.1. Given any set X and $\mathcal{S} \subseteq \mathcal{P}(X)$,

1. $\mathrm{Int}(\mathcal{S})$ represents the graph on $\mathcal{S}$ with edge set $\big\{(T,T') \in \mathcal{S}^2 : T \neq T' \wedge T \cap T' \neq \varnothing\big\}$.
2. $\mathrm{Int}^{\circ}(\mathcal{S})$ denotes the reflexive graph on $\mathcal{S}$ with edge set $\big\{(T,T') \in \mathcal{S}^2 : T \cap T' \neq \varnothing\big\}$.
3. $\mathcal{G}(\mathcal{S})$ denotes $\mathrm{Int}(\mathcal{S})[K_\omega]$, the graph with vertex set $\mathcal{S} \times \omega$ and edge set $\{\{(u,n),(v,m)\} : u \sim v \wedge (u,n) \neq (v,m)\}$.
4. Given $T \in \mathcal{S}$, the symbol K_T denotes the clique $\{T\} \times K_\omega \subseteq \mathcal{G}(\mathcal{S})$.

We use $\sim$ to denote adjacency in graphs and reflexive graphs. Finally, given a set X and a natural number n, $[X]^n$ denotes the set $\{Q \in \mathcal{P}(X) : |Q| = n\}$.

As we will see, strong homogeneity properties in a finite reflexive graph represented as $\mathrm{Int}^{\circ}(\mathcal{S})$ translate into HH-homogeneity of $\mathcal{G}(\mathcal{S})$.

Definition 1. *Let X be a finite set. A* hitting set *for $\mathcal{Y} \subseteq \mathcal{P}(X)$ is $Z \subseteq X$ such that $Z \cap Y \neq \varnothing$ for all $Y \in \mathcal{Y}$. The* hitting number *of $\mathcal{Y}$, denoted by* $\mathrm{hit}(\mathcal{Y})$*, is the cardinality of a hitting set of minimal size.*

Definition 2. *Let X be a finite set and $k \leq |X|$. A k*-capped subset *of $\mathcal{P}(X)$ is any $\mathcal{Q} \subset \mathcal{P}(X)$ such that $[X]^1 \cup [X]^k \subseteq \mathcal{Q}$ and for all $Y \in \mathcal{Q}$ we have $1 \leq |Y| \leq k$.*

Definition 3. *Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive graphs. A (partial)* multimorphism *is a function $m \colon \mathcal{X}' \to \mathcal{P}(\mathcal{Y})$ for some $\mathcal{X}' \subseteq \mathcal{X}$ such that $m(u) \cup m(v)$ is a clique in $\mathcal{Y}$ whenever $u \sim v$.*

Definition 4. *Let $m \colon \mathcal{T} \to \mathcal{P}(\mathcal{T}')$ be a multimorphism. A function $f \colon \mathcal{T} \to \mathcal{T}'$ satisfying $f(t) \in m(t)$ for all $t \in \mathcal{T}$ is called a* selection *of m.*

Lemma 5. *Let X be a set, $\mathcal{S} \subseteq \mathcal{P}(X)$ and $\mathcal{S}' \subseteq \mathcal{S}$. Suppose $m \colon \mathcal{S}' \to \mathcal{P}(\mathcal{S})$ is a multimorphism and $u, v \in \mathcal{S}'$ satisfy $u \sim v$ in $\mathrm{Int}^{\circ}(\mathcal{S})$. If g_1, g_2 are selections of m, then $g_1(u) \sim g_2(v)$. In particular, each selection of m is a homomorphism.*

Notation 2.2. We use π to denote the function $\mathcal{G}(\mathcal{S}) \to \mathcal{S}$ assigning to each v the unique $\pi(v) \in \mathcal{S}$ such that $v \in \{\pi(v)\} \times \omega$. Since $\mathcal{S}$ is the vertex set of $\mathrm{Int}^\circ(\mathcal{S})$, we can think of π as a function $\mathcal{G}(\mathcal{S}) \to \mathrm{Int}^\circ(\mathcal{S})$. Under this interpretation, π is a homomorphism.

Lemma 6. *Let X be a set and $\mathcal{S} \subseteq \mathcal{P}(X)$. Suppose that $h\colon \mathcal{Y} \to \mathcal{Z}$ is a partial homomorphism in $\mathcal{G}(\mathcal{S})$. Then the function $\underline{h}\colon \pi(\mathcal{Y}) \to \mathcal{P}(\mathrm{Int}^\circ(\mathcal{S}))$ given by $\underline{h}(v) = \{\pi(h(y)) : y \in \mathcal{Y} \wedge \pi(y) = v\}$ is a (partial) multimorphism in $\mathrm{Int}^\circ(\mathcal{S})$.*

Proof. Suppose that $u \sim v$ for some $u, v \in \pi(\mathcal{Y})$. This happens if and only if $u \cap v \neq \varnothing$, which means that for any distinct $\hat{u}, \hat{v}$ such that $\pi(\hat{u}) = u$ and $\pi(\hat{v}) = v$, we have $\hat{u} \sim \hat{v}$, which in turn implies $h(\hat{u}) \sim h(\hat{v})$, or equivalently $\pi(h(\hat{u})) \sim \pi(h(\hat{v}))$. Therefore, $\underline{h}(u) \cap \underline{h}(v)$ is a clique and $\underline{h}$ is a multimorphism. □

Notation 2.3. If h is a homomorphism in $\mathcal{G}(\mathcal{S})$, then $\underline{h}$ denotes the multimorphism from Lemma 6.

Lemma 7. *Let X be a finite set and $\mathcal{S} \subseteq \mathcal{P}(X)$. If $m\colon \mathcal{Y} \to \mathcal{P}(\mathrm{Int}^\circ(\mathcal{S}))$ is a partial multimorphism in $\mathrm{Int}^\circ(\mathcal{S})$, then there exists a homomorphism $\hat{m}$ in $\mathcal{G}(\mathcal{S})$ such that $\underline{\hat{m}} = m$.*

Proof. Let $\mathcal{Y}'$ denote $\bigcup\{m(y) : y \in \mathcal{Y}\}$. Define $c\colon \mathcal{Y} \to \omega$ by $c(y) = |m(y)|$. Now for each $y \in \mathcal{Y}$ choose an arbitrary $F_y \subset \{(y, n) : n \in \omega\} \subset \mathrm{Int}^\circ(\mathcal{S})$ such that $|F(y)| = c(y)$, and for each $u \in \mathcal{Y}'$ select a transversal G_u of $\pi^{-1}(m(u))$. Since the preimage under π of each vertex is infinite, we can always arrange for $G_u \cap G_{u'}$ to be disjoint if $u \neq u'$.

Let $\hat{m}_y$ be any bijection $F_y \to G_{m(y)}$. Define $\hat{m}$ as $\bigcup\{\hat{m}_y : y \in \mathcal{Y}\}$. By construction, $\underline{\hat{m}} = m$. It is not hard to prove that $\hat{m}$ is a homomorphism. □

Definition 8. *Let X be a finite set and $\mathcal{S} \subseteq \mathcal{P}(X)$. We say that $\mathcal{S}$ is k-*thin *if for all induced subgraphs $\mathcal{Y}$ of $\mathrm{Int}^\circ(\mathcal{S})$ and multimorphisms $m\colon \mathcal{Y} \to \mathcal{P}(\mathcal{S})$, if $\mathrm{hit}(\mathcal{Y}) \leq k$, then $\mathrm{hit}(\{v \in \mathcal{S} : \exists u \in \mathcal{Y}(v \in m(u))\}) \leq k$.*

For example, if X is finite and $k < |X| < 2k$, then $[X]^1 \cup [X]^k$ is k-thin.

We remind the reader that a set of vertices P has a *cone* in a graph or reflexive graph G if there exists $c \in G$ with $c \sim p$ for all $p \in P$. If the ambient graph G is clear from the context, we often don't mention it.

Lemma 9. *Let X be a finite set and suppose that $\mathcal{S} \subseteq \mathcal{P}(X)$ is k-capped and k-thin for some $k \leq |X|$. Then any finite $\mathcal{Y} \subset \mathcal{G}(\mathcal{S})$ has a cone if and only if $\mathrm{hit}(\pi(\mathcal{Y})) \leq k$.*

Proof. Suppose first that $\mathcal{Y}$ has a cone in G. Then it has a cone c in some K_T where $T \in [X]^k$, and $\pi(c)$ is a hitting set of size k for each of $\pi(y)$ with $y \in \mathcal{Y}$, so $\mathrm{hit}(\pi(\mathcal{Y})) \leq k$.

On the other hand, if $\mathrm{hit}(\pi(\mathcal{Y})) \leq k$ then let P be a hitting set for $\pi(\mathcal{Y})$ of size k. Any element $d \in G \setminus \mathcal{Y}$ with $\pi(d) = P$ is a cone in G over $\mathcal{Y}$. □

Lemma 10. *Let X be a finite set and $2 \leq k \leq |X|$. If $\mathcal{S} \subseteq \mathcal{P}(X)$ is k-thin and k-capped, then $|X| \leq 2k$.*

Proof. Suppose for a contradiction that $\mathcal{S}$ is k-capped and k-thin, but $2k < |X|$.

Without loss, suppose $X = \{1, \ldots, n\}$. Consider $T \coloneqq \{1, \ldots, k\}$, $T' \coloneqq \{k+1, \ldots, 2k\}$, and $P \coloneqq \{W \in [X]^k : W \notin \{T, T'\} \wedge W \subset T \cup T'\}$.

Claim 11. $\mathrm{hit}(P \cup \{T, T'\}) > k$.

Proof. It suffices to prove that no k-element subset of $\{1,\dots,2k\}$ hits $P \cup \{T,T'\}$. If $C \in [\{1,\dots,2k\}]^k$, then $\{1,\dots,2k\} \setminus C$ is an element of $P \cup \{T,T'\}$ with $C \cap C' = \varnothing$. □

Claim 12. *There exists a bijective homomorphism $h\colon P \cup \{\{1\},\{2\}\} \to P \cup \{T,T'\}$.*

Proof. Let h be the function that fixes P pointwise and maps $\{1\} \mapsto T$ and $\{2\} \mapsto T'$. Note that $\{1\},\{2\}$ and T,T' do not form edges in $\mathrm{Int}^\circ(\mathcal{S})$. Since T and T' are connected by edges to all vertices in P (this follows from $2k < |X|$ and our choice for P), h is a homomorphism. □

Note that T hits $P \cup \{\{1\},\{2\}\}$, so $\mathrm{hit}(P \cup \{\{1\},\{2\}\}) \le k$. Since homomorphisms can be interpreted as multimorphisms, h contradicts k-thinness and the result follows. □

Lemma 13. *A countable graph G that satisfies the following two conditions for all finite induced subgraphs A and local homomorphisms h is HH-homogeneous.*

1. *If A has a cone in G, then $h[A]$ has a cone in G.*
2. *If c is a cone over A in G, then there exists $d \in G$ which is cone over $A \cup \{c\}$.*

Proof. Suppose that $h\colon A \to B$ is a surjective homomorphism with finite domain. Enumerate $G = \{g_i : i \in \omega\}$ so that $A = \{g_0,\dots,g_{n-1}\}$ and let $h_0 \coloneqq h$. Suppose that he have succeeded in extending h_0 to a homomorphism h_m with domain $A_m = \{g_0,\dots,g_{n+m-1}\}$. Now we wish to extend h_m to a homomorphism with domain A_{m+1}. There are two possibilities.

First, if $h_m[A_m]$ has a cone c, then we define h_{m+1} as $h_m \cup \{(g_{n+m},c)\}$. And if $h_m[A_m]$ has no cone in G, then let A_m^+ be $N(g_{n+m}) \cap A_m$. If $A_m^+ = \varnothing$, then we can define h_{m+1} as $h_m \cup \{(g_{n+m},q)\}$, where q is any vertex. Finally, if $A_m^+ \neq \varnothing$, then we can use properties 1 and 2 to find a cone d of $h_m[A_m^+]$ not in $h_m[A_m]$ and define h_{m+1} as $h_m \cup \{(g_{n+m},d)\}$. Now let H be $\bigcup\{h_i : i \in \omega\}$. It is easy to verify that H is an endomorphism of G. □

The converse of Lemma 13 holds for countable graphs with vertices of infinite degree, as was proved in [3].

Proposition 14. *Let X be a finite set and suppose that $\mathcal{S} \subseteq \mathcal{P}(X)$ is k-capped and k-thin. Then $\mathcal{G}(\mathcal{S})$ is HH-homogeneous.*

Proof. We show that the two properties from Lemma 13 are satisfied by $\mathcal{G}(\mathcal{S})$.

Suppose that $\mathcal{Y}$ and $\mathcal{Y}'$ are finite induced subgraphs of $\mathcal{G}(\mathcal{S})$, Y has a cone in $\mathcal{G}(\mathcal{S})$, and there exists a surjective homomorphism $h\colon \mathcal{Y} \to \mathcal{Y}'$. By Lemma 9, $\mathrm{hit}(\pi(\mathcal{Y})) \le k$. By k-thinness, $\mathrm{hit}(\pi(\mathcal{Y}')) \le k$, and, again by Lemma 9, $\mathcal{Y}'$ has a cone in $\mathcal{G}(\mathcal{S})$

Next, suppose that c is a cone over $\mathcal{Y}$. Let d be any vertex with $\pi(c) = \pi(d)$, $c \neq d$, and $d \notin \mathcal{Y}$. Since c and d have the same image under π, $c \sim y$ holds if and only if $d \sim y$ for all $y \in \mathcal{Y}$. Moreover, $\pi(c) = \pi(d)$ and $c \neq d$ imply $c \sim d$, so d is a cone over $\mathcal{Y} \cup \{c\}$. □

Proposition 15. *Let X be a finite set and suppose that $\mathcal{S} \subseteq \mathcal{P}(X)$ is k-capped. If $\mathcal{G}(\mathcal{S})$ is HH-homogeneous, then $\mathcal{S}$ is k-thin.*

Proof. We prove the contrapositive. Suppose that $\mathcal{S}$ is not k-thin. Then there exist $\mathcal{Y},\mathcal{Y}' \subseteq \mathcal{S}$ and a surjective multimorphism $h\colon \mathcal{Y} \to \mathcal{Y}'$ such that $\mathrm{hit}(\mathcal{Y}) \le k$ and $\mathrm{hit}(\mathcal{Y}') > k$. By Lemma 7, we can produce from h a local homomorphism $\hat{h}$ such that $\underline{\hat{h}} = h$. Since $\mathrm{hit}(\mathcal{Y}')$ is greater than k, we know that the image of $\hat{h}$ does not have a cone, but its domain does as its image under π is $\mathcal{Y}$. Thus $\mathcal{G}(\mathcal{S})$ cannot be HH-homogeneous. □

Definition 16. *A reflexive graph $\mathcal{T}$ is* multimorphism-homogeneous *if for every finite $\mathcal{A} \subseteq \mathcal{T}$ and every every multimorphism $m\colon \mathcal{A} \to \mathcal{P}(\mathcal{T})$ there exists a multimorphism $\mathcal{T} \to \mathcal{P}(\mathcal{T})$ which extends m.*

Example 1. If $\mathcal{S} \subset \mathcal{P}(5)$ is $[5]^1 \cup [5]^4 \cup \{\{0,4\},\{1,4\}\}$, then $\mathrm{Int}^\circ(\mathcal{S})$ is multimorphism-homogeneous.

Example 2. The pentagon with loops $\mathrm{Int}^\circ(\{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,5\}\})$ is known to be HH-homogeneous (see [4]) but is not multimorphism-homogeneous. For example, the partial multimorphism m mapping $\{1,2\} \mapsto \{\{1,5\},\{4,5\}\}$ and $\{3,4\} \mapsto \{\{2,3\}\}$ cannot be extended to a multimorphism defined on all five vertices.

Theorem 17. *Let $\mathcal{T}$ be a reflexive graph. If $\mathcal{T}$ is multimorphism-homogeneous, then it is HH-homogeneous.*

Proof. Let h be a local homomorphism of $\mathcal{T}$. Then the function h' mapping each $t \in \mathrm{dom}(h)$ to $\{h(t)\}$ is a multimorphism. There exists a multimorphism H defined on $\mathcal{T}$ that extends h'. Any selection of H is an endomorphism of $\mathcal{T}$ that extends h. □

Theorem 18. *Let X be a finite set and $\mathcal{S} \subseteq \mathcal{P}(X)$. Then $\mathcal{G}(\mathcal{S})$ is HH-homogeneous if and only if $\mathrm{Int}^\circ(\mathcal{S})$ is multimorphism-homogeneous.*

Proof (sketch). Suppose first that $\mathcal{G}(\mathcal{S})$ is HH-homogeneous and let $m\colon \mathcal{U} \to \mathcal{V}$ be a partial multimorphism of $\mathrm{Int}^\circ(\mathcal{S})$. By Lemma 7 there is a homomorphism $\hat{m}$ with finite domain such that $\underline{\hat{m}} = m$. By HH-homogeneity of $\mathcal{G}(\mathcal{S})$, there is an endomorphism M that extends $\hat{m}$. Now consider any finite $\mathcal{T} \subset \mathcal{G}(\mathcal{S})$ such that $\pi(\mathcal{T}) = \mathcal{S}$ and $\mathrm{dom}(\hat{m}) \subset \mathcal{T}$; it is easy to verify that $\underline{M \restriction_{\mathcal{T}}}$ is a multimorphism of $\mathrm{Int}^\circ(\mathcal{S})$ that extends m.

Now suppose that $\mathrm{Int}^\circ(\mathcal{S})$ is multimorphism-homogeneous and let $h\colon \mathcal{X} \to \mathcal{Y}$ be a local homomorphism in $\mathcal{G}(\mathcal{S})$. By results in [5] and [1], we can assume that h is a bijection. By Lemma 6 we know that $\underline{h}$ is a partial multimorphism of $\mathrm{Int}^\circ(\mathcal{S})$, which by assumption is a restriction of some multimorphism H defined on all of $\mathrm{Int}^\circ(\mathcal{S})$. We can assume that, for each $w \notin \pi(\mathcal{X})$, $|H(w)| = 1$.

Enumerate $\mathcal{X}$ as $\{x_i : 1 \le i \le |\mathcal{X}|\}$. For each $t \in \pi(\mathcal{X})$, partition K_t into $|\mathcal{X} \cap K_t|$ infinite parts $\{P(t,x_j) : x_j \in K_t\}$, ensuring that $x_j \in P(t,x_j)$. Similarly, for each $y \in \pi(\mathcal{Y})$, partition K_y into $|\mathcal{Y} \cap K_y| + |\{w \in \mathcal{S} \setminus \mathcal{X} : H(w) = \{y\}\}|$, say $\{Q(y,h(x_j)) : h(x_j) \in K_y\} \cup \{Q(y,w) : w \in \mathcal{S} \setminus \mathcal{X} : H(w) = \{y\}\}$, ensuring that each part contains no more than one element of $\mathcal{Y}$. The union of any set of bijections between the cliques that extend h in $\bigcup\{K_u : u \in \mathcal{X}\}$ and map K_w to the appropriate $Q(y,w)$ is an endomorphism of $\mathcal{G}(\mathcal{S})$. □

Problem 19. *Define multimorphism for general relational structures.*

Problem 20. *Characterise k-thin set systems.*

Towards the latter problem, the following definition was suggested by an anonymous reviewer: Given a finite set X, a set system $\mathcal{S} \subseteq \mathcal{P}(X)$, and $Z \in \mathcal{S}$, let $Z^+_{\mathcal{S}}$ denote the subsystem $\{T \in \mathcal{S} : T \cap Z \neq \varnothing\}$. Call $\mathcal{S}$ *k-thin** if for all $Z \subseteq X$ of size at most k and all multimorphisms $m\colon Z^+_{\mathcal{S}} \to \mathcal{P}(\mathcal{S})$ there exists $Y \subseteq X$ of size at most k such that $\{v \in \mathcal{S} : \exists u \in Z^+_{\mathcal{S}}(v \in m(u))\} \subseteq Y^+_{\mathcal{S}}$. It is clear that any k-thin $\mathcal{S}$ is k-thin*. This suggests the following problem, of which a positive solution could be considered a partial solution to Problem 20.

Problem 21. *Is every k-thin* set system k-thin?*

References

[1] Andrés Aranda. The poset of morphism-extension classes of countable graphs. *Mathematica Slovaca*, 72(3):539–554, 2022.

[2] Andrés Aranda and David Hartman. The independence number of HH-homogeneous graphs and a classification of MB-homogeneous graphs. *European Journal of Combinatorics*, 85:103063, 2020.

[3] Peter J Cameron and Jaroslav Nešetřil. Homomorphism-homogeneous relational structures. *Combinatorics, probability and computing*, 15(1-2):91–103, 2006.

[4] Maja Pech. *Local methods for relational structures and their weak Krasner algebras*. PhD thesis, University of Novi Sad (Serbia), 2009.

[5] Momchil Rusinov and Pascal Schweitzer. Homomorphism–homogeneous graphs. *Journal of Graph Theory*, 65(3):253–261, 2010.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Asymptotically Enumerating Independent Sets in Regular k-Partite k-Uniform Hypergraphs*

(Extended abstract)

Patrick Arras† Frederik Garbe† Felix Joos†

Abstract

Counting independent sets in regular bipartite graphs has been a topic of interest since Korshunov and Sapozhenko determined an asymptotic formula for the number of independent sets in the hypercube. A significant subsequent development was the application of the cluster expansion method to this problem due to Jenssen and Perkins. Since then, this approach has been extensively used for graphs, but little is known about analogous questions in the context of hypergraphs.

We develop a polymer model suitable for hypergraphs and apply the cluster expansion method to asymptotically determine the number of independent sets in regular k-partite k-uniform hypergraphs which satisfy natural expansion properties. The resulting formula depends only on the local structure of the hypergraph, making it computationally efficient. In particular, we provide a simple closed-form expression for linear hypergraphs.

1 Introduction

Korshunov and Sapozhenko [26] showed that the number of independent sets of the hypercube is $(1 \pm o(1)) \cdot 2\sqrt{e} \cdot 2^{2^{n-1}}$. Since the hypercube is bipartite with partition classes of order 2^{n-1}, roughly $2 \cdot 2^{2^{n-1}}$ independent sets already arise by only considering the subsets of either partition class. In fact, it turns out that most independent sets are "close" to being a subset of a partition class, that is they have only a very small number of *defect vertices* in the other class.

Over the years, this result has inspired extensive research, leading to significant generalizations of the initial setting. Most often, those rephrase the enumeration of independent sets as an evaluation of the *partition function* of a certain *polymer model*, the *hard-core model* from statistical physics, at *fugacity* 1. Generalizations then arise from varying the fugacity

*This note serves as an extended abstract for [1].

†Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany. E-mails: {arras,garbe,joos}@informatik.uni-heidelberg.de.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through project 428212407 (P. Arras, F. Garbe, F. Joos).

or modifying the polymer model. Translating back to the language of graphs, this amounts to computing a weighted sum of independent sets or counting more general objects such as proper q-colourings or homomorphisms to a fixed graph. The original proof for the hypercube was further developed by the introduction of the influential *graph container method* due to Sapozhenko [31] culminating into more general results due to Galvin [15] while also being considered for general bipartite graphs [16]. Other methods developed in this context are the *entropy method* by Kahn [24] with numerous applications [25, 30] and recently the application of *contour models* from Pirogov-Sinai theory [18]. A pivotal point in the investigation of the hypercube was the application of the so-called *cluster expansion method* due to Jenssen and Perkins [22] substantially improving on the results in [15]. One of the most comprehensive studies to date is due to Jenssen and Keevash [19], who use the cluster expansion method in a very general setting on the class of discrete tori of even sidelength. While many early results concerned rather specific graph classes, it quickly became apparent that only a few properties of the graphs in question are actually relevant, namely regularity, bipartiteness, and a quantifiable (albeit often very small) expansion. Trying to relax the required graph properties as much as possible leads to graphs in which the final count cannot be brought into a closed-form expression. Instead, the statements obtained provide a *fully polynomial-time approximation scheme* (FPTAS) for bipartite expanders [5, 6, 7, 8, 12, 13, 14, 20, 21, 23, 28, 29].

Obtaining similar results for non-bipartite graphs appears to be very difficult, as previous work by the first and third author [2] indicates; Jenssen and Keevash [19] explicitly ask for the number of independent sets in $\mathbb{Z}_3^n$. Naturally, the same question can be asked for hypergraphs. Here, FPTAS for the number of (weak) independent sets are only known if the maximum degree is bounded by a constant depending on the uniformity [10, 17]. Otherwise, finding an approximation is NP-hard [4]. Our results therefore concern counting (weak) independent sets in hypergraphs with the natural generalization of bipartiteness (our hypergraphs are k-uniform and k-partite). As in the graph case, we assume some expansion properties which are also usually assumed for quasirandom graphs. With this, we are able to approximate said number much closer than the best upper bound available in the non-k-partite setting [3, 9].

While our main result (Theorem 2.1) is formulated in the language of cluster expansion and significantly more general, it gives rise to a straightforward method of calculating the number of independent sets if the degree is sufficiently large. Before we can formulate this precisely, we need to address how the aforementioned graph properties look in our setting. A hypergraph is called a *k-graph* if every edge contains exactly k vertices. It is called *k-partite* if its vertex set can be partitioned into k subsets (called *partition classes*) such that every edge contains one vertex of each partition class. We call a hypergraph *r-regular* if every vertex is contained in exactly r edges. For a vertex subset S of a k-graph G, we define its *neighbourhood* by $N(S) := \{v \in V(G) \setminus S \mid \text{there exist } u \in S \text{ and } e \in E(G) \text{ such that } \{u,v\} \subseteq e\}$. We define $\gamma_k := \frac{2^{k-1}}{2^{k-1}-1}$ and note that $\gamma_k > 1$ for $k \geq 2$.

Definition 1.1. Let $k \geq 2$ and G be an r-regular k-partite k-graph with vertex partition $\mathcal{Z}$, where each part has order n. Given $t, b \in \mathbb{N}$ and $\alpha, \beta > 0$, we define the following properties of G:

- Reg(t): $r \geq \frac{1}{t} \log_{\gamma_k} n$.
- Exp$_1$(α): For every $S \subset Z \in \mathcal{Z}$ with $|S| \leq r$, we have $|N(S)| \geq (k-1-\alpha)r|S|$.
- Exp$_2$(β): For every $S \subset Z \in \mathcal{Z}$ with $|S| \leq \beta \frac{n}{r}$, we have $|N(S)| \geq (k-2+\beta)r|S|$.
- Def(b): For every $I \in \mathcal{I}(G)$, there is some $Z \in \mathcal{Z}$ such that $|I \cap Z| \leq b$.

Note that Reg(t) establishes a logarithmic relationship between the degree and the order of the hypergraph. This is the natural threshold at which the structure of a typical independent set changes mirroring the case of discrete tori and the hypercube. Given a subset S of some partition class Z, both expansion conditions Exp_1 and Exp_2 provide a lower bound on $|N(S)|$. Note that the larger the set S, the smaller the required relative expansion. Lastly, property $\mathrm{Def}(b)$ is necessary to allow for the aforementioned identification of defect sets. All of these conditions are typically satisfied in pseudorandom graphs and can also be expected for random regular k-partite k-graphs of logarithmic degree. Now, we denote the set consisting of all independent sets in a graph G by $\mathcal{I}(G)$. To obtain a closed-form expression for $|\mathcal{I}(G)|$, we have to assume G to be *linear*, that is that every pair of distinct edges shares at most one vertex. The following is a corollary of our main result, where α_k, denotes a positive number that only depends on k, which is made explicit later.

Theorem 1.2. *Let $k \geq 3$ and $\beta > 0$. Furthermore, let $b(n)$ be a function with $b(n) = o(n)$. Then for any $\rho > 0$, there is $n_0 \in \mathbb{N}$ such that the following holds for every $n \geq n_0$:*

Let G be a linear r-regular k-partite k-graph with vertex partition $\mathcal{Z}$, where each part has order n. If G satisfies $\mathrm{Reg}(1)$, $\mathrm{Exp}_1(\alpha_k)$, $\mathrm{Exp}_2(\beta)$, and $\mathrm{Def}(b(n))$, then

$$|\mathcal{I}(G)| = (1 \pm \rho) \cdot k \cdot 2^{(k-1)n} \cdot \exp\left(n\gamma_k^{-r}\right) .$$

Note that while our proof requires $k \geq 3$, the same statement is also true for $k = 2$, which follows from a more general result of Jenssen, Perkins, and Potukuchi [23]. Also note that $r \geq \log_{\gamma_k} n$ by $\mathrm{Reg}(1)$ implies that the final exponential term contributes at most a factor of e. Moreover, if r is larger, even if only by a constant factor, the term tends to 1 as $n \to \infty$, which in turn implies that essentially all independent sets have an empty intersection with one partition class. The primary advantage of our main theorem (Theorem 2.1) is that it provides a similar bound while requiring only $\mathrm{Reg}(t)$ for some $t \in \mathbb{N}$. It turns out, however, that the larger t becomes, the more terms we have to include in the argument of the exponential. This reflects the fact that with lower degree, the cost of including a vertex in the independent set (that is, the amount of neighbours now prevented from being included) decreases, so larger defect sets arise.

2 Statement of the main theorem

2.1 Cluster expansion

We use the well-known cluster expansion approach from statistical physics, mostly following the notation from [19]. This section first introduces the general concept and then defines the concrete polymer model for independent sets in hypergraphs. A *polymer model* $(\mathcal{P}, \sim, w)$ consists of a set $\mathcal{P}$ of so-called *polymers*, together with a *compatibility* relation $\sim$ on $\mathcal{P}$ and a weight function $w \colon \mathcal{P} \to [0, \infty)$. The *order* of a polymer S is simply $|S|$. A set $\mathcal{S} \subset \mathcal{P}$ of polymers is called *compatible* if the polymers in $\mathcal{S}$ are pairwise compatible. The *partition function* of the polymer model is then given by

$$\Xi_{(\mathcal{P},\sim,w)} := \sum_{\substack{\mathcal{S} \subset \mathcal{P} \\ \mathcal{S} \text{ compatible}}} \prod_{S \in \mathcal{S}} w(S) . \tag{2.1}$$

For a non-empty vector Γ of not necessarily distinct polymers, we let its *incompatibility graph* H_Γ be the graph with the polymers of Γ as vertices and edges between two polymers

$S, T \in \Gamma$ if $S \nsim T$. A *cluster* is a vector Γ of polymers such that H_Γ is connected. Its *length* is $|\Gamma|$ and its *size* is $\|\Gamma\| := \sum_{S\in\Gamma}|S|$. We denote the infinite set of all clusters by $\mathcal{C}_{(\mathcal{P},\sim,w)}$. Recall now that the *Ursell function* ϕ of a connected graph H is defined by

$$\phi(H) := \frac{1}{|V(H)|!} \sum_{\substack{\text{spanning, connected}\\ \text{subgraphs } F \subset H}} (-1)^{|E(F)|} .$$

With this, we define the weight of a cluster Γ as $w(\Gamma) := \phi(H_\Gamma)\prod_{S\in\Gamma} w(S)$. By [11, Proposition 5.3], we have

$$\log \Xi_{(\mathcal{P},\sim,w)} = \sum_{m=1}^{\infty} \sum_{\substack{\Gamma\in\mathcal{C}_{(\mathcal{P},\sim,w)}\\ \|\Gamma\|=m}} w(\Gamma) \,, \tag{2.2}$$

if the right-hand side converges absolutely. This sum is called the *cluster expansion* of the polymer model $(\mathcal{P}, \sim, w)$ and, provided it converges, can be truncated to yield a good approximation of $\log \Xi_{(\mathcal{P},\sim,w)}$.

2.2 Polymer model

We will use the notion of 2-linkedness: For a k-partite k-graph G and a partition class Z, we consider the 2-graph Z_2 on the vertex set Z with $vu \in E(Z_2)$ if and only if $N_G(\{v\}) \cap N_G(\{u\}) \neq \emptyset$. Furthermore, a non-empty set $S \subset Z$ is called 2*-linked* if $Z_2[S]$ is connected. Also, for a subset S of one partition class, we define the *link graph* of S as the $(k-1)$-graph $L_G(S)$ on the vertex set $N_G(S)$ with edge set $E(L_G(S)) := \{e \setminus S : e \in E(G) \text{ with } e \cap S \neq \emptyset\}$. Now we can describe the polymer model we will use. For a k-partite k-graph G, any partition class Z, and a number $b \in \mathbb{N}$, we consider the polymer model given by

$$\mathcal{P}_{Z,b} := \{S \subset Z : S \text{ is 2-linked and } 1 \le |S| \le b\} \,, \tag{2.3}$$

the compatibility relation defined by $S \sim T$ if and only if $N_G(S) \cap N_G(T) = \emptyset$, and with the polymer weights $w(S) := |\mathcal{I}(L_G(S))| \cdot 2^{-|N_G(S)|}$. We can now set $\mathcal{C}_{Z,b} := \mathcal{C}_{(\mathcal{P}_{Z,b},\sim,w)}$.

2.3 Main theorem

Having introduced the necessary setup we are now able to state our main theorem. We set

$$\alpha_{k,t} := \frac{1}{2} \min\left\{ \frac{\log_2 \gamma_k}{e^{2t}} ; \frac{(k-1)(1-\log 2)\log \gamma_k}{\log(2^{k-1}-1) + \log \gamma_k} \right\} \tag{2.4}$$

and remark that $0 < \alpha_{k,t} < 1$ for every $k \ge 2$ and $t \in \mathbb{N}$.

Theorem 2.1. *Let $k \ge 3$, $t \in \mathbb{N}$, and $\beta > 0$. Furthermore, let $b(n)$ be a function with $b(n) = o(n)$. Then for any $\rho > 0$, there is $n_0 \in \mathbb{N}$ such that the following holds for every $n \ge n_0$: Let G be an r-regular k-partite k-graph with vertex partition $\mathcal{Z}$, where each part has order n. If G satisfies* $\mathrm{Reg}(t)$, $\mathrm{Exp}_1(\alpha_{k,t})$, $\mathrm{Exp}_2(\beta)$, *and* $\mathrm{Def}(b(n))$, *then*

$$|\mathcal{I}(G)| = (1 \pm \rho) \cdot 2^{(k-1)n} \cdot \sum_{Z\in\mathcal{Z}} \exp\left(\sum_{m=1}^{t} \sum_{\substack{\Gamma\in\mathcal{C}_{Z,t}\\ \|\Gamma\|=m}} w(\Gamma) \right) .$$

Notably, neither the parameter β nor the function $b(n)$ occur in the formula of Theorem 2.1, but only influence the choice of n_0, which we do not make explicit. Likewise, both occurrences of the parameter t only determine the cluster size (and consequently, maximum polymer order) at which to truncate the sum in the argument of the exponential. Additional terms are guaranteed to vanish into the error term of $1 \pm \rho$. In other words, Theorem 2.1 asserts that for any hypergraph G satisfying the required properties, the asymptotic number of independent sets is governed solely by the clusters up to size t. The structure of these clusters, however, is a local property of G: Since the polymers in a cluster $\Gamma \in \mathcal{C}_{Z,t}$ are 2-linked and the incompatibility graph H_Γ of Γ is connected, the set $V(\Gamma) := \bigcup_{S \in \Gamma} S$ is 2-linked as well. Moreover, we have $|V(\Gamma)| \leq \|\Gamma\|$. In order to determine all clusters of size at most a constant t that contain some fixed vertex v, it therefore suffices to only consider vertices at distance at most $2(t-1)$ from $v \in Z$, which can be done in time polynomial in n. For $t = 1$, this reduces to only examine the link graph $L_G(\{v\})$. In general r-regular k-partite k-graphs, the vertex v might share multiple edges with any other vertex u. If we require G to be linear, however, this cannot occur and $L_G(\{v\})$ is a perfect matching of r pairwise disjoint edges of uniformity $k-1$. This simple local structure allows for the deduction of Theorem 1.2 from Theorem 2.1.

The proof of Theorem 2.1 expresses $|\mathcal{I}(G)|$ in terms of the partition function and uses (2.2) together with the so-called Kotecký-Preiss condition [27] to truncate the infinite sum.

References

[1] P. Arras, F. Garbe, and F. Joos. Asymptotically enumerating independent sets in regular k-partite k-uniform hypergraphs. *arXiv:2412.14845*, 2024.

[2] P. Arras and F. Joos. Independent sets in discrete tori of odd sidelength. *arXiv:2310.02747*, 2023.

[3] A. E. Balobanov and D. A. Shabanov. On the number of independent sets in simple hypergraphs. *Mat. Zametki*, 1:38–48, 2018. Russian; translation in *Math. Notes*, **1-2** (2018), 33–41.

[4] I. Bezáková, A. Galanis, L. A. Goldberg, H. Guo, and D. Štefankovič. Approximation via correlation decay when strong spatial mixing fails. *SIAM J. Comput.*, 48:279–349, 2019.

[5] S. Cannon and W. Perkins. Counting independent sets in unbalanced bipartite graphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1456–1466. SIAM, Philadelphia, PA, 2020.

[6] C. Carlson, E. Davies, A. Kolla, and A. Potukuchi. Approximately counting independent sets in dense bipartite graphs via subspace enumeration. *arXiv:2307.09533*, 2023.

[7] Z. Chen, A. Galanis, L. A. Goldberg, W. Perkins, J. Stewart, and E. Vigoda. Fast algorithms at low temperatures via Markov chains. *Random Structures Algorithms*, 58:294–321, 2021.

[8] Z. Chen, A. Galanis, D. Štefankovič, and E. Vigoda. Sampling colorings and independent sets of random regular bipartite graphs in the non-uniqueness region. *arXiv:2105.01784*, 2021.

[9] E. Cohen, W. Perkins, M. Sarantis, and P. Tetali. On the number of independent sets in uniform, regular, linear hypergraphs. *European J. Combin.*, 99:Paper No. 103401, 22 pages, 2022.

[10] W. Feng, H. Guo, C. Wang, J. Wang, and Y. Yin. Towards derandomising Markov chain Monte Carlo. *arXiv:2211.03487*, 2022.

[11] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction.* Cambridge University Press, 2017.

[12] T. Friedrich, A. Göbel, M. S. Krejca, and M. Pappik. Polymer dynamics via cliques: New conditions for approximations. *Theoret. Comput. Sci.*, 942:230–252, 2023.

[13] A. Galanis, L. A. Goldberg, and J. Stewart. Fast algorithms for general spin systems on bipartite expanders. *ACM Trans. Comput. Theory*, 13:Art. 25, 18 pages, 2021.

[14] A. Galanis, L. A. Goldberg, and J. Stewart. Fast mixing via polymers for random graphs with unbounded degree. *Inform. and Comput.*, 285:Paper No. 104894, 16 pages, 2022.

[15] D. Galvin. A threshold phenomenon for random independent sets in the discrete hypercube. *Combin. Probab. Comput.*, 20:27–51, 2011.

[16] D. Galvin and P. Tetali. Slow mixing of glauber dynamics for the hard-core model on regular bipartite graphs. *Random Structures Algorithms*, 28(4):427–443, 2006.

[17] K. He, C. Wang, and Y. Yin. Deterministic counting Lovász local lemma beyond linear programming. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3388–3425. SIAM, Philadelphia, PA, 2023.

[18] T. Helmuth, W. Perkins, and G. Regts. Algorithmic Pirogov-Sinai theory. *Probab. Theory Related Fields*, 176:851–895, 2020.

[19] M. Jenssen and P. Keevash. Homomorphisms from the torus. *Adv. Math.*, 430:Paper No. 109212, 89 pages, 2023.

[20] M. Jenssen, P. Keevash, and W. Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM J. Comput.*, 49:681–710, 2020.

[21] M. Jenssen, A. Malekshahian, and J. Park. A refined graph container lemma and applications to the hard-core model on bipartite expanders. *arXiv:2411.03393*, 2024.

[22] M. Jenssen and W. Perkins. Independent sets in the hypercube revisited. *J. Lond. Math. Soc. (2)*, 102:645–669, 2020.

[23] M. Jenssen, W. Perkins, and A. Potukuchi. Approximately counting independent sets in bipartite graphs via graph containers. *Random Structures Algorithms*, 63:215–241, 2023.

[24] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combin. Probab. Comput.*, 10:219–237, 2001.

[25] J. Kahn and J. Park. The number of 4-colorings of the Hamming cube. *Israel J. Math.*, 236:629–649, 2020.

[26] A. D. Korshunov and A. A. Sapozhenko. The number of binary codes with distance 2. *Problemy Kibernet.*, 40:111–130, 1983. Russian.

[27] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Comm. Math. Phys.*, 103:491–498, 1986.

[28] C. Liao, J. Lin, P. Lu, and Z. Mao. Counting independent sets and colorings on random regular bipartite graphs. *arXiv:1903.07531*, 2019.

[29] C. Liao, J. Lin, P. Lu, and Z. Mao. An FPTAS for the hardcore model on random regular bipartite graphs. *Theoret. Comput. Sci.*, 929:174–190, 2022.

[30] R. Peled and Y. Spinka. Rigidity of proper colorings of $\mathbb{Z}^d$. *Invent. Math*, 232:79–162, 2023.

[31] A. A. Sapozhenko. The number of antichains in ranked partially ordered sets. *Diskret. Mat.*, 1:74–93, 1989. Russian; translation in *Discrete Math. Appl.* **1** (1991), 35–58.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Faces of maximal plane graphs without short cycles have bounded degrees

(Extended abstract)

Maria Axenovich* Leon Kießle† Arsenii Sagdeev‡ Maksim Zhukovskii§

Abstract

We consider a family of plane graphs G without cycles of length less than ℓ that are maximal in a sense that adding any new edge to G either makes it non-plane or creates a cycle of length less than ℓ, $\ell \in \mathbb{N}$. We prove that the largest length of a facial cycle $\mathrm{f}_{\max}(\ell)$ of a graph from this family satisfies $\mathrm{f}_{\max}(\ell) = \Theta(\ell)$. As a corollary, we obtain lower bounds on certain planar saturation numbers. We also investigate a similar problem on other surfaces and obtain some general upper and lower bounds on the largest length of facial cycles of the corresponding graphs.

1 Faces of maximal plane graphs

For a given family $\mathcal{F}$ of graphs, what is the largest possible number of edges $\mathrm{ex}(n, \mathcal{F})$ in an n-vertex $\mathcal{F}$-*free* graph, that is, a graph that does not contain any $F \in \mathcal{F}$ as a subgraph? This question originates from the classic works of Mantel [13] and Turán [18] and has been studied extensively ever since, see the surveys [9, 17, 19]. Some special instances of this question are notoriously hard. In particular, for a family $\mathcal{C}_{<\ell} = \{C_3, \ldots, C_{\ell-1}\}$ of cycles of length less than ℓ, we known that $\mathrm{ex}(n, \mathcal{C}_{<\ell}) = O(n^{1+1/\lfloor(\ell-1)/2\rfloor})$. It is widely believed that this upper bound is asymptotically tight, but despite the long history of research, the matching lower bound was only obtained for some small values of ℓ, see [1, 7, 9].

In 2016, Dowden [6] suggested to study planar counterparts of the aforementioned extremal numbers. More formally, for a given family $\mathcal{F}$ of graphs, let $\mathrm{ex}_{\mathcal{P}}(n, \mathcal{F})$ be the largest possible number of edges in an n-vertex plane $\mathcal{F}$-free graph. This question attracted a lot of attention almost instantly: we refer the reader to [5, 10, 11, 12, 15, 16] and the references therein. In a sharp contrast to the general case, the planar extremal number of $\mathcal{C}_{<\ell}$ is not hard to compute:

*Karlsruhe Institute of Technology, Karlsruhe, Germany, maria.aksenovich@kit.edu.
Supported by the DFG grant FKZ AX 93/2-1.
†Karlsruhe Institute of Technology, Karlsruhe, Germany, leon.kiessle@student.kit.edu.
‡Karlsruhe Institute of Technology, Karlsruhe, Germany, sagdeevarsenii@gmail.com.
§School of Computer Science, University of Sheffield, UK, m.zhukovskii@sheffield.ac.uk.

a straightforward application of the Euler formula yields that $\mathrm{ex}_{\mathcal{P}}(n, \mathcal{C}_{<\ell}) \leq \frac{\ell}{\ell-2}(n-2)$, and this upper bound is essentially tight.

In the present paper, we consider another natural parameter of a plane graph G: its largest face length, that we denote by $\mathrm{f}_{\max}(G)$. Of course, the value of $\mathrm{f}_{\max}(G)$ of a plane $\mathcal{C}_{<\ell}$-free graph G can be arbitrary large in terms of ℓ as witnessed by taking $G = C_n$, $n \to \infty$. However, if in addition we assume that G is *maximal plane $\mathcal{C}_{<\ell}$-free graph*, i.e., if adding any new edge to G with both endpoints in G either makes it non-plane or creates a cycle of length less than ℓ, then all sufficiently long cycles are automatically excluded. Another family of plane $\mathcal{F}$-free graphs showing that $\mathrm{f}_{\max}(G)$ can be arbitrarily large is the family of stars. Indeed, a star $G = K_{1,n}$ has only one face of length $2n$. Moreover, adding any new edge to a star creates a triangle, and thus stars are maximal plane $\mathcal{C}_{<\ell}$-free graph for $\ell > 3$. To exclude this family too, we consider only 2-connected plane graphs, namely, plane graphs such that each of their faces is bounded by a cycle. This motivates the following definition:

$$\mathrm{f}_{\max}(\ell) = \max\{\mathrm{f}_{\max}(G) : \ G \text{ is a 2-connected maximal plane } \mathcal{C}_{<\ell}\text{-free graph}\}.$$

By taking $G = C_{2\ell-3}$, we immediately see that $\mathrm{f}_{\max}(\ell) \geq 2\ell - 3$. It is not hard to show that this simple inequality is tight for $\ell = 3, 4, 5$. The case $\ell = 6$ is already substantially harder, but the equality $\mathrm{f}_{\max}(\ell) = 2\ell - 3$ still holds for $\ell = 6$ as was shown by the first author, Ueckerdt, and Weiner [2, Lemma 5]. Perhaps surprisingly, we found a different construction showing that $\mathrm{f}_{\max}(\ell) > 2\ell - 3$ for all $\ell \geq 7$. In addition, we obtained an upper bound of the form $\mathrm{f}_{\max}(\ell) = O(\ell)$.

Theorem 1. *There exist constants $c_1, c_2 \in \mathbb{N}$ such that for all $\ell \geq 3$, we have*

$$3\ell - c_1 \leq \mathrm{f}_{\max}(\ell) \leq 8\ell + c_2.$$

Sketch of the proof of Theorem 1. Our lower bound construction is a subdivided wheel. For the upper bound, we fix a facial cycle C of some 2-connected maximal plane $\mathcal{C}_{<\ell}$-free graph and study the properties of the shortest paths connecting its antipodal vertices. We show that for every pair of antipodal vertices of C, there are two 'short' vertex-disjoint paths between them. For two pairs of antipodal vertices that partite C into 4 roughly equal 'quarters', these shortest paths between them must intersect by planarity. If the cycle C is 'long', then these points of intersection create 4 edge disjoint cycles, one of which is 'short' by the pigeonhole principle, a contradiction. □

Note that in particular, Theorem 1 implies that $\mathrm{f}_{\max}(\ell) < \infty$ for all $\ell \in \mathbb{N}$. It could be interesting to find a short proof if this innocent-looking inequality, but for a moment, we are unable to do so.

2 Planar saturation

Most Turán-type problems have their saturation counterparts, where the goal is somewhat the opposite: one aims to *minimize* the size of the maximal construction, see the survey [8]. In case of the planar extremal numbers discusses in Section 1, the corresponding saturation-type question is as follows. For a given family $\mathcal{F}$ of graphs, what is the largest possible number of edges $\mathrm{sat}_{\mathcal{P}}(n, \mathcal{F})$ in an n-vertex maximal plane $\mathcal{F}$*-free* graph? See also [4, 3] for a slightly different but closely related notion.

Note that if $\mathcal{F} = \mathcal{C}_{<\ell}$ with $\ell > 3$, then every maximal plane $\mathcal{C}_{<\ell}$-free contains at least $n-1$ edges, which is tight as witnessed by stars. Hence, the problem is not very intriguing since $\mathrm{sat}_{\mathcal{P}}(n, \mathcal{C}_{<\ell}) = n-1$. But how does the answer changes if, to exclude this rather degenerate example, we consider only 2-connected graphs? More formally, let

$$\mathrm{sat}_{\mathcal{P}}^{\text{2-con}}(n, \mathcal{C}_{<\ell}) = \min\{|E(G)| : G \text{ is a 2-connected maximal plane } \mathcal{C}_{<\ell}\text{-free graph on } n \text{ vertices}\}.$$

We study the constant $c(\ell)$ such that $\mathrm{sat}_{\mathcal{P}}^{\text{2-con}}(n, \mathcal{C}_{<\ell}) \sim c(\ell)n$ as n tends to infinity (if it exists), and determine its dependence on ℓ asymptotically.

Theorem 2. *There exists positive constants $\alpha_1 < \alpha_2$ such that*

$$1 + \frac{\alpha_1}{\ell} \leq \liminf_{n\to\infty} \frac{\mathrm{sat}_{\mathcal{P}}^{\text{2-con}}(n, \mathcal{C}_{<\ell})}{n} \leq \limsup_{n\to\infty} \frac{\mathrm{sat}_{\mathcal{P}}^{\text{2-con}}(n, \mathcal{C}_{<\ell})}{n} \leq 1 + \frac{\alpha_2}{\ell}.$$

Sketch of the proof of Theorem 2. The upper bound is via an explicit constriction of a 2-connected plane $\mathcal{C}_{<\ell}$-free graph such that almost all of its faces are bounded by cycles of length $2\ell - 3$. For the lower bound, observe that if G is 2-connected maximal plane $\mathcal{C}_{<\ell}$-free graph, then the degree of every face of G is at most $8\ell + O(1)$ by Theorem 1. Now a direct application of the Euler formula completes the proof. $\square$

It would be interesting to close the gap between the bounds in Theorem 2.

3 Faces of maximal surface-embedded graphs

We also consider counterparts of the problem from Section 1 for other surfaces. For a surface Σ and a $\mathcal{C}_{<\ell}$-free graph G embedded in Σ, we say that G is *maximal $\mathcal{C}_{<\ell}$-free graph embedded in Σ* if adding any new edge to G with both endpoints in G creates either a crossing on Σ or a cycle of length less than ℓ. As earlier in the plane, the examples of long cycles and stars show that if a $\mathcal{C}_{<\ell}$-free graph G embedded in Σ is not maximal or has a face not bounded by a cycle, then $\mathrm{f}_{\max}(G)$ may be arbitrarily large in terms of ℓ. This leads to the following definition:

$$\mathrm{f}_{\max}(\ell, \Sigma) = \max\{\mathrm{f}_{\max}(G) : G \text{ is a maximal } \mathcal{C}_{<\ell}\text{-free graph embedded in } \Sigma \text{ such that every face of } G \text{ is bounded by a cycle}\}.$$

First, note that if Σ is a plane, then $\mathrm{f}_{\max}(\ell, \Sigma) = \mathrm{f}_{\max}(\ell)$ because a plane graph G is 2-connected if and only if every face of G is bounded by a cycle. However, for other surfaces, these two conditions may not be equivalent. Second, observe that puncturing a surface and/or cutting out discs from it does not affect the class of graphs that can be embedded in this surface, and thus the value of $\mathrm{f}_{\max}(\ell, \Sigma)$ as well. In particular, in our context, there is no difference between a sphere and a plane, since the latter one is homeomorphic to a punctured sphere. Hence, without loss of generality, we can consider only closed surfaces, namely, compact surfaces without boundary. The celebrated classification theorem, see e.g. [14, Theorem 3.1.3], states that each of these surfaces is homeomorphic to either a sphere with g handles $\mathbb{S}_g$ or to a sphere with g crosscaps $\mathbb{N}_g$ for some $g \geq 0$. Since the case $g = 0$ is covered in Section 1, we assume that $g \geq 1$. Our last result in this paper provides general bounds on $\mathrm{f}_{\max}(\ell, \Sigma)$ for this canonical surfaces.

Theorem 3. *Let $g \geq 1$ and $\ell \geq 3$ be integers and Σ be a surface, $\Sigma \in \{\mathbb{S}_g, \mathbb{N}_g\}$. Then*

$$g(\ell - 5) \leq \mathrm{f}_{\max}(\ell, \Sigma) \leq \left((4g + 4)^2 \ell\right)^{\ell}.$$

Sketch of the proof of Theorem 3. Our lower bound construction is a double subdivided wheel. For the upper bound, we fix a facial cycle C of some maximal $\mathcal{C}_{<\ell}$-free graph embedded in Σ, and study the shortest paths between its vertices. We show that if C is 'long', then the graph formed by these shortest paths contains two *spiders* with different *heads* and 'many' *legs* that 'alternate'. (A *spider* is a tree that is a subdivision of a star, i.e., a tree with exactly one vertex of degree at least three, called the *head* of the spider. A *leg* of a spider is a path with endpoints that are the head and a leaf of the spider.) Since certain bipartite graphs cannot be embedded in Σ, we know that 'many' of the legs of these spiders must intersect. Then the pigeonhole principle yields that one of this legs must contain 'many' different intersections points. On the other hand, this leg is part of a shortest paths and thus must be 'short' itself, a contradiction. $\square$

Note that for every fixed $\ell \geq 6$ and $\Sigma \in \{\mathbb{S}_g, \mathbb{N}_g\}$, Theorem 3 implies that $\Omega(g) = \mathrm{f}_{\max}(\ell, \Sigma) = O(g^{2\ell})$, and thus $\mathrm{f}_{\max}(\ell, \Sigma)$ growths polynomially with g. It would be interesting to determine the correct degree of this polynomial as well as to study the 'opposite' regime when g is fixed and ℓ tends to infinity. It might even be true that there exists a constant $C > 0$ such that $\mathrm{f}_{\max}(\ell, \Sigma) < Cg\ell$ for all positive integers g and ℓ.

References

[1] Noga Alon, Shlomo Hoory, and Nathan Linial. The moore bound for irregular graphs. *Graphs and Combinatorics*, 18:53–57, 2002.

[2] Maria Axenovich, Torsten Ueckerdt, and Pascal Weiner. Splitting planar graphs of girth 6 into two linear forests with short paths. *Journal of Graph Theory*, 85(3):601–618, 2017.

[3] János Barát, Zoltán L Blázsik, Balázs Keszegh, and Zeyu Zheng. On the number of edges in saturated partial embeddings of maximal planar graphs. *arXiv preprint arXiv:2502.05438*, 2025.

[4] Alexander Clifton and Nika Salia. Saturated partial embeddings of planar graphs. *arXiv preprint arXiv:2403.02458*, 2024.

[5] Daniel W Cranston, Bernard Lidickỳ, Xiaonan Liu, and Abhinav Shantanam. Planar turán numbers of cycles: A counterexample. *The Electronic Journal of Combinatorics*, 29(3), 2022.

[6] Chris Dowden. Extremal c4-free/c5-free planar graphs. *Journal of Graph Theory*, 83(3):213–230, 2016.

[7] Ronald D Dutton and Robert C Brigham. Edges in graphs with large girth. *Graphs and Combinatorics*, 7(4):315–321, 1991.

[8] Jill R Faudree, Ralph J Faudree, and John R Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, 1000:DS19–Jul, 2011.

[9] Zoltán Füredi and Miklós Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, pages 169–264. Springer, 2013.

[10] Debarun Ghosh, Ervin Gyori, Ryan R Martin, Addisu Paulos, and Chuanqi Xiao. Planar turán number of the 6-cycle. *SIAM Journal on Discrete Mathematics*, 36(3):2028–2050, 2022.

[11] Ervin Győri, Alan Li, and Runtian Zhou. The planar tur\'an number of the seven-cycle. *arXiv preprint arXiv:2307.06909*, 2023.

[12] Yongxin Lan, Yongtang Shi, and Zi-Xia Song. Extremal h-free planar graphs. *arXiv preprint arXiv:1808.01487*, 2018.

[13] W Mantel. Problem 28, wiskundige opgaven 10 (1907), 60–61.

[14] Bojan Mohar and Carsten Thomassen. *Graphs on surfaces.* Johns Hopkins University Press, 2001.

[15] Ruilin Shi, Zach Walsh, and Xingxing Yu. Dense circuit graphs and the planar turán number of a cycle. *Journal of Graph Theory*, 108(1):27–38, 2025.

[16] Ruilin Shi, Zach Walsh, and Xingxing Yu. Planar turán number of the 7-cycle. *European Journal of Combinatorics*, 126:104134, 2025.

[17] Alexander Sidorenko. What we know and what we do not know about turán numbers. *Graphs and Combinatorics*, 11(2):179–199, 1995.

[18] Paul Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48(436-452):61, 1941.

[19] Jacques Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pages 83–116. Springer, 2016.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Canonical Ramsey numbers for partite hypergraphs

(Extended abstract)

Matías Azócar Carvajal* Giovanne Santos† Mathias Schacht‡

Abstract

We show that canonical Ramsey numbers for partite hypergraphs grow single exponentially for any fixed uniformity.

1 Introduction

Erdős and Rado [7] established the canonical Ramsey theorem, which generalises Ramsey's theorem [13] to an unbounded number of colours. In that work Erdős and Rado characterised all canonical colour patterns that are unavoidable in colourings of edge sets of sufficiently large hypergraphs. The canonical Ramsey number $\mathrm{ER}(K_t^{(k)})$ is the smallest integer n such that any edge colouring $\varphi\colon E(K_n^{(k)}) \longrightarrow \mathbb{N}$ of the complete k-uniform hypergraph on n vertices yields a canonical copy of $K_t^{(k)}$, i.e., a copy which exhibits one of those unavoidable colour patterns. (The precise definition

*Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany. E-mail: matias.azocar.carvajal@uni-hamburg.de.
†Departamento de Ingeniería Matemática, Universidad de Chile, Santiago, Chile. E-mail: gsantos@dim.uchile.cl.
‡Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany. E-mail: schacht@math.uni-hamburg.de.

of these colour patterns is not important at this point, however, we remark that for their definition the underlying vertex set is assumed to be ordered.) We consider quantitative aspects of this theorem. For the Erdős–Rado theorem discussed above it follows from the work of Erdős, Hajnal, and Rado [6, §16.4] (see also reference [5, (4.2)]), the work of Lefmann and Rödl [11], and the work of Shelah [15] that the lower and the upper bound on n grow as $(k-1)$-times iterated exponentials in polynomials of t (see also reference [14, §4]). In other words, in terms of the number of exponentiations the canonical Ramsey number and the non-canonical Ramsey number (for many colours) display the same behaviour.

We study Erdős–Rado numbers for k-partite k-uniform hypergraphs. The extremal problem for k-partite k-uniform hypergraphs is degenerate and as a result the Ramsey number grows much slower. In fact, owing to the work of Kövari, Sós, and Turán [10] and of Erdős [4] those Ramsey numbers for any fixed number of colours grow only exponential and random colourings yield a matching lower bound. Roughly speaking, we show that canonical Ramsey numbers for partite hypergraphs exhibit the same behaviour and, in fact, these extremal results will be crucial in the proof. We recall the definition of canonical colourings for partite hypergraphs.

Definition 1.1. For a k-partite k-uniform hypergraph $H = (V_1 \dot\cup \ldots \dot\cup V_k, E)$, a set $J \subseteq [k]$, and an edge $e \in E$ we write

$$e_J = e \cap \bigcup_{j \in J} V_j$$

for the restriction of e to the vertex classes indexed by J. We say a colouring $\varphi\colon E \longrightarrow \mathbb{N}$ is J*-canonical*, if for all edges e, $e' \in E$ we have

$$\varphi(e) = \varphi(e') \qquad \Longleftrightarrow \qquad e_J = e'_J\,.$$

Moreover, we say the colouring is *canonical* if it is J-canonical for some $J \subseteq [k]$ and a subhypergraph is a *canonical copy*, if the colouring φ restricted to its edges is canonical.

Note that $\varnothing$-canonical colourings are monochromatic and $[k]$-canonical colourings are *rainbow* colourings.

Similarly as above, we define $\mathrm{ER}(K^{(k)}_{t,\ldots,t})$ as the smallest integer n such that every colouring $\varphi\colon E(K^{(k)}_{n,\ldots,n}) \longrightarrow \mathbb{N}$ of the edges of the complete k-partite k-uniform hypergraph with vertex classes of size n yields a canonical copy of $K^{(k)}_{t,\ldots,t}$. It follows from the work of Rado [12] that these numbers exist and a simple probabilistic argument employing a random colouring with $t^k - 1$ colours shows

$$\mathrm{ER}(K^{(k)}_{t,\ldots,t}) \geqslant t^{(1-o(1))t^{k-1}}\,, \tag{1.1}$$

where $o(1) \longrightarrow 0$ as $t \longrightarrow \infty$. We establish a comparable upper bound, which resolves a problem raised by Dobák and Mulrenin [3].

Theorem 1.2. *For sufficiently large t we have $\mathrm{ER}(K^{(2)}_{t,t}) \leqslant t^{3(t+1)}$ and $\mathrm{ER}(K^{(3)}_{t,t,t}) \leqslant t^{30t^3}$. Moreover, for every $k \geqslant 4$ and t sufficiently large the following holds*

$$\mathrm{ER}(K^{(k)}_{t,\dots,t}) \leqslant t^{t^{k^2}}\,.$$

Theorem 1.2 for $k = 2$ is optimal up to the factor 3 in the exponent and similar bounds were obtained by Kostochka, Mubayi, and Verstraëte [9], by Gishboliner, Milojević, Sudakov, and Wigderson [8] and by Dobák and Mulrenin [3]. For $k = 3$ there is a more substantial gap between the lower bound (1.1) and the upper bound provided by Theorem 1.2. In view of that, it would be interesting to decide if the cubic exponent $30t^3$ in the upper bound could be improved to be quadratic.

For larger values of k the gap between the lower and the upper bound widens and in the proof of Theorem 1.2 we made no attempt for obtaining the optimal constant in front of the exponent k^2. However, our method seems to fall short to obtain a factor smaller than 1/2. In particular, we leave it open if the k^2 can be improved to $o(k^2)$ or even to $O(k)$ for $k \longrightarrow \infty$, as suggested by the lower bound (1.1).

2 Preliminaries

The proof of Theorem 1.2 follows the approach of Dobák and Mulrenin [3] and employs an unbalanced variant of Erdős' [4] hypergraph extension of the Kövari–Sós–Turán theorem [10].

Proposition 2.1. *For $k \geqslant 1$ let $H = (V_1 \cupdot \dots \cupdot V_k, E)$ be a k-partite k-uniform hypergraph of density $d = \frac{|E|}{|V_1|\cdots|V_k|}$. If for some positive integers $t_1, \dots, t_k$ we have*

$$\left(\frac{d}{4^{k-1}}\right)^{\prod_{j<i} t_j} |V_i| \geqslant 2t_i \tag{2.1}$$

for every $i \in [k]$, then the number $K^{(k)}_{t_1,\dots,t_k}(H)$ of complete k-partite k-uniform hypergraphs in H with vertex classes $U_j \subseteq V_j$ and $|U_j| = t_j$ for every $j \in [k]$ satisfies

$$K^{(k)}_{t_1,\dots,t_k}(H) > \left(\frac{d}{2^{2k-1}}\right)^{\prod_{j\in[k]} t_j} \prod_{j\in[k]} \binom{|V_j|}{t_j}\,. \tag{2.2}$$

Another key idea, often used to locate rainbow copies, is to consider bounded colourings. For a k-partite k-uniform hypergraph $H = (V_1 \cupdot \ldots \cupdot V_k, E)$ and $J \subseteq [k]$ we write V_J for the set of $|J|$-element vertex sets intersecting the vertex classes indexed by J, i.e.,

$$V_J = \big\{\{v_{i_1}, \ldots, v_{i_{|J|}}\} : v_{i_j} \in V_j \text{ for all } j \in J\big\}.$$

It is clear that $|V_\varnothing| = 1$ and $|V_J| = \prod_{j \in J} |V_j|$.

Definition 2.2. We say a colouring $\varphi \colon E \longrightarrow \mathbb{N}$ of the edge set of a k-partite k-uniform hypergraph $H = (V_1 \cupdot \ldots \cupdot V_k, E)$ is *(δ, J)-bounded* for some $\delta > 0$ and some set $J \subsetneq [k]$ if, for every colour $\ell \in \mathbb{N}$, all but at most $\delta|V_J|$ of the $|J|$-tuples $S \in V_J$ satisfy

$$\big|\{e \in E \colon e_J = S \text{ and } \varphi(e) = \ell\}\big| \leqslant \delta\big|V_{[k] \smallsetminus J}\big|.$$

For $j = 0, \ldots, k-1$ we say the colouring φ is *(δ, j)-bounded*, if it is *(δ, J)-bounded* for every j-element set $J \in [k]^{(j)}$. Moreover, we say φ is *$\boldsymbol{\delta}$-bounded* for some $\boldsymbol{\delta} = (\delta_0, \ldots, \delta_{k-1}) \in (0, 1]^k$ if φ is (δ_j, j)-bounded for every $j = 0, \ldots, k-1$.

Babai [2], Lefmann and Rödl [11] and Alon, Jiang, Miller, and Pritikin [1] have exploited the fact that $\boldsymbol{\delta}$-bounded colourings for sufficiently small choices of $\delta_0, \ldots, \delta_{k-1}$ yield large rainbow subhypergraphs. In fact, being (δ, J)-bounded implies that the number of those obstructions to a rainbow coloured subhypergraph consisting of two edges e, e' of the same colour with $e \cap e' \in V_J$ is at most $\delta_{|J|}|V_J||V_{[k]\smallsetminus J}|^2$.

Proposition 2.3. *For $k \geqslant 2$ and $\boldsymbol{\delta} = (\delta_0, \ldots, \delta_{k-1}) \in (0, 1]^k$ let $\varphi \colon E(K^{(k)}_{V_1,\ldots,V_k}) \longrightarrow \mathbb{N}$ be a $\boldsymbol{\delta}$-bounded colouring with of the complete k-partite k-uniform hypergraph $K^{(k)}_{V_1,\ldots,V_k}$ with vertex partition $V_1 \cupdot \ldots \cupdot V_k$. If for some positive integer $t \leqslant \frac{1}{2} \min\big\{|V_1|, \ldots, |V_k|\big\}$ we have*

$$\delta_j \leqslant \frac{1}{2^{3k-j} \cdot t^{2k-j-1}} \tag{2.3}$$

for every $j = 0, \ldots, k-1$, then φ yields a rainbow copy of the complete k-partite $K^{(k)}_{t,\ldots,t}$.

We shall use a variant of Proposition 2.3, where we move away from complete partite hypergraphs. Instead we start with a partite hypergraph of density d and we are interested in large rainbow subhypergraphs of similar density.

Proposition 2.4. *For $k \geqslant 2$ let $H = (V_1 \cupdot \ldots \cupdot V_k, E)$ be a k-partite k-uniform hypergraph of density $d = \frac{|E|}{|V_1|\cdots|V_k|} > 0$ and for $\boldsymbol{\delta} = (\delta_0, \ldots, \delta_{k-1}) \in (0, 1]^k$ let $\varphi \colon E \longrightarrow \mathbb{N}$ be a $\boldsymbol{\delta}$-bounded colouring. If for some integer m we have*

$$\frac{2^{k+3}}{d} \leqslant m \leqslant \min\big\{|V_1|, \ldots, |V_k|\big\} \quad \text{and} \quad \delta_j < \frac{1}{2^{k+1} \cdot m^{2k-j}} \tag{2.4}$$

for every $j = 0, \dots, k-1$, then H contains a rainbow subhypergraph of density at least $d/2$ with vertex classes $U_j \subseteq V_j$ and $|U_j| = m$ for every $j \in [k]$.

The proof of Proposition 2.4 parallels the proof of Proposition 2.3. Roughly speaking, we show that a randomly chosen subhypergraph inherits the density of H. However, for technical reasons we will refrain from removing vertices from the randomly chosen vertex sets and instead we required smaller values of δ_j in the assumption (2.4) leading to no expected obstructions for the rainbow subhypergraph.

3 Main result

We show a proof of Theorem 1.2 for $k = 3$. Let t be sufficiently large, and set $n = t^{30t^3}$. We fix auxiliary constants

$$\delta_2 = \frac{1}{2^7 t^3}\,, \qquad m_2 = t^{7t}\,, \qquad \text{and} \qquad \delta_1 = \delta_0 = \frac{1}{2^{10} \cdot t^{29t}}$$

Also, we set $\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2)$. We shall show that every colouring $\varphi \colon E(K^{(3)}_{n,n,n}) \longrightarrow \mathbb{N}$ yields a canonical copy of $K^{(3)}_{t,t,t}$. Let $V_1 \,\dot\cup\, V_2 \,\dot\cup\, V_3$ be the vertex partition of $K^{(3)}_{n,n,n}$. Given a colouring φ we consider several cases depending on the 'boundedness properties' of φ.

Note that, if φ is indeed $\boldsymbol{\delta}$-bounded, then our choice of $\boldsymbol{\delta}$ yields a rainbow coloured copy of $K^{(3)}_{t,t,t}$ by Proposition 2.3. Consequently, there exists a minimal index $j_\star \in \{0, 1, 2\}$ such that φ is not $(\delta_{j_\star}, j_\star)$-bounded and let $J_\star \in [3]^{(j_\star)}$ be a set that witnesses this and without loss of generality we may assume $J_\star = \{1, \dots, j_\star\}$.

In case $j_\star = 0$ one of the colours appears at least $\delta_0 n^3$ times. Thus, an application of Proposition 2.1 yields a monochromatic copy of $K^{(3)}_{t,t,t}$.

In case $j_\star = 1$ there is a set $U \subseteq V_1$ of size at least $\delta_1 n$ such that every $u \in U$ is contained in at least $\delta_1 n^2$ edges of the same colour and we denote this colour by $\ell(u)$.

The box principle yields a subset $U_\star \subseteq U$ of size at least $\sqrt{\delta_1 n}$ such that either all colours $\ell(u)$ for $u \in U_\star$ are equal or they are all distinct. We consider the 3-partite 3-uniform hypergraph $H_\star$ with vertex partition $U_\star \,\dot\cup\, V_2 \,\dot\cup\, V_3$ and

$$E(H_\star) = \bigcup_{u \in U_\star} \big\{ e \in V_{[3]} \colon u \in e \text{ and } \varphi(e) = \ell(u) \big\}\,.$$

Since every vertex $u \in U_\star$ has degree at least $\delta_1 n^2$, the hypergraph $H_\star$ has density at least δ_1. Again we apply Proposition 2.1 to $H_\star$, and we obtain either a monochromatic or an $\{1\}$-canonical copy of $K^{(3)}_{t,t,t}$.

It is left to consider the case $j_\star = 2$. Let $U_{J_\star} \subseteq V_{J_\star}$ be a set of size at least $\delta_2 n^2$ such that every 2-tuple $S \in U_{J_\star}$ extends to at least $\delta_2 n$ edges of the same colour and we denote this colour by $\ell(S)$. We consider the bipartite graph G with vertex partition $V_1 \cup V_2$ and $E(G) = U_{J_\star}$. Moreover, we define a colouring $\varphi_G \colon E(G) \longrightarrow \mathbb{N}$ through $\varphi_G(S) = \ell(S)$ for all $S \in E(G)$. Since φ is (δ_j, j)-bounded, for $j \in \{0, 1\}$, the colouring φ_G of the bipartite graph G is $\big(\delta_0/\delta_2, \delta_1/\delta_2\big)$-bounded. Therefore, we may apply Proposition 2.4 with $k = 2, d = \delta_2, (\delta_0/\delta_2, \delta_1/\delta_2)$, and m_2 to G to obtain a rainbow bipartite subgraph $G_\star$ of density at least $\delta_2/2$ with vertex classes $U_j \subseteq V_j$ and $|U_j| = m_2$ for every $j \in [2]$.

Finally, we consider the natural 3-uniform extension $H_\star$ of $G_\star$ on the vertex partition $U_1 \cup U_2 \cup V_3$ with

$$E(H_\star) = \bigcup_{S \in E(G_\star)} \big\{e \in V_{[3]} \colon S \subseteq e \text{ and } \varphi(e) = \ell(S)\big\}$$

and note that the colouring φ restricted to $H_\star$ is $J_\star$-canonical. Moreover, since every edge $S \in E(G_\star)$ extends to at least $\delta_2 n$ distinct 3-tuples of colour $\ell(S)$ under φ, the 3-uniform hypergraph $H_\star$ has density at least $\delta_2^2/2$. Another application of Proposition 2.1 to $H_\star$ yields a $J_\star$-canonical copy of $K^{(3)}_{t,t,t}$. This concludes the proof of Theorem 1.2 for $k = 3$.

For the case $k = 2$ one can check that the same proof works for $n = t^{3(t+1)}$ for sufficiently large t with $\delta_1 = \delta_0 = 2^{-6}t^{-3}$. In fact, for graphs the proof is somewhat simpler, since the case $j_\star \geqslant 2$ does not arise.

For the general case $k \geqslant 4$, we need to deal with $j_\star > 2$. The proof of this case is a natural generalization of the case $j_\star = 2$. We define a $j_\star$-partite $j_\star$-uniform graph G with vertex partition $V_1 \cup \cdots \cup V_{j_\star}$ and $E(G) = U_{J_\star}$. We also define a colouring φ_G of $E(G)$ similarly as before. Owing to the minimal choice of $j_\star$, this colouring is $(\delta_0/\delta_{j_\star}, \ldots, \delta_{j_\star - 1}/\delta_{j_\star})$-bounded. Thus, by choosing suitable auxiliary constants $\delta_0, \ldots, \delta_{k-1}$, and $m_2, \ldots, m_{k-1}$ we can also apply Propositions 2.4 and 2.1 to find a $J_\star$-canonical copy of $K^{(k)}_{t,\ldots,t}$

References

[1] N. Alon, T. Jiang, Z. Miller, and D. Pritikin, *Properly colored subgraphs and rainbow subgraphs in edge-colorings with local constraints*, Random Structures Algorithms **23** (2003), no. 4, 409–433, DOI 10.1002/rsa.10102. MR2016871 ↑2

[2] L. Babai, *An anti-Ramsey theorem*, Graphs Combin. **1** (1985), no. 1, 23–28, DOI 10.1007/BF02582925. MR796179 ↑2

[3] D. Dobák and E. Mulrenin, *Sharp exponents for bipartite Erdős-Rado numbers*, available at arXiv:2410.08982. ↑1, 1, 2

[4] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel J. Math. **2** (1964), 183–190, DOI 10.1007/BF02759942. MR183654 ↑1, 2

[5] P. Erdős and A. Hajnal, *Ramsey-type theorems*, Discrete Appl. Math. **25** (1989), no. 1-2, 37–52, DOI 10.1016/0166-218X(89)90045-0. Combinatorics and complexity (Chicago, 1987). MR1031262 ↑1

[6] P. Erdős, A. Hajnal, and R. Rado, *Partition relations for cardinal numbers*, Acta Math. Acad. Sci. Hungar. **16** (1965), 93–196, DOI 10.1007/BF01886396. MR202613 ↑1

[7] P. Erdős and R. Rado, *A combinatorial theorem*, J. London Math. Soc. **25** (1950), 249–255, DOI 10.1112/jlms/s1-25.4.249. MR37886 ↑1

[8] L. Gishboliner, A. Milojević, B. Sudakov, and Y. Wigderson, *Canonical Ramsey numbers of sparse graphs*, available at arXiv:2410.08644. ↑1

[9] A. Kostochka, D. Mubayi, and J. Verstraëte, *Turán problems and shadows II: Trees*, J. Combin. Theory Ser. B **122** (2017), 457–478, DOI 10.1016/j.jctb.2016.06.011. MR3575215 ↑1

[10] T. Kövari, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloq. Math. **3** (1954), 50–57, DOI 10.4064/cm-3-1-50-57. MR65617 ↑1, 2

[11] H. Lefmann and V. Rödl, *On Erdős–Rado numbers*, Combinatorica **15** (1995), no. 1, 85–104, DOI 10.1007/BF01294461. MR1325273 ↑1, 2

[12] R. Rado, *Direct decomposition of partitions*, J. London Math. Soc. **29** (1954), 71–83, DOI 10.1112/jlms/s1-29.1.71. MR65616 ↑1

[13] F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. (2) **30** (1930), no. 4, 264–286, DOI 10.1112/plms/s2-30.1.264. MR1576401 ↑1

[14] Chr. Reiher, V. Rödl, M. Sales, K. Sames, and M. Schacht, *On quantitative aspects of a canonisation theorem for edge-orderings*, J. Lond. Math. Soc. (2) **106** (2022), no. 3, 2773–2803, DOI 10.1112/jlms.12648. MR4498567 ↑1

[15] S. Shelah, *Finite canonization*, Comment. Math. Univ. Carolin. **37** (1996), no. 3, 445–456. MR1426909 ↑1

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Estimating multicolor ordered Ramsey numbers

(Extended abstract)

Martin Balko* Klára Grinerová*

Abstract

We study multicolor ordered Ramsey numbers, an analog of the classical Ramsey numbers for an arbitrary number of colors and graphs with linearly ordered vertex sets. We generalize upper and lower bounds on two-colored ordered Ramsey numbers of ordered matchings by Conlon, Fox, Lee, and Sudakov to an arbitrary number of colors. Using the extremal theory of matrices, we prove that the q-color ordered Ramsey numbers of ordered matchings on n vertices with interval chromatic number two are in $n^{\Theta(q)}$, which is tight up to a constant in the exponent. We extend this result to ordered graphs with bounded degeneracy and with interval chromatic number two.

1 Introduction

Ramsey theory is a very active and important area of mathematics, whose underlying philosophy states that every sufficiently large system contains a well-organized subsystem. One of the cornerstones of the Ramsey theory is the following result, called the *Ramsey theorem*, which we now state for graphs and $q \geq 2$ colors. For arbitrary graphs $G_1, \ldots, G_q$, there exists a positive integer N such that for every q-coloring χ of the complete graph K_N on N vertices with colors from $[q] = \{1, 2, \ldots, q\}$ there is an $i \in [q]$ such that χ contains a monochromatic copy of G_i in color i as a subgraph. We denote the smallest such integer N as $r(G_1, \ldots, G_q)$ and call it the *multicolor Ramsey number* of $G_1, \ldots, G_q$. For simplicity, if $G_1 = \cdots = G_q$, we write $r(G; q)$ instead of $r(G_1, \ldots, G_q)$. In the case of two colors, we write $r(G)$ instead of $r(G; 2)$.

Obtaining good bounds on Ramsey numbers is a well-known and notoriously difficult problem even for two colors. Until recently, the classical estimates $2^{n/2} \leq r(K_n) \leq 4^n$ on Ramsey numbers of complete graphs [16, 18] were improved only by smaller-order terms [13]. In 2023, Campos, Griffiths, Morris, and Sahasrabudhe [10] came up with an exponential improvement for the upper bound and proved that $r(K_n) \leq (4 - \varepsilon)^n$ for some constant

*Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Czech Republic. E-mail: {balko,klarka}@kam.mff.cuni.cz. M. Balko was supported by grant no. 23-04949X of the Czech Science Foundation (GAČR) and by the Center for Foundations of Modern Computer Science (Charles Univ. project UNCE 24/SCI/008). K. Grinerová was supported by grant no. 23-04949X of the Czech Science Foundation (GAČR).

$\varepsilon > 0$. In the multicolor case, the problem of estimating Ramsey numbers is even less understood. Except for lower-order terms improvements, the best known upper bound for $q \geq 3$ is $r(K_n; q) < q^{qn}$, which can be proved through a modification of the Erdős and Szekeres [18] neighborhood-chasing argument that yields $r(K_n) \leq 4^n$. Very recently, there has been a breakthrough in estimating multicolor Ramsey numbers made by Balister et al. [4] who proved that for each fixed $q \geq 2$ there is some constant $\delta = \delta(q) > 0$ such that

$$r(K_n; q) \leq e^{-\delta n} q^{qn}.$$

For the lower bound, if we repeatedly combine an observation made by Lefmann [24] with the bounds $r(K_n; 2) - 1 \geq 2^{n/2}$ and $r(K_n; 3) - 1 \geq 3^{n/2}$, we obtain

$$r(K_n; 3k) > 3^{kn/2},\ r(K_n; 3k+1) > 2^n 3^{(k-1)n/2},\ r(K_n; 3k+2) > 2^{n/2} 3^{kn/2}.$$

Recently, there was an exciting development in this area and new asymptotically stronger lower bounds on $r(K_n; q)$ were discovered. First, Conlon and Ferber [14] proved that if q is a prime number, then $r(K_n; q+1) > 2^{n/2} q^{3n/8}$. In the cases $q = 2$ and $q = 3$ this yields exponential improvements over the bounds for $r(K_n; 3)$ and $r(K_n; 4)$. Improvements for $q \geq 5$ then follow from repeated applications of Lefmann's observation, yielding

$$r(K_n; 3k) > 2^{7kn/8},\ r(K_n; 3k+1) > 2^{7(k-1)n/8} 3^{3n/8},\ r(K_n; 3k+2) > 2^{7kn/8 + n/2}.$$

Soon after that Wigderson [33] modified the construction of Conlon and Ferber and obtained improved bound $r(K_n; q) \geq \left(2^{\frac{3q}{8} - \frac{1}{4}}\right)^n$. A further slight improvement was then made by Sawin [31], who obtained the estimate $r(K_n; q) \geq 2^{0{,}383796(q-2)n + \frac{n}{2} + o(n)}$.

Motivated by these new breakthroughs, we focus on studying the multicolor *ordered Ramsey numbers*. This is a variant of multicolor Ramsey numbers for graphs with linearly ordered vertex sets. Although estimating ordered Ramsey numbers is a quite active area in extremal combinatorics with applications in discrete geometry and many close connections to famous open problems, see a recent survey [6], they were mostly studied only in the two-color setting. In this paper, we derive the first non-trivial estimates on multicolor ordered Ramsey numbers.

For a positive integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. Throughout the paper, all logarithms are base 2. We omit floor and ceiling signs whenever they are not crucial.

2 Preliminaries

An *ordered graph* $G^<$ on n vertices is a graph whose vertices are labeled with $1, \ldots, n$ and ordered by the standard ordering $<$ of integers. An ordered graph $H^<$ is an *ordered subgraph* of an ordered graph $G^<$ if H is a subgraph of $G^<$ and the vertex ordering of H is a suborder of the ordering $G^<$. The *ordered Ramsey number* $r_<(G^<, H^<)$ of ordered graphs $G^<$ and $H^<$ is the smallest positive integer N such that any red-blue-coloring of $K_N^<$ contains either $G^<$ as a red ordered subgraph or $H^<$ as a blue ordered subgraph.

As we mentioned, the study of ordered Ramsey numbers in the two-color setting is now quite an established part of graph Ramsey theory [5, 7, 8, 9, 15, 23, 30], but there are only a few sporadic results in the multicolor setting [7, 19, 21, 22, 26, 27], mostly for specific classes of ordered graphs. Here we initiate a systematic study of multicolor ordered Ramsey numbers for ordered graphs.

For an integer $q \geq 2$ and arbitrary ordered graphs $G_1^<, \ldots, G_q^<$, the *multicolor ordered Ramsey number* $r_<(G_1^<, \ldots, G_q^<)$ is the smallest positive integer N such that for every q-coloring χ of $K_N^<$ with colors from $[q]$ there is an $i \in [q]$ such that χ contains a monochromatic copy of $G_i^<$ in color i as an ordered subgraph. If $G_1^< = \cdots = G_q^<$, then we simply write $r_<(G^<; q)$ and $r_<(G^<)$ instead of $r_<(G_1^<, \ldots, G_q^<)$ and $r_<(G_1^<, G_2^<)$, respectively. Sometimes, to explicitly point out the number of colors, we use the notation $r_<(G_1^<, \ldots, G_q^<; q)$ instead of $r_<(G_1^<, \ldots, G_q^<)$.

Multicolor ordered Ramsey numbers can be bounded from above and from below by the standard multicolor Ramsey numbers. More precisely, for any integer $q \geq 2$ and ordered graphs $G_1^<, \ldots, G_q^<$, we have

$$r(G_1, \ldots, G_q) \leq r_<(G_1^<, \ldots, G_q^<) \leq r(K_{|G_1|}, \ldots, K_{|G_q|}).$$

In particular, it follows that the multicolor ordered Ramsey numbers are always finite and thus well-defined.

Already in the two-color case, there can be a striking difference between the growth rate of ordered Ramsey numbers and the classical Ramsey numbers. This demonstrates already for ordered matchings, where a *matching* is a graph on an even number of vertices where each vertex has a degree equal to one.

Since unordered matching M on n vertices is a graph with maximum degree 1, it follows from a result by Chvátal, Rödl, Szemerédi, and Trotter [12] that we have $r(M) < cn$ for some positive constant c independent on n. In sharp contrast, there are n-vertex ordered matchings whose ordered Ramsey numbers grow superpolynomially in n as shown by Balko, Cibulka, Král, and Kynčl [7] and independently by Conlon, Fox, Lee, and Sudakov [15], who proved that such a superpolynomial lower bound holds even for almost all ordered matchings.

A *random ordered matching* on $[n]$ is an ordered matching selected uniformly at random from all ordered matchings on $[n]$. We say that an event parametrized by n holds *asymptotically almost surely* if its probability goes to 1 with n going to infinity.

Theorem 1 ([15]). *Let $M^<$ be a random ordered matching on n vertices. Then, asymptotically almost surely,*

$$r_<(M^<) \geq n^{\frac{\log n}{20 \log \log n}}.$$

This lower bound is close to the truth as Conlon, Fox, Lee, and Sudakov [15] proved the following upper bound on ordered Ramsey numbers of ordered graphs with a given interval chromatic number and degeneracy. The *interval chromatic number* $\chi_<(G^<)$ of an ordered graph $G^<$ is the minimum number of intervals into which the vertex set of $G^<$ may be partitioned so that no two vertices in the same interval are adjacent. For a positive integer d, a graph G is *d-degenerate* if every subgraph of G contains a vertex of degree at most d. The *degeneracy* of G is the smallest d for which G is d-degenerate.

Theorem 2 ([15]). *Let $G^<$ be an ordered d-degenerate n-vertex graph with interval chromatic number χ, then*

$$r_<(G^<) \leq n^{32d \log \chi}.$$

Note that for any n-vertex ordered matching $M^<$ we get $r_<(M^<) \leq n^{O(\log n)}$, which apart from the $\log \log n$-factor in the exponent matches the lower bound from Theorem 1. Theorem 2 also implies that there are orderings under which the ordered Ramsey numbers of a d-degenerate graph with n vertices are quasi-polynomial in n. For denser graphs with unbounded interval chromatic number, the ordered Ramsey numbers behave more like the usual Ramsey numbers. In particular, the numbers are exactly the same for complete graphs.

If $M^<$ is an ordered matching on n vertices with interval chromatic number χ, then Conlon, Fox, Lee, and Sudakov [15] proved a stronger estimate

$$r_<(M^<) \leq n^{\lceil \log \chi \rceil + 1} \tag{1}$$

than the one obtained from Theorem 2.

An ordered matching $M^<$ with $\chi_<(M^<) = 2$ on $[2n]$ corresponds to the permutation π on $[n]$ if $\{u, v\} \in E(M^<)$ if and only if $\pi(u) = v - n$. A *random ordered matching with interval chromatic number two* on $[2n]$ is an ordered matching corresponding to a permutation on $[n]$ selected uniformly at random.

Theorem 3 ([8]). *There exists a constant $c > 0$ such that if $M^<$ is a random ordered matching on n vertices with interval chromatic number two, then, asymptotically almost surely,*

$$r_<(M^<) \geq \frac{cn^2}{\log^2 n}.$$

Note that all these results are only in the two-color setting. To our knowledge, there are essentially no analogs of these results for multiple colors. The only such result in the multicolor setting was mentioned by Conlon, Fox, Lee, and Sudakov [15] who mentioned without a proof that $r_<(M^<; q) \leq n^{(2\log n)^{q-1}}$ for every $q \geq 2$ and every ordered matching on n vertices.

3 Our results

In this paper we study the multicolor ordered Ramsey numbers. We first give an upper bound on multicolor ordered Ramsey numbers of ordered matchings with interval chromatic number two. In fact, we prove a stronger statement as we give an estimate on the off-diagonal multicolor ordered Ramsey numbers where one of the ordered matchings is replaced by an ordered complete bipartite graph. For positive integers $n_1, \ldots, n_k$, we use $K^<_{n_1,\ldots,n_k}$ to denote the *trivially ordered* complete k-partite graph $K_{n_1,\ldots,n_k}$ where vertices of each part of size n_i form ith interval in the vertex ordering. For positive integers k and n, we use $K^<_n(k)$ to denote the k-partite $K^<_{n,\ldots,n}$.

Proposition 4. *For every integer $q \geq 2$, let $M^<_1, \ldots, M^<_{q-1}$ be ordered matchings on n vertices with interval chromatic number two and let m be a positive integer. Then*

$$r_<(M^<_1, \ldots, M^<_{q-1}, K^<_{m,m}; q) \leq 2m \left(\frac{n}{2}\right)^{q-1}.$$

In the case $M^<_1 = \cdots = M^<_{q-1} = M^<$ and $m = \frac{n}{2}$, we get the following estimate.

Corollary 5. *For every integer $q \geq 2$, let $M^<$ be an ordered matching on n vertices with interval chromatic number two, we have*

$$r_<(M^<; q) \leq 4 \left(\frac{n}{2}\right)^{q}.$$

For the upper bound, we state the following general result which estimates multicolor ordered Ramsey numbers of ordered matchings.

Theorem 6. *For all integers $k \geq 2$, $m \geq 1$, $n \geq 1$, $q \geq 2$ and $\chi \geq 2$ and all ordered matchings $M_1^<, \ldots, M_{q-1}^<$ on n vertices with interval chromatic number χ, we have*

$$r_<(M_1^<, \ldots, M_{q-1}^<, K_m^<(k); q) \leq n^{(q-1)(\lceil \log \chi \rceil)^{q-2} \lceil \log k \rceil} m.$$

Since every ordered matching on n vertices with interval chromatic number χ is an ordered subgraph of $K_n^<(\chi)$, we immediately get the following upper bound.

Corollary 7. *For all integers $q \geq 2$ and $\chi \geq 2$ and every ordered matching $M^<$ on n vertices with interval chromatic number χ, we have*

$$r_<(M^<; q) \leq n^{(q-1)(\lceil \log \chi \rceil)^{q-1}+1}.$$

In the case $\chi = n$, we get a slightly better upper bound than the bound $r_<(M^<; q) \leq n^{(2 \log n)^{q-1}}$ mentioned by Conlon, Fox, Lee, and Sudakov [15].

Extending the idea from the proof of Theorem 1, we state a lower bound on the diagonal multicolor ordered Ramsey numbers of ordered matchings.

Theorem 8. *For every integer $q \geq 2$, let $M^<$ be a random ordered matching on n vertices. Then, asymptotically almost surely*

$$r_<(M^<; q) \geq n^{\lfloor \frac{q}{2} \rfloor \frac{\log n}{20 \log \log n}}.$$

The ordered matchings that provide the quasipolynomial lower bounds in Theorem 8 have large interval chromatic numbers. Thus, to provide lower bounds on multicolor ordered Ramsey numbers of ordered matchings with interval chromatic number two that match the upper bound in Proposition 4, we need another idea. The following result is a general lower bound on multicolor ordered Ramsey numbers of ordered matchings with interval chromatic number two that are sufficiently jumbled. In this proof, we use the key idea from Fox's [20] work on Stanley-Wilf limits.

Theorem 9. *For every integer $q \geq 2$, there exist $c_q > 0$ and $n_q > 0$ such that for every integer $n \geq n_q$ there is an ordered matching $M^<$ on $2n$ vertices with interval chromatic number two satisfying*

$$r_<(M^<; q) \geq c_q \frac{n^{q/4}}{(\log n)^q}.$$

Note that this lower bound matches the upper bound from Corollary 5 up to a constant in the exponent.

Theorem 10. *For every integer $q \geq 2$, there exist $c_q > 0$ and $n_q > 0$ such that for all integers $n \geq n_q$ and $d \geq 1$ there is an ordered d-regular graph $G^<$ on $2n$ vertices with interval chromatic number two satisfying*

$$r_<(G^<; q) \geq c_q \frac{(dn)^{q/4}}{(\log dn)^q}.$$

Unfortunately, in this case we do not quite have a matching upper bound. The best estimate that we could derive is the following one.

Theorem 11. *Let d, n, q be positive integers with $q \geq 2$ and let $G_1^<, \ldots, G_{q-1}^<$ be ordered d-degenerate graphs, each on n vertices with interval chromatic number two, then*

$$r_<(G_1^<, \ldots, G_{q-1}^<, K_{t,t}^<; q) \leq n^{\frac{2(d+1)^{q-1}-2}{d}} t^{(d+1)^{q-1}}.$$

Since any ordered d-degenerate graph $G^<$ on n vertices with interval chromatic number two is an ordered subgraph of $K^<_{n,n}$, we get, by setting $t = n$, the upper bound $r_<(G^<; q) \leq n^{((d+2)(d+1)^{q-1}-2)/d}$.

4 Open problems

We mainly studied multicolor ordered Ramsey numbers for ordered matchings and ordered graphs with interval chromatic number two and bounded degeneracy. We proved the general upper bound $r_<(M^<; q) \leq n^{(2\log n)^{(q-1)}}$ for any ordered matching $M^<$ on n vertices mentioned by Conlon, Fox, Lee, and Sudakov [15]. They believe that a much stronger upper bound should hold and posed the following problem.

Problem 1 ([15])**.** *For any integer $q \geq 3$, there exists a constant c_q such that $r_<(M^<; q) \leq n^{c_q \log n}$ for any ordered matching $M^<$ on n vertices.*

For ordered matchings on n vertices with interval chromatic number two and $q \geq 2$ colors, we showed that $r_<(M^<; q) \in n^{\Theta(q)}$ in Corollary 5 and Theorem 9. It would be interesting to determine the leading constant in the exponent. We remark that it follows from Theorem 3 and from (1) that the constant is 1 for $q = 2$.

Problem 2. *For any integer $q \geq 3$, determine the constant c such that $r_<(M^<; q) \in n^{cq+o(q)}$ for any ordered matching $M^<$ on n vertices with interval chromatic number two.*

We proved some estimates on multicolor ordered Ramsey numbers of ordered graphs with bounded degeneracy d and interval chromatic number two. In Theorem 11 we got the upper bound $r_<(G^<; q) \in n^{O(d^{q-1})}$ for such ordered graphs on n vertices. On the other hand, Theorem 10 shows that $r_<(G^<; q) \in (dn)^{\Omega(q)}$. Note that the gap between the two bounds is enormous. Thus, it would be interesting to improve these bounds.

Problem 3. *Improve the upper and lower bounds for multicolor ordered Ramsey numbers of d-degenerate graphs with interval chromatic number two.*

We also remark that there are no nontrivial bounds on ordered Ramsey numbers for ordered graphs with bounded degeneracy and bounded interval chromatic number larger than two.

Finally, we mention an interesting open problem about off-diagonal multicolor ordered Ramsey numbers. Conlon, Fox, Lee, and Sudakov [15] considered the numbers $r_<(M^<, K_3^<)$, where $M^<$ is an ordered matching on n vertices. Using the *Lovász local lemma* [3], they showed that $r_<(M^<, K_3^<) \in \Omega((n/\log n)^{4/3})$. On the other hand, they mentioned the upper bound $r_<(M^<, K_3^<) \leq r(K_n, K_3) \in O(n^2/\log n)$, based on a result by Ajtai, Komlós, and Szemerédi [1]. It is not clear where the truth should lie, and Conlon, Fox, Lee, and Sudakov [15] posed the following interesting and still open problem that was recently studied [9, 30].

Problem 4 ([15])**.** *Does there exist an $\varepsilon > 0$ such that any ordered matching $M^<$ on n vertices satisfies $r_<(M^<, K_3^<) = O(n^{2-\varepsilon})$?*

We now introduce a multicolor variant of this problem. That is, we are interested in estimates on $r_<(K_3^<, \dots, K_3^<, M^<; q)$, where $M^<$ is an ordered matching on n vertices and $q \geq 3$ is an integer. Note that we trivially have

$$r_<(K_3^<, \dots, K_3^<, M^<; q) \geq r_<(M^<, K_3^<) \in \Omega((n/\log n)^{4/3})$$

and we could not prove a significantly better bound. On the other hand, we have

$$r_<(K_3^<, \dots, K_3^<, M^<; q) \le r(K_3, \dots, K_3, K_n; q) \in \Theta(n^q \text{ poly} \log n)$$

where the bound on $r(K_3, \dots, K_3, K_n; q)$ was proved by Alon and Rödl [2]. We note that their result solved a longstanding conjecture of Erdős and Sós [17]. Observe that the bounds are far apart in the ordered setting. We are not sure where the truth should lie, so we introduce the following problem.

Problem 5. *Improve the upper and lower bounds on off-diagonal multicolor ordered Ramsey numbers* $r_<(K_3^<, \dots, K_3^<, M^<; q)$.

In the similar spirit of the conjecture by Erdős and Sós [17], we, in particular, ask the following question.

Problem 6. *Are there ordered matchings* $M^<$ *on* n *vertices such that*

$$\lim_{n\to\infty} \frac{r_<(K_3^<, K_3^<, M^<)}{r_<(K_3^<, M^<)} = \infty?$$

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

[2] N. Alon and V. Rödl. Sharp bounds for some multicolor Ramsey numbers. *Combinatorica*, 25(2):125–141, 2005.

[3] N. Alon and J. H. Spencer. *The probabilistic method.* Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[4] P. Balister, B. Bollobás, M. Campos, S. Griffiths, E. Hurley, R. Morris, J. Sahasrabudhe, and M. Tiba. Upper bounds for multicolour Ramsey numbers. `http://arXiv.org/abs/2410.17197`, 2024.

[5] M. Balko. Ramsey numbers and monotone colorings. *J. Combin. Theory Ser. A*, 163:34–58, 2019.

[6] M. Balko. A survey on oredered Ramsey numbers, 2024. submitted.

[7] M. Balko, J. Cibulka, K. Král, and J. Kynčl. Ramsey numbers of ordered graphs. *Electron. J. Combin.*, 27(1):Paper No. 1.16, 32, 2020.

[8] M. Balko, V. Jelínek, and P. Valtr. On ordered Ramsey numbers of bounded-degree graphs. *J. Combin. Theory Ser. B*, 134:179–202, 2019.

[9] M. Balko and M. Poljak. On ordered ramsey numbers of matchings versus triangles. *Electron. J. Combin.*, 31(2), 2024.

[10] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal ramsey. *arXiv preprint arXiv:2303.09521*, 2023.

[11] S. A. Choudum and B. Ponnusamy. Ordered Ramsey numbers. *Discrete Math.*, 247(1-3):79–92, 2002.

[12] V. Chvatál, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr. The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B*, 34(3):239–243, 1983.

[13] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. of Math. (2)*, 170(2):941–960, 2009.

[14] D. Conlon and A. Ferber. Lower bounds for multicolor Ramsey numbers. *Adv. Math.*, 378:Paper No. 107528, 5, 2021.

[15] D. Conlon, J. Fox, C. Lee, and B. Sudakov. Ordered Ramsey numbers. *J. Combin. Theory Ser. B*, 122:353–383, 2017.

[16] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[17] P. Erdős and Vera T. Sós. Problems and results on Ramsey-Turán type theorems (preliminary report). In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State Univ., Arcata, Calif., 1979)*, volume XXVI of *Congress. Numer.*, pages 17–23. Utilitas Math., Winnipeg, MB, 1980.

[18] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[19] V. Falgas-Ravry, E. Räty, and I. Tomon. Dedekind's problem in the hypergrid. `http://arXiv.org/abs/2310.12946`, 2023.

[20] J. Fox. Stanley-Wilf limits are typically exponential. *arXiv preprint arXiv:1310.8378*, 2013.

[21] J. Geneson, A. Holmes, X. Liu, D. Neidinger, Y. Pehova, and I. Wass. Ramsey numbers of ordered graphs under graph operations. `http://arXiv.org/abs/1902.00259`, 2019.

[22] A. Girão, B. Janzer, and O. Janzer. Ordered Ramsey numbers of powers of paths. `http://arXiv.org/abs/2401.02360`, 2024.

[23] L. Gishboliner, Z. Jin, and B. Sudakov. Ramsey problems for monotone paths in graphs and hypergraphs. *Combinatorica*, 44(3):509–529, 2024.

[24] H. Lefmann. A note on Ramsey numbers. *Studia Sci. Math. Hungar.*, 22(1-4):445–446, 1987.

[25] K. G. Milans, D. Stolee, and D. B. West. Ordered Ramsey theory and track representations of graphs. *J. Comb.*, 6(4):445–456, 2015.

[26] G. Moshkovitz and A. Shapira. Ramsey theory, integer partitions and a new proof of the Erdős–Szekeres theorem. *Adv. Math.*, 262:1107–1129, 2014.

[27] D. Mubayi and A. Suk. Off-diagonal hypergraph Ramsey numbers. *J. Combin. Theory Ser. B*, 125:168–177, 2017.

[28] D. Neidinger and D. B. West. Ramsey numbers of interval 2-chromatic ordered graphs. *Graphs Combin.*, 35(5):1065–1076, 2019.

[29] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[30] D. Rohatgi. Off-diagonal ordered Ramsey numbers of matchings. *Electron. J. Combin.*, 26(2):Paper No. 2.21, 18, 2019.

[31] W. Sawin. An improved lower bound for multicolor Ramsey numbers and a problem of Erdős. *J. Combin. Theory Ser. A*, 188:Paper No. 105579, 11, 2022.

[32] J. Spencer. Ramsey's theorem—a new lower bound. *J. Combinatorial Theory Ser. A*, 18:108–115, 1975.

[33] Y. Wigderson. An improved lower bound on multicolor Ramsey numbers. *Proc. Amer. Math. Soc.*, 149(6):2371–2374, 2021.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Curves on the torus with prescribed intersections*

(Extended abstract)

Igor Balla† Marek Filakovský‡ Bartłomiej Kielak§ Daniel Král'¶
Niklas Schlomberg‖

Abstract

Aougab and Gaster [Math. Proc. Cambridge Philos. Soc. 174 (2023), 569–584] proved that any set of simple closed curves on the torus, where any two are non-homotopic and intersect at most k times, has a maximum size of $k + O(\sqrt{k}\log k)$. We determine the maximum size of such a set for every k and in particular show that the maximum size never exceeds $k + 6$.

1 Introduction

We study families of simple closed non-homotopy-equivalent curves on a compact surface such that any two have at most k intersections. Despite its fundamental and elementary nature, determining the maximum size of such families has remained an open problem, even in the

*The work of the first, third and fourth authors was supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University and by the project GA24-11098S of the Czech Science Foundation. The first and fourth authors also acknowledge support by the National Science Foundation under Grant No. DMS-1928930 while the authors were in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Spring 2025 semester. The work of the second author was supported by the OP JAK Project MSCAFellow5_MUNI (CZ.02.01.01/00/22_010/0003229).

†Simons Laufer Mathematical Sciences Institute, 17 Gauss Way, Berkeley, CA 94720. E-mail: iballa1990@gmail.com.

‡Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic. E-mail: filakovsky@fi.muni.cz.

§Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, Germany. Previous affiliation: Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic.

¶Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, and Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany. E-mail: daniel.kral@uni-leipzig.de. Previous affiliation: Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic.

‖Research Institute for Discrete Mathematics and Hausdorff Center for Mathematics, University of Bonn. E-mail: schlomberg@or.uni-bonn.de

case of a torus. We compute the maximum size of such a set of curves on the torus for every $k \in \mathbb{N}$.

Formally, a *k-system* on a surface Σ is defined to be a collection of simple closed curves on Σ such that any two are non-homotopic and intersect at most k times. Determining the maximum size of a k-system on Σ, which is denoted by $N(\Sigma, k)$, has been the subject of an intensive line of research for various surfaces Σ and values k [2–4, 9–12, 14]. A priori, it is not clear whether $N(\Sigma, k)$ is even finite; to this end, Juvan, Malnič and Mohar [11] showed that $N(\Sigma, k)$ is finite for every compact surface Σ and every $k \in \mathbb{N}$. Moreover, when Σ is the closed orientable surface of genus g, Greene [10] showed that $N(\Sigma, k) \leq O(g^{k+1} \log g)$ for any fixed $k \in \mathbb{N}$. In our work, we focus on the case when Σ is the torus $\mathbb{T}^2$; this arguably simplest case turns out to have surprising connections to number theory, which are discussed below.

Juvan, Malnič and Mohar [11] showed that $k + 1 \leq N(\mathbb{T}^2, k) \leq 2k + 3$ and noted that the upper bound can be improved to $\frac{3}{2}k + O(1)$. The connection between this problem and number theory was pointed out by Agol [1], who observed that the size of a k-system on the torus is at most one more than the smallest prime larger than k. This implies that $N(\mathbb{T}^2, k)$ is at most $(1 + o(1))k$ and specifically, using the bound on the size of prime gaps by Baker, Harman and Pintz [6], $N(\mathbb{T}^2, k)$ is at most $k + O(k^{21/40})$. Cramér [7] showed that a positive resolution of the Riemann hypothesis would yield that $N(\mathbb{T}^2, k)$ is at most $k + O(\sqrt{k} \log k)$ and formulated a stronger number-theoretic conjecture that would imply an upper bound of $k + O(\log^2 k)$; we refer to [8, 13] for further discussion including the suspicion that Cramér's error term should actually be $O(\log^{2+\varepsilon} k)$.

Very recently, Aougab and Gaster [5] used combinatorial and geometric arguments in conjunction with estimates from analytic number theory to show that $N(\mathbb{T}^2, k)$, i.e., the maximum size of a k-system on the torus, is at most $k + O(\sqrt{k} \log k)$ (note that this matches Cramér's bound, which is conditioned on a positive resolution of the Riemann hypothesis). Aougab and Gaster also noted that they are not aware of any k-system on the torus whose size exceeds $k + 6$. Our main result actually asserts that the bound of $k + 6$ is tight.

We now state our main result, which determines $N(\mathbb{T}^2, k)$ for all $k \in \mathbb{N}$.

Theorem 1. *Let K_0 be the set containing the 59 integers listed in Table 1. For every $k \in \mathbb{N} \setminus K_0$, it holds that*

$$N(\mathbb{T}^2, k) = \begin{cases} k + 4 & \text{if } k \bmod 6 = 2, \\ k + 3 & \text{if } k \bmod 6 \in \{1, 3, 5\}, \text{ and} \\ k + 2 & \text{otherwise.} \end{cases}$$

The values of $N(\mathbb{T}^2, k)$ for $k \in K_0$ are given in Table 1.

Note that there are only four values of k such that $N(\mathbb{T}^2, k) = k+6$, namely $k \in \{24, 48, 120, 168\}$, and only 13 values such that $N(\mathbb{T}^2, k) = k + 5$.

An immediate corollary of Theorem 1 is the following.

Corollary 2. *There exists $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$, $N(\mathbb{T}^2, k) \leq k + 4$.*

We remark that while the proof of Theorem 1 is computer-assisted (as we discuss further), it is possible to prove Corollary 2 without computer assistance.

k	1	2	19	23	24	25	33	34	37	47
$N(\mathbb{T}^2, k)$	3	4	23	27	30	30	37	38	42	51
$N(\mathbb{T}^2, k) - k$	+2	+2	+4	+4	+6	+5	+4	+4	+5	+4
"pattern"	+3	+4	+3	+3	+2	+3	+3	+2	+3	+3

k	48	49	53	54	55	61	62	63	64	76
$N(\mathbb{T}^2, k)$	54	54	57	59	60	65	67	67	68	80
$N(\mathbb{T}^2, k) - k$	+6	+5	+4	+5	+5	+4	+5	+4	+4	+4
"pattern"	+2	+3	+3	+2	+3	+3	+4	+3	+2	+2

k	83	84	85	89	90	94	113	114	115	118
$N(\mathbb{T}^2, k)$	87	89	89	93	94	98	117	119	119	122
$N(\mathbb{T}^2, k) - k$	+4	+5	+4	+4	+4	+4	+4	+5	+4	+4
"pattern"	+3	+2	+3	+3	+2	+2	+3	+2	+3	+2

k	119	120	121	124	127	139	141	142	143	144
$N(\mathbb{T}^2, k)$	123	126	126	128	132	143	145	147	147	149
$N(\mathbb{T}^2, k) - k$	+4	+6	+5	+4	+5	+4	+4	+5	+4	+5
"pattern"	+3	+2	+3	+2	+3	+3	+3	+2	+3	+2

k	145	154	167	168	169	174	184	204	208	214
$N(\mathbb{T}^2, k)$	149	158	171	174	174	178	188	208	212	217
$N(\mathbb{T}^2, k) - k$	+4	+4	+4	+6	+5	+4	+4	+4	+4	+3
"pattern"	+3	+2	+3	+2	+3	+2	+2	+2	+2	+2

k	234	244	264	274	294	304	324	354	384	
$N(\mathbb{T}^2, k)$	238	247	268	277	297	307	327	357	387	
$N(\mathbb{T}^2, k) - k$	+4	+3	+4	+3	+3	+3	+3	+3	+3	
"pattern"	+2	+2	+2	+2	+2	+2	+2	+2	+2	

Table 1: The values of $N(\mathbb{T}^2, k)$ for $k \in K_0$. The values "pattern" are the additive constants based on $k \bmod 6$ given as in Lemma 7, which determines the values of $N(\mathbb{T}^2, k)$ for all sufficiently large k.

2 Overview of the proof

We provide a brief overview of the proof of Theorem 1, which we find quite accessible as it only utilizes widely known tools from combinatorics, discrete optimization, and geometry, together with simple number-theoretic observations.

We view the torus as $\mathbb{R}^2/\mathbb{Z}^2$. We say that a closed curve is non-trivial if it is not homotopy equivalent to a point. Every non-trivial closed curve in the torus is freely homotopic to a curve $C_{m,n}$ for some non-zero $(m,n) \in \mathbb{Z}^2$, where $C_{m,n}$ is the closed curve parameterized as $(m \cdot t \mod 1, n \cdot t \mod 1)$ for $t \in [0,1]$; if the curve is non-self-intersecting, i.e. simple, then m and n are coprime. The minimum number of crossings of closed simple curves freely homotopic to $C_{m,n}$ and $C_{m',n'}$ is equal to $|mn' - m'n|$ (the minimum is attained by the curves $C_{m,n}$ and $C_{m',n'}$ themselves). This leads us to the following definition. For an integer $k \in \mathbb{N}$, a set $Q \subseteq \mathbb{Z}^2$ is *k-nice* if the following holds:

- Q contains non-zero coprime pairs only,
- Q does not contain both (x,y) and $(-x,-y)$ for any $(x,y) \in \mathbb{Z}^2$, and
- $|xy' - x'y| \leq k$ for all (x,y) and (x',y') contained in Q.

Observe that the maximum size of a k-system on the torus is equal to the maximum size of a k-nice set.

We next analyze combinatorial and geometric properties of k-nice sets, starting with the following property.

Lemma 3. *Let $k \in \mathbb{N}$. If Q is a k-nice set such that $y \geq 0$ for every $(x,y) \in Q$, then the area of the convex hull of Q is at most $\pi k/2$.*

To state the second property, we need to introduce additional notation. We start with a definition: the *height* of a set $Q \subseteq \mathbb{Z}^2$ is the maximum $|y|$ such that $(x,y) \in Q$. We fix $\ell \in \mathbb{N}$ and define

$$\rho_\ell = \prod_{\text{primes } p,\, p|\ell} \left(1 - \frac{1}{p}\right),$$

and α_ℓ as

$$\alpha_\ell = \max_{1 \leq a \leq b \leq 2\ell} |\{z, a \leq z \leq b \text{ and } \gcd(z,\ell) = 1\}| - \rho_\ell(b - a + 1).$$

Observe that any set of n consecutive integers contains at most $\rho_\ell n + \alpha_\ell$ elements that are coprime with ℓ. We further consider the following linear program (LP_ℓ) with 2ℓ variables $\sigma_1, \ldots, \sigma_\ell$ and $\tau_1, \ldots, \tau_\ell$ defined as follows:

$$\begin{aligned}
&\text{maximize} && \sum_{i=1}^{\ell} \rho_i(\tau_i - \sigma_i) && \\
&\text{subject to} && \tau_i \geq \sigma_i \geq 0 && \text{for all } 1 \leq i \leq \ell, \text{ and} \\
& && -1 \leq i\tau_j - j\sigma_i \leq 1 && \text{for all } 1 \leq i, j \leq \ell.
\end{aligned}$$

The objective value of the program (LP_ℓ) is denoted by γ_ℓ for $\ell \in \mathbb{N}$. We remark that it can be shown that $\gamma_\ell < 1$ for every $\ell \geq 4$. Finally, define $\beta_0 = 1$ and $\beta_\ell = \beta_{\ell-1} + \alpha_\ell + \rho_\ell$ for $\ell \in \mathbb{N}$.

Lemma 4. *Let $k \in \mathbb{N}$. For every $h \in \{1, \dots, k\}$, every k-nice set $Q \subseteq \mathbb{Z}^2$ with height exactly h has at most $\gamma_h k + \beta_h$ elements.*

Using Lemmas 3 and 4, we can prove the following (as stated, the lemma is proved with computer assistance, but it can be proved without computer assistance with a worse multiplicative constant next to h).

Lemma 5. *For every $h \geq 41020$ and every $k \geq h$, the maximum size of a k-nice set with height h is at most*

$$\frac{3264\pi}{10255} \cdot k + \frac{4946}{3675} \cdot h + 1.$$

Lemma 5 asserts that if the height of a k-nice set Q is small, then the size of Q is less than k (note that $\frac{3264\pi}{10255} < 1$). Fortunately, it is possible to assume that the height of a maximum size k-nice is sublinear in k.

Lemma 6. *For every $k \in \mathbb{N}$, there exists a k-nice set with maximum size that has height at most $\sqrt{2k}$.*

Since $\gamma_\ell < 1$ for every $\ell \geq 4$, we obtain using Lemmas 4, 5 and 6 that for every $k \geq 3225$, there exists a maximum size k-nice set with height at most 3. Such k-nice sets can be directly analyzed.

Lemma 7. *For every $k \geq 3$, the maximum size of a k-nice set of height at most 3 is*

- $k + 4$ *if* $k \bmod 6 = 2$,
- $k + 3$ *if* $k \bmod 6 \in \{1, 3, 5\}$, *and*
- $k + 2$ *otherwise.*

Hence, we are left with determining the sizes of k-nice sets for $k \leq 3224$. With computer assistance, Lemma 6 can be improved as follows.

Lemma 8. *For every $k \in \{2, \dots, 3224\}$, there exists a k-nice set with maximum size that has height at most $\sqrt{4k/3}$.*

Lemma 8 yields that for every $k \geq 1892$, there exists a maximum size k-nice set with height at most 3. Finally, for $k \in \{3, \dots, 1891\}$, we identified additional structural properties that a maximum size k-nice set can be assumed to possess and prepared a recursive program to search for maximum size k-nice sets with these properties. The output of the program gave a proof of Theorem 1.

References

[1] I. Agol. Bounds on exceptional Dehn filling. *Geometry & Topology*, 4(1):431–449, 2000.

[2] T. Aougab. Constructing large k-systems on surfaces. *Topology and its Applications*, 176:1–9, 2014.

[3] T Aougab. Local geometry of the k-curve graph. *Transactions of the American Mathematical Society*, 370(4):2657–2678, 2018.

[4] T. Aougab, I. Biringer, and J. Gaster. Packing curves on surfaces with few intersections. *International Mathematics Research Notices*, 2019(16):5205–5217, 2017.

[5] T. Aougab and J. Gaster. Curves on the torus intersecting at most k times. *Mathematical Proceedings of the Cambridge Philosophical Society*, 174(3):569–584, 2023.

[6] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, II. *Proceedings of the London Mathematical Society*, 83(3):532–562, 11 2001.

[7] H. Cramér. Some theorems concerning prime numbers. *Arkiv för Matematik, Astronomi och Fysik*, 15(5):33, 1921.

[8] A. Granville. Harald Cramér and the distribution of prime numbers. *Scandinavian Actuarial Journal*, 1995(1):12–28, 1995.

[9] Joshua Evan Greene. On curves intersecting at most once, II. *arXiv preprint arXiv:1811.01413*, 2018.

[10] Joshua Evan Greene. On loops intersecting at most once. *Geometric and Functional Analysis*, 29(6):1828–1843, 2019.

[11] M. Juvan, A. Malnič, and B. Mohar. Systems of curves on surfaces. *Journal of Combinatorial Theory, Series B*, 68:7–22, 1996.

[12] J. Malestein, I. Rivin, and L. Theran. Topological designs. *Geometriae Dedicata*, 168(1):221–233, 2014.

[13] T. Oliveira e Silva, S. Herzog, and S. Pardi. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. *Mathematics of Computation*, 83(288):2033–2060, 2014.

[14] Piotr Przytycki. Arcs intersecting at most once. *Geometric and Functional Analysis*, 25:658–670, 2015.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Ramsey problems for tilings in graphs

(Extended abstract)

József Balogh* Andrea Freschi† Andrew Treglown‡

Abstract

Given a graph H, the Ramsey number $R(H)$ is the smallest $n \in \mathbb{N}$ such that every 2-edge-colouring of K_n yields a monochromatic copy of H. We write mH to denote the union of m vertex-disjoint copies of H. These graphs are also known as H-tilings. A famous result of Burr, Erdős and Spencer states that $R(mK_3) = 5m$ for every $m \geq 2$. On the other hand, Moon proved that every 2-edge-coloured K_{3m+2} yields a mK_3 where each copy of K_3 is monochromatic, for every $m \geq 2$. Crucially, in Moon's result, distinct copies of K_3 might receive different colours.

We investigate the analogous questions where the complete host graph is replaced by a graph of large minimum degree. We determine the largest size of a monochromatic K_3-tiling one can guarantee in any 2-edge-coloured graph of large minimum degree. We also determine the (asymptotic) minimum degree threshold for forcing a K_3-tiling covering a prescribed proportion of the vertices in a 2-edge-coloured graph such that every copy of K_3 in the tiling is monochromatic. These results therefore provide dense generalisations of the theorems of Burr–Erdős–Spencer and Moon.

1 Introduction

Ramsey theory is a central research topic in combinatorics. Ramsey's original theorem [17] asserts that for every graph H, there exists an $n \in \mathbb{N}$ such that every 2-edge-colouring of the complete graph K_n on n vertices yields a monochromatic copy of H. We write $R(H)$ to denote the smallest n for which the above holds.

In general, determining $R(H)$ is a very difficult problem and there are relatively few graphs H for which the exact value of $R(H)$ is known. An interesting class of graphs whose

*Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL, USA. E-mail: jobal@illinois.edu. Research supported in part by NSF grants RTG DMS-1937241, FRG DMS-2152488, the Arnold O. Beckman Research Award (UIUC Campus Research Board RB 24012), the Simons Fellowship and the BRIDGE Seed Fund (University of Birmingham and University of Illinois at Urbana–Champaign).

†UN-REN, Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: freschi.andrea@renyi.hu. Research partially supported by ERC Advanced Grants "GeoScape", no. 882971 and "ERMiD", no. 101054936, and the BRIDGE Seed Fund (University of Birmingham and University of Illinois at Urbana–Champaign).

‡School of Mathematics, University of Birmingham, United Kingdom. Email: a.c.treglown@bham.ac.uk. Research supported by EPSRC grant EP/V048287/1 and the BRIDGE Seed Fund (University of Birmingham and University of Illinois at Urbana–Champaign).

Ramsey behaviour is quite well-understood are so-called tilings. For a fixed graph H, an *H-tiling* is a collection of vertex-disjoint copies of H. For $m \in \mathbb{N}$, we write mH to denote an H-tiling consisting of m copies of H. The following result of Burr, Erdős and Spencer [5] determines the exact value of $R(mK_3)$.

Theorem 1.1 (Burr, Erdős and Spencer [5]). *We have $R(mK_3) = 5m$ for every $m \geq 2$.*

More generally, Burr, Erdős and Spencer [5] proved that, for a fixed graph H without isolated vertices, there exist constants c and m_0 such that $R(mH) = (2|H| - \alpha(H))m + c$ provided $m \geq m_0$, where $\alpha(H)$ is the independence number of H. Burr [4], and subsequently Bucić and Sudakov [3], provided methods for computing c exactly. Bucić and Sudakov [3] also obtained the current best bounds for m_0.

Although not a Ramsey-type question in the classical sense, it is also natural to ask how large a complete 2-edge-coloured graph needs to be to ensure there exists an H-tiling of a given size such that every copy of H is monochromatic. Crucially, in this setting, different copies of H in the tiling are allowed to receive different colours. The following result of Moon [15] settles the $H = K_3$ case of this problem (and was later generalised by Burr, Erdős and Spencer [5] to larger cliques).

Theorem 1.2 (Moon [15]). *For every integer $m \geq 2$, every 2-edge-colouring of K_{3m+2} yields a K_3-tiling consisting of m monochromatic copies of K_3. Furthermore, the term $3m+2$ cannot be replaced by a smaller integer.*

Schelp [18] (see also [14]) proposed the study of Ramsey-type questions where the host graph, rather than being complete, can be any graph satisfying a given minimum degree condition. Various results have been proven in this direction, see for example [1, 8, 9, 13]. Motivated by this line of research, in this extended abstract we consider the natural generalisations of the aforementioned classical Ramsey-type results about tilings to the dense setting. The works of Burr–Erdős–Spencer and Moon suggest the following two problems.

Problem 1.3. *Let H be a fixed graph and $n, r, \delta \in \mathbb{N}$. Determine the largest $m \in \mathbb{N}$ such that any r-edge-coloured n-vertex graph G with minimum degree $\delta(G) \geq \delta$ contains a monochromatic copy of mH.*

Problem 1.4. *Let H be a fixed graph and $n, r, \delta \in \mathbb{N}$. Determine the largest $m \in \mathbb{N}$ such that any r-edge-coloured n-vertex graph G with minimum degree $\delta(G) \geq \delta$ contains an H-tiling consisting of m monochromatic copies of H (and distinct copies of H in the tiling might be coloured differently).*

Various special cases of Problems 1.3 and 1.4 have already been considered and resolved. For example, the $r = 1$ case of both Problems 1.3 and 1.4 is equivalent to determining the largest H-tiling one can guarantee in any n-vertex graph G with $\delta(G) \geq \delta$. An H-tiling in a graph G is *perfect* if it contains all the vertices of G. The minimum degree threshold to force a perfect H-tiling in a graph was determined for $H = K_3$ by Corrádi and Hajnal [7], for $H = K_\ell$ (for any $\ell \in \mathbb{N}$) by Hajnal and Szemerédi [10] and for an arbitrary fixed graph H by Kühn and Osthus [12]. Komlós [11] determined (asymptotically) the minimum degree threshold that guarantees the existence of an H-tiling covering a fixed proportion of the vertices of the host graph, provided the proportion is less than 1, for any fixed graph H. Therefore, the $r = 1$ case of Problems 1.3 and 1.4 is (asymptotically) fully understood.

The $H = K_2$ case of both Problems 1.3 and 1.4 has also been resolved. The case $H = K_2$ of Problem 1.4 is equivalent to determining the largest K_2-tiling in a graph with given minimum degree, and thus it is covered by, for example, the Hajnal–Szemerédi theorem. The case $H = K_2$ of Problem 1.3 has a more interesting history. Given graphs $H_1, \dots, H_r$, we write $R_r(H_1, \dots, H_r)$ to denote the smallest integer n such that any r-edge-colouring of K_n using colours $c_1, \dots, c_r$ yields a monochromatic copy of H_i in colour c_i, for some i. Generalising a result of Cockayne and Lorimer [6], Gyárfás and Sárközy [9] determined $R_3(mK_2, mK_2, S_t)$ for all $t, m \in \mathbb{N}$, where S_t is the star on $t+1$ vertices. The connection of this purely Ramsey-type result to Problem 1.3 is that a red/blue/green edge-coloured K_n which does not contain a green monochromatic copy of S_t can be seen as a red/blue edge-coloured n-vertex graph G with $\delta(G) \geq n - t$. Therefore, Gyárfás and Sárközy's result resolves the case $H = K_2$, $r = 2$ of Problem 1.3. Omidi, Raeisi and Rahimi [16] computed $R_r(mK_2, \dots, mK_2, S_t)$ for all $r, t, m \in \mathbb{N}$, thus resolving the case $H = K_2$ of Problem 1.3 in full.

2 Main results

In this extended abstract, our main focus is to study Problems 1.3 and 1.4 when $H = K_3$ and $r = 2$. Observe that the case $\delta \leq 4n/5$ is uninteresting, as one cannot guarantee a single monochromatic copy of K_3. Indeed, consider a 2-edge-coloured K_5 that does not contain a monochromatic copy of K_3 and blow it up to obtain a 2-edge-coloured balanced complete 5-partite graph G on n vertices. Then $\delta(G) = \lfloor 4n/5 \rfloor$ and G does not contain a monochromatic copy of K_3. For Problem 1.3, the following theorem provides an exact answer when δ is a bit larger than $4n/5$ or a bit smaller than $n-1$.

Theorem 2.1. *Let $n \in \mathbb{N}$ and G be a 2-edge-coloured n-vertex graph. Then G contains a monochromatic copy of mK_3 where m is equal to*

(B.1) $\lfloor(\delta(G)+1)/5\rfloor$ *if* $\frac{65n}{66} \leq \delta(G)$,

(B.2) $\lceil(5\delta(G)-4n)/2\rceil$ *if* $\frac{4n}{5} \leq \delta(G) \leq \frac{5n}{6}$.

Furthermore, parts (B.1) and (B.2) are best possible, in the sense that the statement of the theorem does not hold if m is replaced by a larger number.

The analogous case of Problem 1.4 turns out to be much more tractable. The following theorem provides an (asymptotic) resolution for all values of δ.

Theorem 2.2. *Let $n \in \mathbb{N}$ and G be a 2-edge-coloured n-vertex graph. Then there exists a K_3-tiling in G such that every copy of K_3 is monochromatic and the number of copies of K_3 in the tiling is at least*

(M.1) $\lfloor(2\delta(G)-n)/3\rfloor$ *if* $\frac{7n}{8} \leq \delta(G)$,

(M.2) $\lfloor(4\delta(G)-3n)/2\rfloor - o(n)$ *if* $\frac{5n}{6} \leq \delta(G) \leq \frac{7n}{8}$,

(M.3) $5\delta(G)-4n$ *if* $\frac{4n}{5} \leq \delta(G) \leq \frac{5n}{6}$.

Furthermore, parts (M.1) and (M.3) are best possible and part (M.2) is best possible up to the $o(n)$ term.

Note that Theorems 2.1 and 2.2 can be seen as dense generalisations of the results of Burr–Erdős–Spencer and Moon. Indeed, cases (B.1) and (M.1) imply Theorems 1.1 and Theorem 1.2 respectively.

The following constructions show the sharpness of case (B.1) of Theorem 2.1 and all cases of Theorem 2.2. For brevity, we omit the construction for case (B.2) of Theorem 2.1.

Construction for Theorem 2.1 (B.1). Let $n, \delta \in \mathbb{N}$ with $5 \leq \delta \leq n-1$. Let G be the n-vertex graph where all edges are present except for an independent set S of size $n - \delta \geq 1$. In particular, $\delta(G) = \delta$. Pick a partition $V(G) \setminus S = R\dot{\cup}B$ such that $|R| \leq 3\lfloor(\delta+1)/5\rfloor + 2$ and $|B| \leq 2\lfloor(\delta+1)/5\rfloor + 1$. Assign colour red to all edges that either lie in R or are between S and B. Assign colour blue to all edges that either lie in B or are between R and $S \cup B$.

Observe that there is no monochromatic K_3 intersecting S, i.e., every monochromatic copy of K_3 must lie in $R \cup B$. In particular, a red copy of K_3 must lie completely in R, while a blue copy of K_3 must have at least two vertices in B. Therefore, if there is a monochromatic mK_3 in G then $m \leq \max\{\lfloor |R|/3 \rfloor, \lfloor |B|/2 \rfloor\} \leq \lfloor \frac{\delta+1}{5} \rfloor$, as required.

Construction for Theorem 2.2. Let $n, \delta \in \mathbb{N}$ such that $4n/5 \leq \delta \leq n-1$. Let G be the following n-vertex graph. We have a partition $V(G) = V_0\dot{\cup}V_1\dot{\cup}\cdots\dot{\cup}V_5$ where $|V_i| = n - \delta \geq 1$ for every $i \geq 1$ and $|V_0| = 5\delta - 4n \geq 0$. The sets $V_1, \ldots, V_5$ are independent; all other pairs of vertices form an edge. It is easy to check that $\delta(G) = \delta$. Next, assign colours red and blue to the edges of G as follows. The subgraph $G[V_0 \cup V_1, V_2, V_3, V_4, V_5]$ is a blow-up of a red/blue edge-coloured K_5 that does not contain a monochromatic K_3. Without loss of generality, we may assume that the edges between $V_0 \cup V_1$ and $V_2 \cup V_3$ are blue while the edges between V_2 and V_3 are red. Finally, all edges lying in $V_0 \cup V_1$ are red.

By construction, the subgraph $G[V_0 \cup V_1, V_2, V_3, V_4, V_5]$ does not contain a monochromatic copy of K_3. It follows that every monochromatic copy of K_3 contains an edge lying in $V_0 \cup V_1$, and thus it must be red. In particular, every monochromatic copy of K_3 (i) has at least one vertex in V_0 (since V_1 is independent) and (ii) at least two vertices in $V_0 \cup V_1$. Furthermore, we have that (iii) no red monochromatic copy of K_3 intersects $V_2 \cup V_3$. If there are m vertex-disjoint monochromatic copies of K_3 in G, properties (i), (ii) and (iii) imply that $m \leq \min\{|V_0|, |V_0 \cup V_1|/2, (n - |V_2 \cup V_3|)/3\} = \min\{5\delta - 4n, (4\delta - 3n)/2, (2\delta - n)/3\}$.

3 Proof sketch of Theorems 2.1 and 2.2

For the proof of Theorem 2.2, we employ a common strategy for all cases (M.1)–(M.3): we first find many vertex-disjoint (blow-ups of) cliques in the host graph by combining the Hajnal–Szemerédi theorem [10] with the regularity method,[1] and then find monochromatic vertex-disjoint triangles within each such subgraph. This yields a large K_3-tiling where every copy of K_3 is monochromatic. Note also that case (B.2) of Theorem 2.1 follows immediately from case (M.3) of Theorem 2.2 and the pigeonhole principle. We do not provide further details on these arguments, and instead focus on the more subtle proof of case (B.1) of Theorem 2.1.

A *bowtie* consists of two monochromatic copies of K_3 of different colours which share exactly one vertex. The notion of a bowtie in this context was introduced by Burr, Erdős and Spencer [5], and played a crucial role in their proof of Theorem 1.1. The following new result is a key ingredient of the proof of Theorem 2.1 (B.1).

[1] Formally, the latter is only used in case (M2).

Lemma 3.1. *Suppose a 2-edge-coloured K_7 contains a bowtie. Then there exists another bowtie on a different vertex set.*

The proof of case (B.1) of Theorem 2.1 follows a vertex-switching type argument. Given an n-vertex graph G with $\delta(G) \geq \frac{65n}{66}$, we start by selecting a maximum collection $\mathcal{B}$ of vertex-disjoint copies of K_5, each containing a bowtie. Subject to this, we let $\mathcal{T}$ be a maximum collection of monochromatic triangles, all of the same colour, which are vertex-disjoint from each other as well as from the elements of $\mathcal{B}$. For brevity, we consider the case $\mathcal{T} \neq \emptyset$. Observe that G contains a monochromatic copy of $(|\mathcal{B}| + |\mathcal{T}|)K_3$, and so we may assume for a contradiction that $|\mathcal{B}| + |\mathcal{T}| < \lfloor (\delta(G) + 1)/5 \rfloor$. It follows that strictly less than $\delta(G)$ vertices lie in $\mathcal{B} \cup \mathcal{T}$ and thus there exists an edge e which is not incident to any element in $\mathcal{B} \cup \mathcal{T}$.

Let $T \in \mathcal{T}$ and set $X := T$. By using the fact that $\delta(G) \geq \frac{65n}{66}$, one can argue that there is some element B in $\mathcal{B}$ such that both $B \cup \{e\}$ and $B \cup X$ span two complete subgraphs. In particular, $B \cup \{e\}$ spans a copy of K_7 and so by Lemma 3.1 there exists a bowtie B' and a vertex $x \in V(B)$ such that $V(B') \subseteq V(B \cup \{e\})$ and $x \notin V(B')$. We then modify the collection $\mathcal{B}$ by removing B and adding B' in its place. Crucially, after this modification, the vertex x does not belong to $\mathcal{B}$ anymore and it is adjacent to all vertices of X (since $B \cup X$ spans a clique). We then add x to X. By iterating this procedure, we are able to increase the size of X while ensuring it still spans a clique. Note that the collection $\mathcal{T}$ is not affected at all in this process.

Once X reaches a sufficiently large size, by Moon's result (Theorem 1.2) it must contain two disjoint monochromatic triangles. If these have the same colour as the triangles in $\mathcal{T}$, we can add them to $\mathcal{T}$ in place of T, thus contradicting the maximality of $\mathcal{T}$. Otherwise, one can argue that X contains a bowtie, contradicting the maximality of $\mathcal{B}$.

Data availability statement. A full paper containing the proofs of our results can be found on arXiv [2].

References

[1] J. Balogh, J. Barát, D. Gerbner, A. Gyárfás, and G.N. Sárközy, Partitioning 2-edge-colored graphs by monochromatic paths and cycles. *Combinatorica* **34** (2014), 507–526.

[2] J. Balogh, A. Freschi and A. Treglown, Ramsey-type problems for tilings in dense graphs. arXiv preprint: arXiv:2502.13876.

[3] M. Bucić and B. Sudakov, Tight Ramsey bounds for multiple copies of a graph. *Adv. Combin.* (2023), 22pp.

[4] S.A. Burr, On the Ramsey numbers $r(G, nH)$ and $r(nG, nH)$ when n is large. *Discr. Math.* **65** (1987), 215–229.

[5] S.A. Burr, P. Erdős and J.H. Spencer, Ramsey theorems for multiple copies of graphs. *Trans. Amer. Math. Soc.* **209** (1975), 87–99.

[6] E.J. Cockayne and P.J. Lorimer, The Ramsey number for stripes. *J. Aust. Math. Soc.* **19** (1975), 252–256.

[7] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph. *Acta Math. Hungar.* **14** (1963), 423–439.

[8] L. DeBiasio and L.L. Nelsen, Monochromatic cycle partitions of graphs with large minimum degree. *J. Combin. Theory Ser. B* **122** (2017), 634–667.

[9] A. Gyárfás and G.N. Sárközy, Star versus two stripes Ramsey numbers and a conjecture of Schelp. *Combin. Probab. Comput.* **21** (2012), 179–186.

[10] A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős. *Combin. Theory Appl.* **2** (1970), 601–623.

[11] J. Komlós, Tiling Turán theorems. *Combinatorica* **20** (2000), 203–218.

[12] D. Kühn and D. Osthus, The minimum degree threshold for perfect graph packings. *Combinatorica* **29** (2009), 65–107.

[13] S. Letzter, Monochromatic cycle partitions of 2-coloured graphs with minimum degree $3n/4$. *Electron. J. Combin.* **26** (2019), P1.19.

[14] H. Li, V. Nikiforov and R.H. Schelp, A new class of Ramsey–Turán problems, *Discr. Math.* **310** (2010), 3579–3583.

[15] J.W. Moon, Disjoint triangles in chromatic graphs. *Mathematics Magazine* **39** (1966), 259–261.

[16] G.R. Omidi, G. Raeisi and Z. Rahimi, Stars versus stripes Ramsey numbers. *Eur. J. Comb.* **67** (2018), 268–274.

[17] F.P. Ramsey, On a problem of formal logic. *Proc. Lond. Math. Soc.* **30** (1930), 264–286.

[18] R.H. Schelp, Some Ramsey–Turán type problems and related questions. *Discr. Math.* **312** (2012), 2158–2161.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On the number of edges in saturated partial embeddings of maximal planar graphs

(Extended abstract)

János Barát* Zoltán L. Blázsik† Balázs Keszegh‡ Zeyu Zheng§

Abstract

We investigate the extremal properties of saturated partial plane embeddings of maximal planar graphs. For a planar graph G, the plane-saturation number $\mathrm{sat}_{\mathcal{P}}(G)$ denotes the minimum number of edges in a plane subgraph of G such that the addition of any edge either violates planarity or results in a graph that is not a subgraph of G. We focus on maximal planar graphs and establish an upper bound on $\mathrm{sat}_{\mathcal{P}}(G)$ by showing there exists a universal constant $\epsilon > 0$ such that $\mathrm{sat}_{\mathcal{P}}(G) < (3-\epsilon)v(G)$ for any maximal planar graph G with $v(G) \geq 16$. This answers a question posed by Clifton and Simon. Additionally, we derive lower bound results and demonstrate that for maximal planar graphs with sufficiently large number of vertices, the minimum ratio $\mathrm{sat}_{\mathcal{P}}(G)/e(G)$ lies within the interval $(1/16, 1/9 + o(1)]$.

1 Introduction and Preliminaries

Unless explicitly stated otherwise, all graphs considered in this paper are finite, undirected and simple. Recall from fundamental graph theory that a *planar graph* is a graph that admits

*HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary and University of Pannonia, Veszprém, Hungary. E-mail: barat@renyi.hu. Partially supported by ERC Advanced Grant "GeoScape" No.88271. and NKFIH Grant K.131529.

†Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary. University of Johannesburg Auckland Park, 2006 South Africa. E-mail: blazsik@server.math.u-szeged.hu. Was supported by the ÚNKP-23-4-SZTE-628 New National Excellence Program during the first stages of this research work, and currently is supported by the EKÖP-24-4-SZTE-609 Program. Both of these programs belong to the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

‡HUN-REN Alfréd Rényi Institute of Mathematics and ELTE Eötvös Loránd University, Budapest, Hungary. E-mail: keszegh@renyi.hu. Supported by the National Research, Development and Innovation Office – NKFIH under the grant K 132696 and by the ERC Advanced Grant "ERMiD". This research has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the ELTE TKP 2021-NKTA-62 funding scheme.

§Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: zeyuzhen@andrew.cmu.edu. Supported in part by NSF grant DMS-2154063.

a plane embedding, while a *plane graph* refers to a specific embedding of a planar graph in the plane. For simplicity, we often treat a plane graph as both a graph and its corresponding embedding, depending on the context. We denote the vertex set and the edge set of a graph G by $V(G)$ and $E(G)$, respectively, and the number of vertices and edges by $v(G)$ and $e(G)$.

The degree of a vertex v is denoted by $d(v)$, and the number of degree i neighbors of v is denoted by $d_i(v)$. Vertices of degree 3 play an important role in our subsequent discussion. Their count in G is denoted by $n_3(G)$.

A classic problem in extremal graph theory is the graph saturation problem. Given a graph H, one seeks to determine the minimum number of edges in a graph G such that G does not contain H as a subgraph, but the addition of any edge to G (on the same vertex set) results in a copy of H as a subgraph. For a more in-depth overview of the classic saturation problem, we refer the readers to the survey [4].

The study of the plane saturation problem, initiated by Clifton and Salia [2], is a variant of the saturation problems. A plane graph H is a *plane-saturated subgraph* of G, if adding an extra edge to H either violates the planarity of the embedding, or results in a graph that is not a subgraph of G. We define the *plane-saturation number*, $\mathrm{sat}_{\mathcal{P}}(G)$, as the minimum value of $e(H)$, where H is a plane-saturated subgraph of G.

Partially embedded graphs in the plane were investigated before, for instance in [1, 5]. Jelinek et al. characterized minimal non-extendable plane embeddings [5]. However, for them extendability does not mean adding one edge, but adding all missing edges. Checking their obstructions can tell if one can fully extend to the entire planar graph G, while in our case for saturation we just need that for every edge, this particular edge cannot be added. Angelini et al. studied how to minimize the number of edges of the partial embedding that need to be rerouted to extend it to a planar graph, and argued that this problem is NP-hard [1].

Very recently, parallel to our studies, Clifton and Simon investigated the plane saturation problem for labeled and unlabeled maximal planar graphs in [3]. They demonstrated through construction that there exists an infinite family of maximal planar graphs satisfying $\mathrm{sat}_{\mathcal{P}}(G) \geq 3v(G)/2-3$. They also asked whether there exists some ϵ such that $\mathrm{sat}_{\mathcal{P}}(G) < (3-\epsilon)v(G)$ for any maximal planar graph G. The only previously known result was independently noticed by us and the authors of [3].

Proposition 1.1. *Let G be a maximal planar graph on n vertices, with degree sequence $d_1 \leq d_2 \leq \cdots \leq d_n$. Suppose there exist an index $k \in [n-1]$ and a constant $c > 0$ such that $d_{k+1} - d_k \geq cn$. Then,*

$$\mathrm{sat}_{\mathcal{P}}(G) \leq (3-c)n - 2.$$

2 Main result

In this paper, we address this question and provide an affirmative answer to the general case.

Theorem 2.1. *There exists a universal constant $\epsilon > 0$ such that $\mathrm{sat}_{\mathcal{P}}(G) < (3-\epsilon)v(G)$ holds for any maximal planar graph G on at least 16 vertices.*

For lower bound results, Clifton and Salia [2] showed that any planar graph G without degree 1 or degree 2 twins satisfies $\mathrm{sat}_{\mathcal{P}}(G) > e(G)/16$, where twins refer to a pair of vertices that share the same neighborhood. Since maximal planar graphs do not contain degree 1 or degree 2 twins, the following corollary immediately holds.

Corollary 2.2. *For any maximal planar graph G,* $\mathrm{sat}_{\mathcal{P}}(G) > (3v(G) - 6)/16$.

Clifton and Salia also conjectured in [2] that for planar graphs with minimal degree 3, the ratio $\mathrm{sat}_{\mathcal{P}}(G)/e(G)$ is bounded below by 1/9. We present a construction of an infinite family of maximal planar graphs that achieves this ratio asymptotically, as detailed in the following theorem.

Theorem 2.3. *There exists an infinite family of maximal planar graphs such that each graph G in the family satisfies* $\mathrm{sat}_{\mathcal{P}}(G) \leq \frac{v(G)+82}{3}$.

References

[1] Patrizio Angelini, Giuseppe Di Battista, Fabrizio Frati, Vít Jelínek, Jan Kratochvíl, Maurizio Patrignani, and Ignaz Rutter. Testing planarity of partially embedded graphs. *ACM Trans. Algorithms*, 11(4), April 2015.

[2] Alexander Clifton and Nika Salia. Saturated partial embeddings of planar graphs. *arXiv preprint arXiv:2403.02458*, 2024.

[3] Alexander Clifton and Dániel G Simon. Saturated partial embeddings of maximal planar graphs. *arXiv preprint arXiv:2412.06068*, 2024.

[4] Jill R Faudree, Ralph J Faudree, and John R Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, 1000:DS19–Jul, 2011.

[5] Vít Jelínek, Jan Kratochvíl, and Ignaz Rutter. A kuratowski-type theorem for planarity of partially embedded graphs. *Computational Geometry*, 46(4):466–492, 2013.

[6] Douglas B West. *Combinatorial mathematics.* Cambridge University Press, 2021.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Improved Universal Graphs for Trees

(Extended abstract)

Julian Becker* Konstantinos Panagiotou† Matija Pasch‡

Abstract

A graph G is *universal* for a class of graphs $\mathcal{C}$, if, up to isomorphism, G contains every graph in $\mathcal{C}$ as a subgraph. In 1978, Chung and Graham asked for the minimal number $s(n)$ of edges in a graph with n vertices that is universal for all trees with n vertices. The currently best bounds assert that $n \ln n + O(n) \leq s(n) \leq Cn \ln n + O(n)$, where $C = \frac{5}{\ln 4} \approx 3.607$. Here, we improve the upper bound to $cn \ln n + O(n)$, where $c = \frac{19}{6 \ln 3} \approx 2.882$. We develop in the proof a strategy that, broadly speaking, is based on separating trees in *three* parts, thus enabling us to embed them in a structure that originates from ternary trees.

1 Introduction

A graph G is *universal* for a class of graphs, if, up to isomorphism, G contains every graph in the class as a subgraph. The study of universal graphs was initiated by Rado in 1964 [8] and has flourished over the last six decades, see [4, 6, 1] for some recent work and further references. In this note, we discuss universal graphs with n vertices for the class of trees with n vertices, also known as *tree-complete* graphs [9, 7]. For brevity, we refer to these graphs simply as universal graphs.

Chung and Graham asked for the minimum number $s(n)$ of edges in a universal graph in 1978 [2]. In the same paper, by bounding the degree sequence of any such universal graph, they established the lower bound $s(n) \geq \frac{1}{2} n \ln n - O(n)$. Moreover, using a recursive construction and improving over several previous results, they showed that $s(n) \leq \frac{5}{\ln 4} n \ln n + O(n)$ in [3]. In the following forty years, neither of the bounds was improved. In fact, it was deemed possible that the lower bound is tight [3]. However, quite recently, the lower bound was pushed to $n \ln n - O(n)$ by Győri, Li, Salia and Tompkins [5], where, instead of only considering the degree sequence, the authors counted the edges carefully and jointly for all embeddings. Hence, to date, the best bounds are

$$n \ln n - O(n) \leq s(n) \leq \frac{5}{\ln 4} n \ln n + O(n).$$

*Department of Mathematics, LMU Munich, 80333 Munich, Germany. E-mail: becker@math.lmu.de.
†Department of Mathematics, LMU Munich, 80333 Munich, Germany. E-mail: kpanagio@math.lmu.de.
‡Munich, Germany. E-mail: matija.pasch@gmail.com.

Our main result improves the $5/\ln 4 = 3.606\ldots$ in the upper bound to $19/(6\ln 3) = 2.882\ldots$, which narrows the gap between the two bounds.

Theorem 1. *We have* $s(n) \leq \frac{19}{6\ln 3} n \ln n + O(n)$.

In the proof, we develop a strategy for embedding trees iteratively into a host graph based on ternary trees, which, in its very essence, is inspired by the aforementioned [3]. In particular, given a certain tree, we remove carefully up to two vertices to achieve a partition into components of suitable sizes. Crucially, we are able to embed these components recursively, while preserving certain properties of the host graph. As we will see in Section 2.1, we also simplify along these lines the original strategy presented in [3], which is based on binary trees.

2 Proof strategy

We establish Theorem 1 by constructing explicitly universal graphs with $\frac{19}{6\ln 3} n \ln n + O(n)$ edges. As mentioned in the introduction, we use ternary trees as guiding building structures for the universal graph. In order to provide some intuition and to avoid technical details, we explore the construction on binary trees and outline the strategy regarding ternary trees afterwards. Let us introduce some notation first. For $k \geq 2$, we consider the perfect k-ary tree $T_{h,k}$ of height $h \geq 0$ with levels 0 to h on the vertex set

$$V(T_{h,k}) = \bigcup_{0\leq \ell \leq h} \{1, 2, \ldots, k\}^{\ell}.$$

For a given level $0 \leq \ell < h$ and a vertex $v \in \{1, \ldots, k\}^{\ell}$, the *children* of v are $v1, \ldots, vk \in \{1, \ldots, k\}^{\ell+1}$, where we use an economic notation for elements of $\{1, \ldots, k\}^{\ell}$, e.g. $13231 \in \{1, 2, 3\}^5$. Additionally, for every vertex $v \in V(T_{h,k})$, let D_v denote the set of all *descendants* of v, i.e., all vertices in $V(T_{h,k}) \setminus \{v\}$ with prefix v.

For $a \in \mathbb{N}$, define $v{\pm}a$ as the a-th successor/predecessor of v using the lexicographical order imposed on the level of v, and using the rules $\mathrm{succ}(k \ldots k) := 1 \ldots 1$ and $\mathrm{pred}(1 \ldots 1) := k \ldots k$. For example, if $v = k(k-1)$, then $v{+}1 = kk$ and $v{+}2 = 11$, and if $v = kk$, then $v{-}k = (k-1)k$. We compare vertices at different levels lexicographically by attaching 0's to the shorter one, where 0 is the smallest character in the alphabet. Define the relation " $\succ$ " as the reversed lexicographical order, that is for $x, y \in V(T_{h,k})$, $x \neq y$, set $x \succ y$, if x is lexicographically smaller than y. In the proof of the main statement we will embed multiple trees one after another into a graph G' with vertices in $V(T_{h,k})$. In particular, we place the first tree T' at the smallest $|V(T')|$ vertices of G' (with respect to $\succ$). After that, we place the second tree T'' at the smallest $|V(T'')|$ vertices of $G' \setminus T'$, and so forth. This process can be interpreted as the graph G' being eaten up by the trees, hence we call $\succ$ the *eating order*. Figure 1 shows $T_{3,3}$ and the eating order.

With this notation at hand, we construct for each h a graph $T^*_{h,k}$ on $V(T_{h,k})$ by adding for every $v \in V(T_{h,k})$ all edges that contain v and

(Type 1) every vertex in D_v;

(Type 2) $v - 1, \ldots, v - (k-1)$ and every vertex in $D_{v-1} \cup \cdots \cup D_{v-(k-1)}$;

(Type 3) the half (rounded down) of $w \cup D_w$ that is eaten last, where w is the lexicographically smallest child of $v + 1$.

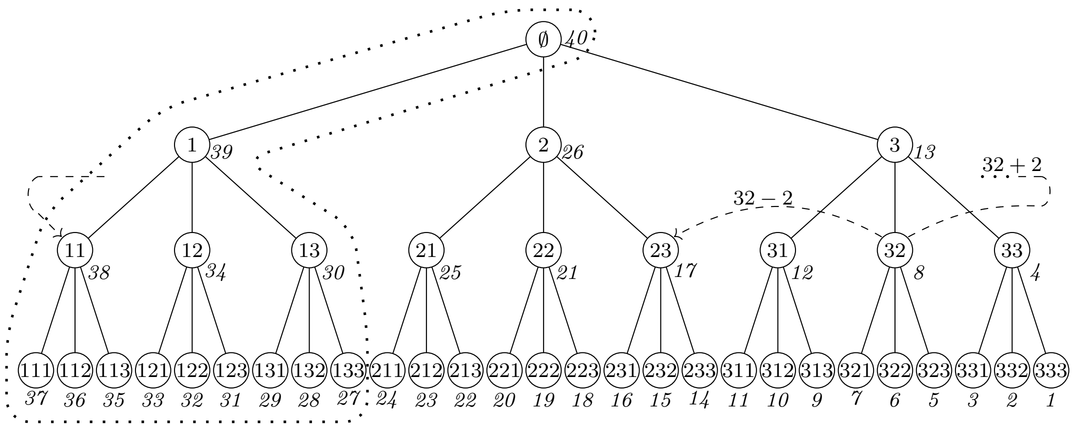


Figure 1: This is the ternary tree $T_{3,3}$ including the eating order. The node labels on level ℓ from $\{1,2,3\}^\ell$ are denoted within the circles, the position of the vertex in the eating order is denoted next to the circle. For example, the vertex with label 32 is the 8-th vertex to be eaten. Further, the arithmetic operations $32+2=11$ and $32-2=23$ on the vertices are indicated by the dashed arrows. The eating order on $V(T_{3,3})$ also determines the vertex sets of the $G_{n,3}$ of height 3. The smallest such graph is $G_{14,3}$ with $V(G_{14,3})$ visualized by the dotted curve (26 vertices have been eaten) and the largest such graph is $G_{40,3}$ with $V(G_{40,3}) = V(T_{3,3})$ (no vertex has been eaten).

These graphs $T^*_{h,k}$ are our basic building blocks for universal graphs. In fact, for an arbitrary $n \in \mathbb{N}$, let $h \geq 0$ be such that $|T^*_{h-1,k}| < n \leq |T^*_{h,k}|$ (where we assume that $T_{-1,k} = \emptyset$), and define $G_{n,k}$ as the induced subgraph of $T^*_{h,k}$ given by the n vertices of $T^*_{h,k}$ that are eaten last. For example, Figure 1 shows the vertex sets of $V(G_{14,3})$ and $V(G_{40,3})$. For both $k=2$ and $k=3$, we show that in this way we indeed obtain universal graphs, see Sections 2.1 and 2.2.

If, for a moment, we ignore the (Type 3) edges, we can readily show by a straightforward, but a bit tedious argument, which is omitted in this extended abstract, that the resulting graphs have $\frac{k}{\ln k} n \ln n + O(n)$ edges. We then also find that $G_{n,3}$ has $(3+\frac{1}{6})\frac{1}{\ln 3} n \ln n + O(n) = \frac{19}{6\ln 3} n \ln n + O(n)$ edges, and since $\frac{19}{6\ln 3} < \frac{k}{\ln k}$ for $k \geq 2$ and $k \neq 3$, it is certainly enough to consider the construction for $k=3$ only. Unfortunately, we did not succeed in creating universal graphs without (Type 3) edges, that is, with only $\frac{3}{\ln 3} n \ln n + O(n)$ edges, as we encountered technical difficulties in the induction step. These additional edges explain the contribution $\frac{1}{6}$ to $3 + \frac{1}{6} = \frac{19}{6}$.

The starting point of the proofs is the following simple and well-known property of separating vertices in trees [3]. Let $|G| = |V(G)|$ be the number of vertices of a graph G.

Lemma 2. *Let $k \in \mathbb{N}_0$ and F be a forest with $|F| \geq k+1$. Then for some vertex s, there is a forest $F' \subset F \setminus s$ such that $k \leq |F'| \leq 2k$.*

Lemma 2 allows to split a forest, by removing one vertex, into two forests containing a suitable number of vertices. This paves the way to construct universal graphs using the binary tree as a base structure and to pursue a divide and conquer approach for the embedding of *any* forest. This is the main approach followed in [3]. We extend this idea by using ternary trees as building blocks for the universal graphs, based on the following enhanced version of Lemma 2.

By a slight abuse of terminology, let (possibly empty) forests $F_1, \dots, F_t$ be a *partition* of a forest F if they are vertex disjoint and $F_1 \cup \dots \cup F_t = F$.

Lemma 3. *Let $0 \leq m$, $2m \leq M$ and F be a forest with at least $M+1$ vertices. Then there exists a vertex $s \in F$ and a partition F_1, F_2, F_3 of $F \setminus s$ such that*

$$m \leq |F_3| \leq M, \quad |F_1| \leq |F| - 1 - M \quad \text{and} \quad |F_2| \leq |F_1|.$$

We use Lemma 3 to iteratively embed forests into certain universal graphs. We also establish that after the embedding, the remainder of these so-called *admissible* graphs, is again admissible.

Definition 4. *Let $k \geq 2$. A non-empty graph A is admissible, if there is a height $h \geq 0$ such that A is the induced subgraph of $T^*_{h,k}$ on the last $|A|$ vertices to be eaten.*

Since we rely on a proof by induction to establish universality, the following equivalent, recursive definition of admissible graphs illustrates the applicability of the induction hypothesis.

Observation 5. *Let $k \geq 2$. The graph A is admissible if it is a subgraph of $T^*_{h,k}$ for some h and if it has one of the following properties.*

1. *A only consists of the root of $T^*_{h,k}$.*
2. *A is recursively constructed, using an admissible subgraph A' of $T^*_{h-1,k}$ (viewed as a subgraph of $T^*_{h,k}$ by adjusting vertex labels), in one of the k possible ways shown in Figure 2.*

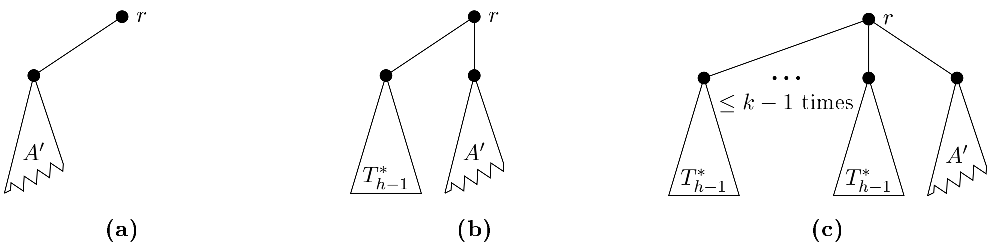


Figure 2: The admissible graph A is recursively constructed using an admissible graph A' of T^*_{h-1}, omitting the subscript k. In a first step, we add 0 to $k-1$ copies of T^*_{h-1} and A' to the root $r = r_A$, so that A' is first in the eating order. In a second step, we add edges as in the construction in Section 2 such that the resulting graph is an induced subgraph of T^*_h.

Note that we will not only apply the induction hypothesis to the intuitive cases in Observation 5. For example, for $k = 3$, the subgraph of $T^*_{3,3}$ induced by the vertices 3, 13, 21, 22 and the descendants of the level 2 vertices is also admissible (cf. Figure 1).

The following result does not only establish that all admissible graphs (and thus in particular the $G_{n,k}$) for $k = 2$ and $k = 3$ are universal, it even ensures the existence of an embedding such that the remainder is still admissible and thereby universal.

Lemma 6. *Let $k \in \{2, 3\}$, G be an admissible graph and F a forest with $|F| < |G|$. Then there exists an embedding $\lambda : V(F) \to V(G)$ of F into G such that $V = \lambda(V(F))$ are the first $|F|$ vertices in the eating order. In particular, $G \setminus V$ is admissible.*

Note that an admissible graph A with n vertices is indeed universal: For any tree T with n vertices and any vertex v in T, we can embed $T \setminus v$ into A using Lemma 6 and then place v at the root r_A, which is connected to *all* of its descendants thanks to the (Type 1) edges.

2.1 $G_{n,2}$ is universal

For $k = 2$, *the (Type 3) edges are not required.* Thus, we disregard these edges in the definition of admissible graphs, the $G_{n,2}$ and in Lemma 6. Since $G_{n,2}$ is admissible, an immediate consequence of Lemma 6 is that $G_{n,2}$ is universal. We proceed to prove Lemma 6 by induction over $|G| \in \mathbb{N}$. Note that G is complete for $h < 2$, so let $h \geq 2$ which implies that we are in the setting of Figure 2. Let F be a forest with $0 < |F| < |G|$, let $V \subset V(G)$ be the first $|F|$ vertices in the eating order and assume that the statement holds for all admissible graphs G' and forests F' with $|F'| < |G'| < |G|$.

First, assume that G is of the type shown in Figure 2a, so that G consists of the root r_G and an admissible subgraph A'. As $|F| < |A'|$, we apply the induction hypothesis to F and A', proving the claim. Otherwise, $|F| = |A'|$. Let $v \in F$ be an arbitrary vertex. We embed $F \setminus v$ into A' using the induction hypothesis and then place v at $r_{A'}$. This yields a proper embedding as discussed below Lemma 6 and leaves us with r_G only.

Let us now assume that G is of the type shown in Figure 2b, so that G consists of the root r_G, one subgraph T^*_{h-1}, and another admissible subgraph A', where A' is first in the eating order. For $|A'| = 1$, we place any vertex v in F at $r_{A'}$ and apply the induction hypothesis to $G \setminus r_{A'}$ and $F \setminus v$. Otherwise, set $N := |T^*_{h-2}| > 0$ and let A'' be the subgraph rooted at the lexicographically largest vertex among the children of $r_{A'}$. If $|F| \leq N + |A''| + 1$, we consider the induced subgraph A given by $r_{A'}$, by A'' and $u = r_{A''} - 1$ with its descendants D_u. Note that the graph A is admissible and that $V \subseteq V(A)$. If the root $r_A = r_{A'}$ is in V, we place an arbitrary vertex v of F at r_A and use the induction hypothesis to embed $F \setminus v$ into A (e.g. like in Figure 3b). If r_A is not in V, we embed F into A (e.g. like in Figure 3a). In any case, we exactly eat V and thus $G \setminus V$ is admissible.

The remaining case is when $|F| > N + |A''| + 1$. Assume that $r_{T^*_{h-1}} \notin V$. Let $R_1 = |A''|$, let $m = \max\{0, |F| - 2N - R_1 - 1\}$ and $M = |F| - N - R_1 - 1$. Note that $2m \leq M$ and thus applying Lemma 3 gives us a separating vertex s and three forests such that

$$m \leq |F_3| \leq M, \quad |F_1| \leq |F| - 1 - M \quad \text{and} \quad |F_2| \leq |F_1|. \tag{1}$$

First, we place s at $r_{A'}$ and embed F_1, F_2, F_3 in that order as follows, see also Figure 3 for a visualization. Here, the conditions from (1) ensure that F_1, F_2, F_3 have the appropriate numbers of vertices. Since $|F_1| \leq N + R_1$, we can embed F_1 using $G_{N+R_1+1,2}$ rooted at $r_{A'}$, leaving an admissible subgraph of size $R_2 + 1 \leq N + 1$. Thus, we embed F_2 using $G_{N+R_2+1,2}$ rooted at $r_{A'}$, leaving an admissible subgraph of size $R_3 + 1 \leq N + 1$. By this procedure, we embed $r_{A'}, F_1, F_2$ at the first $1 + |F_1| + |F_2|$ vertices in the eating order of G. Now, we can embed F_3 in the remaining admissible subgraph rooted at $r_{T^*_{h-1}}$. Therefore, F eats exactly V again. Finally, let $r_{T^*_{h-1}} \in V$. If $r_{A''}$ is the only child of $r_{A'}$, then we proceed almost exactly as above. The difference is that we fix an arbitrary vertex v of F_3, and only embed $F_3 \setminus v$ after F_1 and F_2. We then place v at $r_{T^*_{h-1}}$, similarly to the discussion below Lemma 6. Otherwise, we are exactly in the case of Figure 3a. Here, we have to choose $m = N$ as opposed to $N + 1$, to ensure that $M = 2N + 1 \geq 2N = 2m$, and require $|F_3|$ to be maximal in (1). This implies $|F_3| > m$ and the remainder follows analogously to the previous case.

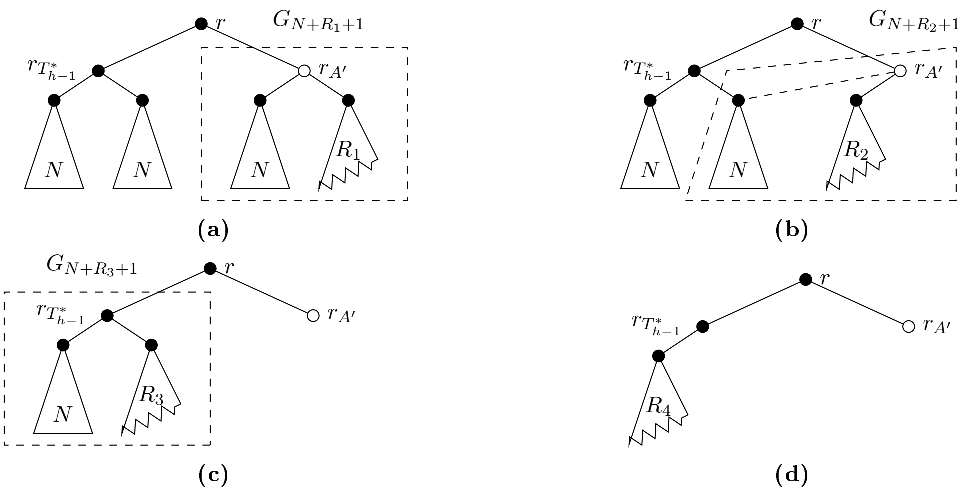


Figure 3: This figure illustrates how we embed up to three forests via the induction hypothesis into $G_{3N+R_1+3,2}$. The separating vertex s is placed at $r_{A'}$, indicated by an unfilled vertex. The dashed region in (a) indicates where we apply the induction hypothesis to embed F_1, leaving an admissible subgraph of size $R_2 + 1$. In (b), the vertices occupied by F_1 are omitted and the dashed region indicates where we apply the hypothesis to embed F_2 (assuming $|F_1| < N + R_1$), leaving an admissible subgraph of size $R_3 + 1$. In (c), the dashed region indicates where we apply the hypothesis to embed F_3 (assuming $|F_1| + |F_2| < 2N + R_1$). It is crucial that each of the dashed regions does not extend over more than two vertices at level 2. Lastly, (d) depicts an admissible subgraph that may remain.

2.2 $G_{n,3}$ is universal

In this section we will omit the subscript $k = 3$ and we will give only the main ideas for establishing the universality of G_n. Let F be a forest with $|F| < |G|$. As in the previous section, we proceed by induction. Also, as for $k = 2$, the base cases and the cases where the induction hypothesis applies directly (e.g. Figure 2a) are immediate. In order to handle the remainder, we shall adapt Lemma 3 in two ways according to the number of vertices in the forest F, which is the case distinction whether we will use one or two separating vertices. Note that we are in a situation as depicted in Figure 2b/2c with root r_G, an admissible subgraph A' (first in the eating order) and up to *two* copies T^R_{h-1} and T^L_{h-1} of T^*_{h-1} (second and third in the eating order). As before, the case $|A'| = 1$ is immediate, so let $|A'| > 1$. Let $N = |T^*_{h-2}| > 0$ and $R = |A''|$, where A'' is the subgraph rooted at the lexicographically largest vertex among the children of $r_{A'}$. We start with the case where we will use one separating vertex, that is, $2N + R + 2 \leq |F| \leq 5N + R + 2$. Again, the idea is to find a separating vertex s, a partition F_1, F_2, F_3 of $F \setminus s$ and to use the embedding strategy discussed in Section 2.1. At this point, it is crucial that none of the embedded F_1, F_2, F_3 extend over more than three of the subgraphs rooted at level 2, which we describe using the following notation. For a forest F^* let $\Delta(F^*) = \Delta_{N,R}(F^*)$ be defined by

$$\Delta(F^*) = N - m \text{ for } |F^*| \geq R \text{ and } \Delta(F^*) = R - m \text{ for } |F^*| < R,$$

where $0 \leq m < N$ is given by the decomposition $|F^*| = m + kN + R$ for $|F^*| \geq R$ and $|F^*| = m$ otherwise. The next result ensures that we can indeed fulfill the size constraints.

Corollary 7. *Let $R > 0$ and $N \geq R$. Let F be a forest with $2N+R+2 \leq |F| \leq 5N+R+2$. Then there exists a vertex $s \in F$ and a partition F_1, F_2, F_3 of $F \setminus s$ such that $F_3 \neq \emptyset$,*

$$|F_1| \leq 2N+R, \quad |F_2| \leq 2N+\Delta(F_1), \quad |F_3| \leq 2N+\Delta(F_1 \cup F_2)+1 \quad \textit{and} \quad |F_1|+|F_2| \geq 2N+R.$$

Notice that for $|F_3| = 2N + \Delta(F_1 \cup F_2) + 1$ we must have $|F| = 5N + R + 2$. As before, in this special case, we will be able to handle the additional vertex. Thus, the first three conditions guarantee that the three forests are not too large. The fourth condition ensures that F_1 and F_2 cover the three subgraphs rooted at level 2 which are eaten first. This is very useful for the case where we need two separating vertices.

In this case, F extends over more than six subgraphs rooted at level 2. The idea is to find a separating vertex s_1 and a partition $F_1, F_2, \overline{F}$ of $F \setminus s_1$. Then, we place s_1 at the lexicographically largest child v of the root (e.g. vertex 3 in Figure 1) and embed (following the eating order) F_1 and F_2. Next, a second separating vertex for $\overline{F}$ is placed at $v-1$ (vertex 2 in Figure 1). Let w be the lexicographically smallest child of v. Recall that there are (Type 3) edges connecting $v-1$ to the half of $w \cup D_w$ eaten last. Thus, in addition to the property that each of the resulting forests does not extend over more than three subgraphs, we also have to ensure that F_1 and F_2 cover at least $\frac{3}{2}N + R$ vertices, i.e. that the second separating vertex connects to all vertices of $\overline{F}$[1].

Corollary 8. *Let $R > 0$ and $N \geq R$. Let F be a forest with $5N+R+3 \leq |F| \leq 8N+R+3$. Let $I_i \in \{0,1\}$ be 1 if $|F| = 8N+R+3$ and $i = 6$. Then there exists a vertex $s_1 \in F$ and a partition $F_1, F_2, F_4, \overline{F}$ of $F \setminus s_1$, a vertex $s_2 \in \overline{F}$ and a partition F_3, F_5, F_6 of $\overline{F} \setminus s_2$ such that $F_6 \neq \emptyset$,*

$$|F_1| \leq 2N+R, \quad |F_1|+|F_2| \geq \frac{3}{2}N+R \quad \textit{and} \quad |F_i| \leq 2N+\Delta\left(\bigcup_{1 \leq j < i} F_j\right)+I_i \quad \textit{for } i > 1.$$

With the two corollaries above it is no longer difficult to show that G_n is universal using almost the same strategy as in Section 2.1 and taking care of some technical details, which we omit here.

Acknowledgments

This project was initiated during a research workshop in Buchboden, June 2024. In particular, we want to thank Angelika Steger for co-organizing this workshop and Kalina Petrova for proposing the problem. Further, we are grateful to Anders Martinsson for valuable discussions.

References

[1] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rodl, A. Ruciński, and E. Szemerédi. Universality and tolerance. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 14–21. IEEE, 2000.

[1] It is convenient to state Corollary 7 for six forests (due to the order on them), but at most five are non-trivial.

[2] F. Chung and R. Graham. On graphs which contain all small trees. *Journal of combinatorial theory, Series B*, 24(1):14–23, 1978.

[3] F. Chung and R. Graham. On universal graphs for spanning trees. *Journal of the London Mathematical Society*, 1983.

[4] A. Ferber, R. Nenadov, and U. Peter. Universality of random graphs and rainbow embedding. *Random Structures & Algorithms*, 48(3):546–564, 2016.

[5] E. Győri, B. Li, N. Salia, and C. Tompkins. A note on universal graphs for spanning trees. *Discrete Applied Mathematics*, 362:146–147, 2025.

[6] D. Johannsen, M. Krivelevich, and W. Samotij. Expanders are universal for the class of all spanning trees. *Combinatorics, Probability and Computing*, 22(2):253–281, 2013.

[7] L. Nebeský. On tree-complete graphs. *Časopis pro pěstování matematiky*, 100(4):334–338, 1975.

[8] R. Rado. Universal graphs and universal functions. *Acta Arithmetica*, 9(4):331–340, 1964.

[9] J. Sedláček. O perfektních a kvaziperfektních grafech. *Časopis pro pěstování matematiky*, 100(2):135–141, 1975.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Borel Local Lemma: Arbitrary Random Variables and Limited Exponential Growth

(Extended abstract)

Anton Bernshteyn* Jing Yu†

Abstract

The Lovász Local Lemma (the LLL for short) is a powerful tool in probabilistic combinatorics that is used to verify the existence of combinatorial objects with desirable properties. Recent years saw the development of various "constructive" versions of the LLL. A major success of this research direction is the Borel version of the LLL due to Csóka, Grabowski, Máthé, Pikhurko, and Tyros, which holds under a subexponential growth assumption. A drawback of their approach is that it only applies when the underlying random variables take values in a finite set. We present an alternative proof of a Borel version of the LLL that holds even if the underlying random variables are continuous and applies to dependency graphs of limited exponential growth.

1 Introduction and main results

Borel combinatorics is an area at the crossroads of combinatorics and descriptive set theory that aims to perform *combinatorial* constructions using *Borel* sets and functions. A recent trend in this area is to design general tools that can be applied to broad problem classes [Ber19; Ber21; Ber23a; Ber23b; BW23; Bra+22; Csó+22; GR21; GR23; QW22; Wei24].

*(AB) Department of Mathematics, University of California, Los Angeles, CA, USA. E-mail: `bernshteyn@math.ucla.edu`. AB's research is partially supported by the NSF CAREER grant DMS-2528522.

†(JY) Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China. E-mail: `jyu@fudan.edu.cn`.

One such tool is the *Lovász Local Lemma* (the *LLL* for short). Here we are interested in the behavior of the LLL in the Borel setting. Specifically, our goal is to generalize and sharpen the Borel LLL due to Csóka, Grabowski, Máthé, Pikhurko, and Tyros [Csó+22]. For our purposes, it will be most convenient to state the LLL using the formalism of *constraint satisfaction problems*.

Definition 1.1 (Constraint satisfaction problems)**.** A *constraint satisfaction problem* (a *CSP* for short) is a tuple $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$, where:

- V, Λ, and $\mathfrak{C}$ are sets whose elements are called *variables*, *labels*, and *constraints* respectively,
- dom assigns to each constraint $\mathfrak{c} \in \mathfrak{C}$ a finite set $\mathsf{dom}(\mathfrak{c}) \subseteq V$ called the *domain* of $\mathfrak{c}$,
- $\mathcal{B}$ assigns to each constraint $\mathfrak{c} \in \mathfrak{C}$ a set $\mathcal{B}(\mathfrak{c}) \subseteq \Lambda^{\mathsf{dom}(\mathfrak{c})}$ of *bad labelings* $\varphi \colon \mathsf{dom}(\mathfrak{c}) \to \Lambda$.

A function $f \colon V \to \Lambda$ is called a *labeling* of V. Such a labeling f *violates* a constraint $\mathfrak{c} \in \mathfrak{C}$ if its restriction to $\mathsf{dom}(\mathfrak{c})$ is in $\mathcal{B}(\mathfrak{c})$; otherwise, f *satisfies* $\mathfrak{c}$. A *solution* to Π is a labeling $f \colon V \to \Lambda$ that satisfies every constraint $\mathfrak{c} \in \mathfrak{C}$. The *dependency graph* of Π, denoted by D_Π, is the graph with vertex set $\mathfrak{C}$ and edge set $\{\{\mathfrak{c}, \mathfrak{c}'\} : \mathfrak{c} \neq \mathfrak{c}', \mathsf{dom}(\mathfrak{c}) \cap \mathsf{dom}(\mathfrak{c}') \neq \varnothing\}$.

Throughout, we shall assume the dependency graph D_Π has finite maximum degree.

Now we are ready to state the Lovász Local Lemma in the CSP framework:

Theorem 1.2 (Lovász Local Lemma [EL75; Spe77; AS16, Corollary 5.1.2])**.** *Let* $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ *be a CSP and let* $\mathbb{P}$ *be a probability measure on* Λ *making* $(\Lambda, \mathbb{P})$ *a standard probability space. If there exist* $p \in [0, 1)$ *and* $d \in \mathbb{N}$ *such that:*

- $\mathcal{B}(\mathfrak{c})$ *is* $\mathbb{P}$*-measurable and* $\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ *for all* $\mathfrak{c} \in \mathfrak{C}$[1] ,
- *the maximum degree of the dependency graph* D_Π *is at most* d*, and*
- $\mathsf{e}\, p\,(d + 1) < 1$ (*where* $\mathsf{e} = 2.71\ldots$ *is the base of the natural logarithm*),

then Π *has a solution* $f \colon V \to \Lambda$.

We are interested in the following general question:

Question 1.3. Let $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ be a CSP such that V and Λ are standard Borel spaces. When can the LLL be invoked to conclude that there exists a *Borel* solution $f \colon V \to \Lambda$ to Π?

All the results of this paper apply to *Borel CSPs*, i.e., CSPs of the form $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ in which V, Λ, and $\mathfrak{C}$ are standard Borel spaces and the assignments dom and $\mathcal{B}$ are Borel in a certain natural sense. Unfortunately, the next example shows the solutions to Borel CSPs produced by the LLL may fail to be Borel.

[1]To unclutter the notation, we shall, when there is no possibility of confusion, use the symbol "$\mathbb{P}$" not only for the given measure on Λ but also for product measures of the form $\mathbb{P}^S$ on Λ^S for some set S.

Example 1.4 (Borel sinkless orientation—a negative result)**.** An orientation of a locally finite graph G is *sinkless* if it has no *sinks*, i.e., no vertices with outdegree 0. The problem of finding a sinkless orientation of G can be encoded by a CSP $\Pi_{\bullet\to}(G)$ as follows. Fix an arbitrary orientation $\vec{G}_0$ of G. We can then identify any orientation $\vec{G}$ of G with a labeling $f\colon E(G) \to \{+,-\}$, where $f(e) = +$ if and only if e is oriented the same in $\vec{G}$ and $\vec{G}_0$. Now we let

$$\Pi_{\bullet\to}(G) \;\coloneqq\; \big(E(G),\, \{+,-\},\, V(G),\, \partial_G,\, \mathcal{B}\big),$$

where for each vertex $v \in V(G)$, $\partial_G(v)$ is the set of all edges of G incident to v and $\mathcal{B}(v)$ contains the unique mapping $\varphi\colon \partial_G(v) \to \{+,-\}$ that makes v a sink. Fix $d \in \mathbb{N}$ and let G be a d-regular graph. It is easy to see that the dependency graph of $\Pi_{\bullet\to}(G)$ is G itself, hence its maximum degree is d. Letting $\mathbb{P}$ be the uniform probability distribution on $\{+,-\}$, we have $\mathbb{P}[\mathcal{B}(v)] = 2^{-d} =: p$ for each $v \in V(G)$, because exactly one out of the 2^d ways to orient the edges incident to v is bad. The inequality $\mathsf{e}\,p\,(d{+}1) \;=\; \mathsf{e}\,2^{-d}\,(d{+}1) \;<\; 1$ is satisfied for all $d \geqslant 4$; moreover, its left-hand side rapidly approaches 0 as d goes to infinity. Nevertheless, no matter how large d is, G may fail to have a Borel sinkless orientation, since Thornton [Tho22, Thm. 3.5], using the determinacy method of Marks [Mar16], constructed for each $d \in \mathbb{N}$ a Borel d-regular graph G with no Borel sinkless orientation.

Nevertheless, the Borel version of the LLL due to Csóka *et al.* [Csó+22] yields a Borel solution to a given CSP Π; to achieve this, they add some extra assumptions on Π, the main one being that the dependency graph D_Π ought to be *of subexponential growth*.

Definition 1.5 (Growth of graphs)**.** The *growth function* $\gamma_G\colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ of a locally finite graph G is given by $\gamma_G(R) \coloneqq \sup_{v\in V(G)} |B_G(v,R)|$, where $B_G(v,R)$ is the *R-ball around v in G*, i.e., the set of all vertices joined to v by a path of at most R edges. The quantity

$$\mathsf{egr}(G) \;\coloneqq\; \lim_{R\to\infty} \sqrt[R]{\gamma_G(R)} \;=\; \inf_{R\geqslant 1} \sqrt[R]{\gamma_G(R)}$$

is called the *exponential growth rate* of G. (The limit exists and is equal to the infimum by Fekete's lemma [SU97, Lem. A.4.2].) We say that G is *of subexponential growth* if $\mathsf{egr}(G) = 1$.

Csóka *et al.* showed that a Borel CSP Π that fulfills the assumptions of the LLL has a Borel solution, provided that its dependency graph is of subexponential growth and, furthermore, the label set Λ and the *order* of Π, $\mathsf{ord}(\Pi) \coloneqq \sup_{\mathfrak{c}\in\mathfrak{C}} |\mathsf{dom}(\mathfrak{c})|$, are finite.

Theorem 1.6 (Csóka–Grabowski–Máthé–Pikhurko–Tyros [Csó+22])**.** *Let* $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ *be a Borel CSP and let* $\mathbb{P}$ *be a Borel probability measure on* Λ*. If there are* $p \in [0,1)$*,* $d \in \mathbb{N}$ *such that:*

(i) $\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ *for all* $\mathfrak{c} \in \mathfrak{C}$*, the maximum degree of* D_Π *is at most* d*, and* $\mathsf{e}\,p\,(d+1) < 1$*,*

(ii) D_Π *is of subexponential growth, and*

(iii) Λ *is a finite set and* $\mathsf{ord}(\Pi) < \infty$*,* *then* Π *has a Borel solution* $f\colon V \to \Lambda$.

While condition (iii) in the above theorem is usually satisfied in combinatorial applications, it is unclear whether its presence is necessary, given that the classical LLL (i.e., Theorem 1.2) is valid even if $(\Lambda, \mathbb{P})$ is a *continuous* probability space. Indeed, the proof of Theorem 1.6 in [Csó+22] uses the finiteness of both Λ and $\mathsf{ord}(\Pi)$ in an essential and apparently unavoidable way. Nevertheless, we develop an alternative proof strategy that completely eliminates condition (iii):

Theorem 1.7. *Let* $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ *be a Borel CSP and let* $\mathbb{P}$ *be a Borel probability measure on* Λ. *Suppose there exist* $p \in [0, 1)$ *and* $d \in \mathbb{N}$ *such that:*

(i) $\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ *for all* $\mathfrak{c} \in \mathfrak{C}$, *the maximum degree of* D_Π *is at most* d, *and* $\mathsf{e}\, p\,(d+1) < 1$,

(ii) D_Π *is of subexponential growth.*

Then Π *has a Borel solution* $f\colon V \to \Lambda$.

Theorem 1.6 is a corollary of [Csó+22, Thm. 4.5], a more general theorem established in [Csó+22], which explicitly invokes the finite values of $|\Lambda|$ and $\mathsf{ord}(\Pi)$. We omit the statement of this theorem due to space constraints. We remark that [Csó+22, Thm. 4.5] holds even if the dependency graph D_Π is *of exponential growth*, as long as its exponential growth rate is sufficiently small. Unfortunately, how small the growth rate has to be significantly depends on $|\Lambda|$ and $\mathsf{ord}(\Pi)$; that is why both Λ and $\mathsf{ord}(\Pi)$ must be finite in Theorem 1.6. In particular, [Csó+22, Thm. 4.5] can only be used when $\mathsf{egr}(D_\Pi) < 2^{2/3} \approx 1.59$. In our main result we generalize it to arbitrary Λ, remove the bound on $\mathsf{ord}(\Pi)$, and make it applicable to dependency graphs whose exponential growth rate may be arbitrarily large (although it must still be bounded in terms of the relationship between p and d):

Theorem 1.8. *Let* $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ *be a Borel CSP and let* $\mathbb{P}$ *be a Borel probability measure on* Λ. *Suppose there exist parameters* $p \in [0, 1)$, $d \in \mathbb{N}$, *and* $s > 1$ *such that:*

(i) $\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ *for all* $\mathfrak{c} \in \mathfrak{C}$ *and the maximum degree of* D_Π *is at most* d,

(ii) $\mathsf{egr}(D_\Pi) < s$, *and*

(iii) $p\,(\mathsf{e}\,(d+1))^s < 1$.

Then Π *has a Borel solution* $f\colon V \to \Lambda$.

Theorem 1.7 is an immediate corollary to Theorem 1.8: if $\mathsf{egr}(D_\Pi) = 1$ and $\mathsf{e}\, p\,(d+1) < 1$, we can find $\varepsilon > 0$ such that $p\,(\mathsf{e}\,(d+1))^{1+\varepsilon} < 1$ and apply Theorem 1.8 with $s = 1 + \varepsilon$.

Example 1.9 (Borel sinkless orientation—a positive result)**.** Consider the sinkless orientation problem $\Pi_{\bullet\to}(G)$ for a d-regular Borel graph G. Since the dependency graph of $\Pi_{\bullet\to}(G)$ is G, by applying Theorem 1.8 with $p = 2^{-d}$, a d-regular Borel graph G has a Borel sinkless orientation provided that $\mathsf{egr}(G) < d/\log_2(e(d+1))$. This extends Thornton's result [Tho22, Thm. 1.5] from graphs of subexponential growth to ones whose exponential growth rate is near-linear in d.

2 Proof ideas

In addition to establishing a more general result, our proof of Theorem 1.8 is also somewhat simpler than the proof of Theorem 1.6 given in [Csó+22], although the two arguments have several common ingredients. The starting point of both approaches is the so-called *Moser–Tardos Algorithm*: a randomized procedure for solving CSPs developed and analyzed by Moser and Tardos in their landmark paper [MT10]. The difficulty in the Borel setting, roughly, is that the Moser–Tardos Algorithm needs a large set of mutually independent random inputs, which cannot be generated in a Borel way. The authors of [Csó+22] deal with this challenge via *randomness conservation*: namely, they reuse the same random input multiple times. Our strategy is to use *probability boosting* instead: thanks to the bound on the exponential growth rate of D_Π, we are able to reduce the given CSP Π to a different CSP Π' whose corresponding value of p is *much* smaller. We then find a Borel solution to Π'—and hence to Π—using the following lemma which is proved via the *method of conditional probabilities*, a standard derandomization technique in computer science [AS16, §16; MR95, §5.6].

Lemma 2.1. *Let $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ be a Borel CSP and let $\mathbb{P}$ be a Borel probability measure on Λ. Suppose there exist parameters $p \in [0, 1)$ and $d \in \mathbb{N}$ such that:*

- *$\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ for all $\mathfrak{c} \in \mathfrak{C}$, the maximum degree of D_Π is at most d, and*

$$p\,(d+1)^{d+1} \;<\; 1.$$

Then Π has a Borel solution $f\colon V \to \Lambda$.

2.1 The Moser–Tardos Algorithm and its local analysis

Throughout §2.1, we fix a CSP $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ and a probability measure $\mathbb{P}$ on Λ making $(\Lambda, \mathbb{P})$ a standard probability space. We also let $D := D_\Pi$ be the dependency graph of Π and assume that D is locally finite. For each $\mathfrak{c} \in \mathfrak{C}$, we write $N[\mathfrak{c}] := B_D(\mathfrak{c}, 1)$ and $N(\mathfrak{c}) := N[\mathfrak{c}]\backslash\{\mathfrak{c}\}$ for the *closed*, resp. *open neighborhood* of $\mathfrak{c}$ in the graph D. Since D is locally finite by assumption, $N[\mathfrak{c}]$ and $N(\mathfrak{c})$ are finite sets.

An important tool in the study of "constructive" aspects of the LLL is the *Moser–Tardos Algorithm* [MT10], (or MTA; see Algorithm 1). MTA takes as input a CSP Π and a map $\tau\colon V \to \Lambda^{\mathbb{N}}$, called a *table* (visualized as a matrix with rows indexed by $\mathbb{N}$, columns indexed by V, and entries from Λ). Given $v \in V$ and $n \in \mathbb{N}$, we write $\tau(v, n)$ for the n-th entry of $\tau(v) \in \Lambda^{\mathbb{N}}$. MTA attempts to solve Π by building a sequence of labelings $f_0, f_1, \ldots : V \to \Lambda$ as follows. Set $f_0(v) := \tau(v, 0)$ for all $v \in V$. At the n-th iteration, let $\mathfrak{C}_n \subseteq \mathfrak{C}$ be the set of all constraints violated by f_n. Pick some D-independent subset $\mathfrak{I}_n \subseteq \mathfrak{C}_n$ and, for each $\mathfrak{c} \in \mathfrak{I}_n$ and $v \in \mathsf{dom}(\mathfrak{c})$, update $f_n(v)$ to be the next value in $\tau(v)$ (tracked by a function $\ell_n\colon V \to \mathbb{N}$). This produces f_{n+1}. Taking a natural limit yields a (possibly partial) labeling f. In our analysis, a sequence $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ of D-independent subsets of $\mathfrak{C}$ is called an MT-*sequence*, and it is said to be *consistent* with (Π, τ) if it can be produced by MTA with input (Π, τ).

Algorithm 1: Moser–Tardos Algorithm (MTA)

Input: A CSP $\Pi = (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B})$ and a table $\tau \colon V \to \Lambda^{\mathbb{N}}$

Initialize $\ell_0(v) := 0$ for all $v \in V$.
for $n = 0, 1, 2, \ldots$ **do**
  Define $f_n(v) := \tau(v, \ell_n(v))$ for all $v \in V$.
  Let $\mathfrak{C}_n := \{\mathfrak{c} \in \mathfrak{C} : f_n \text{ violates } \mathfrak{c}\}$ and pick a D-independent subset $\mathfrak{I}_n \subseteq \mathfrak{C}_n$.
  for $v \in V$ **do**
    if $v \in \mathsf{dom}(\mathfrak{c})$ for some $\mathfrak{c} \in \mathfrak{I}_n$ **then**
      $\ell_{n+1}(v) := \ell_n(v) + 1$
    else
      $\ell_{n+1}(v) := \ell_n(v)$
    end
  end
end
Let $\ell(v) := \lim_{n\to\infty} \ell_n(v) \in \mathbb{N} \cup \{\infty\}$ and define $f(v) := \tau(v, \ell(v))$ whenever $\ell(v) < \infty$.

The *Maximal Moser–Tardos Algorithm*, or MMTA for short, is a variant of MTA with the additional requirement that each set $\mathfrak{I}_n$ be an (*inclusion-*)*maximal* D-independent subset of $\mathfrak{C}_n$ (this is [MT10, Alg. 1.2]). We write $\mathsf{MTA}(\Pi, \tau)$ (resp. $\mathsf{MMTA}(\Pi, \tau)$) for the set of all partial labelings f that can be generated by MTA (resp. MMTA) on input (Π, τ). By definition, $\mathsf{MMTA}(\Pi, \tau) \subseteq \mathsf{MTA}(\Pi, \tau)$.

Lemma 2.2. *Suppose $\mathbb{P}[\mathcal{B}(\mathfrak{c})] < 1$ for all $\mathfrak{c} \in \mathfrak{C}$ and let τ be a table. If a labeling $f \in \mathsf{MMTA}(\Pi, \tau)$ is defined on all of V, then f is a solution to Π.*

It follows from Lemma 2.2 that to solve Π, it is enough to find a table τ such that some labeling $f \in \mathsf{MMTA}(\Pi, \tau)$ is defined everywhere. We shall seek a table with the following stronger property:

Definition 2.3 (Good tables)**.** A table τ is *good* if every $f \in \mathsf{MTA}(\Pi, \tau)$ is defined on all of V.

The following is the main result of [MT10]:

Theorem 2.4 (Moser–Tardos [MT10, Thm. 1.2])**.** *Let p and d be as in Theorem 1.2 (i.e., the LLL). Sample a random table $\tau \colon V \to \Lambda^{\mathbb{N}}$ from the product space $(\Lambda^{V\times\mathbb{N}}, \mathbb{P}^{V\times\mathbb{N}})$. Then for each $v \in V$,*

$$\mathbb{P}[\text{every } f \in \mathsf{MTA}(\Pi, \tau) \text{ is defined on } v] = 1.$$

A random table would typically not be a Borel function. Our approach instead is to treat the problem of finding a good table $\tau \colon V \to \Lambda^{\mathbb{N}}$ as *a CSP in its own right.* The idea is to "localize" the notion of goodness by "zooming in" on r-balls around individual constraints in D. To this end, given $\mathfrak{c} \in \mathfrak{C}$ and $r \in \mathbb{N}$, we define the $(\mathfrak{c}, r)$*-local CSP* $\Pi_{\mathfrak{c},r} := (V, \Lambda, \mathfrak{C}, \mathsf{dom}, \mathcal{B}_{\mathfrak{c},r})$ by letting, for all $\mathfrak{a} \in \mathfrak{C}$,

$$\mathcal{B}_{\mathfrak{c},r}(\mathfrak{a}) := \begin{cases} \mathcal{B}(\mathfrak{a}) & \text{if } \mathfrak{a} \in B_D(\mathfrak{c}, r), \\ \Lambda^{\mathsf{dom}(\mathfrak{a})} & \text{otherwise.} \end{cases}$$

It is clear that, unless $B_D(\mathfrak{c}, r) = \mathfrak{C}$, the CSP $\Pi_{\mathfrak{c},r}$ has no solutions, because the constraints in $\mathfrak{C}\backslash B_D(\mathfrak{c}, r)$ are, by definition, always violated. In particular, the Moser–Tardos Algorithm must fail to generate a solution to $\Pi_{\mathfrak{c},r}$. However, we may ask whether the constraints in $\mathfrak{C}\backslash B_D(\mathfrak{c}, r)$ are in some sense the "main reason" for this failure. Namely, if $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ is an MT-sequence consistent with $(\Pi_{\mathfrak{c},r}, \tau)$, must it be that "many" constraints in the sets $\mathfrak{I}_n$ come from outside $B_D(\mathfrak{c}, r)$? In order to quantify the word "many" in the preceding sentence, we restrict our attention to *finite* MT-sequences, i.e., MT-sequences $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ such that $\sum_{n\in\mathbb{N}} |\mathfrak{I}_n| < \infty$. In other words, $\mathfrak{I}$ is finite if each $\mathfrak{I}_n$ is a finite set and $\mathfrak{I}_n = \varnothing$ for all large enough $n \in \mathbb{N}$. Now, if $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ is a finite MT-sequence consistent with $(\Pi_{\mathfrak{c},r}, \tau)$, we can consider the sets $\mathfrak{I}_n \cap B_D(\mathfrak{c}, r)$ and $\mathfrak{I}_n\backslash B_D(\mathfrak{c}, r)$ for each $n \in \mathbb{N}$, and we want the former ones to not be much larger than the latter. Formally, we say a finite MT-sequence $\mathfrak{I} = (\mathfrak{I}_n)_{n\in\mathbb{N}}$ is $(\mathfrak{c}, r, N, \varepsilon)$-*Følner* for $\mathfrak{c} \in \mathfrak{C}$, r, $N \in \mathbb{N}$, and $\varepsilon \in (0, 1)$ if

$$\sum_{n\in\mathbb{N}} |\mathfrak{I}_n \cap B_D(\mathfrak{c}, r)| \geqslant N \qquad \text{and} \qquad \sum_{n\in\mathbb{N}} |\mathfrak{I}_n\backslash B_D(\mathfrak{c}, r)| < \varepsilon \sum_{n\in\mathbb{N}} |\mathfrak{I}_n|.$$

Definition 2.5 (Locally good tables)**.** Fix $\mathfrak{c} \in \mathfrak{C}$, R, $N \in \mathbb{N}$, and $\varepsilon \in (0, 1)$. A table $\tau \colon V \to \Lambda^{\mathbb{N}}$ is $(\mathfrak{c}, R, N, \varepsilon)$-*locally good* if for all $0 \leqslant r < R$, there is no $(\mathfrak{c}, r, N, \varepsilon)$-Følner MT-sequence consistent with $(\Pi_{\mathfrak{c},r}, \tau)$. A table τ is (R, N, ε)-*locally good* if it is $(\mathfrak{c}, R, N, \varepsilon)$-locally good for all $\mathfrak{c} \in \mathfrak{C}$.

The following lemma is the only place where we use a bound on $\mathsf{egr}(D)$:

Lemma 2.6. *Let $R \in \mathbb{N}$ and $\varepsilon \in (0, 1)$ be such that $\gamma_D(R) < (1 - \varepsilon)^{-R}$, where γ_D is the growth function of D. If a table $\tau \colon V \to \Lambda^{\mathbb{N}}$ is (R, N, ε)-locally good for some $N \in \mathbb{N}$, then τ is good.*

Next we construct a CSP $\mathsf{LG}(R, N, \varepsilon)$ with label set $\Lambda^{\mathbb{N}}$ and constraint set $\mathcal{L}_{R,N,\varepsilon}$ such that a table $\tau \colon V \to \Lambda^{\mathbb{N}}$ is a solution to $\mathsf{LG}(R, N, \varepsilon)$ if and only if τ is (R, N, ε)-locally good.

Then we proceed to bound the maximum degree of the dependency graph of $\mathsf{LG}(R, N, \varepsilon)$:

Lemma 2.7. *For any R, $N \in \mathbb{N}$ and $\varepsilon > 0$, the maximum degree of the dependency graph of $\mathsf{LG}(R, N, \varepsilon)$ is at most $\gamma_D(2R) - 1$.*

The only consequence of Lemma 2.7 we need is that the maximum degree of the dependency graph of $\mathsf{LG}(R, N, \varepsilon)$ can be bounded by a function of R *independent of* N. By contrast, we show that $\mathbb{P}[\mathcal{L}_{R,N,\varepsilon}(\mathfrak{c})]$ is bounded above by a quantity that *goes to* 0 *as* $N \to \infty$:

Lemma 2.8. *Suppose there exist $p \in [0, 1)$, $d \in \mathbb{N}$, $s > 1$, and ε, $\eta \in (0, 1)$ such that:*

(i) *$\mathbb{P}[\mathcal{B}(\mathfrak{c})] \leqslant p$ for all $\mathfrak{c} \in \mathfrak{C}$, the maximum degree of D is at most d, and $p\,(\mathsf{e}\,(d+1))^s < 1$,*

(ii) *$\varepsilon + \dfrac{1}{s} < 1$ and $p^{1-\varepsilon-\frac{1}{s}} \leqslant \dfrac{1-\eta}{1+\eta}$.*

Then for all $\mathfrak{c} \in \mathfrak{C}$ and R, $N \in \mathbb{N}$, $\mathbb{P}[\mathcal{L}_{R,N,\varepsilon}(\mathfrak{c})] \leqslant (1 + \eta)^{-N} F(d, \eta, R)$, where $F(d, \eta, R)$ is independent of N.

2.2 Proof of Theorem 1.8

We start by observing that the Moser–Tardos Algorithm can be used in the Borel setting:

Claim 2.9. *If* $\tau \colon V \to \Lambda^{\mathbb{N}}$ *is a Borel good table, then* Π *has a Borel solution* $f \in \mathsf{MMTA}(\Pi, \tau)$.

Next we show that the CSP $\mathsf{LG}(R, N, \varepsilon)$ is Borel:

Claim 2.10. *For each* R, $N \in \mathbb{N}$ *and* $\varepsilon \in (0, 1)$, *the CSP* $\mathsf{LG}(R, N, \varepsilon)$ *is Borel.*

Take $\varepsilon \in (0, 1)$ such that $\mathsf{egr}(D) < (1 - \varepsilon)^{-1} < s$. By the definition of the exponential growth rate, there is $R \in \mathbb{N}$ such that $\gamma_D(R) < (1-\varepsilon)^{-R}$. On the other hand, the inequality $(1 - \varepsilon)^{-1} < s$ is equivalent to $\varepsilon + 1/s < 1$, and we can take $\eta \in (0, 1)$ so small that $\frac{1-\eta}{1+\eta} > p^{1-\varepsilon-\frac{1}{s}}$. Finally, we take $N \in \mathbb{N}$ so large that $(1 + \eta)^N > F(d, \eta, R)\, \gamma_D(2R)^{\gamma_D(2R)}$, where $F(d, \eta, R)$ is the function from Lemma 2.8. Claim 2.10 and Lemmas 2.7 and 2.8 imply that the CSP $\mathsf{LG}(R, N, \varepsilon)$ satisfies the assumptions of Lemma 2.1. Therefore, $\mathsf{LG}(R, N, \varepsilon)$ has a Borel solution $\tau \colon V \to \Lambda^{\mathbb{N}}$. In other words, τ is a Borel (R, N, ε)-locally good table. By Lemma 2.6, τ is good, so, by Claim 2.9, there exists a Borel solution to Π, and we are done.

References

[AS16] N. Alon and J.H. Spencer. *The Probabilistic Method.* 4th ed. John Wiley & Sons, 2016 (cit. on p.)

[Ber19] A. Bernshteyn. *Measurable versions of the Lovász Local Lemma and measurable graph colorings*, Adv. Math., **353** (2019), 153–223 (cit. on p.)

[Ber21] A. Bernshteyn. *On Baire measurable colorings of group actions*, Ergod. Theory Dyn. Syst., **41** (3) (2021), 818–845 (cit. on p.)

[Ber23a] A. Bernshteyn. *Distributed algorithms, the Lovász Local Lemma, and descriptive combinatorics*, Invent. Math., **233** (2023), 495–542 (cit. on p.)

[Ber23b] A. Bernshteyn. *Probabilistic constructions in continuous combinatorics and a bridge to distributed algorithms*, Adv. Math., **415** (2023), 108895 (cit. on p.)

[BW23] A. Bernshteyn and F. Weilacher. *Borel versions of the Local Lemma and* LOCAL *algorithms for graphs of finite asymptotic separation index*, `https://arxiv.org/abs/2308.14941` (preprint), 2023. To appear in Trans. Amer. Math. Soc. (Cit. on p.)

[Bra+22] S. Brandt, Y.-J. Chang, J. Grebík, C. Grunau, V. Rozhoň, and Z. Vidnyánszky. *Local problems on trees from the perspectives of distributed algorithms, finitary factors, and descriptive combinatorics*, Innovations in Theoretical Computer Science Conference (ITCS), (2022), 29:1–29:26. Full version: `https://arxiv.org/abs/2106.02066` (cit. on p.)

[Csó+22] E. Csóka, Ł. Grabowski, A. Máthé, O. Pikhurko, and K. Tyros. *Moser–Tardos Algorithm with small number of random bits*, `https://arxiv.org/abs/2203.05888` (preprint), 2022 (cit. on p.)

[EL75] P. Erdős and L. Lovász. *Problems and results on 3-chromatic hypergraphs and some related questions*, Infinite and Finite Sets, Colloq. Math. Soc. J. Bolyai, (1975). Ed. by A. Hajnal, R. Rado, and V.T. Sós, 609–627 (cit. on p.)

[GR21] J. Grebík and V. Rozhoň. *Classification of local problems on paths from the perspective of descriptive combinatorics*, Extended Abstracts EuroComb 2021: European Conference on Combinatorics, Graph Theory and Applications, (2021), 553–559. Full version: `https://arxiv.org/abs/2103.14112` (cit. on p.)

[GR23] J. Grebík and V. Rozhoň. *Local problems on grids from the perspective of distributed algorithms, finitary factors, and descriptive combinatorics*, Adv. Math., **431** (2023), 109241 (cit. on p.)

[Mar16] A. Marks. *A determinacy approach to Borel combinatorics*, J. Am. Math. Soc., **29** (2) (2016), 579–600 (cit. on p.)

[MT10] R. Moser and G. Tardos. *A constructive proof of the general Lovász Local Lemma*, J. ACM, **57** (2) (2010) (cit. on p.)

[MR95] R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995 (cit. on p.)

[QW22] L. Qian and F. Weilacher. *Descriptive combinatorics, computable combinatorics, and ASI algorithms*, `https://arxiv.org/abs/2206.08426` (preprint), 2022 (cit. on p.)

[SU97] E.R. Scheinerman and D.H. Ullman. *Fractional Graph Theory.* John Wiley & Sons, 1997. URL: `https://www.ams.jhu.edu/ers/books/fractional-graph-theory-a-rational-approach-to-the-theory-of-graphs/` (cit. on p.)

[Spe77] J. Spencer. *Asymptotic lower bounds for Ramsey functions*, Discrete Math., **20** (1977), 69–76 (cit. on p.)

[Tho22] R. Thornton. *Orienting Borel graphs*, Proc. Am. Math. Soc., **150** (4) (2022), 1779–1793 (cit. on p.)

[Wei24] F. Weilacher. *Computable vs descriptive combinatorics of local problems on trees*, J. Symb. Log., **89** (4) (2024), 1835–1849 (cit. on p.)

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Determining a graph from its reconfiguration graph

(Extended abstract)

Gaétan Berthe * Caroline Brosse † Brian Hearn ‡ Jan van den Heuvel ‡
Pierre Hoppenot § Théo Pierron ¶

Abstract

Given a graph G and a natural number k, the k-recolouring graph $\mathcal{C}_k(G)$ is the graph whose vertices are the k-colourings of G and whose edges link pairs of colourings which differ at exactly one vertex of G. Recently, Hogan et al. proved that G can be determined from $\mathcal{C}_k(G)$ provided k is large enough (quadratic in the number of vertices of G). We improve this bound by showing that $k = \chi(G) + 1$ colours suffice, and provide examples of families of graphs for which $k = \chi(G)$ colours do not suffice.

We then extend this result to k-Kempe-recolouring graphs, whose vertices are again the k-colourings of a graph G and whose edges link pairs of colourings which differ by swapping the two colours in a connected component induced by selecting those two colours. We show that $k = \chi(G) + 2$ colours suffice to determine G in this case.

Finally, we investigate the case of independent set reconfiguration, proving that in only a few trivial cases is one guaranteed to be able to determine a graph G.

1 Single-vertex recolouring

This extended abstract gives an overview of results from our preprint [2].

Let G be a graph and k a positive integer. A (proper) *k-colouring* of G is an assignment of colours from $[k] := \{1, 2, \ldots, k\}$ to the vertices of G such that no adjacent vertices receive the same colour. The *chromatic number* of G, denoted by $\chi(G)$, is the smallest k for which G admits a k-colouring.

* LIRMM, Université de Montpellier, CNRS, Montpellier, France. E-mail: gaetan.berthe@lirmm.fr.
† Université d'Orléans, INSA CVL, LIFO, UR 4022, Orléans, France. E-mail: caroline.brosse@univ-orleans.fr.
‡ Department of Mathematics, London School of Economics & Political Science, Houghton Street, London WC2A 2AE, UK. E-mail: b.hearn@lse.ac.uk, j.van-den-heuvel@lse.ac.uk.
§ Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France. E-mail: pierre.hoppenot@grenoble-inp.fr.
¶ Univ. Lyon, UCBL, CNRS, INSA Lyon, LIRIS, UMR5205, F-69622 Villeurbanne, France. E-mail: theo.pierron@univ-lyon1.fr.

The *k-recolouring graph* $\mathcal{C}_k(G)$ is the graph whose vertices are all the k-colourings of G and whose edges link pairs of k-colourings which differ at exactly one vertex. In this abstract we will make no distinction between a vertex of $\mathcal{C}_k(G)$ and the colouring of G it represents.

Recolouring graphs have been the object of a growing amount of research. One of the first aspects that was studied is their use as a tool for generating k-colourings (almost) uniformly at random. This is done by starting from an initial colouring, and then repeatedly uniformly at random selecting a vertex and trying to recolour it with a uniform at random chosen colour from $[k]$. This process corresponds to following a random walk in $\mathcal{C}_k(G)$. For this process to yield an (almost) uniform distribution within a polynomial number of steps, it is necessary (but not sufficient) for $\mathcal{C}_k(G)$ to be connected and to have polynomial diameter. Therefore, how these two parameters evolve with respect to the number k of allowed colours has been the starting point of quite some research.

Since the question of whether a reconfiguration graph is connected or not is essential for many applications, much research has been done also on the computational complexity of this question for different reconfiguration graphs. Many decision problems involving of this type are now known to be algorithmically hard (often PSPACE-complete). We refer the interested reader to the surveys [3, 5].

Other *reconfiguration graphs* have been studied as well; some of them a long time ago. Reidemeister moves of knot diagrams can be seen as providing edges between different representations of knots. And solving a Rubik's cube consists of finding a path between two configurations in some reconfiguration graph.

By definition, giving G and k completely determines $\mathcal{C}_k(G)$. In early 2024, Asgarli et al. [1] asked if the reverse can also be true. More explicitly, they conjectured that every graph G is uniquely determined by the collection of recolouring graphs $\mathcal{C}_k(G)$, $k \in \mathbb{Z}^+$. Moreover, they conjectured that every graph G is in fact uniquely determined by some finite subcollection of these recolouring graphs. Shortly afterwards, these conjectures were answered in a strong sense by Hogan et al. [4], who proved that a single recolouring graph $\mathcal{C}_k(G)$ suffices to determine G, provided k is large enough.

Theorem 1.1 (Hogan et al. [4])**.**
Let G be an n-vertex graph. If $k > 5n^2$, then G can be determined from the recolouring graph $\mathcal{C}_k(G)$, assuming we know k.

Our first contribution is to improve Theorem 1.1 by first showing that a much smaller k suffices, and by removing the assumption that the value of k is known in advance.

Theorem 1.2.
Let G be a graph. If $k > \chi(G)$, then G can be determined from $\mathcal{C}_k(G)$, even without knowing the exact value of k.

Similarly to the result in [4], the proof of Theorem 1.2 provides a recovery algorithm which runs in polynomial time with respect to the size of $\mathcal{C}_k(G)$ (which in general is exponential in the size of G). However, we can prove that by using more colours, we can improve this recovery algorithm to run in polynomial time with respect to $|V(G)|$.

Theorem 1.3.
Let G be an n-vertex graph of maximum degree Δ. If $k > \min\{n, 2\Delta\}$, then G can be determined from $\mathcal{C}_k(G)$ by considering at most $(kn)^2$ different colourings in $\mathcal{C}_k(G)$, even without knowing the exact value of k.

The proofs of Theorems 1.2 and 1.3 for the most part utilise observations that were already present in the proof of Theorem 1.1 in [4]. However the authors in [4] use a counting argument, whereas we utilise further structural observations.

The bound on the number of colours in Theorem 1.2 is tight; a graph G is not necessarily uniquely determined by its k-recolouring graph $\mathcal{C}_k(G)$ if $k = \chi(G)$. For example, the 2-colouring graph of any connected bipartite graph is the graph with two vertices and no edges. And for any $\chi \geq 2$, the χ-recolouring graph of any $(\chi-1)$-tree (any graph formed by starting with the complete graph on χ vertices, adding a vertex v such that the neighbours of v form a clique on $\chi - 1$ vertices, and repeating this process finitely many times) is the graph with $\chi!$ vertices and no edges.

These examples exploit the fact that some vertices of the graph G are *frozen*, i.e. they cannot be recoloured under any $\chi(G)$-colouring of G. If a graph G contains any frozen vertices, then its recolouring graph is non-unique. To see this, add a new vertex to G with the same set of neighbours as some frozen vertex v. This new graph will have the same $\chi(G)$-recolouring graph as G.

However, we will show that the existence of frozen vertices is not necessary for a graph G to have a non-unique $\chi(G)$-recolouring graph.

Theorem 1.4.
For every $\chi \geq 6$ there exists an arbitrarily large family $\mathcal{F}$ of non-isomorphic graphs, all with chromatic number χ and without frozen vertices, such that $\mathcal{C}_\chi(G) \cong \mathcal{C}_\chi(G')$ for all $G, G' \in \mathcal{F}$.

Note that all of these results can be easily translated to the case of edge-colourings, since colouring the edges of a graph G amounts to vertex-colouring the line graph of G. In particular, one can determine the line graph of G from its k-edge-recolouring graph when k is larger than the chromatic index of G. Using a classical theorem of Whitney [6], this allows one to determine G up to its connected components that are isomorphic to a triangle, a 3-edge star, or a single vertex.

2 Kempe recolouring

In the previous section we considered reconfiguration graphs obtained by changing the colour of only one vertex at the time. There is another operation that is natural to consider in problems involving recolouring: performing a Kempe swap. Given a k-colouring of a graph, a *Kempe swap* involves swapping the colours in a maximal connected subgraph induced by two given colours. Note that recolouring a single vertex is a special case of Kempe swap, called a *trivial Kempe swap.*

The Kempe swap operation leads to the definition of the *k-Kempe-recolouring graph* $\mathcal{K}_k(G)$, whose vertices are all the k-colourings of G and whose edges link k-colourings that differ by a single Kempe swap. Observe that $\mathcal{C}_k(G)$ and $\mathcal{K}_k(G)$ have the same vertex set, while the trivial Kempe swaps show that $E(\mathcal{C}_k(G)) \subseteq E(\mathcal{K}_k(G))$. As before, we make no distinction between a vertex of $\mathcal{K}_k(G)$ and the k-colouring of G it represents.

We extend the previous results to the case of Kempe recolouring. We show that when k is large enough, we can recognise which edges of $\mathcal{K}_k(G)$ correspond to trivial Kempe swaps. In other words, we can recognise which edges of $\mathcal{K}_k(G)$ induce a copy of $\mathcal{C}_k(G)$, and thus apply the machinery from the single-vertex recolouring case.

Theorem 2.1.
Let G be a graph. If $k > \chi(G) + 1$, then G can be determined from $\mathcal{K}_k(G)$, even without knowing the exact value of k.

The same examples used for Theorem 1.4 show that the Kempe-recolouring graph is also not unique if $k = \chi(G)$. This leaves open the case where $k = \chi(G) + 1$.

Question 2.2.
Can every graph G be determined from $\mathcal{K}_{\chi(G)+1}(G)$?

Similarly to the single-vertex recolouring case, our proof of Theorem 2.1 provides a recovery algorithm which runs in polynomial time with respect to the size of $\mathcal{K}_k(G)$, which can be improved to run in polynomial time with respect to $|V(G)|$ if the number of colours used is large enough.

Theorem 2.3.
Let G be an n-vertex graph of maximum degree Δ. If $k > \min\{n, 2\Delta\} + 1$, then G can be determined from $\mathcal{K}_k(G)$ by considering at most $(kn)^2$ different colourings in $\mathcal{K}_k(G)$, even without knowing the exact value of k.

3 Independent set reconfiguration

In this last section we investigate the realm of independent set reconfiguration. An *independent set* in a graph is a set of pairwise non-adjacent vertices. We interpret each independent set as a set of tokens placed on its vertices, and allow tokens to move according to one of three possible rules (so long as the independent set property is not broken):

- *Token Jumping* (TJ for short): a token can be moved to any other vertex.
- *Token Sliding* (TS for short): a token can only be moved to an adjacent vertex.
- *Token Addition and Removal* (TAR for short): a single token can be added or removed.

Given a graph G, we denote by $\mathcal{I}_{\mathrm{J}}^k(G)$ the reconfiguration graph of all independent sets of size k under the TJ rule. In other words, each vertex of $\mathcal{I}_{\mathrm{J}}^k(G)$ corresponds to an independent set of size k in G, and each edge corresponds to a valid token jump. We define $\mathcal{I}_{\mathrm{S}}^k(G)$ similarly.

Note that TJ and TS preserve the number of tokens, but that this is not the case for TAR. Therefore, the definition of a TAR-reconfiguration graph is slightly different: $\mathcal{I}_{\mathrm{AR}}^k(G)$ has a vertex for each independent set of G of size *at least* k, and each edge corresponds to a single token addition or removal.

It turns out that the situation for independent set reconfiguration is quite different than what we saw previously for recolouring. Aside from some trivial cases, one is not guaranteed to be able to determine G from any of the token-reconfiguration graphs defined above. More precisely, we prove the following.

Theorem 3.1.
Let G be a graph. If $\mathrm{R} = \mathrm{S}$ and $k = 1$, or $\mathrm{R} = \mathrm{AR}$ and $k \in \{0, 1\}$, one can determine G from $\mathcal{I}_{\mathrm{R}}^k(G)$. In all other cases, there exist non-isomorphic graphs G with isomorphic independent set reconfiguration graphs.

References

[1] Shamil Asgarli, Sara Krehbiel, Howard W. Levinson, and Heather M. Russell. Counting subgraphs of coloring graphs, 2024.

[2] Gaétan Berthe, Caroline Brosse, Brian Hearn, Jan van den Heuvel, Pierre Hoppenot, and Théo Pierron. Determining a graph from its reconfiguration graph, 2025.

[3] family=Heuvel given=Jan, prefix=van den. The complexity of change. In Simon R. Blackburn, Stefanie Gerke, and Mark Wildon, editors, *Surveys in Combinatorics 2013*, pages 127–160. Cambridge Univ. Press, Cambridge, 2013.

[4] Emma Hogan, Alex Scott, Youri Tamitegama, and Jane Tan. A note on graphs of k-colourings, 2024.

[5] Naomi Nishimura. Introduction to reconfiguration. *Algorithms*, 11(4):Paper No. 52, 25 pp, 2018.

[6] Hassler Whitney. Congruent graphs and the connectivity of graphs. *Amer. J. Math.*, 54(1):150–168, 1932.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On a Property of the 0-1 Subword Order

(Extended abstract)

Kerrin Bielser*

Abstract

The 0-1 subword order $SO(2)$ consists of all 0-1 words, ordered partially and totally (using the subword-relation and the *vip*-order). Among others, Bezrukov [1] proved that this is a Macaulay poset, i.e., a poset with a total order that satisfies a Kruskal-Katona-like condition. In this paper, it is shown that the dual poset $SO(2)^*$ is final shadow increasing, which is interesting in the context of Macaulay posets.

1 Introduction

The study of Macaulay posets was initiated in 1927 by F. S. Macaulay, [7], who proved that products of infinite chains fulfil what was later termed the Macaulay condition. A detailed introduction to the topic can be found in chapter 8 of Engel's book [3] and in the survey by Bezrukov and Leck [2]. The 0-1 subword order is a Macaulay poset (see e.g., [1]), and Leck [6] proved that no other subword order is a Macaulay poset.

According to Engel and Leck [3], some optimisation problems, regarding for example optimal ideals, can be solved in Macaulay posets with specific shadow properties. Leck [5] presents a table of classes of Macaulay posets with their shadow properties. The 0-1 subword order and its dual are of special interest since these are the only known Macaulay posets which are not final shadow increasing and shadow increasing, respectively. With Theorem 3.1, this paper answers the question as to whether the dual of the 0-1 subword order is final shadow increasing. Proofs for other missing entries in the table in [5] have been established by the author and will be presented separately.

2 Preliminaries

For a natural number n ($\in \mathbb{N} := \mathbb{N}_{>0}$), let $[n]$ denote the set $\{1, 2, \dots, n\}$. We are considering totally ordered posets $(P, \leq, \preceq)$, where $\leq$ is the partial and $\preceq$ the total order of the set P. For $x, y \in P$, y *covers* x—written as $x \lessdot y$—if $x < y$ and there is no $z \in P$ with $x < z < y$.

*Institute of Mathematics, Europa-Universität Flensburg, 24943 Flensburg, Germany. E-mail: kerrin.bielser@uni-flensburg.de.

An element $x \in P$ is called maximal if $(y \geq x \Rightarrow y = x)$ holds; minimal elements are defined analogously. For subsets $M, N \subseteq P$, the relation $N \prec M$ means that $x \prec y$ for all $x \in N$ and $y \in M$.

A *rank function* is a function $r : P \to \mathbb{Z}$ with the property that $x \lessdot y$ implies $r(y) = r(x) + 1$ for all $x, y \in P$. Thus, it is possible to partition P into *levels* P_i, consisting of all elements $x \in P$ with $r(x) = i$. A poset is said to be *graded* if minimal or maximal elements exist and have all the same minimal or maximal rank respectively.

The *dual* P^* of a poset P consists of the same set with the reverse orders $x \leq^* y$ iff $y \leq x$ and $x \preceq^* y$ iff $y \preceq x$ for $x, y \in P$. Every subset $S \subseteq P$ of a poset together with the restricted orders $\leq|_S$ and $\preceq|_S$ is called a *subposet*. For clarity, we always consider the rank function of the main poset for subposets. Hasse diagrams are used to represent a structure $(P, \leq, \preceq)$, where $\prec$ is applied within one level from left to right.

2.1 The Dual of the 0-1 Subword Order

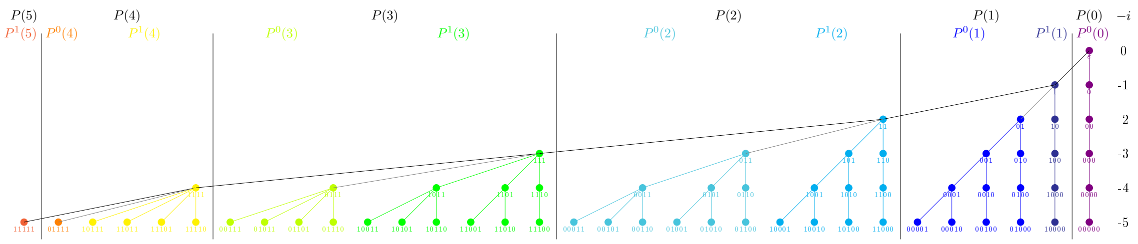


Figure 1: New shadow representation of the upper levels of $SO(2)^*$ (with subposets)

In general, for $z \in \mathbb{N}$, the subword order $SO(z)$ is defined as the poset consisting of all *words* containing *letters* from the *alphabet* $\{0, \ldots, z-1\}$, together with the subword-relation. Thereby, ε denotes the empty word. If a word $\mathbf{x}$ has n letters, the representation $x_1 \cdots x_n$ will be used alternatively. In this paper, we are going to study the dual $SO(2)^*$:

Definition 2.1. The *dual of the 0-1 subword order* is the poset consisting of the set $SO(2)^* := \{x_1 \cdots x_n \mid n \in \mathbb{N}, x_j \in \{0,1\}\ \forall j \in [n]\}$ together with the partial order

$$x_1 \cdots x_n \leq y_1 \cdots y_m :\Leftrightarrow \exists\{i_1, \ldots, i_m\} \subseteq [n] \text{ with } i_1 < \cdots < i_m \text{ and } x_{i_j} = y_j\ \forall j \in [m].$$

Let $w(x_1 \cdots x_n) = \sum_{j=1}^{n} x_j$ denote the number of 1s in a word $x_1 \cdots x_n$. Then the total order $\preceq$ (also known as *vip*-order, referring to the vertex-isoperimetric problem) is

$$\begin{aligned} \mathbf{x} \prec \mathbf{y} :\Leftrightarrow \quad & (1)\ w(\mathbf{x}) > w(\mathbf{y}), \\ \text{or } & (2)\ w(\mathbf{x}) = w(\mathbf{y}) \text{ and } h := \min\{j \mid x_j \neq y_j\} \text{ exists and fulfils } x_h = 0, y_h = 1, \\ \text{or } & (3)\ w(\mathbf{x}) = w(\mathbf{y}), n > m \text{ and } x_j = y_j \quad \forall j \in [m]. \end{aligned}$$

Since there is no minimal but a unique maximal element in this poset, let us consider $r : SO(2)^* \to -\mathbb{N}_0$ as the rank function with $r(x_1 \cdots x_n) = -n$, and $r(\varepsilon) = 0$.

2.2 Macaulay Posets and Some Shadow-Properties

A *segment* S is a subset $S \subseteq P_i$, such that $(s, t \in S, u \in P_i$ and $s \preceq u \preceq t) \Rightarrow u \in S$. The segment is called an *initial* segment if it is minimal or a *final* segment if it is maximal in P_i with respect to $\preceq$. For $\ell \in [|P_i|]$, the notations $C(\ell, P_i)$ and $L(\ell, P_i)$ indicate the initial and final segments in P_i consisting of ℓ elements. Given any $X \subseteq P_i$, its left and right compressions are $C(X) := C(|X|, P_i)$, and $L(X) := L(|X|, P_i)$.

The main subjects of investigation throughout the following sections are the *shadow*—which is $\Delta(x) := \{y \in P \mid y \lessdot x\}$ for $x \in P$, and $\Delta(X) := \bigcup_{x\in X} \Delta(x)$ for $X \subseteq P$—as well as the *new shadow*, which is $\Delta_{\text{new}}(x) := \{y \in \Delta(x) \mid \nexists z \prec x \text{ with } y \lessdot z\}$ for $x \in P$, and $\Delta_{\text{new}}(X) := \bigcup_{x\in X} \Delta_{\text{new}}(x)$ for $X \subseteq P$. A Macaulay poset is a poset such that initial segments are optimal with respect to the shadow minimisation problem, and their shadow sets are also initial segments:

Definition 2.2. Let $(P, \leq)$ be a poset. A total order $\preceq$ on P is a *Macaulay order* if

$$\Delta(C(X)) \subseteq C(\Delta(X)) \text{ for all } X \subseteq P_i, i \in r(P)$$

is fulfilled. Accordingly, $(P, \leq, \preceq)$ is called a *Macaulay poset.*

An important motivation for studying Macaulay posets is the fact that they generalise the Kruskal-Katona theorem, which can be interpreted as the Macaulayness of Boolean lattices with the reverse-lexicographical order (see [2]).

Theorem 2.3 (Leck [6])**.** *The 0-1 subword order $(SO(2), \leq, \prec)$ is a Macaulay poset. Moreover, $SO(z)$ has a Macaulay order only for $z = 2$.*

The following well-known statement is the basis for studying the dual of the 0-1 subword order as a Macaulay poset:

Lemma 2.4 (Engel [3])**.** *A poset $(P, \leq, \prec)$ is a Macaulay poset iff its dual P^* is Macaulay.*

Corollary 2.5 (Leck [5]). $(SO(2)^*, \leq, \prec)$ is a Macaulay poset.

As mentioned, Macaulay posets which additionally ensure the shadow properties defined below are of special interest, e.g., for some optimisation problems [4]. (There are more shadow properties, such as being shadow increasing, which are not a subject of this paper.)

Definition 2.6. A Macaulay poset P is called *additive* iff

$$|\Delta_{\text{new}}(C(S))| \geq |\Delta_{\text{new}}(S)| \geq |\Delta_{\text{new}}(L(S))|$$

for every segment $S \subseteq P_i$ and each $i \in r(P)$. Moreover, it is *final shadow increasing* iff

$$|\Delta_{\text{new}}(L(\ell, P_i))| \leq |\Delta_{\text{new}}(L(\ell, P_{i+1}))|$$

for $i, i+1 \in r(P)$ and $\ell \leq \min\{|P_i|, |P_{i+1}|\}$.

Since $\Delta(C(X)) = \Delta_{\text{new}}(C(X))$, it is sufficient to consider the Hasse diagram in its *new shadow representation* when checking a poset for these shadow properties. The new shadow representation of the Hasse diagram only includes edges between an element and its new shadow.

Lemma 2.7 (Leck [5])**.** *The dual of the 0-1 subword order is additive.*

3 Main result

Theorem 3.1. *The dual of the 0-1 subword order is final shadow increasing., i.e.*

$$|\Delta_{\text{new}}(L(\ell, SO(2)^*_{-i}))| \le |\Delta_{\text{new}}(L(\ell, SO(2)^*_{-i+1}))| \text{ for all } i \in \mathbb{N}, \ell \in [2^{i-1}].$$

The following observations prepare the proof of this theorem.

Definition 3.2. For $k \in \mathbb{N}_0$, define $P(k) \subset SO(2)^*$ as the subposet of all words which contain exactly k ones. Moreover, let $P^1(k) := \{x_1 \cdots x_n \in SO(2)^* \mid x_1 = 1 \text{ and } w(\mathbf{x}) = k\}$ consist of all words in $P(k)$ whose first letter is a one, and let $P^0(k) := P(k) \setminus P^1(k)$ consist of all words of $P(k)$ with a zero at the beginning, as well as—for $k = 0$—the empty word.

Lemma 3.3. *These subposets have the following properties:*

(a) $\dot{\bigcup}_{k \in \mathbb{N}_0} P(k) = SO(2)^*$, *and* $P(k) = P^0(k) \,\dot{\cup}\, P^1(k)$ *for all* $k \in \mathbb{N}_0$,

(b) $P(k+1) \prec P(k)$ *for all* $k \in \mathbb{N}_0$, *and* $P^0(k) \prec P^1(k)$ *for all* $k \in \mathbb{N}$,

(c) $\Delta_{\text{new}}(P(k)_{-i}) = P(k)_{-i-1}$, *and* $\Delta_{\text{new}}(P^e(k)_{-i}) = P^e(k)_{-i-1}$ *for all* $k, i \in \mathbb{N}_0$, $i > k$, $e \in \{0, 1\}$.

Proof. (a) According to its definition, every $\mathbf{x} \in SO(2)^*$ is uniquely assigned to a $P^e(k)$.

(b) Let $\mathbf{x} \in P(k+1)$, $\mathbf{y} \in P^0(k)$, and $\mathbf{z} \in P^1(k)$. Then $w(\mathbf{x}) = k+1 > k = w(\mathbf{y})$ holds, and thus, $\mathbf{x} \prec \mathbf{y}$ according to (1). Furthermore, (2) yields $\mathbf{y} \prec \mathbf{z}$, since $w(\mathbf{y}) = k = w(\mathbf{z})$, $y_1 = 0$ and $z_1 = 1$. So the claim follows from the transitivity of $\preceq$, and (a). For $k = 0$, $P(0) = P^0(0)$.

(c) Using (a) it is sufficient to prove the second part $\Delta_{\text{new}}(P^e(k)_{-i}) = P^e(k)_{-i-1}$. Therefore, let $\mathbf{x} \in P^e(k)_{-i-1}$. Because $i > k$, $x \neq \mathbf{1}_k$—the word of length k without any zero. Thus, it is possible to define $\mathbf{x}'$ as the word which one obtains from $\mathbf{x}$ by deleting the last occurring 0. This word fulfils $\mathbf{x} \lessdot \mathbf{x}'$, it contains i letters, and due to $i > k$, the deleted letter was not the first one. Consequently, $\mathbf{x}' \in P^e(k)_{-i}$, so $\mathbf{x} \in \Delta(\mathbf{x}') \subseteq \Delta(P^e(k)_{-i})$. Let us assume the existence of a $\mathbf{y} \prec P^e(k)_{-i}$ with $\mathbf{x} < \mathbf{y}$. According to the partial orders definition, it is impossible that $w(\mathbf{y}) > k$. Therefore, (b) yields $\mathbf{y} \in P^0(k)_{-i}$ and $\mathbf{x} \in P^1(k)_{-i-1}$ since $\mathbf{y} \prec P^e(k)_{-i}$. Because $w(\mathbf{x}) = k = w(\mathbf{y})$, the word $\mathbf{y}$ cannot be obtained from $\mathbf{x}$ by deleting $x_1 = 1$, and as $y_1 = 0$, this contradicts $\mathbf{x} < \mathbf{y}$. This results in $P^e(k)_{-i-1} \subseteq \Delta_{\text{new}}(P^e(k)_{-i})$ and moreover equality since the subposets are partitioning $SO(2)^*$. □

Remark 3.4. For $\mathbf{1}_n \in P^1(n)$, it holds that both $\mathbf{1}_{n+1} \in \Delta_{\text{new}}(\mathbf{1}_n) \cap P(n+1)$ and $0\mathbf{1}_n \in \Delta_{\text{new}}(\mathbf{1}_n) \cap P^0(n)$. Nevertheless, this will not be a problem during the proof of Theorem 3.1: If $\mathbf{1}_n$ is included in a final segment, this has to equal the whole level. Hence, such a final segment is of length 2^n and can only be compared to a final segment in the level below. It is therefore sufficient to ignore the associated edge in the Hasse diagram since the inequality $|\Delta_{\text{new}}(L(2^n, SO(2)^*_{-n}) \setminus \{\mathbf{1}_{n+1}, 0\mathbf{1}_n\}| \ge |\Delta_{\text{new}}(L(2^n, SO(2)^*_{-n-1})|$ is even stronger than needed for final shadow increase.

Lemma 3.5. *The isomorphisms $P(k) \cong P^0(k)$ and $P(k) \cong P^1(k+1)$ hold in $SO(2)^*$ for every $k \in \mathbb{N}_0$.*

Proof. It is easy to show that the following two functions are both isomorphisms:

$$\varphi_0 : P(k) \to P^0(k), \qquad x_1 \cdots x_n \mapsto 0x_1 \cdots x_n \text{ and}$$
$$\varphi_1 : P(k) \to P^1(k+1), \; x_1 \cdots x_n \mapsto 1x_1 \cdots x_n$$

□

Remark 3.6. According to this lemma, it is possible to split the subposets analogously.

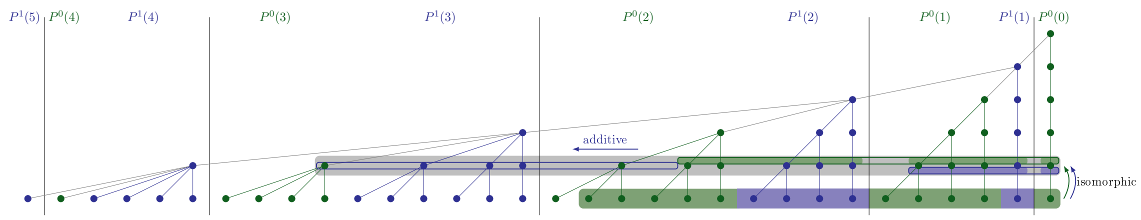


Figure 2: Visualisation of the proof (note that the poset continues infinitely downwards)

Proof of Theorem 3.1. We need to verify $|\Delta_{\text{new}}(L(\ell, SO(2)^*_{-i}))| \le |\Delta_{\text{new}}(L(\ell, SO(2)^*_{-i+1}))|$ for all $i \in \mathbb{N}, \ell \in [2^{i-1}]$. During this proof, let $P := SO(2)^*$, $P^0 := \bigcup_{k\in\mathbb{N}_0} P^0(k)$, and $P^1 := \bigcup_{k\in\mathbb{N}_0} P^1(k)$. Lemma 3.5 implies for all $k \in \mathbb{N}_0, i \in \mathbb{N}$ that $P^0(k)_{-i} \cong P(k)_{-i+1}$ and $P^1(k)_{-i} \cong P(k-1)_{-i+1}$. Furthermore, from Lemma 3.3 (c), we know that the new shadow of every word except $\mathbf{1}_k$ is contained in the same subposet. Since $P^0(0)_{-i} \cong P(0)_{-i+1}$ for every $i \in \mathbb{N}$ and $P^1(0) = \emptyset$,

$$P^0_{-i} \setminus \{0\mathbf{1}_{i-1}\} = P^0(0)_{-i} \cup \bigcup_{k=1}^{i-2} P^0(k)_{-i} \cong P(0)_{-i+1} \cup \bigcup_{k=1}^{i-2} P(k)_{-i+1} = P_{-i+1} \setminus \{\mathbf{1}_{i-1}\} \text{ and}$$
$$P^1_{-i} \setminus \{\mathbf{1}_i\} = \bigcup_{k=1}^{i-1} P^1(k)_{-i} \cong \bigcup_{k=1}^{i-1} P(k-1)_{-i+1} = P_{-i+1} \setminus \{\mathbf{1}_{i-1}\} \text{ follow.}$$

For $i = 1$, the claim is trivial. Therefore, take $i > 1$. We get $L(\ell, P_{-i}) \subseteq P_{-i} \setminus \{0\mathbf{1}_{i-1}, \mathbf{1}_i\}$, because $\ell \le |P_{-i+1}| = 2^{i-1} < 2^i - 2 = |P_{-i}| - 2$. Let $\ell_0 := |L(\ell, P_{-i}) \cap P^0|$ as well as $\ell_1 := |L(\ell, P_{-i}) \cap P^1|$. Since $L(\ell, P_{-i}) \cap P^0$ and $L(\ell, P_{-i}) \cap P^1$ are final segments *within* P^0_{-i} and P^1_{-i} respectively, all things considered

$$\begin{aligned} |\Delta_{\text{new}}(L(\ell, P_{-i}))| &= |\Delta_{\text{new}}(L(\ell, P_{-i}) \cap P^0)| + |\Delta_{\text{new}}(L(\ell, P_{-i}) \cap P^1)| \\ &= |\Delta_{\text{new}}(L(\ell_0, P^0_{-i}))| + |\Delta_{\text{new}}(L(\ell_1, P^1_{-i}))| \\ &= |\Delta_{\text{new}}(L(\ell_0, P_{-i+1}))| + |\Delta_{\text{new}}(L(\ell_1, P_{-i+1}))| \le |\Delta_{\text{new}}(L(\ell, P_{-i+1}))|, \end{aligned}$$

where the last step uses the additivity in $SO(2)^*$ according to Lemma 2.7 (applied to $L(\ell_1, P_{-i+1})$, for example) and $\ell = \ell_0 + \ell_1$. Consequently, $SO(2)^*$ is final shadow increasing. □

References

[1] Sergei L. Bezrukov. Discrete extremal problems on graphs and posets, 1995.

[2] Sergei L. Bezrukov and Uwe Leck. Macaulay posets. *The Electronic Journal of Combinatorics*, (DS12), 2004.

[3] Konrad Engel. *Sperner Theory.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1997.

[4] Konrad Engel and Uwe Leck. Optimal antichains and ideals in macaulay posets. *Graph Theory and Combinatorial Biology*, (7):199–222, 1999.

[5] Uwe Leck. Contributions to the theory of macaulay posets, 2001.

[6] Uwe Leck. Nonexistence of a kruskal–katona type theorem for subword orders. *Combinatorica*, (24):305–312, 2004.

[7] Francis Sowerby Macaulay. Some properties of enumeration in the theory of modular systems. *Proc. London Math. Soc*, (26):531–555, 1927.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Efficient Random-Bit Algorithms for Uniform Involutions

(Extended abstract)

Olivier Bodini* Francis Durand†

Abstract

We succinctly present two algorithms for generating uniform random involutions of size n in $O(n \ln(n))$ time. In the two cases, the samplers achieve asymptotically at the first order optimal random bit consumption. The second algorithm leverages a novel combinatorial decomposition and log-concave distribution sampling of a key parameter, the ghost number. We analyze the asymptotic distribution of this parameter, proving its convergence to a normal distribution. Furthermore, structural properties of involutions are derived directly from the algorithm's dynamics, providing new combinatorial insights into their behavior.

1 Introduction

In this paper, we address the problem of generating uniformly at random an involution of size n with optimal time complexity and random-bit complexity. Involutions, bijective functions that are their own inverses, arise in numerous mathematical and applied contexts. They play a central role in algebra, combinatorics, and theoretical computer science, frequently appearing in classical bijective correspondences with other combinatorial structures. Understanding their properties through efficient random sampling is crucial, both for theoretical advancements and practical applications.

Let I_n denote the number of involutions of $[n]$. Although they are holonomic: $I_n = I_{n-1} + (n-1)I_{n-2} \quad \text{for} \quad n \geq 2$, with initial conditions $I_0 = 1$ and $I_1 = 1$. their efficient uniform sampling is nontrivial, the following description not producing an optimal sampling by the recursive method [11]. The first term in this recurrence accounts for the involutions where n is a fixed point, and the second term corresponds to cases where n is paired with one of the $n-1$ other elements, leaving an involution on the remaining $n-2$ elements. Another classic

*EREN, USPN. E-mail: olivier.bodini@eren.smbh.univ-paris13.fr. Supported by ANR Lambda-Comb.

†LIP6, Sorbonne Université and EREN, USPN. E-mail: francis.durand@ens-paris-saclay.fr. Supported by ANR PAnDAG.

approach is to use the Boltzmann samplers [5], but it is only efficient for approximate-size generation and consumes a significant amount of randomness during the rejection phase.

In combinatorial structures with a large number a_n of objects of size n, the randomness often dominates over rigid structures, making alea the limiting factor. In such cases, Shannon entropy ($\ln_2(a_n)$) becomes the essential quantity to optimize, and this is the guiding principle behind our approach. Here, the Shannon entropy for the involutions is $\left(\frac{\ln(n)}{2} - \frac{1}{2}\right) \cdot \frac{n}{\ln(2)} + \frac{\sqrt{n}}{\ln(2)} + O(1)$. Indeed an involution is a finite union of fixed points (single elements) and 2-cycles (i.e., pairs). This implies that the exponential generating function of the involutions is $I(x) = e^{x+x^2/2}$. A well-known application of the saddle point method gives $I_n \sim c\left(\frac{n}{e}\right)^{n/2} \exp\left(n^{1/2}\right)$, where $c = 2^{-1/2}\exp\left(-\frac{1}{4}\right)$. So, Involutions are a good candidate for our framework.

We propose two time-efficient algorithms that consumes no more than $\frac{n\ln(n)}{2\ln(2)} + O(n)$, which matches well the Shannon entropy for involutions. To the best of our knowledge, these two algorithms are the most efficient ones known to date.

Algorithm 1 (resp. Algorithm 2) operates in the same way, in two stages. The first stage samples a key parameter, κ, that allows partitioning the involutions. This parameter is the number of fixed points (resp. the ghost number, defined later). The second stage distributes values within the created alveoles. The first step employs Devroye's method for sampling log-concave distributions [3, 4], though we do not delve into the details of this aspect. The second stage is carefully designed to minimize unnecessary random-bit consumption, ensuring efficiency in both time and randomness.

Beyond its algorithmic efficiency, our method allows us to investigate structural properties of involutions. These properties are known, but the proofs given are new. Moreover the implementation of our algorithm is available in a Git repository.

Our approach is rooted in the symbolic method, relying on successive pointing-erasing, which naturally leads to a variation of the Hermite polynomials—a connection that sheds new light on the enumeration of involutions. This interplay between algorithmic generation, asymptotic enumeration, and classical combinatorial polynomials provides a richer theoretical framework, making our approach appealing beyond its mere functionality. Moreover, this perspective allows for natural generalizations and alternative interpretations, reinforcing the interplay between combinatorial generation and analytical methods.

2 Efficient Involution Samplers

In this section, we present the two algorithms for the uniform generation of involutions, each offering a distinct perspective on the combinatorial structure of these objects. The first algorithm is an optimized and refined version of a classical recursive approach. It relies on partitioning involutions based on the number of fixed points, enabling efficient selection through a counting and unranking scheme. The second algorithm is an ad hoc approach distinguished by its interesting combinatorial interpretation. Based on an alternative formula for the number of involutions, it provides a different and insightful view of their structure while maintaining simplicity and efficiency comparable to the first method.

2.1 Specifying the Number of Fixed Points

An explicit formula for the number $I_{n,k} = \binom{n}{k} \times \frac{(n-k)!}{2^{(n-k)/2}((n-k)/2)!}$ (with $n-k$ even) of involutions of size n with k fixed points can easily be deduced from the marked generating function $e^{uz+z^2/2}$, which corresponds to the specification $\mathrm{Set}(\mathcal{U} \times \mathcal{Z} + \mathrm{Set}_2(\mathcal{Z}))$

From this partition of involutions by the number of fixed points, one can derive an (effective) random sampling algorithm for involutions.

Algorithm 1 Uniform Random Generation of Involutions

Step 1: Sample the number of fixed points
Sample k from the random variable K, which represents the distribution of the number of fixed points.
Step 2: Generate the elements that serve as the k fixed points.
Directly draw an element for each position or use an entropically optimal unranking operator [8, 7] to select a set of k fixed points uniformly among the $\binom{n}{k}$ possible sets.
Step 3: Pair the remaining elements
Let S be the set of $n-k$ non-fixed elements.
while S is not empty **do**
    Take the smallest non-paired element x from S.
    Randomly select another non-paired element y from S.
    Pair x and y to form a 2-cycle.
    Remove x and y from S.
Return: The generated involution.

Notice that the distribution of the 2-cycles $(n-K)/2$ is log-concave, allowing the use of Devroye's scheme. This step takes $O(\log(n))$ time and entropy on average. Step 2 costs $\ln_2(\binom{n}{k}) + O(n)$ (resp. $+\, O(1)$) random bits depending on the choice that is made [9, 1]. The last step costs $\ln_2(\frac{(n-k)!}{2^{(n-k)/2}((n-k)/2)!}) + O(n)$. This results in an overall $O(n \ln(n))$ time complexity and random-bit consumption.

2.2 Specifying the Number of Ghost Points

Although the first algorithm generates a uniform random involution efficiently, we present here an alternative algorithm that, while maintaining the same simplicity and running time, offers insights into the combinatorial structure of involutions. We present algorithm 2 for generating a uniform random involution of size n.

Theorem 2.1. *The involution, generated by algorithm 2 is a uniform among the involutions of size n.*

Algorithm 2 is based on the following formula for the involutions:

Property 2.2. *The number of involutions I_n of size n is given by:*

$$I_n = \frac{1}{\sqrt{e}} \sum_{2k \geq n} \frac{(2k)^{\underline{n}}}{2^k k!}, \tag{1}$$

where $(2k)^{\underline{n}} = (2k)(2k-1)\dots(2k-n+1)$ is the n-falling factorial of $2k$.

This formula, though not standard, is known and can be found in [10]; however, we derived it independently in our work.

Algorithm 2 Involution Sampling with Ghost Elements

1: **Draw the number of ghost elements** G from the distribution

$$P(G = 2k - n) = \frac{(2k)^{\underline{n}}}{\sqrt{e}2^k k! I_n} \quad (2)$$

2: **Add** G **ghost** elements to the set of n elements, resulting in a total of $G + n$ elements (some real, some ghost).
3: **Generate a uniform random pairing** of the $G + n$ elements.
4: **Transform the pairing into an involution**: For each pair:
- If both elements are real, return the 2-cycle $(a\,b)$.
- If one element is real and the other is ghost, return the fixed point (a).
- If both elements are ghost, discard the pair.

5: **Return** the corresponding involution.

Variation:

1: **Sample** G with the same distribution.
2: **Consider** $\frac{G+n}{2}$, double-slot unit.
3: **Iteratively add all** the elements in one of the $2k$ slots, with at most one element per slot.
4: **Variation**: Analogously transform the double-slot units into an involution.

Remark 2.3. *Algorithm 2 is presented alongside its variation. While the two algorithms are valid, we include both because the main algorithm is more efficient, whereas the variation provides a combinatorial interpretation that is explored in 3.1.*

Generating G_n is not trivial, but it can be achieved using the technique described in [4] for discrete random variables with log-concave distributions. This is justified by the fact that $\frac{G+n}{2}$ has a log-concave distribution (see section A).

Proposition 2.4. *The number of pairs of ghost elements is independent of the generated involution and is a Poisson variable of parameter* $\frac{1}{2}$.

Theorem 2.5. *(Proof is provided in the annex A.) Algorithm 2 runs in* $O(n \ln(n))$ *time and requires, on average,* $\dfrac{n \log_2(n)}{2} + O(n)$ *random bits. The algorithm is quasi-optimal in terms of both runtime and random bit consumption. Specifically:*

- *The runtime* $O(n \log(n))$ *is quasi-optimal, as there exists a lower bound of* $\Omega(n \log(n))$ *for this problem.*
- *The random bit consumption matches the lower bound of* $\dfrac{n \log_2(n)}{2} + \Omega(n)$.

3 Combinatorial Interpretation and Hermite Polynomials

We begin by defining a combinatorial specification: $\mathrm{Set}(\mathcal{U} \times \mathrm{Set}_2(\mathcal{Z} + \mathcal{G}))$. This structure represents a set of pairs, where each pair is marked by a $\mathcal{U}$, and the elements within each pair are either normal $\mathcal{Z}$ or ghost $\mathcal{G}$. Of course, there is a very elementary isomorphism to $\mathrm{Set}(\mathcal{U} \times (\mathcal{Z} \times \mathcal{G} + \mathrm{Set}_2(\mathcal{Z}))) \times \mathrm{Set}(\mathcal{U} \times \mathrm{Set}_2(\mathcal{G}))$. Intuitively, this corresponds to a Cartesian product of involutions, where fixed points are represented by pairs $\{Z, G\}$, and 2-cycles by pairs $\{Z, Z\}$, and additional sets of ghost elements are included. This structure is exactly the one generated by the second algorithm. By disregarding the markers U and the ghost elements G, the structure reduces to a uniform sampling of involutions, as the remaining elements correspond to pairs and singletons of normal elements Z.

But we can also provide a combinatorial interpretation of algorithm 2 using generating functions. Let $\phi(x) = \frac{e^{x^2/2}}{\sqrt{e}} = \frac{1}{\sqrt{e}} \sum_{k\geq 0} \frac{x^{2k}}{2^k k!}$ denote the exponential generating function for configurations of pairs of elements. Additionally, define $G_n(x) = \frac{1}{\sqrt{e}} \sum_{k\geq 0} \frac{(2k)^{\underline{n}} x^{2k-n}}{2^k k!}$ as the generating function G in algorithm 2.

The derivative operator $\frac{d}{dx}$ acts combinatorially as a pointing-erasing operation: applying $\frac{d}{dx}$ to x^n "points" one of the n atoms, putting a label on it and erases it, thus producing nx^{n-1}. This aligns with the interpretation given in [2], where derivative operators correspond to grafting operations on combinatorial structures.

Property 3.1.

$$\left(\frac{d}{dx}\right)^n \phi(x) = G_n(x) = H_n(x)\phi(x), \tag{3}$$

where $H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} I_{n,n-2k} x^{n-2k}$ *are a variant of Hermite polynomials.*

Property 3.1 provides a direct combinatorial interpretation of the variation of algorithm 2:

- **Left-hand side** $((\frac{d}{dx})^n \phi(x))$: Represents n successive pointing-erasing operations on pairs of elements (since it acts on $\phi(x)$). Specifically, it corresponds to iteratively placing the elements of $[n]$ into the slots of the double-slot units.

- **Right-hand side** $(H_n(x)\phi(x))$: Separates the process into two independent components: $H_n(x)$ describes the involutions of size n and $\phi(x)$ describes the empty double-slot units (Poisson-distributed with parameter $\frac{1}{2}$).

References

[1] Axel Bacher, Olivier Bodini, Hsien-Kuei Hwang, and Tsung-Hsi Tsai. Generating random permutations by coin tossing: Classical algorithms, new analysis, and modern implementation. ACM Transactions on Algorithms, 13(2):24:1–24:43, 2017.

[2] Pawel Blasiak and Philippe Flajolet. Combinatorial models of creation-annihilation. arXiv preprint arXiv:1010.0354, 2010.

[3] Luc Devroye. A simple generator for discrete log-concave distributions. Computing, 39(1):87–91, 1987.

[4] Luc Devroye. A note on generating random variables with log-concave densities. Stat. Probab. Lett., 82(5):1035–1039, 2012.

[5] Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. Comb. Probab. Comput., 13(4-5):577–625, 2004.

[6] P Flajolet. Analytic Combinatorics. Cambridge University Press, 2009.

[7] Antoine Genitrini and Martin Pépin. Lexicographic unranking of combinations revisited. Algorithms, 14(3):97, 2021.

[8] Zbigniew Kokosinski. Algorithms for unranking combinations and their applications. In M. H. Hamza, editor, Proceedings of the Seventh IASTED/ISMM International Conference on Parallel and Distributed Computing and Systems, Washington, D.C., USA, October 19-21, 1995, pages 216–224. IASTED/ACTA Press, 1995.

[9] Jérémie Lumbroso. Optimal discrete uniform generation from coin flips, and applications. arXiv preprint arXiv:1304.1916, 2013.

[10] Toufik Mansour, Matthias Schork, and Mark Shattuck. On a new family of generalized stirling and bell numbers. the electronic journal of combinatorics, pages P77–P77, 2011.

[11] Herbert S Wilf. generatingfunctionology. CRC press, 2005.

A Proofs

Proof of 2.2. The result comes from manipulating the generating function of the involutions.

$$e^{x+x^2/2} = \frac{e^{(x+1)^2/2}}{\sqrt{e}} = \frac{1}{\sqrt{e}}\sum_{k\geq 0}\frac{(x+1)^{2k}}{2^k k!} \tag{4}$$

$$= \frac{1}{\sqrt{e}}\sum_{k\geq 0}\sum_{0\leq n\leq 2k}\binom{2k}{n}\frac{x^n}{2^k k!} \tag{5}$$

Hence we can extract: $I_n = n![x^n]e^{x+x^2/2} = \frac{1}{\sqrt{e}}\sum_{2k\geq n}\frac{(2k)!}{(2k-n)!2^k k!}$. □

Proof of 2.1. Denote Σ the involution generated by algorithm 2. Let σ be an involution of size n, we denote c its number of cycles.

If there are $k \geq c$ double-slot units, to make the involution σ, one needs to choose the c units among the k, then to chose which unit has which cycle, then in which slot is the smallest element of each cycle. Once all of those are chosen, there is only one choice among the $k-i+1$ for the i^{th} element that makes σ in the chosen way. So all in all the probability to obtain σ given $G = 2k - n$ is:

$$\mathbb{P}(\Sigma = \sigma | G = 2k-n) = \binom{k}{c}\frac{c!2^c}{(2k)^{\underline{n}}} \tag{6}$$

Hence we can write

$$\mathbb{P}(\Sigma = \sigma) = \sum_{k\geq c}\mathbb{P}(\Sigma = \sigma | G = 2k-n)\mathbb{P}(G = 2k-n) \tag{7}$$

$$= \sum_{k\geq c}\binom{k}{c}\frac{c!2^c}{(2k)^{\underline{n}}}\frac{(2k)^{\underline{n}}}{\sqrt{e}2^k k! I_n} \tag{8}$$

$$= \frac{1}{\sqrt{e}I_n}\sum_{k\geq c}\frac{1}{2^{k-c}(k-c)!} \tag{9}$$

$$= \frac{1}{\sqrt{e}I_n}\sum_{k\geq 0}\frac{(\frac{1}{2})^k}{k!} = \frac{\sqrt{e}}{\sqrt{e}I_n} = \frac{1}{I_n} \tag{10}$$

With eq. (10) coming from the convergent formal series for e^x evaluated at $\frac{1}{2}$. Hence algorithm 2 generates involutions uniformly. $\square$

Proof of 2.4. Let σ an involution and $m \geq 0$ an integer.

$$\begin{aligned}
\mathbb{P}(\Sigma = \sigma \cap W_0 = m) =& \mathbb{P}(G = 2m + 2c - n)\mathbb{P}(\Sigma = \sigma \cap W_0 = m | G = 2m + 2c - n) & (11)\\
=& \mathbb{P}(G = 2m + 2c - n)\mathbb{P}(\Sigma = \sigma | G = 2m + 2c - n) & (12)\\
=& \frac{(2(m+c))^{\underline{n}}}{\sqrt{e} I_n 2^{m+c}(m+c)!}\binom{m+c}{c}\frac{c!2^c}{(2(m+c))^{\underline{n}}} & (13)\\
=& \frac{1}{I_n}\frac{1}{\sqrt{e}2^m m!} & (14)\\
=& \mathbb{P}(\Sigma = \sigma)\mathbb{P}(W_0 = m) & (15)
\end{aligned}$$

With the first equality coming from the fact that to generate σ with m empty pairs, it is needed to have $2m + 2c - n$ ghost elements. The second equality is due to the fact that if $G = 2m + 2c - n$, generating σ implies having m pairs of ghost elements. $\square$

Proof of the log-concavity of $\frac{G+n}{2}$. Let $k \geq \frac{n}{2}$. We call $p_k = \mathbb{P}(\frac{G+n}{2} = k) = \frac{(2k)^{\underline{n}}}{\sqrt{e}2^k k! I_n}$, If $k = \frac{n}{2}$, $p_{k-1} = 0$, so $p_k^2 \geq p_{k-1}p_{k+1}$. Otherwise,

$$\begin{aligned}
\frac{p_k^2}{p_{k+1}p_{k-1}} =& \frac{k+1}{k}\frac{(2k-n+1)(2k-n+2)2k(2k-1)}{(2k-n-1)(2k-n)(2k+1)(2k+2)} & (16)\\
=& \frac{k+1}{k}\frac{(2k-n+1)(2k-1)}{(2k-n-1)(2k+1)}\frac{(2k-n+2)2k}{(2k-n)(2k+2)} & (17)\\
=& \frac{k+1}{k}\frac{4k^2+kn+n-1}{4k^2+kn-n-1}\frac{4k^2-2k(n-2)}{4k^2-2k(n-2)-2n} & (18)\\
\geq& 1 & (19)
\end{aligned}$$

With the inequality coming from the fact all the fraction are of the shape $\frac{f(k)+\alpha}{f(k)}$ with $f(k) > 0$ and $\alpha \geq 0$, it it due to the fact that $2k > n \geq 0$. Hence the distribution is log-concave. $\square$

We denote W_0, W_1 and W_2 the number of pairs with respectively 0, 1 and 2 ghost elements after the insertion of the last element.

Proposition A.1. *As $n \to \infty$:*

- $\frac{G+n}{2}$ *is asymptotically Gaussian with mean* $\frac{n}{2} + \frac{\sqrt{n}}{2}$ *and standard deviation* $\frac{n^{1/4}}{2}$.
- W_2 *is asymptotically Gaussian with mean* $\frac{n}{2} - \frac{\sqrt{n}}{2}$ *and standard deviation* $\frac{n^{1/4}}{2}$.
- W_1 *is asymptotically Gaussian with mean* $\sqrt{n}$ *and standard deviation* $n^{1/4}$.

Proof of A.1. Although the proof is a direct application of the quasi-power theorem [6, Chapter IX], we provide a more elementary proof.

We call $p_k = \mathbb{P}(\frac{G+n}{2} = k) = \frac{(2k)^{\underline{n}}}{\sqrt{e}2^k k! I_n}$. Let $l = O(n^{\frac{1}{4}})$,

$$\frac{p_{\frac{n}{2}+\frac{\sqrt{n}}{2}+l}}{p_{\frac{n}{2}+\frac{\sqrt{n}}{2}}} = \frac{(n+\sqrt{n}+2l)^{\underline{n}}}{2^l\,(n+\sqrt{n})^{\underline{n}}\left(\frac{n}{2}+\frac{\sqrt{n}}{2}+1\right)^{\overline{l}}} \tag{20}$$

$$= \frac{(n+\sqrt{n}+1)(n+\sqrt{n}+3)\dots(n+\sqrt{n}+2l-1)}{(\sqrt{n}+1)\dots(\sqrt{n}+2l)} \tag{21}$$

$$= \frac{(1+\frac{1}{\sqrt{n}}+\frac{1}{n})(1+\frac{1}{\sqrt{n}}+\frac{3}{n})\dots(1+\frac{1}{\sqrt{n}}+\frac{2l-1}{n})}{(1+\frac{1}{\sqrt{n}})\dots(1+\frac{2l}{\sqrt{n}})} \tag{22}$$

$$= \exp\left(\frac{l}{\sqrt{n}} - \frac{1}{\sqrt{n}} - \frac{2}{\sqrt{n}} - \dots - \frac{2l}{\sqrt{n}} + O(n^{-\frac{1}{4}})\right) \tag{23}$$

$$= \exp\left(-\frac{1}{\sqrt{n}} - \frac{2}{\sqrt{n}} - \dots \frac{2l}{\sqrt{n}} + O(n^{-\frac{1}{4}})\right) \tag{24}$$

$$= \exp\left(-\frac{2l(2l+1)}{\sqrt{n}} + O(n^{-\frac{1}{4}})\right) \tag{25}$$

$$= \exp\left(\frac{-(2l)^2}{\sqrt{n}} + O(n^{-\frac{1}{4}})\right) \tag{26}$$

So for any fixed $y \in \mathbb{R}$,

$$\lim_{n\to\infty, l=\frac{n}{2}+\frac{\sqrt{n}}{2}+y\frac{n^{\frac{1}{4}}}{2}} \frac{p_l}{p_{\frac{n}{2}+\frac{\sqrt{n}}{2}}} = e^{\frac{-y^2}{2}}. \tag{27}$$

That is to say, asymptotically, $\frac{G+n}{2}$ is normally distributed, around $\frac{n}{2}+\frac{\sqrt{n}}{2}$, with a standard deviation of $\frac{n^{\frac{1}{4}}}{2}$.

Since, $2W_2 + W_1 = n$, we can write $W_1 = n - 2W_2$. They are $\frac{G+n}{2}$ pairs of elements, each one having 0, 1 or 2 real elements at the end of the execution of the algorithm, so $W_0 + W_1 + W_2 = \frac{G+n}{2}$. This implies that $W_2 = n - \frac{G+n}{2} - W_0$, with W_0 being a Poisson variable of parameter $\frac{1}{2}$ hence W_2 is asymptotically Gaussian with mean $\frac{n}{2} - \frac{\sqrt{n}}{2}$ and standard deviation $\frac{n^{\frac{1}{4}}}{2}$, while W_1 is also asymptotically Gaussian, with mean $\sqrt{n}$ and standard deviation $n^{\frac{1}{4}}$. □

Proof of 2.5. Using the technique described in [4], it is possible to generate $\frac{G+n}{2}$, which has a log-concave distribution, in $O(\ln(\frac{G+n}{2})) = O(\ln(n))$ time and entropy on average. Hence we can generate G in $O(\ln(n))$ time and entropy on average. Once G is generated, generating the pairing is done by iteratively pairing the smallest non paired element to one of the other non paired elements. At step i there are $n+G-2i+1$ other elements. Using the technique developed in [9], one can generate a uniform number in $\{1, ..., k\}$ in $O(\ln(k))$ time and $\log_2(k) + O(1)$ random bits on average. Hence, the whole process of generating $\frac{n+\sqrt{n}}{2} + O(n^{1/4})$ uniform numbers from $\{1, \dots, k\}$ with $k \le n+\sqrt{n}+O(n^{1/4})$ takes $O(n\ln(n))$ time and $\frac{n\log_2(n)}{2}+O(n)$ random bits. These running time and random bit consumption match the beginning of the lower bound $\left(\frac{\ln(n)}{2} - \frac{1}{2}\right) \cdot \frac{n}{\ln(2)} + \frac{\sqrt{n}}{\ln(2)} + O(1)$ given by the Shannon entropy principle. □

Proof. Proof of 3.1 It is clear that applying $\left(\frac{d}{dx}\right)^n$ to every monomial of the series decomposition of $\phi(x)$ gives the series decomposition of $G_n(x)$.

$H_0(x) = 1$ hence $\left(\frac{d}{dx}\right)^0 \phi(x) = H_0(x)\phi(x)$. Since $\frac{d}{dx}\phi(x) = x\phi(x)$ it it clear that the n^{th} derivative of $\phi(x)$ is of the form $H_n(x)\phi(x)$ with $H_n(x)$ a polynomial. In addition, since $\frac{d}{dx}(H_n(x)\phi(x)) = H_{n+1}\phi(x)$, the H_n verify the following recurrence relation: $H_{n+1}(x) = xH_n(x) + H_n'(x)$ Hence since we know the coefficients of $H_n(x)$, the coefficient of degree $n+1-2k$ of $H_{n+1}(x)$ is equal to

$$\frac{n!}{2^k k!(n-2k)!} + \frac{2n!(k+1)}{2^k k!(n-2k+1)!} = \frac{(n+1)!}{2^k k!(n+1-2k)!}. \tag{28}$$

Therefor proving the claim. □

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On graph classes with constant domination-packing ratio

(Extended abstract)

Marthe Bonamy* Mónika Csikós† Anna Gujgiczer‡ Yelena Yuditsky§

Abstract

The dominating number $\gamma(G)$ of a graph G is the minimum size of a vertex set whose closed neighborhood covers all the vertices of the graph. The packing number $\rho(G)$ of G is the maximum size of a vertex set whose closed neighborhoods are pairwise disjoint. In this paper we study graph classes $\mathcal{G}$ such that $\gamma(G)/\rho(G)$ is bounded by a constant $c_{\mathcal{G}}$ for each $G \in \mathcal{G}$. We propose an inductive proof technique to prove that if $\mathcal{G}$ is the class of 2-degenerate graphs, then there is such a constant bound $c_{\mathcal{G}}$. We note that this is the first monotone, dense graph class that is shown to have constant ratio. We also show that the classes of AT-free and unit-disk graphs have bounded ratio. In addition, our technique gives improved bounds on $c_{\mathcal{G}}$ for planar graphs, graphs of bounded treewidth or bounded twin-width. Finally, we provide some new examples of graph classes where the ratio is unbounded.

DOI: `https://doi.org/10.48550/arXiv.2503.05562`

1 Introduction

The duality of packing and covering problems is well known and their relation is widely studied in geometric setups. These two notions can naturally be translated to graphs using the notion of balls.

Balls in graphs. Let $G = (V, E)$ be a simple graph. The *distance* between two vertices $v, u \in V$, denoted $\text{dist}_G(u, v)$, is the length of the shortest path between u and v in G. Let $k \in \mathbb{N}$ and let a *k-ball* centered in $v \in V(G)$ be the set of all vertices at distance at most k from v. For any $v \in V$ let $N_G[v]$ denote the 1-ball centered in v, that is, the closed neighborhood of v in G and let $N_G(v)$ denote $N_G[v] \setminus \{v\}$. The *degree* of v in G, denoted by $\deg_G(v)$ is defined as $|N_G(v)|$. For any $X \subseteq V$ let $N_G[X]$ denote $\bigcup_{v\in X} N_G[v]$ and let $N_G(X)$ denote $\bigcup_{v\in X} N_G(v)$. In the notations above and in subsequent ones, we drop the subscript when G is clear from the context.

*CNRS, LaBRI, Université de Bordeaux, Bordeaux, France. E-mail: `marthe.bonamy@u-bordeaux.fr`.
†CNRS, IRIF, Université Paris Cité, France. E-mail: `csikos@irif.fr`.
‡HUN-REN Alfréd Rényi Institute of Mathematics, MTA–HUN-REN RI Lendület "Momentum" Arithmetic Combinatorics Research Group, Budapest, Hungary. E-mail: `gujgicza@renyi.hu`.
§Université libre de Bruxelles (ULB), Brussels, Belgium. E-mail: `yelena.yuditsky@ulb.be`.

Covering and packing. Covering a graph G with as few balls of radius 1 as possible is commonly known as finding a minimum dominating set of G. Formally, a *dominating set* in G is a collection of vertices such that the union of their closed neighborhoods contains all of $V(G)$. We denote the size of a minimum dominating set in G by $\gamma(G)$. The dual of this problem is to maximize the number of disjoint balls of radius 1 that can be packed in a graph.

Formally, a *packing* in G is a collection of vertices whose closed neighborhoods are pairwise disjoint. In other words, a set of vertices form a packing iff their pairwise distances are at least 3. We let $\rho(G)$ denote the size a maximum packing in G.

A straightforward, general relation between $\gamma(G)$ and $\rho(G)$ is given by the following observation. For any graph G and any packing, a dominating set must have a non-empty intersection with each of the closed neighborhoods of the vertices of the packing, thus $\gamma(G) \geq \rho(G)$. In general those two parameters are not equal. Simple examples where $\gamma(G) > \rho(G)$ include $G = C_4$ or $G = C_5$.

The problems of computing $\gamma(G)$ and $\rho(G)$ have natural formulations as integer programs.

$$\begin{aligned} &\text{minimize} \quad \sum_{v\in V(G)} x_v && \text{maximize} \quad \sum_{v\in V(G)} y_v \\ &\text{s.t.} \quad \sum_{u\in N[v]} x_u \geq 1,\ \forall v \in V(G) && \text{s.t.} \quad \sum_{u\in N[v]} y_u \leq 1,\ \forall v \in V(G), \\ &\qquad x_v \in \{0,1\}, \quad \forall v \in V(G) && \qquad y_v \in \{0,1\}, \quad \forall v \in V(G) \end{aligned} \tag{1}$$

The linear program relaxations of the above integer programs are dual, and their optimums are called fractional dominating and fractional packing numbers, respectively. One of the early papers studying the integrality gap of these problems was due to Lovász [20], who showed that the ratio of integral and fractional dominating numbers is always at most $\log(\Delta(G))$, where $\Delta(G)$ denotes the maximum degree of G.

In this paper, we study the following question.

Question 1.1. Given a graph class $\mathcal{G}$, is there a constant $c_{\mathcal{G}}$ such that

$$\frac{\gamma(G)}{\rho(G)} \leq c_{\mathcal{G}} \quad \text{for each} \quad G \in \mathcal{G} \ \ ?$$

Importantly, a positive answer to this question immediately gives the bound $c_{\mathcal{G}}$ on the integrality gap of both the dominating and the packing problem for any graph $G \in \mathcal{G}$.

2 Preliminaries

There are graph classes for which the ratio γ/ρ is known to be equal to 1, namely, trees [22, Theorem 7], strongly chordal graphs [13, Corollary 3.4], and dually chordal graphs [7, Theorem 3.2]. The ratio is known to be at most 2 for cactus graphs [21, Theorem 8] and connected biconvex graphs [15, Theorem 15]. Recently, it was also shown that for any bipartite cubic graph G, $\gamma(G) \leq \frac{120}{49}\rho(G)$ holds [15, Theorem 5] and for any maximal outerplanar graph H, we have $\gamma(H) \leq 3\rho(H)$ [15, Theorem 6].

2.1 Low-degree graphs.

A general upper bound of $\gamma(G) \leq \Delta(G)\rho(G)$ follows by observing that the union of all neighbors of the vertices in a maximum packing is a dominating set in the graph. This observation was first stated by Henning et al. [16] who also showed that if $\Delta(G) \leq 2$, then $\gamma(G) \leq \rho(G) + 1$ with equality if and only if $G = C_n$ and $n \equiv 1, 2 \mod 3$. Furthermore, they proved that if G is a claw-free graph with $\Delta(G) \leq 3$, then $\gamma(G) \leq 2\rho(G)$ and stated the following conjecture on general subcubic graphs.

Conjecture 2.1. *[16] If G is a connected graph with $\Delta(G) \leq 3$, then $\gamma(G) \leq 2\rho(G) + 1$. The equality holds only for three well-defined graphs, one of which being the Petersen graph.*

2.2 Graphs with no large complete bipartite minors.

A graph H is a *minor* of a graph G if it can be obtained from G by deleting vertices and edges and by contracting edges. Böhme and Mohar [3] studied the more general problem of packing and covering with balls of radius k in $K_{q,r}$-minor free graphs. Their results imply the following on the domination and packing numbers.

Theorem 2.2. *[3, Corollary 1.2] If G does not contain a $K_{q,r}$-minor, then $\gamma(G) < (4r + (q - 1)(r + 1))\rho(G) - 3r + 1$.*

Since graphs embedded in a surface of genus g cannot contain a $K_{k,3}$-minor for any $2g+3 \leq k$, the following bound holds.

Corollary 2.3. *[3, Corollaries 1.3 and 1.4] If G is a graph embedded in a surface of Euler genus g, then $\gamma(G) \leq 4(2g+5)\rho(G) - 9$. In particular, if G is a planar graph, then $\gamma(G) \leq 20\rho(G) - 9$.*

2.3 Graphs with bounded weak coloring numbers.

Theorem 2.2 was generalized in [11] using the *weak* 1*-coloring* and *weak* 2*-coloring numbers* ($\mathrm{wcol}_1(G)$ and $\mathrm{wcol}_2(G)$) of a graph G. Since in this paper we will not work directly with the notion of the weak coloring number, we spare its (lenghty) definition and only give the statement of the theorem.

Theorem 2.4. *[11, Theorem 4] For any G such that $wcol_2(G) \leq c$, we have $\gamma(G) \leq c^2\rho(G)$.*

Remark 2.5. Later, Dvořák proposed the alternative bound of $\gamma(G) \leq 4\mathrm{wcol}_1^4(G)\mathrm{wcol}_2(G)\rho(G)$, which is an improvement for the case where wcol_2 is large enough compared to wcol_1 [12].

Theorem 2.4 implies that the ratio is bounded by some constant for many minor-closed and sparse graph classes, including planar graphs and bounded treewidth graphs. For more information on the weak coloring numbers of minor-closed graph classes, see [17].

2.4 Negative results.

There are much fewer graph classes that were proven to have unbounded ratio. Burger et al. [8] observed that if G is the Cartesian product of two complete graphs on n vertices, then $\gamma(G) = n$ and $\rho(G) = 1$. This implies that there is no constant that bounds the γ/ρ ratio for all bipartite graphs. Another class where the ratio γ/ρ can be arbitrarily large is the class of graphs with arboricity 3 [12].

3 Main result

In this work we extend the lists of graph classes with bounded and unbounded γ/ρ ratios and improve the constant for some graph classes for which some bound was already known. Most of our proofs showing bounded ratios rely on a unified inductive technique which will be outlined in the last section of this extended abstract.

3.1 New graph classes with bounded ratio

Previous results have been suggesting that the key structural property for bounded ratio might be *sparsity*[1]. We provide a positive results for the class of 2-degenerate graphs which are known to be not sparse.

[1]The sparsity we refer to here is one which can, among other ways, be described using the weak coloring numbers.

Theorem 3.1. *For every 2-degenerate graph G, we have $\gamma(G) \leq 7 \cdot \rho(G)$.*

We further extend the list of bounded ratio graph classes with asteroidal triple-free graphs, convex graphs, and intersection graphs of unit disks in the plane.

Theorem 3.2.

1. *For every asteroidal triple-free graph G, we have $\gamma(G) \leq 3 \cdot \rho(G)$.*
2. *For every convex graph G, we have $\gamma(G) \leq 3 \cdot \rho(G)$.*
3. *For every unit-disk graph G, we have $\gamma(G) \leq 32 \cdot \rho(G)$.*

3.2 Improved constants

The classes of graphs with bounded tree-width, planar graphs, and outerplanar graphs have constant weak coloring numbers, therefore by Theorem 2.4, the γ/ρ ratio is bounded by a constant for each of them. However, the precise constants obtained for these classes from the results in [11, 12] are often not optimal for subclasses that have a lot of additional structure. We propose direct proofs for the above classes, resulting in very small constant bounds on the ratio. For graphs with bounded tree-width (and therefore k-trees) we show the following bound.

Theorem 3.3. *Let $k \geq 1$ be an integer. For every graph G with tree-width at most k, we have $\gamma(G) \leq k \cdot \rho(G)$.*

Note that the above bound also improves the bound on γ/ρ shown in [15] from 3 to 2 where G is an outerplanar graph. For planar graphs we show the following bound.

Theorem 3.4. *For every planar graph G, we have $\gamma(G) \leq 10 \cdot \rho(G)$.*

For the case of graphs with bounded twin-width, in [4, Section 7] Bonnet et al. showed that both LP-relaxations (given in Equation (1)) have bounded integrality gap, which implies that the γ/ρ ratio must also be bounded by a constant (but no explicit bound is given there). Here we give direct proof of the following statement.

Theorem 3.5. *Let $k \geq 2$ be an integer. For every graph G with twin-width at most k, we have $\gamma(G) \leq 4k^2 \cdot \rho(G)$.*

3.3 Negative results

We show that the second part of Conjecture 2.1 (stated in [16] and studied in [15]) does not hold.

Theorem 3.6. *There is an infinite family of graphs $\mathcal{G}$ such that for each $G \in \mathcal{G}$, $\Delta(G) \leq 3$ and $\gamma(G) = 2\rho(G) + 1$.*

We show that there is an infinite family of split graphs for which the ratio γ/ρ is not bounded. This in particular implies that the above is not bounded for chordal graphs.

Theorem 3.7. *For every $k \geq 1$, there is a split graph S_k such that $\rho(S_k) = 1$ and $\gamma(S_k) = k$.*

For 3-degenerate graphs we present a construction that shows that the ratio is unbounded in general. This example also shows the existence of graphs with bounded VC-dimension and unbounded ratio.

Theorem 3.8. *For every $k \geq 1$, there is a 3-degenerate graph T_k such that $\rho(T_k) \leq 2$ and $\gamma(T_k) \geq k$.*

4 Outline of the proof technique

One of the difficulties with using induction to prove that γ/ρ is bounded is that deleting vertices can affect the distance (in particular, deleting the common neighbor of two vertices and applying induction may result in them both being in the packing). We circumvent this by working with a stronger induction hypothesis, as follows.

Definition 4.1 ((X, Y)-dominating set of G)**.** *Let G be a graph and let $X, Y \subseteq V(G)$. A set $D \subseteq V(G)$ is an (X, Y)-dominating set of G if $N[D \cup X] \cup Y = V(G)$. We denote by $\gamma_{X,Y}(G)$ the size of a smallest (X, Y)-dominating set of G.*

Definition 4.2 ((X, Y)-packing of G)**.** *Let G be a graph and let $X, Y \subseteq V(G)$. A set $P \subseteq V(G)$ is an (X, Y)-packing of G if any vertex $u \in P$ satisfies the following:*

(a) for any vertex $v \in P \setminus \{u\}$, we have $\mathrm{dist}(u, v) \geq 3$.

(b) for any vertex $x \in X$, we have $\mathrm{dist}(u, x) \geq 2$, or equivalently $P \cap N[X] = \emptyset$.

(c) for any vertex $y \in Y$, we have $\mathrm{dist}(u, y) \geq 1$, or equivalently $P \cap Y = \emptyset$.

We denote by $\rho_{X,Y}(G)$ the size of a largest (X, Y)-packing of G.

While this may seem opaque, the intuition is simple. Informally, X is a set of vertices that will be "for free" in the dominating set and that may have a now-deleted neighbor in the packing, while Y is the next layer: a set of vertices that are already dominated "for free" and that may be at distance two of a now-deleted vertex in the packing.

Proof technique. Let $\mathcal{G}$ be a class of graphs and $C \in \mathbb{N}$. In order to prove that $\frac{\gamma(G)}{\rho(G)} \leq C$ for every graph $G \in \mathcal{G}$, we prove $\frac{\gamma_{X,Y}(G)}{\rho_{X,Y}(G)} \leq C$ for every graph $G \in \mathcal{G}$ and any choice of sets $X, Y \subseteq V(G)$. Note that setting $X = Y = \emptyset$ yields the desired bound on the ratio $\gamma(G)/\rho(G)$. Our inductive proofs have the following main steps.

- Assume that there is a graph $G \in \mathcal{G}$ for which there are sets $X, Y \subseteq V(G)$ such that $\gamma_{X,Y}(G)/\rho_{X,Y}(G) > C$ and consider $G \in \mathcal{G}$ such that it minimizes $|V(G)| + |E(G)|$.
- Choose carefully a $G' \in \mathcal{G}$ that is a subgraph of G and sets $X', Y' \subseteq V(G')$. By the minimality of G, we know that G' has an (X', Y')-dominating set D' and an (X', Y')-packing P' that satisfy $|D'|/|P'| \leq C$.
- Using (D', P'), define an (X, Y)-dominating-packing pair (D, P) for G that satisfies $|D|/|P| \leq C$, reaching a contradiction.

References

[1] Abhijin Adiga and L. Sunil Chandran. Representing a cubic graph as the intersection graph of axis-parallel boxes in three dimensions. *SIAM Journal on Discrete Mathematics*, 28(3):1515–1539, 2014.

[2] Hans-Jürgen Bandelt and Henry Martyn Mulder. Distance-hereditary graphs. *Journal of Combinatorial Theory, Series B*, 41(2):182–208, 1986.

[3] Thomas Böhme and Bojan Mohar. Domination, packing and excluded minors. *Electron. J. Comb.*, 10, 2003.

[4] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: max independent set, min dominating set, and coloring. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 35:1–35:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[5] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: max independent set, min dominating set, and coloring. *SIAM J. Comput.*, 53(5):1602–1640, 2024.

[6] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: Tractable FO model checking. *Journal of the ACM*, 69(1), nov 2021.

[7] Andreas Brandstädt, Victor D. Chepoi, and Feodor F. Dragan. The algorithmic use of hypertree structure and maximum neighbourhood orderings. *Discrete Applied Mathematics*, 82(1):43–77, 1998.

[8] Alewyn P. Burger, Michael A. Henning, and Jan H. van Vuuren. On the ratios between packing and domination parameters of a graph. *Discret. Math.*, 309(8):2473–2478, 2009.

[9] Derek G. Corneil, Stephan Olariu, and Lorna Stewart. Asteroidal triple-free graphs. *SIAM Journal on Discrete Mathematics*, 10(3):399–430, 1997.

[10] Reinhard Diestel. *Graph Theory.* Springer Publishing Company, Incorporated, 5th edition, 2017.

[11] Zdeněk Dvořák. Constant-factor approximation of the domination number in sparse graphs. *European Journal of Combinatorics*, 34(5):833–840, 2013.

[12] Zdeněk Dvořák. On distance r-dominating and $2r$-independent sets in sparse graphs. *Journal of Graph Theory*, 91(2):162–173, 2019.

[13] Martin Farber. Domination, independent domination, and duality in strongly chordal graphs. *Discret. Appl. Math.*, 7(2):115–130, 1984.

[14] Fănică Gavril. The intersection graphs of subtrees in trees are exactly the chordal graphs. *Journal of Combinatorial Theory, Series B*, 16(1):47–56, 1974.

[15] Renzo Gómez and Juan Gutiérrez. Domination and packing in graphs. *Discret. Math.*, 348(1):114393, 2025.

[16] Michael A. Henning, Christian Löwenstein, and Dieter Rautenbach. Dominating sets, packings, and the maximum degree. *Discrete Mathematics*, 311(18):2031–2036, 2011.

[17] Jedrzej Hodor, Hoang La, Piotr Micek, and Clément Rambaud. Weak coloring numbers of minor-closed graph classes. *CoRR*, abs/2407.04588, 2024.

[18] Sang il Oum. Rank-width and vertex-minors. *Journal of Combinatorial Theory, Series B*, 95(1):79–100, 2005.

[19] Hal A. Kierstead and Daqing Yang. Orderings on graphs and game coloring number. *Order*, 20(3):255–264, 2003.

[20] László Lovász. On the ratio of optimal integral and fractional covers. *Discret. Math.*, 13(4):383–390, 1975.

[21] Christian Löwenstein, Dieter Rautenbach, and Friedrich Regen. Chiraptophobic cockroaches evading a torch light. *Ars Combinatoria*, 111:181–192, 2013.

[22] Amram Meir and John W. Moon. Relations between packing and covering numbers of a tree. *Pacific Journal of Mathematics*, 61:225–233, 1975.

[23] Samuel Verblunsky. On the least number of unit circles which can cover a square. *Journal of the London Mathematical Society*, s1-24(3):164–170, 07 1949.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Higher dimensional floorplans and Baxter d-permutations

(Extended abstract)

Nicolas Bonichon* Thomas Muller* Adrian Tanasa*

Abstract

A 2−dimensional mosaic floorplan is a partition of a rectangle by other rectangles with no empty rooms. These partitions (considered up to some deformations) are known to be in bijection with Baxter permutations. A d-floorplan is the generalisation of mosaic floorplans in higher dimensions, and a d-permutation is a $(d-1)$-tuple of permutations. Recently, in N. Bonichon and P.-J. Morel, *J. Integer Sequences* 25 (2022), Baxter d-permutations generalising the usual Baxter permutations were introduced.

In this paper, we consider mosaic floorplans in arbitrary dimensions, and we construct a generating tree for d-floorplans, which generalises the known generating tree structure for 2-floorplans. The encoding of this generating tree appear to be significantly more involved in higher dimensions. Moreover we give a bijection between the 2^{d-1}-floorplans and d-permutations characterized by forbidden vincular patterns. Surprisingly, this set of d-permutations is strictly contained within the set of Baxter d-permutations.

1 Introduction

A *floorplan* of size n is a partition of a rectangle using n interior-disjoint rectangles (called blocks). These combinatorial objects have been studied in various fields of computer science, architecture or discrete geometry.

The boundaries of the blocks in a floorplan define a set of horizontal and vertical edges called *block-edges*. A floorplan is *generic* if the union of horizontal (resp. vertical) block-edges that share the same y-coordinate (resp. x-coordinate) is a single segment. A *border* is a maximal (horizontal or vertical) segment in the inner boundaries of the floorplan. A *mosaic floorplan* is a generic floorplan with no segment crossing. This condition is called the *tatami condition*.

In this paper we investigate a natural generalization of floorplans to higher dimensions. A *d-dimensional floorplan* (or a *box partition*) is a partition of a d-dimensional hyperrectangle

*Laboratoire bordelais de recherche en informatique, Université de Bordeaux, France. E-mails: `nicolas.bonichon@u-bordeaux.fr` , `thomas.muller.1@u-bordeaux.fr`, `ntanasa@u-bordeaux.fr` .

with n interior-disjoint d-dimensional hyperrectangles (called blocks). We call a $(d-1)$-hyperrectangles with fixed coordinate x_i a *facet* of axis i. Given an axis i, the boundary of a d-dimensional hyperrectangle defines two facets of such axis, a lower facet of fixed coordinate $x_i = x_{min,i}$ and a higher one of fixed coordinate $x_i = x_{max,i}$ such that $x_{min,i} < x_{max,i}$. The facets defined by the boundaries of the blocks of a d-floorplan are called *block-facets*. A *d-dimensional floorplan* is *generic* if the set of block-facets that share the same i-th coordinate is also a single facet. A *border* is a maximal facet of the interior of the bounding d-hyperrectangle. Given an inner facet f of a d-dimensional floorplan, we denote by $b(f)$ the border that contains f. Two facets *cross* each other if the intersection of their interiors is non empty A *d-floorplan* is a generic d-dimensional floorplan that has no borders (or equivalently facets) crossing (this being the *tatami condition* in higher dimensions). These objects have already been considered in the literature. In the case $d = 3$, they are called generic boxed Plattenbau in [10].

Two d-floorplans are equivalent if the relative positions (up, down, left, right etc...) of their boxes are the same. When we refer to a d-floorplan, we refer to its whole equivalence class. Mosaic floorplans (or $2-$floorplans) are known to be closely related to pattern-avoiding permutations such as Baxter permutations or separable permutations [1, 3] and other combinatorial objects [5]. The study of the generating functions and the enumeration of families of pattern-avoiding mosaic floorplans is also an active field of research, see for example [2].

Additionally, in [4], a generalization to arbitrary dimensions of a restricted family of mosaic floorplans, called *guillotine partitions*, was considered. The authors find a bijection between $2^{d-1}-$dimensional guillotine partitions and separable d-permutations, which are a higher dimensional generalisation of separable permutations.

In this paper, we exhibit two main results. First, we exhibit a generation tree for d-floorplans. Given a floorplan, we can remove its *top* box and unequivocally fill the resulting empty space in order to obtain a smaller floorplan (using the tatami condition). This gives a natural definition for a generating tree. In the $d = 2$ case, the children of a given floorplan is determined by the number of boxes that touch the top and left boundaries. Moreover, determining these parameters for the children is straightforward [6]. However, for $d \geq 3$, we need to manage more involved parameters. This allows us to enumerate efficiently by an exhaustive generation the set of all d-floorplans for small n.

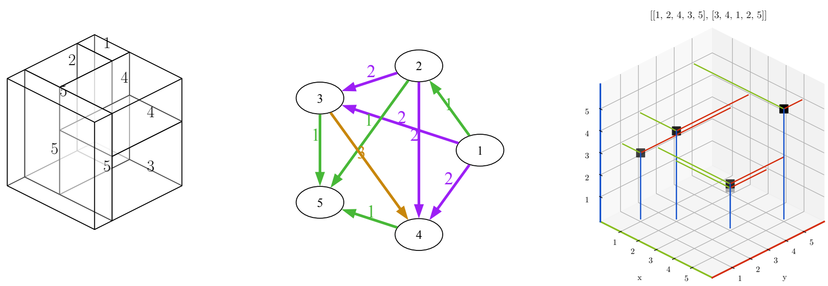


Figure 1: On the left an example of a $3-$floorplan. In the middle the relative order of each block with respect to each direction (x, y, z). On the right the corresponding 3-permutation (considering the 3-floorplan as a 4-floorplan).

The second main contribution is a generalization of the bijection [1] between mosaic floorplans and Baxter permutations. A *d-permutation* (or *multipermutation*) is a tuple of $d-1$ permutations of size n. The presented generalization of the mapping from d-permutation to

2^{d-1}-floorplan is straightforward, but the characterization of the corresponding multipermutation is more involved. These multipermutations are defined by the avoidance of the vincular patterns of Baxter permutations and the dimension 3 patterns of the separable d-permutations. Surprisingly, this set is strictly included in the set of Baxter d-permutations defined in [7]. Moreover, this bijection generalizes the bijection between guillotine 2^{d-1}-floorplans and separable d-permutations of [4].

A summary of the objects and the corresponding pattern avoiding permutations is given in Table 1.

Objects	Permutations	Pattern avoidance
Slicing floorplans	Separable	$\mathrm{Sym}(2413)$
Mosaic floorplans	Baxter	$\mathrm{Sym}(2413\vert_{2,2})$
2^{d-1}-guillotine floorplans	d-Separable	$\mathrm{Sym}(2413), \mathrm{Sym}((312, 213))$
2^{d-1}-Floorplans	sub d-Baxter	$\mathrm{Sym}(2413\vert_{2,2})$, $\mathrm{Sym}((312, 213))$
	$d-$Baxter	$\mathrm{Sym}(2413\vert_{2,2})$, $\mathrm{Sym}((312, 213)\vert_{1,2,.})$, $\mathrm{Sym}((3412, 1432)\vert_{2,2,.})$, $\mathrm{Sym}((2143, 1423)\vert_{2,2,.})$

Table 1: Table of the different class of floorplans and their corresponding permutation classes

Proofs of the results and technical details can be found in [8].

2 A generating tree of d-floorplans

In [1, 11] a block deletion for 2-floorplans operation was introduced. It consists of removing the block incident to a specific corner of the bounding box and then filling the resulting empty space by shifting a block-facet of the deleted block. This operation can be generalized in higher dimensions.

Given a corner q of a d-floorplan $\mathcal{P}$, let B be the block incident to q and let $\bar{q}$ be its opposite corner in B. A *shifting facet* is a block-facet of B incident to $\bar{q}$ such that $b(f) = f$.

A *block deletion* operation with respect to q consists of removing the block B and the facet f. This operation can be seen as shifting the facet f until it reaches q.

Lemma 2.1. *Let B be a block of a d-floorplan $\mathcal{P}$ and q a corner of $\mathcal{P}$. Then the block B has a unique shifting facet.*

Given a d-floorplan $\mathcal{P}$ with more than one block, we define the parent of $\mathcal{P}$ denoted $p(\mathcal{P})$ as the floorplan obtained after a block deletion with respect to the maximal corner $q_{\max}$ of $\mathcal{P}$. This defines a generating tree on d-floorplans whose root is the d-floorplan with only one block. The *children* of a d-floorplan $\mathcal{P}$ are the floorplans $\mathcal{P}'$ such that $p(\mathcal{P}') = \mathcal{P}$. Figure 2 shows the first levels of the generating trees of 3-floorplans induced by the block deletion.

In this generating tree, the children of a d-floorplan can be characterised using the inverse of a block deletion, an operation which we call a *block insertion*.

Before its shifting, a shifting facet f is the lower block-facet of axis i of the block containing q_{max}. Moreover, it is also the union of higher block-facets of axis i of blocks of $\mathcal{P}$. After shifting, f becomes a facet included in the higher facet of the bounding box of $\mathcal{P}$ that contains q_{max}, such that it is the union of higher block-facets of axis i of blocks of $\mathcal{P}$. A *pushable facet* f of axis i is thus a facet of axis i included in the bounding box of $\mathcal{P}$ that is the union of higher

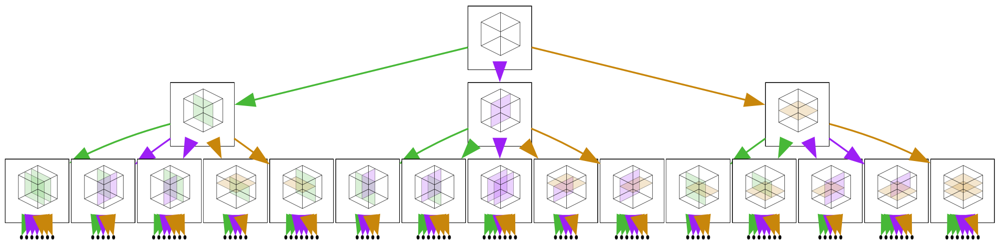


Figure 2: The first levels of the 3-floorplan's generating tree

block-facets of axis i of blocks of $\mathcal{P}$ which contains q_{max}. A *pushable corner* of axis i is a corner q that is the minimal corner of a pushable facet of axis i.

Given a pushable corner q of axis i and its associated facet f_q, we define the *block insertion* associated with q as the d-floorplan obtained by flattening the blocks below f_q in the direction i and inserting a block in the newly created space. This operation defines a mapping between the pushable corners of a d-floorplan and its children. An example of this operation on a 3−floorplan is shown in Figure 3.

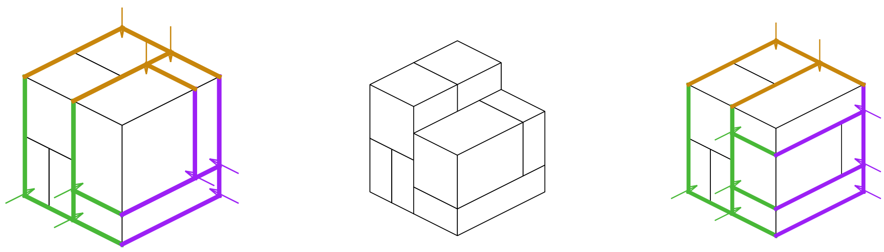


Figure 3: An example of a block insertion on a 3-floorplan.

We say that a pushable corner $q' = (x'_1, \ldots, x'_d)$ of axis i *is shadowed by* a pushable corner $q = (x_1, \ldots, x_d)$ if there exists $j \neq i$ such that $x'_j > x_j$. One can remark that the pushable faces of a given axis are nested. This induces a total order on the pushable corners. The pushable corners of axis i that are shadowed by a pushable corner q of the same axis i are exactly the pushable corners in the interior of the facet f_q. For 2-floorplans, there are no other shadowed pushable corners. As one can see from the example of Figure 3, this is not the case in higher dimensions.

Lemma 2.2. *Let $\mathcal{P}$ be a d-floorplan and q a pushable corner of axis i. The pushable corners of the child $\mathcal{P}'$ of $\mathcal{P}$ with respect to q are the pushable corners of $\mathcal{P}$ that are not shadowed by q plus $\{q^j_{new}, 1 \leq j \leq d\}$.*

By encoding this generating tree, we can enumerate efficiently the set of all d-floorplans. The first entries of the enumeration sequence are given in appendix.

3 Bijection between permutations and floorplans

In higher dimensions, permutations are generalised by *multidimensional permutations*, also called d-permutations [4, 7]. A d-permutation of $[n]$ is a sequence of $d-1$ permutations $\boldsymbol{\pi} = (\pi_1, \dots, \pi_{d-1})$. Given a d-permutation $\boldsymbol{\pi}$, π_0 is the identity permutation on $[n]$. The *diagram* of a d-permutation $\boldsymbol{\pi}$ is the set of points in $P_{\boldsymbol{\pi}} := \{(\pi_0(i), \pi_1(i), \dots, \pi_{d-1}(i)), i \in [n]\}$.

Forbidden patterns: Let a d permutation $\boldsymbol{\pi} = (\pi_1, \dots \pi_{d-1})$ and a sequence of d' indices $\mathbf{i} = i_1 \dots i_{d'}$ in $\{0, \dots, d\}$ $(i_1 < \dots < i_{d'})$. The *projection* of $\boldsymbol{\pi}$ on $\mathbf{i}$ is the d'-permutation defined as $\mathrm{proj}_{\mathbf{i}}(\boldsymbol{\pi}) := (\pi_{i_2}\pi_{i_1}^{-1}, \dots \pi_{i_{d'}}\pi_{i_1}^{-1})$. A $d-$permutation $\boldsymbol{\pi}$ *avoids* a smaller d permutation $\boldsymbol{\pi}'$ (called a *pattern*) if there exists a subset of points of $\boldsymbol{\pi}$ that is order isomorphic to $\boldsymbol{\pi}'$. A $d-$permutation also avoids a pattern of dimension $d' < d$ if all its projections of the same dimension avoid this pattern. A *vincular pattern* denoted $\boldsymbol{\sigma}|_{X_1,\dots,X_d}$ is a pattern $\boldsymbol{\pi}$ and a set of adjacency constraints $X_1, \dots, X_d$. An occurrence of the pattern $\boldsymbol{\sigma}$ in a $d-$permutation is also an occurrence of $\sigma|_{X_1,\dots,X_d}$ if it fulfills the adjacency contraints $X_1, \dots, X_d$ (See [7] for more details). In this paper, we consider F_n^d the class of d-permutations of size n that avoid the patterns Sym $(312, 213))$ and $\mathrm{Sym}(2413|_{2,2})$ (Sym $(\boldsymbol{\sigma})$ denotes the family of patterns obtained by applying the symmetries of the d-cube on $\boldsymbol{\sigma}$) where $\mathrm{Sym}(2413|_{2,2})$ corresponds to the forbidden patterns of Baxter permutations [9, 5].

From floorplans to permutations: The *peeling operation* with respect to a corner q consists of performing recursive block deletions with respect to this corner until it remains a single block. This defines a total order of the blocks that we call *peeling order*, given by the order of deletion of the blocks during the peeling.

The *set of canonical corners* in a 2^{d-1}-floorplan is given by the set of d corners $q_0 \dots q_{d-1}$ for which the coordinates of q_i are given by an alternation of packets of 2^{d-1-i} zeros and ones such that

$$q_i(\mathcal{P}) = \big(\underbrace{0, \dots, 0}_{2^{d-1-i}}, \underbrace{1, \dots, 1}_{2^{d-1-i}}, \underbrace{0, \dots, 0}_{2^{d-1-i}}, \ \dots \ \big) \,. \tag{1}$$

Let $\mathcal{P}$ be a $2^{d-1}-$floorplan, and let σ_{c_i} be the peeling orders of its blocks with respect to the canonical corners. We define the d-permutation $\phi(\mathcal{P})$ as $\pi := (\sigma_{c_1}\sigma_{c_0}^{-1}, \dots, \sigma_{c_{d-1}}\sigma_{c_0}^{-1})$. This defines a mapping ϕ from $2^{d-1}-$floorplans to d-permutations.

From permutations to floorplans: A *direction* f is an element of $\{+1, -1\}^d$. A direction is *positive* if its first element is $+1$. In a d-permutation, there are 2^{d-1} positive directions. If f is a positive direction, the *opposite* of f, denoted $(-f)$, is the direction such that $(-f) = (-1) \times f$. The direction between two points of a $d-$permutation is the direction given by the sign of the difference of their coordinates. These directions give 2^{d-1} independent partial orders of the points of a $d-$permutation. This defines a mapping ψ from d-permutations to set of 2^{d-1} partial orders on the same set of elements.

The second main result of this paper is thus the following theorem:

Theorem 3.1. *The mappings ϕ and ψ define a bijection between F_n^d and the set of $2^{d-1}-$ floorplans with n blocks. One has $\psi = \phi^{-1}$.*

As noted in [4, Corollary 3.3] we can naturally extend the previous theorem to floorplans of arbitrary dimensions (not necessarily a power of 2).

Corollary 3.2. *Let $q \leq 2^{d-1}$. There is a bijection between the set of d-permutations in F_n^d avoiding q (among 2^{d-1}) d-permutations of size 2 with the set of $(2^{d-1} - q)$-dimensional floorplans.*

An example of the bijection showcased here is given in Figure 1

References

[1] Eyal Ackerman, Gill Barequet, and Ron Y Pinter. A bijection between permutations and floorplans, and its applications. *Discrete Applied Mathematics*, 154(12):1674–1684, 2006.

[2] Andrei Asinowski and Cyril Banderier. From geometry to generating functions: rectangulations and permutations, 2024.

[3] Andrei Asinowski, Gill Barequet, Mireille Bousquet-Mélou, Toufik Mansour, and Ron Y Pinter. Orders induced by segments in floorplans and $(2-14-3, 3-41-2)$-avoiding permutations. *Electron. J. Combin.*, 20:P35, 2013.

[4] Andrei Asinowski and Toufik Mansour. Separable d-permutations and guillotine partitions. *Ann. Comb.*, 14:17–43, 2010.

[5] Jean-Christophe Aval, Adrien Boussicault, Mathilde Bouvel, Olivier Guibert, and Matteo Silimbani. Baxter tree-like tableaux, (2021).

[6] Nicolas Bonichon, Mireille Bousquet-Mélou, and Éric Fusy. Baxter permutations and plane bipolar orientations. *Sém. Lothar. Combin.*, 61:B61Ah, 2010.

[7] Nicolas Bonichon and Pierre-Jean Morel. Baxter d-permutations and other pattern-avoiding classes. *Journal of Integer Sequences*, 25(2):3, 2022.

[8] Nicolas Bonichon, Thomas Muller, and Adrian Tanasa. Higher dimensional floorplans and baxter d-permutations, 2025.

[9] Serge Dulucq and Olivier Guibert. Baxter permutations. *Discrete Math.*, 180:143–156, 1998.

[10] Stefan Felsner, Kolja Knauer, and Torsten Ueckerdt. Plattenbauten: touching rectangles in space. In Isolde Adler and Haiko Müller, editors, *International Workshop on Graph-Theoretic Concepts in Computer Science*, volume 12301 of *Lect. Notes in Comp. Sci.*, pages 161–173. Springer, 2020.

[11] Xianlong Hong, Gang Huang, Yici Cai, Jiangchun Gu, Sheqin Dong, Chung Kuan Cheng, and Jun Gu. Corner block list: an effective and efficient topological representation of non-slicing floorplan. In *Proceedings of the 2000 IEEE/ACM international conference on Computer-aided design*, pages 8–12, 2000.

A Appendix

$n \backslash d$	2	3	4	5
1	1	1	1	1
2	2	3	4	5
3	6	15	28	45
4	22	93	244	505
5	92	651	2392	6365
6	422	4917	25204	86105
7	2074	39111	278788	1221565
8	10754	322941	3193204	17932745
9	58202	2742753	37547284	270120905
10	326240	23812341	450627808	4151428385
11	1882960	210414489	5497697848	64839587065
12	11140560	1886358789	67979951368	1026189413865
13	67329992	17116221531	850063243936	
14	414499438	156900657561		
15	2593341586	1450922198319		

Table 2: Values of $|F_n^d|$ for the first values of n.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Immersions of large cliques in graphs with independence number 2 and bounded maximum degree*

(Extended abstract)

Fábio Botler[†] Cristina G. Fernandes[‡] Carla N. Lintzmayer[§] Rui A. Lopes[¶]
Suchismita Mishra[‖] Bruno L. Netto[**] Maycon Sambinelli[††]

Abstract

An immersion of a graph H in a graph G is a minimal subgraph I of G for which there is an injection $\mathrm{i}\colon V(H) \to V(I)$ and a set of edge-disjoint paths $\{P_e : e \in E(H)\}$ in I such that the end vertices of P_{uv} are precisely $\mathrm{i}(u)$ and $\mathrm{i}(v)$. The immersion analogue of Hadwiger Conjecture (1943), posed by Lescure and Meyniel (1985), asks whether every graph G contains an immersion of $K_{\chi(G)}$. Its restriction to graphs with independence number 2 has received some attention recently, and Vergara (2017) raised the weaker conjecture that every graph with independence number 2 has an immersion of $K_{\chi(G)}$. In this paper, we verify Vergara Conjecture for graphs with bounded maximum degree. Specifically, we prove that if G is a graph with independence number 2 and maximum degree less than $19n/29 - 1$, then G contains an immersion of $K_{\chi(G)}$.

*This research has been partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, Finance Code 001.

[†]Universidade de São Paulo, São Paulo, Brazil. E-mail: fbotler@ime.usp.br. Partially supported by CNPq (304315/2022-2) and CAPES (88881.973147/2024-01). CNPq is the National Council for Scientific and Technological Development of Brazil.

[‡]Universidade de São Paulo, São Paulo, Brazil. E-mail: cris@ime.usp.br. Partially supported by CNPq (310979/2020-0 and 404315/2023-2) and by FAPESP (2024/13859-4). FAPESP is the São Paulo Research Foundation, Brazil.

[§]Univ. Federal do ABC, Santo André, Brazil. E-mail: carla.negri@ufabc.edu.br. Partially supported by CNPq (312026/2021-8 and 404315/2023-2) and by L'ORÉAL-UNESCO-ABC For Women In Science.

[¶]Univ. Federal do Rio de Janeiro, Rio de Janeiro, Brazil. E-mail: rui@cos.ufrj.br. Supported by CAPES (88887.843699/2023-00).

[‖]Universidad Andrés Bello, Santiago, Chile. E-mail: suchismitamishra6@gmail.com. Supported by Fondecyt Postdoctoral grant 3220618 of Agencia National de Investigatión y Desarrollo (ANID), Chile.

[**]Univ. Federal do Rio de Janeiro, Rio de Janeiro, Brazil. E-mail: brunoln@cos.ufrj.br. Supported by CAPES (88887.670803/2022-00).

[††]Univ. Federal do ABC, Santo André, Brazil. E-mail: m.sambinelli@ufabc.edu.br. Partially supported by CNPq (407970/2023-1).

1 Introduction

In this paper, every graph is simple. We consider the problem of finding special subgraphs in dense graphs. Specifically, we are interested in a problem related to the following well-known conjecture posed by Hadwiger [14].

Conjecture 1 (Hadwiger, 1943). *Every graph G contains $K_{\chi(G)}$ as a minor.*

Conjecture 1 is still open, but it has been verified in many cases, as for graphs with chromatic number at most 6 [20]. An approach that turned out to be fruitful in exploring Conjecture 1 is to parametrize it by the independence number. Indeed, a graph on n vertices with independence number α has chromatic number at least $\lceil n/\alpha \rceil$, and Conjecture 1 implies that such a graph has a minor of $K_{\lceil n/\alpha \rceil}$. In this direction, Duchet and Meyniel [10] proved that such a graph has a minor of $K_{\lceil n/(2\alpha-1) \rceil}$; and, after several partial results (see [11]), Balogh and Kostochka [3] improved this result by proving that every such graph has a minor of $K_{\lceil n/((2-c)\alpha) \rceil}$, for $c \approx 0.0521$.

As pointed out by Quiroz [19], special attention has been given to the case $\alpha = 2$, and important results explore graphs with small clique covering and fractional clique covering numbers. For example, when G is a graph with an even number n of vertices and independence number 2, Blasiak [4] proved that G contains a minor of $K_{n/2}$ if its fractional clique covering number is less than 3; and Chudnovsky and Seymour [9] proved that G contains a minor of $K_{\lceil n/2 \rceil}$ if G contains a clique of size at least $n/4$.

We are interested in the immersion analogue of Conjecture 1 posed by Lescure and Meyniel [17, Problem 2]. An *immersion* of a graph H in a graph G is a minimal subgraph I of G for which there is an injection $\mathrm{i}\colon V(H) \to V(I)$ and a set of edge-disjoint paths $\{P_e : e \in E(H)\}$ in I such that the end vertices of P_{uv} are precisely $\mathrm{i}(u)$ and $\mathrm{i}(v)$. The vertices of G in the image of i are called the *branch vertices* of I, and we say that such an immersion is *strong* if the internal vertices of the paths P_e are not branch vertices.

Problem 2 (Lescure–Meyniel, 1985). *Does every graph G contain a strong immersion of $K_{\chi(G)}$?*

The weaker version of Problem 2 for (not strong) immersions was posed as a conjecture by Abu-Khzam and Langston [1]. Similarly to Conjecture 1, Problem 2 and the Abu-Khzam–Langston Conjecture received special attention in the case of graphs with independence number 2. In particular, in 2017 Vergara [22] posed the following restriction.

Conjecture 3 (Vergara, 2017). *Every graph G with independence number 2 contains an immersion of $K_{\chi(G)}$.*

Observe that if a graph G on n vertices with independence number 2 contains an immersion of $K_{\chi(G)}$, then it contains an immersion of $K_{\lceil n/2 \rceil}$. In fact, Vergara [22] proved that Conjecture 3 is equivalent to the following.

Conjecture 4 (Vergara, 2017). *Every graph on n vertices with independence number 2 contains an immersion of $K_{\lceil n/2 \rceil}$.*

In this direction, Vergara [22] proved that every graph on n vertices with independence number 2 has a strong immersion of $K_{\lceil n/3 \rceil}$. This was improved by Gauthier, Le, and Wollan [13], who proved that such graphs contain strong immersions of $K_{2\lfloor n/5 \rfloor}$. In 2024, Botler, Jiménez, Lintzmayer, Pastine, Quiroz, and Sambinelli [5] proved that such graphs contain immersions of every complete bipartite graph with $\lceil n/2 \rceil$ vertices (see also [8]); in another direction, Quiroz [18] verified Conjecture 3 for graphs that contain a spanning complete-blow-up of C_5, and Conjecture 4 for graphs with special forbidden subgraphs.

One can expect that vertices of high degree help when looking for large clique immersions. For example, one of the first steps of Vergara's proof of the existence of $K_{\lceil n/3 \rceil}$ immersions in graphs with independence number 2 is to show that their minimum degree is at least $\lfloor 2n/3 \rfloor$, and an important step of the proof of existence of an immersion of $K_{2\lfloor n/5 \rfloor}$ is to prove that every vertex contained in an induced C_5 has degree at least $3n/5$. Similarly, it is not hard to prove that any minimum counterexample to Conjecture 4 has no pair of nonadjacent vertices with at least $\lceil n/2 \rceil - 2$ common neighbors. In this paper, we consider graphs without vertices of large degree. Specifically, we answer Problem 2 positively for graphs with independence number 2 and maximum degree bounded as follows.

Theorem 5. *If G be a graph on n vertices with independence number 2 for which $\Delta(G) < 19n/29 - 1$, then G contains a strong immersion of $K_{\chi(G)}$ and, consequently, a strong immersion of $K_{\lceil n/2 \rceil}$.*

Our proof explores properties of the complement of the studied graph. Specifically, we use the fact that triangle-free graphs with high minimum degree are homomorphic to the well-known Andrásfai graphs (see Section 2). Since C_5 is an Andrásfai graph, our result generalizes the independence-number-2 case of the aforementioned result of Quiroz for graphs containing complements of blow-ups of Andrásfai graphs. Indeed, our result is a consequence of a slightly more general result for graphs with a special 3-clique cover (see Theorem 8 ahead) and the fact that Andrásfai graphs admit a corresponding proper coloring. In fact, we prove the stronger statement that $V(G)$ can be partitioned into two sets A and B such that (i) A induces a clique in G, and (ii) G contains an immersion of a clique whose set of branch vertices is precisely B. One of the main ideas of our proof is to use Hall's Theorem to identify, for each vertex $u \in B$, a vertex $r_u \in A$ such that when u has a missing adjacency, say $uv \notin E(G)$ with $v \in B$, we find a path from u to v through r_v and r_u. The rest of the proof is to show that such paths are edge-disjoint. Due to space constraints, in this paper we present only sketches of some proofs.

2 Andrásfai graphs

Given graphs G and H, a *homomorphism* from G to H is a function $h\colon V(G) \to V(H)$ such that $h(u)h(v) \in E(H)$ for every $uv \in E(G)$. If such a function exists, we say that G is *homomorphic* to H. Let G be a triangle-free graph on n vertices. Andrásfai [2] showed that if $\delta(G) > 2n/5$, then G is bipartite. This result was generalized in many directions, one of which is the following. Häggkvist [15] proved that if $\delta(G) > 3n/8$, then G is 3-colorable, and Jin [16] weakened this minimum degree condition proving that if $\delta(G) > 10n/29$, then G is 3-colorable. Chen, Jin, and Koh [7] strengthened Jin's result showing that G is 3-colorable by exposing the structure of the graph. Specifically, they proved the following result, where Γ_d is the graph (V_d, E_d) for which $V_d = [3d-1]$ and $E_d = \{xy : y = x + i \text{ with } i \in [d, 2d-1]\}$, with arithmetic modulo $3d-1$. The graphs Γ_d for $d \in \mathbb{N}$ are known as the *Andrásfai graphs* (see Figure 1).

Theorem 6 (Chen–Jin–Koh, 1997)**.** *If G is a triangle-free graph on n vertices for which $\delta(G) > 10n/29$, then G is homomorphic to Γ_d for some d.*

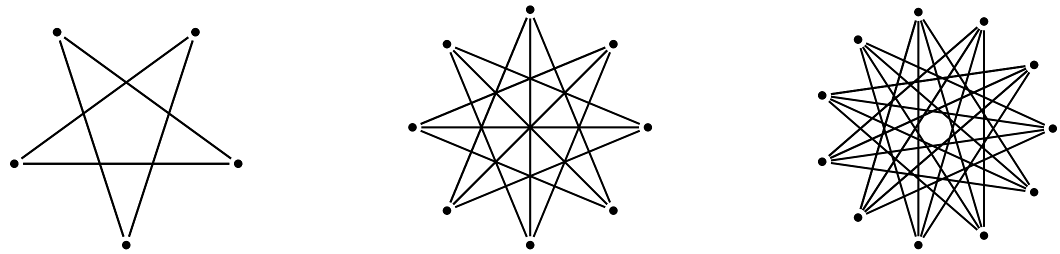


Figure 1: The Andrásfai graphs Γ_2, Γ_3, and Γ_4.

In what follows, given a graph H, we denote by $\overline{H}$ the complement of H. In this paper we use the following property of maximal graphs homomorphic to Γ_d which says, in particular, that an Andrásfai graph with a maximal independent set D admits a 3-coloring having D as one of the color classes.

Lemma 7. *Let G be a maximal graph homomorphic to Γ_d for some $d \in \mathbb{N}$. If I_1 is a maximal independent set of G, then G admits a 3-coloring $\{I_1, I_2, I_3\}$ such that $\overline{G}[I_2 \cup I_3]$ has no induced C_4.*

3 Dense graphs with bounded maximum degree

The proof of Theorem 5 is divided into two steps. First we use that if G is a triangle-free graph on n vertices with independence number 2 and maximum degree less than $19n/29 - 1$, then $\overline{G}$ admits a 3-coloring as in Lemma 7; next, we show that every graph whose complement admits such a 3-coloring contains an immersion of $K_{\chi(G)}$. For that, given a positive integer k, a *k-clique coloring* of a graph G is a partition $\{D_1, \ldots, D_k\}$ of $V(G)$ such that D_i is a clique of G for every $i \in [k]$. For $X, Y \subseteq V(G)$ with $X \cap Y = \emptyset$, we denote by $G[X, Y]$ the bipartite subgraph of G with vertex set $X \cup Y$ and all edges of G between X and Y. Given $X \subseteq V(G)$ and $u \in V(G)$, we use $N_X(u)$ to denote the set of neighbors of u in X. The proof of the next result uses G and its complement $\overline{G}$ at the same time. To avoid confusion, we write $\overline{N}_X(u)$ to refer to the vertices in $X \setminus \{u\}$ that are not adjacent to u in G, and $\overline{N}_X(Y)$ to refer to the union $\bigcup_{u \in Y} \overline{N}_X(u)$, which is the set of vertices in X that are nonadjacent in G to at least one vertex in Y. Observe that $\overline{N}$ is precisely the neighborhood function in $\overline{G}$.

Theorem 8. *Let G be a graph with independence number 2. If G admits a 3-clique coloring $\{D_1, D_2, D_3\}$ such that (i) D_1 is a maximum clique of G; and (ii) $G[D_2 \cup D_3]$ has no induced C_4, then G contains an immersion of a clique whose set of branch vertices is precisely $D_2 \cup D_3$.*

Sketch of the proof. Let $i \in \{2, 3\}$ and let $C \subseteq D_i$. If $|\overline{N}_{D_1}(C)| < |C|$, then the set $(D_1 \setminus \overline{N}_{D_1}(C)) \cup C$ is a clique in G larger than D_1, a contradiction to the maximality of D_1. So, $|\overline{N}_{D_1}(C)| \geq |C|$ for every subset C of D_i. Hence, by Hall's Theorem, there is a matching M_i in $\overline{G}[D_i, D_1]$ that covers D_i. Note that for each vertex $u \in D_2 \cup D_3$ there is precisely one edge in $M_2 \cup M_3$ that contains u, and let $r_u \in D_1$ be the vertex such that $ur_u \in M_2 \cup M_3$.

Note that $r_u \notin N(u)$, and hence, because $\alpha(G) = 2$, r_u is adjacent in G to every non-neighbor of u, i.e., to every vertex in $V(G) \setminus N[u]$. Now, for every $u \in D_2$ and $v \in D_3$ with $uv \notin E' = E\big(G[D_2 \cup D_3]\big)$, let P_{uv} be the path $\langle u, r_v, r_u, v\rangle \subseteq G$. We can prove that the paths P_{uv} with $u \in D_2$, $v \in D_3$, and $uv \notin E'$ are pairwise edge-disjoint. Then, since D_2 and D_3 are cliques, the graph $G[D_2 \cup D_3] \cup \{P_e : e \notin E'\}$ is the desired immersion. □

The next theorem together with Theorem 6 implies Theorem 5. For its proof, we say that a graph G is *k-critical* if $\chi(G) = k$ and $\chi(G - u) < k$, for every $u \in V(G)$.

Theorem 9. *Let G be a graph with independence number at most 2. If the complement of G is homomorphic to Γ_d for some $d \in \mathbb{N}$, then G contains a strong immersion of $K_{\chi(G)}$.*

Sketch of the proof. Let G be a counterexample that minimizes $|V(G)|$, and let $k = \chi(G)$. We can show that G is k-critical and $\overline{G}$ is connected. Then, by a result of Gallai [12] (see [21, Corollary 2]), we conclude that $|V(G)| \geq 2k - 1$. Next we consider a maximal supergraph H of $\overline{G}$ that is homomorphic to Γ_d and a maximum independent set I in H. As $G[I]$ is a clique, we derive that $|I| \leq k - 1$. Applying Lemma 7 on H and I, and then Theorem 8 on $\overline{H}$, which has independence number 2, we can deduce that there is an immersion in G of a clique with $V(G) \setminus I$ as branch vertices. But $|V(G) \setminus I| \geq (2k-1) - (k-1) = k$, a contradiction. □

4 Concluding remarks

In this paper, we explore Conjecture 3 under a maximum degree constraint that allows us to use structural results on triangle-free graphs, and reveals a connection to the chromatic number of their complement. A natural possible improvement in our result is as follows: Brandt and Thomassé [6] proved that triangle-free graphs on n vertices with minimum degree greater than $n/3$ are homomorphic to Vega graphs (which are 4-colorable). This may be used to weaken the bound on $\Delta(G)$ in Theorem 5 to $\Delta(G) < 2n/3 - 1$. We observe that this improvement can be achieved by extending Theorem 8 to graphs whose complement has fractional chromatic number at most 3. Indeed, a triangle-free graph G with minimum degree δ has fractional chromatic number at most n/δ because the family of neighborhoods $\{N(u) : u \in V(G)\}$ with constant weight function $1/\delta$ is a fractional coloring of G. This would reveal a connection to the result of Blasiak [4] mentioned in Section 1. Another interesting direction is to explore larger chromatic numbers of the complement graph by finding longer paths using a broader notion of representatives.

References

[1] Faisal N. Abu-Khzam and Michael A. Langston. Graph coloring and the immersion order. In *International Computing and Combinatorics Conference*, pages 394–403. Springer, 2003.

[2] Béla Andrásfai. Graphentheoretische Extremalprobleme. *Acta Mathematica Hungarica*, 15(3-4):413–438, 1964.

[3] József Balogh and Alexandr V. Kostochka. Large minors in graphs with given independence number. *Discrete Mathematics*, 311(20):2203–2215, 2011.

[4] Jonah Blasiak. A special case of Hadwiger's conjecture. *Journal of Combinatorial Theory, Series B*, 97(6):1056–1073, 2007.

[5] Fábio Botler, Andrea Jiménez, Carla N. Lintzmayer, Adrián Pastine, Daniel A. Quiroz, and Maycon Sambinelli. Biclique immersions in graphs with independence number 2. *European Journal of Combinatorics*, 122:104042, 2024.

[6] Stephan Brandt and Stéphan Thomassé. Dense triangle-free graphs are four-colorable: A solution to the Erdös-Simonovits problem. Available at `https://perso.ens-lyon.fr/stephan.thomasse/liste/vega11.pdf`, 2011.

[7] Chuan-Chong Chen, Guoping P. Jin, and Khee Meng Koh. Triangle-free graphs with large degree. *Combinatorics, Probability and Computing*, 6(4):381–396, 1997.

[8] Rong Chen and Zijian Deng. A simple proof of the existence of complete bipartite graph immersion in graphs with independence number two. Available at SSRN: `https://ssrn.com/abstract=4882255` or `http://dx.doi.org/10.2139/ssrn.4882255`, 2024.

[9] Maria Chudnovsky and Paul Seymour. Packing seagulls. *Combinatorica*, 32:251–282, 2012.

[10] Pierre Duchet and Henry Meyniel. On Hadwiger's number and the stability number. In *North-Holland Mathematics Studies*, volume 62, pages 71–73. Elsevier, 1982.

[11] Jacob Fox. Complete minors and independence number. *SIAM Journal on Discrete Mathematics*, 24(4):1313–1321, 2010.

[12] Tibor Gallai. Kritische graphen II. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*, 8:373–395, 1963.

[13] Gregory Gauthier, Tien-Nam Le, and Paul Wollan. Forcing clique immersions through chromatic number. *European Journal of Combinatorics*, 81:98–118, 2019.

[14] Hugo Hadwiger. Über eine Klassifikation der Streckenkomplexe. *Vierteljschr. Naturforsch. Ges. Zürich*, 88(2):133–142, 1943.

[15] Roland Häggkvist. Odd cycles of specified length in non-bipartite graphs. In *North-Holland Mathematics Studies*, volume 62, pages 89–99. Elsevier, 1982.

[16] Guoping Jin. Triangle-free four-chromatic graphs. *Discrete Mathematics*, 145(1-3):151–170, 1995.

[17] François Lescure and Henry Meyniel. On a problem upon configurations contained in graphs with given chromatic number. *Annals of Discrete Mathematics*, 41:325–331, 1985.

[18] Daniel A. Quiroz. Clique immersions in graphs of independence number two with certain forbidden subgraphs. *Discrete Mathematics*, 344(6):112365, 2021.

[19] Daniel A. Quiroz. Private communication, nov. 2024.

[20] Neil Robertson, Paul Seymour, and Robin Thomas. Hadwiger's conjecture for K_6-free graphs. *Combinatorica*, 13(3):279–361, 1993.

[21] Matěj Stehlík. Critical graphs with connected complements. *Journal of Combinatorial Theory, Series B*, 89(2):189–194, 2003.

[22] Sylvia Vergara. Complete graph immersions in dense graphs. *Discrete Mathematics*, 340(5):1019–1027, 2017.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Extremal Problems on Forest Cuts and Acyclic Neighborhoods in Sparse Graphs*

(Extended abstract)

F. Botler[†] Y. S. Couto[†] C. G. Fernandes[†] E. F. de Figueiredo[‡]
R. Gómez[§] V. F. dos Santos[‡] C. M. Sato[§]

Abstract

Chernyshev, Rauch, and Rautenbach proved that every connected graph G on n vertices for which $e(G) < \frac{11}{5}n - \frac{18}{5}$ has a vertex cut that induces a forest, and conjectured that the same remains true if $e(G) < 3n - 6$ edges. We improve their result by proving that every connected graph on n vertices for which $e(G) < \frac{9}{4}n - \frac{15}{4}$ has a vertex cut that induces a forest. We also study weaker versions of the problem that might lead to an improvement on the bound obtained.

1 Introduction

Let G be a connected graph. A set $S \subset V(G)$ is a *vertex cut* if $G - S$ is disconnected. If $|S| = k$, we say S is a *k-vertex cut*. If S is an independent set, we say S is an *independent cut*. Vertex cuts with special properties have been studied in different contexts. Chen and Yu [1] showed that every connected graph with less than $2n-3$ edges has an independent cut, confirming a conjecture due to Caro. Recently, Chernyshev, Rauch, and Rautenbach proposed the following analogue conjecture, replacing independent set by forest [2, Conjecture 1]. A *forest cut* is a vertex cut that induces a forest.

*This work started during the 6th WoPOCA (Workshop Paulista em Otimização, Combinatória e Algoritmos) in Campinas, Brazil. We thank the organizers for the opportunity to start new collaborations and the agencies that helped making the workshop possible: Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (404315/2023-2) and FAEPEX (2422/23). F. Botler would also like to thank the support of CNPq (304315/2022-2), CAPES (88887.878880/2023-00) and the São Paulo Research Foundation – FAPESP (2024/14906-6). Y.S. Couto would also like to thank the support of FAPESP (2024/18049-0). C.G. Fernandes would also like to thank the support of CNPq (310979/2020-0). V.F. dos Santos would also like to thank the support of CNPq (312069/2021-9, 406036/2021-7 and 404479/2023-5) and FAPEMIG (APQ-01707-21). C.M. Sato would also like to thank the support of CNPq (408180/2023-4). E.F. de Figueiredo would also like to thank the support of CAPES (Finance Code 001).

[†]Institute of Mathematics and Statistics, University of São Paulo, Brazil. E-mail: `{fbotler,cris,yancouto}@ime.usp.br`.

[‡]Department of Computer Science, Federal University of Minas Gerais, Brazil. E-mail: `{ederfgd,viniciussantos}@dcc.ufmg.br`.

[§]Center for Mathematics, Computing and Cognition, Federal University of ABC, Brazil. E-mail: `{gomez.renzo,c.sato} @ufabc.edu.br`.

Conjecture 1 (Chernyshev–Rauch–Rautenbach, 2024)**.** *If G is a connected graph on n vertices with no forest cut, then $e(G) \geq 3n - 6$.*

Chernyshev et al. [2] also showed that Conjecture 1 holds for some classes of graphs. For instance, they showed that a graph G with n vertices has a forest cut if (i) G is a planar graph that is not triangulated; (ii) G has a universal vertex and $e(G) < 3n - 6$; or (iii) G is connected and $e(G) < \frac{11}{5}n - \frac{18}{5}$.

We say a graph is *k-cyclic* if every vertex set of size at most k is dominating or has a cycle in its neighborhood. Note that any (forest) cut disconnects the graph into at least two components, which are not dominating sets, and one of these components has less than $n/2$ vertices. So, Conjecture 1 claims that any $(\frac{n-1}{2})$-cyclic graph has at least $3n - 6$ edges. Moreover, any 2-vertex cut is trivially a forest, so Chernyshev et al. [2] noted that finding good lower bounds for the number of edges on 1-cyclic 3-connected graphs would imply a result towards Conjecture 1, and stated the following.

Conjecture 2 (Chernyshev–Rauch–Rautenbach, 2024)**.** *If G is a 3-connected graph on n vertices such that there is a cycle in the neighborhood of every vertex, then $e(G) \geq \frac{7}{3}n - \frac{7}{3}$.*

The conjecture addresses a proper subclass of 1-cyclic graphs as it requires cycles in the neighborhood of universal vertices. However, it is functionally the same as for 1-cyclic graphs, as even Conjecture 1 holds for graphs with universal vertices [2]. In this paper, we improve the bound from [2] towards Conjecture 1, disprove Conjecture 2, and present lower bounds on the number of edges for 3-connected graphs to be 1-cyclic and 2-cyclic.

Theorem 3. *Let G be a graph on n vertices. Then the following hold. (a) If G is connected and has no forest cut, then $e(G) \geq \frac{9}{4}n - \frac{15}{4}$; (b) If G is 3-connected, 1-cyclic, and $n \geq 6$, then $e(G) \geq \frac{15}{8}n$; (c) If G is 3-connected, 2-cyclic, and $n \geq 6$, then $e(G) \geq 2n$.*

The $n \geq 6$ is necessary in Theorem 3(b) and 3(c) as K_5 minus an edge is 3-connected and 2-cyclic (hence also 1-cyclic), has five vertices and nine edges, but $9 < \frac{15}{8} \cdot 5 = \frac{75}{8} < 10$.

Remark 4. *There are infinite families of (a) 3-connected 1-cyclic graphs on n vertices with exactly $\frac{15n}{8}$ edges and no universal vertices; (b) 4-connected 1-cyclic graphs on n vertices with exactly $2n$ edges; (c) 3-connected 2-cyclic graphs on n vertices with exactly $\frac{9}{4}n$ edges; (d) 4-connected 2-cyclic graphs on n vertices with exactly $\frac{7}{3}n$ edges.*

Remark 4(a) disproves Conjecture 2, proving that Theorem 3(b) is asymptotically tight. For Theorem 3(c), we present a 3-connected 2-cyclic graph and a 4-connected 2-cyclic graph, both with 6 vertices and 12 edges, and, based on Remark 4(d), we pose the following conjecture that would imply an improvement on Theorem 3(a), towards Conjecture 1.

Conjecture 5. *If G is a 4-connected 2-cyclic graph on $n \geq 9$ vertices, then $e(G) \geq \frac{7}{3}n$.*

In Section 2, we prove Theorem 3(a). In Section 3, we prove Theorem 3(b)-(c), and Remark 4. A recent independent work by Li, Tang, and Zhan [3] contains results similar to the ones on 1-cyclic graphs in Section 3. Due to space constraints, we omit a few proofs.

2 Avoiding forest cuts

Chernyshev et al. [2] proved that a connected graph on n vertices with no forest cut must have at least $\frac{11n}{5} - \frac{18}{5}$ edges. For that, they studied properties its counterexamples with a minimum number of vertices. Such properties are in fact shared with a minimum counterexample to Theorem 3(a) and Conjecture 1. To help the exposition, we state a conjecture parameterized by a number α with $2 \leq \alpha \leq 3$.

Conjecture 6 (α-FC Conjecture)**.** If G is a connected graph on n vertices with no forest cut, then $e(G) \geq \alpha(n-3)+3$.

Note that Theorem 3(a) is the same as the $\frac{9}{4}$-FC Conjecture, Chernyshev et al. [2] proved the $\frac{11}{5}$-FC Conjecture and Conjecture 1 is the same as the 3-FC Conjecture. For $2 \leq \alpha \leq 3$, a minimum counterexample to the α-FC Conjecture is a graph G on n vertices with no forest cut, $e(G) < \alpha(n-3)+3$ and n as small as possible. The following lemma is used in the proof of Theorem 3(a).

Lemma 7. *Let G be a minimum counterexample to the α-FC Conjecture, for $2 \leq \alpha \leq 3$. Then (a) G is 4-connected and has at least 8 vertices; (b) no degree-4 vertex in G has a C_4 in its neighborhood; and (c) no two degree-4 vertices are in the same K_4 in G.*

Lemma 7(a) was adapted from the proof of Claim 1 in Chernyshev et al. [2]. They [2, Claim 2] also proved that, in a minimum counterexample to Conjecture 1, every degree-4 vertex has at most two neighbors of degree 4. Lemma 7(b) and 7(c) are strengthenings of this statement. Lemma 7(b) implies that every degree-4 vertex in a minimum counterexample to the α-FC Conjecture lies in a K_4, and we deduce the following from Lemma 7(c).

Corollary 8. *Let G be a minimum counterexample to the α-FC Conjecture, for $2 \leq \alpha \leq 3$. Then the following hold: (a) every degree-4 vertex in G has at most one degree-4 neighbor; and (b) each vertex with degree at least 5 in G has at least two neighbors of degree at least 5.*

Corollary 8(b) is also a strengthening of a result of Chernyshev et al. [2, Claim 3]. We conclude this section with the proof of Theorem 3(a).

Proof of Theorem 3(a). Suppose G is a minimum counterexample to Theorem 3(a), and hence to the $\frac{9}{4}$-FC Conjecture. Let n be the number of vertices of G, and n_i be the number of degree-i vertices in G. By Lemma 7(a), G is 4-connected and $n = \sum_{i=4}^{n-1} n_i \geq 8$. Let F_4 be the set of edges joining degree-4 vertices to vertices with degree at least 5. By Corollary 8(a), we have that $|F_4| \geq 3n_4$. By Corollary 8(b), each degree-j vertex in G with $j \geq 5$ contributes with at most $j-2$ edges to F_4, and hence $|F_4| \leq \sum_{j=5}^{n-1}(j-2)n_j$. Now, since $j-2 \leq 6j-27$ for $j \geq 5$, we have $3n_4 \leq \sum_{j=5}^{n-1}(j-2)n_j \leq \sum_{j=5}^{n-1}(6j-27)n_j = 6(2e(G)-4n_4)-27(n-n_4) = 12e(G)+3n_4-27n$, so $e(G) \geq 9n/4$, a contradiction. □

3 Bounds for 1-cyclic and 2-cyclic graphs

First, we present a family of counterexamples to Conjecture 2 and prove Remark 4(a). Take any 3-connected 3-regular graph (see [4]) with k vertices and replace each vertex with a K_4, connecting each of its neighbors to a distinct vertex in the K_4 and leaving only one vertex of each K_4 with degree 3 (see, e.g., Figure 1). We obtain a 3-connected graph G with precisely $n = 4k$ vertices and $m = \frac{3k}{2} + 6k = \frac{15}{8}n$ edges. Moreover, G is 1-cyclic because each of its vertices is in a K_4.

Remark 4(a) shows that Theorem 3(b) is tight. We denote by $K_s^{\triangle}$ the graph obtained from K_3 by adding s new vertices adjacent to the three vertices of the K_3. The proof of Theorem 3(b) uses the following lemma, whose proof we omit.

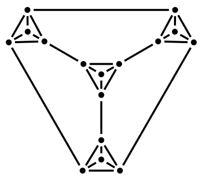


Figure 1: A counterexample to Conjecture 2 built from K_4.

Lemma 9. *If G is a 3-connected 1-cyclic graph on $n \geq 5$ vertices. Then the following hold: (a) every degree-3 vertex has no degree-3 neighbor; and (b) either G is isomorphic to $K_{n-3}^{\triangle}$ or every vertex of G has at least three neighbors of degree at least 4.*

Proof of Theorem 3(b). Let G be a 3-connected 1-cyclic graph on $n \geq 6$ vertices, and n_i be the number of degree-i vertices in G. By Lemma 9(b), either G is isomorphic to $K_{n-3}^{\triangle}$ or every vertex of G has at least three neighbors of degree at least 4. In the former case, as desired, $e(G) = 3n - 6 > \frac{15}{8}n$ as $n \geq 6$. In the latter case, as $4j - 15 \geq j - 3$ for $j \geq 4$, we have $3n_3 \leq \sum_{j=4}^{n-1}(j-3)n_j \leq \sum_{j=4}^{n-1}(4j-15)n_j = 8e(G) - 15n + 3n_3$, i.e., $e(G) \geq \frac{15}{8}n$. □

Note that if we pick an arbitrary 4-connected 4-regular graph and replace each of its vertices by a K_4, leaving all vertices of each K_4 with degree 4, then the graph obtained is 4-connected, 4-regular, and 1-cyclic. Therefore, the lower bound $e(G) \geq 2n$ is best possible for 4-connected 1-cyclic graphs, and proves Remark 4(b). Now, we prove a lower bound on the number of edges for a 3-connected graph to be 2-cyclic. Specifically, we prove Theorem 3(c). We start by proving some properties of 3-connected 2-cyclic graphs.

Lemma 10. *Let G be a 3-connected 2-cyclic graph on $n \geq 6$ vertices. Then every degree-3 vertex has at least two neighbors of degree at least 5.*

Proof. Let v be a degree-3 vertex in G, and x, y, and z be its neighbors. By Lemma 9(a), these three vertices have degree at least 4, and they form a triangle, because $n \geq 5$ and G is 1-cyclic. Suppose, for a contradiction, that x and y have degree 4. Then the neighborhood $N(\{v, x\}) = \{y, z, w\}$, where w is the other neighbor of x. As $n \geq 6$ and G is 2-cyclic, y, x, w form a triangle, and w is also the other neighbor of y. But then $N(\{x, y\}) = \{v, z, w\}$, which must form a cycle because $n \geq 6$. However there is no edge vw, a contradiction. □

Proof of Theorem 3(c). Let n_i be the number of degree-i vertices in G and F be the set of edges joining degree-3 vertices to vertices with degree at least 5. By Lemma 10, we have that $|F| \geq 2n_3$. By Lemma 9(b), either G is isomorphic to $K_{n-3}^{\triangle}$ or every vertex of G has at least three neighbors of degree at least 4. In the former case, G has $3n - 6 \geq 2n$ edges as $n \geq 6$. In the latter case, each degree-j vertex for $j \geq 5$ contributes with at most $j - 3$ edges to F, so $|F| \leq \sum_{j=5}^{n-1}(j-3)n_j$. As $2j - 8 \geq j - 3$ for $j \geq 5$, we have $2n_3 \leq |F| \leq \sum_{j=5}^{n-1}(j-3)n_j \leq \sum_{j=5}^{n-1}(2j-8)n_j = 4e(G) - 8n + 2n_3$, i.e., $e(G) \geq 2n$. □

In Figure 2, on the left, we show two tight examples for Theorem 3(c): the graph $K_3^{\triangle}$, which is 3-connected, and the octahedral graph, which is 4-connected. The third graph in Figure 2 has 9 vertices and 20 edges. Consider the construction illustrated in Figure 1, starting from a 3-connected 3-regular graph on k vertices. If we replace each vertex by an octahedral graph instead of a K_4, we end up with a 3-connected 2-cyclic graph on $6k$ vertices and $\frac{3}{2}k + 12k = \frac{27}{2}k = \frac{9}{4}n$ edges, which proves Remark 4(c). As far as we know, it may hold that $m \geq \frac{9}{4}n$ for the graphs addressed by Theorem 3(c) if $n \geq 10$. The requirement $n \geq 10$ is necessary to exclude the third graph in Figure 2, because $\frac{20}{9} < \frac{9}{4}$.

The lower bound on the number of edges in a 4-connected 2-cyclic graph might be larger. Take a 4-connected 4-regular graph on k vertices, and replace each of its vertices by an octahedral graph, leaving precisely four vertices of each octahedral graph with degree 5. The graph obtained is 4-connected, 2-cyclic, has $6k$ vertices and $m = 2k + 12k = 14k = \frac{7}{3}n$ edges. This proves Remark 4(d), which shows that Conjecture 5 is tight. In Figure 2, on the right, we show a 4-connected 2-cyclic graph on 7 vertices and 16 edges, and two 4-connected 2-cyclic graphs with 8 vertices and 18 edges. Since $\frac{16}{7}$ and $\frac{18}{8}$ are less than $\frac{7}{3}$, these examples justify the condition $n \geq 9$ in Conjecture 5.

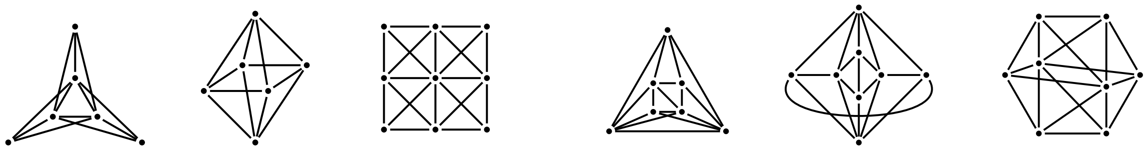


Figure 2: Left: Three 3-connected 2-cyclic graphs, two with 6 vertices and 12 edges and one with 9 vertices and 20 edges. Right: Three 4-connected 2-cyclic graphs, one with 7 vertices and 16 edges, and two with 8 vertices and 18 edges.

4 Final remarks

Several questions remain open. Of course it would be nice to settle Conjecture 1, or to obtain an improvement on Theorem 3(a). Proving Conjecture 5 or finding a family of 4-connected 2-cyclic graphs on n vertices with less than $\frac{7}{3}n$ edges would also be interesting.

The study of k-cyclic graphs with k more than 2 seems to be a possible way to achieve better results towards Conjecture 1. Our exposition points out that we barely use the forest cut requirement for sets larger than 2 in the current results.

References

[1] Guantao Chen and Xingxing Yu. A note on fragile graphs. *Discrete mathematics*, 249(1-3):41–43, 2002.

[2] Vsevolod Chernyshev, Johannes Rauch, and Dieter Rautenbach. Forest cuts in sparse graphs. *Available at arXiv:2409.17724*, 2024.

[3] Chengli Li, Yurui Tang, and Xingzhi Zhan. The minimum size of a 3-connected locally nonforesty graph. *Available at arXiv:2410.23702*, 2024.

[4] Nicholas C. Wormald. The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):156–167, 1981.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Rainbow Path Covers of Sparse Random Graphs

(Extended abstract)

F. Botler* A. Kaique† W. Mendonça† G. O. Mota† C. M. Sato‡

Abstract

We investigate the problem of covering the edges of the random graph $G(n,p)$ with rainbow paths, where the coloring is any proper edge-coloring. We show that for $pn = o((\log n/\log\log n)^2)$, this can be achieved with $O(n)$ rainbow paths with high probability, which is essentially best possible. Our techniques may offer insights for determining minimum rainbow path covers in arbitrary properly-colored graphs.

1 Introduction

A *path decomposition* of a graph G is a family of edge-disjoint paths of G that covers the edge set of G. In 1966, Erdős asked about the minimum size of a path decomposition of a graph G, for which Gallai conjectured that every graph on n vertices admits a path decomposition of size at most $\lceil n/2\rceil$ (see [12]). Several results followed this conjecture: Chung [5] showed a tree decomposition with $\lceil n/2\rceil$ trees; Lovász [12] showed that every graph admits a path decomposition of size at most $n-1$; Donald [8] improved the upper bound to $\lfloor 3n/4\rfloor$; and Dean and Kouider [7] further improved it to $\lfloor 2n/3\rfloor$, which is the best known bound. Other results consider specific graph classes, including planar graphs [1, 3] and graphs with sparse subgraphs induced by even-degree vertices [4, 10, 13].

In order to approach Gallai's Conjecture, which remains open, one may consider path covers instead of path decompositions, where a *path cover* of a graph G is a family of (not necessarily edge-disjoint) paths of G that covers the edge set of G. Such a weakening of Gallai's Conjecture was posed by Chung [6], who conjectured that every graph on n vertices admits a path cover of size at most $\lceil n/2\rceil$. In 1996, Pyber [13] verified this conjecture asymptotically

*Institute of Mathematics and Statistics, University of São Paulo. E-mail: `{fbotler, antoniokaique,mota}@ime.usp.br`. F. Botler was supported by CNPq (304315/2022-2) and CAPES (88887.878880/2023-00). A. Kaique was supported by CAPES (88882.328155/2015-01). G. O. Mota was supported by CNPq (306620/2020-0, 406248/2021-4) and FAPESP (2023/03167-5 and 2024/13859-4).

†Mathematics Department, Federal University of Ceará. E-mail: `walner@mat.ufc.br`. Supported by FAPESP (2023/07695-6).

‡Center for Mathematics, Computing and Cognition, Federal University of ABC. E-mail: `c.sato@ufabc.edu.br`. Supported by FAPESB (EDITAL FAPESB Nº 012/2022 - UNIVERSAL - NºAPP0044/2023). FAPESB is the Bahia Research Foundation.

by proving that every graph can be covered by $n/2 + O(n^{3/4})$ paths; and in 2002, Fan [9] verified it tightly.

We now present an analog of the path cover conjecture concerning the rainbow paths in graphs equipped with a proper edge-coloring. Let G be graph and let $c\colon E(G) \to \mathbb{N}$ be a *proper* edge-coloring of G, i.e., a coloring of $E(G)$ in which $c(e) \neq c(f)$ for every pair of edges incident to the same vertex. We say that a subgraph H of G is *rainbow* if $c(e) \neq c(f)$ for every pair e, f of distinct edges of H. In what follows, to avoid excessive wording, we simply say *colored* instead of properly edge-colored. We say that path cover $\mathcal{P} = \{P_1, \dots, P_k\}$ of a colored graph G is a *rainbow path cover* of G if P_i is a rainbow path (with respect to c) for every $i \in [k] = \{1, \dots, k\}$. We are interested in the following problem posed by Bonamy, Botler, Dross, Naia, Skokan [2], which we present here as a conjecture.

Conjecture 1 (Bonamy–Botler–Dross–Naia–Skokan, 2023)**.** *Every colored graph G admits a rainbow path cover of size $O\big(|V(G)|\big)$.*

Although Conjecture 1 first arose as a possible strategy to handle a problem answered in [2], we believe it is interesting in its own right and it has not been thoroughly explored in the literature yet. A natural first step is thus to find evidence for Conjecture 1 among random graphs. More specifically, we would like to know for which ranges of p does every edge-coloring of $G(n,p)$ asymptotically almost surely (a.a.s.) admit a rainbow path cover of size $O(n)$. In a very sparse regime, Conjecture 1 is naturally true due to the number of edges being linear. On the other hand, Kaique, Hoppen, Mendonça, Mota and Naia [11] verified Conjecture 1 for dense random graphs as follows.

Theorem 2 (Kaique–Hoppen–Mendonça–Mota–Naia, 2025)**.** *Let $\varepsilon > 0$ and $p \geq n^{\varepsilon - 1}$. Then there exists $D = D(\varepsilon) > 0$ such that the following holds with high probability in $G = G(n,p)$. For any proper edge-coloring of G, it is possible to cover $E(G)$ with Dn rainbow paths.*

In this paper, we extend Theorem 2 to a sparse range of p. This range is significant as it is dense enough to require a non-trivial cover while still lacking the high connectivity present in the dense range that aided path finding.

Theorem 3. *There is a constant D so that if $pn = o((\log n/ \log\log n)^2)$ and $G = G(n,p)$, then the following holds with high probability: for any proper edge-coloring of G, it is possible to cover $E(G)$ with Dn rainbow paths.*

In section 2, we present the main tools and techniques used in the proof of Theorem 3 and in section 3, we present an outline of the proof of Theorem 3.

2 Main tools

2.1 Building random rainbow trails

One of the main tools we use to build rainbow paths is a randomized procedure that builds n rainbow trails in parallel while traversing every edge twice. Let $G = (V, E)$ be a graph and c be a proper edge-coloring of G. Let T denote the number of colors used in c and let $t \leq T$ be an integer. The procedure starts by creating trails $W(u) = (u)$ for every $u \in V$. We will also keep track of auxiliary walks that will be useful in the analysis of the procedure: let $W'(u) = (u)$.

Procedure 1: Repeat the following for t steps:

- Choose a color uniformly at random among the colors not chosen so far;
- For each vertex $u \in V$, if there is an edge vv' with the chosen color, where v is the last vertex in $W(u)$, append v' to $W(u)$ and to $W'(u)$; otherwise, append v to $W'(u)$.

For each vertex u, the auxiliary walk $W'(u)$ keeps track of all steps, including the ones where the trail $W(u)$ was not extended. This simple procedure has many nice features: the trails $(W(u))_{u\in V}$ are rainbow; at each step, each of the trails is at a different vertex; and, for $t = T$, each edge is traversed exactly twice. Note that, if the goal was to cover edges with rainbow trails instead of paths, the linear upper bound would follow directly from this procedure. In order to obtain paths, the main obstacle is that the trails may contain cycles. In the next section, we present the properties of short cycles in $G(n, p)$ that we will use in our proof.

2.2 Short cycles in $G(n, p)$

In order to avoid too many cycles in the rainbow trails, which disrupts our construction, our strategy is to bound the number of short cycles in $G(n, p)$, which allows us to remove a small number of edges and increase the girth of the graph. Let Y_ℓ denote the number of cycles of length ℓ in $G(n, p)$. Let $Y_{\leq \ell}$ denote the number of cycles of length at most ℓ in $G(n, p)$. Let $d = pn$. It is well known that $\mathbb{E}(Y_\ell) \sim d^\ell/(2\ell)$ for $d \gg 1$. Using this fact, it is easy to show the following results on short cycles in $G(n, p)$:

Proposition 4. *Let $p = d/n$ with $1 \ll d \ll (\log n/ \log\log n)^2$. We have that $Y_{\leq\ell} \ll n$ a.a.s., for $\ell = \frac{\log n}{\log d}$.*

Proposition 5. *Let $\alpha > 1$ be a constant. Let $p = d/n$ with $1 \ll d \leq \alpha \log n/ \log\log n$. Then $Y_{\leq\ell} \ll n$ a.a.s., for $\ell = d/\alpha$.*

2.3 Large girth graphs

As mentioned before, the rainbow trails built by Procedure 1 can have many cycles, which is an obstacle for building our path cover. To address this, we will slice each trail into shorter segments of length at most $\ell \sim \varepsilon d$, where ε is a small positive constant. This approach means we only need to worry about a trail coming back to itself within a window of at most ℓ steps. For each vertex u, we define the following:

- Let $W(u)$ be the trail built by Procedure 1 starting at u and let $T(u)$ be the number of edges in $W(u)$. Let $(w_j(u))_{0\leq j\leq T(u)} := W(u)$
- Let $W'(u) = (w'_i(u))_{0\leq i\leq T}$ be the auxiliary walk for u.
- For each step i of Procedure 1, let $J(u, i)$ denote the number of edges in the trail $W(u)$ built at that step. This implies that $w'_i(u) = w_{J(u,i)}(u)$.
- For every $i \in \{1, \ldots, T - \ell\}$, let $P_i(u)$ be the trail in $W(u)$ with ℓ edges ending at $w'_i(u)$. More precisely, $P_i(u) := (w_j(u))_{j_0\leq j\leq J(u,i)}$ where $j_0 = \max(0, J(u, i) - \ell)$.
- For every $i \in \{1, \ldots, T - \ell\}$, let $\mathrm{Split}_i(u)$ be the event that $w'_{i+1}(u)$ is different from $w'_i(u)$ and is in $P_i(u)$. We say that $\mathrm{Split}_i(u)$ is a *splitting event.*

One of our main lemmas will be used to show that, for high girth graphs, Procedure 1 causes $o(1)$ splitting events in expectation for each vertex (as long as some conditions are met). This means that we have $o(n)$ splitting events overall, with high probability.

Lemma 6. *Let $G = (V, E)$ be a graph with n vertices and m edges. Let $\bar{d} := 2m/n$ and fix $\ell \sim \varepsilon\bar{d}$. Let $t = t(n) < T$ and $k = k(n)$ be such that $T - t = \omega(k)$. Suppose that the girth of G is greater than k. For every vertex $u \in V$, the expected number of occurrences of Split$_i(u)$ for $i \le t$ is $O\left((\bar{d}/k^2)\log\left(T/(T-t)\right)\right)$.*

Proof sketch. Fix $u \in V$. We will omit u in the notation for simplicity (e.g., $\mathrm{Split}_i(u)$ will be referred to simply as Split_i). We also omit ceiling and floor operators as long as they do not affect the asymptotic bounds. Our goal is to bound $\sum_{i \le t}\Pr[\mathrm{Split}_i]$. We start by partitioning $\{1, \dots, t\}$ into intervals of size $k' := k/2$: for every integer $0 \le j \le t/k'$, define the interval $I_j := [jk', \min((j+1)k' - 1, t)]$. Given $0 \le j \le t/k'$, we show that

$$\textstyle\sum_{i \in I_j} \Pr\left[\mathrm{Split}_i\right] \le \ell/(k'(T - jk' - k')). \tag{1}$$

Equation (1) holds for the following reasons. First, the probability that a splitting step occurs at step i is bounded above by $d(w'_i, P_i)/(T-i)$, where $d(w'_i, P_i)$ is the number of edges from w'_i to P_i. Now we explain the role of the lower bound k on the girth. Let $i_0 = jk'$ be the smallest i in I_j. Let $V_j := \{w'_i\}_{i \in I_j}$ and let $d(V_j, P_{i_0})$ denote the number of edges from V_j to P_{i_0}. Since V_j is connected by a trail of length k', it induces no other edges because that would create a cycle of length less than k. Thus, $\sum_{i \in I_j} d(w'_i, P_i) \le \sum_{i \in I_j} d(w'_i, P_{i_0}) = d(V_j, P_{i_0})$. Since the girth is greater than k, the number of edges from V_j to any k' consecutive vertices in P_{i_0} is at most 1. This implies that $d(V_j, P_{i_0}) \le \ell/k'$ since P_{i_0} has ℓ edges. From these facts, it is easy to derive (1), from which it is straightforward to show that

$$\begin{aligned}
\sum_{i \le t}\Pr[\mathrm{Split}_i] &= \sum_{j=0}^{t/k'}\sum_{i \in I_j}\Pr[\mathrm{Split}_i] \le \sum_{j=0}^{t/k'}\frac{\ell}{k'(T - jk' - k')} = \frac{\ell}{(k')^2}\sum_{j=0}^{t/k'}\frac{1}{T/k' - j - 1} \\
&\le \frac{\ell}{(k')^2}\Big(\mathrm{H}(T/k' - 1) - \mathrm{H}((T-t)/k' - 2)\Big) \le \frac{\ell}{(k')^2}\log\Big(\frac{T}{T-t}\Big)\Theta(1),
\end{aligned}$$

where $\mathrm{H}(x) = \sum_{i=1}^{x} 1/i$ is the x-th Harmonic number.

□

3 Proof overview

In this section we outline the proof of Theorem 3. In the range $pn = O(\log n/\log\log n)$, it is possible to remove $o(n)$ edges from $G(n,p)$ obtain a graph with very large girth a.a.s., preventing splitting events. This can be done by applying Proposition 5 and directly analysing the sparser range where Proposition 5 does not apply. So assume $pn = \Omega(\log n/\log\log n)$. Let $d := pn$ and $G = G(n,p)$. We have that, a.a.s., the number of edges m of $G(n,p)$ satisfies $m \sim \binom{n}{2}p$ and so $\bar{d} := \frac{2m}{n} \sim d$. By Proposition 4, the number of cycles in G of length at most k is $o(n)$ a.a.s., for $k := \log n/\log d$. Sample $G = G(n,p)$ and assume that G satisfies these conditions. Let $c : E(G) \to [T]$ be any proper edge-coloring of G.

Let $E_{\mathrm{cycle}} \subseteq E(G)$ be a set of edges of size $o(n)$ containing one edge from each cycle of length at most k. Let $\mathcal{P}_0$ denote the set of paths such that each path consists of a single edge of E_{cycle}. Clearly, $|\mathcal{P}_0| = |E_{\mathrm{cycle}}| = o(n)$ and $\mathcal{P}_0$ is a collection of rainbow paths.

Let $G' := G - E_{\mathrm{cycle}}$, which has girth greater than k. Let $m' \le m$ denote the number of edges of G' and let $d' = 2m'/n$. The number of colors in G' may be of the same order of k, so we raise the number of colors as follows. Choose $\psi(n)$ edges of G' and assign each of them a new color. (Any large enough $\psi(n) = o(n)$ works here.) Let c' denote the coloring obtained by restricting c to the edges of G' and recoloring $\psi(n)$ edges to distinct new colors, and let $T' \geq \psi(n)$ denote the number of colors used by c'.

Let $\phi = o(1)$. Here we also need ϕ to be large enough so that some properties hold. Let $t' := (1 - \phi/k)T'$. We have that $(T' - t')/k = \omega(1)$ since $T' \geq \psi(n)$. Let $W_1, \ldots, W_n$ be the random trails obtained by Procedure 1 for (G', c', t'). We partition them into trails of length $\ell = \varepsilon d'$, yielding a collection $\mathcal{P}'$ of $O(n)$ rainbow trails. By applying Lemma 6 to graph G' and c', we have that the expected total number of splitting events in the trails in $\mathcal{P}'$ is $O\left((d'n/k^2)\log\left(T'/(T'-t')\right)\right) \ll n$. Thus, by Markov's inequality, the number of splitting events is $o(n)$ a.a.s. By splitting the trails in $\mathcal{P}'$ whenever a splitting event occurs, we obtain $O(n)$ rainbow paths for c'. For each edge with a new color, we may have to split the paths using this edge since they may not be rainbow with respect to c. Since the number of new colors is $\psi(n) = o(n)$ and each edge appears in at most two paths, this causes $o(n)$ splits. Thus, we obtain a collection of $\mathcal{P}_1$ of $O(n)$ rainbow paths for G'.

The expected number of edges E'' not covered by $W_1, \ldots, W_n$ is $o(m'/k^2) = o(n)$ and so $|E''| = o(n)$ a.a.s. Let $\mathcal{P}_2$ denote the set of paths such that each path consists of a single edge of E''. Then $\mathcal{P}_0 \cup \mathcal{P}_1 \cup \mathcal{P}_2$ is a rainbow path cover of size $O(n)$ for G with respect to c.

4 Future directions

There are many directions left to be explored. Naturally, covering the gap between our current bound of $pn = o((\log n/\log\log n)^2)$ and the denser regime of $pn = n^{\varepsilon}$ is a problem of interest, as is determining the exact constant in the asymptotic bound. Additionally, the techniques and properties established in our random graph analysis could provide valuable approaches to prove (or disprove) Conjecture 1, which concerns the minimum number of rainbow paths needed to cover arbitrary properly-colored graphs.

References

[1] A. Blanché, M. Bonamy, and N. Bonichon. Gallai's path decomposition for planar graphs. In *Extended Abstracts EuroComb 2021: European Conference on Combinatorics, Graph Theory and Applications*, pages 758–764. Springer, 2021.

[2] M. Bonamy, F. Botler, F. Dross, T. Naia, and J. Skokan. Separating the edges of a graph by a linear number of paths. *Adv. Comb.*, pages Paper No. 6, 7, 2023.

[3] F. Botler, A. Jiménez, and M. Sambinelli. Gallai's path decomposition conjecture for triangle-free planar graphs. *Discrete Math.*, 342(5):1403–1414, 2019.

[4] F. Botler and M. Sambinelli. Towards Gallai's path decomposition conjecture. *J. Graph Theory*, 97(1):161–184, 2021.

[5] F. R. K. Chung. On partitions of graphs into trees. *Discrete Math.*, 23(1):23–30, 1978.

[6] F. R. K. Chung. On the coverings of graphs. *Discrete Math.*, 30(2):89–93, 1980.

[7] N. Dean and M. Kouider. Gallai's conjecture for disconnected graphs. volume 213, pages 43–54. 2000. Selected topics in discrete mathematics (Warsaw, 1996).

[8] A. Donald. An upper bound for the path number of a graph. *J. Graph Theory*, 4(2):189–201, 1980.

[9] G. Fan. Subgraph coverings and edge switchings. *J. Combin. Theory Ser. B*, 84(1):54–83, 2002.

[10] G. Fan. Path decompositions and Gallai's conjecture. *J. Combin. Theory Ser. B*, 93(2):117–125, 2005.

[11] A. Kaique, W. Mendonça, G. Mota, and T. Naia. In preparation. 2025.

[12] L. Lovász. On covering of graphs. pages 231–236, 1968.

[13] L. Pyber. Covering the edges of a connected graph by paths. *J. Combin. Theory Ser. B*, 66(1):152–159, 1996.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Decomposing random Latin squares into transversals

(Extended abstract)

Candida Bowtell* Richard Montgomery†

Abstract

In 1782, Euler conjectured that no Latin square of order $n \equiv 2 \bmod 4$ has a decomposition into transversals. While confirmed for $n = 6$ by Tarry in 1900, Bose, Parker, and Shrikhande constructed counterexamples in 1960 for each $n \equiv 2 \bmod 4$ with $n \geq 10$. We show that, in fact, counterexamples are extremely common, by showing that if a Latin square of order n is chosen uniformly at random then with high probability it has a decomposition into transversals.

This extended abstract complements [4].

1 Introduction

A *Latin square of order* n is an n by n grid filled with n symbols so that each row and column contains each symbol exactly once. A transversal in a Latin square of order n is a collection of n cells which share no row, column, or symbol. Latin squares have a long history preceding their modern study; for more on this, we recommend the historical survey by Andersen [1], while the broader study of transversals in Latin squares is covered in surveys by Wanless [23] and Montgomery [15].

In 1782, Euler [8] considered: for which n is there a Latin square of order n which can be decomposed into n disjoint transversals? The case $n = 4$ was the topic of an old recreational mathematics problem [17], while Euler was initially particularly interested in the case $n = 6$, considering his famous '36 officers problem'. In this problem, there are 36 officers of 6 different ranks from 6 different regiments, with an officer of each rank in each regiment. Can they stand in a 6 by 6 grid so that each row and each column contains officers of different ranks and different regiments? If there were a solution, then, neglecting the ranks, giving each officer the

*School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK. E-mail: c.bowtell@bham.ac.uk. Supported by Leverhulme Trust Early Career Fellowship ECF–2023–393.

†Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. E-mail: richard.montgomery@warwick.ac.uk. Supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978)

symbol of their regiment will form a Latin square of order 6. For each rank, the set of officers of that rank marks out a transversal, so that this arrangement would give a decomposition of the Latin square of order 6 into 6 disjoint transversals.

Neglecting the regiments and affixing each officer with the symbol of their rank also gives a Latin square (see Figure 1), which is *orthogonal* to the Latin square given by the regiments. That is, all possible n^2 pairs of symbols appear in matching row/column pairs of the two Latin squares. Finding two orthogonal Latin squares of order n is equivalent to the formulation of finding a Latin square of order n which decomposes into transversals. For our work, we use the latter form.

Euler believed there was no solution to his 36 officer's problem, though this was not confirmed until work by Tarry [21] in 1900. More generally, after demonstrating that there are Latin squares of order n which can be decomposed into n disjoint transversals when $n \not\equiv 2 \bmod 4$, Euler conjectured that there are no examples when $n \equiv 2 \bmod 4$. This is true for $n = 2$ and $n = 6$, but, in 1959, Bose and Shrikhande [3] showed that Euler's conjecture is false by constructing counterexamples for $n = 22$ and $n = 50$, before, shortly after, showing with Parker [2] that the conjecture is false for every $n \equiv 2 \bmod 4$ with $n \geq 10$.

The development of the probabilistic method has shown the power of considering random objects as potential counterexamples. It is interesting then, to ask how common counterexamples to Euler's conjecture are, and, in particular, whether a random Latin square of order $n \equiv 2 \bmod 4$ is typically a counterexample? For each $n \in \mathbb{N}$, let $\mathcal{L}(n)$ be the set of Latin squares of order n which use symbols in $[n] = \{1, \dots, n\}$, and let L_n be drawn uniformly at random from $\mathcal{L}(n)$. In 1990, van Rees [22] conjectured that a random Latin square L_n should not have a decomposition into transversals with high probability (whp), however, Wanless and Webb [24] observed in 2006 that numerical calculations suggest that L_n should have such a decomposition with high probability.

It has long been known that, when n is even, a Latin square of order n may not have even a single transversal (as, for example, seen by the canonical example of the addition table for $\mathbb{Z}_{2m}$, for any $m \in \mathbb{Z}$). Recently, Montgomery [14] showed that, for sufficiently large n, every Latin square of order n has a partial transversal with $n - 1$ cells. This is the best known bound towards the well-known Ryser-Brualdi-Stein conjecture [5, 19, 20], with origins from 1967, which suggests that every Latin square of order n should have a transversal when n is odd, and a partial transversal with $n-1$ cells when n is even. Whilst this comes very close to a single transversal, in our work we wish to determine whether, with high probability, we can find n disjoint transversals in a random Latin square. In every Latin square of order n this is not possible. Indeed, clearly for every even order n we have examples of Latin squares with not even a single transversal, and Wanless and Webb [24] confirmed the existence of Latin squares which do not have a decomposition into transversals for every order $n > 3$. However, some approximate version of this is true. In particular, Montgomery, Pokrovskiy and Sudakov [16] showed that every Latin square of order n contains $(1 - o(1))n$ disjoint partial transversals with $(1 - o(1))n$ cells. In other words, every large Latin square has some approximation of the properties we want to find in a random Latin square whp. However, finding these exact properties in a random Latin square whp is surprisingly difficult. For example, it is very challenging to show even that a typical random Latin square contains at least one transversal, and this was proved only in 2020, by Kwan [12]. A significant part of the challenge is finding a way to study a random Latin square. Roughly, this can reasonably be pinned to the rigidity of Latin squares; that is, that it is hard to make small modifications to a Latin square to reach another Latin square.

$0^{\boxed{0}}$	$1^{\boxed{1}}$	$2^{\boxed{2}}$	$3^{\boxed{3}}$	$4^{\boxed{4}}$	$5^{\boxed{5}}$	$6^{\boxed{6}}$	$7^{\boxed{7}}$	$8^{\boxed{8}}$
$1^{\boxed{8}}$	$2^{\boxed{0}}$	$3^{\boxed{1}}$	$4^{\boxed{2}}$	$5^{\boxed{3}}$	$6^{\boxed{4}}$	$7^{\boxed{5}}$	$8^{\boxed{6}}$	$0^{\boxed{7}}$
$2^{\boxed{7}}$	$3^{\boxed{8}}$	$4^{\boxed{0}}$	$5^{\boxed{1}}$	$6^{\boxed{2}}$	$7^{\boxed{3}}$	$8^{\boxed{4}}$	$0^{\boxed{5}}$	$1^{\boxed{6}}$
$3^{\boxed{6}}$	$4^{\boxed{7}}$	$5^{\boxed{8}}$	$6^{\boxed{0}}$	$7^{\boxed{1}}$	$8^{\boxed{2}}$	$0^{\boxed{3}}$	$1^{\boxed{4}}$	$2^{\boxed{5}}$
$4^{\boxed{5}}$	$5^{\boxed{6}}$	$6^{\boxed{7}}$	$7^{\boxed{8}}$	$8^{\boxed{0}}$	$0^{\boxed{1}}$	$1^{\boxed{2}}$	$2^{\boxed{3}}$	$3^{\boxed{4}}$
$5^{\boxed{4}}$	$6^{\boxed{5}}$	$7^{\boxed{6}}$	$8^{\boxed{7}}$	$0^{\boxed{8}}$	$1^{\boxed{0}}$	$2^{\boxed{1}}$	$3^{\boxed{2}}$	$4^{\boxed{3}}$
$6^{\boxed{3}}$	$7^{\boxed{4}}$	$8^{\boxed{5}}$	$0^{\boxed{6}}$	$1^{\boxed{7}}$	$2^{\boxed{8}}$	$3^{\boxed{0}}$	$4^{\boxed{1}}$	$5^{\boxed{2}}$
$7^{\boxed{2}}$	$8^{\boxed{3}}$	$0^{\boxed{4}}$	$1^{\boxed{5}}$	$2^{\boxed{6}}$	$3^{\boxed{7}}$	$4^{\boxed{8}}$	$5^{\boxed{0}}$	$6^{\boxed{1}}$
$8^{\boxed{1}}$	$0^{\boxed{2}}$	$1^{\boxed{3}}$	$2^{\boxed{4}}$	$3^{\boxed{5}}$	$4^{\boxed{6}}$	$5^{\boxed{7}}$	$6^{\boxed{8}}$	$7^{\boxed{0}}$

$0^{\boxed{0}}$	$6^{\boxed{7}}$	$5^{\boxed{8}}$	$4^{\boxed{9}}$	$9^{\boxed{1}}$	$8^{\boxed{3}}$	$7^{\boxed{5}}$	$1^{\boxed{2}}$	$2^{\boxed{4}}$	$3^{\boxed{6}}$
$7^{\boxed{6}}$	$1^{\boxed{1}}$	$0^{\boxed{7}}$	$6^{\boxed{8}}$	$5^{\boxed{9}}$	$9^{\boxed{2}}$	$8^{\boxed{4}}$	$2^{\boxed{3}}$	$3^{\boxed{5}}$	$4^{\boxed{0}}$
$8^{\boxed{5}}$	$7^{\boxed{0}}$	$2^{\boxed{2}}$	$1^{\boxed{7}}$	$0^{\boxed{8}}$	$6^{\boxed{9}}$	$9^{\boxed{3}}$	$3^{\boxed{4}}$	$4^{\boxed{6}}$	$5^{\boxed{1}}$
$9^{\boxed{4}}$	$8^{\boxed{6}}$	$7^{\boxed{1}}$	$3^{\boxed{3}}$	$2^{\boxed{7}}$	$1^{\boxed{8}}$	$0^{\boxed{9}}$	$4^{\boxed{5}}$	$5^{\boxed{0}}$	$6^{\boxed{2}}$
$1^{\boxed{9}}$	$9^{\boxed{5}}$	$8^{\boxed{0}}$	$7^{\boxed{2}}$	$4^{\boxed{4}}$	$3^{\boxed{7}}$	$2^{\boxed{8}}$	$5^{\boxed{6}}$	$6^{\boxed{1}}$	$0^{\boxed{3}}$
$3^{\boxed{8}}$	$2^{\boxed{9}}$	$9^{\boxed{6}}$	$8^{\boxed{1}}$	$7^{\boxed{3}}$	$5^{\boxed{5}}$	$4^{\boxed{7}}$	$6^{\boxed{0}}$	$0^{\boxed{2}}$	$1^{\boxed{4}}$
$5^{\boxed{7}}$	$4^{\boxed{8}}$	$3^{\boxed{9}}$	$9^{\boxed{0}}$	$8^{\boxed{2}}$	$7^{\boxed{4}}$	$6^{\boxed{6}}$	$0^{\boxed{1}}$	$1^{\boxed{3}}$	$2^{\boxed{5}}$
$2^{\boxed{1}}$	$3^{\boxed{2}}$	$4^{\boxed{3}}$	$5^{\boxed{4}}$	$6^{\boxed{5}}$	$0^{\boxed{6}}$	$1^{\boxed{0}}$	$7^{\boxed{7}}$	$8^{\boxed{8}}$	$9^{\boxed{9}}$
$4^{\boxed{2}}$	$5^{\boxed{3}}$	$6^{\boxed{4}}$	$0^{\boxed{5}}$	$1^{\boxed{6}}$	$2^{\boxed{0}}$	$3^{\boxed{1}}$	$8^{\boxed{9}}$	$9^{\boxed{7}}$	$7^{\boxed{8}}$
$6^{\boxed{3}}$	$0^{\boxed{4}}$	$1^{\boxed{5}}$	$2^{\boxed{6}}$	$3^{\boxed{0}}$	$4^{\boxed{1}}$	$5^{\boxed{2}}$	$9^{\boxed{8}}$	$7^{\boxed{9}}$	$8^{\boxed{7}}$

Figure 1: Two Latin squares decomposed into transversals. On the left, the addition group of integers mod 9 is given, which is then decomposed into transversals indicated by integers in the top right, starting with the transversal along the leading diagonal (marked by 0) which is then moved to the right by 1 mod 9 8 times to create 8 new transversals. On the right, Bose, Parker, and Shrikhande's example of a Latin square of order 10 with a transversal decomposition [3].

In [12], Kwan's main focus was the closely related problem of finding a perfect matching in a uniformly random Steiner triple system of order $n \equiv 3 \bmod 6$, using methods that could be adapted for transversals in Latin squares. Ferber and Kwan [9] subsequently showed that a random Steiner triple system of order $n \equiv 3 \bmod 6$ contains disjoint perfect matchings covering all but $o(n^2)$ of its edges. Though they did not do so, similar adaptations to their methods appear capable of showing that a random Latin square of order n has, with high probability, $(1-o(1))n$ disjoint transversals. In this work, we show that, in fact, with high probability a random Latin square contains n disjoint transversals. In particular, then, the proportion of Latin squares of order $n \equiv 2 \mod 4$ which provide a counterexample to Euler's conjecture tends to 1 as n tends to infinity.

Theorem 1.1. *A random Latin square of order n has a decomposition into transversals with probability* $1 - o(1)$.

Since the result of Kwan [12], two alternative methods have been developed to show that a random Latin square of order n has a transversal with high probability, each moreover strengthening this result in different ways. Firstly, Eberhard, Manners, and Mrazović [6] gave a remarkably tight estimate on the number of transversals in a typical random Latin square of order n, using tools from analytic number theory. Then, Gould and Kelly [10] developed techniques from their previous work with Kühn and Osthus [11] to show that a random Latin square is likely to contain a particular type of transversal known as a 'Hamilton transversal', using more combinatorial methods than [6], but which are distinctly different to those in the original approach of Kwan [12]. To prove Theorem 1.1, we also take a combinatorial approach.

2 Proof strategy

2.1 Reformulation in terms of rainbow matchings in $K_{n,n}$

We approach Theorem 1.1 by studying an equivalent formulation in properly coloured graphs. Let $K_{n,n}$ be the complete bipartite graph with vertex classes A and B, where $|A| = |B| = n$. A proper colouring of $K_{n,n}$ is a colouring of the edges so that no two edges which share a vertex have the same colour. An optimal colouring is a proper colouring which uses the minimum number of colours among all proper colourings, which, for $K_{n,n}$, is n. We always assume $K_{n,n}$ is properly coloured using colours from $C := [n] = \{1, \dots, n\}$.

Given a Latin square L of order n whose rows are indexed by A and columns by B, which furthermore uses the symbol set $[n]$, we can define an equivalent optimal colouring of $K_{n,n}$ as follows. For each $a \in A$ and $b \in B$, let the colour of ab, denoted by $c(ab)$, be the symbol in the cell of L whose row corresponds to a and whose column corresponds to b. That a Latin square has n symbols with no symbol appearing twice in any row or any column immediately implies that this colouring uses n colours and is proper, and thus we have an optimal colouring of $K_{n,n}$. Similarly, a Latin square of order n can be constructed from any optimal colouring of $K_{n,n}$, and thus the optimal colourings of $K_{n,n}$ correspond exactly to the Latin squares of order n. Furthermore, it is easy to see that a transversal in a Latin square corresponds exactly under this equivalence to a perfect matching in the corresponding optimally coloured $K_{n,n}$ which has a different colour on each of its edges. We refer to such a matching as a *rainbow perfect matching.*

To prove Theorem 1.1, we therefore, in fact, prove the following equivalent statement.

Theorem 2.1. *Let G be an optimally coloured copy of $K_{n,n}$ chosen uniformly at random from all such colourings. Then, with probability $1-o(1)$, G has a decomposition into rainbow perfect matchings.*

2.2 Proof sketch

We write $\mathcal{G}_n^{\mathrm{col}}$ for the collection of optimally properly coloured copies of $K_{n,n}$ coloured with colour set $C = [n]$ and write $G \sim G_n^{\mathrm{col}}$ when G is selected uniformly at random from $\mathcal{G}_n^{\mathrm{col}}$. Our aim, then, is to show that $G \sim G_n^{\mathrm{col}}$ with high probability has a decomposition into n-edge (perfect) rainbow matchings, $M_1, \dots, M_n$. We refer to these rainbow matchings as our *target* matchings. Using the semi-random method (as, for example, implemented in Latin squares by Montgomery, Pokrovskiy and Sudakov [16]) it can be shown that any $G \in \mathcal{G}_n^{\mathrm{col}}$ contains n disjoint rainbow matchings of size $(1 - o(1))n$. With care, this could be used along with a random partitioning of the remaining edges to show that, with high probability, $G \sim G_n^{\mathrm{col}}$ can be decomposed into n rainbow subgraphs $M_1, \dots, M_n$, one for each of our n target matchings, which each have n edges and are close to perfect matchings, in that they have maximum degree at most 2 and $(1 - o(1)) \cdot 2n$ of the vertices have degree 1. Our aim is to take such a relaxed decomposition, and correct it into n perfect rainbow matchings. To do so, we use methods falling under the overall general technique of 'absorption', as codified by Rödl, Ruciński, and Szemerédi [18] in 2006. The fundamental idea here is that we should prepare for the corrections we will need to make at the end by initially choosing parts of our random subgraphs to allow later corrections to be made. In particular, this preparation and care at the start ensures that we are able to make a large number of different possible corrections, which subsequently leads to more flexibility in completing to a suitable relaxed decomposition, which we then know can be corrected into perfect rainbow matchings due to the care taken at the start.

The absorption strategy that we employ involves a fairly intricate set-up. We first construct a template independent of the Latin square, and then adapt the template to a randomly chosen Latin square using the semi-random method applied in auxiliary hypergraphs, finding the properties we require to hold whp in a random Latin square using the deletion method and (implicitly) the switching method. Often, we require strong recent developments of these techniques, along with further novelties. In what remains of this extended abstract, we highlight some particularly key points about our proof.

The first is that we develop an 'absorption schematic' which gives a sparse set of possible local corrections that together can make any (reasonable) globally-balanced set of corrections. This is a template for building an absorber which is independent of our work in random Latin squares, and thus may be useful elsewhere.

Secondly, the switching method can and has been used directly to find small substructures in random Latin squares (see, for example, [10]), but instead we use the deletion method. This was used by Kwan, Sah, and Sawhney [13] to bound above the likely number of certain substructures, and, as well as developing this, we use the deletion method to bound below the likely number of some particular substructures.

Finally, we note here that a major source of the complexity in finding our required absorption structure in a random Latin square via the semi-random method is that it is found in three applications of the semi-random method to an auxiliary hypergraph, the last of which depends on a previous application. That is, we find part of the absorption structure and require it to satisfy some carefully chosen properties so that we can then apply the semi-random method again to find certain paths connecting up this initial structure. This requires us to use a forbidding list of properties, however all of these properties confirm some simple heuristic. To confirm these, we apply a recent implementation of the semi-random method of Ehard, Glock and Joos [7], using weight functions to record desirable properties.

References

[1] Lars Døvling Andersen. The history of Latin squares. *In Combinatorics: Ancient & Modern (edited by R. Wilson and J.J. Watkins)*, 2013.

[2] Raj Chandra Bose, Ernest T Parker, and Sharadchandra S Shrikhande. Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture. *Canadian Journal of Mathematics*, 12:189–203, 1960.

[3] Raj Chandra Bose and Sharadchandra Shankar Shrikhande. On the falsity of Euler's conjecture about the non-existence of two orthogonal Latin squares of order $4t + 2$. *Proceedings of the National Academy of Sciences*, 45(5):734–737, 1959.

[4] Candida Bowtell and Richard Montgomery. Almost every latin square has a decomposition into transversals. *arXiv preprint arXiv:2501.05438*, 2025.

[5] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory.* Cambridge University Press, 1991.

[6] Sean Eberhard, Freddie Manners, and Rudi Mrazović. Transversals in quasirandom Latin squares. *Proceedings of the London Mathematical Society*, 2023.

[7] Stefan Ehard, Stefan Glock, and Felix Joos. Pseudorandom hypergraph matchings. *Combinatorics, Probability and Computing*, 29(6):868–885, 2020.

[8] Leonhard Euler. Recherches sur un nouvelle espéce de quarrés magiques. *Verhandelingen uitgegeven door het zeeuwsch Genootschap der Wetenschappen te Vlissingen*, pages 85–239, 1782.

[9] Asaf Ferber and Matthew Kwan. Almost all Steiner triple systems are almost resolvable. *Forum of Mathematics, Sigma*, 8:39, 2020.

[10] Stephen Gould and Tom Kelly. Hamilton transversals in random Latin squares. *Random Structures & Algorithms*, 62(2):450–478, 2023.

[11] Stephen Gould, Tom Kelly, Daniela Kühn, and Deryk Osthus. Almost all optimally coloured complete graphs contain a rainbow Hamilton path. *Journal of Combinatorial Theory, Series B*, 156:57–100, 2022.

[12] Matthew Kwan. Almost all Steiner triple systems have perfect matchings. *Proceedings of the London Mathematical Society*, 121(6):1468–1495, 2020.

[13] Matthew Kwan, Ashwin Sah, and Mehtaab Sawhney. Large deviations in random Latin squares. *Bulletin of the London Mathematical Society*, 54(4):1420–1438, 2022.

[14] Richard Montgomery. A proof of the Ryser-Brualdi-Stein conjecture for large even n. *arXiv preprint arXiv:2310.19779*, 2023.

[15] Richard Montgomery. Transversals in Latin squares. *Surveys in Combinatorics*, 2024.

[16] Richard Montgomery, Alexey Pokrovskiy, and Benjamin Sudakov. Decompositions into spanning rainbow structures. *Proceedings of the London Mathematical Society*, 119(4):899–959, 2019.

[17] Jacques Ozanam. *Récréations mathématiques et physiques*, volume 1. 1723.

[18] Vojtěch Rödl, Andrzej Ruciński, and Endre Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[19] H.J. Ryser. Neuere Probleme der Kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, pages 69–91, 1967.

[20] S. K. Stein. Transversals of Latin squares and their generalizations. *Pacific J. Math.*, 59:567–575, 1975.

[21] Gaston Tarry. *Le probléme des 36 officiers.* Secrétariat de l'Association française pour l'avancement des sciences, 1900.

[22] G.H.J. van Rees. Subsquares and transversals in Latin squares. *Ars Combinatoria*, 29:193–204, 1990.

[23] Ian M Wanless. Transversals in Latin squares: a survey. *Surveys in combinatorics*, 392:403–437, 2011.

[24] Ian M Wanless and Bridget S Webb. The existence of Latin squares without orthogonal mates. *Designs, Codes and Cryptography*, 40:131–135, 2006.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Almost-perfect colorful matchings in bipartite graphs

(Extended abstract)

Simona Boyadzhiyska* Micha Christoph† Tibor Szabó‡

Abstract

We prove that, for positive integers n, a_1, a_2, a_3 satisfying $a_1+a_2+a_3 = n-1$, it holds that any bipartite graph G which is the union of three perfect matchings M_1, M_2, and M_3 on $2n$ vertices contains a matching M such that $|M \cap M_i| = a_i$ for $i = 1, 2$, and 3. The bound $n-1$ on the sum is best possible in general and verifies the multiplicity extension of the Ryser-Brualdi-Stein Conjecture, proposed recently by Anastos, Fabian, Müyesser, and Szabó, for three colors.

1 Introduction

An edge-colored graph is *rainbow* if all of its edges have distinct colors. In 1967 Ryser [14] (see also [6]) conjectured that, for every odd n, any proper n-coloring of the complete bipartite graph $K_{n,n}$ contains a rainbow matching of size n. This conjecture proved to be extremely influential in Combinatorics, driving the development of more and more sophisticated approaches and tools. It is still open today, though significant progress has been made through the years.

It is not hard to prove that the statement of the conjecture does not hold for even n, though it has been subsequently conjectured that it only barely does not. More precisely, the Ryser-Brualdi-Stein Conjecture [8, 14, 16] states that any proper edge-coloring of $K_{n,n}$ contains a rainbow matching of size $n-1$. The proof of this conjecture was announced recently by Montgomery [13], following earlier work by several groups of authors (see [7, 9, 10, 11, 12, 15, 17] and the references therein). This takes us deceptively close to the full resolution of the

*HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: simona@renyi.hu. Most of this research was conducted while the author was at the School of Mathematics, University of Birmingham, Birmingham, United Kingdom. The research leading to these results was supported by EPSRC, grant no. EP/V048287/1 and by ERC Advanced Grants "GeoScape", no. 882971 and "ERMiD", no. 101054936. There are no additional data beyond that contained within the main manuscript.

†Department of Mathematics, ETH Zürich, Zürich, Switzerland. E-mail: micha.christoph@math.ethz.ch. The research leading to these results was supported by the SNSF Ambizione Grant No. 216071.

‡Institut für Mathematik, Freie Universität Berlin, Berlin, Germany. E-mail: szabo@mi.fu-berlin.de.

conjecture of Ryser for odd n, though, according to [13], significant novel ideas will likely be necessary to find that last edge completing a rainbow perfect matching.

Building upon earlier work by Arman, Rödl, and Sales [5], Anastos, Fabian, Müyesser, and Szabó [4] investigated what other matching structures, besides rainbow matchings, would be inevitable in a proper edge-coloring of $K_{n,n}$ and conjectured the following extension of Montgomery's Theorem.[1]

Conjecture 1.1 (Multiplicity Ryser-Brualdi-Stein Conjecture [4])**.** *Let G be a complete bipartite graph on $2n$ vertices whose edge set is decomposed into perfect matchings M_i, $i = 1, \dots, n$. Let $a_1, \dots, a_n \in \mathbb{N}_0$ be nonnegative integers such that $\sum_i a_i = n - 1$. Then there exists a matching M in G such that $|M \cap M_i| = a_i$ for every $i \in \{1, \dots, n\}$.*

By setting $a_i = 1$ for all $i \in \{1, \dots, n-1\}$ and $a_n = 0$ in Theorem 1.1, we obtain a strengthening of the Ryser-Brualdi-Stein Conjecture. If there is exactly one nonzero color-multiplicity $a_i = n - 1$, then one can just take a subset of the corresponding perfect matching M_i. It is also easy to confirm Theorem 1.1 when there are two nonzero color-multiplicities, that is, $a_i + a_j = n - 1$. Indeed, the union of the perfect matchings M_i and M_j forms a 2-factor consisting of even cycles, which alternate between the two matchings. To create M, we simply pick edges from M_i component by component, until in some component C we reach the target number a_i of edges designated for M_i. In C, we pick the necessary amount of M_i-edges in a consecutive fashion. Then, after potentially leaving unsaturated the two vertices incident to the first and last selected M_i-edges in C, we can saturate the remaining $2n - 2a_i - 2 = 2a_j$ vertices with M_j-edges.

For three nonzero color-multiplicities a_i, a_j, a_k, Anastos, Fabian, Müyesser, and Szabó [4] proved that a matching with the desired color-multiplicities exists provided that $a_i + a_j + a_k \leq n - 2$. In fact, they showed the result in the more general setting where the three matchings need not form a bipartite graph (and also need not be disjoint); they also proved that in the general case the bound $n - 2$ on the sum $a_i + a_j + a_k$ is best possible (see [4, Remark 2]).

We extend this to a full proof of the Multiplicity Ryser-Brualdi-Stein Conjecture for three nonzero color-multiplicities.

1.1 Main result

Let G be a graph which is the union of three disjoint perfect matchings M_1, M_2, and M_3. For integers $a_1, a_2, a_3 \in \mathbb{N}_0$, a matching M of G is an (a_1, a_2, a_3)*-matching* if $|M \cap M_i| = a_i$ for every $i \in \{1, 2, 3\}$.

Theorem 1.2. *Let G be a bipartite graph on $2n$ vertices which is the union of three disjoint perfect matchings M_1, M_2, and M_3. Then, for any integers $a_1, a_2, a_3 \in \mathbb{N}_0$ satisfying $a_1 + a_2 + a_3 = n - 1$, there exists an (a_1, a_2, a_3)-matching in G.*

As shown in [4, Proposition 1], Theorem 1.2 is best possible: for any n and any integers $0 \leq a_1, a_2, a_3 \leq n - 1$ summing up to n, there exists a bipartite graph on $2n$ vertices which is the disjoint union of three perfect matchings and has no (a_1, a_2, a_3)-matching. As discussed earlier, the analog of Theorem 1.2 is false if we remove the assumption that G is bipartite.

[1] This was also posed independently as a question by Noga Alon [1].

2 Proof ideas

The proof of Theorem 1.2 is elementary, using only augmenting paths, but turns out to be surprisingly delicate.

Let G be a bipartite graph on $2n$ vertices which is the union of three disjoint perfect matchings M_1, M_2, and M_3 and a_1, a_2, a_3 be integers adding up to $n-1$. The proof consists of two main ingredients, which we refer to as the Reduction Lemma and the Switching Lemma. The former allows us to reduce the problem to the case where G is connected. While in many theorems of graph theory it is simple to extend the statement to all graphs once it is known for all *connected* graphs, in our setting this step requires a separate and nontrivial argument. The reason is that, if G is disconnected, we need to find a *perfect* matching in all but one of its components.

Lemma 2.1 (Reduction Lemma)**.** *Let F be a bipartite graph on $2m$ vertices which is the union of three disjoint perfect matchings M_1, M_2, and M_3. Let $b_1, b_2, b_3 \in \mathbb{N}_0$ be integers such that $b_1 + b_2 + b_3 \geq 2m - 1$. Then there exists a perfect matching in F with at most b_i edges of M_i for each $i \in \{1, 2, 3\}$.*

The idea is to apply the Reduction Lemma repeatedly to every component of G except the largest one and then deal with the remaining component using the Switching Lemma.

The Switching Lemma applies only to connected graphs and allows us to "switch" two colors in a matching of size $n-1$. More precisely, it shows that, for any matching of size $n-1$ containing at least one edge of some color j, we can find another matching of size $n-1$ with one more edge of some other color i and one fewer edge of color j.

Lemma 2.2 (Switching Lemma)**.** *Let G be a connected bipartite graph on $2n$ vertices which is the union of three disjoint perfect matchings M_1, M_2, and M_3. Let $a_1, a_2, a_3 \in \mathbb{N}_0$ be integers satisfying $a_3 \geq 1$ and $a_1+a_2+a_3 = n-1$ and suppose that G contains an (a_1, a_2, a_3)-matching. Then G also contains an $(a_1 + 1, a_2, a_3 - 1)$-matching and an $(a_1, a_2 + 1, a_3 - 1)$-matching.*

Starting from an arbitrary $(0, 0, n-1)$-matching in G, we can use the Switching Lemma repeatedly to obtain any (a_1, a_2, a_3)-matching.

2.1 Reduction Lemma

We outline the proof of Theorem 2.1. Assume without loss of generality that $b_1 + b_2 \geq m$. The idea is to build a perfect matching M by considering the graph formed by $M_1 \cup M_2$, which forms a cycle factor in F. We work one cycle at a time, always choosing the shortest available cycle C, and add the edges of either $C \cap M_1$ or $C \cap M_2$ to M in such a way that we never exceed the permitted number of edges from M_1 and M_2. We may only get stuck at the longest cycle. In that case, we show that we can build an almost-perfect matching M', covering all but two vertices of the longest cycle C and using only edges of M_1 and M_2 without exceeding the permitted number thereof. Then, by considering an appropriate M'-augmenting path using edges of M_3, we can obtain a perfect matching intersecting each M_i in a permissible number of edges.

2.2 Switching Lemma

The most technical part of our work is proving Theorem 2.2; we provide a rough sketch of our strategy. Let G, M_1, M_2, and M_3 be as given, a_1, a_2, a_3 be integers satisfying $a_3 \geq 1$ and

$a_1 + a_2 + a_3 = n - 1$, and M be an (a_1, a_2, a_3)-matching in G. Suppose for a contradiction that G contains no $(a_1 + 1, a_2, a_3 - 1)$-matching. The idea is to modify M by switching edges along an appropriate path and find an $(a_1 + 1, a_2, a_3 - 1)$-matching, thus reaching the desired contradiction. As M has size $n - 1$, it leaves precisely two vertices of G unsaturated. We will work with what we call *nearly-M-alternating paths*. Essentially, these are paths that contain the vertices unsaturated by M but are otherwise M-alternating, that is, if we cut a nearly-M-alternating path at the two unsaturated vertices, we will end up with three M-alternating paths. In addition, we will choose these paths in such a way that they interact with the matchings M_1, M_2, and M_3 in a controlled way. Namely, we will use nearly-M-alternating paths that additionally alternate between edges in $M_1 \cup M_3$ and $M_2 \cup M_3$, that is, M_1-edges will only appear in odd positions and M_2-edges only in even ones, or vice versa. We call a path alternating between $M_1 \cup M_3$ and $M_2 \cup M_3$ *good*. This extra control will allow us to "deform" one matching into another along a good nearly-M-alternating path, keeping the number of M_2-edges fixed. This yields the following statement, which is similar in spirit to the Intermediate Value Theorem.

Lemma 2.3. *Let M' be an (a'_1, a'_2, a'_3)-matching of size $n - 1$ in G and P be a good nearly-M'-alternating path. Let M'' be an (a''_1, a'_2, a''_3)-matching of size $n - 1$ in G with $a''_3 \geq a'_3$ such that P is also nearly-M''-alternating. Then, for all $a'_3 \leq a^*_3 \leq a''_3$, there exists an (a^*_1, a'_2, a^*_3)-matching M^* of size $n - 1$.*

The goal is then to find a suitable path P and matchings M' and M'' with $|M' \cap M_2| = |M'' \cap M_2| = a_2$ and $|M' \cap M_3| \leq a_3 - 1 \leq |M'' \cap M_3|$ to which we can apply Theorem 2.3. Roughly speaking, we wish to find such a path connecting the vertices unsaturated by M to an edge in $M \cap M_3$ (which we will then try to "switch out" of the matching). This is possible by connectivity: if u is unsaturated by M, it is possible to reach any other vertex of G on an M-alternating path from u.

3 Multigraphs

It turns out that the proof of our main theorem could be modified to include the case of multigraphs, i.e., when the three matchings M_1, M_2, and M_3 are not necessarily edge-disjoint. Hence we obtain the following stronger result.

Theorem 3.1. *Let G be a bipartite multigraph on $2n$ vertices which is the union of three perfect matchings M_1, M_2, and M_3. Then, for any integers $a_1, a_2, a_3 \in \mathbb{N}_0$ satisfying $a_1 + a_2 + a_3 = n - 1$, the graph G contains an (a_1, a_2, a_3)-matching.*

4 Concluding remarks

In this work, we resolved Theorem 1.1 in the case where there are exactly three nonzero color-multiplicities a_i, a_j, a_k. Naturally, it would be very interesting to extend this to any $k \geq 4$ nonzero color-multiplicities. As the case $k = n - 1$ would imply Montgomery's Theorem, resolving the problem in full generality is expected to be very difficult. Tackling the $k = 4$ case fully already seems to require novel ideas.

Problem 4.1. Let G be a bipartite graph on $2n$ vertices whose edge set is decomposed into perfect matchings M_1, M_2, M_3, and M_4. Let $a_1, a_2, a_3, a_4 \in \mathbb{N}_0$ be nonnegative integers such that $a_1 + a_2 + a_3 + a_4 = n - 1$. Then there exists a matching M in G such that $|M \cap M_i| = a_i$ for $i \in \{1, 2, 3, 4\}$.

A simpler problem would be to show that the constant 1 in Theorem 1.1 can be replaced by some function depending only on k.

Problem 4.2. Determine the smallest function $f : \mathbb{N} \to \mathbb{R}_{\geq 1}$ such that the following is true. Let n be a (sufficiently large) integer, G be a bipartite graph on $2n$ vertices that is the union of k perfect matchings $M_1, \ldots, M_k$, and $a_1, \ldots, a_k \in \mathbb{N}_0$ be integers summing up to $n - f(k)$. Then G contains a matching M such that $|M \cap M_i| = a_i$ for all $i \in \{1, \ldots, k\}$.

In a different direction, it would be interesting to investigate whether Theorem 1.2 is true under weaker conditions. As discussed in [4, Remark 2], if n is even, G is the disjoint union of $n/2$ copies of K_4, and a_1, a_2, a_3 are all odd and $a_1 + a_2 + a_3 = n - 1$, then G has no (a_1, a_2, a_3)-matching. However, this is the only obstruction we are aware of. Taking G to be bipartite automatically eliminates this example, but can we relax this assumption further? We believe that this should be the case and reiterate a conjecture of Anastos, Fabian, Müyesser, and Szabó [4, Conjecture 3].

Conjecture 4.3 ([4])**.** *Let G be a graph on $2n$ vertices which is the union of three disjoint perfect matchings, and suppose that G has a component that is not isomorphic to K_4. Then, for any integers $a_1, a_2, a_3 \in \mathbb{N}_0$ satisfying $a_1 + a_2 + a_3 = n - 1$, G contains an (a_1, a_2, a_3)-matching.*

Acknowledgments

We would like to thank Clément Requilé for many helpful discussions.

References

[1] N. Alon. private communication.

[2] N. Alon. Splitting necklaces. *Adv. Math.*, 63(3):247–253, 1987.

[3] N. Alon and D. B. West. The Borsuk-Ulam theorem and bisection of necklaces. *Proc. Am. Math. Soc.*, 98(4):623–628, 1986.

[4] M. Anastos, D. Fabian, A. Müyesser, and T. Szabó. Splitting matchings and the Ryser-Brualdi-Stein conjecture for multisets. *Electron. J. Combin.*, 30(3):Paper No. 3.10, 6, 2023.

[5] A. Arman, V. Rödl, and M. T. Sales. Colorful matchings. *SIAM J. Discrete Math.*, 37(2):925–950, 2023.

[6] D. Best and I. M. Wanless. What did Ryser conjecture? *arXiv:1801.02893*, 2018.

[7] A. E. Brouwer, A. J. de Vries, and R. M. A. Wieringa. A lower bound for the length of partial transversals in a Latin square. *Nieuw Arch. Wisk. (3)*, 26(2):330–332, 1978.

[8] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory.* Cambridge University Press, 1991.

[9] D. A. Drake. Maximal sets of Latin squares and partial transversals. *J. Statist. Plann. Inference*, 1(2):143–149, 1977.

[10] P. Hatami and P. W. Shor. A lower bound for the length of a partial transversal in a Latin square. *J. Combin. Theory Ser. A*, 115(7):1103–1113, 2008.

[11] P. Keevash, A. Pokrovskiy, B. Sudakov, and L. Yepremyan. New bounds for Ryser's conjecture and related problems. *Trans. Amer. Math. Soc. Ser. B*, 9:288–321, 2022.

[12] K. K. Koksma. A lower bound for the order of a partial transversal in a Latin square. *J. Combinatorial Theory*, 7:94–95, 1969.

[13] R. Montgomery. A proof of the Ryser-Brualdi-Stein conjecture for large even n. *arXiv:2310.19779*, 2023.

[14] H. J. Ryser. Neuere Probleme der Kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, 69(91):35, 1967.

[15] P. W. Shor. A lower bound for the length of a partial transversal in a Latin square. *J. Combin. Theory Ser. A*, 33(1):1–8, 1982.

[16] S. K. Stein. Transversals of Latin squares and their generalizations. *Pacific J. Math.*, 59(2):567–575, 1975.

[17] D. E. Woolbright. An $n \times n$ Latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols. *J. Combinatorial Theory Ser. A*, 24(2):235–237, 1978.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

EPPA witnesses with two blocks of imprimitivity on edges

(Extended abstract)

David Bradley-Williams* Sofia Brenner†

Abstract

A graph H is an *EPPA witness* for its induced subgraph G if every partial isomorphism of G extends to an automorphism of H. In this note, we study graphs with two blocks of imprimitivity on edges, and we determine the subgraphs for which they are EPPA witnesses. This will be applied in the classification of graphs with EPPA witnesses of at most twice the size, which we carry out in a forthcoming paper [1]. As an application, we determine the structure of the induced subgraphs of rook's graphs and Paley graphs for which they are EPPA witnesses.

1 Introduction

A graph H is said to extend all partial automorphisms of an induced subgraph G if, quite literally, every (finite) partial isomorphism of G into G extends to an automorphism of H. For short we say that H is an *EPPA witness* for G. E. Hrushovski in [9] started the entire field of studying EPPA witnesses by establishing that every finite graph G has some finite EPPA witness H, in particular containing G as an induced subgraph. From the outset, the question was posed in [9] of how large H must be to serve as an EPPA witness of G; the optimal number of vertices defines $\mathrm{eppa}(G)$, the *eppa number of* G. This has inspired improvements over the following three decades, most of all in narrowing asymptotic bounds for $\mathrm{eppa}(G)$ in terms of the number of vertices of G, see [2, 4, 7, 8, 10]. However, even verifying particular instances of $\mathrm{eppa}(G)$ has been a neglected open problem, with some recent progress employing group-theoretic methods in [3].

*Computer Science Institute of Charles University, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 118 00 Prague 1, Czech Republic and Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 110 00 Praha 1, Czech Republic. E-mail: williams@math.cas.cz. Supported by Project 24-12591M of the Czech Science Foundation (GAČR).

†University of Kassel, 34132 Kassel, Germany. E-mail: sbrenner@mathematik.uni-kassel.de. Supported by DFG grant 522790373.

We thank Janek Hlavatý for pointing out a mistake in an earlier version of Theorem 3.

A paradigm to follow when studying EPPA witnesses is that effective EPPA witnesses H exhibit strong symmetry properties, and when discovered, such H turn out to be EPPA witnesses for many of their (induced) subgraphs. At the strongest end, if H is its own EPPA witness, then by definition H is an *ultrahomogeneous* graph; note that such an H is then simultaneously an EPPA witness for all of its subgraphs. In this sense, classifications of graphs H that are EPPA witnesses for large subgraphs generalize the classification of finite ultrahomogeneous graphs achieved by A. Gardiner [5]. In our forthcoming paper [1], we make use of this strategy to classify all graphs G which have an EPPA witness of at most twice the order, resolving a problem posed by P. Cameron, J. Hubička, M. Konečný and the first named author in [3].

Now fixing any finite graph G and assuming H is an EPPA witness of minimal order, and among those with a minimal number of edges, then H is edge transitive. So it usually suffices to study edge transitive graphs. Primitivity is a stronger notion than transitivity. Suppose that the edge set $E(H)$ admits a decomposition $E(H) = E_1 \dot\cup \ldots \dot\cup E_k$ into disjoint sets $E_1, \ldots, E_k$ that is preserved by the automorphisms of H. Then $E_1, \ldots, E_k$ are called *blocks* and $\{E_1, \ldots, E_k\}$ is a *block system*. The graph H is *edge primitive* if it is edge transitive and does not preserve any nontrivial block system, i.e. it has no block system with $1 < |E_1| = \cdots = |E_k| < |E(H)|$. Primitivity is a well-studied concept in group theory and algebraic graph theory.

We focus on graphs that are edge transitive, but not edge primitive; they have a nontrivial block system of edges, which we depict and refer to with colours. Thus isomorphisms of the graph can permute the colours, while they must preserve the partition of the edges into colour classes. In this note, we focus on the case that H admits a block system with precisely two blocks. Various natural graph families, such as lattice graphs and certain Paley graphs, have this property. Our main result is the following classification:

Theorem 1. *Let H be a finite edge-transitive graph and assume that there is a block system with precisely two blocks on $E(H)$, which we call* green edges *and* blue edges*. Let G be an induced subgraph of H such that H is an EPPA witness for G. Up to exchanging the colours, one of the following holds:*

1. *If G omits a bichromatic induced path on three vertices, then one of the following holds:*
   - *(a) G is monochromatic, i.e., only contains edges from a single colour class;*
   - *(b) Every connected component of G contains only edges in one of the colours, and all edges of one colour belong to a single component (so G has at most two non-singleton components).*
2. *If G contains a bichromatic induced path on three vertices, then G has at most one non-singleton connected component C, for which one of the following holds:*
   - *(a) C consists of two blue cliques C_1 and C_2, and the green edges in C form a matching. Every green edge has one endpoint in C_1 and the other in C_2.*
   - *(b) C consists of a single blue and a single green clique, and the corresponding subgraphs intersect in precisely one vertex.*
   - *(c) C contains a single blue clique B, and the green edges in C form a matching. Each green edge has one endpoint in B and the other outside of B.*

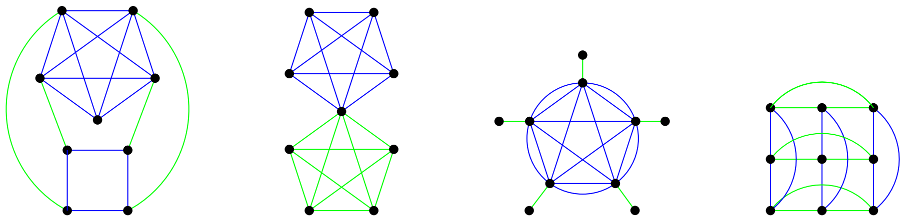


Figure 1: Graphs occuring in Theorem 1.

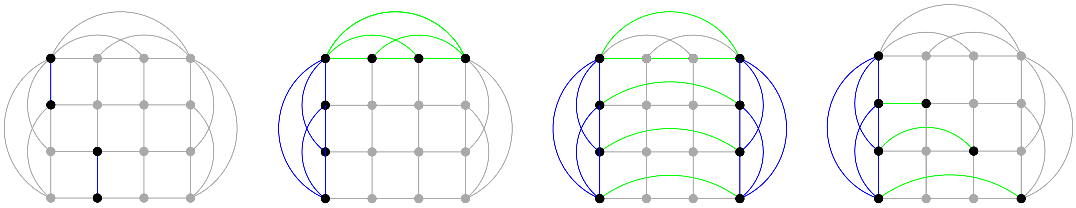


Figure 2: Graphs occurring in Theorem 2. For clarity, not all edges connecting vertices of the same row or column are drawn.

(d) C is an induced subgraph of the 3×3-rook's graph.

See Figure 1 for an illustration of the subcases of the second case. The *$n \times n$-rook's graph* (or square grid graph) is the Cartesian product of two complete graphs K_n. As an application, we study the subgraphs of rook's graphs and Paley graphs for which these graphs are EPPA witnesses. Note that for $n \leq 3$, the rook's graph is ultrahomogeneous and hence an EPPA witness for all of its subgraphs.

Theorem 2. *Let $n \geq 4$ and set $H = K_n \square K_n$ the $n \times n$-rook's graph. Then H is an EPPA witness for an induced subgraph G if and only if G is a monochromatic disjoint union of cliques $C_1, \ldots, C_k$ with $|V(C_1)| + \cdots + |V(C_k)| \leq n$ or an induced subgraph of one of the following graphs: the subgraph of H induced by two distinct columns, by one row and one column, or by the first column together with the diagonal, or the disjoint union of a 3×3-rook's graph and singletons.*

See Figure 2 for an illustration. For a prime power $q > 1$ with $q \equiv 1 \pmod 4$, the vertices of the *Paley graph $P(q)$* are the elements of the field $\mathbb{F}_q$, and $a, b \in \mathbb{F}_q$ are adjacent if $a - b$ is a nonzero square in $\mathbb{F}_q$. If $q \equiv 1 \pmod 8$, then setting $E_B = \{\{a, b\} \in E(P(q)) \colon a - b \text{ is a 4th power in } \mathbb{F}_q\}$ and $E_G = E(P(q)) \setminus E_B$ defines a block system with two blocks on $E(P(q))$.

Theorem 3. *Let $q > 1$ be a prime power with $q \equiv 1 \pmod 8$. Suppose that G is a bichromatic subgraph of the Paley graph $P(q)$ for which $P(q)$ is an EPPA witness (using the colouring described before). Then G is an induced subgraph of the 3×3-rook's graph or one of the graphs depicted in Figure 3. In both cases, we have $|V(G)| \leq 9$.*

The 3×3-rook's graph is ultrahomogeneous and hence an EPPA witness for all of its induced subgraphs. The exceptional graphs depicted in Figure 3 are induced subgraphs of the 4×4-rook's

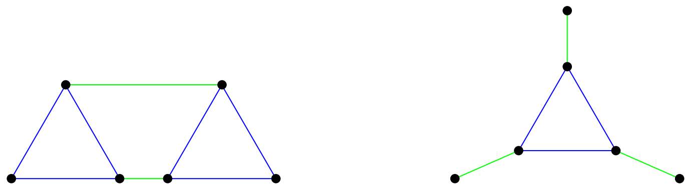


Figure 3: The two exceptional graphs occurring in Theorem 3.

graph, and Theorem 2 shows that the 4×4-rook's graph is an EPPA witness for them. In other words, Theorem 3 asserts that Paley graphs with $q \equiv 1 \pmod 8$ cannot be EPPA witnesses of minimal order of any of their bichromatic induced subgraphs if $q > 16$.

In a forthcoming paper [1], we classify the finite graphs which have an EPPA witness of at most twice their order. This is a natural regime as these EPPA witnesses are rank 3 graphs (see [2, Theorem 1.2]). In this classification, we make use of the classification obtained in Theorem 1. The cases treated in Theorem 2 and 3 occur as subcases.

This note is structured in the following way: in Section 2, we prove Theorem 1. Section 3 contains the application to grid graphs and Paley graphs, i.e., the proofs of Theorems 2 and 3.

2 Classification

In this section, we sketch the proof of the classification given in Theorem 1. Throughout this section, we make the following assumption:

Hypothesis 4. Let H be a finite edge transitive graph such that $E(H) = E_B \dot\cup E_G$ is a block system with two blocks on the edge set $E(H)$ of H. Let G be an induced subgraph of H for which H is an EPPA witness.

As customary, let P_3 denote a path on three vertices and K_l (for $l \in \mathbb{N}$) the complete graph on l vertices.

Blocks as colours For simplicity, we think of blocks as colours, i.e., the edges in E_B are *blue edges* and those in E_G are *green edges.* In contrast to standard edge-coloured graphs, isomorphisms may then *swap* the colours, but they have to respect the partition of edges into colour classes. For a coloured graph K, let $\tilde{K}$ be obtained from K by swapping the colours. We call a graph *monochromatic* if it contains only edges of a single colour, and bichromatic if it contains edges of both colours. Let X and $\tilde{X}$ denote the monochromatic path on three vertices with blue and green edges, respectively, and let Y denote the bichromatic path on three vertices. Moreover, let $F_1, \widetilde{F_1}, F_2, \widetilde{F_2}$ be the bichromatic graphs depicted in Figure 4.

Omitted subgraphs We fix an embedding of G in H. This allows us to think of G as a coloured graph. We say that G *omits* a coloured graph K if there is no colour-preserving isomorphism from K to an induced subgraph of G. Otherwise, G *contains* K. For an uncoloured graph K, we say that G

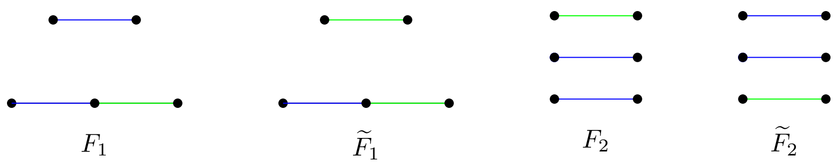


Figure 4: The graphs $F_1, \widetilde{F}_1, F_2, \widetilde{F}_2$.

omits K if G omits all coloured versions of K (or, equivalently, if G, viewed as an uncoloured graph, does not contain an induced subgraph isomorphic to K). For instance, if G omits P_3, then G is a disjoint union of cliques.

We first collect some observations on omitted subgraphs.

Lemma 5. *The graph G omits Y, or both X and $\tilde{X}$. Moreover, G omits F_1, $\widetilde{F}_1$, F_2, and $\widetilde{F}_2$. Furthermore, every triangle in G is monochromatic.*

In the following, let $G_B = G[E_B \cap E(G)]$ and $G_G = G[E_G \cap E(G)]$ be the subgraphs of G induced by the blue and green edges, respectively. We collect some preparatory results for the case that G contains Y.

Lemma 6. *Assume that G contains Y. Then the following hold:*

1. *The graph G has at most one non-singleton component.*
2. *The connected components of G_B and G_G are cliques. If C_B and C_G are connected components of G_B and G_G, respectively, we have $|V(C_B) \cap V(C_G)| \leq 1$.*
3. *Let $c, c' \in \{B, G\}$ with $c \neq c'$. Let C_1 and C_2 be distinct connected components of G_c. Then all edges between C_1 and C_2 are of colour c', and they form a matching.*

In the following, we thus restrict to the case that G is connected. We first consider the following special case.

Lemma 7. *Assume that G is connected and contains Y. If G contains a monochromatic $K_2 \dot\cup K_2$ of colour $c \in \{B, G\}$, then G_c is a disjoint union of cliques C_1 and C_2, and the edges of colour $c' \neq c$ form a matching.*

Proof sketch. Assume $c = B$ and let e be a green edge. Using that G omits the graphs $F_1, F_2, \widetilde{F}_1, \widetilde{F}_2$, we show that e has one endpoint in C_1 and the other in C_2. Since G is connected, there are no additional vertices. $\square$

With this, we are now ready to prove Theorem 1.

Proof sketch of Theorem 1. First assume that G omits Y and that G is not monochromatic. Assume that there exist $u, v, w \in V(G)$ such that $\{u, v\}$ is a blue edge and $\{v, w\}$ is a green edge. As G omits Y, the edge $\{u, w\}$ is present in G. It can neither be blue nor green, as this creates a bichromatic triangle, which is a contradiction by Lemma 5. Thus every connected component of G

is monochromatic. As G omits F_2, G has precisely two non-singleton components, one containing blue and the other green edges.

Now suppose that G contains Y. Then G omits X and $\tilde{X}$ by Lemma 5. By Lemma 6, we may assume that G is connected. Suppose that for some $c \in \{B, G\}$, the graph G_c has distinct connected components C_1 and C_2 with $|V(C_1)| \geq 4$ and $|V(C_2)| \geq 2$. We show that in this case, Lemma 7 applies. Now suppose that, for some $c \in \{B, G\}$, the graph G_c has precisely one non-singleton connected component C. Without loss of generality, assume $c = B$. If also G_G has at most one non-singleton component, then G is the union of two monochromatic cliques intersecting in one vertex as G is connected (see Lemma 6). Thus case 2.(b) applies. If instead G_G has multiple components, we show that each has size at most 2 and thus case 2.(c) applies. If $|V(C)| = 2$, G_G is the disjoint union of two cliques, which are joined by a single blue edge. Then case 2.(a) with interchanged roles of the colours applies.

In the following, we may thus assume that both G_B and G_G have more than one non-singleton connected component, and that each of their components has size at most 3. As the connected components of G_B and G_G are cliques, every vertex in G is incident to at most two blue and two green edges. We may assume that G does not contain a monochromatic $K_2 \dot\cup K_2$ as case 2.(a) applies otherwise. With this, we derive an upper bound on the number of connected components of G_B and G_G, and treat the occurring cases individually. □

3 Applications

In this section, we give some example applications of Theorem 1. We consider rook's graphs and Paley graphs, thus proving Theorems 2 and 3.

Rook's graphs First let H be the $n \times n$-rook's graph. As H is ultrahomogeneous for $n \leq 3$, we may assume $n \geq 4$. Picturing H as an $n \times n$-grid, let E_B denote the vertical edges and E_G the horizontal edges. It is easily seen that $\{E_B, E_G\}$ is a block system with two blocks on $E(H)$. Let G be an induced subgraph of H for which H is an EPPA witness. By Theorem 1, it immediately follows that G has the structure described in Theorem 2. The converse follows from a corollary, observed in [3], of the connected-homogeneity of $K_n \square K_n$ proved in [6] or direct calculation.

Paley graphs Fix a prime $p > 2$ and let $q = p^r$ with $q \equiv 1 \pmod 8$. Set $\mathbb{F}_q^\times = \mathbb{F}_q \setminus \{0\}$. Let $H = P(q)$ be the Paley graph with vertex set $\mathbb{F}_q$. Define E_B as those edges $\{a, b\} \in E(H)$ with $a - b \in (\mathbb{F}_q^\times)^4$, and set $E_G \coloneqq E(H) \setminus E_B$. Note that this is well-defined as $q \equiv 1 \pmod 8$, and thus $a - b \in (\mathbb{F}_q)^4$ if and only if $b - a \in (\mathbb{F}_q)^4$. The automorphism group of H is

$$\mathrm{Aut}(H) = \{x \mapsto ax^{p^i} + b \colon a \in (\mathbb{F}_q^\times)^2,\, b \in \mathbb{F}_q,\, i \in \{0, \dots, r-1\}\}.$$

It is easily seen that $E(H) = E_B \dot\cup E_G$ is a block system on $E(H)$. For $\omega_1, \omega_2 \in \mathbb{F}_q$, let $\mathrm{Aut}(H)_{(\omega_1, \omega_2)}$ denote the elements of $\mathrm{Aut}(H)$ fixing both ω_1 and ω_2. Note that

$$\mathrm{Aut}(H)_{(0,1)} = \{x \mapsto x^{p^i} \colon i \in \{0, \dots, r-1\}\}$$

is a cyclic group of order r. As $P(q)$ is edge transitive, we have $\mathrm{Aut}(H)_{(\omega_1,\omega_2)} \cong \mathrm{Aut}(H)_{(0,1)}$ for all adjacent vertices $\omega_1, \omega_2 \in V(H)$.

Proof sketch of Theorem 3. First let r be odd. Suppose that G contains $K_2 \dot\cup K_2$. Fixing both vertices ω_1, ω_2 of one edge and exchanging the vertices of the other is a partial automorphism of G. By assumption, it extends to an automorphism of even order of H. Thus $\mathrm{Aut}(H)_{(\omega_1,\omega_2)}$ contains an element of even order, which is a contradiction as it is a cyclic group of order r. Hence G omits $K_2 \dot\cup K_2$. With this, we can go through the cases of Theorem 1 to obtain the result. The case that r is even can be treated in a slightly more involved, but similar fashion. $\square$

References

[1] David Bradley-Williams and Sofia Brenner. Graphs with small EPPA witnesses. In preparation.

[2] David Bradley-Williams, Peter Cameron, Jan Hubička, and Matěj Konečný. EPPA numbers of graphs. *Journal of Combinatorial Theory. Series B*, 170:203–224, 2025.

[3] David Bradley-Williams, Peter Cameron, Jan Hubička, and Matěj Konečný. EPPA numbers of graphs II. In preparation.

[4] David M. Evans, Jan Hubička, Matěj Konečný, and Jaroslav Nešetřil. EPPA for two-graphs and antipodal metric spaces. *Proceedings of the American Mathematical Society*, 148:1901–1915, 2020.

[5] Anthony Gardiner. Homogeneous graphs. *Journal of Combinatorial Theory, Series B*, 20(1):94–102, 1976.

[6] Anthony Gardiner. Homogeneity conditions in graphs. *Journal of Combinatorial Theory. Series B*, 24:301–310, 1978.

[7] Stefan Geschke and Menachem Kojman. Symmetrized Induced Ramsey Theory. *Graphs and Combinatorics*, 27(6):851–864, 2011.

[8] Bernhard Herwig and Daniel Lascar. Extending partial automorphisms and the profinite topology on free groups. *Transactions of the American Mathematical Society*, 352(5):1985–2021, 2000.

[9] Ehud Hrushovski. Extending partial isomorphisms of graphs. *Combinatorica*, 12(4):411–416, 1992.

[10] Jan Hubička, Matěj Konečný, and Jaroslav Nešetřil. All those EPPA classes (strengthenings of the Herwig–Lascar theorem). *Transactions of the American Mathematical Society*, 375(11):7601–7667, 2022.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On the distinguishing chromatic number in hereditary graph classes

(Extended abstract)

Christoph Brause* Rafał Kalinowski† Monika Pilśniak‡ Ingo Schiermeyer §

Abstract

The distinguishing chromatic number of a graph G, denoted $\chi_D(G)$, is the minimum number of colours in a proper colouring of vertices of G that is preserved only by the identity automorphism. Collins and Trenk proved that $\chi_D(G) \leq 2\Delta(G)$ for every connected graph G. Moreover, they showed that equality $\chi_D(G) = 2\Delta(G)$ holds only for complete balanced bipartite graphs $K_{p,p}$ and for C_6. In the paper, we show that the upper bound on $\chi_D(G)$ can be substantially reduced if we forbid some small graphs as induced subgraphs of G, that is, we study the distinguishing chromatic number in some hereditary graph classes.

1 Introduction

Let G be a graph, and $c\colon V(G) \to \mathbb{N}$ be a proper vertex colouring. An *automorphism* with respect to the pair (G, c) is a bijective mapping $\varphi\colon V(G) \to V(G)$ such that $c(v) = c(\varphi(v))$ for each $v \in V(G)$, and $vw \in E(G)$ if and only if $\varphi(v)\varphi(w) \in E(G)$ for each $v, w \in V(G)$. The set of automorphisms with respect to (G, c) is denoted by $\mathrm{Aut}(G, c)$. A vertex of G is *fixed* if it is a fixed point of every automorphism of $\mathrm{Aut}(G, c)$. Furthermore, c is *distinguishing* if it fixes every vertex of G. The *distinguishing chromatic number* of a graph G, denoted $\chi_D(G)$, is the minimum number of colours in a proper distinguishing vertex colouring. This concept was introduced by Collins and Trenk [4], who proved the following tight upper bound for $\chi_D(G)$.

Theorem 1. (Collins, Trenk [4]) *If G is a connected graph with maximum degree Δ, then $\chi_D(G) \leq 2\Delta$. The equality holds if and only if $G = K_{\Delta,\Delta}$ or $G = C_6$.*

*TU Bergakademie Freiberg, 09596 Freiberg, Germany and AGH University of Krakow, 30-059 Kraków, Poland. E-mail: brause@math.tu-freiberg.de.de.

†AGH University of Krakow, 30-059 Kraków, Poland. E-mail: kalinows@agh.edu.pl.

‡AGH University of Krakow, 30-059 Kraków, Poland. E-mail: pilsniak@agh.edu.pl.

§TU Bergakademie Freiberg, 09596 Freiberg, Germany and AGH University of Krakow, 30-059 Kraków, Poland. E-mail: ingo.schiermeyer@tu-freiberg.de.

This concept spawned a number of papers. Collins and Trenk [4] showed that $\chi_D(G) = |V(G)|$ if and only if G is a complete multipartite graph. Next, Cavers and Seyffarth [3] characterized the graphs G with $\chi_D(G) \geq |V(G)| - 2$. Laflamme and Seyffarth [10] showed that $K_{\Delta,\Delta-1}$ is the only bipartite graph G with $\chi_D(G) = 2\Delta(G) - 1$. Fijavž, Negami, and Sano [7] proved a notably interesting result that $\chi_D(G) \leq 5$ for every 3-connected planar graph $G \notin \{K_{2,2,2}, C_6 + 2K_1\}$. Cranston [5] confirmed the conjecture of Collins and Trenk that a connected graph $G \neq C_6$ of girth of at least five satisfies $\chi_D(G) \leq \Delta(G) + 1$. Balachandran, Padinhatteer, and Spiga [1] found an infinite family of Cayley graphs G with $\chi_D(G) > \chi(G)$ and relatively small $\mathrm{Aut}(G)$.

Given a family of graphs $H_1, \ldots, H_k$, we say that a graph G is $(H_1, \ldots, H_k)$-*free* if none of the graphs $H_1, \ldots, H_k$ is an induced subgraph of G. For $k = 1$ and $H_1 = H$, we write that G is H-free. The above-mentioned result of Cranston can be stated as follows.

Theorem 2. (Cranston [5]) *If G is a connected (C_3, C_4)-free graph, then $\chi_D(G) \leq \Delta(G) + 1$ unless $G = C_6$.*

In this paper, we provide upper bounds for the distinguishing chromatic number for graphs without one or more induced graphs from the set $\{C_4, 2K_2, K_{1,3}, K_4 - e, K_4\}$. All our bounds are tight, and we characterize graphs that achieve these bounds.

A vertex v of a graph G is called *universal* if $N_G[v] = V(G)$. And a vertex v in a graph G is *simplicial* if $N_G(v)$ is a clique. Given a set $S \subseteq V(G)$, we say that a vertex v is *complete to* S if $S \subseteq N_G(v)$, and *anti-complete to* S if $S \cap N_G(v) = \emptyset$. A set $M \subseteq V(G)$ is a *module of* G if every vertex outside M is complete or anti-complete to M.

If G and H are two vertex-disjoint graphs, then $G + H$ stands for their *join*, that is, the graph with $V(G+H) = V(G) \cup V(H)$ and $E(G+H) = E(G) \cup E(H) \cup \{uv : u \in V(G), v \in V(H)\}$.

It is worth mentioning that our proofs use various new techniques that are completely different from those used in previous papers in the area. None of them are short (except for the one outlined in Section 4), so we omit them in this extended abstract.

2 C_4-free graphs and $2K_2$-free graphs

Our first result is an improvement of Theorem 2 of Cranston.

Theorem 3. *If G is a connected C_4-free graph, then*

$$\chi_D(G) \leq \Delta(G) + 1$$

unless $G = C_6$.

The proof is based on two key lemmas.

Lemma 4. *If G is a connected graph, the set S of simplicial vertices in G is non-empty, and $\chi_D(G) > \chi_D(G - S)$, then either $\chi_D(G) \leq \Delta(G)$ or G is complete multipartite with a universal vertex.*

Lemma 5. *If G is a connected C_4-free graph and $u \in V(G)$ is a vertex such that*

$$\chi_D(G[N_G[u]]) \leq \Delta(G),$$

then

$$\chi_D(G) \leq \Delta(G) + 1.$$

Next, we consider graphs that are both C_4-free and $2K_2$-free, and we prove the following theorem.

Theorem 6. *If G is a connected $(C_4, 2K_2)$-free graph, then*

$$\chi_D(G) \leq \Delta(G) + 1$$

with equality if and only if $G = \alpha(G)K_1 + K_{\omega(G)-1}$ or $G = C_5$.

In the proof, we actually use the result for chordal graphs. We need to define some additional notions beforehand.

A tree T is *symmetric* if all non-leaves have maximum degree, one of these vertices is a central vertex of T, and every leaf has the same distance to the central vertex. We call a graph G *symmetric* if it is a symmetric tree or can be constructed from a symmetric tree T by

A) either adding all edges between all leaves of $N_T(v)$ for each support $v \in V(T)$,

B) or adding all edges between all leaves of $N_T(v)$ and adding a new vertex v' which is adjacent to all leafs of $N_T(v)$ for each support $v \in V(T)$.

A graph G is *chordal* if it is C_k-free for each $k \geq 4$. The following result is used as a lemma for proving Theorem 6 but it may be of interest as such.

Theorem 7. *If G is a connected chordal graph, then*

$$\chi_D(G) \leq \Delta(G) + 1$$

with equality if and only if G is a symmetric graph or $G = \alpha(G)K_1 + K_{\omega(G)-1}$.

At the end of this Section we provide a result on $2K_2$-free graphs.

Theorem 8. *If G is a connected $2K_2$-free graph, then*

$$\chi_D(G) \leq 2\Delta(G) - \omega(G) + 2$$

with equality if and only if G is a complete graph or a balanced complete bipartite graph.

The proof is a very careful analysis of structure of graphs with forbidden $2K_2$ as induced subgraphs. In particular, such a graph G contains a dominating clique of size $\omega(G)$ unless $\omega(G) = 2$. Then we can define a proper colouring step by step, and finally, we show that the colouring is distinguishing.

3 claw-free graphs

In the proofs in this Section we use the following *modular partition* of a graph. Let G be a connected non-complete graph. A non-complete dominating module is *minimal* in G if it cannot be partitioned into two non-complete dominating modules of G. We denote by $p(G)$ the largest integer p such that $V(G)$ can be partitioned into p pairwise disjoint modules $P_1, P_2, \dots, P_p$ of G each of which is non-complete and dominating. We further let $p(G) = 0$ if G is complete.

A star $K_{1,3}$ is called a claw. And, observe that the lexicographic product $K_{n/2}[2K_1]$ is isomorphic to $K_n - M$, where M is a perfect matching in K_n with even n.

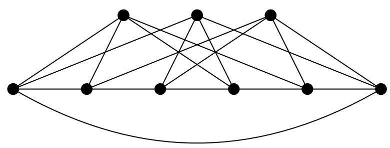


Figure 1: The line graph $L(K_{3,3})$

The main result of this Section follows.

Theorem 9. *If G is a connected claw-free graph of order n, then*

$$\chi_D(G) \leq \Delta(G) + 2$$

with equality if and only if $G = C_6$ or $G = K_{n/2}[2K_1]$.

Actually, this is a corollary of the following general result with a very long technical proof that uses structural Lemma 11.

Theorem 10. *If G is a connected claw-free graph, then*

$$\chi_D(G) \leq \chi(G) + p(G)$$

unless $G \in \{C_6, L(K_{3,3})\}$.

Lemma 11. *If G is a connected non-complete claw-free graph, P is a minimal non-complete dominating module of G, $S \subseteq P$ is a set of vertices which induces a non-complete but connected graph, and $c\colon V(G[P]) \to \mathbb{N}$ is a vertex colouring that fixes all vertices of S, then all vertices of P are fixed.*

4 (claw, diamond)-free graphs

A graph $K_4 - e$ is called a *diamond*. The *distinguishing chromatic index* of a graph G, denoted by $\chi'_D(G)$, is the minimum number of colours in a proper edge-colouring c such that each vertex of G is a fixed point of every automorphism of G that preserves the colouring c.

Theorem 12. *If G is a connected (claw, diamond)-free graph, then*

$$\chi_D(G) \leq \Delta(G) + 1$$

unless $G \in \{C_4, C_6\}$.

We give here an outline of the proof.Every (claw, diamond)-free graph G is a line graph $L(H)$ of some graph H. This is because every Beineke graph (cf. [2]) different from the claw contains a diamond as an induced subgraph. Let $G = L(H)$. Consequently, $\chi_D(G) = \chi'_D(H)$ by the Whitney isomorphism theorem, except for three small graphs (cf. [8]). Kalinowski and Pilśniak [9] proved that $\chi'_D(H) \leq \Delta(H)+1$ except for four small graphs. It is easy to see that $\Delta(G) \geq \Delta(H) - 1$ and the equality holds only if and only if every edge incident to a vertex of degree $\Delta(H)$ is a pendant vertex in H, that is, H is a star $K_{1,n-1}$.

The equality $\chi_D(G) = \Delta(G) + 1$ is achieved for line graphs $G = L(H)$ of Class 2 graphs H.

If $L(H)$ is K_4-free, then $\Delta(H) \leq 3$. Hence, we have the following.

Corollary 13. *If a graph G is (claw, diamond, K_4)-free, then $\chi_D(G) \leq 4$.*

References

[1] N. Balachandran, S. Padinhatteer, and P. Spiga, Vertex transitive graphs G with $\chi_D(G) > \chi(G)$ and small automorphism group. *Ars Math. Contemp.*, 17(1): (2019) 311–318.

[2] L. W. Beineke, Characterizations of Derived Graphs. *J. Combin. Th.*, 9:(1970) 129–135.

[3] M. Cavers and K. Seyffarth, Graphs with large distinguishing chromatic number. *Electron. J. Combin.*, 20(1): (2013) #P19.

[4] K. Collins and A. Trenk, The distinguishing chromatic number. *Electron. J. Combin.*, 13(1): (2006) #R16.

[5] D. Cranston, Proper Distinguishing Colorings with Few Colors for Graphs with Girth at Least 5. *Electron. J. Combin.* 25(3): (2018) #P3.5.

[6] R. Diestel, Graph Theory, 5th Edition. Springer-Verlag, Berlin, Heidelberg, New York 2016.

[7] G. Fijavž, S. Negami, and T. Sano, 3-connected planar graphs are 5-distinguishing colorable with two exceptions. *Ars Math. Contemp.*, 4 (2011) 165–175.

[8] H. A. Jung, Zu einem Isomorphiesatz von H. Whitney für Graphen. *Math. Ann.*, 164 (3): (1966) 270–271.

[9] R. Kalinowski and M. Pilśniak, Distinguishing graphs by edge-colourings. *European J. Combin.*, 45: (2015) 124–131.

[10] C. Laflamme and K. Seyffarth, Distinguishing chromatic numbers of bipartite graphs. *Electron. J. Combin.*, 16(1): (2009) #R76.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Blind cop-width and balanced minors of graphs

(Extended abstract)

Hector Buffière*† Rutger Campbell†‡ Kevin Hendrey†§ Sang-il Oum†‡

Abstract

In this paper, we investigate a pursuit-evasion game where the robber is slow and invisible to the cops. The blind cop-width is the minimum number of cops needed to win the game on a given graph. We link it with other known graph parameters defined in terms of pursuit-evasion games, and show a new lower-bound with respect to the treewidth. The proof introduces the notion of balanced minors, where all bags of a minor model have equal size.

1 Introduction

Pursuit-evasion games on graphs have been studied both from a game-theoretic point of view and through their connections with structural graph theory. The cops & robber games have been identified as an alternative definition of pathwidth and treewidth in the early days of these structural parameters. Both have since played a crucial role in the recent progress of structural graph theory through the *graph minor project* of Robertson and Seymour, and have been used in the development of many parameterized algorithms.

In the classical version of the game, a robber occupies a vertex of a graph and tries to escape a fixed number of cops. At each turn, some cops use helicopters to fly to new vertices, announcing their next positions and leaving the robber the time to move along edges and free vertices. If the robber cannot move to a position that is unoccupied in the next round, the cops catch the robber and win. In this setting, the minimum number of cops needed to be certain to catch the robber is equal to the treewidth of the graph plus one [ST93]. If one restricts the cops to not be able to see the position of the robber, it is equal to the pathwidth plus one [KP85]. One can instead restrict the maximum length of a path the robber is allowed to move along, which can be thought of as considering a robber with finite speed. According to the speed r, these variants are equivalent to other structural parameters, including degeneracy

*Université Paris Cité, CNRS, IRIF, Paris, France and Centre d'Analyse et de Mathématique Sociales, CNRS/EHESS, Paris, France. E-mail: buffiere@irif.fr.
†Supported by the Institute for Basic Science (IBS-R029-C1).
‡Discrete Mathematics Group, Institute for Basic Science (IBS), Daejeon, South Korea. E-mail: rutger@ibs.re.kr, sangil@ibs.re.kr
§School of Mathematics, Monash University, Melbourne, Australia. E-mail: Kevin.Hendrey1@monash.edu

($r = 1$) and the generalized coloring numbers, which were identified by Zhu [Zhu09] as central notions of Nešetřil and Ossona de Mendez's sparsity theory [NO12] (see [Sie25] for a recent survey article about generalized coloring numbers).

In this paper, we consider both conditions at the same time, i.e. the case of a slow and invisible robber, and call the corresponding parameter *blind cop-width*. This work was originally motivated by a paper of Toruńczyk [Tor23], in which cops and robber games were generalized in order to provide a tool suitable for the study of dense graphs, related to the problem of FO model checking on hereditary classes of graphs. Our parameter can be seen as a sparse equivalent to a parameter introduced in this paper, *blind flipwidth*. As a first contribution we answer a conjecture of [Tor23] by showing that every binary tree admits a subdivision where 3 cops have a winning strategy.

The study of pursuit-evasion games an active field by itself (not necessarily motivated by structural problems). As such, quite a number of papers studied game parameters very similar to blind cop-width (see Section 2 for a discussion). Those previous works highlight that this parameter is actually not very well behaved, as exemplified by the fact that it can either increase or decrease after subdividing the graph. In particular, the usual techniques for proving non-trivial lower bounds seem to fail. As the main contribution of this paper, we prove that blind cop-width is functionally larger than treewidth.

Theorem 1. *For every integer k there is an integer $g(k)$ such that every graph of treewidth at least $g(k)$ has blind cop-width at least k.*

In order to prove this result, we introduce the notion of *balanced minors*, which we define as usual graph minors in which each bag of the minor model has an equal number of vertices. We show that balanced minor models of large binary trees have high blind cop-width, and conclude by showing large binary trees can be found as balanced minors in graphs of high treewidth using the celebrated grid theorem. By identifying relationships between blind cop-width and other studied games (made explicit in Section 2), our result answers some problems left open by Bernshteyn and Lee in [BL22]. Namely, we show that both cliques and planar graphs have unbounded *topological inspection number*.

2 Pursuit-evasion games

Following the notation and analysis of [Tor23], we focus on "helicopter" games with varying speed $r \in \mathbb{N} \cup \{\infty\}$ of the robber. We write $\mathrm{copwidth}_r(G)$ for the minimum number of cops needed to catch a robber of speed r on a given graph G. In the case of the blind variant, the associated blind cop-width parameter is written bcw_r. Unlike in the visible setting, the different radii actually yield equivalent parameters, in the sense that one is bounded on a class of graphs if and only if the others are bounded.

Lemma 2. *All blind cop-width parameters with finite radius are functionally equivalent.*

This is shown by noticing that, since the cops are blind, they do not get more information after one turn. Hence, they can simulate two turns at once with twice as many cops, giving the inequality $\mathrm{bcw}_{2r} \leq 2 \cdot \mathrm{bcw}_r$. As a result, in the rest of the paper we focus only on the case $r = 1$ and simply write bcw for it.

While the requirement that cops announce their next positions provides clean structural characterisations, it is perhaps less natural from a game-theoretic point of view, and indeed some related works drop this assumption. Nevertheless, similar strategy simulation arguments show that most variants of the game are linearly tied.

Proposition 3. *Let in() be the inspection number of [BL22]. For every graph G, $in(G) \leqslant bcw(G) \leqslant 2 \cdot in(G)$.*

Such results allow us to extend our results to other parameters, in particular lower-bounds on bcw for some classes of graphs. As such, we also mention the straightforward connection with the variant of the game when cops can only move along vertices.

Proposition 4. *Let $c_0()$ be the parameter of [DDTY15b] and [XYZZ19]. For every graph G, $bcw(G) \leqslant 2 \cdot c_0(G)$.*

3 Blind flip-width

The motivating question of this work was to answer the conjecture in [Tor23] that bounded blind flip-width is equivalent to bounded linear clique-width. We approach this conjecture by its implications in the weakly sparse setting: here by a result of Gurski and Wanke [GW00] we have that on any weakly sparse class of graphs, linear clique-width and pathwidth are functionally equivalent. The straightforward lemma 5 shows that bounded blind cop-width implies bounded blind flip-width. Thus, the conjecture would imply that any weakly sparse class of graphs with bounded blind cop-width also has bounded pathwidth, but we give a counterexample to this property.

Lemma 5. *For every integer r, bfw_r is bounded by a function of bcw.*

The counterexample is a sequence of subdivisions of arbitrary large binary trees (which have unbounded pathwidth) where 3 cops are always sufficient to win the blind cop-width game.

Example 6. Let T_0 be a single vertex, $i \in \mathbb{N}$ and let us assume that $\mathrm{bcw}_1(T_i) \leq 3$. Let l be the smallest integer such that 3 cops can win in T_i in the blind cop-width game in l rounds, we construct T_{i+1} by taking two disjoint copies $T_{i,1}$ and $T_{i,2}$ of T_i along with a new vertex u, and link them with one edge between u and $T_{i,1}$ and a path of length $l+1$ between u and $T_{i,2}$.

A winning blind strategy for 3 cops on T_{i+1} can be achieved in 3 steps. First, the 3 cops clean all the vertices adjacent to edges in $T_{i+1} \setminus (T_{i,1} \cup T_{i,2})$ (this is a path of length $l+2$). Then, they all fly to $T_{i,1}$ and clean it using the strategy in l turns. After these turns, the gas from $T_{i,2}$ has recontaminated the path linking it to u, except for u itself (see Figure 1). The cops then occupy u and its neighbour on the path, and clean the path again. Lastly, the 3 cops fly to $T_{i,2}$ and clean it recursively. △

After using the weakly sparse setting to answer a more general conjecture, a natural question is whether this relationship is a strong as we could expect: since in weakly sparse graphs $\mathrm{copwidth}_r \simeq \mathrm{fw}_r$, do we also have $\mathrm{bcw} \simeq \mathrm{bfw}_r$? However this is false as witnessed by the following example.

Example 7. Let T_i' be obtained from the above defined T_i by adding $|T_i|+1$ children to every leaf. Once the flipper has restrained the robber to a subtree rooted in a leaf of T_i by simulating the strategy the cops have for T_i, he can use one more flip to delete all the new edges adjacent to the leaf, successfully isolating the robber. On the other hand, one can see that T_i' contains a balanced minor model of a binary tree of height $\lfloor i/2 \rfloor$, and as such has blind cop-width at least $\lfloor i/2 \rfloor$ by Proposition 11. △

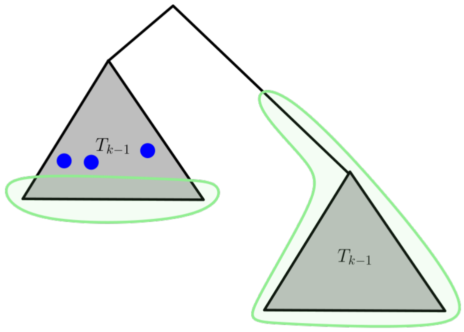


Figure 1: The state of the game when the cops are cleaning the first subtree

As we have seen blind cop-width is closely related but not equivalent to blind flip-width on weakly sparse graphs. However, we can retrieve functional equivalence if we require the stronger condition that the graphs have bounded maximum degree.

Proposition 8. *For every graph G with maximum degree Δ, there exists a function f_Δ depending only on Δ such that $bcw(G) \leqslant f_\Delta(bfw_3(G))$.*

4 Balanced minors

We strengthen the definition of a graph minor by further imposing that each bag of a minor model has equal size, and call such a model a *balanced minor model.* In this setting, it appears that outerplanar graphs play a similar role as planar graphs for classical minors.

Proposition 9. *For every outerplanar graph G there is an integer $f(G)$ such that all graphs containing the $f(G) \times f(G)$ grid as a minor contain G as a balanced minor.*

This result is in fact optimal in the sense that outerplanar graphs are exactly the graphs whose balanced minor containment forbids arbitrary large grid minors.

Proposition 10. *For every graph G that is not outerplanar, there exists arbitrarily large grid minors that do not contain G as a balanced minor.*

We also show that graphs containing large clique as minors contain large cliques as balanced minors, showing that the topological inspection number of cliques is unbounded (see next section for a definition).

5 Bounds for blind cop-width

Few non-trivial lower-bounds have been found so far for blind pursuit-evasion games, and the notion of balanced minors allows us to use counting arguments as well as structural properties to prove the following quite technical result.

Proposition 11. *For all $k \in \mathbb{N}$, there exists $f(k) \in \mathbb{N}$ such that every balanced minor model of a binary tree of height at least $f(k)$ has blind cop-width at least k.*

Since (binary) trees are outerplanar, they appear as balanced minors in every sufficiently large grid minor by Proposition 9, and as such in every graph of large treewidth Robertson and Seymour's grid theorem [RS86].

Theorem 12. *For every integer k there is an integer $g(k)$ such that every graph of treewidth at least $g(k)$ has blind cop-width at least k.*

Proof: Let f_1 obtained from Proposition 11, f_2 from Proposition 9, and f_3 from the grid theorem, then every graph with treewidth at least $f_3(f_2(f_1(k)))$ has blind cop-width at least k. $\square$

The previous theorem shows that a structural necessary condition for having high blind cop-width is high treewidth, as if G contains a graph H of high treewidth then G must have high blind cop-width no matter how much it is subdivided (as subdividing preserves treewidth). However, binary trees of high pathwidth can also cause high blind cop-width (being balanced minors of themselves), even if they can have treewidth 1, which shows that this condition is not sufficient. The case of trees is nonetheless less structurally robust, as those trees of high blind cop-width can always be subdivided to lower the blind cop-width to 3. One can then wonder if this situation always occurs, namely can all graphs of treewidth k be subdivided to graphs of blind cop-width no more than $f(k)$ for some function f? This was considered in [BL22] which introduced the topological inspection number, that can be viewed as a supremum limit of blind cop-width over all subdivisions of a fixed graph. We answer the question positively, essentially proving that the topological inspection number is tied to treewidth.

Theorem 13. *For every graph G of treewidth k, there is a subdivision H of G such that $bcw(H) \leqslant k + 3$.*

References

[AFGP16] Tatjana V. Abramovskaya, Fedor V. Fomin, Petr A. Golovach, and Michal Pilipczuk. How to hunt an invisible rabbit on a graph. *European Journal of Combinatorics*, 52:12–26, 2016.

[BL22] Anton Bernshteyn and Eugene Lee. Searching for an intruder on graphs and their subdivisions. *The Electronic Journal of Combinatorics*, 29(3), 2022.

[CE12] Bruno Courcelle and Joost Engelfriet. *Graph structure and monadic second-order logic: a language-theoretic approach*, volume 138. Cambridge University Press, 2012.

[DDTY15a] Dariusz Dereniowski, Danny Dyer, Ryan M. Tifenbach, and Boting Yang. The complexity of zero-visibility cops and robber. *Theorical Computer Science*, 607:135–148, 2015.

[DDTY15b] Dariusz Dereniowski, Danny Dyer, Ryan M. Tifenbach, and Boting Yang. Zero-visibility cops and robber and the pathwidth of a graph. *Journal of Combinatorial Optimization*, 29(3):541–564, 2015.

[GW00] Frank Gurski and Egon Wanke. The tree-width of clique-width bounded graphs without $k_{n,n}$. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 196–205. Springer, 2000.

[KP85] Lefteris M Kirousis and Christos H Papadimitriou. Interval graphs and searching. *Discrete Mathematics*, 55(2):181–184, 1985.

[NO12] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity - Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012.

[RS83] Neil Robertson and Paul D. Seymour. Graph minors. i. excluding a forest. *Journal of Combinatorial Theory, Series B*, 35(1):39–61, 1983.

[RS86] Neil Robertson and Paul D Seymour. Graph minors. v. excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, 1986.

[Sie25] Sebastian Siebertz. On the generalized coloring numbers. *arXiv preprint arXiv:2501.08698*, 2025.

[ST93] Paul D. Seymour and Robin Thomas. Graph searching and a min-max theorem for tree-width. *Journal of Combinatorial Theory, Series B*, 58(1):22–33, 1993.

[Tor23] Szymon Toruńczyk. Flip-width: Cops and robber on dense graphs. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 663–700, 2023.

[TU15] Satoshi Tayu and Shuichi Ueno. On evasion games on graphs. In *Discrete and Computational Geometry and Graphs - 18th Japan Conference, JCDCGG 2015*, volume 9943, pages 253–264, 2015.

[XYZ22] Yuan Xue, Boting Yang, and Sandra Zilles. A simple method for proving lower bounds in the zero-visibility cops and robber game. *Journal of Combinatorial Optimization*, 43(5):1545–1570, 2022.

[XYZZ19] Yuan Xue, Boting Yang, Farong Zhong, and Sandra Zilles. A partition approach to lower bounds for zero-visibility cops and robber. In *Combinatorial Algorithms - 30th International Workshop, IWOCA 2019*, volume 11638, pages 442–454, 2019.

[Zhu09] Xuding Zhu. Colouring graphs with bounded generalized colouring number. *Discrete Mathematics*, 309(18):5562–5568, 2009.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The Hitting Time Of Nice Factors

(Extended abstract)

Fabian Burghart[*] Marc Kaufmann[†] Noela Müller[‡] Matija Pasch[§]

Abstract

Consider the random u-uniform hypergraph (short: u-graph) process on n vertices, where n is divisible by $r > u \geq 2$. It was recently shown that with high probability, as soon as every vertex is covered by a copy of the complete u-graph K_r on r vertices, the process contains a K_r-factor (RSA, Vol. 65 II, Sept. 2024). This hitting time result is based on a process coupling which builds on a static coupling that was used to establish the corresponding sharp threshold result (RSA, Vol. 61 IV, Dec. 2022). The latter coupling, however, does not only work for *complete* u-graphs, but for the broader class of so-called *nice* u-graphs.

The purpose of this article is to extend the process coupling to nice u-graphs, thereby matching the full scope of the static coupling. As a byproduct, we obtain the extension of the hitting time result to nice (u-graph) factors. Since the relevant combinatorial bounds in the proof for the K_r-case do not generalize easily, we introduce new arguments that do not only apply to nice u-graphs, but make a first step towards the even broader class of strictly 1-balanced u-graphs. Further, we show how the remainder of the process coupling for the K_r-case can be utilized in a black-box manner, for any u-graph. These advances pave the way for future generalizations.

The full proof is available at https://arxiv.org/abs/2409.17764.

1 Introduction

Which properties of a graph guarantee the existence of a perfect matching? A partial answer to this question was given by Petersen over 130 years ago [9]. Sixty years later, in 1947, it was completely resolved by Tutte [11].

[*]Eindhoven University of Technology, Department of Mathematics and Computer Science. E-mail: `f.burghart@tue.nl`. Funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement no. 101034253.

[†]ETH Zürich, Department of Computer Science. E-mail: `marc.kaufmann@inf.ethz.ch`. M. K. gratefully acknowledges support by the Swiss National Science Foundation [grant number 200021_192079].

[‡]Eindhoven University of Technology, Department of Mathematics and Computer Science. E-mail: `n.s.muller@tue.nl`. N.M. is supported by the NWO Gravitation project NETWORKS under grant no. 024.002.003.

[§]Munich, Germany. E-mail: `matija.pasch@gmail.com`.

In the setting of hypergraphs and introducing randomness, the following closely related question arises: Under which conditions are we *likely* to find a perfect matching in the *random r-uniform hypergraph* $H_r(n,\pi)$ with vertex set $[n] = \{1,\dots,n\}$, where each hyperedge is present independently with probability π? For the graph case $r = 2$, this was answered by Erdős and Rényi [2], who observed that $p^* = p^*(n) = \log(n)/n$ is a *sharp threshold* for the existence of a perfect matching — in the following sense. Let $\varepsilon > 0$ be arbitrarily small. Then for any sequence $p \le (1-\varepsilon)p^*$, the *random graph* $G(n,p) = H_2(n,p)$ does not contain a perfect matching with high probability (whp), that is, with probability tending to one as n increases. On the other hand, for $p \ge (1+\varepsilon)p^*$, the random graph $G(n,p)$ does contain a perfect matching whp.

The surprisingly challenging extension of this result to hypergraphs, that is, to $r > 2$, became known as Shamir's problem. Since hypergraphs with isolated vertices cannot contain perfect matchings, and since the sharp threshold for the disappearance of isolated vertices and the sharp threshold for the existence of perfect matchings coincide for graphs, it was conjectured early on that the sharp threshold

$$\pi_0 = \binom{n-1}{r-1}^{-1} \log(n)$$

for the disappearance of isolated vertices in $H_r(n,\pi)$ is also the sharp threshold for the existence of perfect matchings for any $r \ge 2$. It took 40 years until Kahn fully resolved this conjecture in his recent breakthrough papers [7, 8]:

Theorem 1.1 ([8, Theorem 1.4])**.** *For any $r \ge 3$ and $\varepsilon > 0$, whp the r-uniform random hypergraph $H_r(n,(1+\varepsilon)\pi_0)$ contains a perfect matching.*

Shortly after the appearance of [8], Riordan and Heckel [10, 3] were able to employ Theorem 1.1 to obtain the sharp threshold for the existence of K_r-factors in $G(n,p)$. Recall that a K_r-factor is a spanning subgraph composed of vertex-disjoint copies of the complete graph K_r. Clearly, these questions are related: If we replace each hyperedge of a given r-uniform hypergraph with a perfect matching by a copy of K_r, we obtain a graph on the same vertex set that contains a K_r-factor.

Our previous observation regarding isolated vertices and perfect matchings also applies to the new setup: Say a vertex is *K_r-isolated* if it is not contained in a copy of K_r. Then a graph with K_r-isolated vertices cannot contain a K_r-factor. The sharp threshold $p_0 = \pi_0^{1/\binom{r}{2}}$ for the existence of K_r-isolated vertices in $G(n,p)$ is well-known [5, Theorem 3.22]. Guided by the results for perfect matchings, it was conjectured that p_0 also is a sharp threshold for the existence of a K_r-factor in $G(n,p)$. Instead of attempting a laborious adaptation of the proof of Theorem 1.1 to $G(n,p)$, Riordan and Heckel devised an ingenious embedding approach:

Theorem 1.2 ([10, 3])**.** *For any $r > 2$ there exist $\delta, \varepsilon > 0$ such that the following holds. Let $p \le n^{-2/r+\varepsilon}$ and $\pi = (1-n^{-\delta})p^{\binom{r}{2}}$, then there exists a coupling of $G(n,p)$ and $H_r(n,\pi)$ such that whp for every hyperedge h in $H_r(n,\pi)$ the copy of K_r on the vertex set h is contained in $G(n,p)$.*

From this coupling it is not only immediate that $p_0 = \pi_0^{1/\binom{r}{2}}$ is a sharp threshold for the existence of a K_r-factor in $G(n,p)$. It also gives a rigorous foundation for the correspondence of the hypergraph model and the graph model as described above, even though the random graph gives rise to more intricate dependencies between copies of K_r.

1.1 Sharper than sharp thresholds

In 1985, Bollobás and Thomason [1] observed that the close relationship between the disappearance of isolated vertices and the emergence of perfect matchings runs deeper than the mere equality of the thresholds. Let $N_r = \binom{n}{r}$ and let $(H_t^r)_{t=0}^{N_r}$ be the standard *random hypergraph process*, starting with the empty hypergraph H_0 with vertex set $[n]$ and adding one of the hyperedges of size r, which are not already present, uniformly at random in each time step. Let

$$T_H^r = \min\{t : H_t^r \text{ has no isolated vertices}\}$$

be the time t where the last isolated vertex disappears. For the random graph process $(G_t)_t = (H_t^2)_t$, Bollobás and Thomason showed that, whp, as soon as we reach T_H^2, that is, the earliest possible time, the graph $G_{T_H^2}$ contains a perfect matching. Kahn [7] extended this beautiful result to r-uniform hypergraphs for all $r \geq 3$.

Theorem 1.3 ([7, Theorem 1.3]). *Let $r \geq 3$ and $n \in r\mathbb{Z}_+$, then whp $H^r_{T_H^r}$ has a perfect matching.*

The coupling in Theorem 1.2 is one-sided: We 'embed' $H_r(n, \pi)$ into $G(n, p)$. Thus, around the time of the appearance of [10, 3], it seemed unlikely that an analogous coupling approach could be used to translate Theorem 1.3 to K_r-factors. Nevertheless, Heckel, Kaufmann, Müller and Pasch [4] recently provided an extension of the 'static' coupling to a process coupling, thereby enabling a transferral of the hitting time result. Let

$$T_G = \min\{t : \text{ every vertex in } G_t \text{ is contained in at least one copy of } K_r\}.$$

Theorem 1.4 ([4, Theorem 1.6]). *For any $r \geq 3$, there exists a coupling of G_{T_G} and $H^r_{T_H^r}$ such that whp for every hyperedge h of $H^r_{T_H^r}$ the copy of K_r on the vertex set h is contained in G_{T_G}.*

1.2 Nice hypergraphs

Theorems 1.3 and 1.4 establish that, whp, as soon as every vertex is covered by a copy of K_r, a K_r-factor exists. The question arises for which other graphs F such a hitting time result might hold true. Indeed, there is a long-standing conjecture [6, below Conjecture 1.1] that an analogous result does not only hold for K_r-factors, but F-factors within the u-uniform hypergraph process, for any strictly 1-balanced u-uniform hypergraph F. Riordan's coupling result, Theorem 1.2, also provides a comparison tool beyond the case $F = K_r$.

For brevity, we henceforth refer to u-uniform hypergraphs as u-graphs. When we consider more general u-graphs F than complete u-graphs, more than one copy of F might be present on a given vertex set. Hence, in a coupling approach, simple r-uniform hypergraphs cannot adequately encode copies of F anymore. Instead, we resort to *F-graphs*, which are r-uniform labeled multi-hypergraphs with vertices $[n]$, where hyperedges are labelled by the possible copies of F on the hyperedge vertex set. Note that for $F = K_r$, an F-graph is just a simple r-uniform hypergraph (if we ignore the labels). Further, we may define the *random F-graph* $H_F(n, \pi)$ with vertices $[n]$, where we include each F-edge independently with probability π.

We call a u-graph *nice* if it is strictly 1-balanced, 3-connected and, either $u \geq 3$, or $u = 2$ and F cannot be transformed into an isomorphic graph by deleting one edge and adding a different edge.

Examples for the graph case $u = 2$ are complete graphs K_r for $r \geq 4$ and complete bipartite graphs $K_{m,n}$ for $m, n \geq 3$.

Let $v(F)$ and $e(F)$ be the number of vertices and edges of F respectively.

Theorem 1.5 ([10, Theorem 18]). *Let F be a fixed nice u-graph, $u \geq 2$, with $v(F) = r \geq 4$ and $e(F) = s$. Let $d_1 = s/(r-1)$. Then there are positive constants $\varepsilon(F), \delta(F) > 0$ such that, for any $p = p(n) \leq n^{-1/d_1+\varepsilon}$, letting $\pi = (1 - n^{-\delta})p^s$, we may couple the random u-graph $G = H_u(n,p)$ with the random F-graph $H = H_F(n,\pi)$ so that, whp, for every F-edge in H the corresponding copy of F in G is present.*

Given this coupling, the sharp threshold for the existence of F-factors follows as a corollary. Let $\mathrm{aut}(F)$ be the number of automorphisms of F.

Theorem 1.6 ([10, Theorem 10]). *For a fixed nice u-graph F with $v(F) = r$ and $e(F) = s$,*

$$p_0 = ((\mathrm{aut}(F)/r)n^{-r+1}\log n)^{1/s}$$

is a sharp threshold for $H_u(n,p)$ to contain an F-factor.

1.3 Main result

The purpose of this article is to establish the hitting time version of Theorem 1.6, to showcase the adaptability of the coupling approach in [4], and to make a first step towards general strictly 1-balanced u-graphs. Thus, for the random u-graph process $(G_t)_t = (H_t^u)_{t=0}^{N_u}$, let

$$T_G = \min\{t : \text{ every vertex in } G_t \text{ is contained in at least one copy of } F\}$$

be the first time where every vertex is covered by a copy of F. We also need a process version of the random F-graph $H_F(n,\pi)$ in Section 1.2. The total number of F-edges on the vertex set $[n]$ is

$$M = \binom{n}{r}\frac{r!}{\mathrm{aut}(F)}.$$

Now, the *random F-graph process* $(H_t)_{t=0}^{M}$ starts with the empty F-graph H_0 with vertices $[n]$, and in each step we add a new F-edge, uniformly at random. Let

$$T_H = \min\{t : H_t \text{ has no isolated vertices}\}$$

be the time where the last isolated vertex in H_t disappears. Then the following holds.

Theorem 1.7. *Let F be a fixed nice u-graph with $r \geq 4$ vertices. Then we may couple the random u-graph process $(G_t)_{t=0}^{N_u}$ with the random F-graph process $(H_t)_{t=0}^{M}$ so that, whp, for every hyperedge in H_{T_H} the corresponding copy of F is contained in G_{T_G}.*

Theorem 1.3 and Theorem 1.7 together imply that vertices not contained in any copy of F are essentially the only obstruction for an F-factor in the random u-graph process.

Corollary 1.8. *Let F be a fixed nice u-graph with $r \geq 4$ vertices. For $n \in r\mathbb{Z}_+$, whp G_{T_G} contains an F-factor.*

2 Proof Sketch

While we do build on the ideas in [4], we derive a novel bound which can be used in the proof for general strictly 1-balanced u-graphs. Further, we present a short argument which allows to use the last two parts of the proof in [4] in a 'black-box' manner, for any u-graph F. In fact, we obtain all process-related parts in [4] for general u-graphs.

Fix an arbitrary function $g(n) = o(\log n/\log\log n)$ with $g(n) \to \infty$, and define the probabilities

$$\pi_\pm = \frac{\mathrm{aut}(F)}{r!} \cdot \frac{\log n \pm g(n)}{\binom{n-1}{r-1}},$$

which mark the start and the end of the 'critical window' for the disappearance of isolated vertices in $H = H_F(n,\pi)$. Using Theorem 1.6, we let the critical window for the disappearance of vertices not contained in a copy of F in $H_u(n,p)$ be given by $p_\pm = (\pi_\pm/(1-n^{-\delta}))^{1/s}$, where $s = e(F)$.

There are a number of bad events for H that might preclude the success of our coupling approach. These are:

$\mathcal{B}_1(\pi)$: H contains a vertex of degree more than

$$M\pi + \max\left(M\pi, 3\log n\right).$$

$\mathcal{B}_2$: H contains an *avoidable configuration* A, i.e. an F-graph with at most 2^{r+1} edges and nullity $(r-1)e(A) + c(A) - v(A) > 1$, where $c(A)$ is the number of components in A.

$\mathcal{B}_3$: H contains more than $(\log n)^{8g(n)}$ *low-degree vertices*, that is, of degree at most $7g(n)$.

$\mathcal{B}_4$: H contains more than $(\log n)^3$ pairs of *partner hyperedges*, i.e. two F-edges sharing exactly u vertices.

$\mathcal{B}_5$: H contains an isolated vertex.

We denote the union of these events by

$$\mathcal{B} = \mathcal{B}_1(\pi_+) \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4 \cup \mathcal{B}_5.$$

The next result guarantees that by the end of the critical window, we can rule out these bad events whp.

Lemma 2.1. *Let $H = H_F(n,\pi)$ with $\pi \le n^{1-r+o(1)}$, then we have $H \notin \mathcal{B}_1(\pi) \cup \mathcal{B}_2$ whp. Further, we have $H \notin \mathcal{B}$ whp for $\pi = \pi_+$.*

2.1 Riordan's Coupling

Riordan's coupling [10], which couples $G = H_u(n,p)$ and $H = H_F(n,\pi)$ for p and π as in Theorem 1.5, provides the basis for our process coupling. It proceeds by going through an arbitrary, but fixed, order of all potential F-edges, in each step revealing whether the currently considered F-edge is present in H, while not always revealing whether the corresponding copy of F is present in G. After having tested all F-edges of H in this manner, we have constructed H and complete G conditional on the information revealed about it during the testing procedure.

For the convenience of the reader, we reproduce Riordan's coupling algorithm here (details can be found in [10]): Fix an arbitrary order $h_1, \dots, h_M$ of all F-edges. Denote by A_j the event that the corresponding copy of F is in G. For each j from 1 to M:

- Calculate π_j, the conditional probability of A_j given all information revealed so far.
- If $\pi_j \geq \pi$, toss a coin which lands heads with probability π/π_j, independently of everything else. If the coin lands heads, then test whether A_j holds. Include the F-edge h_j in H if and only if the coin lands heads and A_j holds.
- If $\pi_j < \pi$, then toss a coin which lands heads with probability π, independently of everything else, and declare h_j to be present in H if and only if the coin lands heads. In any case, do not test whether A_j holds. If the coin lands heads for any j, we say that the coupling has failed.

The outlined algorithm generates the correct distributions of G and H. Furthermore, if the coupling fails, then $H \in \mathcal{B}_1(\pi) \cup \mathcal{B}_2$, which fortunately happens with probability $o(1)$ for $\pi \leq n^{1-r+o(1)}$ by Lemma 2.1.

2.2 Properties of Riordan's coupling

A natural way to turn $H = H_F(n, \pi_+)$ (such that H contains a perfect matching whp) into a process is the following: Add the F-edges that are present in H in a uniformly random order to obtain the beginning of the random F-graph process, whp slightly beyond the time T_H. Now, if it were possible to use Riordan's coupling to further introduce $G = H_u(n, p)$ and add its u-edges in a uniformly random order such that corresponding copies of F appear in exactly the same order as the F-edges in the F-graph process, and moreover no other copies of F were present in G, we would be in good shape. However, there are two problems with this naive approach: As Riordan's static coupling is one-sided, G may contain *extra copies* of F, i.e. copies of F that are not present as F-edges in H. And secondly, overlapping copies of F do not appear in a uniformly random order in the random u-graph process.

We address the challenge of extra copies in the present section. One can show that the last vertices that are covered by an F-edge in the derived random F-graph process, are whp low-degree vertices of $H = H(n, \pi_+)$. If those vertices were covered by extra cliques in the coupled $G = H_u(n, p_+)$, extra cliques could have a serious impact on the comparability of the hitting times in the F-graph and the u-graph processes. Fortunately, this is not the case:

Lemma 2.2. *Let $G = H_u(n, p_+)$ and $H = H_F(n, \pi_+)$ be coupled via Riordan's coupling. Whp, no low-degree vertex in H is incident with an extra copy of F in G.*

The high-level proof idea of Lemma 2.2 is analogous to the proof of Lemma 3.1 in [4]: First, we bound the probability that an F-edge that is absent in H becomes an extra copy of F in G. To be precise, we show that for a fixed F-graph $H_0 \notin \mathcal{B}_1(\pi_+) \cup \mathcal{B}_2$ and j such that $h_j \notin H_0$, we have

$$\mathbb{P}(A_j \mid H = H_0) \leq \frac{\pi_j^* - \pi_+}{1 - \pi_+}, \qquad \text{where } \pi_j^* = \mathbb{P}(A_j \mid \bigcap_{i: h_i \in H_0} A_i).$$

Lemma 2.2 then follows from a first moment bound on the number of low-degree vertices in H which are also incident to an extra copy in G. The main contribution to the first moment is bounded in the following lemma (for a single vertex).

Lemma 2.3. *For $H_0 \notin \mathcal{B}_1(\pi_+) \cup \mathcal{B}_2$ and any low-degree vertex v of H_0,*

$$\sum_{\substack{j:h_j \notin H_0,\\ v \in h_j}} (\pi_j^* - p_+^s) \leq n^{-\frac{1}{s}+o(1)}. \tag{1}$$

It is in the proof of Lemma 2.3 that we deviate from [4]: Indeed, the present proof of (1) gives a viable alternative to the technical optimization problems in the proofs of Lemma 3.8 and Lemma A.2 in [4], which do not extend beyond the complete u-graphs considered therein.

To sketch the proof, let G_0 be the u-graph obtained from H_0 by replacing every F-edge by the corresponding copy of F, merging any duplicate u-edges. Let $E(G_0)$ denote the u-edges in G_0. Similarly, let E_j denote the u-edges in the copy of F corresponding to h_j. Observe that $\pi_j^* - p_+^s \leq p_+^{|E_j \setminus E(G_0)|}$ and note that we may constrain ourselves to the case $0 < |E_j \setminus E(G_0)| < s$ since F is nice. In this case, consider the u-graph $S = (h_j, E_j \cap E(G_0))$. Writing z for the number of components of S, we obtain from F being strictly 1-balanced that

$$|E_j \setminus E(G_0)| \geq s - d_1(F)(r - z) + \frac{1}{r-1}.$$

Using a combinatorial exploration argument, we then upper-bound the number of potential h_j that contain v and give rise to a u-graph S with z components by $n^{z-1+o(1)}$. Hence, we obtain for (1) that

$$\sum_{\substack{j:h_j \notin H_0,\\ v \in h_j}} (\pi_j^* - p_+^s) \leq \sum_{z=1}^{r-1} n^{z-1+o(1)} p_+^{s-d_1(F)(r-z)+\frac{1}{r-1}} \leq n^{-\frac{1}{s}+o(1)}.$$

2.3 The process coupling

We turn back to the dynamic part of the coupling that we mentioned in Section 2.2, that is, the coupling of the random u-graph process $(G_t)_{t=0}^{N_u}$ and the random F-graph process $(H_t)_{t=0}^{M}$: Starting with the static coupling, we establish a process coupling on a joint timeline under which we have $T'_H = T'_G$ whp, for the hitting time T'_H at which the last isolated vertex in the random F-graph process disappears, and the hitting time T'_G at which the last vertex not contained in a copy of F in the random u-graph process disappears.

In the corresponding part in [4], the random u-graph process $(G_t)_{t=0}^{N_u}$ is coupled with a standard random (simple, unlabeled, r-uniform) hypergraph process $(E_t)_{t=0}^{N_r}$ — since the introduction of F-graphs is not required for the special case $F = K_r$. This coupling was established in three parts, which in particular resolve the second challenge discussed in Section 2.2. In the first step (Section 4, respectively 8.3, in [4]), a suitable coupling of the processes $(G_t)_{t=0}^{N_u}$ and $(E_t)_{t=0}^{N_r}$ on a joint timeline was derived. The crux of this part is to find a way to deal with problematic hyperedges $\mathcal{F}$, which gain a so-called 'partner hyperedge' before the end of the critical window. The second (Propositions 5.1) and third (Proposition 6.1) steps[1]

[1]These steps are only relevant for $u = 2$, both in [4] and for Theorem 1.7.

in [4] establish a coupling that essentially whp embeds a copy of $(E_t)_{t\geq 0}$ into $(E_t)_{t\geq 0}$ itself, while avoiding the problematic hyperedges $\mathcal{F}$.

We adapt the proof from Section 4 (respectively 8.3) in [4] to establish the first step. This is unavoidable due to the extension from complete graphs K_r to more general u-graphs F. Denote by $\mathrm{c_F}(G_{T_G})$ the F-graph obtained from G_{T_G} by using the copies of F in G_{T_G} as labels for the F-edges. Recall that two F-edges are *partners* if they share exactly u vertices.

Proposition 2.4. *We may couple the random u-graph process $(G_t)_{t=0}^{N_u}$ and the random F-graph process $(H_t)_{t=0}^{M}$ so that the following holds. Let $\mathcal{E} = H_{T_H} \setminus \mathrm{c_F}(G_{T_G})$ and let $\mathcal{F}$ be the set of F-edges in H_{T_H} that have a partner F-edge in $H_{T_H+\lfloor g(n)n\rfloor} \setminus H_{T_H}$. Then we have $\mathcal{E} \subset \mathcal{F}$ whp.*

The set $\mathcal{F}$ is empty whp, unless $u = 2$, which we assume in the following. Now, instead of adapting the second and third step as well, we make the following observation, which allows to re-use the coupling in [4] directly. Let

$$T_E = \min\{t : E_t \text{ has no isolated vertices}\}$$

be the hitting time of having no isolated vertices left in the standard random hypergraph process $(E_t)_{t=0}^{N_r}$. Further, let $\tilde{H}_F$ be the multi-hypergraph obtained from H_F by replacing each F-edge by a hyperedge with the same vertex set (i.e. forgetting the labels).

Proposition 2.5. *We may couple the random F-graph process $(H_t)_{t=0}^{M}$ and the random hypergraph process $(E_t)_{t=0}^{N_r}$ such that whp both of the following two properties hold:*

(a) The hitting times at which the last isolated vertex disappears agree: $T_H = T_E$.

(b) The unlabeled processes agree until time $T_H + \lfloor g(n)n\rfloor$, i.e.

$$(\tilde{H}_t)_{t=0}^{T_H+\lfloor g(n)n\rfloor} = (E_t)_{t=0}^{T_H+\lfloor g(n)n\rfloor}.$$

For both covers and factors only the existence of an F-edge on a given vertex set is relevant, but not its actual placement. Thus, for our purposes it indeed suffices to consider the process $(\tilde{H}_t)_t$. Since Proposition 2.5, which holds for *all* u-graphs F, yields a whp equivalence with the process $(E_t)_t$, which was discussed in [4], this argument saves us from the tedious task of adapting the proofs for these steps.

Combining the three steps yields the following coupling.

Proposition 2.6. *There exists a coupling of $(H_t)_t$ and $\mathcal{F}$ to another instance $H'_{T_{H'}}$ of the stopped random F-graph process so that, whp, $H_{T_H} \setminus \mathcal{F} \supseteq H'_{T_{H'}}$.*

Proposition 2.4 and 2.6 yield a chain of couplings that establish Theorem 1.7:

$$H'_{T'_H} \quad \overset{\text{whp}}{\subseteq} \quad H_{T_H} \setminus \mathcal{F} \quad \overset{\text{whp}}{\subseteq} \quad H_{T_H} \setminus \mathcal{E} \quad \subseteq \quad \mathrm{c_F}(G_{T_G}).$$

References

[1] Béla Bollobás and Andrew Thomason. Random graphs of small order. In *Random graphs '83 (Poznań, 1983)*, volume 118 of *North-Holland Math. Stud.*, pages 47–97. North-Holland, Amsterdam, 1985.

[2] Pál Erdős and Alfréd Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar.*, 17:359–368, 1966.

[3] Annika Heckel. Random triangles in random graphs. *Random Structures & Algorithms*, 59(4):616–621, 2021.

[4] Annika Heckel, Marc Kaufmann, Noela Müller, and Matija Pasch. The hitting time of clique factors. *Random Structures & Algorithms*, 65(2):275–312, 2024.

[5] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[6] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[7] Jeff Kahn. Hitting times for Shamir's problem. *Trans. Amer. Math. Soc.*, 375(1):627–668, 2022.

[8] Jeff Kahn. Asymptotics for Shamir's problem. *Advances in Mathematics*, 422:109019, 2023.

[9] Julius Petersen. Die Theorie der regulären graphs. *Acta Math.*, 15:193–220, 1891.

[10] Oliver Riordan. Random cliques in random graphs and sharp thresholds for F-factors. *Random Structures & Algorithms*, 61(4):619–637, 2022.

[11] William T Tutte. The factorization of linear graphs. *Journal of the London Mathematical Society*, 1(2):107–111, 1947.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A limit theorem for Shannon capacity from lattice packings

(Extended abstract)

Pjotr Buys* Sven Polak† Jeroen Zuiddam*

Abstract

We develop a lattice approach to the Shannon capacity problem for fraction graphs and circle graphs. Bohman (2003) proved that, in the limit $p \to \infty$, the Shannon capacity of cycle graphs $\Theta(C_p)$ converges to the fractional clique covering number: $\lim_{p\to\infty} p/2 - \Theta(C_p) = 0$. We strengthen this result by proving that the same is true for all fraction graphs: $\lim_{p/q\to\infty} p/q - \Theta(E_{p/q}) = 0$. Here the fraction graph $E_{p/q}$ is the graph with vertex set $\mathbb{Z}/p\mathbb{Z}$ in which two distinct vertices are adjacent if and only if their distance mod p is strictly less than q.

As a further strengthening, we show that we can obtain the limit via a group-theoretic/lattice approach. In particular, we construct an explicit family of independent sets in fraction graphs leading to the limit as subgroups or lattices. These constructions extend and recover, in a structured and unified manner, many previously known lower bounds in cycle graphs and fraction graphs. The lattice approach is slightly more general than the previously studied linear Shannon capacity, and in particular circumvents known barriers.

1 Introduction

Determining the Shannon capacity of graphs, introduced by Shannon in 1956 to model zero-error communication over noisy channels [Sha56], is a long-standing open problem in discrete mathematics, information theory and combinatorial optimization [KO98, Alo02, Sch03]. In graph-theoretic terms, this parameter measures the rate of growth of the largest independent set in strong powers of a graph.

The Shannon capacity has been studied from many angles, which led to a variety of upper bound methods (e.g., the Lovász theta function [Lov79] and the (fractional) Haemers bound [Hae79, BC19]), lower bound constructions—which have been mostly ad hoc—(e.g., [BMR+71, Boh05, BH03, PS19, RPBN+24]), and structural results [Alo98, AL06, Zui19, Vra21, WZ23].

*University of Amsterdam, The Netherlands. E-mail: pjotr.buys@gmail.com, j.zuiddam@uva.nl. Supported by NWO 639.032.614 and VI.Veni.212.284.

†Tilburg University, The Netherlands. E-mail: s.c.polak@tilburguniversity.edu.

Determining the Shannon capacity is notoriously hard; for all odd cycles of length at least seven, the Shannon capacity is not known, for instance.

Inspired by recent "orbit" constructions for (the best-known) lower bounds on the Shannon capacity of small odd cycles [PS19, dBBZ24], we develop a "group-theoretic approach" to Shannon capacity (which we will see becomes a lattice approach for the class of graphs that we focus on). This approach builds on the idea of constructing independent sets in Cayley graphs that are subgroups. Indeed, Shannon's famous independent set in the square of the five-cycle has such a form: $\{t \cdot (1, 2) : t \in \mathbb{Z}_5\} \subseteq \mathbb{Z}_5^2$ [Sha56]. Where can we find such constructions? Can they give good lower bounds on Shannon capacity? We make progress on both questions.

We focus on a family of Cayley graphs that naturally generalizes odd cycle graphs, namely Cayley graphs for groups $\mathbb{Z}_p$ and generating sets $\{\pm 1, \ldots, \pm(q-1)\}$. Such graphs, which we refer to as fraction graphs, have indeed played a central role in aforementioned orbit constructions for small odd cycles, in particular the seven cycle and fifteen cycle. It will also be natural to work in the infinite version of these graphs, where the group is the circle group, which we refer to as circle graphs.

Subgroup independent sets in fraction graphs and circle graphs correspond, as we will see, precisely to lattices with good size and distance properties, giving rise to a *lattice approach* to Shannon capacity, reminiscent of optimal sphere packing constructions [Via17, CKM$^+$17, CJMS23]. The Shannon capacity problem on fraction graphs can indeed be rephrased as the problem of packing cubes with specified size in a high-dimensional unit-length torus [BMR$^+$71, RS97], and our construction more specifically gives lattice packings in this setting.

Central directions in this work are: (1) methods for proving that a given lattice is good for the lattice approach, (2) a construction of a family of good lattices and as a result many new lower bounds on the Shannon capacity, and (3) a proof that these lattice constructions are optimal in the limit of large fraction graphs.

We summarize our main results here and discuss them in more detail in the rest of the text:

- **Lattice approach.** We develop a lattice approach to the Shannon capacity problem for fraction graphs and circle graphs. Our approach, while similar in spirit to the linear Shannon capacity of Guruswami and Riazanov [GR21], is more general (by identifying equivalent fraction graphs), and in particular avoids related barriers [CFG$^+$93, GR21].

- **New Shannon capacity lower bounds via the lattice approach.** We construct an explicit family of lattices suitable for the lattice approach and thus obtain new lower bounds for the Shannon capacity of fraction graphs. These constructions extend and recover, in a structured and unified manner, previous lower bounds of Polak and Schrijver [Pol19, Theorem 9.3.3] and Baumert, McEliece, Rodemich, Rumsey, Stanley and Taylor [BMR$^+$71].

- **The Bohman limit via lattices.** Using our Shannon capacity lower bounds, we extend the Bohman limit theorem [Boh03, Boh05] from odd cycle graphs to all fraction graphs. In particular, this proves that the lattice approach is strong enough to reach the Shannon capacity in the limit of large cycle graphs and fraction graphs.

2 Shannon capacity, odd cycles, fraction graphs

The Shannon capacity of a graph G is defined as $\Theta(G) = \sup_n \alpha(G^{\boxtimes n})^{1/n}$, where α denotes the independence number and $\boxtimes$ denotes the strong product of graphs.[1] The supremum may equivalently be replaced by a limit (by Fekete's lemma).

For every $p \in \mathbb{N}$ let C_p denote the cycle graph with p vertices. For even p, it is not hard to see that $\Theta(C_p) = p/2$. Shannon [Sha56] determined $\Theta(G)$ for all graphs G with at most six vertices, except for the five cycle C_5, for which he famously proved $\alpha(C_5^{\boxtimes 2}) \geq 5$ (and thus $\Theta(C_5) \geq \sqrt{5}$) using the independent set $\{t \cdot (1, 2) : t \in \mathbb{Z}_5\}$ and $\Theta(C_5) \leq 5/2$ via an upper bound method that we now refer to as the fractional clique covering number (and which we denote by $\overline{\chi_f}$). Lovász [Lov79] in breakthrough work introduced the Lovász theta function (denoted by ϑ), proved that it upper bounds the Shannon capacity and used this to determine $\Theta(C_5) = \sqrt{5}$. For larger odd cycles, the Shannon capacity is not known. In all cases the Lovász theta function gives the best-known upper bound. For instance, for the seven cycle we know $3.2578 \approx 367^{1/5} \leq \Theta(C_7) \leq \vartheta(C_7) = (7\cos(\pi/7))/(1 + \cos(\pi/7)) \approx 3.3177$ [PS19].

For every $p, q \in \mathbb{N}$ we let $E_{p/q}$ be the graph with vertex set $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ in which two distinct vertices are adjacent if and only if their distance mod p is strictly less than q. We refer to these graphs as the *fraction graphs*. Work on these goes back to Vince [Vin88], and later [HN04, Zhu06, BM13, Pol19, PS19]. For example, $E_{p/2} = C_p$ is the cycle graph on p vertices, and $E_{p/1} = \overline{K_p}$ is the graph with no edges on p vertices. For any graph G, H we write $G \leq H$ if there is a cohomomorphism $G \to H$ (a map $V(G) \to V(H)$ that maps non-edges to non-edges). The fraction graphs are totally ordered: $E_{a/b} \leq E_{c/d}$ if and only if $a/b \leq c/d$. In particular, equivalent fractions give equivalent fractions graphs in the cohomomorphism order. It is known that $\overline{\chi_f}(E_{p/q}) = p/q$ for all p, q [HN04] and in particular $\overline{\chi_f}(C_p) = p/2$.

Several upper bounds on the Shannon capacity are known, including the aforementioned (fractional) clique covering number $\overline{\chi_f}$, the Lovász theta function ϑ [Lov79] and the (fractional) Haemers bound [Hae79, BC19]. Theory has been developed about such upper bounds and their structure [WZ23, dBBZ24]. In particular, [dBBZ24] developed a graph limit approach to study Shannon capacity, in which fraction graphs played a central role (to construct converging sequences of graphs). It was also observed there that the best-known constructions for lower bounds on Shannon capacity of small odd cycles can all be obtained using "orbit" constructions in fraction graphs close to those odd cycle graphs.

3 Results: Bohman limit for fraction graphs via lattices

Bohman [Boh03, Boh05] proved that, in the limit $p \to \infty$, the Shannon capacity of cycle graphs $\Theta(C_p)$ converges to the fractional clique covering upper bound:

Theorem 3.1. $\lim_{p\to\infty} p/2 - \Theta(C_p) = 0$.

This was done by constructing a sequence of large independent sets in powers of odd cycles, in the first work by a direct construction based on an earlier construction of Hales, and in the second work using an expansion process of [BMR$^+$71]. We strengthen Theorem 3.1 by proving that the same is true for all fraction graphs:

Theorem 3.2. $\lim_{p/q\to\infty} p/q - \Theta(E_{p/q}) = 0$.

[1]The strong product $G \boxtimes H$ is the graph with vertex set the cartesian product $V(G) \times V(H)$ and edge set given by $E(G \boxtimes H) = \{\{(a, x), (b, y)\} : (a = b \vee \{a, b\} \in E(G)) \wedge (x = y \vee \{x, y\} \in E(H))\}$.

Indeed, Theorem 3.2 implies Theorem 3.1, since $E_{p/2} = C_p$.

Group-theoretic approach to Shannon capacity

We strengthen Theorem 3.2 further by showing that we can obtain the limit via a group-theoretic/lattice approach. In other words, the independent sets leading to the limit we will construct as subgroups. Our approach is slightly more general than the previously studied linear Shannon capacity [CFG$^+$93, GR21], and in particular we circumvent their barrier.

For any Cayley graph H, let $\alpha_{\mathrm{grp}}(H)$ be the size of the largest independent set S of H that is a subgroup, and let $\Theta_{\mathrm{grp}}(H) = \sup_n \alpha_{\mathrm{grp}}(H^{\boxtimes n})^{1/n}$. We will refer to this quantity as the *subgroup Shannon capacity*. We prove the following:

Theorem 3.3. *For every $\varepsilon > 0$, for all $p/q \in \mathbb{Q}$ large enough, there are $a, b \in \mathbb{N}$ such that $a/b = p/q$ and $p/q - \Theta_{\mathrm{grp}}(E_{a/b}) < \varepsilon$.*

We will explain in the rest of the text how we construct the lower bounds on Θ_{grp}. In Figure 1 we plot the lower bounds we obtain from our construction. Our construction encompasses (and extends) several constructions from the literature, in particular from [BMR$^+$71, Pol19, dD22]. We note that those constructions (even combined with "expansion methods" of [BMR$^+$71]) are (to our knowledge) not sufficient to obtain our Theorem 3.3.

Circle graphs

Instead of working with the fraction graphs, our results and constructions can naturally be thought of as taking place in infinite graphs on the circle (of which the fraction graphs are naturally induced subgraphs), as follows. Let $C \subseteq \mathbb{R}^2$ denote a circle with unit circumference. Let $r \in \mathbb{R}_{\geq 2}$. Let E_r° be the infinite graph with vertex set C for which two vertices are adjacent if and only if they have distance strictly less than $1/r$ on C. We refer to these graphs as the (open) *circle graphs*. Note that E_r° is a Cayley graph on the circle group. We prove:

Theorem 3.4. $\lim_{r\to\infty} r - \Theta_{\mathrm{grp}}(E_r^{\circ}) = 0$.

In fact, it can be shown that Theorem 3.4 is equivalent to Theorem 3.3 using that fraction graphs naturally are induced subgraphs of circle graphs.

4 Lattice approach and our construction

We now explain the lattice approach and construction that lead to the above theorems. Given an invertible matrix $A \in \mathbb{R}^{n\times n}$, we let $\mathcal{L}(A)$ denote the lattice generated by the columns of A, that is, $\mathcal{L}(A) = \{Ax : x \in \mathbb{Z}^n\}$. For $v \in \mathbb{Z}^n$ we let $\|v\|_\infty = \max_i |v_i|$. Given a lattice Λ, we define the minimum distance $\lambda_\infty(\Lambda) = \min\{\|v\|_\infty : v \in \Lambda \setminus \{0\}\}$.

Lemma 4.1. *Let $A, B \in \mathbb{Z}^{n\times n}$ and $p, q \in \mathbb{Z}_{\geq 1}$. Suppose $AB = p \cdot I_n$ and that the lattice $\Lambda = \mathcal{L}(A)$ satisfies $\lambda_\infty(\Lambda) = q$. Then $\{v \pmod p : v \in \Lambda\}$ is an independent set of size $|\det(B)|$ in $E_{p/q}^{\boxtimes n}$ and thus $\alpha_{\mathrm{grp}}(E_{p/q}^{\boxtimes n}) \geq |\det(B)|$.*

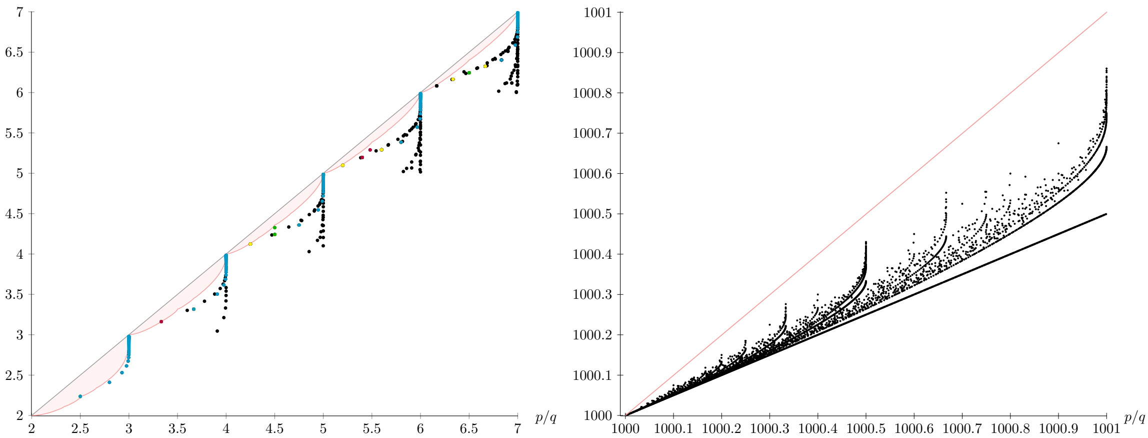


Figure 1: Left: Lower bounds on the Shannon capacity $\Theta(E_{p/q})$ for $p/q \in [2,7]$ obtained via the lattice approach. The red graph is the Lovász theta upper bound; the dots our lower bounds. Previously known points of [BMR+71, Pol19, dD22] are indicated with purple/green, blue and yellow, respectively. Right: Similarly, but for the interval $[1000, 1001]$, showing behaviour for large p/q (again, a subset of these points were known from aforementioned work). Note that here the Lovász theta function $\vartheta(E_{p/q})$ is already very close to p/q.

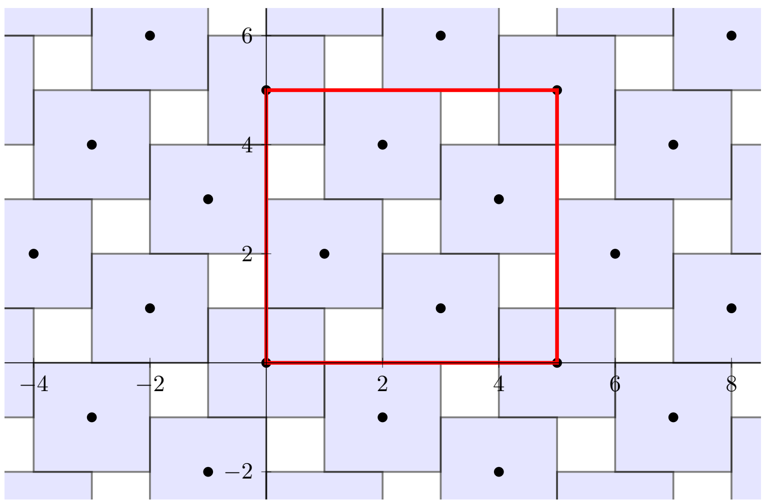


Figure 2: A portion of the lattice $\Lambda = \mathcal{L}(A)$ for $A = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$ is drawn. Letting $B = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$, we have $AB = 5I_2$, $|\det(B)| = 5$ and $\lambda_\infty(\Lambda) = 2$. It follows from Lemma 4.1 that $\alpha_{\mathrm{grp}}(E_{5/2}) \geq 2$. By taking the lattice points inside the red square, one obtains a representation of the independent set $\{v \pmod 5) : v \in \Lambda\}$. The fact that $\lambda_\infty(\Lambda) \geq 2$ can be proven using Lemma 4.2 and is seen visually by the fact that the squares of side-length 2 centered at lattice points do not overlap.

In general it is a difficult problem to determine $\lambda_\infty(\Lambda)$; see [vEB81]. In our proof we use the concept of a P_0-matrix: a real-valued matrix is P_0 if the determinants of all its principle minors are nonnegative; see [JST20]. We prove the following.

Lemma 4.2. *Suppose $X \in \mathbb{R}^{n\times n}$ is a P_0-matrix. Then for every $q > 0$ the matrix $A = X + q\cdot I_n$ is invertible and satisfies $\lambda_\infty(\mathcal{L}(A)) \geq q$.*

To prove Theorem 3.3 we construct a parameterized family of lattices depending on parameters n, k, b, s, and r. We define the following additional quantities.

Definition 4.3. Given $n, k, b \in \mathbb{Z}_{\geq 1}$ and $s, r \in \mathbb{Z}_{\geq 0}$ we define the integers[2] a, p, q:

$$a = k\cdot b^n + s\cdot b + r, \quad p = \frac{s^n r + k a^n}{r + kb^n}, \quad \text{and} \quad q = \frac{s^{n-1} r + kba^{n-1}}{r + kb^n}.$$

and the matrices X, Y, A, B by

$$X_{ij} = \begin{cases} ra^{j-i-1}s^{n-(j-i+1)} & \text{if } i < j, \\ 0 & \text{if } i = j, \\ -ka^{n-(i-j+1)}s^{i-j-1} & \text{if } i > j \end{cases} \quad \text{and} \quad Y_{ij} = \begin{cases} -rb^{j-i-1} & \text{if } i < j, \\ 0 & \text{if } i = j, \\ kb^{n-(i-j+1)} & \text{if } i > j, \end{cases}$$

and $A = qI_n + X$, and $B = \frac{a-r}{b} I_n + Y$.

We consider the lattices $\mathcal{L}(A)$ for various choices of parameters. For $(n, k, b, s, r) = (2, 1, 1, 1, 1)$, for example, one obtains the lattice and independent set depicted in Figure 2. The most challenging part of the proof of Theorem 3.3 is that, under mild conditions on the parameters, the matrix X is a P_0 matrix. We use this to prove the following.

Theorem 4.4. *Given $n, k, b \in \mathbb{Z}_{\geq 1}$ and $s, r \in \mathbb{Z}_{\geq 0}$ with $r \leq b$, we have $\lambda_\infty(A) \geq q$, $AB = pI_n$, and $\det(B) = p$.*

This theorem allows us to apply Lemma 4.1 to the matrices A, B to obtain a parameterized family of independent sets, that are also subgroups, in a range of graphs $E_{p/q}$. The final step in establishing Theorem 3.3 involves proving that these independent sets are sufficiently large and lie sufficiently dense. This is formalized in the following technical lemma.

Lemma 4.5. *Given $\varepsilon > 0$ there is a B_ε such that for all $B_\varepsilon < x$ there are $n, k, b \in \mathbb{Z}_{\geq 1}$ and $s, r \in \mathbb{Z}_{\geq 0}$ with $r \leq b$ such that $p/q < x$ and $x - \varepsilon < p^{1/n}$.*

5 Open problems

We discuss some natural open problems. The first is whether the lattice approach is powerful enough to determine the Shannon capacity of all fractions graphs, in the following sense.

Problem 5.1. Is the Shannon capacity $\Theta(E_{p/q})$ equal to the supremum of the subgroup Shannon capacity $\Theta_{\mathrm{grp}}(E_{a/b})$ over all $a, b \in \mathbb{N}$ such that $a/b = p/q$?

Phrased in terms of the circle graphs, Problem 5.1 asks if the Shannon capacity and subgroup Shannon capacity coincide on circle graphs, that is, if $\Theta(E_r^\circ) = \Theta_{\mathrm{grp}}(E_r^\circ)$ for every $r \in \mathbb{Q}_{\geq 2}$. A related problem is whether subgroup Shannon capacity is strictly increasing:

Problem 5.2. Is $p/q \mapsto \sup_{a,b\in\mathbb{N}, a/b=p/q} \Theta_{\mathrm{grp}}(E_{a/b})$ strictly increasing? Is this the case at $p/q = 2$?

Indeed, [BH03] (strengthened by [Zhu25]) proved that $p/q \mapsto \Theta(E_{p/q})$ is strictly increasing at $p/q = 2$. We do not know whether this is true for $p/q \mapsto \sup_{a,b\in\mathbb{N}, a/b=p/q} \Theta_{\mathrm{grp}}(E_{a/b})$. In fact, given the constructions we have currently found, it seems possible that $p/q \mapsto \sup_{a,b\in\mathbb{N}, a/b=p/q} \Theta_{\mathrm{grp}}(E_{a/b})$ is constant for $p/q \in [2, 2+\varepsilon]$ for some $\varepsilon > 0$.

[2] Even if it is not directly obvious from the definition, it is not hard to show that p and q are in fact integers.

References

[AL06] Noga Alon and Eyal Lubetzky. The Shannon capacity of a graph and the independence numbers of its powers. *IEEE Trans. Inform. Theory*, 52(5):2172–2176, 2006. doi:10.1109/TIT.2006.872856.

[Alo98] Noga Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998. doi:10.1007/PL00009824.

[Alo02] Noga Alon. Graph powers. In *Contemporary combinatorics*, volume 10 of *Bolyai Soc. Math. Stud.*, pages 11–28. János Bolyai Math. Soc., Budapest, 2002. URL: https://web.math.princeton.edu/~nalon/PDFS/cap2.pdf.

[BC19] Boris Bukh and Christopher Cox. On a fractional version of Haemers' bound. *IEEE Trans. Inform. Theory*, 65(6):3340–3348, 2019. arXiv:1802.00476, doi:10.1109/TIT.2018.2889108.

[BH03] Tom Bohman and Ron Holzman. A nontrivial lower bound on the Shannon capacities of the complements of odd cycles. *IEEE Trans. Inform. Theory*, 49(3):721–722, 2003. doi:10.1109/TIT.2002.808128.

[BM13] Sevak H. Badalyan and Stepan E. Markosyan. On the independence number of the strong product of cycle-powers. *Discrete Math.*, 313(1):105–110, 2013. doi:10.1016/j.disc.2012.09.008.

[BMR+71] Leonard D. Baumert, Robert J. McEliece, Eugene Rodemich, Howard C. Rumsey, Jr., Richard Stanley, and Herbert Taylor. A combinatorial packing problem. In *Computers in algebra and number theory (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1970)*, volume IV of *SIAM-AMS Proc.*, pages 97–108. Amer. Math. Soc., 1971. URL: https://math.mit.edu/~rstan/pubs/pubfiles/17.pdf.

[Boh03] Tom Bohman. A limit theorem for the Shannon capacities of odd cycles. I. *Proc. Amer. Math. Soc.*, 131(11):3559–3569, 2003. doi:10.1090/S0002-9939-03-06495-5.

[Boh05] Tom Bohman. A limit theorem for the Shannon capacities of odd cycles. II. *Proc. Amer. Math. Soc.*, 133(2):537–543, 2005. doi:10.1090/S0002-9939-04-07470-2.

[CFG+93] A. Robert Calderbank, Peter Frankl, Ronald L. Graham, Wen-Ching W. Li, and Lawrence A. Shepp. The Sperner capacity of linear and nonlinear codes for the cyclic triangle. *J. Algebraic Combin.*, 2(1):31–48, 1993. doi:10.1023/A:1022424630332.

[CJMS23] Marcelo Campos, Matthew Jenssen, Marcus Michelen, and Julian Sahasrabudhe. A new lower bound for sphere packing, 2023. arXiv:2312.10026.

[CKM+17] Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska. The sphere packing problem in dimension 24. *Ann. of Math. (2)*, 185(3):1017–1033, 2017. doi:10.4007/annals.2017.185.3.8.

[dBBZ24] David de Boer, Pjotr Buys, and Jeroen Zuiddam. The asymptotic spectrum distance, graph limits, and the Shannon capacity, 2024. arXiv:2404.16763.

[dD22] Timo den Dulk. The Shannon capacity on $C(n, k)$, 2022. Thesis, Delft University of Technology. URL: http://resolver.tudelft.nl/uuid:54779ec5-79c5-4481-a0fe-46c3aaaa4aab.

[GR21] Venkatesan Guruswami and Andrii Riazanov. Linear Shannon capacity of Cayley graphs. In *2021 IEEE International Symposium on Information Theory (ISIT)*, page 988–992. IEEE Press, 2021. arXiv:2009.05685, doi:10.1109/ISIT45174.2021.9517713.

[Hae79] Willem Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(2):231–232, 1979. doi:10.1109/TIT.1979.1056027.

[HN04] Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004. doi:10.1093/acprof:oso/9780198528173.001.0001.

[JST20] Charles R. Johnson, Ronald L. Smith, and Michael J. Tsatsomeros. *Matrix positivity*, volume 221 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2020. doi:10.1017/9781108778619.

[KO98] János Körner and Alon Orlitsky. Zero-error information theory. *IEEE Trans. Inform. Theory*, 44(6):2207–2229, 1998. Information theory: 1948–1998. doi:10.1109/18.720537.

[Lov79] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985.

[Pol19] Sven C. Polak. *New methods in coding theory: Error-correcting codes and the Shannon capacity.* PhD thesis, University of Amsterdam, 2019. Chapter 9 contains unpublished joint work with Lex Schrijver. URL: https://hdl.handle.net/11245.1/393eb25b-c1ba-4c6f-89e8-c439a97358b6.

[PS19] Sven C. Polak and Alexander Schrijver. New lower bound on the Shannon capacity of C_7 from circular graphs. *Inform. Process. Lett.*, 143:37–40, 2019. doi:10.1016/j.ipl.2018.11.006.

[RPBN+24] Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi. Mathematical discoveries from program search with large language models. *Nature*, 625(7995):468–475, January 2024. doi:10.1038/s41586-023-06924-6.

[RS97] Andrew V. Reztsov and Ian H. Sloan. On 2D packings of cubes in the torus. *Proc. Amer. Math. Soc.*, 125(1):17–26, 1997. doi:10.1090/S0002-9939-97-03930-0.

[Sch03] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, volume 24B of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.

[Sha56] Claude E. Shannon. The zero error capacity of a noisy channel. *Institute of Radio Engineers Transactions on Information Theory*, IT-2,:8–19, 1956. doi:10.1109/tit.1956.1056798.

[vEB81] Peter van Emde Boas. Another NP-complete partition problem and the complexity of computing short vectors in a lattice. Technical Report 81-04, Mathematisch Instituut, University of Amsterdam, April 1981. URL: https://staff.fnwi.uva.nl/p.vanemdeboas/vectors/mi8104c.html.

[Via17] Maryna S. Viazovska. The sphere packing problem in dimension 8. *Ann. of Math. (2)*, 185(3):991–1015, 2017. doi:10.4007/annals.2017.185.3.7.

[Vin88] Andrew Vince. Star chromatic number. *J. Graph Theory*, 12(4):551–559, 1988. doi:10.1002/jgt.3190120411.

[Vra21] Péter Vrana. Probabilistic refinement of the asymptotic spectrum of graphs. *Combinatorica*, 41(6):873–904, 2021. doi:10.1007/s00493-020-4324-5.

[WZ23] Avi Wigderson and Jeroen Zuiddam. Asymptotic spectra: Theory, applications and extensions, 2023. URL: https://staff.fnwi.uva.nl/j.zuiddam/papers/convexity.pdf.

[Zhu06] Xuding Zhu. Recent developments in circular colouring of graphs. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 497–550. Springer, Berlin, 2006. doi:10.1007/3-540-33700-8_25.

[Zhu25] Daniel G. Zhu. An improved lower bound on the Shannon capacities of complements of odd cycles. *Proc. Amer. Math. Soc.*, 153(4):1751–1759, 2025. arXiv:2402.10025, doi:10.1090/proc/17091.

[Zui19] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity. *Combinatorica*, 39(5):1173–1184, 2019. doi:10.1007/s00493-019-3992-5.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Universality of asymptotic graph cohomomorphism

(Extended abstract)

Pjotr Buys* Anna Luchnikov* Jim Wittebol† Jeroen Zuiddam*

Abstract

The Shannon capacity of graphs, introduced by Shannon in 1956 to model zero-error communication, asks for determining the rate of growth of independent sets in strong powers of graphs. Despite much interest, which has led to a large collection of upper bound methods (e.g., Lovász theta function, projective rank, fractional Haemers bound) and lower bound constructions, much is still unknown about this parameter, for instance whether it is computable. Indeed, results of Alon and Lubetzky (2006) have ruled out several natural routes to such an algorithm.

Recent work has established a dual characterization of the Shannon capacity in terms of the asymptotic spectrum of graphs (a class of well-behaved graph parameters that includes the aforementioned upper bound methods). A core step in this duality theory is to shift focus from Shannon capacity itself to studying the asymptotic relations between graphs, that is, the asymptotic cohomomorphisms: given graphs G and H, is there a cohomomorphism (a map on vertex sets, mapping non-edges to non-edges) from the nth power of G to the $(n + o(n))$th strong power of H? Indeed, Shannon capacity essentially reduces to the case that G has no edges (i.e., is an independent set).

Towards understanding the structural intricacies of the Shannon capacity, we study the "combinatorial complexity" of asymptotic cohomomorphism. As our main result, we prove that the asymptotic cohomomorphism preorder is *universal* for all countable preorders. That is, we prove that any countable preorder can be order-embedded into the asymptotic cohomomorphism preorder (i.e. appears as a sub-order). Previously this was known for (nonasymptotic) cohomomorphism (going back to work of Hedrlín and Pultr), leaving open the possibility that asymptotic cohomomorphism has a simpler structure.

The main strategy of our proof is to construct an order-embedding from an order on certain sets of finite binary strings, that is known to be universal by a result of Hubička and Nešetřil (2005), to asymptotic cohomomorphism. The construction of this embedding relies on results of Vrana on the convex structure of the asymptotic spectrum of graphs, and a new result determining the value of the fractional Haemers bounds on a class of circulant graphs (fraction graphs), extending a result of Bukh and Cox (2018). These ingredients allow us to simulate a certain set of lines with pointwise order as graphs under asymptotic cohomomorphism, which we use to prove our result.

*University of Amsterdam, E-mail: `p.j.w.buys@uva.nl, a.luchnikov@uva.nl, j.zuiddam@uva.nl`
†Vrije Universiteit Brussel, E-mail: `jozef.johannus.wittebol@vub.be`

1 Introduction

Motivated by the Shannon capacity problem and recent approaches to it, we study in this paper a preorder on graphs called asymptotic cohomomorphism, which essentially measures if large powers of a graph allow a cohomomorphism[1] into slightly larger powers of another graph. The homomorphism order (and thus also the cohomomorphism order) has been long known to be be very rich in structure. Indeed it goes back to work of Hedrlín [Hed69, PT80] (see also [HN04a]) that any countable preorder can be order-embedded into it—a property called *universality*—which was followed by results proving even more structure later. Our main result is that asymptotic cohomomorphism, while a much "stronger" preorder than cohomomorphism (i.e. with more relations, which could a priori greatly simplify its structure), is still universal.

Posed by Shannon in 1956 [Sha56], the Shannon capacity problem asks to determine the rate of growth of the independence number of a graph under taking strong graph product powers. This problem is notoriously difficult and has defied many attempts at solving it, even for specific small graphs, and despite much effort. Indeed, a long line of work has led to many important methods and results, in the direction of upper bounds (Lovász theta function [Lov79], fractional Haemers bound [BC19], projective rank [MR16]), lower bounds (e.g., Polak–Schrijver [PS19], de Boer–Buys–Zuiddam [dBBZ24]), and structural results (e.g., Alon [AL06], Alon-Lubetzky [AL06], Zuiddam [Zui19], Vrana [Vra21], Schrijver [Sch23], Wigderson–Zuiddam [WZ23]).

Many questions about Shannon capacity are open. One central open problem is whether Shannon capacity is computable. Alon–Lubetzky [AL06] ruled out several natural approaches to constructing such an algorithm by showing that the "jump" behaviour of the independence number under powers can be very intricate. On the other hand, there are many examples, in several settings in mathematics and computer science, where parameters that are asymptotic or amortized allow a surprisingly simple description. The work in this paper is aimed at a better understanding of the "complexity" of Shannon capacity from a combinatorial point of view.

For this we focus on asymptotic cohomomorphism, an order on graphs that is tightly related to Shannon capacity (essentially replacing the question of embedding large independent sets in powers of a graphs, by embedding large powers of a graph into large powers of another graph) and that was recently introduced in the development of asymptotic spectrum duality [Zui19] (see also [WZ23]), and indeed plays a central role there. Asymptotic spectrum duality gives a dual characterization of Shannon capacity as a minimization problem over a class of functions called the asymptotic spectrum of graphs (which includes aforementioned Lovász theta function, fractional Haemers bound and projective rank, and also fractional clique covering number), but moreover also characterizes asymptotic cohomomorphism.

We summarize here our main results, which we expand on in the rest of the text:

- Motivated by the study of Shannon capacity, we prove that asymptotic graph cohomomorphism is universal for all countable preorders. In other words, every countable preorder appears as a sub-order in it. (In fact, our construction also implies universality of the cohomomorphism order, thus giving a different proof for that.)

- Our proof of universality relies on embedding a preorder on sets of binary strings (that was proven to be universal by Hubička and Nešetřil [HN05]) into graphs. It further relies

[1] A cohomomorphism from G to H is a map $\phi : V(G) \to V(H)$ such that ϕ is a graph homomorphism $\overline{G} \to \overline{H}$. In other words, a cohomomorphism maps non-edges to non-edges.

on the asymptotic spectrum of graphs, its (convex) structure, and a construction of lines and points with special order properties.

- As an important ingredient for our proof, we prove that the fractional Haemers bound (over any field) on fraction graphs coincides with the fractional clique covering number. It is in particular the fact that two different spectral points coincide on fraction graphs that will allow us to use graphs to "simulate" lines with rational coefficients. As another consequence of independent interest, this implies that, for the Shannon capacity on fraction graphs, none of the known upper bound methods can beat the Lovász theta function.

2 Graph cohomomorphism and universality

We denote by $G \boxtimes H$ the strong product of graphs, by $G \sqcup H$ the disjoint union, by $\overline{G}$ the complement graph, by $G + H = \overline{\overline{G} \sqcup \overline{H}}$ the join, by K_n the complete graph on n vertices, and by $E_n = \overline{K}_n$ its complement.

For any graphs G, H, a cohomomorphism $\phi : G \to H$ is a map $\phi : V(G) \to V(H)$ that maps any two distinct nonadjacent vertices to distinct nonadjacent vertices. (In other words, the cohomomorphisms $G \to H$ correspond precisely to the homomorphisms $\overline{G} \to \overline{H}$.) We write $G \leq H$ if there is a graph cohomomorphism $\phi : G \to H$ and we call this the cohomomorhism preorder. The cohomomorphism preorder has been well-studied, and much is known about its rich structure. One such structural property is universality.

Given two preorders $(A, \leq_A)$ and $(B, \leq_B)$ we call a map $\phi : A \to B$ an *order-embedding* if for every $a_1, a_2 \in A$ we have $a_1 \leq_A a_2$ if and only if $\phi(a_1) \leq \phi(a_2)$. We call $(B, \leq_B)$ *countably universal* if for every countable preorder $(A, \leq_A)$ there exists an order-embedding $\phi : A \to B$.

It was established in [Hed69, PT80] that the cohomomorphism preorder is countably universal. This result was proved again in [HN04b] and [HN05]. In fact, in [FHLN17], it is shown that an even stronger "fractal" property holds: For any two graphs $G_1 < G_2$ that are not equivalent to E_1, E_2, any countable preorder can be embedded into the interval $\{H \in \mathcal{G} \mid G_1 < H < G_2.\}$.

3 Result: universality of asymptotic cohomomorphism

For any graph G, the Shannon capacity is defined as $\Theta(G) = \sup_{n \to \infty} \alpha(G^{\boxtimes n})^{1/n}$, where $\boxtimes$ denotes the strong product, α the independence number, and where the supremum may equivalently be replaced by a limit (by Fekete's lemma). Shannon capacity is closely related to a preorder on graphs called asymptotic cohomomorphism: We write $H \lesssim G$ if and only if there exists a function $f : \mathbb{N} \to \mathbb{N}$ satisfying $\lim_{n \to \infty} f(n)/n = 0$, such that for all n, we have $H^{\boxtimes n} \leq G^{\boxtimes (n+f(n))}$. Cohomomorphism implies asymptotic cohomomorphism, but not necessarily the other way around.

How complex is asymptotic cohomomorphism? Could it have a far simpler structure than cohomomorphism? Can it be determined by looking at only few properties of the graphs G and H? Our main result is that, in a specific sense, asymptotic cohomomorphism is as complex as cohomomorphism. Let $(\mathcal{G}, \lesssim)$ denote the set of finite simple graphs equipped with the asymptotic cohomomorphism preorder.

Theorem 3.1. $(\mathcal{G}, \lesssim)$ *is countably universal.*

The high-level idea of the proof of Theorem 3.1 is as follows. (1) The starting point is a preorder $(W, \leq_W)$ on binary strings and a preorder $(\mathcal{W}, \leq_{\mathcal{W}})$ on antichains of elements in W that is known to be countably universal. (2) We embed W in graphs through a construction of lines and points, assigning to $a \in W$ a graph G_a. To ensure this embedding preserves the preorder we need to use tools from asymptotic spectrum duality (in particular convexity) and special knowledge of some elements in the asymptotic spectrum (fractional Haemers bound). (3) We embed then $\mathcal{W}$ in graphs, assigning to $A \in \mathcal{W}$ a graph $\sum_{a\in A} G_a$. We note that our construction in fact also implies the previously known universality of graph cohomomorphism. In the rest of text we will explain these steps in more detail. We start with the necessary ingredients from asymptotic spectrum duality (Section 4) and a new result on the fractional Haemers bound that is of independent interest (Section 5). Then we discuss the construction for Theorem 3.1 (Section 6).

4 Asymptotic spectrum duality

Theorem 4.1 ([Zui19])**.** *Let G, H be graphs. Then $G \lesssim H$ if and only if for all $f \in \Delta(\mathcal{G})$ we have $f(G) \leq f(H)$.*

From Theorem 4.1 it can be shown that $\Theta(G) = \min_{F\in\Delta(\mathcal{G})} F(G)$. Some known elements in $\Delta(\mathcal{G})$ are the fraction clique covering number $\overline{\chi_f}$ [SU97, Proposition 3.2.2], the Lovász theta function ϑ [Lov79], and, for any field $\mathbb{F}$, the fractional Haemers bound $\mathcal{H}_f^{\mathbb{F}}$ [Hae81, Bla13, BC19]. The latter is defined as: $\mathcal{H}_f^{\mathbb{F}}(G) = \inf\{d/r \mid G \text{ has a } (d,r)\text{-subspace representation over } \mathbb{F}\}$, where a (d,r)-subspace representation of a graph G over a field $\mathbb{F}$ is a collection of subspaces $\{S_v \leq \mathbb{F}^d \mid v \in V(G)\}$ such that for each vertex $v \in V(G)$, $\dim S_v = r$ and $S_v \cap \sum_{w\not\sim v} S_w = \{0\}$. It is known that $\overline{\chi_f}$ is the pointwise maximum in $\Delta(\mathcal{G})$. Shannon capacity itself however is not in $\Delta(\mathcal{G})$. Moreover, many incomparability results are known among Lovász theta and the fractional Haemers bounds, and among the fractional Haemers bounds themselves [BC19]. Theorem 3.1 with Theorem 4.1 implies that $\Delta(\mathcal{G})$ is a rich enough class of functions that the evaluation functions $\widehat{G} : f \mapsto f(G)$ with pointwise order are countably universal.

The following two ingredients will play an important role in the proof of Theorem 3.1 as we will explain. The first is stated in [Vra21] and the second can be derived from the convexity results in [Vra21].

Lemma 4.2 ([Vra21])**.** *For any $f \in \Delta(\mathcal{G})$, and graphs G, H, $f(G + H) = \max(f(G), f(H))$.*

Lemma 4.3 (Interpolation)**.** *Let $f_0, f_1 \in \Delta(\mathcal{G})$. For every $0 \leq \lambda \leq 1$, there exists an $f \in \Delta(\mathcal{G})$ such that for every vertex-transitive graph G, we have $f(G) = f_0^{1-\lambda}(G) f_1^{\lambda}(G)$.*

5 Fractional Haemers bound of fraction graphs

For our proof of Theorem 3.1 we employ the so called fraction graphs $E_{p/q}$ defined for $p, q \in \mathbb{N}$. We discuss these here, and discuss a new result about the fractional Haemers bound on them.

Fraction graphs have previously been used to study variations of the chromatic number [Vin88]. They appear in literature also under the names of circular graphs [Pol20], cycle-powers, and (the complement of) rational complete graphs [HN04a] and circular complete graphs [Zhu06].

The graph $E_{p/q}$ has vertex set $\{0, 1, \ldots, p-1\}$, which we identify with $\mathbb{Z}_p$. Write $|x|_p$ for the absolute value of $x \pmod q$. Then, for any distinct $i, j \in \mathbb{Z}_p$, $i \sim j$ in $E_{p/q}$ if and only if

$|i-j|_p < q$. In several ways, these graphs model the preorder $(\mathbb{Q}_{\geq 2}, \leq)$ inside the preorder $(\mathcal{G}, \leq_{\mathcal{G}})$. Indeed, for $p/q, r/s \in \mathbb{Q}_{\geq 2}$, $E_{p/q} \leq E_{r/s}$ if and only if $p/q \leq r/s$ [HN04a]. In fact $E_{p/q} \lesssim E_{r/s}$ if and only if $p/q \leq r/s$, see also [dBBZ24]. In particular, $E_{p/q}$ and $E_{r/s}$ are (asymptotically) cohomomorphically equivalent if and only if $p/q = r/s$. If this is the case, then $\phi(E_{p/q}) = \phi(E_{r/s})$ for all $\phi \in \Delta(\mathcal{G})$.

Bukh and Cox [BC19] proved that for odd cycles the fractional Haemers bounds coincide with the fractional clique covering number: $\mathcal{H}_f^{\mathbb{F}}(C_{2n+1}) = \overline{\chi_f}(C_{2n+1}) = n + 1/2$. We extend this result to all fraction graphs:

Theorem 5.1. *For every $p/q \in \mathbb{Q}_{\geq 2}$, for every field $\mathbb{F}$, $\mathcal{H}_f^{\mathbb{F}}(E_{p/q}) = p/q$.*

We note that Lovász theta $\vartheta(E_{p/q})$ can be strictly smaller than p/q (e.g., $\vartheta(E_{5/2}) = \sqrt{5}$). It follows from Theorem 5.1 that the fractional Haemers bounds can in particular not give upper bounds on Shannon capacity of fraction graphs better than ϑ. It follows from Theorem 5.1 that there are distinct $f_0, f_1 \in \Delta(\mathcal{G})$ such that for all $p/q \in \mathbb{Q}_{\geq 2}$, $f_0(E_{p/q}) = f_1(E_{p/q}) = p/q$. One such spectral point is $\overline{\chi_f}$ [SU97, Proposition 3.2.2], and the Haemers bounds provide more such points. We will need the following slightly stronger statement:

Lemma 5.2. *There exist $f_0, f_1 \in \Delta(\mathcal{G})$ such that $f_0(E_{p/q}) = f_1(E_{p/q}) = p/q$ for all $p/q \geq 2$ and for which there exists a vertex transitive graph G with $f_0(G) \neq f_1(G)$.*

A concrete example of such f_0, f_1 are the fractional Haemers bound over the reals and the fractional clique covering number, where we can take the graph G to be the complement of an orthogonality graph Ω_n, which has vertex set $\{-1, 1\}^n$ and $v \sim w$ if and only if $v^T w = 0$. Indeed, it can be deduced from [WE19] and [MR14] that for suitable n we then have $\mathcal{H}_f^{\mathbb{R}}(\overline{\Omega_n}) < \overline{\chi_f}(\overline{\Omega_n})$.

6 Universality proof sketch

A countably universal preorder on sets of words

To prove Theorem 3.1, we will order-embed a particular preorder $(\mathcal{W}, \leq_{\mathcal{W}})$, known to be countably universal by [HN05], into $(\mathcal{G}, \lesssim)$. Let $W = \{0,1\}^*$ be the set of all finite words over the alphabet $\{0,1\}$. For words w, w' we write $w \leq_W w'$ if and only if w' is a prefix (i.e. initial segment) of w. For example, $011000 \leq_W 011$ and $011111 \not\leq_W 011$. Let $\mathcal{W}$ be the class of all finite subsets A of $W = \{0,1\}^*$ such that no distinct words w, w' in A satisfy $w \leq_W w'$ (i.e. A is an antichain). For $A, B \in \mathcal{W}$, we write $A \leq_{\mathcal{W}} B$ if and only if for each $a \in A$ there exists $b \in B$ such that $a \leq_W b$.

Theorem 6.1 ([HN05])**.** *$(\mathcal{W}, \leq_{\mathcal{W}})$ is countably universal.*

Embedding of words in graphs

Before embedding the entirety of the preorder $(\mathcal{W}, \leq_{\mathcal{W}})$, we first assign graphs G_a to each binary string a in the preorder $(W, \leq_W)$. This assignment will respect the order relations in W, but also an additional property:

Lemma 6.2. *There exists an assignment of graphs G_a to each binary string $a \in W$ in such a way that they satisfy the following properties:*

(1) *For every $a, b \in W$, if $a \leq_W b$, then $G_a \lesssim G_b$.*

(2) *For every $a \in W$, there exists an element $\phi_a \in \Delta(\mathcal{G})$ such that for every $b \in W$, if $a \not\leq_W b$, then $\phi_a(G_a) > \phi_a(G_b)$, and in particular $G_a \not\lesssim G_b$.*

A construction of lines

To construct graphs G_a satisfying the properties of Lemma 6.2, we will "simulate" via asymptotic spectrum duality a certain family of lines (each with an associated point), which are themselves an embedding of the preorder $(W, \leq_W)$, with an additional property. The construction, which is technical, and we will not describe here, gives the following (see also Figure 1).

Lemma 6.3. *Let $1 \leq s < t$. We can assign to each word $v \in W$ a pair of rational numbers (a_w, b_w), a corresponding linear map $\ell_w : x \to a_w x + b_w$ and a "witness value" $r_w \in (s, t)$ such that the following conditions are satisfied:*

(i) *For every $v, w \in W$, if $v \leq_W w$ and $v \neq w$, then $a_v < a_w$ and $b_v < b_w$, and thus in particular $\ell_v < \ell_w$ on (s, t).*

(ii) *For every $v, w \in W$, if $v \not\leq_W w$, then $\ell_v(r_v) > \ell_w(r_v)$, and in particular $\ell_v \not\leq \ell_w$ on (s, t).*

Figure 1: Left: If $v <_W w$, then ℓ_v lies below ℓ_w. Right: The blue lines ℓ_w for $v \not\leq_W w$ have the property that at the checkpoint r_v, ℓ_v lies above all the blue lines.

To give an idea how we derive Lemma 6.2 from Lemma 6.3, for every $w \in W$, the line $\ell_w(x) = a_w x + b_w$ will give a graph $G_w := E_{a_w} \boxtimes G \sqcup E_{b_w}$ for some fixed vertex-transitive graph G as in Lemma 5.2, such that $f_0(G) := s$ and $f_1(G) := t$ are distinct. The point r_w will be related to the witness spectral point ϕ_w of G_w. The latter shall be constructed using Lemma 4.3 and Theorem 5.1 applied to f_0 and f_1, so that $\phi_w(G) = r_w$, $\phi_w(E_{p/q}) = p/q$ for all $p/q \geq 2$ and so $\phi_w(G_v) = a_v r_w + b_v = \ell_v(r_w)$.

Proof of Theorem 3.1

We assign to each $a \in W$ a graph G_a as in Lemma 6.2, and define for all sets $A \in \mathcal{W}$ the graph $G_A = \sum_{a \in A} G_a$, where sum is join. We claim that $A \mapsto G_A$ is an order-embedding.

Suppose $A \leq_{\mathcal{W}} B$. By Theorem 4.1, it is enough to show that for all spectral points f, $f(G_A) \leq f(G_B)$. Let $f \in \Delta(\mathcal{G})$ be any spectral point. Then, by Lemma 4.2, we have $f(G_A) = \max_{w \in A} f(G_w)$. Let $v \in A$ be the word for which this maximum is attained. There exists a word $w \in B$ such that $v \leq_W w$ and thereby by Lemma 6.2, $G_v \lesssim G_w$. Hence, $f(G_v) \leq f(G_w) \leq \max_{w' \in B} f(G_{w'}) = f(G_B)$. Therefore, $G_A \lesssim G_B$.

Next, suppose, $A \not\leq_W B$. Then, there is a word $v \in A$ such that all words w in B satisfy $v \not\leq_W w$. By assumptions of Lemma 6.2, there exists a witness spectral point ϕ_v such that for all $v' \not\geq_W v$, $\phi_v(G_v) > \phi_v(G_{v'})$. Thus, $\phi_v(G_A) \geq \phi_v(G_v) > \max_{w \in B} \phi_v(G_w) = \phi_v(G_B)$. Therefore, we cannot have $G_A \lesssim G_B$.

References

[AL06] Noga Alon and Eyal Lubetzky. The Shannon capacity of a graph and the independence numbers of its powers. *IEEE Trans. Inform. Theory*, 52(5):2172–2176, 2006. doi:10.1109/TIT.2006.872856.

[BC19] Boris Bukh and Christopher Cox. On a fractional version of Haemers' bound. *IEEE Trans. Inform. Theory*, 65(6):3340–3348, 2019. arXiv:1802.00476, doi:10.1109/TIT.2018.2889108.

[Bla13] Anna Blasiak. *A graph-theoretic approach to network coding*. PhD thesis, Cornell University, 2013. URL: https://hdl.handle.net/1813/34147.

[dBBZ24] David de Boer, Pjotr Buys, and Jeroen Zuiddam. The asymptotic spectrum distance, graph limits, and the Shannon capacity, 2024. arXiv:2404.16763.

[FHLN17] Jiří Fiala, Jan Hubička, Yangjing Long, and Jaroslav Nešetřil. Fractal property of the graph homomorphism order. *European Journal of Combinatorics*, 66:101–109, 2017. Selected papers of EuroComb15. doi:10.1016/j.ejc.2017.06.017.

[Hae81] Willem Haemers. An upper bound for the Shannon capacity of a graph. In *Algebraic methods in graph theory, Vol. I, II (Szeged, 1978)*, volume 25 of *Colloq. Math. Soc. János Bolyai*, pages 267–272. North-Holland, Amsterdam-New York, 1981.

[Hed69] Zdeněk Hedrlín. On universal partly ordered sets and classes. *Journal of Algebra*, 11:503–509, 1969. doi:10.1016/0021-8693(69)90089-1.

[HN04a] Pavol Hell and Jaroslav Nešetřil. Graphs and homomorphisms. In *Oxford lecture series in mathematics and its applications*, 2004. doi:10.1093/acprof:oso/9780198528173.001.0001.

[HN04b] Jan Hubička and Jaroslav Nešetřil. Finite paths are universal. *Order*, 21:181–200, 08 2004. doi:10.1007/s11083-004-3345-9.

[HN05] Jan Hubička and Jaroslav Nešetřil. Universal partial order represented by means of oriented trees and other simple graphs. *European Journal of Combinatorics*, 26(5):765–778, 2005. doi:10.1016/j.ejc.2004.01.008.

[Lov79] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985.

[MR14] Laura Mančinska and David E. Roberson. Graph homomorphisms for quantum players. In *9th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014)*, 2014. doi:10.4230/LIPICS.TQC.2014.212.

[MR16] Laura Mančinska and David E. Roberson. Quantum homomorphisms. *Journal of Combinatorial Theory, Series B*, 118:228–267, May 2016. doi:10.1016/j.jctb.2015.12.009.

[Pol20] Sven Polak. New methods in coding theory: Error-correcting codes and the Shannon capacity, 2020. arXiv:2005.02945.

[PS19] Sven C. Polak and Alexander Schrijver. New lower bound on the Shannon capacity of C_7 from circular graphs. *Inform. Process. Lett.*, 143:37–40, 2019. doi:10.1016/j.ipl.2018.11.006.

[PT80] Aleš Pultr and Věra Trnková. *Combinatorial, algebraic and topological representations of groups, semigroups and categories*, volume 22 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-New York, 1980. doi:10.1016/s0924-6509(08)x7022-9.

[Sch23] Alexander Schrijver. On the Shannon capacity of sums and products of graphs. *Indag. Math. (N.S.)*, 34(1):37–41, 2023. doi:10.1016/j.indag.2022.08.009.

[Sha56] Claude E. Shannon. The zero error capacity of a noisy channel. *Institute of Radio Engineers Transactions on Information Theory*, IT-2,:8–19, 1956. doi:10.1109/tit.1956.1056798.

[SU97] Edward R. Scheinerman and Daniel H. Ullman. *Fractional graph theory: A rational approach to the theory of graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1997.

[Vin88] Andrew Vince. Star chromatic number. *J. Graph Theory*, 12:551–559, 1988. doi:10.1002/jgt.3190120411.

[Vra21] Péter Vrana. Probabilistic refinement of the asymptotic spectrum of graphs. *Combinatorica*, 41(6):873–904, 2021. arXiv:1903.01857, doi:10.1007/s00493-020-4324-5.

[WE19] Pawel Wocjan and Clive Elphick. Spectral lower bounds for the orthogonal and projective ranks of a graph. *Electron. J. Combin.*, 26(3):Paper No. 3.45, 14, 2019. arXiv:1806.02734, doi:10.37236/8183.

[WZ23] Avi Wigderson and Jeroen Zuiddam. Asymptotic spectra: Theory, applications and extensions, 2023. URL: https://staff.fnwi.uva.nl/j.zuiddam/papers/convexity.pdf.

[Zhu06] Xuding Zhu. Recent developments in circular colouring of graphs. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 497–550. Springer, Berlin, 2006. doi:10.1007/3-540-33700-8_25.

[Zui19] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity. *Combinatorica*, 39(5):1173–1184, 2019. doi:10.1007/s00493-019-3992-5.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Classification of unstable circulant graphs of twice odd order

(Extended abstract)

Bartłomiej Bychawski*

Abstract

In my talk simple graphs will be the subject of the main focus. A canoncial double cover of a graph $\Gamma = (V, E)$ is the tensor product $\Gamma \times \mathrm{K}_2$. A graph is called stable when the automorphism group of this double cover is as small as possible. More precisely, a graph is stable when $\mathrm{Aut}(\Gamma \times \mathrm{K}_2) \cong \mathrm{Aut}(\Gamma) \times \mathrm{Aut}(\mathrm{K}_2)$ and unstable otherwise. The problem of stability proved to be closely connected with studying symmetries of more general tensor products, hence gained a lot of interests. In the recent years the focus was put on classifying unstable graphs in certain graph families.

The most studied familiy, which still resists classification are Cayley graphs. Even unstable Cayley graphs of cyclic groups called circulants are not classificied yet. In two previous years theory of Schur rings was used to obtain interesting partial results. In my paper I expand on those methods, mixing them with classification of primitive group actions containing cyclic regular action and the theory of Galios cohomology. Such a combination of different techniques allowed me to classify all unstable circulants of twice odd order.

1 Preliminaries and context

In this abstract by *graphs* we understand finite, loopless and undirected graphs. Formally, we denote graphs as pairs (V, E), where V is the set of vertices and E is the edge relation, which is irreflexive and symmetric.

The primary motivation behind the concept of *stability of graphs* is to understand the structure of symmetries of the tensor product $\Gamma \times \Sigma$ with respect to symmetries of graphs Γ and Σ.

Definition 1.1. For two graphs $\Gamma = (V, E)$ and $\Sigma = (W, F)$ by *tensor product of* Γ *and* Σ we understand the graph with vertex set $V \times W$ and edge set

$$\Big\{\big((v_1, w_1), (v_2, w_2)\big) \in (V \times W) \times (V \times W) \;\big|\; (v_1, v_2) \in E \text{ and } (w_1, w_2) \in F\Big\}.$$

*Faculty of Mathematics, Jagiellonian University, 30-348 Kraków, Poland. E-mail: bartlomiej.bychawski@student.uj.edu.pl.

We denote this graph by $\Gamma \times \Sigma$. It is called a tensor product, because if the adjacency matrix of Γ is A and the adjacency matrix of Σ is B, then adjacency matrix of $\Gamma \times \Sigma$ is $A \otimes B$.

Definition 1.2. For a graph $\Gamma = (V, E)$ we call a permutation σ of the set V a *symmetry* or an *automorphism* of Γ when for any pair of vertices $v, w \in V$, $(v, w) \in E$ if and only if $\big(\sigma(v), \sigma(w)\big) \in E$. The set of all automorphisms of Γ forms the group with composition of functions. We denote this group by $\mathrm{Aut}(\Gamma)$.

It is easy to observe, that it always happens that $\mathrm{Aut}(\Gamma) \times \mathrm{Aut}(\Gamma) \leq \mathrm{Aut}(\Gamma \times \Sigma)$, however equality does not have to hold, therefore researchers aimed to describe $\mathrm{Aut}(\Gamma \times \Sigma)$ completely. It was early observed that one should only consider reduced (also known as R-thin or worthy) and connected graphs, as every graph can be reduced and split into its' connected components. In 1974 Dörfler [5, Theorem 8.18] solved the problem assuming both Γ and Σ are non-bipartite.

Then people became interested in multiplying pairs of graphs, among which at least one is bipartite, as it was clear, that they behave somewhat differently. It was early noticed, that if we take Γ non-biaprtite and Σ bipartite, then most symmetries of $\Gamma \times \Sigma$ which were not easily predictable came from symmetries of the graph $\Gamma \times \mathrm{K}_2$.

Definition 1.3. We call a graph Γ *stable*, when $\mathrm{Aut}(\Gamma) \times \mathrm{Aut}(\Gamma) \cong \mathrm{Aut}(\Gamma \times \Sigma)$ and *unstable* otherwise.

The name *stable* comes form the fact, that if Γ is stable, then for Σ from certain large classes of bipartite graphs, all automorphisms of $\Gamma \times \Sigma$ are predictable and easy to describe (cf. [4, Theorem 1.6]).

2 Previous results regarding stability of graphs

It was early noticed, that if a graph Γ is either disconnected, non-reduced or bipartite and symmetric, then it is unstable. For that reason people focused of understanding unstable graphs which does not fall under any of the above categories. These graphs are now called *non-trivially unstable graphs* [17]. In the same work Wilson [17] showed, that each non-trivially unstable graphs can be obtained from one of three general constructions. These constructions were quite general in nature, however he noticed, that for classes of certain well structured graphs all unstable graphs can be described with greater accuracy. In his original paper he pointed out five families of graphs. Those families were Cayley graphs of cyclic groups (also called cicrulants), Generalized Petersen graphs, Rose Window graphs, Toroidal graphs and family of graphs denoted by $V(k, m)$.

Definition 2.1. A Cayley graph of the finite group H and the edge set S (we require $S = S^{-1}$ to make it undirected) is the graph with the vertex set H and an the following edge relation

$$\Big\{(h_1, h_2) \in H \times H \;\big|\; h_2 \cdot h_1^{-1} \in S\Big\}.$$

We denote this graph by $\mathrm{Cay}(H, S)$. By $\mathbb{Z}_n$ we denote the cyclic group of order n represented by a quotient of the group $\mathbb{Z}$ by its subgroup $n\mathbb{Z}$. *Circulants* are therefore graphs, which are of the form $\mathrm{Cay}(\mathbb{Z}_n, S)$ for some $S \subset \mathbb{Z}_n$ such that $0 \notin S$ and $S = -S$.

Unstable graphs belonging to most of these families were later classified [18, 11, 13], however the family of circulant graphs proved itself to be difficult [14, 1, 7]. The original hypothesis given by Wilson turned out to be false, and were later repaired by several authors [14, 7]. Recently partials results started to appear. Fernandez and Hujdurović were the first to show, that there are no non-tivially unstable circulants of odd order [3]. Then Morris [12] found a simpler proof which generalized this result for arbitrary Cayley graphs over abelian groups of odd order. In 2023 Hujdurović and Kovács proved the following result.

Theorem 2.2. ([6, Theorem 1.5]) Let $n = 2p^e$ where p is an arbitraty odd prime and $e \geq 1$. Assume $\Gamma = \mathrm{Cay}(\mathbb{Z}_n, S)$ is connected and non-bipartite. Then Γ is unstable if and only if

i. there exists nonzero $h \in 2\mathbb{Z}_n$ such that $S \cap 2\mathbb{Z}_n + h = S \cap 2\mathbb{Z}_n$;

ii. or $\mathrm{Cay}(\mathbb{Z}_n, S) \cong \mathrm{Cay}(\mathbb{Z}_n, S + p^e)$.

Its proof is much harder and involve study of structure of certain algebras related to groups of symmetries. Those algebras are called *Schur Rings*. The importance of these structures comes from the following observation of Schur.

Observation 2.3. 1([15]) Let H be any finite group, and let G be the group of permutations of set A containing the action of right multiplication by every element of H. Then the linear subspace of the group algebra $\mathbb{Q}[H]$ invariant under the action of the group $\mathrm{stab}_G(e_H)$ is a ring. This ring is called a *transitivity module*.

Definition of a Schur ring is more technical than the one of transitivity module, however it involves only certain partition of the group H. For that reason it is generally easier to work with Schur rings than with transitivity modules. Schur ring are of our interests precisely because they generalize the notion of transitivity modules introduced by Schur. Schur rings over cyclic groups were studied for a long time, and now a lot of structural results are known [9, 8], which makes them an useful tool when working with problems regarding symmetries of circulants or related graphs.

3 Main results and outline of the proof

Main result which will be presented during my talk is the following classification.

Theorem 3.1. Let $n = 2m$ where $m > 1$ is an odd positive integer and let $\Gamma = \mathrm{Cay}(\mathbb{Z}_n, S)$ be a connected and non-bipartite graph. Then Γ is unstable if and only if

i. there exists nonzero $h \in 2\mathbb{Z}_{2m}$ such that $S \cap 2\mathbb{Z}_{2m} + h = S \cap 2\mathbb{Z}_{2m}$;

ii. or $\mathrm{Cay}(\mathbb{Z}_{2m}, S) \cong \mathrm{Cay}(\mathbb{Z}_{2m}, S + m)$.

In previous version of my paper [2] I proved the above statement only for n-s which are square-free. The final result is obtained by extending the methods introduced there. Before we proceed with the sketch of the proof, it is important to notice, that K_2 is a Cayley graph of the group $\mathbb{Z}_2$, hence $\mathrm{Cay}(\mathbb{Z}_n, S) \times \mathrm{K}_2 \cong \mathrm{Cay}(\mathbb{Z}_n \times \mathbb{Z}_2, S \times \{1\})$. From now on we shall use this interpretation often.

The starting point in proving the main theorem is the following lemma of Hujdurović and Kovács. In their paper it is stated in the language of previously mentioned Schur rings, however for simplicity it is formulated here in the language of graphs automorphisms.

Lemma 3.2. ([6, Theorem 1.4]) Let $m > 1$ be an odd integer and let $\Gamma = \mathrm{Cay}(\mathbb{Z}_{2m}, S)$ be connected non-bipartite unstable circulant. Then either

i. there exists nonzero $h \in 2\mathbb{Z}_{2m}$ such that $S \cap 2\mathbb{Z}_{2m} + h = S \cap 2\mathbb{Z}_{2m}$;

ii. or for every $\sigma \in \mathrm{Aut}(\Gamma \times \mathrm{K}_2)$, if $\sigma\big((0,0)\big) = (0,0)$ then $\sigma\big((0,1)\big) \in \big\{(0,1),(m,1)\big\}$.

The first part of the conclusion of the above lemma is precisely the first case from our main theorem. From now on we will assume that Γ is reduced and every $\sigma \in \mathrm{Aut}(\Gamma \times \mathrm{K}_2)$ which fixes the vertex $(0,0)$ satisfies $\sigma\big((0,1)\big) \in \big\{(0,1),(m,1)\big\}$.

This is the moment when the approach of the author diverges significantly from the approach taken in [6]. In the aforementioned paper to finish the proof for $n = 2p^e$ authors constructed the automorphisms using facts which are very specific for Schur rings over groups $\mathbb{Z}_{p^e}$. Here we abandon the purely Schur ring theoretical approach and find a measure of instability for a single symmetry of $\mathrm{Aut}(\Gamma \times \mathrm{K}_2)$, instead of focusing on the global measure of instability of the whole graph, which is best stated using properties of certain transitivity module (cf. [6, Theorem 1.1]).

Definition 3.3. Let $a \in \mathrm{Aut}(\Gamma \times \mathrm{K}_2)$ be a permutation given by the formula $a\big((x,i)\big) = (x, i+1)$. Now let $\mathbb{A}$ be the group of permutations of $\mathbb{Z}_{2m} \times \mathbb{Z}_2$ which map each (x,i) either onto itself or onto $(x+m, i)$. Finally let us define the function

$$\omega : \mathrm{Aut}(\Gamma \times \mathrm{K}_2) \to \mathbb{A} \qquad \sigma \mapsto [\sigma, a] := \sigma \circ a \circ \sigma^{-1} \circ a^{-1}.$$

Above definition makes sense in the case of our interests. Generally the image of ω may not be contained in $\mathbb{A}$, this is a consequence of our assumptions.

Function ω defined above is surprisingly well structured and well suited for our purposes. First of all, preimage of $0 = e_{\mathbb{A}}$ is exactly the standard copy of $\mathrm{Aut}(\Gamma) \times \mathrm{Aut}(\mathrm{K}_2)$ contined in $\mathrm{Aut}(\Gamma \times \mathrm{K}_2)$, hence Γ beeing unstable means that ω is not the 0 function. It is also easy to see that the isomorphism $\mathrm{Cay}(\mathbb{Z}_{2m}, S) \cong \mathrm{Cay}(\mathbb{Z}_{2m}, S + m)$ is equivalent to existence of element $\sigma \in \mathrm{Aut}(\Gamma \times \mathrm{K}_2)$ for which $\omega(\sigma)$ is just an addition of $(m, 0)$ to each element of $\mathbb{Z}_{2m} \times \mathbb{Z}_2$. One can also observe that ω maps the group $G = \mathrm{Aut}(\Gamma \times \mathrm{K}_2)$ into the G-module $\mathbb{A}$, because $\mathbb{A}$ is abelian and G acts on it by conjugation. We can actually say much more, namely that ω satisfies *cocycle equation*, which means that

$$\forall \sigma, \tau \in G \quad \omega(\sigma\tau) = \sigma.\omega(\tau) + \omega(\sigma),$$

where the dot symbol represents the action of G on the G-module $\mathbb{A}$ and the plus symbol represents the group action of $\mathbb{A}$. This fact is the crucial observation, because it allows us to use already well developed *cohomology theory for group modules* (cf. [16, Chapter VII]) also known as *Galois cohomology*.

Definition 3.4. Assume G is a finite group and M is an G-module. A function $\varphi : G \to M$ is called a *cocycle* if it satisfies aforementioned *cocycle equation*. We additionally call φ a *coboundary* if there exists $m \in M$ such that for every $g \in G$, $\varphi(g) = g.m - m$. We denote the group of all cocycles with pointwise addition by $\mathrm{Cocyc}(G, M)$ and its' subgroup made of all coboundries by $\mathrm{Cobund}(G, M)$. Now we are ready to define *the first cohomology group* by

$$H^1(G, M) := \mathrm{Cocyc}(G, M) \Big/ \mathrm{Cobund}(G, M).$$

For groups G with a sufficiently simple structure and reasonable G-modules it is possible to compute this group either via explicit computation or with use of some known exact sequences. Due to the fact that all primitive group actions of G on a given set $\mathbb{X}$ which contain cyclic subgroup acting regularly on $\mathbb{X}$ were classified (cf. [10, Corollary 1.2]), it is possible to calculate cohomology groups for primitive groups which are of interests and suitable group modules.

Definition 3.5. Let G be a finite group acting on a finite set $\mathbb{X}$. Then by $\mathbb{F}_2[\mathbb{X}]$ we denote the linear space over the field $\mathbb{F}_2$ which has a basis $\{\vec{e}_x\}_{x\in\mathbb{X}}$. Then the action of $g \in G$ on $\mathbb{F}_2[\mathbb{X}]$ is defined on a canonical basis by formula $\vec{e}_x \mapsto \vec{e}_{g.x}$ and extended linearly.

To understand why such modules are of interests, observe that $\mathrm{Aut}(\Gamma \times \mathrm{K}_2)$-module $\mathbb{A}$ is naturally isomorphic with $\mathbb{F}_2\left[\mathbb{Z}_{2m} \times \mathbb{Z}_2 \big/ m\mathbb{Z}_{2m} \times 0\right]$ as a group module.

Theorem 3.6. Let standard action of $G \leq \mathrm{Sym}(\mathbb{X})$ on $\mathbb{X} = \mathbb{Z}_k$ be primitive and let the action of addition by elements from $\mathbb{Z}_k$ be a subgroup of G. Under additional technical conditions we can conclude that $H^1(G, \mathbb{F}_2[\mathbb{X}]) \cong 0$ or $H^1(G, \mathbb{F}_2[\mathbb{X}]) \cong \mathbb{F}_2$.

It turns out, that the condition that ω is trivial on the subgroup $(\mathbb{Z}_n)_r$ of G, which consists of actions of adding elements form $\mathbb{Z}_k$ to each point, is enough to identify cocycle just from its' image in the first cohomology group. It allows to obtain the following corollary.

Corollary 3.7. Let standard action of $G \leq \mathrm{Sym}(\mathbb{X})$ on $\mathbb{X} = \mathbb{Z}_k$ be primitive and let the action of addition by elements from $\mathbb{Z}_k$ be a subgroup of G. Under additional technical conditions, if $\omega : G \to \mathbb{F}_2[\mathbb{X}]$ is a nonzero cocycle such that $\omega|_{(\mathbb{Z}_n)_r} \equiv \vec{0}$, then

$$\omega(g) = \begin{cases} \vec{0} & \text{if } g \in G_0 \\ \sum_{x\in X} \vec{e}_x & \text{otherwise} \end{cases}, \text{ where } G_0 \text{ is the unique subgroup of } G \text{ of index 2.}$$

Above corollary motivates us to reduce given problem to the one involving primitive group actions. It is done in three steps.

1. We narrow our view to a certain subgroup of $\mathcal{H} \leq \mathrm{Aut}(\Gamma \times K_2)$ on which function ω is still nonzero, however on every smaller group form a certain family ω already is identically zero
2. We show, that for a certain subgroup of vertices $K \leq \mathbb{Z}_{2m} \times \mathbb{Z}_2$, if we find $\sigma \in \mathcal{H}$ satisfying
$$\omega(\sigma)|_K = +(m,0)|_K, \text{ then we can find } \widetilde{\sigma} \in \mathcal{K} \text{ such that } \omega(\widetilde{\sigma}) = +(m,0).$$
3. We prove that indeed such $\sigma \in \mathcal{H}$ exists by considering the appropriate quotients. The action of $\mathcal{H}$ on the obtained quotient structure is best described by the aforementioned primitive actions. Minimality of $\mathcal{H}$ is needed to ensure that the value of $\omega(\sigma)$ depends only on the action of $\sigma \in \mathcal{H}$ on this quotient structure.

This reduction combined with above described methods of applying Galois cohomology completes the proof of the main theorem.

References

[1] Dave Witte Morris Ademir Hujdurović, Đorđe Mitrović. Automorphisms of the double cover of a circulant graph of valency at most 7. *Algebraic Combinatorics*, 6:1235–1271, 2023.

[2] Bartłomiej Bychawski. Classification of unstable circulants of square-free order. *arXiv:2410.00701*, 2024.

[3] Blas Fernandez and Ademir Hujdurović. Canonical double covers of circulants. *Journal of Combinatorial Theory, Series B*, 154:49–59, 2022.

[4] Yunsong Gan, Weijun Liu, and Binzhou Xia. Unexpected automorphisms in direct product graphs. *Journal of Combinatorial Theory, Series B*, 171:140–164, 2025.

[5] Richard Hammack, Wilfried Imrich, and Sandi Klavzar. *Handbook of Product Graphs, Second Edition*. CRC Press, Inc., 2nd edition, 2011.

[6] Ademir Hujdurović and István Kovács. Stability of cayley graphs and schur rings. *arXiv:2308.01182*, 2023.

[7] Ademir Hujdurović, Đorđe Mitrović, and Dave Morris. On automorphisms of the double cover of a circulant graph. *The Electronic Journal of Combinatorics*, 28, 2021.

[8] Ka Hin Leung and Shing Hing Man. On schur rings over cyclic groups, ii. *Journal of Algebra*, 183(2):273–285, 1996.

[9] Ka Hin Leung and Shing Hing Man. On schur rings over cyclic groups. *Israel Journal of Mathematics*, 106:251–267, 1998.

[10] Cai Li. The finite primitive permutation groups containing an abelian regular subgroup. *Proceedings of the London Mathematical Society*, 87:725 – 747, 2003.

[11] Klavdija Kutnar Milad Ahanjideh, István Kovács. Stability of rose window graphs. *arXiv:2306.01619*, 2023.

[12] Dave Witte Morris. On automorphisms of direct products of cayley graphs on abelian groups. *The Electronic Journal of Combinatorics*, 28, 2021.

[13] Dave Witte Morris. Automorphisms of the canonical double cover of a toroidal grid. *arXiv:2301.05396*, 2023.

[14] Yan-Li Qin, Binzhou Xia, and Sanming Zhou. Stability of circulant graphs. *Journal of Combinatorial Theory, Series B*, 136:154–169, 2019.

[15] Issai Schur. Zur theorie der einfach transitiven permutationgruppen. *S.-B. Preuss. Akad. Wiss. Phys.- Math. Kl.*, 18:598–623, 1933.

[16] Jean-Pierre Serre. *Local Fields*. Graduate Texts in Mathematics. Springer New York, NY, 1979.

[17] Steve Wilson. Unexpected symmetries in unstable graphs. *Journal of Combinatorial Theory, Series B*, 98(2):359–383, 2008.

[18] S. Zhou Y. L. Qin, B. Xia. Canonical double covers of generalized petersen graphs, and double generalized petersen graphs. *J. Graph Theory*, 97:70–81, 2021.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Higher-order Fourier analysis via spectral algorithms

(Extended abstract)

Pablo Candela* Diego González-Sánchez† Balázs Szegedy‡

Abstract

We study an approach to higher-order Fourier analysis based on spectral decomposition of operators. That enables us to establish a new regularity lemma for the U^3 Gowers norm that is completely algorithmic and efficiently computable. The ideas behind this approach and one of our main results are presented in this extended abstract. The paper [4], which will be published elsewhere, contains the full details.

1 Introduction

Higher-order Fourier analysis is a branch of mathematics that has led to groundbreaking results over the past 25 years. Pioneered by Gowers in 2001 [8], this field has resulted in major advancements in areas such as additive combinatorics [14, 22], ergodic theory [19, 33], dynamical systems [17], theoretical computer science [18, 31], and number theory [12] (notably, the Green-Tao theorem). As a generalization of Fourier analysis, it holds promise for applications in the applied sciences, similar to its classical counterpart. However, examples of practical applications remain sparse, with notable exceptions including the work of Tulsiani and Wolf, who introduced quadratic analogues of the Goldreich–Levin

*Instituto de Ciencias Matemáticas, Calle Nicolás Cabrera 13-15, Madrid 28049, Spain E-mail: pablo.candela@icmat.es. Supported by MICINN project PID2020-113350GB-I00

†HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15. Budapest, Hungary, H-1053 E-mail: diegogs@renyi.hu. Supported by MICINN project PID2020-113350GB-I00, and NKFIH projects KPP 133921, Momentum (Lendület) 30003, and the Artificial Intelligence National Laboratory.

‡HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15. Budapest, Hungary, H-1053 E-mail: szegedyb@gmail.com. Supported by MICINN project PID2020-113350GB-I00, and NKFIH projects KPP 133921, Momentum (Lendület) 30003, and the Artificial Intelligence National Laboratory.

algorithm [32], and the recent paper by Kim, Li, and Tidor on the cubic case [21]. Specifically designed for the finite-field context, which is central to theoretical computer science, these algorithms are probabilistic and involve iterative processes [32, See §3]. In this extended abstract, we present a family of algorithms based on the spectral decomposition of operators related to functions. These algorithms are more straightforward (essentially non-iterative) and can be applied to any finite abelian group.

2 Preliminaries

In this note, we will work with functions $f : \mathrm{Z} \to \mathbb{C}$ (or $f \in \mathbb{C}^{\mathrm{Z}}$ for short) defined on finite abelian groups Z. The notation $\mathbb{E}$ represents the average with respect to the counting measure, e.g. for any $f \in \mathbb{C}^{\mathrm{Z}}$ we have $\mathbb{E}_{x\in\mathrm{Z}} f(x) = \frac{1}{|\mathrm{Z}|}\sum_{x\in\mathrm{Z}} f(x)$. The notion of inner product in $\mathbb{C}^{\mathrm{Z}}$ is normalized according to this measure, i.e. $\langle f, g\rangle = \mathbb{E}_{x\in\mathrm{Z}} f(x)\overline{g(x)}$ for any $f, g \in \mathbb{C}^{\mathrm{Z}}$. A Z-*matrix* is a matrix with complex entries whose rows and columns are indexed by elements of Z. The set of Z-matrices is denoted by $\mathbb{C}^{\mathrm{Z}\times\mathrm{Z}}$. The notion of matrix multiplication (and thus also of eigenvalue) is also normalized, i.e. for $M \in \mathbb{C}^{\mathrm{Z}\times\mathrm{Z}}$ and $f \in \mathbb{C}^{\mathrm{Z}}$ we have $(Mf)(x) = \mathbb{E}_{y\in\mathrm{Z}} M(x, y)f(y)$. See [4, §2] for further details.

For any $f \in \mathbb{C}^{\mathrm{Z}}$ and $t \in \mathrm{Z}$, we define the *multiplicative derivative* $\Delta_t f \in \mathbb{C}^{\mathrm{Z}}$ as $\Delta_t f(x) := f(x + t)\overline{f(x)}$. For an integer $k \geq 1$, the Gowers U^k norm on $\mathbb{C}^{\mathrm{Z}}$ is defined as $\|f\|_{U^k}^{2^k} := \mathbb{E}_{x,t_1,\ldots,t_k\in\mathrm{Z}} \Delta_{t_1}\cdots\Delta_{t_k} f(x)$, see [8, Lemma 3.9] and [29, Definition 11.2].[1]

A **regularity lemma** for the U^k norm is a decomposition of any (typically bounded) $f \in \mathbb{C}^{\mathrm{Z}}$ as[2] $f = f_s + f_r$ where $\|f_r\|_{U^k}$ is "small" and f_s is "structured" with respect to the U^k norm, see e.g. [9, Theorem 5.1] and [11, Theorem 1.2]. To detail this a little more and make the connection with (classical) Fourier analysis, let us describe how one such regularity lemma looks like for the U^2 norm.

For any Z, we let $\widehat{\mathrm{Z}} := \{\chi : \mathrm{Z} \to \mathbb{S}^1 \subset \mathbb{C} : \chi \text{ is a homomorphism}\}$ be its dual group. For $f \in \mathbb{C}^{\mathrm{Z}}$ and $\chi \in \widehat{\mathrm{Z}}$, we let $\widehat{f} : \widehat{\mathrm{Z}} \to \mathbb{C}$, defined as $\chi \mapsto \langle f, \chi\rangle$, be the *Fourier transform* of f. Recall that, by the *Fourier inversion formula* [29, (4.4)], we have $f = \sum_{\chi\in\widehat{\mathrm{Z}}} \widehat{f}(\chi)\chi$.

Lemma 2.1 (U^2-regularization)**.** *Let* Z *be a finite abelian group and $f \in \mathbb{C}^{\mathrm{Z}}$ be a function. For any $\varepsilon > 0$, a decomposition of the form $f = f_s + f_r$ where $f_s := \sum_{|\widehat{f}(\chi)|\geq\varepsilon} \widehat{f}(\chi)\chi$ and $f_r := f - f_s$ is a regularity lemma with respect to the U^2 norm.*

Note that in this case, the notion of "structured" with respect to the U^2 norm is that f_s is the sum of "few" Fourier characters.

Analogous regularity lemmas exist for $k \geq 3$, see [11]. A surprising and non-trivial feature of this theory is that to describe f_s for $k \geq 3$, we need to consider also nilpotent structures [1, 6, 13, 14, 20]. A simpler yet powerful way of measuring how structured a function is consists in using the U^k-dual norm, as done in [9]. Such a norm is defined by

[1] The case $k = 1$ yields only a seminorm $\|f\|_{U^1} = |\mathbb{E}_{x\in\mathrm{Z}} f(x)|$, but it is a common abuse of notation to call it the U^1 norm.

[2] Usually for $k \geq 3$ there is an extra summand f_e which is normally bounded and small in L^1 norm.

$\|f\|^*_{U^k} := \sup_{g\in\mathbb{C}^{\mathrm{Z}},\|g\|_{U^k}\leq 1} \mathbb{E}_{x\in\mathrm{Z}} f(x)\overline{g(x)}$, and we will say that $f \in \mathbb{C}^{\mathrm{Z}}$ is U^k *structured* if $\|f\|^*_{U^k}$ is bounded, see [4, §2.3] for further details.

There are various approaches for obtaining such *regularity lemmas* based on additive combinatorics [8, 10, 22, 23], ergodic theory [28, 30], non-standard analysis [7, 14], or combinations of those. However, very few of those approaches have evolved into practical algorithms with the exceptions noted before [21, 32].

3 Outline of the spectral approach

To introduce and motivate our spectral approach, we first show how it can be used to recover 1-st order Fourier analysis (i.e. classical Fourier analysis) from 0-th order Fourier analysis. Our goal is thus to regularize a function $f \in \mathbb{C}^{\mathrm{Z}}$ with respect to the U^2 norm, i.e. recover Lemma 2.1, using spectral analysis and regularization with respect to the U^1 norm.

Lemma 3.1 (U^1-regularization)**.** *Let $f \in \mathbb{C}^{\mathrm{Z}}$. Then $f = f_s + f_r$ where $f_s := \mathbb{E}_{x\in\mathrm{Z}} f(x)$ and $f_r := f - f_s$ is the only decomposition where $\|f_r\|_{U^1} = 0$ and $\|f_s\|^*_{U^1} < \infty$.*

Proof. It follows from [4, Lemma 2.26] that for any $g \in \mathbb{C}^{\mathrm{Z}}$, if $\|g\|^*_{U^1} < \infty$ then g is contant. Hence, f_s must be a constant, and clearly $f_s = \mathbb{E}_{x\in\mathrm{Z}} f(x)$ is the only one satisfying $\|f_r\|_{U^1} = \|f - f_s\|_{U^1} = 0$. □

Now let $f \in \mathbb{C}^{\mathrm{Z}}$ be a function we want to regularize with respect to the U^2 norm. The first step is to turn f into the Z-matrix $f \otimes \overline{f} \in \mathbb{C}^{\mathrm{Z}\times\mathrm{Z}}$, $(x, y) \mapsto f(x)\overline{f(y)}$. Now, for any $t \in \mathrm{Z}$, let us consider $\{(z + t, z) : z \in \mathrm{Z}\}$ which we denote by the Z-*diagonal* at level t of a Z-matrix, see [4, Definition 2.1]. Note that $f \otimes \overline{f}$ evaluated at the Z-diagonal al level t equals the multiplicative derivative $\Delta_t f(x)$. The next step is to replace each Z-diagonal of $f \otimes \overline{f}$ with its U^1-regularized version. That is, let $K : \mathbb{C}^{\mathrm{Z}} \to \mathbb{C}^{\mathrm{Z}}$ be the operator $f \mapsto \mathbb{E}_{x\in\mathrm{Z}} f(x)$ and let $\mathcal{K}(f \otimes \overline{f})$ be the Z-matrix where we have replaced each Z-diagonal of $f \otimes \overline{f}$ by $K(\Delta_t f)$. We are now ready to prove a result that recovers Lemma 2.1 through the spectral decomposition of $\mathcal{K}(f \otimes \overline{f})$.

Lemma 3.2 (Spectral U^2-regularization)**.** *Let $f \in \mathbb{C}^{\mathrm{Z}}$ and $\varepsilon > 0$. Let $K : \mathbb{C}^{\mathrm{Z}} \to \mathbb{C}^{\mathrm{Z}}$ and $\mathcal{K}(f \otimes \overline{f})$ be as above. Then the projection of f to the eigenspaces of $\mathcal{K}(f \otimes \overline{f})$ whose eigenvalues are at least ε^2 equals $\sum_{|\widehat{f}(\chi)|\geq\varepsilon} \widehat{f}(\chi)\chi$.*

Proof. The result follows from the next simple calculation.

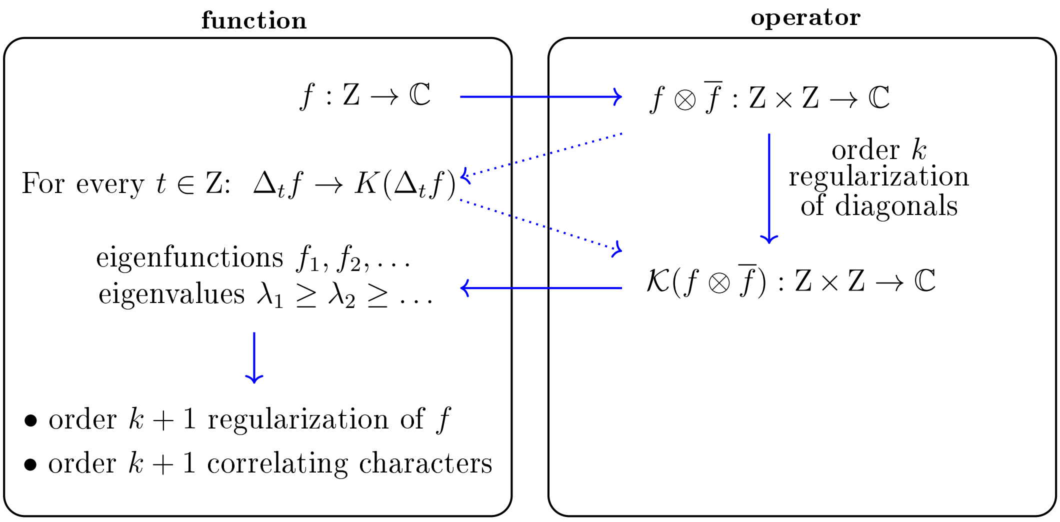
$$\begin{aligned}\mathcal{K}(f \otimes \overline{f})(x, y) &= \mathbb{E}_{z\in\mathrm{Z}} f(x + z)\overline{f(y + z)} = \mathbb{E}_{z\in\mathrm{Z}} \textstyle\sum_{\chi\in\widehat{\mathrm{Z}}} \langle f, \chi\rangle\chi(x + z)\overline{\sum_{\chi'\in\widehat{\mathrm{Z}}} \langle f, \chi'\rangle\chi'(y + z)} \\ &= \textstyle\sum_{\chi,\chi'\in\widehat{\mathrm{Z}}} \langle f, \chi\rangle\overline{\langle f, \chi'\rangle}\mathbb{E}_{z\in\mathrm{Z}}\chi(x + z)\overline{\chi'(y + z)} = \sum_{\chi\in\widehat{\mathrm{Z}}} |\langle f, \chi\rangle|^2\chi(x)\overline{\chi(y)}. \square\end{aligned}$$

It is natural to ask what happens if we repeat the previous procedure, but replace the averaging operator K with an operator that performs U^k regularization. It appears that the spectral information of $\mathcal{K}(f \otimes \overline{f})$ contains crucial information related to $(k + 1)$-th

Fourier analysis. As outlined in [4, §1.1], the following principle encapsulates the key idea behind this approach.

Order increment principle: *If $K : \mathbb{C}^{\mathrm{Z}} \to \mathbb{C}^{\mathrm{Z}}$ maps functions in $\mathbb{C}^{\mathrm{Z}}$ to their k-th order structured parts, then for $f \in \mathbb{C}^{\mathrm{Z}}$ the spectral decomposition of $\mathcal{K}(f \otimes \overline{f})$ can be used to obtain the $(k+1)$-th order structured part (and corresponding useful decomposition) of f.*

In the ultralimit setting, a precise formulation of this principle was identified in [25], where it enables the transformation of k-th order Fourier analysis into $(k+1)$-th order Fourier analysis.



Sketch of the spectral approach to higher-order Fourier analysis from [4, Figure 1].

Our main result, presented below, validates a form of the **Order Increment Principle** in the U^3 case. Before presenting it, let us discuss an important aspect that arises in this setup. Note that, while Lemma 3.1 was "optimal" in the sense that it simultaneously minimizes the U^1 norm of f_r and the U^1-dual norm of f_s, in the U^2 case there is no uniqueness for a U^2-regularization operator $K : \mathbb{C}^{\mathrm{Z}} \to \mathbb{C}^{\mathrm{Z}}$. One possible choice is given by Lemma 2.1, but even in that case, there is a parameter ε that needs to be selected. As discussed in [4, Example 2.45], this operator (referred to as the *Fourier cut-off operator*) is not continuous and has properties that make calculations inconvenient.

Instead, we work with a similar operator which we call the *Fourier denoising operator* [4, (1)]. Fix a constant $\varepsilon > 0$, typically small, and let this operator be

$$K_\varepsilon(f) = \sum_{\{\chi \in \widehat{\mathrm{Z}} : |\widehat{f}(\chi)| \geq \varepsilon\}} \frac{|\widehat{f}(\chi)| - \varepsilon}{|\widehat{f}(\chi)|} \widehat{f}(\chi)\, \chi. \tag{1}$$

This operator reduces the absolute values of the Fourier coefficients using the function $x \mapsto \mathrm{ReLU}(x - \varepsilon)$, while preserving their phases, and crucially zeroes out those smaller than ε. Here by ReLU we mean the map $x \mapsto \frac{x+|x|}{2}$. It is shown that K_ε behaves as a contraction in L^2 space and exhibits other favorable properties, making it closely related to the cut-off operator, as outlined in [4, §2.5]. In the same vein as before, we define $\mathcal{K}_\varepsilon(f \otimes \overline{f})$ as the Z-matrix that results from applying K_ε to each of the Z-diagonals of $f \otimes \overline{f}$.

4 Main result

We can now state one of the main results in [4]. It gives a spectral U^3-regularity result, linking the eigenvalues and eigenvectors of $\mathcal{K}_\varepsilon(f \otimes \overline{f})$ with a quadratic decomposition of $f \in \mathbb{C}^{\mathrm{Z}}$. The theoretical foundation established by this result forms the basis for Algorithm 1 below.

Theorem 4.1. (Spectral U^3-regularization, [4, Theorem 1.1]) *For every $\rho_0 \in [0, 1)$ there exists $\varepsilon_0 > 0$ such that the following holds. Let Z be a finite abelian group and let $f : \mathrm{Z} \to \mathbb{C}$ be a 1-bounded function. Then there exists $\rho \in [\rho_0/2, \rho_0]$ and $\varepsilon \in [\varepsilon_0, 1]$ with the following property. Let f_{reg} be the projection of f to the linear span of the eigenspaces of $\mathcal{K}_\varepsilon(f \otimes \overline{f})$ with corresponding eigenvalues at least ρ. Then $\|f - f_{reg}\|_{U^3} \leq 2\rho^{3/8}$ and there exists $h : \mathrm{Z} \to \mathbb{C}$ such that $\|f_{reg} - h\|_2 \leq \rho$ and $\|h\|^*_{U^3} = O_\rho(1)$.*

The term f_{reg} from Theorem 4.1 can be interpreted as the term f_s of the regularity lemma plus a small L^2 error, see [4, Definition 2.23]. Such errors are common when providing regularity lemmas for the Gowers norms of order $k \geq 3$, see e.g. [9, Theorem 5.1].

Input: $f : \mathrm{Z} \to \mathbb{C}$, $\rho, \varepsilon \in \mathbb{R}_{>0}$
$M \leftarrow f(x) \otimes \overline{f(y)} \in \mathbb{C}^{\mathrm{Z} \times \mathrm{Z}}$
for $t \in \mathrm{Z}$ **do**
  $M(\cdot + t, \cdot) \leftarrow K_\varepsilon(M(\cdot + t, \cdot))$
end
$(\mu_1, v_1), \ldots, (\mu_{|\mathrm{Z}|}, v_{|\mathrm{Z}|}) \leftarrow$ **Eigendecomposition**(M)
Output: $f_{\mathrm{reg}} \leftarrow \sum_{\mu_i \geq \rho} \langle f, v_i \rangle v_i$

Algorithm 1: U^3-regularization algorithm ([4, Algorithm 1]).

Note that the loop in Algorithm 1 can be computed in parallel using the Fast Fourier Transform, thereby reducing the overall computational complexity of the algorithm. Similarly, we do not need to compute all the eigenvectors and eigenvalues of $\mathcal{K}_\varepsilon(f \otimes \overline{f})$, as, in the end, we will only be interested in those eigenspaces whose eigenvalues exceed ρ.

References

[1] O. A. Camarena, B. Szegedy, *Nilspaces, nilmanifolds and their morphisms*, (2010), arXiv:1009.3825.

[2] P. Candela, *Notes on nilspaces: algebraic aspects*, Discrete Anal. 2017, Paper No. 15, 59 pp.

[3] P. Candela, *Notes on compact nilspaces*, Discrete Anal. 2017, Paper No. 16, 57 pp.

[4] P. Candela, D. Gonzalez-Sanchez, B. Szegedy, *Spectral algorithms in higher-order Fourier analysis*, (2025), preprint arXiv:2501.12287.

[5] P. Candela, D. Gonzalez-Sanchez, B. Szegedy, *Free nilspaces, double coset nilspaces and Gowers norms*, (2023), arXiv:2305.11233.

[6] P. Candela, B. Szegedy, *Nilspace Factors for General Uniformity Seminorms, Cubic Exchangeability and Limits*, Mem. Amer. Math. Soc. **287** (2023), no. 1425, v+101 pp.

[7] P. Candela, B. Szegedy, *Regularity and inverse theorems for uniformity norms on compact abelian groups and nilmanifolds*, J. Reine Angew. Math. **789** (2022), 1–42.

[8] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal., **11** (2001), 465–588.

[9] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, Bull. Lond. Math. Soc. **42** (2010), no. 4, 573–606.

[10] W. T. Gowers, L. Milićević, *A quantitative inverse theorem for the U^4 norm over finite fields*, (2020), arXiv:1712.00241.

[11] B. Green, T. Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, in An Irregular Mind, Bolyai Soc. Math. Stud. 21, János Bolyai Mathematical Society, Budapest, 2010, pp. 261–334.

[12] B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), pp. 481–547.

[13] B. Green, T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), no. 2, 465–540.

[14] B. Green, T. Tao, T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372.

[15] Y. Gutman, F. Manners, P. P. Varjú, *The structure theory of nilspaces I*, J. Anal. Math. **140** (2020), no. 1, 299–369.

[16] Y. Gutman, F. Manners, P. P. Varjú, *The structure theory of nilspaces II: Representation as nilmanifolds*, Trans. Amer. Math. Soc. **371** (2019), no. 7, 4951–4992.

[17] Y. Gutman, F. Manners, P. P. Varjú, *The structure theory of nilspaces III: Inverse limit representations and topological dynamics*, Adv. Math. **365** (2020), 107059, 53 pp.

[18] H. Hatami, P. Hatami, S. Lovett, *Higher-order Fourier analysis and applications*, Found. Trends Theor. Comput. Sci. **13** (2019), no. 4, front matter, 247–448.

[19] B. Host, B. Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488.

[20] B. Host, B. Kra, *Parallelepipeds, nilpotent groups, and Gowers norms*, Bull. Soc. Math. France **136** (2008), no. 3, 405–437.

[21] D. Kim, A. Li, J. Tidor, *Cubic Goldreich-Levin*, Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 4846–4892.

[22] J. Leng, A. Sah, M. Sawhney, *Quasipolynomial bounds on the inverse theorem for the Gowers $U^{s+1}[N]$-norm*, (2024), arXiv:2402.17994.

[23] F. Manners, *Quantitative bounds in the inverse theorem for the Gowers U^{s+1}-norms over cyclic groups*, (2018), arXiv:1811.00718.

[24] B. Szegedy, *Higher order fourier analysis as an algebraic theory I*, (2009), arXiv:0903.0897.

[25] B. Szegedy, *Higher order fourier analysis as an algebraic theory II*, (2009), arXiv:0911.1157.

[26] B. Szegedy, *Higher order fourier analysis as an algebraic theory III*, (2010), arXiv:1001.4282.

[27] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[28] T. Tao, *A quantitative ergodic theory proof of Szemerédi's theorem*, Electron. J. Combin. **13** (2006), no. 1, Research Paper 99, 49 pp.

[29] T. Tao, V. Vu, *Additive combinatorics*, Cambridge Stud. Adv. Math., 105 Cambridge University Press, Cambridge, 2010. xviii+512 pp.

[30] T. Tao, T. Ziegler, *The inverse conjecture for the Gowers norm over finite fields via the correspondence principle*, Anal. PDE **3** (2010), no. 1, 1–20.

[31] L. Trevisan, *Additive combinatorics and theoretical computer science*, ACM SIGACT News (2) **40** (2009), 50–66.

[32] M. Tulsiani, J. Wolf, *Quadratic Goldreich–Levin Theorems*, SIAM J. Comput. **43** (2014), no. 2, 730–766.

[33] T. Ziegler, *Universal Characteristic Factors and Furstenberg Averages*, J. Amer. Math. Soc. **20** (2007), 53–97.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Shattering Triples with Six Permutations

(Extended abstract)

Daniela Černá* Bartłomiej Kielak† Jan Volec‡

Abstract

Given six permutations $(\pi_1, \pi_2, \dots, \pi_6)$ of the same ground set $[n]$ and a subset $X \subseteq [n]$ of size 3, we say that the triple X is shattered by $(\pi_1, \pi_2, \dots, \pi_6)$ if every relative order of X is induced by exactly one of the π_i's. A natural extremal question is to determine the maximum number of triples of $[n]$ that six permutations can shatter. Using the flag algebra method, we prove that no six-tuple shatters more than $\frac{1}{2}\binom{n}{3} + O(n^2)$ triples. On the other hand, for every n, we construct six permutations of $[n]$ that shatter at least $\frac{482}{975}\binom{n}{3} \sim 0.4944\binom{n}{3}$ triples. These results improve the previously known bounds of Johnson and Wickes.

1 Introduction

The concept of shattering sets appears in different parts of combinatorics, and it plays a central role in the Vapnik-Chervonenkis theory for statistical learning [VC71, Vap00]. A finite set X is *shattered* by a family $\mathcal{F}$ if the number of different intersections of X with elements of $\mathcal{F}$ is $2^{|X|}$. A natural extremal set theory question is, given integers n and k, to determine the size of the smallest family of subsets of $[n]$ such that every k-element subset of $[n]$ is shattered. In 1973, Kleitman and Spencer [KS73] proved that for any fixed k this number is between $d_1 \cdot 2^k \log n$ and $d_2 \cdot k2^k \log n$ for some positive reals d_1 and d_2.

Shattering has also been studied in the setting of permutations. A k-element subset $X \subseteq [n]$ is shattered by a family of permutations $\mathcal{S} \subseteq S_n$ if the number of different relative orderings of X by the elements of $\mathcal{S}$ is $k!$. In 1971, Spencer [Spe71] showed the existence of a family $\mathcal{S} \subseteq S_n$ of size $\frac{k}{\log_2(k!/(k!-1))} \log_2 n$ such that every k-element subset of $[n]$ is shattered.

*Department of Theoretical Computer Science, Faculty of Information Technology, and Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Trojanova 13, 120 00 Prague 2, Czech Republic. E-mail: `opocedan@cvut.cz`. Supported by the grant 23-06815M of the Grant Agency of the Czech Republic.

†Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, Germany. Previous affiliation: Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic.

‡Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, Prague, 160 00, Czech Republic. E-mail: `jan@ucw.cz`. Supported by the grant 23-06815M of the Grant Agency of the Czech Republic.

On the other hand, Füredi [Fü96] proved that such a family must have size at least $\frac{(k-1)!}{2}\log_2 n$. Radhakrishnan [Rad03] improved the bound of Füredi by a factor of approximately $2/\log_2 e$, and Tarui [Tar08] gave an explicit construction for the case $k = 3$ of size $(2 + o(1))\log_2 n$.

In this work, we focus on a multiplicity version of this problem, where we consider families $\mathcal{S} \subseteq S_n$ of size $k!$ and the aim is to maximize the number of shattered k-element subsets of $[n]$. For the case $k = 3$, i.e., when we shatter triples of $[n]$ with 6 permutations, Johnson and Wickes [JW23] proved the existence of permutations of $[n]$ that shatter at least $\frac{17}{42}\binom{n}{3}$ triples. On the other hand, they proved that if $n \geq 8$ then no $\mathcal{S} \subseteq S_n$ of size 6 can shatter more than $\frac{47}{60}\binom{n}{3}$ triples. Our results are the following improvements of both of the aforementioned bounds.

Theorem 1. *Every six permutations of $[n]$ shatter at most $\frac{1}{2}\binom{n}{3} + O(n^2)$ triples.*

Proposition 2. *For every $n \geq 3$, there exist six permutations of $[n]$ that shatter at least $\frac{482}{975}\binom{n}{3}$ triples.*

2 Definitions and notation

We start with the definition of a permutation pattern. Let $\pi \in S_n$ be a permutation, and $X = \{x_1, x_2, \ldots, x_k\} \subseteq [n]$ such that $x_1 < x_2 < \ldots < x_k$. The *pattern in π induced by X* is the permutation $\tau \in S_k$ satisfying

$$\tau(i) < \tau(j) \Leftrightarrow \pi(x_i) < \pi(x_j)\,.$$

It follows that a family $\mathcal{S} \subseteq S_n$ shatters a k-element subset $X \subseteq [n]$ if and only if

$$\{\tau \text{ pattern in } \pi \text{ induced by } X, \pi \in \mathcal{S}\} = S_k\,.$$

For $\mathcal{S} \subseteq S_n$, we define $F_k(\mathcal{S})$ to be the number of k-element subsets of $[n]$ shattered by $\mathcal{S}$, and for every $n \geq k$, we set

$$F_k(n) := \max_{\substack{\mathcal{S}\subseteq S_n \\ |\mathcal{S}|=k!}} F_k(\mathcal{S}) \quad \text{and} \quad f_k(n) := F_k(n) \Big/ \binom{n}{k}\,.$$

A standard averaging argument yields for every fixed k and every $n \geq k$ that the function $f_k(n)$ is non-increasing, thus the limit $\lim_{n\to\infty} f_k(n)$ exists. For brevity, we write

$$c_k := \lim_{n\to\infty} f_k(n)\,.$$

Remark. *We note that our definition of the pattern in π induced by X differs from the one in [JW23], and if $\mathcal{S} = (\pi_1, \ldots, \pi_{k!})$ is a family that attains $F_k(n)$ in our definition, then $\mathcal{S}' = (\pi_1^{-1}, \ldots, \pi_{k!}^{-1})$ is the corresponding solution in their setting. In particular, the extremal questions are the same.*

3 Proof of Theorem 1

We use a versatile framework of Razborov [Raz07] called flag algebras to show that $c_3 \leq \frac{1}{2}$. The method is robust enough to allow finding an upper bound on c_3 in a straightforward way, however, let us now mention one simple observation that allows reducing the size of the calculations: We may assume that the first permutation in a six-tuple is always the identity. Indeed, the number of shattered triples is the same for $(\pi_1, \pi_2, \ldots, \pi_6)$ and $\left(\pi_1 \circ \pi_1^{-1}, \pi_2 \circ \pi_1^{-1}, \ldots, \pi_6 \circ \pi_1^{-1}\right)$. Therefore, we use flag algebras with five-tuples of permutations, and count the number of triples ordered by those permutations in all the five possible ways that are not monotone increasing.

We apply flag algebras to convergent sequences of five-tuples of permutations, each tuple having permutations of the same order n. The substructure counting is given by picking k elements from $[n]$ uniformly at random, and in each permutation considering the probability distribution on the $k!$ patterns of order k. Let $\mathcal{F}_3$ be the set all of possible outcomes for $k = 3$. Clearly $|\mathcal{F}_3| = (3!)^5 = 7776$. Next, let $\mathcal{H} \subseteq \mathcal{F}_3$ be the set of all 120 five-tuples corresponding to the different orderings of $S_3 \setminus \{\mathrm{id}\}$. Note that the normalized count of shattered triples is equal to the sum of the densities in $\mathcal{H}$.

Our main technical result is the following:

Proposition 3. *There are pattern-density expressions $a_1, a_2, \ldots, a_5$ such that*

$$\left[\!\!\left[\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix}^T \begin{pmatrix} 3 & -1 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 & 1 \\ -1 & -1 & 3 & 1 & -1 \\ -1 & -1 & 1 & 3 & -1 \\ -1 & 1 & -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} \right]\!\!\right] = \sum_{F \in \mathcal{F}_3} d_F \cdot F\,,$$

where $d_F = -1$ for every $F \in \mathcal{H}$ and $d_F \in \left\{-\frac{1}{3}, 0, \frac{1}{3}, 1\right\}$ for every $F \in \mathcal{F} \setminus \mathcal{H}$.

Since the 5×5 matrix is positive definite and $d_F \leq 1$ for all $\mathcal{F} \setminus \mathcal{H}$, we derive the following asymptotic inequality for five-tuples of permutations:

$$\sum_{F \in \mathcal{F}_3 \setminus \mathcal{H}} F - \sum_{H \in \mathcal{H}} H \geq 0\,.$$

In particular, $c_3 \leq \frac{1}{2}$. Moreover, when applying the established "limit" proof to a tuple of permutations on $[n]$, all the involved inequalities are valid up to an error of the order $O\left(\frac{1}{n}\right)$. Therefore, $F_3(n) \leq \frac{1}{2}\binom{n}{3} + O\left(n^2\right)$.

4 Proof of Proposition 2

By monotonicity of $f_3(n)$, it is enough to show that $c_3 \geq \frac{482}{975}$. Firstly, suppose we are able to construct six permutations $(\pi_1, \pi_2, \ldots, \pi_6)$ of $[t]$, for some fixed value of t, that shatter $\alpha\binom{t}{3}$ triples. For each $i \in [6]$, set M_i to be the permutation matrix of π_i, and for a given integer n, let $M_i^{\otimes n}$ denote the n-th Kronecker power of M_i. Now consider the sequence of families of six permutations $\left(\pi_1^{(n)}, \pi_2^{(n)}, \ldots, \pi_6^{(n)}\right)$, where $\pi_i^{(n)}$ is the permutation corresponding to the matrix $M_i^{\otimes n}$. Note that $\pi_i^{(n)}$ is sometimes called the n-th iterated blow-up of π_i. A quick calculation reveals (see also [JW23, Theorem 3.2]) that

$$\lim_{n \to \infty} \frac{F_3\left(\pi_1^{(n)}, \pi_2^{(n)}, \ldots, \pi_6^{(n)}\right)}{\binom{n}{3}} = \alpha \cdot \frac{t-2}{t+1}\,.$$

n	$F_3(n)$	$f_3(n)$	lower bound on c_3
5	8	$\frac{4}{5}$	$\frac{2}{5} \sim 0.4$
6	16	$\frac{4}{5}$	$\frac{16}{35} \sim 0.4571$
7	26	$\frac{26}{35}$	$\frac{13}{28} \sim 0.4643$
8	40	$\frac{5}{7}$	$\frac{10}{21} \sim 0.4762$
9	57	$\frac{19}{28}$	$\frac{19}{40} \sim 0.475$

Table 1: The extremal values for $5 \leq n \leq 9$ and the resulting bound on c_3.

$$\begin{aligned}
\pi_1 &= (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26)\\
\pi_2 &= (3, 2, 1, 22, 23, 21, 25, 24, 26, 11, 10, 17, 16, 15, 14, 13, 12, 20, 19, 18, 4, 6, 5, 8, 9, 7)\\
\pi_3 &= (19, 20, 18, 22, 21, 23, 9, 8, 7, 14, 17, 12, 16, 10, 13, 11, 15, 2, 1, 3, 5, 4, 6, 26, 25, 24)\\
\pi_4 &= (19, 18, 20, 6, 5, 4, 25, 26, 24, 17, 14, 15, 11, 16, 12, 13, 10, 2, 3, 1, 23, 22, 21, 8, 7, 9)\\
\pi_5 &= (18, 21, 20, 16, 11, 14, 6, 10, 7, 4, 25, 24, 2, 23, 1, 26, 3, 17, 22, 19, 13, 15, 12, 9, 5, 8)\\
\pi_6 &= (21, 18, 19, 11, 16, 13, 9, 5, 8, 25, 4, 1, 23, 3, 26, 2, 24, 22, 17, 20, 15, 12, 14, 6, 10, 7)
\end{aligned}$$

Table 2: Six permutations of order 26 that shatter 1446 triples out of 2600.

In particular, it holds that $c_3 \geq f_3(t) \cdot \frac{t-2}{t+1}$ for every $t \geq 3$.

We determined the extremal values for $n \leq 8$ as well as all extremal configurations using an exhaustive computer generation. Moreover, a simple double-counting argument yields that any configuration for $n = 9$ with at least 58 triples contains at least one subconfiguration for $n' = 8$ with either 39 or 40 shattered triples. Since it possible to generate all the configurations for n' with at least 39 shattered triples, inspecting all their possible extensions by one point reveals that $F_3(9) = 57$. These results, together with the corresponding lower bound on c_3, are in Table 1.

Using a heuristic computer search for larger permutations, we have found six permutations of order 26 that shatter 1446 triples out of 2600; see Table 2. In particular, it holds that $f_3(26) \geq \frac{723}{1300}$ and hence $c_3 \geq \frac{482}{975} \sim 0.4944$.

5 Conclusion

In this work, we have shown that $c_3 \in \left[\frac{482}{975}, \frac{1}{2}\right]$, and we conjecture that the upper bound is correct. Apart from our suspicion that asymptotically optimal constructions for $f_3(n)$ are not iterated blow-ups of some finite construction, we have also done some additional (larger) flag algebra computations and they were not able to prove $c_3 < \frac{1}{2}$.

Conjecture 1. *For every n, there exist six permutations of $[n]$ that shatter at least $\frac{1}{2}\binom{n}{3}$ triples.*

We believe it would be interesting to have a better understanding of c_k for $k \geq 4$. Using the result of Spencer [Spe71] and the monotonicity of $f_k(n)$, Johnson and Wickes [JW23] observed that $c_k < 1$. On the other hand, Levenshtein [Lev92] constructed $k!$ permutations of $[k+1]$ that shatter all the k-element subsets; the proof can be also found in [Wic23]. Therefore, the iterated blow-up of the Levenshtein construction yields $c_k \geq \frac{k!}{(k+1)^{k-1}-1}$. So-far, we are not able to say anything substantially better than these two bounds.

Acknowledgements. The third author is very grateful to J. Balogh, A. Clifton, D. Kráľ, B. Lidický, F. Pfender and S. Spiro for many fruitful discussions on the topic.

References

[Fü96] Z. Füredi. Scrambling permutations and entropy of hypergraphs. *Random Structures and Algorithms*, 8(2):97–104, 1996.

[JW23] R. Johnson and B. Wickes. Shattering k-sets with permutations. *Order*, 41(3):419–436, 2023.

[KS73] D. Kleitman and J. Spencer. Families of k-independent sets. *Discrete Mathematics*, 6(3):255–262, 1973.

[Lev92] V. I. Levenshtein. Perfect codes in the metric of deletions and insertions. *Discrete Mathematics and Applications*, 2(3):241–258, 1992.

[Rad03] J. Radhakrishnan. A note on scrambling permutations. *Random Structures and Algorithms*, 22(4):435–439, 2003.

[Raz07] Alexander A. Razborov. Flag algebras. *Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[Spe71] J. Spencer. Minimal scrambling sets of simple orders. *Acta Mathematica Academiae Scientiarum Hungaricae*, 22(3–4):349–353, 1971.

[Tar08] J. Tarui. On the minimum number of completely 3-scrambling permutations. *Discrete Mathematics*, 308(8):1350–1354, 2008.

[Vap00] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer New York, 2000.

[VC71] V. Vapnik and A. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications*, 16(2):264–280, 1971.

[Wic23] B. Wickes. *Covering and Separation for Permutations and Graphs*. PhD thesis, Queen Mary University of London, 2023.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Almost Full Transversals in Equi-n-squares

(Extended abstract)

Debsoumya Chakraborti* Micha Christoph† Zach Hunter‡
Richard Montgomery* Teo Petrov*

Abstract

In 1975, Stein made a wide generalisation of the Ryser-Brualdi-Stein conjecture on transversals in Latin squares, conjecturing that every equi-n-square (an $n \times n$ array filled with n symbols where each symbol appears exactly n times) has a transversal of size $n-1$. That is, it should have a collection of $n-1$ entries that share no row, column, or symbol. In 2017, Aharoni, Berger, Kotlar, and Ziv showed that equi-n-squares always have a transversal with size at least $2n/3$. In 2019, Pokrovskiy and Sudakov disproved Stein's conjecture by constructing equi-n-squares without a transversal of size $n - \frac{\log n}{42}$, but asked whether Stein's conjecture is approximately true. I.e., does an equi-n-square always have a transversal with size $(1-o(1))n$?

We answer this question in the positive. More specifically, we improve both known bounds, showing that there exist equi-n-squares with no transversal of size $n - \Omega(\sqrt{n})$ and that every equi-n-square contains $n - n^{1-\Omega(1)}$ disjoint transversals of size $n - n^{1-\Omega(1)}$.

1 Overview

A *Latin square of order* n is an $n \times n$ array filled with n symbols where each symbol appears exactly once in each row and each column. A *transversal* in the array consists of a collection of cells in the array, no two of which share a row, column or symbol. The *size* of a transversal is its number of cells. In an $n \times n$ array, a transversal of size n is also referred to as a *full transversal*, whereas a smaller transversal is sometimes referred to as a *partial transversal*. The study of Latin squares and transversals dates back to Euler [6] who studied the decomposition of Latin squares into disjoint full transversals. As was known to Euler, a Latin square may have no full transversal.

*Mathematics Institute, University of Warwick, Coventry, United Kingdom.
E-mail: {debsoumya.chakraborti, richard.montgomery, teodor.petrov}@warwick.ac.uk
DC and RM supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978). TP supported by the Warwick Mathematics Institute Centre for Doctoral Training, and by funding from the UK EPSRC (Grant number: EP/W524645/1).
†Department of Computer Science, ETH Zürich, Switzerland.
E-mail: micha.christoph@inf.ethz.ch
Supported by the SNSF Ambizione Grant No. 216071.
‡Department of Mathematics, ETH Zürich, Switzerland.
E-mail: zach.hunter@ifor.math.ethz.ch

However, the prominent Ryser-Brualdi-Stein conjecture [12, 4], originating in the 1960's, suggests that every Latin square of order n should have a transversal of size $n-1$ and, moreover, one of size n if n is odd. This conjecture has seen a lot of activity recently, culminating in the proof by Montgomery [8] that, when n is sufficiently large, every Latin square of order n has a transversal of size $n-1$. For more on this, related results, and the history of the study of Latin squares, see the recent surveys [10, 9].

In 1975, Stein [13] made a series of bold conjectures that, broadly, suggest the Latin square conditions conjectured to guarantee a transversal of order $n-1$ may be overkill. In particular, he conjectured that any *equi-n-square* has a transversal of size $n-1$, where an equi-n-square is an $n \times n$ array filled with n symbols where each symbol appears exactly n times. Thus, a Latin square of order n is an equi-n-square where we additionally require every symbol to appear at most once in each row or column. As some evidence towards his conjecture, Stein [13] used the probabilistic method to show that any equi-n-square contains a transversal of size at least $(1-e^{-1})n$. This bound was the state of the art for some 40 years, until Aharoni, Berger, Kotlar, and Ziv [1] used topological methods to show that any equi-n-square contains a transversal of size at least $2n/3$. Very recently, Anastos and Morris [3] showed that any equi-n-square contains a transversal of size at least $(3/4-o(1))n$. However, in 2019, Pokrovskiy and Sudakov [11] had shown that Stein's conjecture is, indeed, over-ambitious. That is, they constructed equi-n-squares that have no transversal with size larger than $n - \frac{1}{42}\log n$.

While this settles the falsity of Stein's conjecture, the bounds $n - O(\log n)$ and $(3/4-o(1))n$ on the size of the largest transversal that can be guaranteed in any equi-n-square are rather far apart. In particular, Pokrovskiy and Sudakov [11] (see also [10, Problem 4.3]) asked whether Stein's conjecture holds asymptotically, i.e., does an equi-n-square always have a transversal with $(1-o(1))n$ cells? Our main result is to confirm that it not only does, but that it can moreover be almost decomposed into such transversals, with the following result.

Theorem 1.1. *There exists $\varepsilon > 0$ such that every equi-n-square contains at least $n - n^{1-\varepsilon}$ disjoint transversals with size at least $n - n^{1-\varepsilon}$.*

We also modify Pokrovskiy and Sudakov's construction from [11] to show that a related but simpler construction provides equi-n-squares that must omit far more symbols in any transversal, as follows.

Theorem 1.2. *For each $n \in \mathbb{N}$, there is an equi-n-square with no transversal of size $n - \left(\frac{1}{2\sqrt{2}} + o(1)\right)\sqrt{n}$.*

The structure of this extended abstract is as follows. We begin with a proof sketch of the upper bound in Section 2. In Section 3 we give a high-level overview of the proof idea of the lower bound. In section 4 we go more in depth, discussing how to obtain a single large transversal. In particular, in Section 4.2 we consider a specific example, which serves well to demonstrate our main technique for proving Theorem 1.1. For complete details, see [5].

2 Upper bound

In this section we provide brief justification for Theorem 1.2. Let us consider the special case when n is equal to twice a square, i.e. when $n = 2m^2$ for some $m \in \mathbb{N}$. We need to firstly construct an equi-n-square S and then secondly argue it has no transversal of size larger than $n - \Omega(\sqrt{n})$. For convenience, we will refer to the symbols of S as colours instead. The construction of S is as follows. We subdivide the $n \times n$ array into smaller *boxes* of size $m \times m$ and then pair the boxes

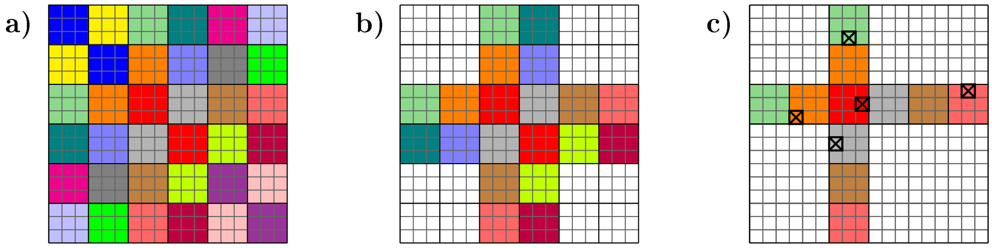


Figure 1: **a)** Our construction in the simplified case when $n = 2m^2$ with $m = 3$. **b)** The boxes comprising the *double-cross* $C_3 \cup C_4$ in the same equi-n-square. **c)** The boxes comprising C_3 with a transversal highlighted – if a red cell appears in C_3 in a transversal then at most 4 of the 5 colours appearing only in C_3 can appear. Thus, at least one colour from $C_3 \cup C_4$ is omitted in any transversal.

and assign the same colour to all cells in both boxes in a single pair, noting that this gives rise to an equi-n-square as a pair of boxes contain exactly $2m^2 = n$ cells. The pairing is done in such a way that resembles a symmetric matrix, as is portrayed for the specific case $m = 3$ in Figure 1a).

After writing down what the equi-n-square S is, as we have just done, the next step of showing that any transversal must have size at most $n - \Omega(\sqrt{n})$ is almost the same as in the proof by Pokrovskiy and Sudakov [11]. In particular, we define a *cross* C_k of S to be the k^{th} box along the leading diagonal together with all the boxes above, below, to the left and to the right of it. The cross C_3 is portrayed in Figure 1c). The symmetry of S implies that the colours appearing in $C_{2k-1} \cup C_{2k}$ do not appear elsewhere in S. Next, we show that any transversal T of S must miss out on a colour from the *double-cross* $C_{2k-1} \cup C_{2k}$. This is because either T contains no cell in the two diagonal boxes, in which case it misses their colour, or T does contain such a cell, say in C_3, as in Figure 1c). So suppose that $k = 2$ and that the colour in this diagonal box is red, so that our description matches the figure. Then if we consider the cross C_3, once again due to the symmetry of S, all non-red colours in C_3 do not appear elsewhere in S and there are $2m - 1$ of them. Then T may contain at most $m - 1$ cells in the *horizontal* boxes, as the red cell of T blocks a row, and T may contain at most $m - 1$ cells in the *vertical* boxes, as once again, the red cell of T blocks a column. It follows that T misses out on a colour from the double-cross $C_{2k-1} \cup C_{2k}$, as claimed above.

Finally, summing over all double-crosses $C_{2k-1} \cup C_{2k}$ and noting that any colour appears in at most two double-crosses, shows that any transversal T of S can have size at most $|T| \leq n - m/2$. The general case can be proved by a slightly modified construction.

3 Lower bound

The data of an equi-n-square S can be encoded into a graph or a hypergraph and this can be particularly useful as then graph theoretic arguments can be used. For example, let $G = K_{n,n}$ be the complete balanced bipartite graph on $2n$ vertices, where the vertex parts correspond the the rows and columns on S. We edge-colour G, by letting $\chi(ij) = s$ where the symbol in the i^{th}-row and j^{th}-column of S is s. In this formulation, a transversal in S corresponds to a rainbow matching in G. Alternatively, one can model S by an n-regular, 3-partite, 3-uniform, $3n$-vertex hypergraph $\mathcal{H}$ with vertex classes $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{A}$ corresponding to the rows, columns and symbols respectively, where $ijs \in \mathcal{H}$ if and only if the symbol in the i^{th}-row and j^{th}-column of S is s. In this setting, a transversal of S corresponds to a matching in $\mathcal{H}$. This latter interpretation is the one we use when proving Theorem 1.1. In particular, the proof idea is as follows. Let S be an equi-n-square and let $\mathcal{H}$ be its auxiliary n-regular, 3-partite, 3-uniform, $3n$-vertex hypergraph as above. We show that one can get rid of a few hyperedges in $\mathcal{H}$ (which corresponds to deleting a few of the entries in S) so that the resultant hypergraph $\mathcal{H}'$ has chromatic index* $\chi'(\mathcal{H}')$ only slightly more than n. As $\mathcal{H}'$ has almost n^2 hyperedges, almost all colour classes in any edge-colouring, using $\chi'(\mathcal{H}')$ colours, must have size almost n. In particular, in $\mathcal{H}'$ we obtain $(1-o(1))n$ disjoint matchings of size $(1-o(1))n$ each. As each matching in $\mathcal{H}'$ corresponds to a transversal in S and as they are all disjoint, this gives us $(1-o(1))n$ disjoint transversals in S of size $(1-o(1))n$ each, as desired.

Unsurprisingly then, the bulk of the argument is in finding a very large subhypergraph of $\mathcal{H}$, whose chromatic index is almost n. As we have just alluded to, we prove a slightly more general theorem about hypergraphs than Theorem 1.1. Let us consider a more detailed exposition between the interplay of the arrays and the auxiliary hypergraphs.

Let us begin with a Latin square L. Then the auxiliary hypergraph $\mathcal{H}(L)$ is an n-regular, 3-partite, 3-uniform, $3n$-vertex hypergraph with vertex parts $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{A}$ corresponding to the sets of rows, columns and symbols of L respectively, where $ija \in \mathcal{H}(L)$ if and only if the the symbol in the i^{th} row and the j^{th} column is a. Since L is a Latin square, $\mathcal{H}(L)$ has the additional property of having codegree at most one. In particular, any two vertices x, y in the same vertex part have codegree zero, and any two vertices x, y in different parts have codegree exactly one. We know that in this case [8] we can guarantee that $\mathcal{H}(L)$ has a matching of size $n-1$ when n is large. To move away from Latin squares to the more general case of equi-n-squares means that we lose the codegree conditions between vertices in $\mathcal{I}$ and $\mathcal{A}$ and also between $\mathcal{J}$ and $\mathcal{A}$, meaning we only have one third of the codegree conditions to work with relative to the Latin square case. Note that it would not be possible to part with all of the codegree conditions, as in general there exist k-regular 3-partite, 3-uniform and $3k$-vertex hypergraphs $\mathcal{H}$ with no matching of size larger than than $8k/9$, as can be shown by adapting a construction of Alon and Kim [2] (for details, see [5]). To obtain a bound on the chromatic index of a hypergraph we make use of the following theorem by Molloy and Reed [7].

Theorem 3.1. *For all k there is a constant C_k such that any k-uniform hypergraph of maximum codegree B and maximum degree Δ has list chromatic index at most* $\left(1 + C_k(B/\Delta)^{1/k}(\log(\Delta/B))^4\right)\Delta$.

Using this, we are able to prove the following hypergraph generalisation of Theorem 1.1.

Theorem 3.2. *There are some $\xi, \eta > 0$ and D_0 such that the following holds for each $D \geq D_0$. Let $\mathcal{H}$ be a 3-uniform hypergraph with maximum degree D and at least $\sqrt{D}$ vertices. Suppose that the graph on $V(\mathcal{H})$ with edges xy present if $\mathrm{cod}_{\mathcal{H}}(x,y) \geq D^{1-\eta}$ is bipartite.*

Then, there is some $\mathcal{H}' \subset \mathcal{H}$ with $e(\mathcal{H}') \geq e(\mathcal{H}) - |\mathcal{H}| \cdot D^{1-\xi}$ and

$$\chi'(\mathcal{H}') \leq (1 + D^{-\xi})D.$$

We note that from Theorem 3.2, one could deduce a more general result than Theorem 1.1. In particular, it allows for an $n \times n$ array A to have n copies of each symbol, n symbols in each row and each column, but additionally it allows for A to contain up to $n^{1-\eta}$ symbols in the same entry, provided they are all distinct.

*The chromatic index of a hypergraph $\mathcal{H}$ is the minimum number of colours needed to colour the edges of $\mathcal{H}$ so that edges of the same colour are vertex-disjoint.

4 One large transversal in equi-n-squares

In this section, we will aim to provide some high-level intuition on why one can expect any equi-n-square to have a single almost-full transversal.

4.1 Low codegree

Since a Latin square L of order n is just an equi-n-square, in which every symbol can appear at most once in any row or column, and since we know that Latin squares always admit a transversal of size $n-1$, it would be natural to transition from the Latin square case to the more general case by allowing symbols to appear at most some bounded number of times in any row or column. Indeed, this case can be easily dealt with via the Rödl nibble, as has been pointed out by Pokrovskiy and Sudakov [11], in the following sense.

Theorem 4.1. *Let S be an equi-n-square in which each symbol appears at most $o(n)$ times in every row and column. Then S has a transversal of size $(1-o(1))n$.*

This shows that Stein's conjecture is asymptotically true whenever the equi-n-square S has 'low codegree'. In order to demonstrate our main proof technique, we consider the following toy example which concerns a certain 'high codegree' case.

4.2 High codegree: a toy example

Suppose that $n = 4m$ and that each colour appears in any column some multiple of $n/4$ times, as is portrayed in Figure 2a). Let us partition the cells of S into a set of *column-blocks* $\mathcal{B}$, where each $B \in \mathcal{B}$ contains $n/4$ cells, all of the same colour and in the same column. We would like an argument which tells us which block to use from which column. This can be done in the following way. Let G be the auxiliary 4-regular bipartite multigraph with vertex classes corresponding to the columns and colours and whose edges are indexed by the blocks $B \in \mathcal{B}$, as depicted in Figure 2b). We can now extract a perfect matching from G, via Hall's condition say, which tells us which block to use for each column. The issue is that it could be the case that all blocks appear in the first $n/4$ rows of S, which would be problematic as the largest transversal obtainable now would be of size $n/4$. To fix this, we aim for a random matching in G, at the cost of it being only an *almost*-perfect matching, as is portrayed in Figure 2d).

We now highlight the importance of having such a random matching M. Let H be the auxiliary bipartite graph with vertex classes corresponding to the rows and columns of S, whose edges correspond to the cells of the blocks in M. Note that since colours have already been matched with columns by M, any matching in H is automatically rainbow. Additionally, since M is random, H should be 'close' to being $n/4$-regular (we expect each row in S to see about $n/4$ entries coming from blocks in M), in which case a defect version of Hall's theorem gives us an almost-perfect matching in H, which translates to an almost-full transversal in S, as desired.

To find the desired random almost-perfect matching M in G we do the following. Firstly, find 4 disjoint perfect matchings M_1, M_2, M_3 and M_4 in G. This can be done, for example, since G is bipartite and regular so we may find one perfect-matching, take it out and proceed by induction. Secondly, we mix M_1 and M_2 together to obtain a new matching M_1' and also M_3 and M_4 to obtain M_2'. We then mix the two new matchings to obtain our final random almost-perfect matching M. The way we mix two matchings is to take their union (which is now a disjoint union of cycles and paths) and for each connected component we take the 'even' or 'odd' edges uniformly at random. If one component is too large, say if $M_1 \cup M_2$ were a Hamiltonian cycle, then there would be too much dependence (half the time we get back M_1 and the other half we get M_2,

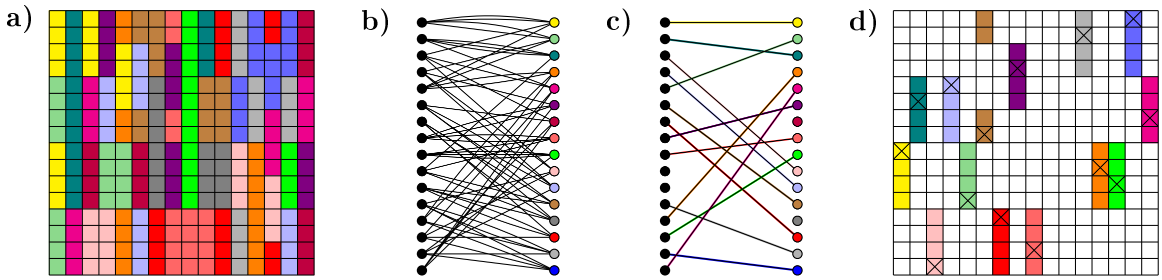


Figure 2: **a)** A subsquare S in which each column can be decomposed into 4 blocks with size 4 which have the same colour (for ease of visualisation, cells of each colour are often drawn next to each other in natural blocks, but this is not required in general). **b)** The multigraph K corresponding to S. **c)** A matching M in K. **d)** The subsquare $S' \subset S$ corresponding to the matching M. Any set of cells which share no row or column must then share no colour and hence be a transversal; an example is marked by crosses.

neither of which were random). To fix this, we break up the large components into smaller ones by deleting a few edges. This step of possibly deleting a few edges is what ensures randomness but also what causes us to lose a few edges as to only obtain an almost-perfect matching.

4.3 General situation

In the general case of equi-n-squares one would have to simultaneously accommodate for column-blocks, row-blocks and instances of low codegree. The way we get around this is to form 'dummy' blocks, so that each cell is contained in exactly 3 blocks. This, along with the bounded-dependence random matching algorithm from the last subsection, helps us find a partially filled $n \times n$ array with an appropriate regularity condition, as well as an 'all or small' codegree condition. We can then convert the 'all' codegrees to 'small' codegrees, say by introducing new colours, and employ the Rödl nibble to find an almost-full transversal. The full details can be found in [5].

5 Concluding remarks

We have shown that the size of a largest transversal that can be guaranteed in any equi-n-square is between

$$n - n^{1-\varepsilon} \quad \text{and} \quad n - \left(\frac{1}{2\sqrt{2}} + o(1)\right)\sqrt{n}.$$

Therefore, we suggest the following to replace Stein's disproved conjecture.

Conjecture 5.1. *There exists a constant $C > 0$ such that every equi-n-square has a transversal with size at least $n - C\sqrt{n}$.*

Furthermore, as our construction in the proof of Theorem 1.2 is very natural, it is also reasonable to suggest that the constant $\frac{1}{2\sqrt{2}}$ is optimal.

References

[1] R. Aharoni, E. Berger, D. Kotlar, and R. Ziv. On a conjecture of Stein. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 87, pages 203–211. Springer, 2017.

[2] N. Alon and J. H. Kim. On the degree, size, and chromatic index of a uniform hypergraph. *Journal of Combinatorial Theory, Series A*, 77(1):165–170, 1997.

[3] M. Anastos and P. Morris. A note on finding large transversals efficiently. *arXiv preprint: 2412.05891*, 2024.

[4] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory.* Cambridge University Press, 1991.

[5] D. Chakraborti, M. Christoph, Z. Hunter, R. Montgomery, and T. Petrov. Almost-full transversals in equi-n-squares. *arXiv preprint arXiv:2412.07733*, 2024.

[6] L. Euler. Recherches sur un nouvelle espéce de quarrés magiques. *Verhandelingen uitgegeven door het zeeuwsch Genootschap der Wetenschappen te Vlissingen*, pages 85–239, 1782.

[7] M. Molloy and B. Reed. Near-optimal list colorings. *Random Structures & Algorithms*, 17(3-4):376–402, 2000.

[8] R. Montgomery. A proof of the Ryser-Brualdi-Stein conjecture for large even n. *arXiv preprint arXiv:2310.19779*, 2023.

[9] R. Montgomery. Transversals in Latin squares. page 131–158, 2024.

[10] A. Pokrovskiy. *Rainbow Subgraphs and their Applications*, page 191–214. London Mathematical Society Lecture Note Series. Cambridge University Press, 2022.

[11] A. Pokrovskiy and B. Sudakov. A counterexample to Stein's equi-n-square conjecture. *Proceedings of the American Mathematical Society*, 147(6):2281–2287, 2019.

[12] H. Ryser. Neuere Probleme der Kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, pages 69–91, 1967.

[13] S. K. Stein. Transversals of Latin squares and their generalizations. 1975.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A Freeable Matrix Characterization of Bipartite Graphs of Ferrers Dimension Three

Parinya Chalermsook [*] Ly Orgo [†] Minoo Zarsav [‡]

Abstract

Ferrers dimension, along with the order dimension, is a standard dimensional concept for bipartite graphs. In this paper, we prove that a graph is of Ferrers dimension three (equivalent to the intersection bigraph of orthants and points in $\mathbb{R}^3$) if and only if it admits a biadjacency matrix representation that does not contain $\Gamma = \begin{pmatrix} * & 1 & * \\ 1 & 0 & 1 \\ 0 & 1 & * \end{pmatrix}$ and $\Delta = \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 1 & 0 & 1 \end{pmatrix}$, where $*$ denotes a zero or one entry.

1 Introduction

Special graph classes (or structured graphs) have played crucial roles in combinatorics and computer science, arising in many application domains. For instance, in the area of algorithm designs, special graph classes have been used, in the past decades, to model real-world data sets. These graphs are often equipped with geometric, topological, or combinatorial properties that can be computationally leveraged.

In this paper, we consider the class of bipartite graphs that have Ferrers dimension three, denoted by CHAIN^3 (see Section 2 for the definition of Ferrers dimension). In a geometric language, it is equivalent to the intersection bigraph of **3D orthants** and **points** in $\mathbb{R}^3$: A bipartite graph $G = (A \cup B, E)$ is of Ferrers dimension three if and only if each vertex $a \in A$ is associated with a point $p_a \in \mathbb{R}^3$ and each vertex $b \in B$ with an orthant $O_b = (-\infty, x_b) \times (-\infty, y_b) \times (-\infty, z_b)$ such that $(a, b) \in E$ if and only if $p_a \in O_b$. The class of 3D orthants arise naturally in data structures (see, e.g., [2]). Ferrers dimension is a standard dimensionality concept closely related to order dimension and interval dimension [7]. Well-known classes of bipartite graphs (e.g., orthogonal ray graphs [3] and grid intersection graphs [8]) have been shown to have constant Ferrers dimension.

It is common for a graph class to admit several equivalent characterizations. The main contribution of this paper is a characterization of CHAIN^3 in terms of a freeable matrix property. Let us start by defining the terminologies. Let P be a 0/1 matrix. We say that matrix M **contains** P if a submatrix of M is equal to P; otherwise when M does not contain P, we say that M is P-**free**. When we allow the entries of P to be $\{0, 1, *\}$, a star can represent

[*]The University of Sheffield, UK. E-mail: `chalermsook@gmail.com`.
[†]Aalto University, Finland. E-mail: `ly.orgo@aalto.fi`.
[‡]Aalto University, Finland. E-mail: `minoo.zarsav@aalto.fi`.

either zero or one, so M is P-free when it is free of all matrices P' obtained by replacing each $*$ of P by either zero or one. A graph G is said to be P-**freeable** if there exists a biadjacency matrix representation of G that is P-free.

Many (geometric) bipartite graph classes are known to admit both intersection bigraph representation and freeable matrix characterization. Table 2 summarizes the existing results for graph classes in the context of our work.

Graph classes	Freeable Matrix	References
CHAIN	$\begin{pmatrix}0 & 1\end{pmatrix}$	[5]
BPG	$\begin{pmatrix}1 & 0\\ * & 1\end{pmatrix}$ $\begin{pmatrix}1 & *\\ 0 & 1\end{pmatrix}$	[4]
CONV	$\begin{pmatrix}1 & 0 & 1\end{pmatrix}$	[1]
CHAIN^2	$\begin{pmatrix}1 & *\\ 0 & 1\end{pmatrix}$	[10]
Chordal Bipartite	$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$	[9]
Stick Graphs	$\begin{pmatrix}* & 1 & *\\ 1 & 0 & 1\end{pmatrix}$ $\begin{pmatrix}1 & *\\ 0 & 1\\ 1 & *\end{pmatrix}$ $\begin{pmatrix}* & 1 & *\\ * & 0 & 1\\ 1 & * & *\end{pmatrix}$	[6]
Segment Ray	$\begin{pmatrix}* & 1 & *\\ 1 & 0 & 1\end{pmatrix}$	[3]
Grid Intersection (GIG)	$\begin{pmatrix}* & 1 & *\\ 1 & 0 & 1\\ * & 1 & *\end{pmatrix}$	[8]
CHAIN^3	$\begin{pmatrix}* & 1 & *\\ 1 & 0 & 1\\ 0 & 1 & *\end{pmatrix}$ $\begin{pmatrix}1 & * & *\\ 0 & 1 & *\\ 1 & 0 & 1\end{pmatrix}$	this paper

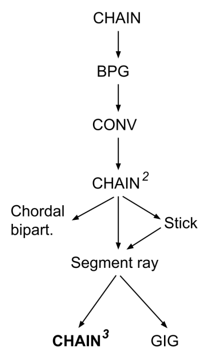


Table 2: Freeable matrix characterizations and (right-hand-side) relations between the corresponding graph classes. The arrow from graph class X to Y denotes that X is a subclass of Y. The graph class CHAIN^d contains all graphs of Ferrers dimension d.

Theorem 1. *A bipartite graph G has Ferrers dimension three if and only if it is Γ and Δ-freeable where* $\Gamma = \begin{pmatrix}* & 1 & *\\ 1 & 0 & 1\\ 0 & 1 & *\end{pmatrix}$ *and* $\Delta = \begin{pmatrix}1 & * & *\\ 0 & 1 & *\\ 1 & 0 & 1\end{pmatrix}$

2 Preliminaries

A bipartite graph $G = (U \cup V, E)$ is a **chain graph** (CHAIN), if its vertices can be linearly ordered as $U = \{u_1, u_2, \ldots, u_{|U|}\}$ and $V = \{v_1, v_2, \ldots, v_{|V|}\}$ so that we have the chains of neighbors, $N(u_1) \subseteq N(u_2) \subseteq \ldots \subseteq N(u_{|U|})$ and $N(v_1) \subseteq N(v_2) \subseteq \ldots \subseteq N(v_{|V|})$. Chain graphs are also called difference graphs, Ferrers bigraphs, and induced $2K_2$-free graphs [8]. The graph class is exactly the intersection bigraph of rays and points in $\mathbb{R}$ (that is, it is representable as points and rays on a real line, such that there is an edge $\{u, v\}$ for some $u \in U$, $v \in V$ if and only if the ray representing v contains the point representing u [3].) Chain graphs are $\begin{pmatrix}1 & 0\end{pmatrix}$-freeable graphs, and their complements are chain graphs [5].

For two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ on the same set of vertices, the *intersection* $G_1 \cap G_2$ is defined as $(V, E_1 \cap E_2)$. We use CHAIN^d to denote the class of all graphs defined by the intersection $\bigcap_{i \in [d]} G_i$, where each $G_i \in \mathsf{CHAIN}$. **Ferrers dimension** of a bipartite graph G is the minimum d, such that $G \in \mathsf{CHAIN}^d$ [8].

Given a matrix A, we write $A[i, j]$ to denote the entry at the i-th row and j-th column. Let A and B be two matrices of the same size. The *Hadamard product* of A and B, denoted by $A \odot B$, is defined by the entry-wise product of the corresponding entries. Note that for two unweighed graphs G_1, G_2 on the same vertex set, the Hadamard product of their adjacency matrices is an adjacency matrix of $G_1 \cap G_2$.

3 Proof of Theorem 1

3.1 “Only if” direction

Let $G = G_1 \cap G_2 \cap G_3$ where each $G_i \in \mathsf{CHAIN}$. For $i \in [3]$, let L_i be an arbitrary linear order of $G_i = (U \cup V, E_i)$, such that for all $u \in U$ and $v \in V$, $u <_{L_i} v$ if and only if $\{u, v\} \in E_i$. Let $L_i[S]$ be the linear order L_i restricted only to the vertices in S. Let A, and $\forall i \in [3]$, A_i be the adjacency matrices of G and G_i. There exists an ordering of rows $L_1[U]$ such that A_1 is $(\begin{smallmatrix}1\\0\end{smallmatrix})$-free. Similarly, Let $L_2[V]$ be a column ordering, such that A_2 is $(\begin{smallmatrix}0 & 1\end{smallmatrix})$-free. We consider all matrices and their Hadamard product $A = A_1 \odot A_2 \odot A_3$ in these row and column orders.

If there is a 0 in $A_1 \odot A_2$, then that 0 must exist in A_1 or A_2. If it exists in A_1, the cells above the 0 in the same column must also be 0. Otherwise, the cells to the right of the 0 in the same row must be 0. It follows that $A_1 \odot A_2$ is $D = (\begin{smallmatrix}1 & *\\0 & 1\end{smallmatrix})$-free.

Note that if there is a 1 in A, then that 1 must also exist in A_1, A_2, and A_3. Assume by contradiction that A contains Γ. Let the indices of these columns in A be $j_1 < j_2 < j_3$ and the indices of the rows be $i_1 < i_2 < i_3$. Since the entries corresponding to indices i_1, i_2, j_2, and j_3 form the submatrix D, which is forbidden in $A_1 \odot A_2$ (yellow highlight on the left in fig. 1), and $A_1 \odot A_2$ must include the 1-s of the submatrix, then $A_1[i_2, j_2]$ and $A_2[i_2, j_2]$ cannot be 0, and therefore $A_3[i_2, j_2] = 0$. Similarly, considering i_2, i_3, j_1, and j_2 implies that $A_3[i_3, j_1] = 0$ (blue highlight on the left in fig. 1). Therefore, the submatrix $F = (\begin{smallmatrix}1 & 0\\0 & 1\end{smallmatrix})$ formed by i_2, i_3, j_1, and j_2 must be entirely in A_3. This is a contradiction because a chain graph cannot contain an induced $2K_2$ [8]. The fact that A avoids Δ can be argued similarly (on the right on fig. 1).

$$\begin{array}{c|ccc|} & j_1 & j_2 & j_3 \\ \hline i_1 & * & 1 & * \\ i_2 & 1 & 0 & 1 \\ i_3 & 0 & 1 & * \\ \hline \end{array} \qquad\qquad \begin{array}{c|ccc|} & j_1 & j_2 & j_3 \\ \hline i_1 & 1 & * & * \\ i_2 & 0 & 1 & * \\ i_3 & 1 & 0 & 1 \\ \hline \end{array}$$

Figure 1: $A_1 \odot A_2$ cannot contain D as a submatrix of Γ or Δ on the left and right respectively.

3.2 “if” Direction

Given a graph $G = (U \cup V, E)$ with an adjacency matrix A that does not contain the submatrices Γ and Δ, we will show that $G \in \mathsf{CHAIN}^3$.

Lemma 1. *Given a graph $G = (U \cup V, E)$ with an adjacency matrix A, that avoids Γ or Δ as a submatrix, there exists a graph $G_{1,2} \in \mathsf{CHAIN}^2$ with an adjacency matrix $A_{1,2}$ such that*

a. $A_{1,2}$ is free from the submarix $(\begin{smallmatrix}1 & *\\0 & 1\end{smallmatrix})$,

b. For all i and j, if $A[i, j] = 1$ then $A_{1,2}[i, j] = 1$,

c. Define $\widetilde{A}$ by replacing each 0-entry in A with $0'$ if the corresponding entry of $A_{1,2}$ is zero; and with 0^* otherwise (so the entries in $\widetilde{A}$ are in $\{1, 0', 0^*\}$.) Then every 0^* in $\widetilde{A}$ is part of some submatrix $\left(\begin{smallmatrix}1 & *\\0^* & 1\end{smallmatrix}\right)$ of $\widetilde{A}$.

Proof. Let the orderings of the rows and columns of A correspond to the order of the nodes in U and V, respectively. We construct graph $G_1 \in \mathsf{CHAIN}$ by describing its point-ray representation: $V = \{v_1, \ldots v_{|V|}\}$ corresponds to points $P^{(1)} = \{p_1, \ldots, p_{|V|}\}$ placed from left to right; U corresponds to the leftward rays $R^{(1)} = \{r_1, \ldots, r_{|V|}\}$. For each $i \in [|V|]$, place the starting point of r_i between p_j and p_{j+1} (if it exists), where j is the maximum integer for which $A[i,j] = 1$. Analogously, we construct $G_2 \in \mathsf{CHAIN}$; let $U = \{u_1, \ldots u_{|U|}\}$ correspond to points $P^{(2)} = \{p'_1, \ldots, p'_{|U|}\}$ placed from top to bottom; V correspond to the downward rays, where each ray r_j starts between p'_{i-1}(if it exists) and p'_i so that i is the least index where $A[i,j] = 1$. Let $A_{1,2}$ be the adjacency matrix of $G_1 \cap G_2$ such that the row and column orders are the same as in A. Since the adjacency matrices of G_1 and G_2 are respectively $\begin{pmatrix} 0 & 1 \end{pmatrix}$- and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$-free, then $A_{1,2}$ is free from submatrix $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ which proves point a of lemma 1.

Now, let $A[i,j] = 1$ for some fixed $i \in [|U|]$ and $j \in [|V|]$. Let j' be the highest index such that $A[i,j'] = 1$. Since $A[i,j] = 1$ then, $j \le j'$ and by definition of leftward rays, we have $A_1[i,j] = 1$. Similarly, for a fixed j, let i' be the least index such that $A[i',j] = 1$. Since $A[i,j] = 1$ then, $i' \le i$, and because of the downward rays, we have $A_2[i,j] = 1$ therefore $A_{1,2}[i,j] = 1$. This proves point b.

To prove c, note that every 0 in $A_{1,2}$ ($0'$ in A) must have 0 in either all the columns to the right of it or in all rows above it because of point a. Let us consider a 0^* in A. Based on the definition, the entry at its position in $A_{1,2}$ must be a 1, and there must exist 1-s above and to the right of it, otherwise, it would be a $0'$. $\square$

Lemma 2. *The matrix $\widetilde{A}$ does not include any of the following matrices as a submatrix:* $\begin{pmatrix} 0^* & 1 \\ 1 & 0^* \end{pmatrix}$, $\begin{pmatrix} 1 & 0^* \\ 0^* & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & * \\ 0' & 1 \end{pmatrix}$, $\begin{pmatrix} 0^* & * \\ 0' & 0^* \end{pmatrix}$

Proof. If $\widetilde{A}$ includes $\begin{pmatrix} 1 & 0^* \\ 0 & 1 \end{pmatrix}$ as a submatrix, then according to Lemma 1, there exists a 1-entry to the right and above the 0^*. These 1-entries with the submatrix $\begin{pmatrix} 1 & 0^* \\ 0 & 1 \end{pmatrix}$, would form the submatrix Γ, which is a contradiction. This means $\widetilde{A}$ is $\begin{pmatrix} 1 & 0^* \\ 0 & 1 \end{pmatrix}$-free and therefore, $\begin{pmatrix} 1 & 0^* \\ 0^* & 1 \end{pmatrix}$-free. With similar reasoning, $\widetilde{A}$ is $\begin{pmatrix} 0^* & 1 \\ 1 & 0^* \end{pmatrix}$-free, or otherwise, it forms the submatrix Δ.

Since $A_{1,2}$ is $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$-free and $0'$, 0^* are defined to be the 0-entries in $\widetilde{A}$, that are 0 and 1 in $A_{1,2}$ respectively, then $\widetilde{A}$ is free of $\begin{pmatrix} 1 & * \\ 0' & 1 \end{pmatrix}$, $\begin{pmatrix} 0^* & * \\ 0' & 0^* \end{pmatrix}$ as well. $\square$

Next, we show an algorithm to sort the columns to exclude $\begin{pmatrix} 1 & 0^* \end{pmatrix}$ as a submatrix. This ordering will give us A_3, such that $A = A_{1,2} \odot A_3$, since 0^* are the zeroes in $\widetilde{A}$ not in $A_{1,2}$

Algorithm 1: Reordering the columns of $\widetilde{A}$

$L_3 \leftarrow \emptyset$
for *cols* $j \leftarrow 1$ **to** $|V|$ **do**
  Add j to the end of L_3
  Let $S_j = \{i : \widetilde{A}[i,j] = 0^*\}$
  if $S_j \neq \emptyset$ **then**
    Let k_j be the least index in L_3 such that $\widetilde{A}[i,k_j] = 1$ for some $i \in S_j$
    In L_3 move j to be right before k_j

Lemma 3. *The adjacency matrix $\widetilde{A}$ with columns sorted according to L_3 is $\begin{pmatrix} 1 & 0^* \end{pmatrix}$-free.*

Proof. We prove it by induction on j. We show that after processing the j-th column, the matrix $\widetilde{A}$ with column indices restricted to $[j]$ in L_3, denoted by A^j, is $\begin{pmatrix} 1 & 0^* \end{pmatrix}$-free.

The statement holds for $j = 1$. Assuming that it is true, for $j - 1$, we show that it holds for j. If $S_j = \emptyset$, the j-th column doesn't move, and the statement holds. So, we consider the case that $S_j \neq \emptyset$. Let L be the order of L_3 after processing j, and suppose by contradiction that A^j contains $\begin{pmatrix} 1 & 0^* \end{pmatrix}$ in the i-th row. Then, there must exist a column $\ell \in [j-1]$, such that $\widetilde{A}[i,j] = 1$, $\widetilde{A}[i,\ell] = 0^*$, and after processing j, the j-th column has been moved somewhere to the left of the ℓ-th column ($j <_L \ell$). Let $i_{min} \in S_j$ be a row such that $\widetilde{A}[i_{min}, j] = 0^*$ and $\widetilde{A}[i_{min}, k_j] = 1$. We have $k_j \leq_L \ell$; otherwise, j and ℓ do not form $\begin{pmatrix} 1 & 0^* \end{pmatrix}$ in the i-th row (Figure 2).

Note that $k_j \neq \ell$, because otherwise columns ℓ, j and rows i, i_{min} form the forbidden submatrix $\begin{pmatrix} 0^* & 1 \\ 1 & 0^* \end{pmatrix}$ or $\begin{pmatrix} 1 & 0^* \\ 0^* & 1 \end{pmatrix}$ depending on the original column and row orders. With the same reasoning, $\widetilde{A}[i_{min}, \ell] \neq 1$, $\widetilde{A}[i, k_j] \neq 0^*$. Additionally, $\widetilde{A}[i_{min}, \ell] = 0^*$ is also a contradiction since by induction hypothesis, A^{j-1} is $\begin{pmatrix} 1 & 0^* \end{pmatrix}$-free, but, row i_{min} with columns k_j, ℓ would form $\begin{pmatrix} 1 & 0^* \end{pmatrix}$. Similarly, We can justify $\widetilde{A}[i, k_j] \neq 1$.

The values that remain possible are $\widetilde{A}[i, k_j] = \widetilde{A}[i_{min}, \ell] = 0'$. Since $\ell, k_j \in [j-1]$, then in the original column ordering ℓ and k_j precede j. We consider two possible cases in the original ordering, either $i < i_{min}$ or $i_{min} < i$. If $i < i_{min}$ then, columns ℓ, j form a forbidden submatrix (Lemma 2), and if $i_{min} < i$, the columns k_j, j do (Figure 3). Both cases lead to a contradiction; the induction statement holds for j. $\square$

Figure 2: On the left, we have j yet to be inserted into the order, and on the right, the order L is depicted.

Figure 3: Depicted are the forbidden submatrices with row orders $i < i_{min}$ and $i_{min} < i$ on the left and right respectively.

Lemma 4. *There exists a chain graph $G_3 \in \mathsf{CHAIN}$ with the adjaciency matrix A_3 (and its extension $\widetilde{A}_3$) based on L_3, such that $A = A_3 \odot A_{1,2}$.*

Proof. Consider the ordering L_3 over V in Algorithm 1, we construct the graph G_3 by ordering points P according to L_3, and for each $u \in U$, we have a rightward ray right before the point corresponding to $v \in N(u)$ that is earliest in L_3 (leftmost 1 on each row). Since we can construct G_3 as an intersection graph of points and rays, then $G_3 \in \mathsf{CHAIN}$. Let A_3 be the adjacency matrix of G_3.

Now, we will show that $A = A_3 \odot A_{1,2}$. According to the construction of G_3, if an entry in A is one, is it in A_3 as well. According to lemma 1, it is also one in $A_{1,2}$. Next, let us consider the 0-s. By definition, each entry $0'$ in $\widetilde{A}$ is zero in $A_{1,2}$ and therefore also in $A_3 \odot A_{1,2}$. If an entry is 0^* in $\widetilde{A}$, then it is 0 in A_3, because the ordering L_3 is $\begin{pmatrix} 1 & 0^* \end{pmatrix}$-free (Lemma 3) and in the construction of G_3, the ray corresponding to each row starts right before the first 1 on that row, leaving out all 0^*-s. We conclude that $A = A_3 \odot A_{1,2}$. $\square$

In summary, we have constructed $G_{1,2} \in \mathsf{CHAIN}^2$ and $G_3 \in \mathsf{CHAIN}$ such that their adjacency matrices satisfy $A = A_{1,2} \odot A_3$. This implies that $G \in \mathsf{CHAIN}^3$, concluding the proof of the "if" direction.

References

[1] A Brandstfidt. Special graph classes a survey. Technical report, Tech. Report SM-DU-199, Universit˜ it Duisburg, 1991.

[2] Timothy M Chan, Kasper Green Larsen, and Mihai Pătraşcu. Orthogonal range searching on the ram, revisited. In *Proceedings of the twenty-seventh annual symposium on Computational geometry*, pages 1–10, 2011.

[3] Steven Chaplick, Pavol Hell, Yota Otachi, Toshiki Saitoh, and Ryuhei Uehara. Intersection dimension of bipartite graphs. In *Theory and Applications of Models of Computation: 11th Annual Conference, TAMC 2014, Chennai, India, April 11-13, 2014. Proceedings 11*, pages 323–340. Springer, 2014.

[4] Lin Chen and Yaacov Yesha. Efficient parallel algorithms for bipartite permutation graphs. *Networks*, 23(1):29–39, 1993.

[5] Sandip Das, M Sen, AB Roy, and Douglas B West. Interval digraphs: An analogue of interval graphs. *Journal of Graph Theory*, 13(2):189–202, 1989.

[6] Felice De Luca, Md Iqbal Hossain, Stephen Kobourov, Anna Lubiw, and Debajyoti Mondal. Recognition and drawing of stick graphs. *Theoretical Computer Science*, 796:22–33, 2019.

[7] Stefan Felsner, Michel Habib, and Rolf H Möhring. On the interplay between interval dimension and dimension. *SIAM Journal on Discrete Mathematics*, 7(1):32–40, 1994.

[8] I Ben-Arroyo Hartman, Ilan Newman, and Ran Ziv. On grid intersection graphs. *Discrete Mathematics*, 87(1):41–52, 1991.

[9] Bettina Klinz, Rüdiger Rudolf, and Gerhard J Woeginger. *Permuting matrices to avoid forbidden submatrices*. Inst. für Math., Techn. Univ. Graz, 1992.

[10] Anish Man Singh Shrestha, Satoshi Tayu, and Shuichi Ueno. On orthogonal ray graphs. *Discrete Applied Mathematics*, 158(15):1650–1659, 2010.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

An Exact Ore-Type condition for Hamilton cycles in oriented graphs

(Extended abstract)

Yulin Chang* Yangyang Cheng† Tianjiao Dai ‡
Qiancheng Ouyang § Guanghui Wang ¶

Abstract

We show that every sufficiently large oriented graph D of order n with $\deg^+(x)+\deg^-(y) \geq \lceil(3n-3)/4\rceil$ whenever D does not have an edge from x to y contains a directed Hamilton cycle. This is best possible and solves a problem of Kelly, Kühn and Osthus from 2012. Our main result generalizes the result of Keevash, Kühn and Osthus [6] and improves the asymptotic bound proved by Kelly, Kühn and Osthus [7].

1 Introduction

Given a fixed graph H, let G be a graph of the same order as H. One may ask under which conditions G has a subgraph that is isomorphic to H. Such problems have received

*Mathematical Institute, Ocean University of China. Email: `ylchang@ouc.edu.cn`. Partially supported by Natural Science Foundation of China (12201352) and Natural Science Foundation of Shandong Province (ZR2022QA083).

†Faculty of Computer Science and Mathematics, University of Passau, Passau, Germany. Email: `yangyang.cheng@uni-passau.de`.

‡School of Mathematics, Shandong University, Jinan, 250100, China. Email: `tianjiao.dai@sdu.edu.cn`. Supported by Natural Science Foundation of Shandong Province(ZR2024QA174)

§Data Science Institute, Shandong University, Jinan, 250100, China. Email: `oyqc@sdu.edu.cn`. Supported by Natural Science Foundation of Shandong Province(ZR2024QA107)

¶School of Mathematics, Shandong University, Jinan, 250100, China. Email: `ghwang@sdu.edu.cn`. Supported by the National Key Research and Development Program (2023YFA1009603) and the Natural Science Foundation of China (12231018).

significant attention throughout the history of graph theory. For most H, deciding whether G contains H as a subgraph is very challenging. For example, when H is a Hamiltonian cycle, this decision problem has been shown to be NP-complete and is, in fact, one of the prototypes of NP-complete problems in complexity theory. Thus, it is interesting to seek sufficient conditions that guarantee a Hamiltonian cycle or even a more general H.

The famous Dirac's theorem [3] asserts that every graph G of order at least 3 with minimum degree $\delta(G) \geq |G|/2$ contains a Hamiltonian cycle. Ore [10] extended this result by showing that every graph G with $|G| \geq 3$ such that for every $x, y \in V(G)$ with $xy \notin E(G)$, $d(x) + d(y) \geq |G|$, also contains a Hamiltonian cycle. A further generalization, due to Chvátal [2], states that if the degree sequence of a graph G is $d_1 \leq d_2 \leq \cdots \leq d_n$, and if $n \geq 3$ and $d_i \geq i+1$ or $d_{n-i} \geq n-i$ for all $i > n/2$, then G contains a Hamiltonian cycle. The degree sequence condition given by Chvátal is best possible in the sense that, for any degree sequence $d_1 \leq d_2 \leq \cdots \leq d_n$ that violates Chvátal's condition, one can always construct a graph H with degree sequence $d'_1 \leq d'_2 \leq \cdots \leq d'_n$ such that $d'_i \geq d_i$ for all $i \in [n]$, which does not contain a Hamiltonian cycle.

Given a digraph D and a vertex $v \in V(D)$, let $d^+(v)$ ($d^-(v)$) denote the number of out-edges (in-edges) incident to v. Define $\delta^+(D) = \min\{d^+(v) \mid v \in V(D)\}$ and $\delta^-(D) = \min\{d^-(v) \mid v \in V(D)\}$. Let $\delta^0(D) = \min\{\delta^+(D), \delta^-(D)\}$. The Hamiltonian cycle problem was generalized to digraphs by Ghouila-Houri [4]:

Theorem 1.1 (Ghouila-Houri [4]). *Every strongly connected digraph D on n vertices with $\delta^+(D) + \delta^-(D) \geq n$ contains a directed Hamilton cycle. In particular, every digraph with $\delta^0(D) \geq n/2$ contains a directed Hamilton cycle.*

For Ore-type conditions, Woodall [12] extended Ore's theorem to the digraph setting:

Theorem 1.2 (Woodall [12]). *Every strongly connected digraph D on n vertices with $d^+(x) + d^-(y) \geq n$ for every pair $x \neq y$ and $xy \notin A(D)$ contains a directed Hamilton cycle.*

However, proving a Chvátal-type condition for digraphs is significantly more challenging. A long-standing conjecture by Nash-Williams proposes a Chvátal-type condition for Hamiltonian cycles (see [9]):

Conjecture 1.3 (Nash-Williams [9]). *Suppose that D is a strongly connected digraph on $n \geq 3$ vertices such that for all $i < n/2$,*

- $d^+_i \geq i+1$ *or* $d^-_{n-i} \geq n-i$;
- $d^-_i \geq i+1$ *or* $d^+_{n-i} \geq n-i$;

then D contains a directed Hamilton cycle.

In [1], Christofides et al. proved an asymptotic version of the Nash-Williams conjecture, but the full conjecture remains open.

An oriented graph is a digraph that contains no 2-cycles, i.e., there is at most one arc between any two vertices. Keevash et al. [6] generalized Dirac's theorem to oriented graphs, proving that there exists an n_0 such that for all $n \geq n_0$, every oriented graph D of order n with $\delta^0(D) \geq \frac{3n-4}{8}$ contains a directed Hamiltonian cycle. Kelly et al. [7] proved an Ore-type condition for oriented graphs: for every $\alpha > 0$, there exists an integer $n_0 = n_0(\alpha)$ such that every oriented graph D of order $n \geq n_0$ satisfying $d^+(x) + d^-(y) \geq (3/4 + \alpha)n$ for each x, y with $xy \notin A(D)$ contains a directed Hamiltonian cycle. Kühn and Osthus [8] proposed the following problem as a tight version of the previous result:

Conjecture 1.4 (Kühn and Osthus [8])**.** *Every oriented graph D of order n satisfying $d^+(x)+d^-(y) \geq \lceil (3n-3)/4 \rceil$ for each x, y with $xy \notin A(D)$ contains a directed Hamiltonian cycle.*

If true, this conjecture would generalize the result in [6]. The main result of our work is to give a full solution to this conjecture when n is sufficiently large:

Theorem 1.5. *There exists an $n_0 > 0$ such that for all $n \geq n_0$, every oriented graph D of order n satisfying $d^+(x) + d^-(y) \geq \lceil (3n-3)/4 \rceil$ for each x, y with $xy \notin A(D)$ contains a directed Hamiltonian cycle.*

2 Proof Strategy

Our proof uses the absorption technique first introduced by Rödl, Szemerédi, and Ruciński [11], as well as stability analysis for the extremal case. First, we define the extremal graph family $\mathcal{F}$ such that each oriented graph D in $\mathcal{F}$ satisfies the property that $d^+(x) + d^-(y) \geq \lceil (3|D|-3)/4 \rceil - 1$ for every pair x, y with $xy \notin A(D)$, and D contains no directed Hamilton cycle. For an appropriate parameter $\delta \in (0, 1)$, we say that an oriented graph D of order n is *δ-extremal* if, by adding or deleting at most δn^2 edges of D, we obtain a graph in $\mathcal{F}$. If D is not *δ-extremal*, then we say D is *δ-stable*.

We first prove that for every δ-stable oriented graph D of order n such that $d^+(x) + d^-(y) \geq \lceil (3n-3)/4 \rceil - 1$ for every pair x, y with $xy \notin A(D)$, the graph D contains a directed Hamilton cycle, where δ is very small and n is sufficiently large compared to $1/\delta$. In this step, we adapt the idea of the absorption technique by building a small absorption structure $\mathcal{A}$ that can "absorb" vertices. Formally, $\mathcal{A}$ consists of disjoint directed paths of length 1 or 3, designed for different purposes, such that we can insert a small set of vertices into $\mathcal{A}$ to transform it into a collection of disjoint directed paths of slightly larger size, while preserving the number of directed paths. The total size of $\mathcal{A}$ is small compared to n, and when we insert the vertices, we do not change the endpoints of the directed paths in $\mathcal{A}$.

Next, we prove a connecting lemma that enables us to join the disjoint paths in $\mathcal{A}$ into a single directed path P_{ab}. The path P_{ab} remains small compared to n, and for every vertex set U that is relatively small compared to $|P_{ab}|$, we can absorb U into P_{ab}, thereby obtaining a larger directed path P' with $V(P') = V(P_{ab}) \cup U$.

In addition to the absorbing path P_{ab}, we also construct a *reservoir set* R, which consists of a collection of disjoint short directed paths and has size much smaller than $|P_{ab}|$. The set R has the property that for every pair of vertices u and v, there exist many disjoint directed paths $P_1, \dots, P_t$ of bounded length within R such that uP_iv is a directed path for each $1 \leq i \leq t$. This allows us to use disjoint short paths from R to connect a collection of disjoint paths into a longer directed path or a directed cycle.

Once the absorbing path P_{ab} and the reservoir set R have been constructed, the remaining task is to cover the vertices in $V(D) \setminus (V(P_{ab}) \cup V(R))$ by disjoint directed paths such that their total number is not too large. In this step, let $D' = D - (V(P_{ab}) \cup V(R))$. We apply the degree form of the Diregularity Lemma to D' and obtain a reduced oriented graph D^{red}. Note that D^{red} has a slightly weaker degree condition than D, but still remains close.

We then show that D^{red} contains a 1-factor; otherwise, D would be δ-extremal. Let F be the 1-factor we find. The final step is to use the Blow-up Lemma to find disjoint directed paths corresponding to F, such that the total number of paths is relatively small. We then use the reservoir set R to connect these directed paths and P_{ab} into a large directed cycle C that is almost Hamiltonian. Note that R may still contain some remaining vertices. The final step is to absorb these vertices into P_{ab}, thereby transforming C into a directed Hamiltonian cycle.

The covering part of our proof is standard and does not require new ideas. Here, we introduce some novel techniques used in constructing the absorbing path P_{ab}. Specifically, we develop a *double-step absorption* method. To illustrate this, we define the following two kinds of absorbing properties:

Definition 2.1. A *strong absorber* of a vertex u (or a pair of distinct vertices (u, v)) is an edge $wz \in E(G)$ such that $wu, uz(vz) \in E(G)$. A vertex u (or a pair of distinct vertices (u, v)) is α_1-*strongly absorbable* if it has at least $\alpha_1 n^2$ strong absorber.

Definition 2.2. An α_1-*weak absorber* of vertex u is a pair of disjoint edges $(ww', z'z)$ such that $wu, uz \in E(G)$ and (w', z') is α_1-strongly absorbable. A vertex u is (α_2, α_1)-*weakly absorbable* if it has at least $\alpha_2 n^4$ of α_1-weak absorber.

Formally, the absorbing structure $\mathcal{A}$ contains two kinds of directed path families, $\mathcal{A}_1$ and $\mathcal{A}_2$, constructed using probabilistic methods. Here, $\mathcal{A}_1$ is a set of disjoint edges designed to absorb vertices (or pairs) that are *strongly absorbable*, while $\mathcal{A}_2$ is used for those that are *weakly absorbable*.

If a vertex v (or a pair (u, v), where u, v are the endpoints of a directed path P with $V(P) \cap V(P_{ab}) = \emptyset$) is strongly absorbable, then the path P_{ab} can absorb it directly by replacing some edge wz in $\mathcal{A}_1$ with wvz (or with wPz), resulting in a new path P' such that $V(P') = V(P_{ab}) \cup \{v\}$ (or $V(P') = V(P_{ab}) \cup V(P)$). This process uses only edges from $\mathcal{A}_1$, and each absorption consumes exactly one edge from $\mathcal{A}_1$.

However, our Ore-type degree condition is not strong enough to guarantee that all vertices are strongly absorbable. The following lemma shows that if a vertex or path is not strongly absorbable, then it is weakly absorbable:

Lemma 2.3. *Let $0 < \alpha_1, \alpha_2 \ll \delta \ll 1$. Let G be an oriented graph satisfying the Ore-type condition. If G is δ-stable, then every vertex $v \in V(G)$ is either α_1-strongly absorbable or (α_2, α_1)-weakly absorbable.*

By this lemma, for any vertex v, if it is not strongly absorbable, then it must be weakly absorbable. Thus, we design $\mathcal{A}_2$ to be a set of disjoint directed paths of bounded length with the form $ww'Pz'z$, where (w', z') is strongly absorbable, which serve as weak absorbers for those vertices that are weakly absorbable.

In practice, if a vertex v is weakly absorbable, we can absorb it into P_{ab} by replacing a path $ww'Pz'z$ from $\mathcal{A}_2$ with wvz, obtaining a new path P_1 such that

$$V(P_1) = V(P_{ab}) \cup \{v\} \setminus (\{w', z'\} \cup \{V(P)\}),$$

where the pair $\{w', z'\}$ is strongly absorbable by definition. Therefore, we can absorb the path $w'z'$ into P_1 again by replacing one edge uv from $\mathcal{A}_1$ by $uw'Pz'v$, resulting in a path P_2 with

$$V(P_2) = V(P_1) \cup \{w', z'\} \cup V(P).$$

The absorbing process described here can be repeated many times, allowing us to absorb a linear number of vertices, provided that their total number is relatively small compared to the size of P_{ab}. According to the strategy outlined earlier, this eventually yields a directed Hamiltonian cycle once we obtain a directed cycle containing P_{ab} with order close to n.

To finish the proof, it remains to consider the case when D is δ-extremal. This means that D is close to an extremal construction. The proof in this case is long and involved, requiring detailed structural analysis, and thus we omit the sketch here.

3 Remarks

It is interesting to seek more applications of the technique we introduced here. Given an oriented graph D, for each $v \in V(D)$, let $d(v) = d^+(v) + d^-(v)$. Define $\delta(D) = \min\{d(v) \mid v \in V(D)\}$ and $\delta^*(D) = \delta(D) + \delta^+(D) + \delta^-(D)$. Häggkvist [5] conjectured the following degree sum condition for Hamiltonian cycles in oriented graphs:

Conjecture 3.1. *Every oriented graph D on n vertices with $\delta^*(D) > (3n-3)/2$ contains a Hamiltonian cycle.*

This conjecture has been confirmed asymptotically in [7]. We believe our methods could be helpful in proving a stable version of the above conjecture, offering hope for a full resolution of the conjecture for large graphs.

References

[1] Demetres Christofides, Peter Keevash, Daniela Kühn, and Deryk Osthus. A semiexact degree condition for Hamilton cycles in digraphs. *SIAM Journal on Discrete Mathematics*, 24(3):709–756, 2010.

[2] Václav Chvátal. On Hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[3] Gabriel Andrew Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[4] Alain Ghouila-houri. Une condition suffisante dexistence dun circuit hamiltonien. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences*, 251(4):495–497, 1960.

[5] Roland Häggkvist. Hamilton cycles in oriented graphs. *Combinatorics, Probability and Computing*, 2(1):25–32, 1993.

[6] Peter Keevash, Daniela Kühn, and Deryk Osthus. An exact minimum degree condition for Hamilton cycles in oriented graphs. *Journal of the London Mathematical Society*, 79(1):144–166, 2009.

[7] Luke Kelly, Daniela Kühn, and Deryk Osthus. A Dirac-type result on Hamilton cycles in oriented graphs. *Combinatorics, Probability and Computing*, 17(5):689–709, 2008.

[8] Daniela Kühn and Deryk Osthus. A survey on Hamilton cycles in directed graphs. *European Journal of Combinatorics*, 33(5):750–766, 2012.

[9] Crispin St John Alvah Nash-Williams. Hamiltonian circuits. *Studies in Math*, 12:301–360, 1975.

[10] Oystein Ore. A note on Hamiltonian circuits. *American Mathematical Monthly*, 67:55, 1960.

[11] Vojtěch Rödl, Endre Szemerédi, and Andrzej Ruciński. An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.

[12] Douglas R Woodall. Sufficient conditions for circuits in graphs. *Proceedings of the London Mathematical Society*, 3(4):739–755, 1972.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Extreme Points of Base Polytope of Submodular Set Function and Limit of Quotient Convergent Graph Sequence

(Extended abstract)

Yaobin Chen* Zhicheng Liu† Yihang Xiao‡ Junchi Zhang§

Abstract

Submodular set functions are of great importance in mathematics and theoretical computer science. In [14], Lovász systematically extended the theory of submodular set functions from finite sets to general set algebras and proposed several open problems including the characterization of extreme points of the base polytope of submodular set functions. We characterize conditions under which the extreme points of the base polytope of a submodular function are *restricting measures* with respect to its *majorizing measure*. Applying this result, we characterize the core of increasing subadditive non-atomic games and provide a positive answer to a question of Kristóf Bérzi, Márton Borbényi, László Lovász and László Márton Tóth [2] regarding the graphing's cycle matroid.

Furthermore, building on the limit theory for set functions introduced by [3], we prove that the limit of convergent sequence of bounded-degree graphs' cycle matroids can be represented as the cycle matroid of a graphing, analogous to the completeness result for local-global convergence in [9].

1 Introduction

Dating back to work by Choquet [6], Edmonds [8] and Lovász [12], *Submodular set functions* are of great importance in mathematics, theoretical computer science and game theory, serving as a fundamental tool in optimization, combinatorics, and economics due to its natural properties and wide-ranging applications. Recently, in [14], Lovász systematically extended

*Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China. Email: ybchen21@m.fudan.edu.cn. Supported by National Natural Science Foundation of China (Grant No. 123B2012).

†Institute of Operations Research and Information Engineering, Beijing University of Technology, Beijing, China. E-mail: manlzhic@emails.bjut.edu.cn.

‡Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China. E-mail: yhxiao23@m.fudan.edu.cn.

§Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China. E-mail: jczhang24@m.fudan.edu.cn.

the theory of submodular set functions from finite sets to general set algebras, establishing a profound connection between analysis and combinatorial theory. In particular, he introduced the concept of continuity of submodular functions and proposed several open problems about the behavior of submodular functions in infinite settings, including the characterization of *extreme points of the base polytope of submodular set functions.*

Given a family $\mathcal{B}$ of subsets of a set Ω, closed under union and intersection of two elements, a set function $\varphi\colon \mathcal{B} \longrightarrow \mathbb{R}$ is called **submodular**, if for any two subsets $X, Y \in \mathcal{B}$, we have

$$\varphi(X) + \varphi(Y) \geq \varphi(X \cup Y) + \varphi(X \cap Y).$$

A set function $\varphi\colon \mathcal{B} \longrightarrow \mathbb{R}$ is called **subadditive**, if for any two disjoint subsets $X, Y \in \mathcal{B}$, we have

$$\varphi(X) + \varphi(Y) \geq \varphi(X \cup Y).$$

Note that if $\varphi(\emptyset) = 0$ and φ is submodular, then φ is subadditive. The function φ is said to be **continuous from below**, if for any chain of Borel subsets

$$S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n \subseteq \ldots, \quad S_i \in \mathcal{B}, \quad S = \bigcup_{n=1}^{\infty} S_n,$$

we have $\varphi(S) = \lim_{n\to\infty} \varphi(S_n)$. A **Polish space** (Ω, τ) is a separable and completely metrizable topological space. A pair $(\Omega, \mathcal{B})$ is called a **standard Borel space** if there exists a topology τ such that (Ω, τ) is a Polish space and the Borel σ-algebra generated by all open sets in Ω is equal to $\mathcal{B}$.

In this paper, we focus on subadditive set function φ on a standard Borel space $(\Omega, \mathcal{B})$ such that φ is increasing (i.e., $\varphi(X) \leq \varphi(Y)$ if $X \subseteq Y$), $\varphi(\emptyset) = 0$ and continuous from below. In this case, we can define the **majorizing measure** of φ as

$$\mu_\varphi\colon \mathcal{B} \longrightarrow \mathbb{R} \cup \{\infty\}, \quad \mu_\varphi(X) = \sup_{X = \bigsqcup_{i=1}^n X_i} \sum_{i=1}^{n} \varphi(X_i), \quad \text{for any } X \in \mathcal{B},$$

where $\{X_1, X_2, \ldots, X_n\}$ ranges over all partitions of X into sets in $\mathcal{B}$. It is proved in [14] that μ_φ is a measure on Ω if φ is increasing, subadditive, $\varphi(\emptyset) = 0$ and continuous from below. A **φ-atom point** of Ω is an element $x \in X$, such that $\varphi(x) \neq 0$. We denote the set of all φ-atom points in Ω by $\mathbf{Atom}(\mu)$. The **base polytope** of φ is defined as

$$\mathbf{bmm}_+(\varphi) = \{\alpha \mid \alpha \text{ is a measure on } \Omega,\ 0 \leq \alpha \leq \varphi,\ \alpha(\Omega) = \varphi(\Omega)\}.$$

The concept of the base polytope plays a central role in combinatorial optimization, matroid theory and submodular function analysis.

For any convex set C in a linear space, an element $x \in C$ is called an **extreme point** of C, if there do not exist $x_1, x_2 \in C \setminus \{x\}$ such that $\alpha = \dfrac{1}{2}(x_1 + x_2)$, denoted by $\mathbf{Ext}(C)$. Given a convex set C of a metric linear space M and n linear functionals $F_1, F_2, \ldots, F_n$ on M, we define $A_0 = C$ and

$$A_i = \{x \in C \mid F_i(x) = \max_{y \in A_{i-1}} F_i(y)\}, \quad 1 \leq i \leq n.$$

If $A_n \neq \emptyset$, then the set A_n is said to be **exposed by** $(F_1, F_2, \ldots, F_n)$ in C. If A_n consists of a single point, we say that this point is **exposed by** $(F_1, F_2, \ldots, F_n)$ in C. The extreme points and exposed points are often used to characterize the whole convex set.

For any Borel space $(\Omega, \mathcal{B})$, we denote by $\mathscr{M}(\Omega)$ the linear space of all finite measures on $(\Omega, \mathcal{B})$. There exists a metric on $\mathscr{M}(\Omega)$ called the Lévy-Prokhorov metric, denoted by d_P, such that $(\mathscr{M}(\Omega), d_P)$ is a metric linear space.

Given a measure μ on some standard Borel space $(\Omega, \mathcal{B})$, for any $Y \in \mathcal{B}$, the **restricting measure** of μ on Y, denoted by μ_Y, is the measure on $(\Omega, \mathcal{B})$ defined as $\mu_Y(Z) = \mu_\varphi(Y \cap Z)$ for any $Z \in \mathcal{B}$. Given two measures μ, ν on $(\Omega, \mathcal{B})$, we say that ν is **absolutely continuous** with respect to μ if, for any $\epsilon > 0$, there exists $\delta > 0$ such that for any $Y \in \mathcal{B}$, if $\mu(Y) < \delta$, then $\nu(Y) < \epsilon$. If ν is absolutely continuous with respect to μ, there exists a Borel function (i.e., inverse of open sets belongs to $\mathcal{B}$) $f \colon \Omega \longrightarrow \mathbb{R}$ such that $\nu(A) = \displaystyle\int_A f(x)\, d\mu(x)$ for all $A \in \mathcal{B}$, and f is called the Radon-Nikodym derivative of ν with respect to μ, denoted by $\dfrac{d\nu}{d\mu}$. Thus, restricting measures of μ are precisely the measures on $(\Omega, \mathcal{B})$ whose Radon-Nikodym derivative f is $\{0, 1\}$-valued μ_φ-almost everywhere (i.e., $\mu_\varphi(\{x \in \Omega \mid f(x) \notin \{0, 1\}\} = 0)$.

We characterize conditions under which extreme points of the base polytope of a submodular function is an restricting measure under the following assumptions.

Theorem 1.1. *Assume that $(\Omega, \mathcal{B})$ is a standard Borel space and φ is an increasing submodular set function on $\mathcal{B}$ with $\varphi(\emptyset) = 0$ and continuous from below. Let μ_φ be the majorizing measure of φ and for any $Y \in \mathcal{B}$, and $\mu_{\varphi,Y}$ be the restricting measure of μ_φ on Y.*

If $\mu_\varphi(\Omega) < \infty$ and $\mathbf{Atom}(\mu_\varphi) = \emptyset$, then the following two statements are equivalent.

1. $\mathbf{Ext}(\mathbf{bmm}_+(\varphi)) = \{\gamma \in \mathbf{bmm}_+(\varphi) \mid \exists Y \in \mathcal{B},\ s.t.\ \gamma = \mu_{\varphi,Y}\}$.

2. *For any finite sequence of linear functionals $F_1, F_2, \ldots, F_n$ on $\mathscr{M}(\Omega)$, the set exposed by $(F_1, F_2, \ldots, F_n)$ in $\mathbf{bmm}_+(\varphi))$ contains an element in*

$$\{\gamma \in \mathbf{bmm}_+(\varphi) \mid \exists Y \in \mathcal{B},\ s.t.\ \gamma = \mu_{\varphi,Y}\}.$$

Remark 1.2. *We provide explanations for Theorem 1.1.*

1. *The condition "φ is submodular" guarantees $\mathbf{bmm}_+(\varphi) \neq \emptyset$. Assuming this fact, we only need φ to be subadditive. Moreover, μ_φ-atom points are actually φ-atom points, and the case μ_φ-atoms exists is not trivial even if Ω is a countable set.*

2. $\mathbf{Ext}(\mathbf{bmm}_+(\varphi)) \supseteq \{\gamma \in \mathbf{bmm}_+(\varphi) \mid \exists Y \in \mathcal{B},\ s.t.\ \gamma = \mu_{\varphi,Y}\}$. *Intuitively, this theorem characterizes when the extreme point set is minimal.*

3. *Exposed sets are easier to consider than extreme points due to availability of a variety of greedy algorithms.*

In the proof of Theorem 1.1, we use a key lemma Theorem 2.3, which says that the "convex hull" of the restricting measure is compact in $\mathscr{M}(\Omega)$. More precisely, assume that $(\Omega, \mathcal{B})$ is a standard Borel space and φ is an increasing submodular set function on $\mathcal{B}$ with $\varphi(\emptyset) = 0$ and continuous from below. Let μ_φ be the majorizing measure of φ and for any $Y \in \mathcal{B}$, let $\mu_{\varphi,Y}$ be the restricting measure of μ_φ on Y. Then the set consisting of all measures $\alpha \in \mathscr{M}(\Omega)$ such that there exists a measure η on the set X consisting of all restricting measures and $\alpha = \displaystyle\int_X \gamma\, d\eta(\gamma)$, is compact. However, when considering X as a metric subspace of the metric linear space $(\mathscr{M}(\Omega), d_P)$, the probability measure space on X is not neacessarily compact, because X itself is not necessarily compact (unless Ω is a finite set), highlighting a key distinction between the finite and infinite settings.

In the following two parts we give some applications of Theorem 1.1 on game theory and matroid theory. And in the third part, we discuss another topic closely related to the second part.

1.1 Core of Non-Atomic Games

In [1], R.J. Aumann (Nobel prize 2005 in Economics) and L.S. Shapley investigate the *non-atomic game*. Let $(\Omega, \mathcal{B})$ be a standard Borel space (often the space is chosen as the interval $[0,1] \subseteq \mathbb{R}$). The points in Ω are players and subsets in $\mathcal{B}$ are coalitions. A **game** on $(\Omega, \mathcal{B})$ is a set function φ on $(\Omega, \mathcal{B})$ with $\varphi(\emptyset) = 0$. A game is called **non-atomic** if there does not exist φ-atoms in $\mathcal{B}$. A game is called subadditive (or superadditive) if, for any two disjoint Borel sets $X, Y \in \mathcal{B}$, we have

$$\varphi(X) + \varphi(Y) \geq (\text{or } \leq)\varphi(X \sqcup Y).$$

Both subadditive games and superadditive games play fundamental roles in game theory. The *core of a game* is a fundamental concept in game theory and has been extensively studied by both game theorists and economists.. Given a subadditive (or superadditive) game φ, the **core** of φ is defined as

$$\mathbf{core}(\varphi) = \{\gamma \mid \gamma \text{ is a finitely additive measure on } \Omega, \gamma \leq (\text{or } \geq)\varphi, \gamma(\Omega) = \varphi(\Omega)\}.$$

Using the above notions, Aumann and Shaply [[1], Theorem I] charaterized the core of superadditive non-atomic games under some particular conditions. By reformulating Theorem 1.1 in the framework of game theory, we give a characterization of the core of increasing subadditive non-atomic games.

Theorem 1.3. *Assume that $(\Omega, \mathcal{B})$ is a standard Borel space and φ is a non-atomic game such that $\alpha \leq \mu$ for some $\mu \in NA^+(\Omega, B)$. If φ is increasing and subadditive, then the extreme points of $\mathbf{core}(\varphi)$ are precisely the restricting measures of the majorizing measure μ_φ if and only if for any finite sequence of linear functionals $F_1, F_2, \ldots, F_n$ on $\mathscr{M}(\Omega)$, the set exposed by $(F_1, F_2, \ldots, F_n)$ contains at least one restricting measure of μ_φ.*

1.2 Cycle Matroid Rank Functions

A **Borel graph** is a triple $(\Omega, \mathcal{B}, E)$, where $(\Omega, \mathcal{B})$ is a standard Borel space and E is a symmetric Borel subset in $\Omega \times \Omega$. A **graphing** is a bounded-degree graph $\mathbf{G} = (\Omega, \mu, E)$, where $(\Omega, \mathcal{B}, E)$ is a Borel graph and μ is an probability measure on $(\Omega, \mathcal{B})$ that satisfies certain invariance properties. The rank function of a graphing $\mathbf{G}$ is defined as

$$\rho_{\mathbf{G}}(F) \triangleq 1 - \mathbb{E}_x[\frac{1}{|V(\mathbf{G}[F]_x)|}], \forall \text{ Borel edge subset } F,$$

where x is a random point from μ, $\mathbf{G}[F]_x$ denotes the connected component containing x in the induced subgraphing $\mathbf{G}[F] = (\Omega, \mathcal{B}, \mu, F)$, and $V(G[F]_x)$ denotes the set of vertices of this component. A graphing $\mathbf{G} = (\Omega, \mathcal{B}, \mu, E)$ is **hyperfinite** if, for every $\varepsilon > 0$, there exists a Borel set $S \subseteq E$ with $\eta_{\mathbf{G}}(S) < \varepsilon$ such that every connected component of $E \setminus S$ is finite. A

forest in $\mathbf{G}$ is an acylic subset $T \subseteq E$. An **essential spanning forest** is a forest that forms a spanning tree in every finite component of $\mathbf{G}$ and contains only infinite T-components in the infinite $\mathbf{G}$-components. For further details on these definitions, we refer to [2].

Theorem 1.4. ([[15], Theorem 4.3]) Given a graphing $\mathbf{G}$, $\rho_{\mathbf{G}}$ is increasing, submodular and $\rho_{\mathbf{G}}(\emptyset) = 0$.

By Theorem 1.4, we can define the base polytope of $\rho_{\mathbf{G}}$. In [2], Bérczi, Borbényi, Lovász and Tóth proved the following characterization of exposed points of $\mathbf{bmm}_+(\rho_{\mathbf{G}})$ and proposed an open question: are all extreme points of $\mathbf{bmm}_+(\rho_{\mathbf{G}})$ exposed, and therefore hyperfinite spanning forests?

Theorem 1.5. ([[2], Theorem 1.2]) Given a graphing $\mathbf{G}$, α is an exposed point of $\mathbf{bmm}_+(\rho_{\mathbf{G}})$ if and only if $\alpha = \tilde{\mu}_F$, the restricting measure for an essential hyperfinte spanning forest F.

By applying Theorem 1.1, we give a positive answer to this question as below.

Theorem 1.6. *Given a graphing* $\mathbf{G}$*,* α *is an extreme point of* $\mathbf{bmm}_+(\rho_{\mathbf{G}})$ *if and only if* $\alpha = \tilde{\mu}_F$*, the restricting measure for an essential hyperfinte spanning forest* F*.*

1.3 Limit for Quotient Qonvergent Graph Sequence

In [3], Bérczi, Borbényi, Lovász and Tóth developed a limit theory for the cycle matroids of graphings as below. For a set function ϕ on a set-algebra $(\mathcal{J}, \mathcal{B})$, a k**-quotient** of φ is the function $\phi \circ F^{-1}$, where $F : J \to [k]$ is a measurable map. We denote by $\mathcal{Q}_k(\phi)$ the set of all k-quotients of ϕ. A sequence of set functions $(\phi_1, \phi_2, \ldots)$ is said to **quotient converge** to ϕ if, for all $k \in \mathbb{N}$, $\mathcal{Q}_k(\phi_n) \to \mathcal{Q}_k(\phi)$ in the Hausdorff distance as $n \to \infty$. In [2], it is shown by Bérczi, Borbényi, Lovász and Tóth that the local-global convergence (we do not explain the definition for brevity) of a sequence of finite graphs to a graphing implies convergence of the corresponding matroids.

Theorem 1.7. ([[2], Theorem 3.2]) Let $(G_i)_{i=1}^{\infty}$ be a sequence of finite graphs with maximal degree bounded by D that converges to a graphing $\mathbf{G}$ in the local-global sense. Then the rank function ρ_i of G_i quotient converges to $\rho_{\mathbf{G}}$.

Moreover, in [9], Hatami, Lovász and Szegedy prove the following completeness result.

Theorem 1.8. ([[9], Theorem 3.2]) Let $(G_i)_{i=1}^{\infty}$ be a local-global convergent sequence of finite graphs with maximal degree bounded by D. Then there exists a graphing $\mathbf{G}$ such that the $(G_i)_{i=1}^{\infty}$ is local-global convergent to $\mathbf{G}$.

Hence, the following natural question is raised: Is the limit of cycle matroids of bounded-degree graph always a cycle matroid of some graphing? We prove it is true as follows.

Theorem 1.9. *Given* $D \in \mathbb{N}^*$*, let* $\{G_i\}_{i=1}^{\infty}$ *be a sequence of connected graphs with maximal degree bounded by* D*. Let* ρ_i *be the rank function of the cycle matroid of* G_i *as a graphing. If* $\{\rho_i\}_{i=1}^{\infty}$ *is quotient convergent, then there exists a graphing* $\mathbf{G}$*, such that the rank function* $\rho_{\mathbf{G}}$ *is a limit of* ρ_i*.*

2 Proof of Theorem 1.1

In the section, we assume that $(\Omega, \mathcal{B})$ is a standard Borel space and φ is an increasing submodular setfunction on $\mathcal{B}$ with $\varphi(\emptyset) = 0$ and continuous from below. Let μ_φ be the majorizing measure of φ. For any $\gamma \in \mathbf{bmm}_+(\varphi)$, we denote by f_γ the Radon-Nikodym derivative $\dfrac{d\gamma}{d\mu_\varphi}$.

First, we show that for the standard Borel space $(\Omega, \mathcal{B})$, there exists a countable set of Borel sets $\{X_i \in \mathcal{B} \mid i \geq 1\}$, such that for any two elements $\alpha_1, \alpha_2 \in \mathbf{bmm}_+(\varphi)$, $\alpha_1 = \alpha_2$ if and only if $\alpha_1(X_i) = \alpha_2(X_i)$ for any $i \geq 1$. This property helps us to transform this "uncountable problem" into a "countable problem" and it is proved in Theorem 2.1.

Lemma 2.1. *Suppose $\mathcal{B} = \sigma(\mathcal{X})$ for a countable family of Borel subsets $\mathcal{X} = \{X_i \in \mathcal{B} \mid i \geq 1\}$. Let $\mathcal{B}_r = \sigma(X_1, \ldots, X_r)$ for $r \geq 1$. Then for any Borel subset $X \in \mathcal{B}$, there exists a sequence of Borel subset $\{X_n \in \mathcal{B}_n \mid n \geq 1\}$, such that $\lim\limits_{n\to\infty} \mu_\varphi(X_n \,\Delta\, X) = 0$. In particular, the map $T : \mathbf{bmm}_+(\varphi) \longrightarrow \mathbb{R}^{\mathbb{N}}$, $\gamma \mapsto (\gamma(X_i))_{i\geq 1}$ is injective.*

Let $\mathscr{M}(X)$ be the linear space of finite measures on X. For any subset of $\mathscr{M}(\Omega)$, including $\mathbf{bmm}_+(\varphi)$, we view it as a metric space with the Lévy-Prokhorov metric d_P. To prove Theorem 1.1, we need to prove that for any $\alpha \in \mathbf{Ext}(\mathbf{bmm}_+(\varphi))$, the Radon-Nikodym derivative $f_\alpha = \dfrac{d\alpha}{d\mu_\varphi}$ is 0-1 valued μ_φ-almost everywhere. Alternatively, we prove that $\mathbf{Ext}(\mathbf{bmm}_+(\varphi))$ is contained in the "convex hull" of elements in $\mathbf{bmm}_+(\varphi)$ whose Radon-Nikodym derivative is 0-1 valued μ_φ-almost everywhere. To clarify what "convex hull" is, we introduce the following definition.

Definition 2.2. *The **set of all Choquet representations** of X in a liear metric space M is defined as*

$$\mathbf{Ch}(X) \triangleq \left\{ y \in M \;\middle|\; \begin{array}{l} \exists \text{ probability measure } \eta \text{ on } X, \text{ such that} \\ f(y) = \displaystyle\int_X f(x)\, d\eta(x), \ \forall \text{ continuous linear functionals } f \text{ on } M \end{array} \right\}.$$

For any Borel subset $X \subseteq \Omega$, we denote the set of all 0-1 valued μ_φ-almost everywhere Borel functions on X by $\mathbf{Ind}(X)$. Then $\{\gamma \in \mathbf{bmm}_+(\varphi) \mid f_\gamma \in \mathbf{Ind}(\Omega)\}$ is a subset of the metric linear space $(\mathscr{M}(\Omega), d_P)$. Using the above definition, to prove Theorem 1.1, we try to prove that

$$\mathbf{Ext}(\mathbf{bmm}_+(\varphi)) \subseteq \mathbf{Ch}(\{\gamma \in \mathbf{bmm}_+(\varphi) \mid f_\gamma \in \mathbf{Ind}(\Omega))\} \subseteq \mathbf{bmm}_+(\varphi).$$

We attempt to apply the Prokhorov's Theorem, which says that the probability measure space of a compact metric space is compact with respect to the Lévy-Prokhorov metric d_P. However, the set $\{\gamma \in \mathbf{bmm}_+(\varphi) \mid f_\gamma \in \mathbf{Ind}(\Omega)\}$ is not necessarily a compact set, and this is the essential gap between the finite case and the infinite case.

As an alternative, We prove the following key lemma.

Lemma 2.3. *For any $a \in (0, \mu_\varphi(\Omega)]$, let $\mathscr{P}_a = \{\gamma \in \mathscr{M}(\Omega) \mid 0 \leq \gamma \leq \mu_\varphi, \gamma(\Omega) = a\}$ and $\mathcal{I}_a = \{\gamma \in \mathscr{M}(\Omega) \mid 0 \leq \gamma \leq \mu_\varphi, \gamma(\Omega) = a, f_\gamma \in \boldsymbol{Ind}(\Omega)\}$.*

If $\mathbf{Atom}(\mu_\varphi) = \emptyset$, then $\mathscr{P}_a = \mathbf{Ch}(\mathcal{I}_a)$ and it is compact. Further,

$$\mathbf{Ch}(\{\gamma \in \mathbf{bmm}_+(\varphi) \mid f_\gamma \in \boldsymbol{Ind}(\Omega))\} = \mathbf{Ch}(\{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi\})$$

is also compact.

We give a sketch of proof of Theorem 2.3 as follows. First, we prove that the space $\mathscr{P}_a$ is a compact subspace of $(\mathscr{M}(\Omega), d_P)$ by Prokhorov's Theorem, and then prove that $\mathscr{P}_a = \mathbf{Ch}(\mathcal{I}_a)$ by the property of Radon-Nikodym derivative and extreme points.

Next, take a countable family of Borel subsets $\mathcal{X} = \{X_i \in \mathcal{B} \mid i \geq 1\}$ such that $\mathcal{B} = \sigma(\mathcal{X})$ and Let $\mathcal{B}_r = \sigma(X_1, \ldots, X_r)$ for $r \geq 1$. We prove that

$$\mathbf{Ch}(\{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi\}) = \bigcap_{n\geq 1} \mathbf{Ch}(\bigcap_{X\in\mathcal{B}_n} \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi(X)\}).$$

To do this, we consider the space of probability measures on $\mathscr{P}_{\varphi(\Omega)}(\Omega)$, and apply Prokhorov's Theorem again to this compact set $\mathscr{P}_{\varphi(\Omega)}(\Omega)$.

Now, we aim to prove that for any $n \geq 1$, $\mathbf{Ch}(\bigcap_{X\in\mathcal{B}_n} \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi(X)\})$ is a closed subset (and thus compact) of the compact set $\mathscr{P}_{\varphi(\Omega)}(\Omega)$ By taking intersections, we complete the proof of Theorem 2.3. To do this, let $\mathcal{A}_n$ be the partition of Ω corresponding to the finite σ-algebra $\mathcal{B}_n$, and we have

$$\mathbf{Ch}(\bigcap_{X\in\mathcal{B}_n} \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi(X)\}) = \mathbf{Ch}(\bigcap_{X\in\mathcal{A}_n} \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi(X)\}).$$

For any Cauchy sequence in this subset, there exists a limit $\beta \in \mathscr{P}_{\varphi(\Omega)}$, and we need to prove that β belongs to this subset. Let $\mathcal{A}_n = \{A_1, A_2, \ldots, A_{h_n}\}$, $a_i = \beta(A_i)$, $Y_i \triangleq \mathcal{I}_{a_i}(A_i)$, $i = 1, 2, \ldots, h_n$, and $Y \triangleq \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma(A_i) = a_i,\ i = 1, 2, \ldots, h_n\}$. The Borel bijection

$$T : Y_1 \times Y_2 \times \cdots \times Y_k \longrightarrow Y,\ \ T((\gamma_1, \gamma_2, \ldots, \gamma_k)) = \gamma,\ \gamma(X) = \sum_{i=1}^{k} \gamma_i(X \cap Z_i),\ \ \forall\, X \in \mathcal{B},$$

and the fact that $\mathscr{P}_{a_i}(A_i) = \mathbf{Ch}(\mathcal{I}_{a_i}(A_i))$ (Here $\mathscr{P}_{a_i}(A_i)$ and $\mathcal{I}_{a_i}(A_i)$ are defined similarly with the ground set Ω replaced by A_i) imply that

$$\beta \in \mathbf{Ch}(Y) \subseteq \mathbf{Ch}(\bigcap_{X\in\mathcal{B}_n} \{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi(X)\}).$$

Finally, given $\alpha \in \mathbf{Ext}(\mathbf{bmm}_+(\varphi))$, we use the condition 2 of Theorem 1.1 to obtain a sequence of elements in $\mathbf{Ch}(\{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi\})$ that converges to α. By Theorem 2.3, there exists a measure η on $\{\gamma \in \mathcal{I}_{\varphi(\Omega)} \mid \gamma \leq \varphi\})$ such that α is the Choquet representation by η. Roughly speaking, this implies that α is in the convex hull of the above set. Combined with the property of extreme points, we conclude that $\alpha \in \mathbf{Ch}(\{\gamma \in \mathbf{bmm}_+(\varphi) \mid f_\gamma \in \mathbf{Ind}(\Omega))\}$. And we complete the proof of the more difficult direction of Theorem 1.1.

3 Proof of Theorem 1.6

Given a graphing $\mathbf{G} = (\Omega, \mathcal{B}, \mu, E)$, a measure η on $\Omega \times \Omega$ is induced by μ, such that for any Borel subsets $A, B \subseteq \Omega$, $\eta(A \times B) = \frac{1}{2}\int_A d_B(x)d\mu(x)$. By Caratheodory's Theorem, there exists a unique measure on $\Omega \times \Omega$ induced by η, and this measure **induces a measure on** E, denoted by $\tilde{\mu}$.

It can be shown that the measure $\tilde{\mu}$ is exactly the majorizing measure of the cycle matroid rank function $\rho_{\mathbf{G}}$. Moreover, the condition 2 is guaranteed by the following theorem.

Theorem 3.1. ([[2], Theorem 4.2, 4.7]) Given a graphing $\mathbf{G} = (\Omega, \mathcal{B}, \mu, E)$ with maximal degree bounded by D, let $\tilde{\mu}$ be the measure on E induced by μ. For any element $\gamma \in \mathbf{bmm}_+(\rho_{\mathbf{G}})$, let f_γ be the Radon-Nikodym derivative of α with respect to $\tilde{\mu}$.

Given $g_1, g_2, \ldots, g_m \in L^\infty(E, \tilde{\mu})$, we define $A_0 = \mathbf{bmm}_+(\rho_{\mathbf{G}})$ and

$$A_i = \{\gamma_0 \in A_{i-1} \mid \phi_{g_i}(\gamma_0) = \max_{\gamma \in A_{i-1}} \phi_{g_i}(\gamma)\}, 1 \leq i \leq m.$$

if f_α is the unique element in A_m, then there exists an essential hyperfinite spanning forest F, such that we can define an element $\gamma \in \mathbf{bmm}_+(\rho_{\mathbf{G}})$ with $\gamma(I) = \tilde{\mu}(I \cap F)$ for any Borel set I and $\gamma \in A_m$.

It remains to show that we can reduce the problem to the case that there exists no $\tilde{\mu}$-atoms. And this follows from the following lemma.

Lemma 3.2. *Given a graphing $\mathbf{G} = (\Omega, \mathcal{B}, \mu, E)$, let $\tilde{\mu}$ be the measure on $(E, \mathcal{B}_E)$ induced by μ. If an edge set $F \subseteq E$ is a $\tilde{\mu}$-atom, then there exists some $e = xy \in E$, such that $\tilde{\mu}(\{e\}) = \tilde{\mu}(F)$. In particular, x, y are μ-atoms. Let $\mathbf{Atom}(\Omega)$ be the set of μ-atoms consisting of a single vertex. Then there exists no edge between $\mathbf{Atom}(\Omega)$ and $\mathbf{Atom}(\Omega)^c$. Further, for the induced graphing $\mathbf{G}[\mathbf{Atom}(\Omega)]$, every connected component is finite and there exist countably many connected components in $\mathbf{G}[\mathbf{Atom}(\Omega)]$.*

4 Proof of Theorem 1.9

We construct an edge colored directed Borel graph $\mathbf{H}^K$ following the method in [9] that is used to prove Theorem 1.8. Briefly, given a sequence of convergent graphs, the authors construct a space consisting of finite vertex-labelled connected graphs, regard these graphs as vertices of a Borel graph and define the edges between these vertices in a particular way, such that every finite graph can be embed into the Borel graph. Moreover, for any convergent graph sequence, they define a sequence of involution measures on the Borel graph to make it isomorphic to the graphs as a graphing respectively, and then use Prokhorov's Theorem to find a weak limit and obtain the required graphing.

In our proof, given a graph sequence whose rank functions are quotient convergent, since what we consider is the rank function of the cycle matroid, we focus on the edges of the graphs rather than the vertices. So we consider the space Ω consisting of all countable rooted directed connected graphs with edge colored by some decoration space K. Here, we carefully choose an injective edge coloring of every finite directed graph to distinguish edges and add an orientation on the edges to distinguish the two vertices of each edge. Thus we can embed any finite directed graph to this Borel graph $\mathbf{H}^K$. For any such graphs, we can give an involution invariant measure on Ω to obtain a graphing isomorphic to this graph as a graphing. We can prove that Ω is a compact space, and thus by Prokhorov's Theorem, the space of probability measures on Ω is compact. So we can choose a weak limit measure μ of some subsequence of the corresponding measures of the graphs in the graph sequence in Theorem 1.9. Then we get a graphing and try to prove that the rank function of this graphing is exactly the limit of the rank functions of the graph sequence.

References

[1] Robert J. Aumann and Lloyd S. Shapley, *Values of non-atomic games*, 1974.

[2] Kristóf Bérczi, Márton Borbényi, László Lovász, and László Márton Tóth, *Cycle matroids of graphings: from convergence to duality*, arXiv preprint arXiv:2406.08945 (2024).

[3] ———, *Quotient-convergence of submodular setfunctions*, arXiv preprint arXiv:2406.08942 (2024).

[4] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication.

[5] Vladimir Igorevich Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[6] Gustave Choquet, *Theory of capacities*, Annales de l'institut Fourier, vol. 5, 1954, pp. 131–295.

[7] John B. Conway, *A course in functional analysis*, vol. 96, Springer, 2019.

[8] Jack Edmonds, *Submodular functions, matroids, and certain polyhedra*, Combinatorial Optimization—Eureka, You Shrink! Papers Dedicated to Jack Edmonds 5th International Workshop Aussois, France, March 5–9, 2001 Revised Papers, Springer, 2003, pp. 11–26.

[9] Hamed Hatami, László Lovász, and Balázs Szegedy, *Limits of locally-globally convergent graph sequences*, Geom. Funct. Anal. **24** (2014), no. 1, 269–296. MR 3177383

[10] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.

[11] Alexander S. Kechris, S. Solecki, and S. Todorcevic, *Borel chromatic numbers*, Adv. Math. **141** (1999), no. 1, 1–44. MR 1667145

[12] László Lovász, *Submodular functions and convexity*, Mathematical Programming The State of the Art: Bonn 1982 (1983), 235–257.

[13] ———, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.

[14] ———, *Submodular setfunctions on sigma-algebras, version 2*, arXiv preprint arXiv:2302.04704 (2023).

[15] ———, *The matroid of a graphing*, J. Combin. Theory Ser. B **169** (2024), 542–560. MR 4792434

[16] James Oxley, *Matroid theory*, second ed., Oxford Graduate Texts in Mathematics, vol. 21, Oxford University Press, Oxford, 2011.

[17] Robert R. Phelps, *Lectures on Choquet's theorem*, second ed., Lecture Notes in Mathematics, vol. 1757, Springer-Verlag, Berlin, 2001.

[18] Oleg Pikhurko, *Borel combinatorics of locally finite graphs*, Surveys in combinatorics 2021, London Math. Soc. Lecture Note Ser., vol. 470, Cambridge Univ. Press, Cambridge, 2021, pp. 267–319.

[19] R. T. Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Reprint of the 1970 original, Princeton Paperbacks.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The burning game on graphs

(Extended abstract)

Nina Chiarelli* Vesna Iršič Chenoweth† Marko Jakovac‡
William B. Kinnersley§ Mirjana Mikalački¶

Abstract

Motivated by the burning and cooling processes, the burning game is introduced. The game is played on a graph G by the two players (Burner and Staller) that take turns selecting vertices of G to burn; as in the burning process, burning vertices spread fire to unburned neighbors. Burner aims to burn all vertices of G as quickly as possible, while Staller wants to prolong the process as much as possible. If both players play optimally, then the number of time steps needed to burn the whole graph G is the game burning number $b_{\rm g}(G)$ if Burner makes the first move, and the Staller-start game burning number $b'_{\rm g}(G)$ if Staller starts.

Here, basic bounds on $b_{\rm g}(G)$ are given and several fundamental properties of the burning game established. An analogue of the burning number conjecture for the burning game is also considered. Finally, it is shown that the problem of determining whether or not $b_{\rm g}(G) \leq k$ is NP-hard.

*FAMNIT and IAM, University of Primorska, Slovenia. E-mail: nina.chiarelli@famnit.upr.si. Supported by the Slovenian Research and Innovation Agency (I0-0035, research program P1-0404, and research projects N1-0210, N1-0370, J1-3003, and J1-4008).

†Faculty of Mathematics and Physics, University Ljubljana and Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia. E-mail: vesna.irsic@fmf.uni-lj.si. Supported by the Slovenian Research and Innovation Agency (Z1-50003, P1-0297, N1-0218, N1-0285, N1-0355) and the European Union (ERC, KARST, 101071836).

‡Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor and Institute of Mathematics, Physics and Mechanics, Ljubljana,Slovenia. E-mail: marko.jakovac@um.si. Supported by the Slovenian Research and Innovation Agency (research core funding No. P1-0297 and projects N1-0285, J1-3002, and J1-4008).

§Department of Mathematics and Applied Mathematical Sciences, University of Rhode Island, USA. E-mail: billk@uri.edu.

¶Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, Serbia. E-mail: mirjana.mikalacki@dmi.uns.ac.rs. Supported partially by the Provincial Secretariat for Higher Education and Scientific Research, Province of Vojvodina (Grant No. 142-451-2686/2021) and Ministry of Science, Technological Development and Innovation of Republic of Serbia (Grants 451-03-137/2025-03/200125 & 451-03-136/2025-03/200125).

1 Introduction

Nowadays, the spread of the social influence is an active research field, as it demonstrates the importance of the underlying network, and no need for in-person communication [11]. The processes of virus propagation in medicine or in computer networks, or the spread of the trends over social networks, are just some examples of various natural and engineered phenomena spread over networks, that are active research topics (see, e.g. [4, 23, 26, 32]). In all the aforementioned problems, the natural question is how quickly the contagion can spread over all the members in the network.

The *burning process* on graphs was introduced in [11] as a simplified deterministic model to analyze this question, inspired by the processes of firefighting [5, 18], graph cleaning [2] and graph bootstrap percolation [3]. (Unknown to these authors, a similar process was actually introduced much earlier in the paper of Alon [1].) Later on, the *cooling process* was introduced in [14] as a dual of the burning process, tending to slow down the burning and modeling the mitigation of infection spread and virus propagation.

The burning and cooling processes are defined as follows (merged together from [11, 14]). Given a finite, simple, undirected graph G, the burning process on G is a discrete-time process. Vertices may be either unburned or burned throughout the process. Initially, in round $t = 0$ all vertices are unburned. At each round $t \geq 1$, one new unburned vertex is chosen to burn, if such a vertex is available. Such a chosen vertex is called a *source*. If a vertex is burned, then it remains in that state until the end of the process. Once a vertex is burned in round t, in round $t+1$ each of its unburned neighbors becomes burned. The process ends in a given round when all vertices of G are burned. The *burning number* of a graph G, denoted by $b(G)$, is the minimum number of rounds needed for the process to end. Analogously, the *cooling number* of G, $CL(G)$, is defined to be the maximum number of rounds for the cooling process to end. The sequence of sources chosen in an instance of the burning process (respectively, cooling process) is referred to as a *burning sequence* (resp. *cooling sequence*. The main difference between the two processes is in the way of choosing the sources. The length of the shortest burning sequence is $b(G)$, and for every graph G it holds that $b(G) \leq CL(G)$. Note that in the burning process, the selection of a new source and the spread of fire to neighboring vertices happen simultaneously. Thus, there is a source selected in every round, even if the selected vertex would have burned in the same round anyway. Note, however, that in the cooling process it can happen that there is no source selected in the last round if all the remaining vertices of the graph already burned.

Although it was only recently introduced, graph burning has stimulated a great deal of research. Much of this research has focused on resolving the so-called *burning number conjecture* posed by Bonato et al. in [11], which asserts that every n-vertex connected graph G satisfies $b(G) \leq \lceil \sqrt{n} \rceil$. Upper bounds on $b(G)$ have been gradually improved over time by several authors (see e.g. [6], [8], [27], and [29]). It is known that if the burning number conjecture holds for trees, then it holds for all connected graphs; thus, several papers have focused on determining or bounding $b(G)$ for various classes of trees, such as spiders ([13], [17]) and caterpillars ([21], [28]). Several authors have also investigated the computational complexity of determining the burning number of a graph; this problem was shown to be NP-complete by Bessy et al. in [7], although polynomial-time approximation algorithms are known for several classes of graphs (see [7], [12], [13]). For more details on previous work in the area, see the recent survey [10].

Motivated by the two aforementioned processes of burning and cooling, we introduce a new graph game - the *burning game*. In the burning game on a graph G, the two players Burner (he/him) and Staller (she/her) take turns selecting vertices of G to burn; as in the burning process, burning vertices spread fire to unburned neighbors. Burner aims to burn all vertices of G as quickly as possible, while Staller wants the process to last as long as possible. The burning game is similar in spirit to several other competitive games based on graph parameters, e.g. domination games [15, 31], the coloring game [9, 33], the competition-independence game [30, 19] and saturation games [20, 24].

Formally, the game is defined as follows. Let G be a finite simple graph. Vertices are burned or unburned, but once burned they stay in this state until the end of the game. At time step $t = 0$, all vertices are unburned. In each time step $t \geq 1$, first all neighbors of burned vertices become burned (*the spreading phase*), and then one of the players burns one unburned vertex in this time step as well (*the selection phase*). The game ends in the first time step t in which all vertices of G are burned. The aim of Burner is to minimize the total number of time steps and the aim of Staller is to maximize it. If both players play optimally, then the number of time steps needed to burn the whole graph G is called the *game burning number* $b_{\mathrm{g}}(G)$ if Burner makes the first move, and the *Staller-start game burning number* $b'_{\mathrm{g}}(G)$ if Staller starts.

One time step represents *one round* in the burning game, whose first part is the spreading phase, and the second part is the selection phase. The first round in the game consists only of the selection phase. However, it is possible that the last round ends after only the spreading phase, if there are no unburned vertices left in the graph.

The burning number of a graph, $b(G)$, can equivalently be viewed as the length of the burning game in which Burner is the only player, while the cooling number $\mathrm{CL}(G)$ can equivalently be seen as the length of the burning game where Staller is the only player. Just note that due to our specification of a round into two phases, the burning game with only Burner playing can slightly differ from the burning processes (in the last round).

We will focus on showing some basic bounds on $b_{\mathrm{g}}(G)$ first, then consider the analogue of the burning number conjecture for the burning game and finally consider the computational complexity of the burning game, and in particular prove that determining whether $b_{\mathrm{g}}(G) \leq k$ is NP-hard. The proofs and additional results on the burning game can be found in our preprint [16].

2 Basic properties

In this section, we establish some basic properties of the burning game, as well as several elementary bounds on b_{g}.

As might be expected, the game burning number of a graph is closely related to its burning and cooling numbers. Note that given a graph G, the square of the graph, G^2, is a graph whose vertices are adjacent if their distance in G is at most 2.

Proposition 1. *If G is a connected graph, then $b(G) \leq b_{\mathrm{g}}(G) \leq \min\{\mathrm{CL}(G), 2b(G^2)-1\}$ and $b(G) \leq b'_{\mathrm{g}}(G) \leq \min\{\mathrm{CL}(G), 2b(G^2)\}$.*

If G is a graph and u is a vertex of G, then the *eccentricity* of u is defined as $\mathrm{ecc}(v) = \max\{d(u,v) : v \in V(G)\}$. The *radius* and *diameter* of G are defined as the minimum and maximum eccentricities, respectively, over all vertices in G.

As with the burning number, the game burning number of a graph G can be bounded above in terms of the radius of G.

Proposition 2. *If G is a connected graph, then $b_{\mathrm{g}}(G) \le \mathrm{rad}(G)+1$ and $b'_{\mathrm{g}}(G) \le \min\{\mathrm{rad}(G)+2, \mathrm{diam}(G)+1\}$.*

When analyzing the burning game, we would often like to consider a game that is "already in progress" – that is, with some vertices already burning. Given a graph G and $B \subseteq V(G)$, we let $G|B$ denote the graph G, with the understanding that the vertices in B are already burning prior to the start of the game. We refer to the burning game on $G|B$ as the burning game *relative to* B, and we denote the number of rounds needed to burn all of $G|B$ – assuming that both players play optimally – by $b_{\mathrm{g}}(G|B)$ if Burner makes the first move, and by $b'_{\mathrm{g}}(G|B)$ if Staller makes the first move.

The following result, known as the *Continuation Principle*[1], formalizes the intuition that starting the game with additional vertices burned can never increase the length of the game.

Theorem 3 (Continuation Principle). *If $A \subseteq B \subseteq V(G)$, then $b_{\mathrm{g}}(G|B) \le b_{\mathrm{g}}(G|A)$ and $b'_{\mathrm{g}}(G|B) \le b'_{\mathrm{g}}(G|A)$.*

Proposition 4. *If G is a connected graph, then $|b_{\mathrm{g}}(G) - b'_{\mathrm{g}}(G)| \le 1$.*

All possibilities from Proposition 4 can be achieved. For $n \ge 2$, $b_{\mathrm{g}}(K_n) = b'_{\mathrm{g}}(K_n) = 2$. For $n \ge 3$, $b_{\mathrm{g}}(K_{1,n}) = 2$ and $b'_{\mathrm{g}}(K_{1,n}) = 3$. For even $n \ge 4$, $b_{\mathrm{g}}(Q_n) = \frac{n}{2} + 2$ and $b'_{\mathrm{g}}(Q_n) = \frac{n}{2} + 1$ (see [16, Theorem 29]).

Theorem 5. *If G is a connected graph and e is any edge in G, then $b_{\mathrm{g}}(G) \le b_{\mathrm{g}}(G-e) \le b_{\mathrm{g}}(G)+2$ and $b'_{\mathrm{g}}(G) \le b'_{\mathrm{g}}(G-e) \le b'_{\mathrm{g}}(G)+2$.*

The following lemma is a useful consequence of Theorem 5.

Lemma 6. *If H is a spanning subgraph of G, then $b_{\mathrm{g}}(G) \le b_{\mathrm{g}}(H)$ and $b'_{\mathrm{g}}(G) \le b'_{\mathrm{g}}(H)$.*

3 The Burning Number Conjecture

As mentioned in the Introduction, the *burning number conjecture*, posed by Bonato et al. [11], has attracted a great deal of attention in recent years. Recall that the burning number conjecture states that for every n-vertex connected graph G, we have $b(G) \le \lceil \sqrt{n} \rceil$ or, equivalently, $b(G) \le b(P_n)$. In this section, we consider the analogous question for the burning game: what is the maximum value of $b_{\mathrm{g}}(G)$ among n-vertex connected graphs G?

It was shown in [11] that for every connected graph G, we have

$$b(G) = \min\{b(T) \colon T \text{ is a spanning tree of } G\}.$$

Unfortunately, this result does not extend to the burning game. Lemma 6 implies that

$$b_{\mathrm{g}}(G) \le \min\{b_{\mathrm{g}}(T) \colon T \text{ is a spanning tree of } G\};$$

however, unlike in the burning process, the equality need not hold.

The first step in establishing an analogue of the burning conjecture for the burning game is determining the game burning number of paths. This turns out to be a nontrivial problem. We are able to obtain a lower and upper bound that differ by at most 1, and can be expressed in a simplified form as follows.

[1] The name for this result is taken from that of an analogous result for the domination game; see [25].

Theorem 7. *For all $n \geq 1$, we have $b_{\mathrm{g}}(P_n) = (1+o(1))\sqrt{2n}$ and $b'_{\mathrm{g}}(P_n) = (1+o(1))\sqrt{2n}$.*

While the burning conjecture is still open, Norin and Turcotte [29] recently showed that it is asymptotically true, i.e. that $b(G) \leq (1+o(1))\sqrt{n}$ for all connected n-vertex graphs G. Using this result together with more technical lemmas, we are able to show the following.

Corollary 8. *If G is a connected graph on n vertices, then*

$$b_{\mathrm{g}}(G) \leq (1+o(1))\sqrt{2n}.$$

Additionally, we observe that the burning number conjecture is related to the general upper bound for the game burning number.

Corollary 9. *If the burning number conjecture is true, then for every connected graph G on n vertices, we have $b_{\mathrm{g}}(G) \leq \lfloor \sqrt{2n} \rfloor + 3$.*

4 Computational Complexity

Finally, we consider the following decision problem:

BurningGame: Given a graph G and positive integer k, is $b_{\mathrm{g}}(G)$ at most k?

As mentioned in the introduction, determining the burning number of graphs is NP-complete even for trees with maximum degree 3 and several other classes of graphs [7]. Consequently, it is natural to suspect that BurningGame is NP-hard (if not NP-complete). We show that this is in fact the case, however, we suspect that BurningGame is actually PSPACE-complete. We prove the NP-hardness of this problem through reduction from the problem 3-SAT, which is well-known to be NP-complete (see, for example, [22]).

Theorem 10. *BurningGame is NP-hard.*

References

[1] Noga Alon. Transmitting in the n-dimensional cube. *Discrete Applied Mathematics*, 37:9–11, 1992.

[2] Noga Alon, Paweł Prałat, and Nicholas Wormald. Cleaning regular graphs with brushes. *SIAM Journal on Discrete Mathematics*, 23(1):233–250, 2009.

[3] József Balogh, Béla Bollobás, and Robert Morris. Graph bootstrap percolation. *Random Structures & Algorithms*, 41(4):413–440, 2012.

[4] Siddhartha Banerjee, Aditya Gopalan, Abhik Kumar Das, and Sanjay Shakkottai. Epidemic spreading with external agents. *IEEE Transactions on Information Theory*, 60(7):4125–4138, 2014.

[5] Amir Barghi and Peter Winkler. Firefighting on a random geometric graph. *Random Structures & Algorithms*, 46(3):466–477, 2015.

[6] Paul Bastide, Marthe Bonamy, Anthony Bonato, Pierre Charbit, Shahin Kamali, Théo Pierron, and Mikaël Rabie. Improved pyrotechnics: closer to the burning number conjecture. *Electronic Journal of Combinatorics*, 30(4):Paper No. 4.2, 12, 2023.

[7] Stéphane Bessy, Anthony Bonato, Jeannette Janssen, Dieter Rautenbach, and Elham Roshanbin. Burning a graph is hard. *Discrete Applied Mathematics*, 232:73–87, 2017.

[8] Stéphane Bessy, Anthony Bonato, Jeannette Janssen, Dieter Rautenbach, and Elham Roshanbin. Bounds on the burning number. *Discrete Applied Mathematics*, 235:16–22, 2018.

[9] Hans L. Bodlaender. On the complexity of some coloring games. *International Journal of Foundations of Computer Science*, 2(2):133–147, 1991.

[10] Anthony Bonato. A survey of graph burning. *Contributions to Discrete Mathematics*, 16(1):185–197, 2021.

[11] Anthony Bonato, Jeannette Janssen, and Elham Roshanbin. How to burn a graph. *Internet Mathematics*, 12(1-2):85–100, 2016.

[12] Anthony Bonato and Shahin Kamali. Approximation algorithms for graph burning. In *Theory and applications of models of computation*, volume 11436 of *Lecture Notes in Computer Science*, pages 74–92. Springer, Cham, 2019.

[13] Anthony Bonato and Thomas Lidbetter. Bounds on the burning numbers of spiders and path-forests. *Theoretical Computer Science*, 794:12–19, 2019.

[14] Anthony Bonato, Holden Milne, Trent G. Marbach, and Teddy Mishura. How to cool a graph. In *Modelling and Mining Networks*, volume 14671 of *Lecture Notes in Computer Science*, pages 115–129. Springer Nature Switzerland, 2024.

[15] Boštjan Brešar, Michael A. Henning, Sandi Klavžar, and Douglas F. Rall. *Domination games played on graphs*. SpringerBriefs in Mathematics. Springer, Cham, 2021.

[16] Nina Chiarelli, Vesna Iršič, Marko Jakovac, William B. Kinnersley, and Mirjana Mikalački. Burning game. *arXiv:2409.11328*, 2024.

[17] Sandip Das, Subhadeep Ranjan Dev, Arpan Sadhukhan, Uma kant Sahoo, and Sagnik Sen. Burning spiders. In *Algorithms and discrete applied mathematics*, volume 10743 of *Lecture Notes in Computer Science*, pages 155–163. Springer, Cham, 2018.

[18] Stephen Finbow and Gary MacGillivray. The firefighter problem: a survey of results, directions and questions. *The Australasian Journal of Combinatorics*, 43:57–78, 2009.

[19] Wayne Goddard and Michael A. Henning. The competition-independence game in trees. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 104:161–170, 2018.

[20] Dan Hefetz, Michael Krivelevich, Alon Naor, and Miloš Stojaković. On saturation games. *European Journal of Combinatorics*, 51:315–335, 2016.

[21] Michaela Hiller, Arie M. C. A. Koster, and Eberhard Triesch. *On the Burning Number of p-Caterpillars*, pages 145–156. Springer International Publishing, Cham, 2021.

[22] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Plenum Press, New York, 1972.

[23] Jeffrey O Kephart and Steve R White. Directed-graph epidemiological models of computer viruses. In *Computation: the micro and the macro view*, pages 71–102. World Scientific, 1992.

[24] Ralph Keusch. Colorability saturation games. *arXiv:1606.09099*, 2016.

[25] William B. Kinnersley, Douglas B. West, and Reza Zamani. Extremal problems for game domination number. *SIAM Journal on Discrete Mathematics*, 27(4):2090–2107, 2013.

[26] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.

[27] Max R. Land and Linyuan Lu. An upper bound on the burning number of graphs. In *Algorithms and models for the web graph*, volume 10088 of *Lecture Notes in Computer Science*, pages 1–8. Springer, Cham, 2016.

[28] Huiqing Liu, Xuejiao Hu, and Xiaolan Hu. Burning number of caterpillars. *Discrete Applied Mathematics*, 284:332–340, 2020.

[29] Sergey Norin and Jérémie Turcotte. The burning number conjecture holds asymptotically. *Journal of Combinatorial Theory. Series B*, 168:208–235, 2024.

[30] James B. Phillips and Peter J. Slater. An introduction to graph competition independence and enclaveless parameters. *Graph Theory Notes of New York*, 41:37–41, 2001.

[31] Julien Portier and Leo V. Versteegen. A Proof of the 3/4-Conjecture for the Total Domination Game. *SIAM Journal on Discrete Mathematics*, 39(1):1–18, 2025.

[32] Everett M Rogers, Arvind Singhal, and Margaret M Quinlan. Diffusion of innovations. In *An integrated approach to communication theory and research*, pages 432–448. Routledge, 2014.

[33] Yu. A. Zuev. Games on graph coloring. *Trudy Instituta Matematiki*, 23(2):56–61, 2015.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Colors of the Pseudotree

(Extended abstract)

David Chodounský * Monroe Eskew † Thilo Weinert ‡

Abstract

We investigate big Ramsey degrees of finite substructures of the universal countable homogeneous meet-tree and its binary variant. We prove that structures containing antichains have infinite big Ramsey degrees, and the big Ramsey degree of a 2-element chain is at least 8 and 7 for the binary variant. We deduce that the generic C-relation does not have finite big Ramsey degrees.

1 Big Ramsey degrees

We will provide examples of Ramsey classes for which the Fraïssé limit does not have finite big Ramsey degrees: the class of finite meet-trees and the class of finite C-relations. This contrasts with the Ramsey property of these classes (proved in [Deu75, BP10, BJP16]). So far only a few examples of this phenomenon have been discovered; the results in [Sau03, HKTZ25] is basically the complete list. Our main technique, which uses counting oscillations of functions, also found application in the upcoming paper [BCC⁺] on big Ramsey degrees of Boolean algebras.

The notion of a big Ramsey degree was first explicitly isolated in [KPT05], and this paper started a surge of results in this area; see [HZ24, Dob23] for recent surveys. Suppose A, B are model-theoretic structures. Denote the set of all substructures of B that are isomorphic to A as $\binom{B}{A}$. For $\ell \in \omega$ write $B \to (B)^A_\ell$ if for every coloring function $c\colon \binom{B}{A} \to \ell$ there exists $C \in \binom{B}{B}$

*Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, Praha 1, Czech Republic, and Department of Applied Mathematics (KAM), Charles University, Malostranské náměstí 25, Praha 1, Czech Republic. E-mail: chodounsky@math.cas.cz. Supported by project 25-15571S of the Czech Science Foundation (GAČR) and by the Czech Academy of Sciences CAS (RVO 67985840).

†Kurt Gödel Research Center, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria. Email: monroe.eskew@univie.ac.at. Supported by Austrian Science Fund (FWF) project P34603.

‡Kurt Gödel Research Center and Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF), Università degli Studi di Udine, via delle Scienze 206, 33100 Udine, Friuli, Italy. Email: aleph_omega@posteo.net. Supported by the Fonds zur Förderung der wissenschaftlichen Forschung, Lise Meitner grant M 3037-N, as well as the Research Project of National relevance "PRIN2022_DIMONTE - Models, sets and classifications(realizzato con il contributo del progetto PRIN 2022 - D.D. n. 104 del 02/02/2022 – PRIN2022_DIMONTE - Models, sets and classifications - Codice 2022TECZJA_003 - CUP N. G53D23001890006. "Finanziato dall'Unione Europea – Next-GenerationEU – M4 C2 I1.1")

such that $c\left[\binom{C}{A}\right] \neq \ell$. We call the smallest ℓ for which $B \to (B)^A_{\ell+1}$ the *big Ramsey degree* of A in B, write $\mathrm{d}(B : A) = \ell$. If no such $\ell \in \omega$ exists, let $\mathrm{d}(B : A) = \infty$.

We use the standard set-theoretic notation. We identify a natural number ℓ with the set $\{0, \dots, \ell - 1\}$. Given set x and function g, we denote the range of $g \upharpoonright x$ as $g[x]$.

2 The structure of universal homogeneous pseudotrees

Definition 1. A *pseudotree* is structure T with an order (reflexive, antisymmetric, transitive) relation $\leq$ and a binary function $\wedge$ satisfying $\sup\{ x \in T \mid x \leq a, x \leq b \} = a \wedge b$ for each $a, b \in T$ (in particular we demand this supremum to exist). Moreover, for each $a \in T$ the set $D_T(a) = \{ x \in T \mid x \leq a \}$ is linearly ordered by $\leq$.

An infinite pseudotree might not to be a tree, since the set $D_T(a)$ does not need to be well-ordered. As usual, we denote the derived strict version of the $\leq$ order by $<$. We say that a pseudotree T is *binary* if for every $a \in T$ and $x, y, z > a$ it is not the case that $a = x \wedge y = x \wedge z = y \wedge z$. In particular, if T is finite, then T is a tree equipped with the meet operation, and T is a binary pseudotree if it is a binary tree.

It is straightforward to check that the class of finite pseudotrees as well as the class of finite binary pseudotrees (with structure respecting embeddings) form a Fraïssé class. We denote the Fraïssé limit of finite pseudotrees by ψ_ω and the Fraïssé limit of finite binary pseudotrees by ψ_2. We write just ψ in cases where the given argument works indifferently for both of these objects, and we use the same letter to also denote the domain of the structure ψ.

These objects are also called universal homogeneous meet-trees in the literature; see, e.g. [KRS21].

Proposition 2. *The pseudotrees ψ_ω and ψ_2 can be up to an isomorphism characterized as a countable pseudotree with the following two properties.*

1. *For each $a \in \psi$ the set $D_\psi(a)$ is order isomorphic to the rational numbers $\mathbb{Q}$.*
2. *(a) For each $a \in \psi_\omega$ there exists a set $\{ x_i \in \psi_\omega \mid a < x_i, i \in \omega \}$ such that $x_i \wedge x_j = a$ for each $i \neq j \in \omega$.*

   (b) ψ_2 is binary and for each $a \in \psi_2$ there exist $x, y > a$ such that $x \wedge y = a$.

Definition 3. We say that $R \subset \psi$ is a *ray* in ψ if either

- $\inf R$ exists, denote it $o(R)$. Then $o(R) \notin R$ and R is a $\subseteq$-maximal linearly ordered subset of ψ with infimum $o(R)$, or
- $\inf R$ does not exist; in that case R is a $\subseteq$-maximal linearly ordered subset of ψ.

We denote the set of all rays in ψ by $\overline{\mathcal{R}}$. In either case, a ray is order-isomorphic to $\mathbb{Q}$.

We will fix a 1-to-1 *enumerating* function $f{:}\,\omega \to \psi \cup \overline{\mathcal{R}}$ with the following properties.

1. $\psi \cap f[n]$ is a pseudotree for each $n \in \omega$,
2. $\psi \subset f[\omega]$.
3. If $f(n) = R \in \overline{\mathcal{R}}$ and $o(R)$ is defined, then $o(R) \in f[n] \cap \psi$.
4. If $f(n) \in \psi$, then $f(n) \in \bigcup (f[n] \cap \overline{\mathcal{R}})$.
5. If $n \neq m$ and $f(n), f(m) \in \overline{\mathcal{R}}$, then $f(n) \cap f(m) = \varnothing$.

Denote the enumerated rays as $\mathcal{R} = f[\omega] \cap \overline{\mathcal{R}}$. Notice that necessarily $f(0) \in \mathcal{R}$ and $f(0)$ is the only element of $\mathcal{R}$ without an infimum. Moreover $\psi = \bigcup \mathcal{R}$. The following lemma has a straightforward proof.

Lemma 4. *There exists an enumerating function f with the stated properties.*

We denote by $e{:}\,\psi \cup \mathcal{R} \to \omega$ the bijection inverse to f. For $a \in \psi$ let $R(a)$ be the unique $R \in \mathcal{R}$ such that $a \in R$, and let $P(a) = \left\{ R \in \mathcal{R} \mid R \cap D_\psi(a) \neq \varnothing \right\}$ and $p(a) = \{ e(R) \mid R \in P(a) \}$. For $a \in \psi$ and $R \in P(a)$, $a \notin R$ define $a \wedge R$ to be the unique element S of $P(a)$ such that $o(S) \in R$.

Lemma 5. *For every $a \in \psi$ is $P(a) \subseteq f[e(a)]$. In particular $P(a)$ is finite.*

Proposition 6. *Suppose $\varphi' \in \binom{\psi}{\psi}$. Then there exists $\varphi \in \binom{\varphi'}{\psi}$ such that for every $R \in \mathcal{R}$ if $R \cap \varphi \neq \varnothing$, then $R \cap \varphi$ is isomorphic to $\mathbb{Q}$.*

Claim. *For every $b \in \varphi'$ there is $R \in \mathcal{R}$ such that $\{ x \in R \cap \varphi' \mid b < x \}$ contains a copy of $\mathbb{Q}$.*

Fix $z \in \varphi'$, $b < z$. Since $\{ x \in \varphi' \mid b < x < z \}$ contains a copy of $\mathbb{Q}$ and $P(z)$ is finite, there must be $R \in P(z)$ as required. ∎

For n such that $f(n) \in \mathcal{R}$ we inductively define order embeddings $g(n){:}\, f(n) \to \varphi'$; we aim for $g = \bigcup\{ g(n) \mid n \in \omega, f(n) \in \mathcal{R} \}$ being an embedding of ψ into φ' such that $\varphi = g[\psi]$ is as required. Start by choosing arbitrary $R \in \mathcal{R}$ such that $\varphi' \cap R$ contains a copy of $\mathbb{Q}$, and let $g(0){:}\, f(0) \to \varphi' \cap R$ be an embedding. Suppose $f(n) \in \mathcal{R}$, and $g(i)$ is defined for all $i < n$, $f(i) \in \mathcal{R}$. Write $h(n) = \bigcup\{ g(i) \mid i < n \}$ and $a' = o(f(n))$. Letting $a = h(n)(a')$, since φ' is isomorphic to ψ there exists $b \in \varphi'$ such that $o(b \wedge R(a)) = a$ and $b \wedge x \leq a$ for all x in the range of $h(n)$. The claim implies that there is $R \in \mathcal{R}$ such that we can fix an embedding $g(n){:}\, f(n) \to \{ x \in R \cap \varphi' \mid b < x \}$. At the end $\bigcup\{ g(n) \mid n \in \omega, f(n) \in \mathcal{R} \}$ is as required. □

3 Indivisibility

A structure is *indivisible* if the big Ramsey degree of singletons is equal to 1. We prove that ψ is indivisible. Let S be the single element pseudotree.

Theorem 7. $\mathrm{d}(\psi : \mathrm{S}) = 1$.

Proof. Suppose $C \subseteq \psi$. We will find $\varphi \in \binom{\psi}{\psi}$ such that either $\varphi \subseteq C$, or $\varphi \cap C = \varnothing$.

Claim. *Either for every $b \in \psi$ there is $R \in \mathcal{R}$ such that $\{ x \in R \cap C \mid b < x \}$ contains a copy of $\mathbb{Q}$, or there is $y \in \psi$ such that for every $b > y$ there is $R \in \mathcal{R}$ such that $\{ x \in R \smallsetminus C \mid b < x \}$ contains a copy of $\mathbb{Q}$.*

Suppose that the first alternative fails and this is witnessed by $y \in \psi$, we will argue that the second alternative holds. Choose any $b > y$, let $R \in \mathcal{R}$ be such that $o(R) = b$. For any $u, v \in R$ with $u < v$, the set $\{ x \mid u < x < v \}$ is isomorphic to $\mathbb{Q}$ and therefore there is $w \in R \smallsetminus C$, $u < w < v$, i.e. $R \smallsetminus C$ is a countable dense linear order and contains a copy of $\mathbb{Q}$. ∎

The rest of the proof follows verbatim the proof of Proposition 6; if the first alternative of the claim occurs, construct $\varphi \subseteq C$. In the second alternative start with choosing $g(0)$ such that $x > y$ for each x in the range of $g(0)$, and construct φ disjoint with C. □

4 A coloring of chains

Suppose $a, b \in \psi$ form a chain, that is $a < b$. We will denote the 2-element pseudotree isomorphic to $\{a, b\}$ as C. We define the value of a coloring function $c(a, b)$ as follows.

Case 1; $R(a) = R(b)$. Let $c(a, b) = 0$ if $e(a) < e(b)$, and let $c(a, b) = 1$ if $e(b) < e(a)$.

Case 2; $R(a) \neq R(b)$. In this case, it must be $e(o(b \wedge R(a))) < e(b \wedge R(a)) \leq e(R(b)) < e(b)$. If $b \wedge R(a) = R(b)$ we let $c(a, b) = \mathrm{u}$. Otherwise, define $c(a, b)$ by distinguishing subcases.

(2) $e(a) < e(o(b \wedge R(a)))$, let $c(a, b) = 2$
(3) $a = o(b \wedge R(a))$, let $c(a, b) = 3$
(e) $e(o(b \wedge R(a))) < e(a) < e(b \wedge R(a))$, let $c(a, b) = \mathrm{e}$
(4) $e(b \wedge R(a)) < e(a) < e(R(b))$, let $c(a, b) = 4$
(5) $e(R(b)) < e(a) < e(b)$, let $c(a, b) = 5$
(6) $e(b) < e(a)$, let $c(a, b) = 6$

Theorem 8. *Let $s = 7$ if ψ is ψ_2, or $s = 7 \cup \{\mathrm{e}\}$ if ψ is ψ_ω. For every $\varphi \in \binom{\psi}{\psi}$ and $c \in s$ there are $a < b \in \varphi$ such that $c(a, b) = c$.*

Proof. We can assume that φ is as in the conclusion of Proposition 6. Fix $a_2 \in \varphi$ arbitrary. Since $R = R(a_2) \cap \varphi$ is isomorphic $\mathbb{Q}$, we can find $a_3, a_\mathrm{e} \in \varphi \cap R$ such that $a_2 < a_3$, $a_\mathrm{e} < a_3$, and $e(a_2) < e(a_3) < e(a_\mathrm{e})$.

There is $x \in \varphi \smallsetminus R$, $a_3 < x$, $o(x \wedge R) = a_3$, and if ψ is ψ_ω, then also $e(x \wedge R) > e(a_\mathrm{e})$. Find $a_4 \in D_\varphi(a_3) \cap R$ such that $e(x \wedge R) < e(a_4)$. Fix $y \in \varphi$, $y > x$ such that $e(R(y)) > e(a_4)$. Find $a_5 \in D_\varphi(a_3) \cap R$ such that $e(R(y)) < e(a_5)$. Finally fix $b \in \varphi \cap R(y)$ such that $e(b) > e(a_5)$, and $a_6 \in D_\varphi(a_3) \cap R$ such that $e(b) < e(a_6)$.

The construction yields $c(a_2, a_3) = 0$, $c(a_3, a_4) = 1$, $c(a_2, b) = 2$, $c(a_3, b) = 3$, $c(a_4, b) = 4$, $c(a_5, b) = 5$, $c(a_6, b) = 6$, and $c(a_\mathrm{e}, b) = \mathrm{e}$ in case ψ is ψ_ω. □

Corollary 9. $\mathrm{d}(\psi_2 : \mathrm{C}) \geq 7$, $\mathrm{d}(\psi_\omega : \mathrm{C}) \geq 8$.

In fact, $\mathrm{d}(\psi_2 : \mathrm{C}) = 7$, $\mathrm{d}(\psi_\omega : \mathrm{C}) = 8$; this will be proved in forthcoming publications, e.g. [CDW25]. However, the methods to prove this are beyond the scope of this abstract.

5 The oscillation on antichains

Denote the three element pseudotree consisting of a root and two incomparable nodes as A. In this section, we prove that $\mathrm{d}(\psi_2 : \mathrm{A}) = \mathrm{d}(\psi_\omega : \mathrm{A}) = \infty$. In order to do this, we use the method of oscillation which has seen widespread use in set theory, see e.g. [Tod88]. Our application was inspired by an argument about the product of Mathias posets; see [BGA17, Observation 7].

Suppose $u_0, u_1 \subset \omega$ are two disjoint finite sets. Let $\simeq$ be the equivalence relation on $u_0 \cup u_1$ defined by declaring $n \simeq m$ if there is $i \in 2$ such that $m, n \in u_i$ and there is no $k \in u_{1-i}$ such that $n < k < m$ or $m < k < n$. Define $\mathrm{osc}(u_0, u_1)$ to be the number of equivalence classes of $\simeq$

$$\mathrm{osc}(u_0, u_1) = |\{ [n]_\simeq \mid n \in u_0 \cup u_1 \}|.$$

For a pair of incomparable elements a, b of ψ define $\mathrm{os}(a, b) = \mathrm{osc}(p(a) \smallsetminus p(b), p(b) \smallsetminus p(a))$. Notice that for any $a, b \in \psi$ the set $p(a) \cap p(b)$ is always an initial subset of $p(a)$.

Theorem 10. *Let $\varphi \in \binom{\psi}{\psi}$. For every $\ell \in \omega \smallsetminus 1$ there are incomparable $a, b \in \varphi$ such that $\mathrm{os}(a, b) = \ell$. In particular, $\mathrm{d}(\psi_2 : \mathrm{A}) = \mathrm{d}(\psi_\omega : \mathrm{A}) = \infty$.*

Proof. We again assume that φ is as in the conclusion of Proposition 6. We prove the theorem by induction on ℓ. Choose any $a_1, b_1, r \in \varphi$ such that $a_1 \wedge b_1 = r < a_1, b_1$, and $R(r) = R(a_1)$. Now $p(a_1)$ is a proper initial subset of $p(b_1)$, i.e. $\mathrm{os}(a_1, b_1) = 1$.

Suppose $a_\ell, b_\ell \in \varphi$ are incomparable and $\mathrm{os}(a_\ell, b_\ell) = \ell$. We may assume without loss of generality $\max(p(a_\ell) \cup p(b_\ell)) \in p(a_\ell)$. Since $R(b_\ell) \cap \varphi$ is isomorphic to $\mathbb{Q}$, there is $y \in \varphi \cap R(b_\ell)$ such that $e(y) > e(a_\ell)$. Let $b_{\ell+1}$ be an element of φ such that $o(b_{\ell+1} \wedge R(b_\ell)) = y$. Then $p(b_{\ell+1})$ is a proper end-extension of $p(b_\ell)$, $(p(b_{\ell+1}) \smallsetminus p(b_\ell)) \cap e(a_\ell) = \varnothing$, and for $a_{\ell+1} = a_\ell$ we get $\mathrm{os}(a_{\ell+1}, b_{\ell+1}) = \ell + 1$. □

6 The generic C-relation

The C-relation introduced in [AN98] is a ternary relation with axioms that describe the behavior of leaves of a finite pseudotree. More precisely, a finite set L equipped with a ternary relation C is a C-relation structure (or just a *C-relation*) if there is a finite pseudotree T such that the set L consists of maximal elements of T, and $C(a; b, c)$ exactly when $a \wedge b < b \wedge c$.

The class of finite C-relations and binary C-relations (i.e. corresponding to binary pseudotrees) are both Fraïssé classes, denote their limits $\mathscr{C}_\omega$ and $\mathscr{C}_2$, both of these objects can be characterized as countable C-relations satisfying certain first order axioms, see [BJP16]. Sam Braunfeld observed [Bra] that our results for ψ imply that $\mathscr{C}_\omega$ and $\mathscr{C}_2$ do not have finite big Ramsey degrees.

We will sketch the argument. Let C_4 be the C-relation on 4 elements $\{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\}$ uniquely determined by declaring $C(\mathrm{a}; \mathrm{c}, \mathrm{d})$, $C(\mathrm{b}; \mathrm{c}, \mathrm{d})$, and $C(\mathrm{c}; \mathrm{a}, \mathrm{b})$.

Theorem 11 (Braunfeld)**.** $\mathrm{d}(\mathscr{C}_2 : \mathrm{C}_4) = \mathrm{d}(\mathscr{C}_\omega : \mathrm{C}_4) = \infty$.

Proof. For $R \in \mathcal{R}$ choose any $r \in R$ and let $D(R) = R \cup D(r)$, and for $R \neq S \in \mathcal{R}$ let $R \wedge S = \max(D(R) \cap D(S))$. For $R, S, T \in \mathcal{R}$ define $C(R; S, T)$ if $R \wedge S < S \wedge T$ or $R \neq S = T$. Using the axioms from [BJP16] it is easy to verify that $(\mathcal{R}, C)$ is isomorphic to $\mathscr{C}_2$ if $\psi = \psi_2$, and $\mathscr{C}_\omega$ if $\psi = \psi_\omega$. Moreover, whenever $\mathcal{S} \in \binom{\mathcal{R}}{\mathcal{R}}$ then $\{ S \wedge R \mid S, R \in \mathcal{S} \} \in \binom{\psi}{\psi}$.

Let $X = \{ a, b, c, d \} \in \binom{\mathcal{R}}{\mathrm{C}_4}$. Define $c(X) = \mathrm{os}(a \wedge b, c \wedge d)$. It follows from Theorem 10 that whenever $\mathcal{S} \in \binom{\mathcal{R}}{\mathcal{R}}$ and $\ell \in \omega \smallsetminus 1$, there is $X \in \binom{\mathcal{S}}{\mathrm{C}_4}$ such that $c(X) = \ell$. □

References

[AN98] S. A. Adeleke and Peter M. Neumann, *Relations related to betweenness: their structure and automorphisms*, Mem. Amer. Math. Soc. **131** (1998), no. 623, viii+125. MR 1388893

[BCC+] Dana Bartošová, David Chodounský, Barbara Csima, Jan Hubička, Matěj Konečný, Joey Lakerdas-Gayle, Spencer Unger, and Andy Zucker, *Oscillating subalgebras of the atomless countable Boolean algebra*, `https://arxiv.org/abs/2505.22603`.

[BGA17] Jörg Brendle and Luz María García Ávila, *Forcing-theoretic aspects of Hindman's theorem*, J. Math. Soc. Japan **69** (2017), no. 3, 1247–1280. MR 3685044

[BJP16] Manuel Bodirsky, Peter Jonsson, and Trung Van Pham, *The reducts of the homogeneous binary branching C-relation*, J. Symb. Log. **81** (2016), no. 4, 1255–1297. MR 3579112

[BP10] Manuel Bodirsky and Diana Piguet, *Finite trees are ramsey under topological embeddings*, 2010, https://arxiv.org/abs/1002.1557.

[Bra] Samuel Braunfeld, personal communication.

[CDW25] David Chodounský, Natasha Dobrinen, and Thilo Weinert, *Big Ramsey degrees and the 2-branching pseudotree*, 2025, https://arxiv.org/abs/2503.22626.

[Deu75] Walter Deuber, *A generalization of Ramsey's theorem for regular trees*, J. Combinatorial Theory Ser. B **18** (1975), 18–23. MR 437370

[Dob23] Natasha Dobrinen, *Ramsey theory of homogeneous structures: current trends and open problems*, ICM—International Congress of Mathematicians. Vol. 3. Sections 1–4, EMS Press, Berlin, [2023] ©2023, pp. 1462–1486. MR 4680287

[HKTZ25] Jan Hubička, Matěj Konečný, Stevo Todorcevic, and Andy Zucker, *On big Ramsey degrees of universal ω-edge-labeled hypergraphs*, 2025, to appear in this volume, https://arxiv.org/abs/2505.22561.

[HZ24] Jan Hubička and Andy Zucker, *A survey on big Ramsey structures*, 2024, https://arxiv.org/abs/2407.17958.

[KPT05] A. S. Kechris, V. G. Pestov, and S. Todorcevic, *Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups*, Geom. Funct. Anal. **15** (2005), no. 1, 106–189. MR 2140630

[KRS21] Itay Kaplan, Tomasz Rzepecki, and Daoud Siniora, *On the automorphism group of the universal homogeneous meet-tree*, J. Symb. Log. **86** (2021), no. 4, 1508–1540. MR 4362923

[Sau03] N. W. Sauer, *Canonical vertex partitions*, vol. 12, 2003, Special issue on Ramsey theory, pp. 671–704. MR 2037079

[Tod88] Stevo Todorčević, *Oscillations of real numbers*, Logic colloquium '86 (Hull, 1986), Stud. Logic Found. Math., vol. 124, North-Holland, Amsterdam, 1988, pp. 325–331. MR 922115

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Universality for transversal Hamilton cycles in random graphs

(Extended abstract)

Micha Christoph* Anders Martinsson† Aleksa Milojević‡

Abstract

A tuple $(G_1, \ldots, G_n)$ of graphs on the same vertex set of size n is said to be Hamilton-universal if for every map $\chi : [n] \to [n]$ there exists a Hamilton cycle such that the i-th edge comes from $G_{\chi(i)}$. Bowtell, Morris, Pehova and Staden proved an analog of Dirac's theorem in this setting, namely that if $\delta(G_i) \geq (1/2 + o(1))n$ then $(G_1, \ldots, G_n)$ is Hamilton-universal. Combining McDiarmid's coupling and a colorful version of the Friedman-Pippenger tree embedding technique, we establish a similar result in the sparse setting of random graphs, showing that there exists C such that if the G_i are independent random graphs sampled from $G(n,p)$, where $p \geq C \log n/n$, then $(G_1, \ldots, G_n)$ is Hamilton-universal with high probability.

1 Introduction

One of the most fundamental questions in graph theory is that of determining whether a given graph contains a Hamilton cycle, i.e. a cycle which contains all vertices of the graph. There is no simple algorithm to check whether a graph contains such a cycle. The problem of finding a Hamilton cycle in a given graph is famously NP-complete. Therefore, researchers focused on finding easy-to-check conditions which ensure that a graph contains such a cycle, and perhaps even allow for finding the cycle in polynomial time. The most famous condition of the kind is Dirac's condition [6], dating back to 1952, which states that every n-vertex graph of minimum degree at least $n/2$ must contain a Hamilton cycle.

Although Dirac's condition is undoubtedly elegant, and tight in the sense that the degree requirement $n/2$ cannot be lowered, it is quite restrictive. For example, it can only be satisfied by dense graphs which have quadratically many edges. It is, therefore, natural to ask what

*Department of Mathematics, ETH Zürich, Switzerland. E-mail: micha.christoph@math.ethz.ch Supported by SNSF Ambizione Grant No. 216071.

†Department of Computer Science, ETH Zürich, Switzerland.
E-mail: anders.martinsson@inf.ethz.ch

‡ETH Zürich, Switzerland. E-mail: aleksa.milojevic@math.ethz.ch. Supported by SNSF Grant 200021-228014.

kind of conditions guarantee the existence of Hamilton cycles in sparser graphs. A common family of sparse graphs to study are the random graphs.

The study of Hamilton cycles in random graphs is almost as old as the study of random graphs. In their original paper from 1960 [7], Erdős and Rényi asked for the number of edges at which one expects to start seeing a Hamilton path in a random graph $G(n, m)$. This question was answered by Posá [16], who showed that one can expect Hamilton cycles in random graphs with $\Omega(n \log n)$ edges, by introducing a very useful and elegant technique now known as the Posá rotation. His argument was later tightened by Komlós and Szemerédi [12], and Bollobás [4] and Ajtai, Komlós and Szemerédi [2], to ultimately attain very good understanding of Hamilton cycles in random graphs. Namely, Bollobás as well as Ajtai, Komlós and Szemerédi showed that if we order the edges of K_n randomly and start adding them one by one to our graph, the time at which the graph becomes Hamiltonian with high probability coincides with the first time at which no vertex has degree less than 2, which is clearly a necessary condition for the graph to be Hamiltonian.

In recent years, the study of Hamilton cycles and other spanning structures in graphs has taken an interesting turn, by considering the colorful variants of classical questions from extremal graph theory. In particular, the setup we will be studying was introduced by Joos and Kim in [11] while answering a question of Aharoni [1], and it goes as follows. Given a collection of m graphs $(G_1, \ldots, G_m)$, which we think of as colors, on the same vertex set V and a graph H with vertices in V, we say that $(G_1, \ldots, G_m)$ contains a *rainbow (or transversal) copy* of H if there exists an injective function $\psi : E(H) \to [m]$ such that the edge e is contained in the graph $G_{\psi(e)}$ for all $e \in E(H)$. In other words, the graph H can be constructed by picking at most one edge from each of the graphs G_i. Joos and Kim showed that if each of the graphs G_i is an n-vertex graph with minimum degree $n/2$ and if $m \geq n$, then there exists a transversal Hamilton cycle in the collection $(G_1, \ldots, G_m)$.

This work inspired a large number of papers on the subject, for example studying the existences of transversal cliques (see [1]), transversal F-factors (see [15]) and transversal subgraphs of random graphs (see [3]). We recommend the survey of Sun, Wang and Wei [17] to the interested reader who would like to learn more about the transversal subgraphs of graph collections. In this paper, we will consider the question of universality for Hamilton cycles in collections of random graphs.

The notion of universality was introduced by Bowtell, Morris, Pehova and Staden in [5]. We say that a collection of graphs $(G_1, \ldots, G_n)$ on the same vertex set of size n is *Hamilton-universal* if for each map $\chi : [n] \to [n]$ there exists a Hamilton cycle whose i-th edge lies in $G_{\chi(i)}$. Bowtell, Morris, Pehova and Staden [5] showed that every collection $(G_1, \ldots, G_n)$ of n graphs on the same vertex set of size n satisfying $\delta(G_i) \geq (1/2 + o(1))n$ is Hamilton-universal. Note that the difference from the result of Joos and Kim is that the edges of the Hamilton cycle must now come from a prescribed member of the collection $(G_1, \ldots, G_n)$, rather than an arbitrary one.

The aim of this paper is to extend the result of Bowtell, Morris, Pehova and Staden to the setting of random graphs, answering a question of Pehova from the 30th British Combinatorial Conference.

Theorem 1.1. *There exists C such that for $p \geq C \log n/n$ the following holds with high probability. Let $(G_1, \ldots, G_n)$ be a tuple of independent random graphs on the same vertex set sampled from $G(n, p)$. Then, $(G_1, \ldots, G_n)$ is Hamilton-universal.*

Note that the above result is tight up to the constant C since G_1 is required to contain a Hamilton cycle for $(G_1, \dots, G_n)$ to be Hamilton-universal. While showing the existence of a Hamilton cycle with any specific coloring is sufficient for Hamilton-universality in the deterministic setting, dealing with the looming union bound over all colorings is the key obstacle to overcome towards Theorem 1.1.

1.1 Proof sketch

We will now go through a short outline of the proof of our result. The tool underlying our proof, called McDiarmid's coupling, is a method which allows us to go from Hamilton cycles in a single random graph to rainbow Hamilton cycles in a collection of n independent random graphs. Introduced by McDiarmid [13] in 1980, the coupling was mostly overlooked in the random graph community until its recent revival by Ferber [8], Ferber and Long [9] and Montgomery [14].

The gist of the method is encapsulated in the following statement, which is essentially due to McDiarmid [13], albeit in a slightly different language. If $\mathcal{F}$ is a family of ordered n-tuples with distinct elements from a ground set E and $S_0, \dots, S_n$ are random subsets of E including each element independently with probability p, then $S_1 \times \cdots \times S_n$ is at least as likely to contain an n-tuple of $\mathcal{F}$ as $S_0 \times \cdots \times S_0$ is. In our setting, one should think of the ground set E as the collection of all edges of the complete graph on n vertices, thus making the random sets $S_0, \dots, S_n$ Erdős-Rényi random graphs, while $\mathcal{F}$ will be a collection of tuples of edges, roughly speaking, representing the Hamilton cycles of the complete graph. However, the actual statement we use is slightly more general, since we need a bit more flexibility.

Lemma 1.2. *Let $n \geq 1$ be a positive integer, $p \in [0,1]$ and E a set. Let $\mathcal{F}$ be a family of ordered n-tuples from $E \cup \{\star\}$ containing each element of E at most once. Let $S'_0, \dots, S'_n$ be i.i.d. random subsets of E, where the element $e \in E$ is included in S'_i with probability p and let $S_i = S'_i \cup \{\star\}$. Then, for any map $\chi : [n] \to [n]$, the probability that $S_0 \times \cdots \times S_0$ contains an n-tuple of $\mathcal{F}$ is smaller or equal to the probability that $S_{\chi(1)} \times \cdots \times S_{\chi(n)}$ contains an n-tuple of $\mathcal{F}$.*

Here is a simple application of Lemma 1.2. Let $\mathcal{C}$ denote the set of edge-ordered Hamilton cycles of the complete graph on n vertices. That is $\mathcal{C}$ is the set of n-tuples of edges $(e_1, \dots, e_n)$ forming a Hamilton cycle in this order. If we apply Lemma 1.2 with $\mathcal{F} = \mathcal{C}$ and any fixed map $\chi : [n] \to [n]$, it follows that a collection of n independent random graphs is at least as likely to contain a χ-colored Hamilton cycle as a single random graph is to contain a Hamilton cycle. Therefore, a collection of n independent random graphs $G(n, p)$ with high probability contains a Hamilton cycle in a given color-pattern already at probability $p \gg \log n/n$.

Note that in order to prove Theorem 1.1, we need to find a Hamilton cycle for *every* color pattern. We would therefore like to apply a union bound over the set of all color patterns. However, the random graph $G(n, C \log n/n)$ fails to have a Hamilton cycle with probability at least n^{-C}, since it has isolated vertices with at least this probability. Therefore, applying McDiarmid's coupling together with a simple union bound is not enough to complete the proof.

The way to circumvent this issue and boost the probability of success is to show that with *very* high probability, one can remove up to εn vertices from $G(n, C \log n/n)$ so that the rest is Hamilton-connected, for some small constant $\varepsilon > 0$. Before we state things more precisely, it will be very convenient for us to partition the set of vertices into L and R with $|L| = \lfloor n/2 \rfloor$ and $|R| = \lceil n/2 \rceil$.

Lemma 1.3. *For every $\varepsilon > 0$ there exists a constant $C > 0$ such that the following holds. If $G \sim G(n,p)$ is a random graph where $p \geq C/n$, with probability at least $1 - \exp(-\Omega_\varepsilon(n^2p))$ we have the following. There exists some set X of at most εn vertices such that for every set $Y \subseteq L$, the induced subgraph of G on $V(G)\backslash(X \cup Y)$ is Hamilton-connected.*

This lemma gives us not only a way of finding almost-spanning paths in G, but we are also allowed to pick and choose which vertices of L we want to include in our path – something we will find very useful later on. However, the almost-spanning path we construct will not be able to cover vertices of X – for that, we need a different argument based on a colorful version of the Friedman-Pippenger tree embedding technique first introduced in [10]. Therefore, we need a final lemma, which allows us to cover sets of at most εn vertices using a path with prescribed edges colors. Given a tuple $(G_1, \dots, G_n)$ of graphs and a map $\chi : [n] \to [n]$, we say that a Hamilton cycle is χ-*colored* if the i-th edge is in $G_{\chi(i)}$ and, similarly, a (not necessarily Hamiltonian) path is χ-colored if the i-th edge is in $G_{\chi(i)}$.

Lemma 1.4. *There exists C such that for $p \geq C \log n/n$ the following holds with high probability. Let $(G_1, \dots, G_n)$ be a tuple of independent random graphs on the same vertex set sampled from $G(n,p)$. Then, for every map $\chi : [n] \to [n]$ and every set X of at most εn vertices, there exists a χ-colored path P with $X \subseteq V(P) \subseteq X \cup L$ such that both endpoints of P are outside X.*

Combining the three lemmas then allows us to prove Theorem 1.1. Instead of sampling the graphs from $G(n,p)$, the idea is to generate two random graphs $G_i, H_i \sim G(n,p')$ with $p' = 1 - \sqrt{1-p} \geq p/2$, independently for each i. Then, the graph $G_i \cup H_i$ follows the distribution of $G(n,p)$. Hence, to prove the statement of the theorem, it suffices to show that the n-tuple of graphs $(G_1 \cup H_1, \dots, G_n \cup H_n)$ is Hamilton-universal with high probability. We will then fix the tuple $(G_1, \dots, G_n)$ and condition on it satisfying the conclusion of Lemma 1.4. Finally, a combination of Lemma 1.2 and 1.3 will show that for every $\chi : [n] \to [n]$, with probability at least $1 - n^{2n}$, the tuple $(H_1, \dots, H_n)$ is such that $(G_1 \cup H_1, \dots, G_n \cup H_n)$ contains a χ-colored Hamilton cycle, which concludes the proof via a union bound over all χ.

2 Concluding remarks

In this paper, we determined the threshold probability p for an n-tuple of independent random graphs $(G_1, \dots, G_n)$ sampled from $G(n,p)$ to be Hamilton-universal. As we explained in the introduction, this result is a variant of a theorem by Bowtell, Morris, Pehova and Staden, who showed that any n-tuple of graphs $(G_1, \dots, G_n)$ with $\delta(G_i) \geq (1/2 + o(1))n$ is Hamilton-universal. The fact that the G_i are sampled independently in the random setting could lead to some intriguing differences.

For example, since the classical Dirac's condition for the containment of Hamilton cycle ($\delta(G) \geq n/2$) is tight, we must require at least the same condition from all graphs of the n-tuple in the deterministic setting. The reason for this is simple – if G does not contain a Hamilton cycle, then $(G, \dots, G)$ does not contain a transversal Hamilton cycle either. Hence, in the transversal setting, the conditions for containment of a given structure cannot be weaker than the corresponding condition in the classical setting.

Once randomness is introduced, this behavior changes dramatically. For example, Anastos and Chakraborti [3] showed that if $(G_1, \dots, G_n)$ is an n-tuple of independent graphs, where $G_i \sim G(n,p)$ with $p \gg \log n/n^2$, then $(G_1, \dots, G_n)$ contains a transversal Hamilton cycle with high probability. This differs starkly from the threshold probability for Hamiltonicity in $G(n,p)$ at $p = (1+o(1))\log n/n$. Therefore, the n-tuple $(G_1,\dots,G_n)$ contains a transversal Hamilton cycle even when no G_i is expected to be Hamiltonian. This shows that looking for transversal structures in the random setting, as opposed to the deterministic setting, can be interesting even when none of the graphs G_i have the required subgraph themselves.

We wonder if we can observe a similar effect in questions regarding universality as well. Though Theorem 1.1 is tight in the stated version, we can not rule out that it can be improved when only considering rainbow Hamilton cycles. More precisely, we say that a tuple $(G_1,\dots,G_n)$ of graphs on the same vertex set of size n is *rainbow Hamilton-universal* if it contains a χ-colored Hamilton cycle for every bijection $\chi : [n] \to [n]$. Note the subtle difference to Hamilton-universality, where χ is not required to be bijective.

Question 2.1. *What is the threshold probability p such that the following holds with high probability? If $(G_1,\dots,G_n)$ is an n-tuple of independent random graphs on the same vertex set sampled from $G(n,p)$, then $(G_1,\dots,G_n)$ is rainbow Hamilton-universal.*

The best lower bound we can prove is of the form $p \geq c/n$, for which we have two different arguments. Firstly, if $c < e$, then the expected number of copies of a Hamilton cycle with a specific rainbow color-pattern is $n!p^n = (1+o(1))\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(\frac{c}{n}\right)^n = o(1)$. Secondly, if $c < \log(2)$ then with high probability there exists a vertex v which is isolated in more than half of the G_i. In this case, there exists a bijection $\chi : [n] \to [n]$ such that v is isolated in $G_{\chi(i)}$ or $G_{\chi(i+1)}$ for all $i \in [n]$ and therefore, $(G_1,\dots,G_n)$ does not contain a χ-colored Hamilton cycle. Curiously, the first lower bound has a better constant than the second, so that the expectation gives a better bound than the threshold for a specific vertex being bad. This is different from the usual behavior of Hamilton cycles in random graphs, where we can expect to find a Hamilton cycle as soon as every vertex has degree at least 2.

It would also be interesting to develop a universal analog to the results of [3].

Problem 2.2. *Show that there exists C such that for $p \geq C\log n/n$ the following holds with high probability. Let $(G_1,\dots,G_n)$ be a tuple of graphs on the same vertex set of size n such that $\delta(G_i) \geq (1/2+o(1))n$. For every i, let $F_i \subseteq G_i$ be obtained by keeping every edge independently with probability p. Then, $(F_1,\dots,F_n)$ is Hamilton-universal.*

References

[1] Ron Aharoni, Matt DeVos, Sebastián González Hermosillo de la Maza, Amanda Montejano, and Robert Šámal. A rainbow version of Mantel's Theorem. *Advances in Combinatorics*, feb 28 2020.

[2] Miklós Ajtai, János Komlós, and E Szemerédi. First occurrence of hamilton cycles in random graphs. *North-Holland Mathematics Studies*, 115(C):173–178, 1985.

[3] Michael Anastos and Debsoumya Chakraborti. Robust hamiltonicity in families of dirac graphs. *arXiv preprint arXiv:2309.12607*, 2023.

[4] Béla Bollobás. The evolution of sparse graphs. *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57, 1984.

[5] Candida Bowtell, Patrick Morris, Yanitsa Pehova, and Katherine Staden. Universality for transversal hamilton cycles. *Bulletin of the London Mathematical Society*, 2024.

[6] Gabriel Andrew Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[7] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[8] Asaf Ferber. Closing gaps in problems related to hamilton cycles in random graphs and hypergraphs. *The Electronic Journal of Combinatorics*, pages P1–61, 2015.

[9] Asaf Ferber and Eoin Long. Packing and counting arbitrary hamilton cycles in random digraphs. *Random Structures & Algorithms*, 54(3):499–514, 2019.

[10] Joel Friedman and Nicholas Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7:71–76, 1987.

[11] Felix Joos and Jaehoon Kim. On a rainbow version of dirac's theorem. *Bulletin of the London Mathematical Society*, 52(3):498–504, 2020.

[12] János Komlós and Endre Szemerédi. Limit distribution for the existence of hamiltonian cycles in a random graph. *Discrete Mathematics*, 43(1):55–63, 1983.

[13] Colin McDiarmid. Clutter percolation and random graphs. *Math. Programming Stud.*, (13):17–25, 1980.

[14] Richard Montgomery. Spanning cycles in random directed graphs. *Random Structures & Algorithms*, 65(3):535–575, 2024.

[15] Richard Montgomery, Alp Müyesser, and Yanitsa Pehova. Transversal factors and spanning trees. *arXiv preprint arXiv:2107.04629*, 2021.

[16] Lajos Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[17] Wanting Sun, Guanghui Wang, and Lan Wei. Transversal structures in graph systems: A survey. *arXiv preprint arXiv:2412.01121*, 2024.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Colouring t-perfect graphs

(Extended abstract)

Maria Chudnovsky* Linda Cook† James Davies‡ Sang-il Oum§
Jane Tan¶

Abstract

Perfect graphs can be described as the graphs whose stable set polytopes are defined by their non-negativity and clique inequalities (including edge inequalities). In 1975, Chvátal defined an analogous class of t-perfect graphs, which are the graphs whose stable set polytopes are defined by their non-negativity, edge inequalities, and odd circuit inequalities. We show that t-perfect graphs are 199053-colourable. This is the first finite bound on the chromatic number of t-perfect graphs and answers a question of Shepherd from 1995. Our proof also shows that every h-perfect graph with clique number ω is $(\omega + 199050)$-colourable.

1 Introduction

Let $G = (V, E)$ be a graph. A *stable* set of G is a set of pairwise non-adjacent vertices. We write $\chi(G)$ for the *chromatic number* of G, that is, the minimum number of colours needed to colour the vertices of G so that no two adjacent vertices have the same colour. Equivalently, $\chi(G)$ is the minimum number of stable sets needed to cover the vertex set of G. A *clique* is a

*Department of Mathematics, Princeton University, Princeton, USA. E-mail: mchudnov@math.princeton.edu. Supported by NSF Grant DMS-2348219 and by AFOSR grant FA9550-22-1-0083.

†Korteweg-de Vries Institute for Mathematics, University of Amsterdam, The Netherlands. E-mail: l.j.cook@uva.nl. Supported by the Gravitation programme NETWORKS (NWO grant no. 024.002.003) of the Dutch Ministry of Education, Culture and Science (OCW) and a Marie Skłodowska-Curie Action of the European Commission (COFUND grant no. 945045).

‡Trinity Hall, University of Cambridge, United Kingdom. E-mail: jgd37@cam.ac.uk.

§Discrete Mathematics Group, Institute for Basic Science (IBS), Daejeon, South Korea. E-mail: sangil@ibs.re.kr. Supported by the Institute for Basic Science (IBS-R029-C1).

¶All Souls College, University of Oxford, United Kingdom. E-mail: jane.tan@maths.ox.ac.uk.

set of vertices that are pairwise adjacent. We write $\omega(G)$ for the size of the largest clique in a graph G, called the *clique number* of G. A graph is *triangle-free* if it does not have a clique of size three.

The stable set problem is the problem of finding the maximum size of a stable set in a graph. Lovász [Lov94] described it as one of the simplest and most fundamental problems concerning graphs. Since the stable set problem is NP-hard in general, it is natural to restrict to input graphs on which the stable set problem can be solved efficiently.

One approach is to consider the polytope generated by stable sets and use techniques from linear programming. For a subset S of V, we write $\chi^S \in \mathbb{R}^V$ to denote its *incidence vector*, that is, a 0-1 vector such that $\chi^S(v) = 1$ if $v \in S$ and $\chi^S(v) = 0$ otherwise. The *stable set polytope* of a graph $G = (V, E)$ is defined as the convex hull of the incidence vectors of the stable sets of G. We denote it by $\mathrm{SSP}(G)$. Since the stable set polytope is a convex hull of a set of points, it can be described by some set of linear inequalities. If we can efficiently identify, for a given point x^* outside the polytope, a linear inequality certifying that x^* is outside the polytope, then by using the ellipsoid method one can solve the maximum weight stable set problem in polynomial time [GLS88]. This problem of identifying such a linear inequality for a polytope is called the *separation problem.* So, if the separation problem for the stable set polytope of a graph can be solved efficiently, then the stable set problem can be solved efficiently for the graph.

Here are some easy inequalities that are satisfied by points $x \in \mathbb{R}^{V(G)}$ in every stable set polytope of a graph G:

(a) *(Nonnegativity)* $x_v \geqslant 0$ for every vertex v of G.

(b) *(Edge inequality)* $x_u + x_v \leqslant 1$ for every edge uv of G.

(c) *(Clique inequality)* $\sum_{v \in K} x_v \leqslant 1$ for every clique K of G.

(d) *(Odd cycle inequality)* $\sum_{v \in V(C)} x_v \leqslant \frac{|V(C)|-1}{2}$ for every odd cycle C of G.

Let $\mathrm{QSTAB}(G)$ be the set of all vectors $x \in \mathbb{R}^{V(G)}$ satisfying (a) and (c). Remarkably, important theorems by Lovász [Lov72] and Fulkerson [Ful72] on perfect graphs, as stated by Chvátal [Chv75], imply that

$$\mathrm{SSP}(G) = \mathrm{QSTAB}(G) \text{ if and only if } G \text{ is perfect.}$$

Perfect graphs were introduced in the 1960s by Berge in terms of chromatic numbers and clique numbers. A graph H is an *induced* subgraph of a graph G if H can be obtained from G by deleting vertices and all incident edges adjacent to them. A graph G is called *perfect* if $\chi(H) = \omega(H)$ for every induced subgraph H of G. In what has since become known as the Strong Perfect Graph Theorem, Chudnovsky, Robertson, Seymour, and Thomas [CRST06] proved a theorem characterising the list of forbidden induced subgraphs for the class of perfect graphs.

Motivated by perfect graphs, Chvátal [Chv75] initiated the study of t-perfect graphs. Let $\mathrm{TSTAB}(G)$ be the set of all vectors $x \in \mathbb{R}^{V(G)}$ satisfying (a), (b), and (d). We define a graph to be *t-perfect* if $\mathrm{SSP}(G) = \mathrm{TSTAB}(G)$.

Let us write $\mathbf{1} \in \mathbb{R}^{V(G)}$ for the vector with all entries equal to 1. It is easy to verify that $\frac{1}{3}\mathbf{1} \in \mathrm{TSTAB}(G)$ for every graph G, from which it follows that K_4 is not t-perfect

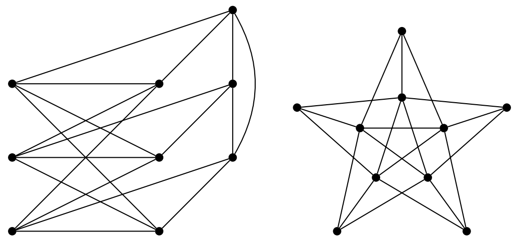


Figure 1. Left: the complement of the line graph of the complement of C_6 (Laurent and Seymour [Sch03, p. 1207]). Right: complement of the line graph of the 5-wheel (Benchetrit [Ben15, Ben16]).

and the fractional chromatic number of any t-perfect graph is at most three. In 1992, Shepherd [JT95, 8.14] asked whether for each t-perfect graph G the polytope $\mathrm{SSP}(G)$ has the *integer decomposition property*, that is, for every positive integer k, every integral vector in $k\,\mathrm{SSP}(G)$ can be written as a sum of k vertices of $\mathrm{SSP}(G)$. If G is t-perfect and $\mathrm{SSP}(G)$ has the integer decomposition property, then $\frac{1}{3}\mathbf{1} \in \mathrm{TSTAB}(G) = \mathrm{SSP}(G)$ should be expressible as a sum of three incidence vectors of stable sets, implying that G is 3-colourable. Indeed, several subclasses of t-perfect graphs are known to be 3-colourable, such as t-perfect graphs that are almost bipartite [FU82], claw-free [BS12], P_5-free [BF17], fork-free [CW22], series-parallel [BU79], and those forbidding certain subdivisions of K_4 [Cat79, GS98].

However, two counterexamples to Shepherd's conjecture, depicted in Figure 1, were found by Laurent and Seymour [Sch03, p. 1207] in 1994 and Benchetrit [Ben15, Ben16] in 2015 respectively, answering Shepherd's question in the negative. These are the only known 4-critical t-perfect graphs. On the other hand, Sebő conjectured that triangle-free t-perfect graphs are 3-colourable (see [BS12]), and this is wide open.

More generally, is every t-perfect graph 4-colourable? This very natural question appears in the problem book of Jensen and Toft [JT95, 8.14], and is attributed to Shepherd from 1994. Reiterating this, Shepherd wrote in the conclusion of his 1995 paper [She95]:

> *For every $k \geqslant 4$, it is not known whether each t-perfect graph is k-colourable.*

Our main result is the first positive answer to this question. We remark that we have optimised the proof for simplicity rather than to optimise the bound.

Theorem 1.1. *Every t-perfect graph is* 199053*-colourable.*

Let $\mathrm{HSTAB}(G)$ be the set of all vectors $x \in \mathbb{R}^{V(G)}$ satisfying (a), (c), and (d). A graph is *h-perfect* if $\mathrm{SSP}(G) = \mathrm{HSTAB}(G)$. By their definitions, we have the relationships $\mathrm{SSP}(G) \subseteq \mathrm{HSTAB}(G) \subseteq \mathrm{TSTAB}(G)$ and $\mathrm{HSTAB}(G) \subseteq \mathrm{QSTAB}(G)$, and t-perfect graphs are precisely h-perfect graphs without K_4 subgraphs. Obviously, every perfect graph is h-perfect and every t-perfect graph is h-perfect.

The study of h-perfect graphs was initiated by Sbihi and Uhry in 1984, who conjectured that every h-perfect graph G with $\omega(G) \geqslant 3$ is $\omega(G)$-colourable [SU84, Conjecture 5.4]. This conjecture is false because of the graphs in Figure 1. However, we prove that it is true up to an additive constant.

Theorem 1.2. *Every h-perfect graph G is $(\omega(G) + 199050)$-colourable.*

We remark that Theorem 1.2 is the first known result bounding the chromatic number of h-perfect graphs in terms of the clique number, establishing that the class of h-perfect graphs is a "χ-bounded" class. The proof of Theorem 1.2 actually reduces to the t-perfect case via a well-known fact, due to Sebő [BS12, Lemma 26], which implies that every h-perfect graph G admits $\omega(G) - 2$ stable sets such that after deleting all these stable sets, we have a triangle-free h-perfect induced subgraph H (which is t-perfect). Thus, we limit the proof outline below to the main theorem for t-perfect graphs.

2 Proof outline

Our first step is to reduce to the case where there are no "short" odd cycles. For this, we use a nice lemma that originated from Sbihi and Uhry [SU84] and was explicitly stated in the PhD thesis of Marcus [Mar96] which says that the fractional chromatic number of a t-perfect graph is tight on its shortest odd cycle. This allows us to show that there is a stable set hitting all shortest odd cycles of a t-perfect graph. Removing such a stable set increases the length of the shortest odd cycle by at least two and decreases the chromatic number by at most one. By doing this repeatedly (four times), we reduce the proof of Theorem 1.1 to proving that every t-perfect graph G with odd girth at least 11 is 199049-colourable.

To proceed, we replace the t-perfect graphs by a superclass that is more amenable to tools from structural graph theory. An *odd wheel* is a graph consisting of an odd cycle and one additional vertex adjacent to every vertex on the cycle. It is an easy consequence of the definition of t-perfection that odd wheels are not t-perfect. Graphs with high chromatic number need not contain any wheels as an induced subgraph [Dav23, PT23], but fortunately for us it is enough to show that G contains an odd wheel in a weaker sense. A *t-contraction* is an operation to contract all edges in $G[N_G[v]]$ for some vertex v whose set of neighbours is stable. A graph H is a *t-minor* of a graph G if H can be obtained from G by a sequence of vertex deletions and t-contractions. Gerards and Shepherd [GS98] showed that every t-minor of a t-perfect graph is t-perfect. Thus, no t-perfect graph contains an odd wheel t-minor, and it therefore suffices to prove that if G has odd girth at least 11 and chromatic number more than 199049, then G contains an odd wheel t-minor.

It is not entirely straightforward to find an odd wheel t-minor, so to do this we introduce a helpful intermediate structure which we find to be of independent interest. An r-arithmetic rope in G is a subgraph consisting of $2r$ paths $Q_{1,1}, Q_{1,2}, \dots, Q_{r,1}, Q_{r,2}$ with ends contained in a vertex set $\{q_1, \dots, q_r\}$ so that (taking indices modulo r)

- for every $1 \leqslant i \leqslant r$, we have that $Q_{i,1}$ and $Q_{i,2}$ are q_1q_2-paths with odd and even length, respectively.
- for every $(h_1, \dots, h_r) \in \{1, 2\}^r$, the graph $Q_{1,h_1} \cup Q_{2,h_2} \cup \cdots \cup Q_{r,h_r}$ is an induced cycle in G, and
- the vertices $q_1, \dots, q_r$ are pairwise at G-distance at least 5.

We refer to Figure 2 for an example with $r = 5$.

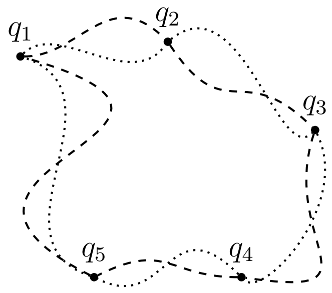


Figure 2. A 5-arithmetic rope. Dotted lines represent odd-length paths and dashed lines represent even-length paths.

The motivation for this structure is that the odd cycles of a graph are critical in determining whether it is t-perfect. By definition, the parity of cycles is preserved under t-contractions, so we need careful control over parity in order to hunt for an odd wheel t-minor in G. This is achieved here by choosing suitable $(h_1, \ldots, h_r) \in \{1, 2\}^r$. The key theorem we prove is that every graph G with odd girth at least 11 and $X \subseteq V(G)$ with $\chi(G[X]) \geqslant 99525$ contains a 5-arithmetic rope in X. Our proof involves iterating an argument similar to that used by Chudnovsky, Scott, Seymour, and Spirkl [CSSS20] to show that the class of graphs without long odd induced cycles is χ-bounded.

Putting together the main argument, we assume that $\chi(G) > 199049$ and that G has odd girth at least 11. Let v be a vertex of G in a connected component with maximum chromatic number, and partition $V(G)$ into levels $L_0 = \{v\}$, L_1, L_2, $\ldots$ where $L_i = N^i(v)$. A standard and simple argument shows that there exists a positive integer i such that $\chi(L_i) \geqslant \lceil 199049/2 \rceil = 99525$. Thus, $G[L_i]$ contains a 5-arithmetic rope $\mathcal{Q}$.

What remains is to extend some odd cycle of our 5-arithmetic rope into an odd wheel t-minor. For each $j \in \{1, 2, \ldots, 5\}$, we choose one vertex x_j in L_{i-1} adjacent to $q_j \in \mathcal{Q}$, and then a vertex y_j in L_{i-2} adjacent to x_j (all 10 vertices are necessarily distinct since the q_j are far apart). A simple lemma will show that $G[L_0 \cup L_1 \cup \cdots \cup L_{i-3} \cup \{x_1, x_2, x_3, x_4, x_5, y_1, y_2, y_3, y_4, y_5\}]$ has a connected bipartite induced subgraph H containing all of the x_j. Three out of these five, say x_a, x_b, and x_c, will be on the same side of H by the pigeonhole principle; delete the other two to obtain a connected bipartite induced subgraph H'. Using the arithmetic rope structure, we can also take suitable paths to obtain an odd induced cycle C in $G[L_i]$ such that q_a, q_b, and q_c split C into three odd-length induced paths.

Consider the subgraph G' of G induced by $V(H') \cup V(C)$. By applying t-contractions to every vertex of H' on the side not containing $\{x_a, x_b, x_c\}$, we obtain a t-minor G'' where all vertices of H' are identified into a single vertex, say w. That is, G'' is a graph consisting of an odd cycle C with an extra vertex w such that w is adjacent to q_a, q_b, and q_c (and possibly more). Let G''' be the graph obtained from G'' by repeatedly applying t-contractions to degree-2 vertices. Since q_a, q_b, and q_c split C into odd-length paths and each t-contraction preserves the parity of these paths in C, we deduce that G''' is an odd wheel with at least three vertices on the rim. This implies that G'' (and hence G) has an odd wheel t-minor, completing the proof.

References

[Ben15] YOHANN BENCHETRIT (May 2015). Propriétés géométriques du nombre chromatique : polyèdres, structures et algorithmes. Theses, Université Grenoble Alpes. URL: `https://theses.hal.science/tel-01292635`.

[Ben16] YOHANN BENCHETRIT (2016). On 4-critical t-perfect graphs. arXiv:1604.02288.

[BF17] HENNING BRUHN and ELKE FUCHS (2017). t-perfection in P_5-free graphs. *SIAM J. Discrete Math.* **31**(3), 1616–1633.

[BS12] HENNING BRUHN and MAYA STEIN (2012). On claw-free t-perfect graphs. *Math. Program.* **133**(1-2), 461–480.

[BU79] MOULOUD BOULALA and JEAN-PIERRE UHRY (1979). Polytope des indépendants d'un graphe série-parallèle. *Discrete Math.* **27**(3), 225–243.

[Cat79] PAUL A. CATLIN (1979). Hajós' graph-coloring conjecture: variations and counterexamples. *J. Combin. Theory Ser. B* **26**(2), 268–274.

[Chv75] VAŠEK CHVÁTAL (1975). On certain polytopes associated with graphs. *J. Combinatorial Theory Ser. B* **18**, 138–154.

[CRST06] MARIA CHUDNOVSKY, NEIL ROBERTSON, PAUL SEYMOUR, and ROBIN THOMAS (2006). The strong perfect graph theorem. *Ann. of Math. (2)* **164**(1), 51–229.

[CSSS20] MARIA CHUDNOVSKY, ALEX SCOTT, PAUL SEYMOUR, and SOPHIE SPIRKL (2020). Induced subgraphs of graphs with large chromatic number. VIII. Long odd holes. *J. Combin. Theory Ser. B* **140**, 84–97.

[CW22] YIXIN CAO and SHENGHUA WANG (2022). On fork-free t-perfect graphs. arXiv:2211.03538.

[Dav23] JAMES DAVIES (2023). Triangle-free graphs with large chromatic number and no induced wheel. *J. Graph Theory* **103**(1), 112–118.

[FU82] JEAN FONLUPT and JEAN-PIERRE UHRY (1982). Transformations which preserve perfectness and h-perfectness of graphs. *Bonn Workshop on Combinatorial Optimization (Bonn, 1980)*, *North-Holland Math. Stud.*, vol. 66, 83–95.

[Ful72] DELBERT R. FULKERSON (1972). Anti-blocking polyhedra. *J. Combinatorial Theory Ser. B* **12**, 50–71.

[GLS88] MARTIN GRÖTSCHEL, LÁSZLÓ LOVÁSZ, and ALEXANDER SCHRIJVER (1988). Geometric algorithms and combinatorial optimization, *Algorithms and Combinatorics: Study and Research Texts*, vol. 2 (Springer-Verlag, Berlin).

[GS98] A. M. H. GERARDS and F. BRUCE SHEPHERD (1998). The graphs with all subgraphs t-perfect. *SIAM J. Discrete Math.* **11**(4), 524–545.

[JT95] Tommy R. Jensen and Bjarne Toft (1995). Graph coloring problems (John Wiley & Sons, Ltd).

[Lov72] László Lovász (1972). Normal hypergraphs and the perfect graph conjecture. *Discrete Math.* **2**(3), 253–267.

[Lov94] László Lovász (1994). Stable sets and polynomials. *Discrete Math.* **124**(1-3), 137–153. Graphs and combinatorics (Qawra, 1990).

[Mar96] Karina Marcus (1996). Multiflots, métriques et graphes h-parfaits : les cycles impairs dans l'optimisation combinatoire. Ph.D. thesis, Université Joseph-Fourier - Grenoble I. URL: `https://theses.hal.science/tel-00005002/`.

[PT23] Pegah Pournajafi and Nicolas Trotignon (2023). Burling graphs revisited, part I: New characterizations. *European J. Combin.* **110**, Paper No. 103686, 24.

[Sch03] Alexander Schrijver (2003). Combinatorial Optimization: Polyhedra and Efficiency (Springer).

[She95] F. Bruce Shepherd (1995). Applying Lehman's theorems to packing problems. *Math. Programming* **71**(3), 353–367.

[SU84] Najiba Sbihi and Jean-Pierre Uhry (1984). A class of h-perfect graphs. *Discrete Math.* **51**(2), 191–205.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

New bounds for proper h-conflict-free colourings

(Extended abstract)

Quentin Chuet* Tianjiao Dai† Qiancheng Ouyang‡ François Pirot§

Abstract

A proper k-colouring of a graph G is called h-conflict-free if every vertex v has at least $\min\{h, \deg(v)\}$ colours appearing exactly once in its neighbourhood. Let $\chi^h_{\mathrm{pcf}}(G)$ denote the minimum k such that such a colouring exists. We show that for every fixed $h \geq 1$, every graph G of maximum degree Δ satisfies $\chi^h_{\mathrm{pcf}}(G) \leq h\Delta + \mathcal{O}(\log \Delta)$. This expands on the work of Cho *et al.*, and improves a recent result of Liu and Reed in the case $h = 1$, which provides further evidence to support a conjecture of Caro, Petruševski and Škrekovski. We conjecture that for every $h \geq 1$ and every graph G of maximum degree Δ sufficiently large, the bound $\chi^h_{\mathrm{pcf}}(G) \leq h\Delta + 1$ should hold, which would be tight. When the minimum degree δ of G is sufficiently large, namely $\delta \geq \max\{100h, 3000 \log \Delta\}$, we show that this upper bound can be further reduced to $\chi^h_{\mathrm{pcf}}(G) \leq \Delta + O(\sqrt{h\Delta})$. This improves a recent bound derived by Kamyczura and Przybyło in the regime $\delta \leq \sqrt{h\Delta}$.

1 Introduction

For a positive integer k, let us denote $[k] := \{1, \ldots, k\}$ the set of the first k integers. A k-*colouring* of a graph G is an assignment $\sigma\colon V(G) \to [k]$. A colouring σ of G is *proper* if $\sigma(u) \neq \sigma(v)$ for every edge $uv \in E(G)$. Given a colouring σ and a vertex $v \in V(G)$, a *witness* of v is a neighbour $u \in N(v)$ such that the colour $\sigma(u)$ appears exactly once in $N(v)$, and we call this colour a *solitary colour*; the colouring σ is *proper conflict-free* (or simply *pcf*) if it is proper and every non-isolated vertex v has a witness. We let $\chi_{\mathrm{pcf}}(G)$ be the smallest integer k such that a pcf k-colouring of G exists. The notion of pcf colouring was introduced by Fabrici *et al.* [7], and was first investigated in the class of graphs of maximum degree Δ by Caro, Petruševski and Škrekovski [2]. Among other results, they obtained the upper bound $\chi_{\mathrm{pcf}}(G) \leq 5\Delta(G)/2$, and conjectured that the optimal upper bound should actually be $\Delta(G) + 1$.

*LISN, Université Paris-Saclay, Gif-sur-Yvette, France. E-mail: quentin.chuet@lisn.fr
†Department of Mathematics, Shandong University, Jinan, China. E-mail: tianjiao.dai@sdu.edu.cn
‡Data Science Institute, Shandong University, Jinan, China. E-mail: oyqc@sdu.edu.cn
§LISN, Université Paris-Saclay, Gif-sur-Yvette, France. E-mail: francois.pirot@lisn.fr

Conjecture 1 (Caro, Petruševski, Škrekovski, 2023 [2, Conjecture 6.4]). *If G is a connected graph of maximum degree $\Delta \geq 3$, then*

$$\chi_{\mathrm{pcf}}(G) \leq \Delta + 1.$$

The condition that G is connected ensures that there is no isolated vertex. The condition $\Delta \geq 3$ is needed to remove cycles from consideration; indeed $\chi_{\mathrm{pcf}}(C_5) = 5 = \Delta(C_5) + 3$, and $\chi_{\mathrm{pcf}}(C_n) = 4 = \Delta(C_n) + 2$ whenever $n > 5$ is not a multiple of 3. The conjectured upper bound is best possible. For instance, if G is a graph obtained by subdividing any subset of edges from the complete graph K_n with $n \geq 5$, then $\chi_{\mathrm{pcf}}(G) = n = \Delta(G) + 1$. In particular, when $G = {K_n}^{1/2}$, the difference between $\chi(G)$ and $\chi_{\mathrm{pcf}}(G)$ is unbounded; we have $\chi(G) = 2$ while $\chi_{\mathrm{pcf}}(G) = n$.

As a first step towards resolving Theorem 1, Cranston and Liu [5] observed that a simple greedy algorithm provides a pcf ($2\Delta+1$)-colouring of any graph G of maximum degree Δ. They moreoever proved that $\chi_{\mathrm{pcf}}(G) < 1.656\Delta$ when Δ is large enough. More recently, Liu and Reed [10] proved that Theorem 1 holds asymptotically, by showing $\chi_{\mathrm{pcf}}(G) \leq \Delta + \mathcal{O}(\Delta^{2/3} \log \Delta)$ as $\Delta \to \infty$. Our first contribution is to improve this upper bound by eliminating the polynomial factor in the second-order term.

Theorem 2. *Let G a graph of maximum degree Δ, then*

$$\chi_{\mathrm{pcf}}(G) \leq \Delta + \mathcal{O}(\log \Delta).$$

Proper conflict-free colourings can be seen as a relaxation of distance-2 colourings: the latter require all neighbours of a vertex to have different colours, whereas the former only ask for one neighbour to have a colour distinct from the others. To interpolate between those two notions, one can introduce the notion of *proper h-conflict-free k-colourings* (or *h-pcf k-colouring* for short), for some integers $h, k \geq 1$. Those are proper k-colourings where every vertex v has at least $\min\{h, \deg(v)\}$ solitary colours. In particular, for a graph G of maximum degree Δ, a 1-pcf k-colouring of G is simply a pcf k-colouring of G, while an h-pcf k-colouring is a distance-2 k-colouring of G whenever $h \geq \Delta - 1$. For every $h \geq 1$, we denote $\chi^h_{\mathrm{pcf}}(G)$ the minimum k such that an h-pcf k-colouring of G exists.

A simple greedy algorithm, akin to the one proposed by Cranston and Liu [5], yields the upper bound $\chi^h_{\mathrm{pcf}}(G) \leq (h+1)\Delta + 1$ for every graph G of maximum degree Δ and every $h \geq 1$. Cho *et al.* [4] improved this upper bound by reducing the number of colours needed by 2.

Theorem 3 (Cho, Choi, Kwon, Park, 2025 [4, Theorem 1.3]). *Let G be a graph of maximum degree Δ, and let $1 \leq h \leq \Delta - 2$ be an integer. Then*

$$\chi^h_{\mathrm{pcf}}(G) \leq (h+1)\Delta - 1.$$

For the specific case $h = \Delta - 2$, they exhibit a graph G of maximum degree Δ arbitrarily large such that $\chi^h_{\mathrm{pcf}}(G) \geq (h+1)(\Delta - 1) = h\Delta + 1$. We consider the opposite end of the spectrum by considering h a fixed constant, and make an asymptotic improvement on Theorem 3 in that case.

Theorem 4. *Let $h \geq 1$ be a fixed integer. If G is a connected graph of maximum degree Δ, then*

$$\chi^h_{\mathrm{pcf}}(G) \leq h\Delta + \mathcal{O}(\log \Delta).$$

The notion of pcf colourings can also be seen as a strengthening of *odd colourings*, cf. [11, 3, 6]. For a graph G and an integer $h \geq 1$, an *h-odd colouring* of G is a proper colouring such that every vertex v observes at least $\min\{h, \deg(v)\}$ colours in its neighbourhood that appear an odd number of times. Let $\chi_o^h(G)$ be the minimum number of colours required for an h-odd colouring of G, and $\chi_o(G) \coloneqq \chi_o^1(G)$. Clearly, $\chi_o^h(G) \leq \chi_{\mathrm{pcf}}^h(G)$ for every $h \geq 1$. A weaker version of Theorem 1 was first formulated in [1] for the odd chromatic number of graphs of maximum degree Δ, and it was shown in [6, Theorem 3] that $\chi_o(G) \leq \Delta + \mathcal{O}(\log \Delta)$. In that sense, Theorem 2 is also a strengthening of this result. If h is a fixed integer and G is a graph of maximum degree Δ, then we have $\chi_o^h(G) \leq h\Delta + \mathcal{O}(\log \Delta)$ as a direct consequence of Theorem 4. The first-order term of this upper bound was shown to be tight in [6, Proposition 28].

Proposition 5 (Dai, Ouyang, Pirot, 2024)**.** *Let $h \geq 1$ be a fixed integer. For infinitely many values of Δ, there exists a graph G of maximum degree Δ and minimum degree $h+1$ such that*

$$\chi_o^h(G) \geq h\Delta + 1.$$

In turn, Theorem 5 demonstrates that Theorem 4 is tight up to the second-order term, i.e. for every fixed integer $h \geq 1$, there exists a graph of maximum degree Δ arbitrarily large such that $\chi_{\mathrm{pcf}}^h(G) \geq h\Delta + 1$. In light of these observations, we propose an extended version of Theorem 1.

Conjecture 6. *Let $h \geq 1$ be an integer, and G be a graph of maximum degree Δ sufficiently large. Then*

$$\chi_{\mathrm{pcf}}^h(G) \leq h\Delta + 1.$$

The lower bound provided by Theorem 5 relies on a graph construction with many vertices of small degree (namely, of degree $h+1$). When there is an additional restriction on the minimum degree of G, it is possible to derive smaller upper bounds on $\chi_{\mathrm{pcf}}^h(G)$. In particular, Liu and Reed [9] have recently proved that Theorem 1 holds for regular graphs in a strong sense. Their result can be stated as follows.

Theorem 7 (Liu, Reed, 2024+ [9, Theorem 1.5])**.** *For every graph G of maximum degree Δ sufficiently large, and minimum degree $\delta > \frac{8000}{8001}\Delta$, letting $h \coloneqq \delta - \frac{8000}{8001}\Delta$, one has*

$$\chi_{\mathrm{pcf}}^h(G) \leq \Delta + 1.$$

The condition on the minimum degree in Theorem 7 is of course very restrictive. One could wonder what happens when this condition is relaxed. This has been considered by Kamyczura and Przybyło [8], who showed the following.

Theorem 8 (Kamyczura, Przybyło, 2024 [8, Theorem 5])**.** *Let G be a graph of maximum degree Δ and minimum degree δ. For every integer h satisfying $20\log\Delta \leq h \leq \delta/75$, one has*

$$\chi_{\mathrm{pcf}}^h(G) \leq \Delta + \mathcal{O}\left(\frac{h\Delta}{\delta}\right).$$

The condition on h implies that $\delta > 1500\log\Delta$. Note that the condition $h \geq 20\log\Delta$ in the above statement is non-restrictive, since $\chi_{\mathrm{pcf}}^h(G)$ is monotonous in h: when $h < 20\log\Delta$, one can use $\chi_{\mathrm{pcf}}^h(G) \leq \chi_{\mathrm{pcf}}^{20\log\Delta}(G)$ to obtain an upper bound on $\chi_{\mathrm{pcf}}^h(G)$ with an application of Theorem 8. If G is almost regular, i.e. $\delta \geq \varepsilon\Delta$ for some fixed $\epsilon > 0$, and if h satisfies $20\log\Delta \leq h \leq \delta/75$, then Theorem 8 implies the upper bound $\chi^h_{\mathrm{pcf}}(G) \leq \Delta + \mathcal{O}(h)$. Dai, Ouyang and Pirot [6] showed that the upper bound $\Delta+\mathcal{O}(h)$ also holds for the h-odd chromatic number without the assumption that G is almost regular.

Theorem 9 (Dai, Ouyang, Pirot, 2024 [6, Theorem 9])**.** *There exists a universal constant C such that, for every graph G of maximum degree Δ and minimum degree δ, if $C\log\Delta \leq h \leq \delta/2$, then*

$$\chi^h_o(G) \leq \Delta + \mathcal{O}(h).$$

We conjecture that a similar statement should hold for the h-pcf chromatic number.

Conjecture 10. *There exists a universal constant C such that, for every graph G of maximum degree Δ and minimum degree δ, if $C\log\Delta \leq h \leq \delta/C$, then*

$$\chi^h_{\mathrm{pcf}}(G) < \Delta + \mathcal{O}(h).$$

We make a step towards resolving this conjecture by proving that $\chi_{\mathrm{pcf}}(G) \leq \Delta+\mathcal{O}\left(\sqrt{h\Delta}\right)$ under those hypotheses, improving over Theorem 8 when $\delta \leq \sqrt{h\Delta}$.

Theorem 11. *Let G be a graph of maximum degree Δ and minimum degree δ. For every integer h satisfying $30\log\Delta \leq h \leq \delta/100$, one has*

$$\chi^h_{\mathrm{pcf}}(G) \leq \Delta + \mathcal{O}\left(\sqrt{h\Delta}\right).$$

2 Methodology

We recall that, in a proper colour σ of a graph G, a solitary colour of a vertex $v \in V(G)$ is a colour of σ having a unique occurrence in $N(v)$, and each $u \in N(v)$ coloured with a solitary colour in $N(v)$ is called a witness of v. The proofs of Theorem 4 and Theorem 11 both use a two-step colouring process. Roughly speaking, the first step constructs a proper k-colouring σ_0 of G which has the additional property that vertices of "small degree" have many solitary colours, while the second step relies on a Rödl Nibble approach to slightly alter σ_0 using m new colours in order to grant solitary colours to vertices of "large degree"; combining these two steps yields a h-pcf $(k+m)$-colouring.

2.1 Overview of the proof of Theorem 4

Let $G = (V, E)$ be a graph of maximum degree Δ sufficiently large, and $h \geq 1$ an integer. Let V^- be the subset of vertices with degree at most $\frac{h}{h+1}\Delta$, and let $V^+ := V \setminus V^-$. Given a subset of vertices $U \subseteq V$, a colouring σ is said to be (U, h)-conflict-free if every vertex $v \in U$ has at least $\min\{h, \deg(v)\}$ solitary colours in σ. In particular, a proper (V, h)-conflict-free colouring of G is a h-pcf colouring. We first exhibit a proper $(h\Delta + 1)$-colouring of G which is almost h-conflict-free.

Claim 12. *There exists a proper $(h\Delta+1)$-colouring σ_0 of G such that σ_0 is (V^-, h)-conflict-free and $(V^+, h-1)$-conflict-free.*

This colouring σ_0 is obtained via the following greedy algorithm: fix an ordering $(v_1, \dots, v_n)$ on the vertices by decreasing degree. For each $v \in V^-$ (resp. V^+), let $S(v)$ be the first $\min\{h, \deg(v)\}$ (resp. $h-1$) neighbours of v in the ordering. Colour sequentially each vertex v_i, for i from 1 to n, with the first colour that does not appear in $N(v_i)$, nor in $S(u)$ for all $u \in N(v_i)$. It is straightforward to check that the colouring σ_0 obtained at the end of this procedure satisfies the statement of Theorem 12 and uses at most $h\Delta + 1$ colours.

Let us fix a colouring σ_0 as provided by Theorem 12. Our goal is now to recolour a fraction of the vertices in order to grant one additional solitary colour to vertices of V^+, but we must be careful of existing solitary colours in doing so. For $v \in V^-$, let $\mathcal{W}(v)$ be an arbitrary subset of $\min\{h, \deg(v)\}$ witnesses of v in σ_0. For $v \in V^+$, let $\mathcal{W}(v)$ be an arbitrary subset of $h-1$ witnesses of v in σ_0. Finally, let $\mathcal{W}(S) \coloneqq \bigcup_{u \in S} \mathcal{W}(u)$. For $u \in V$, let $\mathcal{N}(u) \coloneqq N(u) \cup \mathcal{W}(N(u)) \setminus \{u\}$.

We set $m \coloneqq 4000h \log \Delta$ and $p \coloneqq 1/(h+2)\Delta$. We execute the following procedure, which first initialises subsets $A_1 \dots, A_m$ by randomly sampling V with independent probability p, and then proceeds in m rounds to construct subsets $C_1, \dots, C_m$ which are used to modify the colouring σ_0.

Algorithm: modified Rödl Nibble

```
Initialisation: for i ∈ [m] do
    A_i ← ∅
    for u ∈ V do
        Add u to A_i independently with probability p.
i_0 ← 1
for i from 1 to m do
    B_i ← {u ∈ V : N(u) ∩ A_i ≠ ∅} ∪ ⋃_{j=i_0}^{i-1} C_j;
    C_i ← A_i \ B_i.
    fail-safe: if there exists v ∈ V^+ such that |N(v) ∩ ⋃_{j=i_0}^{i} C_j| > deg(v)/2 then
        i_0 ← i + 1.
for i from i_0 to m do
    Recolour C_i with a new colour.
```

Let $\boldsymbol{\sigma}$ be the (random) colouring obtained at the end of this procedure. This algorithm (and its analysis) is in part inspired by the work of Kamyczura and Przybyło [8], but incorporates new ideas in order to preserve the properties of the colouring σ_0. The sets $\{C_i\}$ are constructed in such a way that a vertex v cannot be included in C_i at the same time as another vertex in $N(v)$ or $\mathcal{W}(N(v))$, which guarantees the following property.

Claim 13. *The colouring $\boldsymbol{\sigma}$ is surely proper, and for every $u \in V$ the vertices of $\mathcal{W}(u)$ are surely witnesses of u.*

Using Chernoff's bounds and the the Lovász Local Lemma, we show that the colouring $\boldsymbol{\sigma}$ is (V^+, h)-conflict-free with non-zero probability, which implies the existence of a h-pcf $(h\Delta + 1 + m)$-colouring of G, and concludes the proof of Theorem 4. Actually, we end up with a bit more than just one additional solitary colour for vertices of V^+.

Claim 14. *With non-zero probability, $\boldsymbol{\sigma}$ is $(V^+, h + 28 \log \Delta)$-conflict-free.*

Note that the fail-safe in the algorithm serves no practical use, but circumvents a technicality that would otherwise hinder the analysis of the algorithm in the proof of Theorem 14: at each round $i \in [m]$ of the algorithm, we want to ensure deterministically that every vertex $v \in V^+$ has sufficiently many neighbours that haven't been selected in $C_1 \cup \cdots \cup C_{i-1}$, so that v has probability $\Omega(1/h)$ of seeing exactly one neighbour in C_i. If the aforementioned condition is not met, we erase all the progress made up to that point (we expect this bad event to happen with low probability).

2.2 Overview of the proof of Theorem 11

Let $G = (V, E)$ be a graph of maximum degree Δ sufficiently large, minimum degree $\delta \geq 3000 \log \Delta$, and h an integer such that $20 \log \Delta \leq h \leq \delta/100$. Set $d := \sqrt{h\Delta}$, let V^- be the subset of vertices with degree at most d, and let $V^+ := V \setminus V^-$.

This time, we start with a proper colouring σ_0 which has twice as many solitary colours for V^- than we need: this enables us the possibility of losing a few solitary colours for V^- during the second step with no consequences.

Claim 15. *There exists a proper $(\Delta + 3d)$-colouring σ_0 such that σ_0 is $(V^-, 2h)$-conflict-free.*

In order to prove Theorem 15, we consider the uniform distribution over all proper $(\Delta+3d)$-colourings of G, and apply a resampling argument together with Chernoff's bounds and the Lopsided version of the Lovász Local Lemma.

Let us fix a colouring σ_0 as provided by Theorem 15. For $v \in V^-$, let $\mathcal{W}(v)$ be an arbitrary subset of $2h$ witnesses of v in σ_0. We set $m := 22d$ and $p := 1/2\Delta$, and we execute the same algorithm as before, except for one small but crucial detail: we replace the term "$\mathcal{N}(u)$" in the definition of B_i by the term "$N(u)$". Let $\boldsymbol{\sigma}$ be the colouring obtained at the end of this modified procedure (which uses $\Delta + 25d$ colours in total). $\boldsymbol{\sigma}$ is still surely a proper colouring, but we no longer ensure that vertices of $\mathcal{W}(u)$ remain witnesses of u in $\boldsymbol{\sigma}$, for $u \in V^-$ — However, we do show that with high probability, at least half of the vertices in $\mathcal{W}(u)$ remain witnesses of u in $\boldsymbol{\sigma}$. We also show that each vertex $u \in V^+$ has at least h witnesses in $\boldsymbol{\sigma}$ with high probability. An application of the Lovász Local Lemma concludes the proof of Theorem 11.

Claim 16. *With non-zero probability, the colouring $\boldsymbol{\sigma}$ is h-pcf.*

References

[1] Y. Caro, M. Petruševski, and R. Škrekovski. Remarks on odd colorings of graphs. *Discrete Applied Mathematics*, 321:392–401, 2022.

[2] Y. Caro, M. Petruševski, and R. Škrekovski. Remarks on proper conflict-free colorings of graphs. *Discrete mathematics*, 346(2):113221, 2023.

[3] E.-K. Cho, I. Choi, H. Kwon, and B. Park. Odd coloring of sparse graphs and planar graphs. *Discrete Mathematics*, 346(5):113305, 2023.

[4] E.-K. Cho, I. Choi, H. Kwon, and B. Park. Brooks-type theorems for relaxations of square colorings. *Discrete Mathematics*, 348(1):114233, 2025.

[5] D. W. Cranston and C.-H. Liu. Proper conflict-free coloring of graphs with large maximum degree. *SIAM Journal on Discrete Mathematics*, 38(4):3004–3027, 2024.

[6] T. Dai, Q. Ouyang, and F. Pirot. New bounds for odd colourings of graphs. *Electronic Journal of Combinatorics*, 31(Issue 4), 2024.

[7] I. Fabrici, B. Lužar, S. Rindošová, and R. Soták. Proper conflict-free and unique-maximum colorings of planar graphs with respect to neighborhoods. *Discrete Applied Mathematics*, 324:80–92, 2023.

[8] M. Kamyczura and J. Przybyło. On conflict-free proper colourings of graphs without small degree vertices. *Discrete Mathematics*, 347(1):113712, 2024.

[9] C.-H. Liu and B. Reed. Peaceful colourings. *arXiv preprint arXiv:2402.09762*, 2024.

[10] C.-H. Liu and B. Reed. Asymptotically optimal proper conflict-free coloring. *Random Structures & Algorithms*, 66(3):e21285, 2025.

[11] M. Petruševski and R. Škrekovski. Colorings with neighborhood parity condition. *Discrete Applied Mathematics*, 321:385–391, 2022.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Fractional domatic number and minimum degree

(Extended abstract)

Quentin Chuet* Hugo Demaret† Hoang La‡ François Pirot§

Abstract

The domatic number of a graph G is the maximum number of pairwise disjoint dominating sets of G. We are interested in the LP-relaxation of this parameter, which is called the fractional domatic number of G. We study its extremal value in the class of graphs of minimum degree d. The fractional domatic number of a graph of minimum degree d is always at most $d+1$, and at least $(1-o(1))\,d/\ln d$ as $d \to \infty$. This is asymptotically tight even within the class of split graphs. Our main result concerns the case $d = 2$; we show that, excluding 8 exceptional graphs, the fractional domatic number of every connected graph of minimum degree (at least) 2 is at least 5/2. We also show that this bound cannot be improved if only finitely many graphs are excluded, even when restricting to bipartite graphs of girth at least 6. This proves in a stronger sense a conjecture by Gadouleau, Harms, Mertzios, and Zamaraev (2024). This also extends and generalises results from McCuaig and Shepherd (1989), from Fujita, Kameda, and Yamashita (2000), and from Abbas, Egerstedt, Liu, Thomas, and Whalen (2016). Finally, we show that planar graphs of minimum degree at least 2 and girth at least g have fractional domatic number at least $3 - O(1/g)$ as $g \to \infty$.

1 Introduction

Given a graph G, a *dominating set* of G is a set $X \subseteq V(G)$ such that $N[X] = V(G)$. Dominating sets are often used to model monitoring problems in networks. A classical problem in Graph Theory is to find the minimum size $\gamma(G)$ of a dominating set of G, called the *domination number* of G [15, 14]. A possible approach to this problem is to study a stronger parameter, the *domatic number* of G, denoted $\mathrm{DOM}(G)$, which is the maximum number of pairwise disjoint dominating sets of G. Since dominating sets are stable through vertex addition, $\mathrm{DOM}(G)$ can equivalently be defined as the maximum size of a partition of $V(G)$ into dominating sets. The latter is captured by the notion of *dominating k-colouring* of G, that

*

†Department of Computer Science, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France. E-mail: hugo.demaret@polytechnique.edu.

‡

§

is a mapping $\phi\colon V(G) \to [k]$ such that $\phi(N[v]) = [k]$ for every vertex $v \in V(G)$ — in other words, every closed neighbourhood spans all the colours in ϕ. Then $\mathrm{DOM}(G)$ is the maximum k such that G has a dominating k-colouring. A straightforward application of the Pigeonhole Principle yields that the minimum colour class in a dominating k-colouring of G has size at most $|V(G)|/k$, hence $\gamma(G) \leq |V(G)|/\,\mathrm{DOM}(G)$. As a consequence, any lower bound on $\mathrm{DOM}(G)$ yields an upper bound on $\gamma(G)$. For instance, Matheson and Tarjan [14] proved that every n-vertex triangulated disc (i.e. a 2-connected planar graph, all internal faces of which are triangles) has domination number at most $n/3$, by showing that its domatic number is at least 3. A parallel can be made with independent sets and the chromatic number of graphs. Remarkably, there is no known proof for the fact that every n-vertex planar graph contains an independent set of size at least $n/4$ that does not rely on the 4-Colour Theorem.

We are interested in the LP-relaxation of the domatic number, called the *fractional domatic number*, which provides a tighter bound on the domination number. This parameter was first formally introduced in [20], although some related notions already appeared in [8]. It has many equivalent definitions; we mention some of them hereafter.

Linear Programming We denote $\mathcal{D}(G)$ the collection of dominating sets of a given graph G. Then the fractional domatic number of G, denoted $\mathrm{FDOM}(G)$, is the solution to the following linear program:

$$\text{Maximise} \sum_{D \in \mathcal{D}(G)} x_D,$$
$$\text{Subject to} \begin{cases} \displaystyle\sum_{\substack{D \in \mathcal{D}(G) \\ v \in D}} x_D \leq 1 & \text{for all } v \in V(G); \\ x_D \geq 0 & \text{for all } D \in \mathcal{D}(G). \end{cases} \tag{1}$$

Probability distribution If $(x_D)_{D \in \mathcal{D}(G)}$ yields a solution to (1), then observe that a renormalisation of the weights x_D yields a probability distribution over $\mathcal{D}(G)$, in such a way that a random dominating set $\mathbf{D}$ drawn according to that distribution satisfies $\mathbb{P}[v \in \mathbf{D}] \leq \frac{1}{\mathrm{FDOM}(G)}$ for every vertex $v \in V(G)$. So the fractional domatic number of G can alternatively be defined as

$$\mathrm{FDOM}(G) = \max\left\{ \frac{1}{p} : \begin{array}{c} \text{there is a random dominating set } \mathbf{D} \text{ of } G \text{ such that} \\ \mathbb{P}[v \in \mathbf{D}] \leq p \text{ for every vertex } v \in V(G). \end{array} \right\}. \tag{2}$$

Dominating sets with bounded overlaps The theory of Linear Programming tells us that the solution to (1) is a rational number. So there is an integer q such that $q \cdot x_D$ is an integer for every $D \in \mathcal{D}(G)$. We construct a multiset $\mathcal{F}$ of dominating sets of G by adding $q \cdot x_D$ copies of D to $\mathcal{F}$ for every $D \in \mathcal{D}(G)$. By doing so, every vertex $v \in V(G)$ appears in at most q dominating sets of $\mathcal{F}$; we call that number of occurrences the *multiplicity of v in $\mathcal{F}$*. Letting $m(\mathcal{F})$ be the maximum multiplicity of a vertex in $\mathcal{F}$, we obtain that

$$\mathrm{FDOM}(G) = \max \frac{\mathcal{F}}{m(\mathcal{F})}, \tag{3}$$

where the maximum is taken over every multiset $\mathcal{F}$ of dominating sets of G.

Dominating $(p:q)$-colourings If a given vertex $v \in V(G)$ appears in fewer than $m(\mathcal{F})$ dominating sets of $\mathcal{F}$, we may add it to some extra ones so that its multiplicity in $\mathcal{F}$ is exactly $m(\mathcal{F})$. This ensures that the maximum in (3) is attained by a multiset $\mathcal{F}$ in which every vertex has equal multiplicity. We represent this with a *dominating $(p:q)$-colouring* of G, that is a mapping $\phi\colon V(G) \to \binom{[p]}{q}$ such that $\bigcup_{u \in N[v]} \phi(u) = [p]$ for every vertex $v \in V(G)$ — in other words, every closed neighbourhood spans all the colours in ϕ. When p and q are not explicit, we say that ϕ is a *fractional dominating colouring* of G. Then we have

$$\mathrm{FDOM}(G) = \max\left\{\frac{p}{q} : \text{there is a dominating } (p:q)\text{-colouring of } G.\right\}. \tag{4}$$

We are interested in determining the extremal value

$$\mathrm{FDOM}(\mathscr{G}, d) := \inf_{\substack{G \in \mathscr{G} \\ \delta(G) \geq d}} \mathrm{FDOM}(G)$$

of the fractional domatic number over graphs of minimum degree at least d within a specific class of graphs $\mathscr{G}$. Our main contributions concern the case $d = 2$ in the class of all graphs, and in the class of planar graphs of large girth.

2 Our contribution

The fractional domatic number of a graph has a high correlation with its minimum degree, as can be observed with the following bounds. The upper bound can be found in [10], and the lower bound can be deduced from the probabilistic definition of the fractional domatic number together with the proof of [3, Theorem 1.2.2].

Proposition 1. *For every graph G of minimum degree d,*

$$\frac{d+1}{1+\ln(d+1)} \leq \mathrm{FDOM}(G) \leq d+1.$$

The upper bound is tight for outerplanar graphs [5], interval graphs [13], and more generally strongly chordal graphs [6] — a graph G is *strongly chordal* if G is chordal and moreover every even cycle of G of length at least 6 has a chord that cuts it into two even cycles.

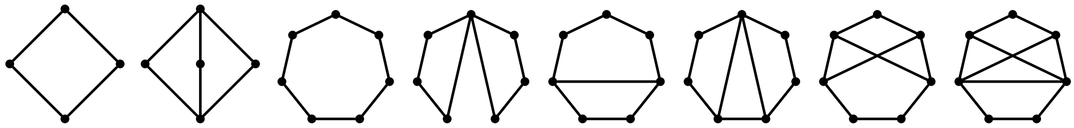
The lower bound is asymptotically tight: in [2], Alon studies the transversal number of random uniform hypergraphs, from which we infer the existence of graphs G of arbitrarily large minimum degree d such that $\mathrm{FDOM}(G) = (1 + o(1))d/\ln d$.

Graphs of minimum degree 2 Theorem 1 is essentially best possible when d is large, yet many questions remain open when d is small. If G has an isolated vertex, then clearly $\mathrm{FDOM}(G) = 1$. Otherwise, G has minimum degree at least 1, and $\mathrm{FDOM}(G) \geq \mathrm{DOM}(G) \geq 2$, as can be observed by taking a maximal independent set and its complement. Recently, Gadouleau, Harms, Mertzios, and Zamaraev [9] gave a simple characterization of every connected graph G such that $\mathrm{FDOM}(G) = 2$: a connected graph G has $\mathrm{FDOM}(G) = 2$ if and only if $\delta(G) < 2$ or $G = C_4$; otherwise, they conjectured that $\mathrm{FDOM}(G) \geq \frac{7}{3}$. This is true when restricting to (induced)$K_{1,6}$-free graphs, which is a consequence of a result of Abbas, Egerstedt, Liu, Thomas, and Whalen [1]. They proved that, except for the family $\mathcal{B}$ of eight graphs depicted in Figure 1, every connected $K_{1,6}$-free graph of minimum degree at least 2 has a dominating $(5:2)$-colouring. It is possible to drop the $K_{1,6}$-free hypothesis if one desires only a bound on the domination number rather than the fractional domatic number: McCuaig and Shepherd [15] showed that every n-vertex connected graph of minimum degree at least 2 not in $\mathcal{B} \setminus \{K_{2,3}\}$ has a dominating set of size at most $2n/5$.

Focusing on regular graphs, Fujita, Yamashita, and Kameda [8] showed that every cubic graph G (every vertex of G has degree 3) has a dominating $(5:2)$-colouring.

Our main contribution is to show a common strengthening of all the results stated above.

Theorem 2. *Let $\mathcal{B}$ be the family of graphs depicted in Figure 1. For every connected graph $G \notin \mathcal{B}$ of minimum degree at least 2, it holds that* $\mathrm{FDOM}(G) \geq \frac{5}{2}$.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 1: The family $\mathcal{B}$ — every graph $G \in \mathcal{B}$ has $\mathrm{FDOM}(G) < 5/2$.

This implies that every connected graph $G \neq C_4$ of minimum degree at least 2 has $\mathrm{FDOM}(G) \geq 7/3$, since for every $G \in \mathcal{B} \setminus \{C_4\}$, one can check that $\mathrm{FDOM}(G) = 7/3$. The value $5/2$ is best possible: one obtains a graph G' with $\mathrm{FDOM}(G') \leq 5/2$ by *gluing* a C_5 to any graph G (see Theorem 4). Moreover, even if one restricts to bipartite graphs of girth 6, it is possible to construct an infinite family of graphs whose fractional domatic number approaches $5/2$ (see Theorem 5).

Planar graphs of large girth One of the key ingredients in our proof of Theorem 2 is to show how to extend a fractional dominating colouring to a long enough *suspended path* — a maximal path with internal vertices of degree 2. This strategy lets us derive a lower bound on the fractional domatic number of planar graphs of minimum degree 2 that approaches the theoretical upper bound 3 as their girth increases.

Theorem 3. *If G is planar graph of minimum degree 2 and girth g, then* $\mathrm{FDOM}(G) = 3 - O(1/g)$.

Terminology Given a graph G and an integer k, a *suspended k-path* in G is a path P of length k with (distinct) end-points of degree at least 3 in G, and inner vertices of degree 2 in G. We denote $G \setminus P$ the graph obtained by removing the edges and inner vertices of P from G. Observe that if G has minimum degree at least 2, then so does $G \setminus P$. Observe also that if G is 2-connected, then every degree-2 vertex of G is part of a suspended path. A *hammock* in G is the union of a suspended 2-path and a suspended 3-path between two non-adjacent vertices.

Tightness The lower bound $5/2$ on the fractional domatic number of connected graphs of minimum degree 2 is best possible up to finitely many exceptional graphs. Indeed, if G_0 is any connected graph of minimum degree 2 with 2 non-adjacent vertices u, v, and G is obtained from G_0 by adding a hammock between u and v, then one can show that $\mathrm{FDOM}(G) \leq 5/2$. So we have the following proposition.

Proposition 4. *There are infinitely many connected graphs G of minimum degree 2 such that* $\mathrm{FDOM}(G) \le 5/2$.

Every graph constructed in the proof of Theorem 4 contains a copy of C_5 as an induced subgraph. So one may wonder how the fractional domatic number is affected if we forbid C_5 as a subgraph. We show that it is not possible to derive a general lower bound better than $5/2$ even if we restrict to bipartite graphs of girth 6.

Proposition 5. *For every $\varepsilon > 0$, there is an infinite family of bipartite connected graphs G of minimum degree 2 and girth 6 such that* $\mathrm{FDOM}(G) \le 5/2 + \varepsilon$.

3 Sketch of proof of Theorem 2

We prove by induction that every connected graph $G \notin \mathcal{B}$ of minimum degree at least 2 has $\mathrm{FDOM}(G) \ge 5/2$. The main inductive step is when G contains a long enough suspended path.

Lemma 6. *Fix an integer $k \ge 2$. Let G be a graph, and let P be a suspended path of G of length at least $3k-2$. Let G' be obtained by removing the internal vertices of P from G. Then*

$$\mathrm{FDOM}(G) \ge \min\left\{\frac{3k-1}{k}, \mathrm{FDOM}(G')\right\}.$$

The base case of the induction is given by the following lemma.

Lemma 7. *Let H be a multigraph of minimum degree at least 3 and multiplicity at most 2. Let G be obtained from H by subdividing each simple edge into a suspended 2-path and each double edge into a hammock. Then*

$$\mathrm{FDOM}(G) \ge \frac{5}{2}.$$

(a) A neighbourhood in H. (b) The corresponding neighbourhood in G after subdivision.

Figure 2: The subdivision rules from H to G in the statement of theorem 7.

Sketch of proof of Theorem 2. Let us assume for the sake of contradiction that there exists a counterexample Γ to the statement of Theorem 2 minimizing the number of edges. We show through the following sequence of claims that Γ must satisfy the conditions of Theorem 7. All of these claims are proved as follows: assume for the sake of contradiction that the structure appears in Γ, remove it and obtain a smaller graph Γ'. If Γ' has no connected component in $\mathcal{B}$, then one obtains by induction a fractional dominating $5/2$-colouring of Γ', which can be extended to Γ. Otherwise, there are only a finite number of possible values for Γ, and we show that $\mathrm{FDOM}(\Gamma) \ge 5/2$ for each of them.

Claim 7.1. Γ *is 2-connected.*

Claim 7.2. Γ *does not contain two adjacent 3^+-vertices.*

Claim 7.3. Γ *has no* C_4 *as a subgraph.*

Claim 7.4. Γ *does not have two suspended twin paths of length* 3*.*

Claim 7.5. *Every suspended* 3*-path of* Γ *is part of a hammock.*

Claim 7.6. *For every* $k \geq 4$*,* Γ *has no suspended* k*-path.*

Combining all the above claims, we conclude that Γ satisfies the hypotheses of Theorem 7, and so $\mathrm{FDOM}(\Gamma) \geq 5/2$, a contradiction. □

4 Planar graphs of large girth

Successive applications of Theorem 6 imply that if one can repetitively extract long suspended path from a graph G of minimum degree 2, until reaching a long cycle, then $\mathrm{FDOM}(G)$ is close to 3. This strategy can be applied when G is a planar graph of large girth, using the following intermediary result. We note that many other classes of graphs have a similar structural property (e.g. graphs of bounded genus and large girth, graphs with maximum average degree $2 + \varepsilon$, ...) and that our result could easily be extended to those.

Lemma 8. *Every* 2*-connected planar graph* G *of girth at least* $5\ell + 1$*, of minimum degree at least* 2*, and of maximum degree at least* 3 *contains a suspended* k*-path for some* $k \geq \ell + 1$*.*

We can now state our main result concerning planar graphs of large girth.

Theorem 9. *For every* $k \geq 2$*, every planar graph* G *of minimum degree at least* 2 *and of girth at least* $15k - 14$ *has fractional domatic number at least* $3 - 1/k$*.*

We note that the girth requirement in Theorem 9 is certainly not optimal: when $k = 2$, it is required that the girth of G is at least 16 to ensure that $\mathrm{FDOM}(G) \geq 5/2$, while we can infer from Theorem 2 that having girth at least 8 suffices to ensure that conclusion, even without the planar hypothesis.

References

[1] Waseem Abbas, Magnus Egerstedt, Chun-Hung Liu, Robin Thomas, and Peter Whalen. Deploying robots with two sensors in $K_{1,6}$-free graphs. *Journal of Graph Theory*, 82(3):236–252, 2016.

[2] Noga Alon. Transversal numbers of uniform hypergraphs. *Graphs and Combinatorics*, 6(1):1–4, March 1990.

[3] Noga Alon and Joel H Spencer. *The probabilistic method.* John Wiley & Sons, 2016.

[4] Béla Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, pages 433–436, 1981.

[5] Ernest J Cockayne and Stephen T Hedetniemi. Towards a theory of domination in graphs. *Networks*, 7(3):247–261, 1977.

[6] Martin Farber. Domination, independent domination, and duality in strongly chordal graphs. *Discrete Applied Mathematics*, 7(2):115–130, 1984.

[7] Uriel Feige, Magnús M. Halldórsson, Guy Kortsarz, and Aravind Srinivasan. Approximating the domatic number. *SIAM Journal on Computing*, 32(1):172–195, 2002.

[8] Satoshi Fujita, Tiko Kameda, and Masafumi Yamashita. A study on r-configurations—a resource assignment problem on graphs. *SIAM Journal on Discrete Mathematics*, 13(2):227–254, 2000.

[9] Maximilien Gadouleau, Nathaniel Harms, George B Mertzios, and Viktor Zamaraev. Graphs with minimum fractional domatic number. *Discrete Applied Mathematics*, 343:140–148, 2024.

[10] Wayne Goddard and Michael A Henning. Fractional domatic, idomatic, and total domatic numbers of a graph. In *Structures of Domination in Graphs*, pages 79–99. Springer, 2020.

[11] Haim Kaplan and Ron Shamir. The domatic number problem on some perfect graph families. *Information Processing Letters*, 49(1):51–56, 1994.

[12] Jan Kratochv, Andrzej Proskurowski, and J Telle. Complexity of graph covering problems. *Nordic Journal of Computing*, 5:173–195, 1998.

[13] Tung-Lin Lu, Pei-Hsin Ho, and Gerard J Chang. The domatic number problem in interval graphs. *SIAM Journal on Discrete Mathematics*, 3(4):531–536, 1990.

[14] Lesley R Matheson and Robert E Tarjan. Dominating sets in planar graphs. *European Journal of Combinatorics*, 17(6):565–568, 1996.

[15] William McCuaig and Bruce Shepherd. Domination in graphs with minimum degree two. *Journal of Graph Theory*, 13(6):749–762, 1989.

[16] Herbert Ellis Robbins. A theorem on graphs, with an application to a problem of traffic control. *The American Mathematical Monthly*, 46(5):281–283, 1939.

[17] Edward R Scheinerman and Daniel H Ullman. *Fractional graph theory: a rational approach to the theory of graphs*. Courier Corporation, 2013.

[18] Alex Scott and Paul Seymour. A survey of -boundedness. *Journal of Graph Theory*, 95(3):473–504, 2020.

[19] Douglas R Stinson. Combinatorial designs: constructions and analysis. *ACM SIGACT News*, 39(4):17–21, 2008.

[20] Jukka Suomela. Locality helps sleep scheduling. In *Working Notes of the Workshop on World-Sensor-Web: Mobile Device-Centric Sensory Networks and Applications*, pages 41–44. Citeseer, 2006.

[21] Jennifer Seberry Wallis. On the existence of hadamard matrices. *Journal of Combinatorial Theory, Series A*, 21(2):188–195, 1976.

[22] Bohdan Zelinka. Domatic number and degrees of vertices of a graph. *Mathematica Slovaca*, 33(2):145–147, 1983.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Path Eccentricity and Forbidden Induced Subgraphs

(Extended abstract)

Sylwia Cichacz* Claire Hilaire† Tomáš Masařík‡ Jana Masaříková§
Martin Milanič¶

Abstract

The path eccentricity of a connected graph G is the minimum integer k such that G has a path such that every vertex is at distance at most k from the path. A result of Duffus, Jacobson, and Gould from 1981 states that every connected $\{\text{claw}, \text{net}\}$-free graph G has a Hamiltonian path, that is, G has path eccentricity 0. Several more recent works identified various classes of connected graphs with path eccentricity at most 1, or, equivalently, graphs having a spanning caterpillar, including connected P_5-free graphs, AT-free graphs, and biconvex graphs. Generalizing all these results, we apply the work on structural distance domination of Bacsó and Tuza [Discrete Math., 2012] and characterize, for every positive integer k, graphs such that every connected induced subgraph has path eccentricity less than k. More specifically, we show that every connected $\{S_k, T_k\}$-free graph has a path eccentricity less than k, where S_k and T_k are two specific graphs of path eccentricity k (a subdivided claw and the line graph of such a graph). As a consequence, every connected H-free graph has path eccentricity less than k if and only if H is an induced subgraph of $3P_k$ or $P_{2k+1} + P_{k-1}$. Our main result also answers an open question of Bastide, Hilaire, and Robinson [Discrete Math., 2025].

*Faculty of Applied Mathematics, AGH University of Krakow, Poland. E-mail:cichacz@agh.edu.pl.

†FAMNIT, University of Primorska, Koper, Slovenia. E-mail: claire.hilaire@upr.si. Supported in part by the Slovenian Research and Innovation Agency (research projects J1-4008 and N1-0370).

‡University of Warsaw, Institute of Informatics, Poland. E-mail: masarik@mimuw.edu.pl. Supported by the Polish National Science Centre SONATA-17 grant number 2021/43/D/ST6/03312.

§University of Warsaw, Institute of Informatics, Poland. E-mail: j.novotna@mimuw.edu.pl. Supported by the Polish National Science Centre Preludium research project no. 2022/45/N/ST6/04232.

¶FAMNIT and IAM, University of Primorska, Koper, Slovenia. E-mail: martin.milanic@upr.si. This work is supported in part by the Slovenian Research and Innovation Agency (I0-0035, research program P1-0285 and research projects J1-3003, J1-4008, J1-4084, J1-60012, and N1-0370) and by the research program CogniCom (0013103) at the University of Primorska.

The full preprint version of this article is available on the arXiv [CHM+25].

1 Introduction

A *Hamiltonian path* in a graph G is a path in G whose vertex set is $V(G)$; a Hamiltonian cycle is defined similarly. The problems of determining whether a given graph has a Hamiltonian path or a Hamiltonian cycle are both NP-complete [GJ79, Kar72], which motivates the search for sufficient conditions for their existence (see, e.g., the survey [Gou14]). One of the early results in the area is the following result of Duffus, Jacobson, and Gould [DJG81] (see also [She91, BD03, Kel06]).

Theorem 1.1. *Every connected {claw, net}-free graph has a Hamiltonian path.*

Here, a *claw* is the complete bipartite graph $K_{1,3}$, the *net* is the graph obtained by appending a pendant edge to each vertex of the complete graph K_3, and, given a set $\mathcal{F}$ of graphs, a graph G is said to be $\mathcal{F}$*-free* if G does not have an induced subgraph isomorphic to any member of $\mathcal{F}$. If $\mathcal{F} = \{F\}$ for some graph F, we also say that G is F*-free*.

The result of Theorem 1.1 is tight, since the claw and the net do not have a Hamiltonian path. Let us remark that, in general, if G is a graph that has a Hamiltonian path and H is a connected induced subgraph of G, then H may fail to have a Hamiltonian path. For example, if H is any n-vertex graph and G is a graph obtained from H by adding to it n new vertices fully adjacent to the vertices of H, then G has a Hamiltonian path, even though H may fail to have one. In particular, this construction could be applied to the claw and the net.

If we instead require that all connected induced subgraphs have a Hamiltonian path, Theorem 1.1 implies the following characterization.

Corollary 1.2. *For every graph G, the following statements are equivalent:*

- *Every connected induced subgraph of G has a Hamiltonian path.*
- *G is {claw, net}-free.*

In 2023, Gómez and Gutiérrez [GG23] generalized the concept of Hamiltonian paths by defining the notion of path eccentricity of graphs. Given a connected graph G and a path P in G, the *eccentricity* of P is the maximum distance from a vertex in G to a vertex in P. Hence, in particular, a path P has eccentricity 0 if and only if P is a Hamiltonian path, and P has eccentricity at most 1 if and only if P is *dominating*, that is, every vertex not in the path has a neighbor on the path. More generally, given an integer $k \geq 0$, a path P is said to be k*-dominating* if it has eccentricity at most k. The *path eccentricity* of a connected graph G is denoted by $\mathbf{pe}(G)$ and defined as the minimum eccentricity of a path in G. Hence, $\mathbf{pe}(G) = 0$ if and only if G has a Hamiltonian path, and $\mathbf{pe}(G) \leq 1$ if and only if G has a dominating path. It is not difficult to see that a graph G has a dominating path if and only if G has a *spanning caterpillar*, that is, a spanning subgraph that is a caterpillar (a tree possessing a dominating path). Hence, understanding graphs that admit dominating paths is motivated by their applications to graph burning and a variant of the cops and robber game. More precisely:

- The burning graph conjecture (see Bonato, Janssen, and Roshanbin [BJR16]) is known to be true for graphs admitting a spanning caterpillar.
- If a graph G admits a spanning caterpillar T, then the minimum number of cops sufficient for the cop player to win on the graph G against the robber moving at the speed of s is at

most ps if G is a subgraph of the graph obtained from T by adding an edge between each pair of vertices that are at distance at most p in T (see Fomin, Golovach, Kratochvíl, Nisse, and Suchan [FGK+10, Lemma 4]).

Since adding a universal vertex to any connected graph H results in a graph G with path eccentricity at most 1, path eccentricity is not monotone under vertex deletion. Nevertheless, and in line with the above discussion preceding Corollary 1.2, which characterizes graphs each connected induced subgraph of which has path eccentricity 0, analogous questions can be addressed for higher values of path eccentricity.

Question 1.3. For a positive integer k, what are the graphs G such that each connected induced subgraph of G has path eccentricity at most k?

For each fixed $k \geq 1$, there are two obstructions to path eccentricity less than k that can be obtained from the claw and the net, as follows. Given a positive integer k, let us denote by S_k the *k-subdivided claw*, that is, the graph obtained from the claw by replacing each of its edges with a path of length k, where the length of a path is the number of its edges (in particular, S_1 is the claw). Furthermore, we denote by T_k the *k-subdivided net*, that is, the graph obtained by appending a path of length k to each vertex of the complete graph K_3 (in particular, T_1 is the net).

It is easy to observe that $\mathbf{pe}(S_k) = \mathbf{pe}(T_k) = k$ for each positive integer k. Consequently, if G is a graph such that each connected induced subgraph of G has path eccentricity less than k, then G is $\{S_k, T_k\}$-free. As the main result of this paper, we show that this necessary condition is in fact also sufficient. We thus obtain the following unifying theorem, valid for all positive integers k (and thus capturing also Corollary 1.2, which corresponds to the case $k = 1$).

Theorem 1.4. *For every integer $k \geq 1$ and every graph G, the following statements are equivalent:*

- *Every connected induced subgraph H of G satisfies $\mathbf{pe}(H) < k$.*
- *G is $\{S_k, T_k\}$-free.*

The case $k = 2$ of Theorem 1.4 characterizes graphs every connected induced subgraph of which has a dominating path, and implies the following.

Corollary 1.5. *Every connected $\{S_2, T_2\}$-free graph has a dominating path.*

Corollary 1.5 unifies and generalizes several results from the literature regarding sufficient conditions for the existence of a dominating path:

1. One of the oldest results along these lines is the result of Bacsó and Tuza in 1990 (see [BT90]) stating that every connected P_5-free graph has a dominating clique or a dominating P_3 (we denote by P_k the k-vertex path). It follows that every connected P_5-free graph has a dominating path.

2. A graph G is said to be *AT-free* if it does not have an independent set I such that $|I| = 3$ and for every vertex $v \in I$, the remaining two vertices in I are in the same component of $G - N[v]$. Every connected AT-free graph has a dominating path (see Corneil, Olariu, and Stewart [COS97], as well as [COS95]).

3. A graph G is *biconvex* if it is bipartite, with parts A and B that can each be linearly ordered so that for each vertex v of G, the neighborhood of v in the part not containing v forms a consecutive segment of vertices with respect to the linear ordering. Gómez and Gutiérrez [GG23] and, independently Antony, Das, Gosavi, Jacob, and Kulamarva [ADG$^+$24], proved that every connected biconvex graph has a dominating path (or, equivalently, a spanning caterpillar).

Indeed, each of the previously stated cases corresponds to a subclass of the class of $\{S_2, T_2\}$-free graphs, hence, Corollary 1.5 implies all these results.

Further consequences of Theorem 1.4 are related to a recent work of Bastide, Hilaire, and Robinson [BHR25]. The authors explored a relation between the path eccentricity of a graph and a certain property of its adjacency matrix (or equivalently on the neighborhoods of its vertices). They introduced the *partially augmented consecutive ones property* (denoted **-C1P* for short): a graph has the *-C1P if there exists an ordering of its vertices sending the open or closed neighborhood of every vertex to a consecutive set. The authors showed that graphs with the *-C1P have path eccentricity at most 2, and conjectured that they actually have path eccentricity at most 1. We show that Theorem 1.4 resolves the conjecture.

Theorem 1.6. *If a graph G has the *-C1P, then* $\mathbf{pe}(G) \leq 1$.

Another result from [BHR25] is related to the following generalization of AT-freeness. A graph G is said to be *k-AT-free* if G does not have an independent set I such that $|I| = 3$ and for every vertex $v \in I$, the remaining two vertices in I are in the same component of the graph $G - \{u\colon u \text{ is at distance at most } k \text{ of } v\}$. In particular, a graph is AT-free if and only if it is 1-AT-free. In [BHR25], the authors proved that for every $k \geq 1$, if a graph G is k-AT-free, then $\mathbf{pe}(G) \leq k$. This result follows from Theorem 1.4. Indeed, it is easy to check that for every $k \geq 2$, the graphs S_k and T_{k-1} are not $(k-1)$-AT-free. Thus, if a graph G is $(k-1)$-AT-free, then G is $\{S_k, T_k\}$-free, which implies by Theorem 1.4 that $\mathbf{pe}(G) < k$.

For the case where a single induced subgraph is excluded, Theorem 1.4 leads to the following characterization. The symbol $+$ denotes the disjoint union of graphs, and for an integer $k \geq 0$ and a graph H, we denote by kH the disjoint union of k copies of H.

Corollary 1.7. *Let H be a graph and $k \geq 1$. Then, the following statements are equivalent:*

- *Every connected H-free graph G has* $\mathbf{pe}(G) < k$.
- *H is an induced subgraph of $3P_k$ or $P_{2k+1} + P_{k-1}$.*

In the case of $3P_2$-free graphs, we obtain a much stronger property. We can find a path with eccentricity at most 1 that is a longest path in the graph. While the proof of this result implies that we can transform a longest path into a longest path that is dominating in polynomial time, it is NP-hard to find a longest path in $3P_2$-free graphs. In fact, determining whether a given graph has a Hamiltonian path is NP-complete in the class of split graphs [Mül96], a subclass of $2P_2$-free graphs.

Theorem 1.8. *Every connected $3P_2$-free graph G has a longest path that is dominating. Moreover, we can find a (not necessarily longest) dominating path in G in time $\mathcal{O}(n^2(n+m))$, where n is the number of vertices and m is the number of edges in G.*

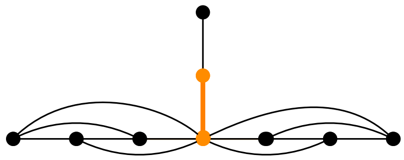


Figure 1: An example of a $\{P_5, 4P_2\}$-free graph no longest path of which is dominating. The vertices depicted in orange need to belong to every dominating path; however, no longest path contains the top orange vertex.

This is a very fine line, as this property is completely violated in the case of $\{S_2, T_2\}$-free graphs or even $\{P_5, 4P_2\}$-free graphs; see Figure 1. It also cannot be generalized to higher eccentricity; see for example a $\{3P_3, P_7 + P_2\}$-free graph on Figure 2. Moreover, in a $3P_2$-free graph there might exist a longest path that is not dominating, as depicted in Figure 3.

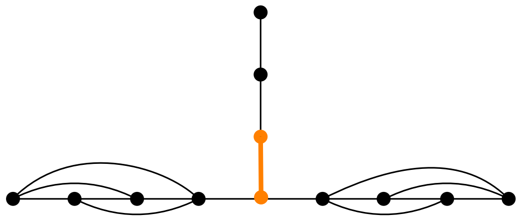


Figure 2: An example of a $\{3P_3, P_7 + P_2\}$-free graph no longest path of which is 2-dominating. The vertices depicted in orange need to belong to every 2-dominating path; however, no longest path contains the top orange vertex.

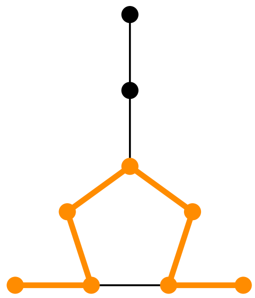


Figure 3: An example of $3P_2$-free graph with one of the longest paths (depicted in orange) that is not dominating.

While Theorem 1.8 actually provides us with a polynomial-time algorithm that finds a dominating path, the algorithm does not necessarily find a *longest* path that is dominating, unless we are given another longest path on the input.

Organization of the paper. In Section 2, we give an overview of known results on structural domination, including a key result of Bacsó and Tuza [BT12] that we use in the proof of our main theorem (Theorem 1.4), which we prove in Section 3. We postpone the proofs of our remaining theorems to the full version of this paper [CHM+25].

2 Background on Structural Domination

Theorem 1.4 is closely related to the framework of *structural domination*. Given a class $\mathcal{D}$ of connected graphs, let us say that a graph G is *hereditarily dominated* by $\mathcal{D}$ if each connected induced subgraph of G contains a dominating set that induces a graph from $\mathcal{D}$. Bacsó, Michalak, and Tuza (see [BMT05]) gave forbidden induced subgraph characterizations of graphs hereditarily dominated by classes of complete bipartite graphs, stars, connected bipartite graphs, and complete k-partite graphs. Similar results were obtained by Michalak (see [Mic07]) for graphs hereditarily dominated by classes of cycles and paths, paths, trees, and trees with bounded diameter. Camby and Schaudt (see [CS16]) proved that for every $k \geq 4$, a graph G is P_k-free if and only if it is hereditarily dominated by the class consisting of all connected P_{k-2}-free graphs along with the k-vertex cycle.

The most general result along these lines was obtained independently by Bacsó and Tuza (see [Tuz08, Bac09]). To state the result, we need one more definition. Given a connected graph H, the *leaf graph* of H is the graph $H^{\Leaf}$ obtained from H by attaching a leaf to each vertex $v \in V(H)$ that is not a cut-vertex (that is, the graph $H - v$ is connected). By definition, these new leaves are pairwise distinct and non-adjacent.

Theorem 2.1 (Bacsó [Bac09] and Tuza [Tuz08])**.** *Let $\mathcal{D}$ be a class of connected graphs that contains the class of all paths, excludes at least one connected graph, and is closed under taking connected induced subgraphs. Let $\mathcal{F}$ be the family of all connected graphs F that do not belong to $\mathcal{D}$, but all proper connected induced subgraphs of F belong to $\mathcal{D}$. Let $\mathcal{F}^{\Leaf}$ be the family of all leaf graphs of graphs in $\mathcal{F}$. Then, a graph G is hereditarily dominated by $\mathcal{D}$ if and only if G is $\mathcal{F}^{\Leaf}$-free.*

The main result of [Tuz08, Bac09] is even more general, without the restriction that $\mathcal{D}$ contains the class of all paths; but since we do not need this result, we omit the precise statement.

A generalization of Theorem 2.1 to domination at distance was proved by Bacsó and Tuza (see [BT12]). A *distance-k dominating set* in a graph G is a set $D \subseteq V(G)$ such that every vertex in G is at distance at most k from some vertex in D. If, in addition, the graph $G[D]$ is connected, then D is said to be a *connected distance-k dominating set.* A graph G is said to be *hereditarily k-dominated* by a class $\mathcal{D}$ of connected graphs if each connected induced subgraph of G admits a distance-k dominating set that induces a graph from $\mathcal{D}$. Given a connected graph H, the *k-leaf graph* of H is the graph $H_k^{\Leaf}$ obtained from H by attaching a pendant path of length k to each vertex $v \in V(H)$ that is not a cut-vertex. Bacsó and Tuza (see [BT12]) generalized Theorem 2.1 as follows.

Theorem 2.2 (Bacsó and Tuza [BT12])**.** *Let $\mathcal{D}$ be a class of connected graphs that contains the class of all paths, excludes at least one connected graph, and is closed under taking connected induced subgraphs. Let $\mathcal{F}$ be the family of all connected graphs F that do not belong to $\mathcal{D}$, but all proper connected induced subgraphs of F belong to $\mathcal{D}$. Let k be a positive integer and let $\mathcal{F}_k^{\Leaf}$ be the family of all k-leaf graphs of graphs in $\mathcal{F}$. Then, a graph G is hereditarily k-dominated by $\mathcal{D}$ if and only if G is $\mathcal{F}_k^{\Leaf}$-free.*

Note that Theorem 2.1 is a special case of Theorem 2.2 for the case $k = 1$.

3 Proof of General Theorem (Theorem 1.4)

Proof of Theorem 1.4. Fix an integer $k \geq 1$. Let G be a graph such that every connected induced subgraph H of G satisfies $\mathbf{pe}(H) < k$. To show that G is $\{S_k, T_k\}$-free, it suffices to show that $\mathbf{pe}(S_k) \geq k$ and $\mathbf{pe}(T_k) \geq k$. Let H be a graph isomorphic to either S_k or T_k and let S be any path in H. We claim that S has eccentricity at least k. Suppose not. Let C be the set of vertices of degree 3 in H. Then, the graph $H - V(C)$ consists of three disjoint paths P, Q, and R, each with k vertices. Moreover, each of the paths P, Q, and R contains a vertex at distance k from $V(C)$. Since the eccentricity of S is less than k, each of the paths P, Q, and R contains a vertex of S. Since S has at most two endpoints, at most two of the paths P,

Q, and R contain an endpoint of P; hence, we may assume without loss of generality that P contains an internal vertex of S but no endpoint of S. However, this implies that two distinct edges of S connect P with the rest of H, a contradiction. This shows that $\mathbf{pe}(H) \geq k$, as claimed.

It remains to prove that if G is $\{S_k, T_k\}$-free, then every connected induced subgraph H of G has path eccentricity less than k, or, equivalently, that every connected $\{S_k, T_k\}$-free graph has path eccentricity less than k. For $k = 1$, the statement coincides with that of Theorem 1.1. Suppose now that $k \geq 2$. Let $\mathcal{F} = \{S_1, T_1\}$ and let $\mathcal{D}$ be the class of all connected $\mathcal{F}$-free graphs. Note that $\mathcal{F}^{\Leaf}_{k-1} = \{S_k, T_k\}$. Consequently, since $\mathcal{D}$ satisfies the hypothesis of Theorem 2.2, it follows that every $\{S_k, T_k\}$-free graph G is hereditarily $(k-1)$-dominated by $\mathcal{D}$. Let H be a connected induced subgraph of G. Then, since G is hereditarily $(k-1)$-dominated by $\mathcal{D}$, there exists a distance-$(k-1)$ dominating set D in H such that the induced subgraph $H[D]$ belongs to $\mathcal{D}$. By Theorem 1.1, the graph $H[D]$ admits a Hamiltonian path P. It follows that P is a path in H with eccentricity less than k, implying that $\mathbf{pe}(H) < k$. □

Acknowledgements. The authors are grateful to Ulrich Pferschy for pointing out reference [ADG+24], which motivated the present research.

References

[ADG+24] Dhanyamol Antony, Anita Das, Shirish Gosavi, Dalu Jacob, and Shashanka Kulamarva. Spanning caterpillar in biconvex bipartite graphs. *Discrete Appl. Math.*, 356:32–36, 2024. doi:10.1016/j.dam.2024.05.016.

[Bac09] Gábor Bacsó. Complete description of forbidden subgraphs in the structural domination problem. *Discrete Math.*, 309(8):2466–2472, 2009. doi:10.1016/j.disc.2008.05.053.

[BD03] Andreas Brandstädt and Feodor F. Dragan. On linear and circular structure of (claw, net)-free graphs. *Discrete Appl. Math.*, 129(2-3):285–303, 2003. doi:10.1016/S0166-218X(02)00571-1.

[BHR25] Paul Bastide, Claire Hilaire, and Eileen Robinson. Path eccentricity of k-AT-free graphs and application on graphs with the consecutive ones property. *Discrete Math.*, 348(7):Paper No. 114449, 2025. doi:10.1016/j.disc.2025.114449.

[BJR16] Anthony Bonato, Jeannette Janssen, and Elham Roshanbin. How to burn a graph. *Internet Math.*, 12(1-2):85–100, 2016. doi:10.1080/15427951.2015.1103339.

[BMT05] Gábor Bacsó, Danuta Michalak, and Zsolt Tuza. Dominating bipartite subgraphs in graphs. *Discuss. Math. Graph Theory*, 25(1-2):85–94, 2005. doi:10.7151/dmgt.1263.

[BT90] Gábor Bacsó and Zsolt Tuza. Dominating cliques in P_5-free graphs. *Period. Math. Hungar.*, 21(4):303–308, 1990. doi:10.1007/BF02352694.

[BT12] Gábor Bacsó and Zsolt Tuza. Distance domination versus iterated domination. *Discrete Math.*, 312(17):2672–2675, 2012. doi:10.1016/j.disc.2011.12.008.

[CHM+25] Sylwia Cichacz, Claire Hilaire, Tomáš Masařík, Jana Masaříková, and Martin Milanič. Path eccentricity and forbidden induced subgraphs, 2025. `arXiv:2503.15747`.

[COS95] Derek G. Corneil, Stephan Olariu, and Lorna Stewart. A linear time algorithm to compute a dominating path in an AT-free graph. *Inform. Process. Lett.*, 54(5):253–257, 1995. `doi:10.1016/0020-0190(95)00021-4`.

[COS97] Derek G. Corneil, Stephan Olariu, and Lorna Stewart. Asteroidal triple-free graphs. *SIAM J. Discrete Math.*, 10(3):399–430, 1997. `doi:10.1137/S0895480193250125`.

[CS16] Eglantine Camby and Oliver Schaudt. A new characterization of P_k-free graphs. *Algorithmica*, 75(1):205–217, 2016. `doi:10.1007/s00453-015-9989-6`.

[DJG81] D. Duffus, M. S. Jacobson, and R. J. Gould. Forbidden subgraphs and the Hamiltonian theme. In *The Theory and Applications of Graphs (Kalamazoo, Mich., 1980)*, pages 297–316. Wiley, New York, 1981.

[FGK+10] Fedor V. Fomin, Petr A. Golovach, Jan Kratochvíl, Nicolas Nisse, and Karol Suchan. Pursuing a fast robber on a graph. *Theoret. Comput. Sci.*, 411(7-9):1167–1181, 2010. `doi:10.1016/j.tcs.2009.12.010`.

[GG23] Renzo Gómez and Juan Gutiérrez. Path eccentricity of graphs. *Discrete Appl. Math.*, 337:1–13, 2023. `doi:10.1016/j.dam.2023.04.012`.

[GJ79] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness.* W. H. Freeman and Co., San Francisco, CA, 1979.

[Gou14] Ronald J. Gould. Recent advances on the Hamiltonian problem: Survey III. *Graphs Combin.*, 30(1):1–46, 2014. `doi:10.1007/s00373-013-1377-x`.

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, The IBM Research Symposia Series, pages 85–103. Plenum, New York-London, 1972. `doi:10.1007/978-1-4684-2001-2_9`.

[Kel06] Alexander Kelmans. On Hamiltonicity of $\{\text{claw}, \text{net}\}$-free graphs. *Discrete Math.*, 306(21):2755–2761, 2006. `doi:10.1016/j.disc.2006.04.022`.

[Mic07] Danuta Michalak. Hereditarily dominated graphs. *Discrete Math.*, 307(7-8):952–957, 2007. `doi:10.1016/j.disc.2005.11.044`.

[Mül96] Haiko Müller. Hamiltonian circuits in chordal bipartite graphs. *Discrete Math.*, 156(1-3):291–298, 1996. `doi:10.1016/0012-365X(95)00057-4`.

[She91] F. Bruce Shepherd. Hamiltonicity in claw-free graphs. *J. Combin. Theory Ser. B*, 53(2):173–194, 1991. `doi:10.1016/0095-8956(91)90074-T`.

[Tuz08] Zsolt Tuza. Hereditary domination in graphs: characterization with forbidden induced subgraphs. *SIAM J. Discrete Math.*, 22(3):849–853, 2008. `doi:10.1137/070699482`.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Graham's rearrangement for dihedral groups

(Extended abstract)

Simone Costa* Stefano Della Fiore † Eva Engel ‡

Abstract

A famous conjecture of Graham asserts that every set $A \subseteq \mathbb{F}_p \setminus \{0\}$ can be ordered so that all partial sums are distinct. Bedert and Kravitz proved in [4] that this statement holds whenever $|A| \leqslant e^{c(\log p)^{1/4}}$.

In this extended abstract, we will outline the proof (we refer to [7] for all the details) of a similar procedure for obtaining an upper bound of the same type in the case of dihedral groups Dih_p.

1 Introduction

Let A be a finite subset of group $(G, \cdot)$. We say that an ordering $a_1, \ldots, a_{|A|}$ of A is *valid* if the partial products (or partial sums in additive notation) $p_1 = a_1, p_2 = a_1 \cdot a_2, \ldots, p_{|A|} = a_1 \cdot a_2 \cdot \cdots \cdot a_{|A|}$ are all distinct. Moreover, this ordering is a sequencing if it is valid and $p_i \neq id$ for any $i \in [1, |A| - 1]$. In this case, we will say that A is sequenceable. In the literature, there are several conjectures about valid orderings and sequenceability. We refer to [9, 13, 15] for an overview of the topic, [1–3, 8] for lists of related conjectures and [5]. Here, we explicitly recall just that of Graham, who conjectured that every set of non-zero elements of $\mathbb{F}_p$ has a valid ordering.

*DICATAM, Sez. Matematica, Università degli Studi di Brescia, Via Branze 43, I 25123 Brescia, Italy, E-mail: simone.costa@unibs,it

†DII, Università degli Studi di Brescia, Via Branze 43, I 25123 Brescia, Italy, E-mail: stefano.dellafiore@unibs.it

‡Princeton University, Department of Mathematics, Class of 2027, Princeton, NJ 08544, USA, E-mail: eva.engel@princeton.edu

Conjecture 1.1 ([11] and [10]). *Let p be a prime. Then every subset $A \subseteq \mathbb{F}_p \setminus \{0\}$ has a valid ordering.*

Until recently, the main results on this conjecture were for small values of $|A|$; in particular, in [8], the conjecture was proved for sets A of size at most 12.

The first result involving arbitrarily large sets A was presented last year by Kravitz [12]. He used a simple rectification argument to show that Graham's conjecture holds for all sets A of size $|A| \leqslant \log p / \log \log p$. A similar argument was also proposed (but not published) by Sawin [16] in a 2015 MathOverflow post.

Then, in [4], Bedert and Kravitz improved this upper bound to the following:

Theorem 1.2 ([4]). *Let p be a large enough prime and let $c > 0$. Then every subset $A \subseteq \mathbb{F}_p \setminus \{0\}$ is sequenceable provided that*

$$|A| \leqslant e^{c(\log p)^{1/4}}.$$

Graham's conjecture over non-abelian groups was first studied by Ollis in [14] over dihedral groups. More recently, the authors in [5] presented a nice connection with rainbow paths and produced the asymptotic result that a subset A of a finite group G admits an ordering in which at least $(1 - o(1))|A|$ many partial products are distinct. Returning to dihedral groups, inspired by the result of Kravitz ([12]) on cyclic groups, Costa and Della Fiore proved in [6] that any subset A of $Dih_p \setminus \{id\}$ (and in general, of a class of semidirect products) is sequenceable, provided that $p > \frac{|A|!}{2}$. Note that this is slightly better than $|A| \leq \log p / \log \log p$. This extended abstract is dedicated to outlining the proof that, with a similar procedure as that of Bedert and Kravitz, subsets of the dihedral group $Dih_p \setminus \{id\}$ are sequenceable if they are not too large with respect to p. Specifically, we use the rectification argument presented in [4] to improve the results of [6] in the case of dihedral groups, ultimately obtaining a super-logarithmic bound on set size with respect to p that is identical to the bound presented by Bedert and Kravitz for sets in $\mathbb{F}_p$. Note that our result implies that every subset $A \subseteq Dih_p \setminus \{id\}$ has a valid ordering given that the stated conditions are met, since sequenceability is more restrictive than having a valid ordering.

Theorem 1.3. *Let p be a large enough prime and let $c > 0$. Then every subset $A \subseteq Dih_p \setminus \{id\}$ is sequenceable provided that*

$$|A| \leqslant e^{c(\log p)^{1/4}}.$$

The main difficulty in adapting the approach of [4] to get this theorem is that the dihedral groups are not abelian and hence the product of the elements of a set is not fixed. For this reason, we have to impose some restrictions on the positions of certain elements.

We let Dih_p denote the dihedral group $\mathbb{F}_p \rtimes_\varphi \mathbb{F}_2$, where the group operation is defined as $(x_1, a_1) \cdot_G (x_2, a_2) = (x_1 + (\varphi_{a_1} x_2), a_1 + a_2)$ where $\varphi_0 = 1$ and $\varphi_1 = -1$. We also define map $\pi_i(x)$ as the projection onto the ith coordinate of $x \in \mathbb{F}_p \rtimes_\varphi \mathbb{F}_2$, where $i \in \{1,2\}$. Given a not necessarily abelian group G, a subset $D = \{d_1, \ldots, d_r\} \subseteq G$ is *dissociated* if $d_{\sigma(1)}^{\epsilon_{\sigma(1)}} \cdot \cdots \cdot d_{\sigma(r)}^{\epsilon_{\sigma(r)}} \neq id$ for all permutations $\sigma \in Sym(r)$ and $(\epsilon_{\sigma(1)}, \ldots, \epsilon_{\sigma(r)}) \in \{-1,0,1\}^r \setminus \{(0,\ldots,0)\}$. The *dimension* of a subset $B \subseteq G$, written $\dim(B)$, is the size of the largest dissociated set contained in B. We define $g = \Theta(f)$ if and only if $f \geq 0$ and $Cf \leq g \leq C'f$ for some constants $C, C' > 0$.

2 Proof sketch

Following [4], we state a lemma that allows us to "rectify" small sets. We define, for each (nonempty) subset $A \subseteq Dih_p$, the parameter

$$R = R(A) := c_1 \max\left((\log p)^{1/2}, \frac{\log p}{\log |A|}\right), \tag{1}$$

where c_1 is a sufficiently small absolute constant. We consider A not fully contained in $\{(x,0) : x \in \mathbb{F}_p\}$, otherwise the statement of Theorem 1.3 follows from that of 1.2.

Lemma 2.1 (See [7] Lemma 3.3). *If $B \subseteq Dih_p$ is a nonempty subset with dimension $\dim(B) < R = R(B)$, then there is some $\phi \in Aut(Dih_p)$ such that the image $\pi_1(\phi(B))$ (where π_1 is the projection over the $\mathbb{F}_p$ component) is contained in $\left(-\frac{p}{90(|B|+1)}, \frac{p}{90(|B|+1)}\right)$.*

Now we show that any subset A of Dih_p can be split in a family of dissociated sets and a remainder E of small dimension such that $\pi_1(E)$ can be embedded in a small interval.

Theorem 2.2 (Structure Theorem — See [7] Theorem 3.5). *Let p be large enough. For every nonempty subset $A \subseteq Dih_p \setminus \{id\}$ not contained in $\{(x,0) : x \in \mathbb{F}_p\}$, there is some $\phi \in Aut(Dih_p)$ such that $\phi(A)$ can be partitioned as*

$$\phi(A) = E \cup \left(\cup_{j=1}^{s} D_j\right),$$

where

(i) $dim(E) < R$ and $E \not\subseteq \{(x,0) : x \in \mathbb{F}_p\}$.

Moreover, if $s > 0$,

(ii) Each D_j is a dissociated set of size $\Theta(R)$ a multiple of eight.

(iii) We can split the interval $[1,s] = L_0 \cup L_1$ so that: 1 and s belong to the same set L_i; for any $j \in L_1$, $D_j \subseteq \{(x,1) : x \in \mathbb{F}_p\}$; and, for any $j \in L_0$, $D_j \subseteq \{(x,0) : x \in \mathbb{F}_p\}$.

(iv) For each $j \in L_1$, we split D_j into D_j^o and D_j^e so that $|D_j^o| = |D_j^e| = |D_j|/2$. Then we set

$$\delta := \left(\prod_{j\in L_1} \prod_{d^o \in D_j^o, d^e \in D_j^e} d^o \cdot d^e\right)\left(\prod_{j \in L_0}\prod_{d \in D_j} d\right)$$

(which does not depend on the orderings on D_j^o and D_j^e). Then $\delta = (z_0, 0)$ where $z_0 \neq 0$ and

$$\pi_1(E \cup \{\delta\}) \subseteq \left(-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)}\right).$$

(v) $D_1 \cup D_s \cup \{\delta\}$ is a dissociated set and $|D_1| = |D_s|$.

Now, the proof proceeds mainly in two steps. Due to the Structure Theorem, we can assume, without loss of generality, that $A = E \cup (\cup_{j=1}^s D_j)$ where each set D_j is dissociated and satisfies the hypotheses of Theorem 2.2, $E \not\subseteq \{(x, 0) : x \in \mathbb{F}_p\}$, and $\pi_1(E) \subseteq \left(-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)}\right)$. To order E, we define

$$P = \left\{x \in E : \pi_2(x) = 0, \text{ and } \pi_1(x) \in \left(0, \frac{p}{90(|E|+1)}\right)\right\},$$

$$N = \left\{x \in E : \pi_2(x) = 0, \text{ and } \pi_1(x) \in \left(-\frac{p}{90(|E|+1)}, 0\right)\right\},$$

$$S = \{x \in E : \pi_2(x) = 1\}.$$

Also we split S, which is non-empty for Theorem 2.2, into $S = S_e \cup S_o$ so that $|S_e| = \lceil |S|/2 \rceil$, $|S_o| = \lfloor |S|/2 \rfloor$ and $\forall x \in S_e, \forall y \in S_o$, $\pi_1(x) > \pi_1(y)$ with respect to the natural ordering on elements in the interval $\left(-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)}\right)$. Finally, given an ordering $\mathbf{x} = x_0, x_1, \ldots, x_m$ we consider the set of partial products $\mathrm{IS}(\mathbf{x}) = \{x_0 \cdot x_1 \cdots x_i : i \in [0, m]\}$. Now we show that, for any family of subsets $Y_1, \ldots, Y_m \subseteq Dih_p$, we can order E as follows:

Proposition 2.3 (See [7] Proposition 4.1)**.** *Let p be large enough. Consider $\delta = (z_0, 0) \in Dih_p$ such that $z_0 > 0$ and $Y_1, \ldots, Y_m$ subsets of Dih_p. Given a finite subset E such that $\pi_1(E) \subseteq \left(-\frac{p}{90(|E|+1)}, \frac{p}{90(|E|+1)}\right) \subseteq Dih_p$ and $E \not\subseteq \{(x, 0) : x \in \mathbb{F}_p\}$, consider P, N, S_e and S_o defined as above. Then there are orderings $\mathbf{p}$ of P, $\mathbf{n}$ of N, $\mathbf{s}_e = (s_0, s_2, \ldots, s_{2\ell})$ of S_e (here $\ell = |S_e| - 1$), and $\mathbf{s}_o = (s_1, s_3, \ldots, s_{2h-1})$ of S_o (here $h = |S_o|$) such that, if we set $\mathbf{x} = \mathbf{p}\ s_0\ \mathbf{n}\ s_1 s_2 \ldots, s_{\ell+h}$ then $\{\delta, \mathbf{x}\}$ is sequenceable and we have*

$$|(\delta \cdot \mathrm{IS}(\mathbf{x})) \cap Y_j| \leqslant 4 \inf_{L \in \mathbb{N}} \left(\frac{|Y_j|}{L} + L + 2\sum_{i=1}^{j-1} |Y_i|\right) \tag{2}$$

for all $1 \leqslant j \leqslant m$.

Here the technical Condition (2) will be crucial in the following computation, once we have chosen suitable Y_js.

Now the proof proceeds by ordering the dissociated sets, with a probabilistic kind of method (inspired by the procedure of [4]). For each D_j with $j \in L_1$ (those whose projection over $\mathbb{F}_2$ components is constantly one), we consider its partition into D_j^o and D_j^e provided by the Structure Theorem 2.2. Here we have that $|D_j^o| = |D_j^e| = \frac{1}{2}|D_j|$. Then due to

Theorem 2.2 (ii) we can randomly split D_j^o into four sets $D_j^{o,(1)} \cup D_j^{o,(2)} \cup D_j^{o,(3)} \cup D_j^{o,(4)}$ and D_j^e into $D_j^{e,(1)} \cup D_j^{e,(2)} \cup D_j^{e,(3)} \cup D_j^{e,(4)}$ where each partition has the same cardinality. Here we set $D_j^{(i)} = D_j^{o,(i)} \cup D_j^{e,(i)}$. For the other dissociated sets, we operate analogously to Lemma 5.1 of [4], i.e. we partition $D_j = \cup_{i=1}^4 D_j^{(i)}$ into four sets of equal size. Then we order these new dissociated sets as follows:

$$D_1^{(1)}, D_1^{(2)}, D_2^{(1)}, D_2^{(2)}, \dots, D_s^{(1)}, D_s^{(2)}, D_1^{(3)}, D_1^{(4)}, D_2^{(3)}, D_2^{(4)}, \dots, D_s^{(3)}, D_s^{(4)}\,. \tag{3}$$

For notational convenience, we rename this new sequence as $T_1, T_2, \dots, T_u$ where $u = 4s$, and we use the same notation T_j^o and T_j^e as used on the D_js.

Now we want to find a sequencing $\mathbf{t}_j$ of the dissociated set T_j such that:

$$\text{If } T_j = T_j^o \cup T_j^e \text{ (or, } \pi_2(T_j) = 1\text{), choose } t_i \in T_j^o \text{ if } i \text{ is odd and } t_i \in T_j^e \text{ if } i \text{ is even.} \tag{4}$$

Here we first prove (see [7], Lemma 5.2) that it is possible to choose the sequence $T_1, \dots, T_u$ so that, defining τ_j as the product of the terms of T_j (which does not depend on the ordering of the elements of T_j), $\tau_1, \dots, \tau_u, \mathbf{x}$ is a sequencing. Then, defining suitable sets Y_j's (see [7], Equation (10)), and exploiting Condition (2) and Property (v) of the Structure Theorem, we prove probabilistically that it is also possible to order each T_j so that the following theorem holds. The following Theorem uses the union bound on the product of terms of the T_js to derive the bound on $|A|$ given in Theorem 1.3.

Theorem 2.4 (See [7] Theorem 5.5)**.** *Let p be large enough and let $c > 0$ and $1 \leq s \leq e^{c(\log p)^{1/4}}$. Consider $D_1, D_2, \dots, D_s \subseteq Dih_p$, dissociated sets, of size $\Theta(R)$ a multiple of eight.*

Defining δ as in Theorem 2.2(v), we assume that D_1, D_s and δ satisfy Property (v) of that theorem. Consider a sequence $\mathbf{x}$ over Dih_p of length at most $e^{c(\log p)^{1/4}}$ such that $\delta, \mathbf{x}$ is a sequencing that satisfies Condition (2) with respect to the sets $Y_j s$ given in [7], Equation (10).

Then it is possible to choose the sequence $T_1, \dots, T_u$ of dissociated sets and the orderings $\mathbf{t}_1, \dots, \mathbf{t}_u$ that satisfy property 4 so that each $\mathbf{t}_i$ is a sequencing and so is $\mathbf{t}_1, \dots, \mathbf{t}_u, \mathbf{x}$.

Given a set $A \subseteq Dih_p \setminus \{id\}$, the statement preceding Lemma 2.1 allows us to assume that up to automorphism of Dih_p, $A = E \cup (\cup_{j=1}^s D_j)$, where E and each D_j satisfy the hypotheses of Theorem 2.2. We can assume that $s > 0$ and, up to changing all the signs of the first (ie. $\mathbb{F}_p$) components, we can assume that $z_0 > 0$ for the $\delta = (z_0, 0)$ defined in Theorem 2.2(v). Here we apply Proposition 2.3 to obtain an ordering $\mathbf{x}$ of E such that $\delta, \mathbf{x}$ is a sequencing satisfying Condition (2) with respect to the sets Y_i mentioned in the paragraph preceding Theorem 2.4. Now Theorem 1.3 follows because, if we assume that $|A| \leqslant e^{c(\log p)^{1/4}}$, Theorem 2.4 shows the existence of a sequencing of A.

Acknowledgements

The first and second authors were partially supported by INdAM–GNSAGA. The authors would also like to thank Noah Kravitz for useful discussions on this topic.

References

[1] B. Alspach, D. L. Kreher, and A. Pastine. The Friedlander–Gordon–Miller conjecture is true, Australas. J. Combin. 67 (2017), 11–24.

[2] B. Alspach and G. Liversidge, On strongly sequenceable abelian groups, *Art Discrete Appl. Math.*, 3 (2020), 19pp.

[3] D. S. Archdeacon, J. H. Dinitz, A. Mattern, and D. R. Stinson, On partial sums in cyclic groups, *J. Combin. Math. Combin. Comput.*, 98 (2016), 327–342.

[4] B. Bedert and N. Kravitz, Graham's rearrangement conjecture beyond the rectification barrier, `arXiv:2409.07403`.

[5] M. Bucić, B. Frederickson, A. Müyesser, A. Pokrovskiy and L. Yepremyan, Towards Graham's rearrangement conjecture via rainbow paths, `arXiv:2503.01825`.

[6] S. Costa and S. Della Fiore, Weak Freiman isomorphisms and sequencings of small sets, `arXiv:2407.15785`

[7] S. Costa, S. Della Fiore and E. Engel, Graham's rearrangement for dihedral groups, `arXiv:2503.18101`.

[8] S. Costa, S. Della Fiore, M. A. Ollis and S. Z. Rovner-Frydman, On Sequences in Cyclic Groups with Distinct Partial Sums, *Electron. J. Combin.*, 29 (2022), #P3.33.

[9] S. Costa, F. Morini, A. Pasotti and M. A. Pellegrini, A problem on partial sums in abelian groups, *Discrete Math.*, 341 (2018), 705–712.

[10] P. Erdős and R. L. Graham, Old and new problems and results in combinatorial number theory. L'Enseignement mathematique (1980), Universite de Geneve.

[11] R. L. Graham, On sums of integers taken from a fixed sequence, J. H. Jordan, W. A. Webb (eds.), Proceedings of the Washington State University Conference on Number Theory, 1971, pp. 22–40.

[12] N. Kravitz, Rearranging small sets for distinct partial sums, *Integers: Electronic Journal of Combinatorial Number Theory*, 24 (2024).

[13] M. A. Ollis, Sequenceable groups and related topics, *Electron. J. Combin.*, DS10 (2002, updated 2013), 34pp.

[14] M. A. Ollis, Sequences in dihedral groups with distinct partial products, *Australas. J. Combin.*, 78 (2020), 35–60.

[15] A. Pasotti and J. H. Dinitz, A survey of Heffter arrays. *Fields Inst. Commun.*, 86 (2024), 353–392.

[16] W. Sawin, comment on the post "Ordering subsets of the cyclic group to give distinct partial sums". MathOverflow (2015), https://mathoverflow.net/q/202857.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Fractional Haemers' bound and the Mycielski construction

(Extended abstract)

Bence Csonka*

Abstract

Let G be a graph, and let $M(G), M_r(G)$ denote the original and the generalized Mycielski construction of G. It is known from previous works that an exact formula is valid for the fractional chromatic number $\chi_f(M(G))$ and more general also for $\chi_f(M_r(G))$ in terms of the parameter $\chi_f(G)$. Furthermore, a similar formula has been derived for the complementary Lovász theta number $\bar{\vartheta}(M(G))$ in terms of $\bar{\vartheta}(G)$. In this paper, we provide an exact formula for the fractional Haemers' bound $\bar{\mathcal{H}}_f(M_r(G)); \mathbb{F})$ for any perfect graph G, and we also present an upper bound for any graph G.

The talk is based on the forthcoming work [3].

1 Introduction

The investigation of parameters that provide upper bounds on Shannon capacity [10] remains an active area in information theory and combinatorics. A special class of these parameters forms the so-called asymptotic spectrum. Notably, Zuiddam [12] demonstrated that for every graph, there exists an element of the asymptotic spectrum that provides the exact value for the Shannon capacity.

Some known elements of the asymptotic spectrum include the fractional chromatic number, the Lovász theta number, and the fractional Haemers' bound. Lovász [8] by defining the Lovász theta number, showed that the Shannon capacity of the cycle of length 5 is $\sqrt{5}$. In the same paper, the question was raised whether $\vartheta(G) = C(G)$ holds for all graphs, where $C(G)$ and $\vartheta(G)$ denote the Shannon capacity and the Lovász theta number of a graph. Haemers [5],[6] provided a negative answer to this question defining the Haemers' bound $\mathcal{H}(G, \mathbb{F})$ that depends on a graph G and a field $\mathbb{F}$. In particular, he showed that for the Haemers' bound and for the complement of the Schläfli graph $\bar{S}$, it holds that $\mathcal{H}(\bar{S}; \mathbb{R}) < \vartheta(\bar{S})$.

Blasiek [1] introduced a generalization of the Haemers' bound, called that was further investigated by Bukh and Cox [2]. The following version of the definition of this parameter can be found in [2].

*Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, E-mail: csonkab@edu.bme.hu

Definition 1. *Let $\mathbb{F}$ be a field, and let G be a graph. A collection $\{X_v \subseteq \mathbb{F}^n : v \in V(G)\}$ is called a dual (n,d)-representation of the graph G over $\mathbb{F}$ if, for any vertex $v \in V(G)$, the following conditions hold:*

1. *X_v is a subspace of $\mathbb{F}^n$;*
2. *$\dim X_v = d$;*
3. *$X_v \cap \sum_{w\in N(v)} X_w = \{0\}$, where the sum denotes the subspace generated by them, and $N(v)$ denotes the neighborhood of vertex v.*

The complementary fractional Haemers' bound is defined as follows:

$$\bar{\mathcal{H}}_f(G;\mathbb{F}) := \inf_{n,d}\left\{\frac{n}{d} : G \text{ has a dual } (n,d)\text{-representation over } \mathbb{F}\right\}.$$

Definition 2 ([9])**.** *Let G be a simple graph and $r \in \mathbb{Z}^+$. Its generalized Mycielskian $M_r(G)$ is defined on the vertices*

$$V(M_r(G)) = V(G) \times \{0,1,...,r-1\} \cup \{z\}$$

with edge set

$$E(M_r(G)) = \{\{(v,0),(w,0)\} : \{v,w\} \in E(G)\} \cup \\ \cup \{\{(v,i),(w,j)\} : \{v,w\} \in E(G) \text{ and } |i-j| = 1\} \cup \{\{(v,r-1),z\} : v \in V(G)\}.$$

Denote $M_2(G)$ simply by $M(G)$.

We are interested in studying the behavior of $\bar{\mathcal{H}}_f$ with respect to the generalized Mycielskian.

2 Preliminaries

Larsen, Propp, and Ullman [7] were the first to establish a formula relating the Mycielski construction to a parameter in the asymptotic spectrum. They showed that for the fractional chromatic number χ_f,

$$\chi_f(M(G)) = \chi_f(G) + \frac{1}{\chi_f(G)}.$$

Later, this result was generalized by Tardif [11] for $M_r(G)$:

$$\chi_f(M_r(G)) = \chi_f(G) + \frac{1}{\sum_{k=0}^{r-1}(\chi_f(G)-1)^k}.$$

In [4] it has been shown that a similar formula can be given for the complementary Lovász theta number. Their result revealed that

$$\bar{\vartheta}\left(M(G)\right) = \frac{4}{3}\bar{\vartheta}(G)\cos\left(\frac{1}{3}\arccos\left(1 - \frac{27}{4\bar{\vartheta}(G)} + \frac{27}{4\bar{\vartheta}^2(G)}\right)\right) - \frac{1}{3}\bar{\vartheta}(G) + 1,$$

where $\bar{\vartheta}$ denotes the complementary Lovász theta number.

It is worth noting that the fractional chromatic number and the complementary Lovász theta number have a duality definition; that is, both χ_f and $\bar{\vartheta}$ have equivalent minimum and maximum formulations, which are used in the proofs of the above mentioned theorems. However, the existence of a dual formulation for $\bar{\mathcal{H}}_f$ is still unknown.

Understanding how various graph operations affect elements of the asymptotic spectrum is a key step toward identifying or constructing parameters that approximate or attain the Shannon capacity. In this regard, studying the effect of the Mycielski construction on fractional Haemers' bound contributes to a broader program in information theory and combinatorics.

3 Main results

Based on the work of Bukh and Cox, we have shown that for the generalized Mycielski construction of complete graphs K_m, the complementary fractional Haemers' bound behaves similarly to the fractional chromatic number.

Theorem 1. *For any field $\mathbb{F}$ and $m \in \mathbb{Z}^+$,*

$$\bar{\mathcal{H}}_f(M_r(K_m); \mathbb{F}) = m + \frac{1}{\sum_{k=0}^{r-1}(m-1)^k}.$$

It follows from the theorem that if G satisfies $\bar{\mathcal{H}}_f(G; \mathbb{F}) = \omega(G)$ (for example, if G is a perfect graph), then $\bar{\mathcal{H}}_f(M_r(G); \mathbb{F})$ can be explicitly expressed as a function of $\omega(G)$.

This leads to the question of whether Tardif's formula holds in a more general sense. We have succeeded in proving that the formula provides an upper bound.

Theorem 2. *For any simple graph G and any field $\mathbb{F}$, the following bound holds:*

$$\bar{\mathcal{H}}_f(M_r(G); \mathbb{F}) \leq \bar{\mathcal{H}}_f(G; \mathbb{F}) + \frac{1}{\sum_{k=0}^{r-1}(\bar{\mathcal{H}}_f(G; \mathbb{F}) - 1)^k}.$$

According to Theorem 2, for the Schläfli graph S, we have $\mathcal{H}_f(M(S); \mathbb{R}) \leq 7 + \frac{1}{7}$ since Haemers proved that $\bar{\mathcal{H}}(S; \mathbb{R}) \leq 7$, hence $\bar{\mathcal{H}}_f(S; \mathbb{R}) \leq 7$.

This raises the following natural question: Is it true that for every graph G and every field $\mathbb{F}$, there is a function $g : \mathbb{R} \to \mathbb{R}$ such that $\bar{\mathcal{H}}_f(M_r(G); \mathbb{F}) = g(\bar{\mathcal{H}}_f(G; \mathbb{F}))$? In particular, does Tardif's formula hold more generally for the complementary fractional Haemers' bound?

Acknowledgment

The author would like to thank his supervisor Gábor Simonyi for useful discussions and guidelines.

References

[1] Anna Blasiak, A graph-theoretic approach to network coding. PhD thesis, *Cornell University*, (2013).

[2] Boris Bukh, Cristopher Cox, On a Fractional Version of Haemers' Bound, *IEEE Trans. Inform. Theory*, **65** (2019), 3340 – 3348.

[3] Bence Csonka, The Fractional Haemers Bound of the Mycielski Construction, paper in preparation, (2025).

[4] Bence Csonka, Gábor Simonyi, Shannon Capacity, Lovász Theta Number and the Mycielski Construction, *IEEE Trans. Inform. Theory*, **70** (2024), 7632 – 7646.

[5] Willem Haemers, On some problems of Lovász concerning the Shannon capacity of a graph, *IEEE Trans. Inform. Theory*, **25** (1979), 231–232.

[6] Willem Haemers, An upper bound for the Shannon capacity of a graph, *Algebraic methods in graph theory*, **25** Vol. I (1978), 267—272.

[7] Michael Larsen, James Propp, Daniel Ullman, The fractional chromatic number of Mycielski graphs, *J. Graph Theory* **19** (1995), 411–416.

[8] László Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* , **25** (1979), 1–7.

[9] Jan Mycielski, Sur le coloriage des graphs, *Colloq. Math.*, **3** (1955), 161–162.

[10] Claude E. Shannon, The zero-capacity of a noisy channel, *IRE Trans. Inform. Theory*, **2** (1956), 8–19.

[11] Claude Tardif, Fractional chromatic numbers of cones over graphs, *J. Graph Theory* **38** (2001), 87–94.

[12] Jeroen Zuiddam, The Asymptotic Spectrum of Graphs and the Shannon Capacity, *Combinatorica*, **39** (2019), 1173—1184.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Beyond the fractional Reed bound for triangle-free graphs

(Extended abstract)

Tianjiao Dai* Qiancheng Ouyang† François Pirot‡
Jean-Sébastien Sereni§

Abstract

Given a graph H, we let $\chi_f(d, H)$ be the supremum of the fractional chromatic numbers over all H-free graphs of maximum degree at most d. We focus on the case $H = K_3$. It has been settled by Dvořák, Sereni, and Volec that $\chi_f(3, K_3) = 2.8$, and the next open case is $d = 4$. In 2002, Molloy and Reed proved that $\chi_f(d, K_3) \leq \frac{d+3}{2}$ for every integer d, which implies that $\chi_f(4, K_3) \leq 3.5$. However, it is conjectured that $\chi_f(4, K_3) = 3.25$. In this paper, we prove that $\chi_f(4, K_3) < 3.4663$, by relying on a methodology introduced by Pirot and Sereni that lets us use a probability distribution over the independent sets of a graph G, and all of its induced subgraphs, in order to obtain a fractional colouring of G.

1 Introduction

Given a graph H, we let $\chi_f(d, H)$ be the supremum of the fractional chromatic numbers over all H-free graphs of maximum degree at most d. When H is a complete graph,

*Department of Mathematics, Shandong University, Jinan, China. E-mail: `tianjiao.dai@sdu.edu.cn`. Supported by Natural Science Foundation of Shandong Province(ZR2024QA174).

†Data Science Institute, Shandong University, Jinan, China. E-mail: `oyqc@sdu.edu.cn`. Supported by Natural Science Foundation of Shandong Province(ZR2024QA107).

‡LISN (Université Paris-Saclay), Gif sur Yvette, France. E-mail: `francois.pirot@upsaclay.fr`. Supported by Natural Science Foundation of Shandong Province().

§Service public français de la recherche, Centre National de la Recherche Scientifique (ICube, CSTB), Strasbourg, France. E-mail: `jean-sebastien.sereni@cnrs.fr`.

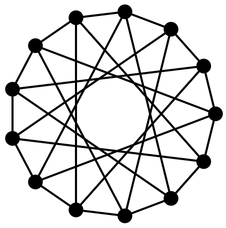


Figure 1: A graph certifying that $\chi_f(4, K_3) \geq 3.25$.

the study of $\chi_f(d, H)$ falls in the domain of Ramsey theory, a domain which emerged in the 1930s following seminal results by van der Waerden [8] and by Ramsey [7], and has attracted a lot of attention ever since. An important result in this case is due to Molloy and Reed [5, Theorem 21.7, p. 244]: known as "the fractional Reed bound", it states that $\chi_f(d, K_n) \leq \frac{d+n}{2}$ for all integers $d, n \geq 2$.

In this paper, we focus on the case $H = K_3$, which is closely related to *off-diagonal Ramsey numbers*. It has been established [3] that $\chi_f(3, K_3) = 14/5$. The same question for larger values of the maximum degree is still open. At one end of the spectrum, we know that $\chi_f(4, K_3)$ lies between 3.25 (see Figure 1) and 3.5 (by the fractional Reed bound). At the other end of the spectrum, one has $\chi_f(d, K_3) \leq (1 + o(1))\, d/\ln d$ as $d \to \infty$, which is a consequence of a result by Molloy [4], and one can infer from a study of random d-regular graphs by Bollobás [1] that $\chi_f(d, K_3) \geq \frac{d}{2 \ln d}$.

A first study of $\chi_f(d, K_3)$ has been made by Pirot and Sereni [6]. They proved the following.

Theorem 1 (Pirot & Sereni, 2021)**.** *For every integer d,*

$$\chi_f(d, K_3) \leq 1 + \min_{k \in \mathbb{N}} \inf_{\lambda > 0} \frac{(1+\lambda)^k + \lambda(1+\lambda)d}{\lambda(1 + k\lambda)}.$$

The upper bound in Theorem 1 which can be effectively computed from this formula improves on the fractional Reed bound as soon as $d \geq 17$. We note that the upper bound $(d + 3)/2$ is a special case of the bound given by Theorem 1 by fixing $(k, \lambda) = (2, +\infty)$, and this is best possible when $d \leq 16$.

2 Our result

Using the same method with a more involved approach, we obtain the following result, in which we investigate $\chi_f(4, K_3)$ and, for the first time in over 20 years, establish an improvement on its previously known bound.

Theorem 2. $\chi_f(4, K_3) < 3.4663$.

2.1 Intuition of our method

The method relies on a Greedy Fractional Colouring Algorithm (GFCA) — which we describe in the next section — that relies on some probability distribution over the independent sets of any given induced subgraph H of G. When this probability distribution has a relatively uniform coverage of the closed neighbourhoods in H — this is referred to as (α, β)-*local occupancy* in works using the GFCA — one can deduce an upper bound on the fractional chromatic number of G, as stated in Lemma 1.

In previous works using the GFCA, the input probability distribution has always been the hard-core distribution on some family of independent sets of H. In this paper, we construct our random independent sets in two steps, where during the second step we lie within the subgraph of H induced by the vertices uncovered by the random independent set constructed so far. The first step is the hard-core distribution with fugacity λ (the value of λ is made explicit at the very end of the method through some numerical optimisation process), and the last step is given by a specific fractional colouring with local demand (see Corollary 1), such that vertices with smaller degree receive a larger proportion of the colours.

3 Prerequisites

In this section we first introduce the notions that will be needed to derive our results.

3.1 Notations

Let G be a given graph. If J is a subset of vertices of G, then we write $N_G(J)$ for the set of vertices that are not in J and have a neighbour in J, while $N_G[J]$ is $N_G(J) \cup J$. We omit the graph subscript when there is no ambiguity, and we write $N(v)$ for $N(\{v\})$. We denote by $\mathscr{I}(G)$ the set of all independent sets of G. If w is a mapping from $\mathscr{I}(G)$ to $\mathbb{R}$, then for every vertex $v \in V(G)$ we set

$$w[v] \coloneqq \sum_{\substack{I \in \mathscr{I}(G) \\ v \in I}} w(I).$$

Further, if $\mathscr{I}$ is a collection of independent sets of G, then $w(\mathscr{I}) \coloneqq \sum_{I \in \mathscr{I}} w(I)$. If I is an independent set of a graph G, a vertex v is *covered* by I if v belongs to I or has a neighbour in I. A vertex that is not covered by I is *uncovered* (by I).

3.2 Greedy fractional colouring algorithm

Our results on fractional colouring are obtained using a greedy algorithm analysed in [2]. This algorithm is a generalisation of an algorithm first described in the book of Molloy and Reed [5, p. 245] for the uniform distribution over maximum independent sets. The setting here is, for each induced subgraph H of the graph we wish to fractionally colour, a probability distribution over the independent sets of H.

Lemma 1 (Davies *et al.*, 2020)**.** *Let G be a graph given with couples (α_v, β_v) for every vertex $v \in V(G)$. For every induced subgraph H of G, let $\mathbf{I}_H$ be a random independent set of H drawn according to a given probability distribution, and assume that*

$$\alpha_v \mathbb{P}[v \in \mathbf{I}_H] + \beta_v \mathbb{E}[|\mathbf{I}_H \cap N(v)|] \geq 1,$$

for every $v \in V(H)$. Then the GFCA produces a fractional colouring w of G which certifies that

$$\chi_f(G) \leq \max_{v \in V(G)} \alpha_v + \beta_v \deg(v).$$

Algorithm 1: The Greedy Fractional Colouring Algorithm (GFCA)

---

for $I \in \mathscr{I}(G)$ **do**
  $w(I) \leftarrow 0$
$H \leftarrow G$
while $|V(H)| > 0$ **do**
  $\iota \leftarrow \min\limits_{v \in V(H)} \dfrac{1 - w[v]}{\mathbb{P}[v \in \mathbf{I}_H]}$
  for $I \in \mathscr{I}(H)$ **do**
    $w(I) \leftarrow w(I) + \mathbb{P}[\mathbf{I}_H = I]\, \iota$
  $H \leftarrow H - \{v \in V(H) : w[v] = 1\}$

This algorithm takes as input a graph G and a probability distribution π on the independent sets of any induced subgraph of G, and returns a fractional colouring of G whose weight is bounded by a function of π and G. We note that in Lemma 1, although there is one probability distribution on each induced subgraph, the couple (α_v, β_v) associated with each vertex is fixed once and for all, which somewhat ties together the different probability distributions involved.

To ensure optimality, our task is to solve the following linear program.

Definition 1. Let G be a graph, and let $\boldsymbol{\varphi}$ map each (induced) subgraph $H \subseteq G$ to a random independent set $\boldsymbol{\varphi}(H) \in \mathscr{I}(H)$. We let $\mathrm{LP}_{\boldsymbol{\varphi}}(G)$, which we call the *linear program associated with* $(G, \boldsymbol{\varphi})$, be defined as follows.

$$\mathrm{LP}_{\boldsymbol{\varphi}}(G)\colon \begin{cases} \text{Minimise} & \alpha + \beta\Delta \\ \text{such that} & \alpha\, \mathbb{P}[v \in \boldsymbol{\varphi}(H)] + \beta\, \mathbb{E}[|\boldsymbol{\varphi}(H) \cap N(v)|] \geq 1, \quad \text{for all } H \subseteq G, v \in V(H); \\ & \alpha, \beta \geq 0. \end{cases}$$

So we need to compute the constraints of the linear program described above. To that end, we define the following.

Definition 2. Let H be a graph and $\mathbf{I}$ be a random independent set of H. For every vertex $v \in V(H)$, we let

$$e_{\mathbf{I}}(v) := \begin{pmatrix} \mathbb{P}\left[v \in \mathbf{I}\right] \\ \mathbb{E}\left[|N(v) \cap \mathbf{I}|\right] \end{pmatrix}$$

be the *constraint* of $\mathbf{I}$ on v.

With the above definition in hand, the constraints of the linear program are all of the form $(\alpha, \beta) \cdot e_{\boldsymbol{\varphi}(H)}(v) \geq 1$, for every $H \subseteq G$ and $v \in V(H)$.

3.3 Hard-core model

The probability distribution that we are going to use as a setting of Lemma 1 uses the hard-core distribution over the independent sets of a graph, which has the Spatial Markov Property. Given a family $\mathscr{I}$ of independent sets of a graph H, and a positive real λ, a random independent set $\mathbf{I}$ drawn according to the hard-core distribution at fugacity λ over $\mathscr{I}$ is such that

$$\mathbb{P}\left[\mathbf{I} = I\right] = \frac{\lambda^{|I|}}{Z_{\mathscr{I}}(\lambda)},$$

for every $I \in \mathscr{I}$, where $Z_{\mathscr{I}}(\lambda) = \sum_{J \in \mathscr{I}} \lambda^{|J|}$ is the *partition function* associated with $\mathbf{I}$.

Among the many consequences of the Spatial Markov Property, one can observe that for every vertex $v \in V(H)$,

$$\mathbb{P}\left[v \in \mathbf{I}\right] = \lambda\, \mathbb{P}\left[v \text{ is uncovered by } \mathbf{I}\right]. \tag{1}$$

4 Outline of the proof

In this section, we focus on optimising the upper bound of $\chi_f(4, K_3)$. To that end, we use a 2-step procedure to construct the random independent sets that feed the GFCA. The first step follows the hard-core distribution, and the second step consists of a fractional colouring with local demand of the uncovered vertices. Before describing the procedure in more detail, we introduce the necessary terminology.

4.1 Setting up the finishing step

Let G be a graph. Given a function $f\colon V(G) \to \mathbb{Q}^+$, an f-fractional colouring of G is a random independent set $\mathbf{I}$ of G such that $\mathbb{P}\left[v \in \mathbf{I}\right] \geq f(v)$ for every vertex $v \in V(G)$. An f-fractional colouring of G is also called a *fractional colouring of G with local demand f*. Given a vertex v, the value of $f(v)$ is called the *demand of v*.

We say that a subgraph H of a given subcubic graph G is *dangerous* if it is isomorphic to C_5 or to K_4^+ (the complete graph on 4 vertices where two non-adjacent edges have been subdivided twice each). A vertex is *dangerous* if it has degree 2 in a dangerous graph. A vertex $v \in V(G)$ is *special* if it belongs to a dangerous subgraph of G, and has degree 2 in

G. Given a subset B of the dangerous vertices of G, we say that a vertex v of degree 2 in a dangerous subgraph H of G is *nailed by* B if either v belongs to B, or v is not special (i.e has degree 3 in G). A *nail* B of G is a subset of the dangerous vertices of G such that for every dangerous subgraph H of G, at least 2 vertices of degree 2 in H are nailed by B.

In order to analyse the final step of our procedure, we use the following result [3].

Theorem 3 (Dvořák, Sereni, Volec; 2014)**.** *Let G be a triangle-free subcubic graph, and let B be a nail of G. For every vertex $v \in B$, we let $f_B(v) = \frac{7-\deg(v)}{14}$, and for every vertex $v \notin B$, we let $f_B(v) = \frac{8-\deg(v)}{14}$. Then G has an f_B-fractional colouring.*

We can show that Theorem 3 has the following result as a corollary.

Corollary 1. *Let G be a triangle-free graph, and let*

$$f_G(v) = \begin{cases} 1 & \text{if } \deg(v) = 0; \\ 1/2 & \text{if } v \text{ belongs to an isolated edge,} \\ 1 - f_G(u) \geq 4/7 & \text{otherwise if } N(v) = \{u\}; \\ 11/28 & \text{if } v \text{ is special and has a special neighbour,} \\ 3/7 & \text{otherwise if } \deg(v) = 2; \\ 5/14 & \text{if } \deg(v) = 3; \\ 0 & \text{if } \deg(v) \geq 4 \end{cases}$$

for every vertex $v \in V(G)$. Then G has an f_G-fractional colouring.

4.2 A description of the probability distribution

We are now ready to describe our fractional colouring procedure for a given triangle-free graph G of maximum degree 4. We will use the GFCA with a probability distribution over the independent sets of any induced subgraph H of G obtained as follows.

(i) We fix some real value $\lambda > 0$ (which will be optimised later on). We let $\mathbf{I}_0$ be drawn according to the hard-core distribution at fugacity λ over the independent sets of H. Let us write this $\mathbf{I}_0 \leftarrow \mathrm{hc}_\lambda(\mathscr{I}(H))$.

(ii) We let $\mathbf{H} \coloneqq H \setminus N[\mathbf{I}_0]$ be the subgraph of H induced by the vertices uncovered by $\mathbf{I}_0$. We apply Corollary 1 to $\mathbf{H}$ in order to obtain an $f_{\mathbf{H}}$-fractional colouring $\mathbf{I}_1$ of $\mathbf{H}$.

(iii) We return the random independent set $\mathbf{I}_0 \cup \mathbf{I}_1$.

It turns out that this procedure is not enough to guarantee that $\chi_f(G) < 3.5$. In order to fall below that threshold, we need to alter our last step and give a non-zero demand $f(v)$ for some vertices v of degree 4, namely the ones which have at least one degree-4 neighbour.

4.3 Giving non-zero demand to (some) degree-4 vertices

Let H be any subgraph of G (which we recall is a triangle-free graph of maximum degree 4). Let $S_{\leq 3} \subseteq V(H)$ be the set of vertices of degree at most 3, and let $S_4 = V(H) \setminus S_{\leq 3}$ be the set of vertices of degree 4. We further partition S_4 into $S_4^0 \cup S_4^+$, where S_4^0 is the set of isolated vertices in $G[S_4]$ (and hence each vertex in S_4^+ has a neighbour in S_4).

Claim 1. *There exists a random set $\mathbf{X} \subseteq S_4$ such that $H \setminus \mathbf{X}$ is deterministically subcubic, and $\mathbb{P}[v \in \mathbf{X}] = 1/2$ for every $v \in S_4^+$.*

Now the procedure of constructing a random independent set of an induced subgraph $H \subseteq G$ by using GFCA will be refreshed as follows.

1. We first draw $\mathbf{I}_0 \leftarrow \mathrm{hc}_\lambda(\mathscr{I}(H))$.
2. We let $\mathbf{H} \coloneqq H \setminus N[\mathbf{I}_0]$, and let $\mathbf{H}' \coloneqq H \setminus \mathbf{X}$ where $\mathbf{X}$ is the random set promised by Claim 1.
3. Let $\mathbf{I}_1$ be an $f_{H'}$-fractional colouring of H'.
4. We return the random independent set $\mathbf{I}_0 \cup \mathbf{I}_1$.

4.4 How to generate the constraints of the linear program

As mentioned earlier, our tasks are to compute all the constraints of vertices in G and to solve the linear program. Notice that if we generate the whole set of constraints $\{e_{\boldsymbol{\varphi}(H)}(v) : H \subseteq G, v \in V(H)\}$, we may end up with a huge number of constraints. It is however possible to considerably reduce the number of constraints by restricting to a specific subset. So we introduce a partial ordering on the set of constraints. In the end, we keep only minimal constraints with respect to that order, and we can argue that this does not affect the result of the linear program.

Given an induced subgraph $H \subseteq G$, let $\mathbf{I}_0$ be drawn by the hard-core distribution at fugacity λ. The random independent set drawn by the hard-core model is usually far from being maximal, in particular with small values of λ. In particular, given a vertex v, there is a non-negligible probability that $\mathbf{I}_0$ does not intersect $N[v]$. In that case, v has a second chance to be drawn during the second step, in $\mathbf{I}_1$. Another interesting case is when $\mathbf{I}_0$ intersects $N(v)$, yet some neighbours of v remain uncovered by $\mathbf{I}_0$. Those will have a second chance to be drawn during the second step, and to take that into account we introduce the notion of "refined constraint" (denoted by $\mathcal{E}^+_{\boldsymbol{\varphi}}(d)$) that takes this into account in a third coordinate. We show how to rely on this set of refined constraints in order to compute a random independent set constructed in two steps as follows.

Lemma 2. *Let G be a triangle-free graph. Let $\mathbf{I}$ be a random independent set of G constructed in two steps, i.e. $\mathbf{I} = \mathbf{I}_0 \cup \boldsymbol{\varphi}(G \setminus N[\mathbf{I}_0])$, where $\mathbf{I}_0$ is drawn from the hard-code distribution at fugacity $\lambda > 0$ from $\mathscr{I}(G)$, and $\boldsymbol{\varphi}$ maps any subgraph $H \subseteq G$ to a random* *independent set of H. Then each constraint $e_{\mathbf{I}}(v)$ is looser than some convex combination of the vectors*

$$\left\{ \frac{1}{\lambda + (1+\lambda)^d} \begin{pmatrix} \lambda + p \\ (1+\lambda)^{d-1}(d\lambda + r) + q - r \end{pmatrix} \;:\; d \in \{0,1,\ldots,\Delta(G)\} \text{ and } \begin{pmatrix} p \\ q \\ r \end{pmatrix} \in \mathcal{E}^+_{\boldsymbol{\varphi}}(d) \right\}.$$

Notice that $\mathcal{E}^+_{\boldsymbol{\varphi}}(d)$ is a notion of *refined constraint*, which will be helpful to compute constraints generated by a random independent set $\mathbf{I} = \mathbf{I}_0 \cup \mathbf{I}_1$ constructed in two steps, where $\mathbf{I}_0$ follows the hard-core distribution.

We are now ready to prove the main result of this section. Let us fix $\lambda = 0.51$. Given an induced subgraph $H \subseteq G$, we let $\boldsymbol{\varphi}_0(H) \coloneqq \mathbf{I}_0 \cup \boldsymbol{\varphi}'(H \setminus N[\mathbf{I}_0])$, where $\mathbf{I}_0 \leftarrow \mathrm{hc}_\lambda(H)$. We construct the linear program $\mathrm{LP}_{\boldsymbol{\varphi}_0}(G)$ with the set of constraints that we obtain by combining Lemma 2 with the values of refined constraints. The solution to that program is reached when $\alpha = 1.8980861$ and $\beta = 0.39205135$; this implies that $\chi_f(4, K_3) < 3.4663$.

References

[1] B. Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, pages 433–436, 1981.

[2] E. Davies, R. de Joannis de Verclos, R. J. Kang, and F. Pirot. Coloring triangle-free graphs with local list sizes. *Random Structures & Algorithms*, 57(3):730–744, 2020.

[3] Z. Dvořák, J.-S. Sereni, and J. Volec. Subcubic triangle-free graphs have fractional chromatic number at most 14/5. *Journal of the London Mathematical Society*, 89(3):641–662, 2014.

[4] M. Molloy. The list chromatic number of graphs with small clique number. *Journal of Combinatorial Theory, Series B*, 134:264–284, 2019.

[5] M. Molloy, M. S. Molloy, and B. Reed. *Graph colouring and the probabilistic method*, volume 23. Springer Science & Business Media, 2002.

[6] F. Pirot and J.-S. Sereni. Fractional chromatic number, maximum degree, and girth. *SIAM Journal on Discrete Mathematics*, 35(4):2815–2843, 2021.

[7] F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc*, 30, 1930.

[8] B. L. van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wiskd., II. Ser.*, 15:212–216, 1927.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Maximum index of connected graphs

(Extended abstract)

Ivan Damnjanović*†

Abstract

The index of a graph is the largest eigenvalue of its adjacency matrix. We solve the index maximization problem on the set of connected simple graphs of order n and size $n-1+e$ for the case when $e \in [0, 85]$ or $n \geq \frac{5}{2}e$.

1 Introduction

We take all graphs to be undirected, finite and simple, and for any graph G, we denote its vertex set and adjacency matrix by $V(G)$ and $A(G)$, respectively. The *index*, or *spectral radius*, of a graph G, denoted by $\rho(G)$, is the largest eigenvalue of $A(G)$. Many researchers have found interest in investigating the spectral radii of graphs; see, e.g., [8]. A binary symmetric matrix with zero diagonal $A \in \mathbb{R}^{n\times n}$ is *stepwise* if, for any $i, j \in [n]$ such that $i < j$ and $A_{ij} = 1$, we have $A_{k\ell} = 1$ for every $k, \ell \in [n]$ such that $k < \ell$, $k \leq i$ and $\ell \leq j$. A graph G is a *threshold graph* if its vertices can be ordered in such a way that the corresponding adjacency matrix is stepwise; see [3, 5].

For any $e \in \mathbb{N}_0$, let k_e be the largest positive integer such that $\binom{k_e}{2} \leq e$ and let $t_e := e - \binom{k_e}{2}$. Note that for any $n \in \mathbb{N}$ and $e \in \mathbb{N}_0$, there is a connected graph of order n and size $n-1+e$ if and only if $n \geq b_e$, where $b_0 := 1$ and $b_e := k_e + 1 + [t_e \neq 0]$ for

*Faculty of Electronic Engineering, University of Niš, Niš, Serbia. E-mail: `ivan.damnjanovic@elfak.ni.ac.rs`. Supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, grant number 451-03-137/2025-03/200102, and the Science Fund of the Republic of Serbia, grant #6767, Lazy walk counts and spectral radius of threshold graphs — LZWK.

†Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia. Email: `ivan.damnjanovic@famnit.upr.si`.

$e \geq 1$. Let $\mathcal{C}_{n,e}$ be the set of all the connected graphs of order n and size $n-1+e$. For each $e \in \mathbb{N}_0$ and $n \geq b_e$, let $\mathcal{D}_{n,e} \in \mathcal{C}_{n,e}$ be the graph $\mathcal{D}_{n,e} := K_n$ if $n < k_e + 2$, and

$$\mathcal{D}_{n,e} := (((K_{k_e - t_e} + K_1) \vee K_{t_e}) + (n - k_e - 2)K_1) \vee K_1$$

if $n \geq k_e + 2$. Also, for each $e \in \mathbb{N}_0$ and $n \geq e + 2$, let $\mathcal{V}_{n,e} \in \mathcal{C}_{n,e}$ be the graph

$$\mathcal{V}_{n,e} := (K_{1,e} + (n - e - 2)K_1) \vee K_1.$$

The index maximization problem on $\mathcal{C}_{n,e}$ was initiated by Brualdi and Solheid in 1986 and later investigated by Cvetković and Rowlinson in 1988, yielding the next three results.

Theorem 1.1 ([2, Theorem 2.1]). *For some $e \in \mathbb{N}_0$ and $n \geq b_e$, suppose that G attains the maximum index on $\mathcal{C}_{n,e}$. Then G is a threshold graph.*

Theorem 1.2 ([2, Theorems 3.1 and 3.2]). *For any $e \in \{0, 1, 2, 3\}$ and $n \geq b_e$, the graph $\mathcal{D}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

Theorem 1.3 ([2, Theorem 3.3], [4]). *For each $e \geq 4$, there is an $f(e)$ such that for every $n \geq f(e)$, the graph $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

The value $f(e)$ from Theorem 1.3 is not trivial to compute and it is above $e^2(e+2)^2$ when $e \geq 7$. In 1991, Bell solved the extremal problem for the case $t_e = 0$ through the following theorem.

Theorem 1.4 ([1]). *For any $\lambda > 3$, let $f(\lambda) = \frac{1}{2}(\lambda+1)(\lambda+6) + 7 + \frac{32}{\lambda-3} + \frac{16}{(\lambda-3)^2}$. Then for each $e \geq 6$ with $t_e = 0$, the following holds:*

(i) *if $b_e \leq n < f(k_e)$, then $\mathcal{D}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$;*

(ii) *if $n = f(k_e)$, then $\mathcal{D}_{n,e}$ and $\mathcal{V}_{n,e}$ are the graphs attaining the maximum index on $\mathcal{C}_{n,e}$;*

(iii) *if $n > f(k_e)$, then $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

In 2002, Olesky, Roy and van den Driessche [6] expanded these results by showing that $\mathcal{D}_{n,e}$ is the unique solution when $t_e = k_e - 1$ and $b_e + 1 \leq n \leq e - 1$. With all of this in mind, it is natural to expect that for any $e \in \mathbb{N}_0$ and $n \geq b_e$, the solution to the index maximization problem on $\mathcal{C}_{n,e}$ is either only $\mathcal{D}_{n,e}$, only $\mathcal{V}_{n,e}$, or both of these graphs. Here, we investigate this extremal problem and resolve the case when $e \in [0, 85]$ or $n \geq \frac{5}{2}e$.

2 Preliminaries

For any threshold graph G of order n, we assume that $V(G) = \{1, 2, \ldots, n\}$ and that $A(G)$ is a stepwise matrix, with the vertices being arranged accordingly. We take all the polynomials and rational functions to be in the variable λ. For any graph G, we use $\mathcal{P}_G(\lambda)$ and $\rho(G)$ to denote the characteristic polynomial and the index of G, respectively. Also, for any nonzero polynomial $P(\lambda)$, we use $\rho(P)$ and $\rho_2(P)$ to denote the largest and second largest real root of P (with multiplicity). In particular, we have $\rho(G) = \rho(\mathcal{P}_G(\lambda))$ for any graph G. For convenience, we let $\rho(P) = -\infty$ if P is a polynomial without real roots and $\rho_2(P) = -\infty$ if P is a polynomial with at most one (simple) real root.

3 T-subgraphs

For any $e \geq 1$ and $n \geq b_e$, let $G \in \mathcal{C}_{n,e}$ be a threshold graph and let $n-1 = d_1 \geq d_2 \geq d_3 \geq \cdots \geq d_n$ be the degree sequence of G. Now, let s be maximal such that $d_s \geq 2$. Then the *T-subgraph* of G, denoted by $\mathcal{T}(G)$, is the subgraph of G induced by the vertices $\{2, 3, \ldots, s\}$. Also, for the sake of brevity, let $\mathcal{T}_1(G) = \mathcal{T}(G) \vee K_1$.

Lemma 3.1. *For some $e \geq 1$ and $n \geq b_e$, let $G \in \mathcal{C}_{n,e}$ be a threshold graph and let $n' = |\mathcal{T}_1(G)|$. Then $\rho(G)$ is the largest real root of $\lambda \, \mathcal{P}_{\mathcal{T}_1(G)}(\lambda) - (n - n') \, \mathcal{P}_{\mathcal{T}(G)}(\lambda)$.*

In view of Lemma 3.1, we introduce the function $\mathcal{R}_G \colon [\rho(\mathcal{T}_1(G)), +\infty) \to \mathbb{R}$ defined by

$$\mathcal{R}_G(\lambda) = \frac{\lambda \, \mathcal{P}_{\mathcal{T}_1(G)}(\lambda)}{\mathcal{P}_{\mathcal{T}(G)}(\lambda)} \quad (\lambda \geq \rho(\mathcal{T}_1(G))),$$

for any $e \geq 1$, $n \geq b_e$ and threshold graph $G \in \mathcal{C}_{n,e}$. From the theory of nonnegative matrices, we have $\rho(\mathcal{T}_1(G)) > \rho(\mathcal{T}(G))$, hence $\mathcal{R}_G$ is well-defined.

Lemma 3.2. *For any $e \geq 1$, $n \geq b_e$ and threshold graph $G \in \mathcal{C}_{n,e}$, the rational function $\mathcal{R}_G$ is a strictly increasing bijection from $[\rho(\mathcal{T}_1(G)), +\infty)$ onto $[0, +\infty)$.*

The next corollary follows directly from Lemma 3.2.

Corollary 3.3. *Suppose that $e \geq 1$ and $n \geq b_e$, and let $G \in \mathcal{C}_{n,e}$ be a threshold graph. Then for any real $\alpha \geq 0$, the polynomial $Q(\lambda) = \lambda \, \mathcal{P}_{\mathcal{T}_1(G)}(\lambda) - \alpha \, \mathcal{P}_{\mathcal{T}(G)}(\lambda)$ satisfies $\rho(Q) = \mathcal{R}_G^{-1}(\alpha)$. Moreover, we have $Q(\lambda) < 0$ for any $\lambda \in (\rho(\mathcal{T}(G)), \rho(Q))$.*

We introduce the polynomial

$$\begin{aligned} \mathcal{Q}_{G_1,G_2}(\lambda) = \lambda &\left(\mathcal{P}_{\mathcal{T}_1(G_1)}(\lambda) \, \mathcal{P}_{\mathcal{T}(G_2)}(\lambda) - \mathcal{P}_{\mathcal{T}_1(G_2)}(\lambda) \, \mathcal{P}_{\mathcal{T}(G_1)}(\lambda)\right) \\ &+ (|\mathcal{T}(G_1)| - |\mathcal{T}(G_2)|) \, \mathcal{P}_{\mathcal{T}(G_1)}(\lambda) \, \mathcal{P}_{\mathcal{T}(G_2)}(\lambda) \end{aligned}$$

for any $e \geq 1$, $n \geq b_e$ and threshold graphs $G_1, G_2 \in \mathcal{C}_{n,e}$. As it turns out, the auxiliary polynomial $\mathcal{Q}_{G_1,G_2}(\lambda)$ can be used while proving the extremal property of $\mathcal{V}_{n,e}$ and $\mathcal{D}_{n,e}$.

Lemma 3.4. *Let $e \geq 4$, $n \geq e + 2$ and let $G \in \mathcal{C}_{n,e}$ be a threshold graph with $G \ncong \mathcal{V}_{n,e}$. Then $\mathcal{Q}_{G,\mathcal{V}_{n,e}}(\lambda)$ is a nonzero polynomial with a positive leading coefficient. Moreover,*

- ***(i)*** *if $\rho(\mathcal{Q}_{G,\mathcal{V}_{n,e}}) < \rho(\mathcal{T}_1(\mathcal{V}_{n,e}))$, then $\rho(\mathcal{V}_{n,e}) > \rho(G)$;*
- ***(ii)*** *if $\rho(\mathcal{Q}_{G,\mathcal{V}_{n,e}}) \geq \rho(\mathcal{T}_1(\mathcal{V}_{n,e}))$ and $n > e + 2 + \mathcal{R}_{\mathcal{V}_{n,e}}(\rho(\mathcal{Q}_{G,\mathcal{V}_{n,e}}))$, then $\rho(\mathcal{V}_{n,e}) > \rho(G)$.*

Lemma 3.5. *Let $e \geq 4$, $t_e \geq 1$, $n \geq b_e$ and let $G \in \mathcal{C}_{n,e}$ be a threshold graph with $G \ncong \mathcal{D}_{n,e}$. Suppose that $\mathcal{Q}_{G,\mathcal{D}_{n,e}}(\lambda) \not\equiv 0$ and let c be the leading coefficient of $\mathcal{Q}_{G,\mathcal{D}_{n,e}}(\lambda)$. Then,*

- ***(i)*** *if $c > 0$ and $\rho(\mathcal{Q}_{G,\mathcal{D}_{n,e}}) < \rho(\mathcal{T}_1(\mathcal{D}_{n,e}))$, then $\rho(\mathcal{D}_{n,e}) > \rho(G)$;*
- ***(ii)*** *if $c < 0$, $\rho_2(\mathcal{Q}_{G,\mathcal{D}_{n,e}}) < \rho(\mathcal{T}_1(\mathcal{D}_{n,e})) < \rho(\mathcal{Q}_{G,\mathcal{D}_{n,e}})$ and $n < b_e + \mathcal{R}_{\mathcal{D}_{n,e}}(\rho(\mathcal{Q}_{G,\mathcal{D}_{n,e}}))$, then $\rho(\mathcal{D}_{n,e}) > \rho(G)$.*

We can now construct an algorithm as follows. The sole input is the argument $e \geq 4$ such that $t_e \geq 1$, and the algorithm then iterates over all the graphs T that appear as the T-subgraph of a threshold graph from $\mathcal{C}_{n,e}$ for some $n \geq b_e$, except for the T-subgraphs of $\mathcal{D}_{n,e}$ and $\mathcal{V}_{n,e}$. We verify the claim that no threshold graph $G \in \mathcal{C}_{n,e}$ with $\mathcal{T}(G) = T$ satisfies both $\rho(G) \geq \rho(\mathcal{D}_{n,e})$ and $\rho(G) \geq \rho(\mathcal{V}_{n,e})$, for any $n \geq b_e$. This is done by way of contradiction, by computing a lower and upper bound on n based on the claims (1) and (2) from Lemmas 3.4 and 3.5. Through a computer-assisted search, we verify that these bounds are indeed contradictory when $e \in [4, 85]$, leading to the following proposition.

Proposition 3.6. *For some $e \in [4, 85]$ and $n \geq b_e$, suppose that G attains the maximum index on $\mathcal{C}_{n,e}$. Then $G \cong \mathcal{D}_{n,e}$ or ($n \geq e + 2$ and $G \cong \mathcal{V}_{n,e}$).*

4 Main results

We proceed with a precise criterion that compares $\rho(\mathcal{D}_{n,e})$ and $\rho(\mathcal{V}_{n,e})$. To this end, we need the auxiliary polynomial

$$\begin{aligned}\Psi_e(\lambda) = {} & \lambda^3(k_e - 1)(k_e - 2)(k_e^2 - 3k_e + 4t_e) \\ & - \lambda^2(k_e^5 - 6k_e^4 + k_e^3(4t_e + 15) - k_e^2(20t_e + 18) + k_e(8t_e^2 + 24t_e + 8) - (4t_e^2 + 12t_e)) \\ & - \lambda(k_e^2 - k_e + 2t_e)(k_e^3 + k_e^2(t_e - 4) - k_e(3t_e - 5) + (4t_e^2 - 2t_e - 2)) \\ & + t_e(k_e - t_e - 1)(k_e^2 - 3k_e + 2t_e)(k_e^2 - k_e + 2t_e)\end{aligned}$$

defined for every $e \geq 4$. Let ψ_e be the largest real root of $\Psi_e(\lambda)$ and let

$$\omega_e = e + 2 + \frac{\psi_e(\psi_e + 1)(\psi_e^2 - \psi_e - 2e)}{\psi_e^2 - e}.$$

The numbers ψ_e and ω_e are well-defined for any $e \geq 4$, as shown by the next two lemmas.

Lemma 4.1. *For any $e \geq 4$, the polynomial $\Psi_e(\lambda)$ is cubic and has a positive leading coefficient.*

Lemma 4.2. *For any $e \geq 4$, we have $\psi_e > k_e + 1$.*

Lemma 4.1 shows that $\Psi_e(\lambda)$ has a real root, hence ψ_e is well-defined. On the other hand, Lemma 4.2 implies that $\psi_e^2 - \psi_e - 2e > 0$. Therefore, ω_e is well-defined and we have $\omega_e > e + 2$ for any $e \geq 4$. We proceed with the following auxiliary lemma.

Lemma 4.3. *For any $e \geq 28$, the function $\lambda \mapsto \Psi_e(\lambda)$ is strictly increasing on $[k_e, +\infty)$.*

With Lemmas 4.1–4.3 in mind, we can apply a strategy inspired by Rowlinson [7] and Bell [1] to compare the indices of graphs $\mathcal{D}_{n,e}$ and $\mathcal{V}_{n,e}$ as follows.

Proposition 4.4. *For any $e \geq 4$ and $n \geq e+2$, exactly one of the next three claims holds:*

(i) *$\rho(\mathcal{V}_{n,e}) < \rho(\mathcal{D}_{n,e}) < \psi_e$;*

(ii) $\rho(\mathcal{V}_{n,e}) = \rho(\mathcal{D}_{n,e}) = \psi_e$;

(iii) $\rho(\mathcal{V}_{n,e}) > \rho(\mathcal{D}_{n,e}) > \psi_e$.

The next proposition can now be proved by using Proposition 4.4 together with the results from Section 3.

Proposition 4.5. *For any $e \geq 4$, we have $\omega_e > e+2$, alongside the following:*

(i) *if $b_e \leq n < \omega_e$, then $\rho(\mathcal{D}_{n,e}) < \psi_e$;*

(ii) *if $n = \omega_e$, then $\rho(\mathcal{D}_{n,e}) = \psi_e$;*

(iii) *if $n > \omega_e$, then $\rho(\mathcal{D}_{n,e}) > \psi_e$.*

By combining Propositions 3.6, 4.4 and 4.5, we solve the index maximization problem on $\mathcal{C}_{n,e}$ for the case $e \in [4, 85]$.

Theorem 4.6. *For any $e \in [4, 85]$ and $n \geq b_e$, we have:*

(i) *if $b_e \leq n < \omega_e$, then $\mathcal{D}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$;*

(ii) *if $n = \omega_e$, then $\mathcal{D}_{n,e}$ and $\mathcal{V}_{n,e}$ are the graphs attaining the maximum index on $\mathcal{C}_{n,e}$;*

(iii) *if $n > \omega_e$, then $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

Finally, by following a strategy inspired by Bell [1], we obtain the following result on the extremality of $\mathcal{V}_{n,e}$.

Theorem 4.7. *Let $e \geq 5$ be such that $t_e \geq 1$ and let $\ell_e = \frac{ek_e}{e-k_e-1}$. Then for any $n > e + 2 + \frac{\ell_e(\ell_e+1)(\ell_e^2-\ell_e-2e)}{\ell_e^2-e}$, the graph $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

The next corollary follows directly from Theorems 1.4 and 4.7.

Corollary 4.8. *For any $e > 85$ and $n \geq \frac{5}{2}e$, the graph $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

Corollary 4.8 extends Theorem 1.3, where the same extremal property of $\mathcal{V}_{n,e}$ was proved with a much larger bound $f(e)$ on n that exceeds $e^2(e+2)^2$ when $e \geq 7$.

5 Conclusion

Theorems 4.6 and 4.7 further substantiate the prediction that for any $e \in \mathbb{N}_0$ and $n \geq b_e$, the solution to the index maximization problem on $\mathcal{C}_{n,e}$ is either only $\mathcal{D}_{n,e}$, only $\mathcal{V}_{n,e}$, or both of these graphs. In view of Propositions 4.4 and 4.5 and Theorem 4.6, we end the paper with the following natural conjecture.

Conjecture 5.1. *For any $e \geq 4$ and $n \geq b_e$, we have:*

(i) *if $b_e \leq n < \omega_e$, then $\mathcal{D}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$;*

(ii) *if $n = \omega_e$, then $\mathcal{D}_{n,e}$ and $\mathcal{V}_{n,e}$ are the graphs attaining the maximum index on $\mathcal{C}_{n,e}$;*

(iii) *if $n > \omega_e$, then $\mathcal{V}_{n,e}$ uniquely attains the maximum index on $\mathcal{C}_{n,e}$.*

References

[1] F. K. Bell. On the maximal index of connected graphs. *Linear Algebra Appl.*, 144:135–151, 1991.

[2] R. A. Brualdi and E. S. Solheid. On the spectral radius of connected graphs. *Publ. Inst. Math.*, 39(53)(59):45–54, 1986.

[3] V. Chvátal and P. L. Hammer. Aggregation of inequalities in integer programming. In P. L. Hammer, E. L. Johnson, B. H. Korte, and G. L. Nemhauser, editors, *Studies in Integer Programming*, volume 1 of *Annals of Discrete Mathematics*, pages 145–162. Elsevier, 1977.

[4] D. Cvetković and P. Rowlinson. On connected graphs with maximal index. *Publ. Inst. Math.*, 44(58):29–34, 1988.

[5] N. V. R. Mahadev and U. N. Peled. *Threshold Graphs and Related Topics*, volume 56 of *Annals of Discrete Mathematics*. Elsevier, 1995.

[6] D. D. Olesky, A. Roy, and P. van den Driessche. Maximal graphs and graphs with maximal spectral radius. *Linear Algebra Appl.*, 346(1–3):109–130, 2002.

[7] P. Rowlinson. On the maximal index of graphs with a prescribed number of edges. *Linear Algebra Appl.*, 110:43–53, 1988.

[8] D. Stevanović. Spectral radius of graphs. In A. M. Encinas and M. Mitjana, editors, *Combinatorial Matrix Theory*, pages 83–130. Springer International Publishing, Cham, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The asymptotic spectrum distance, graph limits, and the Shannon capacity

(Extended abstract)

David de Boer* Pjotr Buys* Jeroen Zuiddam*

Abstract

Determining the Shannon capacity of graphs (Shannon 1956) is a long-standing open problem in information theory, graph theory and combinatorial optimization. Over decades, a wide range of upper and lower bound methods (of combinatorial, algebraic, analytic, and computational nature) have been developed to analyze this problem (e.g., Lovász 1979, Alon 1998, Bohman–Holzman 2003, Bukh–Cox 2019, Guruswami–Riazanov 2021, Polak–Schrijver 2023). However, despite tremendous effort, even small instances of the problem have remained open (e.g., the Shannon capacity of any odd cycle of length at least seven), and central computational and structural questions unanswered.

In recent years, a new dual characterization of the Shannon capacity of graphs, asymptotic spectrum duality, has unified and extended known upper bound methods and structural theorems. This duality originated from the work of Strassen in algebraic complexity theory (Strassen, J. Reine Angew. Math. 1987, 1988, 1991) and has recently seen much activity (Christandl–Vrana–Zuiddam, J. Amer. Math. Soc. 2023, Wigderson–Zuiddam 2024).

In this paper, building on asymptotic spectrum duality, we develop a new theory of graph distance, that we call *asymptotic spectrum distance*, and corresponding limits (reminiscent of, but different from, the celebrated theory of cut-norm, graphons and flag algebras). We propose a graph limit approach to the Shannon capacity problem: to determine the Shannon capacity of a graph, construct a sequence of easier to analyse graphs converging to it.

(1) We give a very general construction of non-trivial converging sequences of graphs (in the asymptotic spectrum distance). As a consequence, we prove new continuity properties of the Shannon capacity, the Lovász theta function and other graph parameters, answering questions of Schrijver and Polak.

(2) We construct Cauchy sequences of finite graphs that do not converge to any finite graph, but do converge to an infinite graph (in the asymptotic spectrum distance). We establish strong connections between convergence questions of finite graphs and the asymptotic properties of Borsuk-like infinite graphs on the circle, and use ideas from the theory of dynamical systems to study these infinite graphs.

*University of Amsterdam, Science Park 107, Amsterdam, Netherlands. E-mail: `d.deboer@uva.nl`, `p.j.w.buys@uva.nl`, `j.zuiddam@uva.nl`. Supported by NWO 613.001.851, 639.032.614, VI.Veni.212.284.

This is an extended abstract. The full paper can be found at `https://arxiv.org/abs/2404.16763`.

(3) We observe that all best-known lower bound constructions for Shannon capacity of small odd cycles can be obtained from a "finite" version of the graph limit approach, in which we approximate the target graph by another graph with a large independent set that is an "orbit" under a natural group action. We develop computational and theoretical aspects of this approach and use these to obtain a new Shannon capacity lower bound for the fifteen-cycle.

The graph limit point-of-view brings many of the constructions that have appeared over time in a unified picture, makes new connections to ideas in topology and dynamical systems, and offers new paths forward. The theory of asymptotic spectrum distance applies not only to Shannon capacity of graphs, but indeed we develop it for a general class of mathematical objects and their asymptotic properties.

1 Introduction

Posed by Shannon in 1956 [Sha56], determining the amount of information that can be transmitted perfectly over a noisy communication channel, the Shannon zero-error capacity, is a longstanding and central open problem in information theory, graph theory and combinatorial optimization [KO98, Alo02, Sch03]. Mathematically, this problem asks for determining the rate of growth of the independence number of powers of graphs.

Over decades, a wide range of upper and lower bound methods (of combinatorial, algebraic, analytic, and computational nature) have been developed to analyze this problem, by Shannon [Sha56], Lovász [Lov79], Haemers [Hae79, Hae81], Alon [Alo98], Alon–Lubetzky [AL06], Polak–Schrijver [PS19, Pol19b], Google DeepMind [RPBN$^+$24] and many others [Ros67, BMR$^+$71, Hal73, CFG$^+$93, BH03, BM13, Bla13, MOr17, GRW18, BC19, ABG$^+$20, Fri21, GR21, Sch23, Zhu24]. However, despite tremendous effort, even small instances of the problem have remained open (in particular, determining the Shannon capacity of any odd cycle of length at least seven), and central computational and structural questions unanswered.

In recent years, a new dual characterization of the Shannon capacity of graphs, called asymptotic spectrum duality [Zui19], has unified and extended known upper bound methods and structural theorems [Vra21, Sim21, Fri21, Sch23]. This duality originates from Strassen's work in algebraic complexity theory on fast matrix multiplication algorithms and tensors [Str87, Str88, Str91, CVZ23] (and goes back to the real representation theorems of Stone and Kadison–Dubois [BS83]) and applies much more generally to asymptotic problems in various fields [JV20, RZ21, Fri23]; see the survey of Wigderson–Zuiddam [WZ23].

In this paper, building on asymptotic spectrum duality, we develop a new theory of graph distance and limits—reminiscent of, but different from, the celebrated theory of cut-norm, graphons and flag algebras in the study of homomorphism densities [LS06, Raz07, BCL$^+$08, Lov12]. The crucial property of this *asymptotic spectrum distance* is that converging graphs have converging Shannon capacity,

$$G_i \to H \;\Rightarrow\; \Theta(G_i) \to \Theta(H),$$

thus suggesting a graph limit approach to the Shannon capacity problem: to determine (or, say, lower bound) the Shannon capacity of a "hard" graph, construct a sequence of easier to analyse (e.g., highly structured) graphs converging to it.

Central results of this paper are (1) how to construct converging sequences of graphs and (2) where to look for graphs that are "easier to analyse". Indeed, a special case of our results is that for any odd cycle graph we can construct many non-trivial converging sequences of very structured graphs converging to it (as illustrated in Figure 1). Using a "finite" version of this idea we will obtain a new lower bound on the Shannon capacity of the fifteen-cycle.

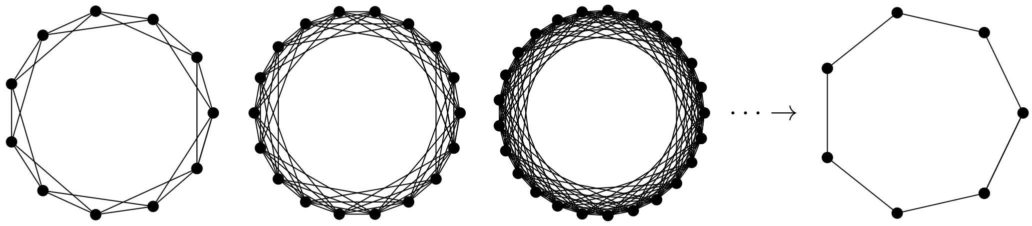


Figure 1: Beginning of non-trivial sequence converging to the seven-cycle (Theorem 3.1).

As we will see, the graph limit point-of-view opens the door to analyzing the Shannon capacity problem from many new angles, including: via infinite graphs, topological and analytical methods, and the theory of dynamical systems. It brings many of the best-known Shannon capacity constructions in a unified picture, and offers new paths forward.

We briefly summarize here our main results and we will expand on these in the rest of the text:

- **Convergence and continuity.** We develop asymptotic spectrum distance, including methods for bounding it. These lead to a very general construction of non-trivial converging sequences of graphs (in a family of circulant graphs that we call fraction graphs, and are also called circular graphs). As a consequence, we prove new continuity properties of the Shannon capacity, the Lovász theta function and other graph parameters. (Section 3)

- **Infinite graphs as limit points.** We construct Cauchy sequences of finite graphs that do not converge to any finite graph, but do converge to an infinite graph. More generally we establish strong connections between convergence questions of (finite) graphs and the asymptotic properties of (Borsuk-like) infinite graphs on the circle, and use ideas from the theory of dynamical systems to study these infinite graphs. (Section 4)

- **Graph limit approach and orbit constructions.** We observe that all best-known lower bound constructions for Shannon capacity of small odd cycles can be obtained from a "finite" version of the graph limit approach, in which we approximate the target graph by another graph with a large independent set that is an "orbit" under a natural group action. We develop computational and theoretical aspects of this approach and use it to obtain a new Shannon capacity lower bound for the fifteen-cycle. (Section 5)

In the rest of this text we discuss basic notions and main results in more detail.

2 Shannon capacity and asymptotic spectrum distance

The Shannon capacity of a graph G is defined as the rate of growth of the independence number under taking large powers of the graph (under the "strong product"), that is, $\Theta(G) = \lim_{n\to\infty} \alpha(G^{\boxtimes n})^{1/n}$ [Sha56]. By Fekete's lemma this equals the supremum $\sup_n \alpha(G^{\boxtimes n})^{1/n}$, so that we may think of Shannon capacity as a maximisation problem. We write $G \leq H$ if there is a cohomomorphism $G \to H$, a map $V(G) \to V(H)$ that maps non-edges to non-edges.

Asymptotic spectrum duality [Zui19] characterizes the Shannon capacity as the minimization $\Theta(G) = \min_{F\in\mathcal{X}} F(G)$, where $\mathcal{X}$, called the asymptotic spectrum of graphs, is a natural family of real-valued functions on graphs (namely, multiplicative, additive, cohomomorphism-monotone, normalized functions). The asymptotic spectrum $\mathcal{X}$ contains the Lovász theta function, fractional Haemers bound, fractional clique covering number, and other well-known upper bounds on Shannon capacity. While we do not have an explicit description of all elements of $\mathcal{X}$, several structural properties are known [Vra21].

We define the asymptotic spectrum distance by $d(G, H) = \sup_{F\in\mathcal{X}} |F(G) - F(H)|$ for any two graphs G, H.[1] From asymptotic spectrum duality it can be seen that converging graphs have converging Shannon capacity. Asymptotic spectrum distance can also be phrased in an operational form using asymptotic spectrum duality. Namely, $d(G, H)$ is small if and only if a "polynomial blow up" of G can be mapped (cohomomorphically) to a slightly larger polynomial blow up of H, and similarly for G and H reversed.[2]

3 Converging sequences

We construct (non-trivial) converging sequences of graphs in the asymptotic spectrum distance. For this we use a natural family of vertex-transitive graphs which we call fraction graphs.[3] For any $p, q \in \mathbb{N}$ with $p/q \geq 2$ we let $E_{p/q}$ be the graph with vertex set $V = \{0, \ldots, p-1\}$ and with distinct $u, v \in V$ forming an edge if $u - v$ or $v - u$ is strictly less than q modulo p.[4]

The graphs $E_{p/q}$, which we call *fraction graphs*, have many nice properties. For instance, $p/q \leq s/t$ (in $\mathbb{Q}$) if and only if $E_{p/q} \leq E_{s/t}$ in the cohomomorphism preorder. That is, fraction graphs are ordered as the rational numbers, and in particular inherit their denseness. (We note that $E_{p/q}$ and $E_{np/nq}$, for $n \in \mathbb{N}_{\geq 2}$, are not isomorphic graphs, but they are equivalent under cohomomorphism.) The fraction graphs contain all cycle graphs as $C_n = E_{n/2}$, and the edgeless graphs as $E_m = E_{m/1} = \overline{K}_m$.

We develop techniques to bound the asymptotic spectrum distance of vertex-transitive graphs and their induced subgraphs, and a vertex removal strategy for fraction graphs. Using these we obtain a general construction of converging sequences:

Theorem 3.1. *For any $a/b \geq 2$, if p_n/q_n converges to a/b from above, then E_{p_n/q_n} converges to $E_{a/b}$.*[5]

Theorem 3.1 solves a problem and extends the work of Schrijver and Polak [Pol19a, Chapter 9]. They constructed (by means of an independent set construction) for every integer $m \in \mathbb{N}$ a sequence of fraction graphs E_{p_n/q_n} that converges to $E_m = E_{m/1}$ from below, which implies that for any sequence p_n/q_n converging to m from below we have that E_{p_n/q_n} converges to E_m.

[1] Strictly speaking, this is not a distance, but it is a distance if we identify graphs under "asymptotic equivalence".

[2] For any $a, b \in \mathbb{N}$ we have $d(G, H) \leq a/b$ if and only if $(E_b \boxtimes G)^{\boxtimes n} \leq ((E_b \boxtimes H) \sqcup E_a)^{\boxtimes(n+o(n))}$ and $(E_b \boxtimes H)^{\boxtimes n} \leq ((E_b \boxtimes G) \sqcup E_a)^{\boxtimes(n+o(n))}$, where $E_n \coloneqq \overline{K}_n$ denotes the graph with n vertices and no edges.

[3] In the literature these have received much attention (under many names), by Vince [Vin88], Schrijver and Polak (as circular graphs) [Pol19a, PS19], as cycle-powers [BM13], and as (the complement of) rational complete graphs [HN04] and circular complete graphs [Zhu06].

[4] Despite the use of the letter p we do not require p to be prime.

[5] We note that, in Theorem 3.1, if the rational numbers p_n/q_n are distinct, then (using the fractional clique covering number) the graphs E_{p_n/q_n} can be seen to be pairwise non-equivalent and even asymptotically non-equivalent, so the sequence is indeed non-trivial.

We leave it as an open problem whether $p_n/q_n \to a/b$ from below also implies $E_{p_n/q_n} \to E_{a/b}$. In the next section we discuss how this problem is closely related to properties of infinite graphs.

Theorem 3.1 can alternatively be phrased in terms of continuity of the graph parameters in the asymptotic spectrum, $F \in \mathcal{X}$. Namely, it says that $p/q \mapsto F(E_{p/q})$ is right-continuous (and the same for $F = \Theta$).

With similar methods we prove:

Theorem 3.2. *For any irrational $r \in \mathbb{R}_{\geq 2}$, if p_n/q_n converges to r, then E_{p_n/q_n} is Cauchy.*

It follows from Theorem 3.2 with a simple argument that there are Cauchy sequences that do not converge to any finite graph. Indeed, the fractional clique cover number of any sequence in Theorem 3.2 converges to the irrational number r, which cannot be the fractional clique cover number of a finite graph. This answers a question raised in [Pol19a, Chapter 9].

There is a natural (albeit abstract) way to complete the space of finite graphs with the asymptotic spectrum distance, by identifying every graph G with its evaluation function $\widehat{G} : \mathcal{X} \to \mathbb{R} : F \mapsto F(G)$. The coarsest topology on $\mathcal{X}$ that makes the $\widehat{G}$ continuous is compact, and thus we can think of graphs as living in the space $C(\mathcal{X})$ of continuous functions on $\mathcal{X}$ endowed with the sup-norm, which is complete.

The above (keeping in mind the analogous theory of cut-norm and graphons) raises the question whether there are concrete, graph-like models for limit points; indeed we prove that infinite graphs provide such models.

4 Infinite graphs as limit points

We prove that certain infinite graphs (on the circle) are limit points for non-converging Cauchy sequences of finite graphs, and investigate the properties of these graphs, and strong connections to the Shannon capacity of finite graphs.

We define two kinds of infinite graphs on the circle (that may be thought of as infinite versions of the fraction graphs). Let $C \subseteq \mathbb{R}^2$ denote a circle with unit circumference. For $r \in \mathbb{R}_{\geq 2}$ we define E_r^{o} to be the infinite graph with vertex set C for which two vertices are adjacent if and only if they have distance strictly less than $1/r$ on C. We define E_r^{c} to be the infinite graph with vertex set C for which two vertices are adjacent if and only if they have distance at most $1/r$ on C. The subtle difference between the "open" and "closed" version of these circle graphs will play an important role.

We naturally extend asymptotic spectrum duality and distance to infinite graphs (with finite clique covering number), and are then able to establish the circle graphs as limit points:

Theorem 4.1. *For any irrational $r \in \mathbb{R}_{\geq 2}$, if p_n/q_n converges to r, then E_{p_n/q_n} converges to E_r^{o}.*

For any two graphs G, H we write $G \lesssim H$ if $G^{\boxtimes n} \leq H^{\boxtimes(n+o(n))}$. We say G and H are asymptotically equivalent, if $G \lesssim H$ and $H \lesssim G$. It follows from our proof of Theorem 4.1 that E_r^{c} is asymptotically equivalent to E_r^{o} for irrational r. For rational r we find:

Theorem 4.2. *Let $r = p/q \in \mathbb{Q}_{\geq 2}$. The following are equivalent:*

(i) *E_r^{c} and E_r^{o} are asymptotically equivalent.*

(ii) *If* $a_n/b_n \to p/q$ *from below, then* $E_{a_n/b_n} \to E_{p/q}$.

Graphs G and H are called equivalent if $G \leq H$ and $H \leq G$. Equivalent graphs are asymptotically equivalent (but asymptotically equivalent graphs may not be equivalent). As a step towards understanding E_r^{c} and E_r^{o} better, we prove that they are not equivalent:

Theorem 4.3. *Let* $r \in \mathbb{R}_{>2}$. *Then* E_r^{c} *and* E_r^{o} *are not equivalent.*

The proof of this result is based on a complete characterization of the self-cohomomorphisms of the circle graphs E_r^{c} and E_r^{o}, for which we use ideas from the theory of dynamical systems.

5 Independent sets from orbit constructions and reductions

Having delved into asymptotic spectrum distance and how to construct converging sequences, we focus in this part on the construction of explicit independent sets, their structure and ways of transforming them between graphs (reductions).

Over time, several structured and concise constructions of independent sets in powers of graphs (say, odd cycles) have been found and investigated, starting with the famous independent set $\{t \cdot (1,2) : t \in \mathbb{Z}_5\}$ in $C_5^{\boxtimes 2}$ [Sha56], and later [BMR+71, GR21, PS19], among others. In the recent work by Google DeepMind [RPBN+24], a large language model (LLM) recovered best-known lower bounds for the Shannon capacity of small odd cycles, and moreover did so in a concise manner (in the sense of Kolmogorov-complexity).

We propose a simple, general framework in which essentially all best-known lower bounds on Shannon capacity of small odd cycles fit—offering an explanation for the aforementioned structure and conciseness. This framework, which can be thought of as a finite version of the graph limit approach, consists of (1) relating the target graph $G = C_n = E_{n/2}$ to another fraction graph $H = E_{p/q}$ (an "intermediate" or "auxilary" graph), and then (2) constructing a large independent set in a power of H using an "orbit" construction. The aforementioned constructions of Shannon [Sha56] and Polak–Schrijver [PS19] are indeed of this form; other bounds (e.g., by Baumert, McEliece, Rodemich, Rumsey, Stanley and Taylor [BMR+71]) can be recovered in this way.

We develop methods for reducing a target graph G to an auxiliary graph and in particular introduce a new "nondeterministic rounding" technique. Using this technique applied to an orbit construction in (a power of) of a fraction graph close to the fifteen-cycle, we find a new Shannon capacity lower bound:

Theorem 5.1. $\Theta(C_{15}) \geq \alpha(C_{15}^{\boxtimes 4})^{1/4} \geq 2842^{1/4} \approx 7.30139$.

Our bound improves the previous bound $\Theta(C_{15}) \geq 7.25584$ which was obtained by Codenotti, Gerace and Resta [CGR03] and independently by Polak and Schrijver (personal communication).

Finally, we investigate independent sets in products of fraction graphs further. We focus on describing, for fixed k, the function $p/q \mapsto \alpha(E_{p/q}^{\boxtimes k})$, and more generally the multivariate version of this function, $(p_1/q_1, \ldots, p_k/q_k) \mapsto \alpha(E_{p_1/q_1} \boxtimes \cdots \boxtimes E_{p_k/q_k})$. For $k = 1$, a simple argument gives $\alpha(E_{p/q}) = \lfloor p/q \rfloor$. For $k = 2$, Hales [Hal73] and Badalyan and Markosyan [BM13] proved

$$\alpha(E_{p_1/q_1} \boxtimes E_{p_2/q_2}) = \min\{\lfloor \lfloor p_1/q_1 \rfloor p_2/q_2 \rfloor, \lfloor \lfloor p_2/q_2 \rfloor p_1/q_1 \rfloor\}.$$

We give a simple proof of this result using orbits. As our main result, we consider $k = 3$ and completely determine $\alpha_3(p_1/q_1, p_2/q_2, p_3/q_3) \coloneqq \alpha(E_{p_1/q_1} \boxtimes E_{p_2/q_2} \boxtimes E_{p_3/q_3})$ for all $p_i/q_i \in [2, 3]$.

References

[ABG+20] Noga Alon, Igor Balla, Lior Gishboliner, Adva Mond, and Frank Mousset. The minrank of random graphs over arbitrary fields. *Israel J. Math.*, 235(1):63–77, 2020. doi:10.1007/s11856-019-1945-8.

[AL06] Noga Alon and Eyal Lubetzky. The Shannon capacity of a graph and the independence numbers of its powers. *IEEE Trans. Inform. Theory*, 52(5):2172–2176, 2006. doi:10.1109/TIT.2006.872856.

[Alo98] Noga Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998. doi:10.1007/PL00009824.

[Alo02] Noga Alon. Graph powers. In *Contemporary combinatorics*, volume 10 of *Bolyai Soc. Math. Stud.*, pages 11–28. János Bolyai Math. Soc., Budapest, 2002. URL: https://web.math.princeton.edu/~nalon/PDFS/cap2.pdf.

[BC19] Boris Bukh and Christopher Cox. On a fractional version of Haemers' bound. *IEEE Trans. Inform. Theory*, 65(6):3340–3348, 2019. arXiv:1802.00476, doi:10.1109/TIT.2018.2889108.

[BCL+08] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008. doi:10.1016/j.aim.2008.07.008.

[BH03] Tom Bohman and Ron Holzman. A nontrivial lower bound on the Shannon capacities of the complements of odd cycles. *IEEE Trans. Inform. Theory*, 49(3):721–722, 2003. doi:10.1109/TIT.2002.808128.

[Bla13] Anna Blasiak. A graph-theoretic approach to network coding. *PhD thesis*, 2013. URL: https://hdl.handle.net/1813/34147.

[BM13] Sevak H. Badalyan and Stepan E. Markosyan. On the independence number of the strong product of cycle-powers. *Discrete Math.*, 313(1):105–110, 2013. doi:10.1016/j.disc.2012.09.008.

[BMR+71] Leonard D. Baumert, Robert J. McEliece, Eugene Rodemich, Howard C. Rumsey, Jr., Richard Stanley, and Herbert Taylor. A combinatorial packing problem. In *Computers in algebra and number theory (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1970)*, volume IV of *SIAM-AMS Proc.*, pages 97–108. Amer. Math. Soc., 1971. URL: https://math.mit.edu/~rstan/pubs/pubfiles/17.pdf.

[BS83] Eberhard Becker and Niels Schwartz. Zum Darstellungssatz von Kadison-Dubois. *Arch. Math. (Basel)*, 40(5):421–428, 1983. doi:10.1007/BF01192806.

[CFG+93] A. Robert Calderbank, Peter Frankl, Ronald L. Graham, Wen-Ching W. Li, and Lawrence A. Shepp. The Sperner capacity of linear and nonlinear codes for the cyclic triangle. *J. Algebraic Combin.*, 2(1):31–48, 1993. doi:10.1023/A:1022424630332.

[CGR03] Bruno Codenotti, Ivan Gerace, and Giovanni Resta. Some remarks on the Shannon capacity of odd cycles. *Ars Combin.*, 66:243–257, 2003. URL: `https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fa75fe8da43bc637170fbeeb58d4c07877ed4197`.

[CVZ23] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal points in the asymptotic spectrum of tensors. *J. Amer. Math. Soc.*, 36(1):31–79, 2023. `doi:10.1090/jams/996`.

[Fri21] Tobias Fritz. A unified construction of semiring-homomorphic graph invariants. *Journal of Algebraic Combinatorics*, 54:693–718, 2021. `arXiv:1901.01090`, `doi:10.1007/s10801-020-00983-y`.

[Fri23] Tobias Fritz. Abstract Vergleichsstellensätze for preordered semifields and semirings I. *SIAM J. Appl. Algebra Geom.*, 7(2):505–547, 2023. `doi:10.1137/22M1498413`.

[GR21] Venkatesan Guruswami and Andrii Riazanov. Linear Shannon capacity of Cayley graphs. In *2021 IEEE International Symposium on Information Theory (ISIT)*, page 988–992. IEEE Press, 2021. `arXiv:2009.05685`, `doi:10.1109/ISIT45174.2021.9517713`.

[GRW18] Alexander Golovnev, Oded Regev, and Omri Weinstein. The minrank of random graphs. *IEEE Trans. Inform. Theory*, 64(11):6990–6995, 2018. `doi:10.1109/TIT.2018.2810384`.

[Hae79] Willem Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(2):231–232, 1979. `doi:10.1109/TIT.1979.1056027`.

[Hae81] Willem Haemers. An upper bound for the Shannon capacity of a graph. In *Algebraic methods in graph theory, Vol. I, II (Szeged, 1978)*, volume 25 of *Colloq. Math. Soc. János Bolyai*, pages 267–272. North-Holland, Amsterdam-New York, 1981.

[Hal73] Raleigh S. Hales. Numerical invariants and the strong product of graphs. *J. Combinatorial Theory Ser. B*, 15:146–155, 1973. `doi:10.1016/0095-8956(73)90014-2`.

[HN04] Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms*, volume 28 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004. `doi:10.1093/acprof:oso/9780198528173.001.0001`.

[JV20] Asger Kjærulff Jensen and Péter Vrana. The asymptotic spectrum of LOCC transformations. *IEEE Trans. Inform. Theory*, 66(1):155–166, 2020. `doi:10.1109/TIT.2019.2927555`.

[KO98] János Körner and Alon Orlitsky. Zero-error information theory. *IEEE Trans. Inform. Theory*, 44(6):2207–2229, 1998. Information theory: 1948–1998. `doi:10.1109/18.720537`.

[Lov79] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, 2012. doi:10.1090/coll/060.

[LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006. doi:10.1016/j.jctb.2006.05.002.

[MOr17] K. Ashik Mathew and Patric R. J. Östergård. New lower bounds for the Shannon capacity of odd cycles. *Des. Codes Cryptogr.*, 84(1-2):13–22, 2017. doi:10.1007/s10623-016-0194-7.

[Pol19a] Sven C. Polak. *New methods in coding theory: Error-correcting codes and the Shannon capacity*. PhD thesis, University of Amsterdam, 2019. URL: https://hdl.handle.net/11245.1/393eb25b-c1ba-4c6f-89e8-c439a97358b6.

[Pol19b] Sven C. Polak. Semidefinite programming bounds for Lee codes. *Discrete Math.*, 342(9):2579–2589, 2019. doi:10.1016/j.disc.2019.05.019.

[PS19] Sven C. Polak and Alexander Schrijver. New lower bound on the Shannon capacity of C_7 from circular graphs. *Inform. Process. Lett.*, 143:37–40, 2019. doi:10.1016/j.ipl.2018.11.006.

[Raz07] Alexander A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007. doi:10.2178/jsl/1203350785.

[Ros67] Moshe Rosenfeld. On a problem of C. E. Shannon in graph theory. *Proc. Amer. Math. Soc.*, 18:315–319, 1967. doi:10.2307/2035288.

[RPBN+24] Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi. Mathematical discoveries from program search with large language models. *Nature*, 625(7995):468–475, January 2024. doi:10.1038/s41586-023-06924-6.

[RZ21] Robert Robere and Jeroen Zuiddam. Amortized circuit complexity, formal complexity measures, and catalytic algorithms. In *62nd IEEE Annual Symposium on Foundations of Computer Science (FOCS 2021)*, pages 759–769. IEEE, 2021. doi:10.1109/FOCS52979.2021.00079.

[Sch03] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, volume 24B of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.

[Sch23] Alexander Schrijver. On the Shannon capacity of sums and products of graphs. *Indag. Math. (N.S.)*, 34(1):37–41, 2023. doi:10.1016/j.indag.2022.08.009.

[Sha56] Claude E. Shannon. The zero error capacity of a noisy channel. *Institute of Radio Engineers Transactions on Information Theory*, IT-2,:8–19, 1956. doi:10.1109/tit.1956.1056798.

[Sim21] Gábor Simonyi. Shannon capacity and the categorical product. *Electron. J. Combin.*, 28(1), 2021. doi:10.37236/9113.

[Str87] Volker Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Math.*, 375/376:406–443, 1987. doi:10.1515/crll.1987.375-376.406.

[Str88] Volker Strassen. The asymptotic spectrum of tensors. *J. Reine Angew. Math.*, 384:102–152, 1988. doi:10.1515/crll.1988.384.102.

[Str91] Volker Strassen. Degeneration and complexity of bilinear maps: some asymptotic spectra. *J. Reine Angew. Math.*, 413:127–180, 1991. doi:10.1515/crll.1991.413.127.

[Vin88] Andrew Vince. Star chromatic number. *J. Graph Theory*, 12(4):551–559, 1988. doi:10.1002/jgt.3190120411.

[Vra21] Péter Vrana. Probabilistic refinement of the asymptotic spectrum of graphs. *Combinatorica*, 41(6):873–904, 2021. doi:10.1007/s00493-020-4324-5.

[WZ23] Avi Wigderson and Jeroen Zuiddam. Asymptotic spectra: Theory, applications and extensions, 2023. URL: https://staff.fnwi.uva.nl/j.zuiddam/papers/convexity.pdf.

[Zhu06] Xuding Zhu. Recent developments in circular colouring of graphs. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 497–550. Springer, Berlin, 2006. doi:10.1007/3-540-33700-8_25.

[Zhu24] Daniel G. Zhu. An improved lower bound on the Shannon capacities of complements of odd cycles, 2024. arXiv:2402.10025.

[Zui19] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity. *Combinatorica*, 39(5):1173–1184, 2019. doi:10.1007/s00493-019-3992-5.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The Zarankiewicz problem on tripartite graphs

(Extended abstract)

Francesco Di Braccio* Freddie Illingworth†

Abstract

In 1975, Bollobás, Erdős, and Szemerédi asked for the smallest τ such that a tripartite graph with n vertices in each part and minimum degree at least $n + \tau$ must contain $K_{t,t,t}$, conjecturing that $\tau = \mathcal{O}(n^{1/2})$ for $t = 2$. We prove that $\tau = \mathcal{O}(n^{1-1/t})$ which confirms their conjecture and is best possible assuming the widely believed conjecture that $\mathrm{ex}(n, K_{t,t}) = \Theta(n^{2-1/t})$. Our proof uses a density increment argument.

We also construct an infinite family of extremal graphs that are pairwise far apart (requiring the change of $\Omega(n^2)$ edges to get between any two).

1 Introduction

Turán-type problems are some of the oldest and most fundamental in Combinatorics. They ask how dense must a graph be in order to force the presence of a particular substructure. The first results were those of Mantel [24] and Turán [26], who determined the maximum number of edges in an n-vertex graph without a triangle or complete graph of size r, respectively. In 1946, Erdős and Stone [12] generalised Turán's theorem by showing that any n-vertex graph with at least $(1 - \frac{1}{r-1} + o(1))\binom{n}{2}$ edges contains a complete r-partite graph with t vertices in each part, denoted $K_r(t)$. The complete multipartite graphs are particularly important since the number of edges required to force the presence of these determines, up to lower order terms, the number of edges required to force the presence of any non-bipartite graph [11].

An old class of Turán-type problems that are still not well-understood asks for density conditions which guarantee a $K_r(t)$ inside a host graph that is itself k-partite (for some $k \geq r$). Questions of this type are often remarkably difficult, even for small values of r, t, and k. For instance, the notorious Zarankiewicz problem [27] from 1951 asks for the maximum number of edges, $z(n; t)$, in a bipartite graph with parts of size n that does not contain $K_{t,t} = K_2(t)$. The celebrated Kővári-Sós-Turán theorem [22] states that $z(n; t) = \mathcal{O}(n^{2-1/t})$. Though this bound is believed to be tight for all $t \geq 2$, no matching lower bound is known for $t \geq 4$ despite receiving considerable attention – for a comprehensive survey, see [14].

*Department of Mathematics, LSE, United Kingdom. E-mail: f.di-braccio@lse.ac.uk.
†Department of Mathematics, UCL, United Kingdom. E-mail: f.illingworth@ucl.ac.uk.

For k-partite host graphs with $k \geq 3$, even less is known. Interest in these problems can be traced back to the seminal 1975 paper of Bollobás, Erdős, and Szemerédi [6]. They initiated the extensive study of multipartite Turán-type problems [20, 25, 18, 23] which has been central to Combinatorics with applications in graph arboricity [1], list colouring [16], defective colouring [19], and strong chromatic numbers [17].

One of their main questions, which has so far remained elusive, concerned the first case, $k = r = 3$. That is, the following tripartite version of the Zarankiewicz problem.

Problem 1 (Bollobás-Erdős-Szemerédi)**.** For each t, what is the smallest $\tau = \tau(n)$ such that any $n \times n \times n$ tripartite graph with minimum degree at least $n + \tau$ contains $K_3(t)$?

In the above, an $n \times n \times n$ *tripartite graph* is a tripartite graph whose parts have size n. The reason to centre the minimum degree at n is that below this threshold the graph can be bipartite, and thus contain no triangles. Bollobás, Erdős, and Szemerédi [6] proved that[1] $\tau = \mathcal{O}(n^{1-1/(3t^2)})$ for $t \geq 2$. They highlighted the case $t = 2$, for which their result yields $\tau = \mathcal{O}(n^{11/12})$, and they conjectured that the correct value is actually $\tau = \Theta(n^{1/2})$. This would be tight, as seen by taking a graph $G = G[V_1, V_2, V_3]$ that contains all the edges in $(V_1 \times V_2) \cup (V_2 \times V_3)$ and adding a $K_{2,2}$-free graph with minimum degree $\Omega(n^{1/2})$ between V_1 and V_3.

Problem 1 saw no progress until a paper of Bhalkikar and Zhao [4] which gives an alternative proof of $\tau = \mathcal{O}(n^{1-1/(3t^2)})$ and provides a second extremal example. Recently, in simultaneous and independent work, Chen, He, Lo, Luo, Ma, and Zhao [9] improved the bound to $\tau = \mathcal{O}(n^{1-1/t(t+1)})$ and introduced yet a new family of extremal examples.

Our main result proves Bollobás, Erdős, and Szemerédi's conjecture that $\tau = \Theta(n^{1/2})$ when $t = 2$ and improves the answer to Problem 1 to $\mathcal{O}(n^{1-1/t})$.

Theorem 2. *For every positive integer t, there is a constant K such that, if G is an $n \times n \times n$ tripartite graph and $\delta(G) \geq n + Kn^{1-1/t}$, then G contains $K_{t,t,t}$.*

If $z(n; t) = \Omega(n^{2-1/t})$, as is widely believed, then this theorem is best possible up to the value of the constant K and so Theorem 2 provides a complete answer to Problem 1.

Our second result shows that there are even more non-isomorphic constructions than previously thought. We exhibit an infinite family of extremal examples which are far from each other (requiring the change of $\Omega(n^2)$ edges to get between any two) and very different from those in [9]. These constructions are based on the classical Andrásfai graphs – we refer the reader to [10, Section 2] for more details.

2 Sketching the proof of Theorem 2

Consider an $n \times n \times n$ tripartite graph G with minimum degree at least $n+\tau$ where $\tau = Kn^{1-1/t}$ for some large constant K. We label the three parts of G as V_1, V_2, V_3. This invokes a natural direction of edges in G: for $u \in V_i$, $v \in V_{i+1}$ (such indices will always be modulo 3) we say that $uv \in E(G)$ is a *forward edge* while $vu \in E(G)$ is a *backward edge*. Each vertex $v \in V_i$ now has a *forward neighbourhood* $N^+(v) \subseteq V_{i+1}$, and a *backward neighbourhood* $N^-(v) \subseteq V_{i-1}$, as well as *forward/backward degrees*, $\deg^+(v)$ and $\deg^-(v)$.

The proof centres around the following special class of edges.

[1] In their paper [6, Thm. 2.6] they state $\tau = \mathcal{O}(n^{1-1/t^2})$. However, as observed by Bhalkikar and Zhao [4], a factor of three is missing from their calculations.

Definition 3 (boosters)**.** Let e be an edge. Let the end-vertices of e be u and v where uv is in the forward direction.

- e is a *booster* if $\deg^+(v) \geq \deg^+(u)$,
- e is an r*-booster* if $\deg^+(v) \geq \deg^+(u) + r$.

One prominent reason why boosters are useful is that they are contained in many triangles.

Observation 4. *Let $\delta(G) \geq n + \tau$. If uv is a booster, then uv is contained in τ triangles.*

Proof. By inclusion-exclusion, the number of triangles containing uv is at least

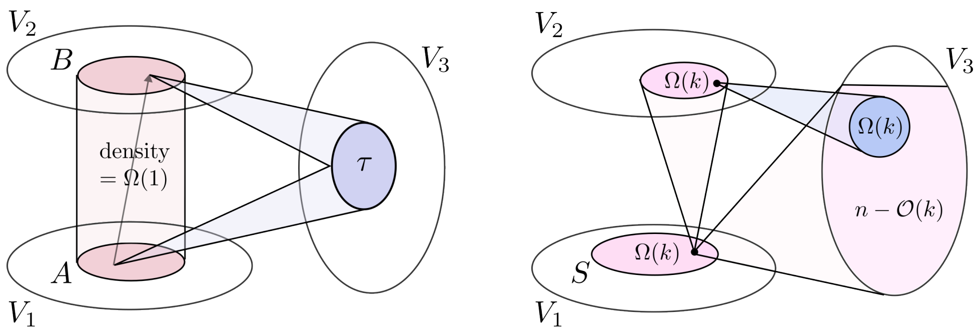
$$\deg^+(v) + \deg^-(u) - n \geq \deg^+(u) + \deg^-(u) - n \geq n + \tau - n = \tau. \qquad \square$$

Note that each booster edge gives $\binom{\tau}{t}$ copies of $K_{1,1,t}$. Thus, if there are sufficiently many booster edges in $V_1 \times V_2$, then by averaging a t-tuple $x_1, \ldots, x_t \in V_3$ is contained in many $K_{1,1,t}$s. If this number exceeds $z(n;t)$, then these $K_{1,1,t}$s form a $K_{t,t}$ in $V_1 \times V_2$, and thus a $K_{t,t,t}$ together with $x_1, \ldots, x_t$. In general there are not enough booster edges to make this simple counting argument work and we will instead need to pin down the structure of the booster edges within G.

Our strategy proceeds as follows. We say an edge uv is r*-heavy* if it is contained in r triangles (recall that boosters are τ-heavy by Observation 4). We first show – in [10, Section 5.1] – that there is some $k = \Omega(\tau)$ such that one of the following two structures must be present in G:

(a) **Dense subgraph of boosters:** a bipartite graph of booster edges with parts of size $\Omega(k)$ which has edge density 1/200;

(b) **Many heavy edges:** $\Omega(k)$ vertices each having forward degree $\mathcal{O}(k)$ and each being incident to $\Omega(k)$ forward edges that are $\Omega(k)$-heavy.



(a) Dense subgraph of boosters **(b)** Many heavy edges

Figure 1. The two possible configurations.

The remainder of the proof splits according to the configuration that we have found: we show that either regime forces the presence of a $K_{t,t,t}$. Both proofs use density increments, albeit in very different ways.

2.1 Dense regime

In this case we have a bipartite graph $H = H[A, B]$ of booster edges with parts of size $\Omega(k) = \Omega(n^{1-1/t})$ where H has density 1/200.

We handle this case through a density increment argument that roughly runs as follows. Note that each edge in $H_0 \coloneqq H$ is contained in at most n triangles. We show that if G is $K_{t,t,t}$-free, then H_0 has a subgraph H_1 containing a positive fraction of the edges of H_0 such that each edge in H_1 is contained in at most $n \cdot n^{-1/(2t^2)}$ triangles. In fact, this procedure can be iterated, thus giving a sequence of bipartite graphs $H_0, H_1, H_2, \ldots$, where the edges of H_{i+1} are contained in a number triangles that is smaller by a factor of $n^{-1/(2t^2)}$ relative to H_i. Since each booster has codegree at least $\Omega(n^{1-1/t})$ by Observation 4, the process can continue at most $2t$ times (by which point H_i still has positive density), and thus eventually we must find a $K_{t,t,t}$.

Let us give a taste of the ideas used in a single step of this procedure. The proof relies on another useful property of booster edges.

Observation 5. *Let $\delta(G) \geq n + \tau$. If a booster edge $uv \in V_i \times V_{i+1}$ is contained in at most d triangles, then $|N^+(v) \triangle N^-(u)| \geq |V_{i+2}| - 2d$.*

Proof. We have

$$\begin{aligned}|V_{i+2} \setminus (N^+(v) \cup N^-(u))| &= |V_{i+2}| - \deg^+(v) - \deg^-(u) + |N^+(v) \cap N^-(u)| \\ &\leq n - \deg^+(u) - \deg^-(u) + d \leq n - (n+\tau) + d = d - \tau.\end{aligned}$$

Thus, there are at most $d + (d - \tau) \leq 2d$ vertices in V_{i+2} that are either adjacent to both u and v or to neither of them. □

In other words, the neighbourhoods of u and v in V_{i+2} approximately partition V_{i+2}, with the approximation error being $2d$. This further means that if uv_1 and uv_2 are booster edges contained in at most d triangles, then v_1 and v_2 share the same neighbourhood up to a now slightly larger error of $4d$. If the edges of H_i are contained in at most d triangles, then any subgraph $H'[A', B']$ in which any two vertices are connected by a short path yields a bipartition $V_{i+2} = R_1 \cup R_2$ such that $G[A', R_1]$ and $G[B', R_2]$ are almost complete (i.e. of density $1 - \mathcal{O}(d/n)$). This observation yields powerful structural information that can be used to find many booster edges whose coneighbourhood is concentrated in a small set, allowing the procedure to continue – we refer the reader to [10, Section 5.2] for the details on how this is carried out.

2.2 Lots of heavy edges

The structure given by case (b) can be summarised through the following notion.

Definition 6 (r-squad)**.** An *r-squad* is a set of $k/800$ vertices in the same part, each having forward degree at most r and being incident to at least $k/800$ forward edges that are $2tr$-boosters.

As a heuristic, one should think of our starting configuration as an $\Omega(k)$-squad (although the definition just given is in fact stronger than case (b) above and it is non-trivial to go from one to the other).

Our density increment argument for this case is considerably more intricate, so we will limit ourselves here to some high level ideas. In a single step of our proof, we consider an r-squad $S \subseteq V_i$ (with $r = \Omega(k)$ at the start) and suppose that many $2tr$-boosters incident with S have their other endpoint in a set $T \subseteq V_{i+1}$. We prove that if G is $K_{t,t,t}$-free then one of the following 'moves' can be made:

(i) **Zoom into a denser sub-pair:** we find subsets $S' \subseteq S$ and $T' \subseteq T$ such that the density of $(2tr)$-boosters in (S', T') is much larger than in (S, T).

(ii) **Switch to an r'-squad with $r' > r$:** we obtain an improvement in the parameter controlling the squad by switching to a squad inside of V_{i+1} (so that the whole picture 'rotates' counterclockwise).

By carefully balancing these two types of moves, we obtain an iterative procedure that eventually terminates and finds a $K_{t,t,t}$ – the precise implementation of this idea can be found in [10, Section 5.3].

References

[1] Noga Alon. The linear arboricity of graphs. *Israel Journal of Mathematics*, 62(3):311–325, 1988.

[2] Noga Alon, Lajos Rónyai, and Tibor Szabó. Norm-graphs: Variations and applications. *Journal of Combinatorial Theory, Series B*, 76(2):280–290, July 1999.

[3] Béla Andrásfai. Über ein Extremalproblem der Graphentheorie. *Acta Mathematica Academiae Scientiarum Hungarica*, 13:443–455, September 1962.

[4] Abhijeet Bhalkikar and Yi Zhao. On subgraphs of tripartite graphs. *Discrete Mathematics*, 346(1):Paper No. 113152, 4, January 2023.

[5] Béla Bollobás. *Extremal Graph Theory.* Dover Publications, Inc., USA, 2004.

[6] Béla Bollobás, Paul Erdős, and Endre Szemerédi. On complete subgraphs of r-chromatic graphs. *Discrete Mathematics*, 13(2):97–107, 1975.

[7] William G. Brown. On graphs that do not contain a Thomsen graph. *Canadian Mathematical Bulletin*, 9(3):281–285, 1966.

[8] Boris Bukh. Extremal graphs without exponentially small bicliques. *Duke Mathematical Journal*, 173(11):2039–2062, August 2024.

[9] Yihan Chen, Jialin He, Allan Lo, Cong Luo, Jie Ma, and Yi Zhao. Complete tripartite subgraphs of balanced tripartite graphs with large minimum degree, November 2024.

[10] Francesco Di Braccio and Freddie Illingworth. The Zarankiewicz problem on tripartite graphs, December 2024.

[11] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:51–57, 1966.

[12] Paul Erdős and Arthur Harold Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52:1087–1091, 1946.

[13] Zoltán Füredi. Turán type problems. In *Surveys in Combinatorics, 1991*, volume 166 of *London Mathematical Society Lecture Note Series*, pages 253–300. Cambridge University Press, 1991.

[14] Zoltán Füredi and Miklós Simonovits. The history of degenerate (bipartite) extremal graph problems. In László Lovász, Imre Z. Ruzsa, and Vera T. Sós, editors, *Erdős Centennial*, pages 169–264. Springer, Berlin, Heidelberg, 2013.

[15] Jie Han and Yi Zhao. Turán number of complete multipartite graphs in multipartite graphs, May 2024.

[16] Penny Haxell. A note on vertex list colouring. *Combinatorics, Probability and Computing*, 10(4):345–347, July 2001.

[17] Penny Haxell. On the strong chromatic number. *Combinatorics, Probability and Computing*, 13(6):857–865, November 2004.

[18] Penny Haxell and Tibor Szabó. Odd independent transversals are odd. *Combinatorics, Probability and Computing*, 15(1–2):193–211, January 2006.

[19] Penny Haxell, Tibor Szabó, and Gábor Tardos. Bounded size components—partitions and transversals. *Journal of Combinatorial Theory, Series B*, 88(2):281–297, July 2003.

[20] Guoping Jin. Complete subgraphs of r-partite graphs. *Combinatorics, Probability and Computing*, 1(3):241–250, September 1992.

[21] János Kollár, Lajos Rónyai, and Tibor Szabó. Norm-graphs and bipartite Turán numbers. *Combinatorica*, 16(3):399–406, September 1996.

[22] Támas Kővári, Vera Sós, and Pál Turán. On a problem of K. Zarankiewicz. *Colloquium Mathematicae*, 3(1):50–57, 1954.

[23] Allan Lo, Andrew Treglown, and Yi Zhao. Complete subgraphs in a multipartite graph. *Combinatorics, Probability and Computing*, 31(6):1092–1101, November 2022.

[24] Willem Mantel. Problem 28: solution by H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh and W. A. Wythoff. *Wiskundige Opgaven*, 10:60–61, 1907.

[25] Tibor Szabó and Gábor Tardos. Extremal problems for transversals in graphs with bounded degree. *Combinatorica*, 26(3):333–351, June 2006.

[26] Pál Turán. Eine Extremalaufgabe aus der Graphentheorie. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.

[27] Kazimierz Zarankiewicz. Problem 101. *Colloquium Mathematicum*, 2:301, 1951.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Sparse volumetric weak ε-nets

(Extended abstract)

Travis Dillon*

Abstract

A "volumetric" weak ε-net theorem states that there are functions $f\colon (0,1]\times\mathbb{N}\to\mathbb{R}_{>0}$ and $v\colon \mathbb{N}\to\mathbb{R}_{>0}$ such that: For any finite collection $\mathcal{F}$ of sets in $\mathbb{R}^d$ with volume at least 1, there is a family $\mathcal{S}$ of at most $f(\varepsilon,d)$ sets, each of volume at least $v(d)$, such that $\operatorname{conv}(\mathcal{G})$ contains a set in $\mathcal{S}$ whenever $\mathcal{G}$ contains at least $\varepsilon|\mathcal{F}|$ sets in $\mathcal{F}$. Previous theorems had $f(\varepsilon,d)=O_d(\varepsilon^{-\alpha})$ with $\alpha=\frac{1}{4}d^2(d+3)^2-1$. We significantly improve this to $\alpha=d+1$. One consequence of this result is a reduced piercing number for volumetric (p,q)-theorems.

1 Introduction

The weak ε-net theorem, first proven in [1], is a gem of combinatorial geometry and has applications throughout discrete and computational geometry (see [14] for a survey). In short, it says that the set of convex hulls of "large" subsets of a point set X in $\mathbb{R}^d$ have a transversal whose size is independent of $|X|$. More precisely:

Theorem (Weak ε-nets). *For every $\varepsilon>0$, there is a constant $c(\varepsilon,d)$ such that for any finite point set $X\subseteq\mathbb{R}^d$, there is a set $P\subseteq\mathbb{R}^d$ of at most $c(\varepsilon,d)$ points such that* $\operatorname{conv}(Y)$ *contains a point of P whenever Y contains at least $\varepsilon|X|$ sets in X.*

Recently, a few papers [11,13,18] have introduced and applied *quantitative* versions of this theorem that produce a transversal by sets of positive volume.

Theorem (Jung–Naszódi [13]). *For every finite family $\mathcal{F}$ of volume-1 convex sets, there is a family $\mathcal{S}$ of at most $O_d(\varepsilon^{-d^2(d+3)^2/4+1})$ sets, each of volume at least d^{-d}, such that* $\operatorname{conv}(\mathcal{G})$ *contains a set in $\mathcal{S}$ whenever $\mathcal{G}$ contains at least $\varepsilon|\mathcal{F}|$ sets in $\mathcal{F}$.*

One focus of research on weak ε-nets is to establish better bounds on $c(d,\varepsilon)$. The first paper [1] proves an upper bound slightly better than $O_d(\varepsilon^{-(d+1)})$; after a series of works [8,16], the current best general upper bound is $c(d,\varepsilon)=o(\varepsilon^{-(d-1/2)})$ [19]. (Somewhat better bounds are

*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, USA. E-mail: `travis.dillon@mit.edu`. Supported by the US National Science Foundation Graduate Research Fellowship under Grant No. 2141064.

known for particular small dimensions.) The current best lower bound is $\Omega_d(\varepsilon^{-1}\log(1/\varepsilon)^{d-1})$ [7].

This paper addresses the same question for quantitative weak ε-nets. In particular, our results improve the bound on $|\mathcal{S}|$ in several regimes. To shorten the theorem statements and highlight their differences, we will write that a family $\mathcal{F}$ *has a weak ε-net of volume v and size k* to mean that

> *There is a family $\mathcal{S}$ of at most k sets, each of volume at least v, such that* $\operatorname{conv}(\mathcal{G})$ *contains a set in $\mathcal{S}$ whenever $\mathcal{G}$ contains at least $\varepsilon|\mathcal{F}|$ sets in $\mathcal{F}$.*

In the first regime, we use a simple parametrization argument to reduce the size of $\mathcal{S}$ at no cost to the volume:

Theorem 1. *Every finite family of volume-1 convex sets has a weak ε-net of volume d^{-d} and size $o_d(\varepsilon^{-(d(d+3)/2-1/2)})$.*

By reducing the volume of the sets in $\mathcal{S}$ only slightly, we can obtain another drastic reduction in size:

Theorem 2. *Every finite family of volume-1 convex sets has a weak ε-net of volume $(5d^3)^{-d}$ and size $O_d(\varepsilon^{-2d})$.*

Finally, if the aim is to minimize $|\mathcal{S}|$, no matter the cost to the volume, we can improve even further:

Theorem 3. *There is a function $v\colon \mathbb{N}\to\mathbb{R}_{>0}$ such that every finite family of volume-1 convex sets has a weak ε-net of volume $v(d)$ and size $O_d(\varepsilon^{-(d+1)})$.*

Our proof gives $v(d) \geq 4^{-d^2(1+o(1))}$. Our proof method extends to quantitative versions of the weak ε-net theorem, also with transversal size $O_d(\varepsilon^{-(d+1)})$, for many parameters besides volume, including diameter, surface area, and mean width. For simplicity, however, we focus only on volume in this abstract. For a more comprehensive account, as well as other related results, see the full paper on which this talk is based [10].

Background and related results

Initial interest in a volumetric weak ε-net theorem was mainly inspired by volumetric versions of the (p,q)-theorem, whose proofs use volumetric weak ε-nets as a crucial ingredient. The classical (p,q)-theorem [2] states that *for any $p \geq q \geq d+1$, there is a constant $c(p,q,d)$ such that for any (p,q)-intersecting family $\mathcal{F}$ of convex sets,*[1] *there is a set X of at most $c(p,q,d)$ points such that every set in $\mathcal{F}$ contains at least one point of X.* A volumetric variant replaces P by a family $\mathcal{S}$ of $c'(p,q,d)$ sets, each of volume at least $v(d)$. Several such theorems have appeared: Rolnick and Soberón [18] proved a version with $v(d) = 1-\delta$ but $q \gg d+1$; more recent work proved volumetric (p,q) theorems for $q \geq 3d+1$ [13] and $q \geq d+1$ [11], though for transversals with much smaller volume $v(d)$. Improving the volumetric weak ε-net yields a direct improvement on the size of the transversal guaranteed by the volumetric (p,q)-theorem. For example, the (p,q)-theorem in [13], whose results are easiest to quantify,

[1] A family $\mathcal{F}$ is *(p,q)-intersecting* if every collection of p sets in $\mathcal{F}$ contains q sets whose intersection is nonempty.

yields a transversal $\mathcal{S}$ with at most $(Cp)^{3d^6/4+O(d^5)}$ sets (for some constant $c > 0$); applying Theorem 3 from this note in their proof improves the transversal size to $(cp)^{3d^3+O(d^2)}$.

This area of volumetric combinatorial geometry originated in a 1982 paper of Bárány, Katchalski, and Pach [5] which proved quantitative versions of Helly's theorem, specifically that *if the intersection of every $2d$ or fewer elements of a finite family $\mathcal{F}$ of convex sets in $\mathbb{R}^d$ has volume greater than or equal to 1, then the volume of $\bigcap \mathcal{F}$ is at least d^{-2d^2}*. They conjectured that $\operatorname{vol}\left(\bigcap \mathcal{F}\right) \geq (Cd)^{-cd}$ for some constants $c, C > 0$. Naszódi proved this conjecture 2016 with $c = 2$ [17] (later improved by Brazitikos to $c = 3/2$ [6]). This breakthrough has sparked a continued interest over the last decade in volumetric and quantitative theorems in combinatorial geometry.

Helly's theorem itself has a much broader reach, and it forms the basis for astounding variety of research in geometry, topology, and combinatorics; for an overview of these many directions, see the surveys [3, 4, 9].

2 Proof outlines

In this section, we prove the theorems introduced in Section 1; to keep the exposition brief, not all details will be written out in full.

John's theorem is a key result used in many proofs in this area; it implies that the largest-volume ellipsoid contained in a convex set C has volume at least $d^{-d}\operatorname{vol}(C)$. Every ellipsoid $E \subseteq \mathbb{R}^d$ is given by the form $A(B^d) + x$ for some positive-definite matrix $A \in \mathbb{R}^{d\times d}$ and vector $x \in \mathbb{R}^d$, where B^d is the d-dimensional unit ball. So we can parametrize ellipsoids via the map $E \mapsto (A, x) \in PD^d \times \mathbb{R}^d$, where PD^d is the set of $d \times d$ positive-definite matrices, which is a convex subset of $\mathbb{R}^{d(d+1)/2}$. To prove Theorem 1, we parametrize the ellipsoids contained the family, thereby obtaining a collection of points in $\mathbb{R}^{d(d+3)/2}$, a space in which we can apply the classical weak ε-net theorem.

Proof of Theorem 1. By John's theorem, each set $F \in \mathcal{F}$ contains an ellipsoid $A_F(B^d)+x_F$ of volume at least d^{-d}; set $y_F := (A_F, x_F) \in PD^d \times \mathbb{R}^d \subset \mathbb{R}^{d(d+1)/2} \times \mathbb{R}^d$. We now apply Rubin's bound on weak ε-nets [19] in the space $\mathbb{R}^{d(d+3)/2}$ to conclude that there is a set $P \subseteq PD^d \times \mathbb{R}^d$ of $o_d(\varepsilon^{d(d+3)/2-1/2})$ points such $P \cap \operatorname{conv}(Y) \neq \emptyset$ for any $Y \subseteq \{y_F\}_{F\in\mathcal{F}}$ of size $|Y| \geq \varepsilon|\mathcal{F}|$. The set $S = \left\{(A, x) : \operatorname{vol}\left(A(B^d) + x\right) \geq d^{-d}\right\}$ is convex; since $\{y_F\}_{F\in\mathcal{F}} \subseteq S$, the set $S \cap P$ is also a weak ε-net. Every point in $S \cap P$ corresponds to an ellipsoid in $\mathbb{R}^d$ of volume at least d^{-d}, and this family of ellipsoids is a volumetric weak ε-net for $\mathcal{F}$. □

The proofs of the remaining weak ε-net results will rely on quantitative versions of the *selection lemma.* The following proof employs volumetric versions of several fundamental theorems in combinatorial geometry; for brevity, we cite the relevant consequences but do not write out the results in full.

Proposition 2.1 (volumetric selection lemma)**.** *There is a function $\beta\colon \mathbb{N} \to \mathbb{R}_{>0}$ such that, for any family $\mathcal{F}$ of convex sets in $\mathbb{R}^d$, each with volume at least 1, there is an ellipsoid of volume at least $(5d^3)^{-d}$ that is contained in $\operatorname{conv}(\mathcal{G})$ for at least $\beta(d)\binom{|\mathcal{F}|}{2d}$ choices of $\mathcal{G} \in \binom{\mathcal{F}}{2d}$.*

Proof. From John's theorem, we know that each set in $\mathcal{F}$ contains an ellipsoid of volume at least d^{-d}. By a parametrization argument similar to the proof of Theorem 1, we can obtain a Tverberg-type theorem for ellipsoids [20]. In particular, any $f(d) = \left(\frac{d(d+3)}{2}+1\right)\left(\frac{d(d+3)}{2}-1\right)+1$

sets in $\mathcal{F}$ can be partitioned into $r := \frac{d(d+3)}{2}$ collections $\mathcal{G}_1, \dots, \mathcal{G}_r$ such that $\bigcap_{i=1}^r \operatorname{conv}(\mathcal{G}_i)$ contains an ellipsoid E of volume d^{-d}.

Given this ellipsoid, take an affine transformation A that maps E to the unit ball B^d. Then for each $i \in [r]$, we have $B^d \subseteq \operatorname{conv}(A(F) : F \in \mathcal{G}_i)$. By a quantitative version of the Steinitz theorem [12], for each $i \in [r]$ there is a set of at most $2d$ points $X_i \subseteq A(\mathcal{G}_i)$ such that $\operatorname{conv}(X_i) \supseteq \frac{1}{5d^2} B^d$.

Now take the inverse affine transformation. Let $\mathcal{G}_i'$ denote a collection of at most $2d$ sets in $\mathcal{G}_i$ that covers $A^{-1}(X_i)$. Then $\operatorname{conv}(\mathcal{G}_i') \supseteq \operatorname{conv}(A^{-1}X_i) \supseteq \frac{1}{5d^2} E$ for each i.

Let $\mathcal{C} = \{\operatorname{conv}(\mathcal{G}) : \mathcal{G} \in \binom{\mathcal{F}}{2d}\}$. The previous argument shows that among any $f(d)$ sets in $\mathcal{F}$, there are r disjoint subfamilies of size $2d$ such that the intersection of their convex hulls contains an ellipsoid of volume at least $(5d^3)^{-d}$. Thus, the number of elements of $\binom{\mathcal{C}}{r}$ whose intersection contains an ellipsoid of volume $(5d^3)^{-d}$ is at least

$$\frac{\binom{|\mathcal{F}|}{f(d)}}{\binom{|\mathcal{F}|-2dr}{f(d)-2dr}} \geq \alpha_d \binom{\binom{|\mathcal{F}|}{2d}}{r} = \alpha_d \binom{|\mathcal{C}|}{r}$$

for some $\alpha_d > 0$. By a volumetric version of the fractional Helly theorem [13, 20], there is a subset $\mathcal{C}' \subseteq \mathcal{C}$ containing at least $\beta(d)\,|\mathcal{C}| = \beta(d)\binom{|\mathcal{F}|}{2d}$ sets such that $\bigcap \mathcal{C}'$ contains an ellipsoid of volume $(5d^{-3})^{-d}$. This is what we wanted to prove. □

Our proof of Theorem 2 follows the original proof for existence of weak ε-nets [1].

Proof of Theorem 2. We construct a weak ε-net via a greedy algorithm, starting from $\mathcal{S} = \emptyset$. At each stage, if some collection $\mathcal{F}' \subseteq \mathcal{F}$ of size $\varepsilon|\mathcal{F}|$ is not yet pierced by an ellipsoid in $\mathcal{S}$, use the selection lemma to choose an ellipsoid of volume $(5d^3)^{-d}$ that is contained in $\operatorname{conv}(\mathcal{G})$ for at least $\beta(d)\binom{|\mathcal{F}'|}{2d}$ choices of $\mathcal{G} \in \binom{\mathcal{F}'}{2d}$; add this ellipsoid to the set $\mathcal{S}$. Every time an ellipsoid is added, the number of sets $\mathcal{G} \in \binom{\mathcal{F}}{2d}$ whose convex hull is pierced by an ellipsoid in $\mathcal{S}$ increases by at least $\beta(d)\binom{\varepsilon|\mathcal{F}|}{2d}$. Therefore, this algorithm adds at most

$$\frac{\binom{|\mathcal{F}|}{2d}}{\beta(d)\binom{\varepsilon|\mathcal{F}|}{2d}} \leq c_d \varepsilon^{-2d}$$

ellipsoids before all $\varepsilon|\mathcal{F}|$-tuples are pierced. □

To prove Theorem 3, we will prove a version of Lemma 2.1 for $(d+1)$-tuples.

Proposition 2.2. *There are functions $v, \beta \colon \mathbb{N} \to \mathbb{R}_{>0}$ such that, for any family $\mathcal{F}$ of convex sets, each with volume at least 1, there is an ellipsoid of volume at least $v(d)$ that is contained in $\operatorname{conv}(\mathcal{G})$ for at least $\beta(d)\binom{|\mathcal{F}|}{d+1}$ choices of $\mathcal{G} \in \binom{\mathcal{F}}{d+1}$.*

To prove Theorem 3, we modify our proof of Lemma 2.1 to find volume in less-structured sets than ellipsoids. The key lemma that carries this out is:

Lemma 2.3. *For any finite sets $X_1, \dots, X_m \subseteq \mathbb{R}^d$ of k points each, such that $\operatorname{vol}\left(\bigcap_{i=1}^m \operatorname{conv}(X_i)\right) \geq 1$, there are subsets $Y_i \subseteq X_i$ of of size $|Y_i| = d+1$ such that $\operatorname{vol}\left(\bigcap_{i=1}^m \operatorname{conv}(Y_i)\right) \geq v(d, m, k) > 0$.*

Proof. Let $\mathcal{H}_i$ denote the set of hyperplanes determined by d points in X_i, and define $\mathcal{H} = \cup_{i=1}^m \mathcal{H}_i$. Then $|\mathcal{H}| \le m\binom{k}{d}$, and these hyperplanes divide $\mathbb{R}^d$ into at most $\left(m\binom{k}{d}\right)^d$ connected regions (see, for example, Proposition 6.1.1 of [15]). Thus, there is a convex set $C \subseteq \bigcap_{i=1}^m \operatorname{conv}(X_i)$ with volume at least $\left(m\binom{k}{d}\right)^{-d}$ whose interior does not intersect any of the hyperplanes in $\mathcal{H}$. By Carathéodory's theorem, the simplices determined by the points of X_j cover the set $\bigcap_{i=1}^m \operatorname{conv}(X_i)$. Because $\int(C)$ intersects no hyperplane determined by X_i, there is a simplex $Y_i \subseteq X_i$ such that $C \subseteq \operatorname{conv}(Y_i)$. Therefore $\operatorname{vol}\left(\bigcap_{i=1}^m \operatorname{conv}(Y_i)\right) \ge \operatorname{vol}(C) \ge \left(m\binom{k}{d}\right)^{-d}$. □

Proof of Lemma 2.2. In the proof of Theorem 2, use Lemma 2.3 to find sets $Y_i \subseteq X_i$ of size $d+1$; then (re)define $\mathcal{G}_i'$ as a collection of $d+1$ sets in $\mathcal{G}_i$ that covers $A^{-1}(Y_i)$. Then there is a set C with $\operatorname{vol}(C) \ge v(d, \frac{1}{2}d(d+3), 2d)\operatorname{vol}(E)$ such that $\operatorname{conv}(\mathcal{G}_i') \supseteq C$ for every i. The rest of the proof may be copied exactly, with every instance of $2d$ replaced by $d+1$, resulting with a final volume of at least $v(d, \frac{1}{2}d(d+3), 2d)\, d^{-d} \ge (cd^3 2^{2d})^{-d}$. □

Proof of Theorem 3. Identical to the proof of Theorem 2, except that Lemma 2.2 replaces Lemma 2.1. □

References

[1] Noga Alon, Imre Bárány, Zoltán Füredi, and Daniel J. Kleitman, *Point selections and weak ε-nets for convex hulls*, Combinatorics, Probability, and Computing **1** (1992), 189–200.

[2] Noga Alon and Daniel J. Kleitman, *Piercing convex sets and the Hadwiger-Debrunner (p, q)-theorem*, Advances in Mathematics **96** (1992), 103–112.

[3] N. Amenta, J. De Loera, and P. Soberón, *Helly's theorem: New variations and applications*, Algebraic and Geometric Methods in Discrete Mathematics (Heather Harrington, Mohamed Omar, and Matthew Wright, eds.), American Mathematical Society, 2017, 55–96.

[4] Imre Bárány and Gil Kalai, *Helly-type problems*, Bulletin of the American Mathematical Society **59** (2022), 471–502.

[5] Imre Bárány, Meir Katchalski, and János Pach, *Quantitative Helly-type theorems*, Proceedings of the American Mathematical Society **86** (1982), 109–114.

[6] Silouanos Brazitikos, *Brascamp–Lieb inequality and the quantitative versions of Helly's theorem*, Mathematika **63** (2017), 272–291.

[7] Boris Bukh, Jiří Matoušek, and Gabriel Nivasch, *Lower bounds for weak epsilon-nets and stair-convexity*, Israel Journal of Mathematics **182** (2011), 199–228.

[8] B. Chazelle, H. Edelsbrunner, M. Grigni, L. J. Guibas, M. Sharir, and E. Welzl, *Improved bounds on weak ε-nets for convex sets*, Discrete & Computational Geometry **13** (1995), 1–15.

[9] L. Danzer, B. Grünbaum, and V. Klee, *Helly's theorem and its relatives*, Convexity, Proceedings of Symposia in Pure Mathematics, vol. 7, American Mathematical Society, 1963, 101–180.

[10] Travis Dillon, *Volumetric selection theorems*, forthcoming (as of May 2025).

[11] Nóra Frankl, Attila Jung, and István Tomon, *The quantitative fractional Helly theorem*, 2024, arXiv:2402.12268v2 [math.CO].

[12] Grigory Ivanov and Márton Naszódi, *Quantitative Steinitz theorem: A polynomial bound*, Bulletin of the London Mathematical Society **56** (2023), 796–802.

[13] Attila Jung and Márton Naszódi, *Quantitative fractional Helly and (p, q)-theorems*, European Journal of Combinatorics **99** (2022), 103424.

[14] Jiří Matoušek, *Epsilon-nets and computational geometry*, New Trends in Discrete and Computational Geometry (János Pach, ed.), Algorithms and Combinatorics, vol. 10, Springer, 1993, 69–89.

[15] Jiří Matoušek, *Lectures on discrete geometry*, Graduate Texts in Mathematics, vol. 212, Springer, 2002.

[16] Jiří Matoušek and Uli Wagner, *New constructions of weak ε-nets*, Discrete & Computational Geometry **32** (2004), 195–206.

[17] Márton Naszódi, *Proof of a conjecture of Bárány, Katchalski and Pach*, Discrete & Computational Geometry **55** (2016), 243–248.

[18] D. Rolnick and P. Soberón, *Quantitative (p, q) theorems in combinatorial geometry*, Discrete Mathematics **340** (2017), 2516–2527.

[19] Natan Rubin, *Stronger bounds for weak ε-nets in higher dimensions*, Proceedings of the Annual Symposium on Foundations of Computer Science (STOC 2021), 2021, 989–1002.

[20] Sherry Sarkar, Alexander Xue, and Pablo Soberón, *Quantitative combinatorial geometry for concave functions*, Journal of Combinatorial Theory, Series A **182** (2021), 105465.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Lazy walk-based bounds on the spectral radius of threshold graphs

(Extended abstract)

Péter Csikvári* Ivan Damnjanović† Dragan Stevanović‡ Stephan Wagner§

Abstract

A threshold graph is a simple graph whose vertices can be ordered in such a way that each vertex is either adjacent or nonadjacent to all of the previous vertices simultaneously. A lazy walk in a graph is a sequence of its vertices such that each two consecutive vertices are either adjacent or equal. Here we give a recurrence formula for the number of particular lazy walks in threshold graphs, which we use to derive both lower and upper bounds on the spectral radius of threshold graphs.

1 Introduction

We consider finite simple graphs. The spectral radius $\rho(G)$ of a graph G is the spectral radius of its adjacency matrix. A threshold graph is a simple graph whose vertices can be ordered as $v_1, v_2, \ldots, v_n$, so that for each $2 \leq i \leq n$, vertex v_i is either adjacent or nonadjacent simultaneously to all of $v_1, v_2, \ldots, v_{i-1}$. This constructive process is naturally described with a generating sequence $a_1 a_2 \cdots a_n$, where $a_i = 1$ (resp. $a_i = 0$) indicates that vertex v_i is adjacent (resp. nonadjacent) to all of $v_1, v_2, \ldots, v_{i-1}$. We will also say that v_i is a type 1 (resp. *type 0*) vertex if $a_i = 1$ (resp. $a_i = 0$).

Brualdi and Hoffman [1, p. 438] posed in 1976 the extremal problem of maximizing the spectral radius over the set of graphs with a prescribed number of edges. Note that this

*HUN-REN Alfréd Rényi Institute of Mathematics and Eötvös Loránd University, Budapest, Hungary. E-mail: peter.csikvari@gmail.com.

†Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia and Faculty of Electronic Engineering, University of Niš, Niš, Serbia and Diffine LLC, San Diego, California, USA. E-mail: ivan.damnjanovic@elfak.ni.ac.rs. Supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, grant number 451-03-137/2025-03/200102, and the Science Fund of the Republic of Serbia, grant #6767, Lazy walk counts and spectral radius of threshold graphs—LZWK.

‡College of Integrative Studies, Abdullah Al-Salem University, Khaldiya, Kuwait. E-mail: dragan.stevanovic@aasu.edu.kw. Supported by the Science Fund of the Republic of Serbia, #6767, Lazy walk counts and spectral radius of threshold graphs—LZWK.

§Institute of Discrete Mathematics, TU Graz, Austria and Department of Mathematics, Uppsala University, Sweden. E-mail: stephan.wagner@tugraz.at.

formulation allows graphs to be disconnected. Brualdi and Hoffman [2] proved in 1985 that each such extremal graph is necessarily a threshold graph, while Rowlinson [3] in 1988 showed that for each $k, t \in \mathbb{N}$ such that $t < k$, the extremal graph with $\binom{k}{2} + t$ edges is the graph whose only nontrivial component arises from K_k by adding a new vertex and then joining it to t vertices of K_k. Interestingly, if we require the graphs to be connected, the corresponding extremal problem is not yet fully resolved. Brualdi and Solheid [4], and independently Simić, Li Marzi and Belardo [5], showed that each extremal graph must still be a threshold graph, with two prominent candidates identified in the literature [6].

Starting with the close relation between the spectral radius and the walk counts, we show here that both lower and upper bounds on the spectral radius of threshold graphs can be obtained from the counts of their particular lazy walks. We first obtain the recurrence relation for the numbers of such lazy walks, and then obtain the bounds by considering particular fragments of this recurrence. Full details are given in [7].

A *lazy walk* of length $k \in \mathbb{N}_0$ in a given graph G is any sequence $v_0 v_1 v_2 \cdots v_k$ of vertices from $V(G)$ such that for each $1 \leq i \leq k$, we have either $v_i = v_{i-1}$ or $v_i \sim v_{i-1}$ in G. Also, an empty sequence is considered to be an *empty lazy walk*. It is clear that for any $k \in \mathbb{N}_0$, the entry $(A(G) + I)^k_{uv}$ contains the number of lazy walks of length k in G whose starting and ending vertex are u and v, respectively. With this in mind, we give the next folklore lemma.

Lemma 1. *Let G be a connected graph and let U_1 and U_2 be any two nonempty subsets of $V(G)$. Furthermore, for each $k \in \mathbb{N}_0$, let Ω_k denote the number of lazy walks of length k in G whose starting and ending vertex belong to U_1 and U_2, respectively. Then we have*

$$\lim_{k \to +\infty} \sqrt[k]{\Omega_k} = 1 + \rho(G).$$

Let G be a connected threshold graph of order $n \geq 4$ and size $n - 1 < m < \binom{n}{2}$ with the generating sequence $a_1 a_2 \cdots a_n$ and the corresponding vertices $v_1, v_2, \ldots, v_n$. For the lazy walk $W\colon v_{i_0} v_{i_1} v_{i_2} \cdots v_{i_k}$, we define its *signature* as $\sigma(W)\colon a_{i_0} a_{i_1} a_{i_2} \cdots a_{i_k}$. Let $c \geq 3$ and $z \in \mathbb{N}$ denote the number of type 1 and type 0 vertices in G, respectively, so that $c + z = n$. We introduce the *backwards zero position* (BZP) sequence as the nonincreasing tuple $(b_1, b_2, \ldots, b_z) \in \mathbb{N}^z$ where b_i denotes the number of type 1 vertices that appear after the i-th type 0 vertex in the generating sequence. The generating sequence, and thus the threshold graph itself, can always be reconstructed from the given BZP sequence and the number of vertices. Also, note that $\binom{c}{2} + \sum_{i=1}^{z} b_i = m$, as well as $1 \leq b_i \leq c - 1$ for any $1 \leq i \leq z$.

2 Recurrence relation for lazy walks in threshold graphs

For $k \geq 1$, let Ψ_k denote the set of all the lazy walks of length $k - 1$ in G whose signature starts and ends with one, and let $\mathrm{LW}_k = |\Psi_k|$. For convenience, let $\mathrm{LW}_0 = 1$.

Let us consider first a lazy walk signature of the form

$$1\,\underbrace{0\cdots0}_{\text{block 1}}\,1\,\underbrace{0\cdots0}_{\text{block 2}}\,1\,\underbrace{0\cdots0}_{\text{block 3}}\,1\cdots1\,\underbrace{0\cdots0}_{\text{block } p}\,1 \tag{1}$$

for some fixed $p \geq 1$. The number of lazy walks with this signature does not depend of the numbers of zeros appearing in each block, since the nonadjacency of type 0 vertices means that all the zeros from the same block correspond to the same type 0 vertex. Thus, we can denote the number of lazy walks with any such signature by F_p. Trivially, $F_0 = c$. Besides, $F_1 = \sum_{i=1}^{z} b_i^2$ since there are b_i^2 ways to choose the starting and the ending type 1 vertex when the i-th type 0 vertex is selected to represent the zero block. Also, note that for any $1 \le i, j \le z$, the i-th and the j-th type 0 vertex have $\min\{b_i, b_j\}$ common neighbors. Since the type 0 vertices representing distinct zero blocks can be selected independently, we conclude that for any $p \ge 2$

$$F_p = \sum_{i_1, i_2, \ldots, i_p = 1}^{z} b_{i_1} \min\{b_{i_1}, b_{i_2}\} \min\{b_{i_2}, b_{i_3}\} \cdots \min\{b_{i_{p-1}}, b_{i_p}\}\, b_{i_p} \tag{2}$$

$$= \sum_{i_1, i_2, \ldots, i_p = 1}^{z} b_{i_1} b_{\max\{i_1, i_2\}} b_{\max\{i_2, i_3\}} \cdots b_{\max\{i_{p-1}, i_p\}}\, b_{i_p}. \tag{3}$$

Actually, if $B \in \mathbb{R}^{z \times z}$ is the matrix defined by $B_{ij} = b_{\max\{i,j\}}$, $1 \le i, j \le z$, so that

$$B = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & \cdots & b_z \\ b_2 & b_2 & b_3 & b_4 & \cdots & b_z \\ b_3 & b_3 & b_3 & b_4 & \cdots & b_z \\ b_4 & b_4 & b_4 & b_4 & \cdots & b_z \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_z & b_z & b_z & b_z & \cdots & b_z \end{bmatrix},$$

then the representation (3) implies that for any $p \ge 1$

$$F_p = w^\intercal B^{p-1} w,$$

where $w = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_z \end{bmatrix}^\intercal$.

We will now rely on the $(F_p)_{p \in \mathbb{N}_0}$ sequence to obtain a recurrence relation for computing the $(\mathrm{LW}_k)_{k \in \mathbb{N}_0}$ sequence. For $W \in \Psi_k$, let $\sigma_1(W)$ be the longest lazy subwalk of W whose signature does not contain two consecutive type 1 vertices, and let $\sigma_2(W)$ denote the remaining part of the walk W. Hence the signature of $\sigma_1(W)$ ends with 1, while the signature of $\sigma_2(W)$ starts with 1 (provided it is not empty).

Further, for $1 \le \ell \le k$, let $\Psi_{k,\ell}$ denote the subset of Ψ_k comprising the lazy walks W where $\sigma_1(W)$ is of length $\ell - 1$. Since

$$\mathrm{LW}_k = |\Psi_k| = \sum_{\ell=1}^{k} |\Psi_{k,\ell}|, \tag{4}$$

it is sufficient to count the walks from these k subsets separately.

Let $W \in \Psi_{k,\ell}$ for some $\ell \le k$. Since any two distinct type 1 vertices are adjacent, the subwalk $\sigma_1(W)$ can be selected independently from the subwalk $\sigma_2(W)$. Obviously, the subwalk $\sigma_2(W)$ can be chosen in $\mathrm{LW}_{k-\ell}$ different ways. On the other hand, the signature of $\sigma_1(W)$ is of the form (1) with a certain number of zero blocks. If $\ell = 1$, there are $F_0 = c$ choices for $\sigma_1(W)$. For any $\ell \ge 2$, there are $\binom{\ell-p-2}{p-1}$ signatures of the form (1) with ℓ vertices and $p \in \mathbb{N}$ zero blocks, with each such signature corresponding to F_p lazy walks. Thus, $|\Psi_{k,1}| = c\mathrm{LW}_{k-1}$, alongside

$$|\Psi_{k,\ell}| = \mathrm{LW}_{k-\ell} \sum_{p \in \mathbb{N}} \binom{\ell - p - 2}{p - 1} F_p \qquad (2 \le \ell \le k).$$

Observe that $|\Psi_{k,2}| = 0$ since $\sigma_1(W)$ cannot consist of two consecutive type 1 vertices. From (4) we now obtain the required recurrence relation

$$\mathrm{LW}_0 = 1, \tag{5}$$

$$\begin{aligned}\mathrm{LW}_k &= c\mathrm{LW}_{k-1} + \sum_{\ell=3}^{k} \mathrm{LW}_{k-\ell} \sum_{p\in\mathbb{N}} \binom{\ell-p-2}{p-1} F_p \\ &= c\mathrm{LW}_{k-1} + \sum_{r=0}^{k-3} \mathrm{LW}_r \sum_{q\in\mathbb{N}_0} \binom{k-3-r-q}{q} F_{q+1} \qquad (k\in\mathbb{N}).\end{aligned} \tag{6}$$

3 Lower bound via lazy walks

Let $(\mathrm{LW}'_k)_{k\in\mathbb{N}_0}$ be the auxiliary sequence defined by

$$\mathrm{LW}'_k = c^k \qquad (0\le k\le 2), \tag{7}$$

$$\mathrm{LW}'_k = c\,\mathrm{LW}'_{k-1} + \sum_{r=0}^{k-3} \mathrm{LW}'_r\, F_1 \qquad (k\ge 3). \tag{8}$$

Comparing (7) and (8) with (5) and (6), one immediately sees that $\mathrm{LW}'_k \le \mathrm{LW}_k$ for $k\in\mathbb{N}_0$. Further, the sequence $(\mathrm{LW}'_k)_{k\in\mathbb{N}_0}$ satisfies the linear recurrence relation

$$\mathrm{LW}'_k - (c+1)\,\mathrm{LW}'_{k-1} + c\,\mathrm{LW}'_{k-2} - F_1\,\mathrm{LW}'_{k-3} = 0. \tag{9}$$

From here and Lemma 1 we easily obtain the following bound.

Theorem 2. *The largest real root of the polynomial $x^3-(c+1)x^2+cx-F_1$ is a lower bound for $1+\rho(G)$.*

This bounds results in the following simple corollary.

Corollary 3. *We have $\rho(G) > c-1+\frac{F_1}{n^2}$.*

Note that the recurrence relation (6) can also be "cut" one step later to obtain another lower bound.

Theorem 4. *The largest real root of $x^5-(c+2)x^4+(2c+1)x^3-(c+F_1)x^2+F_1x-F_2$ is a lower bound for $1+\rho(G)$.*

4 Upper bound via lazy walks

From $\min\{b_i,b_j\}\le\sqrt{b_ib_j}$ one easily gets

$$F_p \le \left(\sum_{i=1}^{z} b_i\sqrt{b_i}\right)^2 \left(\sum_{i=1}^{z} b_i\right)^{p-2} \le \left(\sum_{i=1}^{z} b_i^2\right)\left(\sum_{i=1}^{z} b_i\right)^{p-1} = F_1\left(m-\binom{c}{2}\right)^{p-1}.$$

Hence for the auxiliary sequence $(\mathrm{LW}''_k)_{k\in\mathbb{N}_0}$ defined by

$$\mathrm{LW}''_0 = 1,$$
$$\mathrm{LW}''_k = c\,\mathrm{LW}''_{k-1} + \sum_{r=0}^{k-3} \mathrm{LW}''_r \sum_{q\in\mathbb{N}_0} \binom{k-3-r-q}{q} F_1 \left(m - \binom{c}{2}\right)^q \qquad (k \geq 1),$$

we have that $\mathrm{LW}_k \leq \mathrm{LW}''_k$ for each $k \in \mathbb{N}_0$. Since $(\mathrm{LW}''_k)_{k\in\mathbb{N}_0}$ satisfies the linear recurrence

$$\mathrm{LW}''_k - (c+1)\,\mathrm{LW}''_{k-1} + \left(\binom{c+1}{2} - m\right)\mathrm{LW}''_{k-2} + \left(c\left(m - \binom{c}{2}\right) - F_1\right)\mathrm{LW}''_{k-3} = 0,$$

we get the following upper bound.

Theorem 5. *The largest real root of* $x^3 - (c+1)x^2 + \left(\binom{c+1}{2} - m\right)x + \left(c\left(m - \binom{c}{2}\right) - F_1\right)$ *is an upper bound for* $1 + \rho(G)$.

One can similarly extend this approach to a quintic bound.

Theorem 6. *The largest real root of the polynomial*

$$x^5 - (c+2)x^4 + \left(\binom{c+2}{2} - m\right)x^3 - \left(c - (c+1)\left(m - \binom{c}{2}\right) + F_1\right)x^2$$
$$+ \left(F_1 - c\left(m - \binom{c}{2}\right)\right)x + \left[\left(m - \binom{c}{2}\right)F_1 - \left(\sum_{i=1}^{z} b_i\sqrt{b_i}\right)^2\right]$$

is an upper bound for $1 + \rho(G)$.

References

[1] J.C. Bermond, J.C. Fournier, M. Las Vergnas and D. Sotteau (eds.). Problèmes Combinatoires et Théorie des Graphes. Colloques internationaux du CNRS, Vol. 260, Éditions du CNRS, Paris, 1978.

[2] Richard A. Brualdi and Allan J. Hoffman. On the spectral radius of $(0,1)$-matrices. Linear Algebra Appl. 65:133–146, 1985.

[3] Peter Rowlinson. On the maximal index of graphs with a prescribed number of edges. Linear Algebra Appl. 110:43–53, 1988.

[4] Richard A. Brualdi and E.S. Solheid. On the spectral radius of connected graphs. Publ. Inst. Math. Belgrade NS 39(53):45–54, 1986.

[5] Slobodan Simić, Enzo M. Li Marzi and Francesco Belardo. Connected graphs of fixed order and size with maximal index: structural considerations. Le Mathematiche 59:349–365, 2004.

[6] F.K. Bell. On the maximal index of connected graphs. Linear Algebra Appl. 144:135–151, 1991.

[7] P. Csikvari, I. Damnjanović, D. Stevanović, S. Wagner. Some bounds on the spectral radius of connected threshold graphs. Preprint, arXiv: 2412.16019, 2024.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Meyniel's conjecture for Random Hypergraphs

(Extended abstract)

Gabriel Dias do Couto*

Abstract

Given a k-uniform hypergraph H^k, the Cops and Robbers game consists of a pursuit game in which the cops and the robber move, alternately, from their current vertex to another which is in a common edge. The cops win the game if a cop manages to move to the vertex of the robber in finite time. In [7] the authors conjectured that for any connected k-uniform hypergraph $O(\sqrt{n/k})$ cops are sufficient to guarantee victory. This may be considered a generalization of the well-known Meyniel conjecture for the Cops and Robbers game in graphs. They considered the case that $H^k \sim H^k(n,p)$ is the random hypergraph, in which each edge appears independently with probability p, and proved that $O(\sqrt{n/k}\ \log n)$ cops are sufficient, for a large range of k and p. We show that $O(\sqrt{n/k})$ cops are sufficient in the case of these random hypergraphs $H^k \sim H^k(n,p)$, thus proving the conjecture of [7] in these cases.

1 Introduction

The "Cops and Robbers game" was first introduced by Quilliot [15] and Nowakowski and Winkler [13]. The game takes place in a graph G and is played as follows: firstly the player who will control the cops chooses their initial positions. Then the player who controls the robber chooses, with full information, where she wants to begin. Now the cops player and the robber player move their pieces in alternate turns. All moves must be to a neighboring vertex. The cops player can move as many pieces as he wants (or choose not to). The robber player then makes the choice to move it's piece or stay still. Each player has full information about the graph and the pieces of the other player, at all times. The cops player wins if at some finite time one of its pieces occupies the same vertex as the robber. On the other hand, if the robber player can escape the cops indefinitely, she wins.

Shortly after the work of [13], Aigner and Fromme [2] defined the concept of *cop number*. The cop number of a graph G, $c(G)$, is the minimum number of cops needed for the cop player

*Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil
E-mail: `gdiascouto@aluno.puc-rio.br`. Supported by FAPERJ.

to have a winning strategy in G. The cop number is well defined since n cops, one for each vertex, always catch the robber. They also proved, among other results, that if G is planar, then $c(G) \leqslant 3$.

It can be easily seen that $c(G) = \sum_{G_i \in G} c(G_i)$, where G_i is a connected component of G. For this reason it suffices to study the case where G is connected. The most important conjecture on the game is Meyniel's conjecture (communicated by Frankl [8]).

Conjecture 1. *If G is a connected graph with n vertices, then $c(G) = O(\sqrt{n})$.*

The conjecture is tight in the sense that there are connected graphs such that $\sqrt{n}/2$ cops are not enough to catch the robber (see Bonato and Burgess [4]). We are still very far from the conjecture, since even its weak form, which asserts that $O(n^{1-\varepsilon})$ cops are sufficient (for some $\varepsilon > 0$) is still unknown. The general bounds of $O(n \log\log n/\log n)$ and $O(n/\log n)$ were proved in [8] and [6] respectively. The current best general upper bound is $n2^{-(1+o(1))\sqrt{\log_2 n}}$, which was proved independently by Lu and Peng [12], Scott and Sudakov [16] and Frieze, Krivelevich and Loh [9]. As random structures are often used to try to understand a general scope of a problem, it is natural to ask about the random graph case.

Considering the Erdős-Réyni $G(n,p)$ random graph model, Bonato, Hahn and Wang [5] first showed that if p is constant then $c(G(n,p))$ is logarithmic in n. For $p = n^{\alpha-1}$, $\alpha > 0$, Łuczak and Prałat [11] found evidence to support Meyniel's conjecture, showing a Zig-Zag behaviour of $c(G(n,p))$ depending on α.

Theorem 2. *[11] Let $\alpha > 0$ and $p = n^{\alpha-1+o(1)}$, then with high probability*

$$c(G(n,p)) = \tilde{O}(\sqrt{n}).$$

Furthermore, their result shows that the function $f : (0,1) \to [0,1]$

$$f(\alpha) = \frac{\log(\bar{c}(G(n, n^{\alpha-1})))}{\log n}$$

behaves as in the Figure 1.

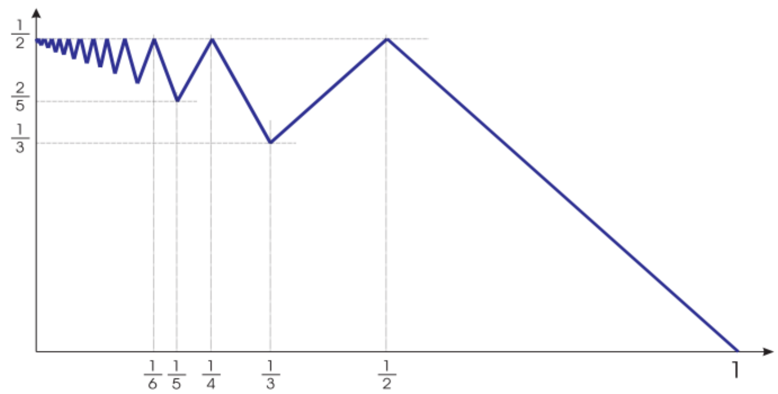


Figure 1: The ZigZag behavior of $f(\alpha)$. Reprinted from [11]

Their strategy consists of using a team of cops to enclose the robber in some neighborhood of size $O(\sqrt{n})$. Bollobás, Kun and Leader proved the same bound, although for sparser regimes of p. Refining the idea of [11], Prałat and Wormald [14] made a multi-team pursuit strategy depending on the density of the random graph and proved Meyniel's conjecture for random graphs for p greater than the connectivity threshold.

Theorem 3. *[14] Let $\varepsilon > 0$ and $p \geqslant (\frac{1}{2} + \varepsilon) \log n/n$, then with high probability*

$$c(G(n,p)) = O(\sqrt{n}).$$

For a hypergraph H, one may similarly define the Cops and Robbers game. The only difference is that one may now move to any vertex which shares an edge with the current vertex. In this way, the game is equivalent to the graph based game in which each edge of H is replaced by a clique. The cop number of H, $c(H)$, is similarly defined as the minimum number of cops needed to ensure that the cop player has a winning strategy in H. This problem was first considered by Gottlob, Leone and Scarcello [10] and Adler [1]. In Erde, Kang, Lehner, Mohar and Schmid [7] the authors introduce the following conjecture, which generalizes Meyniel's conjecture.

Conjecture 4. *[7] If H is a connected k-uniform hypergraph with n vertices, then $c(H) = O(\sqrt{n/k})$.*

The main theorem of [7] is in the setting of random k-uniform hypergraphs $H^k(n,p)$. There they used, as in [11], a team of cops to block the escape of the robber, though with two strategies, known as the edge-surrounding strategy and vertex-surrounding strategy. The strategy used depend on the parameters k and p. This mix of strategies gave them a dual-zigzag result represented in Figure 2. As a consequence of their zigzag result they obtain the following bound.

Theorem 5. *[7] If $k = \omega(\log n)$ and $\frac{n}{k} \geqslant p\binom{n-1}{k-1} = \omega(\log^3 n)$, then with high probability*

$$c(H^k(n,p)) = \tilde{O}\left(\sqrt{n/k}\right).$$

In this work, we combine and adapt ideas from [7] and [14] to remove the log-factor, although we need a slightly bigger k, and therefore prove Conjecture 4 for these random hypergraphs.

Theorem 6. *If $k = \Omega(\log^3 n)$ and $\frac{n}{k} \geqslant p\binom{n-1}{k-1} = \omega(\log^3 n)$, then with high probability*

$$c(H^k(n,p)) = O\left(\sqrt{n/k}\right).$$

2 Techniques

In [7, 11, 14], the authors first prove that for a given class of expanding (hyper)graphs a certain number of cops is sufficient, and then show that the random (hyper)graphs are in the class with high probability. We follow a direction similar to [14], but use, in some cases at the same time, the edge and vertex strategies from [7] with four teams of cops, each with size $O(\sqrt{n/k})$, to catch the robber.

For a vertex $v \in V(H)$ and $r \in \mathbb{N}$, let $N_V(v,r) = \{u \in V(H) : d(u,v) \leqslant r\}$ and $N_E(v,r) = \{e \in E(H) : d(e,v) \leqslant r\}$ be, respectively, the rth vertex and edge neighborhood of v. Also, let $S_V(v,r) = \{u \in V(H) : d(u,v) = r\}$ and $S_E(v,r) = \{e \in E(H) : d(e,v) = r\}$.

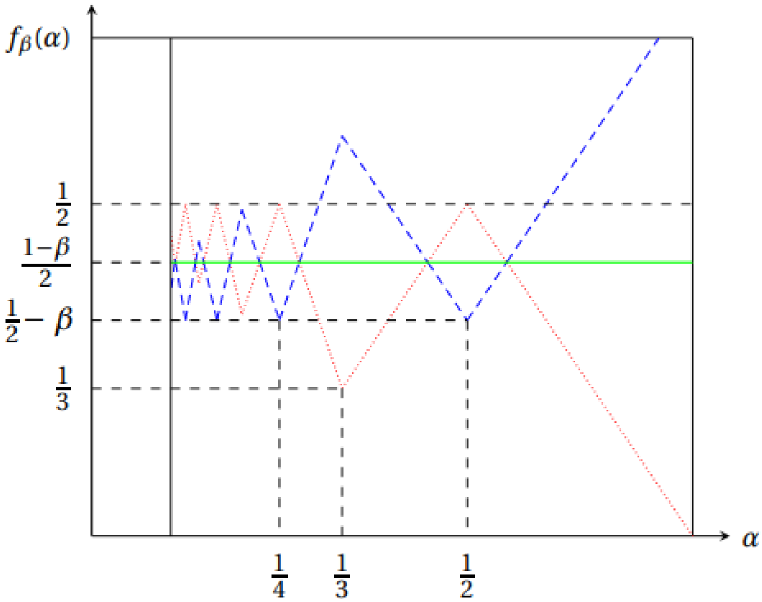


Figure 2: In this setting $p = n^{\alpha-1}$ is as before and $k = n^{\beta}$. We define $f_\beta(\alpha) = \frac{\log(c(H^k(n,p))}{\log n}$ and $\beta = 2/19$ in the graph. The blue dashed line represents the edge-surrounding strategy result and the pink one the vertex-surrounding strategy. The intersection of both strategies gives the worst bounds for the cop number. Reprinted from [7].

The main idea to surround the robber is the following: if the robber starts at $v \in V(H)$ and for a given $u \in S_V(v,r)$ there is a cop in $N_V(u, r+1)$, then the cop player has the option to send this cop to u before the robber can arrive there. Hence, if we can define an injective function $f : S_V(v,r) \to C$, from $S_V(v,r)$ to the set of initial positions of the cops, such that $d(u, f(u)) \leqslant r+1$, the cops can trap the robber in $N_V(v, r-1)$. It is then not difficult to catch the robber.

Another strategy is to surround the robber using its edges-neighborhoods. Therefore, if $e \in S_E(v,r)$ and there is a cop in $N_V(e,r)$, then the cop can reach this edge in round r and the robber will be caught in the next round if she enters that edge. Hence, if we can define an injective function $f : S_E(v,r) \to C$, such that $d(e, f(e)) \leqslant r$, the cops can again win. Note that in the edge strategy you go with distance r instead of $r+1$. This makes it more difficult to guarantee the family of available cops, since you lose a factor of d (the expected degree) in your neighborhood expansion, however, on the other hand, as one starts from an edge, one gains a factor of k to begin with.

Let $d = kp\binom{n-1}{k-1}$. Roughly, if a hypergraph is expanding, then $|N_V(v,r)| = \Theta(d^r)$ and $|N_E(v,r)| = \Theta(d^{r+1}/k)$. Each team of cops will be defined by a $\frac{C}{\sqrt{nk}}$-random subset of $V(H)$ independently. The following lemmas are the key ideas that we use for this result.

Lemma 7 (Vertices)**.** *Let $H^k = (V,E)$ be c-expanding.*

Let $X \subset V$, $|X| \leqslant 2\sqrt{\frac{n}{k}}$. Let $r \in \mathbb{N}$ be such that $d^r \geqslant \sqrt{nk}\log n$. Let $Y \subset V$ be a $\frac{C}{\sqrt{nk}}$-random subset of $V(H^k)$, $C > 0$. Then for sufficiently large C, with probability $1 - o(n^{-2})$, there is an injection $f : X \to Y$ such that $d(x, f(x)) \leqslant r$, $\forall x \in X$.

Lemma 8 (Edges)**.** *Let $H^k = (V, E)$ be c-expanding.*

Let $X \subset E$, $|X| \leqslant 2\sqrt{\frac{n}{k}}$. Let $r \in \mathbb{N}$ be such that $d^r \geqslant \sqrt{n/k}\log n$. Let $Y \subset V$ be a $\frac{C}{\sqrt{nk}}$-random subset of $V(H^k)$, $C > 0$. Then for sufficiently large C, with probability $1 - o(n^{-2})$, there is an injection $f : X \to Y$ such that $d(e, f(e)) \leqslant r$, $\forall e \in X$.

Let $r = \min\{i \in \mathbb{N} : d^i \geqslant \sqrt{nk}\}$. Note that in the lemmas we use that $\sqrt{nk}\log n \cdot \frac{C}{\sqrt{nk}} = \Omega(\log n)$. It may not be the case depending on k and p. Hence, I have to consider the size of d^r, d^r/k and d^{r+1} in order to use the lemmas. The case where $d^r = \sqrt{n/k} \cdot \omega$ and/or $d^{r+1} = \sqrt{nk} \cdot \omega$, with $1 \leqslant \omega \leqslant \log n$, we will require more ideas.

The new idea, inspired by [14], is: if the first team of cops 'densely' covers the vast majority of $S_V(v, r)$ or $S_E(v, r)$, then a second team would need to cover fewer vertices/edges in a second pursuit surrounding the escape routes that were left open by the first team. Hence we need two families of sets:

$$\big\{W(e) \subset S_V(e, r) \,:\, e \in S_E(v, r)\big\} \quad \text{and} \quad \big\{W(u) \subset S_V(u, r+1) \,:\, u \in S_V(v, r)\big\}.$$

We require that both families of subsets have the property of being pairwise disjoint, and that's where the bounds on k and $p\binom{n-1}{k-1}$ are needed. A $1 - \exp(-\Omega(\omega))$ proportion of $W(u)$'s and $W(e)$'s, where ω is as above, will have a cop inside of it, which will give our dense cover. After that we bound the number of escape routes left for the robber to take and surround then with a second team of cops.

The strategy of the second team to cover the gaps relies on the above lemmas to assign cops to vertex and edge strategies appropriately.

We originally mentioned the use of four teams. The other two are used simply to rule out certain "degenerate" strategies of the robber.

References

[1] I. Adler. Marshals, monotone marshals, and hypertree-width. *J. Graph Theory*, 47(4):275–296, 2004.

[2] M. Aigner and M. Fromme. A game of cops and robbers. *Discrete Applied Mathematics*, 8:1–12, 1984.

[3] B. Bollobás, G. Kun, and I. Leader. Cops and robbers in a random graph. *Journal of Combinatorial Theory, Series B*, 103(2):226–236, 2013.

[4] A. Bonato and A. Burgess. Cops and robbers on graphs based on designs. *J. Combin. Designs*, 21:404–418, 2013.

[5] A. Bonato, G. Hahn, and C. Wang. The cop density of a graph. *Contrib. Discrete Math*, 2(2):133–144, 2007.

[6] E. Chiniforooshan. A better bound for the cop number of general graphs. *Journal of Graph Theory*, 58:45–48, 2008.

[7] J. Erde, M. Kang, F. Lehner, B. Mohar, and D. Schmid. Catching a robber on a random k-uniform hypergraph. *arXiv:*, (2307.15512v2), 2024.

[8] P. Frankle. Cops and robbers in graphs with large girth and cayley graphs. *Discrete Applied Mathematics*, 17:301–305, 1987.

[9] A. Frieze, M. Krivelevich, and P. Loh. Variations on cops and robbers. *Journal of Graph Theory*, 69:383–402, 2012.

[10] G. Gottlob, N Leone, and F. Scarcello. Robbers, marshals, and guards: game theoreticand logical characterizations of hypertree width. *Journal of Computer and System Sciences*, 66(4):775–804, 2003.

[11] T. Łuczak and P. Prałat. Chasing robbers on random graphs: zigzag theorem. *Random Structures and Algorithms*, 37:516–524, 2010.

[12] T. Łuczak and P. Prałat. On meyniel's conjecture of the cop number. *Journal of Graph Theory*, 71:192–205, 2012.

[13] R. Nowakowski and P. Winkler. Vertex to vertex pursuit in a graph. *Discrete Mathematics*, 43:230–239, 1983.

[14] P. Prałat and N. Wormald. Meyniel's conjecture holds for random graphs. *Random Structures Algorithms*, 48(2):396–421, 2016.

[15] A. Quilliot. Jeux et pointes fixes sur les graphes. *Ph.D. Dissertation, Universit ́e de Paris VI*, 1978.

[16] A. Scott and B. Sudakov. A bound for the cops and robbers problem. *SIAM J. of Discrete Math*, 25:1438–1442, 2011.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

New constructions of unbalanced $\{C_4, \theta_{3,t}\}$-free bipartite graphs

(Extended abstract)

Baran Düzgün * Ago-Erik Riet † Vladislav Taranchuk ‡

Abstract

In 1979, Erdős conjectured that if $m = O(n^{2/3})$, then $ex(n, m, \{C_4, C_6\}) = O(n)$. This conjecture was disproven by several papers and the current best-known bounds for this problem are

$$c_1 n^{1+\frac{1}{15}} \leq ex(n, n^{2/3}, \{C_4, C_6\}) \leq c_2 n^{1+1/9}$$

for some constants c_1, c_2. A consequence of our work here proves that

$$ex(n, n^{2/3}, \{C_4, \theta_{3,4}\}) = \Theta(n^{1+1/9}).$$

More generally, for each integer $t \geq 2$, we establish that

$$ex(n, n^{\frac{t+2}{2t+1}}, \{C_4, \theta_{3,t}\}) = \Theta(n^{1+\frac{1}{2t+1}})$$

by demonstrating that subsets of points $S \subseteq \mathrm{PG}(n, q)$ for which no $t + 1$ points lie on a line give rise to $\{C_4, \theta_{3,t}\}$-free graphs, where $\mathrm{PG}(n, q)$ is the projective space of dimension n over the finite field of q elements.

*Hacettepe University, Graduate School of Science and Engineering, Beytepe 06800 Ankara, Türkiye. E-mail: baranduzgun@hacettepe.edu.tr. She gratefully acknowledges TÜBİTAK for the support BİDEB-2214-A during her visit at the University of Tartu and BİDEB-2224-A during EuroComb'25 conference.

†Institute of Mathematics and Statistics, University of Tartu, 51009 Tartu, Estonia. E-mail: ago-erik.riet@ut.ee. This work was supported by the Estonian Research Council grant PRG2531.

‡Department of Mathematics: Analysis, Logic and Discrete Mathematics, Ghent University, 9000 Ghent, Belgium. E-mail: vlad.taranchuk@ugent.be.

1 Introduction

Let m, n be positive integers and $\mathcal{F}$ be a family of graphs. The bipartite Turán number $ex(m, n, \mathcal{F})$, is the maximum number of edges in a bipartite graph whose part sizes are m and n and such that it contains no graph in $\mathcal{F}$ as a subgraph. The function $ex(m, n, \mathcal{F})$ has been studied extensively for many different sets $\mathcal{F}$, but many questions remain. See the well-known survey by Füredi and Simonovits [4] for a history of the work done on these types of problems.

One of the most notorious cases of determining $ex(n, m, \mathcal{F})$ is when $\mathcal{F} = C_{2k}$ for some positive integer k. When $m = n$, the order of magnitude of $ex(n, n, \{C_{2k}\})$ is only known for $k = 2, 3, 5$ [4], which coincides with the existence of special finite geometries called generalized polygons. When $m = n^a$ for any $a < 1$, and we allow $n \to \infty$, even less is known. The best constructions in this aforementioned case which yield C_4-free graphs come from the point-block incidence graphs of 2-designs.

In 1979, Erdős conjectured that when $m = O(n^{2/3})$, then $ex(m, n, \{C_4, C_6\}) = O(n)$ [3]. This was disproven first by de Caen and Székely [1] who constructed an infinite family of graphs which yielded

$$cn^{1+\frac{1}{57}+o(1)} \leq ex(n, n^{2/3}, \{C_4, C_6\})$$

for a constant c. Later this lower bound was improved by Lazebnik, Ustimenko and Woldar [6] who constructed an infinite family of graphs which yielded

$$n^{1+\frac{1}{15}} \leq ex(n, n^{2/3}, \{C_4, C_6\}).$$

It is worth mentioning that the class of graphs which is constructed in [6] is very much related to the (q, q^2)-generalized quadrangle. In fact, it seems to have escaped the graph theory community that the existence of such quadrangles also yields the bound obtained in [6] by taking an induced subgraph. The bounds on $ex(n, n^{2/3}, \{C_4, C_6\})$ have not budged in 30 years. In this paper, we establish results which suggest that the upper bound for this problem may be closer to the truth.

A common generalization of the cycle C_{2k}, is the theta graph $\theta_{k,t}$ which is the graph consisting of two vertices joined by t internally vertex-disjoint k-edge paths. While the lower and upper bounds for $ex(n, n, \{C_{2k}\})$ do not have matching orders of magnitude for all $k \neq 2, 3, 5$ [4], it is known that for each k, there exists a (relatively large) constant $t = t(k)$ such that

$$ex(n, n, \theta_{k,t}) = \Theta(n^{1+\frac{1}{k}}).$$

which is the same order of magnitude as the upper bound for $ex(n, n, C_{2k})$ [2]. In [2], Conlon uses a random algebraic method to construct infinite families of graphs not containing $\theta_{k,t}$, where t is fixed, but large relative to k. Therefore, explicit constructions yielding the same bounds and with a smaller t are of great interest.

There has also been some recent progress on determining $ex(n, m, \theta_{t,k})$. Jiang, Ma, and Yepremyan [5] proved that there exists a constant $c = c(k, t)$ such that

$$ex(m, n, \{\theta_{k,t}\}) \leq \begin{cases} c[(mn)^{\frac{k+1}{2k}+m+n}] & \text{if } k \text{ is odd} \\ c[(mn)^{\frac{k+2}{2k}} n^{\frac{1}{2}+m+n}] & \text{if } k \text{ is even} \end{cases}$$

and when $k = 3$, they obtained that $ex(m, n, \{\theta_{3,t}\}) \leq 144t^3((mn)^{\frac{2}{3}} + m + n)$. Theodorakopoulos [8] extended the random algebraic methods used in [2] to prove that for each odd positive integer k and rational number a satisfying $\frac{k-1}{k+1} < a < 1$, there exists a constant $c = c(k)$ such that

$$ex(n, n^a, \{\theta_{k,c_k}\}) = \Theta(\left(n^{1+a}\right)^{\frac{k+1}{2k}})$$

Here we prove that subsets of points S of the projective space $\mathrm{PG}(n, q)$ satisfying the condition that no $t + 1$ points of S lie on a line, produce $\{C_4, \theta_{3,t}\}$-free graphs via their linear representations. In particular, this implies the following theorem.

Theorem 1.1. *Let q be a prime power and t, n be positive integers. Suppose that S is a subset of points of $\mathrm{PG}(n, q)$ satisfying the condition that no $t + 1$ points of S lie on a common line. Then*

$$|S|q^{n+1} \leq ex(q^{n+1}, |S|q^n, \{C_4, \theta_{3,t}\}).$$

We remark that such sets with many elements are known to exist as shown in Lin and Wolf [7]. Consequently, we obtain our result.

Theorem 1.2. *Let $t \geq 2$ be a positive integer. Then*

$$ex(n, n^{\frac{t+2}{2t+1}}, \{C_4, \theta_{3,t}\}) = \Theta(n^{1+\frac{1}{2t+1}}).$$

Finally, our theorem implies the following corollary which suggests that the true value of $ex(n, n^{2/3}, \{C_4, C_6\})$ may be closer to the best-known upper bound.

Corollary 1.3. *We have*

$$ex(n, n^{2/3}, \{C_4, \theta_{3,4}\}) = \Theta(n^{1+\frac{1}{9}}).$$

2 Linear Representations of Point Sets

Let q be a prime power, let $\mathbb{F}_q^{n+1}$ denote the vector space of dimension $n+1$ over the finite field $\mathbb{F}_q$, and let $\mathrm{PG}(n, q)$ be the corresponding projective space.

Definition 2.1. *Let S be a set of points of $\mathrm{PG}(n, q)$ and embed $\mathrm{PG}(n, q)$ as a hyperplane into $\mathrm{PG}(n + 1, q)$. A* linear representation *of S is the geometry whose points are all the points in $\mathrm{PG}(n + 1, q) \setminus \mathrm{PG}(n, q)$ and the lines are all the lines of $\mathrm{PG}(n + 1, q)$ which intersect $\mathrm{PG}(n, q)$ in precisely one point, namely a point of S.*

Remark: Observe that if two lines in $\mathrm{PG}(n+1,q)$ not contained in $\mathrm{PG}(n,q)$ intersect in a point of S, then they are parallel in the linear representation of S.

From this geometry we may build its point-line incidence graph. This graph is bipartite with bipartition classes given by the points and the lines of the geometry. A point will be adjacent to a line in the graph if they are incident in the geometry, i.e. the point is on the line. Denote this graph by $\Gamma_{S,n,q}$. It can easily be verified that $\Gamma_{S,n,q}$ will have the following properties:

1. There are q^{n+1} point vertices, each of degree $|S|$.
2. There are $|S|q^n$ line vertices, each of degree q.

Theorem 2.2. *Let S be a subset of points of* $\mathrm{PG}(n,q)$ *such that any line in* $\mathrm{PG}(n,q)$ *intersects S in at most t points. Then $\Gamma_{S,n,q}$ is $\{C_4, \theta_{3,t}\}$-free.*

Proof. Note that by construction, the linear representation of S is a geometry in which any two lines intersect in at most one point and any two points lie on at most one line. Thus, the graph $\Gamma_{S,n,q}$ is necessarily C_4-free.

Suppose that $\Gamma_{S,n,q}$ contains a $\theta_{3,t}$. This implies there exists a point vertex r and line vertex ℓ between which there are t vertex-disjoint paths of length 3. Geometrically, this implies that the linear representation of S contains a configuration consisting of a point r, t lines which contain r, call them $m_1, m_2, \ldots, m_t$, all of which intersect the line ℓ in our geometry, and such that r is not on ℓ. Note that the set of lines $\ell, m_1, m_2, \ldots, m_t$ all pairwise intersect, and so no two can be parallel. Thus, in $\mathrm{PG}(n+1,q)$, each of the lines $\ell, m_1, \ldots, m_t$ contains a distinct point in S.

Note that all of the lines of this configuration lie in a common plane, Π, the plane in $\mathrm{PG}(n+1,q)$ spanned by r and ℓ. Since Π is not contained in the hyperplane $\mathrm{PG}(n,q)$, it intersects $\mathrm{PG}(n,q)$ in a line, call it ℓ_∞. But this implies that there is a set of $t+1$ points in S (one for each line $\ell, m_1, \ldots, m_t$) which lie on ℓ_∞. Since we assumed any line in $\mathrm{PG}(n,q)$ intersects S in at most t points, this is a contradiction. Thus $\Gamma_{S,n,q}$ is also $\theta_{3,t}$-free. □

Proposition 3 in [7] implies the following result by identifying the affine space $\mathbb{F}_q^{t+1}$ with $\mathrm{PG}(t+1,q)\backslash\mathrm{PG}(t,q)$ inside $\mathrm{PG}(t+1,q)$. For completeness, we give an explicit construction of such a set for all integers $t \geq 2$ and prime powers $q > t$.

Theorem 2.3. *Let q be a prime power and $q > t$ be a positive integer. Then there exists a subset S of points of* $\mathrm{PG}(t+1,q)$ *of size q^t such that no $t+1$ points of S lie on a line.*

It is well-known that the field $\mathbb{F}_{q^t}$ can be viewed as a vector space over $\mathbb{F}_q$. Fix any basis, $\alpha_1, \ldots, \alpha_t$ for $\mathbb{F}_{q^t}$ over $\mathbb{F}_q$. For each x in $\mathbb{F}_{q^t}$, denote by $x|_q$ the vector of the field reduced elements of x. That is, $x|_q = (x_1\alpha_1 + \cdots + x_t\alpha_t)|_q = (x_1, \ldots, x_t)$. Denote by N the norm function from $\mathbb{F}_{q^t}$ to $\mathbb{F}_q$, i.e. $N(x) = x^{(q^t-1)/(q-1)} = x \cdot x^q \cdot x^{q^2} \cdot \ldots \cdot x^{q^{t-1}}$. We note that when $t = 2$, the incidence graph of the linear representation of the set defined in the theorem below is isomorphic to the graph used in [6].

Theorem 2.4. *Let $t \geq 2$ be a positive integer and $q > t$ be a prime power. Then the set*

$$S = \{(1, x|_q, N(x)) : x \in \mathbb{F}_{q^t}\}$$

as a subset of points of $\mathrm{PG}(t+1, q)$ *contains no $t+1$ points on a line.*

Proof. We will omit the notation of field reduction to avoid getting bogged down in notation. Take any two vectors in S, call them $(1, x, N(x))$ and $(1, y, N(y))$. These two points lie on some line in $\mathrm{PG}(t+1, q)$. We will count how many other points $(1, z, N(z))$ in S can lie on this same line. If $(1, z, N(z))$ lies on the same line, it implies that the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ N(x) & N(y) & N(z) \end{bmatrix}$$

has an element in its null space $v = [a, b, c]$ with each $a, b, c \in \mathbb{F}_q^*$. Without loss of generality, we assume $c = -1$. Thus we obtain a linear system

$$\begin{aligned} a + b &= 1, \\ ax + by &= z, \\ aN(x) + bN(y) &= N(z). \end{aligned}$$

This system implies that

$$aN(x) + (1-a)N(y) = N(ax + (1-a)y) \tag{1}$$

Since x and y are fixed, that leaves a as the only variable. Since $a \in \mathbb{F}_q$, then $a^q = a$ and so expanding (1), we obtain a polynomial equation in a of degree t. In particular, the largest degree term a^t has coefficient $N(x-y) \neq 0$. Observe that $a = 0$ and $a = 1$ are both solutions to (1), which are not valid choices of a for us (since $a, b \neq 0$). Consequently, there are at most $t-2$ valid choices of a which solve (1). Each a determines a unique $z = ax + (1-a)y$, so any line contains at most t points of S. □

Consequently, the linear representation of this set produces a biregular, unbalanced bipartite graph with bipartition class sizes $m = q^{t+2}$, $n = q^{2t+1}$, and q^{2t+2} edges which is $\{C_4, \theta_{3,t}\}$-free. By an application of Bertrand's postulate, we obtain that for each t, there exists a constant c_t such that

$$c_t n^{1+\frac{1}{2t+1}} \leq ex(n, n^{\frac{t+2}{2t+1}}, \{C_4, \theta_{3,t}\}).$$

For each such t, the order of magnitude matches the upper bounds given in [5]. Thus we prove Theorem 1.2

$$ex(n, n^{\frac{t+2}{2t+1}}, \{C_4, \theta_{3,t}\}) = \Theta(n^{1+\frac{1}{2t+1}}).$$

References

[1] D. de Caen and L. A. Székely, "The maximum size of 4– and 6–cycle free bipartite graphs on m, n vertices", in *Graphs, Sets and Numbers*, Proc. of the Conference Dedicated to the 60th Birthdays of A. Hajnal and Vera T. Sós, Budapest, (1991).

[2] D. Conlon, "Graphs with few paths of prescribed length between any two vertices", *Bulletin of the London Mathematical Society*, 51(6), (2019), 1015–1021.

[3] P. Erdős, "Some old and new problems in various branches of combinatorics", *Proceedings of the Tenth Southeastern Conference on Combinatorics, Graph Theory and Computing*, Boca Raton, Vol. 1, Congressus Numerantium 23, (1979), 19–38.

[4] Z. Füredi and M. Simonovits, "The History of Degenerate (Bipartite) Extremal Graph Problems", *Erdős Centennial Bolyai Society Mathematical Studies*, 25, (2013).

[5] T. Jiang, J. Ma, and L. Yepremyan, "On Turán exponents of bipartite graphs", *Combinatorics, Probability and Computing*, 31(2), (2022), 333–344.

[6] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar, "New constructions of bipartite graphs on m, n vertices with many edges and without small cycles", *Journal of Combinatorial Theory Series B*, 61(1), (1994), 111–117.

[7] Y. Lin and J. Wolf, "On subsets of $\mathbb{F}_q^n$ containing no k-term progressions", *European Journal of Combinatorics*, 31(5), (2010), 1391–1403.

[8] S. Theodorakopoulos, "Some sharp lower bounds for the bipartite Turán number of theta graphs", , arXiv: 2405.02864, (2024).

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Rainbow Directed C_4's in Arc-colored Digraphs

(Extended abstract)

Mengyu Duan* Binlong Li* Minghui Yu* Shenggui Zhang*

Abstract

A subdigraph in an arc-colored digraph is rainbow if all of its arcs have distinct colors. For a given digraph H, let $rb(n,H)$ be the minimum integer such that every arc-colored complete digraph of order n with at least $rb(n,H)$ colors contains a rainbow copy of H. In this paper, we determine $rb(n,\overrightarrow{C_4})$ and characterize the corresponding extremal arc-colorings of complete digraph of order n. In addition, we prove that an arc-colored digraph D^C on n vertices contains a rainbow $\overrightarrow{C_4}$ if $a(D)+c(D^C)\geq n(n-1)+rb(n,\overrightarrow{C_4})$, where $a(D)$ ($c(D^C)$) is the number of arcs (colors) of D^C. Moreover, we characterize the extremal arc-colored digraphs achieving this bound.

1 Introduction

In this paper, we only consider finite digraphs with no loops and multiple arcs. For terminology and notations not defined here, we refer the reader to [3].

Let D be a digraph with vertex set $V(D)$ and arc set $A(D)$. We use $a(D)$ to denote the number of arcs of D. If $(u,v)\in A(D)$, then we say that u *dominates* v or v is *dominated* by u. For a vertex v of D, the *in-neighborhood* $N_D^-(v)$ of v is the set of vertices

*School of Mathematics and Statistics, Northwestern Polytechnical University, 710129, Xi'an, Shaanxi , China; Xi'an-Budapest Joint Research Center for Combinatorics, Northwestern Polytechnical University, 710129, Xi'an, Shaanxi China. Supported by NSFC (Nos. 12171393, 12071370, 12131013) and Shaanxi Fundamental Science Research Project for Mathematics and Physics (No. 22JSZ009). E-mail: myduan@mail.nwpu.edu.cn (M. Duan), binlongli@nwpu.edu.cn (B. Li), mhyu@mail.nwpu.edu.cn (M. Yu), sgzhang@nwpu.edu.cn (S. Zhang).

dominating v, and the *out-neighborhood* $N_D^+(v)$ of v is the set of vertices dominated by v. We define $N_D(v) = N_D^-(v) \cup N_D^+(v)$. We denote by $d_D^-(v) = |N_D^-(v)|$, $d_D^+(v) = |N_D^+(v)|$, and $d_D(v) = d_D^-(v) + d_D^+(v)$ the *indegree*, *outdegree*, and *degree* of $v \in V(D)$, respectively. When there is no fear of confusion, we write $N^-(v)$, $N^+(v)$, $N(v)$, $d^-(v)$, $d^+(v)$ and $d(v)$ instead of $N_D^-(v)$, $N_D^+(v)$, $N_D(v)$, $d_D^-(v)$, $d_D^+(v)$ and $d_D(v)$, respectively. Two arcs $(u_1, v_1), (u_2, v_2)$ are *nonadjacent* if $\{u_1, v_1\} \cap \{u_2, v_2\} = \emptyset$, otherwise they are *adjacent*. If either $v_1 = u_2$ or $v_2 = u_1$, then they are *consecutive*; and if both $v_1 = u_2$ and $v_2 = u_1$, then they are *symmetric*.

The *symmetric orientation* of a graph G, denoted by $\overleftrightarrow{G}$, is the digraph obtained from G by replacing each edge uv of G with a pair of symmetric arcs (u, v) and (v, u). Let $\overleftrightarrow{K_n}$ be the complete digraph of order n. The *directed path* (or *cycle*) on k vertices is denoted by $\overrightarrow{P_k}$ (or $\overrightarrow{C_k}$). For convenience, we use uv to denote the arc (u, v).

An *arc-coloring* of a digraph D is defined as a mapping $C : A(D) \to \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers. We denote by D^C the digraph D assigned to such an arc-coloring C. We use $C(D^C)$ to denote the set of colors assigned to the arcs of D^C, and we denote by $c(D^C) = |C(D^C)|$ the *color number* of D^C. A subdigraph of D^C is *rainbow* if all its arcs have distinct colors.

The study of the existence of rainbow cycles in edge-colored graphs has a relatively long history, in which short cycles play an important role. Erdős, Simonovits and Sós [6] show that if an edge-colored complete graph G uses at least $|V(G)|$ colors, then it contains a rainbow triangle. Li, Ning, Xu and Zhang [7] show that if the sum of edge number and color number of an edge-colored graph G is at least $|V(G)|(|V(G)|+1)/2$, then it contains a rainbow triangle.

The existence of rainbow C_4's has been widely studied in edge-colored graphs, and sufficient conditions are established for the existence of given edge-colored subgraphs [4, 5, 9]. The strict color number conditions for the existence of rainbow C_4's in edge-colored complete graphs and complete bipartite graphs were given by Alon [1], and Axenovich, Jiang, and Kündgen [2], respectively. Wu, Broersma, Zhang and Li [11] extended these results to general edge-colored graphs and bipartite graphs.

Let D and H be two digraphs. The *rainbow number* $rb(D, H)$ is the minimum integer such that every arc-colored digraph D^C with $c(D^C) \geq rb(D, H)$ contains a rainbow copy of H. If D is the complete digraph $\overleftrightarrow{K_n}$, then we will use $rb(n, H)$ instead of $rb(\overleftrightarrow{K_n}, H)$. Li, Zhang and Li [8] determined the rainbow number of $\overrightarrow{C_3}$ in $\overleftrightarrow{K_n}$ and characterized the corresponding extremal arc-colorings of $\overleftrightarrow{K_n}$. They also gave the sum of arc number and color number to ensure the existence of rainbow $\overrightarrow{C_3}$'s in arc-colored digraphs D^C. Furthermore, they found that the corresponding extremal digraph is $\overleftrightarrow{K_n}$ for $n = 3, 4$. They also conjectured that if an arc-colored digraph D^C of order $n \geq 5$ satisfies $a(D) + c(D^C) \geq n(n-1) + rb(n, \overrightarrow{C_3}) - 1$ and contains no rainbow $\overrightarrow{C_3}$'s, then $D \cong \overleftrightarrow{K_n}$. This conjecture had been disproved by Yang and Wu [12]. Motivated by this result, we give the sum of arc number and color number to ensure the existence of rainbow $\overrightarrow{C_4}$'s in arc-colored digraphs and we characterize the corresponding extremal arc-colored digraphs.

2 Preliminaries

Let D^C be an arc-colored digraph. For a color $i \in C(D^C)$, we use D^i to denote the spanning subdigraph of D induced by all arcs with color i. For a color set $I \subseteq C(D^C)$, we use D^I to denote the spanning subdigraph of D induced by all arcs with colors in I. For a vertex $v \in V(D)$, if all the arcs with color i are incident with v, then we call i a color *saturated* by v. The *color saturated degree* of v in D^C, denoted by $d^s_D(v)$ (or $d^s(v)$ for short), is the number of colors saturated by v. Two arc-colored digraphs $D_1^{C_1}$ and $D_2^{C_2}$ are *isomorphic* if there are bijections $\sigma : V(D_1) \to V(D_2)$ and $\rho : C_1(D_1) \to C_2(D_2)$ such that: (1) $uv \in A(D_1)$ if and only if $\sigma(u)\sigma(v) \in A(D_2)$; and (2) if $uv \in A(D_1)$, then $\rho(C_1(uv)) = C_2(\sigma(u)\sigma(v))$. If two arc-colored digraphs $D_1^{C_1}$ and $D_2^{C_2}$ are isomorphic, then we write $D_1^{C_1} \cong D_2^{C_2}$. For two vertex sets X and Y, an arc *from X to Y* is one with tail in X and head in Y, and an arc *between X and Y* is one either from X to Y or from Y to X. If D_1 and D_2 are two digraphs, then $D_1 \overrightarrow{\nabla} D_2$ denotes the digraph obtained from $D_1 \cup D_2$ by adding all possible arcs from $V(D_1)$ to $V(D_2)$. We will set $D_1 \overrightarrow{\nabla} D_2 \overrightarrow{\nabla} D_3 = D_1 \overrightarrow{\nabla} (D_2 \overrightarrow{\nabla} D_3)$. An *empty digraph* is one in which no two vertices are adjacent, that is, one whose arc set is empty. We denote by Φ_n an empty digraph of order n. Let H be a digraph of order n. We denote by $R(H)$ the arc-colored complete digraphs of order n with a rainbow subdigraph isomorphic to H and all other arcs have a single new color.

Let H_1, H_2, H_3 be three digraphs. We denote by $D(H_1, H_2, H_3)$ the arc-colored digraphs D^C on $V(H_1) \cup V(H_2) \cup V(H_3)$ such that

(1) D contains all possible arcs except that from $V(H_3)$ to $V(H_1)$; and

(2) all the arcs in $H_1 \overrightarrow{\nabla} H_2 \overrightarrow{\nabla} H_3$ have distinct colors and all other arcs (except that from $V(H_3)$ to $V(H_1)$) have a single new color.

We remark that in the above definition H_1 or H_3 is possibly a null digraph (denoted by Φ_0), in which case $D(H_1, H_2, H_3) = R(H_2 \overrightarrow{\nabla} H_3)$ or $R(H_1 \overrightarrow{\nabla} H_2)$.

The Turán number $ex(n, H)$ for a digraph H is defined as the maximum number of arcs in a digraph of order n which is H-free. Let $EX(n, H)$ be the family of all H-free extremal digraphs of order n. Sotteau and Wojda [10] characterize all the extremal digraphs of order $n \geq 5$ without $\overrightarrow{P_4}$'s.

Theorem 1 (Sotteau and Wojda [10]). *For $n \geq 5$, $ex(n, \overrightarrow{P_4}) = \lfloor n^2/3 \rfloor$, and $D \in EX(n, \overrightarrow{P_4})$ if and only if one of the following is true:*

(1) *$n = 5$ and $D \in \{\overleftrightarrow{K_{1,4}},\ \overleftrightarrow{K_2} \cup \overleftrightarrow{K_3},\ \Phi_1 \overrightarrow{\nabla} 2\overleftrightarrow{K_2},\ 2\overleftrightarrow{K_2} \overrightarrow{\nabla} \Phi_1,\ \Phi_3 \overrightarrow{\nabla} \overleftrightarrow{K_2},\ \overleftrightarrow{K_2} \overrightarrow{\nabla} \Phi_3,\ \Phi_2 \overrightarrow{\nabla} (\Phi_1 \cup \overleftrightarrow{K_2}),\ (\Phi_1 \cup \overleftrightarrow{K_2}) \overrightarrow{\nabla} \Phi_2\}$.*

(2) *$n = 6$ and $D \in \{2\overleftrightarrow{K_3},\ \Phi_2 \overrightarrow{\nabla} 2\overleftrightarrow{K_2},\ 2\overleftrightarrow{K_2} \overrightarrow{\nabla} \Phi_2\}$;*

(3) *$n = 7$ and $D \in \{\Phi_3 \overrightarrow{\nabla} 2\overleftrightarrow{K_2},\ 2\overleftrightarrow{K_2} \overrightarrow{\nabla} \Phi_3\}$; or*

(4) *$n \geq 5$ and $D \in \{\Phi_r \overrightarrow{\nabla} \Phi_s \overrightarrow{\nabla} \Phi_t : n = r + s + t$ and $r, s, t \in \{\lfloor n/3 \rfloor, \lceil n/3 \rceil\}\}$.*

3 Main results

In this paper, we obtain the exact value of $rb(n, \overrightarrow{C_4})$.

Theorem 2. *Let $n \geq 4$ be an integer. Then*

$$rb(n, \overrightarrow{C_4}) = \begin{cases} 9, & \text{if } n = 4; \\ 11, & \text{if } n = 5; \\ \lfloor n^2/3 \rfloor + 2, & \text{if } n \geq 6. \end{cases}$$

We also characterize the extremal arc-colorings for rainbow $\overrightarrow{C_4}$'s in $\overleftrightarrow{K_n}$.

Construction 1. *We define arc-colored digraphs R'_4, R'_5, R'_6 as follows. Let*

(1) *R'_4 be the arc-colored complete digraph on $\{x, y_1, y_2, y_3\}$ with two directed triangles $C_1 = y_1y_2y_3y_1$, $C_2 = y_1y_3y_2y_1$ such that $C(a_1) = C(a_2)$ if and only if both a_1 and a_2 are in C_1 or are in C_2;*

(2) *R'_n ($n = 5, 6$) be the arc-colored complete digraph on $X \cup Y$, where $X = \{x_1, x_2\}$, $Y = \{y_1, \ldots, y_{n-2}\}$, with three directed paths $P_i = x_1y_ix_2$, $i = 1, \ldots, n-2$, such that $C(a_1) = C(a_2)$ if and only if both a_1 and a_2 are in P_i for some $i = 1, \ldots, n-2$, or neither a_1 nor a_2 is an arc between X and Y;*

Construction 2. *Let $\mathcal{R}_n$, $n \geq 4$, be the class of arc-colored complete digraphs of order n such that $D^C \in \mathcal{R}_n$ if and only if one of the following is true:*

(1) *$n \in \{4, 5, 6\}$ and $D^C \cong R'_n$; or*

(2) *$n \geq 6$, and $D^C \cong R(H)$ for some $H \in EX(n, \overrightarrow{P_4})$.*

Theorem 3. *Let $f(n) = 8$ if $n = 4$, $f(n) = 10$ if $n = 5$ and $f(n) = \lfloor n^2/3 \rfloor + 1$ if $n \geq 6$. If D^C is an arc-colored complete digraph of order n such that $c(D^C) \geq f(n)$ and D^C contains no rainbow $\overrightarrow{C_4}$'s, then $c(D^C) = f(n)$ and $D^C \in \mathcal{R}_n$.*

Construction 3. *Let $\mathcal{D}_n$, $n \geq 4$, be the class of arc-colored digraphs of order n such that $D^C \in \mathcal{D}_n$ if and only if one of the following is true:*

(1) *$n \in \{4, 5, 6\}$, and $D^C \cong R'_n$;*

(2) *$n = 6$, and $D^C \cong R(2\overleftrightarrow{K_3})$;*

(3) *$n \in \{6, 7\}$, and $D^C \in \{D(\overleftrightarrow{K_2}, \Phi_{n-4}, \overleftrightarrow{K_2})\} \cup \{D(\Phi_{r_1}, 2\overleftrightarrow{K_2}, \Phi_{r_2}) : r_1 + r_2 = n - 4\}$; or*

(4) *$n \geq 6$, and $D^C \in \{D(\Phi_{r_1}, \Phi_s \overrightarrow{\nabla} \Phi_t, \Phi_{r_2}) : r_1 + r_2 + s + t = n$ and $r_1 + r_2,\ s,\ t \in \{\lfloor n/3 \rfloor, \lceil n/3 \rceil\}\}$.*

Theorem 4. *Let D^C be an arc-colored digraph of order $n \geq 4$ such that $a(D) + c(D^C) \geq n(n-1) + rb(n, \overrightarrow{C_4}) - 1$. If D^C contains no rainbow $\overrightarrow{C_4}$'s, then $a(D) + c(D^C) = n(n-1) + rb(n, \overrightarrow{C_4}) - 1$ and $D^C \in \mathcal{D}_n$.*

Sketch of the proof of Theorem 3.

Suppose (for a contradiction) that Theorem 3 is false and choose a counterexample with $c(D^C)$ as small as possible. Firstly, we prove that $c(D^C) = \lfloor n^2/3 \rfloor + 1$. Next we obtain a lower bound of $d^s(v)$ for any vertex $v \in V(D)$ by induction hypothesis. Consequently, we get a contradiction by analysing the construction of $\mathcal{R}_{n-1}$.

Sketch of the proof of Theorem 4.

Suppose (for a contradiction) that Theorem 4 is false and choose a counterexample with the smallest number of vertices, and then with the smallest number of arcs. Firstly, we obtain that D^C contains two nonsymmetric arcs with a same color by $ex(n, \overrightarrow{C_4})$. This implies that we have $a(D) + c(D^C) = a(\overleftrightarrow{K_n}) + rb(n, \overrightarrow{C_4}) - 1$ by induction hypothesis. Next we get a lower bound of $d(v) + d^s(v)$ for any vertex $v \in V(D)$ by induction hypothesis. Consequently, we get a contradiction from $\sum_{v \in V(D)}(d(v) + d^s(v)) \leq 2a(D) + 2c(D^C) - 1$ for $n \geq 7$, and we consider the cases $n = 4, 5, 6$.

References

[1] N. Alon, On a conjecture of Erdős, Simonovits and Sós concerning anti-Ramsey theorems, J. Graph Theory, 7 (1983), 91–94.

[2] M. Axenovich, T. Jiang and A. Kündgen, Bipartite anti-Ramsey numbers of cycles, J. Graph Theory, 47 (2004), 9–28.

[3] J. Bang-Jensen and G. Gutin, Digraphs, Theory, Algorithms and Applications, Springer, (2001).

[4] R. Čada, A. Kaneko, Z. Ryjáček and K. Yoshimoto, Rainbow cycles in edge-colored graphs, Discrete Math., 339 (2016), 1387–1392.

[5] L. Ding, J. Hu, G. Wang and D. Yang, Properly colored short cycles in edge-colored graphs, European J. Combin., 100 (2022), 103436.

[6] P. Erdős, M. Simonovits and V.T. Sós, Anti-Ramsey theorems, Colloquia Mathematica Societatis János Bolyai, 10. Infinite and finite sets, Keszthely (Hungary), (1973).

[7] B. Li, B. Ning, C. Xu and S. Zhang, Rainbow triangles in edge-colored graphs, European J. Combin., 36 (2014), 453–459.

[8] W. Li, S. Zhang and R. Li, Rainbow triangles in arc-colored digraphs, Discrete Appl. Math., 314 (2022), 169–180.

[9] B. Ning and J. Ge, Rainbow C_4's and directed C_4's: The bipartite case study, Bull. Malays. Math. Sci. Soc., 39 (2016), 563–570.

[10] D. Sotteau and A. P. Wojda, Digraphs without directed path of length two or three, Discrete Math., 58 (1986), 105–108.

[11] F. Wu, H. Broersma, S. Zhang and B. Li, Properly colored and rainbow C_4's in edge-colored graphs, J. Graph Theory, 105 (2024), 110–135.

[12] W. Yang and B. Wu, Disproof of a conjecture on the rainbow triangles in arc-colored digraphs, Discrete Appl. Math., 324 (2023), 99–102.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A Purely Geometric Variant of the Gale–Berlekamp Switching Game

(Extended abstract)

Adrian Dumitrescu* Jeck Lim†‡ János Pach§‡ Ji Zeng¶‡

Abstract

We introduce the following variant of the Gale–Berlekamp switching game. Let P be a set of n noncollinear points in the plane, each of them having weight +1 or −1. At each step, we pick a line ℓ passing through at least two points of P, and switch the sign of every point $p \in P \cap \ell$. The objective is to maximize the total weight of the elements of P. We show that one can always achieve that this quantity is at least $n - o(n)$, as $n \to \infty$, and at least $n/3$, for every n. Moreover, these guarantees can be attained by a polynomial time algorithm.

1 Introduction

The Gale–Berlekamp switching game (introduced in the 1960s) can be described as follows. Consider a $\sqrt{n} \times \sqrt{n}$ array of n lights. Each light has two possible states, *on* (or +1) and *off* (−1, respectively). To each row and to each column of the array there is a switch. Turning a switch changes the state of each light in that row or column. Given an initial state of the board, i.e., a certain on or off position for each light in the array, the goal is to turn on as many lights as possible, or equivalently, maximize the number of lights that are on minus the number of lights that are off.

By a *configuration* we shall mean the state of the machine, i.e., precisely which lights are on and off. Following Beck and Spencer [3], define the *signed discrepancy* of a configuration as

*Algoresearch L.L.C., Milwaukee, WI, USA. E-mail: ad.dumitrescu@algoresearch.org. Supported by ERC Advanced Grant no. 882971 "GeoScape".

†California Institute of Technology, Pasadena, CA, USA. E-mail: jlim@caltech.edu. Partially supported by an NUS Overseas Graduate Scholarship.

‡Partially supported by NSF Grant No. DMS-1928930 while in residence at the Simons Laufer Mathematical Sciences Institute in Berkeley, California, during the Spring 2025 semester.

§Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: pach@renyi.hu. Partially supported by NKFIH (National Research, Development and Innovation Office) grant K-131529, ERC Advanced Grant no. 882971 "GeoScape", and by the Eisenbud Professorship.

¶Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: jzeng@ucsd.edu. Supported by ERC Advanced Grants no. 882971 "GeoScape" and no. 101054936 "ERMiD".

the number of lights that are on minus the number of lights that are off. Equivalently, this is just the sum of the weights +1 and −1 corresponding to the on and off lights, respectively.

The asymptotic behavior of the original version of the game, played on the $\sqrt{n} \times \sqrt{n}$ square board, was studied by Brown and Spencer [5] and revisited by Spencer [19]. They showed that one can always achieve a signed discrepancy of $\Omega(n^{3/4})$ and that the order of magnitude of this bound cannot be improved. Interestingly enough, both the upper and the lower bound are obtained by the probabilistic method [1].

In this note, we consider the following, purely geometric, variant of the switching game. The board is any noncollinear n-element point set P in the plane, equipped with a switch for each connecting line. A *configuration* is a pair (P, w), where w is an assignment of ± 1 weights to the elements of P. Denote by $|w|$ the sum of the weights $\sum_{p \in P} w(p)$.

Let $F(P, w_0)$ denote the largest signed discrepancy that can be achieved by applying a series of switches to the *initial configuration* (P, w_0), that is, let

$$F(P, w_0) = \max_{w} |w|, \tag{1}$$

where the maximum is taken over all weight assignments $w : P \to \{+1, -1\}$ that can be obtained starting with the initial assignment w_0. Finally, define

$$F(P) = \min_{w_0} F(P, w_0), \tag{2}$$

where w_0 runs through all initial weight assignments.

We show that, regardless of the initial assignment, if n is large enough, then a skilled player can always achieve a linear signed discrepancy cn, for any constant c arbitrarily close to 1 and independent of P. Further, one could do this by only pulling a linear number of switches out of possibly a quadratic number.

Theorem 1. *For any noncollinear n-element point set P in the plane, consider a board where there is a switch for each connecting line.*

Then we have $F(P) \geq n - o(n)$, i.e., in the corresponding game, one can always turn on at least $n - o(n)$ lights. Moreover, the lower bound can be achieved by an application of $O(n)$ switches.

Our next theorem below gives a concrete bound that holds for every n. Its proof can be found in [9].

Theorem 2. *For any noncollinear n-element point set P in the plane, consider a board where there is a switch for each connecting line.*

Then we have $F(P) \geq n/3$, i.e., in the corresponding game, one can always turn on at least $2n/3 - 1$ lights. Moreover, the lower bound can be achieved by an application of $O(n)$ switches.

Notice that if n is odd and all but one of the points of P are collinear, then we have $F(P) \leq n - 2$. Indeed, if initially an odd number of lights are off (-1), then one cannot turn all lights on, because no switch can change the parity of the off lights. Actually, it is easy to see that for this board, we have $F(P) = n - 2$. On the other hand, if n is even and all but one of the points of P are collinear, then one can turn on all the lights. See Fig. 1.

A key concept in the proofs of our results is that of an *ordinary line*. Given a point set P, a connecting line is called *ordinary* if it contains precisely two points of P. It will be evident

Figure 1: Two warm-up board examples.

from the proof of Theorem 1 that repeatedly computing the "ordinary line graph" of the point set yields a polynomial-time algorithm that achieves $F(P) \geq (1-\varepsilon)n$ for n sufficiently large depending on $\varepsilon > 0$. Similarly, repeatedly computing a line incident to the maximum number of points and reducing the current size by 3 in the proof of Theorem 2 yields a cubic-time algorithm.

Some interesting questions remain:

Problem 3. *Given a noncollinear n-element point set in the plane with assigned ± 1 weights, can one efficiently compute a sequence of switches that produce a weight assignment of maximum signed discrepancy?*

Problem 4. *Is there an absolute constant c such that given any plane noncollinear n-element point set P, one can achieve the maximum signed discrepancy $F(P) \geq n - c$?*

2 Proof of Theorem 1

To prove Theorem 1, we introduce the notion of *ordinary line graph.* The ordinary line graph associated with a point set P is defined as an auxiliary graph G whose vertex set $V(G)$ consists of all points in P, and two points are connected by an edge if and only if they determine an ordinary line in P. The following lemma will be handy for us. The proof is omitted, but can be found in [9].

Lemma 5. *Let G be the ordinary line graph associated to a point set P in the plane, and u be a vertex inside a connected component H of G.*

Then any configuration (P, w) can be transformed into another configuration (P, w') through a series of line switches such that $w'(v) = +1$ for all v in H except u (i.e., $w'(u)$ could be either ± 1), and $w'(v) = w(v)$ for all v not in H. Furthermore, the number of line switches needed is at most $|V(H)| - 1$.

Another ingredient of our proof is the following seminal result of Green and Tao (Theorem 1.5 in [11]). Here, the statement involves the group structure on nonsingular real points on an irreducible cubic curve (see Section 2 in [11]); this group structure, however, will not be needed in our proof of Theorem 1.

Theorem 6 (Green-Tao)**.** *Suppose P is a finite set of n points in the projective real plane $\mathbb{RP}^2$. Let $K > 0$ be a real parameter. Suppose that P spans at most Kn ordinary lines. Suppose also that $n \geq \exp\exp(CK^C)$ for some sufficiently large absolute constant C.*

Then, after applying a projective transformation if necessary, P differs by at most $O(K)$ points (which can be added or deleted) from an example of one of the following three types:

I. $n - O(K)$ points on a line.

II. For some $m = n/2 + O(K)$, the set

$$X_{2m} = \left\{[\cos \frac{2\pi j}{m}, \sin \frac{2\pi j}{m}, 1] : 0 \le j < m\right\} \cup \left\{[-\sin \frac{\pi j}{m}, \cos \frac{\pi j}{m}, 0] : 0 \le j < m\right\}.$$

III. A coset $H + h$, where $3h \in H$, of a finite subgroup H of the nonsingular real points on an irreducible cubic curve, with H having cardinality $n + O(K)$.

Now we are ready to prove Theorem 1 which follows from the next statement.

Proposition 7. *For any real number $\varepsilon > 0$, there exists a constant n_ε such that any initial configuration (P, w_0) on a noncollinear point set P in the plane can be transformed into a configuration (P, w) such that $|w| \ge (1-\varepsilon)|P| - n_\varepsilon$ after $O(|P|)$ line switches.*

Proof. The proof is by induction on $|P|$ denoted as n. For $n < n_\varepsilon$ where n_ε is to be chosen later, this statement is vacuously true. With foresight, we choose $K = \lceil 2/\varepsilon \rceil$. Let G be the ordinary line graph associated with P, then our proof is divided into two cases.

Case 1: the largest connected component H of G has a size at least K. If $P \setminus V(H)$ is not all on a line, we can apply the inductive hypothesis to conclude that after $O(|P| - |V(H)|)$ line switches, the sum of weights of $P \setminus V(H)$ is at least $(1-\varepsilon)(|P| - |V(H)|) - n_\varepsilon$. If $P \setminus V(H)$ is collinear, there exists $p \in P$ not on the unique line determined by $P \setminus V(H)$. Applying switches using lines determined by p and points in $P \setminus V(H)$, the sum of weights of $P \setminus V(H)$ can be $|P| - |V(H)|$, which is even better than the previous case. Now we can apply Lemma 5 to guarantee that all but at most one point in H are assigned with +1 without affecting any points in $P \setminus V(H)$ using at most $|V(H)|$ line switches. By the choice of K, the sum of weights of H is at least $|V(H)| - 2 \ge (1-\varepsilon)|V(H)|$, hence the sum of weights of P is at least $(1-\varepsilon)(|P| - |V(H)|) - n_\varepsilon + (1-\varepsilon)|V(H)| = (1-\varepsilon)|P| - n_\varepsilon$.

Case 2: the largest connected component of G has size less than K. Then, each vertex in G has degree less than K, so the number of edges of G is $< Kn$. Hence, P determines fewer than Kn ordinary lines. Set $n_\varepsilon = \exp\exp(CK^C)$ for C as in Theorem 6. For $n \ge n_\varepsilon$, Theorem 6 implies that, after a projective transformation, $P \subset \mathbb{R}^2 \subset \mathbb{RP}^2$ differs by at most $O(K)$ points (which can be added or deleted) from one of the three types described in Theorem 6.

After deleting $O(K)$ points from P, we can obtain a subset Q satisfying the following properties depending on the types in Theorem 6: In Type I, Q consists of a set of collinear points and one more point not on this line. In Type II, Q consists of a set of collinear points (on the line at infinity) and another set of points on the unit circle. In Type III, any line intersects Q in at most three points (because Q is on an irreducible cubic curve). We conclude the inductive process using the following claim.

Claim 8. *Any initial configuration (Q, w_0) on a noncollinear point set Q, where at most one line intersects Q in more than three points, can be transformed into a configuration (Q, w) with $|w| \ge |Q| - 2$ after $O(|Q|)$ line switches.*

The rest of the proof is devoted to this claim. Let G denote the ordinary line graph associated with the given Q. Our proof of the claim is by induction on the number of connected components of G. In the base case, G has only one component and Lemma 5 gives us what we want.

For the inductive process, we can take the largest component H of G. Notice that $|V(H)| \ge 2$ is guaranteed by the Sylvester–Gallai theorem (see, for example, [11] Theorem 1.1).

Also, there exists a point $u \in V(H)$ not on the potential line intersecting Q in more than three points (otherwise there won't be any edges within H). If $Q \setminus V(H)$ is noncollinear, we can perform $O(|Q| - |V(H)|)$ line switches to make sure that at most one point in $Q \setminus V(H)$ is assigned weight -1 by the inductive hypothesis. If $Q \setminus V(H)$ is a collinear set, we can achieve the same effect by switching the lines passing through a point in $V(H)$ not collinear with $Q \setminus V(H)$. Then, we can apply Lemma 5 to make sure that all points in H other than u are assigned weight +1. Now at most two points in Q could be assigned weight -1, and we only need to deal with the case where there are exactly two such points. In that case, these two points are u and some $v \in Q \setminus V(H)$. Switch the line through u, v. Because this line contains the point u, it only intersects Q in at most three points. Hence this switch reduces the number of -1's assigned to Q to at most one, concluding the proof. □

3 Remarks

1. Recent developments in the theory of ordinary lines along with some interesting applications can be found in [4, 8, 10, 12, 15, 20]. For more details on this subject and its history, the reader can consult [7].

2. The Gale–Berlekamp switching game can be regarded as a problem in coding theory. A configuration is a 0-1 vector c of length n, where each coordinate corresponds to a light. If the light is on, we write a 0, and if it is off, we write a 1. A switch vector s is a 0-1 vector with 1's at the positions corresponding to the lights we want to switch. Let $s_1, \ldots, s_k$ be the switch vectors, and let S be the *linear code* consisting of all partial sums over $\mathbb{Z}_2$. (These partial sums are the code words.) The covering radius of the code is the minimum number $r = r(S)$ such that every configuration c is at a Hamming distance at most r from one of the elements of S. If c is at a distance δ from a code word, it means that starting from the configuration c, we can switch on all but δ lights. For the computational complexity of the Gale–Berlekamp code, see [13, 16]. A generalization of the coding approach to signed graphs was introduced and studied by Solé and Zaslavski [18].

3. Brualdi and Meyer [6] studied the variant of the switching game where each light has k different positions that change cyclically modulo k. The same problem for higher dimensional arrays was addressed in [2]. See also Schauz [17].

4. It is equally natural to consider the version of the Gale–Berlekamp game where we want to *minimize* the absolute value of the discrepancy of the configuration, that is, to balance the numbers of on and off lights as much as possible. Leo Moser conjectured that on a $k \times k$ array, for any initial configuration, there is a sequence of row and column switches that yields a configuration with discrepancy 0, 1, or 2. This was proved by Komlós and Sulyok [14] for all sufficiently large k and then by Beck and Spencer [3] for all k. Similarly, one can ask: What is the minimum absolute value of the discrepancy that can be achieved (for any initial configuration) in our purely geometric variant? We answer this question as follows, whose proof can be found in [9].

Proposition 9. *Any initial configuration (P, w_0) on a noncollinear point set P can be transformed into a configuration (P, w) with $|w| \in \{-2, -1, 0, 1, 2\}$ after $O(|P|)$ line switches in the geometric Gale–Berlekamp game.*

References

[1] Noga Alon and Joel Spencer, *The Probabilistic Method*, fourth edition, Wiley, New York, 2016.

[2] Gustavo Araújo and Daniel Pellegrino, A Gale–Berlekamp permutation-switching problem in higher dimensions, *European Journal of Combinatorics* **77** (2019), 17–30.

[3] József Beck and Joel Spencer, Balancing matrices with line shifts, *Combinatorica* **3(3-4)** (1983), 299–304.

[4] Thomas Boys, Claudiu Valculescu, and Frank de Zeeuw, On the number of ordinary conics, *SIAM Journal on Discrete Mathematics* **30(3)** (2016), 1644-1659.

[5] Thomas A. Brown and Joel H. Spencer, Minimization of ±1 matrices under line shifts, *Colloquium Mathematicum* **1(23)** (1971), 165–171.

[6] Richard A. Brualdi and Seth A. Meyer, A Gale–Berlekamp permutation-switching problem, *European Journal of Combinatorics* **44** (2015), 43–56.

[7] Vašek Chvátal, *The Discrete Mathematical Charms of Paul Erdős*, Cambridge University Press, New York, 2021.

[8] Quentin Dubroff, A better bound for ordinary triangles, Preprint, 2018, `arXiv:1805.06954`.

[9] Adrian Dumitrescu, Jeck Lim, János Pach, and Ji Zeng, A Purely Geometric Variant of the Gale–Berlekamp Switching Game, Preprint, 2025, `arXiv:2502.16305`.

[10] Radoslav Fulek, Hossein N. Mojarrad, Márton Naszódi, József Solymosi, Sebastian Stich, and May Szedlák, On the existence of ordinary triangles, *Computational Geometry: Theory and Applications* **66** (2017), 28–31.

[11] Ben Green and Terence Tao, On sets defining few ordinary lines, *Discrete & Computational Geometry* **50(2)** (2013), 409–468.

[12] Zeye Han, A Note on the weak Dirac conjecture, *Electron. J. Comb.* **24(1)** (2017), #P1.63.

[13] Marek Karpinski and Warren Schudy, Linear time approximation schemes for the Gale–Berlekamp game and related minimization problems, *Proc. 41st annual ACM Symp. Theory of Comp. (STOC 2009)*, 313–322.

[14] János Komlós and Miklós Sulyok, On the sum of elements of ±1 matrices, in: *Combinatorial Theory and Its Applications (Erdős et al., eds.)*, North-Holland 1970, 721–728.

[15] Michael S. Payne and David R. Wood, Progress on Dirac's conjecture, *Electron. J. Comb.* **21(2)** (2014), #P2.12.

[16] Ron M. Roth and Krishnamurthy Viswanathan, On the hardness of decoding the Gale–Berlekamp code, *IEEE Transactions on Information Theory* **54(3)** (2008), 1050–1060.

[17] Uwe Schauz, Anti-codes in terms of Berlekamp's switching game. *Electron. J. Combin.* **19** (2012), #P10.

[18] Patrick Solé and Thomas Zaslavsky, A coding approach to signed graphs, *SIAM Journal on Discrete Mathematics* **7(4)** (1994), 544–553.

[19] Joel Spencer, *Ten Lectures on the Probabilistic Method*, SIAM, 1994.

[20] Frank de Zeeuw, Spanned lines and Langer's inequality, Preprint, 2018, `arXiv:1802.08015`.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The Borodin–Kostochka conjecture and correspondence coloring

(Extended abstract)

Zdeněk Dvořák* Ross J. Kang† David Mikšaník‡

Abstract

Borodin and Kostochka in 1977 conjectured that, if a graph G has maximum degree $\Delta(G) \geq 9$ and its clique number satisfies $\omega(G) \leq \Delta(G) - 1$, then its chromatic number satisfies $\chi(G) \leq \Delta(G) - 1$. We prove this statement with respect to a stronger graph coloring parameter, the correspondence chromatic number χ_{DP}, provided the maximum degree is taken sufficiently large. Specifically, we prove that, if $\Delta(G) \geq 10^{19}$ and $\omega(G) \leq \Delta(G) - 1$, then $\chi_{DP}(G) \leq \Delta(G) - 1$. This strengthens earlier results of Choi, Kierstead and Rabern (2023) and of Reed (1999), which were for the list chromatic number and chromatic number, respectively.

1 Introduction

Here is a common refrain in combinatorics: *How do local constraints influence global structure?*

A classic result of this flavour in graph coloring is the celebrated theorem of Brooks [8], which implies for any graph G that, if $\Delta(G) \geq 3$ and $\omega(G) \leq \Delta(G)$, then $\chi(G) \leq \Delta(G)$, where $\Delta(G)$, $\omega(G)$, and $\chi(G)$ denote the maximum degree, clique number, and chromatic number of G, respectively. In other words, a mild local condition ($\omega(G) \leq \Delta(G)$) gives improved global structure ($\chi(G) \leq \Delta(G)$) over what is known in the general case ($\chi(G) \leq \Delta(G) + 1$). Since $\omega(G) \leq \chi(G)$, this bound on $\chi(G)$ is sharp. Moreover, the condition $\Delta(G) \geq 3$ is best possible due to the odd cycles.

*Computer Science Institute, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. E-mail: rakdver@iuuk.mff.cuni.cz. Supported by the ERC-CZ project LL2328 of the Ministry of Education of Czech Republic.

†Korteweg–de Vries Institute for Mathematics, University of Amsterdam, the Netherlands. E-mail: r.kang@uva.nl. Partially supported by grant OCENW.M20.009 of the Dutch Research Council (NWO) and the Gravitation Programme NETWORKS (024.002.003) of the Dutch Ministry of Education, Culture and Science (OCW).

‡Computer Science Institute, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. E-mail: miksanik@iuuk.mff.cuni.cz. Supported by the the project GAUK 40324 of the Charles University Grant Agency.

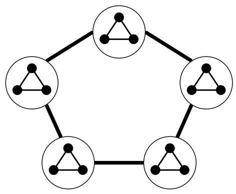


Figure 1: The graph $C_5 \boxtimes C_3$.

Vizing [35] asked for other extremal bounds on $\chi(G)$ given stronger conditions on $\omega(G)$. The extreme case $\omega(G) = 2$ is the most well-studied (see e.g. [6, 11, 29, 25, 30, 4, 17, 24, 27]) and is a difficult and influential challenge. While several other regimes of $\omega(G)$ in terms of $\Delta(G)$ have gained attention (see e.g. [26, 5, 18, 23]), our focus is the other non-trivial extreme case for Vizing's problem, i.e. $\omega(G) = \Delta(G) - 1$. The Borodin–Kostochka conjecture [6] is the main objective in this case.

Conjecture 1 ([6])**.** *If $\Delta(G) \geq 9$ and $\omega(G) \leq \Delta(G) - 1$, then $\chi(G) \leq \Delta(G) - 1$.*

The graph $C_5 \boxtimes C_3$, i.e. the strong product of the cycle on five vertices with the triangle, demonstrates that the condition $\Delta(G) \geq 9$ is necessary; see Figure 1. The most substantial progress is due to Reed [34] who showed that a (much) larger lower bound condition on $\Delta(G)$ is sufficient.

Theorem 1 ([34])**.** *If $\Delta(G) \geq 10^{14}$ and $\omega(G) \leq \Delta(G) - 1$, then $\chi(G) \leq \Delta(G) - 1$.*

One approach towards understanding Conjecture 1 better has been to weaken the statement somewhat. As one example, Cranston and Rabern [16] showed the bound $\chi(G) \leq \Delta(G) - 1$ holds under the assumption that $\Delta(G) \geq 13$ and $\omega(G) \leq \Delta(G) - 3$. As another example, the conjecture has been confirmed under the restriction of G to various non-trivial graph classes (see e.g. [19, 12, 14] to name a few).

Another approach is to consider strengthened or more general versions of Conjecture 1. Since the conjecture is (hypothetically) best possible, this approach leads us to consider stronger coloring parameters, the most natural of which is the *list chromatic number* χ_ℓ (defined below). Choi, Kierstead and Rabern [13] proved a list analogue of Theorem 1, meaning every vertex v of G may draw its color only from its own private list $L(v)$ of $\Delta(G) - 1$ colors, with the assignment L of lists assumed worst-case/adversarial.

Theorem 2 ([13])**.** *If $\Delta(G) \geq 10^{20}$ and $\omega(G) \leq \Delta(G) - 1$, then $\chi_\ell(G) \leq \Delta(G) - 1$.*

We continue along these lines by considering a further strengthened coloring parameter, the *correspondence chromatic number* χ_{DP}, where the list assignment L is "even more adversarial", in the sense that the sets of pairwise conflicts between colors in the lists of adjacent vertices may be assumed worst-case/adversarial. This notion was introduced by Dvořák and Postle [20] to solve a list coloring problem for planar graphs, and it lets us test the extent to which coloring techniques ensure enough global structure so as to be "oblivious" to the colors. We now define correspondence coloring formally, and then (list) coloring as a special case.

Given a graph G, a *correspondence assignment* for G is a pair (L, Π), where $L\colon V(G) \to 2^{\mathbb{N}}$ is a *list assignment* for G and Π is a function assigning to each edge $\{u, v\} \in E(G)$ a matching between $\{u\} \times N(u)$ and $\{v\} \times N(v)$. For $k \in \mathbb{N}$, we say that (L, Π) is *k-correspondence* *assignment* for G if $|L(v)| \geq k$ for every $v \in V(G)$. An (L,Π)-*coloring* of G is a function $\varphi\colon V(G) \to \mathbb{N}$ such that (i) $\varphi(v) \in L(v)$ for every $v \in V(G)$ and (ii) $\{(u,\varphi(u)),(v,\varphi(v))\} \notin \Pi(e)$ for every $e = \{u,v\} \in E(G)$. We say that G is (L,Π)-*colorable* if there exists an (L,Π)-coloring of G. Furthermore, we say G is k-*correspondence colorable* if G is (L,Π)-colorable for every k-correspondence assignment (L,Π) for G. Finally, the *correspondence chromatic number* $\chi_{DP}(G)$ of G is defined as the smallest k such that G is k-correspondence colorable. If we restrict ourselves to the assignments (L,Π) for which $\Pi(e)$ is the identity matching for every $e \in E(G)$, we have list coloring and the list chromatic number $\chi_\ell(G)$ of G. If we furthermore restrict ourselves so that L is a uniform list assignment, we have the classical coloring and the chromatic number $\chi(G)$ of G. It is straightforward to check that $\omega(G) \leq \chi(G) \leq \chi_\ell(G) \leq \chi_{DP}(G) \leq \Delta(G)+1$ for all graphs G.

We first note how these coloring parameters diverge significantly in behavior for graphs G with $\chi(G) = 2$. It is a classic result of Erdős, Rubin and Taylor [21] that $\chi_\ell(K_{n,n}) \sim \log_2 n$ as $n \to \infty$. In contrast, it was observed in [28, 3] that $\chi_{DP}(G) = \Omega(\Delta/\log\Delta)$ for any Δ-regular graph G (see also [10]).

On the other hand, $\chi(G) = \chi_\ell(G) = \chi_{DP}(G)$ for all graphs G with $\omega(G) = \Delta(G)+1$. More interestingly, since the corollary of Brooks' theorem mentioned at the beginning also holds for list and correspondence coloring (see [38] for a short proof), the parameters also coincide for graphs G with $\omega(G) = \Delta(G)$ and $\Delta(G) \geq 3$. The condition $\Delta(G) \geq 3$ is necessary for equality of these parameters in this case, which can be seen by considering the even cycles. The following question, which may also be interpreted as a correspondence coloring analogue of the Borodin–Kostochka conjecture, naturally arises.

Question 1. *What is the smallest number Δ_0 such that every graph G with $\Delta(G) \geq \Delta_0$ and $\omega(G) = \Delta(G) - 1$ satisfies $\chi(G) = \chi_\ell(G) = \chi_{DP}(G)$?*

The following result, our main contribution, implies that $\Delta_0 \leq 10^{19}$.

Theorem 3. *If $\Delta(G) \geq 10^{19}$ and $\omega(G) \leq \Delta(G) - 1$, then $\chi_{DP}(G) \leq \Delta(G) - 1$.*

At a high level, we proceed analogously to the proof of Theorem 1 in [34]. We also use some ideas from the proof of Theorem 2 in [13], which itself shares many similarities with the proof of Theorem 1 in [34]. As in the proofs of the aforementioned theorems, we consider a hypothetical minimal counterexample G to Theorem 3 and an arbitrary $(\Delta(G)-1)$-correspondence assignment (L,Π), and show that G is (L,Π)-colorable, thereby proving that no counterexample to Theorem 3 exists. The basic idea is to partially color G using the so-called *naive coloring procedure*—which colors each vertex $v \in V(G)$ by a uniformly chosen color from $L(v)$ and then resolves all conflicts by uncoloring the vertices involved—and then argue that the resulting partial (L,Π)-coloring of G extends to the whole graph G with positive probability. We carry this out in two steps. In the first step, we obtain structural information about G, which we use in the second step to guide the probability analysis. Our second step is conceptually very similar to the corresponding second step in the proof of Theorem 1, which is considerably easier than the corresponding second step in the proof of Theorem 2, because we have managed to obtain structural information about G analogous to what Reed obtained in his proof of Theorem 1. This can be consider as the main technical contribution of the paper, which we describe in Subsection 2.1. In particular, we show that no vertex outside a clique of size $\Delta(G) - 1$ in G is adjacent to more than 1000 vertices in the clique, whereas Choi, Kierstead and Rabern showed *only* that no such vertex is adjacent to more than 4 vertices with the *same* list in the clique.

Although we work with correspondence coloring, all our arguments hold for list coloring up to minor modifications. Thus, all structural results obtained in our work apply to minimal counterexamples to Theorem 2—some of these results are novel, including the one mentioned in the previous paragraph. Finally, we remark that we did not put much effort in minimizing the lower bound $\Delta(G) \geq 10^{19}$. We provide further proof details in Section 2, but before that we need to introduce more definitions.

More definitions. We write $H \leq G$ to denote that H is an induced subgraph of G, and $H < G$ if $H \leq G$ and $H \neq G$. Given a correspondence assignment (L, Π) for G, $H \leq G$, and $e = \{a, b\} \in E(H)$, we say that a color α at a *blocks* a color β at b if $\{(a, \alpha), (b, \beta)\} \in \Pi(e)$. For a function $\psi\colon V(H) \to \mathbb{N}$ such that $\psi(v) \in L(v)$ for every $v \in V(H)$, instead of writing "$\psi(v)$ at v" we simply write "$\psi(v)$". For example, $\psi(b)$ blocks α at a should be interpreted as $\psi(b)$ at b blocks α at a. Furthermore, we say that $\psi(b)$ *blocks no color at* a if there is no color $\alpha \in L(a)$ such that $\psi(b)$ blocks α at a. Finally, for $\{a, c\} \in E(H)$, we say $\psi(b)$ and $\psi(c)$ *together block at most one color at* a if either $\psi(b)$ or $\psi(c)$ block no color at a, or $\psi(b)$ and $\psi(c)$ block the same color at a, that is, there exists a color $\alpha \in L(a)$ such that $\psi(b)$ blocks α at a and $\psi(c)$ blocks α at a.

Given $H \leq G$, *a partial* (L, Π)*-coloring* of G (on H) is a function $\psi\colon V(H) \to \mathbb{N}$ such that (i) $\psi(v) \in L(v)$ for every $v \in V(H)$ and (ii) $\{(u, \psi(u)), (v, \psi(v))\} \notin \Pi(e)$ for every $e = \{u, v\} \in E(H)$. A partial (L, Π)-coloring ψ of G on H naturally induces a correspondence assignment (L_ψ, Π_ψ) for $G - H$, where $L_\psi(v) := L(v) \setminus \{\alpha \mid \exists u \in N(v) : \psi(u) \text{ blocks } \alpha \text{ at } v\}$ and $\Pi_\psi(\{u, v\})$ is the restriction of $\Pi(\{u, v\})$ to the sets $\{u\} \times L_\psi(u)$ and $\{v\} \times L_\psi(v)$. In this context, we refer to $L_\psi(v)$ as the set of *available* colors for v. Moreover, for $F \leq G$, we say that ψ *extends* to F if there exists a partial (L, Π)-coloring ψ' of G on $H \cup F$ such that $\psi'(v) = \psi(v)$ for every $v \in V(H)$. For a set of vertices $A \subseteq V(G)$, with a slight abuse of notation, we say that ψ extends to A rather than to $G[A]$. Furthermore, when $A = \{v\}$, we simply say that ψ extends to v.

For a graph B with $V(B) \cap V(G) = \emptyset$, $G \vee B$ denotes the complete join of G and B, which is obtained by taking the disjoint union of G and B, and adding edge between every vertex of G and every vertex of B.

2 Overview of proof of Theorem 3

In this section, we provide a more detailed overview of the proof of Theorem 3 (and we defer most of the proof to the full version of the paper). Recall that G is a hypothetical minimal counterexample to Theorem 3 with maximum degree $\Delta := \Delta(G)$, that is, G is not $(\Delta - 1)$-correspondence colorable, but every proper subgraph H of G is $(\Delta(H) - 1)$-correspondence colorable. Furthermore, recall that (L, Π) is an arbitrary $(\Delta - 1)$-correspondence assignment for G. To begin, we construct a partial (L, Π)-coloring ψ of G using the naive coloring procedure:

(i) For each vertex $v \in V(G)$, choose a color $\alpha \in L(v)$ uniformly at random and let $\psi(v) = \overline{\psi}(v) := \alpha$.

(ii) For each edge $e = \{u, v\} \in E(G)$, if $\overline{\psi}(u)$ blocks $\overline{\psi}(v)$, then undefine $\psi(u)$ and $\psi(v)$.

We call $\overline{\psi}(v)$ the *tentative* color of v and say that v *retains its color* if $\psi(v) = \overline{\psi}(v)$.

Our aim is to show that ψ extends to the entire graph G with positive probability, in which case G is $(\Delta - 1)$-correspondence colorable and hence no counterexample to Theorem 3 exists. Deciding whether a partial coloring extends to the entire graph is a difficult problem in general. Fortunately, it turns out that we can show that ψ extends to G greedily with positive probability.

With this in mind, we define a vertex $v \in V(G)$ to be *ψ-good* if $N(v)$ contains two disjoint pairs $\{a, b\}$ of vertices, each retaining its color, such that $\psi(a)$ and $\psi(b)$ together block at most one color at v. Note that if v is uncolored and ψ-good, then ψ extends to v. Indeed, in this case, v has more available colors than the number of uncolored neighbors, and so there exists at least one available color in $L_\psi(v)$ which we can use to extend ψ to v. Thus, if we could prove that all uncolored vertices are ψ-good with positive probability, then we can extend ψ to G greedily. Whether v is ψ-good depends only on the tentative colors of vertices within distance two from v, which means, by the Lovász Local Lemma [22], it suffices to prove that v is ψ-good with sufficiently large probability, namely, at least $1 - 1/4\Delta^4$. Although we are able to prove it only for some of the vertices, this turns out to be enough for extending ψ to G greedily, as we explain next.

Intuitively, a vertex v with many non-edges in $N(v)$ is ψ-good with large probability, as each non-edge in $N(v)$ is a reasonable candidate for a pair of neighbors of v whose retained colors block at most one color at v. In fact, if the number of non-edges in $N(v)$ is at least $\Delta^{1.9}$, we can show that v is ψ-good with probability at least $1 - 1/4\Delta^5$ (we required the exponent to be at least 5 for the reasons explained in the following paragraph). We call vertices with this many non-edges in their neighborhood *sparse*, and the others *dense*.

Dealing with dense vertices is more challenging and for that we need to obtain some structural information about G. One is a *sparse–dense* decomposition of G, which is a partition $\{D_1, D_2, \ldots, D_k, S\}$ of $V(G)$ such that S is a set of sparse vertices and each D_i is a "near-clique" of size almost Δ with the property that no vertex outside of D_i is adjacent to many vertices in D_i. Although we are not able to guarantee that ψ extends to a dense vertex with the necessary probability, we can prove that D_i contains two ψ-good vertices that are adjacent to every vertex in D_i with probability at least $1 - 1/4\Delta^5$. This implies that we can extend ψ to D_i greedily, coloring the two ψ-good vertices last. As before, the Lovász Local Lemma ensures that ψ extends to every sparse vertex in S and every D_i with positive probability. More precisely, we define D_i to be *ψ-happy* if it contains two ψ-good vertices that are adjacent to every vertex in D_i. Since whether D_i is ψ-happy depends on the tentative colors of vertices within distance two from D_i, it suffices to show that every D_i is ψ-happy with probability at least $1 - 1/4\Delta^5$ and every sparse vertex v in S is ψ-good with probability at least $1 - 1/4\Delta^5$. This finishes the overview.

The key place in our proof where we need new ideas to overcome the additional complications that come from the stronger notion of correspondence coloring is in how we prove that D_i is ψ-happy with the necessary probability. As we described in the introduction, we proceed conceptually as in the proof of Theorem 1. In particular, we obtain structural information about minimal counterexamples to Theorem 3 analogous to what Reed obtained in his proof of Theorem 1, which we describe in Subsection 2.1. This allows us to prove that D_i is ψ-happy with probability at least $1 - 1/4\Delta^5$ as in Reed's proof of Theorem 1, but here we have to overcome the following obstacle: changing the tentative color of a single vertex v, say from α to β, can drastically change how many neighbors of v retain their colors, even if the neighbors are colored differently from α and β. Although the probability of this happening is very small, it is a problem when proving concentration results using the standard concentration inequalities. Fortunately, Bruhn and Joos developed a variation of Talagrand's inequality that effectively deals with such situations (see [9, Lemma 17]).

2.1 Structural information about minimal counterexamples to Theorem 3

Since G is a hypothetical minimal counterexample to Theorem 3, for every $H < G$, there exists a partial (L, Π)-coloring ψ of G on H. Observe that $G - H$ is not (L_ψ, Π_ψ)-colorable. In particular, every vertex in G has degree either $\Delta - 1$ or Δ. Moreover, for every $v \in V(G-H)$ we have $|L_\psi(v)| \geq \deg_{G-H}(v) - 1$. This motivates us to make the following definition. A graph H is d_1-*correspondence colorable* if it is (L', Π')-colorable for every correspondence assignment (L', Π') with $|L'(v)| \geq \deg_H(v) - 1$ for every $v \in V(H)$. By the observations above, no induced subgraph of G is d_1-correspondence colorable. For this reason, studying d_1-correspondence colorable graphs is particularly useful. The following lemma, which is crucial to our arguments and may also be of independent interest, gives a useful condition for d_1-correspondence colorability.

Lemma 1. *For $t \geq 6$, if B is a graph with $\omega(B) \leq |B| - 2$, then $K_t \vee B$ is d_1-correspondence colorable.*

We remark that the analogous result is known to hold for list coloring (see [15, Lemma 3.3]).

We now precisely define a *sparse–dense* decomposition of G, following closely the definition in [13]. Recall that vertex $v \in V(G)$ is called *sparse* if $|\overline{E(N(v))}| \geq \Delta^{1.9}$, and *dense* otherwise. A subset of vertices $A \subseteq V(G)$ is a *near-clique* if $\omega(A) \geq |A| - 1$.

Definition 1. *A* sparse–dense decomposition of G *is a partition $\mathcal{D} = \{D_1, D_2, \ldots, D_k, S\}$ of $V(G)$ with the following properties:*

(i) every vertex in S is sparse,

(ii) D_i is a near-clique with $|D_i| \geq \Delta - 8\Delta^{0.9} + 2$ and minimum degree at least $\lceil 3\Delta/4 \rceil$, and

(iii) $|N(v) \cap D_i| \leq \lceil 3\Delta/4 \rceil$ for every $v \in V(G) \setminus D_i$.

Lemma 2. *There exists a sparse–dense decomposition of G.*

The proof of Lemma 2 is almost identical to its counterpart in the proof of Theorem 2 (see Lemma 4.3 in [13]), except that we have to prove the other prerequisite lemmas for correspondence coloring, such as Lemma 1. Let $\mathcal{D} = \{D_1, D_2, \ldots, D_k, S\}$ be a sparse–dense decomposition guaranteed by Lemma 2. Moreover, for each $i \in [k]$, let $C_i \subseteq D_i$ be a maximal clique in D_i, $d_i \in D_i$ such that $D_i = C_i \cup \{d_i\}$, and $K_i := C_i \cap N(d_i)$. Finally, for $A \subseteq V(G)$ and $a \in V(G)$, let $N_A(a) := N(a) \cap A$.

We recall that the primary reason we can show that a sparse vertex is ψ-good with probability at least $1 - 1/4\Delta^5$ is that there are many non-edges in its neighborhood. Perhaps surprisingly, for the same reason, we can show that D_i is ψ-happy with probability at least $1 - 1/4\Delta^5$. Although there might be dense vertices in D_i with only a few non-edges in their neighborhood, Lemma 3 asserts that K_i contains a large set A of vertices, each having a relatively large number of non-edges in its neighborhood. It follows that the probability that $v \in A$ is ψ-good is relatively large, and since there are many vertices in A, we can prove that there are two ψ-good vertices with probability at least $1 - 1/4\Delta^5$, implying that D_i is ψ-happy with probability at least $1 - 1/4\Delta^5$.

Lemma 3. *There exist at least $\Delta/1009$ disjoint triples $\{v, a, b\}$, where $v \in K_i$ and $a, b \in N(v) \setminus C_i$ are two distinct vertices such that $|N_{C_i}(a) \cup N_{C_i}(b)| \leq 1008$.*

We require the triples to be disjoint for the purposes of ensuring concentration.

For $|C_i| \leq \Delta - 3$, Lemma 3 follows from a simple double-counting argument using the following lemma (see Lemma 16 in [34] and its proof therein for an example of such double-counting argument).

Lemma 4. *Let $p \in \mathbb{N}$ such that $|C_i| = \Delta - p$. For every $v \in C_i$, there are at most two distinct vertices $a, b \in N(v) \setminus C_i$ such that $|N_{C_i}(a)| > p + 11$ and $|N_{C_i}(b)| > p + 11$. Furthermore, if $\deg(v) = \Delta - 1$, there is at most one such vertex in $N(v) \setminus C_i$.*

In fact, in the double-counting argument, the following corollary of Lemma 4 is crucial: if $p \geq 3$, then every vertex in C_i has two neighbors outside of C_i, each adjacent to at most $p+11$ vertices in C_i. With additional effort, we prove a slightly weaker assertion for C_i of size $\Delta - 2$ and $\Delta - 1$, but still strong enough for the same double-counting argument. Precisely, we prove that, excluding at most 17 vertices, every vertex in C_i has two neighbors outside C_i that are adjacent to at most 1008 vertices in C_i. In the remainder, we sketch the proof of this assertion for C_i of size $\Delta - 1$. The case where C_i has size $\Delta - 2$ then follows easily.

To the end, suppose that $|C_i| = \Delta - 1$. Every vertex in C_i has either one or two neighbors outside C_i, depending on its degree. Hence vertices of degree $\Delta - 1$ would be problematic if there are many of them. Fortunately, using similar ideas we introduce in the proof of Lemma 6 below, we can prove that there are at most 17 vertices in C_i of degree $\Delta - 1$. For every vertex v in C_i of degree Δ, we have to prove that both of its neighbors outside C_i are adjacent to at most 1008 vertices in C_i. This follows from the following lemmas, which are the main novelties.

Lemma 5. *If $|C_i| = \Delta - 1$, $v \in C_i$ is a vertex of degree Δ, and $a, b \in N(v) \setminus C_i$ are the two neighbors of v outside C_i, then $|N_{C_i}(a)| \leq 8$ or $|N_{C_i}(b)| \leq 8$.*

Lemma 6. *If $|C_i| = \Delta - 1$, then $|N_{C_i}(a)| \leq 1000$ for every vertex $a \in V(G) \setminus C_i$.*

Proof of Lemma 6 (sketch). Let $C := C_i$. For the sake of contradiction, assume that there exists a vertex $a \in V(G) \setminus C$ such that $|N_C(a)| > 1000$. Observe that every vertex $v \in N_C(a)$, except possibly one, has degree Δ. Let $A \subseteq N_C(a) \subseteq C$ be the set of all vertices of degree Δ, and $B := N(A) \setminus (C \cup \{a\})$ be the neighbors of vertices in A outside C different from a. By Lemma 5, every vertex in B is adjacent to at most 8 vertices in C. Hence $|B| \geq |A|/8 \geq 1000/8 \geq 125$. Since $|B| \geq 125$, by a classical Ramsey number result, B contains K_8 (a clique of size 8) or $\overline{K_4}$ (an independent set of size 4). In this proof sketch, we only consider the former case.

With a slight abuse of notation, let K_8 denote some clique of size 8 in B. Moreover, let $V(K_8) := \{y_1, y_2, \ldots, y_8\}$ and x_i be an arbitrary neighbor of y_i in C. Notice that y_i is the only neighbor of x_i in K_8. Suppose first that a is non-adjacent to every vertex in K_8. Let ψ' be a partial (L, Π)-coloring of G on $G - (C \cup \{a\})$, which exists by minimality of G, and ψ be a partial (L, Π)-coloring of G on $G - (C \cup K_8 \cup \{a\})$ obtained from ψ' by uncoloring vertices in K_8. Clearly, $|L_\psi(y_i)| \geq 7$. We claim that in fact $|L_\psi(y_i)| \geq 8$. Let us first show how this claim implies that ψ extends to G, which will be a contradiction. We color y_1 and x_2 so that they block together at most one color at x_1, y_3 and x_4 so that they block together at most one color at x_3, a and some non-neighbor a' of a in C so that they block at most one color at x_1. We color the remaining vertices greedily, leaving x_1 to last and x_3 as the second-to-last.

It remains to show that $|L_\psi(y_i)| \geq 8$ for every $y_i \in V(K_8)$, which is equivalent to the statement that no two neighbors of u_i outside C block at most one color at u_i. We can prove it by contradiction, extending ψ similarly as in the previous paragraph. We also proceed similarly if a is adjacent to some vertex in K_8. $\square$

References

[1] Noga Alon, Stijn Cambie, and Ross J. Kang. Asymmetric list sizes in bipartite graphs. *Ann. Comb.*, 25(4):913–933, 2021.

[2] Noga Alon and Michael Krivelevich. The choice number of random bipartite graphs. *Ann. Comb.*, 2(4):291–297, 1998.

[3] Anton Bernshteyn. The asymptotic behavior of the correspondence chromatic number. *Discrete Mathematics*, 339(11):2680–2692, 2016.

[4] Anton Bernshteyn. The Johansson–Molloy theorem for DP-coloring. *Random Structures Algorithms*, 54(4):653–664, 2019.

[5] Marthe Bonamy, Tom Kelly, Peter Nelson, and Luke Postle. Bounding χ by a fraction of Δ for graphs without large cliques. *J. Combin. Theory Ser. B*, 157:263–282, 2022.

[6] O. V. Borodin and A. V. Kostochka. On an upper bound of a graph's chromatic number, depending on the graph's degree and density. *J. Combinatorial Theory Ser. B*, 23(2-3):247–250, 1977.

[7] Peter Bradshaw, Bojan Mohar, and Ladislav Stacho. Bipartite graphs are $(\frac{4}{5}-\varepsilon)\frac{\Delta}{\log\Delta}$-choosable. *arXiv e-prints*, page arXiv:2409.01513, September 2024.

[8] R. L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2):194–197, 1941.

[9] Henning Bruhn and Felix Joos. A stronger bound for the strong chromatic index. *Combin. Probab. Comput.*, 27(1):21–43, 2018.

[10] Stijn Cambie and Ross J. Kang. Independent transversals in bipartite correspondence-covers. *Canad. Math. Bull.*, 65(4):882–894, 2022.

[11] Paul A. Catlin. A bound on the chromatic number of a graph. *Discrete Math.*, 22(1):81–83, 1978.

[12] Rong Chen, Kaiyang Lan, Xinheng Lin, and Yidong Zhou. Borodin–Kostochka conjecture holds for odd-hole-free graphs. *Graph. Comb.*, 40(2), February 2024.

[13] Ilkyoo Choi, H. A. Kierstead, and Landon Rabern. The list version of the Borodin-Kostochka conjecture for graphs with large maximum degree. *Discrete Math.*, 346(11):Paper No. 113300, 12, 2023.

[14] Daniel W. Cranston and Landon Rabern. Coloring claw-free graphs with $\Delta-1$ colors. *SIAM Journal on Discrete Mathematics*, 27(1):534–549, 2013.

[15] Daniel W. Cranston and Landon Rabern. Coloring a graph with $\Delta-1$ colors: Conjectures equivalent to the Borodin–Kostochka conjecture that appear weaker. *European Journal of Combinatorics*, 44:23–42, 2015.

[16] Daniel W. Cranston and Landon Rabern. Graphs with $\chi=\Delta$ have big cliques. *SIAM Journal on Discrete Mathematics*, 29(4):1792–1814, 2015.

[17] Ewan Davies, Rémi de Joannis de Verclos, Ross J. Kang, and François Pirot. Colouring triangle-free graphs with local list sizes. *Random Structures Algorithms*, 57(3):730–744, 2020.

[18] Ewan Davies, Ross J. Kang, François Pirot, and Jean-Sébastien Sereni. Graph structure via local occupancy. *arXiv e-prints*, page arXiv:2003.14361, March 2020.

[19] Medha Dhurandhar. Validity of borodin and kostochka conjecture for classes of graphs without a single, forbidden subgraph on 5 vertices. *arXiv*, 2101.01354, 2021.

[20] Zdeněk Dvořák and Luke Postle. Correspondence coloring and its application to list-coloring planar graphs without cycles of lengths 4 to 8. *Journal of Combinatorial Theory, Series B*, 129:38–54, 2018.

[21] Paul Erdős, Arthur L. Rubin, and Herbert Taylor. Choosability in graphs. In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State Univ., Arcata, Calif., 1979)*, Congress. Numer., XXVI, pages 125–157, Winnipeg, Man., 1980. Utilitas Math.

[22] Paul Erdős and Lovász László. Problems and results on 3-chromatic hypergraphs and some related questions. *Coll Math Soc J Bolyai*, 10, 01 1974.

[23] Eoin Hurley, Rémi de Joannis de Verclos, and Ross J. Kang. An improved procedure for colouring graphs of bounded local density. *Adv. Comb.*, pages Paper No. 7, 33, 2022.

[24] Eoin Hurley and François Pirot. Uniformly random colourings of sparse graphs. In *STOC'32—Proceedings of the 2023 ACM International Symposium on Theory of Computing*, pages 1357–1370. ACM, New York, 2023.

[25] Anders Johansson. Asymptotic choice number for triangle-free graphs. Technical Report 91-5, DIMACS, 1996.

[26] Anders Johansson. The choice number of sparse graphs. Manuscript, 1996.

[27] Ross J. Kang and Matthieu Rosenfeld. On Vizing's problem for triangle-free graphs. *arXiv e-prints*, page arXiv:2309.10876, September 2023.

[28] Daniel Kráľ, Ondřej Pangrác, and Heinz-Jürgen Voss. A note on group colorings. *J. Graph Theory*, 50(2):123–129, 2005.

[29] Jim Lawrence. Covering the vertex set of a graph with subgraphs of smaller degree. *Discrete Math.*, 21(1):61–68, 1978.

[30] Michael Molloy. The list chromatic number of graphs with small clique number. *J. Combin. Theory Ser. B*, 134:264–284, 2019.

[31] Michael Molloy and Bruce Reed. *Graph Colouring and the Probabilistic Method*, volume 23. 01 2001.

[32] Michael Molloy and Bruce Reed. Colouring graphs when the number of colours is almost the maximum degree. *J. Combin. Theory Ser. B*, 109:134–195, 2014.

[33] Seth Pettie and Hsin-Hao Su. Distributed coloring algorithms for triangle-free graphs. *Information and Computation*, 243:263–280, 2015.

[34] Bruce Reed. A strengthening of brooks' theorem. *Journal of Combinatorial Theory, Series B*, 76(2):136–149, 1999.

[35] V. G. Vizing. Some unsolved problems in graph theory. *Russian Mathematical Surveys*, 23(6):125–141, 1968.

[36] V. G. Vizing. Some unsolved problems in graph theory. *Uspehi Mat. Nauk*, 144(6):117–134, 1968.

[37] V. G. Vizing. Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz*, 29 Metody Diskret. Anal. v Teorii Kodov i Shem:3–10, 101, 1976.

[38] Mariusz Zaąc. A short proof of Brooks' theorem. *arXiv*, 1805.11176, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On a conjecture concerning 4-coloring of graphs with one crossing

(Extended abstract)

Zdeněk Dvořák* Bernard Lidický† Bojan Mohar‡

Abstract

We conjecture that every graph of minimum degree five with no separating triangles and drawn in the plane with one crossing is 4-colorable. We use computer enumeration to show that this conjecture holds for all graphs with at most 28 vertices, explore the consequences of this conjecture and provide some insights on how it could be proved.

1 Introduction

Famously, every planar graph is 4-colorable [1, 2]. It is natural to ask whether this statement can be strengthened. As shown in [5, 6], for every surface Σ other than the sphere, there are infinitely many minimal non-4-colorable (5-*critical*) graphs which can be drawn in Σ (a graph G is k-*critical* if $\chi(G) = k$ and every proper subgraph of G is $(k-1)$-colorable). Still, one could hope that at least for simplest non-trivial surfaces Σ (projective plane, torus), it might be possible to characterize 5-critical graphs drawn in Σ, or at least to develop a polynomial-time algorithm to test 4-colorability of graphs drawn in Σ. However, either of these goals seems far beyond our reach using the current methods.

To obtain at least some intuition, we consider "minimally nonplanar" graphs, or more precisely, graphs that can be drawn in the plane with at most one crossing; let $\mathcal{C}$ denote the class of all such graphs. Let us remark that graphs in $\mathcal{C}$ can be drawn without crossings both in the projective plane and on the torus. It is still easy to construct infinite families of 5-critical graphs in $\mathcal{C}$; however, preliminary computer-assisted investigations suggest that obtaining their characterization could be possible. In this paper, we explore in detail a natural conjecture arising from these investigations.

*Computer Science Institute, Charles University, Prague, Czech Republic. E-mail: rakdver@iuuk.mff.cuni.cz. Supported by the ERC-CZ project LL2328 (Beyond the Four Color Theorem, registered as "Realizace projektu hraničního výzkumu v oblasti teorie grafů – barevnost a návrh algoritmů") of the Ministry of Education of Czech Republic.

†Department of Mathematics, Iowa State University. Ames, IA, USA. E-mail: lidicky@iasate.edu. Supported in part by NSF grant FRG DMS-2152490.

‡Department of Mathematics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. E-mail: mohar@sfu.ca.

Of course, 5-critical graphs have minimum degree at least four. Moreover, there is a standard way to handle a vertex v of degree four, by deleting v and identifying two non-adjacent neighbors of v (which exist unless the graph is K_5). This turns the graph into a smaller graph G' that is non-4-colorable (though not necessarily 5-critical). Assuming that the edges incident with v are non-crossed, this also preserves the existence of a drawing with only one crossing. Thus, G' has a 5-critical subgraph belonging to $\mathcal{C}$. Based on this observation, one could hope to inductively prove statements about the structure of 5-critical graphs in $\mathcal{C}$ of minimum degree four. Indeed, a similar inductive argument used to handle 4-faces has been crucial in the development of the theory of 4-critical triangle-free graphs, which eventually lead to design of a linear-time algorithm for 3-colorability of triangle-free graphs on surfaces [4].

For this argument to take off, one needs to handle the basic case of 5-critical graphs in $\mathcal{C}$ of minimum degree at least five. Based on computer-assisted experiments, we believe that no such graphs exist, and thus K_5 is the sole basic case for the argument described above.

Conjecture 1. *Every 5-critical graph in $\mathcal{C}$ contains a vertex of degree four.*

Indeed, the following stronger statement seems to hold (a subgraph K of a graph G is *separating* if $G - V(K)$ is not connected).

Conjecture 2. *If a graph G does not contain separating triangles and has a drawing in the plane with one crossing such that all vertices not incident with the crossed edges have degree at least five, then G is 4-colorable.*

Let us remark that throughout the paper, the graphs are simple, without loops and parallel edges. It is easy to see that 5-critical graphs do not contain separating cliques, and thus Conjecture 2 implies Conjecture 1. The assumption that G does not contain separating triangles cannot be dropped: Every graph $G \in \mathcal{C}$ has a supergraph of minimum degree at least five in $\mathcal{C}$, obtained by gluing a copy of the icosahedron at each vertex of G. A similar argument proves the following claim.

Lemma 3. *Let $G \in \mathcal{C}$ be a graph without separating triangles, and consider any drawing of G in the plane with at most one crossing. If G is not 4-colorable and Conjecture 2 holds, then there exists a vertex $v \in V(G)$ of degree four such that all incident faces have length three and their boundaries do not contain the crossing. In particular, if Conjecture 2 holds, then every graph in $\mathcal{C}$ that does not contain the 4-wheel as a subgraph is 4-colorable.*

It is well known that it suffices to prove the Four Color Theorem for triangulations of the plane. Similarly, Conjecture 2 has the following equivalent form. Let $\mathcal{C}_0 \subseteq \mathcal{C}$ consist of graphs obtained from 4-connected plane triangulations by choosing distinct facial triangles xuv and yuv sharing an edge uv and adding the edge xy; we say that the vertices x, y, u, and v are *external* and all other vertices *internal*.

Conjecture 4. *Every non-4-colorable graph in $\mathcal{C}_0$ has an internal vertex of degree four.*

Assuming Conjecture 2 holds, we can show that every non-4-colorable graph in $\mathcal{C}_0$ can be obtained from K_5 by a sequence of simple operations (since we need to avoid creating separating triangles, this is not completely straightforward). By *expansion* in the following result, we mean one of the operations illustrated in Figure 1 (we omit the exact definitions from this extended abstract).

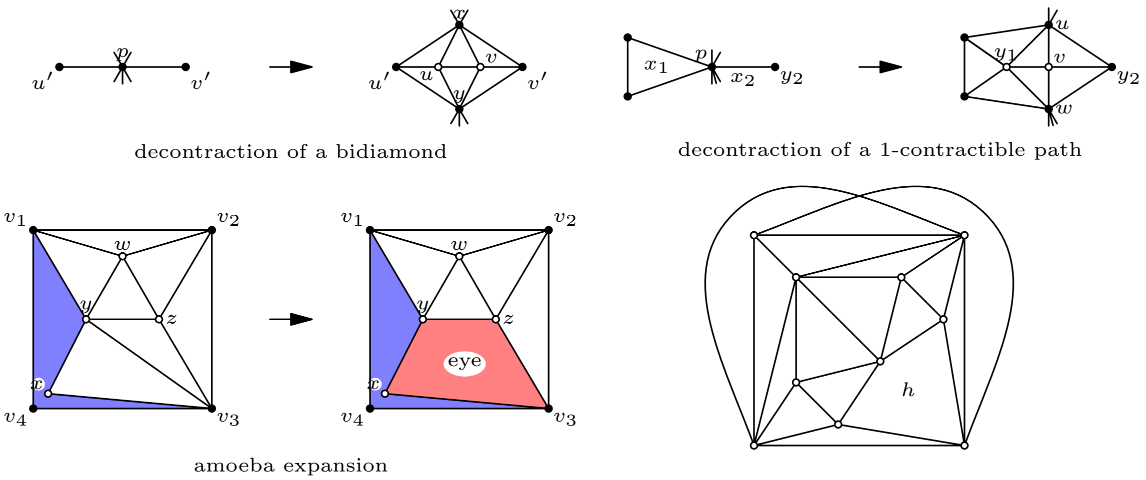


Figure 1: Expansion operations and a 5-critical graph in $\mathcal{C}$ with a 4-face h.

Theorem 5. *If Conjecture 2 holds, then for every non-4-colorable graph $G \in \mathcal{C}_0$, there exists a sequence $K_5 = G_0, G_1, \ldots, G_m = G$ of graphs such that for every $i \in \{1, \ldots, m\}$, the graph $G_i \in \mathcal{C}_0$ is non-4-colorable and is obtained from G_{i-1} by expansion.*

The expansion operation increases the number of vertices, and thus Theorem 5 can be turned into a relatively efficient algorithm to enumerate the non-4-colorable graphs in $\mathcal{C}_0$. However, it needs to be noted that there are 5-critical graphs in $\mathcal{C}$ that do not belong to $\mathcal{C}_0$; see the bottom right of Figure 1 for an example $H \in \mathcal{C}$. Consequently, Theorem 5 is somewhat less explicit than it might appear, since one of the options for the expansion operation (namely, the amoeba expansion) is by necessity essentially "in a subgraph matching the relevant part of H, replace the 4-face by any triangulation without separating triangles".

2 Reduction to precoloring extension

A plane graph H is *rainbow-forbidding* if its outer face is bounded by a 4-cycle $K = v_1v_2v_3v_4$ and no 4-coloring of H satisfies $\varphi(v_1) \neq \varphi(v_3)$ and $\varphi(v_2) \neq \varphi(v_4)$. Similarly, H is *v_1-diagonal-forbidding* if it does not have any 4-coloring satisfying $\varphi(v_1) \neq \varphi(v_3)$ and $\varphi(v_2) = \varphi(v_4)$, and *bichromatic-forbidding* if it does not have any 4-coloring satisfying $\varphi(v_1) = \varphi(v_3)$ and $\varphi(v_2) = \varphi(v_4)$, It is *diagonal-forbidding* if it is v_1-diagonal-forbidding or v_2-diagonal-forbidding. Using the Four Color Theorem, it is easy to see that the following claim holds.

Lemma 6. *A graph G is not 4-colorable and belongs to $\mathcal{C}$ if and only if there exists a rainbow-forbidding plane graph H with the outer face bounded by a 4-cycle $v_1v_2v_3v_4$ such that G is isomorphic to the graph obtained from H by adding the edges v_1v_3 and v_2v_4.*

Thus, the problem of characterizing non-4-colorable graphs in $\mathcal{C}$ is equivalent to the rainbow precoloring extension problem in plane graphs with the outer face of size four. A *canvas* is a plane graph with the outer face bounded by a cycle of length at least four and all other faces of length three. Since adding edges only makes the graph harder to color, we can furthermore reformulate Conjecture 2 in terms of rainbow-forbidding canvases. With some extra care, we can avoid creation of non-facial triangles, thus proving the equivalence with Conjecture 4. More precisely, we obtain the following result (a vertex of a plane graph is *internal* if it is not incident with the outer face, and *external* otherwise).

Theorem 7. *The following claims are equivalent for every positive integer n:*

(i) *Let G be a graph with $|V(G)| = n$ and without separating triangles. If G has a drawing in the plane with one crossing such that all vertices not incident with the crossed edges have degree at least five, then G is 4-colorable.*

(ii) *Let H' be a canvas such that all triangles in H' bound faces and $|V(H')| = n$. If H' is rainbow-forbidding, then H' has an internal vertex of degree at most four.*

(iii) *Let H be a plane graph such that all triangles in H bound faces and $|V(H)| = n$. If H is rainbow-forbidding, then H has an internal vertex of degree at most four.*

(iv) *Every non-4-colorable graph $G \in \mathcal{C}_0$ with n vertices has an internal vertex of degree four.*

In particular, to verify Conjecture 2 for graphs with n vertices, it suffices to verify the claim from Theorem 7(ii) for graphs with n vertices. For small n, this can be done by a computer-assisted enumeration, as we discuss in the following section. Let us remark that Conjecture 2 is also equivalent to the variants of (ii) and (iii) for diagonal-forbidding graphs. Interestingly, the analogous claim for bichromatic-forbidding graphs is false, with infinitely many counterexamples.

3 Generating candidates

A *candidate* is a canvas such that every triangle bounds a face and all internal vertices have degree at least five. A candidate is an *ℓ-candidate* if its outer face has length ℓ. By Theorem 7, in order to verify Conjecture 2 for graphs with at most n vertices, it suffices to enumerate the 4-candidates with at most n vertices and check that none of them is rainbow-forbidding.

We generate the 4-candidates using a variation on the approach of [3]. More precisely, we show that each 4-candidate can be generated from the *diamond* (the 4-cycle with one chord) by a sequence of reverses to the reduction operations depicted in Figure 2, so that all intermediate canvases are also 4-candidates. We implemented an algorithm based on this claim and used it to generate all non-isomorphic 4-candidates with at most 28 vertices, testing more than 10^9 graphs. None of them is rainbow-forbidding, and only the diamond is diagonal-forbidding. This confirms Conjecture 2 for graphs with at most 28 vertices.

4 How to prove Conjecture 2?

It is easy to see that Conjecture 2 is a strengthening of the Four Color Theorem. Up to minor variations, we know only one proof of the Four Color Theorem: Using Kempe chains, prove that many configurations in planar graphs are *reducible*, i.e., cannot appear in a minimal counterexample, then use discharging to show that at least one of them appears in every planar graph. It is natural to ask whether a similar approach can be used to prove Conjecture 2. This indeed seems promising, but there are several issues to overcome. In the rainbow precoloring extension setting, some of the Kempe chains may contain precolored vertices. This limits

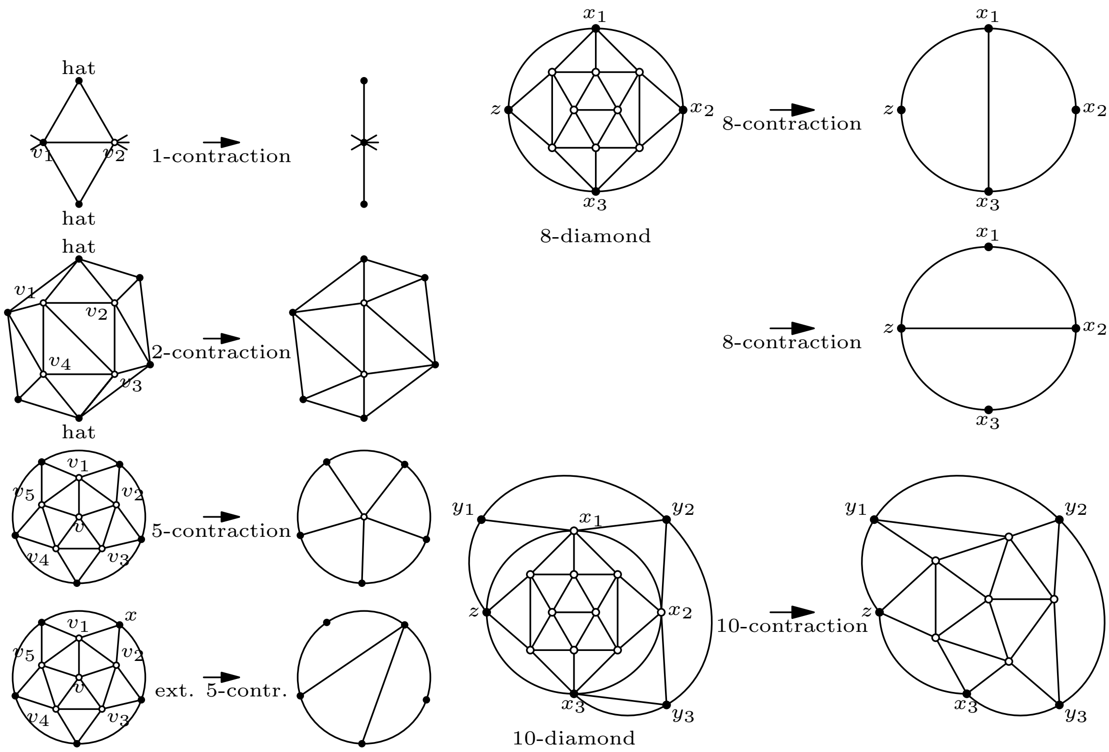


Figure 2: Reduction opertations for 4-candidates

how we can exchange colors on Kempe chains, and results in fewer of the configurations being reducible. Moreover, reducibility arguments often involve replacing a configuration by a smaller one. This could decrease degrees of vertices or create non-facial triangles, and thus we might not obtain a smaller counterexample and fail to reach a contradiction.

Using a Kempe chain argument, we can exclude separating 4-cycles from a minimal counterexample to Conjecture 2. Building further on this argument, we show that we can exclude reducible configurations which only need to be deleted (not replaced by a smaller configuration), called *D-reducible* configurations in [7], even though their deletion decreases the degrees of vertices. As shown in [8], D-reducibility arguments are sufficient to prove the Four Color Theorem. However, many of the configurations that are D-reducible in the setting of the Four Color Theorem are not D-reducible in our setting, since we get extra constraints on switching Kempe chains containing the precolored vertices (in fact, small D-reducible configurations seem to be very rare).

Realistically, to prove the conjecture, we will need to use *C-reducible* configurations, where the configuration is replaced by a smaller one. An issue that proofs of the Four Color Theorem using C-reducible configurations need to overcome is that such a replacement could result in creation of loops, making the reduced graph impossible to color. In our setting, the issue is compounded by the possibility that the reduced graph is not a candidate, i.e., that the replacement of H by H' can create internal vertices of degree at most four or non-facial triangles. One way to deal with the issue would be to find an exact (or at least sufficiently detailed) characterization of non-4-colorable graphs in $\mathcal{C}_0$ based on the assumption that the conjecture holds for graphs with less than $|V(G)|$ vertices, and then use it to argue that the reduced graph cannot be rainbow-forbidding even if it is not a candidate.

References

[1] K. Appel and W. Haken. Every planar map is four colorable, Part I: Discharging. *Illinois J. of Math.*, 21:429–490, 1977.

[2] K. Appel, W. Haken, and J. Koch. Every planar map is four colorable, Part II: Reducibility. *Illinois J. of Math.*, 21:491–567, 1977.

[3] G. Brinkmann and B. D. McKay. Construction of planar triangulations with minimum degree 5. *Discrete Mathematics*, 301:147–163, 2005.

[4] Z. Dvořák, D. Král', and R. Thomas. Three-coloring triangle-free graphs on surfaces VII. A linear-time algorithm. *Journal of Combinatorial Theory, Series B*, 152:483–504, 2022.

[5] S. Fisk. The nonexistence of colorings. *J. Combin. Theory, Ser. B*, 24:247–248, 1978.

[6] J. Hutchinson, B. Richter, and P. Seymour. Colouring Eulerian triangulations. *Journal of Combinatorial Theory, Series B*, 84:225–239, 2002.

[7] N. Robertson, D. P. Sanders, P. Seymour, and R. Thomas. The four colour theorem. *J. Combin. Theory, Ser. B*, 70:2–44, 1997.

[8] J. P. Steinberger. An unavoidable set of D-reducible configurations. *Trans. Amer. Math. Soc*, 362:6633–6661, 2010.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Entropy and the growth rate of universal covering trees

(Extended abstract)

Idan Eisner[*] Shlomo Hoory[†]

Abstract

This work studies the relation between two graph parameters, ρ and Λ. For an undirected graph G, $\rho(G)$ is the growth rate of its universal covering tree, while $\Lambda(G)$ is a weighted geometric average of the vertex degree minus one, corresponding to the rate of entropy growth for the non-backtracking random walk (NBRW). It is well known that $\rho(G) \geq \Lambda(G)$ for all graphs, and that graphs with $\rho = \Lambda$ exhibit some special properties. In this work we derive an easy to check, necessary and sufficient condition for the equality to hold. Furthermore, we show that the variance of the number of random bits used by a length ℓ NBRW is $O(1)$ if $\rho = \Lambda$ and $\Omega(\ell)$ if $\rho > \Lambda$. As a consequence we exhibit infinitely many non-trivial examples of graphs with $\rho = \Lambda$.

1 Introduction

1.1 The Main Players

This paper investigates the relation between two graph parameters Λ and ρ. Given a connected and undirected graph $G = (V, E)$ with minimal degree at least two and maximal degree at least three, we define $\Lambda(G)$ and $\rho(G)$ as follows:

- Λ is a weighted geometric mean of the vertex degrees minus one, where each vertex has weight proportional to its degree:

$$\Lambda(G) = \prod_{v \in V} (\deg(v) - 1)^{\frac{\deg(v)}{2|E|}}.$$

- ρ is the growth rate of the universal cover of G. Equivalently, ρ is defined as the largest eigenvalue of the non-backtracking adjacency matrix B, see [2, 5]. See Section 2 for an exact definition.

[*]Department of Computer Science, Tel-Hai College, Israel. E-mail: `eisnerida@telhai.ac.il`.
[†]Department of Computer Science, Tel-Hai College, Israel. E-mail: `hooryshl@telhai.ac.il`.

One should note that while the graph parameter Λ depends only on the degree distribution of the graph, the parameter ρ depends on the graph structure as well. It is well know that $\rho(G) \geq \Lambda(G)$ for all graphs, see [2, 9]. In this work we give a simple, necessary and sufficient condition for equality to hold. However, before stating the exact result, we begin with some background and motivation.

1.2 Some Motivation

The non-backtracking random walk (NBRW) on the graph G is a process that explores a graph by moving from vertex to vertex, where in each step the walk moves to a randomly chosen neighbor of the current vertex, excluding the one it just visited. It was shown by Alon et. al [1] that this extra bias towards exploration makes the walk mix faster on regular expander graphs, compared to the simple random walk (SRW) that does not employ this extra constraint. Also, it was shown that the SRW as well as the NBRW exhibit a sharp threshold in their convergence to the stationary distribution for regular Ramanujan graphs [11], as well as for various random graph models whp. Lubetzky et al. [12], proved such a result for random d-regular graphs, $\mathcal{G}(n, d)$, Ben-Hamou et al. [4], for random graphs with a prescribed degree distribution in the configuration model, and Conchon-Kerjan [7] for random lifts of a fixed based graph.

The case of a random n-lift of some base graph G is closely related to this work, as the mixing time of the NBRW is $t_s = (1+o(1)) \log_\Lambda n$ with high probability [7], while the diameter of a random n-lift is $t_d = (1 + o(1)) \log_\rho n$ whp [6]. Put differently, performing the NBRW on a random n-lift of G, the first time when all vertices have positive probability of being reached is t_d, while the first time when the walk gets close to the stationary distribution is t_s. Therefore, asking if $\rho(G) = \Lambda(G)$ is the same as asking if the two events occur at the same time, up to a $1 + o(1)$ factor.

A second motivation for the ρ vs. Λ question is the Moore bound for irregular graphs. In [2], Alon et. al proved that any girth g graph with the same degree distribution as G has at least $\Omega(\Lambda(G)^{g/2})$ vertices, while in [9], Hoory proved that any girth g graph covering the base graph G has at least $\Omega(\rho(G)^{g/2})$ vertices. Therefore, when $\rho(G) = \Lambda(G)$ the bounds coincide, and otherwise there is an exponential gap between the two bounds.

1.3 Our Results

For the graph G, define a *suspended path* as a non-backtracking path P in G that has no internal vertices of degree more than two, and where its two endpoints u, v have degree greater than two. The main result of this paper is that a necessary and sufficient condition for G to have $\Lambda = \rho$, is that the following equation holds for all suspended paths P in G:

$$(\deg(u) - 1) \cdot (\deg(v) - 1) = \Lambda(G)^{2|P|}. \tag{1}$$

Graphs satisfying $\rho = \Lambda$ include regular graphs and bipartite-biregular graphs. In addition, any such graph, where all edges are replaced by length k paths for some fixed $k > 1$ also satisfies the condition. It turns out, that the simple combinatorial characterization given in (1), can be used to construct infinitely many graphs with $\rho = \Lambda$ that are not in the above list, such as those in Figure 1.

The third result regards the random variable R_ℓ counting the number of random bits consumed by a length ℓ NBRW starting from the stationary distribution. Namely, the sum of

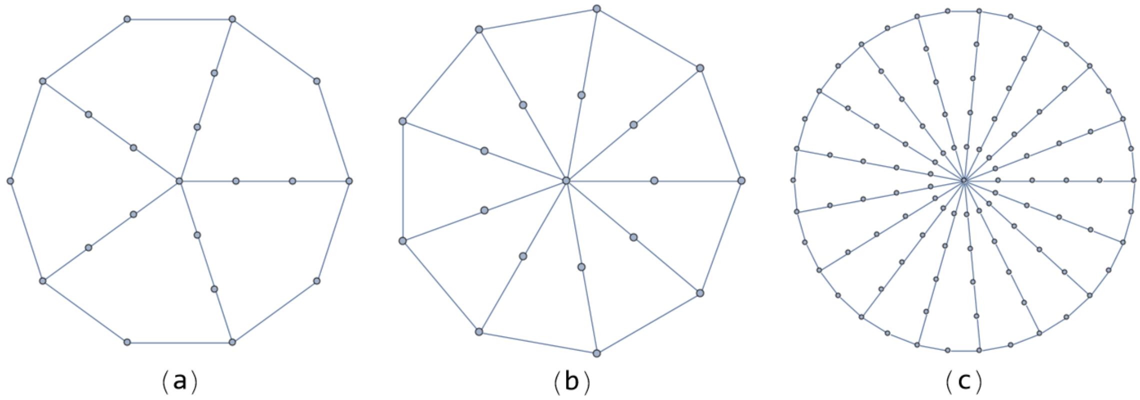


Figure 1: Three graphs with $\rho = \Lambda$ that are not regular, bipartite bi-regular, or obtained from such graphs by replacing each edge by a length k path. It can be verified that the suspended path condition holds with Λ equals $\sqrt{2}$, 2 and $\sqrt{2}$ for the graphs (a), (b) and (c), respectively.

$\log_2(\deg(v) - 1)$ on the vertices visited by the walk. It is not hard to verify that the expected value of R_ℓ is $\ell \log_2 \Lambda$. We prove that the variance of R_ℓ is bounded by a constant independent of ℓ if $\rho(G) = \Lambda(G)$, but grows linearly with ℓ if $\rho(G) > \Lambda(G)$.

2 Preliminaries

Given a simple finite undirected graph $G = (V, E)$, we regard each edge as a pair of directed edges and denote the set of directed edges by $\vec{E}$. For a directed edge $e = (u, v) \in \vec{E}$, we denote its head by $h(e) = v$ and tail by $t(e) = u$. We define $\operatorname{outdeg}(e) = \deg(h(e)) - 1$, $\operatorname{indeg}(e) = \deg(t(e)) - 1$, and the reversed edge $e' = (u, v)' = (v, u)$.

Definition 1. *Let $G = (V, E)$ be a simple undirected graph with minimal degree at least two. For two directed edges e, f, we say that one can transition from e to f if $h(e) = t(f)$ and $e' \neq f$ and denote this relation by $e \to f$. The Non-Backtracking Random-Walk (NBRW) on G is a Markov chain on the state space $\vec{E}$ with the transition matrix Π_G defined by:*

$$\pi_G(e, f) = \begin{cases} \frac{1}{\operatorname{outdeg}(e)}, & \text{if } e \to f \\ 0, & \text{otherwise} \end{cases}$$

We regard the state $e = (u, v)$ as being at v coming from u. Let B denote the non-backtracking adjacency matrix, which is an $\vec{E}$ by $\vec{E}$ 0-1 matrix where $B_{e,f} = 1$ if $e \to f$.

Remark 1. *The definition generalizes to allow for multiple edges and self loops in G, where a self loop can be either a whole-loop or a half-loop. Each edge, except for the half-loops, gives rise to two edges in $\vec{E}$ each being the inverse of the other. A half-loop in G gives rise to a single edge in $\vec{E}$, which is its own inverse.*

Proposition 1. *For any graph $G = (V, E)$ with minimal degree at least two, the uniform edge distribution ν_s defined by $\nu_s(e) = 1/|\vec{E}|$ for all $e \in \vec{E}$ is stationary for the Markov chain Π_G.*

Proposition 2 ([8] proposition 3.3)**.** *The NBRW Markov chain on G as well as the non-backtracking adjacency matrix B are irreducible iff G is connected with* $\operatorname{mindeg}(G) \geq 2$ *and* $\operatorname{maxdeg}(G) > 2$.

We say that a graph G is *NB-irreducible* if it is connected with $\mathrm{mindeg}(G) \geq 2$ and $\mathrm{maxdeg}(G) > 2$.

Given a connected graph G with minimal degree at least 2 and some arbitrary vertex $v_0 \in V(G)$, we define its universal cover $\tilde{G}$. It is an infinite tree, where the vertex set $V(\tilde{G})$ is the set of all finite non-backtracking walks from v_0, and where two vertices of $\tilde{G}$ are adjacent if one walk extends the other by a single step. Its root is the empty walk ϵ, and it can be verified that changing v_0 yields an isomorphic tree.

Definition 2. *Given a connected graph G with minimal degree at least 2 let:*

- $\mathcal{B}_{\tilde{G},r}(v)$ *denote the radius r ball around the vertex v in $\tilde{G}$,*
- $\Omega_{e,\ell}$ *be the set of length ℓ non-backtracking walks in G starting from $e \in \vec{E}$,*
- $\Omega_\ell = \cup_{e\in\vec{E}}\Omega_{e,l}$ *be the set of all length ℓ non-backtracking walks in G.*

The growth rate of $\tilde{G}$, denoted $\rho(G)$, is defined by one of the following alternative definitions, given by the following:

Proposition 3 ([3, 9])**.** *Given an NB-irreducible graph G, the following definitions for $\rho(G)$ are equivalent:*

$$\begin{aligned}\rho(G) &= \lim_{r\to\infty} |\mathcal{B}_{\tilde{G},r}(v)|^{1/r} = \lim_{\ell\to\infty} |\Omega_{e,\ell}|^{1/\ell} = \lim_{\ell\to\infty} |\Omega_\ell|^{1/\ell} \\ &= \lim_{\ell\to\infty} \left(\nu_s \; B^\ell \; \bar{1}\right)^{1/\ell} = \rho(B),\end{aligned}$$

where B denotes the non-backtracking adjacency matrix, and $\rho(B)$ is the largest (Perron) eigenvalue of B. Equality holds regardless of the choice of $v \in V(\tilde{G})$ and $e \in \vec{E}$, where applicable.

Definition 3. *The average growth rate $\Lambda(G)$ is the rate predicted by the stationary distribution ν_s of the NBRW:*

$$\Lambda(G) = \Big(\prod_{e\in\vec{E}} \mathrm{outdeg}(e)\Big)^{1/|\vec{E}|} = \prod_{v\in V(G)} \Big(\deg(v) - 1\Big)^{\deg(v)/|\vec{E}|}. \tag{2}$$

3 The Main Results

Definition 4. *A length ℓ non-backtracking path $P = (e_0, e_1, \ldots, e_{\ell-1})$ in an NB-irreducible graph G is a* **suspended path** *if* $\mathrm{outdeg}(e_i) = \mathrm{indeg}(e_{i+1}) = 1$ *for $i = 0, \ldots, l-2$ and* $\mathrm{indeg}(e_0), \mathrm{outdeg}(e_{l-1}) > 1$*. Denote $|P| = \ell$,* $\mathrm{outdeg}(P) = \mathrm{outdeg}(e_{\ell-1})$ *and* $\mathrm{indeg}(P) = \mathrm{indeg}(e_0)$*.*

Note that the suspended paths of G form a partition of $\vec{E}(G)$.
The following is our main Theorem:

Theorem 1. *For an NB-irreducible graph G the following three conditions are equivalent:*

1. $\rho(G) = \Lambda(G)$.

2. $\text{outdeg}(P) \cdot \text{indeg}(P) = \Lambda(G)^{2|P|}$ *for every suspended path P in G.*

3. $\prod_{e \in C} \text{outdeg}(e) = \Lambda(G)^{|C|}$ *for every non-backtracking cycle C.*

In addition, we have the following result regarding the variance of the number of random bits for the NBRW. Consider the random variable R_ℓ, counting the number of random bits used by a length ℓ NBRW on G, starting from the stationary distribution ν_s. Formally, for $\omega = (e_0, e_1, \ldots, e_\ell) \in \Omega_\ell$ we define $R_\ell(\omega) = \sum_{i=0}^{\ell-1} \log_2 \text{outdeg}(e_i)$. The following theorem states that the variance of R_i exhibits a dichotomy.

Theorem 2. *Given an NB-irreducible graph G, then:*

$$\text{var}[R_\ell] = \begin{cases} O(1) & \text{if } \rho = \Lambda \\ \Theta(\ell) & \text{if } \rho > \Lambda. \end{cases}$$

We end the extended abstract with proof sketches:

Proof sketch for Theorem 1. We first demonstrate that $\rho = \Lambda$ implies the cycle condition. This is done by introducing M_t as an element-wise multiplicative interpolation of Π_G and B, then proving that $\rho(M_t) = \rho^t$. This proof leverages $\rho(M_t)$'s log-convexity, Kingman [10], and an Alon et al. [2] inspired lower bound. Consequently, the log-convexity requirement holding as equality implies, per Nussbaum [13], that $\Pi_G = kD^{-1}BD$ for some positive scalar k and positive diagonal matrix D, yielding the cycle condition.

Next, we prove that violating the suspended path condition implies violating the cycle condition. We define $g(P) = (\log \text{indeg}(P) + \log \text{outdeg}(P))/(2|P|)$ for a suspended path P. The weighted average of g over all suspended paths as well as the expected value of $f(e) = (\log \text{indeg}(e) + \log \text{outdeg}(e))/2$ for a uniformly random edge $e \in \vec{E}$ is shown to be $\log \Lambda$. This enables an iterative procedure that bans suspended paths or connected components with the smallest f-average, ultimately yielding a cycle that violates its condition.

The last step in the proof, is to prove that the suspended path condition implies $\rho = \Lambda$, which is done by showing that the product of the $\text{outdeg}(e)$ along any length ℓ non-backtracking path is Λ^ℓ up to a $\Theta(1)$ multiplicative error. $\square$

Proof sketch for Theorem 2. If $\rho = \Lambda$, we show that the suspended path condition implies that the deviation $(R_\ell(\omega) - \ell \log \Lambda)^2$ is upper bounded by a constant independent of ω, for any length ℓ non-backtracking path ω. Furthermore, $R_\ell(\omega)$ is upper bounded by $O(\ell)$ through the central limit theorem for Markov chains or direct derivation.

The most interesting part of the proof is lower bounding $R_\ell(\omega)$ by $\Omega(\ell)$ when $\rho > \Lambda$. We achieve this by demonstrating that the cycle condition implies the existence of two cycles C_1 and C_2, of the same length, sharing a common edge f, but with different edge out-degree product. We then use an exposure argument revealing a constant portion of the edges of a length ℓ non-backtracking walk ω. Since $R_\ell(\omega)$ conditioned on this exposure is the sum of independent random variables for its segments, $\text{var}[R_\ell]$ is the sum of their variances. We show that, with high probability, a constant portion of these segments are length $|C_1|$ walks starting and ending at f. As the corresponding variables have non-zero variance, the required result follows. $\square$

References

[1] N. Alon, I. Benjamini, E. Lubetzky, and S. Sodin. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics*, 9(04):585–603, 2007.

[2] N. Alon, S. Hoory, and N. Linial. The moore bound for irregular graphs. *Graphs and Combinatorics*, 18(1):53–57, 2002.

[3] O. Angel, J. Friedman, and S. Hoory. The non-backtracking spectrum of the universal cover of a graph. *Transactions of the American Mathematical Society*, 367(6):4287–4318, 2015.

[4] A. Ben-Hamou and J. Salez. Cutoff for nonbacktracking random walks on sparse random graphs. *The Annals of Probability*, 45(3):1752–1770, 2017.

[5] C. Bordenave, M. Lelarge, and L. Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular ramanujan graphs. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1347–1357. IEEE, 2015.

[6] G. Conchon-Kerjan. *Sparse random graphs: from local specifications to global phenomena.* PhD thesis, Université de Paris/Université Paris Diderot (Paris 7), 2021.

[7] G. Conchon-Kerjan. Cutoff for random lifts of weighted graphs. *Annals of Probability*, 50(1):304–338, 2022.

[8] C. Glover. *The Non-Backtracking Spectrum of a Graph and Non-Backtracking PageRank.* Brigham Young University, 2021.

[9] S. Hoory. On the girth of graph lifts. *arXiv preprint arXiv:2401.01238*, 2024.

[10] J. F. Kingman. A convexity property of positive matrices. *The Quarterly Journal of Mathematics*, 12(1):283–284, 1961.

[11] E. Lubetzky and Y. Peres. Cutoff on all ramanujan graphs. *Geometric and Functional Analysis*, 26(4):1190–1216, 2016.

[12] E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Mathematical Journal*, 153(3):475–510, 2010.

[13] R. D. Nussbaum. Convexity and log convexity for the spectral radius. *Linear Algebra and its Applications*, 73:59–122, 1986.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Froböse percolation in the hypercube

(Extended abstract)

Joshua Erde* Max Gutkin† Mihyun Kang‡ Dominik Schmid§

Abstract

Bootstrap percolation is a process in which an initially infected set of vertices in a graph spreads the infection to uninfected vertices according to some rule. In r-neighbour bootstrap percolation, a vertex becomes infected if it has at least r infected neighbours. A central question in the study of bootstrap percolation concerns the likely behaviour of a randomly selected set of initially infected vertices. This version of the problem has been studied extensively in a variety of lattice-like graphs. In particular, Balogh and Bollobás [2] identified a threshold function for 2-neighbour bootstrap percolation in the hypercube Q_n. A related bootstrap percolation variant, Froböse percolation, has also been studied on lattice-like graphs, particularly the grid. In this paper, we examine the process of Froböse percolation in Q_n and use our findings to obtain a threshold result.

1 Introduction

1.1. Motivation and background. Bootstrap percolation, originally introduced in 1979 by Chalupa, Leath, and Reich [9] as a tool in the study of magnetic materials, is known for its wide variety of applications including studying the spread of information in a social network [13], cellular automata [12], and a variety of topics in the physical sciences [1].

In the *r-neighbour bootstrap percolation process* an infection spreads through the vertices of a graph in a series of discrete time steps according to some locally defined update rule. Starting with some set A_0 of *initially infected vertices*, the infected set grows in each time step, giving a sequence $A_0 \subseteq A_1 \subseteq A_2 \subseteq \dots$ where

$$A_i = A_{i-1} \cup \{v \in V(G) \, : \, |N(v) \cap A_{i-1}| \geq r\}.$$

*Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: erde@math.tugraz.at.

†Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: gutkin@math.tugraz.at.

‡Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: kang@math.tugraz.at.

§Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: schmid@math.tugraz.at.

That is, a vertex becomes infected once it has at least r infected neighbours. The process terminates once $A_i = A_{i-1}$ for some $i \in \mathbb{Z}^+$. We say that such a set A_{i-1} is *closed*, and setting $\overline{A} := \cup_{i=0}^{\infty} A_i$, we refer to $\overline{A}$ as the *closure* of A. In the case that $\overline{A} = V(G)$ we say that A *percolates*, or *spans* G.

When considering bootstrap percolation on a graph or family of graphs, it is often of interest to study the 'typical' behaviour of the process. Since the process is monotone, as the initially infected set increases in density, percolation becomes more likely. It is therefore of interest to study the transition of the probability of percolation as the density of the set increases. More precisely, given a graph G and probability $p \in (0,1)$, let $\mathcal{A}_p$ be a randomly chosen p-dense subset of $V(G)$. That is, each vertex of G is included in $\mathcal{A}_p$ independently and with probability p. We are then interested in

$$\mathbb{P}\,(G,p) := \mathbb{P}[\,\mathcal{A}_p \text{ percolates}\,].$$

Given a family of graphs $(G_n)_{n\geq 0}$ and a probability function $p = p(n)$, there is interest in studying the asymptotic behaviour of $\mathbb{P}(G_n, p(n))$ as $n \to \infty$. Note that since the percolation process is monotone, $\mathbb{P}(G_n, p(n))$ is increasing in p. A classic result of Bollobás and Thomason [7] then implies that there must be a *threshold function* for percolation in this model. That is, there exists a function $p_c(n)$ such that $\frac{p(n)}{p_c(n)} = o(1)$ implies $\lim_{n\to\infty} \mathbb{P}(G_n, p(n)) = 0$ and $\frac{p(n)}{p_c(n)} = \omega(1)$ implies $\lim_{n\to\infty} \mathbb{P}(G_n, p(n)) = 1$.

Threshold functions for r-neighbour bootstrap percolation have been studied on a variety of families of graphs, with a significant focus on the d-dimensional grid $[n]^d$ (see, e.g., [3, 4, 8]) and the n-dimensional hypercube Q_n [2, 5]. In particular, Balogh and Bollobás [2] showed that

$$p_c(n) = \frac{1}{n^2\, 2^{2\sqrt{n}}}$$

is a threshold function for 2-neighbour bootstrap percolation on the hypercube Q_n: the graph whose vertices are the set of n-digit binary strings, with two vertices being adjacent if they differ in exactly one coordinate.

In order to obtain their result, Balogh and Bollobás first examined the properties of 2-neighbour percolating sets in hypercubes, then used these structural results to apply probabilistic techniques. However, similar methods have failed to provide results for 3-neighbour bootstrap percolation in hypercubes, although the extremal question regarding the minimum size of r-neighbour percolating sets in the hypercube has been answered asymptotically for all $r \geq 4$ and precisely for $r = 3$ by Morrison and Noel [14].

A related variant of bootstrap percolation, known as Froböse percolation, features a 'simpler' rule for infection that nevertheless requires 3 infected vertices in order to propagate: an uninfected vertex becomes infected if it is the sole uninfected vertex in some 4-cycle. Froböse initially introduced Froböse percolation for use in studying fluid flows via the hard-square lattice gas [10]. Since then it has been studied in a variety of contexts: extremal sets for Froböse-like update rules on the grid $[n]^d$ were studied in [6], and threshold functions for Froböse percolation have also been studied in [11] on the two dimensional grid $[n]^2$.

On the hypercube Q_n, Froböse percolation arises naturally due to its geometric interpretation: viewing Q_n as a polytope, the Froböse infection rule states that whenever a 2-dimensional face contains 3 (out of a possible 4) infected vertices, the remaining vertex then becomes infected. As the infection condition for Froböse percolation can only be satisfied when a vertex has at least two infected neighbours, Froböse bootstrap percolation is strictly stronger than

2-neighbour bootstrap percolation, in the sense that given a graph G and a subset $A \subseteq V(G)$, if A percolates via Froböse percolation then A also percolates via 2-neighbour percolation. Thus, any threshold function $p_c(n)$ for Froböse percolation in the hypercube Q_n satisfies $p_c(n) = \Omega(n^{-2}\, 2^{-2\sqrt{n}})$. In this paper, we study Froböse percolation directly to show that $p_c(n) = n^{-1}\, 2^{-\sqrt{2n}}$ is a threshold function for Froböse percolation in Q_n. In fact, we prove a stronger result.

1.2. Main result and key proof ideas. We will consider the asymptotic behaviour of Froböse percolation on the n-dimensional hypercube Q_n as $n \to \infty$, and will say that an event $E = E(n)$ holds with high probability (*whp* for short) if $\lim_{n\to\infty} \mathbb{P}[E] = 1$. Given $n \geq 0$ and a probability function $p = p(n)$, set $\Phi(Q_n, p)$ to be the probability that a randomly chosen p-dense subset of $V(Q_n)$ percolates via Froböse percolation.

Theorem 1.1. Let

$$p_* = p_*(n) := \frac{1}{n\, 2^{\sqrt{2n}}}.$$

Then there exist positive constants $c_1 < c_2 \in \mathbb{R}^+$ such that the following hold.

(1) If $p \leq c_1\, p_*$, then $\lim_{n\to\infty} \Phi(Q_n, p) = 0$.

(2) If $p \geq c_2\, p_*$, then $\lim_{n\to\infty} \Phi(Q_n, p) = 1$.

Recall that for each $d \in \mathbb{N}$ the d-dimensional hypercube is denoted by Q_d. We obtain Theorem 1.1 by first analysing the structure of Froböse percolating sets in Section 2. In particular, we classify precisely which subsets of $V(Q_n)$ form closed sets for Froböse percolation. This allows us to demonstrate that any Froböse percolating set $A \subseteq V(Q_n)$ corresponds to a sequence of nested subcubes $Q_0 = Q_{t_0} \subseteq Q_{t_1} \subseteq \ldots \subseteq Q_{t_\alpha} = Q_n$, each of which is *internally spanned* by A; that is, $A \cap Q_{t_i}$ percolates on Q_{t_i}. Furthermore, we identify Froböse percolating sets of minimum size and obtain a lower bound on the number of such sets in Q_n. Using this information about the structure of Froböse percolating sets, in Section 3 we prove Theorem 1.1 using a two-step process for each statement. The first step in both cases is to establish bounds on the probability that there exists an internally spanned subcube of dimension $d = o(n)$. To show the existence of c_1, we then use a first moment method argument to show that the requisite sequence of internally spanned subcubes does not exist with high probability. To show the existence of c_2, we use a second moment method argument to obtain an internally spanned subcube of dimension d, and then a sprinkling argument to show that the necessary sequence of internally spanned subcubes exists with high probability.

2 Structural results: minimum percolating sets

We begin by providing a detailed description of several properties of Froböse percolating sets in Q_n. In particular, we are interested in identifying and counting minimum percolating sets. In order to do so, we begin by identifying the sets where the percolation process stabilises.

Lemma 2.1. A subset $S \subseteq V(Q_n)$ is closed if and only if $S = S_1 \cup \ldots \cup S_k$ where each S_i is a cube, and for $i \neq j$, $\mathrm{d}(S_i, S_j) \geq 2$.

This can be seen by observing that the infection rule for the Froböse percolation model can be equivalently described as ‘a vertex v becomes infected if it is a contained in a copy of a subgraph isomorphic to Q_2 in which all other vertices are infected.’ As the intersection of two hypercubes is a hypercube (if non-empty), it follows that if a set is not closed then it contains a pair of adjacent hypercubes. Moreover, if a pair of adjacent hypercubes do not themselves form a higher dimensional hypercube, then there exists a copy of Q_2 containing exactly three vertices from said hypercubes.

Lemma 2.1 implies that if Q_k and Q_m are adjacent hypercubes, their closure under Froböse percolation $\overline{Q_k \cup Q_m}$ is itself a hypercube; in fact, their closure is the minimal hypercube containing both Q_k and Q_m. Such a hypercube has dimension at most $k + m + 1$. Suppose now that A is a Froböse percolating set for Q_n. Rather than considering how the process evolves step by step, let us consider an equivalent process where in each step a pair of adjacent hypercubes merge into a single hypercube of larger dimension. By the end of the process, the dimension of the largest dimensional subcube has increased by an average of at most 1 with each step. We therefore obtain Lemmas 2.2 and 2.3.

Lemma 2.2. If $A \subseteq V(Q_n)$ and A percolates via Froböse percolation, then $|A| \geq n + 1$.

Lemma 2.3. Let A be a percolating set for Q_n. Then there exists a sequence of subcubes $Q_0 = Q_{t_0} \subseteq Q_{t_1} \subseteq \ldots \subseteq Q_{t_\alpha} = Q_n$ such that the following hold:

1. Each Q_{t_i} is internally spanned by A.
2. For each $1 \leq i \leq \alpha$, $t_i \leq 2t_{i-1} + 1$.
3. For each $1 \leq i \leq \alpha$, $Q_{t_i} = \overline{Q_{t_{i-1}} \cup Q_k}$ for some internally spanned Q_k, where $k \leq t_{i-1}$ and $Q_k \neq Q_{t_i}$ for any i.

Lastly, we obtain a lower bound on the number of Froböse spanning sets of Q_n. In order to do so, we count the number of ways to construct a tuple $(v_0, v_1, \ldots, v_n)$ of vertices that form such a set. We first select an initial copy of Q_2 and choose 3 of its vertices to be v_0, v_1, and v_2. For each $i \geq 3$, we choose v_i to be adjacent to the cube spanned by $\{v_0, \ldots, v_{i-1}\}$. In order to avoid overcounting, we also require that v_i is not adjacent to the cube spanned by $\{v_0, \ldots, v_{i-2}\}$. Arguing in this manner, we obtain the following lower bound on the number of Froböse spanning sets:

Lemma 2.4. For all $n \geq 2$, the hypercube Q_n contains at least $\Omega((n-3)!\,2^{\binom{n}{2}-1})$ distinct Froböse-percolating sets of size $n + 1$.

We turn our attention now towards the application of these structural results in a probabilistic setting.

3 Probabilistic results: proof of Theorem 1.1

Let $F(d, p)$ denote the probability that a random p-dense subset $\mathcal{A}_p$ of $V(Q_d)$ Froböse percolates on Q_d. We will prove each statement of Theorem 1.1 separately, as the two cases are quite distinct.

First we will prove that if $p \leq c_1\, p_*$, then whp the Froböse percolation process fails on Q_n. The first step is to establish an upper bound on $F(d, p)$ using Lemmas 2.2 and 2.3.

Theorem 3.1. If $p = 2^{-s}$, $d < s$, and $s > 5$, then $F(d,p) \leq 2^{d\log d + \binom{d}{2} - (d+1)s}$.

The proof is inductive. For small d, the bound on $F(d,p)$ can be seen directly by counting the number of $(d+1)$-sets of Q_d. For larger d, by Lemma 2.3 Q_d is the closure of the union of two smaller subcubes. The inductive step then involves weighing the number of possible pairs of such subcubes against the probability that both are internally spanned.

Given Theorem 3.1 and taking $p = c_1\, p_*$ for a sufficiently small constant c_1, we can now show that whp no sequence as specified by Lemma 2.3 exists in Q_n, and hence Q_n is not spanned. To do so, assume that such a sequence exists, choose one of maximal length, and let Q_ℓ be the smallest subcube in said sequence whose dimension satisfies $\ell > \lceil\sqrt{2n}\rceil$. Let Q_k be the subcube in the sequence immediately preceding Q_ℓ. By Lemma 2.3 there exists a subcube Q_m not contained in the sequence such that $\overline{Q_k \cup Q_m} = Q_\ell$ and $m \leq k \leq \sqrt{2n} < \ell \leq 2\sqrt{2n}+1$. Using Theorem 3.1, we show that the expected number of triples (Q_ℓ, Q_k, Q_m) of such subcubes is $o(1)$, and hence whp no such sequence exists.

Next we will prove that if $p \geq c_2\, p_*$, then whp the Froböse percolation process percolates on Q_n. As before, we begin by bounding $F(d,p)$, this time from below. The following bound follows directly from Lemma 2.4:

Theorem 3.2. If $d \geq 2$, then $F(d,p) \geq (d-3)!\, 2^{\binom{d}{2}-1} p^{d+1}(1-p)^{2^d-(d+1)}$.

Using Theorem 3.2, a second moment argument shows that if $p \geq \frac{c_2}{2}\, p_*$ for sufficiently large c_2, then whp Q_n contains an internally spanned subcube of dimension $d = \lfloor\sqrt{2n} + \frac{\log n}{2}\rfloor$. Using a multi-round sprinkling argument with probabilities $(p_i)_{i=1}^{n-d}$ satisfying $\sum_{i=1}^{n-d} p_i < p$, we then grow the internally spanned Q_d by one dimension at a time, using a vertex infected with probability p_i to reach dimension $d+i$. As subcubes of dimension at least d have many neighbours, the sprinkling process reaches Q_n whp, completing the proof.

4 Further Study

Given the similarities between the results presented here on Froböse percolation and those of Balogh and Bollobás [2] concerning 2-neighbour percolation, two immediate questions present themselves.

Question 1. Is there a sharp threshold for Froböse percolation in the hypercube Q_n and if so, what is its value?

In particular, we are interested in exploring whether the techniques used in [5] to obtain a sharp threshold for 2-neighbour percolation in Q_n can be extended to Froböse percolation.

In addition, there are natural higher-dimensional generalisations of the Froböse percolation process, in which a vertex becomes infected if it is the sole uninfected vertex in some copy of Q_ℓ, which we call *ℓ-dimensional Froböse percolation*. Note that if a set is percolating for 3-dimensional Froböse percolation, then it is also percolating for 3-neighbour bootstrap percolation. As of yet, the question of threshold functions for 3-neighbour percolation in Q_n remains an important open question in the field. Considering the similarities in the behaviour of 2-neighbour and Froböse percolation on Q_n, and the fact that Froböse percolation permits a simpler analysis than 2-neighbour percolation, the following question might be a first step towards understanding 3-neighbour bootstrap percolation in Q_n.

Question 2. What is the threshold behaviour of 3-dimensional Froböse percolation in the hypercube Q_n?

Acknowledgements

The authors are supported in part by the Austrian Science Fund (FWF) [10.55776/P36131] and [10.55776/DOC183]. The second author is supported in part by NSERC PGS-D [588633-2024]. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] Joan Adler and Uri Lev. Bootstrap percolation: visualizations and applications. *Brazillian Journal of Physics*, 33(3), 2003.

[2] József Balogh and Béla Bollobás. Bootstrap percolation on the hypercube. *Probab. Theory Related Fields*, 134(4):624–648, 2006.

[3] József Balogh, Béla Bollobás, Hugo Duminil-Copin, and Robert Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364(5):2667–2701, 2012.

[4] József Balogh, Béla Bollobás, and Robert Morris. Bootstrap percolation in three dimensions. *Ann. Probab.*, 37(4):1329–1380, 2009.

[5] József Balogh, Béla Bollobás, and Robert Morris. Bootstrap percolation in high dimensions. *Combin. Probab. Comput.*, 19(5-6):643–692, 2010.

[6] József Balogh, Béla Bollobás, Robert Morris, and Oliver Riordan. Linear algebra and bootstrap percolation. *J. Combin. Theory Ser. A*, 119(6):1328–1335, 2012.

[7] B. Bollobás and A. Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

[8] Raphaël Cerf and Emilio N. M. Cirillo. Finite size scaling in three-dimensional bootstrap percolation. *Ann. Probab.*, 27(4):1837–1850, 1999.

[9] J. Chalupa, P. L. Leath, and G. R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C Solid State Physics*, 12(1):L31–L35, January 1979.

[10] K. Froböse. Finite-size effects in a cellular automaton for diffusion. *J. Statist. Phys.*, 55(5-6):1285–1292, 1989.

[11] Janko Gravner, Alexander E. Holroyd, and Robert Morris. A sharper threshold for bootstrap percolation in two dimensions. *Probab. Theory Related Fields*, 153(1-2):1–23, 2012.

[12] Ivailo Hartarsky. Bootstrap percolation, probabilistic cellular automata and sharpness. *J. Stat. Phys.*, 187(3):Paper No. 21, 17, 2022.

[13] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Theory Comput.*, 11:105–147, 2015.

[14] Natasha Morrison and Jonathan A. Noel. Extremal bounds for bootstrap percolation in the hypercube. *J. Combin. Theory Ser. A*, 156:61–84, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The Erdős-Hajnal graphs and reflection of infinite chromatic numbers

(Extended abstract)

Monroe Eskew* Yair Hayut †

Abstract

Using large cardinals, we produce a model of the axioms of set theory that answers several open questions in infinite combinatorics, and in particular shows that the original upper bounds on the chromatic numbers of certain universal graphs on the ω_n's given by Erdős and Hajnal were optimal. This leads to a reflection property of uncountable chromatic numbers that gives an optimal analogue to the De Bruijn–Erdős compactness theorem for finite chromatic number.

In recent work [2], we constructed a model of the ZFC axioms of set theory that answers several questions about chromatic numbers of uncountable graphs. The De Bruijn–Erdős Theorem says that for any graph G, if there exists a natural number n such that every finite subgraph H of G has chromatic number $\chi(H) \leq n$, then so does the whole graph G. This is a kind of compactness property for finite chromatic number. In contrast, Erdős and Hanjal [1] showed that there are always counterexamples to this kind of compactness for infinite chromatic number:

Theorem 1 (Erdős-Hajnal). *For every infinite cardinal κ and every $n < \omega$, there exists a graph G of size $\beth_n(\kappa)^+$ such that $\chi(G) > \kappa$, but $\chi(H) \leq \kappa$ for every subgraph of G of size $\leq \beth_n(\kappa)$.*

Here, we use the standard notation that $\beth_0(\kappa) = \kappa$, $\beth_{n+1}(\kappa) = 2^{\beth_n(\kappa)}$, and λ^+ is the next cardinality above λ. In particular, there exists a graph on the powerset of the reals such that such that every subgraph of size at most continuum has countable chromatic number, yet the whole graph has uncountable chromatic number.

In [1], Erdős and Hajnal also introduced their universal graph $\mathsf{EH}(\kappa, \lambda)$, defined by taking the vertex set as the set of all functions $f : \kappa \to \lambda$, and connecting f, g with an edge when $|\{\alpha : f(\alpha) = g(\alpha)\}| < \kappa$. (So two functions are connected when they *dis*agree almost everywhere.)

*Kurt Gödel Research Center, University of Vienna, Vienna, Austria. Partially supported by Austrian Science Fund projects P34603 and PIN1355423.

†Einstein Institute of Mathematics, Edmond J. Safra Campus, The Hebrew University of Jerusalem, Givat Ram. Jerusalem, 9190401, Israel. Partially supported by the Israel Science Fund 1967/21.

They showed that for every graph G of size κ, if every strictly smaller subgraph H has $\chi(H) \leq \lambda$, then there is an edge-preserving homomorphism $f : G \to \mathsf{EH}(\kappa, \lambda)$. This implies that $\chi(G) \leq \chi(\mathsf{EH}(\kappa, \lambda))$. By Theorem 1, we have that for every infinite κ and $n < \omega$, $\chi(\mathsf{EH}(\beth_n(\kappa)^+, \kappa)) > \kappa$.

The chromatic numbers of these graphs can vary between models of set theory. For example, [6] shows that the value of $\chi(\mathsf{EH}(\omega_2, \omega))$ is independent of the Generalized Continuum Hypothesis (GCH). However, the range of consistent values has largely remained a mystery. Because of their usefulness in computing the chromatic numbers of infinite graphs in general, several authors, such as Todorčević [8], have urged their computation. We show that the lower bounds computed by Erdős and Hajnal are optimal:

Theorem 2. *If ZFC is consistent with a huge cardinal, then there is a model of ZFC + GCH in which for all $n \leq m < \omega$, $\chi(\mathsf{EH}(\omega_m, \omega_n)) = \omega_{n+1}$.*

Our model also satisfies many instances of the Foreman-Laver graph reflection property of [3, 5]. Recall that $[\kappa, \lambda] \twoheadrightarrow [\kappa', \lambda']$ means that every graph with κ vertices and chromatic number λ has a subgraph with κ' vertices and chromatic number λ'.

Theorem 3. *It is consistent relative to a huge cardinal that for all natural numbers k, m, n with $n \leq m$, $[\omega_{m+k}, \omega_m] \twoheadrightarrow [\omega_{n+k}, \omega_n]$.*

This property can consistently fail rather dramatically. For example, Shelah [7] showed that if we assume Gödel's axiom of constructiblity $V = L$, then for every successor cardinal κ, there is a graph of size κ such that every small subgraph has countable chromatic number, yet the whole graph has chromatic number κ.

The constructions behind Theorem 1 show that, under GCH, we cannot have $[\omega_{m+k_1}, \omega_m] \twoheadrightarrow [\omega_{n+k_0}, \omega_n]$ for $0 < n \leq m$ and $k_0 < k_1$. Thus we achieve an optimal approximation to compactness for graphs on the cardinals below $\aleph_\omega$—Although we cannot generally reflect the chromatic number of an infinite graph into a small subgraph, we can consistently reflect the *gap* between the chromatic number and the size of the graph into small subgraphs.

Our method is roughly as follows. Starting from a model of ZFC with a huge cardinal, we build, via Cohen's method of forcing, a model of ZFC in which for every finite n, the cardinal ω_n carries an ideal with strong combinatorial properties that reflects that it was a very large cardinal in the original model. From the existence of such ideals, we deduce the Foreman-Woodin transfer property [4] that moves structure on ω_n to ω_{n+1}. In particular, this allows us to find uniform ultrafilters on each ω_m that yield ultrapowers of structures of size ω_n, for $n < m$, of the minimal possible size, ω_{n+1}. This itself answers questions from model theory dating back to the 1960s. These small ultrapowers give optimal colorings of the graphs $\mathsf{EH}(\omega_m, \omega_n)$.

In future work, we hope to find small consistent values for Erdős-Hajnal graphs on more cardinals and to investigate similar reflection phenomena. We conjecture that it is consistent with the axioms of set theory that for every graph G, if all small subgraphs have countable chromatic number, then G has chromatic number at most ω_1.

References

[1] P. Erdős and A. Hajnal. On chromatic number of infinite graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 83–98. Academic Press, New York-London, 1968.

[2] Monroe Eskew and Yair Hayut. Dense ideals. *arXiv preprint arXiv:2410.14359*, 2024.

[3] Matthew Foreman. Potent axioms. *Trans. Amer. Math. Soc.*, 294(1):1–28, 1986.

[4] Matthew Foreman. An $\aleph_1$-dense ideal on $\aleph_2$. *Israel J. Math.*, 108:253–290, 1998.

[5] Matthew Foreman and Richard Laver. Some downwards transfer properties for $\aleph_2$. *Adv. in Math.*, 67(2):230–238, 1988.

[6] P. Komjáth. The chromatic number of some uncountable graphs. In *Sets, graphs and numbers (Budapest, 1991)*, volume 60 of *Colloq. Math. Soc. János Bolyai*, pages 439–444. North-Holland, Amsterdam, 1992.

[7] Saharon Shelah. Incompactness for chromatic numbers of graphs. In *A tribute to Paul Erdős*, pages 361–371. Cambridge Univ. Press, Cambridge, 1990.

[8] Stevo Todorcevic. Combinatorial dichotomies in set theory. *Bull. Symbolic Logic*, 17(1):1–72, 2011.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 – 29, 2025

Local resilience of random geometric graphs with respect to connectivity*

(Extended abstract)

Alberto Espuny Díaz† Lyuben Lichev‡ Alexandra Wesolek§

Abstract

Given an increasing graph property $\mathcal{P}$, a graph G is α-resilient with respect to $\mathcal{P}$ if, for every spanning subgraph $H \subseteq G$ where each vertex keeps more than a $(1-\alpha)$-proportion of its neighbours, H has property $\mathcal{P}$. This notion naturally extends the study of sufficient minimum-degree conditions for $\mathcal{P}$ to subgraphs of arbitrary graphs. We study this notion of local resilience with $\mathcal{P}$ being connectivity and G being a random geometric graph $G_d(n,r)$ obtained by embedding n vertices independently and uniformly at random in $[0,1]^d$, and connecting two vertices by an edge if the distance between them is at most r.

We show that, for every $\varepsilon > 0$, if r is a constant factor above the sharp threshold for connectivity of $G_d(n,r)$, the random geometric graph is $(1/2-\varepsilon)$-resilient with respect to connectivity, where the constant $1/2$ cannot be improved. However, contrary to binomial random graphs, for sufficiently small $\varepsilon > 0$, connectivity is not born $(1/2-\varepsilon)$-resilient in the 2-dimensional random geometric graph process. Our results are thus tight up to the dependence of ε on the other parameters.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-000

*The full version of this paper can be found in [6].

†Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany. E-mail: espuny-diaz@informatik.uni-heidelberg.de. Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through projects no. 447645533 and 513704762.

‡Technische Universität Wien, Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria. E-mail: lyuben.lichev@tuwien.ac.at. Supported by the Austrian Science Fund (FWF) grant no. 10.55776/ESP624.

§Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany. E-mail: alexandrawesolek@gmail.com. Supported by the DFG under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

1 Introduction

Given a graph G that satisfies a certain property $\mathcal{P}$, a general question in graph theory asks how much one needs to modify G so that it stops satisfying $\mathcal{P}$. This general question can be interpreted from different points of view, and has motivated the study of the *resilience* of graphs with respect to different properties. The following localised version of this question has received special attention when considering spanning properties. Given a graph G and some $\alpha \in [0,1]$, we say that a subgraph $H \subseteq G$ is an *α-subgraph* of G if for every $v \in V(G)$ we have $d_H(v) \geq \alpha d_G(v)$. If G satisfies an increasing property $\mathcal{P}$, we define the *local resilience* of G with respect to $\mathcal{P}$ as the supremum of the values $\alpha \in [0,1]$ such that *every* $(1-\alpha)$-subgraph of G satisfies $\mathcal{P}$. Alternatively, one may think that an adversary is allowed to remove edges from G subject to the condition that the proportion of edges incident to any given vertex cannot be decreased by more than a factor of α. Then, the local resilience of G with respect to $\mathcal{P}$ is the minimum α such that the adversary can produce a $(1-\alpha)$-subgraph of G without property $\mathcal{P}$.

Many natural statements can be expressed in the language of local resilience. For example, it is a simple observation that every n-vertex graph G with minimum degree $\delta(G) > n/2-1$ is connected, and that this is best possible. This statement can be rephrased in the language of local resilience of the complete graph. Indeed, fix $\alpha^* := (\lfloor n/2 \rfloor - 1)/(n-1)$. Then, for every $\alpha > \alpha^*$, every α-subgraph of K_n is connected and this is best possible, implying that the local resilience of K_n with respect to connectivity is $1-\alpha^* \sim 1/2$. Likewise, all classical extremal results with minimum degree conditions guaranteeing some property $\mathcal{P}$ can be seen as statements about the local resilience of K_n with respect to $\mathcal{P}$.

One of the main avenues for research in this context has been the local resilience of random graphs with respect to different properties. The binomial random graph $G(n,p)$, sampled by including each of the possible $\binom{n}{2}$ edges on vertex set $[n] := \{1,\dots,n\}$ independently with probability p, has received particular attention since the systematic study of local resilience was introduced by Sudakov and Vu [18]. In their pioneering paper, among other results, they considered the local resilience of $G(n,p)$ with respect to Hamiltonicity and the containment of perfect matchings. After further consideration by different authors [1, 2, 7], Lee and Sudakov [11] finally showed that, for $p \gg \log n/n$ and any fixed constant $\varepsilon > 0$, asymptotically almost surely (that is, with probability tending to 1 as n grows to infinity, which we abbreviate as a.a.s.) the local resilience of $G(n,p)$ with respect to Hamiltonicity is at least $1/2-\varepsilon$. This result has been further sharpened later, both into a 'degree sequence' version [5] and to a 'hitting time' result [12, 13].

Local resilience has also been studied in random regular graphs. Given integers $1 \leq d < n$, a random d-regular graph $G_{n,d}$ is obtained by sampling an element from the set of all n-vertex d-regular graphs uniformly at random. Improving on earlier work of Ben-Shimon, Krivelevich and Sudakov [1], Condon, Espuny Díaz, Girão, Kühn and Osthus [4] proved a result analogous to that of Lee and Sudakov [11] for random regular graphs: for every $\varepsilon \in (0,1/2]$ and all $d \gg 1/\varepsilon$, a.a.s. every $(1/2+\varepsilon)$-subgraph of $G_{n,d}$ is Hamiltonian.

Naturally, the results above imply that both $G(n,p)$ and $G_{n,d}$ are a.a.s. $(1/2-\varepsilon)$-resilient with respect to connectivity. It is worth noting that the constant $1/2$ in these results is best possible in great generality: this is a consequence of a result of Stiebitz [17], who showed that the vertices of every graph can be partitioned into two sets in such a way that every vertex v has at least $d(v)/2 - 1$ neighbours in the part to which it belongs. In particular, for any n-vertex graph G whose minimum degree tends to infinity with n, we may consider such a partition and delete all edges between the two parts to obtain a $(1/2 - o(1))$-subgraph which is not connected. This implies that the local resilience of G with respect to any property which necessitates connectivity is at most $1/2 + o(1)$.

In this extended abstract, we focus on the local resilience of *random geometric graphs*, which were first introduced by Gilbert [8]. Given positive integers d and n and a real number $r > 0$, the d-dimensional random geometric graph $G_d(n, r)$ is a graph on vertex set $[n]$ whose edges are generated as follows. Consider n mutually independent uniform random variables $X_1, \ldots, X_n$ on $[0, 1]^d$. For each pair of distinct elements $i, j \in [n]$, the edge ij is added to the graph if and only if $\|X_i - X_j\| \le r$, where $\|\cdot\|$ denotes the Euclidean distance.

2 Results

We begin a systematic analysis of the local resilience of random geometric graphs with respect to different properties. In this note we present our results about connectivity. We obtain a result analogous to those for $G(n, p)$ and $G_{n,d}$.

Theorem 1. *For every $\varepsilon \in (0, 1/2]$ and integer $d \ge 1$, there exists a constant $C > 0$ such that, for every $r \ge C(\log n/n)^{1/d}$, a.a.s. every $(1/2 + \varepsilon)$-subgraph of $G_d(n, r)$ is connected.*

Definition 1 provides a local resilience analogue of classical results of Godehardt and Jaworski [9] and Penrose [14, 15] on the threshold for connectivity of random geometric graphs. Moreover, we can extend this result to k-connectivity for a large range of values of k. Together with the aforementioned result of Stiebitz [17], Definition 1 implies that a.a.s. the local resilience of random geometric graphs $G_d(n, r)$ with respect to connectivity is $1/2 \pm o(1)$ whenever $r = \omega((\log n/n)^{1/d})$. If we restrict our attention to $(1/2 - \varepsilon)$-subgraphs of $G_d(n, r)$, we can say more about the size of a largest component.

Proposition 2. *For every $\varepsilon \in (0, 1/2]$, there exists a constant $C^* > 0$ such that, for every integer $d \ge 1$ and $r \ge C^*(\log n/n)^{1/d}$, a.a.s. there exists a $(1/2 - \varepsilon)$-subgraph of $G_d(n, r)$ with components of order at most $5rn$.*

This can be shown in a straightforward way by considering a partition of $[0, 1]^d$ into subregions defined by hyperplanes parallel to $x = 0$ such that all regions have width roughly $3r$ (say). One can then show that a.a.s. the graph obtained by removing all edges of $G_d(n, r)$ which are not contained in one of the regions is a $(1/2 - \varepsilon)$-subgraph of $G_d(n, r)$.

Going back to Definition 1, we remark that the expression $C(\log n/n)^{1/d}$ must be larger than the (sharp) connectivity threshold of $G_d(n, r)$. This implies that, for fixed ε, the constant $C = C(d, \varepsilon)$ needs to grow at speed $\Omega(\sqrt{d})$ as a function of d (see, e.g., [16, Theorem 13.2]). In fact, a careful proof shows that it suffices to have a constant C which grows with speed $\Theta(\sqrt{d})$ (see [6, Theorem 3.1]). The behaviour of C with respect to ε is more mysterious. While we are uncertain if C can be chosen universally for all $\varepsilon \in (0, 1/2]$ as a function of d when $d \geq 3$, we can prove that this is not the case when $d \in \{1, 2\}$.

Theorem 3. *The following statements hold.*

(i) *For every $\varepsilon \in (0, 1/2]$, if $r \leq \log n/4\varepsilon n$, then a.a.s. there exists a disconnected $(1/2 + \varepsilon)$-subgraph of $G_1(n, r)$.*

(ii) *For every $C > 0$, there exists $\varepsilon \in (0, 1/2]$ such that, for every $r \leq C(\log n/n)^{1/2}$, a.a.s. there exists a disconnected $(1/2 + \varepsilon)$-subgraph of $G_2(n, r)$.*

This result has one important consequence. It is now well known that, throughout the random graph process (which we do not define here), as one increases the density of random graphs, a.a.s. they become connected as soon as their last isolated vertex disappears [3], and the same is true when increasing the density in the geometric setting [14, 15] when $d \geq 2$ (these are called 'hitting time' results). Haller and Trujić [10] showed that a local resilience version of this holds in the random graph process. In contrast to this, Definition 3 (ii) shows that a local resilience version of the geometric hitting time result fails when $d = 2$. It would be interesting to determine whether this is also the case for larger d or, on the contrary, a hitting time version holds for sufficiently large d.

We include below a short proof of Definition 1. A tighter proof, as well as proofs of all other results, can be found in [6], where we also obtain results about the local resilience of $G_d(n, r)$ with respect to Hamiltonicity and the containment of long cycles.

3 Proof of Definition 1

In this section, we provide a simple proof of Definition 1. The proof relies mostly on the fact that, for a suitable choice of the parameters and a sufficiently small $\delta = \delta(d, \varepsilon) > 0$, all pairs of vertices at Euclidean distance at most δr from each other have many common neighbours in any $(1/2+\varepsilon)$-subgraph H of $G_d(n, r)$. Thus, the square H^2 of H (obtained by adding an edge between any pair of vertices at distance at most 2 in H) contains $G_d(n, \delta r)$, which is a.a.s. connected for all sufficiently large r. In fact, this same approach yields higher connectivity: one can adapt the proofs below to show that one retains $cr^d n$-connectivity, for a suitable constant $c > 0$.

Our proof of Definition 1 will use the following properties of a random point set in $[0, 1]^d$, which follow from standard applications of Chernoff bounds (so we omit the details).

Lemma 4. *For every fixed integer $d \geq 1$ and $\varepsilon \in (0, 1/2]$, there exist constants $\delta, C_1 > 0$ such that the following holds. Let V be a set of n points sampled independently and uniformly from $[0, 1]^d$. Then, for any $r \in [C_1(\log n/n)^{1/d}, \sqrt{d}]$,*

(a) *a.a.s., for every $v \in V$, the ball $B(v, r)$ contains at most $(1 + \varepsilon)|B(v, r) \cap [0, 1]^d|n$ points of V, and*

(b) *a.a.s., for every pair of points $u, v \in V$ such that $\|u - v\| \leq \delta r$, the region $B(u, r) \cap B(v, r) \cap [0, 1]^d$ contains at least $(1 - \varepsilon)|B(v, r) \cap [0, 1]^d|n$ points of $V \setminus \{u, v\}$.*

In order to prove Definition 1, we first prove the following auxiliary lemma.

Lemma 5. *For every fixed integer $d \geq 1$ and $\varepsilon \in (0, 1/2]$, there exist constants $\delta, C_1 > 0$ such that, for any $r \in [C_1(\log n/n)^{1/d}, \sqrt{d}]$, a.a.s. the following holds: every pair of vertices of $G_d(n, r)$ at Euclidean distance at most δr have a common neighbour in each $(1/2 + \varepsilon)$-subgraph of $G_d(n, r)$.*

Proof. Let $G \sim G_d(n, r)$ and condition on the event that $V(G)$ satisfies the properties described in Definition 4 (a) and (b), which holds a.a.s. By Definition 4 (a), every vertex $v \in V(G)$ has degree at most $(1 + \varepsilon)|B(v, r) \cap [0, 1]^d|n$ in G. By Definition 4 (b), every pair of vertices $u, v \in V(G)$ such that $\|u - v\| \leq \delta r$ have at least $(1 - \varepsilon)|B(v, r) \cap [0, 1]^d|n$ common neighbours in G. Therefore, if $H \subseteq G$ is a $(1/2 + \varepsilon)$-subgraph of G, for any pair of vertices $u, v \in V(G)$ such that $\|u - v\| \leq \delta r$, we have that

$$|N_H(u) \cap N_H(v)| \geq |N_G(u) \cap N_G(v)| - (1/2 - \varepsilon)(d_G(u) + d_G(v)) > 0. \qquad \square$$

We will also use the fact that sufficiently dense random geometric graphs are connected, which is a particular case of a result of Penrose [16, Theorem 13.2]. Here, we present a simple proof of the weaker result that we need.

Lemma 6. *For each $d \geq 1$, there exists a constant $C_2 > 0$ such that, if $r \geq C_2(\log n/n)^{1/d}$, then a.a.s. $G_d(n, r)$ is connected.*

Proof. Consider a tessellation $\mathcal{Q}$ of $[0, 1]^d$ with cubic cells of side length $s := \lceil \sqrt{d + 3}/r \rceil^{-1}$. Let Γ be an auxiliary graph where two cells are joined by an edge whenever they share a $(d - 1)$-dimensional face. Observe that, for any cells q and q' with $qq' \in E(\Gamma)$ and any pair of points $p, p' \in q \cup q'$, the choice of s guarantees that $\|p - p'\| \leq r$. In particular, all vertices of the graph $G \sim G_d(n, r)$ contained in two such cells form a complete graph. Moreover, note that $r/2\sqrt{d + 3} \leq s \leq r/\sqrt{d + 3}$ and so, in particular, the volume of each cell is $s^d \geq 2^{-d}(d + 3)^{-d/2}r^d$. By choosing C_2 sufficiently large, it follows from a Chernoff bound and a union bound that a.a.s. every $q \in \mathcal{Q}$ contains at least one vertex of G. Since vertices in neighbouring cells are joined by edges and a.a.s. each cell contains at least one vertex, the connectivity of $G_d(n, r)$ follows from the connectivity of Γ. $\square$

Proof of Theorem 1. Let δ and C_1 be the constants given by Definition 5 applied with d and ε, and let C_2 be the constant given by Definition 6 with d as input. Fix $C := \max\{C_1, C_2/\delta\}$. Couple $G_d(n, \delta r)$ and $G_d(n, r)$ in a way that their vertices have the same positions in $[0, 1]^d$ and fix a $(1/2 + \varepsilon)$-subgraph H of $G_d(n, r)$. Moreover, condition on the events that $G_d(n, \delta r)$ is connected and that, for every edge $uv \in E(G_d(n, \delta r))$, the vertices u and v have at least one common neighbour in H (note that both events hold a.a.s. by Definitions 5 and 6 with our choice of C). This implies that $G_d(n, \delta r) \subseteq H^2$ and hence H^2 is also connected, which can only occur if H is connected, as desired. $\square$

References

[1] S. Ben-Shimon, M. Krivelevich and B. Sudakov, Local resilience and Hamiltonicity maker-breaker games in random regular graphs. *Comb. Probab. Comput.* 20.2 (2011), 173–211, DOI: `10.1017/S0963548310000453`.

[2] ———, On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs. *SIAM J. Discrete Math.* 25.3 (2011), 1176–1193, DOI: 10.1137/110821299.

[3] B. Bollobás and A. Thomason, Random graphs of small order. M. Karoński and A. Ruciński (eds.), *Random graphs '83 (Poznań, 1983)*, *North-Holland Math. Stud.*, vol. 118, 47–97, North-Holland, Amsterdam (1985), DOI: 10.1016/S0304-0208(08)73612-0.

[4] P. Condon, A. Espuny Díaz, A. Girão, D. Kühn and D. Osthus, Dirac's theorem for random regular graphs. *Comb. Probab. Comput.* 30.1 (2021), 17–36, DOI: 10.1017/S0963548320000346.

[5] P. Condon, A. Espuny Díaz, J. Kim, D. Kühn and D. Osthus, Resilient degree sequences with respect to Hamilton cycles and matchings in random graphs. *Electron. J. Comb.* 26.4 (2019), paper no. P4.54, 22 pages, DOI: 10.37236/8279.

[6] A. Espuny Díaz, L. Lichev and A. Wesolek, On the local resilience of random geometric graphs with respect to connectivity and long cycles. *arXiv e-prints* (2024). arXiv: 2406.09921.

[7] A. Frieze and M. Krivelevich, On two Hamilton cycle problems in random graphs. *Isr. J. Math.* 166 (2008), 221–234, DOI: 10.1007/s11856-008-1028-8.

[8] E. N. Gilbert, Random plane networks. *J. Soc. Ind. Appl. Math.* 9 (1961), 533–543, DOI: 10.1137/0109045.

[9] E. Godehardt and J. Jaworski, On the connectivity of a random interval graph. *Random Struct. Algorithms* 9.1-2 (1996), 137–161, DOI: 10.1002/(SICI)1098-2418(199608/09)9:1/2<137::AID-RSA9>3.0.CO;2-Y.

[10] L. Haller and M. Trujić, On resilience of connectivity in the evolution of random graphs. *Electron. J. Comb.* 26.2 (2019), paper no. p2.24, 15 pages, DOI: 10.37236/7885.

[11] C. Lee and B. Sudakov, Dirac's theorem for random graphs. *Random Struct. Algorithms* 41.3 (2012), 293–305, DOI: 10.1002/rsa.20419.

[12] R. Montgomery, Hamiltonicity in random graphs is born resilient. *J. Comb. Theory, Ser. B* 139 (2019), 316–341, DOI: 10.1016/j.jctb.2019.04.002.

[13] R. Nenadov, A. Steger and M. Trujić, Resilience of perfect matchings and Hamiltonicity in random graph processes. *Random Struct. Algorithms* 54.4 (2019), 797–819, DOI: 10.1002/rsa.20827.

[14] M. D. Penrose, The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.* 7.2 (1997), 340–361, DOI: 10.1214/aoap/1034625335.

[15] ———, On k-connectivity for a geometric random graph. *Random Struct. Algorithms* 15.2 (1999), 145–164, DOI: 10.1002/(SICI)1098-2418(199909)15:2<145::AID-RSA2>3.0.CO;2-G.

[16] ———, *Random geometric graphs*, *Oxf. Stud. Probab.*, vol. 5. Oxford: Oxford University Press (2003).

[17] M. Stiebitz, Decomposing graphs under degree constraints. *J. Graph Theory* 23.3 (1996), 321–324, DOI: 10.1002/(SICI)1097-0118(199611)23:3<321::AID-JGT12>3.0.CO;2-H.

[18] B. Sudakov and V. H. Vu, Local resilience of graphs. *Random Struct. Algorithms* 33.4 (2008), 409–433, DOI: 10.1002/rsa.20235.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Factors and powers of cycles in the budget-constrained random graph process*

(Extended abstract)

Alberto Espuny Díaz[†] Frederik Garbe[†] Tássio Naia[‡] Zak Smith[†]

Abstract

We consider the budget-constrained random graph process introduced by Frieze, Krivelevich and Michaeli: a player, *Builder*, is presented with t distinct edges of K_n one by one, uniformly at random. Builder may *purchase* at most b of these edges, and must (irrevocably) decide whether to purchase each edge as soon as it is offered. Her goal is to construct a graph which a.a.s. satisfies a certain property; we investigate the properties of containing F-factors or powers of Hamilton cycles.

We obtain general lower bounds on the budget b, as a function of t, required for Builder to obtain partial F-factors, for arbitrary F, which in particular imply bounds for complete F-factors and powers of Hamilton cycles; this gives a negative answer to a question of Frieze, Krivelevich and Michaeli, in a strong sense. Conversely, we exhibit simple strategies for constructing partial F-factors, complete F-factors for strictly-1-balanced F, and the k-th power of a Hamilton cycle, the first of which shows that our general lower bound is best possible up to constant factors.

---

*This extended abstract is based on the results in [5].

[†]Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany. E-mails: {espuny-diaz,garbe,smith}@informatik.uni-heidelberg.de.
This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through projects no. 513704762 (A. Espuny Díaz) and 428212407 (F. Garbe, Z. Smith).

[‡]Centre de Recerca Matemàtica, 08193 Bellaterra, Spain. E-mail: tnaia@member.fsf.org.
Supported by the Grant PID2023-147202NB-I00 funded by MICIU/AEI/10.13039/501100011033. This work is also partially supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M)

1 Introduction

Given a positive integer n, consider the *random graph process*, that is, a random sequence of graphs $G_0 \subseteq G_1 \subseteq \ldots \subseteq G_{\binom{n}{2}}$ obtained by letting G_0 be the empty graph on a (labelled) set of n vertices and, for each $i \in [\binom{n}{2}]$, letting $G_i = G_{i-1} \cup \{e_i\}$ for an edge e_i chosen uniformly at random among the edges not in G_{i-1}. For a non-trivial increasing graph property $\mathcal{P}$, there exists a unique $\tau_{\mathcal{P}} \in [\binom{n}{2}]$ such that $G_{\tau_{\mathcal{P}}} \in \mathcal{P}$ but $G_{\tau_{\mathcal{P}}-1} \notin \mathcal{P}$; the random variable $\tau_{\mathcal{P}}$ is called the *hitting time* for $\mathcal{P}$ in the random graph process, and is closely related to the threshold for $\mathcal{P}$ in the binomial random graph $G(n,p)$.

In general, even graph properties which in principle require 'few' edges to occur often have comparably 'large' hitting times, meaning that, in order for a random graph to satisfy the desired property, one usually needs to have 'many more' random edges than the minimum number of edges required for the property. For example, Hamilton cycles and triangle factors both consist of linearly many edges but typically do not appear in the random graph process until it has around $n \log n/2$ and $\Omega(n^{4/3} \log^{1/3} n)$ edges, respectively [11, 14]. Moreover, as edges arrive one by one throughout the random graph process, it is generally hard to know whether an edge will be crucial for the desired property. These facts motivated Frieze, Krivelevich and Michaeli [7] to introduce an 'online' model for constructing graphs following the random graph process, but with a constraint on the 'budget', that is, on the number of edges that the final constructed graph is allowed to have.

Formally, we define the *budget-constrained random graph process* to be a one-player game in which a uniformly random sequence of $t \in [\binom{n}{2}]$ (distinct) edges of K_n are presented, one edge at a time, to the player, **Builder**. Each time she is presented with an edge, she must immediately and irrevocably decide whether or not to 'purchase' it, without knowing the order of the remaining edges. We write $B_i \subseteq G_i$ for the graph of all edges purchased up to time i, for each $i \in [t]$. Builder's goal is to construct a graph which satisfies a desired (monotone graph) property $\mathcal{P}$ by time t, while buying at most $b \in [t]$ edges. We call a function defining the choices Builder makes when confronted with any possible sequence of t edges a *(t,b)-strategy* if it respects the budget restriction $|E(B_t)| \leq b$, and we say it is *successful for $\mathcal{P}$* if asymptotically almost surely (a.a.s.) B_t satisfies $\mathcal{P}$.

The general question of interest, given a monotone increasing property $\mathcal{P}$, is: for which pairs (t,b) does Builder have a successful (t,b)-strategy for $\mathcal{P}$? In particular, how close can she get to the trivial bounds $b \geq \min_{G\in\mathcal{P}} |E(G)|$ and $t \geq \tau_{\mathcal{P}}$?

2 Previous results

The budget-constrained random graph process has already been well-studied for various natural properties. For the property of having minimum degree at least some constant $k \geq 1$, Frieze, Krivelevich and Michaeli [7] described a strategy which is asymptotically optimal in both time and budget. Moreover, at the hitting time τ_k itself, they achieved budget $C_k n$, for which the constant C_k was improved further by Katsamaktsis and Letzter [13]. The strategy of [7] also yields a k-connected graph at the hitting time for

$k \geq 3$, and, for all $k \geq 2$, Lichev [16] showcased an asymptotically optimal strategy for k-connectivity. Observe that in the case $k = 1$, a very easy strategy suffices for exactly optimal time and budget: simply purchase any edge that does not create a cycle.

Frieze, Krivelevich and Michaeli [7] also proposed a successful (τ_2, Cn)-strategy for Hamiltonicity (with $C > 1$ constant), and Anastos [2] showed conversely that (exactly at the hitting time) C must be bounded away from 1, while on the other hand providing an asymptotically optimal strategy. Similar results hold for perfect matchings [2, 7]. Finally, the asymptotics are known for constructing copies of (fixed, small) trees and cycles [7], as well as the diamond and any number of triangles sharing a unique vertex (Iľkovič, León and Shu [10]), but all other fixed graphs remain an open problem.

In view of their results for spanning properties, Frieze, Krivelevich and Michaeli [7] raised the following question.

Question 2.1 ([7])**.** *Let H be an n-vertex graph with bounded maximum degree (so note $|E(H)| = O(n)$). Does Builder have a successful (t, b)-strategy for containing a copy of H with t 'close' to the hitting time for H and b linear?*

As a concrete example, they asked whether this can be achieved for the square of a Hamilton cycle (for $k \in \mathbb{N}$, the *k-th power* of a graph G is the graph obtained from G by adding an edge between any pair of vertices whose graph distance in G is at most k.)

Question 2.2 ([7])**.** *Does there exist a successful $(O(n^{3/2}), O(n))$-strategy for the square of a Hamilton cycle?*

Our work [5] shows that the answer to these questions is negative in general and that, in some cases, close to the hitting time, the budget cannot be asymptotically smaller than the cost of the trivial strategy of purchasing every edge presented to Builder. We showcase this by considering *graph factors*. More generally, we obtain new results about the interplay between t and b required to have successful strategies for constructing graphs which contain different graph factors, including powers of a Hamilton cycle. We achieve this by proving lower bounds for b as a function of t, as well as showing that these are essentially tight by providing (almost) matching successful (t, b)-strategies.

3 Our results

Let F be a fixed graph. We say that a graph G on n vertices (where $|V(F)|$ divides n) contains an *F-factor* if its vertex set can be partitioned into $n/|V(F)|$ sets of equal size in such a way that the graph induced by G on each of the sets contains a copy of F. In more generality (and removing the restriction on the divisibility of n), for any $\alpha \in (0, 1)$, we say that G contains an *α-F-factor* (or just *partial F-factor*) if it contains a set of at least $\alpha n/|V(F)|$ vertex-disjoint copies of F.

Finding F-factors in graphs is one of the classical problems in graph theory, and has received much attention in the context of binomial random graphs. The general case was first considered independently by Ruciński [20] and by Alon and Yuster [1]. The threshold

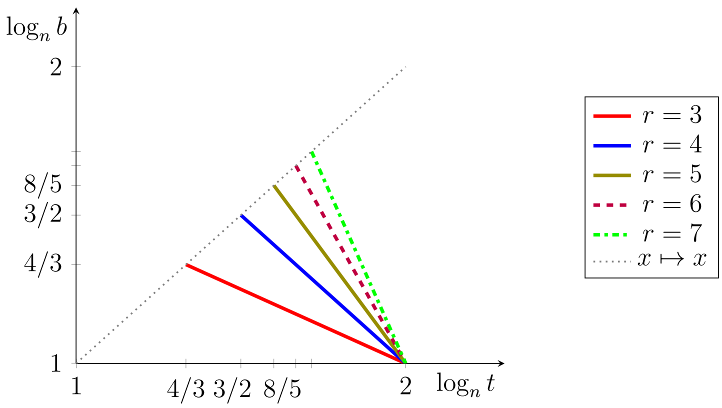


Figure 1: Optimal budget for successful (t, b)-strategies for K_r-factors for $r \in \{3, 4, 5, 6, 7\}$.

for (complete) F-factors for all strictly 1-balanced F was determined about 15 years later in the influential paper of Johansson, Kahn and Vu [11], and there have been many further developments, including hitting time results, since then [3, 4, 8, 9]. Given this wealth of research, it is very natural to consider F-factors in the budget-constrained random graph process. For simplicity, we consider here the particular case that $F = K_r$ with $r \geq 2$, for which a.a.s. the hitting time satisfies $\tau_{K_r} = \Theta(n^{-2/r} \log^{1/\binom{r}{2}} n)$ [11]. We prove the following lower bound on the budget required by any successful strategy.

Theorem 3.1. *Let $r \geq 2$ and $\alpha \in (0, 1)$. Let $t = t(n)$ and $b = b(n)$ be such that there exists a successful (t, b)-strategy for an α-K_r-factor. Then we must have $b = \Omega(n^{r-1}/t^{r/2-1})$.*

In particular, 'close' to the hitting time τ_{K_r} (that is, if $t \leq n^{2-2/r+o(1)}$), the budget cannot be improved by a polynomial factor from the trivial strategy (that is, $b \geq n^{2-2/r-o(1)}$). Moreover, as shown in Figure 1, when considering $\log_n b$ as a function of $\log_n t$, we observe a linear interpolation between the trivial strategy at the hitting time and the trivial strategy that uses linear budget with quadratic time (namely, buy only edges belonging to a fixed K_r-factor). For $r \geq 3$, this provides a negative answer to Theorem 2.1 in a strong sense: for some graphs H, a linear budget cannot suffice unless t is quadratic.

The proof of Theorem 3.1 is based upon simple but crucial ways of regarding the restrictions on budget and time, respectively. Indeed, the budget b imposes a maximum degree constraint upon a 'large' set V of *normal* vertices, which intuitively allows us to bound the number of copies of spanning trees of K_r contained in V in the bought graph B_t. Given such a spanning tree, the time t allowed represents a bound on the probability that the remaining $\binom{r}{2} - r + 1$ edges required to complete the tree to a copy of K_r are ever presented. The proof itself requires a little more subtlety, as formalising this argument depends upon the order in which the edges of K_r appear in the process; the details may be found in our paper [5].

In fact, our result is not limited to clique factors, and this same proof naturally gives a lower bound on b for any successful strategy that creates (partial) F-factors, for any fixed graph F. This extends the negative answer to Theorem 2.1 to a large class of graphs. Indeed, for any bounded-degree graph H which itself contains a (partial) F-factor for some fixed graph F satisfying a mild condition, our main result guarantees that there are no successful (t, b)-strategies for H with linear budget when t is 'close' to the hitting time for H, and in fact that a linear budget cannot suffice unless t is quadratic.

In addition, we complement Theorem 3.1 by showcasing a successful (t, b)-strategy matching the lower bound on the budget up to a constant factor, demonstrating that our results are tight.

Theorem 3.2. *Let $r \geq 2$ and $\alpha \in (0, 1)$. For every $\delta > 0$, there exists some $K > 0$ such that, if $t = t(n) \geq Kn^{2-2/r}$ with $t \leq n^{2-\delta}$, there exists a $b = b(n) = O(n^{r-1}/t^{r/2-1})$ such that there exists a successful (t, b)-strategy for an α-K_r-factor.*

The strategy is in fact very simple: we fix a partition of the vertex set into a suitable number of equally-sized parts, and buy only edges contained within each part for the allocated time t. Observe that it suffices to find a (partial) factor in each part. The smaller we choose the parts, the further we can reduce the budget, but we must choose them large enough that a.a.s. the desired factors appear within time t; this bound yields a budget matching Theorem 3.1 (up to a constant factor).

The same strategy works for complete factors, at the expense of increasing the time and budget by a logarithmic factor. Since Theorem 3.1 immediately implies a lower bound for complete factors, this provides an essentially tight result in this case too. Moreover, as with the lower bounds, our results here are much more general, and hold for (partial) F-factors for large families of graphs F; again, see [5] for the precise statements.

As mentioned, our main result yields lower bounds for building any graph containing linearly many vertex-disjoint cycles of constant length, such as powers of Hamilton cycles; this provides a negative answer to Theorem 2.2 in a strong sense. There has been substantial research into the containment of the k-th power of a Hamilton cycle in random graphs [6, 12, 15, 18, 19], including very recent work determining sharp thresholds [17, 21], and the hitting time is known to be $\Theta(n^{2-1/k})$ a.a.s. for all $k \geq 2$. As with factors, these are therefore a very natural object of study in the budget-constrained random graph process. We adapt a method of Kühn and Osthus [15] to obtain a successful strategy showing that our lower bound is essentially tight, as in the following theorem.

Theorem 3.3. *Fix an integer $k \geq 2$. For every $t = \omega(n^{2-1/k})$, there exists some $b = n^{2k-1+o(1)}/t^{k-1}$ with a successful (t, b)-strategy for the k-th power of a Hamilton cycle. Conversely, for any successful (t, b)-strategy to exist, we must have $b \geq n^{2k-1-o(1)}/t^{k-1}$.*

The proof of the upper bound extends the ideas used for K_r-factors into a multi-stage strategy, and relies on the absorption method to show that Builder indeed constructs the k-th power of a Hamilton cycle. The formalisation of the proof is much more complex than the previous, and relies on several technical probabilistic arguments. All details can be found in [5].

References

[1] N. Alon and R. Yuster, Threshold functions for H-factors. *Comb. Probab. Comput.* 2.2 (1993), 137–144, DOI: 10.1017/S0963548300000559.

[2] M. Anastos, Constructing Hamilton cycles and perfect matchings efficiently. *arXiv e-prints* (2022). arXiv: 2209.09860.

[3] F. Bughart, A. Heckel, M. Kaufmann, N. Müller and M. Pasch, Sharp Thresholds for Factors in Random Graphs. *arXiv e-prints* (2024). arXiv: 2411.14138.

[4] F. Bughart, M. Kaufmann, N. Müller and M. Pasch, The hitting time of nice factors. *arXiv e-prints* (2024). arXiv: 2409.17764.

[5] A. Espuny Díaz, F. Garbe, T. Naia and Z. Smith, Graph factors and powers of Hamilton cycles in the budget-constrained random graph process. *arXiv e-prints* (2025). arXiv: 2502.20870.

[6] M. Fischer, N. Škorić, A. Steger and M. Trujić, Triangle resilience of the square of a Hamilton cycle in random graphs. *J. Comb. Theory, Ser. B* 152 (2022), 171–220, DOI: 10.1016/j.jctb.2021.09.005.

[7] A. Frieze, M. Krivelevich and P. Michaeli, Fast construction on a restricted budget. *arXiv e-prints* (2022). arXiv: 2207.07251.

[8] S. Gerke and A. McDowell, Nonvertex-balanced factors in random graphs. *J. Graph Theory* 78.4 (2015), 269–286, DOI: 10.1002/jgt.21805.

[9] A. Heckel, M. Kaufmann, N. Müller and M. Pasch, The hitting time of clique factors. *Random Struct. Algorithms* 65.2 (2024), 275–312, DOI: 10.1002/rsa.21218.

[10] D. Iľkovič, J. León and X. Shu, Multi-cyclic graphs in the random graph process with restricted budget. *arXiv e-prints* (2024). arXiv: 2412.17620.

[11] A. Johansson, J. Kahn and V. Vu, Factors in random graphs. *Random Struct. Algorithms* 33.1 (2008), 1–28, DOI: 10.1002/rsa.20224.

[12] J. Kahn, B. Narayanan and J. Park, The threshold for the square of a Hamilton cycle. *Proc. Amer. Math. Soc.* 149.8 (2021), 3201–3208, DOI: 10.1090/proc/15419.

[13] K. Katsamaktsis and S. Letzter, Building graphs with high minimum degree on a budget. *European J. Comb.* (in press), 104 119, DOI: 10.1016/j.ejc.2025.104119.

[14] J. Komlós and E. Szemerédi, Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.* 43 (1983), 55–63, DOI: 10.1016/0012-365X(83)90021-3.

[15] D. Kühn and D. Osthus, On Pósa's conjecture for random graphs. *SIAM J. Discrete Math.* 26.3 (2012), 1440–1457, DOI: 10.1137/120871729.

[16] L. Lichev, d-connectivity of the random graph with restricted budget. *J. Graph Theory* 108.2 (2025), 293–312, DOI: 10.1002/jgt.23180.

[17] T. Makai, M. Pasch, K. Petrova and L. Schiller, Sharp thresholds for higher powers of Hamilton cycles in random graphs. *arXiv e-prints* (2025). arXiv: 2502.14515.

[18] R. Nenadov and N. Škorić, Powers of Hamilton cycles in random graphs and tight Hamilton cycles in random hypergraphs. *Random Struct. Algorithms* 54.1 (2019), 187–208, DOI: 10.1002/rsa.20782.

[19] O. Riordan, Spanning subgraphs of random graphs. *Comb. Probab. Comput.* 9.2 (2000), 125–148, DOI: 10.1017/S0963548399004150.

[20] A. Ruciński, Matching and covering the vertices of a random graph by copies of a given graph. *Discrete Math.* 105.1-3 (1992), 185–197, DOI: 10.1016/0012-365X(92)90141-2.

[21] M. Zhukovskii, Sharp thresholds for spanning regular subgraphs. *arXiv e-prints* (2025). arXiv: 2502.14794.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Simple Graph Covers of Dipoles with Loops

(Extended abstract)

Filip Filipi*

Abstract

Motivated by the notion of a covering space in topology, a graph A covers a graph B if there is a locally bijective homomorphism from A to B. The H-Cover problem asks whether an input graph covers a graph H. According to the current trends, graphs having multiple edges, loops and semi-edges can be considered but the Strong Dichotomy Conjecture of Bok et al. [2022] asserts that the complexity of this problem is already determined by input graphs which are simple (contain no multiple edges, loops or semi-edges). A quasi-order on connected graphs was introduced in [Computational Complexity of Covering Disconnected Multigraphs. FCT 2021]: A connected graph A is stronger than a connected graph B if every simple graph covering A also covers B.

In [Kratochvíl: Towards strong dichotomy of graph covers, GROW 2022 - Book of open problems, p. 10, https://grow.famnit.upr.si/GROW-BOP.pdf], it was conjectured that if A has no semi-edges, then A is stronger than B if and only if A covers B. In the talk of Kratochvíl and Nedela presented at EUROCOMB'23, this conjecture was proved for any A provided that B is a cubic graph on a single vertex, and for any B provided that A is a graph on two vertices joined by some number of parallel edges.

In a follow-up to their work, we completely characterize the being stronger relation in the cases when A is any graph on a single vertex and when A is a graph on two vertices with no semi-edges. As a consequence, we have proved the conjecture stated above for any graph A on at most two vertices. Our methods can be further generalized to graphs with semi-edges to provide simple regular graphs with certain structural restrictions. Such graphs can be viewed as generalized snarks.

1 Preliminaries

As is very well described in [4], there are two main approaches to defining a multigraph with loops and semi-edges - the standard approach via vertices and generalized edges called links, and via darts and partitions on them. Both of these formalisms have their advantages, but here we will stick to the standard one.

*Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. E-mail: `filip.filipi@matfyz.cuni.cz`.

Definition 1. *A (finite)* graph *is a finite set of* vertices *and a finite set of* links *such that each link is either*

- *an* edge*, then it is incident with two different vertices and it adds one to their degrees,*
- *a* loop*, then it is incident with a single vertex and it adds two to its degree,*
- *a* semi-edge*, then it is incident with a single vertex and it adds one to its degree.*

For a graph G, by $V(G), \Lambda(G)$ we denote its set of vertices and links, respectively. Moreover, G is called simple *if it has no loops, no semi-edges, and for each pair of different vertices there is at most one edge incident with both of them.*

A graph H is a *subgraph* of a graph G if $V(H) \subseteq V(G)$ and $\Lambda(H) \subseteq \Lambda(G)$ in such a way that the type and incidence of each link of H are inherited from G. Given a set of vertices $U \subseteq V(G)$, by $G[U]$ we denote the *induced subgraph* of the graph G on U, i.e., $G[U]$ is the subgraph of G such that $V(G[U]) = U$ and each link of G which has all its incident vertices in U is contained in $\Lambda(G[U])$. We say that a set F of links in G is *incident* with a vertex u in G if there is at least one link in F incident with u. We say that the set F *spans* the subgraph H of G if $F \subseteq \Lambda(H)$, and F is incident with every vertex of H. We say that sets of links $F_1, \ldots, F_m$ in G are (pairwise) *vertex-disjoint* if for every vertex in G, at most one of the sets F_i is incident with it.

For the graph G, we now briefly denote its substructures used in this abstract. A *matching* is a set of edges in G such that no two edges share an incident vertex. Similarly, a *semi-matching* is a set of edges and/or semi-edges in G such that no two of these links share an incident vertex. A matching/semi-matching is called *perfect* in G whenever it spans G.

If $n \in \mathbb{N}$, $v_1, \ldots, v_n$ are pairwise different vertices in G, and $e_1, \ldots, e_{n-1}$ are edges in G such that for each $i = 1, \ldots, n-1$, the edge e_i is incident with the vertices v_i and v_{i+1}, then we say that $P := \{e_i \mid i = 1, \ldots, n-1\}$, the set of these edges, is a *closed path* in G with *endpoint vertices* v_1 and v_n (it may happen that $v_1 = v_n$ when $n = 1$). Moreover, if $\lambda_n \notin P$ is an edge or a loop in G incident with vertices v_1 and v_n, then we say that $P \cup \{\lambda_n\}$ is a *cycle* in G. Note that both a single loop and a pair of parallel edges also form cycles. We say that the graph G is *connected*, if for any pair of its vertices u, v, there exists a closed path in G with endpoints u and v.

An elegant definition of a covering projection is provided by the formalism of darts that we chose not to use. Using the standard model, the formal definition is rather technical, so in Definition 2 we define it by the means of preimages via conditions given by [4, Proposition 4]. Note that the omitted incidence preservation is a consequence of the other conditions[1]. The perspective provided by this definition is the one we use most often when verifying that something really is a covering projection.

Definition 2. *A* covering projection $\mathbf{f}$ *from a graph G to a connected graph H is a pair of surjective*[2] *maps $\mathbf{f}_V : V(G) \to V(H)$ and $\mathbf{f}_\Lambda : \Lambda(G) \to \Lambda(H)$ such that the preimage $\mathbf{f}_\Lambda^{-1}$ of*

- *a semi-edge incident with a vertex v is a semi-matching spanning the graph $G\left[\mathbf{f}_V^{-1}[v]\right]$,*

[1] Proof of the contrapositive: If $\lambda := \mathbf{f}_\Lambda(\mu)$ is not incident with $\mathbf{f}_V(u)$, then $\mu \in \mathbf{f}_\Lambda^{-1}[\lambda] \subseteq \Lambda\left(G\left[\mathbf{f}_V^{-1}[U]\right]\right)$ for U the set of incident vertices of λ. Since $\mathbf{f}_V(u) \notin U$, we get that $u \notin \mathbf{f}_V^{-1}[U]$; thus, μ is not incident with u.

[2] Surjectivity of these maps is implied by the other conditions of the definition whenever G is non-empty.

- *a loop incident with a vertex v is a union of vertex-disjoint cycles and it spans the graph $G\left[\mathbf{f}_V^{-1}[v]\right]$,*
- *an edge incident with two different vertices u, v is a matching which spans the graph $G\left[\mathbf{f}_V^{-1}[u] \cup \mathbf{f}_V^{-1}[v]\right]$ in such a way that each of its edges is incident with one vertex of $\mathbf{f}_V^{-1}[u]$ and one vertex of $\mathbf{f}_V^{-1}[v]$.*

Although it would make sense to omit the connectedness assumption, we will not do that because there are multiple different notions of covering projections with disconnected target [4], and this would be just one of them.

Covering projections are locally bijective, incidence preserving, and degree preserving. By $G \to H$, we shall denote that there exists a covering projection from a graph G to a connected graph H, and we will say that G *covers* H. If the connected graph H is fixed, the H-Cover decision problem asks whether a given input graph covers the graph H.

The notion of the covering projection was used, for example, to construct highly symmetric graphs with certain properties [2]. The computational complexity of the H-Cover problem for graphs with multiple edges and loops was first studied in [1]. Semi-edges were later considered in [3]. During further research on the H-Cover problem in [4], the following notion emerged.

Definition 3. *Let A and B be connected graphs. We say that A is* stronger *than B, denoted by $A \rhd B$, if it holds that every simple graph G covering A also covers B.*

The relation $\rhd$ is transitive, and so it defines a quasi-order on connected graphs. Moreover, since the composition of covering projections is again a covering projection, for every simple graph A it holds that $A \rhd B$ if and only if $A \to B$. A conjecture was proposed by Kratochvíl in [5].

Conjecture 1. *Let A and B be connected graphs. If A has no semi-edges, then $A \rhd B$ if and only if $A \to B$.*

2 Main results

We adopt the following notation from [6].

Notation. By $F(s,\ell)$, we denote the graph on a single vertex with $s \in \mathbb{N}_0$ semi-edges and $\ell \in \mathbb{N}_0$ loops, and by $W(s_1,\ell_1,b,\ell_2,s_2)$, we denote the graph on two vertices formed by joining $F(s_1,\ell_1)$ and $F(s_2,\ell_2)$ using $b \in \mathbb{N}_0$ parallel edges. Whenever $b \geq 1$, the graph $W(s_1,\ell_1,b,\ell_2,s_2)$ is connected and will be called a (generalized) *dipole*. Moreover, the number b will be referred to as the number *bars* of this dipole.

At EUROCOMB'23 conference, Kratochvíl and Nedela [6] presented the following progressions towards Conjecture 1. If B is a cubic graph on a single vertex, that is $B = F(1,1)$ or $B = F(3,0)$, then for any connected graph A with no semi-edges it holds that $A \rhd B$ if and only if $A \to B$. On the other hand, if $A = W(0,0,b,0,0)$ for some $b \in \mathbb{N}$, then for any connected graph B it holds that $A \rhd B$ if and only if $A \to B$.

In Theorem 3, we extend the latter result to all dipoles with no semi-edges. Additionally, in Theorem 4, we fully describe how the relation $A \rhd B$ behaves for the general case of $A = F(s,\ell)$. By doing this, we especially prove that Conjecture 1 holds for all connected graphs A on at most two vertices.

Theorem 1 (Generalization of [6, Theorem 1])**.** *Let A and B be connected graphs such that $A \triangleright B$. Then*

1. *$|V(B)|$ divides $2\,|V(A)|$. Moreover, if B has a semi-edge, then $|V(B)|$ divides $|V(A)|$.*
2. *if A has no semi-edges, then $|V(B)|$ divides $|V(A)|$. Moreover, if B has a semi-edge, then $2\,|V(B)|$ divides $|V(A)|$.*

The following theorem is a crucial tool used in all the upcoming results.

Theorem 2. *Let A be a connected graph and let there be an odd set of vertices $S \subseteq V(A)$ such that $A[S]$ has no semi-edges, and there are only k edges in A with one incident vertex in S and the other one in $V(A) \setminus S$. Then there is a simple graph which covers A, and it has at most k pairwise disjoint perfect matchings.*

We now verify that Conjecture 1 holds for all connected graphs A on two vertices.

Theorem 3. *Let A and B be connected graphs. If A is a dipole with no semi-edges, then $A \triangleright B$ if and only if $A \to B$.*

Outline of the proof. We use Theorem 1 to limit the size of B such that $A \triangleright B$. If B has only one vertex, then Theorem 2 and the degree preservation of covering projections suffice. Otherwise, B has to be a dipole having no semi-edges. According to the following lemma, together with the fact that covering projections preserve degrees, by constructing a simple graph G which covers A but $|V(G)|$ is not divisible by 4, we show that $A = B$, in particular, $A \to B$.

Lemma 1. *Let A and B be dipoles having different number of bars. If G is a simple graph such that there are covering projections $\mathbf{f} : G \to A$ and $\mathbf{g} : G \to B$, then for any vertex u in A, and any vertex v in B, it holds that*

$$\left|\mathbf{f}_V^{-1}[u] \cap \mathbf{g}_V^{-1}[v]\right| = \frac{|V(G)|}{4}.$$

□

For graphs A on a single vertex (with possible semi-edges), a complete characterization of the relation $A \triangleright B$ for arbitrary connected graphs B can be stated as follows. As its consequence, Conjecture 1 also holds for all graphs A on a single vertex.

Theorem 4. *Let A and B be connected graphs. If A has only one vertex, then $A \triangleright B$ if and only if one of the following conditions is satisfied.*

- *$A = F(s, \ell)$ and $B = F(s', \ell')$ in such a way that $s + 2\ell = s' + 2\ell'$ and $s' \leq s$,*
- *$A = F(1, 0)$ and $B = W(0, 0, 1, 0, 0)$,*
- *$A = F(2, 0)$ and $B = W(0, 0, 2, 0, 0)$.*

Outline of the proof. We first use Theorem 1 to limit the size of B such that $A \triangleright B$. If B has only one vertex and $A = F(s, \ell)$, then we use the degree preservation of covering projections to show that the degrees of the vertices in A and B are equal. Then we apply Theorem 2 onto $A' := W(0, \ell, s, \ell, 0)$, a graph which covers A, to limit the number of semi-edges in B. The remaining cases are given by Theorem 4, and for those, $A \triangleright B$ holds. Otherwise, B has to be a dipole with no semi-edges, and we make use of an inductive argument provided by the following lemma.

Lemma 2. *If* $F(s,\ell) \not\triangleright W(s_1,\ell_1,b,\ell_2,s_2)$*, then also* $F(s,\ell+1) \not\triangleright W(s_1,\ell_1+1,b,\ell_2+1,s_2)$*.*

We then figure out what the initial steps of these inductions are, and construct ad hoc witnesses for them. This reduces the problem to the cases listed in Theorem 4 for which we prove that $A \triangleright B$ truly holds. $\square$

3 Conclusion and Final Remarks

We have proved that Conjecture 1 holds for connected graphs A (with no semi-edges) on at most two vertices. In addition, for graphs A on a single vertex, we have fully characterized all pairs A,B such that $A \triangleright B$. The methods used have proven to be useful even in cases when A contains semi-edges. For example, applying Theorem 2 to a graph A' constructed by taking two disjoint copies of A and pairing all the corresponding semi-edges into edges, we obtain a simple graph which covers A'. Since $A' \to A$, it will also cover A. We do hope for the full characterization of pairs $A \triangleright B$ for graphs A on two vertices with possible semi-edges, but for the time being, this is a work in progress.

4 Acknowledgment

The findings presented here are part of the results of my Master's thesis, and I am grateful to Jan Kratochvíl for his dedicated supervision.

References

[1] James Abello, Michael R. Fellows, and John C. Stillwell. On the complexity and combinatorics of covering finite complexes. *Australian Journal of Combinatorics*, 4:103–112, 1991.

[2] Norman Biggs. *Algebraic graph theory.* Cambridge University Press, 1974.

[3] Jan Bok, Jiří Fiala, Petr Hliněný, Nikola Jedličková, and Jan Kratochvíl. Computational complexity of covering multigraphs with semi-edges: Small cases. In *46th International Symposium on Mathematical Foundations of Computer Science (MFCS 2021)*, pages 21:1–21:15. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2021.

[4] Jan Bok, Jiří Fiala, Nikola Jedličková, Jan Kratochvíl, and Michaela Seifrtová. Computational complexity of covering disconnected multigraphs. *Discrete Applied Mathematics*, 359:229–243, 2024.

[5] Jan Kratochvíl. Towards strong dichotomy of graph covers. *GROW 2022 - Book of Open Problems*, 2022.

[6] Jan Kratochvíl and Roman Nedela. Graph covers and generalized snarks. In *European Conference on Combinatorics, Graph Theory and Applications*, pages 681–687, 2023.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Helly-type theorems for monotone properties of boxes

(Extended abstract)

Nóra Frankl* Attila Jung†

Abstract

We present a unified approach to prove Helly-type theorems for monotone properties of boxes such as having large volume or containing points with integer coordinates. As a corollary, we obtain short new proofs for several earlier results regarding specific monotone properties.

Helly's theorem [Hel23] states that if in a finite family of convex sets in $\mathbb{R}^d$ the intersection of any $d+1$ members is nonempty, then the intersection of the whole family is nonempty. This is a cornerstone result in Discrete Geometry with many generalisations and extensions, see for example the survey [BK22].

For special families of convex sets often stronger results hold. Among such families, axis parallel boxes (called simply *boxes* from now on) received particular attention [Kat80, DG82, Eck88, Eck91, ES24]. It is folklore that requiring non-empty pairwise intersections in a family of boxes guarantees a common point in the whole family, regardless of the dimension [Eck88]. Indeed, project the family to each coordinate axes, and apply Helly's theorem in each of the d families of intervals obtained this way.

The quantitative volume theorem of Bárány, Katchalski and Pach [BKP82] states that if in a finite family $\mathcal{F}$ of convex sets the intersection of any $2d$ members has volume at least 1, then the intersecion of the whole family is of volume at least $v(d)$, for some positive function $v(d)$. Naszódi proved that $v(d)$ can be chosen to be of order d^{-cd} [Nas16]. Although the number $2d$ is best possible even for families of boxes, the volume bound can be improved to $v(d)=1$ in this case. This is a direct corollary of our first observation.

Observation 1. *For any finite family $\mathcal{F}$ of boxes in $\mathbb{R}^d$, there is a subfamily $\mathcal{F}' \subset \mathcal{F}$ of size at most $2d$ such that $\cap\mathcal{F} = \cap\mathcal{F}'$.*

*School of Mathematics and Statistics, The Open University, UK. E-mail: nora.frankl@open.ac.uk. Partially supported by ERC Advanced Grant no. 882971 "GeoScape" at the beginning of this work.

†ELTE Eötvös Loránd University and HUN-REN Alfréd Rényi Institute of Mathematics, Hungary. Email: jungattila@gmail.com. Supported by the ERC Advanced Grant "ERMiD", by the EXCELLENCE-24 project no. 151504 of the NRDI Fund, and by the NKFIH grants TKP2021-NKTA-62, FK132060 and SNN135643.

We call a property P of boxes a *monotone property* if $B \in P$ and $B \subset C$ implies $C \in P$. Observation 1 implies results not only with respect to the volume, but also with respect to any monotone property.

Corollary 1. *For any monotone property P and any finite family $\mathcal{F}$ of boxes in $\mathbb{R}^d$ if the intersection of every subfamily of $\mathcal{F}$ of size $2d$ has property P, then the intersection of all the members of $\mathcal{F}$ has property P.*

If P is the property of containing a point from a given finite set, we obtain discrete Helly theorems for boxes proved by Halman [Hal08] and in more general forms by Edwards and Soberón [ES24]. Proofs of various earlier Helly-type results for boxes were specific to given monotone properties. In this note we present results similar to Observation 1 from which many of these results follow at once.

We call a family $\mathcal{B}$ of boxes *P-intersecting* if the intersection of its members has property P. In this language Corollary 1 states that if every subfamily of size $2d$ is P-intersecting, then the whole family is P-intersecting.

Colourful results

Let $\mathcal{F}_1, \ldots, \mathcal{F}_{d+1}$ be $d+1$ finite families of convex sets. The Colourful Helly theorem of Lovász (first published by Bárány [Bár82]) states that if for any choice $C_i \in \mathcal{F}_i$ we have $\cap_{i=1}^{d+1} C_i \neq \emptyset$, then there exists an $i \in [d+1]$ such that $\cap \mathcal{F}_i \neq \emptyset$. As the families are not necessarily distinct, this result generalises Helly's theorem. The Colourful Helly theorem is optimal in the sense that an analogous statement is not true for fewer than $d+1$ families. This can be shown by a construction using hyperplanes in general position.

We show the following strong intersection property of boxes, which generalises Observation 1 and implies a Colourful Helly theorem for monotone properties.

Theorem 2. *Let $\mathcal{B}_1, \ldots, \mathcal{B}_{2d}$ be finite families of boxes in $\mathbb{R}^d$. Then there is a selection $B_i \in \mathcal{B}_i$ for each $i \in [2d]$ and an index $\ell \in [2d]$ such that $\cap_{i=1}^{2d} B_i \subset \cap \mathcal{B}_\ell$.*

Corollary 3. *Let P be any monotone property of boxes and let $\mathcal{B}_1, \ldots, \mathcal{B}_{2d}$ be finite families of boxes in $\mathbb{R}^d$. If for every choice $B_i \in \mathcal{B}_i$ the family $\{B_1, \ldots, B_{2d}\}$ is P-intersecting, then there exists an $\ell \in [2d]$ such that $\mathcal{B}_\ell$ is P-intersecting.*

The following construction shows that the parameter $2d$ is best possible in Corollary 3 when P is the property of having volume at least 1, hence best possible in Theorem 2. Let $\mathcal{F}$ be a family of $2d$ halfspaces whose intersection is a box of volume $\varepsilon < 1$. Replacing these halfspaces with their intersections with a large cube, we obtain a family of $2d$ boxes such that the intersection of any $2d-1$ of them has volume at least 1, but the intersection of all of them has volume $\varepsilon < 1$.

Damásdi, Földvári and Naszódi [DFN21] proved a volumetric version of the Colourful Helly theorem for $3d$ families of convex sets. If property P is having volume at least 1, Corollary 3 implies a volumetric Colourful Helly theorem for $2d$ families of boxes with volume bound $v(d) = 1$. If P is containing at least n points from a given finite set, we obtain a result of Edwards and Soberon [ES24, Theorem 1.2].

Fractional results

The fractional Helly theorem of Katchalski and Liu [KL79] generalises Helly's theorem by showing that if $\mathcal{F}$ is a finite family of convex sets from $\mathbb{R}^d$ such that for some $\alpha > 0$, at least $\alpha\binom{|\mathcal{F}|}{d+1}$ of the $(d+1)$-tuples of $\mathcal{F}$ have a nonempty intersection, then there is a subfamily $\mathcal{F}' \subset F$ of size $|\mathcal{F}'| \geq \beta(\alpha, d)|\mathcal{F}|$ with nonempty intersection. Variants for general convex sets include a version with respect to containing lattice points by Bárány and Matoušek [BM03], or having large volume first proved by Sarkar , Xue and Soberón [SXS21] for intersecting $\frac{d(d+3)}{2}$-tuples and improved in [FJT24] for intersecting $(d+1)$-tuples.

We prove two variants for boxes with respect to monotone properties. The first of these is stronger in the sense that it assumes that a positive fraction of the $d+1$ tuples has property P, as opposed to assuming it for $2d$ tuples in the second. However, in the second version β is more optimal.

Theorem 4. *Let P be a monotone property of boxes. For every dimension d and every real number $\alpha > 0$ there exists $\beta > 0$ such that the following holds. Let $\mathcal{B}$ be a family of boxes in $\mathbb{R}^d$. If at least $\alpha\binom{|\mathcal{B}|}{d+1}$ of the $(d+1)$-tuples in $\mathcal{B}$ are P-intersecting, then there exists a P-intersecting subfamily $\mathcal{B}' \subseteq \mathcal{B}$ of size at least $\beta|\mathcal{B}|$.*

If P is the property of containing a point from a fixed set, Theorem 4 implies Theorem 1.3 of [ES24] with a shorter proof. For the same property, Theorem 3.2 of [ES24] is a fractional result with parameter $2d$. Although the bound for β in our Theorem 5 is slightly weaker, it still converges to 1 if α does so.

Theorem 5. *Let P be a monotone property of boxes, and let $\mathcal{B}$ be a finite family of boxes in $\mathbb{R}^d$ such that at least $\alpha\binom{|\mathcal{B}|}{2d}$ of the $2d$-tuples of $\mathcal{B}$ are P-intersecting. Then for $\beta = 1 - 2d(1-\alpha)^{1/(2d+1)}$ there is a P-intersecting subfamily $\mathcal{B}' \subseteq \mathcal{B}$ of cardinality at least $\beta|\mathcal{B}|$.*

Unlike in the case of Helly's theorem for boxes, the parameter $(d+1)$ in the fractional version is optimal for boxes for a general α. Indeed, let $\mathcal{F}_i$ be a family of $\lfloor n/d \rfloor$ or $\lceil n/d \rceil$ hyperplanes orthogonal to the i-th coordinate axis. Intersecting the union of these families with a large cube, we obtain a family of n boxes where $d+1$ cannot be lowered for $\alpha = \binom{n}{d}/(n/d)^d$. By using thick hyperplanes, we obtain a construction with proper d-dimensional boxes. However, Katchalski [Kat80] showed that when α is sufficiently large, we can decrease $d+1$ to 2. Going beyond this, Eckhoff [Eck88], confirming a conjecture of Kalai [Kal84], determined for each $1 \leq k \leq d+1$ the smallest α for which if $\alpha\binom{|\mathcal{F}|}{k}$ many k-tuples of members of a family of boxes $\mathcal{F}$ intersect, then there is a subset $\mathcal{F}' \subseteq \mathcal{F}$ with $|\mathcal{F}'| \geq \beta(\alpha, d)|\mathcal{F}|$ such that $\cap\mathcal{F}' \neq \emptyset$. As a first step towards generalising Eckhoff's result to monotone properties, we prove an analogue of Katchalki's result.

Theorem 6. *Let P be a monotone property of boxes. For every dimension d there exists a $c_d > 0$ such that for every $\alpha \in (1 - c_d, 1]$ there exists a $\beta > 0$ such that the following holds.*

Let $\mathcal{B}$ be a family of boxes in $\mathbb{R}^d$. If at least $\alpha\binom{|\mathcal{B}|}{2}$ of the pairs in $\mathcal{B}$ are P-intersecting, then there exists a P-intersecting subfamily $\mathcal{B}' \subseteq \mathcal{B}$ of size at least $\beta|\mathcal{B}|$.

We note that in Katchalski's result the values c_d and β are best possible, whereas we don't have any reason to think that they are optimal in our result.

During the preparation of our manuscript, a paper of Eom, Kim and Lee appeared where they prove Theorem 6 independently in the case when P is the property that a box contains an element of a fixed finite sets $S \subset \mathbb{R}^d$ [EKL25].

1 Proofs and proof sketches

Every box B in $\mathbb{R}^d$ is the intersection of $2d$ halpfspaces $H_1^B, \ldots, H_{2d}^B$. We index the halfspaces so that for any two boxes B_1 and B_2 and for any index i, the halfspaces $H_i^{B_1}$ and $H_i^{B_2}$ are translates of each other. Our proofs rely on introducing $2d$ orderings on the set of boxes $<_1, \ldots, <_{2d}$ defined by these halfspaces as follows. For two boxes B_1, B_2 we set $B_1 <_i B_2$ if $H_i^{B_1} \subseteq H_i^{B_2}$. Using these orderings, the proofs of Theorem 2 and Theorem 5 are very short.

Proof of Theorem 2. Iteratively define a permutation $\pi \in S_{2d}$ such that $\mathcal{B}_{\pi(i)}$ contains a minimal element $B_{\pi(i)}$ of $\cup_{j=1}^{2d}\mathcal{B}_j \setminus (\cup_{j=1}^{i-1}\mathcal{B}_{\pi(j)})$ according to $<_i$. Then $\cap_{i=1}^{2d} B_{\pi(i)} \subseteq \cap\mathcal{B}_{\pi(2d)}$, so we can choose $\{B_{\pi(1)}, \ldots, B_{\pi(2d)}\}$ and $\ell = \pi(2d)$. □

Proof of Theorem 5. Let $\gamma = (1-\alpha)^{1/(2d+1)}$. For each i let $\mathcal{A}_i \subseteq \mathcal{B}$ be the set of the first $\gamma|\mathcal{B}|$ boxes according to $<_i$. Then $|\{\{A_1, \ldots, A_{2d}\} : A_i \in \mathcal{A}_i\}\}| \geq \frac{1}{(2d)!}|\mathcal{A}_1| \cdot (|\mathcal{A}_2|-1)\ldots(|\mathcal{A}_{2d}| - 2d) = \binom{\gamma|\mathcal{B}|}{2d} > (1-\alpha)\binom{|\mathcal{B}|}{2d}$ if $|\mathcal{B}|$ is large enough[1], thus by assumption, there is a box A in $\{\cap_{i=1}^{2d} A_i : A_i \in \mathcal{A}_i\}$ that has property P. By the properties of the orderings, $A \subseteq B$ for any $B \in \mathcal{B}' := \mathcal{B} \setminus (\cup_{i=1}^{2d}\mathcal{A}_i)$, and $|\mathcal{B}'| \geq (1-2d\gamma)|\mathcal{B}| = (1-2d(1-\alpha)^{1/(2d+1)})|\mathcal{B}|$. □

The framework of the proof of our Fractional result with parameter $d+1$ is based on ideas from a paper pf Bárány and Matousek [BM03]. The proof of Theorem 4 follows from the next result via supersaturation as in [BM03].

Theorem 7. *[Weak Colourful Helly for boxes] Let $\mathcal{B}_1, \ldots, \mathcal{B}_{d+1}$ be finite families of boxes in $\mathbb{R}^d$ with $|\mathcal{B}_1| = \ldots = |\mathcal{B}_{d+1}| \geq 2^{2d^2+1}$. Then there is a selection $B_i \in \mathcal{B}_i$ for each $i \in [d+1]$ and an index $\ell \in [d+1]$ such that there is a subset $\mathcal{B}'_\ell \subseteq \mathcal{B}_\ell$ with $|\mathcal{B}'_\ell| \geq 2^{-2d^2-1}|\mathcal{B}_\ell|$ and $\cap_{i=1}^{d+1} B_i \subset \cap\mathcal{B}'{}_\ell$.*

We say that two ordered pairs of boxes (B_1, B_2) and (B'_1, B'_2) are *ordered consistently*, if for each $k \in [2d]$ we have $B_1 <_k B_2$ holds if and only if $B'_1 <_k B'_2$. An ordered set $\mathcal{B} = (B_1, \ldots, B_k)$ of boxes is ordered consistently, if for any $i < j$ and $i' < j'$ the ordered pairs (B_i, B_j) and $(B_{i'}, B_{j'})$ are ordered consistently. An ordered pair $(\mathcal{F}_1, \mathcal{F}_2)$ of sets of boxes is ordered consistently, if for any $B_1, B'_1 \in \mathcal{F}_1$, $B_2, B'_2 \in \mathcal{F}_2$ the pairs $(B_1, B_2), (B'_1, B'_2)$ are ordered consistently. The proof of Theorem 7 uses the following simple claim whose proof we omit.

Claim 8. *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two families boxes in $\mathbb{R}^d$ with $|\mathcal{B}_1| = m$, $|\mathcal{B}_2| = n$. Then for any $k \in [2d]$ there exist $\mathcal{B}'_1 \subseteq \mathcal{B}_1$ and $\mathcal{B}'_2 \subseteq \mathcal{B}_2$ with $|\mathcal{B}'_1| = \lfloor \frac{m}{2} \rfloor$, $|\mathcal{B}'_2| = \lfloor \frac{n}{2} \rfloor$ such that we either have $B_1 <_k B_2$ for all $B_1 \in \mathcal{B}'_1, B_2 \in \mathcal{B}'_2$, or we have $B_2 <_k B_1$ for all $B_1 \in \mathcal{B}'_1, B_2 \in \mathcal{B}'_2$.*

Proof of Theorem 7. By discarding at most half of the elements from each family $\mathcal{B}_i$, we may assume that the size of each of the families is a large power of 2 so that we can omit floors when halving them. By applying Claim 8 successively to each pair $\mathcal{B}_i, \mathcal{B}_j$ and each relation $<_k$, we obtain subfamilies $\mathcal{B}'_1 \subseteq \mathcal{B}_1, \ldots, \mathcal{B}'_{d+1} \subseteq \mathcal{B}_{d+1}$ with $|\mathcal{B}'_i| \geq 2^{-2d^2-1}|\mathcal{B}_i|$ such that each pair $(\mathcal{B}'_i, \mathcal{B}'_j)$ is ordered consistently. By discarding the excess, we may assume that $|\mathcal{B}'_i| = 2^{-2d^2-1}|\mathcal{B}_i|$.

Let $\mathcal{B}' = \cup_{i=1}^{d+1}\mathcal{B}'_i$. For each $k \in [2d]$ there is an i such that $\mathcal{B}'_i$ consists of exactly the first $\frac{1}{d+1}|\mathcal{B}'|$ members of $\mathcal{B}'$ according to $<_k$. Since there are $2d$ orderings and $d+1$ families,

[1] Note that in the lower bound we only count $\{A_1, \ldots, A_{2d}\}$ if it has size $2d$.

there is an $\ell \in [d+1]$ such that $\mathcal{B}_\ell$ is the family consisting of the first $\frac{1}{d+1}|\mathcal{B}'|$ members of $\mathcal{B}'$ according to at most one ordering, say $<_1$. Let $B_\ell \in \mathcal{B}_\ell$ be the minimal member of $\mathcal{B}_\ell$ according to $<_1$, and for all $i \neq \ell$ let $B_i \in \mathcal{B}_\ell$ be an arbitrary box. Since for all $k \in [2d]$, $k \neq 1$ there is an i such that all members of $\mathcal{B}_i$ are smaller than all members of $\mathcal{B}_\ell$ according to $<_k$, and B_ℓ is the minimal element of $\mathcal{B}_\ell$ according to $<_1$, we have $\cap_{i=1}^{d+1} B_i \subset \cap \mathcal{B}_\ell$. □

Proof of Theorem 6. By applying the Erdős-Szekeres theorem succesively for each of the $2d$ ordering of boxes, there exists an $N = N(d)$ such that among every N boxes there are $2d$ which are ordered consistently. Thus, there are at least

$$\frac{\binom{|\mathcal{B}|}{N}}{\binom{|\mathcal{B}|-2d}{N-2d}} = \frac{(N-2d)!(2d)!}{N!}\binom{|\mathcal{B}|}{2d}$$

distinct subfamilies of $\mathcal{B}$ of size $2d$ which are ordered consistently.

Since there can only be $(1-\alpha)\binom{n}{2}\binom{n-2}{2d-2} \leq (1-\alpha)2d^2\binom{n}{2d}$ different $2d$-tuples which contain a non-P-intersecting pair, by setting $c_d = \frac{(N-2d)!(2d)!}{2d^2 N!}$ and $\alpha' = \frac{(N-2d)!(2d)!}{N!} - (1-\alpha)2d^2 \geq 0$, there are at least $\alpha'\binom{n}{2d}$ many consistently ordered $2d$ tuples among which any pair is P-intersecting. The following proposition, whose proof we omit, implies that these $2d$-tuples are P-intersecting as well.

Proposition 9. *If $\mathcal{B} = \{B_1 <_1 \ldots <_1 B_k\}$ are ordered consistently, then $\cap_{i=1}^{k} B_i = B_1 \cap B_k$.*

Thus, the intersection of $\alpha'\binom{|\mathcal{B}|}{2d}$ of the $2d$-tuples of $\mathcal{B}$ are P-intersecting and we can apply Theorem 5. □

References

[Bár82] Imre Bárány. A generalization of Carathéodory's theorem. *Discrete Mathematics*, 40(2-3):141–152, 1982.

[BK22] Imre Bárány and Gil Kalai. Helly-type problems. *Bulletin of the American Mathematical Society*, 59(4):471–502, 2022.

[BKP82] Imre Bárány, Meir Katchalski, and János Pach. Quantitative Helly-type theorems. *Proceedings of the American Mathematical Society*, 86(1):109–114, 1982.

[BM03] Imre Bárány and Jiří Matoušek. A fractional helly theorem for convex lattice sets. *Advances in Mathematics*, 174(2):227–235, 2003.

[DFN21] Gábor Damásdi, Viktória Földvári, and Márton Naszódi. Colorful Helly-type theorems for the volume of intersections of convex bodies. *Journal of Combinatorial Theory, Series A*, 178:105361, 2021.

[DG82] Ludwig Danzer and Branko Grünbaum. Intersection properties of boxes in R^d. *Combinatorica*, 2:237–246, 1982.

[Eck88] Jürgen Eckhoff. Intersection properties of boxes. Part I: An upper-bound theorem. *Israel Journal of Mathematics*, 62:283–301, 1988.

[Eck91] Jürgen Eckhoff. Intersection properties of boxes part II: Extremal families. *Israel Journal of Mathematics*, 73:129–149, 1991.

[EKL25] Taehyun Eom, Minki Kim, and Eon Lee. Fractional discrete Helly for pairs in a family of boxes. *arXiv preprint arXiv:2503.11997*, 2025.

[ES24] Timothy Edwards and Pablo Soberón. Extensions of discrete helly theorems for boxes. *arXiv preprint arXiv:2404.14308*, 2024.

[FJT24] Nóra Frankl, Attila Jung, and István Tomon. The Quantitative Fractional Helly theorem. *arXiv preprint arXiv:2402.12268*, 2024.

[Hal08] Nir Halman. Discrete and lexicographic Helly-type theorems. *Discrete & Computational Geometry*, 39:690–719, 2008.

[Hel23] Ed Helly. Über Mengen konvexer Körper mit gemeinschaftlichen Punkten. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 32:175–176, 1923.

[Kal84] Gil Kalai. Intersection patterns of convex sets. *Israel Journal of Mathematics*, 48:161–174, 1984.

[Kat80] Meir Katchalski. Boxes in R^n—a 'fractional' theorem. *Canadian Journal of Mathematics*, 32(4):831–838, 1980.

[KL79] Meir Katchalski and Andy Liu. A problem of geometry in $\mathbb{R}^n$. *Proceedings of the American Mathematical Society*, 75(2):284–288, 1979.

[Nas16] Márton Naszódi. Proof of a conjecture of Bárány, Katchalski and Pach. *Discrete & Computational Geometry*, 55:243–248, 2016.

[SXS21] Sherry Sarkar, Alexander Xue, and Pablo Soberón. Quantitative combinatorial geometry for concave functions. *Journal of Combinatorial Theory, Series A*, 182:105465, 2021.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Metric properties of electrical networks and the graph reconstruction problems

(Extended abstract)

V. G. Gorbounov [*]   A.A. Kazakov [†]

Abstract

Using the generalized Temperley trick, we demonstrate the explicit embedding of circular electrical networks into totally non-negative Grassmannians. Building on this result, we show that the effective resistances between boundary nodes of circular electrical networks satisfy the Kalmanson property, and we provide the full characterization of planar electrical Kalmanson metrics. Also, we outline a graph reconstruction algorithm, which can be applied in phylogenetic network analysis and can be used for the numerical solution of the Calderón problem.

1 Introduction

Electrical network theory is now a well-established applied field, originated in the mid-19th to early 20th centuries in the work of Kirchhoff, Maxwell, Kennelly, Norton and many other physicists and engineers. From a mathematical point of view, this theory is a great source of a large number of highly non-trivial combinatorial and geometric results and structures, which appear and find their applications in the wide variety of mathematical fields from exactly solvable models in statistical physics [14], [11] from the famous Calderon problem [5] and inverse problems in phylogenetic network theory [13].

2 Preliminaries

Definition 2.1. An **electrical network** $\mathcal{E}(G, w)$ is a planar graph $G(V, E)$, embedded into a disk and equipped with a conductivity function $w : E(G) \to \mathbb{R}_{\geq 0}$, which together satisfy the following conditions:

[*]Faculty of Mathematics, National Research University Higher School of Economics, Usacheva 6, 119048 Moscow, Russia. E-mail: `vgorb10@gmail.com`.

[†]Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Russia, 119991, Moscow, GSP-1, 1 Leninskiye Gory, Main Building; Kazan Federal University, N.I. Lobachevsky Institute of Mathematics and Mechanics, Kazan, 420008, Russia. Supported by RSF (project No. 24-21-00158).

- All nodes are divided into the set of inner nodes V_I and the set of boundary nodes V_B;
- An edge weight $w(v_iv_j) = w_{ij}$ denotes the conductivity of this edge.

An electrical network is called **circular** if its boundary nodes lie on the boundary circle and are enumerated clockwise from 1 to $|V_B| := n$ (inner nodes are enumerated arbitrarily from $n+1$ to $|V|$).

Consider an electrical network $\mathcal{E}(G, w)$ and apply voltages $\mathbf{U} : V_B \to \mathbb{R}$ to its boundary nodes V_B. Then these boundary voltages induce the unique harmonic extension on all vertices $U : V \to \mathbb{R}$, which might be found out by Ohm's and Kirchhoff's laws:

$$\sum_{j \in V} w_{ij}\big(U(i) - U(j)\big) = 0, \ \forall i \in V_I.$$

One of the main objects associated with each harmonic extension of boundary voltages $\mathbf{U}$ are boundary currents $\mathbf{I} = \{I_1, \dots, I_n\}$ running through boundary nodes:

$$I_k := \sum_{j \in V} w_{ij}\big(U(k) - U(j)\big), \ k \in \{1, \dots, n\}.$$

Theorem 2.2. [2] *Consider an electrical network $\mathcal{E}(G, w)$. Then, there is a matrix $M_R(\mathcal{E}) = (x_{ij}) \in Mat_{n\times n}(\mathbb{R})$ such that the following holds:*

$$M_R(\mathcal{E})\,\boldsymbol{U} = \boldsymbol{I}.$$

*This matrix is called the **response matrix** of a network $\mathcal{E}(G, \omega)$.*

Taking the «inverse» of a matrix $M_R(\mathcal{E})$, we can obtain another important electrical network characteristic:

Definition 2.3. Let $\mathcal{E}(G, \omega)$ be a connected electrical network with n boundary nodes, and let the boundary voltages $U = (U_1, \dots, U_n)$ be such that

$$M_R(\mathcal{E})U = -e_i + e_j,$$

where $e_k, \ k \in \{1, \dots, n\}$ is the standard basis of $\mathbb{R}^n$.

Let us define the **effective resistance** R_{ij} between nodes i and j as follows $|U_i - U_j| := R_{ij}$. Effective resistances are well-defined and $R_{ij} = R_{ji}$.

We will organize the effective resistances R_{ij} in an **effective resistances matrix** $R(\mathcal{E})$ setting $R_{ii} = 0$ for all i.

A well-known result, with broad applications in applied mathematics and organic chemistry, states that:

Theorem 2.4. [3] *Let $\mathcal{E}(G, \omega)$ be a connected electrical network then for any of its three boundary nodes k_1, k_2 and k_3 the triangle inequality holds:*

$$R_{k_1k_3} + R_{k_2k_3} - R_{k_1k_2} \geq 0.$$

Hence the set of all R_{ij} defines a metric on the boundary nodes of G.

3 Main results

Kirchhoff's classical matrix tree theorem helps us to calculate R_{ij} using partition functions of spanning trees. Building upon this, the generalized Temperley trick [8] allows us to alternatively calculate R_{ij} with dimer (almost perfect matching) partition functions. The last combinatorial statement is the key to explicitly relating electrical network theory with the geometry of totally non-negative Grassmannians $Gr_{\geq 0}(n-1, 2n)$:

Theorem 3.1. [1], [13] *Let $\mathcal{E}(G, \omega)$ be a connected electrical network with n boundary nodes. Using its effective resistances matrix $R(\mathcal{E})$, let us define a matrix:*

$$\Omega_R(\mathcal{E}) = \begin{pmatrix} 1 & m_{11} & 1 & -m_{12} & 0 & m_{13} & 0 & \dots \\ 0 & -m_{21} & 1 & m_{22} & 1 & -m_{23} & 0 & \dots \\ 0 & m_{31} & 0 & -m_{32} & 1 & m_{33} & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{1}$$

where

$$m_{ij} = -\frac{1}{2}(R_{i,j} + R_{i+1,j+1} - R_{i,j+1} - R_{i+1,j}).$$

Then, $\Omega_R(\mathcal{E})$ defines a point $\mathcal{L}(\mathcal{E})$ in the totally non-negative part of $Gr(n-1, 2n)$, it means that:

- *The dimension of the row space of $\Omega_R(\mathcal{E})$ is equal to $n-1$;*
- *Each $n-1 \times n-1$ minors of $\Omega_R(\mathcal{E})$ is non-negative;*
- *Plucker coordinates of the point of $\mathcal{L}(\mathcal{E})$ correspond to $n-1 \times n-1$ minors of the matrix $\Omega'_R(\mathcal{E})$ obtained from $\Omega_R(\mathcal{E})$ by deleting the last row.*

Many interesting inequalities involving R_{ij} follow from the positivity of the Plucker coordinates of the point represented by $\Omega_R(\mathcal{E})$. Some of them have an elegant interpretation:

Theorem 3.2. [13] *Let $\mathcal{E}(G, \omega)$ be a connected circular electrical network and let i_1, i_2, i_3, i_4 be any four nodes in the circular order. Then the Kalmanson inequalities hold:*

$$R_{i_1 i_3} + R_{i_2 i_4} \geq max(R_{i_2 i_3} + R_{i_1 i_4}, R_{i_1 i_2} + R_{i_3 i_4}).$$

Since, R_{ij} can be considered as the Kalmanson metric on boundary nodes.

Additionally, Theorem 3.1 allows us to obtain the exhaustive characterization of planar electrical Kalmanson metrics:

Theorem 3.3. [13] *Let $D = (d_{ij}) \in Mat_{n \times n}(\mathbb{R})$ be a matrix of a Kalmanson metric with respect of a cyclic numeration $1, \dots, n$, then D is the effective resistance matrix of a connected circular electrical network $\mathcal{E}$ with n boundary nodes if and only if the matrix Ω_D constructed from D according to the formula* (1) *defines a point X in $\mathrm{Gr}_{\geq 0}(n-1, 2n)$ and the Plucker coordinate $\Delta_{24\dots 2n-2}(X)$ does not vanish.*

The last theorem can be elegantly reformulated in the following equivalent form:

Theorem 3.4. [13] *Let $D = (d_{ij}) \in Mat_{n \times n}(\mathbb{R})$ be a matrix of a Kalmanson metric with respect to a cyclic numeration $1, \dots, n$ and let us consider a matrix $M(D) = (m_{ij})$, $m_{ij} = \frac{1}{2}(d_{i,j} + d_{i+1,j+1} - d_{i,j+1} - d_{i+1,j})$ of the coefficients in its circular split decomposition, see [9]. Then D is the effective resistance matrix of a connected circular electrical network $\mathcal{E}$ with n boundary nodes if and only if the matrix $-M(D)$ is a response matrix $M_R(\mathcal{E}^*)$ of a network $\mathcal{E}^*$. Moreover, $\mathcal{E}^*$ is the dual network to $\mathcal{E}$.*

4 Applications

A notoriously difficult problem in electrical network theory is to recover a graph of an electrical network from its effective resistance matrix, see [2]. Beyond its fundamental applications in physics and engineering (see [7], [5], [12]), solutions to this problem can also be used for the reconstruction of phylogenetic networks (see [10], [4], [13]), owing to the fact that the effective resistance metric of circular networks satisfies the Kalmanson property. We will focus on presenting a sketch of a new algorithm for reconstructing the topology of a graph of a circular electrical network from its effective resistance matrix.

Problem 4.1. Consider an electrical network $\mathcal{E}(G, w)$ on an unknown graph G. It is required to reconstruct (up to electrical transformations, see [2]) a graph G by a known effective resistance matrix $R_{\mathcal{E}}$.

The ability to solve the last problem is revealed by the following results.

Definition 4.2. [6] A circular electrical network is called **minimal** if the strands of its median graph do not have self-intersections; any two strands intersect at most one point, and the median graph has no loops or lenses.

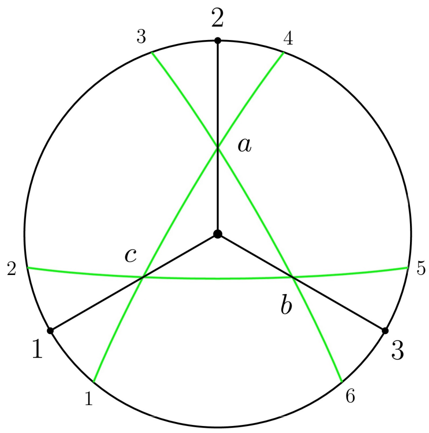


Figure 1: Star-shape network, its median graph and the strand permutation $\tau(\mathcal{E}) = (14)(36)(25)$

Theorem 4.3. [2] *Each minimal circular electrical network $\mathcal{E}(G, \omega)$ is unique up to electrical transformations defined by its strand permutations $\tau(\mathcal{E})$.*

Denote by A_i the columns of the matrix $\Omega_R(\mathcal{E})$ and define the column permutation $g(\mathcal{E})$ as follows: $g(\mathcal{E})(i) = j$, if j is the minimal number such that $A_i \in \operatorname{span}(A_{i+1}, \ldots, A_j)$, where the indices are taken modulo $2n$.

Theorem 4.4. [13] *Up to the star-triangle transformations, a topology of each minimal electrical network $\mathcal{E}(G, w)$ with n boundary nodes can be uniquely recovered by a column «rank-patterns» of $\Omega_R(\mathcal{E})$:*

$$g(\mathcal{E}) + 1 = \tau(\mathcal{E}) \mod 2n.$$

Remark 4.5. A more advanced technique called the generalized chamber ansatz applied to $\Omega_R(\mathcal{E})$ provides an algorithm for recovering the conductivity function ω as well, see [5].

References

[1] L. Guterman A. Kazakov B. Bychkov, V. Gorbounov. Symplectic geometry of electrical networks. *Journal of Geometry and Physics*, 207:105323, 2005.

[2] J. A. Morrow B. Curtis. Inverse problems for electrical networks. *Inverse Problems*, 13, 2000.

[3] M. Randić D.J. Klein. Resistance distance. *J Math Chem*, 12:81–95, 1993.

[4] S. Forcey. Circular planar electrical networks, split systems, and phylogenetic networks. *SIAM Journal on Applied Algebra and Geometry*, 7(1):49–76, 2023.

[5] A. A. Kazakov. Inverse problems related to electrical networks and the geometry of non-negative grassmannians. *arXiv preprint arXiv:2502.16710*, 2025.

[6] R. W. Kenyon. The laplacian on planar graphs and graphs on surfaces. *arXiv:1203.1256*, 2012.

[7] F.G. Vasquez L. Borcea, V. Druskin. On the structure of linear electrical impedance tomography with resistor networks. *Inverse Problems*, 24(3):035013, 2008.

[8] T. Lam. Electroid varieties and a compactification of the space of electrical networks. *Advances in mathematics*, 338:549–600, 2018.

[9] R. Weismantel M. Deza, M. Laurent. Geometry of cuts and metrics. *Berlin : Springer*, 2, 1997.

[10] D. Scalzo S. Forcey. Phylogenetic networks as circuits with resistance distance. *Frontiers in Genetics*, 11:586664, 2020.

[11] D. V. Talalaev. Vertex electrical model: Lagrangian and nonnegativity properties. *Theoretical and Mathematical Physics*, 210(2):216–223, 2022.

[12] G. Uhlmann. Electrical impedance tomography and calderón's problem. *Inverse problems*, 25(21):123011, 2009.

[13] A. Kazakov V. Gorbounov. Electrical networks and data analysis in phylogenetics. *Data Analytics and Topology*, 1(1):33–45, 2025.

[14] D. Talalaev V. Gorbounov. Electrical varieties as vertex integrable statistical models. *Journal of Physics A: Mathematical and Theoretical*, 53(45):454001, 2020.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Separating Path Systems for 2-Degenerate Graphs

(Extended abstract)

Cristina G. Fernandes* Carlos Hoppen† George Kontogeorgiou‡
Guilherme Oliveira Mota§ Danni Peng¶

Abstract

We prove that any connected 2-degenerate graph G with n vertices admits a family $\mathcal{P}$ of at most n paths that *strongly separates* $E(G)$, which means that for any pair of distinct edges e, f of G there is a path in $\mathcal{P}$ containing e and not f, and another path containing f and not e. Using this we show that n paths are also enough to strongly separate the edges of any n-vertex outerplanar graph and any connected subcubic graph except for K_4. For bipartite planar graphs with n vertices, we show that there is a separating family of at most $3n/2$ paths.

1 Introduction

Given a collection $\mathcal{P}$ of paths in a graph G, we say that two edges e, f of G are *separated* by $\mathcal{P}$ (and $\mathcal{P}$ *separates* e, f) if there are two paths P_e and P_f in $\mathcal{P}$ such that P_e contains e but not f, and P_f contains f but not e. We say $\mathcal{P}$ is a *strong separating path system* for G if $\mathcal{P}$ separates any pair of edges of G. If for any pair of edges e, f of G there is a path in $\mathcal{P}$ containing only one of e, f, then we say that $\mathcal{P}$ is a *weak separating path system* for G. The sizes of a minimum strong and a minimum weak separating path system for G are denoted, respectively, by $\mathrm{ssp}(G)$ and $\mathrm{wsp}(G)$.

*Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil. E-mail: `cris@ime.usp.br`. Partially supported by CNPq (310979/2020-0 and 404315/2023-2).

†Instituto de Matemática e Estatística, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. E-mail: `choppen@ufrgs.br`. Supported by FAPERGS (21/2551-0002053-9) and CNPq (315132/2021-3 and 408180/2023-4).

‡Centro de Modelamiento Matemático, Universidad de Chile, Santiago, Chile. E-mail: `gkontogeorgiou@dim.uchile.cl`. Supported by ANID grant CMM Basal FB210005.

§Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil. E-mail: `mota@ime.usp.br`. Supported by CNPq (306620/2020-0, 406248/2021-4) and FAPESP (2023/03167-5 and 2024/13859-4).

¶Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, Brazil. E-mail: `danni.peng@impa.br`.

Falgas-Ravry, Kittipassorn, Korándi, Letzter, and Narayanan [4] conjectured that there are weak separating path systems with $O(n)$ paths for every n-vertex graph, and obtained interesting results for particular cases. Balogh, Csaba, Martin, and Pluhár [1] strengthened the conjecture to also apply to strong separating systems. Both groups observed that the $O(n \log n)$ bound holds for any n-vertex graph. A substantial improvement was obtained by Letzter [9], who proved that $\mathrm{ssp}(G) = O(n \log^* n)$ for any n-vertex graphs G. This conjecture was confirmed by Bonamy, Botler, Dross, Naia, and Skokan [2], who showed that $\mathrm{ssp}(G) \leq 19n$ for any n-vertex graph G.

The bound of $19n$ is not tight and, in fact, it could be true that $\mathrm{ssp}(G) \leq (1 + o(1))n$ for every connected n-vertex graph G, where the connectivity condition is necessary as the graph composed by $n/4$ disjoint copies of K_4 requires $5n/4$ paths to be strongly separated. Recently, the first and fourth authors together with Sanhueza-Matamala [5, 6] proved that $\mathrm{ssp}(K_n) = (1 + o(1))n$ and obtained upper bounds for $\mathrm{ssp}(G)$ for certain n-vertex αn-regular graphs. For example, they showed that $\mathrm{ssp}(K_{n/2,n/2}) \leq (\sqrt{5/2} - 1 + o(1))n$, which is best-possible up to the $o(1)$ term. More recently, the third author together with Stein [7] obtained weak and strong separating path systems for K_n with at most $n+1$ and $n+9$ paths respectively. From now on, we consider only strong separating path systems.

We say a subcubic graph is *strictly subcubic* if it is not cubic. A graph G is 2-*degenerate* if every subgraph of G contains a vertex of degree at most 2. We contribute to this line of research by proving that $\mathrm{ssp}(G) \leq n$ for every 2-degenerate n-vertex graph G. If all components of G contain at least three vertices, we remark that our separating family has the additional property that every edge is contained in exactly two paths, and that every vertex is the endpoint of exactly two paths. The class of 2-degenerate graphs includes, for instance, all connected series-parallel graphs, all connected strictly subcubic graphs, and all connected planar graphs of girth at least 6. Using this result, we also prove that $\mathrm{ssp}(G) \leq n$ for every connected subcubic n-vertex graph G except for K_4, that $\mathrm{ssp}(G) \leq 3n/2$ for every bipartite planar graph G, and that $\mathrm{ssp}(G) \leq 2n$ for every Hamiltonian planar graph (every 4-connected planar graph is Hamiltonian).

2 2-Degenerate graphs

A graph is *planar* if it can be drawn in the plane in such a way that its edges intersect only at their ends. Such a drawing is called a *planar drawing*. An interesting class of planar graphs are the *outerplanar*, which are those that admit a planar drawing for which all of its vertices lie in the outer face. A *k-degenerate graph G* is a graph such that every subgraph of G contains a vertex of degree at most k. Equivalently, there is an ordering of the vertices of G such that any vertex has at most k neighbours that appear earlier in the ordering.

In this section, we prove that $\mathrm{ssp}(G) \leq n$ for every 2-degenerate n-vertex graph G. Then we will obtain upper bounds for outerplanar, subcubic, and bipartite planar graphs. Before going into our results, note that, since $\mathrm{ssp}(G) = \sum_{i=1}^{k} \mathrm{ssp}(G_i)$, where $G_1, \ldots, G_k$ are the connected components of G, if $\mathrm{ssp}(G_i) \leq |V(G_i)|$ for $i = 1, \ldots, k$, then $\mathrm{ssp}(G) \leq n$. Thus, we may restrict attention to connected graphs.

We refer to a path just by its sequence of vertices. That is, for a path with vertex set $\{v_1, \ldots, v_\ell\}$ where $v_i v_{i+1}$ is an edge for every $1 \leq i \leq \ell - 1$, we write $P = (v_1, \ldots, v_\ell)$.

Theorem 2.1. *For every connected 2-degenerate n-vertex graph G, we have* $\mathrm{ssp}(G) \leq n$.

In fact, we prove a stronger statement which will be useful for our other results and gives an extra structural property for our strong path system for 2-degenerate graphs.

Theorem 2.2. *For every connected 2-degenerate graph G on $n \geq 3$ vertices, there is a strong separating path system for G with n paths and the following two properties: (i) every edge lies in exactly two paths; (ii) there are exactly two paths ending at each vertex of G.*

Proof. The proof is by induction on n. If $n = 3$, then G is either a length-2 path or a triangle. If G is a length-2 path, say (v_1, v_2, v_3), the family of the three possible paths in G, by which we mean $\{(v_1, v_2, v_3), (v_1, v_2), (v_2, v_3)\}$, is a strong separating path system for G satisfying (i) and (ii). If G is a triangle with vertices $\{v_1, v_2, v_3\}$, then our strong separating path system satisfying (i) and (ii) is $\{(v_1, v_2, v_3), (v_2, v_3, v_1), (v_3, v_1, v_2)\}$. For the induction step, consider $n \geq 4$, and assume that the result holds for any connected 2-degenerate graph with fewer than n vertices.

First suppose that G has a vertex v of degree at most two for which $G - v$ is connected. If v has degree one, consider a separating path system $\mathcal{P}'$ of $G - v$ satisfying (i) and (ii), of size $|\mathcal{P}'| = n-1$, and let P_u be one of the paths in $\mathcal{P}'$ that ends at the neighbour u of v. Let P be the path that extends P_u to v. See Figure 1(a). The path family $\mathcal{P} = (\mathcal{P}' \setminus \{P_u\}) \cup \{P, (u, v)\}$ is a path system such that every edge lies in exactly two paths and there are exactly two paths ending at each vertex of G, that is, it satisfies (i) and (ii). It is separating because $e = (u, v)$ is a path in $\mathcal{P}$ and, by induction, any edge in $G - v$ lies in a path that is distinct from P and therefore avoids e. If v has degree two, let u and w be its neighbours, and let $\mathcal{P}'$ be a strong separating path system of $G - v$ satisfying (i) and (ii), of size $|\mathcal{P}'| = n-1$. Let P_u be one of the paths in $\mathcal{P}'$ that ends in u and let P_w be a path in $\mathcal{P}'$ that ends in w with the property that $P_w \neq P_u$. Let P_1 and P_2 be the paths that extend P_u and P_w to v using the edges uv and vw, respectively. See Figure 1(b). The path family $\mathcal{P} = (\mathcal{P}' \setminus \{P_u, P_w\}) \cup \{P_1, P_2, (u, v, w)\}$ is a path system with the property that every edge lies in exactly two paths and there are exactly two paths ending at each vertex of G. Note that P_1 and P_2 strongly separate uv from vw. The edges uv and vw are separated from the remaining edges by (u, v, w). Finally, by induction, any edge e' in G' lies in a path other than P_u (and so corresponds to a path in $\mathcal{P}$ that avoids uv) and lies in a path other than P_w (and so corresponds to a path in $\mathcal{P}$ that avoids vw).

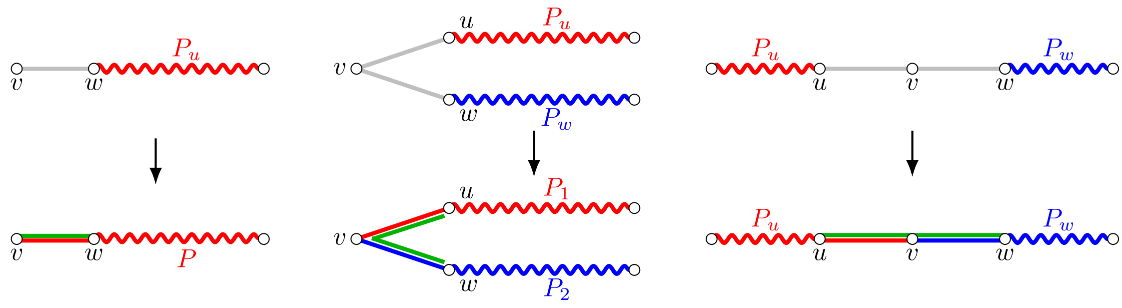


Figure 1: Illustration for the cases in the proof of Theorem 2.2.

Next suppose that the removal of any vertex v in G with degree at most two disconnects the graph. This means that the minimum degree of G is two and that the removal of a vertex v of degree two produces two components G_1 and G_2 of size n_1 and n_2, where $n_1 + n_2 = n - 1$

and $n_1, n_2 \geq 3$. By induction, let $\mathcal{P}_1$ and $\mathcal{P}_2$ be strong separating path systems for G_1 and G_2, respectively, with the required properties. Let u and w be the neighbours of v in G_1 and G_2, respectively, and consider paths P_u in $\mathcal{P}_1$ ending at u and P_w in $\mathcal{P}_2$ ending at w. Let P_1 and P_2 be the paths that extend P_u and P_w to v using the edges uv and vw, respectively. See Figure 1(c). Consider the path family $\mathcal{P} = (\mathcal{P}_1 \cup \mathcal{P}_2 \setminus \{P_u, P_w\}) \cup \{P_1, P_2, (u, v, w)\}$. It is easy to check that it is a strong separating path system for G satisfying properties (i) and (ii). $\square$

Consider a connected 2-degenerate graph G on $n \geq 3$ vertices, with non-neighbours u and v with $d_G(u) = d_G(v) = 2$. If u and v are the last two vertices considered in the induction above, from the last paragraph of the proof, we derive the following.

Fact 2.3. *Let G be a graph on $n \geq 3$ vertices whose components are strictly subcubic. If there are non-neighbours u and v in G with $d_G(u) = d_G(v) = 2$, then there is a strong separating path system for G that contains two length-2 paths, one with internal vertex u and the other with internal vertex v.*

Proof. Let $u, v \in V(G)$ be non-neighbours with $d_G(u) = d_G(v) = 2$. Denote the neighbours of u and v respectively by u_1, u_2 and v_1, v_2, and let G' be the graph obtained by removing u and v. Observe that G' is 2-degenerate because every subgraph of G' is also a subgraph of G. Apply Theorem 2.2 to each component of G' with at least three vertices. Using the obtained path systems, together with trivial path systems for possible components of G' with at most two vertices, one can obtain a strong separating path system $\mathcal{P}'$ for G' with $n-2$ paths such that there are two paths ending at u_1 and two paths ending at u_2. Then, we can take two distinct paths P_{u_1} and P_{u_2} ending respectively at u_1 and u_2. Similarly, let P_{v_1} and P_{v_2} be two distinct paths ending respectively at v_1 and v_2. Then, by extending P_{u_1} and P_{u_2} to u, and P_{v_1} and P_{v_2} to v, we just need to add the length-2 paths (u_1, u, u_2) and (v_1, v, v_2) to obtain the desired strong separating path system. $\square$

3 Corollaries of the result for 2-degenerate graphs

In this section we obtain some results that follow from the results in Section 2.

Theorem 3.1. *If G is a connected cubic n-vertex graph and it is not K_4, then $\operatorname{ssp}(G) \leq n$.*

Proof. As K_4 is the only connected cubic graph in which every edge lies in a triangle, we can take an edge $e = uv$ of G that belongs to no triangle. Then, each component of the graph $H = G - e$ is strictly subcubic, and u and v have disjoint neighbourhoods in H, which are denoted $\{u_1, u_2\}$ and $\{v_1, v_2\}$, respectively. From Fact 2.3, there is a strong separating path system $\mathcal{P}'$ for H with n paths that contains two paths, P_u and P_v, of length 2 with internal vertex u and v, respectively. Our aim is to show that these paths of length 2 may be re-routed as paths of length 3 using e in a way that preserves strong separation. See Figure 2. Recall that $\mathcal{P}'$ may be chosen so that every edge of H lies in exactly two paths and every vertex of H is the end of exactly two paths. Consider the two paths P_1 and P_2 starting at u, and assume without loss of generality that P_1 uses the edge uu_1 and P_2 uses the edge uu_2. Let Q_1 and Q_2 be the corresponding paths for v. If $P_1 = Q_1$, then $P_1 \neq Q_2$ and $P_2 \neq Q_1$. In this case, interchange Q_1 and Q_2 so that $P_1 \neq Q_1$ and $P_2 \neq Q_2$. Consider the path system $\mathcal{P} = \mathcal{P}' \setminus \{P_u, P_v\} \cup \{(u_1, u, v, v_1), (u_2, u, v, v_2)\}$. We claim that $\mathcal{P}$ is strong separating for G. It is easy to see that uv is strongly separated from any edge of H. Moreover, any two edges that were strongly separated in H are clearly strongly separated by the same paths (or by the paths that replaced them), except for the pairs $\{uu_1, vv_1\}$ and $\{uu_2, vv_2\}$, which were separated by P_u and P_v in $\mathcal{P}'$, but lie on the same new paths. However, our choice of re-routing guarantees that the pair $\{uu_1, vv_1\}$ is strongly separated by P_1 and Q_1, while the pair $\{uu_2, vv_2\}$ is strongly separated by P_2 and Q_2. □

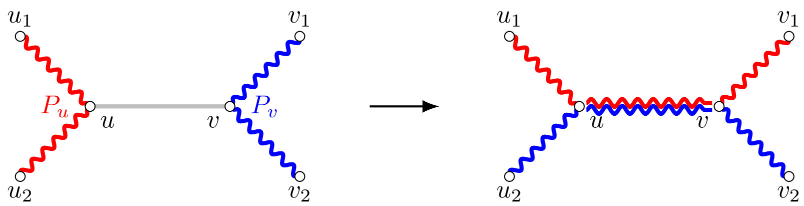


Figure 2: Illustration for the proof of Theorem 3.1.

We obtain the following direct corollary from Theorem 3.1.

Corollary 3.2. *If G is a subcubic n-vertex graph with k connected components isomorphic to K_4, then* $\mathrm{ssp}(G) \leq n + k$.

Proof. Since K_4 requires 5 paths to be strongly separated, we need $5k$ paths for the components isomorphic to K_4. Furthermore, in view of Theorem 3.1, we can strongly separate the remaining edges with $n - 4k$ paths, for a total of $5k + (n - 4k) = n + k$ paths. □

Outerplanar graphs always contain a vertex of degree at most 2, and their class is closed under (induced) subgraphs, so they are 2-degenerate. Therefore, the following is a direct corollary of Theorem 2.1.

Corollary 3.3. *If G is an outerplanar n-vertex graph, then* $\mathrm{ssp}(G) \leq n$.

The well-known $(4, 3)$-Conjecture by Chartrand, Geller, and Hedetniemi [3] states that every planar graph has an edge partition into two outerplanar graphs. This is known to hold for Hamiltonian planar graphs: one possible partition is the edges on the inside of a Hamilton cycle versus the edges on the outside (the edges on the cycle may be added to any of the parts). Tutte [10] proved that every 4-connected planar graph is Hamiltonian, so the $(4, 3)$-Conjecture holds for 4-connected planar graphs.

Corollary 3.4. *If G is a planar n-vertex graph that satisfies the* $(4, 3)$*-Conjecture, then* $\mathrm{ssp}(G) \leq 2n$.

Proof. Let $G = (V, E)$. Then its edge set may be split as $E = E_1 \cup E_2$, where $H_1 = (V, E_1)$ and $H_2 = (V, E_2)$ are outerplanar. By Corollary 3.3, there are strong separating path systems $\mathcal{P}_1$ and $\mathcal{P}_2$ of size n for H_1 and H_2, respectively. The path system $\mathcal{P}_1 \cup \mathcal{P}_2$ is clearly a strong separating path system for G. □

From [8], in every bipartite planar n-vertex graph, there are at most $n/2$ edges whose removal results in a 2-degenerate graph. This and Theorem 2.1 imply the following.

Corollary 3.5. *If G is a bipartite planar n-vertex graph, then* $\mathrm{ssp}(G) \leq 3n/2$.

Krisam [8] conjectured that, in every bipartite planar n-vertex graph G, the removal of $n/4$ edges would be enough to obtain a 2-degenerate graph, which would imply that $\mathrm{ssp}(G) \leq 5n/4$.

Acknowledgment. This research was partly supported by a joint project FAPESP and ANID (2019/13364-7), and by CAPES (Finance Code 001).

References

[1] József Balogh, Béla Csaba, Ryan R. Martin, and András Pluhár. On the path separation number of graphs. *Discrete Applied Mathematics*, 213:26–33, 2016.

[2] Marthe Bonamy, Fábio Botler, François Dross, Tássio Naia, and Jozef Skokan. Separating the edges of a graph by a linear number of paths. *Advances in Combinatorics*, 2023. 7pp.

[3] Gary Chartrand, Dennis Geller, and Stephen Hedetniemi. Graphs with forbidden subgraphs. *Journal of Combinatorial Theory, Series B*, 10(1):12–41, 1971.

[4] Victor Falgas-Ravry, Teeradej Kittipassorn, Dániel Korándi, Shoham Letzter, and Bhargav P. Narayanan. Separating path systems. *Journal of Combinatorics*, 5(3):335–354, 2014.

[5] Cristina G. Fernandes, Guilherme O. Mota, and Nicolás Sanhueza-Matamala. Separating path systems in complete graphs. In *Latin American Symposium on Theoretical Informatics*, volume 14578 of *LNCS*, pages 98–113. Springer, 2024.

[6] Cristina G. Fernandes, Guilherme O. Mota, and Nicolás Sanhueza-Matamala. Separating path systems in complete graphs. *Random Structures & Algorithms*, 66(3):e70006, 2025.

[7] George Kontogeorgiou and Maya Stein. An exact upper bound for the minimum size of a path system that weakly separates a clique, 2024, arXiv:2403.08210.

[8] Nadine D. Krisam. Maximal k-degenerate spanning subgraphs. Master's thesis, Karlsruhe Institute of Technology (KIT), 2021.

[9] Shoham Letzter. Separating paths systems of almost linear size, 2022, arXiv:2211.07732.

[10] William T. Tutte. A theorem on planar graphs. *Transactions of the American Mathematical Society*, 82:99–116, 1950.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Graph reconstruction from queries on triples

(Extended abstract)

Florian Galliot * Hoang La † Raphaëlle Maistre ‡ Matthieu Petiteau §
Dimitri Watel ¶

Abstract

Graph reconstruction has been studied for a long time. The famous Kelly-Ulam conjecture [3][6] states that any graph of order $n \geqslant 3$ is uniquely reconstructible (up to isomorphism) from the multiset of its subgraphs of order $n-1$. One way to approach this conjecture is to study reconstruction from other partial information. Let $k \leqslant n$ be an integer, and let $\mathbf{\Pi}$ be a partition of the set of graphs of order k. We are given a graph G with known vertices but unknown edges. We consider labelled graphs: the vertices are labelled, and we are looking for exact reconstruction as opposed to reconstruction up to isomorphism. For each induced subgraph of G of order k, we are given the element of $\mathbf{\Pi}$ to which it belongs, up to isomorphism. As such, $\mathbf{\Pi}$ defines a query on the graph G. For instance, in the case $k = 3$ and $\mathbf{\Pi} = \{\{\therefore, \dot{\underline{\ \ }}\}, \{\wedge, \triangle\}\}$, we are given a binary information on each subgraph H of order 3 of G: either H is isomorphic to $\therefore$ or $\dot{\underline{\ \ }}$, or it is isomorphic to $\wedge$ or $\triangle$. This example corresponds to graph reconstruction from the list of all connected triples of vertices in G. This problem has recently been studied by Bastide et al. [1], who provided a polynomial-delay algorithm to enumerate all graphs that are consistent with a given list of connected triples, and by Qi [5], who gave a structural characterization of graphs that can be uniquely reconstructed from their connected triples.

In this paper, we investigate all other possible partitions $\mathbf{\Pi}$ for $k = 3$. For each partition, we provide a structural characterization of graphs that are uniquely reconstructible from the information provided by the query, as well as a polynomial-delay algorithm to enumerate all graphs that are consistent with given query answers.

*Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France. E-mail : `florian.galliot@univ-amu.fr`

†LISN, Université Paris-Saclay, CNRS, Gif-sur-Yvette, France. E-mail : `hoang.la.research@gmail.com`

‡LISN, Université Paris-Saclay, CNRS, Gif-sur-Yvette, France. E-mail : `raphael.maistre-matus@universite-paris-saclay.fr`

§Univ. Grenoble Alpes, CNRS, Institut Fourier, 38000 Grenoble, France. E-mail : `matthieu.petiteau@univ-grenoble-alpes.fr`

¶ensIIE, 91000 Evry and SAMOVAR, 91000 Evry. E-mail : `dimitri.watel@ensiie.fr`

1 Introduction

The most famous conjecture on graph reconstruction has been stated independently by Kelly [3] in 1942 and Ulam [6] in 1960.

Conjecture 1 (Kelly-Ulam)**.** *Every graph of order $n \geqslant 3$ is uniquely reconstructible (up to isomorphism) from the multiset of its subgraphs of order $n-1$.*

We introduce a new framework to consider reconstruction from partial information. A graph on the vertex set $V = \{1, \ldots, n\}$ is called a labelled n-graph. Consider an integer $k \leqslant n$ and a partition $\mathbf{\Pi}$ of the set of graphs of order k. A query $Q_{\mathbf{\Pi}}$ is a function which, for each labelled n-graph G and each k-subset $S \subseteq V$, provides the set of graphs $Q_{\mathbf{\Pi}}(G)(S) \in \mathbf{\Pi}$ to which $G[S]$ belongs, up to isomorphism.

Problem: Labelled Graph Reconstruction
Given an integer n and a partition $\mathbf{\Pi}$ of the set of graphs of order $k \leqslant n$:

- **Uniqueness:** Structural characterization of the labelled n-graphs G that are uniquely reconstructible from $\mathbf{\Pi}$, meaning that, for every labelled n-graph $H \neq G$, there exists a k-subset $S \subseteq V$ such that $Q_{\mathbf{\Pi}}(H)(S) \neq Q_{\mathbf{\Pi}}(G)(S)$.
- **Enumeration:** Polynomial-delay enumeration of all graphs that are consistent with some input information, i.e., given some function $f : \binom{V}{k} \to \mathbf{\Pi}$, enumerate all labelled n-graphs G such that, for all k-subset $S \subseteq V$, $Q_{\mathbf{\Pi}}(G)(S) = f(S)$.

The partition $\mathbf{C}_3 = \{\{\therefore, \dot{-}\}, \{\wedge, \triangle\}\}$ corresponds to graph reconstruction from connected triples, for which the enumeration problem has been addressed by Bastide et al. [1] and the uniqueness problem has been solved by Qi [5]. In this paper, we solve these problems for all other possible queries on triples.

2 Main results

Our results are summed up in Table 1. Some of them extend to larger values of k (for partitions that admit a natural generalization to any k, such as $\mathbf{E}$ which generalizes to the query on the number of edges, or $\mathbf{K}$ which generalizes to the clique query).
The **complement** of a partition $\mathbf{\Pi}$ is the partition $\overline{\mathbf{\Pi}}$ defined as $\overline{\mathbf{\Pi}} = \{\{\bar{H} \mid H \in \pi\} \mid \pi \in \mathbf{\Pi}\}$. Whenever two partitions are complementary, we need only address one of the two. Indeed, G is uniquely reconstructible from $\mathbf{\Pi}$ if and only if $\overline{G}$ is uniquely reconstructible from $\overline{\mathbf{\Pi}}$.

3 Uniqueness theorems

We now state the uniqueness theorem associated to each partition. For reasons of space, we only give a hint of the main idea used for some of the results. For some partitions, we also make an observation regarding the enumeration result.

Theorem 2 ($\mathbf{E} = \{\{\triangle\}, \{\wedge\}, \{\dot{-}\}, \{\therefore\}\}$: query on the number of edges)**.**
Let n be an integer such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from $\mathbf{E}$ if and only if G is not isomorphic to: P_3, $\overline{P_3}$, P_4, C_4, or $\overline{C_4}$.

Notation	Query	Uniqueness	Enumeration	$k \geqslant 4$
-	$\{\{\triangle, \wedge, \dot{-}, \therefore\}\}$	-	-	-
E	$\{\{\triangle\}, \{\wedge\}, \{\dot{-}\}, \{\therefore\}\}$	✓[3]	✓[3]	✓[3]
K	$\{\{\triangle\}, \{\therefore, \dot{-}, \wedge\}\}$	✓[3]	✓[3]	✓[3]
$\overline{\mathbf{K}}$	$\{\{\therefore\}, \{\triangle, \wedge, \dot{-}\}\}$	✓[3]	✓[3]	✓[3]
L	$\{\{\wedge\}, \{\therefore, \dot{-}, \triangle\}\}$	✓[3]	✓[3]	
$\overline{\mathbf{L}}$	$\{\{\dot{-}\}, \{\triangle, \wedge, \therefore\}\}$	✓[3]	✓[3]	
C	$\{\{\triangle, \wedge\}, \{\dot{-}, \therefore\}\}$	✓[1]	✓[2]	
J	$\{\{\triangle, \dot{-}\}, \{\wedge, \therefore\}\}$	✓[3]	✓[3]	
A	$\{\{\triangle, \therefore\}, \{\wedge, \dot{-}\}\}$	✓[3]	✓[3]	
B	$\{\{\triangle\}, \{\therefore\}, \{\wedge, \dot{-}\}\}$	✓[3]	✓[3]	
F	$\{\{\wedge\}, \{\dot{-}\}, \{\triangle, \therefore\}\}$	✓[3]	✓[3]	
D	$\{\{\triangle\}, \{\wedge\}, \{\therefore, \dot{-}\}\}$	✓[3]	✓[3]	
$\overline{\mathbf{D}}$	$\{\{\therefore\}, \{\dot{-}\}, \{\triangle, \wedge\}\}$	✓[3]	✓[3]	
M	$\{\{\therefore\}, \{\wedge\}, \{\triangle, \dot{-}\}\}$	✓[3]	✓[3]	
$\overline{\mathbf{M}}$	$\{\{\triangle\}, \{\dot{-}\}, \{\therefore, \wedge\}\}$	✓[3]	✓[3]	

[1]Qi (2023) [2]Bastide et al. (2023) [3]GLMPW (2025+)

Table 1: All 15 possible queries on triples of vertices. A check in the column "$k \geqslant 4$" means the results extend to the generalization of the partition to larger values of k.

Sketch of proof. More generally, consider the query for general $k \leqslant n-1$ which returns the number of edges in the induced subgraph. For $k \leqslant n-2$, we can query all the tuples to compute the number of edges in every induced subgraph of G of order at least $n-2$, and then retrieve the edge set. For $k = n-1$, G is uniquely reconstructible if and only if it is uniquely characterized by its degree sequence. □

Theorem 3 ($\mathbf{F} = \{\{\wedge\}, \{\dot{-}\}, \{\triangle, \therefore\}\}$)**.**
Let n be an integer such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from **F** *if and only if G is neither a complete graph or a null graph and G is uniquely reconstructible from* **E**.

Theorem 4 ($\mathbf{J} = \{\{\triangle, \dot{-}\}, \{\wedge, \therefore\}\}$: query on the parity of the number of edges)**.**
Let n be an integer such that $n \geqslant 3$. No graph is uniquely reconstructible from **J**. *More precisely, for every labelled n-graph G, there exist exactly 2^{n-1} graphs consistent with $Q_{\mathbf{J}}(G)$ including G itself.*

Sketch of proof. The main argument is that $Q_{\mathbf{J}}(G) = Q_{\mathbf{J}}(H) \iff Q_{\mathbf{J}}(G \triangle H) = Q_{\mathbf{J}}(I_n)$ where $G \triangle H$ is the symmetric difference operator on edge sets and I_n is the null graph of order n. Note that $Q_{\mathbf{J}}(G) = Q_{\mathbf{J}}(I_n)$ if and only if G is a complete bipartite graph. □

Theorem 5 ($\mathbf{K} = \{\{\triangle\}, \{\therefore, \dot{-}, \wedge\}\}$: clique query)**.**
Let n be integers such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from **K** *if and only if:*

- *For every edge $uv \in E(G)$, there is a triangle in G containing uv.*

- *For every non-edge $uv \notin E(G)$, there is a triangle in $G + uv$ containing uv.*

Observation. The enumeration with polynomial delay is obtained using a branching tree where each node marks a choice between keeping the edge or not. Using the answers to the queries, we can cut a large number of branches and achieve polynomial delay.

Definition 6 (NEP graph)**.**
A graph $G = (V, E)$ is a ***non-edge partition*** *graph (NEP for short) if and only if there exists a proper partition $V = V_1 \uplus V_2$ such that for every pair $uv \notin E$ of vertices in V_1 (respectively V_2), $N(u) \cap V_2$ and $N(v) \cap V_2$ form a partition of V_2 (respectively $N(u) \cap V_1$ and $N(v) \cap V_1$ form a partition of V_1).*

Theorem 7 ($\mathbf{L} = \{\{\wedge\}, \{\therefore, \overset{\cdot}{\_}, \triangle\}\}$)**.**
A labelled n-graph G is uniquely reconstructible from $\mathbf{L}$ if and only if all connected components of G are not NEP *and at most one of them has order* 1*.*

Sketch of proof. We prove that any connected labelled n-graph G containing a uniquely reconstructible induced subgraph is uniquely reconstructible. From that, we can prove that any connected labelled n-graph containing an independent set of size 3 is uniquely reconstructible. Note that such graphs cannot be NEP. We then prove that any connected graph with no such independent set is uniquely reconstructible if and only if it is not NEP. ☐

Theorem 8 ($\mathbf{M} = \{\{\wedge\}, \{\therefore\}, \{\triangle, \overset{\cdot}{\_}\}\}$)**.**
Let n be an integer such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from $\mathbf{M}$ if and only if it is uniquely reconstructible from $\mathbf{L}$.

Theorem 9 ($\mathbf{A} = \{\{\triangle, \therefore\}, \{\wedge, \overset{\cdot}{\_}\}\}$)**.**
Let n be an integer such that $n \geqslant 3$. No graph is uniquely reconstructible from $\mathbf{A}$.

Proof. $\overline{G}$ is consistent with $Q_{\mathbf{A}}(G)$ for every labelled n-graph G. ☐

Definition 10. *A (p, q)-**perfect-dominating pair** in a graph G is a pair (S, T) of disjoint subsets of $V(G)$ such that $|S| = p$, $|T| = q$, and every $x \in V(G) \backslash (S \cup T)$ satisfies $N(x) \cap (S \cup T) \in \{S, T\}$.*

Theorem 11 ($\mathbf{B} = \{\{\triangle\}, \{\therefore\}, \{\wedge, \overset{\cdot}{\_}\}\}$)**.**
Let n be an integer such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from $\mathbf{B}$ if and only if it contains no (p, q)-perfect-dominating pair (S, T) with $p, q \in \{1, 2\}$ such that $G[S \cup T]$ is K_3-free and I_3-free.

Observation. The enumeration with polynomial delay is achieved by reducing the partition to a 2-CNF formula, where the variables correspond to pairs of vertices and $True$ means that the pair is an edge. We then refer to Feder [2] algorithm to enumerate all satisfying assignments of the 2-CNF formula.

Theorem 12 ($\mathbf{C} = \{\{\triangle, \wedge\}, \{\overset{\cdot}{\_}, \therefore\}\}$)**.** Refer to the result of Qi [5, Theorem 4.12].

Observation. This partition corresponds to the reconstruction problem from connected triples addressed by Bastide et al. [1] for the enumeration and Qi [5] for the uniqueness. The case $k \geqslant 4$ has also been explored in [1] and by Kluk et al. in [4].

Theorem 13 ($\mathbf{D} = \{\{\triangle\}, \{\wedge\}, \{\cdot^{\cdot}\cdot, \overset{\cdot}{\_}\}\}$).
Let n be an integer such that $n \geqslant 3$. A labelled n-graph G is uniquely reconstructible from $\mathbf{D}$ *if and only if G has at most one isolated vertex and has no connected component inducing a P_2, P_3, P_4, or C_4.*

Sketch of proof. We first prove that every connected graph of order 5 is uniquely reconstructible. As $\mathbf{D}$ is a refinement of $\mathbf{L}$, we also have the property that any connected graph G containing a uniquely reconstructible induced subgraph is uniquely reconstructible. $\square$

4 Perspective

From these results, several questions arise, suggesting multiple directions for future research. We employed various techniques to address both uniqueness and enumeration problems for each partition. It remains unknown whether a more unified approach could be developed. Additionally, we are interested in extending our results to larger values of k and exploring whether these results can be generalized to other types of queries.

Furthermore, the concept of partition raises several questions regarding graph reconstruction. Some partitions seem to provide more information than others, and it would be valuable to develop a method to quantify the amount of information conveyed by each partition. This could lead to a deeper understanding of the relationship between different partitions and their effectiveness in graph reconstruction. Specifically, some partitions rarely achieve uniqueness for any query, while for others, it is not even necessary to have information on every k-subset to achieve uniqueness. This also asks the question of the minimal number of k-subsets needed to achieve uniqueness for various partitions.

Given a partition $\mathbf{\Pi}$, we are also interested in understanding which properties of the input function $f : \binom{V}{k} \to \mathbf{\Pi}$ allow for the existence of graphs that are consistent with f. If no graph is consistent with f, can we determine which graphs are closest for some adequate notion of distance, e.g. which graphs maximize the number of satisfied k-subsets ? If f has been obtained by adding noise to some f_0 such that there exists a unique graph consistent with f_0, when is it possible to recover this graph from f depending on the nature of the noise ?

References

[1] Paul Bastide, Linda Cook, Jeff Erickson, Carla Groenland, Marc van Kreveld, Isja Mannens, and Jordi L. Vermeulen. Reconstructing graphs from connected triples. In Daniël Paulusma and Bernard Ries, editors, *Graph-Theoretic Concepts in Computer Science*, pages 16–29, Cham, 2023. Springer Nature Switzerland.

[2] Tomás Feder. Network flow and 2-satisfiability. *Algorithmica*, 11:291–319, 1994.

[3] Paul Joseph Kelly. *On isometric transformations.* Ph.D. thesis, University of Wisconsin, 1942.

[4] Kacper Kluk, Hoang La, and Marta Piecyk. Graph reconstruction with connectivity queries. In *Graph-Theoretic Concepts in Computer Science: 50th International Work-*

shop, WG 2024, Gozd Martuljek, Slovenia, June 19–21, 2024, Revised Selected Papers, pages 343–357. Springer-Verlag, 2024.

[5] Yaxin Qi. Graph reconstruction from connected triples. 2023.

[6] Stanislaw Ulam. *A Collection of Mathematical Problems*, volume 8 of *Interscience Tracts in Pure and Applied Mathematics*. Interscience Publishers, New York, NY, 1960.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A unified convention for achievement positional games

(Extended abstract)

Florian Galliot* Jonas Sénizergues†

Abstract

We introduce achievement positional games, a convention for positional games which encompasses the Maker-Maker and Maker-Breaker conventions. We consider two hypergraphs, one red and one blue, on the same vertex set. Two players, Left and Right, take turns picking a previously unpicked vertex. Whoever first fills an edge of their color, blue for Left or red for Right, wins the game (draws are possible). We establish general properties of such games. In particular, we show that a lot of principles which hold for Maker-Maker games generalize to achievement positional games. We also study the algorithmic complexity of deciding whether Left has a winning strategy as first player when all blue edges have size at mot p and all red edges have size at most q. This problem is in P for $p, q \leq 2$, but it is NP-hard for $p \geq 3$ and $q = 2$, coNP-complete for $p = 2$ and $q \geq 3$, and PSPACE-complete for $p, q \geq 3$. A consequence of this last result is that, in the Maker-Maker convention, deciding whether the first player has a winning strategy on a hypergraph of rank 4 after one round of (non-optimal) play is PSPACE-complete. A full version of this paper is available at [6].

1 Introduction

Positional games. *Positional games* have been introduced by Hales and Jewett [7] and later popularized by Erdős and Selfridge [3]. The game board is a hypergraph $H = (V, E)$, where V is the vertex set and $E \subseteq 2^V$ is the edge set. Two players take turns picking a previously unpicked vertex of the hypergraph, and the result of the game is defined by some *convention*. The two most popular conventions are called *Maker-Maker* and *Maker-Breaker*. As they revolve around trying to fill an edge *i.e.* pick all the vertices of some edge, they are often referred to as "achievement games". In the Maker-Maker convention, whoever first fills an edge wins (draws are possible), whereas in the *Maker-Breaker* convention, Maker aims at filling an edge while Breaker aims at preventing him from doing so (no draw is possible). For

*Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France. E-mail : florian.galliot@univ-amu.fr
†Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400 Talence, France. E-mail : jonas.senizergues@u-bordeaux.fr

all conventions, the main question is the result of the game with optimal play (who wins, or is it a draw), including the complexity of the associated algorithmic problems.

The Maker-Maker convention was the first one to be introduced, in 1963 by Hales and Jewett [7]. The game of *Tic-Tac-Toe* is a famous example. As a simple *strategy-stealing* argument [7] shows that the second player cannot have a winning strategy, the question is whether the given hypergraph H is a first player win or a draw with optimal play. This decision problem is trivially tractable for hypergraphs of rank 2 *i.e.* whose edges have size at most 2, but it is PSPACE-complete for hypergraphs that are 6-uniform *i.e.* whose edges have size exactly 6 (by a combination of [8] and [1]). Maker-Maker games are notoriously difficult to handle since both players must manage offense and defense at the same time.

The Maker-Breaker convention was introduced for that reason, in 1978 by Chvátal and Erdős [2]. It is by far the most studied, as it presents some convenient additional properties compared to the Maker-Maker convention thanks to the players having complementary goals, the most crucial one being subhypergraph monotonicity. The problem of deciding which player has a winning strategy when, say, Maker starts, is tractable for hypergraphs of rank 3 [5] but PSPACE-complete for 5-uniform hypergraphs [8] (a very recent improvement on the previously known result for 6-uniform hypergraphs [10]).

Unified achievement games. We introduce *achievement positional games.* Such a game is a triple $\mathcal{G} = (V, E_L, E_R)$, where (V, E_L) and (V, E_R) are hypergraphs which we see as having *blue edges* and *red edges* respectively. There are two players, taking turns picking a previously unpicked vertex: Left aims at filling a blue edge, while Right aims at filling a red edge. Whoever reaches their goal first wins the game, or we get a draw if this never happens. Achievement positional games include all Maker-Maker and Maker-Breaker games. Indeed, Maker-Maker games correspond to the case $E_L = E_R$, while Maker-Breaker games correspond to the case $E_R = \varnothing$ (or $E_L = \varnothing$) or equivalently to the case where each of E_L and E_R is the set of all minimal transversals of the other (a *transversal* of a set of edges E is a set of vertices that intersects each element of E). That last interpretation of Maker-Breaker games puts their asymmetrical nature into question, which is another motivation for the introduction of a unifying convention.

Objectives and results. We first establish elementary properties of achievement positional games in general. In particular, we look at some general principles which hold in the Maker-Maker convention to see if they generalize to achievement positional games. For all those that we consider, we show that this is indeed the case, emphasizing the fact that most properties of Maker-Maker games come from their "achievement" nature rather than symmetry. Our second objective is the study of the algorithmic complexity of the game. We get results for almost all edge sizes, which are summed up in Table 1. As a corollary, we also show that deciding whether the next player has a winning strategy for the Maker-Maker game on a hypergraph of rank 4 after one round of (non-optimal) play is PSPACE-complete.

2 Preliminaries

In this paper, a *hypergraph* is a pair (V, E) where V is a finite *vertex set* and $E \subseteq 2^V \setminus \{\varnothing\}$ is the *edge set.* An *achievement positional game* is a triple $\mathcal{G} = (V, E_L, E_R)$ where (V, E_L) and (V, E_R) are hypergraphs. The elements of E_L and E_R are called *blue edges* and *red edges* respectively. Two players, Left and Right, take turns picking a vertex in V that has not been picked before. We say a player *fills* an edge if that player has picked all the vertices of that

$q \backslash p$	0 , 1	2	3	4	5+
0 , 1	LSPACE [trivial]	LSPACE [9]	P [5]	open	PSPACE-c [8]
2	LSPACE [trivial]	P [Th. 4.1]	NP-hard [Th. 4.3]	NP-hard [Th. 4.3]	PSPACE-c [8]
3+	LSPACE [trivial]	coNP-c [Th. 4.2]	PSPACE-c [Th. 4.4]	PSPACE-c [Th. 4.4]	PSPACE-c [8]

Table 1: Algorithmic complexity of deciding whether Left has a winning strategy as first player, for blue edges of size at most p and red edges of size at most q.

edge. The blue and red edges can be seen as the winning sets of Left and Right respectively, so that the result of the game is determined as follows:

- If Left fills a blue edge before Right fills a red edge, then Left wins.
- If Right fills a red edge before Left fills a blue edge, then Right wins.
- If none of the above happens before all vertices are picked, then the game is a draw.

For algorithmic considerations, we introduce the problem AchievementPos(p,q) which consists in deciding, given an achievement positional game $\mathcal{G}$ such that all blue edges have size at most p and all red edges have size at most q, whether Left has a winning strategy on $\mathcal{G}$ as first player.

Like all positional games, AchievementPos(p,q) is in PSPACE as the game cannot last more than $|V|$ moves. We can also notice that, for all k, AchievementPos(0,k) and AchievementPos(1,k) are trivial problems. Meanwhile, for all k, AchievementPos(k,0) and AchievementPos(k,1) are equivalent to the Maker-Breaker game played on hypergraphs of rank k, so the literature provides some results.

Proposition 2.1. AchievementPos(k,0) *and* AchievementPos(k,1) *are in* LSPACE *for* $k \geq 2$*, in* P *for* $k = 3$*, but are* PSPACE*-complete for* $k \geq 5$*. Moreover,* AchievementPos(0,k) *and* AchievementPos(1,k) *are in* LSPACE *for all* k*.*

3 General results

A lot of convenient properties of Maker-Maker games generalize to achievement positional games. For instance, this is the case for the well-known *strategy-stealing* argument [7] which ensures that the second player can never have a winning strategy in the Maker-Maker convention.

Lemma 3.1 (Strategy Stealing)**.** *Let* $\mathcal{G} = (V, E_L, E_R)$ *be an achievement positional game. If there exists a bijection* $\sigma : V \to V$ *such that* $\sigma(e) \in E_L$ *and* $\sigma^{-1}(e) \in E_L$ *for all* $e \in E_R$*, then Left has a non-losing strategy on* $\mathcal{G}$ *as first player.*

Pairing strategies are an important tool in both Maker-Breaker and Maker-Maker conventions [7]. A *complete pairing* of a hypergraph H is a set Π of pairwise disjoint pairs of vertices such that every edge of H contains some element of Π. If H admits a complete pairing, then the outcome is a Breaker win for the Maker-Breaker game or a draw for the Maker-Maker game, as picking one vertex from each pair prevents the other player from filling an edge. We observe that, in general achievement positional games, pairing strategies may still be used as non-losing strategies which block the opponent.

Let us also mention the following monotonicity property: adding or shrinking blue edges cannot harm Left, and adding or shrinking red edges cannot harm Right.

4 Complexity results

Theorem 4.1. AchievementPos(2,2) *is in* P.

Sketch of the proof. After a series of forced moves, we get a situation where all edges have size exactly 2. At this point, the player who is next to play can be assumed not to have a P_3 (path on 3 vertices) of their color, as they would have a winning strategy in two moves. In particular, assume that Right is next to play, otherwise Right has a non-losing pairing strategy. Right must force all of Left's moves until she has broken every blue P_3. Any move u by Right triggers a sequence of forced moves, corresponding to an alternating red-blue path $P(u)$ which is easy to compute. If $P(u)$ ends with a red edge for some u, then we can assume that Right picks u and all forced moves along $P(u)$ are played, as Right keeps the initiative. However, if $P(u)$ ends with a blue edge for all u, then Right avoids a loss if and only if she can trigger one last sequence of forced moves after which every blue P_3 is broken. □

Theorem 4.2. AchievementPos(2,3) *is* coNP*-complete.*

Proof. We consider the complement of this problem, or rather an equivalent version of it. We show that it is NP-complete to decide whether Left has a non-losing strategy as first player on an achievement positional game where all blue (resp. red) edges have size at most 3 (resp. at most 2).

Let us first show membership in NP. Consider the following strategy $\mathcal{S}$ for Right: pick some u that wins in one move if such u exists, otherwise pick some v that prevents Left from winning in one move if such v exists, otherwise pick some w at the center of an intact red P_3 is one exists, otherwise (Left has a non-losing pairing strategy) pick an arbitrary vertex. Clearly, if Right has a winning strategy on $\mathcal{G}$ as second player, then $\mathcal{S}$ is one. Moreover, the move prescribed by $\mathcal{S}$ in any given situation is easily computed in polynomial time. Therefore, a polynomial certificate for Left's non-losing strategy is simply the sequence of all of Left's moves, assuming that Right plays according to $\mathcal{S}$.

We now reduce 3-SAT to our problem. Consider an instance ϕ of 3-SAT, with a set of variables V and a set of clauses C. We build a game $\mathcal{G} = (V, E_L, E_R)$ as follows (see Figure 1 for a visual example):

- For all $x \in V$, we define $V_x = \{x, \neg x\}$.
- For all $c = \ell_1 \vee \ell_2 \vee \ell_3 \in C$, with literals ℓ_1, ℓ_2, ℓ_3, we define $V_c = \left\{c_{\ell_1}, c_{\ell_2}, c_{\ell_3}, c'_{\ell_1}, c'_{\ell_2}, c'_{\ell_3}\right\}$.
- $V = \bigcup_{x\in V} V_x \cup \bigcup_{c\in C} V_c \cup \{\omega, \check{\omega}, \hat{\omega}\}$.
- $E_L = \bigcup_{x\in V} \{\{x, \neg x\}\} \cup \bigcup_{c=\ell_1\vee\ell_2\vee\ell_3\in C} \left\{\left\{\ell_1, c_{\ell_1}, c'_{\ell_1}\right\}, \left\{\ell_2, c_{\ell_2}, c'_{\ell_2}\right\}, \left\{\ell_3, c_{\ell_3}, c'_{\ell_3}\right\}\right\}$.
- $E_R = \bigcup_{c=\ell_1\vee\ell_2\vee\ell_3\in C} \left\{\left\{c_{\ell_1}, c_{\ell_2}\right\}, \left\{c_{\ell_2}, c_{\ell_3}\right\}, \left\{c_{\ell_3}, c_{\ell_1}\right\}\right\} \cup \left\{\left\{\omega, \check{\omega}\right\}, \left\{\omega, \hat{\omega}\right\}\right\}$.

Since there are multiple pairwise vertex-disjoint red P_3's, every move from Left must threaten to win on the next move until the last red P_3 is broken.

As such, Left must start by picking $\ell \in \{x, \neg x\}$ for some $x \in V$, which forces Right to pick the other one since $\{x, \neg x\} \in E_R$. After that, it can easily be shown that it is optimal for Left to pick c_ℓ for each clause c which contains the literal ℓ, as it forces Right to pick c'_ℓ in response since $\{\ell, c_\ell, c'_\ell\} \in E_R$. This breaks every red P_3 in the clause gadgets corresponding to clauses containing ℓ.

Left must repeat this process of picking a literal and then breaking all clause gadgets of clauses containing that literal, until he has picked at least one of c_{ℓ_1}, c_{ℓ_2} or c_{ℓ_3} for each clause $c = \ell_1 \vee \ell_2 \vee \ell_3 \in C$.

If there exists a valuation μ which satisfies ϕ, then Left succeeds in doing so, by picking x if $\mu(x) = \mathsf{T}$ or $\neg x$ if $\mu(x) = \mathsf{F}$, for all $x \in V$. After that, he can simply pick ω, thus ensuring not to lose the game. If such a valuation does not exist, then Left will have to play a move that does not force Right's answer while leaving at least one red P_3 intact, thus losing the game. All in all, Left has a non-losing strategy on $\mathcal{G}$ as first player if and only if ϕ is satisfiable, which concludes the proof. □

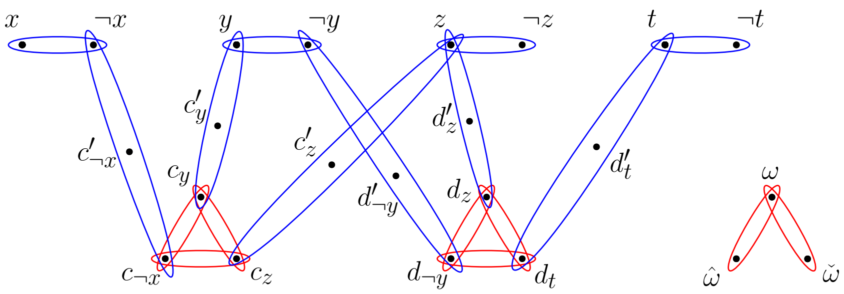


Figure 1: The full gadget from the proof of Theorem 4.2 for a set of two clauses $c = \neg x \vee y \vee z$ and $d = \neg y \vee z \vee t$.

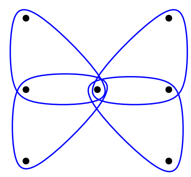


Figure 2: A blue butterfly.

Theorem 4.3. AchievementPos(3,2) *is* NP-*hard.*

Proof. We use the same construction as in the previous proof (see Figure 1), except we add two copies of the blue butterfly gadget (see Figure 2) to transform draws into wins for Left. Indeed, the game unfolds in the same way until/if Left breaks the last red P_3, after which Left will be able to play the first move in one of the butterflies to win the game. □

Theorem 4.4. AchievementPos(3,3) *is* PSPACE-*complete.*

Sketch of the proof. We perform a reduction from 3-QBF[11]. The input is a logic formula ϕ in CNF form, with clauses of size exactly 3 and variables $x_1, \ldots, x_{2n}$. Two players, Satisfier and Falsifier, take turns setting the variables $x_1, \ldots, x_{2n}$ (in that order) to T or F. Satisfier starts the game, and wins if ϕ is satisfied, otherwise Falsifier wins. Given ϕ, we build an instance $\mathcal{G}$ of AchievementPos(3,3) such that Left has a winning strategy on $\mathcal{G}$ as second player if and only if Falsifier has a winning strategy on ϕ.

The variable gadget associated to x_1 is pictured in Figure 3. Its edges of size 2 are the only ones in the entire game, so Right is forced to pick either $t_{1,R}$ (interpreted as $\mu(x_1) = \mathsf{T}$) or $f_{1,R}$ ($\mu(x_1) = \mathsf{F}$). This triggers a forced sequence of moves on the four vertices at the top. These moves update the variable gadget associated to x_2, which becomes as pictured in Figure 4. Left is forced to pick either $t_{2,L}$ ($\mu(x_2) = \mathsf{T}$) or $f_{2,L}$ ($\mu(x_2) = \mathsf{F}$), etc. until all variable gadgets have been played in. This marks the end of Phase 1, with half the vertices in $U = \bigcup_{1 \leq i \leq 2n} \{t_{i,R}, f_{i,R}\}$ having been picked by Right (by genuine choice for odd i, by forced choice for even i). The (blue) clause-edges are defined using U and parities, *e.g.* a clause $c = x_1 \vee x_2 \vee \neg x_3$ yields a blue edge $\{t_{1,R}, f_{2,R}, f_{3,R}\}$. During Phase 2, we use blue butterflies to allow Left to pick the remaining half of U, thus filling a clause-edge if and only if μ does not satisfy ϕ. □

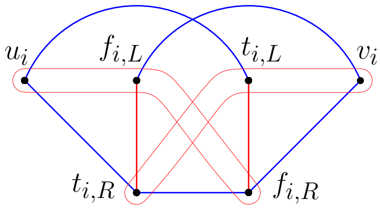


Figure 3: The variable gadget for odd i.

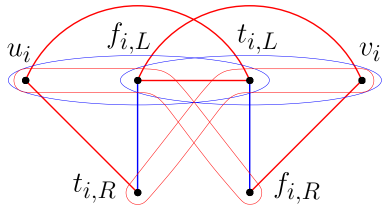


Figure 4: The variable gadget for even i.

Corollary 4.5. *Deciding whether the first player has a winning strategy for the Maker-Maker game on a hypergraph of rank 4 with one round having already been played is* PSPACE*-complete.*

Proof. Let $\mathcal{G}$ be an instance of AchievementPos(3,3). We add a vertex u_L to each $e \in E_L$ and a vertex u_R to each $e \in E_R$, then we forget about the colors. We get a hypergraph H such that, after one round of the Maker-Maker game on H where the players pick u_L and u_R respectively, we get precisely $\mathcal{G}$. □

5 Conclusion

We have introduced achievement positional games, a new convention for positional games where the players try to fill different edges. We have established some of their general properties (see [6] for all the results), which are not any weaker compared to the subfamily of Maker-Maker games, and obtained complexity results for almost all edge sizes. A corollary is that the Maker-Maker convention is PSPACE-complete for positions that can be obtained from a hypergraph of rank 4 after just one round of play, which is the first known complexity result on this convention for edges of size between 3 and 5.

We have not determined the exact complexity of the cases $(p, q) \in \{(3, 2), (4, 2)\}$, even though we know they are NP-hard. The commonplace intuition within the community is that Maker-Breaker games on hypergraphs of rank 4 are PSPACE-complete, which would imply that AchievementPos(4,2) also is. As for AchievementPos(3,2), a proof of either membership in NP or PSPACE-hardness would be compelling.

A natural prospect would be to define avoidance positional games, where whoever first fills an edge of their color loses. Since the case $E_L = E_R$ (*Avoider-Avoider convention*) is already PSPACE-complete for edges of size 2 [4], the complexity aspects would not be as interesting. However, an analogous study to that of Section 3 could be performed to better understand general properties of avoidance games.

Acknowledgments

We thank Sylvain Gravier for inspiring the topic of this paper and providing useful feedback. This research was partly supported by the ANR project P-GASE (ANR-21-CE48-0001-01).

References

[1] Jesper Makholm Byskov. Maker-Maker and Maker-Breaker games are PSPACE-complete. *BRICS Report Series*, 11(14), 2004.

[2] V. Chvátal and P. Erdős. Biased positional games. *Ann. Disc. Math.*, 2:221–229, 1978.

[3] Paul Erdös and J.L Selfridge. On a combinatorial game. *Journal of Combinatorial Theory, Series A*, 14, 1973.

[4] Stephen A. Fenner, Daniel Grier, Jochen Messner, Luke Schaeffer, and Thomas Thierauf. Game values and computational complexity: An analysis via black-white combinatorial games. In Khaled Elbassioni and Kazuhisa Makino, editors, *Algorithms and Computation*, pages 689–699, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[5] Florian Galliot, Sylvain Gravier, and Isabelle Sivignon. Maker-Breaker is solved in polynomial time on hypergraphs of rank 3. 2022.

[6] Florian Galliot and Jonas Sénizergues. A unified convention for achievement positional games. 2025.

[7] A. W. Hales and R. I. Jewett. Regularity and positional games. *Trans. Amer. Math. Soc.*, 106:222–229, 1963.

[8] Finn Orson Koepke. Solving Maker-Breaker games on 5-uniform hypergraphs is PSPACE-complete. 2025.

[9] Md Lutfar Rahman and Thomas Watson. Complexity of unordered CNF games. *ACM Trans. Comput. Theory*, 12(3), June 2020.

[10] Md Lutfar Rahman and Thomas Watson. 6-uniform Maker-Breaker game is PSPACE-complete. In *38th International Symposium on Theoretical Aspects of Computer Science (STACS 2021)*, volume 187 of *LIPIcs*, pages 57:1–57:15, Dagstuhl, Germany, 2021. Schloss Dagstuhl, Leibniz-Zentrum für Informatik.

[11] Larry J. Stockmeyer and Albert R. Meyer. Word problems requiring exponential time (preliminary report). In *Proceedings of the Fifth Annual ACM Symposium on Theory of Computing*, STOC '73, pages 1—9, New York, NY, USA, 1973. Association for Computing Machinery.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Minimum blocking sets for families of partitions

(Extended abstract)

Guillermo Gamboa Quintero* Ida Kantor†

Abstract

A 3-partition of an n-element set V is a triple of pairwise disjoint nonempty subsets X, Y, Z such that $V = X \cup Y \cup Z$. We determine the minimum size $\varphi_3(n)$ of a set $\mathcal{E}$ of triples such that for every 3-partition X, Y, Z of the set $\{1, \ldots, n\}$, there is some $\{x, y, z\} \in \mathcal{E}$ with $x \in X$, $y \in Y$, and $z \in Z$. In particular,

$$\varphi_3(n) = \left\lceil \frac{n(n-2)}{3} \right\rceil.$$

For $d > 3$, one may define an analogous number $\varphi_d(n)$. We determine the order of magnitude of $\varphi_d(n)$, and prove the following upper and lower bounds, for $d > 3$:

$$\frac{2n^{d-1}}{d!} - o(n^{d-1}) \leq \varphi_d(n) \leq \frac{0.86}{(d-1)!}n^{d-1} + o(n^{d-1}).$$

1 Blocking all partitions of a set

In this paper, we are interested in the following extremal problem: given n, find a smallest possible family $\mathcal{E}$ of triples of elements of $\{1, \ldots, n\}$ such that whenever we partition $\{1, \ldots, n\}$ into three nonempty sets X, Y, Z, there is some triple $\{x, y, z\} \in \mathcal{E}$ for which $x \in X$, $y \in Y$ and $z \in Z$. We call such family a 3-*blocking set*.

*Computer Science Institute of Charles University, Prague, Czechia. E-mail: gamboa@iuuk.mff.cuni.cz. Supported by Charles University project PRIMUS/24/SCI/012.

†Computer Science Institute of Charles University, Prague, Czechia. E-mail: ida@iuuk.mff.cuni.cz. Supported by GAČR grant 22-19073S.

Given a hypergraph $\mathcal{H}$, a *transversal* (or a *vertex cover*, or a *blocking set*) is a set of vertices that intersects all hyperedges of $\mathcal{H}$. Naturally, one is looking for a tranversal of minimum possible size. Our problem is an instance of this: consider the hypergraph where the vertices are all triples of numbers in $\{1, \ldots, n\}$, and each hyperedge corresponds to the partition X, Y, Z of $\{1, \ldots, n\}$ into three nonempty parts: the corresponding hyperedge contains all triples that have one vertex in each of the three parts. For hypergraphs that are well-behaved, the asymptotic solution is known (see [3, 7], and many later papers). However, these results do not apply to our hypergraph, as it does not satisfy the conditions (it is far from regular, and the maximum codegree is close to the maximum degree).

Our problem also has a strong connection to Turán theory of hypergraphs. If $\mathcal{F}$ is a family of hypergraphs, then the Turán number $\mathrm{ex}(n, \mathcal{F})$ is the maximum number of edges in an n-vertex hypergraph that does not contain any hypergraph in $\mathcal{F}$ as a subgraph. Turán numbers of hypergraphs are notoriously hard to handle (see for example the nice surveys [4, 5]) but in a lot of particular cases, a progress has been made. In our case, if B is a 3-blocking set, its complement (i.e., the family of triples that are not in B) is a hypergraph that has no spanning complete 3-partite subhypergraph. That is, we are looking for $\mathrm{ex}(n, \mathcal{F}_n)$, where $\mathcal{F}_n$ is the family of all spanning 3-partite 3-uniform hypergraphs on n vertices. We have $\binom{n}{3} - \mathrm{ex}(n, \mathcal{F}_n) = \Theta(n^2)$, but we are interested in determining this number exactly. Usually in Turán theory problems, the family $\mathcal{F}$ is fixed and the same for all n, while in our case, the forbidden hypergraphs are spanning. Examples where the configuration in question is spanning include, e.g., theorems that specify conditions (in graphs or hypergraphs) that guarantee a Hamiltonian cycle, or a packing by copies of a particular graph (see, e.g., [6, 9]).

We may define a d-blocking set analogously, as follows.

Definition 1. *If V is a finite set and $d \in \mathbb{N}$, a* d-partition *of V is a d-element family of sets $\{X_1, \ldots, X_d\}$ that are all nonempty and pairwise disjoint, such that $V = X_1 \cup \cdots \cup X_n$.*

Definition 2. *Let d be a positive integer. If $\mathcal{E} \subseteq \binom{V}{d}$, we say that $\mathcal{E}$* blocks *the d-partition $\{X_1, \ldots, X_d\}$ if there is a d-tuple $\{x_1, \ldots, x_d\} \in \mathcal{E}$ such that $x_i \in X_i$ for all i. We call $\mathcal{E}$ a* d-blocking set *if it blocks all d-partitions of V. Denote by $\varphi_d(n)$ the size of the smallest d-blocking set for an n-element set V.*

We determine the minimum size of a 3-blocking set exactly, for all n.

Theorem 1. *For all n, we have $\varphi_3(n) = \left\lceil \frac{n(n-2)}{3} \right\rceil$.*

For $d > 3$, we determine the order of magnitude for the minimum size of a d-blocking set, and present nontrivial bounds on the multiplicative coefficient.

Theorem 2. *For $d > 3$, we have $\frac{2n^{d-1}}{d!} - o(n^{d-1}) \le \varphi_d(n) \le \frac{0.86}{(d-1)!} n^{d-1} + o(n^{d-1})$.*

2 Observations for general d

For $d = 1$, there is only one 1-partition of V, with a single class equal to V. Trivially, $\varphi_1(n) = 1$.

For $d = 2$, we want a set $\mathcal{E}$ of pairs of elements of V such that for every partition of V into disjoint sets X and Y, there is some $\{x, y\} \in \mathcal{E}$ with $x \in X$ and $y \in Y$. That is, $\varphi_2(n)$ is the smallest number edges in a connected graph on n vertices, i.e., $\varphi_2(n) = n - 1$.

If we select a point x and take all d-tuples that contain x, this is a d-blocking set of size $\binom{n-1}{d-1}$, giving the trivial upper bound (for a fixed d and for n tending to ∞)

$$\varphi_d(n) \leq \frac{n^{d-1}}{(d-1)!} - o(n^{d-1}).$$

Several nice recurrences and inequalities can be obtained easily, but the following one is the only one that we ended up using (in establishing a stronger upper bound).

Lemma 1. *For every d and $n > d$, we have $\varphi_d(n) \leq \varphi_{d-1}(n-1) + \varphi_d(n-1)$.*

Proof. Select a point x from the underlying set. Let $\mathcal{A}$ be a $(d-1)$-blocking set on the underlying set $[n] \setminus \{x\}$. If we add x to each set in $\mathcal{A}$, the resulting family $\mathcal{A}'$ blocks all d-partitions where one of the parts is equal to $\{x\}$. Let $\mathcal{B}$ be a d-blocking set on the underlying set $[n] \setminus \{x\}$. This set also blocks all partitions of $[n]$ where $\{x\}$ is not a part of its own. Together, $\mathcal{A}' \cup \mathcal{B}$ is a d-blocking set of $[n]$ of size at most $\varphi_{d-1}(n-1) + \varphi_d(n-1)$. $\square$

3 Basic tool for $d = 3$, and a lower bound

Definition 3. *If $\mathcal{E}$ is a system of triples of elements of V and if X is a set, we define $G_{\mathcal{E}}(X)$ to be the graph with vertex set $V \setminus X$ and edges being all pairs $\{y, z\}$ such that $\{x, y, z\} \in \mathcal{E}$ for some $x \in X$.*

The following theorem is our main tool for proving that the set that we construct is indeed a blocking set.

Theorem 3. *$\mathcal{E}$ is a blocking set if and only if for each nonempty $X \subseteq V$, the graph $G_{\mathcal{E}}(X)$ is connected.*

Proof. Suppose that $\mathcal{E}$ is a blocking set and $X \subseteq V$. Suppose for contradiction that $V \setminus X$ can be decomposed into disjoint nonempty subsets Y, Z such that there are no edges of $G_{\mathcal{E}}(X)$ going from Y to Z. Then $\{X, Y, Z\}$ is a 3-partition that is not blocked.

Now suppose that for each $X \subseteq V$, the graph $G_{\mathcal{E}}(X)$ is connected, and let $\{X, Y, Z\}$ be a 3-partition. Since $G_{\mathcal{E}}(X)$ is connected, there is some $y \in Y$ and $z \in Z$ such that $\{y, z\}$ is an edge of $G_{\mathcal{E}}(X)$. That is, $\{x, y, z\} \in \mathcal{E}$ for some $x \in X$. Since $\{X, Y, Z\}$ was an arbitrary 3-partition, $\mathcal{E}$ is a blocking set. $\square$

In particular, if $\mathcal{E}$ is a blocking set, then $G_{\mathcal{E}}(\{x\})$ is connected for all x. In this case, for each $x \in V$, $G_{\mathcal{E}}(\{x\})$ has at least $n - 2$ edges (where n is the number of vertices), so at least $n - 2$ triples contain x. By double counting the number of pairs (x, e) where e is a triple in $\mathcal{E}$ that contains x, we get

$$\varphi(n) \geq \left\lceil \frac{n(n-2)}{3} \right\rceil. \tag{1}$$

It is easy to find, e.g., constructions with $\binom{n-1}{2}$ triples, but our goal is to match the lower bound (1) exactly. A theorem similar to Theorem 3 characterizing the d-blocking sets in terms of $(d-1)$-blocking sets holds for all d, but with increasing d, the situation seems to be too complex for such theorem to be useful.

4 Construction of a minimum blocking set for $d = 3$

Suppose now that A and B are two k-element sets, with $A = \{a_0, \ldots, a_{k-1}\}$ and $B = \{b_0, \ldots, b_{k-1}\}$. Let us define a hypergraph $\mathcal{H}(A, B)$ on the vertex set $A \cup B$. Its edge set consists of all triples $\{a_i, a_j, b_{i+j}\}$ and all triples $\{a_i, a_j, b_{i+j+1}\}$. Here and in the rest of the text, sequences are understood as circular and arithmetic operations on the indices are taken modulo k.

For each $n \in \mathbb{N}$, we define an n-vertex set V as follows. For $k = \lfloor n/3 \rfloor$, take the union of three disjoint k-element sets $A_0 = \{u_0, \ldots, u_{k-1}\}$, $A_1 = \{v_0, \ldots, v_{k-1}\}$, and $A_2 = \{w_0, \ldots, w_{k-1}\}$, and if n is congruent to 1 or 2 (modulo 3), add also either the element ∞ or two elements ∞_1, ∞_2, respectively.

Now let us define a set of triples $\mathcal{E}$. We start with all triples that are in the hypergraphs $\mathcal{H}(A_0, A_1)$, $\mathcal{H}(A_1, A_2)$, $\mathcal{H}(A_2, A_0)$. If n is divisible by 3, we add the k triples $\{u_i, v_i, w_i\}$ for $i = 0, \ldots, k-1$. If $n \equiv 1 \pmod 3$, add the triples $\{\infty, u_i, v_i\}$, $\{\infty, v_i, w_i\}$, $\{\infty, w_i, u_{i+1}\}$ for all i, and if $n \equiv 2 \pmod 3$, add the triples $\{\infty_1, u_i, v_i\}$, $\{\infty_1, v_i, w_i\}$, $\{\infty_2, w_i, u_{i+1}\}$, $\{\infty_2, u_i, v_{i+1}\}$, $\{\infty_1, \infty_2, w_i\}$ for all i.

The size of the set $\mathcal{E}$ matches the lower bound (1). To prove Theorem 1, we need to show that in each of the three cases, the set $\mathcal{E}$ is a blocking set. That is, using Theorem 3, we need to verify that for each $X \subseteq V$, the graph $G_{\mathcal{E}}(X)$ is connected. This is nontrivial (in fact, the proof takes multiple pages in the full version of our paper) and outside the scope of this extended abstract.

5 Lower and upper bound for $d > 3$

To prove a lower bound on $\varphi_d(n)$ for general $d > 3$, we use a similar (albeit more complicated) argument to the one we used for the lower bound for $d = 3$. We get the following:

Theorem 4. *For $d \geq 3$ and $n \geq d$, we have*

$$\varphi_d(n) \geq \frac{n(n-1)\cdot\dots\cdot(n-(d-3))(n-(d-1))}{d\cdot(d-1)\cdot\dots\cdot 3} = \frac{2n^{d-1}}{d!} - o(n^{d-1}).$$

As we have mentioned in Section 2, a simple upper bound is $\varphi_d(n) \leq \frac{n^{d-1}}{(d-1)!} - o(n^{d-1})$. Using the following construction, we can prove a better upper bound.

Construction for $d > 3$: Suppose that we have $n = 2k$. Partition our point set into two k-elements sets, A and B. For each $i \in \{0, \ldots, d-1\}$, let $S_i = \binom{A}{i}$ and let T_i be a $(d-i)$-blocking set of size $\varphi_{d-i}(k)$ for the set B (i.e., consisting of $(d-i)$-element subsets of B). Define $W_i = \{s_i \cup t_i | s_i \in S_i \text{ and } t_i \in T_i\}$. Now let $W = \cup_{i=0}^{d} W_i$.

We claim that W is a d-blocking set. Let $\{X_1, \ldots, X_d\}$ be a d-partition of $A \cup B$. Let i be the number of indices j such that $X_j \subseteq A$. If take all X_j such that $X_j \subseteq A$ and we pick an element from each of them, we will get an i-tuple, let us call it s_i. The remaining $d-i$ sets X_j, when restricted to B, form a $(d-i)$-partition of B. So T_i contains a $(d-i)$-tuple t_i blocking this partition. The union $s_i \cup t_i$ blocks the partition $\{X_1, \ldots, X_d\}$. But this set belongs to W_i, and thus to W. It follows that W is a blocking set.

Now let us get an estimate on the size of W. We would like to find some constant γ_d (dependent on d) such that, essentially, $\varphi_d(n) \leq \gamma_d n^{d-1}$. The set W_i has size $\binom{k}{i} \cdot \varphi_{d-i}(k)$, so

$$\varphi_d(2k) \leq |W| \leq \varphi_d(k) + \sum_{i=1}^{d-1} \binom{k}{i} \varphi_{d-i}(k). \tag{2}$$

Using this, it turns out that the following recursively defined γ_d works: we set $\gamma_1 = \gamma_2 = 1$, $\gamma_3 = \frac{1}{3}$, and for $d > 3$ put

$$\gamma_d = \frac{1}{2^{d-1}-1} \cdot \sum_{i=1}^{d-1} \frac{\gamma_{d-i}}{i!}.$$

Theorem 5. *For each d there is a constant β_d such that for every n with $n \geq d$ we have*

$$\varphi_d(n) \leq \gamma_d \cdot n^{d-1} + \beta_d \cdot n^{d-2}.$$

It is not hard to see (using induction and the binomial theorem) that this bound is no worse than the trivial upper bound on the constant, i.e., $\frac{1}{(d-1)!}$. The following theorem shows that it is strictly better.

Theorem 6. *For all $d \geq 3$, we have $\gamma_d < \frac{0.86}{(d-1)!}$.*

6 Remarks and open problems

This problem appeared naturally during our investigation of lines in finite metric spaces. See [2] for the results of this investigation, and [1] for more general context. However, it is clear that this extremal problem is interesting in its own right.

For $d > 3$ we were not able to find the asymptotic value of $\varphi_d(n)$. In this paper, we showed that there are constants c_1 and c_2 (dependent on d) such that $c_1 \cdot n^{d-1} \leq \varphi_d(n) \leq c_2 \cdot n^{d-1}$. Our feeling is that the lower bound might be tight, but that the upper bound is nowhere close to the truth. The ultimate goal is to close the gap between the constants. But even establishing the existence of the limit $\lim_{n\to\infty} \frac{\varphi_d(n)}{n^{d-1}}$ (without finding its value) would be interesting. A full paper containing the proofs of our results can be found on arXiv [8].

References

[1] Vašek Chvátal. A de Bruijn-Erdős theorem in graphs? In *Graph theory—favorite conjectures and open problems. 2*, Problem Books in Mathematics, pages 149–176. Springer, Cham, 2018.

[2] Vašek Chvátal, Noé de Rancourt, Guillermo Gamboa Quintero, Ida Kantor, and Péter G. N. Szabó. Metric spaces in which many triangles are degenerate. *Discrete Applied Mathematics*, 350:97–99, 2024.

[3] P. Frankl and V. Rödl. Near perfect coverings in graphs and hypergraphs. *European Journal of Combinatorics*, 6(4):317–326, 1985.

[4] Zoltán Füredi. Turán type problems. In *Surveys in combinatorics, 1991 (Guildford, 1991)*, volume 166 of *London Mathematical Society Lecture Note Series*, pages 253–300. Cambridge University Press, Cambridge, 1991.

[5] Peter Keevash. Hypergraph Turán problems. In *Surveys in combinatorics 2011*, volume 392 of *London Mathematical Society Lecture Note Series*, pages 83–139. Cambridge University Press, Cambridge, 2011.

[6] Daniela Kühn and Deryk Osthus. Embedding large subgraphs into dense graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Mathematical Society Lecture Note Series*, pages 137–167. Cambridge University Press, Cambridge, 2009.

[7] Nicholas Pippenger and Joel Spencer. Asymptotic behavior of the chromatic index for hypergraphs. *Journal of Combinatorial Theory Series A*, 51(1):24–42, 1989.

[8] Guillermo Gamboa Quintero and Ida Kantor. Minimum blocking sets for families of partitions. *arXiv preprint, arXiv:2505.15362*, 2025.

[9] Vojtech Rödl and Andrzej Ruciński. Dirac-type questions for hypergraphs—a survey (or more problems for Endre to solve). In *An irregular mind*, volume 21 of *Bolyai Society Mathematical Studies*, pages 561–590. János Bolyai Mathematical Society, Budapest, 2010.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Turán-type problems and the abstract chromatic number

(Extended abstract)

Dániel Gerbner* Hilal Hama Karim† Gaurav Kucheriya‡

Abstract

In 2020, Coregliano and Razborov introduced a general framework to study limits of combinatorial objects, using logic and model theory. They introduced the abstract chromatic number and proved/reproved multiple Erdős-Stone-Simonovits-type theorems in different settings. In 2022, Coregliano extended this by showing that similar results hold when we count copies of K_t instead of edges.

Our aim is threefold. First, we provide a purely combinatorial approach. Second, we extend their results by showing several other graph parameters and other settings where Erdős-Stone-Simonovits-type theorems follow. Third, we go beyond determining asymptotics and obtain corresponding stability, supersaturation, and sometimes even exact results.

1 Introduction

One of the most fundamental results in extremal combinatorics is the theorem of Turán [37], which determines the maximum number of edges among n-vertex graphs that do not contain K_{k+1} as a subgraph, in other words, K_{k+1}-free graphs. More generally, given a graph F, let $\mathrm{ex}(n, F)$ denote the largest $|E(G)|$ among n-vertex F-free graphs G. Turán's theorem [37] states that $\mathrm{ex}(n, K_{k+1}) = |E(T(n, k))|$, where $T(n, k)$ is the complete k-partite graph with each part of order $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. The celebrated Erdős-Stone-Simonovits (ESS) theorem [12, 14] is the most general result in the area, which states that the same holds if we forbid another graph with chromatic number $k + 1$, apart from an error term $o(n^2)$, i.e., for any graph F we

*HUN-REN Alfréd Rényi Institute of Mathematics, E-mail: gerbner@renyi.hu. Research supported by the National Research, Development and Innovation Office - NKFIH under the grants FK 132060 and KKP-133819.

†Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary. E-mail: hilal.hamakarim@edu.bme.hu. Research supported by the National Research, Development and Innovation Office - NKFIH under the grant FK 132060.

‡Department of Applied Mathematics, Charles University, Prague, Czechia, Email: gaurav@kam.mff.cuni.cz. Supported by GAČR grant 22-19073S and GAUK project 92125.

For proofs please refer to [19].

have $\mathrm{ex}(n, F) = |E(T(n, \chi(F) - 1))| + o(n^2)$. Note that this determines the asymptotics of $\mathrm{ex}(n, F)$, if F is not bipartite.

There are several different generalizations of the Turán problem. One line of research is to consider graphs with some extra structure. Various examples include vertex-ordered [33, 36], cyclically ordered [4] and edge-ordered graphs [21]. Probably the main common theme in all of these studies is obtaining an analogous result to the Erdős-Stone-Simonovits theorem. The essence of this is to find an appropriate notion of "chromatic number" in those contexts so that it can play the role of the usual chromatic number in the ESS result.

There have been attempts to study these types of problems in a unified and general way. Coregliano and Razborov [9] introduced a general model theoretic framework to study limits of combinatorial objects. They define *abstract chromatic number* of "open interpretations" on theories of graphs to capture such different notions of chromatic numbers in a unified way. They obtained the ESS result for the density of the edges in this general setting and Coregliano [8] extended this to the density of cliques.

Both in the ad hoc manner and in the unified approach, the general aim is to determine the objective extremum (e.g. the maximum number of edges or cliques) among the set of all graphs that can be underlying graphs of the graphs with the extra structure that possess some desired properties (e.g. not containing certain forbidden configurations). For example, in vertex-ordered graphs, we are interested in the maximum number of edges of an n-vertex vertex-ordered graph G avoiding F in an ordered sense. The ordering does not play any role in counting the edges; thus, we can think of this as counting the edges of the *underlying graph* of G, i.e., the ordinary graph we obtain from G by simply ignoring the ordering. The problem then reduces to finding the largest number of edges among graphs that can be underlying graphs of F-free vertex-ordered graphs. This way, the family of all graphs is partitioned into a family $\mathcal{A}(F)$ of *allowed graphs* and a family $\mathcal{F}(F)$ of forbidden graphs. In each case of graphs with an extra structure, and in the model theoretic approach, the corresponding chromatic number is defined in a way that if its value for a graph F is k, then we have $T(n, k-1)$ is among $\mathcal{A}(F)$ and for n large enough $T(n, k)$ is in $\mathcal{F}(F)$. This is the core idea in the proofs for the ESS-like results.

In this paper, we introduce a general, unified and yet purely combinatorial approach. We consider partitions $(\mathcal{A}, \mathcal{F})$ of the family of all graphs into $\mathcal{A}$ and $\mathcal{F}$, and define the "abstract chromatic number" of such partitions. Let $K_k(n)$ denote $T(nk, k)$ and let $T(n, \infty) = K_n$.

Definition 1.1. We say that a partition $(\mathcal{A}, \mathcal{F})$ is *Turán-suitable* if for all sufficiently large n one of the following condition holds:

- $K_n \in \mathcal{A}$.
- There exists an integer k such that each complete $(k-1)$-partite graph with each part of order at least n is in $\mathcal{A}$ but no $G \in \mathcal{A}$ contains $T(n, k)$ as a subgraph.

For simplicity, we will say *suitable* instead of Turán-suitable for the rest of this paper.

Next, we define the abstract chromatic number of a suitable partition, which coincides with the definition given in [9] in the simple specific cases they present as examples.

Definition 1.2. Given a suitable partition, its *abstract chromatic number* is ∞ if $K_n \in \mathcal{A}$ for all sufficiently large n. Otherwise, the abstract chromatic number is the largest k such that every complete $(k-1)$-partite graph with each part of order at least n is in $\mathcal{A}$, for every sufficiently large n.

Note that k is the same as in the definition of suitable partitions. Also note that our approach is in some sense stronger than that of [9]. They deal only with finitely axiomatizable theories (although mention that it is easy to extend their results). For example, the case that $\mathcal{A}$ is the family of bipartite graphs does not fit into their setting but is handled by our approach.

Let us consider some graph parameter $h(G)$, where h is a function from the finite graphs to the real numbers. Let

$$g(n, F) = g_h(n, F) = \max\{h(G) : G \text{ is an } n\text{-vertex } F\text{-free graph}\}.$$

Then we say that g is a *Turán-type function*. For instance, in the classical Turán problem $h(G) = |E(G)|$. For other examples, see Section 2.

We extend this to suitable partitions as follows.

$$g(n, (\mathcal{A}, \mathcal{F})) := \max\{h(G) : G \text{ is an } n\text{-vertex graph in } \mathcal{A}\}.$$

Note that if $(\mathcal{A}, \mathcal{F})$ is a monotone partition, then $g(n, (\mathcal{A}, \mathcal{F}))$ is simply $g(n, \mathcal{F})$.

As we mentioned above, an essential result in those generalizations of the Turán problem is the ESS theorem. Therefore, we define the notion of k-ESS for the Turán type functions.

Definition 1.3. Let T be an n-vertex complete $(k-1)$-partite graph and $h(T)$ be some real-valued graph parameter. We say that $g = g_h$ is *weakly k-ESS* if for any graph F with chromatic number k, $g(n, F) = (1 + o(1))h(T)$. We say that g is *strongly k-ESS* if the above holds with T being the Turán graph $T(n, k-1)$.

We say that g is *weakly k-ESS with respect to a partition* $(\mathcal{A}, \mathcal{F})$ if $g(n, (\mathcal{A}, \mathcal{F})) = (1 + o(1))h(T)$ for a complete $(k-1)$-partite graph T on n vertices. We say that g is *strongly k-ESS with respect to* $(\mathcal{A}, \mathcal{F})$ if the above holds with T being the Turán graph $T(n, k-1)$.

Theorem 1.1. *If g is a weakly (resp. strongly) k-ESS Turán-type function, then g is also weakly (resp. strongly) k-ESS with respect to any suitable partition with abstract chromatic number k.*

We say that a Turán-type function $g = g_h$ is *weakly k-ESS-stable*, if g is weakly k-ESS and for any graph F with chromatic number k, any F-free n-vertex G with $h(G) \geq (1-o(1))g(n, F)$ can be turned to an n-vertex complete $(k-1)$-partite graph T by adding and/or deleting $o(n^2)$ edges.

We can define *strongly k-ESS-stable* functions and weakly/strongly k-ESS-stable functions with respect to partitions analogously to the k-ESS functions.

Theorem 1.2. *If g is a weakly (resp. strongly) k-ESS-stable Turán-type function, then g is also weakly (resp. strongly) k-ESS-stable with respect to any suitable partition with abstract chromatic number k.*

Given a graph F of chromatic number k, we let $\sigma(F)$ denote the smallest color class among all possible proper k-colorings of F. Given a family $\mathcal{F}$ of graphs with smallest chromatic number k, we let $\sigma(\mathcal{F})$ be the smallest $\sigma(F)$ among k-chromatic elements of $\mathcal{F}$.

We say that a Turán-type function $g = g_h$ is *weakly k-ESS-sigma* if $g(n, F) = h(T)$ for some n-vertex graph T that we obtain from an $(n - \sigma(F) + 1)$-vertex complete $(k-1)$-partite graph by adding $\sigma(F)-1$ vertices and joining each of them to each other vertex. We can define *strongly k-ESS-sigma* functions and weakly/strongly k-ESS-sigma functions with respect to partitions analogously to the k-ESS functions.

Theorem 1.3. *If g is a weakly (resp. strongly) k-ESS-sigma Turán-type function, then g is also weakly (resp. strongly) k-ESS-sigma with respect to any suitable partition with abstract chromatic number k.*

We say that a Turán-type function $g = g_h$ is *k-ESS-supersat* if for any $\varepsilon > 0$ there is $\delta > 0$ such that for any sufficiently large n, any n-vertex graph G with $h(G) > (1+\varepsilon)g(n, F)$ we have that G contains at least $\delta n^{|V(F)|}$ copies of F, for any k-chromatic graph F. Again, we can define *k-ESS-supersat with respect to a partition* $(\mathcal{A}, \mathcal{F})$ analogously.

Theorem 1.4. *If g is a k-ESS-supersat Turán-type function, then g is also k-ESS-supersat with respect to any suitable partition with abstract chromatic number k.*

2 Turán-type functions and suitable partitions

Let us list some Turán-type functions that satisfy the requirements of some of our theorems. We start with k-ESS functions.

Counting edges, cliques. The examples in [8] and [9]. The Erdős-Stone-Simonovits theorem [12, 14] itself shows that counting edges is strongly k-ESS, and a theorem of Alon and Shikhelman [1] shows that counting K_t is strongly k-ESS if $k > t$.

Counting asymptotically (weakly) Turán-good graphs. Given a k-chromatic graph F, a graph H is *k-Turán-good* [22] if $\mathrm{ex}(n, H, F) = \mathcal{N}(H, T(n, k-1))$ and *weakly F-Turán-good* if $\mathrm{ex}(n, H, F) = \mathcal{N}(H, T)$ for some complete $(k-1)$-partite graph, where $\mathcal{N}(H, T)$ denotes the number of copies of H in G. We only need an asymptotic version.

Functions of degree sequences. Let f be a function and $h(G) := \sum_{v \in V} f(d(v))$. Pikhurko and Taraz [34] showed that g_h is weakly k-ESS for a large class of functions. The study of the special case $f(n) = n^r$, r is an integer was initiated by Caro and Yuster [7], who conjectured that this function is weakly k-ESS. It was proved for any real $r \geq 1$ by Bollobás and Nikiforov [2]. They showed in [3] that if $r \leq k$, then this function is strongly k-ESS.

Some topological indices. There are several topological indices of the form $h(G) = \sum_{uv \in E(G)} f(d(u), d(v))$. They are used in chemical graph theory. Gerbner [17] showed that is f is a monotone increasing polynomial, then g_h is weakly 3-ESS, moreover, weakly 3-ESS-stable.

Spectral radius. Let $h(G)$ denote the spectral radius of the adjacency matrix of G. Nikiforov [31] showed that g_h is strongly k-ESS.

p-spectral radius. Kang and Nikiforov [25] initiated the study of Turán-type problems for the p-spectral radius. This is defined as $h(G) = \max\{2\sum_{uv \in E(G)} x_u x_v : x_1, \dots, x_n \in \mathbb{R}, |x_1|^p + \cdots + |x_n|^p = 1\}$. Li and Peng [28] showed that g_h is strongly k-ESS.

Higher order spectral radius. The t-clique tensor of a graph G is an order t dimension n tensor, with entries $a_{i_1 i_2 \cdots i} = 1/(t-1)!$ if $v_{i_1}, \dots, v_{i_t}$ form a clique in G, and 0 otherwise. Let $h(G)$ denote the spectral radius of this tensor. Lu, Zhou and Bu [29] showed that g_h is strongly $(t+1)$-ESS.

Local density. Let $h(G) = h_\alpha(G)$ denote the smallest number of edges spanned by αn vertices of G. Keevash and Sudakov [27] showed that if $1 - 1/2(k-1)^2 \leq \alpha \leq 1$, then g_h is strongly k-ESS.

Small perturbations and combinations of k-ESS functions.

Let us continue with some k-ESS-stable functions.

Counting edges. The well-known Erdős-Simonovits stability [10, 11, 35] means that $\mathrm{ex}(n, F)$ is k-ESS-stable.

Counting (weakly) F-Turán-stable graphs. The first stability result concerning $\mathrm{ex}(n, H, F)$ is due to Ma and Qiu [30], who showed that counting cliques K_t is k-ESS-stable if $k > t$. Several other results followed, and in fact by now in most cases when we know that counting H is k-ESS, we also know that it is k-ESS-stable. Highlights include paths [24], complete t-partite graphs with $t < k$, and every graph if k is large enough [18]. Several other results can be found in [15].

Spectral radius. Nikiforov [32] showed that the spectral radius is strongly k-ESS-stable.

p-spectral radius. Li and Peng [28] showed that g_h is strongly k-ESS-stable.

Let us continue with a strongly k-ESS-sigma and a k-ESS-supersat function.

Counting large complete balanced $(k-1)$-partite graphs. Gerbner [16] showed that if H is the complete $(k-1)$-partite graph $K_{a,\dots,a}$ and a is large enough, then counting H is strongly k-ESS-sigma.

Counting subgraphs. It was shown in [13] that $\mathrm{ex}(n, F)$ is k-ESS-supersat. It was extended to every subgraph of chromatic number at most k by Halfpap and Palmer [23].

Let us continue with listing some suitable partitions.

Edge-ordered, vertex-ordered, cyclically ordered graphs. These were the main examples of graphs with extra structure in [8, 9]. The interested reader may find more details [4, 36, 21].

Forbidden induced family of graphs. This is another important example from [9].

Rainbow Turán. Keevash, Mubayi, Sudakov and Verstraëte [26] introduced the following problem. What is the maximum number of edges in an n-vertex graph that has a proper edge-coloring without a rainbow copy of F? Here rainbow copy of F means that each edge gets a distinct color. Counting other subgraphs in this setting was initiated in [20].

Finally, we mention that there are several similar results that can be obtained by modifying our definitions a little bit. An example is adding some assumption on h. Another is when we do not require every complete $(k-1)$-partite graph T to be in $\mathcal{A}$, only that there is a graph in $\mathcal{A}$ that is close to T. This can be applied e.g. for regular Turán problems [5, 6].

References

[1] N. Alon, C. Shikhelman. Many T copies in H-free graphs. *Journal of Combinatorial Theory, Series B*, **121**: 146–172, 2016.

[2] B. Bollobás, V. Nikiforov. Degree powers in graphs with forbidden subgraphs. *Electron. J. Combin.*, **11**, R42, 2004.

[3] B. Bollobás, V. Nikiforov. Degree powers in graphs: The Erdős-Stone Theorem *Combinatorics, Probability and Computing*, **21**, 89–105, 2012.

[4] P. Brass, G. Károlyi, P. Valtr. A Turán-type extremal theory of convex geometric graphs. *Discrete and Computational Geometry. Algorithms and Combinatorics*, **25**, 275–300, 2003.

[5] S. Cambie, R. de Joannis de Verclos, R. J. Kang. Regular Turán numbers and some Gan-Loh-Sudakov-type problems. *Journal of Graph Theory*, **102**(1), 67–85, 2023.

[6] Y. Caro and Zs. Tuza. Regular Turán numbers. *The Australasian Journal of Combinatorics*, **78**, 133–144, 2020.

[7] Y. Caro, R. Yuster. A Turán type problem concerning the powers of the degrees of a graph, *Electron. J. Combin.*, **7**, R47, 2000.

[8] L. N. Coregliano. On the abstract chromatic number and its computability for finitely axiomatizable theories. *Journal of Combinatorial Theory, Series B*, **154**: 175–210, 2022.

[9] L. N. Coregliano, A.A. Razborov. Semantic limits of dense combinatorial objects. *Usp. Mat. Nauk*, **75** (4): 45–152, 2020.

[10] P. Erdős. Some recent results on extremal problems in graph theory, *Theory of Graphs* (Internl. Symp. Rome), 118–123, 1966.

[11] P. Erdős. On some new inequalities concerning extremal properties of graphs, *Theory of Graphs* (Proc. Colloq., Tihany), 77–81, 1966.

[12] P. Erdős, M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.* **1**, 51–57, 1966.

[13] P. Erdős, M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, **3**(2), 181–192, 1983.

[14] P. Erdős, A. H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society* **52**, 1087–1091, 1946.

[15] D. Gerbner. Some stability and exact results in generalized Turán problems, *Studia Sci. Math. Hungarica*, **60**(1), 16–26, 2023.

[16] D. Gerbner. On the extremal graphs in generalized Turán problems. *Discrete Mathematics*, **347**(6), 114021, 2024.

[17] D. Gerbner. On extremal values of some degree-based topological indices with a forbidden or a prescribed subgraph. *Discrete Applied Mathematics*, **360**, 459–466, 2025.

[18] D. Gerbner and H. Hama Karim. Stability from graph symmetrization arguments in generalized Turán problems. *Journal of Graph Theory*, **107**(4), 681–692, 2024.

[19] D. Gerbner, H. Hama Karim and G. Kucheriya. On Turán-type problems and the abstract chromatic number. arXiv preprint arXiv:2410.09617.

[20] D. Gerbner, T. Mészáros, A. Methuku, C. Palmer, Generalized Rainbow Turán Numbers. *Electron. J. Combin.*, **29**(2), P2.44, 2022.

[21] D. Gerbner, A. Methuku, D.T. Nagy, D. Pálvölgyi, G. Tardos, M. Vizer: Turán problems for edge-ordered graphs, *Journal of Combinatorial Theory, Series B*, **160**, 66–113, 2023.

[22] D. Gerbner, C. Palmer, Some exact results for generalized Turán problems, *European Journal of Combinatorics*, **103**, 103519, 2022.

[23] A. Halfpap, C. Palmer. On supersaturation and stability for generalized Turán problems. *Journal of Graph Theory*, **97**(2), 232–240, 2021.

[24] D. Hei, X. Hou, B. Liu. Some exact results of the generalized Turán numbers for paths, *European Journal of Combinatorics*, **110**, 103682, 2023.

[25] L. Kang, V. Nikiforov, Extremal problem for the p-spectral radius of graphs, *Electron. J. Combin.*, **21**(3), P3.21, 2014.

[26] P. Keevash, D. Mubayi, B. Sudakov, J. Verstraëte. Rainbow Turán problems. *Combinatorics, Probability and Computing* **16**(1), 109–126, 2007.

[27] P. Keevash, B. Sudakov. Local density in graphs with forbidden subgraphs. *Combinatorics, Probability and Computing*, **12**(2), 139–153, 2003.

[28] Y. Li and Y. Peng. New proofs of stability theorems on spectral graph problems. arXiv preprint arXiv:2203.03142, 2022.

[29] C. Liu, J. Zhou, C. Bu, An Erdős-Stone type result for high-order spectra of graphs, arXiv preprint arXiv:2311.13368, 2023.

[30] J. Ma, Y. Qiu. Some sharp results on the generalized Turán numbers, *European Journal of Combinatorics*, **84**, 103026, 2018.

[31] V. Nikiforov. A spectral Erdős-Stone-Bollobás theorem. *Combinatorics, Probability and Computing*, **18**, 455–458, 2009.

[32] V. Nikiforov, Stability for large forbidden subgraphs, *Journal of Graph Theory*, **62**(4), 362–368, 2009.

[33] J. Pach, G. Tardos. Forbidden paths and cycles in ordered graphs and matrices. *Isr. J. Math.*, **155**, 359–380, 2006.

[34] O. Pikhurko and A. Taraz. Degree sequences of F-free graphs. *Electron. J. Combin.*, **12**, R69, 2005.

[35] M. Simonovits. Extremal graph problems with symmetrical extremal graphs. Additional chromatic conditions, *Discrete Mathematics*, **7**, 349–376, 1974.

[36] G. Tardos. Extremal theory of ordered graphs. *In Proceedings of the International Congress of Mathematics*, **3**, 3219–3228, 2018.

[37] P. Turán. On an extremal problem in graph theory (in Hungarian). *Matematikai és Fizikai Lapok* **48**, 436–452, 1941.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Facial diagrams and cycle double cover

(Extended abstract)

Babak Ghanbari* Robert Šámal†

Abstract

We approach the cycle double cover conjecture by looking for a circular 2-cell embedding of cubic graphs on an arbitrary surface. It is easy to see that if such an embedding exists, we can get to it from an arbitrary starting 2-cell embedding by repeating "twists of an edge". We study this twisting operation in detail and deduce bounds on the number of singular edges (edges where a face meets itself).

1 Introduction

The cycle double cover conjecture claims that for every bridgeless graph there is a collection of cycles covering each edge exactly twice. See [Zha12] for details and partial results. It is known that it is enough to show the conjecture for 3-regular graphs and that for such graphs it is equivalent to finding a circular 2-cell embedding on some surface. In other words, we are looking for an embedding such that the dual graph is loopless. Our goal in this paper is to represent the dual graph in a way that makes it easier to study the effects of a single twist operation.

In an embedding of a graph G, an edge e is called *singular* if there exists a facial walk F that traverses e twice. Otherwise, e is called *regular*. If a singular edge e is traversed twice in the same direction by F we call it *good singular*, and if e traversed twice in opposite directions by F we call it *bad singular*. Note that in orientable embeddings all singular edges are bad singular. For graph embeddings, we follow the notation of [MT01], in particular we will use pair (π, λ) to denote the combinatorial embedding of a graph.

2 Twist operation

We study the operation on graph embeddings called the **twist** of an edge e, i.e., the change of $\lambda(e)$. The following proposition shows the effect of twisting an edge common to two faces,

*Computer Science Institute of Charles University, Prague, Czech Republic. E-mail: babak@iuuk.mff.cuni.cz. Supported by the project GAUK182623 of the Charles University Grant Agency.

†Computer Science Institute of Charles University, Prague, Czech Republic. E-mail: samal@iuuk.mff.cuni.cz. Supported by grant 25-16627S of the Czech Science Foundation.

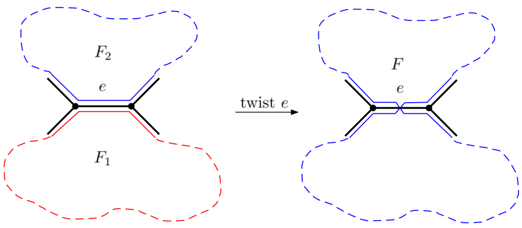


Figure 1: Twist of an edge e between facial walks F_1 and F_2

see Figure 1. Our goal in this paper is to study this in more detail with the motivation by the Cycle Double Cover conjecture.

Proposition 2.1. *Let F_1 and F_2 be distinct facial walks and let $e \in E(F_1) \cap E(F_2)$ be regular. Then, by the twist of e, F_1 and F_2 will become one facial walk.*

As mentioned in [MT01], each local change of rotation around a vertex v changes the signs of all three edges incident with v, which is the same as the twist of all three edges incident with v. Therefore, we have the following lemma.

Lemma 2.1. *Let (π, λ) be an embedding of a bridgeless cubic graph G and let (π', λ) be another embedding resulting from (π, λ) by a sequence of local changes of rotations. Then, one can find an embedding (π, λ') with a sequence of local twists of the edges in (π, λ) such that (π, λ') and (π', λ) have the same set of facial walks.*

In other words, one can start from any embedding in the graph and find all the other embeddings of the graph by twisting the edges. So, if the CDC conjecture is true, then given any embedding (π, λ) of a bridgeless cubic graph G, we can always find some (actually, every) CDC by a sequence of twist operations on (π, λ).

3 Facial diagram

In this section, we introduce another representation of facial walks that helps us in our study of the twist operation; we call it *facial diagram*. Let (π, λ) be an embedding for a bridgeless cubic graph G with facial walks $F_1, F_2, \ldots, F_k$ where $F_i = \{(v_0^i, e_1^i, v_1^i, e_2^i, \ldots, v_{t-1}^i, e_t^i)\}$. A *facial diagram*, or *FD* for G is a cubic graph H with its vertices (here we say node) and edges defined as follows. H has nodes $V(H) = \{e_j^i : e_j^i \in F_i\}$. Since an embedding covers every edge of G twice, we have $|V(H)| = 2|E(G)|$. H has the following three types of edges. (See Figures 2 and 3.)

1. **Facial link**: $v_j^i = (e_j^i, e_{j+1}^i)$ where $e_j^i, v_j^i, e_{j+1}^i \in F_i$. These are exactly the copies of the vertices of G that appear on the facial walks of the embedding.
2. **Singular link**: (e_j^i, e_k^i) where $e_j^i, e_k^i \in F_i$ and e_j^i and e_k^i are the same edge in G (but different nodes in H).
3. **Regular link**: (e_k^i, e_l^j) where $e_k^i = e_l^j$ as edges of G, $e_k^i \in F_i$, and $e_l^j \in F_j$.

Since there is a 1-1 correspondence between the set of edges of G and the set of singular and regular links of H, in the rest of the paper, if there is no confusion, we simply write a singular

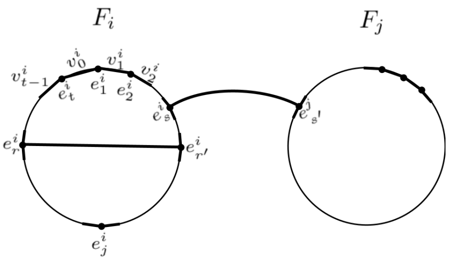
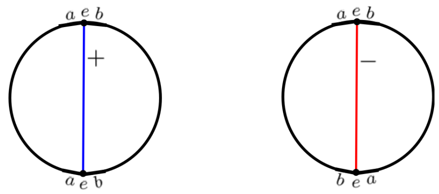


Figure 2: Representation of facial walks F_i and F_j in the facial diagram of a graph G with a singular link $(e_r^i, e_{r'}^i)$ and a regular link $(e_s^i, e_{s'}^j)$.

link $(e_j^i, e_k^i) \in F_i$ as e_j (its label in E(G)) and a regular link (e_k^i, e_l^j) between F_i and F_j as e_k (corresponding edge label in G); See Figure 3 as an example. We say that two singular links e_1 and e_2 are *crossing*, if they appear as $(\dots, e_1, \dots, e_2, \dots, e_1, \dots, e_2, \dots)$ in a facial walk F of the facial diagram.

Next, we define a signature for every singular link in H. Let $e = (a, b)$ be a singular edge in a facial walk F. If e is a bad singular edge i.e. $(\dots, a, e, b, \dots, b, e, a, \dots) \in F$, then we give the singular link e in H a $+$ sign and we call it a *bad singular link.* Otherwise, if e is a good singular edge i.e. $(\dots, a, e, b, \dots, a, e, b, \dots) \in F$, then we give the singular link e in H a $-$ sign and we call it a *good singular link.* See the following pictures.



Lemma 3.1. *Let e_1, and e_2 be two crossing singular links in a facial walk F.*

1. *If e_1 has $-$ sign. Then, by twist of e_1, both e_1, and e_2 turn to a regular link.*
2. *If e_1 has $+$ sign. Then twist of e_1 changes the sign of e_2 while the sign of e_1 remains unchanged.*

By Lemma 3.1, if we twist a $-$ link the number of singular edges of the new corresponding embedding reduces at least by 1. If we twist a $+$ link, the number of singular links remains unchanged. However, after the twist of e, the order of nodes and links on one side of e will change. We will use this property of the $+$ links to change the sign of some of the other singular links. See also Figure 5.

3.1 Properties of the facial diagram

The facial diagram has the following nice properties. We include their proofs in the full paper. Let G be a bridgeless cubic graph, F be a facial walk in an embedding of G, and H be a facial diagram of the same embedding. Let $e_1 = (a, b)$, $e_2 = (b, c)$, and $e_3 = (b, d) \in E(F)$. Then,

1. $(\dots, a, e_1, b, e_2, c, \dots, a, e_1, b, e_2, c, \dots) \notin F$.

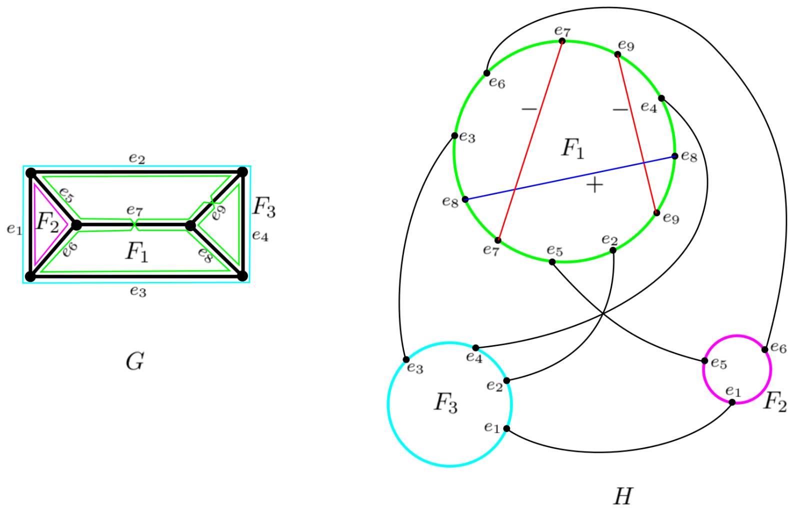


Figure 3: The facial diagram H of an embedding of a cubic graph G where e_8 is a bad singular link and e_7 and e_9 are good singular links. For simplicity, facial links' labels of H and vertex labels of G are omitted.

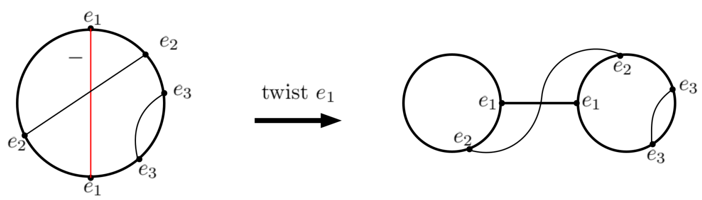


Figure 4: Twist of a $-$ link e_1. After the twist, e_2 turns to regular but e_3 remains singular.

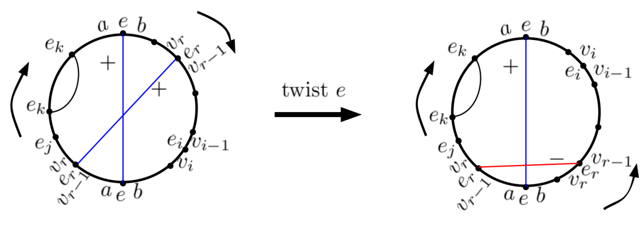
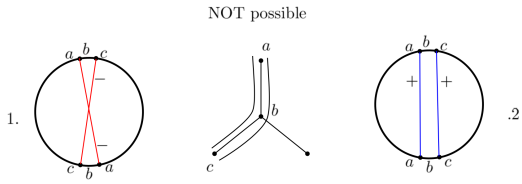


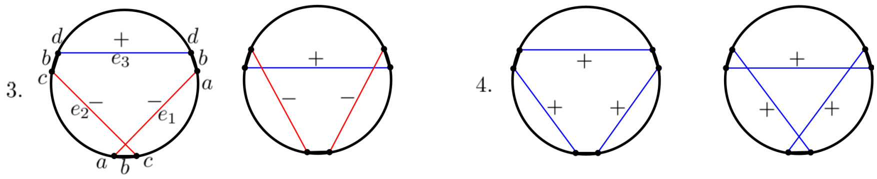
Figure 5: Twist of a $+$ link e. Observe how the order of nodes and facial links on one side of e changes. As a result sign of e_r changes to $-$ and then twist of e_r changes both e and e_r to regular edges.

2. $(\ldots, a, e_1, b, e_2, c, \ldots, c, e_2, b, e_1, a, \ldots) \notin F$.



3. Suppose that e_1 and e_2 are singular and have sign $-$. If $(\ldots, a, e_1, b, e_2, c, \ldots) \in F$ then, either $(\ldots, a, e_1, b, e_2, c, \ldots, a, e_1, b, e_3, d, \ldots, d, e_3, b, e_2, c, \ldots) \in F$, or $(\ldots, a, e_1, b, e_2, c, \ldots, d, e_3, b, e_2, c, \ldots, a, e_1, b, e_3, d, \ldots) \in F$

4. Suppose that e_1 and e_2 are singular and have sign $+$. If $\{\ldots, a, e_1, b, e_2, c, \ldots\} \in F$ then, either $\{\ldots, a, e_1, b, e_2, c, \ldots, c, e_2, b, e_3, d, \ldots, d, e_3, b, e_1, a, \ldots\} \in F$, or $\{\ldots, a, e_1, b, e_2, c, \ldots, d, e_3, b, e_1, a, \ldots, c, e_2, b, e_3, d, \ldots\} \in F$



5. We call a vertex v *saturated* if all the edges incident with v are singular. The number of $+$ edges incident with a saturated vertex is either 1 or 3.

6. Consider an embedding of a bridgeless cubic graph G with the minimum number of singular edges. The facial diagram of this embedding has no crossing links.

7. Let $e \in E(F_1) \cap E(F_2)$ be a regular link in a facial diagram H, If we twist e, the number of singular links in the new facial diagram H' is the number of singular links in H plus $|E(F_1) \cap E(F_2)|$.

8. Let e be a regular link. Then twist of e is a good singular link and therefore has sign $-$ in the facial diagram.

9. Twist of a regular link e does not change the sign of any singular links.

4 Random embedding

In this section, we approximate the number of singular edges by using random embeddings. Let (π, λ) be a random embedding of a cubic graph G. Results of several computer experiments counting the number of singular edges in a random embedding of a cubic graph indicate that the expected number of bad singular edges is $\frac{m}{3}$, where m is the number of edges in G. Also, in expectation, this number seems to be the same for good singular edges and regular edges. Therefore, we propose the following conjecture.

Conjecture 1. *In a random embedding of a bridgeless cubic graph G,*

1. *The expected number of bad singular edges is $\frac{m}{3}$.*
2. *The expected number of good singular edges is $\frac{m}{3}$.*
3. *The expected number of regular edges is $\frac{m}{3}$.*

As an application, we have the following theorem.

Theorem 4.1. *Suppose that Conjecture 1 is true and let G be a bridgeless cubic graph. Then, there exists an embedding of G with at most $\frac{m}{3}$ singular edges where $m = |E(G)|$.*

Proof. If conjecture 1 holds, then there exists an embedding and therefore a facial diagram where the number of singular links with sign $+$ is at most $\frac{m}{3}$. If every other link is regular, there is nothing to prove and the number of singular links (and therefore singular edges of the embedding) is at most $\frac{m}{3}$. Otherwise, start with any $-$ link e and apply Lemma 3.1 to the facial diagram. As a result, e and every link crossing with e changes to a regular link in the new facial diagram. Since it does not increase the number of $+$ links, one can repeat this process until there are no $-$ links without increasing the number of $+$ links. Therefore, there exists a facial diagram with at most $\frac{m}{3}$ bad singular links. $\square$

A referee suggested to us that the above result actually has a simpler proof: take any perfect matching, and extend the complementary 2-factor into a facial double cover. Then, only the edges of the matching can be singular (see [GŠ24]).

References

[GŠ24] Babak Ghanbari and Robert Šámal. Approximate cycle double cover. In Adele Anna Rescigno and Ugo Vaccaro, editors, *Combinatorial Algorithms*, pages 421–432, Cham, 2024. Springer Nature Switzerland.

[MT01] Bojan Mohar and Carsten Thomassen. *Graphs on Surfaces.* Johns Hopkins series in the mathematical sciences. Johns Hopkins University Press, 2001.

[Zha12] Cun-Quan Zhang. *Circuit double cover.* London Mathematical Society Lecture Note Series. Cambridge University Press, 2012.

On the largest k-product-free subsets of the Alternating Groups

(Extended abstract)

Anubhab Ghosal* Peter Keevash†

Abstract

A subset A of $\mathrm{Alt}(n)$ is k-product-free if for all $a_1, a_2, \dots, a_k \in A$, $a_1 a_2 \dots a_k \notin A$. We determine the largest 3-product-free and 4-product-free subsets of $\mathrm{Alt}(n)$ for sufficiently large n. We also obtain strong stability results and results on multiple sets with forbidden cross products. The principal technical ingredient in our approach is the theory of hypercontractivity in S_n.

1 Introduction

The problem of determining the size of the largest product-free subsets of groups was first considered by Babai and Sós [1] in 1985, who asked if there is some constant $c > 0$ such that any finite group G contains a product-free subset of density at least c. Gowers [3] answered this question in the negative, showing that in fact the maximum density is governed by the 'quasi-randomness' of the group, i.e. the smallest dimension among its non-trivial representations.

1.1 Largest product-free subsets of $\mathrm{Alt}(n)$

In particular, Gowers obtained the upper bound $O(n^{-\frac{1}{3}})$ for the alternating group $\mathrm{Alt}(n)$. In the other direction, Crane (reported by Green [4]) and Kedlaya [6] constructed the following product-free subset of $\mathrm{Alt}(n)$ which implies a lower bound of order $n^{-\frac{1}{2}}$ on $m(\mathrm{Alt}(n))$; diagrammatically, F_I^x can be visualised as $x \to I \not\to I$.

Definition 1.1. For $x \in [n]$ and $I \subseteq [n]$, define $F_I^x := \{\pi \in \mathrm{Alt}(n) : \pi(x) \in I, \pi(I) \cap I = \emptyset\}$.

Eberhard [2] improved the upper bound on $m(\mathrm{Alt}(n))$ to $n^{-\frac{1}{2}} \log^{O(1)}(n)$. Keevash, Lifshitz and Minzer completely resolved the question for sufficiently large n, in fact, specifying the structure of the extremal families.

Theorem 1.2 ([7]). *Suppose n is sufficiently large and A is a product-free subset of $\mathrm{Alt}(n)$ of maximal size. Then A or A^{-1} is some F_I^x.*

*Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK. E-mail: `ghosal@maths.ox.ac.uk`. Supported by the Clarendon Fund.

†Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK. E-mail: `keevash@maths.ox.ac.uk`. Supported by ERC Advanced Grant 883810.

1.2 Largest k-product-free subsets of $\mathrm{Alt}(\mathbf{n})$

We now consider the following natural generalisation of product-freeness to longer products.

Definition 1.3. A subset A of G is said to k-product-free if for all $a_1, a_2, \ldots, a_k \in A$, $a_1 a_2 \ldots a_k \notin A$. Let $m_k(G)$ denote the maximal density of a k-product-free subset of G.

The problem of determining $m_k(G)$ has been studied for various groups G. For the integers or the free group on a finite alphabet, Łuczak and Schoen [8] and Illingworth, Michel, and Scott [5], respectively, showed that $m_k(G)$ is $\rho(k)^{-1}$, where $\rho(k) := \min\{l \in \mathbb{N} : l \nmid k-1\}$. For $G = \mathrm{Alt}(n)$, the study of $m_k(G)$ is a case of [7, Problem 7.1]. We will report on some new results below; we will sketch the proof of one of these results but omit discussion of the rest due to lack of space.

The 'standard' construction, analogous to those considered in the previous literature, is to consider all permutations in $\mathrm{Alt}(n)$ that permute $\{1, 2, \ldots, \rho(k)\}$ cyclically. This set is easily seen to be k-product-free and gives us the lower bound $m_k(A_n) \geqslant n^{-\rho(k)}$.

On the other hand, we can show the following upper bound, which is fairly sharp for large even k, for which we have $n^{-2} \leqslant m_k(\mathrm{Alt}(n)) \leqslant O(n^{-2+\frac{4}{k+1}})$.

Theorem 1.4. *For $k \geqslant 6$, $m_k(\mathrm{Alt}(n)) = O(n^{-2+\frac{4}{k+1}})$.*

For $k = 3, 4$ we find an unexpected phenomenon: there are better constructions than the standard one, which we can show are optimal.

Definition 1.5. For $x, y \in [n]$, define $G^x_y := \{\pi \in \mathrm{Alt}(n) : \pi(x) = y, \pi(y) \neq x\}$.

Diagrammatically, G^x_y can be visualised as $x \to y \not\to x$. Note that G^x_y is 3-product-free and that $\mu(G^x_y) = (1 - o(1))n^{-1}$. We show that these are the largest 3-product-free sets for sufficiently large n, and, in fact, stability holds.

Theorem 1.6. *The following holds for n sufficiently large. Suppose A is a 3-product-free subset of $\mathrm{Alt}(n)$ satisfying $\mu(A) \geqslant (1 - 2^{-17})n^{-1}$. Then, A is a subset of some G^x_y.*

Next, we describe our candidates for maximal 4-product-free sets.

Definition 1.7. For distinct $x, y \in [n]$ and disjoint subsets S and T of $[n]$, define $H^{x,y}_{S,T} := \{\pi \in \mathrm{Alt}(n) : \pi(x) = y, \pi^{-1}(x) \in S, \pi(y) \in T\}$.

Diagrammatically, we can denote $H^{x,y}_{S,T}$ as $S \to x \to y \to T$. The density is maximised at $(0.25 - o(1))n^{-1}$ when S and T have size $(0.5 - o(1))n$ each. We will show that $H^{x,y}_{S,T}$ are the extremal 4-product-free subsets of $\mathrm{Alt}(n)$ for large enough n and that stability holds over a large regime of sizes.

Theorem 1.8. *Suppose n is large enough and that A is a 4-product-free subset of $\mathrm{Alt}(n)$ with $\mu(A) > n^{-1.19}$. Then, A is a subset of some $H^{x,y}_{S,T}$.*

1.3 k-product-free tuples

We also consider the cross problem, both for its own sake and as a tool in the proof for the original problem.

Definition 1.9. A tuple $\mathrm{A} = (A_0, A_1, \dots, A_k)$ of subsets of $\mathrm{Alt}(n)$ is said to be k-product-free if there are no solutions to the equation $a_1 a_2 \cdots a_k = a_0$, where $a_l \in A_l$ for all $l \in [k]_0$.

We consider the problem of maximising $\min\{\mu(A_0), \dots, \mu(A_k)\}$ among product-free tuples $(A_0,A_1,\dots, A_k)$. For $k = 2$, Eberhard [2] proved an upper bound of $\frac{\log^{O(1)}(n)}{\sqrt{n}}$ which is tight up to polylogarithmic factors; the problem was fully resolved in [7].

Note that $(\mathbb{1}_{k \not\to 0}, \mathbb{1}_{1\to 0}, \mathbb{1}_{2\to 1}, \dots, \mathbb{1}_{k\to k-1})$ is k-product-free in $\mathrm{Alt}(n)$. On the other hand, for $k = 3$ we have the following bound which matches this construction up to polylogarithmic factors.

Theorem 1.10. *Suppose (A_0, A_1, A_2, A_3) is 3-product free in $\mathrm{Alt}(n)$. Then,*

$$\min\{\mu(A_0), \mu(A_1), \mu(A_2), \mu(A_3)\} \leqslant \frac{\log^{O(1)}(n)}{n}.$$

Corollary 1.10.1. *Suppose $(A_0, \dots, A_k)$ is k-product free in $\mathrm{Alt}(n)$ with $A_l \neq \emptyset$ for all n. Then, the $(k-2)^{th}$ smallest among $\mu(A_0), \dots, \mu(A_k)$ is less than $\frac{\log^{O(1)}(n)}{n}$.*

Proof. Suppose $\mu(A_0)$, $\mu(A_1)$, $\mu(A_2)$ and $\mu(A_3)$ are the largest 4. Pick $a_i \in A_i$ for $i > 3$. Observe that $(A_0, A_1, A_2, A_3 \prod_{i>3} a_i)$ is 3-product-free and so we are done by Theorem 1.10. The other cases are similar. □

Corollary 1.10.2. *Suppose $A_0,\dots, A_k$ are non-empty subsets of $\mathrm{Alt}(n)$ so that the $(k-2)^{th}$ smallest among $\mu(A_0),\dots, \mu(A_k)$ is at least $\frac{\log^{O(1)}(n)}{n}$. Then $A_0 A_1 \dots A_k = \mathrm{Alt}(n)$.*

Proof. Let $g \in \mathrm{Alt}(n)$ and note that $(A_0^{-1} g, A_1, \dots, A_k)$ can't be k-product-free by Corollary 1.10.1. □

2 Proof of Theorem 1.10

In the rest of this extended abstract, we will sketch the key ideas required to prove Theorem 1.10.

2.1 Level decomposition and the weighted count of k-products

From now on, we will equip S_n with the uniform measure μ. Further, we will equip $\mathbb{R}^{S_n}$ with the expectation inner product and, to emphasise this, write $\mathbb{R}^{S_n}$ as $L^2(S_n)$. We will also consider expectation q-norms for functions in $L^2(S_n)$. Finally, we equip matrices in $\mathbb{R}^{n\times n}$ with the Frobenius inner product and the Frobenius norm.

For ordered sets $I = (i_1, \ldots, i_d), J = (j_1, \ldots, j_d) \in [n]_d$, let $\mathbb{1}_{I \to J} := \{\sigma \in S_n : \sigma(i_k) = j_k \text{ for all } k \in [d]\}$. A d-umvirate is any set of the above form. We will write $x_{I \to J}$ to mean the indicator function of $\mathbb{1}_{I \to J}$. Let W_d denote the linear span of the indicators of the d-umvirates and let $V_{=d} := W_d \cap W_{d-1}^{\perp}$. Note that $W_n = W_{n-1}$, and so we get an orthogonal decomposition $L^2(S_n) = \bigoplus_{d=0}^{n-1} V_{=d}$. For $f \in L^2(S_n)$, and $d \in [n-1]_0$, define $f^{=d}$ as the projection of f onto $V_{=d}$.

Definition 2.1. For $I, J \in [n]_d$ and $f \in L^2(S_n)$, define $\mathbb{E}\left[f_{I \to J}\right] := \mathbb{E}_{\sigma \sim S_n}\left[f \mid \mathbb{1}_{I \to J}\right]$.

Definition 2.2. $T(f_0, f_1, \ldots, f_k) = \mathbb{E}_{\pi_1, \ldots, \pi_k \sim S_n}[f_0(\pi_1 \ldots \pi_k) f_1(\pi_1) \ldots f_k(\pi_k)]$.

The next lemma states that for f_i supported on $\mathrm{Alt}(n)$, the weighted count T of k-products distributes over the levels, with the higher degree terms contributing error terms to the sum. The former is a consequence of the representation theory of S_n, whereas the latter follows from the trace method. A proof for the case $k = 2$ can be found in [7, Lemmas 2.12 and 2.13] which straightforwardly generalises to arbitrary k.

Lemma 2.3. *Suppose $f_l \in L^2(S_n)$ are supported on $\mathrm{Alt}(n)$ with $\|f_l\|_2^2 = \alpha_l$ for $l \in [k]_0$ and that $d + 1 < \frac{n}{10}$. Then,*

$$T(f_0, f_1, \ldots, f_k) = 2 \cdot \sum_{d=0}^{\tilde{d}} T(f_0^{=d}, f_1^{=d}, \ldots, f_k^{=d}) + O_{\tilde{d}}\left(n^{-\frac{(k-1)(d+1)}{2}} \left(\prod_{l=0}^{k} \alpha_l\right)^{\frac{1}{2}}\right). \tag{1}$$

In particular, for $k = 3$, and $\alpha_l \gg n^{-1}$, one obtains that

$$T(f_0, \ldots, f_3) > \prod_{l=0}^{3} \alpha_l + T(f_0^{=1}, \ldots, f_3^{=1}) \tag{2}$$

2.2 Analysis of Linear Terms and Density Increment

We now recall the analysis of functions in $V_{=1}$ from [7]. One can write $f^{=1}$ as $\sum_{i,j} a(i,j) x_{i \to j}$, where $a(i,j) = \left(1 - \frac{1}{n}\right)\left(\mathbb{E}\left[f_{i \to j}\right] - \mathbb{E}\left[f\right]\right)$. Writing M_f for the matrix $(a(i,j))$, it further turns out that $T(f_0^{=1}, f_1^{=1}, \ldots, f_k^{=1}) = \frac{1}{(n-1)^k} \left\langle \prod_{l=k}^{1} M_{f_l}, M_{f_0} \right\rangle$ Therefore, it holds that

$$|T(f_0^{=1}, \ldots, f_3^{=1})| \leqslant \frac{1}{(n-1)^3} \prod_{l=0}^{3} \|M_{f_l}\|. \tag{3}$$

We now recall the pseudorandom level-1 inequality from [7] and use it to prove a key density increment proposition.

Lemma 2.4. *[7, Lemma 3.6] Let $\varepsilon \in \left(0, \frac{1}{2}\right)$ and $f \colon S_n \to \{0,1\}$ with $\mathbb{E}\left[f\right] \leqslant \frac{1}{2}$. Write $f^{=1}$ in normalised form with coefficient matrix M. Denote $\varepsilon'' = \max\left(\varepsilon, \mathbb{E}\left[f\right]\right)$ and $M^{\varepsilon} := (a(i,j) \mathbb{1}_{|a(i,j)| < \varepsilon})$. Then, $\frac{1}{n-1}\|M^{\varepsilon}\|_2^2 \leqslant \mathbb{E}\left[f\right] \varepsilon'' \log^{O(1)}\left(\frac{1}{\varepsilon''}\right)$.*

Proposition 2.5. *Suppose* $(A_0, \cdots, A_3)$ *is* 3*-product-free in* $\mathrm{Alt}(n)$ *with* $\mu(A_l) = \alpha_l \gg n^{-1}$. *Then, there exist* $l \in [k]_0$ *and* $i, j \in [n]^2$ *so that* $\mu(A_l^{i\to j}) \geqslant \alpha_l n^{0.48}$.

Proof. Let $f_l = \mathbb{1}_{A_l}$. Suppose, for contradiction, that $\mu(A_l^{i\to j}) \leqslant \alpha_l n^{0.49}$, for all i, j and l. Then, the entries of M_{f_l} are bounded in absolute value by $\varepsilon_l := \alpha_l n^{0.48}$ we can apply Lemma 2.4 to get that $\|M_{f_l}\| \leqslant \alpha_l n^{0.74}$. Combined with Equations (2) and (3), this implies $T(f_0, \ldots, f_3) > 0$, a contradiction. □

The following fairness proposition ensures that most restrictions maintain the density of a set. It is a straightforward deduction from Lemma 2.4.

Proposition 2.6. *Let* A *be a subset of* S_n *with* $n^{-3} < \mu(A) < \frac{1}{2}$. *Then, there are less than* $n \log^{O(1)}(n)$ *pairs* $(i, j) \in [n]^2$ *with* $\mu(A_{i\to j}) < 0.98\mu(A)$.

Proof of Theorem 1.10. Suppose that $(A_0, \ldots, A_3)$ is 3-product-free and that

$$\mu(A_l) \geqslant \frac{\log^K(n)}{n} \text{ for } l \in [3]_0.$$

We want to arrive at a contradiction for large enough K. We will apply Proposition 2.5 to get a l and i, j so that

$$\mu(A_l^{i\to j}) \geqslant \mu(A_l) n^{0.48} \tag{4}$$

We will now remove the umvirate $\mathbb{1}_{i\to j}$ from A_l and continue. We can keep iterating this process until one of $\mu(A_l)$ becomes less than n^{-1}. As at each step, we remove a restriction of size at most n^{-1}, this process runs for at least $\log^K(n) - 1$ steps. At the end, we can find an l and at least $\frac{\log^K(n)-1}{4}$ directed edges $i \to j$ so that Equation (4) holds. In this directed graph, there is either a matching or a star or an inverse star of size at least $\log^{\frac{K}{4}}(n)$. Noting that $(A_0, \ldots, A_3)$ is 3-product-free iff $(A_3^{-1}, A_0^{-1}, A_1, A_2)$ is 3-product-free and so on, we can assume without loss of generality that $l = 3$. Suppose first that we have a large star. Label the centre of the star by 1. We therefore get edges $(1, j_1), (1, j_2), \cdots, (1, j_t)$ where $t \geqslant \log^{\frac{K}{4}}(n)$ with density of A_3 incremented along these edges as in Equation (4).

Consider the set $\mathcal{C} = \{c \in [n] : \mu(A_0^{1\to c}) \geqslant 0.5\mu(A_0)\}$. By Markov's inequality, there are atleast $0.5\mu(A_0)n \geqslant \log^{\frac{K}{4}}(n)$ elements in $\mathcal{C}$. Pick r u.a.r. from $[t]$, c u.a.r. from $\mathcal{C}$ and u u.a.r. from $[n]$, all independently and consider the diagram $1 \to j_r \to u \to c$. Applying Proposition 2.6, one gets that w.h.p., $\mu(A_2^{j_r\to u}) \geqslant 0.98\mu(A_2)$ and $\mu(A_1^{u\to c}) \geqslant 0.98\mu(A_1)$, provided K is large enough. Now, let $A_0' = (cn)A_0^{1\to c}(1n)$, $A_1' = (cn)A_1^{u\to c}(un)$, $A_2' = (un)A_1^{j_r\to u}(j_r n)$ and $A_3' = (j_r n)A_0^{1\to j_r}(1n)$. Then, w.h.p., $(A_0', \ldots, A_3')$ is a 3-product-free tuple in $\mathrm{Alt}(n-1)$, with atleast one of the densities incremented by $n^{0.48}$, while the others fall by at most a factor of 0.5.

The case where we have an inverse star or a matching is very similar. In conclusion, provided K is a large enough constant, we can always find 3-product-free tuples in $\mathrm{Alt}(n-1)$ with density incremented as above. We can now iterate this process, which leads to a contradiction after at most 100 steps as densities are bounded above by 1. □

References

[1] László Babai and Vera T Sós. Sidon sets in groups and induced subgraphs of cayley graphs. *European Journal of Combinatorics*, 6(2):101–114, 1985.

[2] Sean Eberhard. Product mixing in the alternating group. *Discrete Analysis*, 2(2016), 2016.

[3] W.T. Gowers. Quasirandom groups. *Combinatorics, Probability and Computing*, 17(3):363–387, 2008.

[4] B.J. Green. 100 open problems. *manuscript*, 2024.

[5] Freddie Illingworth, Lukas Michel, and Alex Scott. The structure and density of k k-product-free sets in the free semigroup and group. *Journal of the London Mathematical Society*, 111(1):e70046, 2025.

[6] Kiran S. Kedlaya. Product-free subsets of groups. *The American Mathematical Monthly*, 105(10):900–906, 1998.

[7] Peter Keevash, Noam Lifshitz, and Dor Minzer. On the largest product-free subsets of the alternating groups. *Inventiones mathematicae*, 237(3):1329–1375, 2024.

[8] Tomasz Łuczak and Tomasz Schoen. On infinite sum-free sets of natural numbers. *Journal of Number Theory*, 66(2):211–224, 1997.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25–29, 2025

Blowups of triangle-free graphs

(Extended abstract)

António Girão* Zach Hunter† Yuval Wigderson‡

Abstract

A highly influential result of Nikiforov states that if an n-vertex graph G contains at least γn^h copies of a fixed h-vertex graph H, then G contains a blowup of H of order $\Omega_{\gamma,H}(\log n)$. While the dependence on n is optimal, the correct dependence on γ is unknown; all known proofs yield bounds that are polynomial in γ, but the best known upper bound, coming from random graphs, is only logarithmic in γ. It is a major open problem to narrow this gap.

We prove that if H is triangle-free, then the logarithmic behavior of the upper bound is the truth. That is, under the assumptions above, G contains a blowup of H of order $\Omega_H(\log n/\log(1/\gamma))$. This is the first non-trivial instance where the optimal dependence in Nikiforov's theorem is known.

As a consequence, we also prove an upper bound on multicolor Ramsey numbers of blowups of triangle-free graphs, proving that the dependence on the number of colors is polynomial once the blowup is sufficiently large. This shows that, from the perspective of multicolor Ramsey numbers, blowups of fixed triangle-free graphs behave like bipartite graphs.

1 Background and main results

Given an integer k and a graph H, its *blowup* $H[k]$ is the graph obtained from H by replacing every vertex by an independent set of order k, and every edge by a copy of the

*Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK. Research supported by ERC Advanced Grant no. 883810. E-mail: `girao@maths.ox.ac.uk`

†Department of Mathematics, ETH Zürich, 8092 Zürich, Switzerland. Email: `zach.hunter@math.ethz.ch`

‡Institute for Theoretical Studies, ETH Zürich, 8006 Zürich, Switzerland. Research supported by Dr. Max Rössler, the Walter Haefner Foundation, and the ETH Zürich Foundation. Email: `yuval.wigderson@eth-its.ethz.ch`

complete bipartite graph $K_{k,k}$. Blowups are fundamental objects in graph theory, and many important results in extremal graph theory concern the problem of finding large H-blowups in graphs with certain properties. For example, the Erdős–Stone theorem [12] states that given any graph H, integer k, and parameter $\varepsilon > 0$, any sufficiently large graph G with edge density at least $1 - \frac{1}{\chi(H)-1} + \varepsilon$ contains $H[k]$ as a subgraph. Much of the subsequent work in extremal graph theory, culminating in the Chvátal–Szemerédi theorem [7], has been focused on determining the optimal "sufficiently large" condition in this theorem.

A closely related line of research was initiated by Nikiforov [21, 22], who proved the following remarkable theorem.

Theorem 1.1 (Nikiforov [21, 22]). *Let H be an h-vertex graph, let $\gamma > 0$, and let n be sufficiently large. If G is an n-vertex graph with at least γn^h copies of H, then G contains an H-blowup $H[k]$, where*

$$k \geq c_H(\gamma) \log n,$$

for some constant $c_H(\gamma) > 0$ depending only on γ and H.

This result gives the best possible dependence on n, since a standard computation shows that a random n-vertex graph has, with high probability, $\Omega(n^h)$ copies of H and no copy of $H[k]$ for any $k \geq 2 \log n$. We remark too that one natural approach to proving such a theorem—passing to an auxiliary h-uniform hypergraph whose edges are the H-copies in G—does not work, and can only prove a bound of $k = \Omega((\log n)^{\frac{1}{h-1}})$. Thus, Theorem 1.1 is one of many theorems (in addition to the famous (6,3) theorem [25], among others) capturing the idea that the H-copies in a graph G have extra structure beyond what is found in a general h-uniform hypergraph.

While Theorem 1.1 gives the optimal n-dependence, the optimal dependence on γ and H is unknown. Nikiforov proved that, for any fixed graph H, one can take $c_H(\gamma) = \Omega(\gamma^h)$ if $H = K_h$ is a complete graph, and $c_H(\gamma) = \Omega(\gamma^{h^2})$ in general. However, this is very far from the best known upper bound $c_H(\gamma) = O_H(1/\log \frac{1}{\gamma})$, which again comes from considering a random graph of the appropriate edge density, namely $\gamma^{1/e(H)}$.

Theorem 1.1 is an extremely useful result with many applications (e.g. [13, 16, 23, 24]), and as such, there have been several attempts to improve the bounds on $c_H(\gamma)$. Rödl and Schacht [24] proved that we may take $c_{K_h}(\gamma) \geq \gamma^{1+o(1)}$ when H is complete, and Fox–Luo–Wigderson [13] improved this to $c_H(\gamma) \geq \gamma^{1-1/e(H)+o(1)}$ for all H. However, just as in Nikiforov's original argument, all of these bounds are still polynomial in γ, whereas the best known upper bound is logarithmic. The only case where the truth is known is when H is bipartite (which is a degenerate case of the problem); in this case, the Kővári–Sós–Turán theorem [17] immediately implies that $c_H(\gamma) = \Omega_H(1/\log \frac{1}{\gamma})$, matching the upper bound up to a constant factor.

Our main result proves the same bound for all triangle-free graphs H, yielding the first non-trivial case where the optimal bound in Theorem 1.1 is known.

Theorem 1.2. *For every triangle-free graph H on h vertices, there exists a constant $\alpha_H > 0$ such that the following holds for all $0 < \gamma \leq \frac{1}{2}$ and all n. If G is an n-vertex graph with at least γn^h copies of H, then $H[k] \subseteq G$, where*

$$k \geq \alpha_H \frac{\log n}{\log \frac{1}{\gamma}}.$$

It is natural to conjecture that the same result is true for all graphs H. However, it appears that proving this, even in the simplest case of $H = K_3$, would require substantial new techniques.

As a consequence of Theorem 1.2, we obtain a surprising result about multicolor Ramsey numbers. Recall that, for a graph F and an integer $q \geq 2$, the *Ramsey number* $r(F; q)$ is defined as the least integer N such that every q-coloring of $E(K_N)$ contains a monochromatic copy of F. In general, our understanding of $r(F; q)$ is rather poor; for example, it is a major open problem [11, 14, 20] to determine whether $r(K_3; q)$ grows exponentially or super-exponentially with q, and an even more major open problem [5, 8, 26, 27] to determine the growth rate of $r(K_k; 2)$ as $k \to \infty$. However, for complete bipartite graphs, our understanding is fairly complete, and it is known [6] that

$$q^{ck} \leq r(K_{k,k}; q) \leq q^{Ck} \tag{1}$$

for all $q, k \geq 2$, where $C > c > 0$ are absolute constants. Here, the lower bound follows from a random coloring, and the upper bound follows from the Kővári–Sós–Turán theorem. In particular, this implies that for fixed F, $r(F; q)$ grows polynomially in q if F is bipartite, whereas it is easy to see[1] that $r(F; q)$ grows at least exponentially in q if not. However, our next result shows that if F is a large blowup of a fixed triangle-free graph, then the dependence on q *does* eventually become polynomial.

Theorem 1.3. *Let H be an h-vertex triangle-free graph, and let $q \geq 2$ be an integer. If $k \geq 100^h q^{4h^2}$, then*

$$r(H[k]; q) \leq q^{\Lambda_H k},$$

where $\Lambda_H > 0$ is a constant depending only on H.

This result is best possible up to the constant Λ_H, since a random coloring again witnesses that $r(H[k]; q) \geq q^{ck}$ for every non-empty graph H and some absolute constant $c > 0$. Moreover, as discussed above, the assumption that k is sufficiently large is also necessary, since the dependence on q is super-polynomial if k is fixed and q is large whenever H is non-bipartite. In fact, this argument shows that the "sufficiently large" condition on k in Theorem 1.3 is nearly best possible in terms of the q-dependence, in that k must be at least of order $q^{\Omega_H(1)}$ for such a statement such as Theorem 1.3 to be true.

[1] Indeed, since K_{2^q} is the edge-union of q bipartite graphs, we have $r(F; q) > 2^q$ in case F is non-bipartite.

2 Discussion and corollaries

One interesting feature of Theorem 1.2 is that we do not assume that n is sufficiently large with respect to γ, in contrast to previous results on this topic. In particular, Theorem 1.2 gives a non-trivial result even when γ is a small negative power of n, as stated in the following result.

Corollary 2.1. *Let H be an h-vertex triangle-free graph, and let G be an n-vertex graph. If G contains at least $n^{h-\alpha_H/k}$ copies of H, then $H[k] \subseteq G$.*

Indeed, Theorem 2.1 follows immediately from Theorem 1.2 by plugging in $\gamma = n^{-\alpha_H/k}$. This result can be equivalently stated in the language of *generalized extremal numbers* [1], where we recall that $\mathrm{ex}(n, H, F)$ denotes the maximum number of copies of H that can appear in an n-vertex F-free graph. In this language, Theorem 2.1 states that

$$\mathrm{ex}(n, H, H[k]) < n^{h-\alpha_H/k}$$

for all triangle-free H. For general graphs H, the best known upper bound for this problem follows by a reduction to a hypergraph extremal problem, which yields the bound

$$\mathrm{ex}(n, H, H[k]) = O_H\left(n^{h-1/k^{h-1}}\right) \tag{2}$$

for any h-vertex graph H. Recently, several authors [2, 3, 19] have attempted to improve this bound; in particular, the results of [2, 3] imply that $\mathrm{ex}(n, H, H[k]) = o(n^{h-1/k^{h-1}})$. However, their proof techniques rely on the (hyper)graph removal lemma, and therefore give only a very slight improvement over (2). In contrast, Theorem 2.1 gives a power-savings improvement over (2) whenever k is sufficiently large in terms of H; to the best of our knowledge, this is the first example of such a power-savings improvement for a non-bipartite graph H. We remark that proving an analogue of Theorem 2.1 in the first non-trivial case of $H = K_3$ would, in particular, make progress towards a conjecture of Fox, Sankar, Simkin, Tidor, and Zhou [15, Conjecture 6.4] on the extremal numbers of Latin squares.

More generally, one can let $\gamma = \gamma(n)$ decay to 0 at some rate as $n \to \infty$, and Theorem 1.2 can still yield useful results. Results along these lines for other choices of $\gamma(n)$ have implications for certain hypergraph extremal problems. For example, Theorem 1.2 shows that if G has at least $\exp(-\sqrt{\log n})n^h$ copies of a triangle-free graph H, then it contains a blowup of H of order $\Omega(\sqrt{\log n})$. Rödl and Schacht [24, Problem 3] showed that such a statement for $H = K_3$ would yield a result like Theorem 1.1 in hypergraphs, which remains a major open problem. Similarly, Conlon, Fox, and Sudakov [9, Theorem 1.1] proved a certain analogue of the Erdős–Hajnal conjecture for 3-uniform hypergraphs, but conjectured [9, Conjecture 1; 10, Conjecture 3.16] that their result could be quantitatively strengthened. The main barrier to improving their result is strengthening a key technical lemma [9, Lemma 3.3], which would roughly boil down to proving that an n-vertex graph with $\Omega(n^3/\log n)$ triangles contains a copy of $K_3[k]$ with $k \geq (\log n)^{1-o(1)}$. Again, Theorem 1.2 implies that such a result holds if we replace H by any triangle-free graph.

While blowups are interesting in their own right, results about graph blowups are generally of great utility, and often immediately imply results about other graphs. For example, the following is an immediate consequence of Theorem 1.3, combined with a deep result of Łuczak [18] on the *homomorphism threshold* of triangle-free graphs.

Corollary 2.2. *For every $\beta > 0$, there exists $C_\beta > 0$ such that the following holds for all $q \geq 2$ and all sufficiently large k. If F is a k-vertex triangle-free graph with minimum degree at least $(\frac{1}{3} + \beta)k$, then*

$$r(F; q) \leq q^{C_\beta k}.$$

Indeed, the result of Łuczak [18] mentioned above states that there exists a triangle-free graph[2] H_β, depending only on β, such that any graph F satisfying the assumption of Theorem 2.2 is a subgraph of $H_\beta[k]$, and thus Theorem 2.2 is an immediate consequence of Theorem 1.3.

As discussed above, it is natural to conjecture that Theorem 1.2 holds for all H. If true, this would imply that Theorem 1.3 holds for all H; in particular, the $H = K_h$ case of this result would imply that $r(F; q)$ is polynomial in q whenever F is a sufficiently large graph of bounded chromatic number. We believe that such a statement is interesting in its own right, as it shows that the bipartite behavior carries through whenever $\chi(F)$ is bounded. Moreover, proving such a statement may be easier than extending Theorem 1.2 to all H. Even the $H = K_3$ case seems interesting and challenging.

Conjecture 2.3. *For all $q \geq 2$ and all sufficiently large k, we have*

$$r(K_{k,k,k}; q) \leq q^{Ck},$$

where $C > 0$ is an absolute constant.

References

[1] N. Alon and C. Shikhelman, Many T copies in H-free graphs, *J. Combin. Theory Ser. B* **121** (2016), 146–172.

[2] J. Balogh, S. Jiang, and H. Luo, On the maximum number of r-cliques in graphs free of complete r-partite subgraphs, 2024. Preprint available at arXiv:2402.16818.

[3] A. Basu, V. Rödl, and Y. Zhao, The number of cliques in hypergraphs with forbidden subgraphs, 2024. Preprint available at arXiv:2405.07763.

[4] S. Brandt and S. Thomassé, Dense triangle-free graphs are four-colorable: A solution to the Erdős-Simonovits problem, 2011. Preprint available at https://perso.ens-lyon.fr/stephan.thomasse/liste/vega11.pdf.

[2] In [4, Corollary 4.3(3)], a precise description of these graphs H_β is given; they are the so-called *Vega graphs*.

[5] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe, An exponential improvement for diagonal Ramsey, 2023. Preprint available at arXiv:2303.09521.

[6] F. R. K. Chung and R. L. Graham, On multicolor Ramsey numbers for complete bipartite graphs, *J. Combin. Theory Ser. B* **18** (1975), 164–169.

[7] V. Chvátal and E. Szemerédi, On the Erdős-Stone theorem, *J. London Math. Soc. (2)* **23** (1981), 207–214.

[8] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. of Math. (2)* **170** (2009), 941–960.

[9] D. Conlon, J. Fox, and B. Sudakov, Erdős–Hajnal-type theorems in hypergraphs, *J. Combin. Theory Ser. B* **102** (2012), 1142–1154.

[10] D. Conlon, J. Fox, and B. Sudakov, Recent developments in graph Ramsey theory, in *Surveys in combinatorics 2015*, *London Math. Soc. Lecture Note Ser.*, vol. 424, Cambridge Univ. Press, Cambridge, 2015, 49–118.

[11] P. Erdős, Some unsolved problems, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** (1961), 221–254.

[12] P. Erdös and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.

[13] J. Fox, S. Luo, and Y. Wigderson, Extremal and Ramsey results on graph blowups, *J. Comb.* **12** (2021), 1–15.

[14] J. Fox, J. Pach, and A. Suk, Bounded VC-dimension implies the Schur-Erdős conjecture, *Combinatorica* **41** (2021), 803–813.

[15] J. Fox, M. Sankar, M. Simkin, J. Tidor, and Y. Zhou, Ramsey and Turán numbers of sparse hypergraphs, 2023. Preprint available at arXiv:2401.00359.

[16] N. Kamčev and A. Müyesser, Unavoidable patterns in locally balanced colourings, *Combin. Probab. Comput.* **32** (2023), 796–808.

[17] T. Kövari, V. T. Sós, and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.

[18] T. Łuczak, On the structure of triangle-free graphs of large minimum degree, *Combinatorica* **26** (2006), 489–493.

[19] D. Mubayi and S. Mukherjee, Triangles in graphs without bipartite suspensions, *Discrete Math.* **346** (2023), Paper No. 113355, 19pp.

[20] J. Nešetřil and M. Rosenfeld, I. Schur, C. E. Shannon and Ramsey numbers, a short story, *Discrete Math.* **229** (2001), 185–195.

[21] V. Nikiforov, Graphs with many copies of a given subgraph, *Electron. J. Combin.* **15** (2008), Note 6, 6pp.

[22] V. Nikiforov, Graphs with many r-cliques have large complete r-partite subgraphs, *Bull. Lond. Math. Soc.* **40** (2008), 23–25.

[23] V. Nikiforov and C. C. Rousseau, Ramsey goodness and beyond, *Combinatorica* **29** (2009), 227–262.

[24] V. Rödl and M. Schacht, Complete partite subgraphs in dense hypergraphs, *Random Structures Algorithms* **41** (2012), 557–573.

[25] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, *Colloq. Math. Soc. János Bolyai*, vol. 18, North-Holland, Amsterdam-New York, 1978, 939–945.

[26] A. Sah, Diagonal Ramsey via effective quasirandomness, *Duke Math. J.* **172** (2023), 545–567.

[27] J. Spencer, Ramsey's theorem—a new lower bound, *J. Combin. Theory Ser. A* **18** (1975), 108–115.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Defect and transference versions of the Alon–Frankl–Lovász theorem

(Extended abstract)

Lior Gishboliner* Stefan Glock† Peleg Michaeli‡ Amedeo Sgueglia†

Abstract

Confirming a conjecture of Erdős on the chromatic number of Kneser hypergraphs, Alon, Frankl and Lovász proved that in any q-colouring of the edges of the complete r-uniform hypergraph, there exists a monochromatic matching of size $\lfloor \frac{n+q-1}{r+q-1} \rfloor$. Here, we obtain a transference version of this theorem. More precisely, for fixed q and r, we show that with high probability, a monochromatic matching of approximately the same size exists in any q-colouring of a random hypergraph, already when the average degree is a sufficiently large constant. In fact, our main new result is a defect version of the Alon–Frankl–Lovász theorem for almost complete hypergraphs. From this, the transference version is obtained via a variant of the weak hypergraph regularity lemma. The proof of the defect version uses tools from extremal set theory developed in the study of the Erdős matching conjecture.

1 Introduction

A flourishing trend in combinatorics has been showing that classical theorems concerning dense graphs (or hypergraphs) have corresponding analogues in the setting of (sparse)

*Department of Mathematics, University of Toronto, Canada. *Email*: lior.gishboliner@utoronto.ca.

†Fakultät für Informatik und Mathematik, Universität Passau, Germany. *Email*: stefan.glock@uni-passau.de, amedeo.sgueglia@uni-passau.de. SG is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — 542321564, AS is funded by the Alexander von Humboldt Foundation.

‡Mathematical Institute, University of Oxford, UK. *Email*: peleg.michaeli@maths.ox.ac.uk. Research supported by ERC Advanced Grant 883810.

random graphs. Such results are usually known as *transference* theorems and include, among many others, the breakthroughs of Rödl and Ruciński [15] on the transference of Ramsey's theorem, and of Conlon and Gowers [5] and Schacht [16] on the transference of Turán's theorem. Moreover, the study of combinatorial theorems for random graphs has generated several exciting recent developments in probabilistic and extremal combinatorics, including the sparse regularity method, hypergraph containers, the KŁR conjecture and the absorption method. We refer to the surveys of Conlon [4] and Böttcher [2] (and the references therein) for more details.

Here we are interested in a transference version of the celebrated result of Alon, Frankl and Lovász [1] concerning the chromatic number of Kneser hypergraphs. In 1955, Kneser [11] conjectured that if the r-subsets of a $(2r+q-1)$-element set are split into q classes, then one of the classes will contain two disjoint r-subsets. The conjecture remained open for 23 years, until Lovász [13] gave a topological proof using the Borsuk–Ulam theorem. His contribution is often considered to be the start of the field of topological combinatorics, and we refer to the book of Matoušek [14] for more examples. More generally, given $n,r,k \in \mathbb{N}$ with $n \geq kr$, the *Kneser hypergraph* $KG^k(n,r)$ is the k-uniform hypergraph (short: *k-graph*) where the vertices are all the r-subsets of $[n]$ and a collection of k vertices forms an edge if the corresponding r-sets are pairwise disjoint. Then, Kneser's conjecture is equivalent to the statement that $KG^2(n,r)$ is not q-colourable if $n \geq 2r+q-1$. As a generalization of this, Erdős [8] conjectured even further that $KG^k(n,r)$ is not q-colourable if $n \geq (q-1)(k-1)+kr$. The case $k=2$ corresponds to Kneser's original conjecture. Moreover, the validity of the case $r=2$ is a classical result by Cockayne and Lorimer [3]. The conjecture of Erdős was finally resolved by Alon, Frankl and Lovász [1], also using topological methods. We state their result in an equivalent form concerning the size of a monochromatic matching that we can guarantee in any colouring of the edges of a complete hypergraph.

Theorem 1.1 (Alon, Frankl and Lovász [1])**.** *Let $n,r,q \in \mathbb{N}$ with $r,q \geq 2$. Then any q-colouring of the edges of the complete n-vertex r-graph $K_n^{(r)}$ contains a monochromatic matching of size at least $\lfloor \frac{n+q-1}{r+q-1} \rfloor$.*

The bound $k := \lfloor \frac{n+q-1}{r+q-1} \rfloor$ on the size of the matching is best possible as shown by the following construction. Partition $[n]$ into q sets $V_1,\dots,V_q$ such that V_i has size at most k for each $1 \leq i \leq q-1$ and V_q has size at most $rk+r-1$. Given an edge $e \in K_n^{(r)}$, for $1 \leq i \leq q-1$, we assign colour i to e if and only if e intersects V_i, and we assign colour q to e if and only if e is completely contained in V_q. Note that an edge might be assigned several colours $1 \leq i \leq q-1$, and in this case we choose one of these colours arbitrarily. Then this yields an extremal q-colouring with respect to Theorem 1.1. Indeed, for $1 \leq i \leq q-1$, every edge with colour i has to intersect V_i and hence any matching of colour i has size at most $|V_i| \leq k$. Moreover, every edge with colour q is completely contained in V_q and hence the size of any matching of colour q is at most $\lfloor |V_q|/r \rfloor \leq k$.

Here, we obtain a transference version of Theorem 1.1. The random hypergraph model we consider is the binomial random r-graph $\mathbb{G}^{(r)}(n,p)$, which has n vertices, and where each r-set of vertices forms an edge independently with probability p. We show that if $p \gg n^{-r+1}$, then $\mathbb{G}^{(r)}(n,p)$ typically contains a monochromatic matching of asymptotically the same size as what is guaranteed to exist in $K_n^{(r)}$ by Theorem 1.1.

Theorem 1.2 (Transference version of the AFL Theorem)**.** *For all $r, q \in \mathbb{N}$ with $r, q \geq 2$, and all $\mu > 0$, there exists $C > 0$ such that, provided $p \geq Cn^{-r+1}$, w.h.p. the following holds for $G \sim \mathbb{G}^{(r)}(n,p)$: For any q-colouring of the edges of G, there exists a monochromatic matching of size at least $(1-\mu)\frac{n}{r+q-1}$.*

The size of the matching is asymptotically best possible since one cannot do better even in the complete r-graph, as explained above. In fact, it is also necessary that C is sufficiently large given μ, since G must contain at most $O(\mu n)$ isolated vertices. Indeed, by the optimality of Theorem 1.1, if G is an n-vertex r-graph with n' isolated vertices, then there is a q-colouring of its edges whose largest monochromatic matching has size at most $\lfloor \frac{(n-n')+q-1}{r+q-1} \rfloor$.

The graph case $r = 2$ (i.e. a transference version of the Cockayne–Lorimer theorem) was already proved by Gishboliner, Krivelevich and Michaeli [10], with earlier results of Letzter [12] and Dudek and Prałat [6] implying the cases $(r,q) = (2,2)$ and $(r,q) = (2,3)$, respectively.

Theorem 1.2 can be proved by combining the sparse hypergraph regularity method with the following "defect" version of the Alon–Frankl–Lovász theorem, which shows that, for large n, the conclusion of Theorem 1.1 approximately holds even for edge-colourings of *almost* complete r-uniform hypergraphs.

Theorem 1.3 (Defect version of the AFL Theorem)**.** *For all $r, q \in \mathbb{N}$ with $r, q \geq 2$, and all $\mu > 0$, there exists $\varepsilon > 0$ such that the following holds for all sufficiently large n. Let G be an n-vertex r-graph whose edges are q-coloured and assume $e(G) \geq (1-\varepsilon)\binom{n}{r}$. Then G contains a monochromatic matching of size at least $(1-\mu)\frac{n}{r+q-1}$.*

The deduction of Theorem 1.2 from Theorem 1.3 via the regularity method works in a similar way as in the graph case. However, finding a strategy to prove Theorem 1.3 presents several challenges. First, the topological proof of the Alon–Frankl–Lovász theorem does not seem to be robust enough to generalise to hypergraphs that miss a significant proportion of the edges. Moreover, the proof of the graph case ($r = 2$) from [10] relies on a good understanding of matchings in graphs in form of the Tutte–Berge formula, no analogue of which is available for hypergraphs. Finally, one could try to prove the existence of certain small coloured configurations ("gadgets") that can be repeatedly removed until the remaining hypergraph has a very specific structure. For instance, in the case of two colours, in [9] it is implicitly proved that either there exist two edges of different colour that share $r-1$ vertices, or the colouring is almost monochromatic. For general q, one possible gadget would be a set of $r+q-1$ vertices that contains an edge of each colour. However, there are constructions where all colour classes are large yet there is no such gadget, so this approach seems infeasible as well. We instead follow a new approach and make use of tools from extremal set theory developed in the study of the Erdős matching conjecture (see Section 2 for more details).

2 Proof overview

As already mentioned, the transference theorem follows from our defect theorem via a standard use of the regularity method. Therefore, we focus on the latter.

Let G be an n-vertex q-edge-coloured r-graph G with $e(G) \geq (1 - o(1))\binom{n}{r}$, and recall that we want to show that G contains a monochromatic matching of size $(1 - o(1))\frac{n}{r+q-1}$. Our first idea is to fix a large $k \in \mathbb{N}$ (independent of n), and consider the family $\mathcal{F}$ of the k-subsets F of $V(G)$ for which $G[F]$ is a complete r-graph. Observe that since G is almost-complete, almost all k-subsets of $V(G)$ belong to $\mathcal{F}$, i.e., $|\mathcal{F}| \geq (1 - o(1))\binom{n}{k}$. It may seem at this stage that we have not really made any progress as, thinking of $\mathcal{F}$ as a k-uniform hypergraph on $V(G)$, we still have that $\mathcal{F}$ is only almost-complete. However, we have gained the flexibility of choosing k, and our argument relies on choosing k to be large enough in terms of r, q.

For each $F \in \mathcal{F}$, we can apply the Alon–Frankl–Lovász theorem as a blackbox to $G[F]$ and get a monochromatic matching M_F of $G[F]$ of size roughly $\frac{k}{r+q-1}$. Assign to F the colour of the monochromatic matching M_F, giving a q-colouring of $\mathcal{F}$. Let blue be the most popular colour in this colouring of $\mathcal{F}$, and let $\mathcal{F}' := \{F \in \mathcal{F} : M_F \text{ is blue}\}$; so $|\mathcal{F}'| \geq (1 - o(1))\binom{n}{k}/q$. Our goal is now to find a certain structure in $\mathcal{F}'$ which would translate to a large blue matching in the original hypergraph G.

A naive approach would be to look for a large blue matching in $\mathcal{F}'$. More precisely, we need a matching of size $(1 - o(1))\frac{n}{k}$ in $\mathcal{F}'$ to obtain a matching of size $(1 - o(1))\frac{n}{r+q-1}$ in G. With this approach, the relevant question is what size of a matching is guaranteed to exist in an n-vertex k-graph with a given number of edges. This is a classical problem of Erdős [7], known as the Erdős matching conjecture, asking, given $n, k, t \in \mathbb{N}$, to determine the maximum number of edges of an n-vertex k-graph which does not contain a matching of size $t+1$. If $n < k(t+1)$, the problem is trivial as the family $\binom{[n]}{k}$ itself does not contain a matching of size $t+1$. When $n \geq k(t+1)$, Erdős conjectured that the answer is $\max\{\binom{n}{k} - \binom{n-t}{k}, \binom{k(t+1)-1}{k}\}$. The two bounds correspond to two natural extremal constructions: a star-like construction, obtained by taking all edges intersecting $[t]$ in at least one vertex, and a clique-like construction, obtained by taking all edges completely contained in $[k(t+1)-1]$. However, even if the conjecture was known to be true, it would not imply the desired bound on the size of the matching. Indeed, for $t = (1-\varepsilon)\frac{n}{k}$, the clique-like construction has density roughly $(1-\varepsilon)^k$, and, for large k, the star-like construction has density roughly $1 - e^{-(1-\varepsilon)}$. Hence, the density of $\mathcal{F}'$, which is about $1/q$, is not enough to guarantee the existence of an almost perfect matching.

Our second idea is that instead of a matching, we look in $\mathcal{F}'$ for a collection of edges which are allowed to overlap (only) mildly. More precisely, our goal is to find $F_1, \dots, F_s \in \mathcal{F}'$ with $s = (1 - o(1))\frac{n}{k}$ such that the set W of vertices appearing in more than one of the sets $F_1, \dots, F_s$ is small. Perhaps surprisingly, such $F_1, \dots, F_s$ always exist provided that k is large enough as a function of $\binom{n}{k}/|\mathcal{F}'|$. Then, removing all edges in $\bigcup_{i=1}^{s} M_{F_i}$ which intersect W, we obtain a blue matching in G of the desired size $(1 - o(1))\frac{n}{r+q-1}$.

We finally state the result which allows to find the desired sets $F_1, \dots, F_s \in \mathcal{F}'$. Roughly speaking, it says that in a hypergraph with large uniformity a constant density suffices to guarantee an “almost-perfect almost-cover”.

Theorem 2.1. *Let $1/n \ll 1/k \ll 1/C \ll \alpha, \beta$ and $\mathcal{F} \subseteq \binom{[n]}{k}$ with $|\mathcal{F}| \geq \beta\binom{n}{k}$. Then there exist $s := \lceil (1-\alpha)\frac{n}{k} \rceil$ sets $F_1, \ldots, F_s \in \mathcal{F}$ with $|F_1 \cup F_2 \cup \cdots \cup F_s| \geq (k-C)s$.*

The proof of Theorem 2.1 uses tools from extremal set theory, developed in the context of the Erdős matching conjecture

3 Concluding remarks

We obtained a transference and a defect version of the Alon–Frankl–Lovász theorem. It would be very interesting to characterize the extremal colourings for the AFL Theorem, that is, those q-colourings for which the bound in Theorem 1.1 is tight. In the graph case $r = 2$, the extremal colourings were characterized in [17] using the Gallai–Edmonds decomposition theorem. Closely related to this, it would be desirable to have a stability version, which should say that any q-colouring for which the largest monochromatic matching has size at most $(1+\mu)\frac{n}{r+q-1}$ must be ε-close to one of the extremal examples, that is, by recolouring at most εn^r edges, we obtain one of the extremal examples.

We point out that the colouring described after Theorem 1.1 is definitely not the only extremal example. For instance, let $x_1, \ldots, x_q$ be any positive integers such that $x_1 + \cdots + x_q = r+q-1$. (The choice $x_1 = \cdots = x_{q-1} = 1$ and $x_q = r$ corresponds to the construction in Section 1.) Then partition $[n]$ into sets $V_1, \ldots, V_q$ such that $|V_i| = x_i \cdot \frac{n}{r+q-1}$ (we are ignoring rounding issues here). For any edge $e \in K_n^{(r)}$, there must be $i \in [q]$ such that $|e \cap V_i| \geq x_i$. (If not, $|e| \leq \sum_{i=1}^{q} |e \cap V_i| \leq \sum_{i=1}^{q}(x_i - 1) = r + q - 1 - q < r$.) If there are multiple such i, just pick one arbitrarily. Then colour e with colour i. Observe that, for every $i \in [q]$, every edge with colour i intersects V_i in at least x_i vertices and thus any matching in colour i has size at most $|V_i|/x_i = \frac{n}{r+q-1}$.

References

[1] N. Alon, P. Frankl, and L. Lovász. The chromatic number of Kneser hypergraphs. *Trans. Amer. Math. Soc.*, 298(1):359–370, 1986.

[2] J. Böttcher. Large-scale structures in random graphs. In *Surveys in combinatorics 2017*, volume 440 of *London Math. Soc. Lecture Note Ser.*, pages 87–140. Cambridge Univ. Press, Cambridge, 2017.

[3] E. J. Cockayne and P. J. Lorimer. The Ramsey number for stripes. *J. Austral. Math. Soc.*, 19:252–256, 1975.

[4] D. Conlon. Combinatorial theorems relative to a random set. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pages 303–327. Kyung Moon Sa, Seoul, 2014.

[5] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. *Ann. of Math. (2)*, 184(2):367–454, 2016.

[6] A. Dudek and P. Prałat. On some multicolor Ramsey properties of random graphs. *SIAM J. Discrete Math.*, 31(3):2079–2092, 2017.

[7] P. Erdős. A problem on independent r-tuples. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 8:93–95, 1965.

[8] P. Erdős. Problems and results in combinatorial analysis. In *Colloquio Internazionale sulle Teorie Combinatorie*, pages 3–17, Rome, 1973, 1976. Accad. Naz. Lincei.

[9] L. Gishboliner, S. Glock, and A. Sgueglia. Tight Hamilton cycles with high discrepancy. Combin. Probab. Comput., to appear.

[10] L. Gishboliner, M. Krivelevich, and P. Michaeli. Color-biased Hamilton cycles in random graphs. *Random Structures Algorithms*, 60(3):289–307, 2022.

[11] M. Kneser. Aufgabe 360. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 58:27, 1955.

[12] S. Letzter. Path Ramsey number for random graphs. *Combin. Probab. Comput.*, 25(4):612–622, 2016.

[13] L. Lovász. Kneser's conjecture, chromatic number, and homotopy. *J. Combin. Theory Ser. A*, 25(3):319–324, 1978.

[14] J. Matoušek. *Using the Borsuk-Ulam theorem.* Universitext. Springer-Verlag, Berlin, 2003. Lectures on topological methods in combinatorics and geometry, Written in cooperation with Anders Björner and Günter M. Ziegler.

[15] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[16] M. Schacht. Extremal results for random discrete structures. *Ann. of Math. (2)*, 184(2):333–365, 2016.

[17] C. Xu, H. Yang, and S. Zhang. On characterizing the critical graphs for matching Ramsey numbers. *Discrete Appl. Math.*, 287:15–20, 2020.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

On the maximum diameter of d-dimensional simplicial complexes

(Extended abstract)

Stefan Glock* Olaf Parczyk† Silas Rathke† Tibor Szabó†

Abstract

We consider the problem of Santos about finding the largest possible diameter of a strongly connected d-dimensional simplicial complex on n vertices. For each d, we find the optimum value for every n that is large enough using the technique of absorption. The underlying theorem can be seen as a generalization of finding a tight Euler trail in a d-uniform hypergraph.

The Hirsch Conjecture from the 1950s postulated that the maximum diameter of a d-dimensional polytope with n facets is $n - d$. In 2012, Santos [9] disproved the Hirsch Conjecture by constructing a 43-dimensional polytope with 86 facets and diameter at least 44. Subsequently several further constructions were found, with increasing gap between their diameter and $n - d$, most recently also with the help of AI [11]. Most recently a new infinite sequence of counterexamples was found where the diameter is $(1 + c)(n - d)$ [5], $c > 0$. In this direction, the most important open question remains: is the diameter of every d-dimensional polytope with n facets bounded by a polynomial in n and d? If this did not turn out to be the case, then the simplex algorithm for solving linear programs would not be polynomial for any pivot rule.

While studying the Polynomial Hirsch Conjecture, Santos [10] suggested to consider the maximum possible diameter of simplicial d-complexes, a generalization of polytopes. A *(simplicial) complex* on n vertices is a family $\mathcal{C}$ of subsets of $[n]$ which is closed under taking subsets. The maximal elements of $\mathcal{C}$ are called *facets*. If every facet is of size $d+1$, $\mathcal{C}$ is called a *simplicial d-complex*. The *dual graph* $G(\mathcal{C})$ of a simplicial d-complex $\mathcal{C}$ has the set of facets as vertex set and two facets are connected by an edge if their intersection has size d. The *diameter* of $\mathcal{C}$ is the diameter of its dual graph $G(\mathcal{C})$. If $G(\mathcal{C})$ is connected, $\mathcal{C}$ is called *strongly*

*Fakultät für Informatik und Mathematik, Universität Passau, Innstraße 41, 94032 Passau, Germany.
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 542321564
E-mail: `stefan.glock@uni-passau.de`.

†Fachbereich Mathematik und Informatik, Freie Universität Berlin, Arnimallee 3, 14195 Berlin, Germany.
E-mails: `parczyk|szabo@mi.fu-berlin.de, s.rathke@fu-berlin.de`.

connected. Santos [10] defined $H_s(n,d)$ to be the maximum diameter of a strongly connected d-complex on $[n]$ and proved for fixed $d \geq 2$

$$\Omega\left(n^{\frac{2d+2}{3}}\right) \leq H_s(n,d) \leq \frac{1}{d}\binom{n}{d}. \tag{1}$$

The upper bound here is in fact quickly justified. On a shortest path between any two vertices in the dual graph, the first vertex (corresponding to a $(d+1)$-set) contains $d+1$ sets of size d, while each subsequent vertex contains d further d-sets that are not contained in any previous vertex, as this would create a shortcut. Hence, if ℓ is the number of vertices of the path, then there must exist at least $d \cdot \ell + 1$ sets of size d. Thus $\ell \leq \frac{1}{d}\left(\binom{n}{d} - 1\right)$ and for the largest diameter we obtain the slightly improved upper bound

$$H_s(n,d) \leq \left\lfloor \frac{1}{d}\binom{n}{d} - \frac{d+1}{d} \right\rfloor. \tag{2}$$

For the lower bound, Criado and Santos [4] gave an explicit algebraic construction of simplicial d-complexes using finite fields, whose diameter matched the order of magnitude of the upper bound for every fixed d and an infinite sequence of n. Criado and Newman [3] used a probabilistic construction with the Lovász Local Lemma to establish the existence of a construction for any fixed $d \geq 3$ and all n, that also significantly reduced the gap between the bounds from a factor exponential in d to $\mathcal{O}(d^2)$. Most recently Bohman and Newman [1] managed to pin down the precise asymptotics for every fixed $d \geq 2$ using the differential equations method to track the evolution of a random greedy algorithm:

$$\left(\frac{1}{d} - (\log n)^{-\varepsilon}\right)\binom{n}{d} \leq H_s(n,d)\,, \tag{3}$$

where $\varepsilon < 1/d^2$ and n is sufficiently large. We note that an earlier result of Dębski, Lonc, and Rzążewski [6] on harmonious colorings of fragmentable k-uniform hypergraphs also yield an asymptotically precise lower bound of $\left(\frac{1}{d} - o(1)\right)\binom{n}{d}$ on $H_s(n,d)$.

At the end of their paper, Bohman and Newman remarked that any improvement of their lower bound would be interesting. In [8], the last three authors gave explicit constructions to determine $H_s(n,2)$ for every n showing that the upper bound (2) can be achieved for $d = 2$ and all n except $n = 6$:

$$H_s(n,2) = \begin{cases} \left\lfloor \frac{1}{2}\binom{n}{2} - \frac{3}{2} \right\rfloor & n \neq 6 \\ 5 = \left\lfloor \frac{1}{2}\binom{6}{2} - \frac{3}{2} \right\rfloor - 1 & n = 6. \end{cases} \tag{4}$$

In [8] it is conjectured that the simple upper bound (2) can also be achieved for all other d as long as n is large enough. Here we prove this conjecture.

Theorem 1. *For every positive integer $d \geq 2$, there exists a positive integer n_0 such that for all $n > n_0$,*

$$H_s(n,d) = \left\lfloor \frac{1}{d}\binom{n}{d} - \frac{d+1}{d} \right\rfloor.$$

Constructing (the d-skeleton of) a simplicial d-complex on vertex set $[n]$ with diameter t is equivalent to finding a sequence of $t+1$ distinct subsets of $[n]$ of size $d+1$ such that two consecutive sets intersect in exactly d vertices and any other pair intersects in at most $d-1$

vertices. It is exactly then that the dual graph of the simplicial complex induced by these $(d+1)$-sets is a path of length t.

A special, *vertex-sequential* way of constructing such a sequence of $(d+1)$-sets is via a sequence $(x_1, \dots, x_{t+1+d})$ of (not necessarily distinct) elements from $[n]$, where one takes the $t+1$ sets $\{x_i, x_{i+1}, \dots, x_{i+d}\}$, $i = 1, \dots, t+1$, formed by $d+1$ consecutive entries of the sequence. If these sets are all distinct and every d-subset of $[n]$ occurs at most once among $d+1$ consecutive elements of the sequence, then the dual graph of the obtained simplicial complex is a path, as desired. We call simplicial complexes created this way *straight*. The simplicial complexes created in [3, 4] and the most recent in [1] are of this special form, while [10] is not. The extremal 2-complexes in [8] and the extremal d-complexes for Theorem 1 also contain significant straight portions, but are not entirely of that type either.

The key property of the above vertex-sequential construction, that any d-set should appear at most once among $(d+1)$ consecutive elements of the sequence, reminds one of the classic notion of universal cycle, where d-sets supposed to appear uniquely among d consecutive elements. A *universal cycle* for $[n]$ is a cyclic sequence with $\binom{n}{d}$ elements in which every d-set of $[n]$ appears exactly once consecutively. In 1992, it was conjectured by Chung, Diaconis, Graham [2] that there is a universal cycle for $[n]$ whenever $\binom{n-1}{d-1}$ is divisible by d and n is large enough. This was confirmed by Glock, Joos, Kühn, and Osthus [7] in 2020.

For our second theorem, we prove the analogue of this result for the cyclic variant of the straight construction. We say that a cyclic sequence with elements from $[n]$ is an *extra-tight universal cycle* for $[n]$ if every d-subset of $[n]$ appears exactly once among $d+1$ consecutive elements of the sequence. Note that every $d+1$ consecutive sequence elements form a $(d+1)$-set and these sets form the d-skeleton of a simplicial d-complex whose dual graph is a cycle. Observe that each element of the sequence appears in d^2 of these d-sets. Thus, an extra-tight universal cycle can only exist if $\binom{n-1}{d-1}$ is divisible by d^2. We show that this divisibility condition is also sufficient for large enough n.

Theorem 2. *For every positive integer $d \geq 2$, there exists a positive integer n_0 such that for all $n > n_0$, the following holds. If d^2 divides $\binom{n-1}{d-1}$, then there is an extra-tight universal cycle for $[n]$.*

Theorem 1 and Theorem 2 follow from a more general result which we state in the language of d-uniform hypergraphs. There, the previous definition of extra-tightness translate as follows:

Definition 3. Given a d-graph H, an *extra-tight trail* T in H is a sequence $(v_1, \dots, v_k)$ of vertices of H such that the sets $e_{i,j} \coloneqq \{v_i, \dots, v_{i+d}\} \setminus \{v_{i+j}\}$ with $i \in [k-d]$, $j \in [d]$ and the set $e_{k-d+1,d} \coloneqq \{v_{k-d+1}, \dots, v_d\}$ are all pairwise distinct and in $E(H)$. The *ends* of T are $(v_1, \dots, v_d)$ and $(v_{k-d+1}, \dots, v_k)$. An *extra-tight tour* C in H is a vertex sequence $(v_1, \dots, v_k)$ such that the edges $e_{i,j} \coloneqq \{v_i, \dots, v_{i+d}\} \setminus \{v_{i+j}\}$ with $i \in [k]$, $j \in [d]$ are pairwise distinct and all in H (where $v_{k+i} \coloneqq v_i$ for $i \in [d]$). We will often identify extra-tight trails resp. tours with the hypergraph containing all the edges of the extra-tight trail resp. tour. If there is an extra-tight trail resp. tour C in H with $E(C) = E(H)$, then we say that H has an *extra-tight Euler trail* resp. *tour*.

Note that a sequence $(x_1, \dots, x_k)$ with $x_i \in [n]$ defines a straight simplicial d-complex if and only if $(x_1, \dots, x_k)$ is an extra-tight trail in $K_n^{(d)}$ with vertex set $[n]$ and an extra-tight universal cycle for $[n]$ is an extra-tight Euler tour in $K_n^{(d)}$.

We establish Theorem 2 in a much more general form, concerning extra-tight Euler tours in d-uniform hypergraphs. It turns out that in addition to the necessary divisibility condition

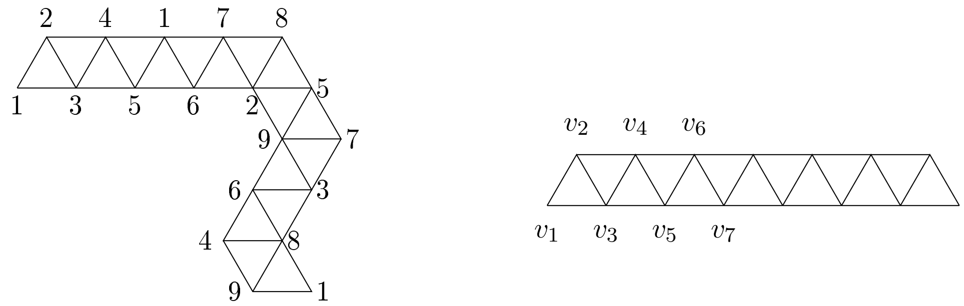


Figure 1: On the left: A strongly connected simplicial 2-complex on [9] with maximum diameter. The facets are given by the triangles. To ensure that the dual graph is a path, each pair in $\binom{[9]}{2}$ appears at most once as an edge. The diameter is maximum because only the pair $\{4, 7\}$ does not appear as an edge. On the right: A straight simplicial 2-complex.

on the vertex degrees, it is sufficient to assume that the minimum $(d-1)$-degree $\delta(G)$ is large, that is every $(d-1)$-set is contained in sufficiently many edges of the underlying hypergraph G.

Theorem 4. *For every integer $d \geq 2$ there are an $\alpha > 0$ and a positive integer n_0 such that the following holds for every integer $n \geq n_0$. If G is an n-vertex d-graph with $\delta(G) \geq (1-\alpha)n$ for which all vertex degrees are divisible by d^2, then it has an extra-tight Euler tour.*

For extra-tight Euler trails, the divisibility constraint on the 1-degrees is a bit more complicated for the vertices at the ends of the trail, but again we can show that these, together with a high minimum $(d-1)$-degree is sufficient:

Theorem 5. *For every integer $d \geq 2$ there are an $\alpha > 0$ and a positive integer n_0 such that the following holds for every integer $n \geq n_0$. Let G be an n-vertex d-graph with $\delta(G) \geq (1-\alpha)n$ and $\{v_1, \dots, v_d\}, \{v'_1, \dots, v'_d\}$ disjoint edges in G such that*

$$\deg_G(v) \equiv \begin{cases} i(d-1)+1 \pmod{d^2} & \exists\, i \in [d] : v \in \{v_i, v'_i\} \\ 0 \pmod{d^2} & \textit{else}. \end{cases} \tag{5}$$

Then there exists an extra-tight Euler trail with the ends $(v_1, \dots, v_d)$ and $(v'_d, \dots, v'_1)$.

Here Theorem 1 is not an immediate consequence, as the condition on the 1-degrees might not be satisfied. We solve this issue by using the fact that a simplicial d-complex is not required to be straight but can make "turns" (cf. Figure 1). We build some of these turns at the beginning of the simplicial complex such that the d-graph of the d-sets that are not covered by this initial segment satisfies the divisibility conditions of Theorem 5. Then we can apply Theorem 5 to find the rest of the simplicial d-complex.

For the proof of Theorem 5, on a very high level, we adapt the approach of Glock, Joos, Kühn, and Osthus [7] from the paper in which they prove the conjecture of Chung, Diaconis, Graham [2] that $K_n^{(d)}$ has a tight d-uniform Euler tour when $\binom{n-1}{d-1}$ is divisible by d and n is large enough. There, they proceed by first constructing a trail in which every ordered $(d-1)$-set of vertices appears at least once consecutively. Afterwards, they use the existence of designs to cover the remaining edges with tight cycles, which can then be included by rerouting the trail.

This approach of merging cycles into the trail does not translate to extra-tight cycles and trails. The problem is illustrated in Figure 2 for $d = 2$. There, the black and red edges form an extra-tight trail $(\dots, v_1, v_2, v_3, v_4, v_5, \dots)$ and an extra-tight cycle $(\dots, v_1', v_2', v_3, v_4', v_5', \dots)$ intersecting in one vertex v_3. If one attempted to merge these two into one extra-tight trail $(\dots, v_1, v_2, v_3, v_4', v_5', \dots, v_1', v_2', v_3, v_4, v_5, \dots)$, the red edges would no longer be used, and the blue edges would potentially be not available.

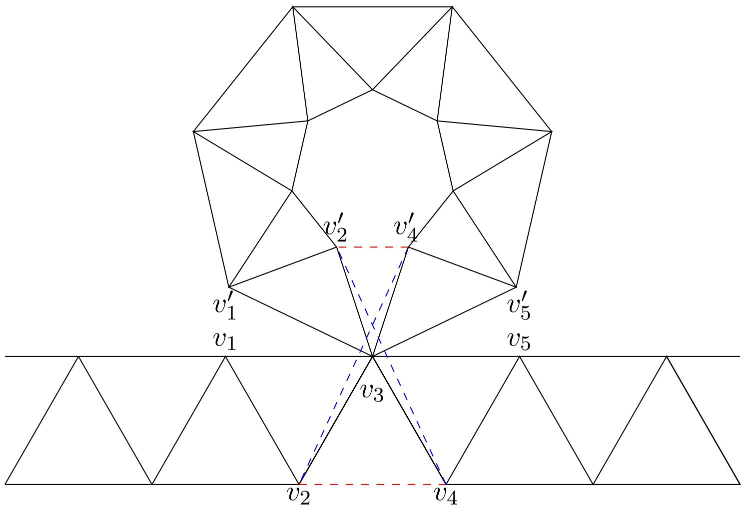


Figure 2: Merging an extra-tight tail and an extra-tight cycle into one extra-tight trail

The main novel ingredient in our approach is the construction of *switchers*. These are flexible structures that we can build into an initial extra-tight trail and allow us to "switch" between the red edges and the blue edges so that the high-level approach discussed earlier becomes feasible. Another challenge to tackle is that we cannot find switchers for all possible "gluing" locations, but have to reduce their number.

References

[1] Tom Bohman and Andrew Newman, *Complexes of nearly maximum diameter*, arXiv:2204.11932 (2022).

[2] Fan Chung, Persi Diaconis, and Ron Graham, *Universal cycles for combinatorial structures*, Discrete Mathematics **110** (1992), no. 1, 43–59.

[3] Francisco Criado and Andrew Newman, *Randomized construction of complexes with large diameter*, Discrete & Computational Geometry **66** (2021), 687–700.

[4] Francisco Criado and Francisco Santos, *The maximum diameter of pure simplicial complexes and pseudo-manifolds*, Discrete & Computational Geometry **58** (2017), 643–649.

[5] Alex Davies, Prateek Gupta, Sébastien Racanière, Grzegorz Swirszcz, Adam Zsolt Wagner, Theophane Weber, and Geordie Williamson, *Drums of high width*, Proceedings of the London Mathematical Society **130** (2025), no. 3, e70031.

[6] Michał Dębski, Zbigniew Lonc, and Paweł Rzążewski, *Harmonious and achromatic colorings of fragmentable hypergraphs*, European Journal of Combinatorics **66** (2017), 60–80.

[7] Stefan Glock, Felix Joos, Daniela Kühn, and Deryk Osthus, *Euler tours in hypergraphs*, Combinatorica **40** (2020), no. 5, 679–690.

[8] Olaf Parczyk, Silas Rathke, and Tibor Szabó, *The maximum diameter of* 2*-dimensional simplicial complexes*, submitted, 2025.

[9] Francisco Santos, *A counterexample to the Hirsch conjecture*, Ann. of Math. **176** (2012), no. 1, 383–412.

[10] ———, *Recent progress on the combinatorial diameter of polytopes and simplicial complexes*, Top **21** (2013), no. 3, 426–460.

[11] Grzegorz Swirszcz, Adam Zsolt Wagner, Geordie Williamson, Sam Blackwell, Bogdan Georgiev, Alex Davies, Ali Eslami, Sebastien Racaniere, Theophane Weber, and Pushmeet Kohli, *Advancing geometry with ai: Multi-agent generation of polytopes*, arXiv preprint arXiv:2502.05199 (2025).

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

Strong odd coloring in minor-closed classes

(Extended abstract)

Miriam Goetze* Fabian Klute† Kolja Knauer‡ Irene Parada§
Juan Pablo Peña¶ Torsten Ueckerdt‖

Abstract

We show that the strong odd chromatic number on any proper minor-closed graph class is bounded by a constant. We almost determine the smallest such constant for outerplanar graphs.

1 Strong Odd Colorings of Graphs

Recently, Kwon and Park [14] introduced the notion of *strong odd colorings* of graphs. Here, a proper vertex-coloring $\varphi\colon V(G) \to [t]$ of a graph G is *strong odd* if for every vertex $v \in V(G)$ and every color $i \in [t]$ the quantity $|N(v) \cap \varphi^{-1}(i)|$ is either zero or odd. In other words, among the neighbors of any vertex every color appears an odd number of times or not at all. The *strong odd chromatic number* $\chi_{\mathrm{so}}(G)$ of G is the minimum t such that G admits a strong odd coloring with t colors. See Fig. 1 for some examples.

*Karlsruhe Institute of Technology, Germany. E-mail: `miriam.goetze@kit.edu`. Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 520723789.

†Universitat Politècnica de Catalunya, Spain. E-Mail: `fabian.klute@upc.edu`. Supported by María Zambrano grant 2022UPC-MZC-94041 funded by the Spanish Ministry of Universities and the European Union (NextGenerationEU) and by grant PID2023-150725NB-I00 funded by MICIU/AEI/10.13039/501100011033.

‡Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Spain. E-mail: `kolja.knauer@ub.edu`. Partially supported by the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M) and the *Ministerio de Economía, Industria y Competitividad* through grant PID2022-137283NB-C22.

§Universitat Politècnica de Catalunya, Spain. E-Mail: `irene.parada@upc.edu`. I. P. is a Serra Húnter Fellow, partially supported by grant PID2023-150725NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by the UPC grant ALECTORS.

¶DIM-CMM, Universidad de Chile, Chile E-Mail: `juan.pena@dim.uchile.cl`. Support by Basal program FB210005 and Doctoral Fellowship grant 21211955, ANID, CHILE.

‖Karlsruhe Institute of Technology, Germany. E-mail: `torsten.ueckerdt@kit.edu`. Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 520723789.

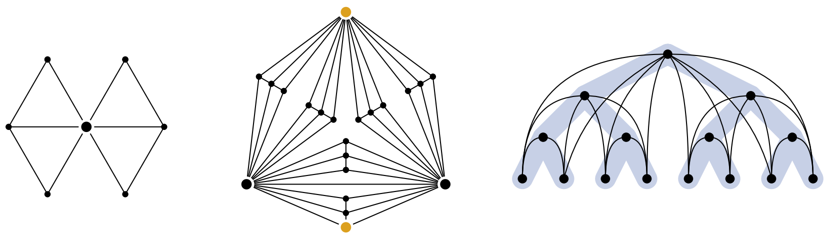


Figure 1: An outerplanar graph G' with $\chi_{\text{so}}(G') = 7$ (left), a planar graph G'' with $\chi_{\text{so}}(G'') = 20$ (middle) and the graph G_3 from Theorem 1 with $\chi_{\text{so}}(G_3) > 2^3$ (right). The black vertices in G' and G'' have pairwise distinct colors in every strong odd coloring.

Related Colorings. Strong odd colorings are closely related to a number of similar concepts. A proper vertex-coloring of G is *odd* if in the neighborhood of every non-isolated vertex some color appears an odd number of times. The corresponding parameter $\chi_{\text{o}}(G)$ is called the *odd chromatic number*. A proper vertex-coloring of G is *conflict-free* if in the neighborhood of every non-isolated vertex some color appears exactly once. The corresponding parameter is denoted by $\chi_{\text{pcf}}(G)$. Both, χ_{o} and χ_{pcf} have received plenty of attention in recent years, see e.g. [1, 3, 4, 6, 7, 10–13, 15–17].

If we require that among the neighbors of every vertex, every color appears exactly once or not at all, this corresponds to coloring the *square* G^2 of G. Here, G^2 has the same vertex set as G with two vertices being adjacent in G^2 if their distance in G is either 1 or 2. The chromatic number of G^2 is studied in particular for planar graphs with respect to Wegner's Conjecture [18], see also [5]. From the above definitions, we immediately get the following inequalities.

$$\chi(G) \;\leq\; \chi_{\text{o}}(G) \begin{array}{c} \leq \\ \leq \end{array} \begin{array}{c} \chi_{\text{so}}(G) \\ \\ \chi_{\text{pcf}}(G) \end{array} \begin{array}{c} \leq \\ \leq \end{array} \chi(G^2) \;\leq\; \Delta(G)^2 + 1 \tag{1}$$

Here, $\Delta(G)$ denotes the maximum degree of G. The last inequality can also be reversed in some sense, since trivially $\chi(G^2) > \Delta(G)$ for all graphs G. However, all remaining inequalities in (1) can be arbitrarily far apart. Petruševski and Škrekovski [16] show that there is no function f such that $\chi_{\text{o}}(G) \leq f(\chi(G))$ for all G. It is also known [12] that there is no function f such that $\chi_{\text{pcf}}(G) \leq f(\chi_{\text{o}}(G))$ for all G. Further, since $\Delta(G) < \chi(G^2)$, any graph class with unbounded degree but bounded χ_{pcf} shows that there is no function f such that $\chi(G^2) \leq f(\chi_{\text{pcf}}(G))$ for all G, e.g. bipartite permutation graphs, see [12]. Similarly, Kwon and Park [14] note that there is no function f such that $\chi(G^2) \leq f(\chi_{\text{so}}(G))$ for all G, while they show graphs with $\chi_{\text{so}}(G) \in \Omega(\chi_{\text{o}}(G)^2)$. Here, we improve the latter.

Observation 1. *There is no function f such that $\chi_{\text{so}}(G) \leq f(\chi_{\text{o}}(G))$ for all graphs G.*

Proof. For every integer $k \geq 1$ let G_k be the bipartite graph whose vertex set consists of the vertices of the full rooted binary tree T_k of height k, i.e., the distance of the root to every leaf is k. For every leaf v of T_k, in G_k we put all edges between v and its ancestors, see the right of Fig. 1 for an example. It is known that $\chi_{\text{o}}(G_k) \leq 4$ for all k [12].

We show by induction on k, that in order to color the leaves such that all colors appearing in the neighborhood of any internal vertex appear an odd number of times, at least 2^k colors are needed. This clearly holds for $k = 1$. For $k \geq 2$, take such a coloring φ of the leaves of G_k. Note that G_k consists of two disjoint copies G', G'' of G_{k-1} plus the root vertex r adjacent to all leaves in G' and G''. Further, restricting φ to the leaves of G' and G'' gives feasible colorings of G' and G''. Hence, each color appearing on the leaves of G' appears an odd number of times on the leaves of G' and similarly for G''. Hence, no color can be used on the leaves of G' and on the leaves of G''. Thus, we need twice as many colors as for G_{k-1}, i.e., by induction φ uses at least $2 \cdot 2^{k-1} = 2^k$ colors on the leaves.

Finally, as the root of T_k is connected to all leaves, we get that $\chi_{\mathrm{so}}(G_k) \geq 2^k + 1$. □

Planar and Outerplanar Graphs. Kwon and Park [14] asked whether the strong odd chromatic number is bounded on the class $\mathcal{P}$ of all planar graphs. This was answered affirmatively by Caro, Petruševski, Škrekovski, and Tuza [2] who show that for

$$c_{\mathcal{P}} = \max\{\chi_{\mathrm{so}}(G) \mid G \text{ planar}\} \qquad \text{we have} \qquad 12 \leq c_{\mathcal{P}} \leq 388.$$

We increase the lower bound to 20, which gives a negative answer to [2, Problem 3.6.1].

Observation 2. *For the planar graph G in the middle of Fig. 1 we have $\chi_{\mathrm{so}}(G) = 20$.*

For the class $\mathcal{O}$ of all outerplanar graphs, the authors [2] show that for

$$c_{\mathcal{O}} = \max\{\chi_{\mathrm{so}}(G) \mid G \text{ outerplanar}\} \qquad \text{we have} \qquad 7 \leq c_{\mathcal{O}} \leq 30.$$

The lower bound of 7 is given by the outerplanar graph on the left of Fig. 1, and suspected to be the correct value [2, Problem 3.6.2]. We reduce the upper bound to 8.

Proposition 3. *For every outerplanar graph G we have $\chi_{\mathrm{so}}(G) \leq 8$.*

Proper Minor-Closed Classes. The class $\mathcal{P}$ of all planar graphs and the class $\mathcal{O}$ of all outerplanar graphs are closed under taking minors. Our main result is that the strong odd chromatic number is bounded on every proper[1] minor-closed graph class. Hence, this is a far-reaching generalization of the results above, concerning only the classes $\mathcal{P}$ and $\mathcal{O}$.

Theorem 4. *For every proper minor-closed graph class $\mathcal{G}$, there exists a constant $c_{\mathcal{G}}$ such that for every graph $G \in \mathcal{G}$ we have $\chi_{\mathrm{so}}(G) \leq c_{\mathcal{G}}$.*

Our proof of Theorem 4 goes in three steps.

(1) Bounded Treewidth. First, we look at graphs of bounded treewidth and prove that there is a function f_1 such that $\chi_{\mathrm{so}}(G) \leq f_1(\mathrm{tw}(G))$ for all G. Given any graph G with $\mathrm{tw}(G) = k$, we consider any k-tree[2] H with $G \subseteq H$ and use a BFS-layering $L_0, L_1, \ldots$ of H, where $L_d = \{v \in V(H) \mid \mathrm{dist}_H(v, r) = d\}$ is the set of vertices at distance d to a fixed root vertex r. Crucially, for each $d \geq 0$ we get induced subgraphs $G[L_d] \subseteq H[L_d]$ of treewidth at most $k-1$, which allows us to do induction on k.

The hardest part is to account for edges between layers, i.e., to ensure that for every vertex $v \in L_d$ and every color i used on L_{d+1} the quantity $|N(v) \cap \varphi^{-1}(i)|$ is zero or odd. We do this with a stronger claim involving some additional subsets $M_1, \ldots, M_\ell \subseteq L_{d+1}$

[1] different from the class of all graphs
[2] edge-maximal graph of treewidth k

that must intersect every color class oddly or not at all. Then, reusing colors on every third layer, we construct a proper $f_1(k)$-coloring φ of H that is strong odd for G. Our function $f_1(k)$ is exponential in k, which is however unavoidable, as for example, G_k from Theorem 1 has $\operatorname{tw}(G_k) \leq k$ and $\chi_{\mathrm{so}}(G_k) > 2^k$.

We remark that outerplanar graphs have treewidth at most 2. And indeed our proof of Theorem 3 (outerplanar graphs have strong odd 8-colorings) also works along a BFS-layering of a 2-tree H with $G \subseteq H$, but extra care is needed to only use 8 colors.

(2) Bounded Row-Treewidth. The next step are graphs of bounded row-treewidth, that is graphs G with $G \subseteq H \boxtimes P$ where H has small treewidth and P is a path. The *strong product* $G' \boxtimes G''$ of two graphs is the graph with vertex set $V(G') \times V(G'')$ having an edge between two distinct vertices (u,p) and (v,q) if and only if ($uv \in E(G')$ or $u = v$) and ($pq \in E(G'')$ or $p = q$).

The *row-treewidth* $\operatorname{rtw}(G) = \min\{k \mid G \subseteq H \boxtimes P, \operatorname{tw}(H) = k\}$ has gained a lot of prominence around the so-called Graph Product Stucture [9]. Using the layer-structure of $H \boxtimes P$ together with the bound $\chi_{\mathrm{so}}(H) \leq f_1(\operatorname{tw}(H))$ above, we prove that $\chi_{\mathrm{so}}(G) \leq f_2(\operatorname{rtw}(G))$ for a universal (again exponential) function f_2 and all G.

(3) Clique Sums. Third, we extend our results to subgraphs of so-called (w,k,t)-sums. A *$\leq w$-clique-sum* is obtained from two graphs G', G'' by identifying a clique C' of size at most w in G' with a clique C'' of the same size in G''. For integers $w,k,t \geq 1$, a *(w,k,t)-sum* is a graph G that can be obtained by taking $\leq w$-clique-sums of graphs $G_1, \ldots, G_n$ where each G_i is obtained from $H_i \boxtimes P$ for some k-tree H_i after adding t universal vertices. A result of Dujmović, Joret, Micek, Morin, Ueckerdt, and Wood [8, Theorem 41] states that for every proper minor-closed graph class $\mathcal{G}$ there are constants k and t, such that every $G \in \mathcal{G}$ is a subgraph of a $(2(k+1)+t,k,t)$-sum.

The iterative clique-sums give rise to yet another layering structure (similar to BFS-layerings). Using this, we give a universal function f_3 such that $\chi_{\mathrm{so}}(G) \leq f_3(w,k,t)$ whenever G is a subgraph of a (w,k,t)-sum, which concludes the proof of Theorem 4.

Bounded Row-Treewidth. Let us discuss step **(2)** in more detail. The vertices of P correspond to *layers* of $H \boxtimes P$, which are vertex-disjoint copies of a k-tree H. Edges run only inside a layer or between consecutive layers. The plan is to use for each layer L a strong odd coloring φ of $G \cap L$ with $f_1(k)$ colors, and reuse the same color-set on every third layer. However, we must also ensure that for every vertex v on a layer L and every color i on a neighboring layer L' the quantity $|N(v) \cap \varphi^{-1}(i)|$ is either zero or odd. To this end, we actually consider already in step **(1)** a more general setup with directed graphs.

A proper t-coloring $\varphi\colon V(\vec{G}) \to [t]$ of a directed graph $\vec{G}$ is *strong odd* if for every $v \in V(\vec{G})$ with out-neighborhood $N^+(v)$ and every $i \in [t]$ we have that $|N^+(v) \cap \varphi^{-1}(i)|$ is zero or odd. That is, each color appears among the out-neighbors of each vertex either not at all, or an odd number of times. (In case for every $\vec{uv} \in E(\vec{G})$ we also have $\vec{vu} \in E(\vec{G})$, this coincides with the strong odd coloring of the underlying undirected graph G of $\vec{G}$.)

Proposition 5 (Bounded Treewidth with Directed Graphs)**.**
For all $k, \ell \in \mathbb{N}$, there exists a constant $f_1(k,\ell)$ such that for every undirected k-tree H and every collection $\vec{G}_1, \ldots, \vec{G}_\ell$ of directed subgraphs of H, there is a proper coloring $\varphi\colon V(H) \to [f_1(k,\ell)]$ of H that restricts to a strong odd coloring of $\vec{G}_i$ for each $i \in [\ell]$.

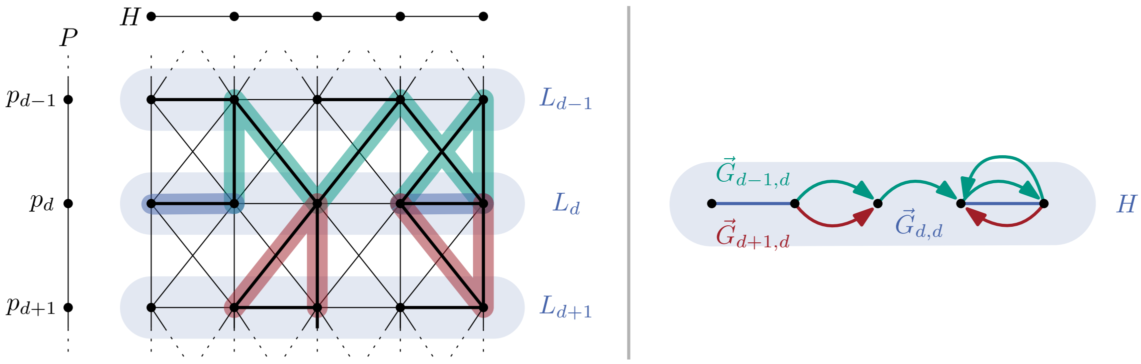


Figure 2: Left: A subgraph G (thick edges) on three layers L_{d-1}, L_d, L_{d+1} of $H \boxtimes P$. Right: The corresponding three directed subgraphs $\vec{G}_{d,d}$ (blue), $\vec{G}_{d-1,d}$ (green), $\vec{G}_{d+1,d}$ (red) of H.

Having Theorem 5 at hand, now assume that G is an (undirected) subgraph of $H \boxtimes P$ for some k-tree H. Let $p_1, p_2, \dots$ be the vertices of P and $L_1, L_2, \dots$ be the corresponding layers. For every layer L_d, we define three directed subgraphs $\vec{G}_{d,d}, \vec{G}_{d-1,d}$ and $\vec{G}_{d+1,d}$ of H, see Fig. 2 for an example. These three graphs model the edges of G within L_d, between L_{d-1} and L_d, and between L_{d+1} and L_d, respectively. Formally, for every (undirected) edge $\{(u, p_i), (v, p_j)\}$ in $G \subseteq H \boxtimes P$ with $u \neq v$, we put a directed edge $\vec{uv}$ in $\vec{G}_{i,j} \subseteq H$.

For each layer L_d, let $\varphi_d \colon V(H) \to [f_1(k, 3)]$ be the coloring of the k-tree H given by Theorem 5 with respect to three directed subgraphs $\vec{G}_{d,d}, \vec{G}_{d-1,d}, \vec{G}_{d+1,d}$. Finally, we define the coloring φ of $H \boxtimes P$ by setting $\varphi(v, p_i) = (\varphi_i(v), i \bmod 3)$. Observing that φ is a strong odd coloring of G with $3 \cdot f_1(k, 3)$ colors, we can conclude the desired result.

Theorem 6 (Bounded Row-Treewidth)**.**
There exists a function f_2 such that for every graph G we have $\chi_{\mathrm{so}}(G) \leq f_2(\mathrm{rtw}(G))$.

Acknowledgments

Part of this work was conducted at the 12th Annual Workshop on Geometry and Graphs held at the Bellairs Research Institute in February 2025. We are grateful to the organizers and participants for providing an excellent research environment.

References

[1] Yair Caro, Mirko Petruševski, and Riste Škrekovski. Remarks on proper conflict-free colorings of graphs. *Discrete Mathematics*, 346(2):14, 2023. `doi:10.1016/j.disc.2022.113221`.

[2] Yair Caro, Mirko Petruševski, Riste Škrekovski, and Zsolt Tuza. On strong odd colorings of graphs. arXiv preprint, 2024. `doi:10.48550/arXiv.2410.02336`.

[3] Eun-Kyung Cho, Ilkyoo Choi, Hyemin Kwon, and Boram Park. Brooks-type theorems for relaxations of square colorings. *Discrete Mathematics*, 348(1):12, 2025. `doi:10.1016/j.disc.2024.114233`.

[4] Eun-Kyung Cho, Ilkyoo Choi, Hyemin Kwon, and Boram Park. Proper conflict-free coloring of sparse graphs. *Discrete Applied Mathematics*, 362:34–42, 2025. doi:10.1016/j.dam.2024.11.016.

[5] Daniel W. Cranston. Coloring, list coloring, and painting squares of graphs (and other related problems). *Electronic Journal of Combinatorics*, 30(2), 2023. doi:10.37236/10898.

[6] Daniel W. Cranston and Chun-Hung Liu. Proper conflict-free coloring of graphs with large maximum degree. *SIAM Journal on Discrete Mathematics*, 38(4):3004–3027, 2024. doi:10.1137/23M1563281.

[7] Tianjiao Dai, Qiancheng Ouyang, and François Pirot. New bounds for odd colourings of graphs. *Electronic Journal of Combinatorics*, 31(4), 2024. doi:10.37236/12110.

[8] Vida Dujmović, Gwenaël Joret, Piotr Micek, Pat Morin, Torsten Ueckerdt, and David R. Wood. Planar graphs have bounded queue-number. *Journal of the ACM*, 67(4):38, 2020. doi:10.1145/3385731.

[9] Zdeněk Dvořák, Tony Huynh, Gwenael Joret, Chun-Hung Liu, and David R. Wood. Notes on graph product structure theory. In *2019–20 MATRIX annals*, pages 513–533. Cham: Springer, 2021. doi:10.1007/978-3-030-62497-2_32.

[10] Igor Fabrici, Borut Lužar, Simona Rindošová, and Roman Soták. Proper conflict-free and unique-maximum colorings of planar graphs with respect to neighborhoods. *Discrete Applied Mathematics*, 324:80–92, 2023. doi:10.1016/j.dam.2022.09.011.

[11] Robert Hickingbotham. Odd colourings, conflict-free colourings and strong colouring numbers. *Australasian Journal of Combinatorics*, 87:160–164, 2023. URL: ajc.maths.uq.edu.au/pdf/87/ajc_v87_p160.pdf.

[12] Andrea Jiménez, Kolja Knauer, Carla Negri Lintzmayer, Martín Matamala, Juan Pablo Peña, Daniel A. Quiroz, Maycon Sambinelli, Yoshiko Wakabayashi, Weiqiang Yu, and José Zamora. Boundedness for proper conflict-free and odd colorings. arXiv preprint, 2023. doi:10.48550/arXiv.2308.00170.

[13] Mateusz Kamyczura and Jakub Przybyło. On conflict-free proper colourings of graphs without small degree vertices. *Discrete Mathematics*, 347(1):6, 2024. doi:10.1016/j.disc.2023.113712.

[14] Hyemin Kwon and Boram Park. Strong odd coloring of sparse graphs. arXiv preprint, 2024. doi:10.48550/arXiv.2401.11653.

[15] Chun-Hung Liu. Proper conflict-free list-coloring, odd minors, subdivisions, and layered treewidth. *Discrete Mathematics*, 347(1):16, 2024. doi:10.1016/j.disc.2023.113668.

[16] Mirko Petruševski and Riste Škrekovski. Colorings with neighborhood parity condition. *Discrete Applied Mathematics*, 321:385–391, 2022. doi:10.1016/j.dam.2022.07.018.

[17] Tao Wang and Xiaojing Yang. On odd colorings of sparse graphs. *Discrete Applied Mathematics*, 345:156–169, 2024. doi:10.1016/j.dam.2023.11.039.

[18] Gerd Wegner. Graphs with given diameter and a coloring problem. Technical report, University of Dortmund, 1977. `doi:10.17877/DE290R-16496`.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Wild generalised truncation of infinite matroids

(Extended abstract)

J. Pascal Gollin* Attila Joó†

Abstract

For $n \in \mathbb{N}$, the n-truncation of a matroid M of rank at least n is the matroid whose bases are the n-element independent sets of M. One can extend this definition to negative integers by letting the $(-n)$-truncation be the matroid whose bases are all the sets that can be obtained by deleting n elements of a base of M. If M has infinite rank, then for distinct $m, n \in \mathbb{Z}$ the m-truncation and the n-truncation are distinct matroids.

Inspired by the work of Bowler and Geschke on infinite uniform matroids, we provide a natural definition of generalised truncations that encompasses the notions mentioned above. We call a generalised truncation wild if it is not an n-truncation for any $n \in \mathbb{Z}$ and we prove that, under Martin's Axiom, any finitary matroid of infinite rank and size of less than continuum admits $2^{2^{\aleph_0}}$ wild generalised truncations.

1 Introduction

Searching for a concept of infinite matroids with duality was initiated by Rado [10]. Rado's search inspired Higgs, Oxley and others to investigate possible definitions [7, 9]. The theory of infinite matroids gained a new momentum when Bruhn et al. [4] rediscovered independently and axiomatised the same infinite matroid concept that was found by Higgs in the late 1960s. Going beyond the work of Higgs, they gave five sets of cryptomorphic axiomatisation. Their axiomatisation in terms of bases reads as follows.

A set $\mathcal{B} \subseteq \mathcal{P}(E)$ is the *set of basis of a matroid* M on a given ground set E if

(B1) $\mathcal{B} \neq \emptyset$;

(B2) For all $B_0, B_1 \in \mathcal{B}$ and all $x \in B_0 \setminus B_1$ there exists an element $y \in B_1 \setminus B_0$ such that $(B_0 \setminus \{x\}) \cup \{y\} \in \mathcal{B}$;

*FAMNIT, University of Primorska, 6000 Koper, Slovenia. E-mail: pascal.gollin@famnit.upr.si. Supported by the Slovenian Research and Innovation Agency (research project N1-0370).

†Department of Mathematics, University of Hamburg, Hamburg, Germany. E-mail: attila.joo@uni-hamburg.de. Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Grant No. 513023562.

(BM) For every $X \subseteq E$, the set of maximal elements of the poset $(\{X \cap B \colon\ B \in \mathcal{B}\}, \subseteq)$ form a cofinal subset.

The *truncation* of a matroid M of non-zero rank is the matroid on the same ground set E whose bases are all sets that can be obtained by the deletion of one element of a base of M (see [2, Definition 3.1]). If the rank of M is at least $n \in \mathbb{N} \setminus \{0\}$, then truncation can be iterated n times starting with M. Let us call the resulting matroid the $(-n)$-*truncation* of M. Note that if the rank of M is infinite, then so is the rank of its $(-n)$-truncation. For $n \in \mathbb{N}$, the n-*truncation* of a matroid M of rank at least n is the matroid on the same ground set whose bases are the n-element independent sets of M. Clearly, if M has infinite rank and $m, n \in \mathbb{Z} \setminus \{0\}$ with $m \neq n$, then the m-truncation and n-truncation of M are different matroids.

Our aim of this paper is to find a natural common generalisation of these concepts that allows for more flexibility 'in between' these concepts in a similar sense as Bowler and Geschke [3] generalised the concept of uniform matroids, as we will discuss further below. Let $E(M)$, $\mathcal{I}(M)$, and $\mathcal{B}(M)$ stand for the ground set, independent sets, and the bases of matroid M respectively. We propose the following definition.

Definition 1.1. A matroid N is a *generalised truncation* of matroid M if

(I) $E(N) = E(M)$,

(II) $\mathcal{I}(N) \subseteq \mathcal{I}(M)$,

(III) for all $I \in \mathcal{I}(N) \setminus \mathcal{B}(N)$ and all $e \in E \setminus I$, if $I \cup \{e\} \in \mathcal{I}(M)$, then $I \cup \{e\} \in \mathcal{I}(N)$.

Note that every matroid is a generalised truncation of itself which we call the *trivial* generalised truncation. One can ask if there are non-trivial generalised truncations other than the n-truncation for $n \in \mathbb{Z}$. The generalised truncations of free matroids are exactly the uniform matroids (in the sense of the definition of Bowler and Geschke [3, Definition 2]). A uniform matroid U is *wild* if neither U nor its dual has finite rank. Using our terminology, a uniform matroid is wild if it is neither a free matroid nor an n-truncation of a free matroid for suitable $n \in \mathbb{Z}$. Wild uniform matroids were constructed by Bowler and Geschke in [3] under Martin's Axiom. It is unknown if their existence can be proved in ZFC alone. We will call a non-trivial generalised truncation *wild* if it is not an n-truncation for any $n \in \mathbb{Z}$. The main result of this paper reads as follows.

Theorem 1.2. *Under Martin's Axiom, every finitary matroid M of infinite rank on a ground set E with $|E| < 2^{\aleph_0}$ admits a wild generalised truncation.*

This article is an extended abstract of [5] in which many proofs have been omitted.

2 Preliminaries

Infinite matroids. For $X \subseteq E(M)$, let $M \restriction X$ be the matroid on X where $\mathcal{B}(M \restriction X)$ consists of the maximal elements of $\{X \cap B \colon\ B \in \mathcal{B}(M)\}$. It is known that $M \restriction X$ is indeed a matroid, and it is called the *restriction* of M to X. Similarly, $M.X$, is the matroid on X where $\mathcal{B}(M.X)$ consists of the minimal elements of $\{X \cap B \colon\ B \in \mathcal{B}(M)\}$, and it is called the *contraction* of M to X. We write $M \setminus X$ and M/X for $M \restriction (E \setminus X)$ and $M.(E \setminus X)$ respectively. Their respective names are the *deletion* and the *contraction* of X in M.

The *rank* $r(M)$ of matroid M is $n \in \mathbb{N}$ if it has a base of size n and ∞ if its bases are infinite. For $X, Y \subseteq E(M)$, we write $r_M(X)$ for $r(M \restriction X)$ and call it the *rank* of X. Moreover,

$r_M(X|Y)$ stands for $r_{M/Y}(X \setminus Y)$ and it is called the *relative rank* of X with respect to Y. A set $X \subseteq E$ *spans* $e \in E$ if $r_M(\{e\}|X) = 0$. A matroid M is *finitary* if whenever all finite subsets of a set $X \subseteq E$ are independent in M, then so is X. For more about infinite matroids we refer to [1].

Martin's Axiom. Let $(P, \leq)$ be a partial order. A set $D \subseteq P$ is *dense* if every $p \in P$ has a lower bound in D, that is there exists a $d \in D$ with $d \leq p$. A set $A \subseteq P$ is a *strong antichain* if no two distinct elements of A have a common lower bound in P. We say that P satisfies the *countable chain condition* (or, ccc, for short) if every strong antichain in P is countable. A non-empty set $F \subseteq P$ is a *filter* if

- F is downward directed, that is any two distinct elements of F have a common lower bound in F, and
- F is upwards closed, that is for every $f \in F$ and every $p \in P$, if $f \leq p$, then $p \in F$.

Let $\mathfrak{c}$ denote $2^{\aleph_0}$, the size of the continuum.

Martin's Axiom. *For every partial order $(P, \leq)$ that satisfies the* ccc*, every set I with $|I| < \mathfrak{c}$, and every family $\langle D_i : i \in I \rangle$ of dense subsets of P there exists a filter F on P such that $F \cap D_i$ is non-empty for every $i \in I$.*

Theorem 2.1 ([8])**.** *Under Martin's Axiom, $2^\kappa = \mathfrak{c}$ for every cardinal κ with $\aleph_0 \leq \kappa < \mathfrak{c}$.*

3 Preparations

We give a characterisation for a set $\mathcal{F}$ to be the set of bases of a generalised truncation.

Lemma 3.1. *A set $\mathcal{F}$ is the set of the bases of a generalised truncation of a matroid M if and only if it satisfies the following conditions:*

(1) $\emptyset \neq \mathcal{F} \subseteq \mathcal{I}(M)$;

(2) If $B \in \mathcal{F}$ and $B' \in \mathcal{I}(M)$ with $|B \setminus B'| = |B' \setminus B| < \aleph_0$, then $B' \in \mathcal{F}$;

(3) If $B, B' \in \mathcal{F}$, then no proper subset of B spans B' in M;

(4) For every $I, J \in \mathcal{I}(M)$ with $I \subseteq J$, if there is a $B \in \mathcal{F}$ with $I \subseteq B$, then there is a $B' \in \mathcal{F}$ such that either $I \subseteq B' \subseteq J$ or $B' \supseteq J$.

For the remainder of this section, we fix a matroid M of infinite rank and set $E := E(M)$ and $\mathcal{I} := \mathcal{I}(M)$. For $I, J \in \mathcal{I}$, we say J *almost spans* I, and write $I \trianglelefteq J$, if $r_M(I|J) < \infty$. This defines a pre-order $\trianglelefteq$ on $\mathcal{I}$. For $I, J \in \mathcal{I}$, we say I and J are *weakly equivalent* if $I \trianglelefteq J$ and $J \trianglelefteq I$. Moreover, we say I and J are *strongly equivalent*, and write $I \sim J$, if $r_M(I|J) = r_M(J|I) < \infty$. Clearly the relation of strong equivalence is reflexive and symmetric, and it is, indeed, an equivalence relation (we omit the proof).

We observe the following characterisation of strong equivalence with finite difference.

Lemma 3.2. *If $I, J \in \mathcal{I}$ with $|I \setminus J| < \aleph_0$, then $I \sim J$ if and only if $|I \setminus J| = |J \setminus I|$.*

For $I \in \mathcal{I}$, let $[I]$ denote the equivalence class of I with respect to $\sim$. Strong equivalence is clearly a refinement of the weak equivalence, thus it is compatible with almost spanning in the following sense.

Observation 3.3. *Let $I, I', J, J' \in \mathcal{I}$. If $I \sim I'$, $J \sim J'$, and $I \trianglelefteq J$, then $I' \trianglelefteq J'$.*

Hence, the pre-order of almost spanning extends to a pre-order on the set of equivalence classes of $\sim$. We abuse the notation by denoting this by $\trianglelefteq$ as well.

Note that for a finite $I \in \mathcal{I}$ and every $J \in \mathcal{I}$, we have that J almost spans I, and so $[I] \trianglelefteq [J]$. Similarly, if $I \in \mathcal{I}$ with $r(M/I) < \infty$, then $[J] \trianglelefteq [I]$ for every $J \in \mathcal{I}$.

4 Main result

We actually prove the following slight strengthening of Theorem 1.2.

Theorem 4.1. *Let M be a finitary matroid of infinite rank on a ground set E with $|E| < \mathfrak{c}$, and let $\mathcal{F}_0$ be the union of less than $\mathfrak{c}$ many pairwise $\trianglelefteq$-incomparable equivalence classes of $\sim$. If Martin's Axiom holds, then there exists a generalised truncation N of M such that $\mathcal{B}(N) \supseteq \mathcal{F}_0$.*

Proof sketch. Without loss of generality, we may assume that $\mathcal{F}_0$ is non-empty. If $\mathcal{F}_0$ contains a finite set I, it is not hard to show that $[I]$ is the set of bases of the $|I|$-truncation of M. Similarly, if $\mathcal{F}_0$ contains an I with $r(M/I) = n < \infty$, then it is not hard to show that $[I]$ is the set of bases of the $(-n)$-truncation of M.

Suppose now that $\mathcal{F}_0$ contains neither finite sets nor co-finite subsets of bases. We define an increasing continuous sequence $\langle \mathcal{F}_\alpha \colon\ \alpha < \mathfrak{c}\rangle$ with the intension that $\mathcal{F} := \bigcup_{\alpha<\mathfrak{c}} \mathcal{F}_\alpha$ is the set of bases of a generalised truncation N of M. Let $\mathcal{T}$ consists of the ordered pairs (I, J) of M-independent sets such that I is a co-infinite subset of J. We think of a pair $(I, J) \in \mathcal{T}$ as the "task" that, if the bases of N we constructed so far contains a superset B of I, to add a B' for which either $I \subseteq B' \subseteq J$ or $B' \supseteq J$ in order to not violate (4) of Theorem 3.1. Let $\langle (I_\alpha, J_\alpha) \colon\ \alpha < \mathfrak{c}\rangle$ be a sequence with range $\mathcal{T}$ in which every $(I, J) \in \mathcal{T}$ appears unbounded often. For $\alpha < \mathfrak{c}$, we maintain:

(i) $\mathcal{F}_\alpha \subseteq \mathcal{I}$;

(ii) $\mathcal{F}_\alpha$ satisfies the restriction of Theorem 3.1(4) to the pairs $\{(I_\beta, J_\beta) \colon\ \beta < \alpha\}$;

(iii) if $B, B' \in \mathcal{F}_\alpha$ with $B \not\sim B'$, then B and B' are $\trianglelefteq$-incomparable; and

(iv) $\mathcal{F}_{\alpha+1} \setminus \mathcal{F}_\alpha$ is either empty or an equivalence class of $\sim$.

Clearly, $\mathcal{F}_0$ is suitable and these conditions cannot fail at limit steps. Suppose that $\mathcal{F}_\alpha$ is defined for $\alpha < \mathfrak{c}$ and satisfies the conditions. If $\mathcal{F}_\alpha$ satisfies the restriction of Theorem 3.1(4) to the pairs $\{(I_\beta, J_\beta) \colon\ \beta < \alpha + 1\}$, then set $\mathcal{F}_{\alpha+1} := \mathcal{F}_\alpha$. Suppose it does not.

Let $P := \mathsf{Fn}(J_\alpha \setminus I_\alpha, 2)$, i.e. the poset of functions whose domain is a finite subset of $J_\alpha \setminus I_\alpha$ and whose range is a subset of $\{0, 1\}$, ordered by $\supseteq$. We pick a transversal $\mathcal{R}$ of the equivalence classes included in $\mathcal{F}_\alpha$. (iv) ensures that $|\mathcal{R}| < \mathfrak{c}$. Let

$$\mathcal{R}_{I_\alpha} := \{B \in \mathcal{R} \colon\ I_\alpha \trianglelefteq B\} \quad \text{and} \quad \mathcal{R}^{J_\alpha} := \{B \in \mathcal{R} \colon\ B \trianglelefteq J_\alpha\}.$$

For $B \in \mathcal{R}_{I_\alpha}$ and $n < \omega$, let $C_{B,n} := \{p \in P \colon\ r_M(p^{-1}(1)|B) \geq n\}$. For $B \in \mathcal{R}^{J_\alpha}$ and $n < \omega$, let $D_{B,n} := \{p \in P \colon\ r_M(B|J_\alpha \setminus p^{-1}(0)) \geq n\}$.

Claim 1. *Each element of $\mathcal{D} := \{C_{B,n}, D_{B',n} \colon\ n < \omega, B \in \mathcal{R}_{I_\alpha}, B' \in \mathcal{R}^{J_\alpha}\}$ is dense in P.*

Since $|\mathcal{D}| < \mathfrak{c}$ as $|\mathcal{R}| < \mathfrak{c}$ and P is known to be ccc, Martin's Axiom guarantees that there exists a filter F in P such that each $p \in F$ has non-empty intersection with every element of $\mathcal{D}$. We set $B_\alpha := I_\alpha \cup \bigcup_{p\in F} p^{-1}(1)$ and $\mathcal{F}_{\alpha+1} := \mathcal{F}_\alpha \cup [B_\alpha]$.

Claim 2. *$\mathcal{F}_{\alpha+1}$ maintains the conditions (i)-(iv).*

Proof. We only show that (iii). We may assume that $\mathcal{F}_{\alpha+1} \neq \mathcal{F}_\alpha$, since otherwise there is nothing to prove. Then $\mathcal{F}_{\alpha+1} \setminus \mathcal{F}_\alpha = [B_\alpha]$ by construction. Let $B \in \mathcal{F}_\alpha$ be given. We may assume that $B \in \mathcal{R}$ because the relation $\trianglelefteq$ is compatible with $\sim$ (Theorem 3.3).

First we show that $B_\alpha \ntrianglelefteq B$. If $B \in \mathcal{R} \setminus \mathcal{R}_{I_\alpha}$, then, by definition, $I_\alpha \ntrianglelefteq B$ and hence $B_\alpha \ntrianglelefteq B$ because $B_\alpha \supseteq I_\alpha$. If $B \in \mathcal{R}_{I_\alpha}$, then let $n < \omega$ be arbitrary and let $p \in F \cap C_{B,n}$. Then $r_M(B_\alpha|B) \geq r_M(p^{-1}(1)|B) \geq n$. Since $n < \omega$ was arbitrary, $r_M(B_\alpha|B) = \infty$ follows, which means $B_\alpha \ntrianglelefteq B$. The proof that $B \ntrianglelefteq B_\alpha$ is similar. □

Claim 3. *$\mathcal{F} \coloneqq \bigcup_{\alpha<\mathfrak{c}} \mathcal{F}_\alpha$ is the bases of a generalized truncation of M.*

Proof. We check the condition given in Theorem 3.1. Clearly, $\mathcal{F} \neq \emptyset$ because we assumed that $\mathcal{F}_0 \neq \emptyset$. We have $\mathcal{F} \subseteq \mathcal{I}$ by construction. Therefore, (1) holds. If $B, B' \in \mathcal{I}$ with $|B \setminus B'| = |B' \setminus B| < \aleph_0$, then $B \sim B'$ by Theorem 3.2. The set $\mathcal{F}$ is closed under strong equivalence because $\mathcal{F}_0$ is and we maintained (iv). Thus Theorem 3.2 implies that in particular (2) holds. Preserving (iii) implies that (3) is satisfied. To check (4), let $I, J \in \mathcal{I}$ with $I \subseteq J$ and suppose that there is a $B \in \mathcal{F}$ with $I \subseteq B$. If $|J \setminus I| < \aleph_0$, then Theorem 3.2 provides a B' with $B' \sim B$ such that either $I \subseteq B' \subseteq J$ or $B' \supsetneq J$ depending on if $|B \setminus I| \leq |J \setminus I|$ or $|B \setminus I| > |J \setminus I|$. By (2), we have $B' \in \mathcal{F}$. Suppose that $|J \setminus I| \geq \aleph_0$. Let $\alpha < \mathfrak{c}$ be an ordinal such that $B \in \mathcal{F}_\alpha$. Then, by construction, there is a $\beta > \alpha$ such that $(I_\beta, J_\beta) = (I, J)$. But then (ii) ensures that $\mathcal{F}_{\beta+1}$ contains a B' such that either $I \subseteq B' \subseteq J$ or $B' \supsetneq J$. □

This concludes the sketch of the proof of the theorem. □

Theorem 1.2 then follows by taking an infinite and co-infinite subset B of a base of M and applying Theorem 4.1 with $\mathcal{F}_0 \coloneqq [B]$. Indeed, for the resulting matroid N, $B \in \mathcal{B}(N)$ ensures that N is a wild generalised truncation of M.

Corollary 4.2. *If Martin's Axiom holds, then every finitary matroid M of infinite rank on a ground set E with $|E| < \mathfrak{c}$ has exactly $2^{\mathfrak{c}}$ pairwise non-isomorphic wild truncations.*

Conjecture 4.3. *It is consistent relative to ZFC that every finitary matroid of infinite rank admits a wild generalised truncation.*

Our proof of Theorem 4.1 relies heavily on the assumption that matroid M is finitary. We ask if it possible to prove something for general matroids in a suitable setting.

Question 4.4. *Is it consistent relative to ZFC that every matroid of infinite rank admits a wild generalised truncation?*

References

[1] Nathan Bowler. Infinite matroids. Habilitation Thesis, University of Hamburg, 2014.

[2] Nathan Bowler and Johannes Carmesin. Matroids with an infinite circuit-cocircuit intersection. *J. Combin. Theory Ser. B*, 107:78–91, 2014.

[3] Nathan Bowler and Stefan Geschke. Self-dual uniform matroids on infinite sets. *Proc. Amer. Math. Soc.*, 144(2):459–471, 2016.

[4] Henning Bruhn, Reinhard Diestel, Matthias Kriesell, Rudi Pendavingh, and Paul Wollan. Axioms for infinite matroids. *Adv. Math.*, 239:18–46, 2013.

[5] J. Pascal Gollin and Attila Joó. Wild generalised truncations of infinite matroids. Preprint, 2025.

[6] Denis Arthur Higgs. Equicardinality of bases in B-matroids. *Canad. Math. Bull.*, 12:861–862, 1969.

[7] Denis Arthur Higgs. Matroids and duality. *Colloq. Math.*, 20:215–220, 1969.

[8] Thomas Jech. *Set theory.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 3rd millennium ed. edition, 2003.

[9] James G. Oxley. Infinite matroids. *Proc. London Math. Soc. (3)*, 37(2):259–272, 1978.

[10] Richard Rado. Abstract linear dependence. *Colloq. Math.*, 14:257–264, 1966.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25-29, 2025

Partial quasigroup digraphs

(Extended abstract)

Manuel González-Regadera* Raúl M. Falcón[†] Daniel Kotlar[‡]

Abstract

In the literature, one may find different generalizations of the notion of a Cayley digraph (such as quasigroup digraphs or quasi-Cayley digraphs, among others), all of them arisen from an algebraic structure satisfying certain conditions and a subset of its elements. In this paper, we introduce partial quasigroup digraphs (PQDs) as a natural generalization of Cayley digraphs. Unlike any previous approach, our proposal does not arise from a set of elements of the underlying algebraic structure (in our case, a quasigroup), but from a partial Latin square embedded in its multiplication table. We prove that every digraph is a PQD whose topology only depends on the partial Latin square under consideration. Then, we characterize those completable partial Latin squares that give rise to simple, undirected, regular, vertex-transitive, and strongly connected PQDs.

1 Introduction

The *Cayley digraph* $\mathrm{Cay}(G, S)$ of a group G, with *connection set* $S \subseteq G$, is a digraph whose set of vertices is formed by the elements of G, so that there is an arc from a vertex $a \in G$ to a vertex $b \in G$ whenever $b \in aS$. This digraph is regular and vertex-transitive. If $1 \notin S$, then $\mathrm{Cay}(G, S)$ is simple. If $S = S^{-1}$, then it is undirected. Finally, if the set S generates the group G, then $\mathrm{Cay}(G, S)$ is strongly connected.

*Department of Applied Mathematics I, Universidad de Sevilla, Spain. E-mail: mgonzalez34@us.es.
Supported by the Research and Innovation Project PPIT-FEDER 2023 *"Modeling small-world, scale-free networks from combinatorial designs based on quasigroup digraphs"*, co-financed by the EU - Ministry of Finance and Public Administration - European Funds - Andalusian Regional Government - Ministry of University, Research and Innovation.
[†]Department of Applied Mathematics I, Universidad de Sevilla, Spain. E-mail: rafalgan@us.es.
[‡]Computer Science Department, Tel-Hai College, Israel. E-mail: dannykot@telhai.ac.il.

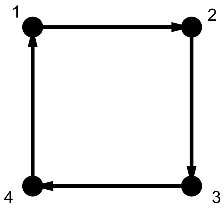
Generalizing Cayley digraphs from groups to more general algebraic structures is an active area of research [1, 2, 5, 9]. In 1969, Teh [8] introduced *quasigroup digraphs* (QDs) as a natural generalization of Cayley digraphs in which the underlying group is replaced by a quasigroup. The latter is a pair $(Q,*)$ formed by a set Q endowed with a binary operation $*$ such that both equations $a*x=b$ and $y*a=b$ have unique solutions $x,y\in Q$ for all $a,b\in Q$. More precisely, $x=a\backslash b$ and $y=b/a$, which correspond, respectively, to the left and right divisions of the quasigroup $(Q,*)$. The latter is identified with its multiplication table $L_{(Q,*)}:=\left(L_{(Q,*)}[a,b]\right)_{a,b\in Q}$, with $L_{(Q,*)}[a,b]:=ab$. (From now on, we omit the product $*$ when there is no risk of confusion.) It is a *Latin square*, in which no two cells in the same row or in the same column contain the same symbol. The Latin square $L_{(Q,*)}$ is uniquely determined by its set of entries $\mathrm{Ent}(L_{(Q,*)}):=\{(a,b,ab):\ a,b\in Q\}$.

A relevant aspect here is that a digraph is regular if and only if it is a QD (see [3]). However, unlike Cayley digraphs, a QD may not be vertex-transitive. To guarantee this property, Shee [7] imposed the condition of *right-associativity* on the subset $S\subseteq Q$. That is, $(ab)S=a(bS)$ for all $a,b\in Q$. Imposing additional conditions on the set S has turned out to be an excellent way to generalize some other basic properties of Cayley digraphs. In this regard, Gauyacq [4] introduced *quasi-Cayley digraphs* as a QD over a right-associative generating subset $S\subseteq Q$, with right unit $e\in Q\backslash S$ and such that $e\in sS$ for all $s\in S$. Under these conditions, one may ensure that every quasi-Cayley digraph is undirected, connected, and simple. More generally, Mwambene [6] proved that a quasigroup digraph $\mathrm{Cay}(Q,S)$ is simple whenever $a\notin aS$ for all $a\in Q$, while it is undirected whenever $a\in(as)S$ for all $(a,s)\in Q\times S$. He termed *Cayley set* to any subset $S\subseteq Q$ satisfying these two conditions.

In this paper, we discuss this topic by showing that these additional conditions in S can be imposed much more generally on the entries related to that set. Note here that the existence of an arc $a\to b$ in $\mathrm{Cay}(Q,S)$ is ensured by the entry $(a,a\backslash b,b)\in\mathrm{Ent}(L_{(Q,*)})$, whenever $a\backslash b\in S$. Note also that every subset $A\subseteq\mathrm{Ent}(L_{(Q,*)})$ uniquely determines a *partial Latin square*. That is, an $n\times n$ array such that each symbol in Q appears at most once per row and at most once per column. It is in turn the multiplication table of a *partial quasigroup* $(Q,*_A)$, in which both equations $a*_A x=b$ and $y*_A a=b$ have at most one solution $x,y\in Q$ for all $a,b\in Q$. With a slight abuse of notation, we indicate this fact as $A\subseteq L_{(Q,*)}$.

In the QD $\mathrm{Cay}(Q,S)$, we are therefore only interested in the partial Latin square embedded in our quasigroup $(Q,*)$ whose nonempty entries are those in $L_{(Q,*)}$ that appear in the columns indexed by S. In fact, every Latin square containing these entries gives rise to the same QD, regardless of the remaining entries. Thus, for example, the cyclic digraph in Figure 1 is a QD associated with the Cayley set $S=\{2\}$ and any of the two Latin squares L_1 and L_2 indicated therein. (Here, we highlight the entries related to the Cayley set.) Based on all these remarks, we introduce the following concept.

Definition 1. *The* partial quasigroup digraph *of the quasigroup* $(Q,*)$*, with* connection set $S\subseteq L_{(Q,*)}$, *is the digraph* $\mathrm{Cay}_*(Q,S)$ *(or simply* $\mathrm{Cay}(Q,S)$ *if there is no risk of confusion) whose set of vertices is formed by all elements of* Q*, and so that there is an arc* (a,b) *from an initial vertex* $a\in Q$ *to a terminal vertex* $b\in Q$ *if and only if* $(a,a\backslash b,b)\in\mathrm{Ent}(S)$.



$L_1 \equiv$

1	**2**	3	4
2	**3**	4	1
3	**4**	1	2
4	**1**	2	3

$L_2 \equiv$

3	**2**	1	4
1	**3**	4	2
2	**4**	3	1
4	**1**	2	3

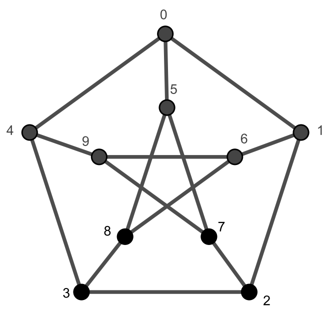
$L_3 \equiv$

3	**2**	1	4
1	4	**3**	2
4	1	2	3
2	3	4	**1**

Figure 1: A cyclic digraph as some PQDs.

Every digraph can therefore be associated with different partial Latin squares embedded in distinct Latin squares, although all of them contain the same subset of symbols in each given row. Thus, for example, the cyclic digraph in Figure 1 is a PQD associated with each of the three Latin squares L_1, L_2, and L_3, and the partial Latin squares highlighted therein.

The main aim of this paper is to describe the properties that a partial Latin square must satisfy so that its related PQD has the usual properties of Cayley digraphs and QDs. In order to understand the potential of this approach, note here that every digraph can be represented by a partial quasigroup embedded in a group. Thus, for example, the Petersen graph, which is not a Cayley digraph, can be embedded in the cyclic group $\mathbb{Z}_{10}$ as indicated in Figure 2.



0	**1**	2	3	**4**	**5**	6	7	8	9
1	**2**	3	4	5	**6**	7	8	9	**0**
2	**3**	4	5	6	**7**	8	9	0	**1**
3	**4**	5	6	7	**8**	9	0	1	**2**
4	5	6	7	8	**9**	**0**	1	2	**3**
5	6	**7**	**8**	9	**0**	1	2	3	4
6	7	**8**	**9**	0	**1**	2	3	4	5
7	8	**9**	0	1	**2**	3	4	**5**	6
8	9	0	1	2	**3**	4	**5**	**6**	7
9	0	1	2	3	**4**	5	**6**	**7**	8

Figure 2: The Petersen graph as a PQD.

The paper is organized as follows. Section 2 describes some preliminary concepts on Latin squares that are used throughout the paper. Then, in Section 3 we characterize those subsets of entries which give rise to PQDs that are simple, undirected, regular, vertex-transitive or strongly connected. Moreover, we show how these characterizations generalize in a natural way all the conditions that are commonly imposed on the set S in any quasigroup digraph $\mathrm{Cay}(Q, S)$ satisfying the same properties.

2 Preliminaries

Let $\mathcal{L}(Q)$ be the set of partial Latin squares of order $|Q|$, with entries in Q. The *shape* of $L \in \mathcal{L}(Q)$ is the binary square array $\mathrm{Sh}(L)$ of order $|Q|$ containing a zero in a given cell iff the same cell in L is empty, and 1, otherwise. L is *k-row-regular* if each row contains exactly k nonempty entries. It is *k-symbol-regular* if each symbol appears exactly k times.

Let $\mathcal{S}_3$ be the symmetric group on the set $\{1,2,3\}$, and let $\pi \in \mathcal{S}_3$. The *π-conjugate* $L^\pi \in \mathcal{L}(Q)$ is defined so that $\mathrm{Ent}(L^\pi) = \big\{\big(a_{\pi(1)}, a_{\pi(2)}, a_{\pi(3)}\big) : \ (a_1,a_2,a_3) \in \mathrm{Ent}(L)\big\}$. Furthermore, let $\mathcal{S}_Q$ be the symmetric group on Q. A *right isotopism* from L to $L' \in \mathcal{L}(Q)$ is any pair $(f,g) \in \mathcal{S}_Q \times \mathcal{S}_Q$ such that $L'[f(a),g(b)] = f(L[a,b])$ for all $a,b \in Q$. If $L = L'$, then it is a *right autotopism* of L. The set $\mathrm{RAut}(L)$ formed by the right autotopisms of L has a group structure. We say that L is *rs-transitive* under the action of $\mathrm{RAut}(L)$ if, for all $a,b \in Q$, there is a right autotopism $(f,g) \in \mathrm{RAut}(L)$ such that $f(a) = b$.

3 Characterizing connection sets of PQDs

Let $\mathrm{Cay}(Q,S)$ be the PQD of a quasigroup $(Q,*)$, with connecton set $S \subseteq L_{(Q,*)}$. The concept of PQD generalizes in a natural way that of QD. To see it, we define the set $S_{\mathrm{col}} := \{s \in Q\colon \ (a,s,as) \in \mathrm{Ent}(S), \text{ for some } a \in Q\}$.

Lemma 2. *If* $|S| = |Q| \cdot |S_{\mathrm{col}}|$, *then* $\mathrm{Cay}(Q,S) = \mathrm{Cay}(Q,S_{\mathrm{col}})$.

Unlike QDs, every digraph is a PQD, regardless of the quasigroup under consideration. Its topology only depends on the shape of the (23)-conjugate of its connection set.

Lemma 3. *Every digraph of order $|Q|$ is a PQD for exactly one connection set embedded in $L_{(Q,*)}$. In particular,* $\mathrm{Cay}(Q, L_{(Q,*)})$ *is the complete graph of order $|Q|$.*

Lemma 4. *A digraph is isomorphic to* $\mathrm{Cay}(Q,S)$ *iff its adjacency matrix is permutation-similar to* $\mathrm{Sh}\left(S^{(23)}\right)$.

We establish the following equivalence relation among completable partial Latin squares.

Definition 5. *Let $S_1, S_2 \in \mathcal{L}(Q)$ be two completable partial Latin squares. They are* pqd-compatible *if both* $\mathrm{Sh}\left(S_1^{(23)}\right)$ *and* $\mathrm{Sh}\left(S_2^{(23)}\right)$ *are permutation-similar.*

Proposition 6. *Two PQDs are isomorphic iff their connection sets are pqd-compatible.*

Theorem 7. *If the connection sets of two PQDs are right isotopic, then both digraphs are isomorphic.*

Corollary 8. *Isomorphic completable partial Latin squares describe isomorphic PQDs.*

Hence, PQDs enable one to establish a natural translation among notions and properties of completable partial Latin squares and digraphs. As an illustrative example, we characterize those completable partial Latin squares that give rise to regular, simple, undirected, strongly connected or vertex-transitive PQDs.

Lemma 9. $\mathrm{Cay}(Q,S)$ *is k-regular iff* $S^{(23)}$ *is k-regular.*

Lemma 10. $\mathrm{Cay}(Q,S)$ *is simple iff* $(a, a\backslash a, a) \notin \mathrm{Ent}(S)$ *for all* $a \in Q$. *Furthermore,* $\mathrm{Cay}(Q,S)$ *is undirected iff* $(b, b\backslash a, a) \in \mathrm{Ent}(S)$ *whenever* $(a, a\backslash b, b) \in \mathrm{Ent}(S)$.

We say that $S \in \mathcal{L}(Q)$ is *Cayley* iff both conditions in Lemma 10 hold.

Proposition 11. *If* $|S| = |Q| \cdot |S_{\mathrm{col}}|$, *then S is Cayley iff* S_{col} *is a Cayley set.*

A completable partial Latin square $S \in \mathcal{L}(Q)$ *generates* the set Q iff there is an element $r \in Q$ such that, for every $a \in Q$, there is a sequence $(s_0, \ldots, s_m) \in Q^{m+1}$, for some $m \geq 0$, with $s_0 = r$ and $s_m = a$, satisfying $(s_k, s_k\backslash s_{k+1}, s_{k+1}) \in \mathrm{Ent}(S)$ for all $k < m$. We denote this fact by $\langle S \rangle = Q$. In addition, we call r a *root* of S.

Proposition 12. $\mathrm{Cay}(Q,S)$ *is strongly connected iff every element in Q is a root of S.*

Proposition 13. *If* $\mathrm{Cay}(Q,S)$ *is vertex-transitive, then it is strongly connected iff* $\langle S \rangle = Q$.

The next lemma characterizes those PQDs that are vertex-transitive.

Lemma 14. $\mathrm{Cay}(Q,S)$ *is vertex-transitive iff, for each* $a, b \in Q$, *there is* $f \in \mathcal{S}_Q$, *with* $f(a) = b$, *and* $(c, c\backslash d, d) \in \mathrm{Ent}(S) \Leftrightarrow (f(c), f(c)\backslash f(d), f(d)) \in \mathrm{Ent}(S)$, *for all* $c, d \in Q$.

Proposition 15. $\mathrm{Cay}(Q,S)$ *is vertex-transitive whenever S is rs-transitive under* $\mathrm{RAut}(S)$.

We say that a partial Latin square $S \subseteq L_{(Q,*)}$ is *right associative* if $(a, a\backslash b, b) \in \mathrm{Ent}(S) \Leftrightarrow (ca, (ca)\backslash(cb), cb) \in \mathrm{Ent}(S)$ for all $c \in Q$. It is *right commutative* if $(a, a\backslash b, b) \in \mathrm{Ent}(S) \Leftrightarrow (ac, (ac)\backslash(bc), bc) \in \mathrm{Ent}(S)$ for all $c \in Q$.

Proposition 16. *A PQD is vertex-transitive if its connection set is either right associative or right commutative.*

Proposition 17. *Every PQD with a right commutative connection set is a PQD with a right associative connection set.*

Proposition 18. *If* $|S| = |Q| \cdot |S_{\mathrm{col}}|$, *then* S_{col} *is right associative or commutative iff S is.*

Theorem 19. *Every connected PQD with either a right associative or a right commutative connection set is quasi-Cayley.*

4 Conclusions and further work

In this paper, we have introduced the partial quasigroup digraph $\mathrm{Cay}(Q,S)$ associated with a quasigroup $(Q,*)$ and a partial Latin square $S \subseteq L_{(Q,*)}$. It constitutes a natural generalization of quasigroup digraphs (and hence Cayley digraphs), which arise whenever the nonempty entries of S are those in a set of columns in the multiplication table of the quasigroup. We have characterized those partial Latin squares that give rise to simple, undirected, regular, strongly connected, and vertex-transitive partial quasigroup digraphs.

As indicated in the Introduction, every digraph can be represented by a partial quasigroup embedded in a group. An open question to deal with consists of studying under which conditions it can be represented by a partial group, so that the associative property holds for the partial binary operation under consideration.

References

[1] Al-Tahan M, Davvaz B. Generalized Cayley digraphs over hypergroups and their graph product. *Discrete Math. Algorithms Appl.* **17** (2025), Paper No. 2450022.

[2] Bussaban L, Kaewkhao A, Suantai S. Cayley digraphs of gyrogroups.*Quasigroups Related Systems* **27**(2019), 25–32.

[3] Dörfler W. Every regular graph is a quasigroup digraph. *Discrete Math.* **10** (1974), 181–183.

[4] Gauyacq G. On quasi-Cayley digraphs. *Discrete App. Math.* **77** (1997), 43–58.

[5] Knauer K., Surroca GP. On monoid graphs. *Mediterr. J. Math.* **20** (2023), Paper No. 26, 24 pp.

[6] Mwambene E. Characterisation of regular graphs as loop graphs. *Quaest. Math.* **28** (2005), 243–250.

[7] Shee SC. On a class of point-symmetric graphs induced by quasigroups. *Nanta Math.* **5** (1971), 92–107.

[8] Teh HH, An algebraic representation theorem of finite regular graphs. *Nanta Math.* **3** (1969), 88–94.

[9] Zhu XM, Liu W, Yang X. The isomorphism of generalized Cayley digraphs on finite non-abelian simple groups. *Discrete Math.* **346** (2023), Paper No. 113292, 10 pp.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Distinguishing Graphs and Cost Number using Automorphism Representations

(Extended abstract)

Alexa Gopaulsingh* Zalán Molnár †

Abstract

A *distinguishing coloring* of a graph is a vertex coloring such that only the identity automorphism of the graph preserves the coloring. A *2-distinguishable graph* is a graph which can be distinguished by a coloring using 2 colors. The *cost* $\rho(G)$ of a 2-distinguishable graph is the smallest size of a color set of a distinguishing coloring of G. The *determining number* of a graph, $Det(G)$, is the minimum number of nodes, which if fixed by a coloring, would ensure that the coloring distinguishes the entire graph.

Boutin posed an open problem in [8] which asks if $\rho(G)$ and $Det(G)$ can be arbitrarily far apart. It is trivial that it cannot be so for the case $Det(G) = 1$ but the answer was unknown for $Det(G) \geq 2$. We solve this problem for the case $Det(G) = 2$. We show that for the case $Det(G) = 2$, that not only is the cost bounded but in fact it takes small values with $\rho(G) = 2$, 3 or 4. In order to establish this, the concept of the *automorphism representation* of a graph is introduced. Graphs having equivalent automorphism representations implies that they have the same distinguishing number (note that just having isomorphic automorphism groups is not enough for this to hold). This prompts a factoring of graphs by which two graphs are *distinguishably equivalent* iff they have equivalent automorphism representations.

1 Introduction

The notion of distinguishing colorings was introduced in 1977 by László Babai under the name of asymmetric colorings, see [4]. This is a coloring of the nodes of a graph, not necessarily proper, such that only the identity automorphism preserves the coloring. Following this, in 1996, Albertson and Collins introduced in [3] the **distinguishing number** of a graph G, $D(G)$, as the least integer d such that G admits a vertex coloring which is a distinguishing coloring of G with d colors. A graph which can be distinguished by a coloring using d colors is said to be **d-distinguishable**. There are several papers finding the distinguishing numbers of various classes of graphs, see [1, 11, 18, 15].

*Deaprtment of Logic, Eötvös Loránd University, Hungary. E-mail: `alexa279e@gmail.com`.
†Department of Logic, Eötvös Loránd University, Hungary. E-mail: `mozaag@gmail.com`.

The anecdotal setup introducing distinguishing colorings is one of seeing the nodes of a graph as identical looking keys placed in a graph structure. Then we ask, what is the minimum number of colors needed to be able to tell which key is which? For example, if the keys are connected in a line by a string, we need two colors, say red and blue, to differentiate, for example the keys on the different ends. In [6], Boutin noticed that when 2-distinguishable graphs are considered, we can *really* use one color and consider the uncolored class the remaining color. In our chain of keys example, say we have 10 keys or nodes. It is enough to only use red and color all but, say 3 keys red and now each key will have a unique place in this structure. The natural question, then became what is the minimum number of nodes that need to be painted in order to distinguish the graph, see [10]? This motivates the definition of the **cost of a 2-distinguishable graph**, $\rho(G)$, as the minimum number of nodes needed to be colored, red say, to distinguish the graph, see Definition 2.16. For the 10-node path in the given example, coloring all but 3 nodes red is wasting paint and it is enough to color 1 end node red to distinguish it for any P_n, where $n \geq 2$. For the cycles C_n, where $n \geq 6$, it can be shown that only 3 nodes need to be painted red in order to make the structure rigid. The difference in the number of nodes of the graph and cost can be vast. Boutin in [6] gave the example that the hypercube Q_{16} has 2^{16} vertices and $16! \times 2^{16}$ automorphisms yet one only needs to color just 7 nodes red to distinguish it. For $n \geq 5$, Boutin [6] showed that for hypercubes Q_n, $\lceil log_2 n \rceil + 1 \leq \rho(Q_n) \leq 2\lceil log_2 n \rceil - 1$.

Recent work shows that there are several classes of graphs that are 2-distinguishable. These classes include hypercubes Q_n, where $n \geq 4$, see [5], Cartesian powers G^n of a connected graph $G \neq K_2, K_3$ where $n \geq 2$, see [15], Kneser graphs $K_{n,k}$ where $n \geq 6$, and $k \geq 2$, see [2], and 3-connected planar graphs (excluding seven small graphs), see [13]. In fact, it is conjectured that almost all finite graphs are 2-distinguishable, see [17].

A significant concept used for finding the distinguishing and cost numbers of a graph, is the *determining number* of a graph G, denoted by $Det(G)$ (see Definition 2.16). This is the minimimum size of a set of nodes S, which is such that if φ is a non-trivial automorphism of G, then φ moves at least one node of S. This concept was also introduced in the literature as the *fixing number* of a set, see [12], [14]. In [9], considering 2-distinguishable graphs Boutin observed that $\rho(G) \geq Det(G)$. Now, if $Det(G) = 1$, i.e. only one node needs to be fixed to distinguish G, then coloring that node red and everything else blue would distinguish it and so its cost would also be 1. However, it was not clear what the cost would be of $Det(G) \geq 2$, prompting Boutin to ask in [8], the question:

Open question: Find graphs for which $\rho(G)$ is arbitrarily larger than $Det(G)$.

We will show in this paper that if $Det(G) = 2$, then one cannot expect to find such graphs. Indeed, not only is the cost bounded but in this case it takes small values and here $\rho(G) = 2, 3$ or 4, no matter how large the graph is. Note however, that the increase in complexity of the solution from $Det(G) = 1$ to 2 is drastic as will be witnessed by the complexity of the proof required to establish the result. It is currently unknown what the situation is for $Det(G) \geq 3$, even whether this condition bounds the cost at all. Moreover, if it does, then what is the bound?

A concept developed, in order to show the main result is that of automorphism representations of a graph. This is essentially the automorphism group of a graph but represented as a permutation of its nodes (modulo different possible labellings of the graph). Note that, two graphs having isomorphic automorphism groups is not enough to conclude that they have the same distinguishing number. Similar observations were made by Tymoczko in [18]. It turns out that graphs having equivalent automorphism group representations *is* enough to conclude that they have the same distinguishing number. Here, we developed this direction and under the appropriate labelling correspondence, the very same colorings which can be used to distinguish one graph can be used to distinguish the other. This leads to a factoring of graphs under the equivalence relation of two graphs being equal iff they have equivalent automorphism representations. These graphs can be considered to be *distinguishably equivalent* and to give a k-distinguishing coloring of any graph is always to implicitly distinguish all of the graphs in its equivalence class. Next, we will give definitions and examples building up this notion which we will then use to solve the main result.

2 Main result

Definition 2.1. Let G be a graph. A **labelling** of G is an injective map, $l : V(G) \to S$, naming the nodes of G. For each automorphism α of G that is expressed in a cycle notation, $(v_{11}v_{12}...v_{1k_1})...(v_{m1}v_{m2}...v_{mk_n})$, let $l(\alpha)$ denote $(l(v_{11})l(v_{12})...l(v_{1k_1}))...(l(v_{m1})l(v_{m2})...l(v_{mk_n}))$. We define $Aut(G, V(G)_l) := \{l(\alpha) : \alpha \in Aut(G)\}$ as **the automorphism representation of G under the labelling l.** Note that any permutation can be represented as a product of disjoint cycles and we will always represent them this way throughout this paper unless stated otherwise.

Definition 2.2. We define an equivalence relation on the class of graphs and labelled representations: $G_1 \equiv G_2$ and $Aut(G_1, V(G_1)_{l_i}) \equiv Aut(G_2, V(G_2)_{l_j})$ iff there is $l_1 : V(G_1) \to S$ and $l_2 : V(G_2) \to S$ such that $Aut(G_1, V(G_1)_{l_1}) = Aut(G_2, V(G_2)_{l_2})$. Then G_1 and G_2 are called **distinguishably equivalent** and the equivalence class of $Aut(G, V(G)_{l_i})$ is called the **automorphism representation** of G and is denoted by $Aut(G, V(G))$.

Proposition 2.3. *The following holds:*

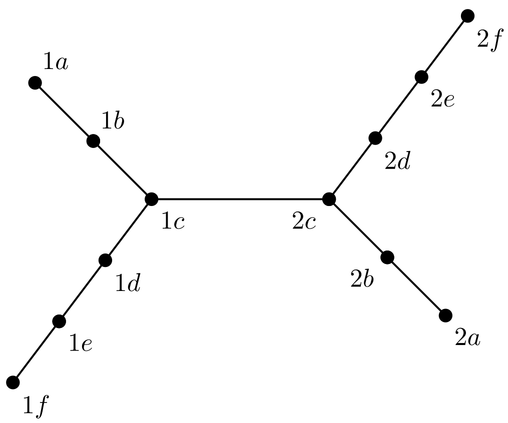
(i) $Aut(G_1, V(G_1)) = Aut(G_2, V(G_2)) \Rightarrow Aut(G_1) \cong Aut(G_2)$,

(ii) $Aut(G_1) \cong Aut(G_2) \not\Rightarrow Aut(G_1, V(G_1)) = Aut(G_2, V(G_2))$,

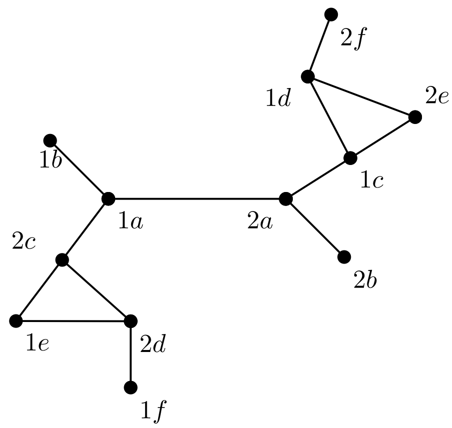
(iii) $Aut(G_1, V(G_1)) = Aut(G_2, V(G_2)) \Rightarrow |V(G_1)| = |V(G_2)|$,

(iv) $|V(G_1)| = |V(G_2)| \not\Rightarrow Aut(G_1, V(G_1)) = Aut(G_2, V(G_2))$.

Example 2.4. Consider the graphs G_1 and G_2, as shown in the following diagram. These are distinguishably equivalent graphs as witnessed by the following labellings and the automorphism representations corresponding to these:



G_1



G_2

$$Aut(G_1, V(G)_{l_1}) = Aut(G_2, V(G)_{l_2}) = \{e, (1a, 2a)(1b, 2b)(1c, 2c)(1d, 2d)(1e, 2e)(1f, 2f)\}$$

Definition 2.5. Let $c : V(G) \to C$ be a c-coloring of G. An automorphism $\varphi \in Aut(G, V(G)_l)$ is **non-monochromatically colored by c** or **broken by c** if at least one of the cycles in its disjoint cycle representation, is non-monochromatically colored. In particular, let

$$\varphi = (n_{11}, n_{12} \dots n_{1k_1})(n_{21}, n_{22} \dots n_{2k_2}) \dots (n_{m1}, n_{m2} \dots n_{mk_m}) \in Aut(G, V(G)_l),$$

where $n_{ij} = l(v_x)$ for some labelling $l : V(G) \to S$ of G. Then, φ is broken by c if there is some $(n_{i1}, n_{i2} \dots n_{ik_i})$, such that $c(n_{ij}) \neq c(n_{ij+1})$ for some j or $c(n_{ik_i}) \neq c(n_{i1})$ and $k_i \geq 2$.

Proposition 2.6. *Let $l_1 : V(G) \to S_1$, $l_2 : V(G) \to S_2$ be two labellings of G, let $c : V(G) \to C$ be a coloring of G and let $\varphi \in Aut(G)$. Then, $l_1(\varphi) \in Aut(G, V(G)_{l_1})$ is broken by c iff $l_2(\varphi) \in Aut(G, V(G)_{l_2})$ is broken by c.*

Remark 2.7. Proposition 2.6 indicates that if we consider one graph only, then the labelling does not matter. However, to determine if two graphs are equivalent, one needs to find corresponding labellings for which their representations are equal. In Example 2.4, if we switch the labels of say $1c$ and $1d$ in G_1, leaving the remaining labels the same for G_1 and G_2, then these labellings do not witness their equivalence.

Definition 2.8. An **automorphism** $\varphi \in Aut(G)$ is said to be broken by a coloring $c : V(G) \to C$, if φ's representation in $Aut(G, V(G)_l)$ is broken by c for any labelling of the nodes l. An **automorphism representation of a graph**, $Aut(G, V(G))$, is said to be **non-monochromatically colored** by a coloring $c : V(G) \to C$ iff for any labelling $l : V(G) \to S$, every non-identity $\varphi \in Aut(G, V(G)_l)$ is broken or non-monochromatically colored.

Lemma 2.9. *A graph G is distinguished by a coloring $c : V(G) \to C$, iff for any labelling, $l : V(G) \to S$, $Aut(G, V(G)_l) - \{e\}$ is non-monochromatically colored by c.*

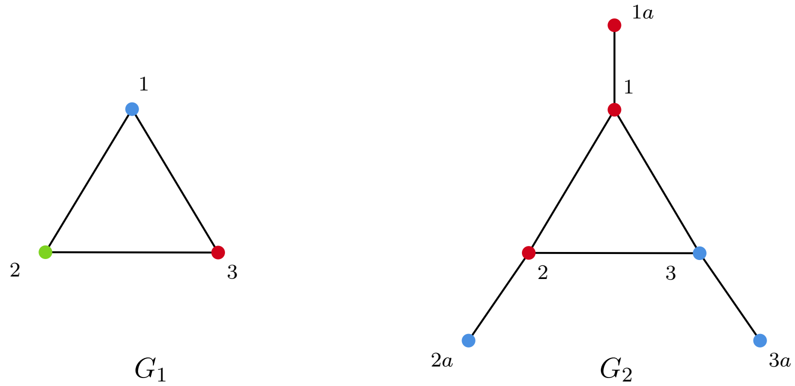
Theorem 2.10. *Let G_1 and G_2 be two graphs. Then,*

$$Aut(G_1, V(G_1)) = Aut(G_2, V(G_2)) \Rightarrow D(G_1) = D(G_2).$$

Remark 2.11. Using Proposition 2.6 and Theorem 2.10, we see that in order to distinguish a graph, it is enough to give it an arbitrary labelling l and find a non-monochromatic coloring of $Aut(G, V(G)_l)$.

Remark 2.12. While Theorem 2.10 states that the distinguishing numbers of G_1 and G_2 are equal if $Aut(G_1, V(G_1)) = Aut(G_2, V(G_2))$, note that is not true if simply $Aut(G_1) \cong Aut(G_2)$.

Example 2.13. Let G_1 be a triangle and G_2 be a triangle with nodes of degree 1 attached to each node of the triangle, as below. Then, $Aut(G_1) \cong Aut(G_2) \cong S_3$ which is the symmetric group on 3 elements but $D(G_1) = 3$ and $D(G_2) = 2$. However, as $|V(G_1)| \neq |V(G_2)|$, using Proposition 2.3 (iii), $Aut(G_1, V(G_1)) \neq Aut(G_2, V(G_2))$ and $G_1 \not\cong G_2$.



$$
\begin{aligned}
Aut(G_1, V(G_1)l_1) &= \{e, (1,2), (1,3), (2,3), (1,3,2), (1,2,3)\} \\
Aut(G_2, V(G_2)l_2) &= \{e, (1,2)(1a,2a), (1,3)(1a,3a), (2,3)(2a,3a), \\
&\qquad (1,2,3)(1a,2a,3a), (1,3,2)(1a,3a,2a)\}.
\end{aligned}
$$

Notice that the distinguishing colorings of G_1 and G_2 above induce non-monochromatic colorings of all of the non-identity permutations in their automorphism representations (recall that these are colorings where at least one (but not necessarily all)) of the cycles in a permutation are colored non-monochromatically). Conversely, any non-monochromatic coloring of a graph's labelled automorphism representation induces a distinguishing coloring of the graph.

We note that if H is an induced subgraph of G, then a distinguishing coloring of G when restricted to H need not be a distinguishing coloring of H. However, when the nodes of H have the same neighbours in $G - H$, then this holds. This is shown next.

Lemma 2.14. *Let G be a graph and H an induced subgraph of G such that for any h_1, $h_2 \in H$, $N(h_1) - V(H) = N(h_2) - V(H)$. If a coloring $k : V(G) \to C$ is a distinguishing coloring of G, then $k|_{V(H)}$ is a distinguishing coloring of H, where $N(h_i) = \{z : (h_i, z) \in E(G)\}$.*

Proof. Let φ be a non-identity automorphism of H. Let $\varphi_G : V(G) \to V(G)$ defined by $\varphi_G(v) = \varphi(v)$, if $v \in V(H)$ and $\varphi_G(v) = v$, if $v \in V(G) - V(H)$.

Claim 2.15. *φ_G is an automorphism of G.*

Proof. To see this, note that there are 3 types of edges and non-edges in G: (i) Edges and non-edges in H, (ii) Edges and non-edges in $G - H$ and (iii) Edges and non-edges between H and G. The first two types of edge and non-edges are clearly preserved by φ_G. The third type is preserved since $N(h_1) - V(H) = N(h_2) - V(H)$ for any $h_1,\ h_2 \in V(H)$, implies that for $g \in V(G-H)$, $(h_1, g) \in E(G)$ iff $(h_2, g) \in E(G)$. Since φ_G preseves all edges and non-edges of G, it is an automorphism of G and the claim is shown. □

Let $l : V(G) \to S$ be an injective map labelling the nodes of G and let l_H be this map restricted to the nodes of H. Let $l(V(G) - V(H)) = \{g_1, \ldots g_k\}$ be the l-labelling of the nodes in $G - H$. Then $\varphi \in Aut(H, V(H)_l)$ implies $\varphi_G = (g_1)\ldots(g_k)\varphi \in Aut(G, V(G)_l)$ by the previous claim. By Lemma 2.9, as k is a distinguishing coloring of G, it must non-monochromatically color φ_G. Since φ_G has the same disjoint cycle representation as φ apart from single cycles, k non-monochromatically colors φ_G iff k non-monochromatically color φ. As k is a distinguishing coloring of G, it non-monochromatically colors all elements in $Aut(G, V(G)_l) - \{e\}$ by Lemma 2.9. In particular, k non-monochromatically colors all elements in $\{\varphi_G : \varphi \neq e \text{ and } \varphi \in Aut(H, V(H)_{l_H})\}$. Hence, k non-monochromatically colors all elements in $Aut(H, V(H)_{l_H}) - \{e\}$ and we get that $k|_{V(H)}$ is a distinguishing coloring of H by Lemma 2.9. □

Lemma 2.14, along with several other technical lemmas will be used to prove the main result.

Definition 2.16. Let G be a graph. A subset $A \subseteq V(G)$ is said to be a *determining set* for G if whenever $\varphi \in Aut(G)$ so that $\varphi(x) = x$ for all $x \in A$, then $\varphi = e$, the identity. The **determining number**, $Det(G)$, of G is the minimum size of a determining set for G. If G is a 2-distinguishable graph, call a color class in a 2-distinguishing coloring of G a **distinguishing class** and the size of the smaller color class, **the cost of the coloring**. If G is a *2-distinguishable* graph, then the minimum size of distinguishing class of G is called the **cost number** of G and is denoted by $\rho(G)$.

Boutin showed the following in [8]:

Lemma 2.17. *A subset of vertices S is a distinguishing class for G iff S is a determining set for G with the property that every automorphism that fixes A setwise, also fixes it pointwise.*

Our main result answering the question of Boutin asked in [8] for $Det(G) = 2$, is the following:

Theorem 2.18 (Main Theorem)**.** *Let G be a graph with $Det(G) = 2$ and $D(G) = 2$, then $\rho(G) = 2, 3,$ or 4.*

Outline of the proof: Let $\{x, y\}$ be a minimum sized determining set of G. Consider the permutation $(xy)\alpha$, where α is a product of zero or more cycles encoding the images of the permutation for $V(G) - \{x, y\}$. We showed that α must consist of 2-cycles and single-cycles only. Now if $(xy)\alpha \notin Aut(G)$, then only the identity automorphism fixes the set and by Lemma 2.17 $\{x, y\}$ is a distinguishing class of G. Therefore, we can color nodes x and y red and the remaining nodes blue and this will be a distinguishing coloring of G. Else, $(xy)\alpha \in Aut(G)$. This will break down into cases.

We will show that α must contain at least one transposition of nodes, which must be a neighbour-non-neighbour pair of the nodes x and y. This is defined to be a pair of nodes (d_1, d_2) say, such that, $(x, d_1) \in E(G)$, $(x, d_2) \notin E(G)$, $(y, d_2) \in E(G)$ and $(y, d_1) \notin E(G)$. Here, either $(xd_1)(y)\beta_1 \in Aut(G)$ or not. In the second case we will show that $\{x, y, d_1\}$ forms a distinguishing class. The first case will break down to two cases: Either $(xy)(d_1d_2) \in Aut(G)$ or $(xy)(d_1d_2)(d_3d_4)\beta_2 \in Aut(G)$. For the former situation, we will show that the graph is not

2-color distinguishable. The case of $(xy)(d_1d_2)(d_3d_4)\beta_2 \in Aut(G)$, breaks down into two further cases. Either there exists a transposition as part of this permutation which is also a neighbour-non-neighbour pair of x and y and is not equal to (d_1, d_2) or there exists a neighbour-non-neighbour pair of d_1 and d_2 in this permutation. We examined all of the possibilities in these cases which sometimes lead to $\{x, y, d_1\}$ being a distinguishing class or $\{x, y, d_1, d_3\}$ being a distinguishing class. For the latter case of a 4-element distinguishing class, the help of several technical lemmas was needed to show that no non-identity permutation leaves this set invariant. $\square$

3 Open problems and future works

1. For $Det(G) \geq 3$, can $\rho(G)$ be arbitrarily far from G?

2. We proved that for a graph G with $D(G) = 2$ and $Det(G) = 2$, that $\rho(G) = 2, 3$ or 4. It is easy to give examples for $\rho(G)$ being 2 or 3, however we could not find an example where $\rho(G) = 4$. Find such an example or show that one does not exist.

3. Classify graphs which are distinguishably equivalent. For example, $Aut(G, V(G)_l) = \{e\}$, where $e = (1)\dots(n)$ for some l labelling, which labels the nodes $\{1 \dots n\}$, corresponds to graphs which are the asymmetric graphs on n nodes. Hence, while a graph and its complement are always distinguishably equivalent, many more possibilities exist, see also Example 2.4.

References

[1] Michael O. Albertson and Debra L. Boutin. Distinguishing geometric graphs. *J. Graph Theory*, 53(2):135–150, 2006.

[2] Michael O. Albertson and Debra L. Boutin. Using determining sets to distinguish kneser graphs. *Electron. J. Comb.*, 14(1), 2007.

[3] Michael O. Albertson and Karen L. Collins. Symmetry breaking in graphs. *Electron. J. Comb.*, 3(1), 1996.

[4] László Babai. Asymmetric trees with two prescribed degrees. *Acta Math. Hung.*, 29(1-2):193–200, 1977.

[5] Bill Bogstad and Lenore J. Cowen. The distinguishing number of the hypercube. *Discrete Math.*, 283(1-3):29–35, 2004.

[6] Debra L. Boutin. Small label classes in 2-distinguishing labelings. *Ars Math. Contemp.*, 1(2):154–164, 2008.

[7] Debra L. Boutin. The cost of 2-distinguishing cartesian powers. *Electron. J. Combin.*, 20(1), 2013.

[8] Debra L. Boutin. The cost of 2-distinguishing selected kneser graphs and hypercubes. *J. Comb. Math. Comb. Comput.*, 85:161–171, 2013.

[9] Debra L. Boutin. The cost of 2-distinguishing hypercubes. *Discrete Math.*, 344(9), 2021.

[10] Debra L. Boutin. Paint cost and the frugal distinguishing number. *Art Discret. Appl. Math.*, 6(2), 2021.

[11] Melody Chan. The distinguishing number of the direct product and wreath product action. *J. Algebr. Comb.*, 24:331–345, 2006.

[12] David Erwin and Frank Harary. Destroying automorphisms by fixing nodes. *Discrete Math.*, 306(24):3244–3252, 2006.

[13] Tomotaka Fukuda, Seiya Negami, et al. 3-connected planar graphs are 2-distinguishable with few exceptions. *Yokohama Math. J.*, 54(2):1–11, 2008.

[14] Courtney R. Gibbons and Joshua D. Laison. Fixing numbers of graphs and groups. *Electron. J. of Combin.*, 16(1), 2009.

[15] Wilfried Imrich and Sandi Klavžar. Distinguishing cartesian powers of graphs. *J. Graph Theory*, 53(3):250–260, 2006.

[16] Wilfried Imrich, Simon M. Smith, Thomas W. Tucker, and Mark E. Watkins. Infinite motion and 2-distinguishability of graphs and groups. *J. Algebr. Comb.*, 41(1):109–122, 2015.

[17] Rafał Kalinowski and Monika Pilśniak. Distinguishing graphs by edge-colourings. *Eur. J. Comb.*, 45:124–131, 2015.

[18] Julianna S. Tymoczko. Distinguishing numbers for graphs and groups. *Electron. J. Combin.*, 11(1), 2004.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Near Triple Arrays

(Extended abstract)

Alexey Gordeev* Klas Markström† Lars-Daniel Öhman‡

Abstract

We introduce *near triple arrays* as binary row-column designs with at most two consecutive values for the replication numbers of symbols, for the intersection sizes of pairs of rows, pairs of columns and pairs of a row and a column. Near triple arrays form a common generalization of such well-studied classes of designs as triple arrays, (near) Youden rectangles and Latin squares.

We enumerate near triple arrays for a range of small parameter sets and show that they exist in the vast majority of the cases considered. As a byproduct, we obtain the first complete enumerations of 6×10 triple arrays on 15 symbols, 7×8 triple arrays on 14 symbols and 5×16 triple arrays on 20 symbols.

Next, we give several constructions for families of near triple arrays, and e.g. show that near triple arrays with 3 rows and at least 6 columns exist for any number of symbols. Finally, we investigate a duality between row and column intersection sizes of a row-column design, and covering numbers for pairs of symbols by rows and columns. These duality results are used to obtain necessary conditions for the existence of near triple arrays. This duality also provides a new unified approach to earlier results on triple arrays and balanced grids.

1 Introduction

The study of experimental designs that allow for eliminating the influence of multiple factors on an experiment began with the early works of Fisher and was developed further by among others Agrawal in the 1950:s and 60:s. Agrawal [2] introduced a class of experimental designs that would later be known as *triple arrays*. A triple array is a binary (no repeated symbols in any row or column), equireplicate (each symbol occurs the same number of times) array such that (RC) any row and column have a constant number of common symbols, (RR) any pair of rows has a constant number of common symbols, and (CC) any pair of columns has a constant

*Department of Mathematics and Mathematical Statistics, Umeå University, Sweden. E-mail: gordalserg@gmail.com. Supported by Kempe foundation grant JCSMK23-0058.

†Department of Mathematics and Mathematical Statistics, Umeå University, Sweden. E-mail: klas.markstrom@umu.se.

‡Department of Mathematics and Mathematical Statistics, Umeå University, Sweden. E-mail: lars-daniel.ohman@umu.se.

number of common symbols. Bagchi and Shah [3] showed that triple arrays are statistically optimal among all binary equireplicate arrays with regard to a large class of optimality criteria.

Unfortunately, the intersection conditions (RC), (RR), (CC) lead to some very restrictive constraints on the possible sizes of such designs. In the more recent literature, for example McSorley et al. [16], these types of designs have mainly been studied without any relaxations on the three conditions stated above. When relaxed versions have been studied, it has been done by completely removing one of the three conditions. For example, Bailey et al. [4] studied *sesqui arrays*, which are arrays with condition (CC) removed. This sacrifices the possibility of eliminating one factor affecting the experiment, while unfortunately not expanding the range of possible sizes very much.

In previous work by some of the present authors [10], relaxations of *Youden rectangles*, which likewise suffer from a restricted number of possible sizes, were studied. Youden rectangles can be seen as $r \times v$ triple arrays on v symbols. Note that for Youden rectangles conditions (RC) and (RR) hold trivially. In [10], *near Youden rectangles* were defined by relaxing condition (CC) to allow two consecutive column intersection sizes. It turns out that such arrays exist for a large majority of small sizes.

Our general approach, similar to that of [10], is to define *near triple arrays* by relaxing all three intersection properties to allow two values concentrated around the average intersection size, and also allowing two consecutive values as replication numbers. This definition includes Latin squares, (near) Youden rectangles and triple arrays, which all are near triple arrays with additional restrictions placed on the number of rows, columns and symbols: for example, Latin squares are precisely $n \times n$ near triple arrays on n symbols.

0	1	2	3	4	5	6	7	8
1	2	3	4	5	9	7	10	11
4	6	0	9	10	7	11	8	2
11	9	10	8	6	0	3	1	5

0	1	2	3	4	5
1	2	3	6	7	8
4	0	6	8	5	7
6	7	5	0	3	1

Figure 1: A 4×9 triple array on 12 symbols and a 4×6 near triple array on 9 symbols.

The *column design* of a triple array is a block design with points and blocks corresponding to, respectively, columns and symbols of the array, with a point appearing in a block when the corresponding column contains the corresponding symbol. The *row design* can be defined similarly. In the column design of a triple array, all blocks have the same size and each pair of points is covered by the same number of blocks, so it is a *balanced incomplete block design (BIBD)*. The column design of a near triple array belongs to a more general class of *maximally balanced maximally uniform designs (MBMUDs)*, which were studied by Bofill and Torras [5], or, in the equireplicate case, to the more well-known class of *regular graph designs*, introduced by John and Mitchell [12]. Our relaxation of triple arrays is thus also parallel to earlier efforts to generalize BIBDs. We give a necessary condition for the existence of the column design of a near triple array in Theorem 5.1 and use it to show that there are no near triple arrays for several families of parameter sets.

Near triple arrays will by definition minimize the deviation of the size of the pairwise column intersections from their average. In [10] it was shown that for equireplicate arrays with the same number of columns and symbols this will also make the covering numbers, i.e. the number of columns which contain a given pair of symbols, deviate as little as possible from their average. For near triple arrays in general this is no longer true. Given that many classical block designs are defined in terms of their covering numbers, it also becomes natural to consider arrays where the covering numbers are restricted to two consecutive values. When taking into account covering symbol pairs by both columns and rows, this leads to a family of arrays we call *near balanced grids*.

Near balanced grids are in turn a generalization of *balanced grids*, introduced by McSorley et al. [16]. Balanced grids are binary arrays in which the number of columns and rows containing a given pair of symbols is constant, i.e. does not depend on the choice of the pair. They showed that any $r \times c$ triple array on $r + c - 1$ symbols is a balanced grid, and later McSorley [14] showed that the converse also holds. We give an alternative proof of this result and show that it is in fact a special case of a more general relationship between near triple arrays and near balanced grids. More precisely, in Theorem 6.1 we show that for any given numbers of rows, columns and symbols, either near triple arrays on these parameters are precisely near balanced grids, or only one of these two classes of designs are possible for these parameters.

In addition to the results on the theory of near triple arrays, we also completely enumerate near triple arrays for a range of small parameter sets, and show that they exist for a vast majority of combinations of small sizes of an array and the number of symbols. We also extend previous enumerative results for triple arrays, (near) Youden rectangles, sesqui arrays and several other types of row-column designs.

The rest of the paper is structured as follows. In Section 2, we give the central definitions. In Section 3, we discuss the algorithm we used to enumerate near triple arrays as well as the results of the enumeration. In Section 4, we give general constructions and existence proofs. In Section 5, we study row and column designs of near triple arrays and derive non-existence conditions. In Section 6, we investigate the relationship between near triple arrays and near balanced grids. The complete proofs of all results stated below are available in the full version of this article [9].

2 Definitions and basic properties

An $r \times c$ *row-column design* on v symbols is a two-dimensional array with r rows, c columns and each cell filled with one of v symbols. It is *binary* if no symbol appears more than once in any row or column. We restrict our investigation to binary arrays with each of v symbols used at least once, that is, $\max(r, c) \leq v \leq rc$. To avoid rather trivial examples, we also generally assume $r, c \geq 3$.

For a real number x, denote $x^- := \lfloor x \rfloor$ and $x^+ := \lceil x \rceil$. The *average replication number* of a row-column design is $e := rc/v$. If e is an integer and every symbol occurs e times in the array, the row-column design is called *equireplicate* with *replication number* e. We will call a row-column design *near equireplicate* if e is not an integer and every symbol occurs in the array either e^- or e^+ times.

For a binary $r \times c$ row-column design on v symbols, we denote by λ_{rc}, λ_{rr}, λ_{cc} the average number of common symbols between a row and a column, two rows, and two columns, respectively. The next lemma can be proved via careful double counting.

Lemma 2.1. *For a binary (near) equireplicate $r \times c$ row-column design on v symbols,*

$$\lambda_{rc} = e + \frac{(e^+ - e)(e - e^-)}{e}, \quad \lambda_{rc}^- = e^-, \quad \lambda_{rc}^+ = e^+, \quad \lambda_{rr} = \frac{c(\lambda_{rc} - 1)}{r - 1}, \quad \lambda_{cc} = \frac{r(\lambda_{rc} - 1)}{c - 1}.$$

Definition 2.2. *An $(r \times c, v)$-near triple array is a binary (near) equireplicate $r \times c$ row-column design on v symbols in which (1) any row and column have either λ_{rc}^{-} or λ_{rc}^{+} common symbols, (2) any two rows have either λ_{rr}^{-} or λ_{rr}^{+} common symbols, (3) any two columns have either λ_{cc}^{-} or λ_{cc}^{+} common symbols.*

To avoid having to calculate the quantities involved in the definition explicitly, it will often be convenient to instead use the following alternative characterization of near triple arrays.

Proposition 2.3. *A binary $r \times c$ row-column design on v symbols in which, for some integers x, x_{rc}, x_{rr} and x_{cc}, (1) any symbol occurs x or $x+1$ times, (2) any row and column have either x_{rc} or $x_{rc}+1$ common symbols, (3) any two rows have either x_{rr} or $x_{rr}+1$ common symbols, (4) any two columns have either x_{cc} or $x_{cc}+1$ common symbols, is an $(r \times c, v)$-near triple array. Conversely, any near triple array satisfies all these conditions.*

Note that $(r \times c, v)$-triple arrays are precisely $(r \times c, v)$-near triple arrays with integer e, λ_{rc}, λ_{rr}, λ_{cc}. For a near triple array in general, each of λ_{rc}, λ_{rr} and λ_{cc} can either be or not be an integer. Lemma 2.1 implies that λ_{rc} is an integer precisely when the design is equireplicate.

3 Enumeration

For a complete description of our enumeration algorithm and the measures we took to verify the correctness of the computations, see the full version of this article [9]. The approach we used is often called *orderly generation*, and was introduced independently by Faradžev [7] and Read [18]. See chapter 4 of Kaski and Östergård's book [13] for an overview of such algorithms in a more general setting. Similar to many other design enumeration problems, the number of partial objects during our search often exceeds the number of complete near triple arrays by many orders of magnitude. This has been a major bottleneck in several recent enumeration efforts for other row column designs [10, 11]. We address this problem in two ways. First, a major advantage of our approach is that different branches of the search tree are completely independent. In particular, no additional cross-checks of the results to remove isomorphic copies of the same object are required, so there is no need to simultaneously store all partial objects obtained at any given step. This means that even though the number of partial objects may be huge, we are not limited by storage space constraints. The independence of different branches of the search tree also makes parallelization straightforward. Second, we generate near triple arrays by filling one cell at a time. This is contrary to a more common approach, when row-column designs, or, more generally, incidence matrices of combinatorial structures, are generated one row at a time. Making the steps of the generation procedure as small as possible allows us to detect and reject non-completable partial objects much sooner, reducing the total size of the search tree.

We have completely enumerated all near triple arrays on $r = 3, 4, 5, 6, 7$ rows and $c \leq 15$ columns with $rc \leq 50$, except for a few cases where we only established existence. In addition, we have been able to complete the enumeration of triple arrays for some parameter sets beyond this range and have determined that, up to isotopism (applying permutations to the sets of rows, columns and symbols), there are precisely 684782 $(7 \times 8, 14)$-triple arrays, 270119 $(6 \times 10, 15)$-triple arrays and 26804 $(5 \times 16, 20)$-triple arrays. We have determined that there are no near triple arrays on 33 parameter sets out of the 812 parameter sets considered. In every such case, we were able to find examples of arrays in which just one of the three intersection conditions was further relaxed to allow three consecutive values for the size of an intersection.

We have also used a modified version of our algorithm to independently verify and extend earlier results on the enumeration of (near) Youden rectangles, sesqui arrays and other types of row-column designs by Jäger et al. [10], [11].

The most symmetric triple array we found is the $(6 \times 10, 15)$-triple array with autotopism group size 720. This design was described by Nilson in [17], and it has the exceptional property that any pair of occurrences of a symbol lies in an *intercalate*, that is, a 2×2 Latin subsquare. We give two further rather symmetric examples in Figure 2, both of which have similar properties. In the $(6 \times 10, 15)$-triple array with autotopism group size 120, each cell belongs to exactly one intercalate, that is, 15 intercalates form a partition of the table. In the $(7 \times 8, 14)$-triple array with autotopism group size 168, any two occurrences of a single symbol lie in a 3×2 Latin subrectangle, and any pair of distinct symbols that appear together in some column likewise appear in one common 3×2 Latin subrectangle.

0	1	2	3	4	5	6	7	8	9
1	0	3	4	5	2	10	11	12	13
6	11	0	8	14	13	9	12	3	2
8	13	14	11	0	6	5	4	7	10
10	9	7	14	12	1	11	2	5	8
12	7	10	1	9	14	3	6	13	4

0	1	2	3	4	5	6	7
1	8	3	9	5	10	7	11
2	12	13	8	3	1	11	6
7	5	9	2	6	12	13	10
8	0	5	7	13	11	4	12
10	6	11	4	8	2	9	0
13	9	0	12	10	4	1	3

Figure 2: Triple arrays with large autotopism groups.

The algorithms used were implemented in C++, and the source code is available at [8]. The summary of all enumerative results is available in the full version of this article [9]. With some exceptions due to size restrictions, the data we generated is available at [1].

4 Constructions and existence

In this section we show that near triple arrays on 3 rows and at least 6 columns exist for any number of symbols. In the full version of this article [9] we give more explicit constructions, including a more direct construction of $(3 \times c, v)$-near triple arrays with $c \geq 6$, $v \geq 3c/2$, and constructions from Latin squares and (near) Youden rectangles.

Lemma 4.1. *There exists an $(r \times c, v)$-near triple array whenever $v \geq rc - c/2$.*

Proof sketch. There are so few repeated symbols that one just needs to distribute them in such a way that row-row intersections are balanced. □

According to our computational results, there are no $(3 \times 3, 6)$, $(3 \times 4, 6)$ or $(3 \times 5, 8)$-near triple arrays, but there are examples of $(3 \times c, v)$-near triple arrays in all considered cases with $c \geq 6$. Existence turns out to hold for all larger c as well.

Theorem 4.2. *There exist $(3 \times c, v)$-near triple arrays for any $c \geq 6$ and $v \geq c$.*

Proof sketch. Cases with $6 \leq c \leq 13$ form the base of induction, with the corresponding examples of arrays found by computer search. For $c \geq 14$, a $(3 \times c, v)$-near triple array for any v either exists due to Lemma 4.1, or can be constructed by concatenating two near triple arrays with parameters $(3 \times 6, 9)$ and $(3 \times (c - 6), v - 9)$, or $(3 \times 7, 7)$ and $(3 \times (c - 7), v - 7)$, which exist by the induction hypothesis. □

5 Row and column designs

For an integer n and a real m, define $S(n,m) := n\binom{m}{2} + n(m-m^-)(m^+-m)/2$, where $\binom{m}{2}$ for a possibly non-integer m is defined as $m(m-1)/2$. The next theorem gives a necessary condition for the existence of the column design of a near triple array, and thus for the existence of the near triple array itself. It implies non-existence of $(k \times (\binom{k}{2}-s), \binom{k}{2}+1)$-near triple arrays for positive integers k,s with $1 \le s \le \frac{(k-1)(k-2)}{2k}$, further examples of such families of parameter sets are available in the full version of the article [9].

Theorem 5.1. *Let T be an $(r\times c, v)$-near triple array with parameters $e, \lambda_{rc}, \lambda_{rr}, \lambda_{cc}$. Denote the average number of times a pair of symbols appears together in a column by $\mu_c := \binom{r}{2}c/\binom{v}{2} = \frac{e(r-1)}{v-1}$. Then $S(\binom{c}{2}, \lambda_{cc}) \ge S(\binom{v}{2}, \mu_c)$, with equality if and only if each pair of symbols is covered by μ_c^- or μ_c^+ columns.*

Proof sketch. One can count in two ways the number of times a pair of symbols appears together in a pair of columns and show that one of these counts is the left hand side of the inequality and the other is at least its right hand side. □

We have computationally showed that there are no $(4 \times 11, 13)$-near triple arrays. In Theorem 5.2, we generalize this non-existence result to a larger family of parameter sets.

Theorem 5.2. *For $r \ge 4$, there are no $(r \times (r(r-1)-1), r(r-1)+1)$-near triple arrays.*

Proof sketch. One can show by the pigeonhole principle that in such an array there would always be two columns with at least two common symbols. □

In Theorem 4.2, we proved that for $c \ge 6 = 3(3-1)$ there exist $(3 \times c, v)$-near triple arrays for any $v \ge c$. Theorem 5.2 shows that the bound $c \ge r(r-1)$ is actually a necessary condition for any $r \ge 4$ for the existence of $(r \times c, v)$-near triple arrays for all $v \ge c$.

6 Near balanced grids

For a binary $r \times c$ row-column design T on v symbols, let μ be the average number of times a pair of symbols appears together in a row or a column, that is

$$\mu := \frac{\binom{c}{2}r + \binom{r}{2}c}{\binom{v}{2}} = \frac{c(c-1)r + r(r-1)c}{v(v-1)} = \frac{e(r+c-2)}{v-1}.$$

T is called a *balanced grid* if μ is an integer and for each pair of symbols the number of rows and columns containing the pair is equal to μ. Balanced grids were introduced by McSorley et al. in [16]. In particular, they showed that any balanced grid is equireplicate. We define an $(r \times c, v)$-*near balanced grid* as a binary (near) equireplicate $r \times c$ row-column design on v symbols in which the number of times each pair of symbols appears in a row or a column is μ^- or μ^+. Clearly, a balanced grid is a near balanced grid.

Theorem 6.1. *Consider a $(r \times c, v)$ binary row column design, and define $S_{\mathrm{NBG}} := S(\binom{v}{2}, \mu)$ and $S_{\mathrm{NTA}} := S(\binom{c}{2}, \lambda_{cc}) + S(\binom{r}{2}, \lambda_{rr}) + S(rc, \lambda_{rc})$. Then (a) if $S_{\mathrm{NTA}} < S_{\mathrm{NBG}}$, then there are no $(r \times c, v)$-near triple arrays; (b) if $S_{\mathrm{NTA}} > S_{\mathrm{NBG}}$, then there are no $(r \times c, v)$-near balanced grids; (c) if $S_{\mathrm{NTA}} = S_{\mathrm{NBG}}$, then any $(r \times c, v)$-near triple array is an $(r \times c, v)$-near balanced grid and vice versa.*

Proof sketch. One can count in two ways the number of times a pair of symbols appears together in a pair of rows and/or columns and show that one of the counts is at least S_{NTA}, and is equal to S_{NTA} for near triple arrays, and the other count is at least S_{NBG}, and is equal to S_{NBG} for near balanced grids. $\square$

Theorem 6.1 implies, for example, that for $k \geq 3$ and $s \geq k-1$, $(k \times (k-1)s, ks)$-near triple arrays are $(k \times (k-1)s, ks)$-near balanced grids and vice versa, and that there are no $(k \times 2(k-1), 2k)$-near triple arrays for $k \geq 4$. We give further examples of such families of parameter sets in the full version of the article [9].

Corollary 6.2. *Any $(r \times c, v)$-triple array with $v > \max(r, c)$ has $v \geq r+c-1$. Any $(r \times c, v)$-balanced grid has $v \leq r+c-1$. Any $(r \times c, r+c-1)$-triple array is an $(r \times c, r+c-1)$-balanced grid and vice versa.*

Proof sketch. Define $S_{\text{TA}} := \binom{c}{2}\binom{\lambda_{cc}}{2} + \binom{r}{2}\binom{\lambda_{rr}}{2} + rc\binom{e}{2}$, $S_{\text{BG}} := \binom{v}{2}\binom{\mu}{2}$. It can be shown by elementary algebraic manipulations that $S_{\text{TA}} - S_{\text{BG}} = \frac{(v-(r+c-1))(r+c-2)(v-c)(v-r)e^2}{4(v-1)(r-1)(c-1)}$. In particular, if $v > \max(r, c)$, then the sign of $S_{\text{TA}} - S_{\text{BG}}$ coincides with the sign of $v-(r+c-1)$. The claims then follow from the facts that $S_{\text{NTA}} \geq S_{\text{TA}}$ with equality precisely for triple arrays and that $S_{\text{NBG}} \geq S_{\text{BG}}$ with equality precisely when μ is integer. $\square$

Corollary 6.2 corresponds to Theorems 3.2, 4.2, 6.1 in [16] and Theorem 2.5 in [14] combined. We note that previous proofs of the first two claims of Corollary 6.2 used linear algebra. Our approach to all three claims is uniform and thus more directly highlights a certain duality between triple arrays and balanced grids.

One parameter set with $S_{\text{NTA}} = S_{\text{NBG}}$, $(6 \times 6, 9)$, is of particular interest. Balanced grids with $r = c$ are also known as *binary pseudo-Youden designs*. McSorley and Phillips [15] determined that there are 696 non-isotopic $(6 \times 6, 9)$ binary pseudo-Youden designs. Our search likewise found 696 non-isotopic near triple arrays on this parameter set, and Theorem 6.1 explains this. This duality does not hold for binary pseudo-Youden designs in general: Cheng [6] gave a construction of $(\binom{4k+4}{2} \times \binom{4k+4}{2}, (4k+3)^2)$ binary pseudo-Youden designs whenever $4k+3$ is a prime of a prime power. Theorem 6.1 implies that this construction produces near triple arrays only when $k = 0$, that is, for the parameter set $(6 \times 6, 9)$.

Acknowledgements

The computational work was performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at High Performance Computing Center North (HPC2N).

References

[1] Near triple arrays, triple arrays, sesqui arrays, mono arrays and AO-arrays are available at `http://abel.math.umu.se/~klasm/Data/ARRAYDESIGNS/`.

[2] H. Agrawal. Some methods of construction of designs for two-way elimination of heterogeneity, 1. *Journal of the American Statistical Association*, 61(316):1153–1171, 1966.

[3] S. Bagchi and K. R. Shah. On the optimality of a class of row-column designs. *Journal of Statistical Planning and Inference*, 23(3):397–402, 1989.

[4] R. A. Bailey, P. J. Cameron, and T. Nilson. Sesqui-arrays, a generalisation of triple arrays. *The Australasian Journal of Combinatorics*, 71(3):427–451, 2018.

[5] P. Bofill and C. Torras. MBMUDs: a combinatorial extension of BIBDs showing good optimality behaviour. *Journal of Statistical Planning and Inference*, 124(1):185–204, 2004.

[6] C.-S. Cheng. A family of pseudo Youden designs with row size less than the number of symbols. *Journal of Combinatorial Theory, Series A*, 31(2):219–221, 1981.

[7] I. Faradžev. Constructive enumeration of combinatorial objects. In *Problèmes combinatoires et théorie des graphes*, pages 131–135, 1978.

[8] A. Gordeev. *The source code for the enumeration algorithm used in the article "Near triple arrays"*, 2025. Zenodo. `https://doi.org/10.5281/zenodo.14925326`.

[9] A. Gordeev, K. Markström, and L.-D. Öhman. Near Triple Arrays. *arXiv preprint arXiv:2503.07166*, 2025.

[10] G. Jäger, K. Markström, D. Shcherbak, and L.-D. Öhman. Small Youden rectangles, near Youden rectangles, and their connections to other row-column designs. *Discrete Mathematics & Theoretical Computer Science*, vol. 25:1(Combinatorics), 2023.

[11] G. Jäger, K. Markström, D. Shcherbak, and L.-D. Öhman. Enumeration of row-column designs. *arXiv preprint arXiv:2406.16444*, 2024.

[12] J. A. John and T. J. Mitchell. Optimal incomplete block designs. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 39(1):39–43, 1977.

[13] P. Kaski and P. R. J. Östergård. *Classification Algorithms for Codes and Designs.* Springer, Berlin, 2006.

[14] J. P. McSorley. Double arrays, triple arrays and balanced grids with $v = r + c - 1$. *Designs, Codes and Cryptography*, 37(2):313–318, 2005.

[15] J. P. McSorley and N. C. Phillips. Complete enumeration and properties of binary pseudo-Youden designs. *Journal of Statistical Planning and Inference*, 137(4):1464–1473, 2007.

[16] J. P. McSorley, N. C. K. Phillips, W. D. Wallis, and J. L. Yucas. Double arrays, triple arrays and balanced grids. *Designs, Codes and Cryptography*, 35(1):21–45, 2005.

[17] T. Nilson. Intercalates in double and triple arrays. *J. Comb. Des.*, 30(3):135–151, 2022.

[18] R. C. Read. Every one a winner or how to avoid isomorphism search when cataloguing combinatorial configurations. In *Algorithmic Aspects of Combinatorics*, volume 2 of *Annals of Discrete Mathematics*, pages 107–120. Elsevier, 1978.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A Christofides-based approach to the travelling salesman problem in the unit cube

(Extended abstract)

Alexey Gordeev*

Abstract

In 1992, Bollobás and Meir showed that for any n points in the k-dimensional unit cube $[0,1]^k$, $k \geq 2$, one can always find a tour $x_1, \ldots, x_n$ through these n points with $(\sum_{i=1}^{n} |x_i - x_{i+1}|^k)^{1/k} \leq c_k$, where $|x - y|$ is the Euclidean distance between x and y, $x_{n+1} = x_1$, and c_k is an absolute constant depending only on k. They further conjectured that the best possible constant for every $k \geq 2$ is $c_k = 2^{1/k}\sqrt{k}$. This is only known to be true for $k = 2$ due to Newman. Recently, Balogh, Clemen and Dumitrescu disproved the conjecture for $k = 3$ by showing that $c_3 \geq 2^{7/6}$. They also gave the best currently known bounds $c_k \leq 6.709(\frac{2}{3})^{1/k}\sqrt{k}$ and $c_k \leq 2.91\sqrt{k}(1 + o_k(1))$.

We prove that for any even-sized set of n points in $[0,1]^k$ one can always find a perfect matching $x_1y_1, \ldots, x_{n/2}y_{n/2}$ on these n points with $(\sum_{i=1}^{n/2} |x_i - y_i|^k)^{1/k} \leq 2^{1/k}\sqrt{2k}$. We combine this result with ideas from the famous Christofides algorithm, while also making several improvements to a ball packing argument used in the earlier results, improving the bounds further to $c_k \leq 5.059 \cdot (1.28)^{1/k}\sqrt{k}$ and $c_k \leq 2.65\sqrt{k}(1 + o_k(1))$.

1 Introduction

Throughout the text we assume $k \geq 2$. For a point $a = (a_1, \ldots, a_k) \in \mathbb{R}^k$ we denote by $|a| := \sqrt{\sum_{i=1}^{k} a_i^2}$ the Euclidean length of a. For a graph $G = (V, E)$ with $V \subseteq \mathbb{R}^k$ and its edge $e = ab \in E$ we denote by $|e| := |a - b|$ the Euclidean distance between its endpoints. Let

$$s_k(G) := \left(\sum_{e \in E} |e|^k \right)^{1/k}.$$

Let $X \subseteq [0,1]^k$ be a set of n points in the k-dimensional unit cube. Let $s_k^{\mathrm{HC}}(X) := \min_G s_k(G)$, where the minimum is over all Hamiltonian cycles $G = (X, E)$. Let

$$s_k^{\mathrm{HC}}(n) := \sup\{s_k^{\mathrm{HC}}(X) \mid X \subseteq [0,1]^k, |X| = n\}.$$

*Department of Mathematics and Mathematical Statistics, Umeå University, Sweden. E-mail: gordalserg@gmail.com. Supported by Kempe foundation grant JCSMK23-0058.

Finally, let $s_k^{\mathrm{ST}}(X)$, $s_k^{\mathrm{ST}}(n)$ and $s_k^{\mathrm{PM}}(X)$, $s_k^{\mathrm{PM}}(n)$ be defined similarly for, respectively, spanning trees and perfect matchings instead of Hamiltonian cycles, with $s_k^{\mathrm{PM}}(X)$ and $s_k^{\mathrm{PM}}(n)$ defined only for sets X of even size and for even n. By clustering points of X around two opposite vertices of $[0,1]^k$, we see that

$$s_k^{\mathrm{HC}}(n) \geq 2^{1/k}\sqrt{k}, \quad s_k^{\mathrm{ST}}(n) \geq \sqrt{k}, \quad s_k^{\mathrm{PM}}(n) \geq \sqrt{k}$$

for all $k \geq 2$ and $n \geq 2$. In 1992, Bollobás and Meir [3] proved that $s_k^{\mathrm{HC}}(n) \leq 9\left(\frac{2}{3}\right)^{\frac{1}{k}} \cdot \sqrt{k}$ and $s_k^{\mathrm{ST}}(n) \leq 3\sqrt{k}$. Note that these bounds do not depend on n. They further conjectured that $s_k^{\mathrm{HC}}(n) = 2^{1/k}\sqrt{k}$ for all $n \geq 2$ and $k \geq 2$. In the case $k = 2$ it has been shown to be true by Newman, see Problem 57 in [5].

Recently, Balogh, Clemen, and Dumitrescu [2] disproved this conjecture for $k = 3$ by presenting the set $X = \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\} \subseteq [0,1]^3$. The set X is a binary code of length 3 with minimum Hamming distance 2, so $s_k^{\mathrm{HC}}(X) = 2^{7/6} > 2^{1/3}\sqrt{3}$. They also improved the upper bounds on s_k^{HC} and s_k^{ST} and gave a bound on s_k^{PM} as a corollary:

$$\begin{aligned}
s_k^{\mathrm{HC}}(n) &\leq 3\sqrt{5}\left(\frac{2}{3}\right)^{1/k} \cdot \sqrt{k} \leq 6.709 \cdot \left(\frac{2}{3}\right)^{1/k} \cdot \sqrt{k},\\
s_k^{\mathrm{ST}}(n) &\leq \sqrt{5k} \leq 2.237 \cdot \sqrt{k},\\
s_k^{\mathrm{PM}}(n) &\leq 3\sqrt{5}\left(\frac{1}{3}\right)^{1/k} \cdot \sqrt{k} \leq 6.709 \cdot \left(\frac{1}{3}\right)^{1/k} \cdot \sqrt{k}.
\end{aligned}$$

Further, they gave asymptotic bounds for s_k^{HC} and s_k^{ST}, the second of which is optimal:

$$\begin{aligned}
s_k^{\mathrm{HC}}(n) &\leq 2.91 \cdot \sqrt{k}(1 + o_k(1)),\\
s_k^{\mathrm{ST}}(n) &\leq \sqrt{k}(1 + o_k(1)).
\end{aligned}$$

In the case of spanning trees, the bounds on $s_k^{\mathrm{ST}}(n)$ in [3, 2] are obtained by a ball packing argument: given a minimum spanning tree (in terms of the sum of Euclidean lengths of all edges) on a set of n points, one can put a ball at the center of each edge of the tree with a radius proportional to the length of the edge, such that all these balls are disjoint and also all covered by one large ball. The resulting inequality on the sum of volumes of these balls then implies a bound on $s_k^{\mathrm{ST}}(n)$.

We improve on this argument in two ways. First, we show that replacing the balls with half-balls oriented towards the center of the unit cube allows to cover all of them by one ball of a much smaller radius, which already improves the bound to $s_k^{\mathrm{ST}}(n) \leq 2^{1/k} \cdot 2\sqrt{k}$. Second, by carefully moving the balls away from the center and toward one of the endpoints of their edges, we can slightly increase the radius of each ball so that they are still disjoint. This further improves the bound on $s_k^{\mathrm{ST}}(n)$.

Bounds on $s_k^{\mathrm{HC}}(n)$ are obtained in [3, 2] from bounds on $s_k^{\mathrm{ST}}(n)$ by using the following idea, first developed by Sekanina [6, 7]: one can convert an Eulerian tour of a minimum spanning tree into a Hamiltonian cycle by contracting some paths into edges in such a way that each edge of the resulting cycle corresponds to a contraction of at most three edges of the tree, and each edge of the tree is used at most twice. We note that in the classic 2-approximation algorithm for the metric travelling salesman problem a solution is similarly obtained by contracting paths in an Eulerian tour of a minimum spanning tree.

The famous Christofides algorithm [4] achieves $\frac{3}{2}$-approximation for the metric travelling salesman problem by instead of a spanning tree using the union of a minimum spanning tree and a minimum matching on the set of vertices with odd degree in the tree. In the present article, we use a similar approach to improve the bound on $s_k^{\mathrm{HC}}(n)$. We first show a better bound on $s_k^{\mathrm{PM}}(n)$ by adapting the ball packing argument to the case of perfect matchings. Then we prove an analogue of the result of Sekanina for the Eulerian tour of the union of a spanning tree and a matching. Combining all new results together, we improve the bounds to

$$
\begin{aligned}
s_k^{\mathrm{PM}}(n) &\leq 2^{1/k}\sqrt{2k} && \leq 1.415 \cdot 2^{1/k} \cdot \sqrt{k} && \text{(Theorem 3.4)},\\
s_k^{\mathrm{ST}}(n) & && \leq 1.823 \cdot 2^{1/k} \cdot \sqrt{k} && \text{(Theorem 5.2)},\\
s_k^{\mathrm{HC}}(n) & && \leq 5.059 \cdot (1.28)^{1/k} \cdot \sqrt{k} && \text{(Theorem 6.1)},\\
s_k^{\mathrm{HC}}(n) & && \leq 2.65 \cdot \sqrt{k}(1+o_k(1)) && \text{(Theorem 6.2)}.
\end{aligned}
$$

The rest of the paper is structured as follows. In Section 2, we give the improved half-ball packing argument. In Section 3, we apply it to the case of perfect matchings. In Section 4, we derive the bound on s_k^{HC} from bounds on s_k^{ST} and s_k^{PM}. In Section 5, we improve the ball packing for spanning trees. Finally, in Section 6, we prove new bounds on $s_k^{\mathrm{HC}}(n)$.

2 Half-ball packing

Lemma 2.1. *Let $u, v \in \left[-\frac{1}{2}, \frac{1}{2}\right]^k$, $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq \sqrt{\alpha(1-\alpha)}$. Then*

$$\sqrt{|\alpha u + (1-\alpha)v|^2 + \beta^2 |u-v|^2} \leq \frac{\sqrt{k}}{2},$$

and this inequality is the best possible.

Proof sketch. One can show that

$$|\alpha u + (1-\alpha)v|^2 + \beta^2|u-v|^2 = p\,|w|^2 + (1-p)|u|^2$$

for some $w \in \left[-\frac{1}{2}, \frac{1}{2}\right]^k$ and $0 \leq p \leq 1$. This implies

$$|\alpha u + (1-\alpha)v|^2 + \beta^2|u-v|^2 \leq (p + (1-p)) \sum_{i=1}^{k} \frac{1}{4} = \frac{k}{4}. \qquad \square$$

Lemma 2.2. *Let $G = (V, E)$ be a graph with $V \subseteq [0,1]^k$, let $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq \sqrt{\alpha(1-\alpha)}$. For each edge $e = ab \in E$ let $m_e := \alpha a + (1-\alpha)b$ and*

$$H_e := \{w \in \mathbb{R}^k \mid |w - m_e| < \beta|a-b|,\ (w, m_e) < |m_e|^2\},$$

i.e. H_e is the open half-ball of radius $\beta|a-b|$ centered at m_e and directed towards the origin. If the half-balls $\{H_e\}_{e\in E}$ are pairwise disjoint, then

$$s_k(G) \leq 2^{\frac{1}{k}} \frac{\sqrt{k}}{2\beta}.$$

Proof. For any $w \in H_e$, $w = m_e + z$ for some $z \in \mathbb{R}^k$ with $|z| < \beta|a-b|$ and $(m_e + z, m_e) < (m_e, m_e)$, i.e. $(z, m_e) < 0$. It follows that

$$|w| = \sqrt{|m_e + z|^2} = \sqrt{|m_e|^2 + |z|^2 + 2(z, m_e)} < \sqrt{|m_e|^2 + \beta^2|a-b|^2}.$$

Thus, by Lemma 2.1, for each $e \in E$ we have $H_e \subseteq B$, where B is the open ball of radius $\frac{\sqrt{k}}{2}$ centered at the origin. Since $\{H_e\}_{e\in E}$ are pairwise disjoint, the sum of their volumes is at most the volume of B. Denoting by V_k the volume of the k-dimensional unit ball, we have

$$\sum_{e\in E} \frac{V_k}{2}\beta^k|e|^k \leq V_k \left(\frac{\sqrt{k}}{2}\right)^k, \text{ so } s_k(G) \leq 2^{\frac{1}{k}}\frac{\sqrt{k}}{2\beta}. \qquad \square$$

As we can see, by using half-balls instead of balls, we lose just half of their volume but are now able to cover all of them by a ball of radius $\frac{\sqrt{k}}{2}$ instead of $\frac{\sqrt{5k}}{4}$ in Lemma 2.4 in [2].

3 Perfect matchings

Lemma 3.1. *Let $a, b, c, d \in \mathbb{R}^k$, $0 \leq \alpha \leq 1$, $m_{ab} = \alpha a + (1-\alpha)b$ and $m_{cd} = \alpha c + (1-\alpha)d$. Then*

$$|m_{ab} - m_{cd}|^2 = \alpha^2|a-c|^2 + (1-\alpha)^2|b-d|^2 + \alpha(1-\alpha)\left(|a-d|^2 + |b-c|^2 - |a-b|^2 - |c-d|^2\right).$$

Proof. Both sides are equal to

$$\alpha^2(|a|^2 + |c|^2) + (1-\alpha)^2(|b|^2 + |d|^2) - 2\alpha^2(a,c) - 2(1-\alpha)^2(b,d) \\ - 2\alpha(1-\alpha)\left((a,d) + (b,c) - (a,b) - (c,d)\right). \qquad \square$$

By setting $\alpha = \frac{1}{2}$, we get the classical formula for the length of a bimedian of a tetrahedron (see, for example, page 56 in [1]):

Corollary 3.2. *For any $a, b, c, d \in \mathbb{R}^k$,*

$$\left|\frac{a+b}{2} - \frac{c+d}{2}\right|^2 = \frac{1}{4}\left(|a-c|^2 + |b-d|^2 + |a-d|^2 + |b-c|^2 - |a-b|^2 - |c-d|^2\right).$$

We use this formula to show that if one places the ball of radius $\frac{|e|}{2\sqrt{2}}$ on the center of each edge e of a minimum perfect matching, these balls turn out to be disjoint. Curiously, the matching is required to minimize not the sum of edge lengths but the sum of squares of edge lengths.

Lemma 3.3. *Let $V \subseteq \mathbb{R}^k$ with $|V| = n$ even, and let $M = (V, E)$ be a perfect matching with $\sum_{e\in E}|e|^2$ minimal possible. For each edge $e = ab \in E$, let B_e be the open ball of radius $\frac{|a-b|}{2\sqrt{2}}$ centered at $\frac{a+b}{2}$. Then the balls $\{B_e\}_{e\in E}$ are pairwise disjoint, and the factor $\frac{1}{2\sqrt{2}}$ is as large as possible.*

Proof. Let $e = ab, e' = cd \in E$ be any pair of distinct edges of the matching M. Since $\sum_{e\in E}|e|^2$ is minimal over all perfect matchings on V, we have

$$|a-c|^2 + |b-d|^2 \geq |a-b|^2 + |c-d|^2 \text{ and } |a-d|^2 + |b-c|^2 \geq |a-b|^2 + |c-d|^2.$$

Then, by Corollary 3.2 and the Cauchy-Schwarz inequality, the distance between the centers of the balls B_e and $B_{e'}$ is at least

$$\left|\frac{a+b}{2}-\frac{c+d}{2}\right|^2 \geq \frac{1}{4}\left(|a-b|^2+|c-d|^2\right) \geq \frac{1}{8}\left(|a-b|+|c-d|\right)^2,$$

thus

$$\left|\frac{a+b}{2}-\frac{c+d}{2}\right| \geq \frac{|a-b|}{2\sqrt{2}}+\frac{|c-d|}{2\sqrt{2}}.$$

The inequality is tight if $e = ab$ and $e' = cd$ are two opposite edges of a regular tetrahedron. $\square$

Theorem 3.4. *For any $V \subseteq [0,1]^k$ with $|V| = n$ even there exists a perfect matching $M = (V, E)$ with*

$$s_k(M) \leq 2^{\frac{1}{k}}\sqrt{2k}.$$

Proof. Let $M = (V, E)$ be a perfect matching on V which minimizes $\sum_{e\in E}|e|^2$. Due to Lemma 3.3, the half-balls $\{H_e\}_{e\in E}$ from Lemma 2.2 with $G = M$, $\alpha = \frac{1}{2}$, $\beta = \frac{1}{2\sqrt{2}}$ are disjoint, thus

$$s_k(M) \leq 2^{\frac{1}{k}}\frac{\sqrt{k}}{2/(2\sqrt{2})} = 2^{\frac{1}{k}}\sqrt{2k}.$$ $\square$

4 A Christofides-based approach

The k-th power of a (multi)graph $G = (V, E)$ is the (multi)graph $G^k := (V, E_k)$ with each path consisting of at most k edges in G corresponding to an edge in E_k connecting the endpoints of that path.

Lemma 4.1. *Let G be an Eulerian (multi)graph on $n \geq 3$ vertices which can be decomposed into a spanning tree T and a matching M, i.e. $E_G = E_T \sqcup E_M$. Then G^3 contains a Hamiltonian cycle H such that each edge of G is used exactly once by H, and, moreover, each edge of H uses at most 2 edges of T and at most 1 edge of M.*

Proof sketch. Fix any vertex as the root of T and then split an Eulerian cycle $e_1, \ldots, e_m$ of G into paths by making a cut between e_m and e_1, and also making a cut between e_i and e_{i+1} if the common vertex of e_i and e_{i+1} is the child in the rooted tree of one of their other endpoints. Contracting the resulting paths into edges gives the desired H. $\square$

Lemma 4.2. *Let G be an Eulerian (multi)graph on a finite set of points $V \subseteq \mathbb{R}^k$, $|V| = n \geq 2$, which can be decomposed into a spanning tree T and a matching M, i.e. $E_G = E_T \sqcup E_M$. Let $s_k(T) \leq c_T$ and $s_k(M) \leq c_M$. Then there exists a Hamiltonian cycle H on V such that*

$$s_k(H) \leq \left(\frac{c_T + c_M}{2c_T + c_M}\right)^{\frac{1}{k}} \cdot (2c_T + c_M).$$

Proof sketch. Bound the length of every edge of H from Lemma 4.1 by the sum of lengths of corresponding edges of G and use Hölder's inequality. $\square$

5 Improved ball packing for spanning trees

Lemma 5.1. *Let $V \subseteq \mathbb{R}^k$ with $|V| = n$, let $T = (V, E)$ be a spanning tree with $\sum_{e \in E} |e|$ minimal possible, let $\frac{1}{2} < \alpha < 1$ and*

$$\beta(\alpha) := \begin{cases} \frac{\sqrt{3\alpha^2 - 3\alpha + 1}}{2}, & \frac{1}{2} < \alpha < \frac{\sqrt{5}-1}{2}, \\ (1-\alpha)\sqrt{\frac{\alpha(2\alpha-1)}{3\alpha^2-3\alpha+1}}, & \frac{\sqrt{5}-1}{2} \le \alpha < 1. \end{cases}$$

Fix any vertex $r \in V$ to be the root of T. For each edge $e = ab \in E$, where b is the parent of a, let B_e be the open ball of radius $\beta(\alpha)|a - b|$ centered at $m_e := \alpha a + (1 - \alpha)b$. Then the balls $\{B_e\}_{e \in E}$ are pairwise disjoint, and function $\beta(\alpha)$ is the best possible.

Proof sketch. For any pair of distinct edges e_1 and e_2, at least one of the balls B_{e_1}, B_{e_2} is closer to the endpoint of its edge that is farther away from the root of T. The result follows by a careful application of Lemma 3.1 and Hölder's inequality. □

Theorem 5.2. *For any $V \subseteq [0, 1]^k$ with $|V| = n$ there exists a spanning tree $T = (V, E)$ with*

$$s_k(T) \le \frac{1}{2(1-\alpha_0)}\sqrt{\frac{3\alpha_0^2 - 3\alpha_0 + 1}{\alpha_0(2\alpha_0 - 1)}} \cdot 2^{\frac{1}{k}}\sqrt{k} \le 1.823 \cdot 2^{\frac{1}{k}}\sqrt{k},$$

where $\alpha_0 = 0.645\ldots$ is the largest of two real roots of the polynomial $12\alpha^4 - 21\alpha^3 + 17\alpha^2 - 7\alpha + 1$.

Proof. Let $T = (V, E)$ be a spanning tree which minimizes $\sum_{e \in E} |e|$. Due to Lemma 5.1, the half-balls $\{H_e\}_{e \in E}$ from Lemma 2.2 with $G = T$, $\frac{1}{2} < \alpha < 1$ and $\beta = \beta(\alpha)$ from Lemma 5.1, are disjoint, thus

$$s_k(T) \le 2^{\frac{1}{k}} \frac{\sqrt{k}}{2\beta(\alpha)}.$$

The function $\beta(\alpha)$ achieves its maximum on $\frac{1}{2} < \alpha < 1$ at α_0 (polynomial $12\alpha^4 - 21\alpha^3 + 17\alpha^2 - 7\alpha + 1$ is the numerator of the derivative of $\beta(\alpha)$) and $\alpha_0 > \frac{\sqrt{5}-1}{2}$. □

6 Hamiltonian cycles

Theorem 6.1. *For any $V \subseteq [0, 1]^k$ with $|V| = n$ there exists a Hamiltonian cycle $H = (V, E)$ with*

$$s_k(H) \le \left(\frac{c + \sqrt{2}}{2c + \sqrt{2}}\right)^{\frac{1}{k}} \cdot (2c + \sqrt{2}) \cdot 2^{\frac{1}{k}}\sqrt{k} \le (0.64)^{\frac{1}{k}} \cdot 5.059 \cdot 2^{\frac{1}{k}}\sqrt{k},$$

where

$$c = \frac{1}{2(1-\alpha_0)}\sqrt{\frac{3\alpha_0^2 - 3\alpha_0 + 1}{\alpha_0(2\alpha_0 - 1)}} \le 1.823,$$

$\alpha_0 = 0.645\ldots$ is the largest of two real roots of the polynomial $12\alpha^4 - 21\alpha^3 + 17\alpha^2 - 7\alpha + 1$.

Proof. Let $T = (V, E_T)$ be a spanning tree from Theorem 5.2, and let $V_{odd} \subseteq V$ be the set of vertices that have odd degree in T. Let $M = (V_{odd}, E_M)$ be a perfect matching from Theorem 3.4. Applying Lemma 4.2 to the Eulerian (multi)graph $G = (V, E_T \sqcup E_M)$ gives the desired bound. □

Theorem 6.2. *For any $V \subseteq [0,1]^k$ with $|V| = n$ there exists a Hamiltonian cycle $H = (V, E)$ with*

$$s_k(H) \leq 2.65 \cdot \sqrt{k}(1 + o_k(1)).$$

Proof sketch. Repeat the proof of Theorem 1.6 in [2], but instead of Lemma 2.1 from [2] use the new Lemma 5.1. Then the constant

$$3 \cdot \frac{4}{\sqrt{2\pi e}} \leq 2.91$$

gets replaced with (α_0 is the same as in Theorem 6.1)

$$3 \cdot \frac{1}{1-\alpha_0} \sqrt{\frac{3\alpha_0^2 - 3\alpha_0 + 1}{\alpha_0(2\alpha_0 - 1)2\pi e}} \leq 2.65. \qquad \square$$

References

[1] N. Altshiller-Court. *Modern Pure Solid Geometry.* Macmillan, 1935.

[2] J. Balogh, F. C. Clemen, and A. Dumitrescu. On a Traveling Salesman Problem for Points in the Unit Cube. *Algorithmica*, 86(9):3054–3078, 2024.

[3] B. Bollobás and A. Meir. A travelling salesman problem in the k-dimensional unit cube. *Operations Research Letters*, 11(1):19–21, 1992.

[4] N. Christofides. Worst-case analysis of a new heuristic for the traveling salesman problem. *Report 388, Graduate School of Industrial Administration, Carnegie Mellon University*, 1976.

[5] D. J. Newman. *A Problem Seminar.* Springer, New York, 1982.

[6] M. Sekanina. On an ordering of the vertices of a graph. *Časopis pro pěstování matematiky*, 088(3):265–282, 1963.

[7] M. Sekanina. On an Algorithm for Ordering of Graphs. *Canadian Mathematical Bulletin*, 14(2):221–224, 1971.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On almost Gallai colourings in complete graphs

(Extended abstract)

Alexandr Grebennikov* Letícia Mattos† Tibor Szabó‡

Abstract

For $t \in \mathbb{N}$, we say that a colouring of $E(K_n)$ is *almost t-Gallai* if no two rainbow t-cliques share an edge. Motivated by a lemma of Berkowitz on bounding the modulus of the characteristic function of clique counts in random graphs, we study the maximum number $\tau_t(n)$ of rainbow t-cliques in an almost t-Gallai colouring of $E(K_n)$. For every $t \geq 4$, we show that $n^{2-o(1)} \leq \tau_t(n) = o(n^2)$. For $t = 3$, surprisingly, the behaviour is substantially different. Our main result establishes that

$$\left(\frac{1}{2} - o(1)\right) n \log n \leq \tau_3(n) = O\big(n^{\sqrt{2}} \log n\big),$$

which gives the first non-trivial improvements over the simple lower and upper bounds. Our proof combines various applications of the probabilistic method and a generalisation of the edge-isoperimetric inequality for the hypercube.

1 Introduction

A colouring of the edges of a graph is called *Gallai colouring* if it admits no *rainbow triangle*, i.e. a triangle whose edges have pairwise distinct colours. This term was introduced by Gyárfás and Simonyi [17] due to the close connection of these colourings to the seminal work of Gallai [12] on comparability graphs, where he obtained a structural classification of all colourings of the complete graph that avoid rainbow triangles. Since its introduction, Gallai colourings have garnered significant attention and have been explored in various contexts, including Ramsey-type problems [6, 7, 10, 15, 16], extremal graph theory [3, 8, 9], and graph entropy [20].

We introduce a natural extension of Gallai colourings, in which rainbow triangles, or more generally rainbow t-cliques, are allowed if they are edge-disjoint. We will then be interested in maximising their number in this scenario. We describe our original motivation for this concept in Theorem 3.

*Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria. E-mail: aleksandr.grebennikov@ist.ac.at.

†Institut für Informatik, Universität Heidelberg, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany. E-mail: mattos@uni-heidelberg.de.

‡Institut für Mathematik, Freie Universität Berlin, Arnimallee 6, D-14195 Berlin, Germany. E-mail: szabo@math.fu-berlin.de.

Yet we find the arising extremal problem attractive in its own right, making it the main focus of the current paper.

For $t, n \in \mathbb{N}$ and $t \geq 3$, we say that a function $c : E(K_n) \to \mathbb{N}$ is an *almost t-Gallai colouring* if no two rainbow t-cliques in it share an edge. Here a clique is called *rainbow* if its edges have pairwise distinct colours. We denote by $\tau_t(n)$ the maximum number of rainbow t-cliques that an almost t-Gallai colouring can contain. Observe that $\tau_t(n) \geq \lfloor \frac{n}{t} \rfloor$, as one may take $\lfloor \frac{n}{t} \rfloor$ vertex-disjoint t-cliques, each coloured arbitrarily in a rainbow fashion, and colour all the remaining edges with the same (arbitrary) colour. The trivial upper bound on $\tau_t(n)$ is $M(n,t)$: the maximum number of edge-disjoint t-cliques in K_n. Note that $M(n,t)$ is at most $\binom{n}{2}/\binom{t}{2}$, while a classical result of Rödl [23] on the Erdős–Hanani conjecture assures that, for fixed t, $M(n,t) = (1+o(1))\binom{n}{2}/\binom{t}{2}$.

Our first theorem shows that for $t \geq 4$ neither of these simple bounds is tight. In particular, we show that $\tau_t(n)$ is subquadratic and determine its behaviour up to a factor of $n^{o(1)}$.

Theorem 1. *We have*

(i) $\tau_t(n) = o(n^2)$ *for* $t \geq 3$, *and*

(ii) $\tau_t(n) \geq n^{2-o(1)}$ *for* $t \geq 4$.

Moreover, the construction for the lower bound uses only $\binom{t}{2}$ *colours.*

For the upper bound, we use the Graph Removal Lemma, proved independently by Alon, Duke, Lefmann, Rödl and Yuster [1], and Füredi [11]. For the lower bound, we employ the construction of Kovács and Nagy [21] based on the $(3,h)$-gadget-free sets introduced by Alon and Shapira [2], which in turn relies on a variant of the large 3-AP-free sets of Behrend.

Interestingly, for $t = 3$ the lower bound construction fails to work. In our main theorem we show that this is not a coincidence and improve the upper bound by a polynomial factor. We also improve the trivial linear lower bound by a logarithmic factor. All logarithms in this paper are in base 2.

Theorem 2. *We have*

$$\left(\frac{1}{2} - o(1)\right) \cdot n \log n \leq \tau_3(n) \leq O(n^{\sqrt{2}} \log n).$$

Moreover, the construction for the lower bound uses only three colours.

For our lower bound we construct an almost 3-Gallai colouring of K_n containing $(1/2-o(1))n\log n$ rainbow triangles by embedding one of the colours into the hypercube graph of dimension $\log n$. For the upper bound we show that in any almost 3-Gallai colouring of K_n there are $O(n^{\sqrt{2}} \log n)$ rainbow triangles. The proof combines various applications of the probabilistic method and a generalisation of the edge-isoperimetric inequality for the hypercube due to Bernstein [5], Harper [18], Hart [19] and Lindsey [22].

We believe that our lower bound lies closer to the true value of $\tau_3(n)$, and therefore we conjecture that the stronger upper bound of $O(n \log n)$ should hold.

Remark. It is worth noting that when the host graph is not complete, the maximum number of rainbow triangles in an almost 3-Gallai colouring can differ significantly from the bounds in Theorem 2. For instance, consider the tripartite graph H of Ruzsa and Szemerédi [24], on n vertices with $n^{2-o(1)}$ edges such that each edge appears in exactly one triangle. Colouring the edges of every triangle of H in three different colours, we obtain an almost 3-Gallai coloring of H. Indeed, all the triangles in H are rainbow and also pairwise edge-disjoint.

Finally, we demonstrate an application of Theorems 1 and 2, which in fact was our original motivation for the concept of almost Gallai colourings. We obtain an upper bound on the modulus of the characteristic function of clique counts in the binomial random graph $G(n,p)$, for constant p. This improves a result of Berkowitz [4, Lemma 18].

Theorem 3. *Let $p \in (0,1)$ and $t \geq 4$ be constants, and let X_t be the number of t-cliques in $G(n,p)$. Then, for $s \in [-\pi, \pi]$ we have*

$$|\mathbb{E}\left(e^{isX_3}\right)| \leq \exp\left(-\Omega(s^2 n \log n)\right)| \qquad \text{and} \qquad |\mathbb{E}\left(e^{isX_t}\right)| \leq \exp\left(-\Omega(s^2 n^{2-o(1)})\right)$$

for some $c(t) > 0$, for all $t \geq 4$.

This theorem implies that $|\mathbb{E}\left(e^{isX_t}\right)|$ is exponentially small for $s \geq n^{-1+\varepsilon}$ and $t \geq 4$. Such estimates (combined with estimates for other ranges of s obtained by different methods, see [4]) are commonly used to prove anticoncentration results for the random variable X_t. The idea of the proof is to combine our constructions for the lower bounds on $\tau_t(n)$ given by Theorems 1 and 2 and Berkowitz's decoupling trick [4].

2 Sketch of proof – Theorem 2

From this point onward, we will refer to almost 3-Gallai colourings simply as almost Gallai colourings. Note that the definition of an almost Gallai colouring places no restriction on the number of colours used. Let $g(n)$ denote the maximum number of rainbow triangles in an almost Gallai 3-colouring of K_n. One can show that for every $n \in \mathbb{N}$ we have

$$g(n) \leq \tau_3(n) \leq \frac{9}{2} g(n).$$

Thus, it suffices to provide upper and lower bounds for $g(n)$.

2.1 The lower bound

In this sketch we consider only the case when $n = 2^m + m$, for some $m \in \mathbb{N}$. Divide the vertex set of K_n into two parts: R, indexed by the elements of $[m] := \{1, \ldots, m\}$, and L, indexed by $\{0,1\}^m$, the set of 0-1 vectors of length m. Then we colour the edges of K_n as follows:

1. For $u \in L = \{0,1\}^m$ and $i \in R = [m]$, colour the edge ui blue if $u_i = 1$ and red otherwise.
2. For $u, v \in L$, colour the edge uv green if the Hamming distance $\sum_{i=1}^m |u_i - v_i|$ between them equals 1.
3. Colour all the remaining edges red.

Figure 1 below illustrates this colouring for $m = 2$, where there are $n = 6$ vertices.

One can show that our colouring is almost 3-Gallai and has exactly one rainbow triangle for each edge on the green Hamming cube in L, which implies that

$$\tau_3(n) \geq m2^{m-1} = (1/2 + o(1))n \log n.$$

The general case can be handled using arguments analogous to those employed in the proof of the edge-isoperimetric inequality for the hypercube. For a complete proof, see [14].

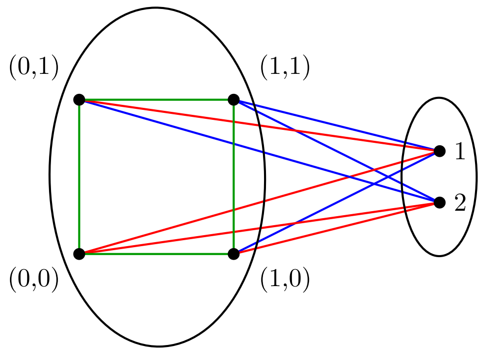


Figure 1: The colouring for $m = 2$. Only the edges of rainbow triangles are shown.

2.2 The upper bound

We will use the three colours red (R), blue (B) and green (G). When the colouring is clear from the context, we will denote the red, blue and green neighbourhoods of a vertex v by $N_R(v)$, $N_B(v)$ and $N_G(v)$, respectively. The sizes of these neighbourhoods will be denoted by $d_R(v)$, $d_B(v)$ and $d_G(v)$, respectively.

The first lemma gives an $O(n \log n)$ upper bound on the number of rainbow triangles which are not entirely contained in any of the three monochromatic neighbourhoods of a given vertex. If there is a vertex v such that none of its three colour degrees is "large", we can apply the induction hypothesis to each of these neighbourhoods and combine it with the bound of the lemma to obtain an efficient upper bound on the total number of rainbow triangles.

Lemma 4. *There exists an absolute constant $C_1 > 0$ such that the following holds for all $n \in \mathbb{N}$. For any almost Gallai 3-colouring $c : E(K_n) \to \{R, B, G\}$ of K_n and any vertex $v \in V(K_n)$, the number of rainbow triangles of c which are not fully contained in any of the sets $N_R(v)$, $N_B(v)$ or $N_G(v)$ is at most $C_1 n \log n$.*

Lemma 4 is effective if there is a vertex whose neighbourhoods in each of the three colours are not very large. Otherwise, every vertex has a "dominant colour" on its incident edges, and the inductive bound is not sufficient. In this case, it turns out that most vertices share the same dominant colour. Our second lemma provides an estimate on the number of rainbow triangles inside a set of vertices, which contains relatively few edges in one of the colours.

Lemma 5. *There exists an absolute constant $C_2 > 0$ such that the following holds for all $D, n \geq 3$. Let $c : E(K_n) \to \{R, B, G\}$ be an almost Gallai 3-colouring of K_n such that the green degree $d_G(v) \leq D$ for all $v \in K_n$. Then, the number of rainbow triangles in c is at most $C_2 n \sqrt{g(D) \log n}$.*

We define $F(x) = x^{\sqrt{2}} \log x$ for $x > 0$, and extend it continuously by setting $F(0) = 0$. Then it suffices to prove that $g(n) \leq CF(n)$ for some absolute constant C. We show this by induction on n.

Let $c : E(K_n) \to \{R, B, G\}$ be an arbitrary almost Gallai 3-colouring of K_n. We divide the proof into two cases based on whether every vertex in c has a dominant colour or not. Set $D = n^{2-\sqrt{2}} \leq n$ and first suppose that there is a vertex v such that its degree in each colour in c is at most $n - D$. Each rainbow triangle is either fully contained in one of the neighbourhoods $N_R(v)$, $N_B(v)$, and $N_G(v)$, or it is not. Hence, by Lemma 4, the colouring c contains at most

$$g(d_R(v)) + g(d_B(v)) + g(d_G(v)) + C_1 n \log n \leq F(D) + F(n - D) + C_1 n \log n \leq CF(n)$$

rainbow triangles. One can obtain the first inequality by using the induction hypothesis and the convexity of F. The second inequality follows by choosing C to be an absolute constant much larger than C_1.

Now we are in the situation when each vertex has a "dominant" colour, namely such that the degree in this colour is at least $n - D$. By using double-counting arguments, one may show that in this case there exists a colour and a set, say red and V_R, respectively, such that the following holds. The set V_R has size at least $n - 2D$, and every vertex in V_R has red as its dominant colour. The restriction of our colouring to V_R satisfies the assumption of Lemma 5, and hence we can conclude that in this case the number of rainbow triangles in c is at most

$$C_2\sqrt{C} \cdot nD^{\frac{\sqrt{2}}{2}} \log n + 2D^2 \leq CF(n).$$

The last inequality is satisfied by taking C to be an absolute constant sufficiently large with respect to C_2. For a complete proof, see [14].

References

[1] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, The algorithmic aspects of the Regularity Lemma, *J. Algorithms*, **16** (1994), 80–109.

[2] N. Alon and A. Shapira, Linear equations, arithmetic progressions and hypergraph property testing, *Theory Comput.*, **1** (2005), 177–216.

[3] J. Balogh and L. Li, The typical structure of Gallai colorings and their extremal graphs, *SIAM J. Discrete Math.*, **33** (2019), 2416–2443.

[4] R. Berkowitz, A local limit theorem for cliques in $G(n, p)$, arXiv:1811.03527 (2018).

[5] A. Bernstein, Maximally connected arrays on the n-cube, *SIAM J. Appl. Math.*, **15** (1967), 1485–1489.

[6] K. Cameron, J. Edmonds, and L. Lovász, A note on perfect graphs, *Period. Math. Hungar.*, **17** (1986), 173–175.

[7] F. Chung and R. Graham, Edge-colored complete graphs with precisely colored subgraphs, *Combinatorica*, **3** (1983), 315–324.

[8] J. de Oliveira Bastos, F. Siqueira Benevides, and Jie Han, The number of Gallai k-colorings of complete graphs, *J. Combin. Theory Ser. B*, **144** (2020), 1–13.

[9] J. Fox, A. Grinshpun, and J. Pach, The Erdős–Hajnal conjecture for rainbow triangles, *J. Combin. Theory Ser. B*, **111** (2015), 75–125.

[10] S. Fujita, C. Magnant, and K. Ozeki, Rainbow generalizations of Ramsey theory: a survey, *Graphs Combin.*, **26** (2010), 1–30.

[11] Zoltán Füredi, Extremal hypergraphs and combinatorial geometry, In *Proc. Internat. Congr. Math.* (1995), 1343–1352.

[12] T. Gallai, Transitiv orientierbare graphen, *Acta Math. Acad. Sci. Hungar.*, **18** (1967), 25–66.

[13] J. Gilmer and S. Kopparty, A local central limit theorem for triangles in a random graph, *Random Structures Algorithms*, **48** (2016), 732–750.

[14] A. Grebennikov, L. Mattos and T. Szabó, On almost Gallai colourings in complete graphs, arXiv:2503.17334 (2025).

[15] A. Gyárfás and G. Sárközy, Gallai colorings of non-complete graphs, *Discrete Math.*, **310** (2010), 977–980.

[16] A. Gyárfás, G. Sárközy, A. Sebő, and S. Selkow, Ramsey-type results for Gallai colorings, *J. Graph Theory*, **64** (2010), 233–243.

[17] A. Gyárfás and G. Simonyi, Edge colorings of complete graphs without tricolored triangles, *J. Graph Theory*, **46** (2004), 211–216.

[18] L. Harper, Optimal assignments of numbers to vertices, *J. Soc. Ind. Appl. Math.*, **12** (1964), 131–135.

[19] S. Hart, A note on the edges of the n-cube, Discrete Math., **14** (1976), 157–163.

[20] J. Korner and G. Simonyi, Graph pairs and their entropies: modularity problems, *Combinatorica*, **20** (2000), 227–240.

[21] B. Kovács and Z. Nagy, Multicolor Turán numbers II. – a generalization of the Ruzsa–Szemerédi theorem and new results on cliques and odd cycles, *J. Graph Theory*, **107** (2024), 629–641.

[22] J. Lindsey, Assignment of numbers to vertices, *Amer. Math. Monthly*, **71** (1964), 508–516.

[23] V. Rödl, On a packing and covering problem, *European J. of Comb.*, **6** (1985), 69–78.

[24] I. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, *Combinatorics (Proc. 5th Hungar. Colloq.)*, **18** (1978), 939–945.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Generalized Turán problem for directed cycles*

(Extended abstract)

Andrzej Grzesik† Justyna Jaworska‡ Bartłomiej Kielak§ Piotr Kuc¶
Tomasz Ślusarczyk‖

Abstract

For integers $k, \ell \geq 3$, let $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ denote the maximum number of directed cycles of length k in any oriented graph on n vertices which does not contain a directed cycle of length ℓ. We establish the order of magnitude of $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for every k and ℓ and determine its value (up to a lower order error term) when $k \nmid \ell$ and ℓ is large enough. Additionally, we calculate the value of $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for some other specific pairs (k, ℓ) showing that a diverse class of extremal constructions can appear for small values of ℓ.

The Turán problem [16] is a fundamental subject in extremal graph theory. In its simplest form, for a given graph F, it asks for the maximum number of edges in an n-vertex graph which is F-free, meaning it does not contain F as a subgraph. The asymptotic value of this number is known for many classes of graphs, in particular all nonbipartite graphs, but remains unknown for many bipartite graphs, e.g. even cycles [14].

One straightforward variation of the above is the *generalized Turán problem* [1]: determining the maximum number of copies of a given graph H in an n-vertex graph which does not contain a forbidden graph F as a subgraph. We denote this quantity as $\mathrm{ex}(n, H, F)$. Particular attention was paid to the case when both H and F are cycles – see e.g. [3, 6, 7, 8, 10, 11].

In this work we consider an analogous question in the setting of oriented graphs (directed graphs without cycles of length 2). Let $\overrightarrow{C_i}$ denote a directed cycle on i vertices. We focus on

*The first, second and fourth authors were supported by the National Science Centre grant 2021/42/E/ST1/00193.

†Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: andrzej.grzesik@uj.edu.pl.

‡Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: justynajoanna.jaworska@student.uj.edu.pl.

§Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. Current affiliation: Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, Germany. E-mail: bkielak@alumni.uj.edu.pl.

¶Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: piotr.kuc@student.uj.edu.pl.

‖Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: tomaszsl@mit.edu.

$\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ – the maximum number of directed cycles of length k in an n-vertex oriented graph which does not contain a subgraph isomorphic to a directed cycle of length ℓ.

We start by establishing the order of magnitude of $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for every k and ℓ.

Theorem 1. *Let $k, \ell \geq 3$ be different integers. If $k \nmid \ell$, then $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \Theta(n^k)$, while if $k \mid \ell$, then $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \Theta(n^{k-1})$.*

The distinction between the cases $k \mid \ell$ and $k \nmid \ell$ comes from the fact that if $k \mid \ell$, then $\overrightarrow{C_k}$ is a homomorphic image of $\overrightarrow{C_\ell}$ and therefore a $\overrightarrow{C_\ell}$-free graph cannot have positive density of $\overrightarrow{C_k}$ (consequence of the directed analogue of the blow-up lemma [12]). We refer to the case $k \mid \ell$ as the 'sparse setting' and to the case $k \nmid \ell$ as the 'dense setting'.

In the dense setting, a natural candidate for the extremal construction is a balanced blow-up of $\overrightarrow{C_k}$. A *blow-up* of a graph H on m vertices $v_1, v_2, \ldots, v_m$ is any graph G on a vertex set partitioned into disjoint independent sets $V_1, V_2, \ldots, V_m$ (called *blobs*), in which there is an arc uv for $u \in V_i$ and $v \in V_j$ if and only if there is an arc $v_i v_j$ in H. A blow-up is *balanced* if $||V_i| - |V_j|| \leq 1$ for any i, j. Balanced blow-up of $\overrightarrow{C_k}$ has many copies of $\overrightarrow{C_k}$ and does not contain cycles of length not divisible by k, in particular $\overrightarrow{C_\ell}$. This construction can be further improved by reducing the number of blobs to the smallest divisor of k greater than 2 which does not divide ℓ (call this number d). We show that for many pairs (k, ℓ), the balanced blow-up of $\overrightarrow{C_d}$ is indeed the optimal construction.

Theorem 2. *Let k and ℓ be integers satisfying $k \geq 3$, $\ell \geq 2(k-1)^2$, and $k \nmid \ell$. Denote by d the smallest integer greater than 2 that divides k but does not divide ℓ. If $2 \nmid k$, $2 \mid \ell$ or $d \leq 4$, then $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \frac{n}{k}\left(\frac{n}{d}\right)^{k-1} + o(n^k)$.*

Why are these additional divisibility conditions needed? If k is even and ℓ is odd, then we can consider a construction which imitates a "blow-up of $\overrightarrow{C_2}$" – a random orientation of a complete balanced bipartite graph. This graph contains many copies of $\overrightarrow{C_k}$ when $2 \mid k$ and no $\overrightarrow{C_\ell}$ when $2 \nmid \ell$. However, it contains less copies of $\overrightarrow{C_k}$ than a balanced blow-up of $\overrightarrow{C_d}$ for $d \leq 4$, which is why this construction is better only if $d \geq 5$ in the first place. We show that under such divisibility conditions this construction is optimal for large enough ℓ.

Theorem 3. *Let k and ℓ be integers satisfying $k \geq 3$, $\ell \geq 33k^2$, and $k \nmid \ell$. If $2 \mid k$, $4 \nmid k$, $2 \nmid \ell$, and $3 \nmid k$ or $3 \mid \ell$, then $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell}) = \frac{2}{k}\left(\frac{n}{4}\right)^k + o(n^k)$.*

Both theorems have an assumption that ℓ is large enough comparing to k. Although it might be possible to improve these bounds, it turns out they cannot be dropped entirely. We determine $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ for $k \in \{3, 4, 5\}$ and all ℓ not divisible by k, except one pair, showing that when ℓ is small, various extremal constructions may appear.

Theorem 4. *(i) $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell}) = \lceil \frac{n}{3} \rceil \lceil \frac{n-1}{3} \rceil \lceil \frac{n-2}{3} \rceil$ for $\ell = 4$ or $\ell = 5$.*

(ii) $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_\ell}) = \frac{n^3}{27} + \mathcal{O}(n^2)$ for all $\ell > 6$, $3 \nmid \ell$

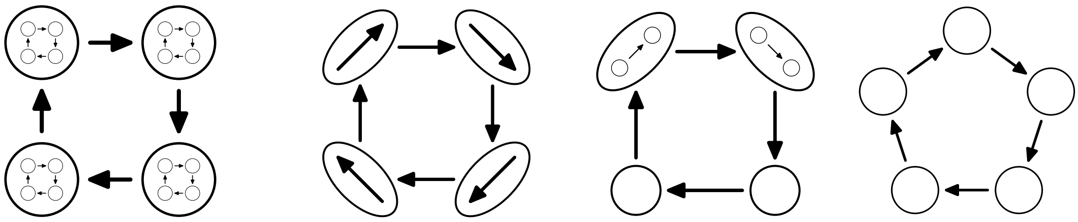
Theorem 5. *(i) $\mathrm{ex}(n, \overrightarrow{C_4}, \overrightarrow{C_3}) = \frac{n^4}{4^4-1} + o(n^4)$.*

(ii) $\mathrm{ex}(n, \overrightarrow{C_4}, \overrightarrow{C_\ell}) = \left(\frac{n}{4}\right)^4 + o(n^4)$ for all $\ell > 4$, $4 \nmid \ell$.

Theorem 6. *(i) $\mathrm{ex}(n, \overrightarrow{C_5}, \overrightarrow{C_3}) = \frac{1}{512} n^5 + o(n^5)$.*

(ii) $\mathrm{ex}(n, \overrightarrow{C_5}, \overrightarrow{C_7}) = \frac{27}{16}\left(\frac{n}{5}\right)^5 + o(n^5)$.

(iii) $\mathrm{ex}(n, \overrightarrow{C_5}, \overrightarrow{C_\ell}) = \left(\frac{n}{5}\right)^5 + o(n^5)$ *for all* $\ell > 5,\ 5 \nmid \ell,\ \ell \neq 7$.



(a) An iterated blow-up of $\overrightarrow{C_4}$ optimal for $\mathrm{ex}(n, \overrightarrow{C_4}, \overrightarrow{C_3})$, which is also maximizing the number of induced copies of $\overrightarrow{C_4}$ [4].

(b) Extremal constructions in Theorem 6: (i) a balanced blow-up of $\overrightarrow{C_4}$ with transitive tournaments inside blobs; (ii) an unbalanced blow-up of $\overrightarrow{C_4}$ with a balanced complete bipartite graph orientation inside two blobs; (iii) a balanced blow-up of $\overrightarrow{C_5}$.

Figure 1: Summary of known extremal constructions in the dense case.

The corresponding extremal constructions are: a balanced blow-up of $\overrightarrow{C_k}$ (Theorems 4, 5.(ii), 6.(iii)), an iterated blow-up of $\overrightarrow{C_k}$ (Theorem 5.(i), see Figure 1a) and other modifications of blow-ups (Theorem 6, see Figure 1b). The missing value $\mathrm{ex}(n, \overrightarrow{C_5}, \overrightarrow{C_4})$ is not achieved in any of the listed constructions. The level of diversity of extremal graphs appears to grow with k.

Lastly, we present one optimal bound in the sparse setting.

Theorem 7. $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_6}) = \frac{1}{4}n^2 + o(n^2)$.

An extremal construction achieving this bound is an unbalanced blow-up of a $\overrightarrow{C_3}$ in which one blob has size 1 (thus preventing existence of $\overrightarrow{C_6}$) and the two others are equal (see Figure 2).

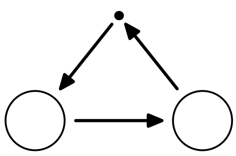


Figure 2: An extremal construction for $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_6})$.

Proof outlines

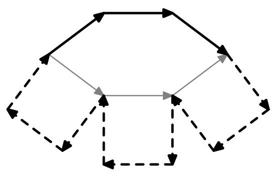
Complete proofs of the presented results are contained in [9], together with remarks on a similar problem in the setting of directed graphs (with cycles of length 2 allowed) and for other orientations of cycles as forbidden graphs.

Theorem 1

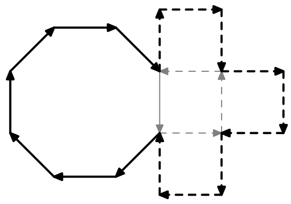
The first part of the theorem is immediate as the number of copies of $\overrightarrow{C_k}$ is upper bounded by n^k and a balanced blow-up of $\overrightarrow{C_k}$ contains $\mathcal{O}(n^k)$ copies of $\overrightarrow{C_k}$ and is $\overrightarrow{C_\ell}$-free.

Note that a blow-up of $\overrightarrow{C_k}$ with one blob of size 1 and remaining blobs balanced does not contain a $\overrightarrow{C_\ell}$, but has $\left(\frac{n-1}{k-1}\right)^{k-1}$ copies of $\overrightarrow{C_k}$, which gives the lower bound for the second part.

It remains to show that $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_{kt}}) = \mathcal{O}(n^{k-1})$ for $k \geq 3$, $t \geq 2$. We use induction on t. The high-level idea is as follows. Take a $\overrightarrow{C_{kt}}$-free graph G. Successively remove from G arcs which are contained in Cn^{k-3} copies of $\overrightarrow{C_k}$ for some carefully chosen value $C = C(k,t)$. This will remove $\mathcal{O}(n^{k-1})$ copies of $\overrightarrow{C_k}$. Now, if $t < k$, we claim that every two directed paths of length t with common endpoints share at least one internal vertex (Figure 3a). This bounds the number of directed paths of length t between some fixed $v, u \in V(G)$ by $(t-1)^2 n^{t-2}$, and consequently, the number of copies of $\overrightarrow{C_k}$ is at most $n(n-1) \cdot (t-1)^2 n^{t-2} \cdot n^{k-t-1} = \mathcal{O}(n^{k-1})$ (because each $\overrightarrow{C_k}$ contains a path of length t as a subgraph). For $t \geq k$, we show that G is $\overrightarrow{C_{k(t-k+2)}}$-free for the inductive step. Assume there is a $\overrightarrow{C_{k(t-k+2)}}$ in G. Since each arc has many copies of $\overrightarrow{C_k}$ containing it, we can extend this $\overrightarrow{C_{k(t-k+2)}}$ into a forbidden $\overrightarrow{C_{kt}}$ as in Figure 3b, contradiction.



(a) Two disjoint paths of length t force a $\overrightarrow{C_{kt}}$



(b) The presence of $\overrightarrow{C_{k(t-k+2)}}$ forces the existence of $\overrightarrow{C_{kt}}$

Figure 3: Forbidden structures in the proof of Theorem 1

Theorem 2

The lower bound is achieved by a balanced blow-up of $\overrightarrow{C_d}$. To upper bound $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$, we first prove the following auxiliary lemma, whose proof follows from the *cycle weighing* technique developed by Král', Norin, and Volec in [13].

Lemma 8. *Let G be a graph satisfying the following property: for any copy C of $\overrightarrow{C_k}$, each vertex of G has at most $\frac{2k}{d}$ neighbors in C. Then the number of copies of $\overrightarrow{C_k}$ in G is at most $\frac{n}{k}\left(\frac{n}{d}\right)^{k-1}$.*

In order to prove Theorem 2, take any extremal n-vertex graph G. Firstly, using Szemerédi Regularity Lemma, we remove $o(n^k)$ copies of $\overrightarrow{C_k}$ to delete from G all closed directed walks of length ℓ (we denote their family by $\overrightarrow{\mathcal{W}_\ell}$ for short). We can also assume that every arc of G is contained in some copy of $\overrightarrow{C_k}$.

Assuming our 'cleared' graph G violates conditions of Lemma 8, we derive a contradiction by showing that G contains a member of $\overrightarrow{\mathcal{W}_\ell}$. This reasoning is split into several cases, let us consider the case $d = 3$ as an example.

It is easy to observe that if some vertex v has more than $\frac{2k}{3}$ neighbors on a k-cycle C, then G must contain a $\overrightarrow{T_3}$ (a transitive triangle). Now consider a $\overrightarrow{C_k}$ containing the arc from the source to the sink of this $\overrightarrow{T_3}$. Together with the two other arcs of $\overrightarrow{T_3}$ it forms intersecting closed directed walk of length $k+1$ and $\overrightarrow{C_k}$. This structure contains closed directed walks of length $x(k+1)+yk$ for arbitrary integers $x, y \geq 0$. Now, since $\ell > k(k-1)$, using Sylvester's result ([15]) on the Frobenius Coin Problem, we conclude that G contains a closed directed walk of length ℓ, as desired.

Theorem 3

The proof of Theorem 3 is split into two lemmas.

Lemma 9. *Let k and ℓ be integers satisfying $k \geq 3$, $\ell > 33k^2$, and $k \nmid \ell$. If $2 \mid k$, $4 \nmid k$, $2 \nmid \ell$, and $3 \nmid k$ or $3 \mid \ell$, then $\mathrm{ex}(n, \overrightarrow{C_k}, \overrightarrow{C_\ell})$ is achieved (up to a lower order error term) in an orientation of a complete bipartite graph.*

A general idea of the proof of the lemma is to consider the extremal graph G and show that the underlying unoriented graph $\widetilde{G}$ has large minimum degree and does not contain short odd cycle. Then the bipartiteness of G follows from the Andrásfai-Erdős-Sós theorem [2].

Claim 10. $\delta(\widetilde{G}) \geq \frac{2}{23}n$.

To show this, we first bound the number of directed paths of given length in G. This implies that a hypothetical vertex of a small degree in $\widetilde{G}$ appears in a limited number of copies of $\overrightarrow{C_k}$. But then we could replace this vertex with a copy of a different one to build a better construction.

Claim 11. *Graph G does not contain any orientation of C_m where $m \leq 23$ is odd.*

Similarly as in Figure 3a, we extend the hypothetical orientation of C_m with a forward arcs and $m-a$ backward arcs into a member of $\overrightarrow{\mathcal{W}}_{a+(m-a)(k-1)}$. Since every vertex appears in a large number of copies of $\overrightarrow{C_k}$ (as in Claim 10), we obtain another closed directed walk of a bounded length. Combining it with the closed directed walk of length $a+(m-a)(k-1)$ and the k-cycle, we can build a forbidden closed directed walk of length ℓ and reach a contradiction.

Lemma 12. *Let G be an orientation of a complete bipartite graph on n vertices. If $2 \mid k$ and $4 \nmid k$, then G contains at most $\frac{2}{k}\left(\frac{n}{4}\right)^k$ copies of $\overrightarrow{C_k}$.*

This proof uses a spectral approach. The number of copies of $\overrightarrow{C_k}$ is upper bounded by $\frac{1}{k}\mathrm{tr}M^k = \frac{1}{k}\sum_{i=1}^{n}\lambda_i^k$, where M is the adjacency matrix of G and $(\lambda_i)_{i\in[n]}$ are its eigenvalues. Using inequalities on complex numbers, we bound $\sum_{i=1}^{n}\lambda_i^k$ by $2(\sum_{i=1}^{\lfloor n/2\rfloor}\mathrm{Re}\,\lambda_i)^k$ and then apply the following proposition.

Proposition 13. *(Ky Fan, [5]) If $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_n$ are the eigenvalues of $\frac{M+M^T}{2}$, then $\sum_{i=1}^{\lfloor n/2\rfloor} \mathrm{Re}\,\lambda_i \leq \sum_{i=1}^{\lfloor n/2\rfloor}\rho_i$.*

Observe that $M+M^T$ is the adjacency matrix of $K_{n,m}$, whose spectrum is known. This gives the wanted bound.

Theorems 4-6

Proofs of these theorems employ Szemerédi Regularity Lemma in the directed version (except for Theorem 4 where sharper bounds were obtained), the cycle weighing technique mentioned earlier, flag algebras [17], and various other combinatorial techniques.

Theorem 7

The extremal example achieving $\frac{1}{4}n^2 + o(n^2)$ copies of $\overrightarrow{C_3}$ was already shown in Figure 2.

Let G be an extremal oriented graph on n vertices. We may assume that each arc of G is contained in at least one copy of $\overrightarrow{C_3}$. We say that an arc of G is *thin* if it is contained in exactly one $\overrightarrow{C_3}$, otherwise we say that it is *thick*.

The crucial observation is that each copy of $\overrightarrow{C_3}$ in G contains at least one thin arc (otherwise, a $\overrightarrow{C_6}$ is easily formed), so $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_6})$ is upper bounded by the number of thin arcs in G. It turns out that the unoriented graph on thin arcs ignoring their orientation, G', cannot contain a blow-up of a triangle – it would either violate the definition of a thin arc or create a $\overrightarrow{C_6}$ in G. Hence, it can be made triangle-free by removing $o(n^2)$ edges and a triangle-free graph has at most $\frac{1}{4}n^2$ edges due to Mantel's theorem. This implies the wanted bound on $\mathrm{ex}(n, \overrightarrow{C_3}, \overrightarrow{C_6})$.

References

[1] Noga Alon and Clara Shikhelman. Many T copies in H-free graphs. *Journal of Combinatorial Theory, Series B*, 121:146–172, 2016.

[2] Béla Andrásfai, Paul Erdős, and Vera T. Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Mathematics*, 8(3):205–218, 1974.

[3] Csongor Beke and Oliver Janzer. On the generalized Turán problem for odd cycles. *SIAM Journal on Discrete Mathematics*, 38(3):2416–2428, 2024.

[4] Łukasz Bożyk, Andrzej Grzesik, and Bartłomiej Kielak. On the inducibility of oriented graphs on four vertices. *Discrete Mathematics*, 345(7):112–874, 2022.

[5] Ky Fan. On a theorem of Weyl concerning eigenvalues of linear transformations: II. *Proceedings of the National Academy of Sciences*, 36(1):31–35, 1950.

[6] Dániel Gerbner, Ervin Győri, Abhishek Methuku, and Máté Vizer. Generalized Turán problems for even cycles. *Journal of Combinatorial Theory, Series B*, 145:169–213, 2020.

[7] Lior Gishboliner and Asaf Shapira. A generalized Turán problem and its applications. *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 760–772, 2018.

[8] Andrzej Grzesik. On the maximum number of five-cycles in a triangle-free graph. *Journal of Combinatorial Theory, Series B*, 102(5):1061–1066, 2012.

[9] Andrzej Grzesik, Justyna Jaworska, Bartłomiej Kielak, Piotr Kuc, and Tomasz Ślusarczyk. Generalized turán problem for directed cycles. *arXiv:2505.22189*, 2025.

[10] Andrzej Grzesik and Bartłomiej Kielak. On the maximum number of odd cycles in graphs without smaller odd cycles. *Journal of Graph Theory*, 99(2):240–246, 2022.

[11] Hamed Hatami, Jan Hladký, Daniel Král', Serguei Norine, and Alexander Razborov. On the number of pentagons in triangle-free graphs. *Journal of Combinatorial Theory, Series A*, 120(3):722–732, 2013.

[12] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. Blow-up lemma. *Combinatorica*, 17:109–123, 1997.

[13] Daniel Král', Sergey Norin, and Jan Volec. A bound on the inducibility of cycles. *Journal of Combinatorial Theory, Series A*, 161:359–363, 2019.

[14] Oleg Pikhurko. A note on the Turán function of even cycles. *Proceedings of the American Mathematical Society*, 140(11):3687–3692, 2012.

[15] James J. Sylvester. On subvariants, i.e. semi-invariants to binary quantics of an unlimited order. *American Journal of Mathematics*, 5(1):79–136, 1882.

[16] Paul Turán. On an extremal problem in graph theory (in Hungarian). *Középiskolai Matematikai és Fizikai Lapok*, 48:436–452, 1941.

[17] Emil Vaughan. Flagmatic software package, further developed by Jakub Sliačan, available at `https://github.com/jsliacan/flagmatic`.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Outdegree conditions forcing directed cycles*

(Extended abstract)

Andrzej Grzesik† Justyna Jaworska‡ Piotr Kuc§

Abstract

Kelly, Kühn and Osthus [J. Combin. Theory Ser. B 100 (2010), 251–264] made a conjecture on the minimum semidegree which forces an oriented graph to contain a directed cycle of a given length at least 4. The conjecture was proven by its authors for cycles of length not divisible by 3 and by Grzesik and Volec [Int. Math. Res. Not. IMRN 2023 (2023), 9711–9253] for other cycles. In this paper we prove a stronger statement determining the minimum outdegree which forces an oriented graph to contain a directed cycle of a given large enough length.

Introduction

One of the most known and often studied open problems on oriented graphs is the following conjecture from 1978.

Conjecture 1 (Caccetta, Häggkvist [3]). *Every n-vertex oriented graph with $\delta^+(G) > \frac{n}{\ell}$ contains a directed cycle of length at most ℓ.*

Here by an *oriented graph* we understand a directed graph without loops and multiple edges, while $\delta^+(G)$ and $\delta^-(G)$ denotes the *minimum outdegree*, respectively *minimum indegree*, of the oriented graph G.

There are many partial results on Conjecture 1, including proofs with an additive error term in the bound on the cycle length, a multiplicative error term in the minimum outdegree assumption, or solutions with an additional assumption on forbidden subgraphs. For more results and problems related to the Caccetta–Häggkvist conjecture, see a summary in [12].

*This work was supported by the National Science Centre grant 2021/42/E/ST1/00193.
†Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: andrzej.grzesik@uj.edu.pl.
‡Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: justynajoanna.jaworska@student.uj.edu.pl.
§Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: piotr.kuc@student.uj.edu.pl.

Even if we strengthen the assumptions in Conjecture 1 by assuming the *minimum semidegree* $\delta^{\pm}(G) = \min(\delta^{+}(G), \delta^{-}(G))$ instead of just the minimum outdegree, the essential difficulty seems to persist and the conjecture is also open.

Motivated by this conjecture, Kelly, Kühn and Osthus [8] proposed in 2010 the following problem: for a given integer $\ell \geq 4$, determine the minimum semidegree enforcing an oriented graph to contain a directed cycle of length *exactly* ℓ. Despite a similar formulation, this problem is substantially different, and the answer depends on divisibility conditions of ℓ. Indeed, Kelly, Kühn and Osthus conjectured the following.

Conjecture 2 (Kelly, Kühn, Osthus [8])**.** *Let $\ell \geq 4$ be an integer and k be the smallest integer greater than 2 that does not divide ℓ. Then there exists n_0 such that every oriented graph G on $n \geq n_0$ vertices with minimum semidegree $\delta^{\pm}(G) \geq \lfloor \frac{n}{k} \rfloor + 1$ contains a directed cycle of length ℓ.*

Kelly, Kühn and Osthus [8] proved Conjecture 2 in the case $k = 3$, that is for cycles of lengths divisible by 3. They also showed an asymptotic statement (with the semidegree bound $(1+\varepsilon)\frac{n}{k}$) in the cases $k = 4$ and $k = 5$. Later, Kühn, Osthus and Piguet [10] proved a similar asymptotic statement for other values of k and $\ell > 10^7 k^6$. Conjecture 2 was proved in 2023 with a slight adjustment in the semidegree bound.

Theorem 3 (Grzesik, Volec [5])**.** *Let $\ell \geq 4$ be an integer and k be the smallest integer greater than 2 that does not divide ℓ. Then there exists n_0 such that for every oriented graph G on $n \geq n_0$ vertices:*

- *if $\ell \not\equiv 3 \bmod 12$ and $\delta^{\pm}(G) \geq \frac{n}{k} + \frac{k-1}{2k}$, then G contains $\overrightarrow{C_\ell}$,*
- *if $\ell \equiv 3 \bmod 12$ and $\delta^{\pm}(G) \geq \frac{n}{4} + \frac{1}{4}$, then G contains $\overrightarrow{C_\ell}$.*

Moreover, the above bounds are best possible.

Similarly as in the Caccetta-Häggkvist conjecture, Conjecture 2 seems to hold in a stronger version, when we assume the minimum outdegree instead of the minimum semidegree. However, techniques in the aforementioned results cannot be easily adopted to the stronger version as they heavily rely on bounding the directed diameter, which cannot be done without the indegree assumption.

The first result regarding the stronger version of Conjecture 2 is the following theorem.

Theorem 4 (Czygrinow et al. [4])**.** *For every fixed integer $\ell \geq 4$, there exists n_0 such that every oriented graph G on $n \geq n_0$ vertices with minimum outdegree $\delta^{+}(G) \geq \frac{n}{3} + \frac{1}{3}$ contains a directed cycle of length ℓ.*

The above outdegree bound is tight for every ℓ not divisible by 3 with the extremal construction being the balanced blow-up of $\overrightarrow{C_3}$. We are thus left with the case $3|\ell$.

We prove a similar result in the remaining cases for cycles of large enough length.

Theorem 5. *Let ℓ be an integer divisible by 3 and k be the smallest integer greater than 2 that does not divide ℓ. If $\ell \geq 20k^2$, then there exists n_0 such that every oriented graph G on $n \geq n_0$ vertices with $\delta^{+}(G) \geq \frac{n}{k} + \frac{k-1}{k}$ contains $\overrightarrow{C_\ell}$.*
Moreover, this outdegree bound is best possible if $\ell \not\equiv 3, 6 \bmod 12$ and $\ell \not\equiv 24, 48 \bmod 60$.

Note that when $k > 9$, the inequality $\ell \geq 20k^2$ becomes redundant as every ℓ satisfies it. In the special cases, where the outdegree bound in Theorem 5 is not tight, we provide constructions witnessing that the optimal outdegree bound is within an interval of length at most $\frac{2}{k}$. We proved the following theorem showing that the optimal outdegree bound may indeed be a bit smaller.

Theorem 6. *Let $\ell \geq 150$ be an integer divisible by 3. There exists n_0 such that for every oriented graph G on $n \geq n_0$ vertices:*

- *if $\ell \equiv 6 \bmod 12$ and $\delta^+(G) \geq \frac{n}{4} + \frac{2}{4}$, then G contains $\overrightarrow{C_\ell}$,*
- *if $\ell \equiv 24 \bmod 60$ and $\delta^+(G) \geq \frac{n}{5} + \frac{3}{5}$, then G contains $\overrightarrow{C_\ell}$.*

Moreover, these outdegree bounds are best possible.

Notation

We denote by $\overrightarrow{C_k}$ a directed cycle on k vertices and by $\overrightarrow{\mathcal{W}_k}$ the family of closed directed walks of length k. We define the *distance* from a vertex u to a vertex v as the smallest length of a directed path from u to v; if there is no such path, we say this distance is infinite. For an oriented graph F, we say that an oriented graph is *F-free* if it does not contain F as a subgraph. Similarly, for a family $\mathcal{F}$, a graph if $\mathcal{F}$-free if it does not contain any element of F as a subgraph.

A *blow-up* of an oriented graph H on k vertices $v_1, v_2, \ldots, v_k$ is any oriented graph on a vertex set partitioned into disjoint independent sets $V_1, V_2, \ldots, V_k$, in which there is an edge uv for $u \in V_i$ and $v \in V_j$ if and only if there is an edge v_iv_j in H. The sets V_i are called *blobs*. A blow-up is *balanced* if $||V_i| - |V_j|| \leq 1$ for any i, j. Note that an oriented graph is homomorphic to H if and only if it is a subgraph of a blow-up of H.

Constructions

Before we proceed to the proofs, we present constructions showing that the outdegree bounds in Theorem 5 and Theorem 6 are best possible.

For $r \in \{3, 4, \ldots, k-1\}$, we define an *$r$-maneuver* as the following operation (that could be repeated multiple times) on a balanced blow-up of $\overrightarrow{C_k}$ with blobs $V_1, \ldots, V_k$ (see Figure 1):

- pick a *special* vertex v from an arbitrary blob V_i, which was not chosen before;
- add edges from all vertices in V_{i+r-1} to the special vertex v;
- remove from V_{i+r} a vertex, which was not special in any previous maneuver.

Note that applying an r-maneuver does not change the minimum outdegree of the graph and decreases the number of vertices by 1. However, when traversing our graph, we can choose a shortcut once and return to the starting blob after r steps instead of k. In particular, after m applications of an r-maneuver, the graph contains only directed cycles of lengths with remainder modulo k in the set $\{0, r, 2r, 3r, \ldots, mr\}$.

In Theorem 6, our goal is to present, for some arbitrarily large n, an n-vertex $\overrightarrow{C_\ell}$-free oriented graph G with outdegree $\delta^+(G) = \frac{n+k-3}{k}$. Consider an integer n such that $n + k - 3$

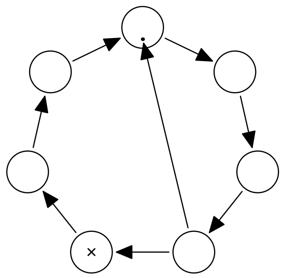


Figure 1: Example of a 4-maneuver.

is divisible by k and a balanced blow-up of $\overrightarrow{C_k}$ on $n+k-3$ vertices. Applying $k-3$ times a 3-maneuver we obtain an n-vertex oriented graph with outdegree $\frac{n+k-3}{k}$ and no cycle of length ℓ, as wanted.

Similarly, in Theorem 5, we consider an integer n such that $n+k-2$ is divisible by k and apply the following maneuvers on a balanced blow-up of $\overrightarrow{C_k}$ on $n+k-2$ vertices:

- for $k=4$ and $\ell \equiv 1 \bmod 4$ – apply twice a 3-maneuver;
- for $k=5$ and $\ell \equiv 1$ or $2 \bmod 5$ – apply three times a 4-maneuver, or a 3-maneuver;
- for odd $k>5$ – apply $a-1$ times a $\left(\frac{k-1}{2}\right)$-maneuver and $k-a-1$ times a $\left(\frac{k+1}{2}\right)$-maneuver, where $a \in \{1,2,\dots,k-1\}$ is such that $a\frac{k-1}{2} \equiv \ell \bmod k$;
- for even $k>5$ – apply $a-1$ times a 3-maneuver and $k-a-1$ times a $(k-3)$-maneuver, where $a \in \{1,2,\dots,k-1\}$ is such that $a \cdot 3 \equiv \ell \bmod k$.

Proof outline

We present only an overview of the proof of Theorem 5, as Theorem 6 is proven in a similar way, but using a more detailed analysis. The major step of the proof is to show a stability lemma providing an approximate structure of a $\overrightarrow{\mathcal{W}_\ell}$-free oriented graph with a large minimum outdegree. Its statement depends on the parity of ℓ, so we split the proof into two cases.

Even case

We start with presenting the needed stability lemma and sketching its proof. Note that even ℓ means that either $\ell \equiv 2 \bmod 4$, or $k \geq 5$.

Lemma 7. *Let $\ell \geq 4$ be an even integer divisible by 3 and k be the smallest integer greater than 2 that does not divide ℓ. If $\ell \geq 20k^2$, then for every $\varepsilon > 0$ there exist $\gamma > 0$ and n_0 such that every $\overrightarrow{\mathcal{W}_\ell}$-free oriented graph H on $n \geq n_0$ vertices with $\delta^+(H) \geq \frac{n}{k}(1-\gamma)$ contains an induced subgraph H' homomorphic to $\overrightarrow{C_k}$ with $\delta^\pm(H') \geq \frac{n}{k}(1-\varepsilon)$.*

To prove Lemma 7, we firstly modify the proof of Claim 4.1 in [5] to obtain that every vertex $v \in H$ has distance at most $5\lfloor\frac{k-1}{3}\rfloor$ to more than half of the other vertices. This implies that there exists a vertex v_0 such that, for any vertex $v \neq v_0$, there is a directed path of length at most $10\lfloor\frac{k-1}{3}\rfloor$ from v to v_0. Then, we sequentially construct H' by considering a sequence of sets V_d ($d = 1,2,\dots,k$) of vertices at distance d from v_0. On each step we show that V_i are independent and there are no edges from V_j to V_i for $i < j$, except $i = 1$, $j = k$, because otherwise we would find two short intersecting closed directed walks, which can be combined to obtain the forbidden $\overrightarrow{\mathcal{W}_\ell}$ by reduction to the classical Frobenius coin problem. Given this, we infer that each set V_d is of size at least $\delta^+(H)$. The lower bound on $\delta^+(H')$ is a consequence of the order of H'. Lastly, we prove the upper bound on the maximum indegree of H', which implies, using an edge counting argument, that there are very few vertices with a low indegree. Removing them we obtain the wanted induced subgraph of a high semidegree.

We may now proceed to the proof of Theorem 5. Assuming there exists a counterexample G, we apply Szemerédi Regularity Lemma in the degree version for oriented graphs [1] to obtain a $\overrightarrow{\mathcal{W}_\ell}$-free oriented subgraph with a large outdegree, for which we use Lemma 7. This gives that G consists of a subgraph H' of a blow-up of $\overrightarrow{C_k}$ with $\delta^\pm(H') \geq \frac{n}{k}(1-\varepsilon)$ and some additional edges and vertices.

Let us first tackle the additional edges between the vertices of H'. Each such edge constitutes a *shortcut* one could take when traversing the blobs of H'. Those shortcuts can then be combined to form various cycle lengths. It can be shown that a set of $k-1$ disjoint shortcuts already implies existence of a $\overrightarrow{C_\ell}$. This implies that there exists a set of at most $2k-4$ vertices which are incident to all the shortcuts. By removing them from $V(H')$, we may assume that H' is an induced subgraph of G homomorphic to $\overrightarrow{C_k}$.

We are thus left with a set of additional vertices $T = V(G) \setminus V(H')$. We follow the approach used in the proof of Lemma 4.3 in [5] with modifications forced by the lack of the minimum indegree assumption. Every vertex $v \in T$ with non-negligible inneighborhood in $z \geq 2$ blobs forms $z-1$ non-zero *sidewalks*, which, similarly to shortcuts, can be combined into cycles of various lengths. We consider the largest possible set of compatible sidewalks not creating $\overrightarrow{C_\ell}$ and apply a corollary of Kneser's Theorem [9] proved in [5], to obtain that the sum $\sum(z-1)$ is less than $k-1$. But then the outdegree bound forces at least one more sidewalk disjoint from the others, which completes the proof by contradiction.

Odd case

For odd ℓ, consequently $k = 4$, the thesis of Lemma 7 does not hold. For example, any balanced orientation of a complete bipartite graph is $\overrightarrow{\mathcal{W}_\ell}$-free and has a large minimum outdegree, but almost certainly does not contain an induced subgraph on $n - o(n)$ vertices homomorphic to $\overrightarrow{C_4}$. Nevertheless, in this case we are able to prove the following stability lemma.

Lemma 8. *For any odd integer $\ell \geq 45$ divisible by 3 and $\varepsilon > 0$ there exist $\gamma > 0$ and n_0 such that every $\overrightarrow{\mathcal{W}_\ell}$-free oriented graph H on $n \geq n_0$ vertices with $\delta^+(H) \geq \frac{n}{4}(1-\gamma)$ contains an induced bipartite subgraph H' with $\delta^\pm(H') \geq \frac{n}{4}(1-\varepsilon)$.*

The proof of Lemma 8 follows similar lines as the proof of Lemma 7. Modifying the proof of Claim 5.3 in [5], we obtain that there exists a vertex v_0 such that for any vertex $v \neq v_0$ there is a directed path of length at most 6 from v to v_0. Then, we construct the sets V_1, V_2, V_3 and V_4, but now we only show that there are no edges between V_i and V_j if $i-j$ is even. Combining bounds on the number of edges incident to appropriately chosen subsets of $V_1 \cup V_2 \cup V_3 \cup V_4$, we prove that the sum of their sizes is large, implying the needed outdegree bound of H'. Finally, similarly as in Lemma 7, we show that there are very few vertices with a low indegree, whose removal returns the wanted bipartite induced subgraph of a high semidegree.

Lemma 8 allows us to prove the following lemma, providing stronger structural information on an oriented graph with large minimum outdegree.

Lemma 9. *For any odd integer $\ell \geq 45$ divisible by 3 and $\varepsilon > 0$ there exists n_0 such that every $\overrightarrow{C_\ell}$-free oriented graph G on $n \geq n_0$ vertices with $\delta^+(G) \geq \frac{n+1}{4}$ contains a subgraph H' homomorphic to $\overrightarrow{C_4}$ with $\delta^\pm(H') \geq \frac{n}{4}(1-\varepsilon)$.*

To prove Lemma 9, we infer by the aforementioned Szemerédi Regularity Lemma that G contains a $\overrightarrow{\mathcal{W}_\ell}$-free subgraph H satisfying the assumptions of Lemma 8, so G contains a bipartite subgraph H' with $\delta^\pm(H') \geq \frac{n}{4}(1-\varepsilon)$. Let A and B be the subsets of $V(G)$ from the bipartition of H'. The key idea now is to use the minimum outdegree assumption of G to find two large subsets $V_1 \subset A$ and $V_2 \subset B$ without edges from V_2 to V_1. To do this, we use a vertex in $V(G) \setminus (A \cup B)$ to find a directed path of length $\ell - 3$ that starts in B and ends in A and define V_1 as the inneighbourhood of the starting vertex and V_2 as the outneighbourhood of the terminal one. By the semidegree bound of H', they are both large. Finally, we set $V_3 = A \setminus V_1$, $V_4 = B \setminus V_2$, and cyclically bound the number of edges from V_{i+1} to V_i, using the bound on the number of edges from V_i to V_{i-1}.

Lemma 9 gives us insight into the structure of a hypothetical counterexample to Theorem 5. The further proof resembles the proof for the previously considered even case, but with a more involved argument allowing to remove all sidewalks not changing the parity of attainable cycles.

References

[1] Noga Alon and Asaf Shapira. Testing subgraphs in directed graphs. *Journal of Computer and System Sciences*, 69:354–382, 2004.

[2] John Adrian Bondy. Counting subgraphs: a new approach to the caccetta-häggkvist conjecture. *Discrete Mathematics*, 165-166:71–80, 1997.

[3] Louis Caccetta and Roland Häggkvist. On minimal digraphs with given girth. *Congressus Numeratum*, 21:181–187, 1978.

[4] Andrzej Czygrinow, Theodore Molla, Brendan Nagle, and Roy Oursler. On even rainbow or nontriangular directed cycles. *Journal of Combinatorics*, 12(4):589–662, 2021.

[5] Andrzej Grzesik and Jan Volec. Degree conditions forcing directed cycles. *International Mathematics Research Notices*, 2023(11):9711–9753, 2023.

[6] Peter Hamburger, Penny Haxell, and Alexandr Kostochka. On directed triangles in digraphs. *The Electronic Journal of Combinatorics*, 14:N19, 2007.

[7] Jan Hladkỳ, Daniel Král', and Sergey Norin. Counting flags in triangle-free digraphs. *Combinatorica*, 37(1):49–76, 2017.

[8] Luke Kelly, Daniela Kühn, and Deryk Osthus. Cycles of given length in oriented graphs. *Journal of Combinatorial Theory, Series B*, 100(3):251–264, 2010.

[9] Martin Kneser. Summenmengen in lokalkompakten abelschen Gruppen. *Mathematische Zeitschrift*, 66:88–110, 1956. (in German).

[10] Daniela Kühn, Deryk Osthus, and Diana Piguet. Embedding cycles of given length in oriented graphs. *European Journal of Combinatorics*, 34(2):495–501, 2013.

[11] Jian Shen. Directed triangles in digraphs. *Journal of Combinatorial Theory, Series B*, 74:405–407, 1998.

[12] Blair D. Sullivan. A summary of results and problems related to the Caccetta-Häggkvist conjecture. *arXiv:math/0605646*, 2006.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Antidirected paths in oriented graphs*

(Extended abstract)

Andrzej Grzesik† Marek Skrzypczyk‡

Abstract

We show that for any integer $k \geq 4$, every oriented graph with minimum semidegree bigger than $\frac{1}{2}k + \frac{1}{2}\sqrt{k}$ contains an antidirected path of length k. Consequently, every oriented graph on n vertices with more than $(k + \sqrt{k})n$ edges contains an antidirected path of length k. This asymptotically proves conjectures of Stein and of Addario-Berry, Havet, Linhares Sales, Reed and Thomassé, respectively.

1 Introduction

There are classic results of Dirac [4] and Erdős and Gallai [5] that determine optimal bounds on the minimum degree or the number of edges that forces a graph to contain a path of a given length. In this paper, we continue a recent line of research on similar conditions forcing an oriented graph to contain a path oriented in alternating directions. This is the most natural and interesting orientation of a path for which no optimal bounds are known.

An *oriented graph* G is a graph with an orientation on every edge. In particular, for every $u, v \in V(G)$ at most one of uv and vu is an edge of G. If there is an edge uv in $E(G)$, then we say that it is directed from u to v, v is an *out-neighbour* of u and u is an *in-neighbour* of v. The *in-degree* of v, denoted by $d^-(v)$, is the number of in-neighbours of v, while the *out-degree* of v, denoted by $d^+(v)$, is the number of out-neighbours of v. The *minimum semidegree* of an oriented graph G is $\delta^\pm(G) = \min\{d^-(v), d^+(v) : v \in V(G)\}$. We also

*This work was supported by the National Science Centre grant 2021/42/E/ST1/00193.

†Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: andrzej.grzesik@uj.edu.pl.

‡Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: marek.skrzypczyk@doctoral.uj.edu.pl.

define the *minimum pseudo-semidegree* $\bar{\delta}^{\pm}(G)$ of G as follows. If G has no edges, then $\bar{\delta}^{\pm}(G) = 0$, otherwise $\bar{\delta}^{\pm}(G)$ is the largest integer d such that each vertex has out-degree either 0 or at least d, and in-degree either 0 or at least d. Obviously always $\bar{\delta}^{\pm}(G) \geq \delta^{\pm}(G)$.

An *antidirected path*, for short *antipath* (*antidirected cycle*, in short *anticycle*), is an orientation of a path (cycle) such that for each vertex v we have either $d^{-}(v) = 0$ or $d^{+}(v) = 0$. The *length* of an oriented path (cycle) is the number of its edges. Note that there is only one antipath of a given odd length, whereas there are two non-isomorphic antipaths of any given even length. Clearly, each anticycle has even length.

Stein [11] conjectured the following condition on the minimum semidegree forcing an oriented graph to contain any given orientation of a path.

Conjecture 1 (Stein [11])**.** *For any integer $k \geq 1$, every oriented graph G with $\delta^{\pm}(G) > \frac{1}{2}k$ contains each oriented path of length k.*

The above threshold, if true, is tight. This can be seen by noticing that a blow-up of any directed cycle with $\frac{1}{2}k$ vertices in each blob has semidegree equal to $\frac{1}{2}k$ and does not contain an antipath of length k. The conjecture is known to be true for directed paths [6] and for $k \leq 4$ [9]. Since the antipath is the most natural orientation for which Conjecture 1 is open, and known extremal examples witness that not containing an antipath seems to be the most restrictive condition, attention was recently drawn to the antipaths. In particular, Stein and Zárate-Guerén [13] proved that an approximate version of Conjecture 1 holds for antipaths in large oriented graphs on n vertices if k is linear in n.

An interesting observation is that if we replace the minimum semidegree condition in Conjecture 1 for antipaths by the weaker condition on the minimum pseudo-semidegree, then we obtain an equivalent conjecture. It is so, because it is easy to combine copies of an oriented graph G (and copies with reversed orientations) to create an oriented graph G' with $\delta^{\pm}(G') = \bar{\delta}^{\pm}(G)$ without creating new antipaths. Therefore, Conjecture 1 for antipaths is equivalent to the following problem.

Conjecture 2. *For any integer $k \geq 1$, every oriented graph G with $\bar{\delta}^{\pm}(G) > \frac{1}{2}k$ contains each antipath of length k.*

Klimošová and Stein [9] proved that for $k \geq 3$ the stronger bound $\bar{\delta}^{\pm}(G) > \frac{3}{4}k - \frac{3}{4}$ implies that G contains each antipath of length k. This was further improved by Chen, Hou and Zhou [3] to approximately $\frac{2}{3}k$ and by Skokan and Tyomkyn [10] to $\frac{5}{8}k$.

In this paper, we improve the minimum pseudo-semidegree bound to a value that agrees with the conjectured $\frac{1}{2}k$ threshold up to a $\mathcal{O}(\sqrt{k})$ error term.

Theorem 3. *For any integer $k \geq 1$, every oriented graph G with $\bar{\delta}^{\pm}(G) > \frac{1}{2}k + \frac{1}{2}\sqrt{k}$ contains each antipath of length k.*

Addario-Berry, Havet, Linhares Sales, Thomassé and Reed [1] stated the following conjecture related to Conjecture 2.

Conjecture 4 (Addario-Berry et al. [1])**.** *For any integer $k \geq 1$, every oriented graph G on n vertices with more than $(k-1)n$ edges contains each antidirected tree on $k+1$ vertices.*

Stein and Zárate-Guerén [13] showed that any oriented graph on n vertices with more than cn edges contains an oriented graph G with $\bar{\delta}^{\pm}(G) > \frac{1}{2}c$. This means that any improvement in the bound for the minimum pseudo-semidegree forcing an oriented graph to contain antipaths is providing an improvement on the bound for the edge number forcing to contain antipaths. Therefore, Theorem 3 implies the following corollary.

Corollary 5. *For any integer $k \geq 1$, every oriented graph G on n vertices with more than $(k + \sqrt{k})n$ edges contains each antipath of length k.*

2 Overview of the proof

We make use of the following lemmas proved by Klimošová and Stein [9].

Lemma 6 (Klimošová, Stein [9])**.** *Let $k \geq 1$ be an integer and G be an oriented graph with $\bar{\delta}^{\pm}(G) \geq \frac{1}{2}k$ having a longest antipath of length m. If $m < k$, then m is odd.*

Lemma 7 (Klimošová, Stein [9])**.** *Let $k \geq 1$ be an integer and G be an oriented graph with $\bar{\delta}^{\pm}(G) > \frac{1}{2}k$ containing an anticycle of length $m+1$. If $m < k$, then G contains an antipath of length $m+1$.*

In order to prove Theorem 3, we consider $k \geq 1$ and any oriented graph G satisfying $\bar{\delta}^{\pm}(G) > \frac{1}{2}k + \frac{1}{2}\sqrt{k}$. Let $A = v_0 v_1 v_2 \dots v_m$ be a longest antipath in G. If $m \geq k$, then we are done, so we may assume that $m < k$. By Lemma 6 m is odd. If we show that G contains an anticycle of length $m+1$, then by Lemma 7 there is a longer antipath, which gives a contradiction to the maximality. Therefore, we may assume that G does not contain an anticycle of length $m+1$.

Denote by $E = \{v_0, v_2, \dots, v_{m-1}\}$ the set of vertices in $V(A)$ with even index and by $O = \{v_1, v_3, \dots, v_m\}$ the set of vertices in $V(A)$ with odd index. From symmetry we may assume that every edge of A is directed from E to O.

If $v_0 v_{2i+1} \in E(G)$ for some positive $i < \frac{1}{2}m$, then $v_{2i} v_{2i-1} \dots v_0 v_{2i+1} v_{2i+2} \dots v_m$ is an antipath on all vertices of A with all edges directed from E to O. Similarly, if $v_{2i} v_m \in E(G)$, then also $v_0 v_1 \dots v_{2i} v_m v_{m-1} \dots v_{2i+1}$ is such an antipath. Let $S \subseteq O$ and $P \subseteq E$ be the sets of vertices that can be the second and the penultimate vertices in some antipath on all vertices of A that can be obtained from A by the above operations. In particular, $v_1 \in S$ and $v_{m-1} \in P$.

Claim 8. *There is an edge from v_0 to P or from S to v_m.*

Proof. For the sake of contradiction assume that there are no such edges. This means that vertex v_0 has at most $\frac{m-1}{2} - |P|$ out-neighbours in E, so at least $\bar{\delta}^{\pm}(G) - \frac{m-1}{2} + |P|$ out-neighbours in O. If v_{2i+1} is an out-neighbour of v_0 in O for $i > 0$, then $v_{2i-1} \in S$. Therefore, $|S| \geq \bar{\delta}^{\pm}(G) - \frac{m-1}{2} - 1 + |P| > |P|$. A symmetric argument from the perspective of v_m leads to $|P| > |S|$, which is a contradiction. □

By Claim 8 and symmetry, we may assume that G contains an edge v_1v_m. This means that vertex v_1 has simultaneously positive in-degree and out-degree. In particular, it has at least $2\bar{\delta}^{\pm}(G) - m + 1 \geq 2\bar{\delta}^{\pm}(G) - k + 2$ neighbours in the set $R = V(G) \setminus \{v_2, v_3, \ldots, v_m\}$. We may assume that v_1 has at least $\bar{\delta}^{\pm}(G) - \frac{1}{2}k + 1$ in-neighbours in R, because each out-neighbour $w \in R$ of v_1 creates an antipath $v_2v_3 \ldots v_mv_1w$ of length m with v_1 as the penultimate vertex, which is a symmetric situation. Let F be the set of those at least $\bar{\delta}^{\pm}(G) - \frac{1}{2}k + 1 \geq 2$ in-neighbours of v_1 in R. Note that each vertex w in F is the first vertex on an antipath $wv_1v_2 \ldots v_m$ of length m, in particular $v_0 \in F$.

Claim 9. *For any different vertices $v, w \in F$ and a positive integer $i < \frac{1}{2}m$, if $vv_{2i} \in E(G)$ and $wv_{2i+1} \in E(G)$, then $wv_{2i-1} \notin E(G)$ and $wv_{2i} \notin E(G)$.*

Proof. The claim follows easily from the maximality of the antipath A. If $wv_{2i-1} \in E(G)$, then $v_{2i}vv_1v_2 \ldots v_{2i-1}wv_{2i+1}v_{2i+2} \ldots v_m$ is a longer antipath, while if $wv_{2i} \in E(G)$, then $v_{2i-1}v_{2i-2} \ldots v_1vv_{2i}wv_{2i+1}v_{2i+2} \ldots v_m$ is a longer antipath. □

Claim 10. *For every integer $i \leq \frac{m-3}{4}$ there are at most $2|F| + 1$ edges from F to the set $\{v_{4i}, v_{4i+1}, v_{4i+2}, v_{4i+3}\}$.*

Proof. Let $Q = \{v_{4i}, v_{4i+1}, v_{4i+2}, v_{4i+3}\}$. If there is a vertex in F with 4 out-neighbours in Q, then by Claim 9 every other vertex in F does not have an edge to v_{4i}, v_{4i+2} and simultaneously to v_{4i+1} and v_{4i+3}, so there are at most $|F| + 3 \leq 2|F| + 1$ edges from F to Q. If two different vertices $v, w \in F$ have 3 out-neighbours in Q, then by Claim 9 either $vv_{4i+2} \notin E(G)$ or $wv_{4i+3} \notin E(G)$, similarly, either $wv_{4i+2} \notin E(G)$ or $vv_{4i+3} \notin E(G)$. But this implies that v and w have edges to both v_{4i} and v_{4i+1}, which also contradicts Claim 9. Therefore, there is at most 1 vertex in F with 3 out-neighbours in Q and all the other vertices in F have at most 2 out-neighbours in Q. This implies that there are at most $2|F| + 1$ edges from F to Q. □

Note that there are no edges from F to v_m as otherwise it creates the forbidden anticycle of length $m+1$. Hence, there are at most $|F|$ edges from F to the set $\{v_{m-1}, v_m\}$. Therefore, Claim 10 implies that there are at most

$$\frac{m+1}{2}|F| + \frac{m+1}{4} \leq \frac{k}{2}|F| + \frac{k}{4}$$

edges from F to $V(A)$. On the other hand, since all out-neighbours of the vertices in F are in $V(A)$ as otherwise we have a longer antipath in G, there are at least $\bar{\delta}^{\pm}(G)|F|$ edges from F to $V(A)$. This means that

$$\frac{k}{4} \geq \left(\bar{\delta}^{\pm}(G) - \frac{k}{2}\right)|F| \geq \left(\bar{\delta}^{\pm}(G) - \frac{k}{2}\right)\left(\bar{\delta}^{\pm}(G) - \frac{k}{2} + 1\right) > \frac{\sqrt{k}}{2} \cdot \frac{\sqrt{k}+1}{2} > \frac{k}{4},$$

which is a contradiction concluding the proof of Theorem 3.

We remark that a bit more detailed analysis of the above proof shows that the weaker bound $\bar{\delta}^{\pm}(G) > \frac{1}{2}(k + \sqrt{k-3} - 1)$ is already sufficient. By additionally considering proper rounding to the nearest integers in the used bounds, one can observe that the presented proof shows that Conjecture 2 is true for $k \leq 11$.

References

[1] L. Addario-Berry, F. Havet, L. Linhares Sales, B. Reed, and S. Thomassé. Oriented trees in digraphs. *Discrete Mathematics*, 313:967–974, 2013.

[2] S. A. Burr. Subtrees of directed graphs and hypergraphs. In *Proceedings of the Eleventh Southeastern Conference on Combinatorics, Graph Theory and Computing, Boca Raton, Congr. Numer*, volume 28, pages 227–239, 1980.

[3] B. Chen, X. Hou, and H. Zhou. Long antipaths and anticycles in oriented graphs. *Discrete Mathematics*, 348:114412, 2025.

[4] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of The London Mathematical Society*, 2:69–81, 1952.

[5] P. Erdős and T. Gallai. On maximal paths and circuits of graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 10:337–356, 1959.

[6] B. Jackson. Long paths and cycles in oriented graphs. *Journal of Graph Theory*, 5:145–157, 1981.

[7] P. Keevash, D. Kühn, and D. Osthus. An exact minimum degree condition for hamilton cycles in oriented graphs. *Journal of the London Mathematical Society*, 79, 2008.

[8] L. Kelly. Arbitrary orientations of hamilton cycles in oriented graphs. *Electronic Journal of Combinatorics*, 18, 2009.

[9] T. Klimošová and M. Stein. Antipaths in oriented graphs. *Discrete Mathematics*, 346:113515, 2023.

[10] J. Skokan and M. Tyomkyn. Alternating paths in oriented graphs with large semidegree. *arXiv:2406.03166*, 2024.

[11] M. Stein. Tree containment and degree conditions. *Discrete Mathematics and Applications*, 165:459–486, 2020.

[12] M. Stein. Oriented trees and paths in digraphs. *Surveys in Combinatorics 2024. London Mathematical Society Lecture Note Series. Cambridge University Press*, pages 271–295, 2024.

[13] M. Stein and C. Zárate-Guerén. Antidirected subgraphs of oriented graphs. *Combinatorics, Probability and Computing*, 33:1–21, 2024.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

SPECTRAL APPROACHES FOR d-IMPROPER CHROMATIC NUMBER

(EXTENDED ABSTRACT)

Krystal Guo* Ross J. Kang* Gabriëlle Zwaneveld*

Abstract

In this paper, we explore algebraic approaches to d-improper and t-clustered colourings, where the colouring constraints are relaxed to allow some monochromatic edges. Bilu [J. Comb. Theory Ser. B, 96(4):608-613, 2006] proved a generalization of the Hoffman bound for d-improper colourings. We strengthen this theorem by characterizing the equality case. In particular, if the Hoffman bound is tight for a graph G, then the d-improper Hoffman bound is tight for the strong product $G \boxtimes K_{d+1}$. Moreover, we prove d-improper analogous for the inertia bound by Cvetkovíc and the multi-eigenvalue lower bounds of Elphick and Wocjan.

We conjecture an equality between the chromatic number of a graph G and the d-improper chromatic number of its strong product with a complete graph, $G \boxtimes K_{d+1}$, and prove the conjecture in special graph classes, including perfect graphs and graphs with chromatic number at most 4. Other supporting evidence for the conjecture includes a fractional analogue, a clustered analogue, and various spectral relaxations of the equality.
Keywords: graph colouring, eigenvalue bounds, improper colouring, clustered colouring

Improper colouring is a natural relaxation of proper colouring where we allow monochromatic edges but only in some controlled fashion. This type of colouring, also called defective colouring, permits a prescribed maximum degree in the monochromatic subgraphs. Throughout the paper, we use $\chi^d(G)$ to denote the *d-improper chromatic number* of G, the least number of parts in a partition of the vertex set of G so that each part induces maximum degree at most d. Note that $\chi^0(G)$ coincides with the usual chromatic number $\chi(G)$ of G.

The idea of "tolerating" a controlled amount of error in proper colourings arises naturally in certain optimization settings. For example, it occurs in the applied contexts of ad hoc radio communication networks [10] and circuit design [1]. Within graph theory, such colourings have been studied going back to a paper of Lovász in 1966 [15]. They have been considered in structural contexts [6, 11] and probabilistic/extremal contexts [2, 13, 9]. See [17] for a comprehensive recent survey. Colourings and graph products have also been studied in [8].

*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands. (`{k.guo,j.r.kang,g.c.zwaneveld}@uva.nl`)

We leverage the use of algebraic methods for d-improper colourings. In Section 1, we generalize other spectral bounds to d-improper colourings. Moreover, we look at when equality holds in a generalization by Bilu [3] of the well-known chromatic number bound of Hoffman [12]. We prove in Lemma 4 that $G \boxtimes K_{d+1}$ admits a d-improper Hoffman colouring when G admits a proper Hoffman colouring. Incidentally, we also find that for such graphs G we have $\chi(G) = \chi^d(G \boxtimes K_{d+1})$. In Section 2, we dwell on the tantalizing prospect that equality holds for all G as stated in the next conjecture:

Conjecture 1. *For every graph G and every integer $d \geq 0$, we have*

$$\chi(G) = \chi^d(G \boxtimes K_{d+1}). \tag{1}$$

1 Spectral bounds for d-improper chromatic number

An important goal in graph theory is to develop good bounds on the chromatic number, which is NP-hard to compute, using graph invariants that are easier to compute, such as the degrees or graph eigenvalues. In this section, we generalize some well-known spectral lower bounds for the chromatic number to the d-improper chromatic number. Let λ_i be the ith largest eigenvalue of a graph G.

Bilu [3] has generalized the well-known chromatic number bound of Hoffman [12] to also hold for d-improper colourings (the classic result is the $d = 0$ specialization below).

Theorem 2 (Bilu [3]). *For every integer $d \geq 0$ and every graph G, we have*

$$\chi^d(G) \geq \frac{\lambda_1 - \lambda_n}{d - \lambda_n} = 1 - \frac{\lambda_1 - d}{\lambda_n - d}.$$

To grasp the Bilu bound better, we define a *d-improper Hoffman colouring* to be a vertex partition of G into parts inducing maximum degree d having $1 - \frac{\lambda_1 - d}{\lambda_n - d}$ parts, so that G exactly attains the bound. In this paper, we have found necessary and sufficient conditions for when equality holds in the generalized Hoffman bound. We use $m = \chi^d(G)$, then

Theorem 3. *For every integer $d \geq 0$ and every graph G, we have*

$$\chi^d(G) \geq \frac{\lambda_1 - \lambda_n}{d - \lambda_n} = 1 - \frac{\lambda_1 - d}{\lambda_n - d}.$$

Furthermore, if equality holds for some graph G, then the following hold:

(a) the multiplicity of λ_n in G is at least $m - 1$;

(b) if G is connected, then in any d-improper Hoffman colouring all colour classes are d-regular and the partition into colour classes is weight-regular with respect to the weights u_v where u is the Perron eigenvector of A;

(c) if G has a unique Hoffman colouring up to permutation of the colour classes, then λ_n has multiplicity exactly $m - 1$; and

(d) if G is regular, then the partition into $\chi^d(G)$ colour classes is equitable, and every vertex has d neighbours in its colour class and $d - \lambda_n$ neighbours in every other colour class.

Notice that the Hoffman bound for $\chi(G)$ and Bilu's bound for $\chi^d(G \boxtimes K_{d+1})$ are the same. We use this to construct d-improper Hoffman colourable graphs from Hoffman colourable graphs.

Lemma 4. *If G admits a Hoffman colouring, then $G \boxtimes K_{d+1}$ admits a d-improper Hoffman colouring.*

Proof. If G admits a Hoffman colouring, then $\chi(G) = 1 - \frac{\lambda_1(G)}{\lambda_n(G)}$. Moreover, every eigenvalue of $G \boxtimes K_{d+1}$ is either -1 or $(d+1)\lambda_i(G) + d$. Since -1 is not largest nor smallest eigenvalue of $G \boxtimes K_{d+1}$, we have $\lambda_1(G \boxtimes K_{d+1}) = (d+1)\lambda_1(G) + d$ and $\lambda_n(G \boxtimes K_{d+1}) = (d+1)\lambda_n(G) + d$. Thus,

$$\begin{aligned} 1 - \frac{\lambda_1(G)}{\lambda_n(G)} &= \chi(G) \geq \chi^d(G \boxtimes H) \geq \frac{\lambda_1(G \boxtimes K_{d+1}) - \lambda_n(G \boxtimes K_{d+1})}{d - \lambda_n(G \boxtimes K_{d+1})} \\ &= \frac{(d+1)(\lambda_1(G) - \lambda_n(G))}{d - (d+1)\lambda_n(G) - d} = 1 - \frac{\lambda_1(G)}{\lambda_n(G)}. \end{aligned}$$

□

Next, we generalize the inertia bound of Cvetković [5] to d-improper colourings. This bound uses the number of positive and negative eigenvalues of various matrices compatible with G, including its adjacency matrix. In the case of d-improper chromatic number, we use the number of eigenvalues of a matrix A greater than d is denoted by $n_d^+(A)$ and the number of eigenvalues smaller than $-d$ denoted by $n_d^-(A)$. Lastly, $\alpha_d(G)$ denotes the size of the maximal induced subgraph of G with maximum degree at most d.

Theorem 5 (d-improper inertia bound)**.** *Let G be a graph and let A be a symmetric matrix such that $|A_{u,v}| \leq 1$ for all u and v and $A_{u,v} = 0$ if $u \not\sim v$ in G. Then $\alpha^d(G) \leq \min\{n - n_d^+(A), n - n_d^-(A)\}$.*

Corollary 6. *Let G be a graph and let A be a symmetric matrix such that $|A_{u,v}| \leq 1$ for all u and v and $A_{u,v} = 0$ if $u \not\sim v$. Then $\chi^d(G) \geq \left\lceil \max\left\{\frac{n}{n - n_d^+(A)}, \frac{n}{n - n_d^-(A)}\right\}\right\rceil$*

Lastly, we generalize the multi-eigenvalue lower bounds of Elphick and Wocjan [7] to also hold for d-improper colourings. Let D be the diagonal matrix containing the degrees of the vertices of G, so $D_{uu} = d(u)$. Then $L = D - A$ is called the *Laplacian of G* and $Q = D + A$ is called the *signless Laplacian of G*. Suppose the eigenvalues of L are denoted by $\mu_1 \geq \ldots \geq \mu_n = 0$ and the eigenvalues of Q by $\theta_1 \geq \ldots \geq \theta_n \geq 0$.

Theorem 7. *For all graphs G and positive integers d and $m \leq n$, the d-improper chromatic number of G satisfies*

$$\chi^d(G) \geq 1 + \frac{-dm + \sum_{i=1}^m \lambda_i}{dm - \sum_{i=1}^m \lambda_{n+1-i}} \tag{2}$$

$$\chi^d(G) \geq 1 + \frac{-dm + \sum_{i=1}^m \lambda_i}{dm + \sum_{i=1}^m (\mu_i - \lambda_i)} \tag{3}$$

$$\chi^d(G) \geq 1 + \frac{-dm + \sum_{i=1}^m \lambda_i}{dm + \sum_{i=1}^m (\lambda_i + \mu_i - \theta_i)} \tag{4}$$

$$\chi^d(G) \geq 1 + \frac{-dm + \sum_{i=1}^m \lambda_i}{dm + \sum_{i=1}^m (\lambda_i + \mu_{n+1-i} - \theta_{n+1-i})} \tag{5}$$

Plugging in $m = 1$ in Equation 2 gives the d-improper Hoffman bound. Similarly, $m = 1$ in Equation 3 corresponds to a generalization of Nikoforov's bound [16] to the d-improper chromatic number and $m = 1$ in Equations 4 and 5 correspond to the d-improper versions of Kolotilina bounds [14].

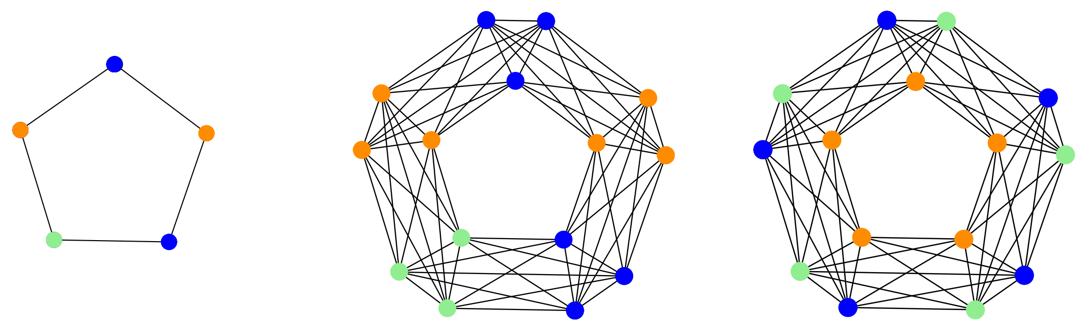
2 The d-improper chromatic number of $G \boxtimes K_{d+1}$

In Lemma 4, we saw that $G \boxtimes K_{d+1}$ admits a d-improper Hoffman colouring when G admits a proper Hoffman colouring. Moreover, the eigenvalues of the adjacency matrix of $G \boxtimes K_{d+1}$ are -1 and $(d+1)\lambda_i(G) + d$ where $\lambda_i(G)$ is an eigenvalue of the adjacency matrix of G. Hence, Bilu's bound for these graphs equals

$$\chi^d(G \boxtimes K_{d+1}) \geq \frac{(d+1)\lambda_1 + d - ((d+1)\lambda_n + d)}{d - ((d+1)\lambda_n + d)} = \frac{\lambda_1 - \lambda_n}{-\lambda_n} = 1 - \frac{\lambda_1}{\lambda_n},$$

which is the same as the Hoffman bound for $\chi(G)$. Consequently, for such graphs G, we have $\chi(G) = \chi^d(G \boxtimes K_{d+1})$. Could this be true for all graphs G, rather than just Hoffman colourable graphs? This is the assertion in Conjecture 1.

The difficulty of this conjecture is to prove $\chi^d(G \boxtimes K_{d+1}) \geq \chi(G)$. To see the converse inequality $\chi(G) \geq \chi^d(G \boxtimes K_{d+1})$, consider any proper colouring $c : V(G) \to \mathbb{N}$ of G and construct a d-improper colouring of $G \boxtimes K_{d+1}$ by using $c(v)$ on (v, i), for each i see Figure 1. Since not every d-improper colouring of $G \boxtimes K_{d+1}$ comes from a proper colouring of G, it is far from obvious how a similar argument could be used for the converse.



(a) Colouring of C_5

(b) The lifted 2-improper colouring of $C_5 \boxtimes K_3$ from the colouring of C_5 in (a).

(c) 2-improper colouring of $C_5 \boxtimes K_3$ that is not lifted from any colouring of C_5

Figure 1: Colourings of G can be lifted to d-improper colourings of $G \boxtimes K_t$, but not all d-improper colourings of $G \boxtimes K_{d+1}$ are lifted from some colouring of G.

In this section, we will present some evidence in support of our conjecture. In particular, we have shown the conjecture to be true in some special cases.

For $d = 0$ the statement follows from the fact that 0-improper colourings coincide with proper colourings and that $G = G \boxtimes K_1$ for all graphs G. Moreover, since every connected component with maximum degree 1 has size at most 2 the 1-improper and 2-clustered chromatic number are the same. Hence, the result for $d = 1$ follows from Corollary 11.

Moreover, we prove that Conjecture 1 holds for all graphs G satisfying $\chi(G) \leq 4$, using a neat argument of Buys [4]. This implies that our conjecture holds for all planar graphs.

Theorem 8. *For every graph G with $\chi(G) \leq 4$ and every integer $d \geq 0$, we have*

$$\chi(G) = \chi^d(G \boxtimes K_{d+1}).$$

Furthermore, using similar arguments as for Hoffman colourable graphs, we can prove that the conjecture holds for all perfect graphs and all graphs G attaining equality in either Nikoforov's bound [16] or Kolotilina's bound [14].

We also proved two weaker forms of Conjecture 1. First, we proved a fractional analogue, where χ_f^d stands for the fractional d-improper chromatic number.

Theorem 9. *For every graph G and all positive integers d, we have*

$$\chi_f^d(G \boxtimes K_{d+1}) = \chi_f(G).$$

Secondly, we prove the clustered analogue of Conjecture 1. A colouring is *t-clustered* if every monochromatic component contains at most t vertices. The *t-clustered chromatic number* of a graph G, denoted $\chi^{\underline{t}}(G)$, is the least number of colours in a t-clustered colouring of G. Since every monochromatic component of t vertices has maximum degree $t-1$, every t-clustered colouring is a $t-1$-improper colouring, and so, for all graphs G, $\chi^{\underline{t}}(G) \geq \chi^{t-1}(G)$. In our paper, we give the analogous statement of Lemma 4 for t-improper colourings.

In our paper, we prove that, starting with an ℓt-clustered colouring of $G \boxtimes K_t$, we are able to obtain an ℓ-clustered colouring of G such that the colour of $v \in V(G)$ is the colour of one of the vertices $(v, i) \in V(G \boxtimes K_t)$.

Theorem 10. *For every graph G and positive integers t and ℓ, we have*

$$\chi^{\underline{\ell}}(G) = \chi^{\underline{\ell t}}(G \boxtimes K_t).$$

Since a 1-clustered colouring is a proper colour, we derive an immediate consequence the clustered analogue of Conjecture 1.

Corollary 11. *For every graph G and positive integer t, we have*

$$\chi(G) = \chi^{\underline{t}}(G \boxtimes K_t).$$

References

[1] Dehia Ait-Ferhat. *Design of exact solutions for the manufacturing of "vias" using DSA technology. (Conception de solutions exactes pour la fabrication de "vias" en utilisant la technologie DSA).* PhD thesis, Grenoble Alpes University, France, 2018.

[2] Noga Alon, Guoli Ding, Bogdan Oporowski, and Dirk Vertigan. Partitioning into graphs with only small components. *J. Combin. Theory Ser. B*, 87(2):231–243, 2003.

[3] Yonatan Bilu. Tales of Hoffman: Three extensions of Hoffman's bound on the graph chromatic number. *Journal of Combinatorial Theory, Series B*, 96(4):608–613, 2006.

[4] Pjotr Buys, 2023. Personal communication.

[5] Dragoš Cvetkovíc. Chromatic number and the spectrum of a graph. *Publ. Inst. Math.(Beograd)*, 14(28):25–38, 1972.

[6] Katherine Edwards, Dong Yeap Kang, Jaehoon Kim, Sang-il Oum, and Paul Seymour. A relative of Hadwiger's conjecture. *SIAM J. Discrete Math.*, 29(4):2385–2388, 2015.

[7] Clive Elphick and Pawel Wocjan. Unified spectral bounds on the chromatic number. *Discussiones Mathematicae Graph Theory*, 35(4):773–780, 2015.

[8] Louis Esperet and David R Wood. Colouring strong products. *European Journal of Combinatorics*, 121:103847, 2024.

[9] António Girão, Freddie Illingworth, Alex Scott, and David R. Wood. Defective coloring of hypergraphs. *Random Structures Algorithms*, 64(3):663–675, 2024.

[10] Frédéric Havet, Ross J. Kang, and Jean-Sébastien Sereni. Improper coloring of unit disk graphs. *Networks*, 54(3):150–164, 2009.

[11] Jan van den Heuvel and David R. Wood. Improper colourings inspired by Hadwiger's conjecture. *J. Lond. Math. Soc. (2)*, 98(1):129–148, 2018.

[12] Alan J. Hoffman. On eigenvalues and colorings of graphs. In *Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., 1969)*, pages 79–91. Academic Press, New York-London, 1970.

[13] Ross J. Kang and Colin McDiarmid. The t-improper chromatic number of random graphs. *Combin. Probab. Comput.*, 19(1):87–98, 2010.

[14] L Kolotilina. Inequalities for the extreme eigenvalues of block-partitioned hermitian matrices with applications to spectral graph theory. *Journal of Mathematical Sciences*, 176(1), 2011.

[15] L. Lovász. On decomposition of graphs. *Studia Sci. Math. Hungar.*, 1:237–238, 1966.

[16] Vladimir Nikiforov. Chromatic number and spectral radius. *Linear algebra and its applications*, 426(2-3):810–814, 2007.

[17] David R. Wood. Defective and clustered graph colouring. *Electron. J. Combin.*, #DS23, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On graphs without cycles of length 0 modulo 4

(Extended abstract)

Ervin Győri[*] Binlong Li[†] Nika Salia[‡] Casey Tompkins[§] Kitti Varga[¶]
Manran Zhu[‖]

Abstract

Bollobás proved that for every k and ℓ such that $k\mathbb{Z}+\ell$ contains an even number, an n-vertex graph containing no cycle of length ℓ mod k can contain at most a linear number of edges. The precise (or asymptotic) value of the maximum number of edges in such a graph is known for very few pairs ℓ and k. In this work we precisely determine the maximum number of edges in a graph containing no cycle of length 0 mod 4.

1 Introduction

It is well-known that n-vertex graphs containing no even cycles can contain at most $\lfloor \frac{3}{2}(n-1)\rfloor$ edges. On the other hand, if only a set of odd cycles are forbidden, then taking a balanced complete bipartite graph yields $\lfloor \frac{n^2}{4}\rfloor$ edges, and this is sharp for sufficiently large n [16]. Given these observations, it was natural to consider the extremal problem where for natural numbers k and ℓ such that $k\mathbb{Z}+\ell$ contains an even number, all cycles of length ℓ mod k are forbidden. It was conjectured by Burr and Erdős [8] that such a graph could contain at most a linear number of edges. This conjecture was proved by Bollobás [3].

Given the result of Bollobás, it is interesting to determine the smallest constant $c_{\ell,k}$ (where $k\mathbb{Z}+\ell$ contains an even number) such that every n-vertex graph with $c_{\ell,k}n$ edges must contain

[*]Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: gyori.ervin@renyi.hu. Supported by NKFIH, grant K132696.

[†]School of Mathematics and Statistics, Northwestern Polytechnical University and Budapest Joint Research Center for Combinatorics, Northwestern Polytechnical University, Xi'an, 710072, China. E-mail: binlongli@nwpu.edu.cn. Supported by NSFC (12171393) and Shaanxi Fundamental Science Research Project for Mathematics and Physics (22JSZ009).

[‡]King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. E-mail: salianika@gmail.com. Supported by NKFIH, grant K132696.

[§]Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: ctompkins496@gmail.com. Supported by NKFIH, grants K132696 and K135800.

[¶]Budapest University of Technology and Economics and MTA-ELTE Egerváry Research Group, Budapest, Hungary. E-mail: vkitti@renyi.hu.

[‖]Corvinus University, Center for Collective Learning, Budapest, Hungary. E-mail: manran.zhu@uni-corvinus.hu. Supported by HORIZON, grant 101086712

a cycle of length $\ell \bmod k$. The problem of finding such an optimal $c_{\ell,k}$ was mentioned by Erdős in [9]. Various improvements to the general bounds on $c_{\ell,k}$ have been obtained [18, 19, 20] culminating in a recent result of Sudakov and Verstraëte [17] showing that for $3 \leq \ell < k$, the value of $c_{\ell,k}$ is within an absolute constant of the maximum number of edges in a k-vertex C_ℓ-free graph divided by k. Thus, for even $\ell \geq 4$ the general problem of determining $c_{\ell,k}$ is at least as difficult as determining the Turán number of C_ℓ (for which we only know the order of magnitude when $\ell \in \{4, 6, 10\}$).

While the focus of the presentation is to determine the number of edges (or equivalently the average degree) which forces an n-vertex graph to contain a cycle from a given modulo class, we mention that analogous results for other parameters including minimum degree, connectivity and chromatic number have been heavily investigated. See Liu and Ma [12] and Gao, Huo, Liu and Ma [10] for the solution of many conjectures in this direction, as well as Lyngsie and Merker [14] in the 3-connected case.

The precise value of $c_{\ell,k}$ is known for very few pairs ℓ and k. As mentioned above it is well-known that $c_{0,2} = \frac{3}{2}$. It was proved that $c_{0,3} = 2$ by Chen and Saito [4], which resolved a conjecture of Barefoot et al [2]. The n-vertex graph avoiding all cycles of length $0 \bmod 3$ with the maximum number of edges is the complete bipartite graph $K_{2,n-2}$. In fact Chen and Saito [4] proved a stronger result (also conjectured by Barefoot et al [2]) that a graph of minimum degree at least 3 contains a cycle of length $0 \bmod 3$, which implies the aforementioned results.

Dean, Kaneko, Ota and Toft [6] (see also Saito [15]) showed that every n-vertex 2-connected graph of minimum degree at least 3 either contains a cycle of length $2 \bmod 3$ or is isomorphic to K_4 or $K_{3,n-3}$. From this result it is easily deduced that for n sufficiently large, $K_{3,n-3}$ maximizes the number of edges in a graph not containing a cycle of length $2 \bmod 3$. Consequently, $c_{2,3} = 3$.

We now discuss the case of cycles of length $1 \bmod 3$ which has been resolved only recently. Dean, Kaneko, Ota and Toft [6] proved that every 2-connected graph of minimum degree at least 3 and with no cycle of length $1 \bmod 3$ is a Petersen graph. This result was strengthened by Lu and Yu [13] who showed that every connected graph of minimum degree at least 3 and with no cycle of length $1 \bmod 3$ contains a Petersen graph as an induced subgraph. However, it is not clear how one could derive a result on the maximum number of edges from these results. Nonetheless, $c_{1,3} = \frac{5}{3}$ was recently obtained by Bai, Li, Pan and Zhang [1]. A general estimate of $c_{\ell,2} \leq \ell + 2$ was given in the original paper of Erdős [8].

Gao, Li, Ma and Xie [11] proved that an n-vertex graph G with at least $\frac{5}{2}(n-1)$ edges contains two consecutive even cycles unless $4 \mid (n-1)$ and every block of G is isomorphic to K_5. This result settled the $k = 2$ case of conjecture of Verstraëte [21] about the maximum number of edges in graphs avoiding cycles of k consecutive lengths. As a consequence of this result Gao, Li, Ma and Xie proved that $c_{2,4} = \frac{5}{2}$.

We will consider the problem of maximizing edges in a graph containing no cycle of length $0 \bmod 4$. This is the last remaining class modulo 4 since the others contain only odd numbers. An extensive investigation of such graphs was undertaken by Dean, Lesniak and Saito [7]. They proved, among several other results, that $c_{0,4} \leq 2$.

Our main result is an exact determination of $c_{0,4}$. In fact we determine a sharp upper bound on the number of edges in a graph containing no cycle of length $0 \bmod 4$, and as a consequence we obtain $c_{0,4} = \frac{19}{12}$.

Theorem 1. *Let G be an n-vertex graph. If $e(G) > \lfloor \frac{19}{12}(n-1) \rfloor$, then G contains a cycle of length* $0 \bmod 4$.

Constructions attaining this upper bound for every $n \geq 2$ are given in the following section.

2 Extremal graphs

Define L_8 and L_{13} to be the graphs show in Figure 1.

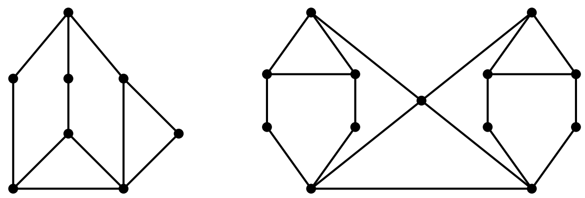


Figure 1: The graphs L_8 and L_{13}.

For $n \geq 2$, we define the graph G_n as follows: Let

$$n-1 = 12q_1 + r_1,\ 0 \leq r_1 \leq 11;$$
$$r_1 = 7q_2 + r_2,\ 0 \leq r_2 \leq 6;$$
$$r_2 = 2q_3 + r_3,\ 0 \leq r_3 \leq 1.$$

Let G_n be a connected graph consisting of q_1 blocks isomorphic to L_{13}, q_2 blocks isomorphic to L_8, q_3 blocks isomorphic to K_3 and r_3 blocks isomorphic to K_2. One can compute that G_n contains no (0 mod 4)-cycle and $e(G_n) = \lfloor \frac{19}{12}(n-1) \rfloor$.

Let $\mathcal{C}_{0 \bmod 4}$ be the set of all (0 mod 4)-cycles. We have

$$\mathrm{ex}(n, \mathcal{C}_{0 \bmod 4}) = \left\lfloor \frac{19}{12}(n-1) \right\rfloor.$$

3 Key Notions

In this section we briefly outline some basic notions and lemmas involved in the proof.

Let G be a graph and $x, y \in V(G)$. A path from x to y is called an *(x,y)-path.* If X, Y are two subgraphs of G or subsets of $V(G)$, then a path from X to Y is an (x,y)-path with $x \in X$, $y \in Y$, and all internal vertices in $V(G) \backslash (X \cup Y)$. A path (cycle) is *even* (*odd*) if its length is even (odd). The graph consisting of an odd cycle C, a path P_1 from x to C and a path P_2 from C to y with $V(P_1) \cap V(P_2) = \emptyset$ (not excluding the case that P_1 and/or P_2 are trivial), is called an *adjustable path* from x to y (or briefly, an *adjustable (x,y)-path*). Notice that an adjustable (x,y)-path contains both an even (x,y)-path and an odd (x,y)-path. For a path P or a cycle C, we denote by $|P|$ or $|C|$ its length. We write $\mathrm{end}(P) = \{x, y\}$ if P is a path or adjustable path from x to y.

Denote by Θ a graph consisting of three internally-disjoint paths from a vertex x to a vertex y, and denote by Θ^e such a graph where all three paths are even. For $k = 3, 4$, define H_k^o (respectively H_k^e) to be a subdivision of K_4 such that each edge of some k-cycle in the K_4 corresponds to an odd path (respectively, even path). Define the *odd necklace* N^o to be a graph consisting of an adjustable (x_1, x_2)-path R_1, an adjustable (x_2, x_3)-path R_2, an adjustable (x_3, x_1)-path R_3, such that R_1, R_2, R_3 are pairwise internally-disjoint.

Lemma 1. *Each of* Θ^e, N^o, H_3^e, H_4^o, H_4^e *contains a* $(0 \bmod 4)$*-cycle.*

Lemma 2. *Every non-planar graph contains a* $(0 \bmod 4)$*-cycle.*

Lemma 3. *Let* C *be an even cycle and* P_i, $i = 1, 2, 3$, *be even bridges of* C.
(1) If P_1 *has an even span, then* $C \cup P_1$ *contains a* $(0 \bmod 4)$*-cycle.*
(2) If P_1, P_2 *are crossed on* C, *then* $C \cup P_1 \cup P_2$ *contains a* $(0 \bmod 4)$*-cycle.*
(3) If P_1, P_2, P_3 *are pairwise internally-disjoint, then* $C \cup P_1 \cup P_2 \cup P_3$ *contains a* $(0 \bmod 4)$*-cycle.*

Lemma 4. *Let* C *be an even cycle,* P_1, P_2 *be crossed bridges of* C, *and* R *be an adjustable path from* $P_2 - C$ *to* C, *such that* P_1 *is even and* P_1, R *are internally-disjoint. Then* $C \cup P_1 \cup P_2 \cup R$ *contains a* $(0 \bmod 4)$*-cycle.*

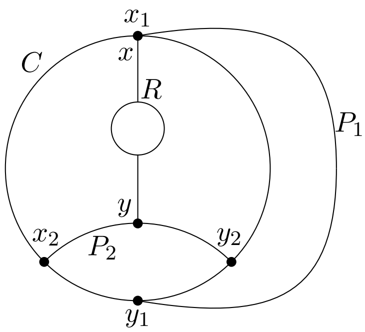


Figure 2: Construction of Lemma 4.

Lemma 5. *Let* C *be an even cycle,* P_1, P_2 *be two vertex-disjoint bridges of* C *with even spans, and* R *be an adjustable path from* $P_1 - C$ *to* $P_2 - C$, *such that* C *and* R *are vertex-disjoint. Then* $C \cup P_1 \cup P_2 \cup R$ *contains a* $(0 \bmod 4)$*-cycle.*

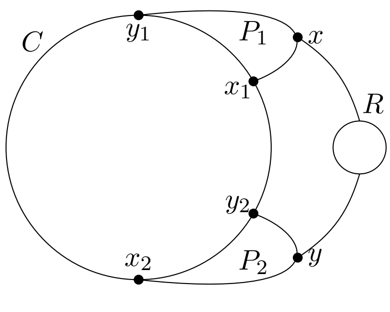


Figure 3: Construction of Lemma 5.

Lemma 6. *Let* C_1, C_2 *be odd cycles with* $|C_1| \equiv |C_2| \bmod 4$, *and* P_1, P_2, P_3 *be vertex-disjoint paths from* C_1 *to* C_2.
(1) If C_1, C_2 *are vertex-disjoint, and* $|P_1| + |P_2|$ *even, then* $C_1 \cup C_2 \cup P_1 \cup P_2$ *contains a* $(0 \bmod 4)$*-cycle.*

(2) If $V(C_1) \cap V(C_2) = \{x\}$*,* P_1 *is even and* $x \notin V(P_1)$*, then* $C_1 \cup C_2 \cup P_1$ *contains a* $(0 \bmod 4)$*-cycle.*
(3) If C_1, C_2 *are vertex-disjoint, then* $C_1 \cup C_2 \cup P_1 \cup P_2 \cup P_3$ *contains a* $(0 \bmod 4)$*-cycle.*

Lemma 7. *Let* C_1, C_2, C_3 *be three odd cycles with* $|C_1| \equiv |C_2| \equiv |C_3| \bmod 4$ *such that they pairwise intersect at a vertex* x*. Let* P_i *be a path from* C_i *to* C_{i+1} *that is vertex-disjoint with* C_{i+2}*,* $i = 1, 2, 3$ *(the subscripts are taken modulo 3), such that* P_1, P_2, P_3 *are pairwise internally-disjoint. Then* $C_1 \cup C_2 \cup C_3 \cup P_1 \cup P_2 \cup P_3$ *contains a* $(0 \bmod 4)$*-cycle.*

Lemma 8. *Let* C_1, C_2, C_3 *be three odd cycles with* $|C_1| \equiv |C_2| \equiv |C_3| \bmod 4$ *such that they pairwise intersect at a vertex* x*. Let* P_i *be a path from a vertex* y *to* $C_i - x$*,* $i = 1, 2, 3$*, where* $y \notin V(C_1) \cup V(C_2) \cup V(C_3)$*, such that* P_1, P_2, P_3 *are internally-disjoint with* C_1, C_2, C_3 *and are pairwise internally-disjoint. Then* $C_1 \cup C_2 \cup C_3 \cup P_1 \cup P_2 \cup P_3$ *contains a* $(0 \bmod 4)$*-cycle.*

Lemma 9. *If* G *is a bipartite graph of order* $n \geq 4$ *containing no* $(0 \bmod 4)$*-cycle, then* $e(G) \leq \lfloor \frac{3}{2}(n-2) \rfloor$*.*

Let $\{x, y\}$ be a cut of G, and H be a component of $G - \{x, y\}$. The graph G' obtained from G by first removing all the edges between $\{x, y\}$ and H, and then adding the edges in $\{xz : yz \in E(G), z \in V(H)\} \cup \{yz : xz \in E(G), z \in V(H)\}$, is called a *switching* of G at $\{x, y\}$.

Lemma 10. *If* G *has a 2-cut* $\{x, y\}$ *and* G' *is a switching of* G *at* $\{x, y\}$*, then* $e(G') = e(G)$ *and* G' *has a* $(0 \bmod 4)$*-cycle if and only if so does* G*.*

References

[1] Yandong Bai, Binlong Li, Yufeng Pan, and Shenggui Zhang. On graphs without cycles of length 1 modulo 3. *arXiv preprint arXiv:2503.03504*, 2025.

[2] CA Barefoot, LH Clark, Jack Douthett, RC Entringer, and MR Fellows. Cycles of length 0 modulo 3 in graphs. *preprint*, 1991.

[3] Bela Bollobás. Cycles modulo k. *Bulletin of the London Mathematical Society*, 9(1):97–98, 1977.

[4] GT Chen and Akira Saito. Graphs with a cycle of length divisible by three. *Journal of Combinatorial Theory, Series B*, 60(2):277–292, 1994.

[5] CRJ Clapham, A Flockhart, and John Sheehan. Graphs without four-cycles. *Journal of Graph theory*, 13(1):29–47, 1989.

[6] N. Dean, A. Kaneko, K. Ota, and B. Toft. Cycles modulo 3. *Dimacs Technical Report*, 91(32), 1991.

[7] Nathaniel Dean, Linda Lesniak, and Akira Saito. Cycles of length 0 modulo 4 in graphs. *Discrete Mathematics*, 121(1-3):37–49, 1993.

[8] Paul Erdős. Some recent problems and results in graph theory, combinatorics and number theory. *Proceedings of the Seventh Southeastern Conference on Combinatorics, Graph Theory, and Computing (Louisiana State Univ., Baton Rouge, La., 1976), Congress. Numer. XVII*, pages 3–14, 1976.

[9] Paul Erdős. Some of my favourite problems in number theory, combinatorics, and geometry. *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo*, 2(2):165–186, 1995.

[10] Jun Gao, Qingyi Huo, Chun-Hung Liu, and Jie Ma. A unified proof of conjectures on cycle lengths in graphs. *International Mathematics Research Notices*, 2022(10):7615–7653, 2022.

[11] Jun Gao, Binlong Li, Jie Ma, and Tianying Xie. On two cycles of consecutive even lengths. *Journal of Graph Theory*, 106(2):225–238, 2024.

[12] Chun-Hung Liu and Jie Ma. Cycle lengths and minimum degree of graphs. *Journal of Combinatorial Theory, Series B*, 128:66–95, 2018.

[13] Mei Lu and Zhengguang Yu. Cycles of length 1 modulo 3 in graph. *Discrete Applied Mathematics*, 113(2-3):329–336, 2001.

[14] Kasper S Lyngsie and Martin Merker. Cycle lengths modulo k in large 3-connected cubic graphs. *arXiv preprint arXiv:1904.05076*, 2019.

[15] Akira Saito. Cycles of length 2 modulo 3 in graphs. *Discrete Mathematics*, 101(1-3):285–289, 1992.

[16] Miklós Simonovits. Extremal graph problems with symmetrical extremal graphs. additional chromatic conditions. *Discrete Mathematics*, 7(3-4):349–376, 1974.

[17] Benny Sudakov and Jacques Verstraëte. The extremal function for cycles of length ℓ. *The Electronic Journal of Combinatorics*, 24(1), 2017.

[18] Carsten Thomassen. Graph decomposition with applications to subdivisions and path systems modulo k. *Journal of Graph Theory*, 7(2):261–271, 1983.

[19] Carsten Thomassen. *Paths, circuits and subdivisions.* Danmarks Tekniske Højskole. Matematisk Institut, 1986.

[20] Jacques Verstraëte. On arithmetic progressions of cycle lengths in graphs. *Combinatorics, Probability and Computing*, 9(4):369–373, 2000.

[21] Jacques Verstraëte. Extremal problems for cycles in graphs. In *Recent Trends in Combinatorics*, pages 83–116. Springer, 2016.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On the geometric k-colored crossing number

(Extended abstract)

Benedikt Hahn* Bettina Klinz† Birgit Vogtenhuber*

Abstract

We study the *geometric k-colored crossing number* of complete graphs $\overline{\overline{\mathrm{cr}}}_k(K_n)$, which is the smallest number of monochromatic crossings in any k-edge colored straight-line drawing of K_n. We substantially improve asymptotic upper bounds on $\overline{\overline{\mathrm{cr}}}_k(K_n)$ for $k = 2, \ldots, 10$ by developing a procedure for general k that derives k-edge colored drawings of K_n for arbitrarily large n from initial drawings with a low number of monochromatic crossings. We obtain the latter by heuristic search, employing a MAX-k-CUT-formulation of a subproblem in the process.

1 Introduction

A *drawing* Γ *of a graph* G is a representation of G in $\mathbb{R}^2$ where vertices are represented as distinct points, edges are represented as simple continuous curves connecting their endpoints, and no curve passes through the representation of a vertex. For simplicity, we assume that no two curves share more than finitely points or a tangent point and that no three edges share a point in their relative interior. A *crossing* in Γ is a point in the relative interior of two curves. For brevity, we will mostly refer to the elements of Γ as vertices and edges.

Crossing minimization for non-planar graphs is of great interest from both a theoretical and a practical point of view. The *crossing number* $\mathrm{cr}(G)$ of G is the minimum number of crossings $\mathrm{cr}(\Gamma)$ in any drawing Γ of G. A plethora of variants of crossing numbers have been studied; see for example the survey of Schaefer [24]. Despite intensive research, various important problems — such as determining $\mathrm{cr}(K_n)$ — still remain open for the "original" crossing number, and the same holds true for many relevant variants. One such variant and the topic of this work is the *geometric k-colored crossing number* $\overline{\overline{\mathrm{cr}}}_k(G)$. It is defined as

$$\overline{\overline{\mathrm{cr}}}_k(G) \coloneqq \min_{\Gamma} \min_{G=G_1\cup\ldots\cup G_k} \sum_{i=1}^{k} \mathrm{cr}(\Gamma|_{G_i}), \tag{1}$$

*Institute of Algorithms and Theory, Graz University of Technology, Austria.
E-mail: benedikt.hahn@tugraz.at, birgit.vogtenhuber@tugraz.at
†Institute of Discrete Mathematics, Graz University of Technology, Austria.
E-mail: klinz@math.tugraz.at

where Γ ranges over all *straight-line drawings* of G (i.e., drawings in which the edges are straight-line segments). Equivalently, $\overline{\overline{\mathrm{cr}}}_k(G)$ is the minimum number of *monochromatic crossings* in any *straight-line k-edge-colored drawing of G*. Straight-line drawings are also called geometric graphs, which motivates the name *geometric k-colored crossing number* [1]. The geometric k-colored crossing number is closely related to *geometric thickness*, which is the minimum k such that $\overline{\overline{\mathrm{cr}}}_k(G) = 0$.

For $k = 1$, $\overline{\overline{\mathrm{cr}}}_k(G)$ is the classical *rectilinear crossing number* of a graph (mostly denoted by $\overline{\mathrm{cr}}(G)$). Determining $\overline{\overline{\mathrm{cr}}}_1(G)$ is $\exists\mathbb{R}$-hard [8, 12] and exact values for $\overline{\overline{\mathrm{cr}}}_k(G)$ are known only for few graph classes. In particular, despite intensive research, $\overline{\overline{\mathrm{cr}}}_1(K_n)$ is still unknown for $n \geq 31$ and there is no candidate for a closed formula. On the positive side, the *rectilinear crossing constant* $\overline{\overline{\mathrm{cr}}}_1 := \lim_{n\to\infty} \overline{\overline{\mathrm{cr}}}_1(K_n)/\binom{n}{4}$ is known to exist (see for example [21]). Its bounds meanwhile have been narrowed to

$$0.37997 \leq \overline{\overline{\mathrm{cr}}}_1 \leq 0.38045 \tag{2}$$

(see [1] and the arXiv-version of [4]). The upper bound employs a construction from [2] that generates drawings of K_n with arbitrarily large n and small rectilinear crossing number given an initial drawing Γ with few crossings and a so-called *halving matching* of Γ (a matching between vertices with incident halving edges).

For fixed $k \geq 2$, the *geometric k-colored crossing constant* $\overline{\overline{\mathrm{cr}}}_k := \lim_{n\to\infty} \overline{\overline{\mathrm{cr}}}_k(K_n)/\binom{n}{4}$ exists as well (with an identical proof as for $\overline{\overline{\mathrm{cr}}}_1$ in [21]). For $k = 2$, the previously best known bounds on $\overline{\overline{\mathrm{cr}}}_2$ are

$$\frac{1}{33} = 0.0\bar{3} \leq \overline{\overline{\mathrm{cr}}}_2 \leq 0.11798016, \tag{3}$$

both of which were shown in [5]. The upper bound is obtained by a generalized notion of halving matchings and a construction based on the approach from [2], but is specifically tailored to two colors.

For $k \geq 3$, bounds on $\overline{\overline{\mathrm{cr}}}_k$ are derived by the following: For any $k \geq 1$ and any graph G, $\overline{\overline{\mathrm{cr}}}_k(G)$ is bounded from below by the *k-colored crossing number* cr_k which is defined as in (1) but with Γ ranging over all possible drawings of K_n instead of only straight-line drawings. On the other hand, $\overline{\overline{\mathrm{cr}}}_k(G)$ is bounded from above by the *k-page book crossing number* $\mathrm{bkcr}_k(G)$, also defined as in (1) but with Γ restricted to drawings of G with the n vertices in convex position. From the existence of $\mathrm{cr}_k := \mathrm{cr}_k(K_n)/\binom{n}{4}$ and $\mathrm{bkcr}_k := \mathrm{bkcr}_k(K_n)/\binom{n}{4}$, combined with the best known asymptotic bounds on $\mathrm{cr}_k(K_n)$ and $\mathrm{bkcr}_k(K_n)$, we obtain

$$\frac{3}{29k^2} \leq \mathrm{cr}_k \leq \overline{\overline{\mathrm{cr}}}_k \leq \mathrm{bkcr}_k \leq \frac{2}{k^2} - \frac{1}{k^3}. \tag{4}$$

The lower bound in (4) stems from [26] and is via an application of the Crossing Lemma [3]. The upper bound follows independently from two different constructions [10, 25] as noticed in [11].

Our Contribution. In this work, we develop a technique to improve the upper bounds of $\overline{\overline{\mathrm{cr}}}_k$ for any fixed $k \geq 2$ by generalizing the approach from [5] to $k \geq 2$ and by improving it for $k = 2$. To this end, we also find provably *optimal matchings* for any $k \geq 2$ (for $k = 1$, the matchings from [2] were known to be optimal; the ones used in [5] were not). We exemplify our approach

[1] We use the notation of Schaefer [24] for $\overline{\overline{\mathrm{cr}}}_k(G)$, but a different name than in previous literature for the following reasons: We do not write "geometric k-planar crossing number" as in [20] to avoid confusion with the concept of k-planar graphs, and we do not write "rectilinear k-colored crossing number" as in [5, 13] to avoid confusion with the related but different "rectilinear k-planar crossing number" and to highlight the relation to geometric thickness.

for $k = 2, \ldots, 10$ (using heuristic search methods for initial drawings and colorings) and obtain substantially improved upper bounds for $\overline{\overline{\mathrm{cr}}}_2, \ldots \overline{\overline{\mathrm{cr}}}_{10}$.

Further related work. The geometric k-colored crossing number first appeared in [20]. In the literature, the geometric k-colored crossing number has also been treated in the context of the *k-colored crossing ratio of a drawing* Γ, that is, the ratio between the k-colored crossing number $\mathrm{cr}_k(\Gamma)$ of Γ and $\mathrm{cr}(\Gamma)$. The authors of [5] proved that there is some constant $c > 0$ such that for all large enough n and all straight-line drawings Γ of K_n, $\mathrm{cr}_k(\Gamma)/\mathrm{cr}(\Gamma) \leq 1/2 - c$. In [13] this is generalized to any k and to dense graphs. In [9] it is shown that for n points chosen uniformly at random from a unit square, the induced straight-line drawing Γ of K_n has $\mathrm{cr}_k(\Gamma)/\mathrm{cr}(\Gamma) \leq 1/2 - 7/50$ in expectation.

The crossing properties of Γ are captured by the *crossing graph* of Γ, whose vertices are the edges of Γ and in which two vertices are adjacent if they cross. As the total number of uncolored crossings in Γ is fixed, a k-edge-coloring of Γ realizing $\mathrm{cr}_k(\Gamma)$ is equivalent to a k-vertex-coloring of the crossing graph of Γ that maximizes adjacencies between differently colored vertices, i.e., a *maximum k-cut.* The problem MAX-k-CUT is $\mathcal{NP}$-hard and also hard to approximate in general [14, 17]. Moreover, it remains $\mathcal{NP}$-hard for segment intersection graphs [19] (and hence for crossing graphs of drawings) [19]. On the other hand, there is a PTAS for MAX-k-CUT for dense graphs [7] (and hence for crossing graphs of drawings of K_n).

2 The doubling construction

We consider straight-line drawings of K_n given by some set of points $P \subseteq \mathbb{R}^2$ in general position with a k-edge-coloring χ. We denote the number of monochromatic crossings in the resulting drawing by $\mathrm{cr}_k(P; \chi)$.

We work with matchings, which match each vertex with an incident edge such that no edge is matched twice. Formally, a *matching* is a map $m : P \to P$ with $m(p) \neq p$, $m(m(p)) \neq p$ for all $p \in P$. We call $pm(p)$ the *matching edge of* p and think of it as being oriented away from p. We denote the color of the matching edge of p as $\overline{c}(p) \coloneqq \chi(pm(p))$. For each color c, the number of edges incident to p with color c that lie to the left (respectively right) of the line spanned by the matching edge of p is denoted as $S_c^\ell(p)$ (respectively $S_c^r(p)$).

A *χ-halving matching* as defined in [5] is a matching with the additional property that for each point p, a color c with the maximum number of incident edges at p fulfills $\left|S_c^\ell(p) - S_c^r(p)\right| \leq 1$.

Given a point set P_0, a k-edge-coloring χ_0 and matching m_0, we construct a point set P_1 with $|P_1| = 2|P_0|$ together with a k-edge-coloring χ_1 and a matching m_1 in a way that can be iteratively repeated to obtain χ_t, P_t, and m_t for any $t \in \mathbb{N}$ with few monochromatic crossings in the following way.

Point set: We replace each point $p \in P_0$ by the two points with distance ε to p on the line spanned by the edge $pm_0(p)$ for a sufficiently small positive ε (i.e., such that no smaller ε changes the order type of the point set). The resulting points are the *children of p.* We denote them by p_1 and p_2 such that p_1 is further from $m_0(p)$ than p_2. In turn, p is the *parent* of p_1 and p_2. We further denote the left and right child of m_0 from the perspective of p as $m_0(p)_\ell$ and $m_0(p)_r$.

Coloring and matching: We choose $\chi_1(p_i q_j) = \chi_0(pq)$ if $p \neq q$. Independently for each vertex p, we decide on $c'(p) \coloneqq \chi_1(p_1 p_2)$ and $m_1(p_1), m_1(p_2)$ and call these choices the *details (at p).* We restrict the choice of $m_1(p_1)$ and $m_1(p_2)$ to the children of $m_0(p)$ and p_1, p_2 and disallow $m_1(p_2) = p_1$, in order to preserve the rough structure of m_0; see Figure 1 for an example. We do not enforce any canonical method to choose the details but will later describe how details that optimize the asymptotic number of crossings can be found. In contrast, the authors of [5] choose the details at a vertex p according to a case distinction on the color of $m_0(p)$ and the values of $S_c^\ell(p)$ and $S_c^r(p)$, which is not always optimal.

Figure 1: One step in the doubling procedure at a vertex p. Matching edges are drawn bold and with an arrowhead. Dashed lines are the extensions of matching edges along which the vertices are split.

Iterated application: Given P_i, χ_i, and m_i from iteration i, we construct P_{i+1}, χ_{i+1}, and m_{i+1} analogously to the first step. In particular, each vertex in P_i is a parent of two vertices in P_{i+1}. Calling a vertex in some P_i a *descendant* of $p \in P_0$ if they are transitively related by the parent relation, the descendants of p form an infinite full binary tree T_p rooted at p. We set the left child of p to p_1 and the right child to p_2. We denote by p_j^i the vertex on level i (thus in P_i) of T_p at position $j \in \{1, \dots, 2^i\}$ from left to right. In this notation, $p = p_1^0$, $p_1 = p_1^1$, $p_2 = p_2^1$, and the two children of p_j^i are p_{2j-1}^{i+1} and p_{2j}^{i+1}.

For each descendant $p_j^i \in P_i$ of $p \in P_0$, we choose the details at p_j^i identically to those at p: $\chi_{i+1}(p_{2j-1}^{i+1}p_{2j}^{i+1}) = \chi_1(p_1p_2)$ and if $m_1(p_1) = m_0(p)_\ell$ then $m_{i+1}(p_{2j-1}^{i+1}) = m_i(p_j^i)_\ell$, if $m_1(p_1) = m_0(p)_r$ then $m_{i+1}(p_{2j-1}^{i+1}) = m_i(p_j^i)_r$, and if $m_1(p_1) = p_2$ then $m_{i+1}(p_{2j-1}^{i+1}) = p_{2j}^{i+1}$ (analogously for p_{2j}^{i+1}).

3 Analysis

The following theorem counts the number of crossings after t step of the doubling construction. The proof is based on similar counting arguments as in [5, Claim 1] and [2, Lemma 3]. It can be found in the extended arXiv-version of this paper at [16] but is omitted here due to space restrictions.

Theorem 1. *Given a point set P_0, a k-edge-coloring χ_0, a matching m_0, and details at all vertices of P_0, the number of monochromatic crossings after t iterations of the doubling construction is*

$$\begin{aligned}
\mathrm{cr}_k(P_t;\chi_t) \;&=\; 16^t\,\mathrm{cr}_k(P_0;\chi_0) + \sum_{i=0}^{t-1} 16^{t-i-1}\left[\binom{2^i|P_0|}{2} - 2^i|P_0|\right] \\
&\quad + 4\sum_{p\in P_0}\sum_{c=1}^{k}\sum_{i=0}^{t-1} 16^{t-i-1}\sum_{j=1}^{2^i}\left[\binom{S_c^\ell(p_j^i)}{2} + \binom{S_c^r(p_j^i)}{2}\right] \\
&\quad + 2\sum_{p\in P_0}\sum_{i=0}^{t-1} 16^{t-i-1}\sum_{j=1}^{2^i}\left[S_{\bar{c}(p)}^\ell(p_j^i) + S_{\bar{c}(p)}^r(p_j^i)\right].
\end{aligned} \tag{5}$$

To determine the asymptotics of (5) for $t \to \infty$, we consider the values of $S_c^d(p_j^i)$. We reason that there exist offsets $o_1, o_2 \in \{0, 1, 2\}$ depending only on $p \in P_0$, $c \in \{1, \dots, k\}$, and $d \in \{\ell, r\}$, such that

$$S_c^d(p_{2j-1}^{i+1}) = 2 \cdot S_c^d(p_j^i) + o_1 \qquad \text{and} \qquad S_c^d(p_{2j}^{i+1}) = 2 \cdot S_c^d(p_j^i) + o_2$$

for all i, j. As the doubling procedure behaves identically for all iterations, it is sufficient to make the following arguments for $i = 0$. The factor 2 appears because each edge pq, $q \neq m_0(p)$ gives rise to two edges incident to q_1 and two edges incident to q_2 with the same color as pq and on the same side of the respective halving edges. The offsets o_1 and o_2 stem from the six edges p_1p_2, p_1q_1, p_1q_2 at p_1 and p_2p_1, p_2q_1, p_2q_2 at p_2. Two of these are the matching edges of p_1 and p_2 and do not count towards $S_c^d(p_1)$ or $S_c^d(p_2)$ for any c and d, so $o_1, o_2 \leq 2$. Further, $o_2 \leq 1$, as no choice for the matching edge of p_2 leaves the remaining two edges on the same side. Finally, if $o_1 = 2$, then $o_2 \neq 0$ as is apparent from a short case distinction. The various values of o_1 and o_2 give rise to the following five closed formulas for $S_c^d(p_j^i)$. The correctness of these formulas, which we denote as $f_{(o_1,o_2)}(S_c^d(p), i, j)$, can be proven by induction on i, see the appendix of [16].

o_2 \ o_1	0	1	2
0	$2^i S_c^d(p)$	$2^i S_c^d(p) + 2^i - j$	-
1	$2^i S_c^d(p) + j - 1$	$2^i S_c^d(p) + 2^i - 1$	$2^i S_c^d(p) + 2^{i+1} - j - 1$

Plugging these into (5), we obtain

$$\begin{aligned} \mathrm{cr}_k(P_t;\chi_t) \;=\; & 2^{4t}\,\mathrm{cr}_k(P_0;\chi_0) + \left[\sum_{i=0}^{t-1} 16^{t-i-i}\left(\binom{2^i|P_0|}{2} - 2^i|P_0|\right)\right] \\ & + 4\sum_{p\in P_0}\sum_{c=1}^{k}\sum_{d\in\{\ell,r\}}\left[\sum_{i=0}^{t-1} 16^{t-i-i}\sum_{j=1}^{2^i}\binom{f_{(o_1,o_2)}(S_c^d(p),i,j)}{2}\right] \\ & + 2\sum_{p\in P_0}\sum_{d\in\{\ell,r\}}\left[\sum_{i=0}^{t-1} 16^{t-i-i}\sum_{j=1}^{2^i} f_{(o_1,o_2)}(S_{\overline{c}(p)}^d(p),i,j)\right], \end{aligned} \tag{6}$$

where (o_1, o_2) depend on the current p, c and d. Let us denote the bracketed terms in (6) by $A(|P_0|)$, $B_{(o_1,o_2)}(S_c^d(p))$, and $C_{(o_1,o_2)}(S_c^d(p))$, respectively. Straightforward but long computations yield closed formulas for these eleven functions (again found in the appendix of [16]), each of which is of the form $2^{4t}p_4(x) + 2^{3t}p_3(x) + 2^{2t}p_2(x) + 2^t p_1(x)$ for polynomials p_1, p_2, p_3, p_4. From this we obtain the following theorem.

Theorem 2. *Given a non-empty point set P_0, $|P_0| \geq 3$ and a k-edge-coloring χ_0, a matching m_0, and details at all vertices of P_0, there are $\alpha, \beta, \gamma, \delta \in \mathbb{R}, \alpha > 0, \beta < 0, \alpha + \beta + \gamma + \delta = \mathrm{cr}_k(P_0; \chi_0)$ such that for any $t \in \mathbb{N}_0$*

$$\mathrm{cr}_k(P_t; \chi_t) = \alpha 2^{4t} + \beta 2^{3t} + \gamma 2^{2t} + \delta 2^t.$$

Proof. The existence of such $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ is a direct consequence of Equation (6) being a finite sum over the functions $A(|P_0|)$, $B_{(o_1,o_2)}(S_c^d(p))$, and $C_{(o_1,o_2)}(S_c^d(p))$, each of which has the desired form. Then, $\alpha > 0$ and $\beta < 0$ follow from the signs of the relevant coefficients in the closed formulas when $|P_0| \geq 3$. Finally, setting $t = 0$ implies $\alpha + \beta + \gamma + \delta = \mathrm{cr}_k(P_0; \chi_0)$. □

Asymptotics and Optimal Matchings. With Theorem 2 we can now bound the geometric k-colored crossing constant by

$$\overline{\overline{\mathrm{cr}}}_k \leq \lim_{t\to\infty} \frac{\mathrm{cr}_k(P_t;\chi_t)}{\binom{|P_t|}{4}} = \lim_{t\to\infty} \frac{\alpha 2^{4t} + \mathcal{O}(2^{3t})}{\frac{|P_0|^4}{24} 2^{4t} + \mathcal{O}(2^{3t})} = \frac{24\alpha}{|P_0|^4}.$$

To compute α, one only needs to determine $S_c^d(p)$ and (o_1, o_2) for each (p, c, d) and sum the contributions to α involving the A, B and C-terms from (6). We call the sum of these terms involving $S_c^d(p)$ for fixed p the *local α (at p)*.

Given P_0, χ_0 and m_0, the details that minimize the total α result from choosing, at each $p \in P_0$, the details that minimize the local α. Due to this independence, even if only P_0 and χ_0 are given, the optimal matching can be found efficiently: Define weights w on P_0^2 by setting $w(p, q)$ to the minimum local α at p over all details at p if $m_0(p) = q$ was fixed. Further, let $H = (P_0 \cup \binom{P_0}{2}, \{(p, pq) \mid p, q \in P_0, p \neq q\})$ be the bipartite graph with edge weights $w_H(p, pq) = w(p, q)$. An optimum matching which minimizes α in the doubling procedure corresponds to a P_0-saturating matching (in the usual sense) in H with minimum weight, which can be determined in polynomial time [18].

4 Computational results for $k \in \{2, \ldots 10\}$

We focussed on obtaining upper bounds for the geometric k-colored crossing constant $\overline{\overline{\mathrm{cr}}}_k$ for $k = 2\ldots, 10$. For $k = 2$, the authors of [5] provide a set of 135 points $\mathcal{P}_2'$ with a 2-edge-coloring χ_2' such that $\mathrm{cr}_2(\mathcal{P}_2'; \chi_2') = 1470756$. Using their doubling procedure via halving matchings on this instance, they obtain $\overline{\overline{\mathrm{cr}}}_2 < 0.11798016$. For the same instance, using our doubling procedure we obtain a better bound of $\overline{\overline{\mathrm{cr}}}_2 < 0.11750015$ given by the optimum (non-halving) matching found by the bipartite-matching approach described above.

For $k \geq 3$ the best known upper bounds on $\overline{\overline{\mathrm{cr}}}_k$ in (4) are from the book-crossing number.

To improve these bounds, we searched for point sets $\mathcal{P}_k$ and k-edge-colorings χ_k for $k \geq 2$ with small $\mathrm{cr}_k(\mathcal{P}_k; \chi_k)$. To this end, we employed various MAX-k-CUT heuristics implemented in [23]. Starting with $\mathcal{P}_2'$, we ran the heuristics on the crossing graph of the induced drawing to find a k-edge-coloring with few monochromatic crossings. Keeping the coloring fixed, we further reduced the number of monochromatic crossing by perturbing the points. Iterating these two steps, we obtained point sets $\mathcal{P}_k$ and k-edge-colorings χ_k for $k = 2, \ldots, 10$. Finding optimum matchings and details at every vertex via the bipartite matching approach, the upper bounds on $\overline{\overline{\mathrm{cr}}}_k$ detailed in Table 1 are obtained. These improve the bounds from the k-page book crossing number by a factor of about 3. The point sets $\mathcal{P}_k$ and their edge-colorings and matchings together with a Python script that determines the derived upper bound on $\overline{\overline{\mathrm{cr}}}_k$ can be found in [15]. Let us note that while the optimum matchings for our instances are not halving, $S_c^\ell(p)$ and $S_c^r(p)$ tend to have similar values for most points p and colors c.

k	$\mathrm{cr}_k(\mathcal{P}_k; \chi_k)$	UB on $\overline{\overline{\mathrm{cr}}}_k$ from $\mathcal{P}_k$	UB on $\overline{\overline{\mathrm{cr}}}_k$ from [5]	UB on $\overline{\overline{\mathrm{cr}}}_k$ from bkcr_k	Improvement factor
2	1468394	0.117 314 12	0.117 980 16	0.375 000 00	1.006
3	732746	0.060 624 66	-	0.185 185 19	3.032
4	413342	0.035 721 51	-	0.109 375 00	3.062
5	264459	0.023 893 26	-	0.072 000 00	3.013
6	183248	0.017 260 49	-	0.050 925 93	2.950
7	133405	0.013 140 79	-	0.037 900 87	2.884
8	99638	0.010 283 34	-	0.029 296 88	2.849
9	78269	0.008 453 39	-	0.023 319 62	2.759
10	60922	0.006 926 71	-	0.019 000 00	2.743

Table 1: Upper bounds on the geometric k-colored crossing constant for $k = 2, \ldots, 10$.

5 Conclusion

We introduced a procedure that creates k-edge-colored straight-line drawings of K_n for large n with few monochromatic crossings, given an initial drawing for small n with this property. By finding k-edge-colored drawings with few monochromatic crossings, we kickstarted this procedure to improve the upper bounds on the geometric k-colored crossing constant for $k = 2, \ldots, 10$. While our method is applicable for larger k, gaining on the upper bound from the book crossing number in [10, 25] for all k at once is still open. We believe this is possible with a construction that does not arrange the points in convex position.

The lower bound on $\overline{\overline{\text{cr}}}_k$ in (4) is unlikely to be improved using the Crossing Lemma (unless a better one is found). A more promising avenue could be the study of ℓ-edges and $\leq \ell$-edges in a similar fashion as in [1] and previously [6] for the (non-colored) rectilinear crossing constant.

Acknowledgements. We thank the authors of [23] for providing us with their source code by making it publicly available at [22] and Oswin Aichholzer for providing us with source code used for computational results in [5].

References

[1] Bernardo M. Ábrego, Mario Cetina, Silvia Fernández-Merchant, Jesús Leaños, and Gelasio Salazar. On $\leq k$-Edges, Crossings, and Halving Lines of Geometric Drawings of K_n. *Computing Research Repository - CORR*, 48, 2011. `DOI:10.1007/s00454-012-9403-y`.

[2] Bernardo M. Ábrego and Silvia Fernández-Merchant. Geometric drawings of K_n with few crossings. *Journal of Combinatorial Theory, Series A*, 114(2):373–379, 2007. `DOI:10.1016/j.jcta.2006.05.003`.

[3] Eyal Ackerman. On topological graphs with at most four crossings per edge. *Computational Geometry*, 85:101574, 2019. `DOI:10.1016/j.comgeo.2019.101574`.

[4] Oswin Aichholzer, Frank Duque, Ruy Fabila-Monroy, Oscar García-Quintero, and Carlos Hidalgo-Toscano. An Ongoing Project to Improve the Rectilinear and the Pseudolinear Crossing Constants. *Journal of Graph Algorithms and Applications*, 24(3):421–432, 2020. arXiv Version: `https://arxiv.org/abs/1907.07796`. `DOI:10.7155/jgaa.00540`.

[5] Oswin Aichholzer, Ruy Fabila-Monroy, Adrian Fuchs, Carlos Hidalgo-Toscano, Irene Parada, Birgit Vogtenhuber, and Francisco Zaragoza. On the 2-Colored Crossing Number. In *Graph Drawing and Network Visualization*, volume 11904, pages 87–100. Springer International Publishing, 2019. `DOI:10.1007/978-3-030-35802-0_7`.

[6] Oswin Aichholzer, Jesus Garcia, David Orden, and Pedro Ramos. New Lower Bounds for the Number of $(\leq k)$-Edges and the Rectilinear Crossing Number of K_n. *Discrete & Computational Geometry*, 38(1):1–14, 2007. `DOI:10.1007/s00454-007-1325-8`.

[7] Sanjeev Arora, David Karger, and Marek Karpinski. Polynomial Time Approximation Schemes for Dense Instances of $\mathcal{NP}$-Hard Problems. *Journal of Computer and System Sciences*, 58(1):193–210, 1999. `DOI:10.1006/jcss.1998.1605`.

[8] Daniel Bienstock. Some provably hard crossing number problems. *Discrete & Computational Geometry*, 6(3):443–459, 1991. DOI:10.1007/BF02574701.

[9] Sergio Cabello, Éva Czabarka, Ruy Fabila-Monroy, Yuya Higashikawa, Raimund Seidel, László Székely, Josef Tkadlec, and Alexandra Wesolek. A note on the 2-colored rectilinear crossing number of random point sets in the unit square. *Acta Mathematica Hungarica*, 173(1):214–226, 2024. DOI:10.1007/s10474-024-01436-9.

[10] Ernesto Damiani, Ottavio D'Antona, and Payam Salemi. An upper bound to the crossing number of the complete graph drawn on the pages of a book. In *Journal of Combinatorics, Information & System Sciences*, volume 19, pages 75–84. MD Publications Pvt, 1994. ISSN: 0250-9628.

[11] Etienne de Klerk, Dmitrii V. Pasechnik, and Gelasio Salazar. Improved Lower Bounds on Book Crossing Numbers of Complete Graphs. *SIAM Journal on Discrete Mathematics*, 27(2):619–633, 2013. DOI:10.1137/120886777.

[12] Nathaniel Dean. Mathematical programming formulation of rectilinear crossing minimization. Technical Report 12, DIMACS, 2002.

[13] Ruy Fabila-Monroy. A note on the k-colored crossing ratio of dense geometric graphs. *Computational Geometry*, 124–125:102123, 2025. DOI:10.1016/j.comgeo.2024.102123.

[14] Michael R. Garey, David S. Johnson, and Larry J. Stockmeyer. Some simplified *NP*-complete graph problems. *Theoretical Computer Science*, 1(3):237–267, 1976. DOI:10.1016/0304-3975(76)90059-1.

[15] Benedikt Hahn, Bettina Klinz, and Birgit Vogtenhuber. Instances with low geometric k-colored crossing number, 2025. DOI:10.3217/tw2dv-87653.

[16] Benedikt Hahn, Bettina Klinz, and Birgit Vogtenhuber. On the geometric k-colored crossing number of K_n, 2025. DOI:10.48550/arXiv.2505.18014.

[17] Viggo Kann, Sanjeev Khanna, Jens Lagergren, and Alessandro Panconesi. On the Hardness of Approximating Max k -Cut and Its Dual. In *Proceedings of the Fourth Israel Symposium on Theory of Computing and Systems*, pages 61–67. IEEE Computer Society, 1996.

[18] Harold W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955. DOI:10.1002/nav.3800020109.

[19] Sumio Masuda, Kazuo Nakajima, Toshinobu Kashiwabara, and Toshio Fujisawa. Crossing minimization in linear embeddings of graphs. *IEEE Transactions on Computers*, 39(1):124–127, 1990. DOI:10.1109/12.46286.

[20] János Pach, László A. Székely, Csaba D. Tóth, and Géza Tóth. Note on k-planar crossing numbers. *Computational Geometry*, 68:2–6, 2018. DOI:10.1016/j.comgeo.2017.06.015.

[21] R. Bruce Richter and Carsten Thomassen. Relations Between Crossing Numbers of Complete and Complete Bipartite Graphs. *The American Mathematical Monthly*, 104(2):131–137, 1997. DOI:10.1080/00029890.1997.11990611.

[22] Vilmar Jefté Rodrigues de Sousa. Max-k-Cut branching algorithm and heuristics, 2020. URL: `https://github.com/vilmarjet/maxKcut`, visited on March 7, 2025.

[23] Vilmar Jefté Rodrigues de Sousa, Miguel F. Anjos, and Sébastien Le Digabel. Computational study of a branching algorithm for the maximum k-cut problem. *Discrete Optimization*, 44:100656, 2022. DOI:10.1016/j.disopt.2021.100656.

[24] Marcus Schaefer. The Graph Crossing Number and its Variants: A Survey. *The Electronic Journal of Combinatorics*, page DS21, 2024. DOI:10.37236/2713.

[25] Farhad Shahrokhi, László A. Székely, Ondrej Sýkora, and Imrich Vrt'o. The book crossing number of a graph. *Journal of Graph Theory*, 21(4):413–424, 1996. DOI:10.1002/(SICI)1097-0118(199604)21:4<413::AID-JGT7>3.0.CO;2-S.

[26] Alireza Shavali and Hamid Zarrabi-Zadeh. On the Biplanar and k-Planar Crossing Numbers. In *Proceedings of the 34th Canadian Conference on Computational Geometry*, pages 293–297, 2022.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Proper Rainbow Saturation Numbers for Cycles

(Extended abstract)

Anastasia Halfpap* Bernard Lidický† Tomáš Masařík‡

Abstract

We say that an edge-coloring of a graph G is *proper* if every pair of incident edges receive distinct colors, and is *rainbow* if no two edges of G receive the same color. Furthermore, given a fixed graph F, we say that G is *(proper) rainbow F-saturated* if G admits a proper edge-coloring which does not contain any rainbow subgraph isomorphic to F, but the addition of any edge to G makes such an edge-coloring impossible. The maximum number of edges in a (proper) rainbow F-saturated graph is the rainbow Turán number, whose study was initiated in 2007 by Keevash, Mubayi, Sudakov, and Verstraëte. Recently, Bushaw, Johnston, and Rombach introduced study of a corresponding saturation problem, asking for the *minimum* number of edges in a (proper) rainbow F-saturated graph. We term this minimum the *proper rainbow saturation number* of F, denoted $\mathrm{sat}^*(n, F)$. We asymptotically determine $\mathrm{sat}^*(n, C_4)$, answering a question of Bushaw, Johnston, and Rombach. We also exhibit constructions which establish upper bounds for $\mathrm{sat}^*(n, C_5)$ and $\mathrm{sat}^*(n, C_6)$.

1 Introduction

A central problem in extremal graph theory is to understand the set of n-vertex graphs G which do not contain some forbidden subgraph F. Formally, given graphs G, F, we say that G *contains* a copy of F (or F*-copy*) if G contains a subgraph (not necessarily induced) isomorphic to F; if G does not contain a copy of F, we say that G is F*-free*. Note that if G is F-free, then all subgraphs of G are also F-free; thus, it is natural to restrict our attention to edge-maximal F-free n-vertex graphs, as these contain all F-free n-vertex graphs. We use $V(G), E(G)$ to denote the vertex and edge sets of G, and for $x, y \in V(G)$ such that $xy \notin E(G)$, we denote by

*Department of Mathematics, Iowa State University, Ames, IA, USA. E-mail: ahalfpap@iastate.edu
†Department of Mathematics, Iowa State University, Ames, IA, USA. E-mail: lidicky@iastate.edu Supported in part by NSF grant FRG DMS-2152490.
‡University of Warsaw, Institute of Informatics, Warszawa, Poland. E-mail: masarik@mimuw.edu.pl Supported by Polish National Science Centre SONATA-17 grant number 2021/43/D/ST6/03312.

The full preprint version of this article is available on the arXiv [7].

$G + xy$ the graph on vertex set $V(G)$ with edge set $E(G) \cup \{xy\}$. A graph G is *F-saturated* if G is F-free but, for any $e \notin E(G)$, the graph $G + e$ contains a copy of F.

Further restricting our focus, we may ask which F-saturated n-vertex graphs are somehow extremal. That is, in the set of n-vertex F-saturated graphs, which elements optimize some graph parameter? This question yields two natural avenues of research. The *Turán number* $\mathrm{ex}(n, F)$ is the maximum number of edges among all n-vertex, F-saturated graphs. Famously first considered by Mantel [6] in the case $F = K_3$, and for general cliques by Turán [11], the study of $\mathrm{ex}(n, F)$ remains a vibrant area of study in its own right, as well as giving rise to a variety of natural generalizations and variations. On the other hand, the *saturation number* $\mathrm{sat}(n, F)$ is the *minimum* number of edges among all n-vertex, F-saturated graphs. The study of saturation numbers is also well-established (see, for instance, [4]) and, like Turán problems, it is natural to generalize saturation problems to a variety of contexts.

This paper concerns saturation problems in an edge-colored setting. An *edge-coloring* of a graph G is a function $c : E(G) \to \mathbb{N}$. We say that $c(e)$ is the *color* of e, and that c is a *proper* edge-coloring if for any two incident edges e, f, we have $c(e) \neq c(f)$. An edge-colored graph is *rainbow* if all of its edges receive different colors. Given fixed graphs G, F, and a proper edge-coloring c of G, we say that G is *rainbow-F-free under c* if G does not contain any copy of F which is rainbow with respect to c. Moreover, we say that G is *(properly) rainbow F-saturated* if G satisfies the following conditions.

1. There exists a proper edge-coloring c of G such that G is rainbow-F-free under c;
2. For any edge $e \notin E(G)$, any proper edge-coloring of $G + e$ contains a rainbow F-copy.

Motivated by a problem in additive number theory, Keevash, Mubayi, Sudakov, and Verstraëte [8] introduced the *rainbow Turán number* $\mathrm{ex}^*(n, F)$ in 2007, which is the maximum number of edges in an n-vertex, rainbow F-saturated graph. Following the analogy between $\mathrm{ex}(n, F)$ and $\mathrm{ex}^*(n, F)$, it is natural to also consider the rainbow counterpart to $\mathrm{sat}(n, F)$. Bushaw, Johnston, and Rombach [3] recently initiated a study of this rainbow version of the saturation number, denoted $\mathrm{sat}^*(n, F)$, the minimum number of edges in an n-vertex rainbow F-saturated graph. We call $\mathrm{sat}^*(n, F)$ the *proper rainbow saturation number* of F, since all edge-colorings in this setting are proper. While slightly lengthy, this terminology distinguishes $\mathrm{sat}^*(n, F)$ from an already-studied function which has been termed the rainbow saturation number in the literature (see, e.g., [1], [2], [5]), and which does not assume a setting of proper edge-colorings.

Given that consideration of $\mathrm{sat}^*(n, F)$ is extremely new, few results have been established in the area, and the general behavior of $\mathrm{sat}^*(n, F)$ remains unclear. For instance, while it is simple to observe that $\mathrm{ex}(n, F) \leq \mathrm{ex}^*(n, F)$ for all F, it is not obvious whether we have $\mathrm{sat}(n, F) \leq \mathrm{sat}^*(n, F)$ for all F. The following theorem illustrates that we sometimes have $\mathrm{sat}(n, F) < \mathrm{sat}^*(n, F)$, and in fact $\mathrm{sat}^*(n, F)$ may differ from $\mathrm{sat}(n, F)$ by a multiplicative factor. Here and throughout, we denote by P_ℓ the path on ℓ *vertices* (that is, on $\ell - 1$ edges).

Theorem 1.1 ([3, Theorem 3.5]). *For each $n \geq 16$, we have*

$$\lfloor \frac{4n}{5} \rfloor \leq \mathrm{sat}^*(P_4, n) \leq \frac{4}{5}n - \frac{17}{10}c,$$

where $0 \leq c \leq 4$ and $c \equiv -n \mod 5$.

For contrast, $\mathrm{sat}(n, P_4)$ is approximately $\frac{n}{2}$.

Theorem 1.2 ([9, Proposition 5])**.**

$$\operatorname{sat}(n, P_4) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{n+3}{2} & \text{if } n \text{ is odd} \end{cases}$$

Apart from $\operatorname{sat}^*(P_4, n)$, we do not have tight bounds on any proper rainbow saturation number, except for trivial cases where every proper coloring of F is rainbow and thus $\operatorname{sat}(n, F) = \operatorname{sat}^*(n, F)$. The goal of this paper is to contribute to the understanding of the proper rainbow saturation numbers of cycles, in particular by determining $\operatorname{sat}^*(n, C_4)$ asymptotically. The previous best known bounds on $\operatorname{sat}^*(n, C_4)$ are due to Bushaw, Johnston, and Rombach [3].

Theorem 1.3 ([3, Theorem 3.6])**.** *For $n \geq 4$, we have $n \leq \operatorname{sat}^*(n, C_4) \leq 2n - 2$.*

For comparison, the ordinary saturation number of C_4 is known exactly.

Theorem 1.4 ([10])**.** $\operatorname{sat}(n, C_4) = \lfloor \frac{3n-5}{2} \rfloor$.

We contribute the following bounds on $\operatorname{sat}^*(n, C_4)$, which asymptotically determine its value and show that $\operatorname{sat}^*(n, C_4)$ is separated from $\operatorname{sat}(n, C_4)$ by a constant multiplicative factor.

Theorem 1.5. *For $n \geq 7$,*

$$\operatorname{sat}^*(n, C_4) \leq \frac{11}{6}n + O(1).$$

Theorem 1.6. *Let $\frac{11}{45} > \varepsilon > 0$ be given. There exists $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$,*

$$\operatorname{sat}^*(n, C_4) > \left(\frac{11}{6} - \varepsilon\right)n.$$

For cycles of length greater than 4, little is known. While it is currently known that $\operatorname{sat}^*(n, F)$ is always linear in n [2, Section 6.1], Bushaw, Johnston, and Rombach [3] described a class of graphs with linear proper rainbow saturation number. This provides an upper bound on $\operatorname{sat}^*(n, F)$ for F being an odd cycle, but the constant c given may be very large.

Theorem 1.7 ([3, Theorem 4.2])**.** *Suppose that F contains no induced even cycle. Then there is a constant c depending only on F such that $\operatorname{sat}^*(n, F) \leq cn$.*

There are no published bounds on $\operatorname{sat}^*(n, C_{2\ell})$ for $\ell > 2$. We contribute constructions that improve this state of affairs for C_5 and C_6. For C_5, we obtain a single bound regardless of the parity of n.

Theorem 1.8. *For $n \geq 9, \operatorname{sat}^*(n, C_5) \leq \lfloor \frac{5n}{2} \rfloor - 4$.*

For C_6, our bound varies slightly with the congruence class of n modulo 3. To avoid separate cases, we allow a constant error term which absorbs this discrepancy.

Theorem 1.9. $\operatorname{sat}^*(n, C_6) \leq \frac{7}{3}n + O(1)$.

Using the elementary observation that, for $n > 2$, a rainbow C_ℓ-saturated graph contains no acyclic component (since the addition of an edge either within an acyclic component yields a component containing at most one C_ℓ copy, which can be properly colored to avoid a rainbow C_ℓ-copy, while an edge between distinct components creates no new cycles at all), we have the immediate lower bound $\mathrm{sat}^*(n, C_\ell) \geq n$ for all ℓ and all $n > 2$. Thus, the bounds given in Theorems 1.8 and 1.9 seem reasonable, although we do not attempt to find matching lower bounds.

We present only a construction and a part of the proof of Theorem 1.5 in the extended abstract. The full proofs are available in the preprint version on the arXiv [7].

Notation. We denote *degree* of a vertex v in a graph G by $d_G(v)$ and the *minimum degree* of a vertex v in a graph G by $\delta(G)$. If G is clear from context we omit the subscript and simply write $d(v)$ for the degree of v in G. Given vertices u, v we denote *distance* by $d(u, v)$. That is, $d(u, v)$ is the minimum number of edges on a path from u to v. We use $N[v]$ to denote the *closed neighbourhood* of a vertex v and we let $N(v) := N[v] \setminus \{v\}$. For $S \subseteq V(G)$, we use $N(S)$ (resp. $N[S]$) as a shortcut for $\bigcup_{v \in S} N(v)$ (resp. $\bigcup_{v \in S} N[v]$). Given a graph G and $S \subset V(G)$, we use $G[S]$ to denote the subgraph of G induced on S, that is, the graph with vertex set S and edge set $E(G[S]) = \{uv \in E(G) : u, v \in S\}$. Let $S_{a,b,c}$ denote a claw graph where each edge is subdivided $a - 1$, $b - 1$, and $c - 1$ times, respectively.

2 Bounds for C_4

We begin by improving the upper bound on $\mathrm{sat}^*(n, C_4)$, with a construction showing that $\mathrm{sat}^*(n, C_4) \leq \frac{11n}{6} + O(1)$. Before stating the construction, we establish a variety of facts about properly rainbow C_4-saturated graphs, which will be useful throughout. We begin with a proposition collecting a few elementary observations.

Proposition 2.1. *Let G be a rainbow C_4-saturated graph. Then the following hold.*

1. *G contains at most one vertex of degree 1;*
2. *For any vertices $u, v \in V(G)$, $d(u, v) \leq 3$;*
3. *For any vertices $u, v \in V(G)$, $|N(u) \cap N(v)| < 4$.*

Next, we prove the following key lemma, which will be required to demonstrate that our construction is properly rainbow C_4-saturated. Its proof is available in a preprint on the arXiv [7].

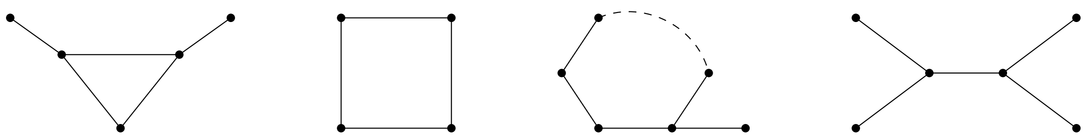


Figure 1: Subgraphs not appearing in $N(v)$

Lemma 2.2. *Let G be a graph and $v \in V(G)$. If there exists a proper edge-coloring of G which is rainbow C_4-free, then the subgraph of G induced on $N(v)$ does not contain the following subgraphs (not necessarily induced); see Figure 1:*

1. *A copy of K_3 with pendant edges from two vertices;*
2. *C_4;*
3. *A copy of C_k with a pendant edge, for any $k \geq 5$;*
4. *The double star $D_{2,2}$, or any subdivision thereof.*

With the above lemma established, we can quickly prove that the following construction is properly rainbow C_4-saturated.

Construction 1. Suppose $n \equiv i \mod 6$ with $n \geq 7$. For convenience, if n is divisible by 6, we shall set $i = 6$, not $i = 0$. Let G_n be the graph consisting of a universal vertex u whose neighborhood induces $\lfloor \frac{n-1}{6} \rfloor - 1$ copies of $S_{1,2,2}$ and one copy of $S_{1,\lceil \frac{6+i-2}{2} \rceil, \lfloor \frac{6+i-2}{2} \rfloor}$.

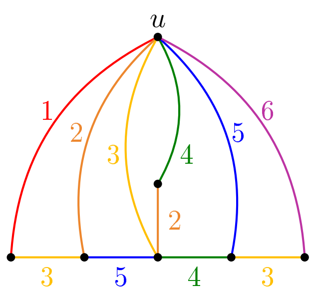


Figure 2: A copy of $S_{1,2,2}$ in $N(u)$, colored to avoid a rainbow C_4-copy

Theorem 1.5. *For $n \geq 7$,*

$$\mathrm{sat}^*(n, C_4) \leq \frac{11}{6}n + O(1).$$

Proof. Observe that Construction 1 has $\frac{11}{6}n + O(1)$ edges. Observe also that in Construction 1, every copy of C_4 intersects precisely one component of the graph induced on $N(u)$. Thus, to verify that a coloring of Construction 1 is rainbow-C_4-free, it suffices to verify that each component C of the graph induced on $N(u)$ can be colored so that $\{u\} \cup V(C)$ is rainbow-C_4-free. In Figure 2, given a copy of $S_{1,2,2}$ in $N(u)$ and colors for those edges incident to u, we exhibit a proper coloring of $S_{1,2,2}$ so that the graph induced on $\{u\} \cup V(S_{1,2,2})$ is rainbow-C_4-free. We can similarly, for any $i \in \{1, \ldots, 6\}$, color a copy of $S_{1,\lceil \frac{6+i-2}{2} \rceil, \lfloor \frac{6+i-2}{2} \rfloor}$ in $N(u)$ so as to avoid a rainbow C_4-copy. Thus, Construction 1 admits a rainbow-C_4-free proper edge coloring.

We now show that Construction 1 is rainbow C_4-saturated. Since u is a universal vertex, any edge added to the construction is contained within $N(u)$. Label the components of the graph induced on $N(u)$ as $C_1, C_2, \ldots, C_k$. If an edge is added between C_i and C_j, then, since C_i and C_j each contain a vertex of degree 3, this added edge will either create a subdivision of $D_{2,2}$ within $N(u)$ or create a vertex $v \in N(u)$ with $|N(u) \cap N(v)| = 4$. By Lemma 2.2 and Proposition 2.1, either outcome implies that the resulting graph does not admit a rainbow-C_4-free proper edge coloring. If an edge is added between non-adjacent vertices of a single component C_i, then we can quickly verify that this addition will create either a triangle with pendant edges from two vertices, a copy of C_4, or a copy of C_k with a pendant edge for some $k \geq 5$. In any case, by Lemma 2.2, the addition of an edge to $N(u)$ must yield a graph not admitting a rainbow-C_4-free proper edge coloring. □

References

[1] Michael Barrus, Michael Ferrara, Jennifer Vandenbussche, and Paul Wenger. Colored saturation parameters for rainbow subgraphs. *Journal of Graph Theory*, 86:375–386, March 2017. doi:10.1002/jgt.22132.

[2] Natalie Behague, Tom Johnston, Shoham Letzter, Natasha Morrison, and Shannon Ogden. The rainbow saturation number is linear. *SIAM Journal on Discrete Mathematics*, 38(2):1239–1249, 2024. doi:10.1137/23M1566881.

[3] Neal Bushaw, Daniel Johnston, and Puck Rombach. Rainbow saturation. *Graphs and Combinatorics*, 38(5), September 2022. doi:10.1007/s00373-022-02566-z.

[4] Bryan L. Currie, Jill R. Faudree, Ralph J. Faudree, and John R. Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, pages DS19:1–DS19:98, October 2021. doi:10.37236/41.

[5] António Girão, David Lewis, and Kamil Popielarz. Rainbow saturation of graphs. *Journal of Graph Theory*, 94:421–444, October 2017. doi:10.1002/jgt.22532.

[6] H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh, and W.A. Wythoff. Problem 28. *Wiskundige Opgaven met de Oplossingen*, 10:60–61, 1910.

[7] Anastasia Halfpap, Bernard Lidický, and Tomáš Masařík. Proper rainbow saturation numbers for cycles, 2024. arXiv:2403.15602.

[8] Peter Keevash, Dhruv Mubayi, Benny Sudakov, and Jacques Verstraëte. Rainbow Turán problems. *Combinatorics, Probability, and Computing*, 16(1):109–126, 2007. doi:10.1017/S0963548306007760.

[9] László Kászonyi and Zsolt Tuza. Saturated graphs with minimal number of edges. *Journal of Graph Theory*, 10:203–210, June 1986. doi:10.1002/jgt.3190100209.

[10] L. Ollmann. $K_{2,2}$-saturated graphs with a minimal number of edges. In *Proceedings of the 3rd Southeastern Conference on Combinatorics, Graph Theory, and Computing*, pages 367–392, 1972.

[11] Paul Turán. Eine Extremalaufgabe aus der Graphentheorie. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Reconstructing graphs and their connectivity using graphlets

(Extended abstract)

David Hartman*† Aneta Pokorná*† Daniel Trlifaj†

Abstract

An analytical procedure that generalizes the degree distribution, called graphlet degree distribution, has recently been applied in complex networks. A graphlet is a graph rooted in one of its vertices. $(\leq n-1)$- degree sequence (gds) of vertex v in a graph G of size n is the sequence of numbers of embeddings of each possible graphlet of size max $n-1$ into G. We study a version of the reconstruction conjecture using $(\leq n-1)$-gds for all vertices called $(\leq n-1)$-graphlet degree distribution (gdd). Our main result shows a reconstruction for 2-connected graphs having vertex-deleted asymmetric subgraph, given that the deleted vertex is either unique or all such vertices form twins. Furthermore, we provide a simple proof reconstructing trees from their $(\leq n-1)$-gdd. A similar result holds for the classical reconstruction from subgraphs. Still, the proof is more complicated than for reconstruction from graphlets, which shows the potential strength of graphlet degree distribution in reconstructing graphs. We also mention several relations between the form of graphlet degree distribution and the sizes and structure of the cutsets of the given graph.

1 Introduction

Graphlet degree sequence is a tool for complex network analysis that describes the local topology of a vertex by counting the number of induced occurrences of all (usually) small rooted subgraphs called graphlets. This term has been introduced by Przulj [13, 14]. Initially, graphlets were utilized to improve the random model of protein-protein interaction networks. Later, graphlets have been applied, for example, to identify cancer gene candidates [11], to understand the topology of social networks [8], to predict connections of microRNAs with diseases [2], or to analyze the structure of the human brain [4]. Even though the graphlet degree sequences are highly applicable, there are few theoretical explanations for their strength.

*Institute of Computer Science, Charles University, Malostranské nám. 25, 118 00, Praha, Czechia. E-mail: hartman@iuuk.mff.cuni.cz, pokorna@iuuk.mff.cuni.cz.

†Computer Science institute, Czech Academy of Sciences, Pod Vodárenskou věží, Praha, Czechia. E-mail: hartman@cs.cas.cz, pokorna@cs.cas.cz, trlifaj@cs.cas.cz. DH, AP, and DT were supported by the Czech Science Foundation Grant No. 23-07074S.

The graphlet analysis lies in counting ($\leq k$)-graphlet degree sequence of every vertex v, that is, the counts of induced occurrences of graphlets of sizes at most k where the root is matched to v. The matrix containing ($\leq k$)-sequence of every vertex in the given graph is called its ($\leq k$)-*graphlet degree distribution (gdd).*

The question of whether the structure of a graph can be recovered from the multiset of its vertex-deleted subgraphs, called the *deck*, known as the reconstruction conjecture (RC), is one of the long-standing unresolved conjectures in graph theory. We examine an alternative version of this conjecture, asking whether a graph is reconstructible given its ($\leq n-1$)-*graphlet degree distribution* (gdd). The intuition behind this effort is that the information about the number of rooted subgraphs for respective vertices seems to provide richer information compared to the deck. We can support this intuition by showing that for the 2-connected case, the deck can be easily obtained from the ($\leq n-1$)-gdd.

Originally, it was assumed that solving RC for asymmetric graphs would be easier than for symmetric ones ([1], page 446). However, there are no results related to the standard reconstruction conjecture using asymmetry. Hence, the following theorem shows a difference between ($\leq n-1$)-graphlet degree distribution and the reconstruction deck.

Theorem 1. *Let H be a 2-vertex connected graph containing a vertex $v \in V(H)$ for which $H \setminus \{v\}$ has no non-trivial automorphism, i.e. is rigid, and one of the following holds*

1. *(unique) There is no vertex $v' \in V(H) \setminus \{v\}$ such that $H \setminus \{v\} \cong H \setminus \{v'\}$*
2. *(twins) If there are vertices $v^1, \ldots, v^k$, $k \geq 2$, such that $H \setminus \{v^i\} \cong H \setminus \{v\}$ $\forall i \in [k]$, then also $N_H(v^i) = N_H(v)$ $\forall i \in [k]$.*

Then G is reconstructible from its $(n-1)$-graphlet degree distribution.

In other words, a 2-connected graph can be reconstructed from its $(n-1)$-gdd if it has a rigid vertex-deleted subgraph H_{asym}, which can be either obtained only by a removal of a unique vertex, or if all vertices whose removal leads to H_{asym} share the same neighborhood. Note that our proof implies that there may be at most two such twin vertices, i.e. $k \leq 2$.

This result covers many graphs - by the well-known fact that almost all finite graphs are asymmetric [3] - hence vertex-deleted asymmetric subgraphs should not be rare. Note that the validity of conditions 1. and 2. of Theorem 1 can be easily assessed given the ($\leq n-1$)-graphlet degree distribution of H.

We say that a pair of vertices $u, v \in V(H)$ is *pseudosimilar* whenever $G \setminus \{u\} \cong G \setminus \{v\}$, but there is no automorphism mapping u to v [5]. Theorem 1 can be rephrased using this definition as follows.

Theorem 1 (Alternative formulation)**.** *Let H be a 2-vertex connected graph containing a vertex $v \in V(H)$ such that $H \setminus \{v\}$ is rigid and one of the following holds*

1. *There is no vertex $v' \in V(H)$ pseudosimilar to v.*
2. *All vertices $v^1, \ldots, v^k$, $k \geq 2$ pseudosimilar to v have the same neighbourhood as v.*

Then G is reconstructible from its $(n-1)$-graphlet degree distribution.

2 Preliminaries

Throughout this work, $H = (V, E)$ denotes an unoriented graph with vertices V and edges $E \subseteq \binom{V}{2}$. The fact that an edge connects vertices u and v is denoted as $\{u, v\} \in E$ or, more simply, $uv \in E$. Unless stated otherwise, $n = |V|$. For $v \in V$, $N_H(v)$ is the *neighbourhood* of v, i.e. $\{w \mid vw \in E\}$. For this and the following definitions, the subscript H is omitted whenever H is clear from the context. Then $\deg_H(v) = |N_H(v)|$ is the *degree* of a vertex v. The distance $d_H(u, v)$ between two vertices $u, v \in V(H)$ is the length of the shortest path between u and v. The *eccentricity* $\epsilon(v)$ is $\max_{u \in V(H)}\{d(u, v)\}$. The symbol $\cong$ stands for graph isomorphism. The set of all non-isomorphic graphs is denoted by $\mathcal{G}$, $\mathcal{G}_n = \{G \in \mathcal{G} |\ |V(G)| = n\}$ and $\mathcal{G}_{\leq n} = \{G \in \mathcal{G} |\ |V(G)| \leq n\}$.

A *graphlet* rooted in r is a pair (G, r) where G is a subgraph of H with $|V(G)| < n = |V(H)|$ and $r \in V(G)$. We call the graph G the *underlying graph* of the graphlet (G, r) and assume that there exists a function $U((G, r)) = G$. Let $\mathcal{G}'$ denote the set of all non-isomorphic graphlets, ${\mathcal{G}'}_n = \{G \in \mathcal{G}' |\ |V(G)| = n\}$, ${\mathcal{G}'}_{\leq n} = \{G \in \mathcal{G}' |\ |V(G)| \leq n\}$ and ${\mathcal{G}'}_{[H]}$ is the set of all graphlets with underlying graph H. Graphlets (G, r) and (H, s) are *isomorphic*, denoted as $(G, r) \cong (H, s)$, if $G \cong H$ and the respective isomorphism maps r to s. We use a fixed ordering γ of $\mathcal{G}$, such that $\gamma(G) < \gamma(H)$ whenever $|V(G)| < |V(H)|$ and $\gamma(G) = \gamma(H)$ iff $G \cong H$ and a fixed ordering ϑ of $\mathcal{G}'$, such that $\vartheta((G, r)) < \vartheta((H, s))$ whenever $\gamma(G) < \gamma(H)$ and $\vartheta((G, r)) = \vartheta((H, s))$ iff $(G, r) \cong (H, s)$. The graphlet (G, r) with $\vartheta((G, r)) = i$ is denoted by $\mathbf{G}^i$ and its underlying graph is $G = U(\mathbf{G}^i)$.

There are several choices of the ordering ϑ. For graphlets up to size 5, it was given explicitly by [14]. However, Hasan et al. [7] gave a similar ordering based on adjacency matrix, which is deterministically defined for all graphlets.

We say that a vertex $v \in V(H)$ *touches* a graphlet (G, r), if there exists an embedding e of G inside H such that $e(r) = v$. The *graphlet degree* $deg_H((G, r), v)$ of a vertex $v \in V(H)$ and graphlet (G, r) in H is the number of times v touches (G, r). The $\leq$ *k-graphlet degree sequence* (shortly $\leq k$-gds) of $v \in V(H)$ is a vector of values $(deg_H(\mathbf{G}^i, v) \mid i \in \{\vartheta({\mathcal{G}'}_{\leq k})\})$. Finally, $\leq$ *k-graphlet degree distribution* (or shortly $\leq k$-gdd) of a graph H is a matrix D with dimensions $|V(H)| \times |\vartheta({\mathcal{G}'}_{\leq k})|$ where $D_{i,j} = deg_H(\mathbf{G}^j, i)$.

3 Relationship with graph connectivity

The following lemma shows that H must be 2-vertex connected if we do not want to lose any information when moving from H to its $(n-1)$-graphlet degree distribution. Otherwise, some of the vertex-deleted graphs become disconnected, split into a larger number of smaller graphlets and are not present in the $(n-1)$-graphlet degree distribution of H.

Lemma 2. *Consider a connected graph, H, on n vertices. If H is 2-vertex connected, $(n-1)$-gdd of H determines the $(\leq n-2)$-gdd of H.*

Remark. This result has an analog in the RC, known as Kelly's lemma. It states that the number of induced subgraphs of a graph G can be easily obtained from the deck of G [9].

Proof. If H is 2-vertex connected, then any graphlet $\mathbf{G}^i$ of size $\ell < (n-1)$ is induced exactly $(n-\ell)$-times in the $(n-1)$-graphlet degree distribution of H, each time obtained by removal of a different vertex from $V(H) \setminus V(U(\mathbf{G}^i))$. For each vertex v and graphlet $\mathbf{G}^i$ of size less than $(n-1)$, we obtain the entry $D_{v,i}$ of the $(\leq n-2)$-graphlet degree sequence by simply counting how many times v touches $\mathbf{G}^i$ in $(n-1)$-sized graphlets of H rooted at v and divide this number by $(n - |V(U(\mathbf{G}^i))|)$ to obtain $D_{v,i}$. □

Note it has been shown that if the reconstruction conjecture holds for all 2-connected graphs, then it holds for all graphs [15]. Lemma 2 can be easily generalized.

Proposition 3. *Consider a connected graph, H, on n vertices. If H is k-vertex connected, $(n-k+1)$-graphlet degree sequence of H determines $(\leq n-k)$-graphlet degree sequence of H.*

We also proved the following relation with the connectivity of a graph.

Proposition 4. *Let H be a graph, $|V(H)| = n$. Then H is k-vertex connected if and only if $\sum_{(G,j)\in \mathcal{G}'_{n-k+1}} deg_H((G,j),v) = \binom{n-1}{k-1}$ for every $v \in V(H)$.*

Corollary 5. *Let H be a connected graph on n vertices with its $(n-1)$-graphlet degree sequence $D^{n\times m}$. The property $\sum_{i=1}^{m} D_{v,i} = n-1$ (1) holds for zero, one or all vertices of H. These cases correspond to H having more different articulations, precisely one articulation and no articulation, respectively. Moreover, if there is precisely one articulation, then it is the vertex having property* (1).

Corollary 6. *Let H be a $(k-1)$-vertex connected graph on n vertices with its $(n-k+1)$-graphlet degree sequence $D^{n\times m}$. Let p be the number of vertices that satisfy $\sum_{i=1}^{m} D_{v,i} = \binom{n-1}{k-1}$. Then $p \in \{0, 1, \ldots, k-1, n\}$. Furthermore, $p \leq k-1$ occurs when there exist p vertices lying in every vertex cut of size k and these vertices are identified by satisfying* (1). *Finally, $p = n$ holds if H is $(k+1)$-connected.*

4 Reconstructions using graphlet degree distribution

When given a $(n-1)$-graphlet degree sequence of a 2-vertex connected graph H on n vertices (and thus indirectly given the whole $(\leq n-1)$-gds of H by Lemma 2), we can determine the deck of H and thus all graphs reconstructible from their deck are also reconstructible from their $(n-1)$-gdd.

The reconstruction conjecture has already been proved for trees [10]. The proof of reconstruction from graphlet degree distribution is much easier, which suggests potentially higher descriptive power of $(\leq n-1)$-gdd.

Observation 7. *Let us have a tree T on $3 \leq n$ vertices. Given $(\leq n-1)$-gds of T, we can uniquely reconstruct the graph T.*

Idea of the proof. At first, we identify the central vertex c of T, i.e., the vertex that lies in the center of the longest path of T. We say that a graphlet $\mathbf{G}^i$ is *path-like*, if $U(\mathbf{G}^i) \cong P_k$ for some k and if the root of $\mathbf{G}^i$ lies at the endpoint of the path. Let us denote $lp(v)$ the length of the longest path-like graphlet with its end rooted in $v \in V(T)$. Note that $lp(v)$ is easily obtained from $(\leq n-1)$-gdd of T as we know the ordering of graphlets ϑ including the path-like ones. Observe that $lp(v) = \epsilon(v)$. Then the central vertex c is the vertex with minimal $lp(v)$. It is well known that a tree has at most two possible central vertices. Choose an arbitrary one.

Next, we proceed by examining only the $(\leq n-1)$-gds of c in T. We denote a graphlet (G, c) as a *trunked tree* rooted in c if G is a tree and $deg_G(c) = 1$. Starting from the maximal ones, we will process all trunked trees rooted in c from the given graphlet degree sequence of c. Observe that each trunked tree rooted in c identifies a tree of the forest induced by $V(T) \setminus \{c\}$. We construct the structure of T by first adding vertex c and continue by adding maximum trunked trees as follows. Each time we add the maximum trunked tree T_m to T, we also subtract the $(\leq n-1)$-gds of c in T_m from the processed $(\leq n-1)$-gds of c in T. The last step ensures that we do not account for the same subtree (nor its subset) twice. $\square$

The proof of Theorem 1 follows.

Proof. At first, from the 2-connectivity of H and Lemma 2 we obtain the $(\leq n-1)$-gdd of H. Let $\mathcal{H}'_{\text{n-1}}$ be the set of graphlets of size $n-1$ touching at least one vertex of H, i.e. $\sum_{k=1}^{n} D_{k,i} \geq 0$. By the assumption, there is a non-empty set $\mathcal{H}_{\text{sat}}$ of asymmetric graphs satisfying either condition 1 or 2 in $U(\mathcal{H}'_{\text{n-1}})$.

Note that $\mathcal{H}_{\text{sat}}$ can be identified using only graphlet degree distribution of H and the knowledge of the graphlet indexing function ϑ. A graph F is asymmetric (or rigid) iff $\mathcal{G}'_{[F]}$ contains $|V(F)|$ distinct graphlets. Condition 1 is satisfied if $\sum_{k=1}^{n} D_{k,i} = 1$ for every $i = \vartheta(\mathbf{F})$, $\mathbf{F} \in \mathcal{G}'_{[F]}$. Moreover, for $F \in U(\mathcal{H}'_{\text{n-1}})$ such that $F \cong H \setminus \{v\}$, $N_H(v) = \{x \in V(F) \mid \deg_H(x) \neq \deg_F(x)\}$, where $deg_H(x) = D_{\vartheta((K_2,x)),x} = D_{0,x}$ and $\deg_F(x)$ is the same as the degree of the root in $\vartheta((F,x))$, which can be inferred from the structure of (F,x).

We fix a graph $H_{\text{asym}} = H \in \mathcal{H}_{\text{sat}}$ such that $\min\{\vartheta(\mathbf{H}) \mid \mathbf{H} \in \mathcal{G}'_{[H]}\}$ is minimized.

Observation 8. *H_{asym} has orbits of size one, so every occurrence of H_{asym} in H adds 1 to every column indexed by $p \in \vartheta(\mathcal{G}'_{[H_{asym}]})$, i.e. adds an occurrence of every graphlet with H_{asym} as underlying graph, and to every row indexed by $V(H) \setminus \{v\}$, where $H \setminus \{v\} \cong H_{asym}$, i.e. adds every possible rooting of H_{asym}.*

The next step is to identify (or fix if there are multiple choices) the vertex $v \in V(H)$ in $V(H) \setminus V(H_{\text{asym}})$.

If condition 1 holds, then for every $w \in V(H) \setminus \{v\}$, $D_{w,p}$ is nonzero (specifically equal to one) for exactly one $p \in \vartheta(\mathcal{G}'_{[H_{\text{asym}}]})$, that is, there is exactly one graphlet rooted at w with the underlying graph isomorphic to H_{asym}. In contrast, there is no such entry for v.

If condition 2 holds, then for each $i, j \in [k]$, there exists an automorphism $f_{i,j}$ of H such that $f(v^i) = v^j$ and $f(u) = u \; \forall u \in V(H) \setminus \{v^i, v^j\}$.

Now if $k > 2$, there exist vertices v^h, v^i, v^j such that $H \setminus \{v^h\} \cong H \setminus \{v^i\} \cong H \setminus \{v^j\} \cong H_{\text{asym}}$. But then $f_{i,j}$ reduced to $V(H) \setminus \{v^h\}$ is a non-trivial automorphism of $H \setminus \{v^h\}$, contradicting $H \setminus \{v^h\} \cong H_{\text{asym}}$. So $k = 2$ and we have $G \cong H \setminus \{v^1\} \cong H \setminus \{v^2\} \cong H_{\text{asym}}$. The submatrix of D induced by the columns indexed by $\vartheta(\mathcal{G}'_{[G]})$ has the following values:

For every $w \in V(H) \setminus \{v^1, v^2\}$, $D_{w,p} = 2$ for $p = \vartheta((G,w))$ and $D_{w,q} = 0 \; \forall q \in \vartheta(\mathcal{G}'_{[G]}), q \neq p$.

For v^1, v^2, $D_{v^1,p} = 1$ if $p \in \vartheta((G, v^1))$ and $D_{v^1,q} = 0 \; \forall q \in \vartheta(\mathcal{G}'_{[G]}), q \neq p$ and symmetrically for values in $D_{v^2,*}$.

Finally, we reconstruct H by taking H_{asym} and introducing a new vertex v. Because H_{asym} is asymmetric, each vertex $w \in V(H_{\text{asym}})$ has its own automorphism orbit, for which we can identify its adjacency to v by the following process.

Suppose $D_{p,w} \neq 0$ for $p \in \{\vartheta(\mathcal{G}'_{[H_{\text{asym}}]}\}$, hence p is the unique index of the graphlet $(H \setminus \{v\}, w)$. Now if $\deg_H(w) = \deg_H(\mathbf{G}^0, w) = D_{0,w}$, where $U(\mathbf{G}^0) = K_2$ is the same as the degree of the root in (H_{asym}, w), then $wv \notin E(H)$, otherwise $wv \in E(H)$.

$\square$

Acknowledgements

The authors were supported by the Czech Science Foundation Grant No. 23-07074S.

References

[1] K. J. Asciak, M.A. Francalanza, J. Lauri, and W. Myrvold. A survey of some open questions in reconstruction numbers. *Ars Combinatoria*, 097:443–456, 2010.

[2] Xing Chen, Na-Na Guan, Jian-Qiang Li, and Gui-Ying Yan. Gimda: Graphlet interaction-based mirna-disease association prediction. *Journal of Cellular and Molecular Medicine*, 22(3):1548–1561, 2018.

[3] P. Erdős and A. Rényi. Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungaricae*, 14:295–315, 1963.

[4] P. Finotelli, C. Piccardi, E. Miglio, and P. Dulio. A graphlet-based topological characterization of the resting-state network in healthy people. *Frontiers in neuroscience*, 15:665544–, 2021.

[5] F. Harary and Palmer E. On similar points of a graph. *Journal of Mathematics and Mechanics*, 15(4):623–630, 1966.

[6] F. Harary and Palmer E. On similar points of a graph. *Journal of Mathematics and Mechanics*, 15(4):623–630, 1966.

[7] Adib Hasan, Po-Chien Chung, and Wayne Hayes. Graphettes: Constant-time determination of graphlet and orbit identity including (possibly disconnected) graphlets up to size 8. *PLOS ONE*, 12(8):1–12, 08 2017.

[8] J. Janssen, M. Hurshman, and N. Kalyaniwalla. Model Selection for Social Networks Using Graphlets. *Internet Mathematics*, 8(4):338 – 363, 2012.

[9] P.J. Kelly. *On Isometric Transformations*. University of Wisconsin–Madison, 1942.

[10] J. Lauri. Proof of harary's conjecture on the reconstruction of trees. *Discrete Mathematics*, 43(1):79–90, 1983.

[11] T. Milenković, V. Memišević, A. K. Ganesan, and N. Pržulj. Systems-level cancer gene identification from protein interaction network topology applied to melanogenesis-related functional genomics data. *Journal of The Royal Society Interface*, 7(44):423–437, 2010.

[12] Václav Nýdl. Finite graphs and digraphs which are not reconstructible from their cardinality restricted subgraphs. *Commentationes Mathematicae Universitatis Carolinae*, 022(2):281–287, 1981.

[13] N. Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):e177–e183, 01 2007.

[14] N. Pržulj, D. G. Corneil, and I. Jurisica. Modeling interactome: scale-free or geometric? *Bioinformatics*, 20(18):3508–3515, 07 2004.

[15] Yang Yongzhi. The reconstruction conjecture is true if all 2-connected graphs are reconstructible. *Journal of Graph Theory*, 12(2):237–243, 1988.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

All Vertices Of Nowhere Dense Modelings Are Traceable

(Extended abstract)

David Hartman* † Tomáš Hons* Jaroslav Nešetřil*

Abstract

We consider the problem within the field of graph limit theory of tracing vertices of the limit object to the converging sequence. Let $\{G_n\}$ be a sequence of finite graphs that FO-converges to a modeling limit L and let a be an arbitrary vertex of L. Under the assumption that L is nowhere dense, we prove that the vertex a is traceable. Moreover, we show that vertices of every bounded degree graphing are traceable.

1 Introduction

One of the goals of graph limit theory is to find a convenient limit notion of a sequence of graphs [18]. A large range of approaches exists in the literature, focusing on various properties of the graph sequence, such as subgraph densities [17], neighborhood densities [3], or adjacency operators [2].

A strong convergence notion is the first-order convergence, also called structural convergence, that is based on solution densities of first-order formulas [22]. The first-order convergence is a general framework of convergence for relational structures. Although, we will be interested in the case of graphs with additional unary predicates and constants, we define the notions in full generality.

Fix a signature relational signature λ, possibly with constants. The *solution density* of a λ-formula ϕ and in a finite λ-structure A, denoted by $\langle \phi, A \rangle$, is the probability that ϕ is satisfied by a tuple of vertices of A selected uniformly at random (for ϕ sentence, we set $\langle \phi, A \rangle = 1$ if $A \models \phi$, and $\langle \phi, A \rangle = 0$ otherwise). Let $\mathrm{FO}(\lambda)$ be the set of all first-order λ-formulas. For a set $X \subseteq \mathrm{FO}(\lambda)$, we say that a sequence of finite A-structures $\{A_n\}$ is X*-convergent* if the sequence $\{\langle \phi, A_n \rangle\}$ converges for each formula $\phi \in X$. The limit structure L, called *modeling*,

*Computer Science Institute (IÚUK, MFF), Charles University, Prague, Czech Republic.
†Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic.
Emails: {hartman, honst, nesetril}@iuuk.mff.cuni.cz. The work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Synergy Grant DYNASNET, grant agreement No 810115). TH is further supported by Project 24-12591M of the Czech Science Foundation (GAČR).

is a λ-structure on a standard Borel space with a measure ν satisfying that all the first-order definable sets of L are measurable. The value $\langle\phi, L\rangle$ is defined as the measure of the set $\phi(L)$, the set of solutions of ϕ in L, using the appropriate power of the measure ν (corresponding to the number of free variables of ϕ). All our modelings are of size continuum. Nevertheless, note that for a finite λ-structure A considered as a modeling on a finite space with the uniform measure both definitions of $\langle\phi, A\rangle$ coincide. A modeling L is an X-limit of an X-convergent sequence $\{A_n\}$ if $\lim_n\langle\phi, A_n\rangle = \langle\phi, L\rangle$ for each $\phi \in X$.

Here we focus on the case of graphs that may contain additional marks on vertices. That is, the signature λ contains a single binary symbol $\sim$ that is realized by a symmetric and irreflexive relation, and possibly finitely many unary symbols and constants. We call these structures *colored graphs*. A graph without additional symbols is called a *simple graph*. We mostly speak about graphs if their exact nature is understood from the context.

The set of first-order formulas in the language of simple graphs is denoted by FO($\sim$) instead of FO($\{\sim\}$) for brevity. The strength of the first-order convergence given by the expressiveness of the first-order logic causes that not every FO($\sim$)-convergent sequence of graphs admits a modeling limit. However, there are general theorems asserting existence of modeling limits under further assumptions [21][12]. In particular, is is known that modeling limits exist for all sequences of graphs from a monotone class $\mathcal{C}$ if and only if $\mathcal{C}$ is a nowhere dense class [21].

The nowhere dense property [19] is central to our result. We say that a graph class $\mathcal{C}$ is *nowhere dense* if there is a function $t : \mathbb{N} \to \mathbb{N}$ such that for each $d \in \mathbb{N}$ no graph in $\mathcal{C}$ contains a d-subdivision of the complete graph $K_{t(d)}$ as a subgraph. A d-subdivision of $K_{t(d)}$ in the graph obtained by subdividing each edge of $K_{t(d)}$ by at most d new vertices. We call an infinite graph G nowhere dense if there is $t : \mathbb{N} \to \mathbb{N}$ such that for each $d \in \mathbb{N}$ a d-subdivision of the graph $K_{t(d)}$ is not a subgraph of G. We say that the minimal such function $t : \mathbb{N} \to \mathbb{N}$ is the *sparsity function* of G. A graph modeling L is nowhere dense if the underlying infinite graph is nowhere dense. Note that for a fixed d and n the property whether a graph contains a d-subdivision of K_n as a subgraph can be expressed by a first-order sentence. This implies that a if graph sequence $\{G_n\}$ FO($\sim$)-converges to a modeling L, then $\{G_n\}$ is nowhere dense (as a graph class) if and only if L is nowhere dense.

Here we consider a problem by Nešetřil and Ossona de Mendez, which is related to the area of network archaeology [13][6][5]. If a is a vertex of a graph G with signature λ, we denote by (G, a) the expansion of G to a $(\lambda \cup \{c\})$-structure, where c is a new constant, interpreting c as the vertex a. We say that (G, a) is the graph G rooted at the vertex a. Upon proving that if L is a modeling and $a \in V(L)$, then (L, a) is also a modeling [22, Lemma 3.2], Nešetřil and Ossona de Mendez asked the following [22, Problem 3.1].[1]

Problem 1. *Suppose that $\{G_n\}$ be a sequence of finite graphs that* FO(λ)*-converges to a modeling L. Let a be a vertex of L. Is there a sequence of vertices $\{a_n\}, a_n \in V(G_n)$, such that the rooted graphs $\{(G_n, a_n)\}$* FO($\lambda \cup \{c\}$)*-converge to the rooted modeling (L, a)?*

If Problem 1 has positive answer for a given sequence $\{G_n\}$, modeling L and its vertex a, we say that the vertex a is *traceable in the sequence* $\{G_n\}$. If a vertex a is traceable in any sequence $\{G_n\}$ that FO(λ)-converges to the modeling L, we call the vertex *a traceable*. Moreover, we extend the definition of traceable vertices to finite tuples. A tuple of vertices $\boldsymbol{a} = (a^1, \dots, a^k)$ of L is traceable in $\{G_n\}$ if there are tuples $\boldsymbol{a}_n = (a^1_n, \dots, a^k_n), \boldsymbol{a}_n \in V^k(G_n)$

[1]The problem was originally asked for general relational structures. Here we consider only the case of graphs.

such that $\{(G_n, \boldsymbol{a}_n)\}$ FO-converges to $(L, \boldsymbol{a})$ in the corresponding signature $\lambda \cup \{c_1, \dots, c_k\}$. Note that assuming that the tuple $(a^1, \dots, a^k)$ is traceable is stronger than assuming that all the individual vertices $a^i, i \in [k]$ are traceable.

It was proved by Christofides and Král' that the answer to Problem 1 is negative in general, even when restricting to the case of simple graphs. However, they also proved that almost all vertices of L are traceable in $\{G_n\}$ [7]. A particular kind of traceable vertices, the algebraic vertices of L, was identified by the authors in [14].

As our main result, we prove that all vertices of the nowhere dense modelings are traceable.

Theorem 1. *Let $\{G_n\}$ be an $\mathrm{FO}(\lambda)$-convergent sequence of graphs with a nowhere dense modeling limit L. Then for any vertex $a \in V(L)$ there is a sequence $\{a_n\}$, $a_n \in V(G_n)$, such that $\{(G_n, a_n)\}$ $\mathrm{FO}(\lambda \cup \{c\})$-converges to (L, a).*

Since a rooted nowhere dense modeling remains a nowhere dense modeling, a repeated application of Theorem 1 shows that all finite tuples of vertices of L are traceable. Moreover, the proof technique used for a certain special case shows that vertices of a residual modeling [20] (with any countable relational signature) are traceable. These include bounded degree graphings [10][22, Lemma 3.43].

2 Preliminaries

Structures All our structures contain a single binary relation, which is symmetric and irreflexive, and possibly finitely many additional unary relations and constants (roots). We call these structures (colored) graphs. The vertex set of a graph G is denoted by $V(G)$. Let $\lambda \subseteq \sigma$ be two signatures of colored graphs. A λ-graph A can be *expanded* into a σ-graph B be realizing the additional symbols in σ. Then B is a *σ-expansion of* A, while A is called the *λ-reduct* of B. Our main results, as well as the proof technique, is concerned with expansion of λ-graphs into richer σ-graphs and its influence on the first-order convergence.

Note that there is a close tie between first-order convergence of λ-graphs and its σ-expansions, so that we can usually omit the explicit mentions of the signature of the convergence. Indeed, it is clear that if the sequence $\{B_n\}$ of σ-graphs is $\mathrm{FO}(\sigma)$-convergent, it is also $\mathrm{FO}(\lambda)$-convergent as $\mathrm{FO}(\lambda) \subseteq \mathrm{FO}(\sigma)$. Moreover, the sequence $\{B_n\}$ is $\mathrm{FO}(\lambda)$-convergent if and only if the sequence $\{A_n\}$ of λ-reducts is $\mathrm{FO}(\lambda)$-convergent as these sequences are identical from the perspective of λ-formulas. Thus, the only non-trivial direction is the transition from $\mathrm{FO}(\lambda)$-convergence of the λ-graphs $\{A_n\}$ to $\mathrm{FO}(\sigma)$-convergence of their σ-expansions $\{B_n\}$. Consequently, we mostly omit the explicit mention of the signature and speak about FO-convergence of a sequence, which is understood as the convergence with respect to their full signature.

Formulas A sentence is a formula without free variables. The set of λ-sentences is denoted by $\mathrm{FO}_0(\lambda)$. The quantifier rank of a formula ϕ, denoted by $\mathrm{qrank}(\phi)$, is the maximal depth of nested quantifiers. The set of all λ-formulas with quantifier rank at most k is denoted by $\mathrm{FO}^{k\text{-qrank}}(\lambda)$. The d-neighborhood of vertex a in the graph G, that is the set $\{v \in V(G) : \mathrm{dist}(v, a) \leq d\}$, is denoted by $N_G^d(a)$. For $d = 1$, we write simply $N_G(a)$. Given a set $S \subseteq V(G)$, we denote by $G[S]$ the subgraph of G induced by S. A formula $\phi(x_1, \dots, x_k)$ is d-local if its satisfaction depends only on the d-neighborhood of the free variables of ϕ and the roots of G, say $r_1, \dots, r_\ell$. That is, for any vertices $a_1, \dots, a_k$, we have

$$G \models \phi(a_1, \dots, a_\ell) \text{ if and only if } G[S] \models \phi(a_1, \dots, a_\ell), \text{ where } S = \bigcup_{i \in [k]} N^d_G(a_i) \cup \bigcup_{i \in [\ell]} N^d_G(r_i).$$

(We can view the roots of G as additional free variables c_i of ϕ with the fixed valuation $c_i \mapsto r_i$.) The set of all d-local λ-formulas is denoted by $\mathrm{FO}^{d\text{-local}}(\lambda)$, and $\mathrm{FO}^{\mathrm{local}}(\lambda) = \bigcup_{d \in \mathbb{N}} \mathrm{FO}^{d\text{-local}}(\lambda)$ is the set of all local formulas.

Recall the Gaifman locality theorem, which states that every first-order formula ϕ can be expressed as a boolean combination of local formulas and sentences of quantifier-rank bounded by a fixed function of qrank(ϕ) [11]. Consequently, the $(\mathrm{FO}_0 \cup \mathrm{FO}^{\mathrm{local}})$-convergence implies the full FO-convergence [22, Theorem 2.23]. More generally, the $(\mathrm{FO}_0 \cup \mathrm{FO}^{d\text{-local}})$-convergence implies the $\mathrm{FO}^{f(d)\text{-qrank}}$-convergence, where $f : \mathbb{N} \to \mathbb{N}$ is a fixed unbounded monotone function, which can be deduced from Gaifman's theorem.

Statements about traceability Recall the theorem by Christofides and Král', which will be one of our main tools [7].

Theorem 2. *Suppose that G_n $\mathrm{FO}(\lambda)$-converges to a modeling L with the measure ν. Then a ν-random vertex of L is traceable in G_n with probability* 1.

We say that a formula ϕ is *algebraic* in L, if its solution set $\phi(L)$ is finite. A vertex a is said to be algebraic in L if it is a solution of a formula $\phi(x)$, which is algebraic in L. The following was proved by the authors in [14].

Theorem 3. *Each algebraic vertex of L is traceable.*

Unlinking and linking We define two graph operations Unlink and Link. The former takes a graph (G, a) and creates the graph $\mathrm{Unlink}(G, a) = (H, a, M)$ such that it removes all edges adjacent to the vertex a from G and marks its former neighbors, the set $N_G(a)$, by a new unary symbol M. The linking operation is the exact converse. That is, if a is an isolated vertex in H and $a \notin M$, then $\mathrm{Link}(H, a, M)$ returns the graph (G, a), where a is connected to all the vertices of M, and forgets the mark M. It is immediate that the operations of unlinking and linking compose to identity. Formally, both operations can be expressed as a quantifier-free interpretation of dimension 1 [22, Definition 2.42], which justifies the following claim.

Lemma 4. *Suppose that (G_n, a_n) $\mathrm{FO}^{k\text{-qrank}}$-converges to (L, a). Then $\mathrm{Unlink}(G_n, a_n)$ $\mathrm{FO}^{k\text{-qrank}}$-converges to $\mathrm{Unlink}(L, a)$ in the respective signature.*

Conversely, suppose that (H_n, a_n, M_n) $\mathrm{FO}^{k\text{-qrank}}$-converges to (L, a, M). Then $\mathrm{Link}(H_n, a_n, M_n)$ $\mathrm{FO}^{k\text{-qrank}}$-converges to $\mathrm{Link}(L, a, M)$ in the respective signature.

3 Proof of Theorem 1

Given a sequence $\{G_n\}$ that FO-converges to a nowhere dense modeling L and a vertex a of L, our goal is to find a sequence $\{a_n\}$, $a_n \in V(G_n)$, such that $\{(G_n, a_n)\}$ FO-converges to (L, a). To motive our further approach, we first partially solve the problem for special vertices a of L whose close neighborhood has measure 0. Note that for a fixed $d \in \mathbb{N}$, the set $N^d_L(a)$ is definable in (L, a) and hence is measurable.

Lemma 5. *Fix $d \in \mathbb{N}$ and let a be a vertex of L whose $6d$-neighborhood has measure 0. Then there are vertices $a_n \in G_n$ such that $\{(G_n, a_n)\}$ converges to (L, a) with respect to all sentences and d-local formulas in the extended signature.*

Sketch. Since the measure of the $2d$-neighborhood of a is 0, the probability that the d-neighborhoods of a and a randomly chosen vertex from L intersect is 0. Therefore, for a d-local formula $\phi(x_1, \dots, x_k)$ in the extended signature of (L, a), the value $\langle \phi, (L, a) \rangle$ can be computed from the local properties of the randomly chosen vertices $x_1, \dots, x_k$ and the local properties of the vertex a. Consequently, to ensure that $\{(G_n, a_n)\}$ $\mathrm{FO}^{d\text{-local}}$-converges to (L, a), it is enough that the vertices a_n approximate the local properties of a and that the measure of their $2d$-neighborhood tends to 0. Furthermore, to ensure the FO_0-convergence, the vertices a_n must approximate the type of the vertex a (which also entails the approximation of the local properties of a).

More precisely, let $\phi_1'(x), \phi_2'(x), \dots$ be an enumeration of formulas satisfied by a in L and let $\phi_k(x) = \bigwedge_{i=1}^{k} \phi_i'(x)$. Then we want to find vertices a_n satisfying that

(i) for each $k \in \mathbb{N}$, there is $N \in \mathbb{N}$ such that for all $n \geq N$ we have $G_n \models \phi_k(a_n)$,

(ii) the measure of the $2d$-neighborhood of a_n, the value $\nu_{G_n}(N^{2d}_{G_n}(a_n))$, tends to 0.

To do so, it is sufficient to find for each fixed $k \in \mathbb{N}$ a sequence of vertices $a_n^k \in G_n$ such that from certain index N on, we have $G_n \models \phi_k(a_n^k)$, while also $\nu_{G_n}(N^{2d}_{G_n}(a_n^k)) \to 0$. Then we can select the desired vertices $a_n \in V(G_n)$ by diagonalization of the sequences $\{a_n^k\}$.

A natural strategy is then choose as a_n^k the vertex v of G_n that minimalizes $\nu_{G_n}(N^{2d}_{G_n}(v))$ among the solutions of the formula $\phi_k(x)$ (disregarding at most finitely many graphs G_n without a solution of ϕ_k). We claim that this yields the required property $\nu_{G_n}(N^{2d}_{G_n}(a_n^k)) \to 0$, otherwise we find a contradiction. That is, suppose for contradiction that

$$\limsup_n \min_{v \in \phi(G_n)} \nu_{G_n}(N^{2d}_{G_n}(v)) > 0.$$

We consider the formula $\xi(x_1, \dots, x_\ell)$, ℓ is sufficiently large, stating that “there is a vertex satisfying $\phi(x)$ with distance $> 6d$ from each x_i.” Then we can show that $\liminf_n \langle \xi, G_n \rangle < 1$, while $\langle \xi, L \rangle = 1$, reaching a contradiction with the FO-convergence of $\{G_n\}$ to L. In particular, the second part $\langle \xi, L \rangle = 1$ is witnessed by the vertex a with $\nu_L(N^{6d}_L(a)) = 0$, while the first part requires further justification. □

This argument shows that we can solve the case of vertices which are far from the support of ν_L. Moreover, this generalizes for vertices that can separated from the support in a convenient way. That is, in our case, by the Unlink operation.

Our proof of Theorem 1 is essentially algorithmic. Suppose that there is $d \in \mathbb{N}$ such that the d-neighborhood of a in L has positive measure. We will iteratively root traceable vertices $r^1, r^2, \dots$ in the d-neighborhood of a. By Theorem 2, we can obtain each r^i by a random draw from the set. In the i-th step, upon marking the new vertex r^i, we try to apply the algebraicity criterion from Theorem 3 to locate additional traceable vertices in $N^d_L(a)$ (vertices may become algebraic in the expanded structure). Then we apply the unlink operation to the vertex r^i (and the approximating sequence a_n^i) as well as to the additional algebraic traceable vertices (and their respective sequences), obtaining a modeling L^i with $N^d_{L^i}(a) \subset N^d_{L^{i-1}}(a) \subset \cdots \subset N^d_L(a)$. We continue with further iterations until $N^d_{L^i}(a) = 0$ for some i.

We claim that the number of iterations is bounded in terms of the sparsity function $t : \mathbb{N} \to \mathbb{N}$ of L.

Claim 6. *We have $N^d_{L^i}(a) = 0$ for some $i \le g(t,d)$, where g is a fixed function.*

With Claim 6 in hand, the rest of the proof uses only standard techniques. We apply Lemma 5 to find vertices $a^d_n \in G_n$ such that $\{(G^i_n, a^d_n)\}$ converges to (L^i, a) with respect the set $(\mathrm{FO}_0 \cup \mathrm{FO}^{d/6\text{-local}})$, thus also with respect to the set $\mathrm{FO}^{f(d/6)\text{-qrank}}$, due to Gaifman's theorem. By Lemma 4, the linking operation preserves this partial convergence to the original graphs $\{(G_n, a^d_n)\}$ and (L, a). That is, we can ensure convergence of every particular formula ϕ with $\mathrm{qrank}(\phi) \le f(d/6)$. Since the function f is unbounded, each formula converges for a sufficiently large value of d. Then we can diagonalize the sequences $\{a^d_n\}$ to find the desired sequence $\{a_n\}$, ensuring convergence of all formulas simultaneously. That is, satisfying the conclusion of Theorem 1.

In the rest, we focus on the combinatorially more interesting justification of Claim 6.

3.1 Proof of Claim 6

Fix $d \in \mathbb{N}$ and let $t : \mathbb{N} \to \mathbb{N}$ be the sparsity function of L. Our goal is to prove that after at most $g(t,d)$ iterations of rooting vertices r_i, rooting algebraic vertices, and their further unlinking, we obtain a modeling where the d-neighborhood of a has measure 0.

Let us start with a simple Ramsey-type statement.

Lemma 7. *Suppose that G is a graph on $X \dot\cup Y \dot\cup Z$ containing a path P_{xy} of length at most d for each $x \in X, y \in Y$ such that $V(P_{xy}) \cap X = \{x\}, V(P_{xy}) \cap Y = \{y\}$. Moreover, suppose that*

$$\text{for every } x \in X \text{ and } y, y' \in Y, y \ne y', \text{ it holds } V(P_{xy}) \cap V(P_{xy'}) = \{x\}. \qquad (*)$$

If $|Y| \ge 4d|X|^2$, then G contains a $(2d-1)$-subdivision of $K_{|X|}$ as a subgraph.

We can state the condition $(*)$ in a succinct way that the paths $\{P_{xy} : y \in Y\}$ form a subdivision of a star centered at x. The core argument behind Claim 6 is the following.

Lemma 8. *Suppose that a is the center of $(d-1)$-subdivided star with $\ell \ge 4d \cdot t(2d-1)^2$ leaves that form a traceable tuple $\boldsymbol{r}$. Then a is traceable.*

Proof. There is a formula $\phi(x)$, which uses the constants $\boldsymbol{c}$ of roots $\boldsymbol{r}$, that asserts the existence of such a collection of paths forming a $(d-1)$-subdivided star centered in x with leaves $\{r_i : i \in [\ell]\}$. Clearly, $(L, \boldsymbol{r}) \models \phi(a)$. We claim that $\phi(x)$ is algebraic in $(L, \boldsymbol{r})$. If not, we use Lemma 7 for $Y = \boldsymbol{r}$ and X an arbitrary set of $t(2d-1)$ solutions of $\phi(x)$ to find a $(2d-1)$-subdivided clique of size $t(2d-1)$ as a subgraph of L, which is a contradiction. Therefore, $\phi(x)$ is algebraic in $(L, \boldsymbol{r})$ and we conclude that a is traceable by Theorem 3. □

Now we can sketch the proof of Claim 6.

Sketch of Claim 6. We iteratively build a tree T as a collection of paths in the modeling L from the traceable vertices r^i to a (oriented towards a). We start with an empty tree. Upon adding the root r^i to the d-neighborhood of a in the modeling L^{i-1}, we consider an arbitrary shortest path P from r^i to a that does not create a cycle when adding it to T (there is always such a path). Note that P is also a path in the original modeling L. We can assign to each vertex of T a level to be its distance to a in T. Clearly, the maximal level is d.

Once an internal vertex v of T has sufficiently many children in T, we deduce by Lemma 8 that it is traceable (a vertex with ℓ children in T is the center of a star with at least ℓ traceable leaves) and we unlink the vertex v when creating the modeling L^i. Consequently, the vertex v cannot acquire new descendants in T as it does belong to the d-neighborhood of a in L^i and further modelings. It follows that vertices of T have bounded degrees, which together with the bounded depth of T implies that T can have only boundedly many leaves, say at most $g(t,d) = \prod_{i=1}^{d} h(t,i)$, where $h(t,i) = 4i \cdot t(2i-1)^2$ is the function from Lemma 8. Since we add a new leaf to T in each iteration, the process may continue for at most $g(t,d)$ many steps before the vertex a is detected as algebraic and subsequently unlinked, ensuring $\nu_{L^i}(N^d_{L^i}(a)) = 0$. □

Note that the process may terminate anytime sooner when $\nu_{L^i}(N^d_{L^i}(a)) = 0$. That is, we do not necessarily reach the state when the vertex a is algebraic and unlinked.

4 Comments

We remark that Lemma 5 applies for general relational structures, where the distances are measured in the corresponding Gaifman graphs. Therefore, any vertex of a general modeling L whose connected component has measure 0 is traceable. This is the case of residual modelings [20], which include bounded degree graphings [10][22, Lemma 3.43].

A growing body of evidence suggests that the right model-theoretic generalization of nowhere dense to the hereditary setting is given by the notion of monadic dependence [1][9][24]. We conjecture that monadically dependent graph modelings are traceable.

Conjecture 1. *Let L be a monadically dependent graph modeling. Then all vertices of L are traceable.*

As a first step towards Conjecture 1 may be to tackle the better understood case of monadically stable graphs [8][9].

References

[1] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European Journal of Combinatorics*, 36:322–330, February 2014.

[2] Ágnes Backhausz and Balázs Szegedy. Action convergence of operators and graphs. *Canadian Journal of Mathematics*, 74(1):72–121, September 2020.

[3] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6(none), jan 2001.

[4] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Counting Graph Homomorphisms. In *Topics in Discrete Mathematics*, pages 315–371. Springer Berlin Heidelberg, 2006.

[5] Anna M. Brandenberger, Luc Devroye, and Marcel K. Goh. Root estimation in galton–watson trees. *Random Structures & Algorithms*, 61(3):520–542, January 2022.

[6] Sébastien Bubeck, Luc Devroye, and Gábor Lugosi. Finding adam in random growing trees. *Random Structures & Algorithms*, 50(2):158–172, November 2016.

[7] Demetres Christofides and Daniel Král'. First-order convergence and roots. *Combinatorics, Probability and Computing*, 25(2):213–221, feb 2015.

[8] Jan Dreier, Nikolas Mählmann, and Sebastian Siebertz. First-order model checking on structurally sparse graph classes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC '23, pages 567–580. ACM, June 2023.

[9] Jan Dreier, Nikolas Mählmann, and Szymon Toruńczyk. Flip-breakability: A combinatorial dichotomy for monadically dependent graph classes. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC '24, pages 1550–1560. ACM, June 2024.

[10] Gábor Elek. Note on limits of finite graphs. *Combinatorica*, 27(4):503–507, jul 2007.

[11] Haim Gaifman. *On Local and Non-Local Properties*, pages 105–135. Elsevier, 1982.

[12] Andrzej Grzesik, Daniel Král', and Samuel Mohr. *Strong Modeling Limits of Graphs with Bounded Tree-Width*, pages 273–279. Springer International Publishing, 2021.

[13] John Haigh. The recovery of the root of a tree. *Journal of Applied Probability*, 7(1):79–88, April 1970.

[14] David Hartman, Tomáš Hons, and Jaroslav Nešetřil. Structural convergence and algebraic roots. *Combinatorics, Probability and Computing*, pages 1–9, December 2024.

[15] P. E. Haxell. A note on vertex list colouring. *Combinatorics, Probability and Computing*, 10(4):345–347, July 2001.

[16] Wilfrid Hodges. *Model Theory*. Cambridge University Press, mar 1993.

[17] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, nov 2006.

[18] László Lovász. *Large networks and graph limits*. American Mathematical Society, 2012.

[19] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity*. Springer Berlin Heidelberg, 2012.

[20] Jaroslav Nešetřil and Patrice Ossona de Mendez. Cluster analysis of local convergent sequences of structures. *Random Structures & Algorithms*, 51(4):674–728, apr 2017.

[21] Jaroslav Nešetřil and Patrice Ossona de Mendez. Existence of modeling limits for sequences of sparse structures. *The Journal of Symbolic Logic*, 84(02):452–472, mar 2019.

[22] Jaroslav Nešetřil and Patrice Ossona de Mendez. A unified approach to structural limits and limits of graphs with bounded tree-depth. *Memoirs of the American Mathematical Society*, 263(1272), jan 2020.

[23] Katrin Tent and Martin Ziegler. *A Course in Model Theory*. Cambridge University Press, mar 2012.

[24] Szymon Toruńczyk. Flip-width: Cops and robber on dense graphs. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 663–700. IEEE, November 2023.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Quickly excluding an apex forest

(Extended abstract)

Jędrzej Hodor[*] Hoang La[†] Piotr Micek[‡] Clément Rambaud[§]

Abstract

We give a short proof that for every apex-forest X on at least two vertices, graphs excluding X as a minor have layered pathwidth at most $2|V(X)| - 3$. This improves upon a result by Dujmović, Eppstein, Joret, Morin, and Wood (SIDMA, 2020). Our main tool is a structural result about graphs excluding a forest as a rooted minor, which is of independent interest. We develop similar tools for treedepth and treewidth. We discuss implications for Erdős-Pósa properties of rooted models of minors in graphs.

1 Introduction

Within the seminal *Graph minors* series, spanning from 1983 to 2010, Robertson and Seymour described the structure of graphs excluding a graph as a minor. One of many key insights of this series is the interplay between forbidding graphs as minors and treewidth or pathwidth. Indeed, excluding a planar graph as a minor is equivalent to having bounded treewidth, which follows from the Grid Minor Theorem [12]. Similarly, excluding a forest

[*]Theoretical Computer Science Department, Faculty of Mathematics and Computer Science and Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland. E-mail: `jedrzej.hodor@gmail.com`. Supported by a Polish Ministry of Education and Science grant (Perły Nauki; PN/01/0265/2022).

[†]LISN, Université Paris-Saclay, CNRS, Gif-sur-Yvette, France. E-mail: `hoang.la.research@gmail.com`.

[‡]Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland. E-mail: `piotr.micek@uj.edu.pl`. Supported by the National Science Center of Poland under grant UMO-2023/05/Y/ST6/00079 within the Weave-UNISONO program.

[§]Université Côte d'Azur, CNRS, Inria, I3S, Sophia-Antipolis, France. E-mail: `clement.rambaud@inria.fr`. Supported by the French Agence Nationale de la Recherche under contract Digraphs ANR-19-CE48-0013-01.

as a minor is equivalent to having bounded pathwidth, which was proved in the first paper of the series [11]. Another relevant statement following this pattern is that excluding a path as a minor is equivalent to having bounded treedepth; see e.g. [10, Chapter 6].

In this paper, we study analogous statements for excluding apex-type graphs as minors. Recall that a graph is an *apex graph* if it can be made planar by the removal of at most one vertex, and a graph is an *apex-forest* if it can be made acyclic by the removal of at most one vertex. It turns out that forbidding apex-type graphs as minors interplays with the layered versions of treewidth, pathwidth, and treedepth. Dujmović, Morin, and Wood [6] proved that a minor-closed class of graphs excludes an apex graph if and only if it has bounded layered treewidth. Similarly, Dujmović, Eppstein, Joret, Morin, and Wood in [4], proved that a minor-closed class of graphs excludes an apex-forest if and only if it has bounded layered pathwidth. Our first contribution is a short and simple proof of the latter statement with an explicit and much better bound on layered pathwidth. In what follows, for a graph G, we denote by $\operatorname{tw}(G)$, $\operatorname{pw}(G)$, $\operatorname{td}(G)$, and $\operatorname{lpw}(G)$ the treewidth, pathwidth, treedepth, and layered pathwidth of G respectively.

Before stating the main theorem, let us recall the definition of layered pathwidth. The width of a path decomposition of a graph with respect to a given layering[1] is the maximum size of the intersection of a bag in the path decomposition and a layer. The *layered pathwidth* of a graph is the minimum possible width of a path decomposition of the graph with respect to a layering of the graph.

Theorem 1. *For every apex-forest X with at least two vertices, and for every graph G, if G is X-minor-free, then* $\operatorname{lpw}(G) \leq 2|V(X)| - 3$.

A graph is a *fan* or (an *apex-path*) if it becomes a path by the removal of at most one vertex. We introduce the concept of layered treedepth mimicking other layered parameters. Recall that treedepth of a graph is the minimum depth of an elimination tree of the graph. The depth of an elimination tree of a graph with respect to a given layering is the maximum size of the intersection of a root-to-leaf path in the elimination tree and a layer. The *layered treedepth* of a graph is the minimum possible depth of an elimination tree of the graph with respect to a layering of the graph.

It is immediate that fans may have arbitrarily large layered treedepth. Conversely, we prove that excluding a fan as a minor implies having bounded layered treedepth.

Theorem 2. *For every fan X with at least three vertices, and for every graph G, if G is X-minor-free, then* $\operatorname{ltd}(G) \leq \binom{|V(X)|-1}{2}$.

The next two statements will follow immediately from the definitions of layered treedepth and pathwidth, Theorem 2 and Theorem 1, respectively. Recall that the *diameter* of a graph G, denoted by $\operatorname{diam}(G)$, is the maximal distance between two vertices in G taken over all pairs of vertices in G.

[1]A *layering* of a graph is a partition of vertices into a sequence of sets (called *layers)* such that for every edge, its both endpoints lie either in one of the sets in subsequent sets.

Corollary 3. *For every fan X with at least two vertices, and for every connected graph G, if G is X-minor-free, then* $\mathrm{td}(G) \leq \binom{|V(X)|-1}{2}(\mathrm{diam}(G)+1)$.

Corollary 4. *For every apex-forest X with at least two vertices, and for every connected graph G, if G is X-minor-free, then* $\mathrm{pw}(G) \leq (2|V(X)|-3)(\mathrm{diam}(G)+1)-1$.

Note that Corollary 3 and Corollary 4 are both optimal in the following sense. There are fans of diameter 2 and unbounded treedepth and there are apex-forests of diameter 2 and unbounded pathwidth. We also give a construction showing that the upper bound in Corollary 4 is tight up to a multiplicative constant.

A natural strengthening of Theorem 1 is the following (false) product structure statement: for every apex-forest X, there is a constant c_X such that for every X-minor-free graph G, we have $G \subseteq H \boxtimes P^2$ for some graph H with $\mathrm{pw}(H) \leq c_X$ and some path P. The statement is false even for $X = K_3$, as Bose, Dujmović, Javarsineh, Morin, and Wood [2] proved that trees do not admit such a product structure. Since K_3 is a fan, the construction in [2] also shows that the analogous strengthening of Theorem 2 does not hold.

2 Graph parameters focused on a given set

The key tools that we use in the proofs of our main results are variants of graph decompositions (and parameters) focused on a given set of vertices. We believe that the parameters and their properties are of independent interest. In particular, we establish connections between the new parameters and rooted models in graphs. We start a discussion with the new version of treedepth, and afterward, we discuss the new versions of pathwidth and treewidth.

Let G be a graph and let S be a fixed subset of vertices of G. We say that a model of a graph H in G is *S-rooted* if every branch set of the model intersects S.

An *elimination forest* of (G, S) is an elimination forest F of H, an induced subgraph of G such that S is contained in $V(H)$ and for every component C of $G - V(H)$, there is a root-to-leaf path in F containing all the neighbors of $V(C)$ in G. The *treedepth* of (G, S), denoted by $\mathrm{td}(G, S)$, is the minimum vertex-height of an elimination forest of (G, S). Recall that if a graph G has no model of P_ℓ, then $\mathrm{td}(G) < \ell$. We prove an analogous result within the setting of S-rooted models of paths.

Theorem 5. *For every positive integer ℓ, for every graph G, and for every $S \subseteq V(G)$, if G has no S-rooted model of P_ℓ, then* $\mathrm{td}(G, S) \leq \binom{\ell}{2}$.

Theorem 5 is the main ingredient of the proof of Theorem 2. Actually, the intuition standing behind this is very simple. For a vertex u in a graph G, we set $S = N(u)$. Now, if $G - u$ has a S-rooted model of a path P_ℓ, then G has a model of P_ℓ with a universal vertex added, and so G has a model of every fan on $\ell + 1$ vertices.

---

[2]The *strong product* $G_1 \boxtimes G_2$ of two graphs G_1 and G_2 is the graph with vertex set $V(G_1 \boxtimes G_2) := V(G_1) \times V(G_2)$ and that contains the edge with endpoints (v, x) and (w, y) if and only if $vw \in E(G_1)$ and $x = y$; or $v = w$ and $xy \in E(G_2)$; or $vw \in E(G_1)$ and $xy \in E(G_2)$.

Let G be a graph and let $S \subseteq V(G)$. A *tree decomposition* (resp. *path decomposition*) of (G,S) is a tree decomposition (resp. path decomposition) $\mathcal{B}$ of H, an induced subgraph of G such that S is contained in $V(H)$, and for every component C of $G - V(H)$, there exists a bag of $\mathcal{B}$ containing all the neighbors of $V(C)$ in G. The *treewidth* (resp. *pathwidth*) of (G,S), denoted by $\mathrm{tw}(G,S)$ (resp. $\mathrm{pw}(G,S)$), is the minimum width of a tree decomposition (resp. path decomposition) of (G,S).

Bienstock, Robertson, Seymour, and Thomas [1] first proved that if a graph G has no model of a forest F, then $\mathrm{pw}(G) \leq |V(F)| - 2$. Note that the first bound on $\mathrm{pw}(G)$ in terms of $|V(F)|$ was proved by Robertson and Seymour in [11]. On the other hand, Diestel gave a beautiful and short proof of the above in [3]. We prove an analogous result within the setting of S-rooted models of forests.

Theorem 6. *For every forest F with at least one vertex, for every graph G, and for every $S \subseteq V(G)$, if G has no S-rooted model of F, then* $\mathrm{pw}(G,S) \leq 2|V(F)| - 2$.

Again, Theorem 6 is a key ingredient of the proof of Theorem 1. The proof of Theorem 6 follows closely the argument by Diestel mentioned before.

Finally, we discuss briefly $\mathrm{tw}(G,S)$. Note that this notion already appeared in [8], and similar notions appeared in [13] and [7]. Let G be a graph and let $S \subseteq V(G)$. We say that a model of a plane graph H in G is *S-outer-rooted* if every branch set corresponding to a vertex in the outer face of H intersects S. The Grid Minor Theorem can be generalized to the setting of S-outer-rooted models as follows.

Theorem 7. *For every plane graph H, there exists a positive integer c_H such that, for every graph G, and for every $S \subseteq V(G)$, if G has no S-outer-rooted model of H, then* $\mathrm{tw}(G,S) \leq c_H$.

The hard part of the proof was in fact done by Marx, Seymour, and Wollan [9]. They studied tangles of graphs focused on a given subset of vertices. We just show that their variant of tangles is functionally equivalent to the discussed variant of treewidth. See also the full version of the paper where we explicitly give the constant c_H.

3 Erdős-Pósa property

Finally, we discuss the applications of our techniques to Erdős-Pósa properties for rooted models. A classical result by Robertson and Seymour states that families of connected subgraphs in graphs of bounded treewidth admit the Erdős-Pósa property. We show the analog version for treewidth focused on a prescribed set of vertices, which turns out to be a useful tool.

Lemma 8. *Let G be a graph, let $S \subseteq V(G)$, let $\mathcal{W} = (T, (W_x \mid x \in V(T)))$ be a tree decomposition of (G,S), and let $\mathcal{F}$ be a family of connected subgraphs of G each of them intersecting S. For every positive integer k, either*

1. *there are k pairwise vertex-disjoint subgraphs in $\mathcal{F}$ or*

2. *there is a set* $Z \subseteq V(G)$ *that is the union of at most* $k-1$ *bags of* $\mathcal{W}$ *such that* $V(F) \cap Z \neq \emptyset$ *for every* $F \in \mathcal{F}$.

Lemma 8 with Theorem 7 yield that outer-rooted models of a fixed connected plane graph admit the Erdős-Pósa property.

Corollary 9. *For every connected plane graph* H*, there exists a function* $d_H \colon \mathbb{N} \to \mathbb{N}$ *such that for every graph* G*, for every* $S \subseteq V(G)$*, and for every positive integer* k*, either*

1. G *has* k *vertex-disjoint* S*-outer-rooted models of* H *or*

2. *there exists a set* $Z \subseteq G$ *such that* $|Z| \leq d_H(k-1)$ *and* $G - Z$ *has no* S*-outer-rooted model of* H.

Recently, Dujmović, Joret, Micek, and Morin [5] showed that for every tree T, for every graph G, for every positive integer k, either G has k disjoint models of T, or there is a set Z of at most $|V(T)|(k-1)$ vertices such that $G - Z$ is T-minor-free. Theorem 6 and Lemma 8 imply the following Erdős-Pósa property for rooted models of trees.

Corollary 10. *For every tree* T*, for every graph* G*, for every* $S \subseteq V(G)$*, and for every positive integer* k*, either*

1. G *has* k *vertex-disjoint* S*-rooted models of* T *or*

2. *there exists a set* $Z \subseteq G$ *such that* $|Z| \leq (2k|V(T)|-1)(k-1)$ *and* $G - Z$ *has no* S*-rooted model of* T.

4 Open problems

Some of the bounds that we provided are not tight, we summarize potential improvements below.

Problem A. Within Theorem 5, we show that for every positive integer ℓ, for every graph G, and for every $S \subseteq V(G)$, if G has no S-rooted model of P_ℓ, then $\mathrm{td}(G,S) \leq \binom{\ell}{2}$. Is there a better bound? Perhaps linear in ℓ?

Problem B. Within Theorem 6 we show that for every forest F with at least one vertex, for every graph G, for every $S \subseteq V(G)$, if G has no S-rooted model of F, then $\mathrm{pw}(G,S) \leq 2|V(F)| - 2$. Is there a better bound? Perhaps $|V(F)| - 2$?

Problem C. Within Corollary 10 we show the Erdős-Pósa property for S-rooted trees with a bound $(2k|V(T)|-1)(k-1) = \mathcal{O}(k^2)|V(T)|$. Perhaps $\mathcal{O}(k)|V(T)|$?

References

[1] Dan Bienstock, Neil Robertson, Paul Seymour, and Robin Thomas. Quickly excluding a forest. *Journal of Combinatorial Theory, Series B*, 52(2):274–283, 1991.

[2] Prosenjit Bose, Vida Dujmović, Mehrnoosh Javarsineh, Pat Morin, and David R. Wood. Separating layered treewidth and row treewidth. *Discrete Mathematics & Theoretical Computer Science*, 24(1), 2022. arXiv:2105.01230v3.

[3] Reinhard Diestel. Graph Minors I: A Short Proof of the Path-width Theorem. *Combinatorics, Probability and Computing*, 4(1):27–30, 1995.

[4] Vida Dujmović, David Eppstein, Gwenaël Joret, Pat Morin, and David R. Wood. Minor-closed graph classes with bounded layered pathwidth. *SIAM Journal on Discrete Mathematics*, 34(3):1693–1709, 2020. arXiv:1810.08314.

[5] Vida Dujmović, Gwenaël Joret, Piotr Micek, and Pat Morin. Tight bound for the Erdős-Pósa property of tree minors. *arXiv preprint*, 2024. arXiv:2403.06370.

[6] Vida Dujmović, Pat Morin, and David R. Wood. Layered separators in minor-closed graph classes with applications. *Journal of Combinatorial Theory, Series B*, 127:111–147, 2017. arXiv:1306.1595.

[7] Eduard Eiben, Robert Ganian, Thekla Hamm, and O joung Kwon. Measuring what matters: A hybrid approach to dynamic programming with treewidth. *Journal of Computer and System Sciences*, 121:57–75, 2021. arXiv:1908.10132.

[8] Bart M.P. Jansen and Céline M.F. Swennenhuis. Steiner Tree Parameterized by Multiway Cut and Even Less. In Timothy Chan, Johannes Fischer, John Iacono, and Grzegorz Herman, editors, *32nd Annual European Symposium on Algorithms (ESA 2024)*, volume 308 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 76:1–76:16, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. LIPIcs.ESA.2024.76.

[9] Daniel Marx, Paul Seymour, and Paul Wollan. Rooted grid minors. *Journal of Combinatorial Theory, Series B*, 122:428–437, 2017. arXiv:1307.8138.

[10] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012.

[11] Neil Robertson and Paul Seymour. Graph minors. I. Excluding a forest. *Journal of Combinatorial Theory, Series B*, 35(1):39–61, 1983.

[12] Neil Robertson and Paul Seymour. Graph minors. V. Excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, 1986.

[13] Ignasi Sau, Giannos Stamoulis, and Dimitrios M. Thilikos. *Parameterizing the quantification of CMSO: model checking on minor-closed graph classes*, pages 3728–3742. SIAM, 2025. arXiv:2406.18465.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Finding large k-colorable induced subgraphs in (bull, chair)-free and (bull,E)-free graphs

(Extended abstract)

Nadzieja Hodur* Monika Pilśniak† Magdalena Prorok‡ Paweł Rzążewski§

Abstract

We study the Max Partial k-Coloring problem, where we are given a vertex-weighted graph, and we ask for a maximum-weight induced subgraph that admits a proper k-coloring. For $k = 1$ this problem coincides with Maximum Weight Independent Set, and for $k = 2$ the problem is equivalent (by complementation) to Minimum Odd Cycle Transversal. Furthermore, it generalizes k-Coloring and List-k-Coloring.

We show that Max Partial k-Coloring on n-vertex instances with clique number ω can be solved in time

- $n^{\mathcal{O}(k\omega)}$ if the input graph excludes the bull and the chair as an induced subgraph,
- $n^{\mathcal{O}(k\omega \log n)}$ if the input graph excludes the bull and E as an induced subgraph.

This implies that k-Coloring can be solved in polynomial time in the former class, and in quasipolynomial-time in the latter one.

1 Introduction

Graph coloring is undoubtedly one of the most studied notions in graph theory, both from the structural and from the algorithmic point of view. For a positive integer k, in the k-Coloring problem we are given a graph G, and we ask whether G admits a proper k-coloring of G, i.e., an assignment of labels from $\{1, \ldots, k\}$ to the vertices of G so that adjacent vertices receive distinct labels. It is well-known that for each $k \geq 3$, the k-Coloring problem is NP-hard [32]. However, such a hardness result is typically just a start of further research questions: What makes the problem hard? Which instances are actually hard? This motivates the study of the complexity of the problem for restricted graph classes, with the hope of understanding the boundary between tractable and intractable cases.

*AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland. nhodur@agh.edu.pl.
†AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland. pilsniak@agh.edu.pl.
‡AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland. prorok@agh.edu.pl.
§Warsaw University of Technology & University of Warsaw, Poland. pawel.rzazewski@pw.edu.pl.

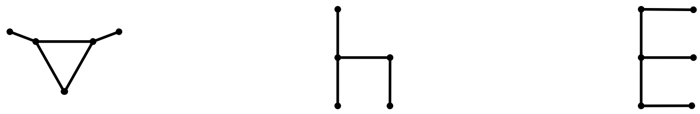


Figure 1: Considered forbidden subgraphs: bull (left), chair (middle), and E(right).

A natural way of obtaining such graph classes is by forbidding certain substructures. For graphs G, F, we say that G is *F-free* if F is not an induced subgraph of G. In other words, we cannot obtain F from G by deleting vertices. For a family $\mathcal{F}$ of graphs, we say that G is *$\mathcal{F}$-free* if it is F-free for every $F \in \mathcal{F}$. If $\mathcal{F} = \{F_1, \dots, F_k\}$, we usually write $(F_1, \dots, F_k)$-free instead of $\{F_1, \dots, F_k\}$-free.

Note that for each $\mathcal{F}$, the family of $\mathcal{F}$-free graphs is *hereditary*, i.e., closed under vertex deletion. Conversely, every hereditary family can be equivalently defined as $\mathcal{F}$-free graphs, for some (possibly infinite) $\mathcal{F}$.

The study of k-Coloring in hereditary graph classes is an active and fruitful area of research. It is known that for the case of F-free graphs, i.e. if we forbid just one induced subgraph, the problem remains NP-hard for every $k \geq 3$ unless F is a linear forest (i.e. a forest of paths) [24, 15, 27]. Let us focus on the case where F is connected, i.e. it is a path on t vertices, denoted by P_t. If $t \leq 5$, then k-Coloring is polynomial-time solvable for every k [22]. Furthermore, we know that if $t = 6$, then k-Coloring is polynomial-time solvable for $k \leq 4$ [30, 11, 12] and NP-hard otherwise [25]. It is also known that for P_7-free graphs, k-Coloring is polynomial-time solvable for $k = 3$ [3] and NP-hard for $k \geq 4$ [25]. Finally, for P_8-free graphs, k-Coloring is NP-hard for all $k \geq 4$ [6].

This leaves open the complexity of 3-Coloring in P_t-free graphs for $t \geq 8$, which is one of the notorious open problems in algorithmic graph theory. Interestingly, it was shown that 3-Coloring in P_t-free graphs can be solved in *quasipolynomial time* for any fixed t [29]. This is a strong indication that none of the remaining cases is NP-hard.

Let us remark that almost all mentioned algorithmic results work for the more general List-k-Coloring problem, where each vertex v of the input graph G is equipped with a *list* $L(v) \subseteq \{1, \dots, k\}$, and we ask for a proper coloring assigning to each vertex a color from its list. The only exception is the case of $k = 4$ and $t = 6$. Indeed, List-4-Coloring is NP-hard for P_6-free graphs [20].

There are also some results concerning (List-)k-Coloring of $\mathcal{F}$-free graphs, when $\mathcal{F}$ contains more graphs [4, 10, 9, 26, 18, 14, 5]; see also survey papers by Schiermeyer [31] and Golovach, Johnson, Paulusma, and Song [19].

Our main motivation was the recent paper of Hodur, Pilśniak, Prorok, and Schiermeyer [23]. Among other results, they proved that 3-Coloring can be solved in polynomial time in the class of (bull, E)-free, where E is a subdivided star $S_{1,2,2}$ and bull is a K_3 with additional pendant edges at two vertices of the triangle (see Fig. 1).[1] The goal of this paper is to generalize their result to (List-)k-Coloring, for any $k \geq 3$.

In fact, we solve an even further generalization of k-Coloring, defined as follows.

[1]Actually, they showed a stronger result, as they characterized all minimal (bull, E)-free graphs that *not* 3-colorable.

Max Partial k-Coloring

Input: a graph G, a revenue function $\mathsf{rev} : V(G) \times \{1, \ldots, k\} \to \mathbb{Q}_{\geq 0}$

Task: a set X and a proper k-coloring c of $G[X]$, such that $\sum_{v \in X} \mathsf{rev}(v, c(v))$ is maximum possible

We can think of the value of $\mathsf{rev}(v, c)$ as the prize we get for coloring a vertex with color c. We aim to find a partial coloring that maximizes the total prize.

Clearly, this formalism captures k-Coloring: it is sufficient to set $\mathsf{rev}(v, c) = 1$ for every v and c, and check whether the maximum revenue is equal to the number of vertices. If we additionally want to express the lists $L : V(G) \to \{1, \ldots, k\}$, i.e., capture List-$k$-Coloring, we can set $\mathsf{rev}(v, c) = 0$ if $c \notin L(v)$ and $\mathsf{rev}(v, c) = 1$ otherwise, and again ask if there is a solution with total revenue $|V(G)|$. However, Max Partial k-Coloring generalizes more problems. If $k = 1$, then the problem is equivalent to Max Weight Independent Set problem. If $k = 2$, then we ask for a maximum-weight induced bipartite subgraph, which is by complementation equivalent to Odd Cycle Transversal. Both these problems received considerable attention from the algorithmic graph theory community [17, 16, 29, 28, 21, 1, 13, 8, 2].

We consider classes of (bull, chair)-free and (bull, E)-free graphs; here chair is a subdivided star $S_{1,1,2}$. Note that chair is an induced subgraph of E and thus (bull, chair)-free graphs form a proper subclass of (bull, E)-free graphs. The following two theorems are the main results of our paper.

Theorem 1. *For every $k \geq 1$, Max Partial k-Coloring on (bull, chair)-free instances with n vertices and clique number ω can be solved in time $n^{\mathcal{O}(k\omega)}$.*

Theorem 2. *For every $k \geq 1$, Max Partial k-Coloring on (bull, E)-free instances with n vertices and clique number ω can be solved in time $n^{\mathcal{O}(k\omega \log n)}$.*

While the running time in Theorem 2 is not polynomial, but quasipolynomial in n, it still gives a strong evidence that the problem is not NP-hard. Indeed, in such a case all problems in NP can be solved in quasipolynomial time, which is unlikely according to our current understanding of complexity theory.

Interestingly, the algorithms in Theorems 1 and 2 are exactly the same; the only difference is the complexity analysis. We start with a careful analysis of the structure of (bull, E)-free graphs. We observe that after exhaustively guessing a constant number of vertices and their color, we can decompose the input graph into parts that (1) are "simpler" and (2) the connections between the parts are "well-structured." The first property allows us to call the algorithm recursively for each part, in order to obtain their corresponding partial solutions. Then the second property is used to combine these partial solutions into the solution of the input instance. By "simpler" we typically mean that the clique number of a part is smaller than the clique number of the graph itself, but in one case (for (bull, E)-free graphs) "simpler" means just "multiplicatively smaller." This explains the running time in Theorems 1 and 2. Let us remark that the idea of using the clique number to bound the complexity of an algorithm already appears in the literature [8, 7].

Notice that if we are only interested in solving List-k-Coloring, we can safely assume that the clique number of each instance is at most k: otherwise we can safely reject. Thus we immediately obtain the following corollaries.

Corollary 3. *For every $k \geq 3$,* LIST k-COLORING *on n-vertex (bull, chair)-free graphs can be solved in time $n^{\mathcal{O}(k^2)}$.*

Corollary 4. *For every $k \geq 3$,* LIST k-COLORING *on n-vertex (bull, E)-free graphs can be solved in time $n^{\mathcal{O}(k^2 \log n)}$.*

Furthermore, using the win-win approach of Chudnovsky et al. [8], we can show that for every k, the MAX PARTIAL k-COLORING problem can be solved in *subexponential time* in (bull, E)-free graphs, with no restrictions on the clique number.

Corollary 5. *For every $k \geq 1$,* MAX PARTIAL k-COLORING *on (bull, E)-free instances with n vertices can be solved in time $2^{\mathcal{O}(k \cdot \sqrt{n} \log^{3/2} n)}$.*

We believe that the problem is actually polynomial-time solvable, and we state this as a conjecture.

Conjecture 6. *For every $k \geq 1$,* MAX PARTIAL k-COLORING *on (bull, E)-free graphs can be solved in polynomial time.*

2 Sketch of the proof.

Here we sketch the proof of our main results; for simplicity, we focus on the case of (bull, chair)-free graphs. Let (G, rev) be an input instance of MAX PARTIAL k-COLORING, where G is an (bull, chair)-free graph on n vertices and clique number of G is ω. The proof proceeds by induction on ω.

We start by greedily building an inclusion-wise maximal induced path P. We consider two cases, depending on the number of vertices of P.

Case A: P has at most 6 vertices. Let us denote the consecutive vertices of P by $x_1, \ldots, x_p$. For $j \in [p]$, let A_j be the set of vertices that are adjacent to x_j, but not to $x_{j'}$ for any $j' < j$. Assuming (X, c) is the solution sought, for each color $i \in [k]$ and each $j \in [p]$, we can define S_j^i as an inclusion-wise minimal subset of $X \cap A_j \cap c^{-1}(i)$ with the property that $N(X \cap A_j \cap c^{-1}(i)) \cap \bigcup_{j'>j} A_{j'} = N(S_j^i) \cap \bigcup_{j'>j} A_{j'}$. Using the fact that G is (bull, chair)-free, we can prove that $|S_j^i| \leq 2$ for each $i \in [k]$. Thus, the number of possible choices of family of sets $\{S_j^i, i \in [k], j \in [p]\}$ is bounded from above by n^{2pk}. We proceed as follows.

We exhaustively guess the set $X \cap P$, its coloring and the family of sets S_j^i for $i \in [k]$ and $j \in [p]$. It results with at most $(k+1)^6 n^{2pk}$ branches. We reject all branches that produce an immediate contradiction, in particular, for each $i \in [k]$, the set $\bigcup_{j \in [p]} S_j^i$ should be independent, should not have edges to vertices from $P \cap X$ colored i, and every $v \in \bigcup_{j \in [p]} S_j^i$ should satisfy $\mathsf{rev}(v, i) > 0$. Consider one such guessed branch.

We adjust the revenue function. For any vertex $v \in X \cap P$ we forbid color $c(v)$ to any neighbor of v. For each vertex $v \in S_j^i$ we forbid color i to any neighbor of v and any color but i to v. Finally, for any vertex $v \in \bigcup_{j'>j} A_{j'}$, if v is not adjacent to any vertex in S_j^i, we forbid i to any neighbor of v in A_j. Now, the solutions obtained on subinstances cannot be incompatible.

We call the algorithm independently for subinstances $(G[A_i], \mathsf{rev})$ for $i \in [p]$ and (C, rev) for any connected component C of $G \setminus N[P]$. We observe that each of the subinstances has clique number at most $\omega - 1$, so we can proceed inductively.

Case B: P has at least 7 vertices. Using the fact that G is a (bull, chair)-free, we can extend P to a "fat path" or a "fat cycle" R – the graph obtained from a path or a cycle by substituting each vertex with an arbitrary graph, and keeping all edges between graphs corresponding to consecutive vertices. Furthermore, every vertex of G that has a neighbor in R (call this set D), is actually adjacent to all R. Finally, for every component C of $T = G - (R \cup D)$ there exists a vertex of D that is adjacent to every vertex of C. As before, the minimal subsets S^i of $X \cap D$, satisfying $N(S^i) \cap T = N(D \cap X \cap c^{-1}(i)) \cap T$, are small: $|S^i| \leq 2$ for any $i \in [k]$. We proceed as follows.

We exhaustively guess the set A of colors that can appear on the vertices of $X \cap R$ and the family of minimal subsets S^i for each color $i \in [k]$ (rejecting immediately contradictory branches). This results in at most $2^k n^{2k}$ branches.

Having guessed the set A of colors and sets S^i, we adjust the revenue function: we forbid colors from A to the vertices of D and any color except i to vertices of S^i, for any $i \in [k]$. Finally, for any vertex $v \in T$, if v is not adjacent to any vertex in S^i_j, we forbid i to any neighbor of v in D.

Finally, we call the algorithm for the simpler subinstances: $(G[R], \mathsf{rev}), (G[D], \mathsf{rev})$ and (C, rev) for any connected component C of T. The instances $(G[D], \mathsf{rev})$ and (C, rev) have clique number at most $\omega - 1$, so we can apply the inductive assumption. Now let us focus on $G[R]$ and recall that it can be decomposed into a (possibly cyclic) sequence of graphs, each of clique number at most $\omega - 1$, where any two consecutive graphs are fully adjacent to each other. These observations allow us to solve $(G[R], \mathsf{rev})$, combining a simple dynamic programming with inductive calls on graphs with smaller clique number.

References

[1] T. Abrishami, M. Chudnovsky, M. Pilipczuk, P. Rzążewski, and P. D. Seymour. Induced subgraphs of bounded treewidth and the container method. *SIAM J. Comput.*, 53(3):624–647, 2024.

[2] A. Agrawal, P. T. Lima, D. Lokshtanov, P. Rzążewski, S. Saurabh, and R. Sharma. Odd cycle transversal on P_5-free graphs in polynomial time. *ACM Trans. Algorithms*, 21(2), Jan. 2025.

[3] F. Bonomo, M. Chudnovsky, P. Maceli, O. Schaudt, M. Stein, and M. Zhong. Three-coloring and list three-coloring of graphs without induced paths on seven vertices. *Comb.*, 38(4):779–801, 2018.

[4] F. Bonomo-Braberman, M. Chudnovsky, J. Goedgebeur, P. Maceli, O. Schaudt, M. Stein, and M. Zhong. Better 3-coloring algorithms: Excluding a triangle and a seven vertex path. *Theor. Comput. Sci.*, 850:98–115, 2021.

[5] C. Brause, I. Schiermeyer, P. Holub, Z. Ryjáček, P. Vrána, and R. Krivos-Bellus. 4-colorability of P_6-free graphs. *Electron. Notes Discret. Math.*, 49:37–42, 2015.

[6] H. Broersma, P. A. Golovach, D. Paulusma, and J. Song. Updating the complexity status of coloring graphs without a fixed induced linear forest. *Theor. Comput. Sci.*, 414(1):9–19, 2012.

[7] M. Chudnovsky, J. Czyzewska, K. Kluk, M. Pilipczuk, and P. Rzazewski. Sparse induced subgraphs in P_7-free graphs of bounded clique number. *CoRR*, abs/2412.14836, 2024.

[8] M. Chudnovsky, J. King, M. Pilipczuk, P. Rzążewski, and S. Spirkl. Finding large H-colorable subgraphs in hereditary graph classes. *SIAM J. Discret. Math.*, 35(4):2357–2386, 2021.

[9] M. Chudnovsky, P. Maceli, J. Stacho, and M. Zhong. 4-coloring P_6-free graphs with no induced 5-cycles. *J. Graph Theory*, 84(3):262–285, 2017.

[10] M. Chudnovsky, S. Spirkl, and M. Zhong. List 3-coloring P_t-free graphs with no induced 1-subdivision of $K_{1,s}$. *Discret. Math.*, 343(11):112086, 2020.

[11] M. Chudnovsky, S. Spirkl, and M. Zhong. Four-coloring P_6-free graphs. I. Extending an excellent precoloring. *SIAM J. Comput.*, 53(1):111–145, 2024.

[12] M. Chudnovsky, S. Spirkl, and M. Zhong. Four-coloring P_6-free graphs. II. Finding an excellent precoloring. *SIAM J. Comput.*, 53(1):146–187, 2024.

[13] K. K. Dabrowski, C. Feghali, M. Johnson, G. Paesani, D. Paulusma, and P. Rzążewski. On cycle transversals and their connected variants in the absence of a small linear forest. *Algorithmica*, 82(10):2841–2866, 2020.

[14] K. K. Dabrowski and D. Paulusma. On colouring $(2P_2, H)$-free and (P_5, H)-free graphs. *Inf. Process. Lett.*, 134:35–41, 2018.

[15] T. Emden-Weinert, S. Hougardy, and B. Kreuter. Uniquely colourable graphs and the hardness of colouring graphs of large girth. *Comb. Probab. Comput.*, 7(4):375–386, 1998.

[16] P. Gartland and D. Lokshtanov. Independent set on P_k-free graphs in quasi-polynomial time. In S. Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 613–624. IEEE, 2020.

[17] P. Gartland, D. Lokshtanov, T. Masařík, M. Pilipczuk, M. Pilipczuk, and P. Rzążewski. Maximum weight independent set in graphs with no long claws in quasi-polynomial time. In B. Mohar, I. Shinkar, and R. O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 683–691. ACM, 2024.

[18] S. Gaspers, S. Huang, and D. Paulusma. Colouring square-free graphs without long induced paths. *J. Comput. Syst. Sci.*, 106:60–79, 2019.

[19] P. A. Golovach, M. Johnson, D. Paulusma, and J. Song. A survey on the computational complexity of coloring graphs with forbidden subgraphs. *J. Graph Theory*, 84(4):331–363, 2017.

[20] P. A. Golovach, D. Paulusma, and J. Song. Closing complexity gaps for coloring problems on H-free graphs. *Inf. Comput.*, 237:204–214, 2014.

[21] A. Grzesik, T. Klimošová, M. Pilipczuk, and M. Pilipczuk. Polynomial-time algorithm for maximum weight independent set on P_6-free graphs. *ACM Trans. Algorithms*, 18(1):4:1–4:57, 2022.

[22] C. T. Hoàng, M. Kamiński, V. V. Lozin, J. Sawada, and X. Shu. Deciding k-colorability of P_5-free graphs in polynomial time. *Algorithmica*, 57(1):74–81, 2010.

[23] N. Hodur, M. Pilśniak, M. Prorok, and I. Schiermeyer. On 3-colourability of $(bull, H)$-free graphs. Preprint available at `https://arxiv.org/abs/2404.12515`, 2024.

[24] I. Holyer. The NP-completeness of edge-coloring. *SIAM J. Comput.*, 10(4):718–720, 1981.

[25] S. Huang. Improved complexity results on k-coloring P_t-free graphs. *Eur. J. Comb.*, 51:336–346, 2016.

[26] V. Jelínek, T. Klimošová, T. Masařík, J. Novotná, and A. Pokorná. On 3-coloring of $(2P_4, C_5)$-free graphs. *Algorithmica*, 84(6):1526–1547, 2022.

[27] D. Leven and Z. Galil. NP completeness of finding the chromatic index of regular graphs. *J. Algorithms*, 4(1):35–44, 1983.

[28] D. Lokshtanov, M. Vatshelle, and Y. Villanger. Independent set in P_5-free graphs in polynomial time. In C. Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 570–581. SIAM, 2014.

[29] M. Pilipczuk, M. Pilipczuk, and P. Rzazewski. Quasi-polynomial-time algorithm for independent set in P_t-free graphs via shrinking the space of induced paths. In H. V. Le and V. King, editors, *4th Symposium on Simplicity in Algorithms, SOSA 2021, Virtual Conference, January 11-12, 2021*, pages 204–209. SIAM, 2021.

[30] B. Randerath and I. Schiermeyer. 3-colorability in P for P_6-free graphs. *Discret. Appl. Math.*, 136(2-3):299–313, 2004.

[31] B. Randerath and I. Schiermeyer. Vertex colouring and forbidden subgraphs - A survey. *Graphs Comb.*, 20(1):1–40, 2004.

[32] L. J. Stockmeyer. Planar 3-colorability is polynomial complete. *SIGACT News*, 5(3):19–25, 1973.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Approximate Itai-Zehavi conjecture for random graphs

(Extended abstract)

Lawrence Hollom* Lyuben Lichev† Adva Mond‡ Julien Portier§
Yiting Wang¶

Abstract

We show that the celebrated Itai-Zehavi conjecture holds asymptotically for Erdős-Rényi random graphs $G(n,p)$ when $np = \omega(\log n)$, and for random regular graphs $G(n,d)$ when $d = \omega(\log n)$. Moreover, we confirm the conjecture up to a constant factor for sparser random regular graphs. This answers positively a question of Draganić and Krivelevich.

1 Introduction

Covering and packing problems belong to the interface between combinatorics and discrete geometry, and have received considerable attention from both communities. Given a combinatorial host structure, it is often of interest to find disjoint or in some sense independent substructures equipped with certain additional properties. The following celebrated conjecture of Itai and Zehavi [16] is part of the said framework.

Conjecture 1.1 (Itai and Zehavi [16]). *For every $d \geq 1$, every d-connected graph G and every vertex r in G, G contains d spanning trees $T_1, \dots, T_d$ such that, for every vertex $v \in V(G) \setminus \{r\}$, the paths from v to r in $T_1, \dots, T_d$ are pairwise internally vertex-disjoint.*

Observe the dependency between connectivity and the number of spanning trees in Theorem 1.1 is optimal. Indeed, if G is a d-regular d-connected graph, a simple counting argument shows there are at most d such trees. Although the case $d = 1$ of Conjecture 1.1 is immediate, it was only confirmed for $d \leq 4$ [12, 3, 16, 4].

*Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, CB3 0WA, United Kingdom
†Institute of Statistics and Mathematical Methods in Economics, TU Wien, Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria
‡Department of Mathematics, King's College London, Strand, London, WC2R 2LS, United Kingdom
§Ecole Polytechnique Federale de Lausanne (EPFL), Rte Cantonale, 1015 Lausanne, Switzerland
¶Institute of Science and Technology Austria (ISTA), 3400 Klosterneuburg, Austria

While Theorem 1.1 is widely open for general d, weaker lower bounds have been derived. For convenience of discussion, we call spanning trees satisfying the condition in Theorem 1.1 *independent spanning trees* (abbreviated ISTs). The best known general bound is due to Censor-Hillel, Ghaffari and Kuhn [2], who studied a related notion called connected dominating sets (abbreviated CDS). They proved that, in a d-connected graph G, there are at least $\Omega(d/\log^2 n)$ disjoint CDSs, implying the existence of $\Omega(d/\log^2 n)$ ISTs. Recently, Draganić and Krivelevich [5] proved that (n,d,λ)-graphs with $d/\lambda = \Omega(1)$ contain $\Omega(d/\log d)$ disjoint CDSs, and thus $\Omega(d/\log d)$ ISTs. They also observed that this bound is optimal up to a constant factor, implying that disjoint CDSs are too restrictive to shed light on Theorem 1.1 even for the special case of (n,d,λ)-graphs.

In the same paper, Draganić and Krivelevich [5] asked if one can prove Theorem 1.1 for random graphs or, more modestly, find $\Theta(d)$ ISTs in the random regular graph $G(n,d)$, and $\Theta(np)$ ISTs in the Erdős-Rényi random graph $G(n,p)$ with $np = \omega(\log n)$. The main result of our work answers their question positively.

Theorem 1.2 ([9]). (a) *Fix $p = p(n) \in [0,1]$ with $np = \omega(\log n)$. Then, for every vertex r, the Erdős-Rényi graph $G(n,p)$ contains $(1-o(1))np$ ISTs rooted at r whp*[1].

(b) *Fix $d = d(n) \in [4, n-1]$. Then, for $n-o(n)$ vertices r, the random regular graph $G(n,d)$ contains $\lfloor d/4 \rfloor$ ISTs rooted at r whp.*

We observe that the constant 1/4 in Theorem 1.2 might possibly be improved to 1/3 at the price of a more technical presentation. Since we are not able to come closer to 1 with our approach, we stick to the current cleaner version.

The proofs of parts (a) and (b) in Theorem 1.2 use significantly different ideas. We will give a fairly complete sketch for (a) and a (much) coarser sketch for (b).

2 Proof sketch of Theorem 1.2 part (a) for $G(n,p)$

It suffices to show that, for every $\varepsilon \in (0,1)$, there exists a constant $C = C(\varepsilon)$ such that, for every vertex r, there are at least $(1-\varepsilon)np$ ISTs rooted at r in $G(n, C/n)$ with probability $1 - o(1/n)$.

The proof starts with splitting $G(n,p)$ into two parts, $G_1 \sim G(n,p_1)$ and $G_2 \sim G(n,p_2)$, with p_1 much smaller than p and $p_2 \approx p$. Before any of G_1, G_2 becomes exposed, we fix a root vertex r and find $k = (1-\varepsilon)np$ of its neighbours in G, namely $v_1,\dots,v_k$. By using a breadth-first search (BFS) exploration process in G_1, we consecutively construct disjoint *core sets* $C_1,\dots,C_k$. Roughly speaking, in the ISTs $T_1,\dots,T_k$ produced from our construction, the set C_i will consist of the vertices in T_i which are neither leaves nor parents of leaves. Vertices in C_i will be expected to have a significant number of descendants and, for this reason, will be attached as leaves in each of the trees $T_j \neq T_i$.

We construct the core sets consecutively as follows. For every $i \in [k]$, upon having $C_1,\dots,C_{i-1}$ already built, we start a BFS exploration process in $G_1 \setminus (C_1 \cup \dots \cup C_{i-1} \cup \{r, v_{i+1},\dots,v_k\})$ from v_i and run it until ε/p vertices have been explored. The vertices explored during this process form our new set C_i. In particular, at the time when the BFS exploration away from v_i stops, there exists a set $B_i \subseteq C_i$ such that none of the edges between B_i and $G_1 \setminus (C_1 \cup \dots \cup C_i \cup \{r, v_{i+1},\dots,v_k\})$ has been explored at this moment.

[1] We abbreviate "with high probability" as whp. A sequence of events $\mathcal{E}_n$ holds whp if $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$.

After $C_1, \ldots, C_k$ have been constructed, the final phase of our construction is to attach the vertices remaining outside $(C_i \cup \{r\})_{i=1}^k$ via paths of length at most two to the core sets. To this end, denote by $N_1, \ldots, N_k$ the neighbourhoods of $B_1, \ldots, B_k$ in G_1 outside the core sets: these random sets are not explored right away, as their randomness will be needed later. We will refer to N_i as the *connecting layer* of the tree T_i. For every $i \in [k]$, the idea is to connect vertices outside $\{r\} \cup C_i \cup N_i$ via edges to N_i. In fact, given that vertices in the core set C_i have their neighbourhood in G_1 explored and must still attach to $T_j \neq T_i$, this is the moment when the fresh randomness of G_2 is used. Observe that, thanks to the disjointness of the core sets, the only obstruction to the independence of the produced trees would be if a vertex u attaches to the same parent $v = N_i \cap N_j$ in T_i and T_j. We are able to avoid this by using a classic result for the existence of a matching in balanced bipartite Erdős-Rényi graphs. The said theorem is applied for an auxiliary graph with parts $X = [k]$ and $Y = N_{G_2}(u) \setminus (\{r\} \cup C_1 \cup \ldots \cup C_k)$ where $x \in X$ is adjacent to $y \in Y$ if $y \in N_x$; note that the randomness reserved by not exposing $N_1, \ldots, N_k$ is used to show that edges in this graph actually appear independently of each other.

3 Proof sketch of Theorem 1.2 part (b) for $G(n, d)$

If $d = \omega(\log n)$, the result follows from combining part (a) with a sandwiching theorem of Gao, Isaev and McKay [8]. We now focus on the regime $d = O(\log n)$

The strategy from part (a) where randomness is exposed in two stages does not have an immediate analogue for random regular graphs. We first restrict our considerations to even n; the odd n case is addressed thereafter. To overcome this inconvenience, our first step consists of defining an alternative random graph obtained as a union of d uniformly chosen perfect matchings on the same set of n vertices conditioned on being edge-disjoint. Formally, in [10], we proved the following “one-sided contiguity” result, which may be of independent interest:

Theorem 3.1 (Theorem 1.7 from [10])**.** *Let $d = d(n) \geq 3$ with $d = O(n^{1/10})$. Let μ_d be the uniform measure on the set of all d-regular graphs on n vertices and let ν_d be a measure on the same set, but with measure proportional to the number of 1-factorisations. For any sequence of events $(A_n)_{n=1}^\infty$ with $A_n \subseteq \mathcal{G}_d(n)$ for all $n \geq 1$, if $\lim_{n\to\infty} \nu_d(A_n) = 0$, then $\lim_{n\to\infty} \mu_d(A_n) = 0$.*

Note that, for $d = O(1)$, the converse is also known to hold (see [15]). That proof is based on the small subgraph conditioning method, and does not extend all the way to $d = O(\log n)$. Our proof of Theorem 3.1 uses estimates of McKay [14] for the number of copies of sparse graphs in a dense host graph, a convenient consequence of Strassen's theorem observed by Isaev, McKay, Southwell and Zhukovskii [11] and builds on an approach of Gao [6, 7] for the counting of triangles and perfect matchings in random regular graphs.

Equipped with Theorem 3.1, we show part (a) in the alternative matching model where we have d edge-disjoint union of perfect matchings. We partition the matchings into groups of 4. The first three matchings in group i form a 3-regular graph G_i, and the fourth is denoted by $\mathcal{M}_i$. While G_i is not quite a uniformly random 3-regular graph (it needs to be edge-disjoint from $\bigcup_{j\neq i} G_j$ and $\bigcup_j M_j$), we prove that it is very close to one.

By choosing a random vertex r, we form trees $T_1, \ldots, T_k$ by a BFS exploration of the graphs $G_1, \ldots, G_k$ away from r. These trees are typically almost ISTs, meaning that there are very few vertices v such that the interiors of the paths from v to r in two distinct trees

$T_i \neq T_j$ contain a common vertex: such vertices are called *bad*. We show that for any bad vertex v, typically all issues caused by v are solved by rerouting v and its descendants in T_i (or in T_j), meaning that each descendant u of v exchanges the edge to its parent in T_i (resp. in T_j) with the edges containing it in $\mathcal{M}_i$ (resp. in $\mathcal{M}_j$).

When analysing the rerouting process, we need to compute (among other things) probabilities for two types of events:

(i) There is a vertex w appearing on the path from v to r in each of the trees T_i and T_j.

(ii) A vertex w is matched to a vertex u in the matching $\mathcal{M}_i$, possibly given that $(\bigcup_{j<i} G_j) \cup (\bigcup_{j<i} \mathcal{M}_j)$ and potentially some edges of $\mathcal{M}_i$ are already exposed.

For (i), we first introduce an notion of random overlay of the graphs $G_1, \ldots, G_k$. We call the unlabelled copy of a graph $G \subseteq K_n$ obtained by erasing the labels of $V(G)$ its *skeleton*. Fix $m = O(1)$ and any unlabelled 3-regular graphs $H_1, \ldots, H_m$ on n vertices (in fact, any m graphs of bounded degree). We randomly overlay $H_1, \ldots, H_m$ in the sense that, for each $i \in [m]$, we form a labelled graph $\sigma_i(H_i) \subseteq K_n$ where σ_i is a random bijection between vertices of H_i and K_n. The bijections $\sigma_1, \ldots, \sigma_m$ are selected independently. By a computation of moments, we show that $\sigma_1(H_1) \cup \ldots \cup \sigma_m(H_m)$ does not have multiple edges with probability bounded away from 0 (as a function of m). From here, we deduce that an event depending on a bounded number of graphs among $G_1, \ldots, G_k$ holds whp in the original model if it does so in the random overlay model. The computation in (ii) relies on the already mentioned counting results of McKay [14].

For odd n (and even d), we reduce the problem to the (even) $n-1$ case with additional constraints. To transition between $G(n-1, d)$ and $G(n, d)$, we design an operation *op* which transforms a d-regular graph on $n-1$ vertices containing an induced matching M_d on $d/2$ edges into a d-regular graph on n vertices. More precisely, *op* removes the edges in M_d from G_{n-1} and introduces a new vertex v which connects to all vertices in M_d. Conversely, to go from a d-regular graph G_n on n vertices to a graph on $n-1$ vertices and a matching on $d/2$ edges, one can remove a vertex from G_n whose neighbourhood forms an independent set and add an arbitrary perfect matching on $N(v)$. We show that applying *op* to the random graph $G(n-1, d)$ and a random induced matching M_d on $d/2$ edges in it produces an approximately uniform element of $\mathcal{G}_d(n)$. Finally, by adapting (parts of) the proof of Theorem 1.2 for even n , we construct k ISTs $T_1, \ldots, T_k$ in G_{n-1} which avoid M_d and additionally allow v to be added as a leaf to each of $T_1, \ldots, T_k$ without violating the assumption that paths from v to r are internally disjoint. This finishes the proof in the case of odd n.

4 Concluding remarks

We have established that the Itai-Zehavi conjecture holds asymptotically for Erdős-Rényi graphs random graphs and sparse random regular graphs. For dense random regular graphs, we prove it approximately. Both results leave some natural open questions.

For Erdős-Rényi graphs, our result works in the regime of $p = \omega(\log n/n)$. One may ask if the result still holds in sparser regimes:

Question 4.1. *Does the conclusion of Theorem 1.2 (a) hold for all p above the sharp threshold for k-connectivity in $G(n,p)$ (which is $(\log n + (k-1)\log\log n + \omega(1))/n$)?*

While degrees are less concentrated in this regime, the required number of ISTs is also much smaller. Techniques of Krivelevich and Samotij [13] might be useful for this question.

For random d-regular graphs, the best one can get using our method is $\lfloor d/3 \rfloor$. Recall that we group 3 perfect matchings into a graph G_i and find an IST in each. The 3 cannot be replaced by 2 since we need each G_i to have diameter $O(\log n)$ and so that each tree T_i has height $O(\log n)$. This diameter condition only holds for random d-regular graphs with $d \geq 3$ (a classic result established by Bollobás and Fernandez de la Vega [1]). A natural next step is to find d or at least $d - o(d)$ of ISTs in random d-regular graphs.

Question 4.2. *Can one show the same conclusion of Theorem 1.2 (b) with $\lfloor d/4 \rfloor$ replaced by $d - o(d)$?*

References

[1] B. Bollobás and W. Fernandez de la Vega. The diameter of random regular graphs. *Combinatorica*, 2:125–134, 1982.

[2] K. Censor-Hillel, M. Ghaffari, G. Giakkoupis, B. Haeupler, and F. Kuhn. Tight bounds on vertex connectivity under sampling. *ACM Transactions on Algorithms (TALG)*, 13(2):1–26, 2017.

[3] J. Cheriyan and S. Maheshwari. Finding nonseparating induced cycles and independent spanning trees in 3-connected graphs. *Journal of Algorithms*, 9(4):507–537, 1988.

[4] S. Curran, O. Lee, and X. Yu. Finding four independent trees. *SIAM Journal on Computing*, 35(5):1023–1058, 2006.

[5] N. Draganić and M. Krivelevich. Disjoint connected dominating sets in pseudorandom graphs. *arXiv preprint arXiv:2410.16072*, 2024.

[6] P. Gao. The number of perfect matchings, and the nesting properties, of random regular graphs. *Random Structures and Algorithms*, 62(4):935–955, 2023.

[7] P. Gao. Triangles and subgraph probabilities in random regular graphs. *The Electronic Journal of Combinatorics*, paper P1.2, 2024.

[8] P. Gao, M. Isaev, and B. D. McKay. Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms*, pages 690–701. SIAM, Philadelphia, PA, 2020.

[9] L. Hollom, L. Lichev, A. Mond, J. Portier, and Y. Wang. Approximate Itai-Zehavi conjecture for random graphs. In preparation.

[10] L. Hollom, L. Lichev, A. Mond, J. Portier, and Y. Wang. Monotonicity and decompositions of random regular graphs. In preparation.

[11] M. Isaev, B. D. McKay, A. Southwell, and M. Zhukovskii. Sprinkling with random regular graphs. *Electronic Journal of Probability*, 30:Paper No. 1, 2025.

[12] A. Itai and M. Rodeh. The multi-tree approach to reliability in distributed networks. *Information and Computation*, 79(1):43–59, 1988.

[13] M. Krivelevich and W. Samotij. Optimal packings of Hamilton cycles in sparse random graphs. *SIAM Journal on Discrete Mathematics*, 26(3):964–982, 2012.

[14] B. D. McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria*, 19:15–25, 1985.

[15] N. C. Wormald. Models of random regular graphs. *London Mathematical Society lecture note series*, pages 239–298, 1999.

[16] A. Zehavi and A. Itai. Three tree-paths. *Journal of Graph Theory*, 13(2):175–188, 1989.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

Counting big Ramsey degrees of the homogeneous and universal $\mathbf{K}_4$-free graph

(Extended abstract)

Jan Hubička* Matěj Konečný† Štěpán Vodseďálek‡ Andy Zucker§

Abstract

Big Ramsey degrees of Fraïssé limits of finitely constrained free amalgamation classes in finite binary languages have been recently fully characterised by Balko, Chodounský, Dobrinen, Hubička, Konečný, Vena, and Zucker. A special case of this characterisation is the universal homogeneous $\mathbf{K}_4$-free graph. We give a self-contained and relatively compact presentation of this case and compute the actual big Ramsey degrees of small graphs.

1 Introduction

Given graphs $\mathbf{G}$ and $\mathbf{H}$, we denote by $\mathrm{Emb}(\mathbf{G}, \mathbf{H})$ the set of all embeddings $\mathbf{G} \to \mathbf{H}$. If $\mathbf{H}'$ is another graph and $\ell \leq k < \omega$, we write $\mathbf{H}' \longrightarrow (\mathbf{H})^{\mathbf{G}}_{k,\ell}$ to denote the following statement: For every colouring $\chi \colon \mathrm{Emb}(\mathbf{G}, \mathbf{H}') \to \{1, \dots, k\}$ with k colours, there exists an embedding $f \colon \mathbf{H} \to \mathbf{H}'$ such that the restriction of χ to $\mathrm{Emb}(\mathbf{G}, f(\mathbf{H}))$ takes at most ℓ distinct values.

For a countably infinite graph $\mathbf{H}$ and a finite induced subgraph $\mathbf{G}$ of $\mathbf{H}$, the *big Ramsey degree of* $\mathbf{G}$ *in* $\mathbf{H}$ is the least number $D \in \omega$ (if it exists) such that $\mathbf{H} \longrightarrow (\mathbf{H})^{\mathbf{G}}_{k,D}$ for every $k \in \omega$. We say that $\mathbf{H}$ *has finite big Ramsey degrees* if the big Ramsey degree of every finite subgraph $\mathbf{G}$ of $\mathbf{H}$ exists. Big Ramsey degrees of other kinds of structures (orders, hypergraphs, ...) are defined in a complete analogy, see recent surveys for details [7, 12, 11].

*Department of Applied Mathematics (KAM), Charles University, Malostranské náměstí 25, Praha 1, Czech Republic E-mail: hubicka@kam.mff.cuni.cz. Supported by project 25-15571S of the Czech Science Foundation (GAČR)

†Institute of Algebra, TU Dresden, Dresden, Germany. E-mail: matej.konecny@tu-dresden.de. Supported by a project that has received funding from the European Union (Project POCOCOP, ERC Synergy Grant 101071674). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

‡Charles University, Malostranské náměstí 25, Praha 1, Czech Republic E-mail: stepanvodsedalek@gmail.com. Also supported by project 25-15571S.

§Department of Pure Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada. E-mail: a3zucker@uwaterloo.ca. Supported by NSERC grants RGPIN-2023-03269 and DGECR-2023-00412.

The concept of big Ramsey degrees, isolated by Kechris, Pestov, and Todorcevic [13], originated in the study of colourings of subsets of the order of rationals $(\mathbb{Q}, \leq)$. In 1969, Laver introduced a rather general proof technique to obtain upper bounds on big Ramsey degrees of $(\mathbb{Q}, \leq)$ [19]. In 1979, Devlin determined the precise big Ramsey degrees proving, somewhat suprisingly, that the big Ramsey degree of a chain with n elements in $(\mathbb{Q}, \leq)$ is precisely the *n-th odd tangent number*: the $(2n-1)$-th derivative of $\tan(x)$ evaluated at 0, the sequence A000182 in the On-line Encyclopedia of Integer Sequences (OEIS) [4, 19].

Graph $\mathbf{H}$ is *homogeneous* if every isomorphism between finite induced subgraphs of $\mathbf{H}$ extends to an automorphism of $\mathbf{H}$. The *Rado graph* $\mathbf{R}$ is the (up to isomorphism) unique countable homogeneous graph which is *universal*, that is, every countable graph can be embedded to $\mathbf{R}$. Similarly, for every $k > 2$ there exists an (up to isomorphism) unique countable homogeneous $\mathbf{K}_k$-free graph $\mathbf{R}_k$ such that every countable $\mathbf{K}_k$-free graph can be embedded to $\mathbf{R}_k$. We call $\mathbf{R}_k$ the *countable homogeneous $\mathbf{K}_k$-free graph*. See e.g. [9].

Laver's proof can be adapted to the graph $\mathbf{R}$, and in 2006 this was refined by Laflamme, Sauer, and Vuksanovic [15] to precisely characterise its big Ramsey degrees. Big Ramsey degrees of cliques and anticliques are again the odd tangent numbers, and Larson [16] used a Maple program to compute, for a given n, the sum of big Ramsey degrees of all graphs with n vertices, yielding a sequence A293158 in OEIS.

In 2020, Dobrinen developed new techniques to prove finiteness of big Ramsey degrees of $\mathbf{R}_3$ [5] (see [10] for a simpler proof) and later of all graphs $\mathbf{R}_k$, $k \geq 3$ [6]. Zucker simplified and further generalized Dobrinen's proof to Fraïssé limits of finitely constrained free amalgamation classes in finite binary languages [21] and in 2024, Balko, Chodounský, Dobrinen, Hubička, Konečný, Vena, and Zucker gave a precise characterisation [2]. In this generality, even the statement of the characterization is very technically challenging and definitions of [2] need a careful analysis of every specific case they are applied to. The big Ramsey degrees are determined by a number of special trees called *diaries*. To understand them, the reader needs to internalize approximately 21 definitions up to page 22 of [2]. A short and self-contained description of big Ramsey degrees of $\mathbf{R}_3$ appears in [1]. In this note we give a similar description of diaries of $\mathbf{R}_4$ with the aim to count them.

2 Diaries of $\mathbf{K}_4$-free graphs

We first present the definition and then discuss the intuition behind it. We fix an *alphabet* $\Sigma = \{0, 1, 2\}$, denote by Σ^* the set of all finite words in the alphabet Σ, and by $|w|$ the length of the word w. Given $i < |w|$ we denote by w_i the letter of word at index i. Indices start by 0. For $S \subseteq \Sigma^*$, we let $\overline{S}$ be the set S extended by all prefixes of words in S. Given $\ell \geq 0$, we put $\overline{S}_\ell = \{w \in \overline{S} : |w| = \ell\}$. A word $w \in S$ is a *leaf* of S if there is no $w' \in S$ extending w. Given a word w and a letter $c \in \Sigma$, we denote by $w^\frown c$ the word obtained by adding c to the end of w. We also set $S^\frown c = \{w^\frown c : w \in S\}$.

Given distinct $u, v, w \in \Sigma^*$ with $|u| = |v| = |w| = \ell$, we define the following predicates:

$$\mathbf{1}(u) \equiv \exists_{i<\ell} : u_i = 1 \qquad \mathbf{2}(u) \equiv \exists_{i<\ell} : u_i = 2$$
$$\mathbf{11}(u, v) \equiv \exists_{i<\ell} : u_i = v_i = 1 \qquad \mathbf{22}(u, v) \equiv \exists_{i<\ell} : u_i = v_i = 2$$
$$\mathbf{111}(u, v, w) \equiv \exists_{i<\ell} : u_i = v_i = w_i = 1 \qquad u \perp v \equiv \neg\mathbf{1}(u) \text{ or } \neg\mathbf{1}(v) \text{ or } \mathbf{22}(u, v)$$

Definition 2.1 ($\mathbf{K}_4$-free diaries)**.** A set $S \subseteq \Sigma^*$ is called a $\mathbf{K}_4$*-free-diary* if no member of S extends any other and precisely one of the following seven conditions is satisfied for every i with $0 \leq i < \sup_{w \in S} |w|$:

1. **Splitting (possibly with new 1):** There is $w \in \overline{S}_i$ such that $\overline{S}_{i+1} = \overline{S}_i{}^\frown 0 \cup \{w\}^\frown 1$.
2. **New 1:** There is $w \in \overline{S}_i$ such that $\neg\mathbf{1}(w)$ and $\overline{S}_{i+1} = (\overline{S}_i \setminus \{w\})^\frown 0 \cup \{w\}^\frown 1$.
3. **New 2:** There is $w \in \overline{S}_i$ such that $\mathbf{1}(w)$, $\neg\mathbf{2}(w)$ and $\overline{S}_{i+1} = (\overline{S}_i \setminus \{w\})^\frown 0 \cup \{w\}^\frown 2$.
4. **New 11:** There are distinct words $v, w \in \overline{S}_i$ with $\mathbf{1}(v)$, $\mathbf{1}(w)$ and $\neg\mathbf{11}(v,w)$ such that $\overline{S}_{i+1} = (\overline{S}_i \setminus \{v,w\})^\frown 0 \cup \{v,w\}^\frown 1$.
5. **New 22:** There are distinct words $v, w \in \overline{S}_i$ with $\mathbf{2}(v)$, $\mathbf{2}(w)$, $\mathbf{11}(v,w)$, $\neg\mathbf{22}(v,w)$ and $\mathbf{111}(u,v,w)$ for every $u \in \overline{S}_i$ satisfying $\mathbf{11}(u,v)$ and $\mathbf{11}(u,w)$ such that $\overline{S}_{i+1} = (\overline{S}_i \setminus \{v,w\})^\frown 0 \cup \{v,w\}^\frown 2$.
6. **New 111:** There are distinct words $u, v, w \in \overline{S}_i$ with $\mathbf{11}(u,v)$, $\mathbf{11}(u,w)$, $\mathbf{11}(v,w)$, $\neg\mathbf{22}(u,v)$, $\neg\mathbf{22}(u,w)$, $\neg\mathbf{22}(v,w)$ and $\neg\mathbf{111}(u,v,w)$ such that $\overline{S}_{i+1} = (\overline{S}_i \setminus \{u,v,w\})^\frown 0 \cup \{u,v,w\}^\frown 1$.
7. **Leaf:** There is $w \in \overline{S}_i$ with $\mathbf{2}(w)$ satisfying:
   (a) **No new 11:** For every distinct $u, v \in \{z \in \overline{S}_i \setminus \{w\} : z \not\perp w\}$ it holds that $\mathbf{11}(u,v)$.
   (b) **No new 111:** For every distinct $u, v, v' \in \{z \in \overline{S}_i \setminus \{w\} : z \not\perp w\}$ satisfying $u \not\perp v$, $v \not\perp v'$ and $u \not\perp v'$ it holds that $\mathbf{111}(u,v,v')$.
   (c) **No new 2:** For every $u \in \{z \in \overline{S}_i \setminus \{w\} : z \not\perp w\}$ and $v \in S$, $|v| < i$ such that $w_{|v|} = u_{|v|} = 1$ it holds that $\mathbf{2}(u)$.
   (d) **No new 22:** For every distinct $u, v \in \{z \in \overline{S}_i \setminus \{w\} : z \not\perp w\}$ such that $\mathbf{111}(u,v,w)$ it holds that $\mathbf{22}(u,v)$. Moreover for every distinct $u, u' \in \{z \in \overline{S}_i \setminus \{w\} : z \not\perp w\}$ and $v \in S$, $|v| < i$ such that $w_{|v|} = u_{|v|} = u'_{|v|} = 1$ it holds that $\mathbf{22}(u,u')$.

   Moreover: $\overline{S}_{i+1} = \{z \in S_i \setminus \{w\} : z \perp w\}^\frown 0 \cup \{z \in S_i \setminus \{w\} : z \not\perp w\}^\frown 1$.

If S is a $\mathbf{K}_4$-free-diary then by $\mathbf{G}(S)$ we denote the graph with vertex set S with $u, v \in S$, $|u| < |v|$ forming an edge if and only if $v_{|u|} = 1$. Given a $\mathbf{K}_4$-free graph $\mathbf{G}$, we denote by $T(\mathbf{G})$ the set of all $\mathbf{K}_4$-free diaries S for which $\mathbf{G}(S)$ is isomorphic to $\mathbf{G}$.

Theorem 2.2. *For every finite $\mathbf{K}_4$-free graph $\mathbf{G}$, the big Ramsey degree of $\mathbf{G}$ in $\mathbf{R}_4$ equals $|T(\mathbf{G})| \cdot |\mathrm{Aut}(\mathbf{G})|$.*

A few known values are listed in Table 2. It is rather surprising to see such a complex structure of diaries arise from three very natural concepts of homogeneity, forbidding a clique, and the big Ramsey degree. Big Ramsey theorems always fix an enumeration of vertices and the structure arises from the tree of types which we describe now.

A graph is *enumerated* if its vertex set is ω. Given an enumerated graph $\mathbf{H}$ and $\ell \in \omega$, we call a finite graph $\mathbf{X}$ a *type of* $\mathbf{H}$ *on level* ℓ if the vertex set of $\mathbf{X}$ is $\{0, 1, \dots, \ell-1, t\}$ (where t is called the *type vertex*) and the graph created form $\mathbf{X}$ by removing t is an induced subgraph of $\mathbf{H}$. Types on level ℓ thus correspond to one vertex extensions of $\mathbf{H} \restriction_{\{0,1,\dots,\ell-1\}}$. Graph $\mathbf{X}$ is called a *type of* $\mathbf{H}$ if it is a type of $\mathbf{H}$ on level ℓ for some $\ell \in \omega$.

We denote by $\mathbb{T}_\mathbf{H}$ the set of all types of $\mathbf{H}$. Given $\mathbf{X}, \mathbf{X}' \in \mathbb{T}_\mathbf{H}$ we put $\mathbf{X} \subseteq \mathbf{X}'$ and call $\mathbf{X}'$ *a successor of* $\mathbf{X}$ if $\mathbf{X}$ is an induced subgraph of $\mathbf{X}'$. A successor is *immediate* it it differs by

	$\mathbf{K}_1$	$\mathbf{K}_2$	$\overline{\mathbf{K}}_2$	$\mathbf{K}_3$	$\overline{\mathbf{K}}_3$	$\mathbf{P}_2$	$\overline{\mathbf{P}}_2$
$\mathbf{R}$	1	$2 \cdot 2$ [17]	$2 \cdot 2$ [15]	$16 \cdot 3!$ [15]	$16 \cdot 3!$ [15]	$40 \cdot 2$ [16]	$40 \cdot 2$ [16]
$\mathbf{R}_3$	1 [14]	$2 \cdot 2$ [18]	$5 \cdot 2$ [2]	0	$161 \cdot 3!$ [2, 1]	$50 \cdot 2$ [2, 1]	$128 \cdot 2$ [2, 1]
$\mathbf{R}_4$	1 [8]	$36 \cdot 2$	$23 \cdot 2$	$22658 \cdot 3!$	$197613 \cdot 3!$	$160488 \cdot 2$	$267900 \cdot 2$

Table 1: Big Ramsey degrees of small graphs in $\mathbf{R}$, $\mathbf{R}_3$ and $\mathbf{R}_4$. $\overline{\mathbf{G}}$ denotes the complement of $\mathbf{G}$. $\mathbf{P}_2$ is a path with 2 edges. Big Ramsey degrees are often defined with respect to copies while we use embeddings. The difference between these two values is the size of the automorphism group of the graph. To prevent misunderstandings, we list values with respect to copies explicitly multiplied by the size of the automorphism group.

	$\mathbf{R}$ (A000182 in OEIS)	$\mathbf{R}_3$	$\mathbf{R}_4$
$\overline{\mathbf{K}}_1$	$1 \cdot 1!$	$1 \cdot 1!$	$1 \cdot 1!$
$\overline{\mathbf{K}}_2$	$2 \cdot 2!$	$5 \cdot 2!$	$23 \cdot 2!$
$\overline{\mathbf{K}}_3$	$16 \cdot 3!$	$161 \cdot 3!$	$197\,613 \cdot 3!$
$\overline{\mathbf{K}}_4$	$272 \cdot 4!$	$134\,397 \cdot 4!$	$*\ 272\,252\,729\,538\,223 \cdot 4!$
$\overline{\mathbf{K}}_5$	$7\,936 \cdot 5!$	$7\,980\,983\,689 \cdot 5!$	$*\ 43\,391\,315\,736\,159$ $690\,773\,738\,687\,637\,585 \cdot 5!$
$\overline{\mathbf{K}}_6$	$353\,792 \cdot 6!$	$45\,921\,869\,097\,999\,781 \cdot 6!$	$*\ 1\,075\,426\,511\,374\,671\,039\,522\,386$ $376\,330\,779\,194\,191\,609$ $662\,344\,240\,057\,102\,999 \cdot 6!$
$\overline{\mathbf{K}}_7$	$22\,368\,256 \cdot 7!$	$35\,268\,888\,847\,472\,944\,795\,910\,097 \cdot 7!$	?
$\overline{\mathbf{K}}_8$	$1\,903\,757\,312 \cdot 8!$	$4\,885\,777\,205\,485\,902\,177$ $648\,027\,702\,583\,670\,093 \cdot 8!$	?
$\overline{\mathbf{K}}_9$	$209\,865\,342\,976 \cdot 9!$	$159\,271\,391\,109\,084$ $147\,116\,751\,767\,705\,171$ $032\,995\,283\,089\,412\,057 \cdot 9!$	?
$\overline{\mathbf{K}}_{10}$	$29\,088\,885\,112\,832 \cdot 10!$	$1\,546\,604\,163\,029$ $698\,823\,334\,234\,758\,731$ $306\,633\,891\,622\,324\,147$ $639\,816\,544\,352\,644\,405 \cdot 10!$	?

Table 3: Big Ramsey degrees of anti-cliques. Values denoted by $*$ have not yet been verified by an independent implementation and should thus be considered preliminary [20].

only one vertex. Notice that every type has at most two immediate successors and $\mathbb{T}_{\mathbf{H}}$ can be viewed as an infinite tree rooted in the unique type on level 0. If $\mathbf{X}$ is on level ℓ and $\ell' \leq \ell$, we denote by $\mathbf{X}|_{\ell'}$ the unique type $\mathbf{X}' \in \mathbb{T}_{\mathbf{H}}$ on level ℓ satisfying $\mathbf{X}' \subseteq \mathbf{X}$.

Given $n \in \omega$ an *n-labeled graph* $\mathbf{G}$ is a graph we also denote by $\mathbf{G}$ along with a function $\chi_{\mathbf{G}}$ assigning every vertex v of $\mathbf{G}$ a *label* $\chi_{\mathbf{G}}(v) \in \{0, 1, \ldots, n-1\}$. To simplify the discussion below, we will additionally require the vertex sets of n-labeled graphs to be disjoint from ω. Given $n \in \omega$, an n-labeled graph $\mathbf{G}$, and types $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{n-1}$ of $\mathbf{H}$, all on the same level $\ell \in \omega$, we denote by $\mathbf{G} \oplus (\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$ the unique (non n-colored) graph $\mathbf{G}'$ extending $\mathbf{G}$ by vertices $0, 1, \ldots, \ell - 1$ such that for every vertex $v \in G$ it holds that the subgraph induced by $\mathbf{G}'$ on $\{0, 1, \ldots, \ell-1, v\}$ is isomorphic to $\mathbf{X}_{\chi_{\mathbf{G}}(v)}$ by renaming t to v. Given a tuple $(\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$ of types of a graph $\mathbf{H}$, all on the same level, we denote by $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$ the set of all finite n-labeled graphs $\mathbf{G}$ such that $\mathbf{G} \oplus (\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{n-1})$ has an embedding to $\mathbf{H}$.

Given an enumerated graph $\mathbf{H}$ and its vertex v, we denote by $\mathrm{Tp}_{\mathbf{H}}(v)$ the *type of v in $\mathbf{H}$* created from $\mathbf{H} \restriction \{0, 1, \ldots, v\}$ by renaming v to t. Given an enumerated $\mathbf{K}_4$-free graph $\mathbf{H}$, Zucker's theorem can colour only those subgraphs $\mathbf{A}$ of $\mathbf{H}$ which are simultaneously:

1. **Meet-closed**: $\max\{\ell < u : \mathrm{Tp}_{\mathbf{H}}(u)|_\ell = \mathrm{Tp}_{\mathbf{H}}(v)|_\ell\} \in A$ for every $u < v \in A$.

2. **Closed for age-changes**: For every $u_0, u_1, \ldots, u_{n-1} \in A$ and every $\ell < \min\{u_0, u_1, \ldots u_{n-1}\}$ such that $\mathrm{Age}_{\mathbf{H}}(\mathrm{Tp}_{\mathbf{H}}(u_0)|_{\ell+1}, \mathrm{Tp}_{\mathbf{H}}(u_1)|_{\ell+1}, \ldots, \mathrm{Tp}_{\mathbf{H}}(u_{n-1})|_{\ell+1}) \neq \mathrm{Age}_{\mathbf{H}}(\mathrm{Tp}_{\mathbf{H}}(u_0)|_\ell, \mathrm{Tp}_{\mathbf{H}}(u_1)|_\ell, \ldots, \mathrm{Tp}_{\mathbf{H}}(u_{n-1})|_\ell)$ we have $\ell \in A$.

Given an arbitrary subgraph $\mathbf{A}$ of $\mathbf{H}$, its *closure* is the (unqiue) inclusion minimal subgraph $\mathbf{B}$ of $\mathbf{H}$ which contains $\mathbf{A}$ and satisfies the conditions above.

Designing a diary for $\mathbf{R}_4$ corresponds to finding an enumerated $\mathbf{K}_4$-free graph $\mathbf{H}$ and an embedding $\varphi\colon \mathbf{R}_4 \to \mathbf{H}$ which minimizes, for every finite $\mathbf{K}_4$ graph $\mathbf{A}$, the number of order-preserving-isomorphism types of closures of graphs $\varphi(f(\mathbf{A}))$, $f \in \mathrm{Emb}(\mathbf{A}, \mathbf{R}_4)$ which then corresponds exactly to its big Ramsey degree.

Methods for minimizing the number of meets were introduced by Devlin, so the main difficulty is to minimize the number of ways age-changes can occur. Let $\mathbf{H}$ be a universal $\mathbf{K}_4$-free graph. Given a type $\mathbf{X}$ of $\mathbf{H}$ there are three possible sets $\mathrm{Age}_{\mathbf{H}}(\mathbf{X})$. If the vertex t is isolated then $\mathrm{Age}_{\mathbf{H}}(\mathbf{X})$ consists of all finite (1-coloured) $\mathbf{K}_4$-free graphs. If the neighborhood of t contains no edges then $\mathrm{Age}_{\mathbf{H}}(\mathbf{X})$ consists of all finite $\mathbf{K}_3$-free graphs. Finally, if the neighborhood of t contains a triangle then $\mathrm{Age}_{\mathbf{H}}(\mathbf{X})$ contains only graphs with no edges. The second resp. third case corresponds to the predicate **1** resp. **2**.

Similarly, given types $\mathbf{X}_0$ and $\mathbf{X}_1$ of $\mathbf{H}$ on the same level ℓ, $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_0, \mathbf{X}_1)$ inherits the structure of $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_2)$ on vertices of label 0 and $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_2)$ on vertices of label 1. There are three options for structures spanning both labels. Either $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_0, \mathbf{X}_1)$ contains triangles with both labels, or it contains only edges with both labels (if there exists $i < \ell$ such that t is connected to i in both $\mathbf{X}_0$ and $\mathbf{X}_1$) or it contains no 2-labeled edges (if there exist $i < j$ connected by an edge in $\mathbf{H}$ such that t is connected to both i and j in both $\mathbf{X}_0$ and $\mathbf{X}_1$). Again, the second resp. third case corresponds to the predicate **11** resp. **22**. See Example 4.3.5 of [3] for details.

Finally, given types $\mathbf{X}_0$, $\mathbf{X}_1$, and $\mathbf{X}_2$ on level ℓ, $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2)$ inherits the structure of ages of each of the pair of types considered. Moreover, it is possible that $\mathrm{Age}_{\mathbf{H}}(\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2)$ contains triangles spanning all three labels. These triangles are blocked if either there exists i such that t is connected to i in all three types (this is captured) by predicate **111**), or the age of one of the pairs already forbids the edge spanning the two labels.

If $\mathbf{H}$ is enumerated, every $\mathbf{X} \in \mathbb{T}_{\mathbf{H}}$ is determined by its level and the neighborhood of t. We can describe this by a word in the alphabet Σ. Given a word w and $\ell < |w|$, put $i(w, \ell) = \ell + \{j < i : w_j = 2\}$. It describes a type $\mathbf{X}_w$ on level $i(w, |w|)$ constructed as follows. If $w_i = 0$ then $i(w, \ell)$ and t are not adjacent. If $w_i = 1$ then $i(w, \ell)$ is adjacent to t. If $w_i = 2$ then there are adjacent vertices $i(w, \ell)$ and $i(w, \ell + 1)$ (in $\mathbf{H}$ as well as in $\mathbf{X}_w$) both adjacent to t.

Given a $\mathbf{K}_4$-free diary S, every word $w \in S$ corresponds to a type of some $\mathbf{K}_4$-free graph $\mathbf{H}$ in which $\mathbf{R}_4$ is embedded. Leaves correspond to types of vertices of $\mathbf{R}_4$ while non-leaves represent gadgets which are used to reduce the number of closures of subgraphs. These gadgets are of two kinds: a vertex or an edge connected to certain types. Each gadget represents either a meet (splitting) or an age-change, and every change is minimal (so, for example, **1** must happen before **2**). Finally, the conditions on leaves signify the fact that ages of every other type with the leaf vertex should already be minimized.

References

[1] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluís Vena, and Andy Zucker. Characterisation of the big Ramsey degrees of the generic partial order. Submitted, arXiv:2303.10088, 2023.

[2] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluís Vena, and Andy Zucker. Exact big Ramsey degrees for finitely constrained binary free amalgamation classes. *Journal of the European Mathematical Society*, August 2024.

[3] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluis Vena, and Andy Zucker. Exact big ramsey degrees for finitely constrained binary free amalgamation classes, 2023. Manuscript with more discussion and examples than published version, arXiv:2110.08409v2.

[4] Denis Devlin. *Some partition theorems and ultrafilters on ω.* PhD thesis, Dartmouth College, 1979.

[5] Natasha Dobrinen. The Ramsey theory of the universal homogeneous triangle-free graph. *Journal of Mathematical Logic*, 20(02):2050012, 2020.

[6] Natasha Dobrinen. The Ramsey theory of Henson graphs. *Journal of Mathematical Logic*, 23(01):2250018, 2023.

[7] Natasha Dobrinen. Ramsey theory of homogeneous structures: current trends and open problems. In *International Congress of Mathematicians*, pages 1462–1486. EMS Press, 2023.

[8] Mohamed El-Zahar and Norbert W. Sauer. The indivisibility of the homogeneous K_n-free graphs. *Journal of Combinatorial Theory, Series B*, 47(2):162–170, 1989.

[9] Wilfrid Hodges. *Model theory*, volume 42. Cambridge University Press, 1993.

[10] Jan Hubička. Big Ramsey degrees using parameter spaces. *Advances in Mathematics*, 2025. To appear, arXiv:2009.00967.

[11] Jan Hubička and Matěj Konečný. Twenty years of Nešetřil's classification programme of Ramsey classes. *arXiv:2501.17293*, 2025.

[12] Jan Hubička and Andy Zucker. A survey on big Ramsey structures. *arXiv:2407.17958*, 2025.

[13] Alexander S. Kechris, Vladimir G. Pestov, and Stevo Todorčević. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geometric and Functional Analysis*, 15(1):106–189, 2005.

[14] Péter Komjáth and Vojtech Rödl. Coloring of universal graphs. *Graphs and Combinatorics*, 2(1):55–60, 1986.

[15] Claude Laflamme, Norbert W. Sauer, and Vojkan Vuksanovic. Canonical partitions of universal structures. *Combinatorica*, 26(2):183–205, 2006.

[16] Jean A. Larson. Counting canonical partitions in the random graph. *Combinatorica*, 28(6):659–678, 2008.

[17] Maurice Pouzet and Norbert Sauer. Edge partitions of the Rado graph. *Combinatorica*, 16:505–520, 12 1996.

[18] Norbert W. Sauer. Edge partitions of the countable triangle free homogeneous graph. *Discrete Mathematics*, 185(1-3):137–181, 1998.

[19] Stevo Todorcevic. *Introduction to Ramsey spaces*, volume 174. Princeton University Press, 2010.

[20] Štěpán Vodseďálek. Counting big Ramsey degrees. Bachelor's thesis (in preparation), Charles University, 2025.

[21] Andy Zucker. On big Ramsey degrees for binary free amalgamation classes. *Advances in Mathematics*, 408:108585, 2022.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Big Ramsey degrees of universal ω-edge-labeled hypergraphs

(Extended abstract)

Jan Hubička[*] Matěj Konečný[†] Stevo Todorcevic[‡] Andy Zucker[§]

Abstract

We show that the big Ramsey degrees of every countable universal u-uniform ω-edge-labeled hypergraph are infinite for every $u \geq 2$. Together with a recent result of Braunfeld, Chodounský, de Rancourt, Hubička, Kawach, and Konečný this finishes full characterisation of unrestricted relational structures with finite big Ramsey degrees.

1 Introduction

Let A be a set and let u be a positive integer. We denote by $\binom{A}{u}$ the set of all u-element subsets of A. Given a countable set L of *labels*, an *L-edge-labeled u-uniform hypergraph* (or simply an *edge-labeled hypergraph*) is a pair $\mathbf{A} = (A, e_\mathbf{A})$, where $e_\mathbf{A}$ is a function $e_\mathbf{A} \colon \binom{A}{u} \to L$. We call A the *vertex set* of $\mathbf{A}$ and consider only finite and countably infinite vertex sets. We say that $\mathbf{A}$ is *finite* if A is finite. We will view u-uniform hypergraphs as $\{0, 1\}$-edge-labeled u-uniform hypergraphs (where the label 0 represents non-edges) and *graphs* as $\{0, 1\}$-edge-labeled 2-uniform hypergraphs.

Given L-edge-labeled u-uniform hypergraphs $\mathbf{A} = (A, e_\mathbf{A})$ and $\mathbf{B} = (B, e_\mathbf{B})$, an *embedding* $f \colon \mathbf{A} \to \mathbf{B}$ is an injective function $f \colon A \to B$ such that for every $E \in \binom{A}{u}$ we have $e_\mathbf{A}(E) = e_\mathbf{B}(f[E])$, where $f[E] = \{f(v) : v \in E\}$. If $A \subseteq B$ and the inclusion map is an

[*]Department of Applied Mathematics (KAM), Charles University, Malostranské náměstí 25, Praha 1, Czech Republic E-mail: hubicka@kam.mff.cuni.cz. Supported by project 25-15571S of the Czech Science Foundation (GAČR)

[†]Institute of Algebra, TU Dresden, Dresden, Germany. E-mail: matej.konecny@tu-dresden.de. Supported by a project that has received funding from the European Union (Project POCOCOP, ERC Synergy Grant 101071674). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

[‡]Department of Mathematics, University of Toronto, 40 St George St., Toronto, ON, Canada. E-mail: stevo@math.toronto.edu. Supported by NSERC grants RGPIN-2019-455916 and RGPIN-2025-04386.

[§]Department of Pure Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada. E-mail: a3zucker@uwaterloo.ca. Supported by NSERC grants RGPIN-2023-03269 and DGECR-2023-00412.

embedding, we call $\mathbf{A}$ a *substructure* of $\mathbf{B}$. We say that $\mathbf{A}$ is *homogeneous* if every isomorphism between finite substructures of $\mathbf{A}$ extends to an automorphism of $\mathbf{A}$, and $\mathbf{A}$ is *universal* if every countable L-edge-labeled u-uniform hypergraph embeds into $\mathbf{A}$. It is a well-known consequence of the Fraïssé theorem [9] that for every finite integer u and finite or countable set L there exists an up-to-isomorphism unique universal and homogeneous L-edge-labeled hypergraph $\mathbf{R}^u_L$. Equivalently, $\mathbf{R}^u_L$ can be characterised by the *extension property*: For every u-uniform L-edge-labeled hypergraph $\mathbf{B}$ and its finite substructure $\mathbf{A}$, every embedding $\mathbf{A} \to \mathbf{R}^u_\omega$ extends to an embedding $\mathbf{B} \to \mathbf{R}^u_\omega$, see e.g. [10]. If μ is a probability measure on L with full support, then letting $e_\mu \colon \binom{\omega}{u} \to L$ be randomly generated according to μ, the structure (ω, e_μ) is with probability 1 isomorphic to $\mathbf{R}^u_L$, and thus hypergraphs $\mathbf{R}^u_L$ can be called *random* countable edge-labeled hypergraphs. $\mathbf{R}^2_{\{0,1\}}$ is known as the *random graph* or *Rado graph* [5]. Given edge-labeled hypergraphs $\mathbf{A}$ and $\mathbf{B}$, we denote by $\mathrm{Emb}(\mathbf{A},\mathbf{B})$ the set of all embeddings from $\mathbf{A}$ to $\mathbf{B}$. If $\mathbf{C}$ is another edge-labeled hypergraph and $\ell \leq k < \omega$, we write $\mathbf{C} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,\ell}$ to denote the following statement:

> For every colouring $\chi \colon \mathrm{Emb}(\mathbf{A},\mathbf{C}) \to \{1,\dots,k\}$ with k colours, there exists an embedding $f \colon \mathbf{B} \to \mathbf{C}$ such that the restriction of χ to $\mathrm{Emb}(\mathbf{A}, f(\mathbf{B}))$ takes at most ℓ distinct values.

For a countably infinite edge-labeled hypergraph $\mathbf{B}$ and a finite substructure $\mathbf{A}$ of $\mathbf{B}$, the *big Ramsey degree of* $\mathbf{A}$ *in* $\mathbf{B}$ is the least number $D \in \omega$ (if it exists) such that $\mathbf{B} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,D}$ for every $k \in \omega$. We say that $\mathbf{B}$ *has finite big Ramsey degrees* if the big Ramsey degree of every finite substructure $\mathbf{A}$ of $\mathbf{B}$ exists.

In 1969 Laver introduced a proof technique which shows that $\mathbf{R}^2_L$ has finite big Ramsey degrees for every finite set L [7, 8, 17]. This was refined by Laflamme, Sauer, and Vuksanovic [14] to precisely characterise the big Ramsey degrees of these structures. Finiteness of big Ramsey degrees of $\mathbf{R}^3_{\{0,1\}}$ was announced at Eurocomb 2019 by Balko, Chodounský, Hubička, Konečný, and Vena [1] with a proof published in 2020 [2]. In 2024, Braunfeld, Chodounský, de Rancourt, Hubička, Kawach, and Konečný [4] extended the proof to arbitrary finite $u > 0$ and finite L and generalised the setup to model-theoretic L'-structures where L' is a (possibly infinite) relational language containing only finitely many relations of every given arity $a > 1$. Answering Question 7.5 of [4] we show that the assumption about finiteness of L as well as the above assumption about language L' is necessary:

Theorem 1.1. *Let $u > 1$ be finite and let $\mathbf{A}$ be any ω-edge-labeled u-uniform hypergraph with 2 vertices. Then $\mathbf{A}$ does not have finite big Ramsey degree in $\mathbf{R}^u_\omega$.*

It is known that the big Ramsey degrees of $\mathbf{R}^1_\omega$ are finite [4]. It is also easy to show:

Theorem 1.2. *Let $u > 1$ be finite and $\mathbf{A}$ be the ω-edge-labeled u-uniform hypergraph with 1 vertex. Then the big Ramsey degree of $\mathbf{A}$ in $\mathbf{R}^u_\omega$ is 1.*

Consequently, our result concludes the characterisation of unrestricted structures with finite big Ramsey degrees (see [4] for precise definitions). Our proof introduces a new technique that complements the existing arguments for infinite lower bounds which can be divided into three types: Counting number of oscillations of monotone functions assigned to subobjects [6, 16, 3], study of the partial order of ages (ranks or orbits) of vertices [15], and arguments based on the distance and diameter in metric spaces [14].

Solving the question about finiteness of big Ramsey degrees of $\mathbf{R}^2_\omega$ suggests the following question about its reduct, which forgets the actual labels of edges and only records information about pairs of vertices with equivalent labels:

Problem 1.3. Let L be a relational language with a single quaternary relation R and $\mathcal{K}$ the class of all finite L-structures $\mathbf{A}$ such that

1. for every $(a,b,c,d) \in R^{\mathbf{A}}$ it holds that $a \neq b$, $c \neq d$ and $(c,d,a,b) \in R^{\mathbf{A}}$,
2. for pair of distinct vertices a, b of $\mathbf{A}$ it holds that $(a,b,a,b), (a,b,b,a) \in R^{\mathbf{A}}$,
3. whenever (a,b,c,d) and (c,d,e,f) is in $R^{\mathbf{A}}$ then also $(a,b,e,f) \in R^{\mathbf{A}}$.

(In other words, $R^{\mathbf{A}}$ defines an equivalence on 2-element subsets of vertices of A.) Does the Fraïssé limit of $\mathcal{K}$ have finite big Ramsey degrees?

It is known that $\mathcal{K}$ has a precompact Ramsey expansion [12, 11] (fixing a linear ordering of vertices as well as a linear ordering of equivalence classes) and thus finite small Ramsey degrees. However, the question about the finiteness of big Ramsey degrees is fully open.

2 Compressed tree of types

We devote the rest of this abstract to a discussion of the proof of Theorem 1.1. Toward that, we fix $u > 1$ and a hypergraph $\mathbf{R}^u_\omega$ with vertex set ω (that is, we work with an arbitrary but fixed enumeration of $\mathbf{R}^u_\omega$). We will construct explicit colourings which contradict the existence of big Ramsey degrees in $\mathbf{R}^u_\omega$.

Our construction is based on ideas used for analyzing structures which *do* have finite big Ramsey degrees. This is done using Ramsey-type theorems working with the so-called *tree of types*, see e.g. [13]. The main difficulty of applying this technique to $\mathbf{R}^u_\omega$ is the fact that the tree of types of $\mathbf{R}^u_\omega$ is infinitely branching. We overcome this problem by using a related tree which is finitely branching but the number of immediate successors of a vertex grows very rapidly. This lets us reverse the argument and instead of showing that big Ramsey degrees are finite, we obtain enough structure to show that they are infinite.

Let us introduce the key definitions. Put $L = \omega \cup \{\star\}$ where $\star$ will play the role of a special label which intuitively means that the information is "missing".

Definition 2.1 (f-type)**.** Let $f\colon \omega \to \omega \cup \{\omega\}$ be an arbitrary function. We call an L-edge-labeled u-uniform hypergraph $\mathbf{X}$ an *f-type of level ℓ* if:

1. The vertex set of $\mathbf{X}$ is $X = \{0, 1, \ldots, \ell - 1\} \cup \{t\}$ where t is a special vertex, called the *type vertex.*
2. For every $E \in \binom{X}{u}$ with $e_{\mathbf{X}}(E) \neq \star$ it holds that $t \in E$ and $e_{\mathbf{X}}(E) < f(\max(E \setminus \{t\}))$.
3. For every $E \in \binom{X}{u}$ with $t \in E$ such that $f(\max(E \setminus \{t\})) = \omega$ it holds that $e_{\mathbf{X}}(E) \neq \star$.

We also call a hyper-graph $\mathbf{X}$ simply an *f-type* if it is an f-type of level ℓ for some $\ell \in \omega$. In this situation we put $\ell(\mathbf{X}) = \ell$.

Definition 2.2 (Tree of f-types)**.** Let $f\colon \omega \to \omega \cup \{\omega\}$ be an arbitrary function. By T_f we denote the set of all f-types. We will view T_f as a (set-theoretic) tree equipped with a partial order $\sqsubseteq$ and operation $\wedge$ (*meet*) defined as follows: Given f-types $\mathbf{X}, \mathbf{Y} \in T_f$ we put $\mathbf{X} \sqsubseteq \mathbf{Y}$ if and only if $\mathbf{X}$ is an (induced) sub-structure of $\mathbf{Y}$. By $\mathbf{X} \wedge \mathbf{Y}$ we denote the (unique) f-type $\mathbf{Z} \in T_f$ such that $\mathbf{Z} \sqsubseteq \mathbf{X}$, $\mathbf{Z} \sqsubseteq \mathbf{Y}$ of largest level among all f-types with this property. Finally, given integer ℓ, we put $T_f(\ell) = \{\mathbf{X} \in T_f : \ell(\mathbf{X}) = \ell\}$ and call it the *level* ℓ of T_f. We call $\mathbf{X} \in T_f$ an *immediate successor* of $\mathbf{Y} \in T_f$ if and only if $\mathbf{Y} \sqsubseteq \mathbf{X}$ and $\ell(\mathbf{X}) = \ell(\mathbf{Y}) + 1$.

The usual tree of types corresponds to using the constant function f^ω where $f^\omega(i) = \omega$ for every $i \in \omega$. Every f^ω-type $\mathbf{X}$ of level ℓ can be thought of as a one vertex extension of some ω-edge-labeled u-uniform hypergraph $\mathbf{A}$ with vertex set $\{0, 1, \ldots, \ell - 1\}$. For this reason we put $e_{\mathbf{X}}(E) = \star$ for every $E \in \binom{\{0,1,\ldots,\ell-1\}}{u}$ since this label is determined by $\mathbf{A}$. We will consider functions f with $\mathrm{Im}(f) \subseteq \omega$ and then f-types capture only partial information about these one vertex extensions. We make this explicit as follows:

Definition 2.3 (f-type of a vertex)**.** Given $v \in \mathbf{R}^u_\omega$, the *$f$-type of* v, denoted by $\mathrm{Tp}_f(v)$, is an f-type $\mathbf{X}$ of level v where given $E \in \binom{X}{u}$, and writing $E' = (E \setminus \{t\}) \cup \{v\}$, we have

$$e_{\mathbf{X}}(E) = \begin{cases} e_{\mathbf{R}^u_\omega}(E') & \text{if } t \in E \text{ and } e_{\mathbf{R}^u_\omega}(E') < f(\max(E \setminus \{t\})) \\ \star & \text{otherwise.} \end{cases}$$

Notice that for every choice of f it follows by universality and homogeneity of $\mathbf{R}^u_\omega$ that for every f-type $\mathbf{X}$ there exist infinitely many vertices v of $\mathbf{R}^u_\omega$ satisfying $\mathrm{Tp}_f(v) \sqsupseteq \mathbf{X}$.

3 Persistent colouring of $\mathbf{R}^u_\omega$

If function $f\colon \omega \to \omega \cup \{\omega\}$ is fixed then every vertex v of $\mathbf{R}^u_\omega$ is associated with the f-type $\mathrm{Tp}_f(v) \in T_f$. Given two vertices of $\mathbf{R}^u_\omega$, we can then study their iterated meet closure in the tree T_f defined as follows.

Definition 3.1. Given a pair of nodes $\mathbf{X}, \mathbf{Y} \in T_f$, its *$f$-height*, denoted by $\mathrm{height}_f(\mathbf{X}, \mathbf{Y})$, is the number of repetitions of the following procedure:

1. Put $\mathbf{Z} = \mathbf{X} \wedge \mathbf{Y}$.
2. If $\mathrm{Tp}_f(\ell(\mathbf{Z})) = \mathbf{Z}$ terminate.
3. Repeat from step 1 with $\mathbf{X} = \mathrm{Tp}_f(\ell(\mathbf{Z}))$ and $\mathbf{Y} = \mathbf{Z}$.

Given vertices $v, w \in R^u_\omega$ we also put $\mathrm{height}_f(v, w) = \mathrm{height}_f(\mathrm{Tp}_f(v), \mathrm{Tp}_f(w))$.

Theorem 3.2. *Assume that $f(\ell)\colon \omega \to \omega$ is a function satisfying*

$$f(\ell) \geq \prod_{u-2 \leq i < \ell} (f(i) + 1)^{\binom{i}{u-2}}$$

for every $\ell \in \omega$. Then for every embedding $\varphi\colon \mathbf{R}^u_\omega \to \mathbf{R}^u_\omega$ there exists integer m such that for every $n > m$ there exist vertices $v, w \in \varphi[R^u_\omega]$ satisfying

1. *if $u = 2$ then $e_{\mathbf{R}^u_\omega}(\{v, w\}) = 0$ and,*

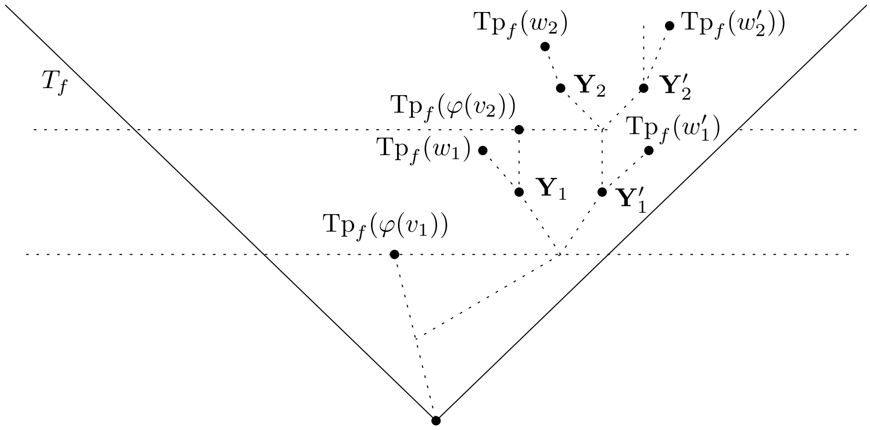


Figure 1: Configuration of tree nodes used in the proof of Theorem 3.2.

2. $\mathrm{height}_f(v, w) = n$.

Notice that Theorem 3.2 immediately implies Theorem 1.1. Let $\mathbf{A}$ be as in Theorem 1.1 and assume $A = \{0, 1\}$. If $u = 2$, without loss of generality we can also assume that $e_{\mathbf{A}}(\{0, 1\}) = 0$. Given finite $n > 1$, we define colouring $\chi_n\colon \mathrm{Emb}(\mathbf{A}, \mathbf{R}^{\omega}_{u}) \to n$ by putting $\chi_n(h) = \mathrm{height}_f(h(0), h(1)) \mod n$ for every $h \in \mathrm{Emb}(\mathbf{A}, \mathbf{R}^u_\omega)$. By Theorem 3.2, for every embedding $\varphi\colon \mathbf{R}^u_\omega \to \mathbf{R}^u_\omega$ there are copies of $\mathbf{A}$ in every colour showing that the big Ramsey degree of $\mathbf{A}$ is greater than n.

Proof of Theorem 3.2 (sketch). Fix f and embedding $\varphi\colon \mathbf{R}^u_\omega \to \mathbf{R}^u_\omega$ as in the statement. The rapid growth of f ensures that for every f-type $\mathbf{X}$ it holds that number of immediate successors of $\mathbf{X}$ is greater than number of nodes of T_f of level $\ell(\mathbf{X})$. This makes it possible to obtain for every vertex $v \in R^u_\omega$ (up to $u-1$ exceptions) vertices $v_+, v'_+ \in R^u_\omega$ with the property that $\varphi(v) = \ell(\mathrm{Tp}_f(\varphi(v))) = \ell(\mathrm{Tp}_f(\varphi(v_+)) \wedge \mathrm{Tp}_f(\varphi(v'_+))$.

Using a technique inspired by Lachlan, Sauer, and Vuksanovic [14] we obtain vertices $v_0, v_1, \ldots \in R^u_\omega$, $w_1, w_2, \ldots \in \varphi[R^u_\omega]$, $w'_1, w'_2, \ldots \in \varphi[R^u_\omega]$ and nodes $\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}'_0, \mathbf{Y}'_1, \ldots$ in configuration as depicted in Figure 1. Then it follows that for every $i > 1$ we get $\mathrm{height}_f(w_{i+1}, w'_{i+1}) = \mathrm{height}_f(w_i, w'_i) + 1$. □

4 Acknowledgements

The initial variant of this construction was obtained during the Thematic Program on Set Theoretic Methods in Algebra, Dynamics and Geometry at Fields Institute. We would like to thank the organizers Michael Hrušák, Kathryn Mann, Justin Moore and Stevo Todorcevic, for such a great program.

References

[1] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, and Lluís Vena. Big Ramsey degrees of 3-uniform hypergraphs. *Acta Mathematica Universitatis Comenianae*, 88(3):415–422, 2019. Extended abstract for Eurocomb 2019.

[2] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, and Lluís Vena. Big Ramsey degrees of 3-uniform hypergraphs are finite. *Combinatorica*, 42(2):659–672, 2022.

[3] Dana Bartošová, David Chodounský, Barbara Csima, Jan Hubička, Matěj Konečný, Joey Lakerdas-Gayle, Spencer Unger, and Andy Zucker. Oscillating subalgebras of the atomless countable Boolean algebra. Preprint, arXiv:2505.22603.

[4] Samuel Braunfeld, David Chodounský, Noé de Rancourt, Jan Hubička, Jamal Kawach, and Matěj Konečný. Big Ramsey Degrees and Infinite Languages. *Advances in Combinatorics*, 2024:4, 2024. 26pp.

[5] Peter J Cameron. The random graph. *The Mathematics of Paul Erdös II*, pages 333–351, 1997.

[6] David Chodounský, Monroe Eskew, and Thilo Weinert. Colors of the pseudotree. To appear as Extended abstract for Eurocomb 2025, arXiv:2503.18727, 2025.

[7] Denis Devlin. *Some partition theorems and ultrafilters on ω.* PhD thesis, Dartmouth College, 1979.

[8] Paul Erdős and András Hajnal. Unsolved and solved problems in set theory. In *Proceedings of the Tarski Symposium (Berkeley, Calif., 1971), Amer. Math. Soc., Providence*, volume 1, pages 269–287, 1974.

[9] Roland Fraïssé. Sur certaines relations qui généralisent l'ordre des nombres rationnels. *Comptes Rendus de l'Academie des Sciences*, 237:540–542, 1953.

[10] Wilfrid Hodges. *Model theory*, volume 42. Cambridge University Press, 1993.

[11] Jan Hubička and Matěj Konečný. Twenty years of Nešetřil's classification programme of Ramsey classes. *arXiv:2501.17293*, 2025.

[12] Jan Hubička and Jaroslav Nešetřil. All those Ramsey classes (Ramsey classes with closures and forbidden homomorphisms). *Advances in Mathematics*, 356C:106791, 2019.

[13] Jan Hubička and Andy Zucker. A survey on big Ramsey structures. *arXiv:2407.17958*, 2025.

[14] Claude Laflamme, Norbert W. Sauer, and Vojkan Vuksanovic. Canonical partitions of universal structures. *Combinatorica*, 26(2):183–205, 2006.

[15] Norbert W Sauer. Canonical vertex partitions. *Combinatorics, Probability and Computing*, 12(5-6):671–704, 2003.

[16] Stevo Todorcevic. Oscillations of real numbers. In *Logic colloquium '86 (Hull, 1986)*, volume 124 of *Stud. Logic Found. Math.*, pages 325–331. North-Holland, Amsterdam, 1988.

[17] Stevo Todorcevic. *Introduction to Ramsey spaces*, volume 174. Princeton University Press, 2010.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Hypertree shrinking avoiding low degree vertices

(Extended abstract)

Karolína Hylasová* Tomáš Kaiser†

Abstract

The shrinking operation converts a hypergraph into a graph by choosing, from each hyperedge, two endvertices of a corresponding graph edge. A hypertree is a hypergraph which can be shrunk to a tree on the same vertex set. Klimošová and Thomassé [J. Combin. Theory Ser. B 156 (2022), 250–293] proved (as a tool to obtain their main result on edge-decompositions of graphs into paths of equal length) that any rank 3 hypertree T can be shrunk to a tree where the degree of each vertex is at least $1/100$ times its degree in T. We prove a stronger bound that moreover applies to hypergraphs of any rank, replacing the constant $1/100$ with $1/2k$ when the rank is k. In place of entropy compression (used by Klimošová and Thomassé), we use a hypergraph orientation lemma combined with a characterisation of edge-coloured graphs admitting rainbow spanning trees.

1 Introduction

A hypergraph H is a *hypertree* if it is possible to choose two vertices from each hyperedge in such a way that the chosen pairs, viewed as edges of a graph on vertex set $V(H)$, form a tree.

The above operation that produces a graph from a hypergraph by choosing a pair of vertices in each hyperedge will be called *shrinking*. Thus, a hypertree is a hypergraph which can be shrunk to a tree.

Klimošová and Thomassé [3] derived a result on hypertree shrinking that preserves vertex degree up to a constant factor, and used it as one of the tools needed to obtain their main result about decompositions of 3-edge-connected graphs into paths of equal length. Their lemma on shrinking [3, Lemma 22] is as follows:

*Department of Mathematics and European Centre of Excellence NTIS (New Technologies for the Information Society), University of West Bohemia in Pilsen, Pilsen, Czech Republic. E-mail: khylas@kma.zcu.cz.

†Department of Mathematics and European Centre of Excellence NTIS (New Technologies for the Information Society), University of West Bohemia in Pilsen, Pilsen, Czech Republic. E-mail: kaisert@kma.zcu.cz.

Lemma 1. *Let H be a hypertree with hyperedges of size at most three. It is possible to shrink H to a tree T such that for the degrees of any vertex v of H, we have*

$$d_T(v) \geq \frac{d_H(v)}{100}.$$

Lemma 1 is proved using entropy compression, a method originally devised to prove an algorithmic version of the Lovász Local Lemma [4].

In this paper, we use different methods to strengthen Lemma 1 in two ways: first, our version ensures a stronger degree bound, and second, it applies to hypergraphs with an arbitrary size of the hyperedges. In addition, the proof is conceptually simpler. We prove:

Theorem 2. *Let H be a hypertree with hyperedges of size at most k. It is possible to shrink H to a tree T such that for every vertex v of H,*

$$d_T(v) \geq \max\left\{ \left\lfloor \frac{d_H(v)}{k} \right\rfloor, 1 \right\}.$$

In particular, we have $d_T(v) \geq d_H(v)/2k$ for such a tree T.

2 Preliminaries

We review some of the basic definitions on hypergraphs. A *hypergraph* is a pair $H = (V, E)$, where V is a finite set and E is a set of subsets of V. The elements of V are the *vertices* of H and the elements of E are the *hyperedges* of H. The *rank* of a hypergraph H is the maximum size of a hyperedge of H. The number of hyperedges of H containing a vertex $v \in V$ is the *degree* of v and is denoted by $d_H(v)$.

It will be useful in our argument to consider directed hypergraphs, consisting of a set of hyperarcs on a finite vertex set. A *hyperarc* is a hyperedge e together with a designated *head*. The other vertices of e are called the *tails* of e. The *indegree* of vertex v in a directed hypergraph $\vec{H}$, denoted by $d^{IN}_{\vec{H}}(v)$, is the number of hyperarcs whose head is v. Similarly, the *outdegree* of v (denoted by $d^{OUT}_{\vec{H}}(v)$) is the number of hyperarcs in which v is a tail. Observe that a hyperarc of size k contributes to the indegree of exactly one vertex and to the outdegree of $k-1$ vertices.

3 Tools

In the proof of Theorem 2, we will need two main tools. The first one is an orientation lemma for hypergraphs with lower bounds on the indegrees. The second one is a characterisation of edge-coloured graphs admitting rainbow spanning trees, described in the last few paragraphs of this section. Throughout this section we define $f(X)$ as $\sum_{v \in X} f(v)$.

We begin with the topic of hypergraph orientation. A result of Frank, Királyi and Királyi [2, Lemma 3.3] gives a necessary and sufficient condition for the existence of an orientation of a hypergraph with prescribed indegrees.

What we actually need is a sufficient condition for the existence of an orientation where the indegree of each vertex satisfies a given lower bound. We originally proved the following result by modifying the proof of [1, Lemma 3], but (as pointed out by a reviewer) it is a simple consequence of Hall's Marriage Theorem.

Lemma 3. *Let $H = (V, E)$ be a hypergraph and let $f : V \to \mathbb{Z}_{\geq 0}$ be a mapping of the vertex set V of H into the set of non-negative integers. Assume that for every $F \subseteq V$,*

$$f(F) \leq e^*(F), \tag{1}$$

where $e^*(F)$ denotes the number of hyperedges incident with F. Then there is an orientation $\vec{H}$ of H such that

$$d^{IN}_{\vec{H}}(v) \geq f(v) \tag{2}$$

for every $v \in V$.

The second main tool which we will use to prove Theorem 2 is a necessary and sufficient condition for an edge-coloured graph to contain a rainbow spanning tree.

Let G be a multigraph with a (not necessarily proper) edge colouring. A subgraph of G is *rainbow* if it does not contain two edges with the same color. The following necessary and sufficient condition for the existence of a rainbow spanning tree has been derived in [6] and, independently, in [5, Section 41.1a]. We remark that although [6] states the result for simple graphs only, is is established in [5] for multigraphs, i.e., graphs in which parallel edges are allowed.

Theorem 4 ([6])**.** *Let G be a (possibly improperly) edge-coloured multigraph of order n. There exists a rainbow spanning tree of G if and only if for any set of r colours ($0 \leq r \leq n-2$), the removal of all edges coloured with these r colours from G results in a graph with at most $r+1$ components.*

4 Shrinking hypertrees

In this section, we outline the proof of Theorem 2. We begin with an application of Lemma 3 which yields a lower bound for the indegrees in an orientation of a hypergraph obtained by setting $f(v) = \lfloor d_H(v)/k \rfloor$ in Lemma 3:

Lemma 5. *Let $H = (V, E)$ be a hypergraph of rank at most k. There exists an orientation $\vec{H}$ of H in which*

$$d^{IN}_{\vec{H}}(v) \geq \left\lfloor \frac{d_H(v)}{k} \right\rfloor$$

for every vertex $v \in V$.

Suppose that H is a hypertree of rank k. Let G_H be a multigraph obtained by adding, for each hyperedge e of H, a complete graph on the vertex set of e. We colour each of these complete graphs by different colours. Since H is a hypertree, G_H has a rainbow spanning tree and consequently, the condition from Theorem 4 holds.

Let $\vec{H}$ be the orientation from Lemma 5. We construct an edge-coloured multigraph $G(\vec{H})$ on the vertex set of $\vec{H}$ by the following rule: for each hyperarc $\vec{e}$ of $\vec{H}$, add to G a star whose center is the head of e and whose leaves are the tails of e. Furthermore, all the edges of this star have the same colour, which differs from the colours used for the other stars.

We claim that $G(\vec{H})$ has a rainbow spanning tree T. Unlike the multigraph G_H, in the construction of $G(\vec{H})$ we used stars instead of the complete graphs, but this replacement has no effect on the validity of the condition of Theorem 4, since the number of components does not change.

For each hyperedge e of H, T has one edge e' corresponding to e, namely the edge selected from the star corresponding to an orientation $\vec{e}$ of e. Moreover, the head of $\vec{e}$ is one of the endvertices of e'. This means that each star increases the degree of its center in T by one.

It follows that if v is a vertex of H, then the degree of v in T is at least the indegree of v in $\vec{H}$ — that is, at least $\lfloor d_H(v)/k \rfloor$. At the same time, the degree of each vertex is at least 1 because T is a tree. It is easy to see that this implies the second assertion of Theorem 2.

Acknowledgements

We would like to thank the reviewers for their valuable comments and suggestions.

References

[1] Y. Caro and A. Hansberg. Degrees in oriented hypergraphs and ramsey p-chromatic number. *The Electronic Journal of Combinatorics*, 19:P16, 2012.

[2] A. Frank, T. Király, and Z. Király. On the orientation of graphs and hypergraphs. *Discrete Applied Mathematics*, 131:385–400, 2003.

[3] T. Klimošová and S. Thomassé. Edge-partitioning 3-edge-connected graphs into paths. *Journal of Combinatorial Theory, Series B*, 156:250–293, 2022.

[4] R. A. Moser and G. Tardos. A constructive proof of the general lovász local lemma. *Journal of the ACM*, 57, 2010.

[5] A. Schrijver. *Combinatorial Optimization.* Springer, 2003.

[6] K. Suzuki. A necessary and sufficient condition for the existence of a heterochromatic spanning tree in a graph. *Graphs and Combinatorics*, 22:261–269, 2006.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

An improved hypergraph Mantel's Theorem

(Extended abstract)

Daniel Iľkovič*   Jun Yan†

Abstract

In a recent paper [1], Chao and Yu used an entropy method to show that the Turán density of a certain family $\mathcal{F}$ of $\lfloor r/2 \rfloor$ triangle-like r-uniform hypergraphs is $r!/r^r$. Later, Liu [7] determined for large n the exact Turán number $\mathrm{ex}(n, \mathcal{F})$ of this family, and showed that the unique extremal graph is the balanced complete r-partite r-uniform hypergraph. These two results together can be viewed as a hypergraph version of Mantel's Theorem. Building on their methods, we improve both results by showing that they still hold with a subfamily $\mathcal{F}' \subset \mathcal{F}$ of size $\lceil r/e \rceil$ in place of $\mathcal{F}$.

1 Introduction

Given a family of r-uniform hypergraphs $\mathcal{F}$, an r-uniform hypergraph H is said to be *$\mathcal{F}$-free* if H does not contain any $F \in \mathcal{F}$ as a subgraph. The famous Turán problem studies the following two quantities.

- The *Turán number* $\mathrm{ex}(n, \mathcal{F})$, which is the maximum number of edges in an $\mathcal{F}$-free r-uniform hypergraph on n vertices.
- The *Turán density* $\pi(\mathcal{F})$, which is defined to be $\lim_{n\to\infty} \mathrm{ex}(n, \mathcal{F})/\binom{n}{r}$.

In the case when $r = 2$, or equivalently on graphs, the Turán problem is well-understood. For every non-bipartite graph F, the Erdős-Stone Theorem determines $\pi(F)$ exactly and

*Faculty of Informatics, Masaryk University, Brno, Czech Republic; Mathematics Institute and DIMAP, University of Warwick, Coventry, UK.
†Mathematics Institute, University of Warwick, Coventry, UK. Supported by the Warwick Mathematics Institute Centre for Doctoral Training, and by funding from the UK EPSRC (Grant number: EP/W523793/1).

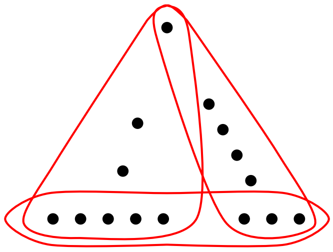


Figure 1: The $(5,3)$-tent $\Delta_{(5,3)}$

$\mathrm{ex}(n,F)$ asymptotically. In the special case when $F=K_3$, Mantel's Theorem states that $\pi(K_3)=1/2$ and $\mathrm{ex}(n,K_3)=\lfloor n^2/4\rfloor$, with the balanced complete bipartite graph being the unique extremal graph.

In contrast, when $r\geq 3$, the hypergraph Turán problem is notoriously difficult with very few known results. For a comprehensive overview on known results in hypergraph Turán problems, we refer the readers to Keevash's survey [6].

One general result states that for every $r\geq 2$ and every family $\mathcal{F}$ of r-uniform hypergraphs, $\pi(\mathcal{F})\geq r!/r^r$ if and only if no $F\in\mathcal{F}$ is r-partite, and $\pi(F)=0$ otherwise. Note that $\pi(K_3)=1/2=2!/2^2$, so a natural avenue of research is to generalise Mantel's Theorem to $r\geq 3$, by finding a family $\mathcal{F}$ of hypergraph analogues of the triangle that satisfies $\pi(\mathcal{F})=r!/r^r$.

One such attempt uses the r-uniform hypergraph $\mathbb{T}_r$, which is isomorphic to the hypergraph $\Delta_{(r-1,1)}$ that will be defined below. Frankl and Füredi showed that $\pi(\mathbb{T}_3)=3!/3^3=2/9$ in [2], and Pikhurko showed that $\pi(\mathbb{T}_4)=4!/4^4=3/32$ in [9]. However, Frankl and Füredi [3] also showed that $\pi(\mathbb{T}_5)>5!/5^5$ and $\pi(\mathbb{T}_6)>6!/6^6$, so forbidding just $\mathbb{T}_r$ does not generalise Mantel's Theorem in general.

Instead, for every $1\leq i\leq\lfloor r/2\rfloor$, define the $(r-i,i)$-*tent* $\Delta_{(r-i,i)}$ to be the r-uniform hypergraph with vertex set $[2r-1]$ and edge set (see Figure 1)

$$\{\{1,2,\dots,r\},\{1,\dots,i,r+1,\dots,2r-i-1,2r-1\},\{i+1,\dots,r,2r-i,\dots,2r-1\}\}.$$

For $1\leq k\leq\lfloor r/2\rfloor$, let $\mathcal{F}_{r,k}=\{\Delta_{(r-i,i)}\mid 1\leq i\leq k\}$. Recently, Chao and Yu found the following generalisation of Mantel's Theorem in their breakthrough paper [1].

Theorem 1.1 ([1, Theorem 1.4])**.** *For every $r\geq 2$, $\pi(\mathcal{F}_{r,\lfloor r/2\rfloor})=r!/r^r$.*

Chao and Yu's proof of Theorem 1.1 used a novel entropy method, which reduced this Turán density problem to an optimisation problem whose optimal value is an upper bound on $\pi(\mathcal{F}_{r,\lfloor r/2\rfloor})$. They then showed that the optimal value is $r!/r^r$, which combined with the fact that no $F\in\mathcal{F}_{r,\lfloor r/2\rfloor}$ is r-partite gives the result.

Soon after, using a powerful framework developed by Liu, Mubayi and Reiher [8] and refined by Hou, Liu and Zhao [4] to prove strong stability results in Turán problems, Liu [7] proved the following exact version of Theorem 1.1.

Theorem 1.2 ([7, Theorem 1.2]). *For every $r \geq 2$ and n sufficiently large,* $\mathrm{ex}(n, \mathcal{F}_{r,\lfloor r/2 \rfloor}) = |E(T^r(n))|$*, with the unique extremal graph being $T^r(n)$, the balanced complete r-partite r-uniform hypergraph on n vertices.*

In our paper [5], we improve both Theorem 1.1 and Theorem 1.2. First, we show that forbidding a smaller family of tents still gives Turán density $r!/r^r$. Note that we start with $r = 4$ as $r = 3$ does not satisfy $\lceil r/e \rceil \leq \lfloor r/2 \rfloor$.

Theorem 1.3. *For every $r \geq 4$, let $k = \lceil r/e \rceil$, then $\pi(\mathcal{F}_{r,k}) = r!/r^r$.*

Our proof follows Chao and Yu's entropy method, with most of the effort devoted to showing that the corresponding optimisation problem obtained using $\mathcal{F}_{r,k}$ instead of $\mathcal{F}_{r,\lfloor r/2 \rfloor}$ still has the same optimal value $r!/r^r$. Moreover, we show that the value $k = \lceil r/e \rceil$ cannot be improved using this method as the optimal value can exceed $r!/r^r$ if $k < \lceil r/e \rceil$.

We then proceed as in [7] to get an exact Turán number result from this density result.

Theorem 1.4. *For every $r \geq 4$, let $k = \lceil r/e \rceil$, then* $\mathrm{ex}(n, \mathcal{F}_{r,k}) = |E(T^r(n))|$ *for every sufficiently large n, with $T^r(n)$ being the unique extremal graph.*

In Section 2 and 3, we give sketches of our proofs of Theorem 1.3 and Theorem 1.4, respectively. We end by mentioning a few related problems in Section 4. For more details we refer the readers to our paper [5].

2 The density result

Let $\mathcal{F}$ be a family of r-uniform hypergraphs. An r-uniform hypergraph H is *$\mathcal{F}$-hom-free* if there is no homomorphism from any $F \in \mathcal{F}$ to H. It is known that $\pi(\mathcal{F})$ is the supremum of the blowup density $b(H)$, taken over all $\mathcal{F}$-hom-free r-uniform hypergraphs H.

To study the blowup density using entropy, Chao and Yu defined the following concept of entropic density in [1]. Here, a tuple $(X_1, \ldots, X_r)$ of random variables is *symmetric* if the distributions of $(X_{\sigma(1)}, \ldots, X_{\sigma(r)})$ are the same for all permutations σ of $[r]$.

Definition 2.1 (Entropic density, ratio sequence). Let H be an r-uniform hypergraph.

- A tuple of random vertices $(X_1, \ldots, X_r) \in V(H)^r$ is a *random edge with uniform ordering* on H if $(X_1, \ldots, X_r)$ is symmetric and $\{X_1, \ldots, X_r\}$ is always an edge of H.
- The *entropic density* of H, denoted by $b_{\mathrm{entropy}}(H)$, is the maximum of $2^{\mathbb{H}(X_1,\ldots,X_r) - r\mathbb{H}(X_1)}$ taken over all random edges $(X_1, \ldots, X_r)$ with uniform ordering on H.
- The *ratio sequence* of a random edge $(X_1, \ldots, X_r)$ with uniform ordering is the sequence $(x_1, \ldots, x_r)$ given by $x_i = 2^{\mathbb{H}(X_i | X_{i+1}, \ldots, X_n) - \mathbb{H}(X_i)}$ for each $i \in [r]$.

A key result of Chao and Yu [1] states that these two notions of densities are equivalent.

Proposition 2.2 ([1, Proposition 5.4]). $b(H) = b_{\mathrm{entropy}}(H)$ *for any r-uniform hypergraph H.*

The following lemma follows directly from Chao and Yu's proof of Lemma 7.2 in [1].

Lemma 2.3. *Let $1 \le k \le \lfloor r/2 \rfloor$, let H be an $\mathcal{F}_{r,k}$-hom-free r-uniform hypergraph, and let $(X_1, \dots, X_r)$ be a random edge with uniform ordering on H whose ratio sequence is $0 < x_1 \le \dots \le x_r = 1$. Then, $x_i + x_j \le x_{i+j}$ for every $i \in [k]$ and $i + j \le n$.*

With Lemma 2.3 in mind, for every $1 \le k \le \lfloor r/2 \rfloor$, define $\mathcal{X}_{r,k} \subset [0,1]^r$ to be

$$\mathcal{X}_{r,k} = \{0 < x_1 \le \dots \le x_r = 1 \mid x_i + x_j \le x_{i+j} \text{ for every } i \in [k] \text{ and } i + j \le n\}.$$

It now suffices to prove the following optimisation result.

Theorem 2.4. *For every $r \ge 4$, let $k = \lceil r/e \rceil$, then*

$$\max\left\{\prod_{i=1}^{r} x_i \;\middle|\; (x_1, \dots, x_r) \in \mathcal{X}_{r,k}\right\} = \frac{r!}{r^r},$$

and the unique tuple achieving the maximum is given by $x_i = i/r$ for every $i \in [r]$.

Indeed, we first show that Theorem 2.4 implies Theorem 1.3.

Proof of Theorem 1.3. Let H be an $\mathcal{F}_{r,k}$-hom-free r-uniform hypergraph. Let $(X_1, \dots, X_r)$ be any random edge with uniform ordering on H. Using well-known properties of entropy, one can verify that the ratio sequence satisfies $\prod_{i=1}^{r} x_i = 2^{\mathbb{H}(X_1,\dots,X_r) - r\mathbb{H}(X_1)}$. It then follows from the three results above that $b(H) \le r!/r^r$, so $\pi(\mathcal{F}_{r,k}) \le r!/r^r$. The reverse inequality follows as the r-uniform hypergraph consisting of just one edge has blowup density $r!/r^r$. □

We now provide a sketch of the proof of Theorem 2.4.

Proof of Theorem 2.4 (sketch). Since $(x_1, \dots, x_r) \in \mathcal{X}_{r,k}$, we have $rx_1 \le x_r = 1$, so say $x_1 = (1 - \varepsilon)/r$ for some $\varepsilon > 0$. Also, for every $k + 1 \le i \le r$, we have $x_i \le 1 - (r - i)x_1 = (i + (r - i)\varepsilon)/r$. If additionally $x_i = (1 - \varepsilon)i/r$ for every $i \in [k]$, then we can use $\log(1 + t) \le t$ and a Riemann sum bound to show that

$$\log\left(\frac{\prod_{i=1}^{r} x_i}{\prod_{i=1}^{r} \frac{i}{r}}\right) \le \varepsilon\left(-k + r\int_{\frac{k}{r}}^{1} \frac{1-t}{t}\mathrm{d}t\right) = \varepsilon r(\log r - \log k - 1) \le 0,$$

with equality if and only if $\varepsilon = 0$, as required.

Otherwise, there exists $j \in [k]$ with $x_j > jx_1 = (1 - \varepsilon)j/r$. Then, for some small $\delta > 0$, we carefully construct a new sequence $(x_1', \dots, x_r')$ with $|x_i' - x_i| \le \delta$ for all $i \in [r]$ and $\prod_{i=1}^{r} x_i' > \prod_{i=1}^{r} x_i$. With a technical case analysis, we also show that $(x_1', \dots, x_r') \in \mathcal{X}_{r,k}$, thus $(x_1, \dots, x_r)$ is not optimal. □

Additionally, the choice of $k = \lceil r/e \rceil$ in Theorem 2.4 is essentially best possible.

Lemma 2.5. *For every $r \ge 2$, if $1 \le k < \lfloor r/e \rfloor$, then there exists $(x_1, \dots, x_r) \in \mathcal{X}_{r,k}$ satisfying $\prod_{i=1}^{r} x_i > r!/r^r$.*

Proof (sketch). Let $\varepsilon > 0$ be sufficiently small. Let $x_i = (1 - \varepsilon)i/r$ for every $i \in [k]$, and $x_i = (i + (r - i)\varepsilon)/r$ for every $k + 1 \le i \le r$. It is easy to verify that $(x_1, \dots, x_r) \in \mathcal{X}_{r,k}$, and we use a similar integral bound as in the proof above to show that $\prod_{i=1}^{r} x_i > r!/r^r$. □

3 The exact Turán number result

Our proof of Theorem 1.4 follows Liu's proof of Theorem 1.2 in [7] and uses the powerful framework introduced in [8] and later refined in [4] to prove strong stability results in Turán problems. For brevity, we highlight only on the following concept of degree-stability.

Definition 3.1. Let $\mathcal{F}$ be a family of r-uniform hypergraphs with $\pi(\mathcal{F}) > 0$, and let $\mathfrak{H}$ be a family of $\mathcal{F}$-free r-uniform hypergraphs. $\mathcal{F}$ is *degree-stable* with respect to $\mathfrak{H}$ if there exist $\varepsilon > 0$ and N such that every $\mathcal{F}$-free r-uniform hypergraph H on $n \geq N$ vertices with $\delta(H) \geq (\pi(\mathcal{F})/(r-1)! - \varepsilon)\, n^{r-1}$ is in $\mathfrak{H}$.

The key intermediate result is the following, which we obtain by adapting Liu's proof in [7] of his corresponding result with $\mathcal{F}_{r,\lfloor r/2 \rfloor}$ in place of $\mathcal{F}_{r,k}$. Since $\mathcal{F}_{r,k}$ is not blowup-invariant, we need to prove the following result for a broader family $\mathcal{T}_{r,k}$ first, then apply the Hypergraph Removal Lemma to complete the proof.

Proposition 3.2. *For every $r \geq 4$ and $k = \lceil r/e \rceil$, $\mathcal{F}_{r,k}$ is degree-stable with respect to $\mathfrak{K}_r^r$, the family of r-partite r-uniform hypergraphs.*

Now we sketch how Theorem 1.4 follows from this degree-stability result.

Proof of Theorem 1.4 (sketch). Let H be an $\mathcal{F}_{r,k}$-free r-uniform hypergraph on n vertices with $|E(H)| \geq |E(T^r(n))|$ and n sufficiently large. Repeatedly delete vertices with low degrees as long as they exist. This process cannot go on for too long as otherwise we obtain an $\mathcal{F}_{r,k}$-free hypergraph $\overline{H}$ on $\overline{n}$ vertices with $|E(\overline{H})|$ significantly larger than $|E(T^r(\overline{n}))|$, contradicting Theorem 1.3. Suppose H' on n' vertices is what we obtain at the end of this process, then using the degree-stability given by Proposition 3.2, H' must be r-partite, so $|E(H')| \leq |E(T^r(n')|$. But this can only happen if $H = H'$, so H itself is r-partite and $|E(H)| \leq |E(T^r(n))|$, with equality if and only if $H = T^r(n)$. □

4 Concluding Remarks

By Lemma 2.5, it is not possible to use this entropy method to show $\pi(\mathcal{F}_{r,k}) = r!/r^r$ for $k < \lfloor r/e \rfloor$. It is natural then to ask the following question.

Question 4.1. *Is $\pi(\mathcal{F}_{r,k}) > r!/r^r$ if $k < \lfloor r/e \rfloor$?*

Our main result also applies to the following two related problems.

There is a more general definition of a λ-*tent* Δ_λ for every partition λ of r, with our earlier definition corresponding to partitions of size 2. Under this definition, it follows from Theorem 1.3 that for every $r \geq 4$ and $k = \lceil r/e \rceil$, $\pi(\Delta_{(r-k,1,\ldots,1)}) = r!/r^r$.

For $L \subset \{0, 1, \ldots, r-1\}$, an r-uniform hypergraph H is *L-intersecting* if $|e \cap e'| \in L$ for any distinct edges e, e' in H. As a consequence of Theorem 1.3, for every $r \geq 4$, if $k = \lceil r/e \rceil$ and $L \subset \{0, 1, \ldots, r-k-1\}$, then $b(H) \leq r!/r^r$ for every L-intersecting r-uniform hypergraph H. This notion also provides a potential approach to Question 4.1, as if $k < \lfloor r/e \rfloor$, $L = \{0, 1, \ldots, r-k-1\}$ and there exists an L-intersecting r-uniform hypergraph H with $b(H) > r!/r^r$, then this would imply that $\pi(\mathcal{F}_{r,k}) \geq b(H) > r!/r^r$.

References

[1] Ting-Wei Chao and Hung-Hsun Hans Yu. When entropy meets Turán: new proofs and hypergraph Turán results. *arXiv:2412.08075*, 2024.

[2] Peter Frankl and Zoltán Füredi. A new generalization of the Erdős-Ko-Rado theorem. *Combinatorica*, 3(3-4):341–349, 1983.

[3] Peter Frankl and Zoltán Füredi. Extremal problems whose solutions are the blowups of the small Witt-designs. *Journal of Combinatorial Theory. Series A*, 52(1):129–147, 1989.

[4] Jianfeng Hou, Xizhi Liu, and Hongbin Zhao. A criterion for Andrásfai–Erdős–Sós type theorems and applications. *arXiv:2401.17219*, 2024.

[5] Daniel Iľkovič and Jun Yan. An improved hypergraph Mantel's Theorem. *arXiv:2503.14474*, 2025.

[6] Peter Keevash. Hypergraph Turán problems. *Surveys in Combinatorics*, 392:83–140, 2011.

[7] Xizhi Liu. On a hypergraph Mantel theorem. *arXiv:2501.19229*, 2025.

[8] Xizhi Liu, Dhruv Mubayi, and Christian Reiher. A unified approach to hypergraph stability. *Journal of Combinatorial Theory. Series B*, 158(2):36–62, 2023.

[9] Oleg Pikhurko. An exact Turán result for the generalized triangle. *Combinatorica*, 28(2):187–208, 2008.

MULTI-CYCLIC GRAPHS IN THE RANDOM GRAPH PROCESS WITH RESTRICTED BUDGET

DANIEL IĽKOVIČ, JARED LEÓN, AND XICHAO SHU

ABSTRACT. We study a controlled random graph process introduced by Frieze, Krivelevich, and Michaeli. In this model, the edges of a complete graph are randomly ordered and revealed sequentially to a builder. For each edge revealed, the builder must irrevocably decide whether to purchase it. The process is subject to two constraints: the number of observed edges t and the builder's budget b. The goal of the builder is to construct, with high probability, a graph possessing a desired property.

Previously, the optimal dependencies of the budget b on n and t were established for constructing a graph containing a fixed tree or cycle, and the authors claimed that their proof could be extended to any unicyclic graph. The problem, however, remained open for graphs containing at least two cycles, the smallest of which is the graph K_4^- (a clique of size four with one edge removed).

In this paper, we provide a strategy to construct a copy of the graph K_4^- if $b \gg \max\left\{n^6/t^4, n^{4/3}/t^{2/3}\right\}$, and show that this bound is tight, answering the question posed by Frieze et al. concerning this specific graph. We also give a strategy to construct a copy of a graph consisting of k triangles intersecting at a single vertex (the k-fan) if $b \gg \max\left\{n^{4k-1}/t^{3k-1}, n/\sqrt{t}\right\}$, and also show that this bound is tight. These are the first optimal strategies for constructing a multi-cyclic graph in this random graph model.

1. INTRODUCTION

Random graph processes have garnered significant attention in recent years, as they offer a natural framework for modeling the evolution of complex networks over time. These processes do not only provide a straightforward, and in some cases the only, method for constructing counterexamples in extremal graph theory but also exhibit strong connections to algorithmic graph theory. In particular, the analysis of random graph processes can lead to valuable insights into the typical running times of algorithms.

Random graph processes can be defined in various ways, with one of the most well-known and the simplest being the Erdős–Rényi random graph process, introduced by Erdős and Rényi [11, 12]. This process begins with an empty graph on n vertices. At each step, an edge is chosen uniformly at random from the set of non-edges and is added to the graph.

At any fixed time t, the random graph process is distributed as a uniform random graph $G(n,t)$. A *graph property* is a family of graphs that is invariant under isomorphisms. Such a property is called *monotone* if it remains preserved under the addition of edges. The *hitting time* for a monotone property P is defined as the random variable $t_P =$

FACULTY OF INFORMATICS, MASARYK UNIVERSITY, BOTANICKÁ 68A, 602 00 BRNO, CZECH REPUBLIC.

MATHEMATICS INSTITUTE AND DIMAP, UNIVERSITY OF WARWICK, COVENTRY, UK

E-mail addresses: 493343@mail.muni.cz, jared.leon@warwick.ac.uk, 540987@mail.muni.cz.

(DI, XS): Research supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University.

(JL): Research supported by the Warwick Mathematics Institute Centre for Doctoral Training.

$\min\{t \in [N] : G(n,t) \text{ satisfies } P\}$. A significant body of research focuses on identifying the hitting time for various monotone graph properties in the random graph process.

1.1. **Random graph process with restricted budget.** Beyond the classical binomial random graph model, also known as the Erdős–Rényi random graph, numerous fascinating variants have been investigated. Examples include the H-free process [9, 13, 22, 27], with particular attention given to the triangle-free process [7, 10, 14], the more recently studied model of multi-source invasion percolation on the complete graph [1], Achlioptas processes proposed by Dimitris Achlioptas [6, 8, 24, 25, 26], and the semi-random graph process suggested by Peleg Michaeli [5], and studied recently in [3, 4, 16, 17, 18, 19, 23].

Recently, Frieze, Krivelevich, and Michaeli [15] introduced the following controlled random graph process. Consider a Builder who observes the edges of a random permutation $e_1, \ldots, e_N$ one at a time and must make an irrevocable decision whether to select e_i upon seeing it. A (t,b)-strategy is an online algorithm (either deterministic or randomized) that Builder employs in this setting, where he only observes the first t edges in the permutation and is allowed to select at most b of them. The Builder's goal is to build a graph that has a specified graph property. For instance, there exists a simple $(t_{\text{con}}, n-1)$-strategy ensuring that Builder's graph becomes connected at the hitting time for connectivity: the Builder selects an edge only if it reduces the number of connected components in the graph.

In this model, the problems of constructing spanning structures such as a perfect matching, a Hamilton cycle, a graph with a prescribed minimum degree, or a graph with a fixed level of connectivity have been extensively studied in [2, 15, 20, 21]. It was demonstrated that for these properties, a linear edge budget suffices to achieve the desired structure at or shortly after the hitting time. This contrasts sharply with the superlinear hitting time required so that there are no isolated vertices in the graph, which, as is well known, occurs at $\left(\frac{1}{2}+o(1)\right) n \log n$.

Much less is known about the construction of small subgraphs, specifically those whose number of vertices remains bounded independently of the size of the host graph. The problem of constructing trees and unicyclic graphs was addressed in [15]. Notably, Frieze et al. [15] established the following results.

Theorem 1.1. *Let $k \geqslant 3$ be an integer and let T be a k-vertex tree. If $t \geqslant b \gg \max\{(n/t)^{k-2}, 1\}$ then there exists a (t,b)-strategy B of Builder such that*

$$\lim_{n\to\infty} \mathbb{P}(T \subseteq B_t) = 1$$

and if $b \ll (n/t)^{k-2}$ then for any (t,b)-strategy B of Builder,

$$\lim_{n\to\infty} \mathbb{P}(T \subseteq B_t) = 0.$$

Theorem 1.2. *Let $k \geqslant 1$ be an integer and let $H = C_{2k+1}$ or $H = C_{2k+2}$. Write $b^* = b^*(n,t,k) = \max\{n^{k+2}/t^{k+1}, n/\sqrt{t}\}$. If $t \gg n$ and $b \gg b^*$ then there exists a (t,b)-strategy B of Builder such that*

$$\lim_{n\to\infty} \mathbb{P}(H \subseteq B_t) = 1,$$

and if $t \ll n$ or $b \ll b^*$ then for any (t,b)-strategy B of Builder,

$$\lim_{n\to\infty} \mathbb{P}(H \subseteq B_t) = 0.$$

Theorems 1.1 and 1.2 analyze the budget thresholds required for constructing trees and cycles. Extending these results to any fixed unicyclic graph is straightforward: once the cycle is formed, the remaining forest can be efficiently built with a constant budget. Consequently, the smallest graph not covered by their results is the graph on four vertices and five edges, K_4^- (the diamond). A notable challenge is that several natural strategies lead to different upper bounds, each one optimal within distinct time regimes. A similar issue arises in the case of constructing cycles.

1.2. **Our results.** In this paper, we discuss optimal strategies for constructing small subgraphs with multiple cycles. Firstly, we present an optimal strategy for the first unresolved case outlined in [15], the graph K_4^-.

Theorem 1.3. *Let $b^* = b^*(n,t,k) = \max\{n^6/t^4,\ n^{4/3}/t^{2/3}\}$. If $t = \omega\left(n^{6/5}\right)$ and $b = \omega\left(b^*\right)$, then there exists a (t,b)-strategy B of Builder such that*

$$\lim_{n\to\infty} \mathbb{P}\left(K_4^- \subset B_t\right) = 1,$$

and if $t = o\left(n^{6/5}\right)$ or $b = o\left(b^*\right)$ then for any (t,b)-strategy B of Builder,

$$\lim_{n\to\infty} \mathbb{P}\left(K_4^- \subset B_t\right) = 0.$$

Proof sketch. We begin by showing the first part of the theorem by designing the algorithm. Firstly, let us consider the case where $n^{7/6} \ll t = O\left(n^{7/5}\right)$. In this case we have that $b \gg b^* = n^6/t^4$. Let $T = t/3$, and let $r = \omega\left(n^7/T^5\right)$ such that $rt/n = o(b)$.

The strategy in this case is to select a set of r vertices and then reveal the first T edges of the graph, buying every edge that intersects these selected vertices as long as the neighborhoods have size at most $\Theta(t/n)$. Thus, we obtain a set of neighborhoods that cover a large number of edges of the graph, having purchased at most $o(b)$ edges. We then reveal the next T edges and buy at most $b/2$ edges contained in these neighborhoods, obtaining $\omega\left(n^2/T\right)$ triangles. In the last step, we reveal the remaining T edges, at least one of which will, with high probability, complete the graph K_4^-.

In the second case, we now assume that $t = \omega\left(n^{7/5}\right)$. For this case we have that $b \gg b^* = n^{4/3}/t^{2/3}$. Let $T = t/2$.

Similarly to the previous case, we now fix a vertex and, while revealing the first T edges, we buy each incident edge, obtaining a neighborhood of size $O(T/n)$ and having purchased at most $b/2$ edges. We then reveal the remaining T edges, and buy every edge contained in the neighborhood; if with high probability two of them are non-disjoint, we construct a copy of K_4^-.

Now we proceed to analyze the second part of the theorem. If $t \ll n^{6/5}$, the graph G_t w.h.p. does not contain the graph K_4^-. Therefore, suppose that $t \gg n^{6/5}$, and that $b \ll b^*$, where $b^* = \max\{n^6/t^4, n^{4/3}/t^{2/3}\}$. As in the 1-statement, we will split the analysis into two cases depending on t.

The whole analysis of the 0-statement is based on the following claim.

Claim 1.4. *With high probability, G contains at most cbt^2/n^3 triangles, for some constant c.*

Let us first consider the case where $n^{6/5} \ll t \ll n^{7/5}$. In this case we have that $b \ll b^* = n^6/t^4$. We will now prove that there is no (t,b)-strategy that generates the graph K_4^-.

Clearly, we may assume that $b = \omega\left(t/n\right)$, as increasing the value of b makes the claim stronger. Let $G \subset G_t$ be a spanning subgraph with b edges of the random graph G_t. We will first prove that, w.h.p. the number of copies of C_4 and K_3^+ in G is $O(bt^3/n^4)$. Assuming this, note that these are the only graphs that can be formed by taking the graph K_4^- and removing one edge from it. Then, G contains $O(bt^3/n^4)$ candidate pairs of vertices that generate the graph K_4^-, and the probability of hitting one of these edges by revealing only one edge is at most $O\left(bt^3/n^6\right) = o\left(1/t\right)$. Hence, by revealing $O(t)$ additional edges, w.h.p. none of these edges hit any of the candidate pairs, and the claim follows. Thus, it only remains to show that w.h.p. the number copies of C_4 and K_3^+ in G is $O(bt^3/n^4)$.

We start by upper-bounding the number of copies of the graph C_4 in G. We start by noticing that there are at most $O(bn/t) \cdot O(t^2/n^2) = O(bt/n)$ copies of P_3 in G. We can only generate a copy of P_4 by extending a copy of P_3 or by joining two disjoint edges with a third. We analyze these two cases separately.

For a vertex v, let Y_v be the random variable that counts the number of copies of P_3 with end-vertex v. We have that $\mathbb{E}Y_v = \Theta((bt/n)/n) = \Theta(bt/n^2)$. Let $X_1, \ldots, X_t$ be independent copies of a random variable that, when revealing a random edge on top of G (possibly an existing one), counts the number of new copies of the graph P_4 formed by extending a copy of P_3. Given that any copy of P_3 can be extended to a copy of P_4 by adding an edge to either of its end-vertices, we have $\mu = \mathbb{E}X_i = \Theta(\mathbb{E}Y_v)$. We will use Bennet inequality to bound the number of copies of P_4. Let $S_t = \sum_{i=1}^t X_i - \mu$. Then, since every vertex is an end-vertex of at most $O(t^2/n^2)$ copies of P_3, we have that, for all i, $|X_i - \mu| \leqslant ct^2/n^2$, for some constant c. Let $a = ct^2/n^2$. Given that, w.h.p. the random variable X_i is bounded from both sides, we have that $\sigma^2 = \sum_{i=1}^t \mathbb{E}(X_i - \mu)^2 \leqslant at\mu$. Thus we arrive to the conclusion that w.h.p. we have that $S_t < c_4 t\mu$, for a large enough constant c_4.

Let $q = \sum_{i=1}^t X_i$. Given that $c_4 t\mu > S_t = q - t\mu$, we have that $q < (c_4+1)t\mu$. Therefore, there are at most $(c_4+1)t\mu = (c_4+1) \cdot bt^2/n^2$ copies of P_4 that extend a copy of P_3. We now consider copies of P_4 created by joining two disjoint edges with a third edge.

The expectation of generating a copy of P_4 by joining two edges can be upper bounded by the square of the number of endpoints of the present edges. Similarly as in the previous case, using Bennet inequality, we get that the number of P_4 generated in this case is negligible compared to the previous case. We conclude that G contains at most $O(bt^2/n^2)$ copies of P_4. Finally, we let $Z_1, \ldots, Z_t$ be independent copies of the same random variable that, when revealing a random edge on top of G, counts the number of new copies of the graph C_4. Note that each copy of P_4 generates exactly one possible candidate pair of vertices for generating a copy of C_4. Since $\mu_3 = \mathbb{E}Z_i \leqslant c_5 bt^2/n^2 \cdot 1/n^2 = c_5 bt^2/n^4$, after applying Bennet inequality, there are at most $c_6 t\mu_3 = O(bt^3/n^4)$ copies of C_4 in G.

Now we bound the number of copies of K_3^+ in G. Since $b = \omega(n^3/t^2)$, by Claim 1.4 the number of copies of K_3 in G is at most $O(bt^2/n^3)$. Then, as every vertex of G has degree $O(t/n)$, every copy of a triangle is part of at most $O(t/n)$ copies of K_3^+. Hence, G contains $O(bt^3/n^4)$ copies of K_3^+.

In the second case of the 0-statement we have that $t = \omega\left(n^{7/5}\right)$. In this case, we also have that $b \ll b^* = n^{4/3}/t^{2/3}$. Let G be a spanning subgraph of G_t with b edges. Similar to the previous case, we bound the number of copies of C_4 and K_3^+, arriving at the desired conclusion. □

We also consider another kind of multi-cyclic graph. For an integer $k \geqslant 1$, let T_k be the graph consisting on k triangles intersecting in a single vertex. We obtained the following result.

Theorem 1.5. *Let $k \geqslant 1$ be an integer, and let $b^* = b^*(n,t,k) = \max\{n^{4k-1}/t^{3k-1}, n/\sqrt{t}\}$. If $t = \omega\left(n^{\frac{4k-1}{3k}}\right)$ and $b = \omega(b^*)$, then there exists a (t,b)-strategy B of Builder such that*

$$\lim_{n\to\infty} \mathbb{P}(T_k \subset B_t) = 1,$$

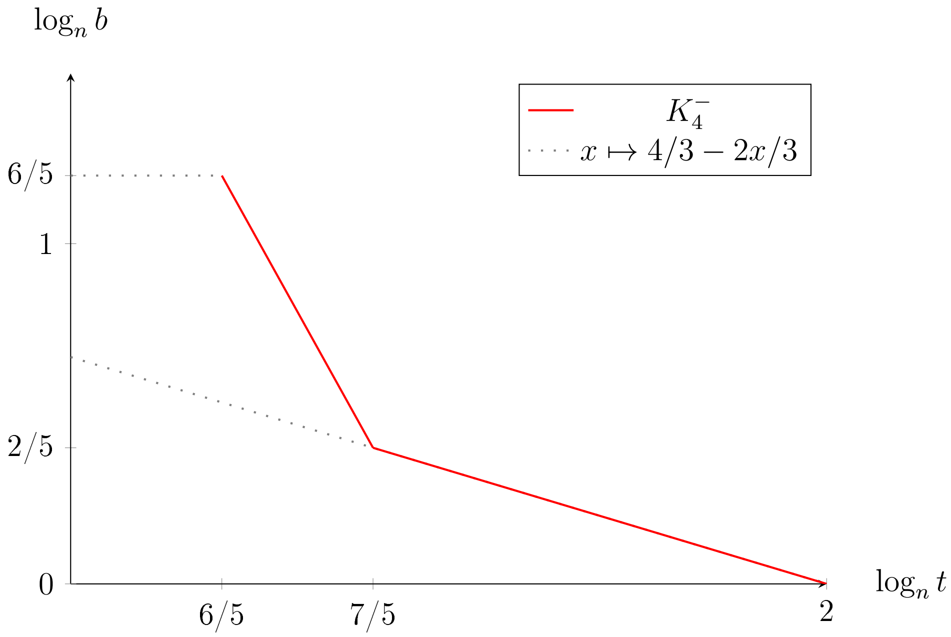


FIGURE 1. Budget threshold for K_4^-.

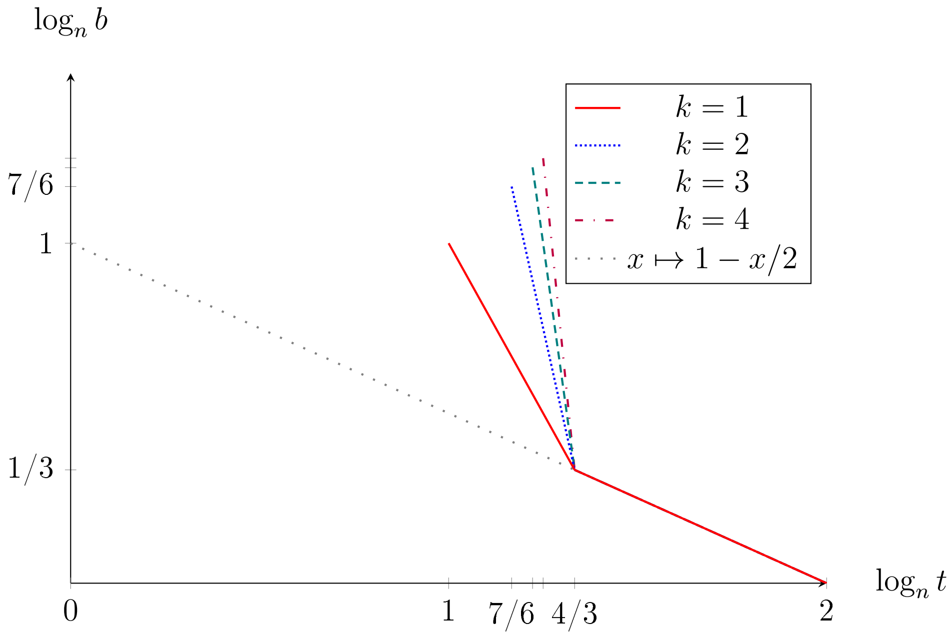


FIGURE 2. Budget thresholds for T_k.

and if $t = o\left(n^{\frac{4k-1}{3k}}\right)$ *or* $b = o\left(b^*\right)$ *then for any* (t, b)*-strategy* B *of Builder,*

$$\lim_{n\to\infty} \mathbb{P}\left(T_k \subset B_t\right) = 0.$$

We present in Figures 1 and 2 visualizations of the dependencies between the time t and minimum threshold for a budget b in terms of n. For better presentation, we scale them logarithmically. Also, notice that for the case $k = 1$, the graph T_k is a triangle, and our formulas generalize the result in [15] for this case.

References

[1] L. Addario-Berry and J. Barrett: *Multisource invasion percolation on the complete graph*, Ann. Probab. **51** (2023), 2131–2157.

[2] M. Anastos: *Constructing hamilton cycles and perfect matchings efficiently* (2022), 2209.09860.

[3] N. Behague, T. G. Marbach, P. Prałat and A. Ruciński: *Subgraph games in the semi-random graph process and its generalization to hypergraphs*, Electron. J. Combin. **31** (2024), Paper No. 3.30, 18.

[4] O. Ben-Eliezer, L. Gishboliner, D. Hefetz and M. Krivelevich: *Very fast construction of bounded-degree spanning graphs via the semi-random graph process*, Random Structures Algorithms **57** (2020), 892–919.

[5] O. Ben-Eliezer, D. Hefetz, G. Kronenberg, O. Parczyk, C. Shikhelman and M. s. Stojaković: *Semi-random graph process*, Random Structures Algorithms **56** (2020), 648–675.

[6] S. Bhamidi, A. Budhiraja and X. Wang: *The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs*, Probab. Theory Related Fields **160** (2014), 733–796.

[7] T. Bohman: *The triangle-free process*, Adv. Math. **221** (2009), 1653–1677.

[8] T. Bohman and A. Frieze: *Avoiding a giant component*, Random Structures Algorithms **19** (2001), 75–85.

[9] T. Bohman and P. Keevash: *The early evolution of the H-free process*, Invent. Math. **181** (2010), 291–336.

[10] T. Bohman and P. Keevash: *Dynamic concentration of the triangle-free process*, Random Structures Algorithms **58** (2021), 221–293.

[11] P. Erdős and A. Rényi: *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.

[12] P. Erdős and A. Rényi: *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61.

[13] P. Erdős, S. Suen and P. Winkler: *On the size of a random maximal graph*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993) (1995), 309–318.

[14] G. Fiz Pontiveros, S. Griffiths and R. Morris: *The triangle-free process and the Ramsey number $R(3,k)$*, Mem. Amer. Math. Soc. **263** (2020), v+125.

[15] A. Frieze, M. Krivelevich and P. Michaeli: *Fast construction on a restricted budget* (2023), 2207.07251.

[16] D. Gamarnik, M. Kang and P. Prałat: *Cliques, chromatic number, and independent sets in the semi-random process*, SIAM J. Discrete Math. **38** (2024), 2312–2334.

[17] P. Gao, B. Kamiński, C. MacRury and P. Prałat: *Hamilton cycles in the semi-random graph process*, European J. Combin. **99** (2022), Paper No. 103423, 24.

[18] P. Gao, C. MacRury and P. Prałat: *A fully adaptive strategy for Hamiltonian cycles in the semi-random graph process*, in: Approximation, randomization, and combinatorial optimization. Algorithms and techniques, *LIPIcs. Leibniz Int. Proc. Inform.*, volume 245 (2022), Art. No. 29, 22.

[19] S. Gilboa and D. Hefetz: *Semi-random process without replacement*, J. Comb. **14** (2023), 167–196.

[20] K. Katsamaktsis and S. Letzter: *Building graphs with high minimum degree on a budget* (2024), 2401.15812.

[21] L. Lichev: *d-connectivity of the random graph with restricted budget*, J. Graph Theory **108** (2025), 293–312.

[22] D. Osthus and A. Taraz: *Random maximal H-free graphs*, Random Structures Algorithms **18** (2001), 61–82.

[23] P. Prałat and H. Singh: *Power of k choices in the semi-random graph process*, Electron. J. Combin. **31** (2024), Paper No. 1.11, 23.

[24] O. Riordan and L. Warnke: *Achlioptas process phase transitions are continuous*, Ann. Appl. Probab. **22** (2012), 1450–1464.

[25] O. Riordan and L. Warnke: *The phase transition in bounded-size achlioptas processes* (2017), 1704.08714.

[26] J. Spencer and N. Wormald: *Birth control for giants*, Combinatorica **27** (2007), 587–628.

[27] L. Warnke: *Dense subgraphs in the H-free process*, Discrete Math. **311** (2011), 2703–2707.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Hitting times and the power of choice for random geometric graphs

(Extended abstract)

Dawid Ignasiak* Lyuben Lichev†

Abstract

We consider a random geometric graph process where random points $(X_i)_{i\geq 1}$ are embedded consecutively in the d-dimensional unit torus $\mathbb{T}^d$, and every two points at distance at most r form an edge. As $r \to 0$, we confirm that well-known hitting time results for connectivity and Hamiltonicity in the Erdős-Rényi graph process also hold for its geometric analogue. Moreover, we exhibit a sort of probabilistic monotonicity for each of these properties.

We also study a geometric analogue of the power of choice where, at each step, an agent is shown two independent random points sampled from $\mathbb{T}^d$ and is allowed to choose one of them. When the agent is allowed to make their choice with the knowledge of the entire sequence of random points (offline 2-choice), we show that they can construct a connected graph at the first time t when none of the first t pairs contains two isolated vertices in the graph induced by $(X_i)_{i=1}^{2t}$, and maintain connectivity thereafter. We also derive similar results for Hamiltonicity. This shows that each of the two properties can be attained two times faster (time-wise) and with four times fewer points in the offline 2-choice model compared to its 1-choice analogue.

In the online version where the agent only knows the process until the current time step, we exhibit rather sharp thresholds for connectivity and Hamiltonicity. The online 2-choice model does not significantly accelerate the two properties (time-wise) but allows to realise them on twice as few points compared to its 1-choice analogue.

1 Introduction

The study of dynamic random graphs and networks is a central topic in probability theory and combinatorics. In the world of binomial random graphs where randomness is carried

*TU Wien, Faculty of Mathematics and Geoinformation, Institute of Statistics and Mathematical Methods in Economics, Research Unit of Mathematical Stochastics, Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria. E-mail: dawid.ignasiak@tuwien.ac.at.

†TU Wien, Faculty of Mathematics and Geoinformation, Institute of Statistics and Mathematical Methods in Economics, Research Unit of Mathematical Stochastics, Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria. E-mail: lyuben.lichev@tuwien.ac.at. This research was supported by the Austrian Science Fund (FWF) grant No. 10.55776/ESP624.

by the set of edges, the corresponding random process is defined naturally via ordering the edges uniformly at random and sequentially including them in the graph. A classic geometric analogue of the above process consists in sampling independently n random points in a compact metric space (usually the unit hypercube $[0,1]^d$ or its boundary-free analogue, the unit torus $\mathbb{T}^d$), ordering edges according to their length and including them sequentially in the graph, starting with the shortest one. Penrose [6, 7] showed that the dynamic random geometric graph on $[0,1]^d$ or $\mathbb{T}^d$ for $d \geq 1$ becomes connected when the last isolated vertex disappears. Díaz, Mitsche, and Pérez [2] further derived a sharp threshold result for Hamiltonicity for the random geometric graph on $[0,1]^2$. Their theorem was later strengthened to a hitting time result in every dimension $d \geq 1$ by Balogh, Bollobás, Krivelevich, Müller and Walters [1] and Muller, Pérez-Giménez and Wormald [5]. For many more related results, we recommend the book of Penrose [7].

In this work, we take a different perspective and define the dynamic random geometric graph via a sequence of random points $(X_i)_{i\geq 1}$ embedded consecutively in the torus $\mathbb{T}^d$.[1] Contrary to [1, 2, 5, 6], our asymptotic parameter is the radius of connectivity $r \to 0$. This alternative point of view allows to see the model as a discrete-time stochastic process and brings in related natural questions. Some of the latter may be seen as 'dual' to the questions considered in the dynamic model on a fixed point set: for example, we provide analogues of the hitting time results for connectivity and Hamiltonicity from [1, 5, 6] for every dimension $d \geq 1$. We observe that, unlike in the dynamic model on a fixed point set, connectivity and Hamiltonicity are *not* monotone properties in our case: for example, an isolated point can appear in the presence of a connected or a Hamiltonian graph. We exhibit a sort of probabilistic monotonicity showing that the said scenario is atypical.

Other natural questions seem to have no clear analogues in the model on a fixed point set. For example, the adopted discrete-time point of view allows us to analyse the model through the lens of the classic *power of choice* paradigm introduced by Mitzenmacher [4]. More precisely, we consider a setting where an *agent* aims to construct a graph with a certain property. At each round, the agent is presented with two uniformly chosen points and is allowed to include one of them in the graph either based on complete information for the sequence $(X_i)_{i\geq 1}$ (offline 2-choice) or only based on the history of the process (online 2-choice). In this work, we analyse how fast the agent can attain a connected or a Hamiltonian geometric graph in the above setting compared to the original (1-choice) process. In the offline 2-choice process, we show that with high probability (whp), the agent is able to build a connected graph by seeing roughly half of the points needed in the 1-choice process, and using roughly a quarter of the points in the 1-choice process. The online 2-choice process improves the number of points required by a factor of 2 but the agent needs to play the game essentially until the connectivity threshold for the 1-choice process. Similar results are derived for Hamiltonicity.

Notation. Here, we gather relevant notation. Fix $d \geq 1$, a sequence $(X_i)_{i=1}^{\infty}$ of independent random points distributed uniformly in the d-dimensional unit torus $\mathbb{T}^d$ and a *radius of connectivity* $r > 0$. For every $i \geq 1$, the points X_{2i-1} and X_{2i} are called a *partner pair*. For a finite set of points S, we denote by $G(S, r)$ (or simply $G(S)$, when the radius is clear from the context) the geometric graph on vertex set S where every pair of points at (Euclidean) dis-

[1] Analogues of our results also hold for the unit cube $[0,1]^d$ with some modifications taking boundary effects into account.

tance at most r forms an edge. We also write G_t for $G((X_i)_{i=1}^t, r)$. Our asymptotic parameter is r: in particular, we work in the regime $r \to 0$.

Conditionally on the sequence $(X_i)_{i=1}^\infty$, a *choice set* is a subset U of the positive integers containing exactly one point from each partner pair. Fix $\mathbb{T}^* = \bigcup_{k=0}^\infty (\mathbb{T}^d)^k$. We consider two variations of the online 2-choice process. In the *simultaneous* variation, the points in every pair are revealed at once. Then, the agent is allowed to add one of the points in a partner pair x, y to the already constructed set W according to a *simultaneous choice function*, that is, a measurable function $f : \mathbb{T}^* \times (\mathbb{T}^d)^2 \to \{1, 2\}$ where $f(W, x, y) = 1$ indicates adding x to W and $f(W, x, y) = 2$ indicates adding y. In the *consecutive* 2-choice process, the agent needs to take a decision based on the already constructed set W and the first point in the current pair x, y according to a *consecutive choice function*, that is, a measurable function $g : \mathbb{T}^* \times \mathbb{T}^d \to \{1, 2\}$ where $g(W, x) = 1$ indicates adding x to W and $g(W, x) = 2$ indicates adding y. For every $t \geq 1$, a choice set U and a choice function f, we denote $\mathbf{X}_t(U) = \{X_i : 1 \leq i \leq 2t\} \cap U$ and $\mathbf{X}_t(f)$ the subset of $\{X_i : 1 \leq i \leq 2t\}$ produced by the choice function f.

2 Main results

Our first result analyses the time when the dynamic random geometric graph becomes connected or Hamiltonian. Define the hitting times

$$\tau_1^c = \tau_1^c(r) := \min\{t \geq 1 : G_t \text{ has no isolated vertex}\},$$
$$\tau_1^h = \tau_1^c(r) := \min\{t \geq 1 : G_t \text{ has no vertex of degree 0 or 1}\}.$$

Theorem 1. *Whp each of the graphs G_t with $t \geq \tau_1^c$ is connected, and each of the graphs G_t with $t \geq \tau_1^h$ is Hamiltonian.*

Our second result provides an analogue of Theorem 1 for the offline 2-choice process. Define $\tau_2^c = \tau_2^c(r)$ to be the smallest $t \geq 1$ such that, for every $i \in \{1, \ldots, t\}$, X_{2i-1} and X_{2i} are not simultaneously isolated in G_{2t}. Similarly, define $\tau_2^h = \tau_2^h(r)$ to be the smallest $t \geq 1$ such that, for every $i \in \{1, \ldots, t\}$, at least one of X_{2i-1} and X_{2i} has degree at least 2 in G_{2t}. A moment thought shows that the graph $G(\mathbf{X}_t(U))$ cannot be connected for any choice set U when $t < \tau_2^c$, and cannot be Hamiltonian when $t < \tau_2^h$. We show that these necessary conditions are actually sufficient.

Theorem 2. *Whp there is a choice set U such that the graph $G(\mathbf{X}_t(U))$ is connected for every $t \geq \tau_2^c$. Similarly, whp there is a choice set U such that the graph $G(\mathbf{X}_t(U))$ is Hamiltonian for every $t \geq \tau_2^h$.*

We turn to the online 2-choice process. First, we observe that a standard computation shows that whp $\tau_1^c \approx \tau_1^h$ are both equal to $C_d r^{-d} \log(1/r) + O(r^{-d} \log\log(1/r))$ for a dimension-dependant constant $C_d > 0$. The following sharp threshold result indicates that the qualitative behaviours of the simultaneous and the consecutive online 2-choice process do not differ much.

Theorem 3. *There exists a dimension-dependant constant C'_d such that whp, for every simultaneous choice function f, $G(\mathbf{X}_t(f))$ is disconnected for every $t \leq \tau_1^c - C'_d r^{-d} \log\log(1/r)$. Moreover, whp there is a consecutive choice function g such that $G(\mathbf{X}_t(g))$ is Hamiltonian for every $t \geq \tau_1^h + C'_d r^{-d} \log\log(1/r)$.*

Analogues of the above results for k-connectivity with $k \geq 2$ are provided in our original work [3] but related explanations are omitted for the sake of brevity.

3 Outline of some proof ideas

First of all, we outline the ideas behind Theorem 1; these bear certain similarities to the approach in [1, 5]. We concentrate on the connectivity property for the sake of cleaner exposition. First of all, by using our good knowledge for the value of τ_1^c, we stop the construction of the graph at time $t_1^c = \mathbb{E}[\tau_1^c] - h(r)/r^d$ for a function $h(r)$ tending to infinity suitably slowly. It is not hard to show that whp, for every $t \le t_1^c$, the graph G_t had an isolated vertex, and $G_{t_1^c}$ still has $\omega(1)$ of them. Moreover, by a result of Penrose [7] and our choice of h, whp the only vertices outside the giant component are isolated.

It remains to ensure that isolated vertices connect to the giant by time τ_1^c, and that no new isolated vertices ever appear. The first task is rather standard but the second is less routine: it requires to track the region which is out of reach for the giant component (later called *dangerous* region), and it could have a rather unpleasant shape. By using a tessellation $\mathcal{T}$ of the torus $\mathbb{T}^d$ with mesh $\Theta((r/\log(1/r))^d)$, we are able to show that whp the volume of the dangerous region is of the same order as the expected proportion of isolated vertices, and can be covered by a 'small' number of cubes in $\mathcal{T}$ (*dangerous* cubes). Finally, we guarantee that the isolated points and the dangerous cubes are typically far from each other and are surrounded by many vertices in the giant connected component at distance between r and $r + o(r)$ from themselves. As a result, whp isolated vertices become absorbed by the giant component by time τ_1^c, and the entire dangerous area becomes within reach of the giant before being visited for the first time, implying the connectivity statement in Theorem 1.

The proof of Theorem 2 carries some complications since $\tau_2^* \approx \tau_1^*/2$ for $* \in \{c, h\}$, meaning that there remain various small components disconnected from or only weakly connected to the giant at time τ_2^*. Luckily, these components have nice common features: they are well-clustered, have small diameter and entirely contain only a few partner pairs.

Again, we restrict our attention to the analysis for connectivity. We first consider a tessellation $\mathcal{T}$ of $\mathbb{T}^d$ into cubes of edge length cr for a small constant $c = c(d)$, and form an auxiliary graph H on $\mathcal{T}$ where cubes q, q' form an edge when $|x - y| \le r$ for all $x \in q$ and $y \in q'$. We call a cube *full* if it contains at least M points for some large constant $M = M(c, d)$: in particular, full cubes in H induce a giant connected component $\mathcal{C}$ covering a region R, and a small number of components of bounded size. Call $\bar{R}^c$ the region at distance at most r from the complement of R. By choice of M, we manage to find a set of points S such that every full square in $\mathcal{C}$ contains a unique point in S and every partner pair x, y with $x \in S$ satisfies $y \notin S \cup \bar{R}^c$. The set S is connected, well-spread and including it in the choice set U makes our decision for most other partner pairs x, y simple: namely, if $x \notin \bar{R}^c$ (so x is 'in the middle' of the well-connected set S), the position of y is enough to determine whether it should be added to U or not. By consecutively adding points from unused pairs at distance at most r from S to this set, and updating the component $\mathcal{C}$ (and the region R) by adding new visited cubes, we gradually shrink the 'bad region' $\bar{R}^c$. In turn, this simplifies the decision for a growing number of partner pairs. This procedure is iterated for a large but bounded number of steps. Finally, by starting from single partner pairs and tracking their way to the (initially fixed) set S, we manage to show that one can find a consistent family of paths ending in S and intersecting every partner pair exactly once.

Finally, the proof of Theorem 3 relies on two main ideas. First, we focus on the lower bound, denoted $t_0(r) = \tau_1^c - C_d' r^{-d} \log\log(1/r)$. Consider the 1-choice process at time $t_1(r) = r^d/\log\log(1/r)$. At this point, it is possible to find a set $S' \subseteq \{X_i : i \in [t_1]\}$ containing a strictly positive fraction of partner pairs X_{2i-1}, X_{2i} and such that the balls $(B(X_s, r))_{s\in S'}$ are pairwise disjoint. Moreover, we show that typically $\omega(1)$ of the pairs in S' have exactly one isolated vertex in the graph G_{t_0}. Conditionally on this fact and on the set of partner pairs in S' satisfying this condition, each point in such a pair has probability $1/2$ to be the isolated one, and this choice is done independently for different pairs. As a result, for any choice function f, the probability that it removes all the vertices isolated in G_{t_0} is at most $2^{-\omega(1)} = o(1)$.

The upper bound comes down to analysing a particular (consecutive) choice function g constructed as follows. First, g outputs 1 at each of the first $C''_d r^{-d} \log\log(1/r)$ steps for a suitable constant C''_d. Then, similarly to the 1-choice setting, we identify a set of full squares in a tessellation forming a giant component in an analogously defined auxiliary graph. This implies the existence of a dense, well-connected component where a Hamilton cycle can be constructed and flexibly modified to absorb new vertices, and some small 'islands' of total volume $(\log(1/r))^{-5}$, say, where vertices need to be absorbed 'from the outside'. From this point on, the function g includes the first point only if it belongs to some of these sparsely visited islands, and the second point in the partner pair otherwise. The upper bound follows from analysing g using tools from the proof of Theorem 1.

References

[1] J. Balogh, B. Bollobás, M. Krivelevich, T. Müller, and M. Walters. Hamilton cycles in random geometric graphs. *The Annals of Applied Probability*, pages 1053–1072, 2011.

[2] J. Díaz, D. Mitsche, and X. Pérez. Sharp threshold for Hamiltonicity of random geometric graphs. *SIAM J. Discrete Math.*, 21(1):57–65, 2007.

[3] D. Ignasiak and L. Lichev. Hitting times and the power of choice for random geometric graphs. in preparation.

[4] M. Mitzenmacher. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems*, 12(10):1094–1104, 2001.

[5] T. Müller, X. Pérez-Giménez, and N. Wormald. Disjoint hamilton cycles in the random geometric graph. *Journal of Graph Theory*, 68(4):299–322, 2011.

[6] M. D. Penrose. The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.*, 7(2):340–361, 1997.

[7] Mathew Penrose. *Random geometric graphs.* Oxford studies in probability. Oxford Univ. Press, Oxford, reprint. edition, 2004.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Enumeration of regular multipartite hypergraphs

(Extended abstract)

Mikhail Isaev* Tamás Makai† Brendan D. McKay ‡

Abstract

We determine the asymptotic number of regular multipartite hypergraphs, also known as multidimensional binary contingency tables, for all values of the parameters.

1 Introduction

Let $[n] := \{1, \ldots, n\}$ be a vertex set partitioned into r disjoint classes $V_1, \ldots, V_r$. We consider multipartite r-uniform hypergraphs such that every edge has exactly one vertex in each class and there are no repeated edges. We call such hypergraphs (r, r)-*graphs*. Note that $(2, 2)$-graphs are just bipartite graphs. These objects are also known as r-dimensional binary contingency tables.

The *degree* of a vertex is the number of edges that contain it. If every vertex has degree d then the hypergraph is called d-*regular*, which implies (unless there are no edges) that all the classes have the same size. We are interested in the number of labelled d-regular (r, r)-graphs with $n = mr$ vertices where every partition class contains exactly m vertices. We denote this number by $H_r(d, m)$.

In order to motivate our answer, we start with a non-rigorous argument. Consider a (r, r)-graph G with classes of size m and md edges, created by choosing uniformly at random md distinct edges out of the m^r available. Let $\mathcal{R}$ be the event that G is d-regular. Then

$$H_r(d, m) = \binom{m^r}{md} \Pr(\mathcal{R}).$$

*School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, Australia. E-mail: m.isaev@unsw.edu.au. Supported by the Australian Research Council grant DP250101611.

†Department of Mathematics, LMU Munich, Munich 80333, Germany. E-mail makai@math.lmu.de. Supported by the Australian Research Council grant DP190100977.

‡School of Computing, Australian National University, Canberra, ACT 2601, Australia. E-mail: brendan.mckay@anu.edu.au. Supported by the Australian Research Council grant DP250101611.

To estimate $\Pr(\mathcal{R})$, we first look at one class V_i, $i \in [r]$. The probability of the event $\mathcal{R}_i$ that every vertex in V_i has degree d is

$$\Pr(\mathcal{R}_i) := \frac{\binom{m^{r-1}}{d}^m}{\binom{m^r}{md}},$$

since $\binom{m^{r-1}}{d}$ is the number of ways to choose d edges of a (r,r)-graph incident to one vertex in V_i and these choices are independent. If the events $\mathcal{R}_i$ were also independent, we would have $\Pr(\mathcal{R}) = \prod_{i=1}^r \Pr(\mathcal{R}_i)$, providing the estimate

$$\hat{H}_r(d,m) := \binom{m^r}{md} \prod_{i\in[r]} \Pr(\mathcal{R}_i) = \frac{\binom{m^{r-1}}{d}^{rm}}{\binom{m^r}{md}^{r-1}}.$$

Of course, the events $\mathcal{R}_i$ are not independent, but comparing $\hat{H}_r(d,m)$ with the correct value $H_r(d,m)$ will be instructive.

In the case of $(2,2)$-graphs, that is for d-regular bipartite graphs, we have as $n = 2m \to \infty$ that

$$H_2(d,m) = (1+o(1))\, e^{-1}\hat{H}_2(d,m)$$

except in the trivial cases $d = 0$ and $d = m$. This is the result of three previous investigations. The sparse range was solved by McKay [7] (see also [4]), the intermediate range of densities by Liebenau and Wormald [6], and the dense range by Canfield and McKay [3].

In this paper we show that, for $r \geq 3$, the estimate $\hat{H}_r(d,m)$ is asymptotically correct.

Theorem 1. *Let $n = rm \to \infty$, where $m = m(n)$ and $r = r(n) \geq 3$. Then for any $0 \leq d \leq m^{r-1}$ we have*

$$H_r(d,m) = (1+o(1))\, \hat{H}_r(d,m).$$

First note that the statement holds when $m = 1$, and we can assume that $m \geq 2$ for the remainder of the proof. In addition, by considering the complement of individual hypergraphs we have $H_r(m^{r-1}-d,m) = H_r(d,m)$ and a quick calculation gives $\hat{H}_r(m^{r-1}-d,m) = \hat{H}_r(d,m)$. Therefore we may restrict ourselves to the case where $d \leq m^{r-1}/2$.

Consider the following three regions.

(a) $d = o(m)$ for $r = 3$;

(b) $rd^2 = o(m^{r-2})$ for $r \geq 3$;

(c) $d = \Omega(r^{16}m)$.

For $r = 3$ these three regions trivially cover every possible choice of d and m. While for $r > 3$ this holds if $m^{(r-2)/2}r^{-1/2} = \Omega(r^{16}m)$ which can be seen from the equivalent form

$$r^{33}m^{4-r} = \exp(33\log r + (4-r)\log m) = O(1),$$

by considering the $r = O(1), m = \omega(1)$ and $r = \omega(1)$ cases separately.

For the sparse regimes (regions (a) and (b)) we employ a combinatorial model introduced in 1972 by Békéssy, Békéssy and Komlós [1] and later developed under the name "configurations" by Bollobás and others [2, 8]. For region (b) this model produces a uniform random d-regular (r,r)-graph with probability $1-o(1)$. However, the method turns out to be insufficient to cover all of region (a). In order to extend the result to the remainder of region (a) we apply the method of switchings [7].

Furthermore, for the dense regime (region (c)) we use the complex-analytic approach, relying on the machinery developed by Isaev and McKay [5].

2 Sparse regimes

In this section we consider regions (a) and (b). Attach d "spines" to each vertex. A *configuration* is a partition of the spines into md sets of size r such that each set contains a spine attached to a vertex in each vertex class. Each configuration yields a d-regular multi-(r,r)-graph, possibly with repeated edges, if each spine set is equated with an edge containing the vertices to which the spines in the set are attached. A simple counting argument gives that the number of d-regular multi-(r,r)-graphs is $((md)!)^{r-1}/(d!)^{mr}$. Now we only need to establish the probability that a randomly chosen configuration corresponds to a simple (r,r)-graph, denoted by $P_r(d,m)$.

In region (b) a straightforward calculation gives that $P_r(d,m) = 1 - o(1)$. On the other hand for region (a) we show $P_r(d,m) = \exp(-d^2/2m+o(1))$, for which we require the following switching argument.

For the switching argument we start by showing that the probability of having more than M edges or having an edge with multiplicity larger than 2 is $o(1)$, for some sufficiently large M. Denote by $\mathcal{T}(\ell)$ the set of d-regular multi-(r,r)-graphs with exactly ℓ double edges and no edge with multiplicity larger than 2. The key idea is to establish the values $|\mathcal{T}(\ell+1)|/|\mathcal{T}(\ell)|$, which leads to $|\mathcal{T}(\ell+1)|/|\mathcal{T}(0)|$ using a telescoping product and finally to $\sum_{\ell=0}^{M} |\mathcal{T}(\ell)|/|\mathcal{T}(0)|$ which asymptotically matches the inverse of the required value. In particular, we use Corollary 4.5 from Greenhill, McKay, Wang [4], which establishes this result when sufficiently tight estimates exist for $|\mathcal{T}(\ell+1)|/|\mathcal{T}(\ell)|$.

It only remains to establish $|\mathcal{T}(\ell+1)|/|\mathcal{T}(\ell)|$, which we achieve by using the following switching (see Figure 1):

$$\mathcal{F}_1 = \big\{\{a,b,c\},\{a_1,b_5,c_3\},\{a_2,b_4,c_6\},\{a_3,b_1,c_5\},\{a_4,b_6,c_2\},\{a_5,b_3,c_1\},\{a_6,b_2,c_4\}\big\}$$
$$\mathcal{F}_2 = \big\{\{a,b_2,c_1\},\{a_1,b,c_2\},\{a_2,b_1,c\}\big\} \cup \big\{\{a_j,b_j,c_j\} : 3 \le j \le 6\big\},$$

where $\{a,a_j : 1 \le j \le 6\}$, $\{b,b_j : 1 \le j \le 6\}$ and $\{c,c_j : 1 \le j \le 6\}$ are sets of spines where each set is associated with a different vertex class.

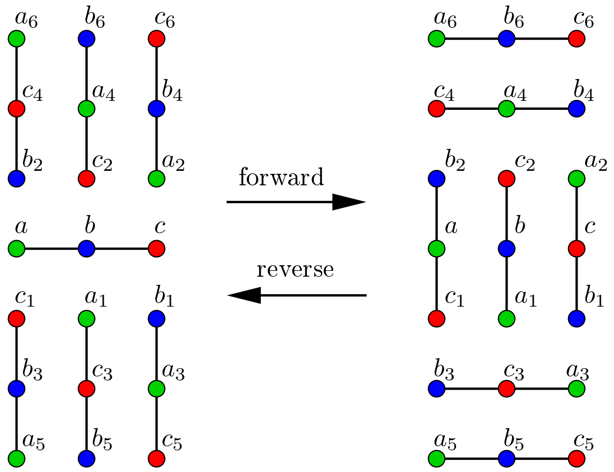


Figure 1: Switching operations for $r = 3$.
The colour indicates the vertex class associated with the spine.

Assume that the spines a', b', c' form an edge, and that spines a, b, c are attached to the same vertices as the spines a', b', c' respectively. Then, under some additional conditions, the forward switching decreases the number of double edges by exactly one, and the reverse switching increases the number of double edges by exactly one, while the number of edges with multiplicity larger than 2 remains unchanged.

The number of forward switchings from $\mathcal{T}(\ell)$ to $\mathcal{T}(\ell+1)$ is roughly $2\ell m^6 d^6$, as we can select one half of a double edge in 2ℓ ways and the remaining 6 edges can be chosen in roughly $m^6 d^6$ ways. In addition the number of reverse switchings is roughly $m^5 d^5 (d-1)^3$, as we can choose a simple edge a', b', c' in roughly md ways, the spines a, b, c so that they are on the same vertices as a', b', c' respectively in $(d-1)^3$ ways. This has now established the 3 edges containing a, b, c and there are roughly $m^4 d^4$ choices for the remaining 4 edges.

Using a double counting argument, this gives us an estimate on $|\mathcal{T}(\ell+1)|/|\mathcal{T}(\ell)|$ of $(d-1)^3/(2\ell dm)$, and consequently $P_r(d,m) = \exp(-d^2/2m + o(1))$.

3 Dense range

In this regime we consider region (c) using the complex-analytic approach. We establish a generating function for (r,r)-graphs by degrees, then extract the required coefficient via Fourier inversion and perform asymptotic analysis on the resulting multidimensional integral.

Let $\mathcal{S}_r(m)$ denote the set of all possible edges of (r,r)-graphs and clearly $|\mathcal{S}_r(m)| = m^r$. The generating function for (r,r)-graphs by degree sequence is

$$\prod_{e\in\mathcal{S}_r(m)} \Big(1 + \prod_{j\in e} x_j\Big).$$

Using Cauchy's coefficient formula, the number of d-regular (r,r)-graphs is

$$H_r(d,m) = [x_1^d \cdots x_n^d] \prod_{e\in\mathcal{S}_r(m)} \Big(1+\prod_{j\in e} x_j\Big) = \frac{1}{(2\pi i)^{rm}} \oint \cdots \oint \frac{\prod_{e\in\mathcal{S}_r(m)}(1+\prod_{j\in e} x_j)}{\prod_{j\in[rm]} x_j^{d+1}}\, d\boldsymbol{x}.$$

Let $\lambda = d/m^{r-1}$ then considering the contours $x_j = \left(\dfrac{\lambda}{1-\lambda}\right)^{1/r} e^{i\theta_j}$ for $j \in [n]$, the value of $H_r(d,m)$ can be determined by evaluating the integral

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{\prod_{e\in\mathcal{S}_r(m)}\big(1+\lambda(e^{i\sum_{j\in e}\theta_j}-1)\big)}{\exp\big(id\sum_{j\in[rm]}\theta_j\big)}\, d\boldsymbol{\theta}.$$

The integrand $F(\boldsymbol{\theta})$ has an inherent symmetry which occurs when for any $c_1 + \ldots + c_m = 0$ we replace θ by $\theta + c_i$ for every component of $\boldsymbol{\theta}$ corresponding to a vertex in V_i and for every $1 \le i \le r$. By integrating $F(\boldsymbol{\theta})$ only over

$$\{\boldsymbol{\theta} \in U_n(\pi) : \theta_{2m} = \theta_{3m} = \cdots = \theta_{rm} = 0\},$$

we eliminate this symmetry (at the cost of a multiplicative factor) and as a consequence the integrand has only one maximum point in the range, namely the null vector. This is advantageous as the value of the integral is essentially determined by evaluating it over a small region around each maximum, while the contribution to the integral from the remaining area is insignificant.

Next we evaluate the integral in a small box $\mathcal{B}$ around the maximum. Taking the Taylor expansion of $\sum_{e\in\mathcal{S}_r(m)}\log(1+\lambda(\exp(i\sum_{j\in e}\theta_j)-1))$, the first order term is $id\sum_{j\in[rm]}\theta_j$, which cancels with the denominator of $F(\boldsymbol{\theta})$. In addition, the second order term can be expressed as $\boldsymbol{\theta}^t A\boldsymbol{\theta}$, for some matrix A that has rank $n-r+1$.

The function will resemble a truncated multivariate Gaussian distribution with covariance matrix A. We apply the theoretical framework developed by Isaev and McKay in [5, Section 4] for asymptotically estimating such integrals, including cases where the covariance matrix is not of full rank.

Outside $\mathcal{B}$, we consider $|F(\boldsymbol{\theta})|$. Note that the absolute value of the denominator is 1 and by Taylor's theorem for any $x\in\mathbb{R}$ and $\lambda\in[0,1]$, we have

$$|1+\lambda(e^{ix}-1)|=\sqrt{1-2\Lambda(1-\cos x)}\le\exp\left(-\frac{\Lambda}{2}\left(1-\frac{|x|_{2\pi}^2}{12}\right)|x|_{2\pi}^2\right),$$

where $|x|_{2\pi}=\min_{k\in\mathbb{Z}}|x-2\pi k|$.

We split the remaining regime into 3 overlapping regimes. For the first 2 cases it suffices to show that $\sum_{e\in\mathcal{S}_r(m)}\theta_e^2$ is large, where $\theta_e=\left|\sum_{j\in e}\theta_j\right|_{2\pi}$. First we consider the case when there exists a vertex class such that each component θ_j corresponding to the vertex class is far from at least m/r other components in $\boldsymbol{\theta}$ corresponding to the same vertex class. If the components of $\boldsymbol{\theta}$ corresponding to a pair of vertices a,b in the same vertex class are far from each other, then for any $e,f\in\mathcal{S}_r(m)$, satisfying $e\triangle f=\{a,b\}$ we can establish a lower bound on either θ_e^2 or θ_f^2. The case is completed as we have many such pairs even after considering potential repetitions.

Next we consider the case where for each $1\le j\le r$ there is a θ_{k_j} close to at least $m-m/r$ components of $\boldsymbol{\theta}$ in their vertex class and in addition $|\theta_{k_1}+\dots+\theta_{k_r}|_{2\pi}$ is large. This ensures that for every $e\in\mathcal{S}_r(m)$ consisting only of vertices whose components are close to the respective θ_{k_j} we can bound θ_e^2 from below, and the case is completed as there are many such edges.

Due to our definition of closeness, the first two cases cover areas where there is a vertex class such that at least m/r components in $\boldsymbol{\theta}$ are outside $\mathcal{B}$, allowing us to establish a sufficiently small uniform upper bound on $F(\boldsymbol{\theta})$ in these cases.

The last remaining case considers what happens if there is at least 1 component in $\boldsymbol{\theta}$ outside $\mathcal{B}$, and at most m/r in every vertex class. We split the integration regime into further parts, depending on which components of $\boldsymbol{\theta}$ are outside $\mathcal{B}$. Assume that this set is $\mathcal{O}$ then we are interested in the integral of $|F(\boldsymbol{\theta})|$, where the elements of $\mathcal{O}$ are integrated outside $\mathcal{B}$ while the remaining elements are integrated inside $\mathcal{B}$ in the respective dimensions. Recall that

$$|F(\boldsymbol{\theta})|=\prod_{e\in\mathcal{S}_r(m)}\left|1+\lambda\left(e^{i\sum_{j\in e}\theta_j}-1\right)\right|,$$

and we can split this product according to whether an edge contains a vertex in $\mathcal{O}$ or not. Denote the part of the product disjoint from $\mathcal{O}$ by $\hat{F}(\hat{\boldsymbol{\theta}})$, after further components have been removed so that the number of comoponents remaining in each vertex class is the same. First we integrate $\hat{F}(\hat{\boldsymbol{\theta}})$ (over the lower dimensional equivalent of $\mathcal{B}$). Similarly to before, this function resembles a truncated multivariate Gaussian distribution and a repeated application of the approximation from Isaev and McKay [5, Section 4] provides the integral. Next consider only the edges which contain exactly one vertex in $\mathcal{O}$. The value of θ_e corresponding to these edges turns out to be large. This provides a uniform upper bound on $|F(\theta)|/\hat{F}(\hat{\boldsymbol{\theta}})$, which we integrate over the remaining components of $\boldsymbol{\theta}$. Summing the result over every choice of $\mathcal{O}$ completes this case and the proof.

References

[1] A. Békéssy, P. Békéssy and J. Komlós, Asymptotic enumeration of regular matrices, *Studia Sci. Math. Hungar.*, **7** (1972) 343–353.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *Europ. J. Combin.*, **1** (1980) 311–316.

[3] E. R. Canfield and B. D. McKay, Asymptotic enumeration of dense 0-1 matrices with equal row sums and equal column sums, *Electronic J. Combin.*, **12** (2005) #R29.

[4] C. Greenhill, B. D. McKay and X. Wang, Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums, *J. Combin. Th., Ser. A*, **113** (2006) 291–324.

[5] M. Isaev and B. D. McKay, Complex martingales and asymptotic enumeration, *Random Structures Algorithms*, **52** (2018) 617–661.

[6] A. Liebenau and N. Wormald, Asymptotic enumeration of digraphs and bipartite graphs by degree sequence, *Random Structures Algorithms*, **62** (2023) 259–286.

[7] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums, in *Enumeration and Design*, (Academic Press, 1984) 225–238.

[8] N. C. Wormald, Models of random regular graphs, in *Surveys in Combinatorics*, 1999 (eds. J. D. Lamb and D. A. Preece), London Mathematical Society Lecture Note Series, vol. 267, Cambridge University Press, Cambridge (1999), 239–298.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Tight bounds for intersection-reverse sequences, edge-ordered graphs and applications

(Extended abstract)

Barnabás Janzer* Oliver Janzer† Abhishek Methuku‡ Gábor Tardos§

Abstract

In 2006, Marcus and Tardos proved that if $A^1, \dots, A^n$ are cyclic orders on some subsets of a set of n symbols such that the common elements of any two distinct orders A^i and A^j appear in reversed cyclic order in A^i and A^j, then $\sum_i |A^i| = O(n^{3/2}\log n)$. This result is tight up to the logarithmic factor and has since become an important tool in Discrete Geometry. In this paper we improve this to the optimal bound $O(n^{3/2})$, and use our results to resolve several open problems in Discrete Geometry and Extremal Graph Theory as follows.

(i) We prove that every n-vertex topological graph that does not contain a self-crossing four-cycle has $O(n^{3/2})$ edges. This resolves a problem of Marcus and Tardos from 2006.

(ii) We show that n pseudo-circles in the plane can be cut into $O(n^{3/2})$ pseudo-segments, which, in turn, implies new bounds on the number of point-circle incidences and on other geometric problems. This goes back to a problem of Tamaki and Tokuyama from 1998 and improves several results in the area.

(iii) We also prove that the edge-ordered Turán number of the four-cycle C_4^{1243} is $\Theta(n^{3/2})$. This gives the first example of an edge-ordered graph whose Turán number is known to be $\Theta(n^\alpha)$ for some $1 < \alpha < 2$, and answers a question of Gerbner, Methuku, Nagy, Pálvölgyi, Tardos and Vizer.

Using different methods, we determine the largest possible extremal number that an edge-ordered forest of order chromatic number two can have. Kucheriya and Tardos

*Department of Mathematics, ETH Zürich, Switzerland. Research supported by SNSF grant 200021_228014 and by a fellowship at Magdalen College, University of Oxford. Email: barnabas.janzer@math.ethz.ch

†Institute of Mathematics, EPFL, Lausanne, Switzerland. Email: oliver.janzer@epfl.ch.

‡Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL, USA. Email: abhishekmethuku@gmail.com

§HUN-REN, Alfréd Rényi Institute of Mathematics, Budapest, Hungary. Research supported by the National Research, Development and Innovation Office projects K-132696 and SNN-135643 and by the ERC Advanced Grants "ERMiD" and "GeoScape." Email: tardos@renyi.hu

showed that every such graph has extremal number at most $n2^{O(\sqrt{\log n})}$, and conjectured that this can be improved to $n(\log n)^{O(1)}$. We disprove their conjecture in a strong form by showing that for every $C > 0$, there exists an edge-ordered tree of order chromatic number two whose extremal number is $\Omega(n2^{C\sqrt{\log n}})$.

1 Introduction

A major line of research at the intersection of Discrete Geometry and Graph Theory is the study of classical problems of the latter field for graphs that have some geometric or algebraic structure. Recent highlights of this subject include a semi-algebraic version of Zarankiewicz's problem [5], which has applications to point-variety incidence bounds, and a proof that string graphs have the Erdős–Hajnal property [21].

A topological graph is a representation of a graph in the plane such that the vertices of the graph are distinct points on the plane and the edges correspond to Jordan arcs joining the corresponding pairs of points. We assume that no edge passes through a vertex other than its endpoints and every two edges have a finite number of common interior points and they properly cross at each of these points. A geometric graph is a topological graph in which the edges are represented by straight line segments.

In the '60s, Avital and Hanani [3] as well as Erdős and Perles (see [13]) initiated, and later Kupitz [11] and many others continued the systematic study of extremal problems for geometric and topological graphs. In particular, they posed the following general question, which is a geometric analogue of the classical Turán problem. Given a collection of forbidden geometric configurations, what is the maximum number of edges that an n-vertex topological graph can have without containing any of the forbidden configurations? A well-studied instance of this general problem is the following question, posed by Pach, Pinchasi, Tardos and Tóth [14] in 2004.

Problem 1.1 (Pach–Pinchasi–Tardos–Tóth [14])**.** *Let K be a fixed abstract graph. What is the maximum number $\mathrm{ex}_{\mathrm{cr}}(n, K)$ of edges that a geometric graph on n vertices can have if it contains no self-intersecting copy of K?*

Trivially, we have $\mathrm{ex}_{\mathrm{cr}}(n, K) \geq \mathrm{ex}(n, K)$, and in particular, the answer is quadratic when K is not bipartite. Moreover, if K is not planar, then any copy of K in a geometric (or topological) graph is self-intersecting, so $\mathrm{ex}_{\mathrm{cr}}(n, K) = \mathrm{ex}(n, K)$. Hence, Problem 1.1 is the most interesting when K is planar and bipartite. Pach, Pinchasi, Tardos and Tóth [14] showed that $\mathrm{ex}_{\mathrm{cr}}(n, P_3) = \Theta(n \log n)$, where P_3 denotes the path with three edges. On the other hand, their best bound for the maximum number of edges in a *topological* graph on n vertices without a self-crossing P_3 is only $O(n^{3/2})$. Tardos [20] constructed, for each positive integer k, a geometric graph with a superlinear number of edges and no self-crossing path of length k.

One of the first results for topological graphs concerning Problem 1.1 was due to Pinchasi and Radoičić [18], who showed that the maximum number of edges in an n-vertex topological graph without a self-crossing four-cycle is $O(n^{8/5})$. Their proof relies on the study of intersection-reverse cyclic orders.

Definition 1.2. *Let $A^1, \ldots, A^n$ be cyclic orders of subsets of a finite alphabet. We say that A^i and A^j are* intersection-reverse *if their common elements appear in reverse order in the two cyclic orders. We say that $A^1, \ldots, A^n$ are* pairwise intersection-reverse *if for each $1 \leq i < j \leq n$, A^i and A^j are intersection-reverse.*

In what follows, for a cyclic or linear order A, we write $|A|$ for the number of elements ordered by A. For a vertex v of a topological graph G, let $L_G(v)$ be the list of its neighbours, ordered cyclically counterclockwise according to the initial segment of the connecting edge. Pinchasi and Radoičić observed that if the lists $L_G(u)$ and $L_G(v)$ are not intersection-reverse for two distinct vertices u and v of the topological graph G, then G contains a self-crossing four-cycle. Furthermore, they proved that if $A^1, \ldots, A^n$ are pairwise intersection-reverse cyclic orders on subsets of size d of a set of n symbols, then $d = O(n^{3/5})$. Applying this result to the cyclic orders $L_G(u)$, $u \in V(G)$, they deduced that every n-vertex topological graph without a self-crossing four-cycle has $O(n^{8/5})$ edges.

The upper bound for the size of intersection-reverse cyclic orders (and hence also for the maximum number of edges in topological graphs without a self-crossing four-cycle) was greatly improved by Marcus and Tardos.

Theorem 1.3 (Marcus–Tardos [12])**.** *Let $A^1, A^2, \ldots, A^n$ be a collection of pairwise intersection-reverse cyclic orders on subsets of a set of n symbols. Then $\sum_{i=1}^n |A^i| = O(n^{3/2} \log n)$.*

Note that this result is tight up to the logarithmic factor. Indeed, it is well-known that $\mathrm{ex}(n, C_4) = \Theta(n^{3/2})$ (see, e.g., [6]). This implies that there are examples of cyclically ordered lists $A^1, A^2, \ldots, A^n$ with $\sum_{i=1}^n |A^i| = \Theta(n^{3/2})$ where no distinct A^i and A^j share more than one symbol (so they are trivially intersection-reverse). Consequently, Marcus and Tardos raised the natural question of determining if the logarithmic factor is needed in Theorem 1.3. We answer this question by showing that the logarithmic factor can be completely removed.

Theorem 1.4. *Let $A^1, \ldots, A^n$ be a collection of pairwise intersection-reverse cyclic orders on some subsets of a set of n symbols. Then $\sum_{i=1}^n |A^i| = O(n^{3/2})$.*

In fact, we deduce Theorem 1.4 from a stronger result which concerns linear orders rather than cyclic orders.

Theorem 1.5. *If $A^1, \ldots, A^n$ are linear orders on some subsets of a set of n symbols such that no three symbols appear in the same order in any two distinct linear orders A^i and A^j, then $\sum_{i=1}^n |A^i| = O(n^{3/2})$.*

As demonstrated in [12], Theorem 1.3 has far-reaching consequences in Discrete Geometry. In the next two subsections, we discuss some of the geometric consequences of our main results. In Section 1.3, we present another application of our main results in the extremal theory of edge-ordered graphs. For an outline of the ideas required to prove our theorems above, see Section 2.1 of the arXiv version of this paper [8].

1.1 Self-crossing four-cycles in topological graphs

As mentioned above, using Theorem 1.3, Marcus and Tardos [12] showed that every n-vertex topological graph without a self-crossing four-cycle has $O(n^{3/2} \log n)$ edges. As we noted earlier, there exist C_4-free graphs on n vertices with $\Theta(n^{3/2})$ edges, so this bound is tight up to the logarithmic factor. Using Theorem 1.4, we obtain a tight bound for the maximum number of edges in a topological graph without a self-crossing four-cycle.

Corollary 1.6. *Every n-vertex topological graph without a self-crossing four-cycle has $O(n^{3/2})$ edges.*

Of course, this implies that every n-vertex geometric graph without a self-crossing four-cycle has $O(n^{3/2})$ edges. Furthermore, Corollary 1.6 implies improved bounds for other interesting geometric problems as well, for example, the number of tangencies between families of curves (see [15, 9]).

1.2 Number of incidences between pseudo-circles and points

Perhaps the most important consequence of our result concerns cutting pseudo-circles and pseudo-parabolas into pseudo-segments, which are defined as follows. A collection of *pseudo-circles* is a collection of simple closed Jordan curves, any two of which intersect at most twice, with proper crossings at each intersection. A collection of *pseudo-parabolas* is a collection of graphs of continuous real functions defined on the entire real line such that any two intersect at most twice and they properly cross at these intersections. A collection of *pseudo-segments* is a collections of curves, any two of which intersect at most once. Furthermore, for a collection $\mathcal{C}$ of pseudo-circles, the *cutting number* $\tau(\mathcal{C})$ is the minimum number of cuts that transforms $\mathcal{C}$ into a collection of pseudo-segments.

In 1998, Tamaki and Tokuyama [19] considered the problem of cutting pseudo-parabolas into pseudo-segments. Such results are very useful since pseudo-segments are much easier to work with compared to pseudo-parabolas and pseudo-circles. Furthermore, as we will see shortly, bounds on the cutting number of pseudo-circles directly translate into bounds on the number of incidences between pseudo-circles and points. Marcus and Tardos [12] used Theorem 1.3 to improve upon earlier results, including those of Tamaki and Tokuyama [19], Aronov and Sharir [2], and Agarwal, Aronov, Pach, Pollack and Sharir [1], to show that a collection of n pseudo-parabolas or n pseudo-circles can be cut into $O(n^{3/2}\log n)$ pseudo-segments. Our Theorem 1.4 implies the following improved bound (using the reduction from [12]).

Corollary 1.7. *Let $\mathcal{C}$ be a collection of n pseudo-parabolas or a collection of n pseudo-circles. Then $\mathcal{C}$ can be cut into $O(n^{3/2})$ pseudo-segments.*

A result of Agarwal et al. [1] states that if $\mathcal{C}$ is a collection of n pseudo-circles and $\mathcal{P}$ is a set of m points, then the number of incidences between $\mathcal{C}$ and $\mathcal{P}$ satisfies $I(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + n + \tau(\mathcal{C}))$, where $\tau(\mathcal{C})$ is the cutting number of $\mathcal{C}$. In particular, by the result of Marcus and Tardos [12], $I(\mathcal{C}, P) = O(m^{2/3}n^{2/3} + m + n^{3/2}\log n)$. By Corollary 1.7, we obtain (see [1, 2, 12]) the following polylogarithmically improved bounds for the number of incidences between points and (pseudo-)circles.

Corollary 1.8. *If $\mathcal{C}$ is a collection of n pseudo-circles and a $\mathcal{P}$ is a set of m points in the plane, then the number of incidences is*

$$I(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m + n^{3/2}).$$

If, in addition, $\mathcal{C}$ is a collection of circles (not just pseudo-circles), then

$$I(\mathcal{C},\mathcal{P}) = O(m^{2/3}n^{2/3} + m^{6/11}n^{9/11} + m + n).$$

For more geometric consequences of our results, we refer the reader to [12].

1.3 Edge-ordered graphs

In this subsection, we give two new results about the extremal numbers of edge-ordered graphs. While the first result is independent of the subject of intersection-reverse sequences, the second one will rely on our Theorem 1.5.

A systematic study of the extremal numbers of edge-ordered graphs was initiated by Gerbner, Methuku, Nagy, Pálvölgyi, Tardos and Vizer [7], although specific problems of similar kind had been considered much earlier (see, e.g., [4]). Formally, an edge-ordered graph is a finite simple graph $G = (V, E)$ with a linear order on its edge set E. A subgraph of an edge-ordered graph is itself an edge-ordered graph with the induced edge-order. We say that the edge-ordered graph G contains another edge-ordered graph H if H is isomorphic to a subgraph of G, where the isomorphism must respect the edge-order. For a positive integer n and an edge-ordered graph H, let the extremal number $\mathrm{ex}_<(n, H)$ of H be the maximal number of edges in an edge-ordered graph on n vertices that does not contain H. Note that trivially we have $\mathrm{ex}_<(n, H) \geq \mathrm{ex}(n, H)$, where, with a slight abuse of notation, we used H for both an edge-ordered graph and for its underlying unordered graph.

Gerbner et al. [7] defined the so-called *order chromatic number* $\chi_{\mathrm{or}}(H)$ of an edge-ordered graph H and used it to establish a version of the celebrated Erdős–Stone–Simonovits theorem for edge-ordered graphs. For non-empty edge-ordered graphs H, this parameter takes values in $\mathbb{Z}_{\geq 2} \cup \{\infty\}$. Their result then states that $\mathrm{ex}_<(n, H) = \binom{n}{2}$ if $\chi_{\mathrm{or}}(H) = \infty$ and $\mathrm{ex}_<(n, H) = \left(1 - \frac{1}{\chi_{\mathrm{or}}(H)-1} + o(1)\right)\binom{n}{2}$ otherwise.

Analogously to the case of classical (unordered) extremal graph theory, this shows that the most interesting case is where $\chi_{\mathrm{or}}(H) = 2$, as in that case the above result does not determine the asymptotics of $\mathrm{ex}_<(n, H)$. A simple but important dichotomy in classical extremal graph theory states that if H is a forest, then $\mathrm{ex}(n, H) = O(n)$, whereas if H contains a cycle, then $\mathrm{ex}(n, H) = \Omega(n^{1+\varepsilon})$ for some $\varepsilon > 0$ which can depend on H. Using the inequality $\mathrm{ex}_<(n, H) \geq \mathrm{ex}(n, H)$, it is still true that if H is an edge-ordered graph containing a cycle, then $\mathrm{ex}_<(n, H) = \Omega(n^{1+\varepsilon})$ for some $\varepsilon > 0$, but there exist edge-ordered forests with order chromatic number greater than two, so they can have extremal number $\Theta(n^2)$. The natural analogues of unordered forests in this context are therefore forests with order chromatic number two, and the extremal numbers of these graphs are of great interest. Gerbner et al. [7] studied the extremal numbers of certain short edge-ordered paths with order chromatic number two and showed that the extremal number can be $\Omega(n \log n)$. In the other direction, Kucheriya and Tardos [10] proved that for every edge-ordered forest H with order chromatic number two, we have $\mathrm{ex}_<(n, H) \leq n2^{O(\sqrt{\log n})}$, and conjectured that the stronger bound $\mathrm{ex}_<(n, H) \leq n(\log n)^{O(1)}$ should hold. This was the edge-ordered analogue of a similar conjecture on *vertex-ordered graphs*: Pach and Tardos [16] conjectured that vertex-ordered forests of interval chromatic number two have extremal number $n(\log n)^{O(1)}$. Recently, Pettie and Tardos [17] refuted the conjecture of Pach and Tardos. Building upon their result, we also disprove the conjecture of Kucheriya and Tardos, and completely settle this problem, by showing the following result.

Theorem 1.9. *For any $C > 0$, there exists an edge-ordered tree H with order chromatic number two such that $\mathrm{ex}_<(n, H) = \Omega(n2^{C\sqrt{\log n}})$.*

This matches the upper bound $\mathrm{ex}_<(n, H) \leq n2^{O(\sqrt{\log n})}$ proved by Kucheriya and Tardos [10] for edge-ordered forests of order chromatic number two. Note that while our result completely settles the problem of how large the extremal number of an edge-ordered forest of order chromatic number two can be, the analogous question for vertex-ordered graphs is wide open. Indeed, while [17] proves the same lower bound for extremal functions of some vertex-ordered forests of interval chromatic number two, the upper bound analogous to the one proved for edge-ordered graphs in [10] is entirely missing (and seems to be hard) for vertex-ordered graphs.

We know much less about the extremal numbers of edge-ordered graphs containing a cycle compared to that of edge-ordered forests. In fact, prior to our work, the only edge-ordered graphs H for which the order of magnitude of $\mathrm{ex}_<(n, H)$ were known are some small forests and graphs with order chromatic number greater than two. In particular, there was no edge-ordered graph H for which we knew that $\mathrm{ex}_<(n, H) = \Theta(n^\alpha)$ for some $\alpha \in (1, 2)$. A result of Gerbner et al. [7] came close to showing the existence of such a graph. Let C_4^{1243} be the four-cycle $abcd$ whose edges are ordered as $ab < bc < da < cd$. (We remark that there are two other edge-orderings of C_4, but they both have order chromatic number ∞, so their extremal number is $\binom{n}{2}$.) Gerbner et al. [7] proved that $\mathrm{ex}_<(n, C_4^{1243}) = O(n^{3/2} \log n)$, which comes close to the trivial lower bound $\Omega(n^{3/2})$. They asked whether $\mathrm{ex}_<(n, C_4^{1243}) = \Theta(n^{3/2})$. We use our Theorem 1.5 to answer this question affirmatively.

Theorem 1.10. $\mathrm{ex}_<(n, C_4^{1243}) = \Theta(n^{3/2})$.

As mentioned above, this is the first edge-ordered graph of order chromatic number two which is not a forest and whose extremal number is known (up to a constant factor).

Acknowledgements. We thank David Conlon, António Girão, Gaurav Kucheriya and Van Magnan for interesting discussions related to the topic of this paper.

References

[1] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *Journal of the ACM*, 51(2):139–186, 2004.

[2] B. Aronov and M. Sharir. Cutting circles into pseudo-segments and improved bounds for incidences and complexity of many faces. *Discrete & Computational Geometry*, 28:475–490, 2002.

[3] S. Avital and H. Hanani. Graphs. *Gilyonot Lematematika*, 3:275–287 (in Hebrew), 1966.

[4] V. Chvátal and J. Komlós. Some combinatorial theorems on monotonicity. *Canadian Mathematical Bulletin*, 14(2):151–157, 1971.

[5] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl. A semi-algebraic version of Zarankiewicz's problem. *Journal of the European Mathematical Society*, 19(6):1785–1810, 2017.

[6] Z. Füredi and M. Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, pages 169–264. Springer, 2013.

[7] D. Gerbner, A. Methuku, D. T. Nagy, D. Pálvölgyi, G. Tardos, and M. Vizer. Turán problems for edge-ordered graphs. *Journal of Combinatorial Theory, Series B*, 160:66–113, 2023.

[8] B. Janzer, O. Janzer, A. Methuku, and G. Tardos. Tight bounds for intersection-reverse sequences, edge-ordered graphs and applications. *arXiv preprint arXiv:2411.07188*, 2024.

[9] B. Keszegh and D. Pálvölgyi. The number of tangencies between two families of curves. *Combinatorica*, 43(5):939–952, 2023.

[10] G. Kucheriya and G. Tardos. A characterization of edge-ordered graphs with almost linear extremal functions. *Combinatorica*, 43(6):1111–1123, 2023.

[11] Y. S. Kupitz. Extremal problems in combinatorial geometry. *Aarhus University Lecture Notes Series*, 53, 1979.

[12] A. Marcus and G. Tardos. Intersection reverse sequences and geometric applications. *Journal of Combinatorial Theory, Series A*, 113(4):675–691, 2006.

[13] J. Pach. Geometric graph theory. In *Surveys in Combinatorics, 1999*, London Mathematical Society Lecture Note Series, page 167–200. Cambridge University Press, 1999.

[14] J. Pach, R. Pinchasi, G. Tardos, and G. Tóth. Geometric graphs with no self-intersecting path of length three. *European Journal of Combinatorics*, 25(6):793–811, 2004.

[15] J. Pach, A. Suk, and M. Treml. Tangencies between families of disjoint regions in the plane. *Discrete & Computational Geometry*, 45:131–138, 2012.

[16] J. Pach and G. Tardos. Forbidden paths and cycles in ordered graphs and matrices. *Israel Journal of Mathematics*, 155:359–380, 2006.

[17] S. Pettie and G. Tardos. A refutation of the Pach–Tardos conjecture for 0–1 matrices. *arXiv preprint arXiv:2407.02638*, 2024.

[18] R. Pinchasi and R. Radoičić. On the number of edges in geometric graphs with no self-intersecting cycle of length 4. In *Towards a Theory of Geometric Graphs*, volume 342 of *Contemporary Mathematics*, pages 233–243. American Mathematical Society, Providence, RI, 2004.

[19] H. Tamaki and T. Tokuyama. How to cut pseudo-parabolas into pseudosegments. *Discrete & Computational Geometry*, 19:265–290, 1998.

[20] G. Tardos. Construction of locally plane graphs with many edges. In *Thirty Essays on Geometric Graph Theory*, pages 541–562. Springer, New York, 2013.

[21] I. Tomon. String graphs have the Erdős–Hajnal property. *Journal of the European Mathematical Society*, 26(1):275–287, 2023.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Chromatic number and regular subgraphs

(Extended abstract)

Barnabás Janzer* Raphael Steiner* Benny Sudakov*

Abstract

In 1992, Erdős and Hajnal posed the following natural problem: Does there exist, for every $r \in \mathbb{N}$, an integer $F(r)$ such that every graph with chromatic number at least $F(r)$ contains r edge-disjoint cycles on the same vertex set? We solve this problem in a strong form, by showing that there exist n-vertex graphs with fractional chromatic number $\Omega\left(\frac{\log\log n}{\log\log\log n}\right)$ that do not even contain a 4-regular subgraph. This implies that no such number $F(r)$ exists for $r \geq 2$. We show that assuming a conjecture of Harris, the bound on the fractional chromatic number in our result cannot be improved.

1 Introduction

A great deal of attention in graph theory has been devoted to the following fundamental meta-question: Given a graph G of huge chromatic number, what kind of subgraphs and substructures can we guarantee to find in it? Over the years, a rich and interconnected array of problems of this type has been discussed in the literature, see, e.g., the recent survey article [14] by Scott for an overview (with a particular focus on induced subgraphs). Despite the popularity of these problems, many of them remain notoriously difficult to attack, in part because of a lack of tools that exist for finding structures in graphs with a large (but constant) chromatic number.

Several natural problems of this type were raised by Erdős and his collaborators, the most famous of which is a conjecture due to Erdős and Hajnal [3], stating that for all $k, g \in \mathbb{N}$ there exists an integer $f(k, g)$ such that every graph of chromatic number at least $f(k, g)$ has a subgraph of chromatic number k and girth at least g. The existence of $f(k, 4)$ for all $k \in \mathbb{N}$ was proved using a beautiful argument by Rödl [13], but progress has been lacking since. In this paper, we shall be concerned with the following problem raised by Erdős and Hajnal [6, Problem 6] dating back at least to 1992, that regards the structure of cycles in graphs of large chromatic number.

Problem 1.1 (Erdős and Hajnal 1992). *Is it true that for every $r \in \mathbb{N}$ there exists a number $F(r)$ such that every graph of chromatic number at least $F(r)$ contains r edge-disjoint cycles on the same vertex set?*

*Department of Mathematics, ETH Zürich, Zurich, Switzerland.
E-mail: { `barnabas.janzer,raphaelmario.steiner,benny.sudakov` }`@math.ethz.ch`. B. J. and B. S. supported by SNSF grant 200021-228014. R.S. Supported by the SNSF Ambizione Grant No. 216701.

This problem was repeated by Erdős [7, Problem 12] in a paper published in 1997, and is also included in the 1998 book [2] by Chung and Graham on Erdős's legacy of problems, see also the problem entries on the corresponding websites maintained by Thomas Bloom[1] and Fan Chung[2].

Since a graph with large chromatic number must contain many edges, Problem 1.1 is closely related to another question of Erdős [4]. In 1976 he asked for the maximum number of edges in an n-vertex graph that does not contain r edge-disjoint cycles on the same vertex set. In a recent breakthrough on this problem, Chakraborti, Janzer, Methuku and Montgomery [1] proved an upper bound of the form $n \cdot \mathrm{polylog}(n)$ for every fixed r, improving the previous best bound of $O(n^{3/2})$. This directly implies (via degeneracy) that every n-vertex graph without r edge-disjoint cycles on the same vertex set has chromatic number at most $\mathrm{polylog}(n)$. The best known lower bound on the number of edges in an n-vertex graph without r edge-disjoint cycles on the same vertex set (for $r \geq 2$) is of the form $\Omega(n \cdot \log\log n)$. This follows from a famous construction due to Pyber, Rödl and Szemerédi [12] which is known at least since 1985 (see [11]). Concretely, the authors of [12] showed that there exist n-vertex graphs with $\Omega(n \log\log n)$ edges and no k-regular subgraph for any $k \geq 3$, thus in particular containing no edge-disjoint cycles with the same vertex set. However, this construction is inherently bipartite and can thus not directly be used to construct graphs of large chromatic number. In fact, perhaps this feature of Pyber, Rödl and Szemerédi's construction was what inspired Erdős and Hajnal to pose Problem 1.1 above.

As the main result of this paper, we give a strong negative answer to Problem 1.1, by showing the following result. Here, $\chi_f(G)$ denotes the well-known *fractional chromatic number*.

Theorem 1.2. *For some $c > 0$ and all sufficiently large $n \in \mathbb{N}$, there exists an n-vertex graph G such that $\chi_f(G) \geq c\frac{\log\log n}{\log\log\log n}$ and G contains no 4-regular subgraph.*

In particular, since $\chi(G) \geq \chi_f(G)$ for every graph G, this shows the existence of graphs with arbitrarily large chromatic number that do not even contain two edge-disjoint cycles with the same vertex set. Hence, none of the numbers $F(r)$ in the problem of Erdős and Hajnal can exist for any r greater than 1 (trivially, we have $F(1) = 3$). The proof of Theorem 1.2 involves carefully analyzing a randomly constructed multi-partite variant of the construction of Pyber, Rödl and Szemerédi mentioned above. By slightly modifying our arguments we also obtain a much simpler proof for the existence of dense graphs without 3-regular subgraphs (the original analysis by Pyber, Rödl and Szemerédi [12] involved heavy computations).

The problem of determining the asymptotic behaviour of the maximum number of edges of graphs without a k-regular subgraph (for some $k \geq 3$) was posed by Erdős-Sauer [5] in 1975 and attracted a lot of attention in the last 40 years. It was recently fully resolved by Janzer and Sudakov [9], who showed that the answer is $\Theta(n \log\log n)$, thus matching the lower-bound construction of Pyber, Rödl and Szemerédi.

Theorem 1.3 ([9])**.** *For every integer $k \geq 3$ there exists some $C_k > 0$ such that for sufficiently large n every n-vertex graph without a k-regular subgraph has at most $C_k n \log\log n$ edges.*

In the light of this discussion, it seems natural to ask for the asymptotic behaviour of the maximum chromatic number $g_k(n)$ of an n-vertex graph without a k-regular subgraph. Theorems 1.2 and 1.3 imply that $\Omega\left(\frac{\log\log n}{\log\log\log n}\right) \leq g_k(n) \leq O(\log\log n)$ for every fixed $k \geq 4$. Given the local sparsity of graphs without a regular subgraph, we suspect that the lower bound gives the truth for every $k \geq 3$. Supporting this claim, we can alsow show (see Section 3 for the detailed statement) that every n-vertex graph without a k-regular subgraph has fractional chromatic number at most $O\left(\frac{\log\log n}{\log\log\log n}\right)$.

---

[1]See `https://www.erdosproblems.com/641`

[2]See `https://mathweb.ucsd.edu/~erdosproblems/erdos/newproblems/ManyEdgeDisjointCycles.html`

Notation and Preliminaries. Given a graph G, we denote by $V(G)$ its vertex set, by $E(G)$ its edge set, and by $e(G)$ its number of edges. For a subset $X \subseteq V(G)$, we denote by $G[X]$ the induced subgraph of G with vertex set X, and for two disjoint sets $U, V \subseteq V(G)$ we denote by $e_G(U,V)$ the number of edges in G with one endpoint in U and one endpoint in V. Given a graph G, its *fractional chromatic number* $\chi_f(G)$ is defined as the optimal value of the following linear program (by $\mathcal{I}(G)$ we denote the collection of independent sets in G):

$$\begin{aligned} \min \sum_{I \in \mathcal{I}(G)} x_I & \\ \text{s.t.} \sum_{I \in \mathcal{I}(G): v \in I} x_I &\geq 1 \quad (\forall v \in V(G)), \\ x_I &\geq 0 \quad (\forall I \in \mathcal{I}(G)). \end{aligned}$$

Note that the optimal value of the corresponding integer program (requiring $x_I \in \mathbb{Z}$) is exactly the chromatic number $\chi(G)$ of the graph, and thus $\chi(G) \geq \chi_f(G)$ for every graph G. By linear programming duality, we can also express $\chi_f(G)$ as the optimal value of

$$\begin{aligned} \max \sum_{v \in V(G)} w_v & \\ \text{s.t.} \quad \sum_{v \in I} w_v &\leq 1 \quad (\forall I \in \mathcal{I}(G)), \\ w_v &\geq 0 \quad (\forall v \in V(G)). \end{aligned}$$

This implies that a graph G satisfies $\chi_f(G) \leq k$ for some real number $k > 1$ if and only if for every weighting $w : V(G) \to \mathbb{R}_{\geq 0}$ there exists an independent set I in G such that $\sum_{v \in I} w(v) \geq \frac{1}{k} \sum_{v \in V(G)} w(v)$.

2 Proof of Theorem 1.2

We define a random graph as follows. Set $\epsilon = \epsilon_n = \frac{1}{\sqrt{\log n}}$. Let $C = \frac{1}{10} \log \log n$, and take disjoint sets $B_1, \ldots, B_C$ such that $|B_i| = n^{1-20^i \epsilon}$. (We will omit floors and ceilings.) The vertex set of our graph is $V = \bigcup_{i=1}^{C} B_i$ (and each B_i will be an independent set). For each $i \in [C]$, each vertex $v \in B_i$, and each $j \in [C]$ with $j > i$, we independently and uniformly at random pick a vertex $w_{v,j} \in B_j$ to be the unique neighbour of v in B_j. Then our random graph G has edge set given by all such pairs $vw_{v,j}$. Note that $|V| < n$ for large n.

Lemma 2.1. *With probability* $1 - o(1)$, *G does not contain a* 4*-regular subgraph.*

Proof. Assume that G has a 4-regular subgraph H. Let H have $s > 0$ vertices, and, for convenience, set $B_{C+1} = \emptyset$ and $B_0 = V$. Let $i \in [C] \cup \{0\}$ such that $|B_{i+1}| < s/1000 \leq |B_i|$.

Claim 2.2. *For some $x > 0$, H has a subgraph H'' on x vertices such that $V(H'') \subseteq \bigcup_{j=1}^{i-1} B_j$ and $e(H'') \geq 1.1x$.*

Proof. First note that $\sum_{j \geq i+1} |B_j| \leq |B_{i+1}| \sum_{j \geq i+1} n^{-(20^j - 20^{i+1})\epsilon} \leq 2|B_{i+1}| \leq \frac{1}{500} s$. Since H is 4-regular, at most $\frac{1}{125} s$ edges of H have an endpoint in $\bigcup_{j > i} B_j$. Let H' denote the subgraph of H induced by $V(H) \cap \bigcup_{j=1}^{i} B_j$, and H'' the subgraph induced by $V(H) \cap \bigcup_{j=1}^{i-1} B_j$. So our previous observations give $e(H') \geq (2 - \frac{1}{100}) s$.

Let $X = V(H'') = V(H) \cap \bigcup_{j=1}^{i-1} B_j$ and $Y = V(H) \cap B_i$. Since H is 4-regular, we have $e_H(X,Y) \leq 4|Y|$. Also, since H is a subgraph of G, we have $e_H(X,Y) \leq |X|$. It follows that $e_H(X,Y) \leq \frac{4}{5}(|X| + |Y|) \leq \frac{4}{5} s$, and hence $e(H'') \geq (2 - \frac{1}{100}) s - \frac{4}{5} s > 1.1 s \geq 1.1|X|$. □

For each $i \in [C] \cup \{0\}$, let $\mathcal{A}_i$ be the event that G has a subgraph H'' on vertex set $V(H'') \subseteq \bigcup_{j=1}^{i-1} B_j$ such that for some $0 < x \leq 1000|B_i|$ we have $|V(H'')| = x$ and $e(H'') \geq 1.1x$. Clearly, $\mathcal{A}_i$ can only hold if $i \geq 2$. Note that, if n is large enough, for any fixed i and x, the probability that G has such a subgraph H'' is at most

$$\begin{aligned}\binom{n}{x}\binom{x^2/2}{\lceil 1.1x\rceil}\left(\frac{1}{|B_{i-1}|}\right)^{\lceil 1.1x\rceil} &\leq \left(\frac{en}{x}\right)^x\left(\frac{ex^2/2}{|B_{i-1}|\lceil 1.1x\rceil}\right)^{\lceil 1.1x\rceil}\\ &\leq \left(\frac{en}{x}\right)^x\left(\frac{ex}{|B_{i-1}|}\right)^{1.1x}\\ &\leq \left(10\frac{nx^{0.1}}{|B_{i-1}|^{1.1}}\right)^x\\ &\leq \left(100\frac{n|B_i|^{0.1}}{|B_{i-1}|^{1.1}}\right)^x\\ &= \left(100n^{-\frac{9}{10}20^{i-1}\epsilon}\right)^x\\ &\leq e^{-\frac{1}{2}x\sqrt{\log n}}.\end{aligned}$$

It follows that $\Pr[\mathcal{A}_i] \leq \sum_{x\geq 1} e^{-\frac{1}{2}x\sqrt{\log n}} \leq 2e^{-\frac{1}{2}\sqrt{\log n}}$ (if n is large enough). Hence, with probability $1-o(1)$, for all $i \in [C]\cup\{0\}$, $\mathcal{A}_i$ does not hold. The result follows using Claim 2.2. $\square$

Remark. Our arguments can be used to give a significantly shorter proof of the result of Pyber, Rödl and Szemerédi [12] about the existence of graphs with n vertices, $\Theta(n\log\log n)$ edges and no 3-regular subgraphs – the construction is (essentially) the same, but our calculations are much simpler. Indeed, let G' denote the subgraph of G obtained by keeping only the edges which have an endpoint in B_1 and let H be a 3-regular subgraph of G' with s vertices and $1.5s$ edges. Since G' is bipartite and H is regular we have exactly $s/2$ vertices of H in B_1. Using the same notation as in the above proof, we have $e(H') \geq 1.49s$, $e_H(X,Y) \leq 3|Y|$ and $e_H(X,Y) \leq |X \cap B_1| = s/2$. If X has size at most $0.9s$, then $e(H'') = e(H') - e_H(X,Y) \geq 0.99s \geq 1.1|X|$. Otherwise $|Y| \leq 0.1s$, $e_H(X,Y) \leq 0.3s$ and again $e(H'') = e(H') - e_H(X,Y) \geq 1.19s > 1.1|X|$. The rest of the proof works the same way as in the above.

Lemma 2.3. *With probability* $1-o(1)$, $\chi_f(G) = \Omega(\frac{\log\log n}{\log\log\log n})$.

Proof. We work with the dual formulation for $\chi_f(G)$. assign weight $1/|B_i|$ to each vertex in B_i. Since the total weight is $C = \frac{1}{10}\log\log n$, it suffices to show that, with probability $1-o(1)$, each independent set has weight at most $10\log C$.

Let $\mathcal{E}_i$ be the event that G contains an independent set $I \subseteq \bigcup_{j\geq i} B_j$ with total weight at least $9\log C$ such that $|I \cap B_i|/|B_i| \geq \frac{\log C}{C}$. Clearly, if G has an independent set of weight at least $10\log C$, then $\mathcal{E}_i$ holds for some $i \in [C]$. Thus, we will upper bound the probability of $\mathcal{E}_i$ for each i.

Given some i and a positive integer $t \geq \frac{\log C}{C}|B_i|$, we consider the probability that there is some $I \subseteq \bigcup_{j\geq i} B_j$ with total weight at least $9\log C$ and $|I \cap B_i| = t$ such that I is independent in G. Fix such a set I, and let $p_j = |I \cap B_j|/|B_j|$ for all $j \geq i$, and note that $\frac{\log C}{C} \leq p_i \leq 1$. The probability that there are no edges in G between $I \cap B_i$ and $I \cap \bigcup_{j>i} B_j$ is

$$\prod_{j>i}(1-p_j)^{p_i|B_i|} \leq e^{-p_i|B_i|\sum_{j>i}p_j} \leq e^{-p_i|B_i|\cdot 8\log C}.$$

Moreover, the number of such sets I is at most

$$\binom{|B_i|}{p_i|B_i|}2^{\sum_{j>i}|B_j|}.$$

Note that $\sum_{j>i} |B_j| \leq 2|B_{i+1}| = 2|B_i|n^{-(20^{i+1}-20^i)\epsilon} \leq |B_i|e^{-5\sqrt{\log n}}$, and $\binom{|B_i|}{p_i|B_i|} \leq \left(\frac{e}{p_i}\right)^{p_i|B_i|} \leq e^{(1+\log(1/p_i))p_i|B_i|}$. Thus, by the union bound, the probability that such an independent set exists in G is at most

$$\begin{aligned}
e^{-p_i|B_i|\cdot 8\log C} \cdot e^{(1+\log(1/p_i))p_i|B_i|} \cdot 2^{|B_i|e^{-5\sqrt{\log n}}} &\leq \exp\left(|B_i|\left(-8p_i\log C + p_i(1+\log(1/p_i)) + e^{-5\sqrt{\log n}}\right)\right) \\
&\leq \exp\left(|B_i|\left(-8p_i\log C + 2p_i\log C + e^{-5\sqrt{\log n}}\right)\right) \\
&\leq \exp\left(-|B_i|\left(6(\log C)^2/C - e^{-5\sqrt{\log n}}\right)\right) \\
&= \exp\left(-\Omega(|B_i|/\log\log n)\right).
\end{aligned}$$

The calculation above shows that for any fixed value of $|I \cap B_i|$, the probability of such an I existing is $e^{-\Omega(|B_i|/\log\log n)}$. Since there are at most $|B_i|$ choices for the value of $|I \cap B_i|$, this gives

$$\Pr[\mathcal{E}_i] \leq |B_i|e^{-\Omega(|B_i|/\log\log n)} \leq e^{-n^{1/2}}$$

if n is large enough. Hence, with probability $1-o(1)$, none of the events $\mathcal{E}_i$ (for $i \in [C]$) hold, and hence G has no independent set of weight at least $10\log C$. □

Theorem 1.2 now follows by combining Lemma 2.1 and Lemma 2.3, noting that for sufficiently large n the constructed graph G has $< n$ vertices. We can then simply fill up G with isolated vertices to make the number of vertices match exactly n, while leaving the fractional chromatic number unchanged.

Remark. One can show that the graph G constructed above indeed has chromatic number $\chi(G) = \Theta\left(\frac{\log\log n}{\log\log\log n}\right)$, i.e., the lower bound on $\chi(G)$ given by the fractional chromatic number $\chi_f(G)$ is tight up to a constant factor.

3 Concluding remarks

We can also show that assuming the following conjecture of Harris about the fractional chromatic number of triangle-free graphs, the lower bound on the fractional chromatic number in our main result, Theorem 1.2, cannot be improved.

Conjecture 3.1 ([8])**.** *There exists an absolute constant $K > 0$ such that for every sufficiently large $d \in \mathbb{N}$, every triangle-free d-degenerate graph G satisfies $\chi_f(G) \leq K\frac{d}{\log d}$.*

The aforementioned result reads as follows.

Proposition 3.2. *If Conjecture 3.1 holds, then for every $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that for sufficiently large $n \in \mathbb{N}$ every n-vertex graph G without a k-regular subgraph satisfies*

$$\chi_f(G) \leq C_k \frac{\log\log n}{\log\log\log n}.$$

Harris' conjecture was recently confirmed by Martinsson [10], making the conclusion of Proposition 3.2 unconditionally true. The proof of the proposition is based on a subsampling-trick. We have to omit it due to the lack of space.

References

[1] Debsoumya Chakraborti, Oliver Janzer, Abhishek Methuku, and Richard Montgomery. Edge-disjoint cycles with the same vertex set, 2024. arXiv preprint, arXiv:2404.07190.

[2] Fan Chung and Ron Graham. *Erdős on graphs: His legacy of unsolved problems.* AK Peters/CRC Press, 1998.

[3] Paul Erdős. Problems and results in combinatorial analysis and graph theory. In *Proof techniques in graph theory*, pages 27–35. Academic Press, New York, 1969.

[4] Paul Erdős. Problems and results in graph theory and combinatorial analysis. In *Proc. British Comb. Conf.*, number 5, pages 169–192, 1976.

[5] Paul Erdős. Some recent progress on extremal problems in graph theory. *Congr. Numer.*, 14:3–14, 1975.

[6] Paul Erdős. Some of my favourite problems in various branches of combinatorics. *Le Matematiche*, 47(2):231–240, 1992.

[7] Paul Erdős. Some recent problems and results in graph theory. *Discrete Mathematics*, 164(1-3):81–85, 1997.

[8] David G Harris. Some results on chromatic number as a function of triangle count. *SIAM Journal on Discrete Mathematics*, 33(1):546–563, 2019.

[9] Oliver Janzer and Benny Sudakov. Resolution of the Erdős–Sauer problem on regular subgraphs. *Forum of Mathematics, Pi*, 11:e19, 2023.

[10] Anders Martinsson. Triangle-free d-degenerate graphs have small fractional chromatic number, 2025.

[11] László Pyber. Regular subgraphs of dense graphs. *Combinatorica*, 5(4):347–349, 1985.

[12] László Pyber, Vojtech Rödl, and Endre Szemerédi. Dense graphs without 3-regular subgraphs. *Journal of Combinatorial Theory, Series B*, 63(1):41–54, 1995.

[13] Vojtech Rödl. On the chromatic number of subgraphs of a given graph. *Proceedings of the American Mathematical Society*, 64(2):370–371, 1977.

[14] Alex Scott. Graphs of large chromatic number. In *Proceedings of the International Congress of Mathematicians*, volume 6, pages 4660–4681, 2022.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The interpolation theorem for asymmetric trees

(Extended abstract)

Jaroslav Nešetřil* Jialu Zhu†

Abstract

In this paper, we prove that interpolation theorem holds for asymmetric trees. That is, for any two asymmetric trees T' and T with $T' \subset T$, there is a copy of T', we say T'', such that there is a sequence of vertices $v_1, v_2, \ldots, v_t$ of $V(T - T'')$ such that the induced subtree $T[V(T) - \{v_1, v_2, \ldots v_i\}]$ is asymmetric and v_{i+1} is a leaf of $T[V(T) - \{v_1, v_2, \ldots v_i\}]$. Moreover, T'' is necessary, i.e., we can not find a sequence of $V(T - T')$ such that the above holds. Moreover, we show that interpolation theorem does not holds for other class of graphs. This theorem strengths the result independently proved by Nešetřil [3], Robertson and Zimmer [5], Sheehan [7], in which they proved the special case $T' = T_{1,2,3}$ and $T_{1,2,3}$ is the smallest asymmetric tree.

1 Introduction

For a graph G, an *automorphism* of G is a mapping f from $V(G)$ to $V(G)$ such that $uv \in E(G)$ if and only if $f(u)f(v) \in E(G)$. It is obvious that every graph G has a trivial automorphism f where $f(v) = v$ for any vertex $v \in V(G)$ and we denote f by Id. Let $\mathrm{Aut}(G)$ be the *automorphism group* of G which consist of all automorphism of G. The study of automorphism groups has occupied an important place in the development of algebraic graph theory. A classical result proved by Frucht [2] showed that any finite group can be realized as a automorphism group of some graph. This result connects group theory with graph theory.

Theorem 1. *For any finite group* Γ*, there is a graph* G *such that* $\mathrm{Aut}(G) = \Gamma$.

There are two extreme cases for $\mathrm{Aut}(G)$. One is that every mapping of $V(G)$ is a automorphism of graph G, i.e., $\mathrm{Aut}(G) = S_n$. The other one is that G only have one trivial automorphism, i.e., $\mathrm{Aut}(G) = \{Id\}$.

*Computer Science Institute, Faculty of Mathematics and Physics, Charles University, Czech Republic. E-mail: `nesetril@kam.mff.cuni.cz`.

†Faculty of Mathematics and Physics, Charles University, Czech Republic and School of Mathematical Sciences, Zhejiang Normal University, China. E-mail: `jialuzhu@iuuk.mff.cuni.cz`.

If $\mathrm{Aut}(G) = S_n$, then for any two vertices u and v, there is an automorphism f of G such that $f(u) = v$. This shows that every vertices in G are equivalent. Furthermore, there are only two graphs G with $\mathrm{Aut}(G) = S_n$: complete graph K_n and independent set $\overline{K_n}$. Therefore, it is straightforward to characterize graph with $\mathrm{Aut}(G) = S_n$.

If $\mathrm{Aut}(G) = \{Id\}$, then each vertex in G has its own unique property. We call such a graph an *asymmetric graph*. If $|\mathrm{Aut}(G)| \geq 2$, i.e., there exists a non-trivial automorphism of G, then we say that G is a *symmetric graph*.

Similarly to the characterization of graphs satisfying $\mathrm{Aut}(G) = S_n$, someone may expect to find a characterization for asymmetric graphs. However, a classical result proved by Erdös and Rényi [1] shows that almost all finite graphs are asymmetric.

Definition 1. *For an asymmetric graph G, if all proper induced subgraphs of G with at least 2 vertices are symmetric, then we say G is a* minimal asymmetric graph.

In 1988, Nešetřil conjectured at an Oberwolfach Seminar that there are only 18 minimal asymmetric graphs which is confirmed by Schweitzer and Schweitzer [6]. It follows from the definition of minimal asymmetric graph that every asymmetric graph contains a minimal asymmetric graph as subgraph. Therefore, there is a natural question as follows.

Question 1. *Assume that G is an asymmetric connected graph and H is an asymmetric connected subgraph of G, is it possible to delete vertices in $V(G-H)$ one by one such that every resulting graph is connected and still asymmetric?*

2 Main Result

For any positive integer $n_1 \leq n_2 \leq \ldots \leq n_k$, let $T_{n_1,n_2,\ldots,n_k}$ be the tree obtained from k paths $P_i = (x_{i,0}, x_{i,1}, \ldots, x_{i,n_i})$ of length n_i by identifying $\{x_{i,0} : 1 \leq i \leq k\}$ into a single vertex x. For example, the smallest asymmetric tree is $T_{1,2,3}$.

Unfortunately, the answer to the above question is negative for asymmetric graphs, and even for asymmetric trees: Let T be an asymmetric tree $T_{1,2,3,\ldots,k}$ with $d(x) = k \geq 4$. Let

$$T' = T[\{x, x_{1,1}, x_{2,1}, x_{2,2}, x_{3,1}, x_{3,2}, x_{3,3}\}].$$

Then T' is isomorphic to $T_{1,2,3}$. Observe that for any vertex $v \in V(T - T')$, if $T' + v$ is connected, then $T' + v$ is symmetric and hence we cannot find the such sequence of vertices. But we can find a copy of T' in T, we say T'', such that the sequence of $V(T - T'')$ exist.

Therefore, the interpolation theorem for asymmetric graphs is defined as follows.

Definition 2. *Given an asymmetric graph G and an asymmetric proper subgraph G' of G, if we can find a copy G'' of G' such that there exists a sequence of vertices $v_1, v_2, \ldots, v_t$ in $V(G - G'')$ satisfying the following conditions (let $G_i = G[V(G) - \bigcup_{j=1}^{i}\{v_j\}]$):*

- $t = |V(G)| - |V(G')|$,
- *v_i is not cut-vertex of G_{i-1}, for any $1 \leq i \leq t$,*
- *G_i is asymmetric, for any $1 \leq i \leq t$,*

then we say that interpolation theorem holds for (G', G).

Nešetřil [3], Robertson and Zimmer [5], Sheehan [7] have independently proved that interpolation theorem holds for $(T_{1,2,3}, T)$ where $T_{1,2,3}$ is the smallest asymmetric tree and T is any asymmetric tree.

The main result of this paper is that the interpolation theorem holds for asymmetric trees. Moreover, this result is tight. That is, this conclusion cannot be further extended to asymmetric graphs with cycles.

Theorem 2. *For any asymmetric tree T' and T with $T' \subseteq T$, interpolation theorem holds for (T', T).*

Figure 1 shows a simple example of Theorem 2.

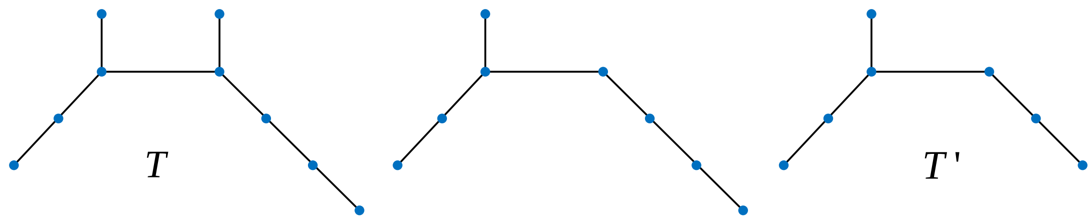


Figure 1: Example of a sequence of asymmetric subtrees between an asymmetric tree T and $T' = T_{1,2,3}$.

3 Sketch Proof

For any asymmetric tree T, let $A(T)$ be the set of all maximal asymmetric proper subtrees of T. Here we say T' is a *maximal asymmetric proper subtree* of T, if

(1) $T' \subset T$ is asymmetric and

(2) there is no asymmetric tree $T'' \subset T$ such that $T' \subset T''$.

Nešetřil [4] proved that asymmetric trees T can be determined by $A(T)$.

Theorem 3. *Assume T and S are any two asymmetric trees. Then $T \cong S$ if and only if $A(T) \cong A(S)$.*

Here note that $A(T)$ is the multi-set and $A(T) \cong A(S)$ means that there is a one–to–one mapping f from $A(T)$ to $A(S)$ such that $T' \cong f(T')$. Moreover, Nešetřil proved the following result which is used for proving main theorem in [4].

Theorem 4. *For any asymmetric tree T, there exists $T' \in A(T)$ such that $|T'| = |T| - 1$.*

In this paper, we defined strongly maximal asymmetric proper subtree which is stronger than maximal asymmetric proper subtree.

Definition 3. *We say T' is a* strongly maximal asymmetric proper subtree *of T if*

(1) $T' \subset T$ is asymmetric and

(2) there is no asymmetric tree $T'' \subset T$ such that T'' containing a copy of T'.

The condition (2) in definition of maximal asymmetric proper subtree only consider the subtree T'' contains T' as a subtree, but the second condition in definition of strongly one consider all subtrees T'' contain a copy of T' which is more larger. So we say it is strongly maximal asymmetric proper subtree.

Definition 4. *For any asymmetric tree T, let $A^\star(T)$ be the set of all strongly maximal asymmetric proper subtrees of T.*

As T' is a subtree of T'' implies that T'' contains a copy of T', we can conclude that $T' \in A^\star(T)$ implies that $T' \in A(T)$ and hence

$$A^\star(T) \subseteq A(T).$$

Observation 5. *For an asymmetric tree T, $A^\star(T) = \varnothing$ if and only if T is minimal asymmetric tree, i.e., $T = T_{1,2,3}$.*

Indeed, the main result proved in this paper is that all strongly maximal asymmetric proper subtrees are obtained from T by deleting one vertex.

Theorem 6. *For any asymmetric tree T, if $T' \in A^\star(T)$, then $|T'| = |T| - 1$.*

We confirm Theorem 2 by proving it is equivalent to Theorem 6. In the following proof, we show that Theorem 2 is true if and only if Theorem 6 holds. To prove it, we need to use the following statement as a bridge:

$$\text{For any asymmetric tree } T' \text{ and } T \text{ with } T' \subset T, \text{ if } T' \in A^\star(T), \text{ then } (T', T) \text{ holds for the interpolation theorem.} \qquad (\star)$$

The following observation will be used frequently in the following proof.

Observation 7. *If (T', T'') and (T'', T) both hold for interpolation theorem, then (T', T) also holds for interpolation theorem.*

Proof. As (T'', T) holds for interpolation theorem, let $v_1, v_2, \ldots, v_s$ be the sequence of vertices in $V(T - T_1'')$ (where T_1'' is a copy of T'' in T) such that $T - \{v_1, v_2, \ldots, v_i\}$ is asymmetric for any $1 \le i \le s$.

Note that (T', T'') holds for interpolation theorem and $T_1'' \cong T''$, we know that (T', T_1'') holds for interpolation theorem. Hence there is a sequence $v_{s+1}, v_{s+2}, \ldots, v_t$ of vertices in $V(T_1'' - T_1')$ (where T_1' is a copy of T' in T_1'') such that $T_1'' - \{v_1, v_2, \ldots, v_i\}$ is asymmetric for any $s + 1 \le i \le t$.

Then T_1' is a copy of T' in T and there is a sequence $v_1, v_2, \ldots, v_t$ of vertices in $V(T_1' - T)$ such that $T - \{v_1, v_2, \ldots, v_i\}$ is asymmetric for any $1 \le i \le t$. This implies that (T', T) holds for interpolation theorem. □

Theorem 8. *Theorem 2, $(\star)$ and Theorem 6 are equivalent.*

Proof. It suffice to prove that Theorem 2 is equivalent to $(\star)$ and $(\star)$ is equivalent to Theorem 6.

Claim 9. *Theorem 2 is equivalent to $(\star)$*

Proof. Note that Theorem 2 asserts that (T', T) holds for interpolation theorem For any asymmetric tree T' and T with $T' \subseteq T$. Observe that all elements of $A^\star(T)$ is an asymmetric subtree of T. Thus $(\star)$ is a special case of Theorem 2. So it is trivial that Theorem 2 can imply $(\star)$. Therefore it suffice to prove that $(\star)$ implies Theorem 2.

Assume to the contrary that there are two asymmetric trees T' and T with $T' \subseteq T$ and $T' \notin A^\star(T)$ such that (T', T) does not hold for interpolation theorem. Choose such T and T' with minimum $|T| - |T'|$.

If $T = T_{1,2,3}$, then $T' = T$ (as $T_{1,2,3}$ is the smallest asymmetric tree) and hence it is trivial that Theorem 2 holds for (T', T), a contradiction. So we can assume that $T \neq T_{1,2,3}$ and hence $A^\star(T) \neq \varnothing$ by Observation 5.

As $T' \notin A^\star(T)$, it follows from the definition of $A^\star(T)$ that there exists an asymmetric tree $T'' \in A^\star(T)$ such that T'' contains a copy of T'. Combining with $(\star)$ and the minimality of $|T| - |T'|$, we know that (T', T'') and (T'', T) hold for interpolation theorem. By Observation 7, (T', T) holds for interpolation theorem. □

Claim 10. $(\star)$ *is equivalent to Theorem 6.*

Proof. Firstly, we shall show that Theorem 6 can imply $(\star)$.

Recall that Theorem 6 says that all strongly maximal asymmetric proper subtrees of T have size $|T| - 1$. Hence if $T' \in A^\star(T)$, then there is a vertex $x \in T$ such that $T - x \cong T'$. It easy to see that T and $T - x$ is the sequence we desired and hence Theorem 6 can imply $(\star)$ easily.

Next we shall show that $(\star)$ implies Theorem 6.

Assume to the contrary that there are asymmetric trees T' and T with $T' \in A^\star(T)$ such that $|T'| \leq |T| - 2$.

It follows from $(\star)$ that there is a sequence of asymmetric trees $T_1, T_2, \ldots, T_t$ where $t = |V(T)| - |V(T')| + 1$ such that $T' \sqsubseteq T_i \sqsubseteq T_{i-1}$ and $|T_i| = |T_{i-1}| - 1$.

But on the other hand, by the definition of $A^\star(T)$, we know that for any proper subtree T'' of T such that $T' \sqsubset T''$, T'' is symmetric. This means that there is no asymmetric subtree of T of size $\geq |T'| + 1$ containing a copy of T', which contradicts to the existence of T_{t-1}. □

Theorem 8 follows from Claim 9 and Claim 10 and hence this completes the whole proof of Theorem 8. □

As a consequence, we confirm Theorem 2, i.e., interpolation theorem holds for asymmetric trees.

References

[1] Paul Erdös and Alfréd Rényi. Asymmetric graphs. *Acta Mathematica. Academiae Scientiarum Hungaricae*, 14:295–315, 1963.

[2] Robert Frucht. Herstellung von Graphen mit vorgegebener abstrakter Gruppe. *Compositio Math.*, 6:239–250, 1939.

[3] Jaroslav Nešetřil. The structure of asymmetric graphs. Master's thesis, McMaster University, 1969.

[4] Jaroslav Nešetřil. A congruence theorem for asymmetric trees. *Pacific Journal of Mathematics*, 37:771–778, 1971.

[5] Neil Robertson and Jack Adrian Zimmer, Jr. Automorphisms of subgraphs obtained by deleting a pendant vertex. *Journal of Combinatorial Theory. Series B*, 12(2):169–173, 1972.

[6] Pascal Schweitzer and Patrick Schweitzer. Minimal asymmetric graphs. *Journal of Combinatorial Theory. Series B*, 127:215–227, 2017.

[7] John Sheehan. Fixing subgraphs and ulam's conjecture. *Journal of Combinatorial Theory. Series B*, 14(2):125–131, 1973.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A NOTE ON INFINITE VERSIONS OF (p, q)-THEOREMS

(EXTENDED ABSTRACT)

Attila Jung* Dömötör Pálvölgyi†

Abstract

We prove that fractional Helly and (p, q)-theorems imply $(\aleph_0, q)$-theorems in an entirely abstract setting. We give a plethora of applications, including reproving almost all earlier $(\aleph_0, q)$-theorems about geometric hypergraphs that were proved recently. Some of the corollaries are new results, for example, we prove that if $\mathcal{F}$ is an infinite family of convex compact sets in $\mathbb{R}^d$ and among every $\aleph_0$ of the sets some $d+1$ contain a point in their intersection with integer coordinates, then all the members of $\mathcal{F}$ can be pierced by finitely many points with integer coordinates.

1 Introduction

Let $\mathcal{K}_d$ be the hypergraph whose vertices are the compact convex sets in $\mathbb{R}^d$, and edges represent intersecting families of convex sets; note that the edges form a downwards closed set system. Many results of combinatorial convexity can be stated as properties of this hypergraph, called nerve complex in topology.

For a hypergraph $\mathcal{H}$, denote its vertex set by $V(\mathcal{H})$, the number of its edges by $e(\mathcal{H})$, and the q-uniform part, consisting of the edges that contain exactly q vertices, by $\mathcal{H}^{(q)}$. For an $S \subset V(\mathcal{H})$ vertex set, let $\mathcal{H}[S]$ be the subhypergraph spanned by S, which consists of the edges contained entirely in S.

According to the celebrated Alon-Kleitman (p, q)-theorem, for every $p \geq d+1$, for every family $\mathcal{F}$ of compact convex sets in $\mathbb{R}^d$, if among every p members of $\mathcal{F}$ some $d+1$ are intersecting (i.e., they all have a point in common), then all the members of $\mathcal{F}$ can be pierced by $C = C(p, d)$ points. In our language, this can be stated as follows.

Theorem 1 (Alon and Kleitman [2]). *For every finite $p \geq d+1$ there exists a $C < \infty$ with the property that if $S \subset V(\mathcal{K}_d)$ is such that $\mathcal{K}_d^{(d+1)}[S]$ does not contain independent sets of size p, then S can be covered with C edges of $\mathcal{K}_d$.*

*ELTE Eötvös Loránd University and HUN-REN Alfréd Rényi Institute of Mathematics. E-mail: jungattila@gmail.com. Supported by the ERC Advanced Grant "ERMiD", by the EXCELLENCE-24 project no. 151504 of the NRDI Fund, and by the NKFIH grants TKP2021-NKTA-62, FK132060 and SNN135643.

†ELTE Eötvös Loránd University and HUN-REN Alfréd Rényi Institute of Mathematics. E-mail: domotor.palvolgyi@ttk.elte.hu. Supported by the ERC Advanced Grant "ERMiD", by the EXCELLENCE-24 project no. 151504 of the NRDI Fund, and by the NKFIH grants TKP2021-NKTA-62.

One of the main ingredients in its proof, the fractional Helly theorem of Katchalski and Liu, can be phrased as follows.

Theorem 2 (Katchalski and Liu [8]). *If $S \subset V(\mathcal{K}_d)$ is a finite subset and $e(\mathcal{K}_d^{(d+1)}[S]) \geq \alpha\binom{|S|}{d+1}$ for some $\alpha > 0$, then there exists an edge of $\mathcal{K}_d[S]$ of size $\beta|S|$ where $\beta = \beta(\alpha, d) > 0$ depends only on α and d.*

More generally, we say that a q-uniform (infinite)[1] hypergraph $\mathcal{H}$ satisfies the fractional Helly property, if for all $\alpha > 0$ there exists a $\beta > 0$ such that if $e(\mathcal{H}[S]) \geq \alpha\binom{|S|}{q}$ for some finite $S \subset V(\mathcal{H})$, then there exists a (q-uniform) clique of $\mathcal{H}[S]$ of size $\beta|S|$. Because of Helly's theorem, Theorem 2 is equivalent to that $\mathcal{K}_d^{(d+1)}$ satisfies the fractional Helly property.

For $0 \leq k < d$, let $\mathcal{B}_{d,k}$ be the hypergraph whose vertices are compact balls from $\mathbb{R}^d$, and edges represent families of balls which can be pierced by a single k-flat (k-dimensional affine subspace). Keller and Perles proved an infinite variant of the Alon-Kleitman theorem for k-flats intersecting Euclidean balls, that states that if we are given a collection of closed balls S such that among any infinite subcollection of S there are $k+2$ that can be pierced by a single k-flat, then there are finitely many k-flats that pierce all balls of S. In our language, this can be stated as follows.

Theorem 3 (Keller and Perles [9]). *If $S \subset V(\mathcal{B}_{d,k})$ is such that $\mathcal{B}_{d,k}^{(k+2)}[S]$ has no infinitely large independent set, then S can be covered with a finite number of edges of $\mathcal{B}_{d,k}$.*

Theorem 3 was proved for $k = 0$ for general balls, and for $k > 0$ for only unit balls in the first version of [9] that appeared in SoCG 2022, and later for all $0 \leq k < d$ for general radius balls that can be found in their arXiv preprint. We prove that such an infinite variant of the Alon-Kleitman theorem follows from the corresponding finite version and a fractional Helly theorem. In fact, we prove that if our hypergraph satisfies the fractional Helly property, then the condition of the infinite variant of the Alon-Kleitman theorem implies the condition of the finite version with some finite p. We state this in the contrapositive form as follows.

Theorem 4. *If a q-uniform hypergraph satisfies the fractional Helly property and has arbitrarily large finite independent sets, then it has an infinitely large independent set.*

Theorem 4 is proved in Section 2. In the rest of the introduction, we list a few corollaries of it, all analogs of the result of Keller and Perles.

Combining Theorem 4 with Theorems 1 and 2, we get the following result.

Corollary 5. *Let $\mathcal{F}$ be a family of compact convex sets in $\mathbb{R}^d$. If among every $\aleph_0$ members of $\mathcal{F}$ some $d+1$ are intersecting, then all the members of $\mathcal{F}$ can be pierced by finitely many points.*

In our language, this can be stated as follows. If $S \subset V(\mathcal{K}_d)$ is such that $\mathcal{K}_d^{(d+1)}[S]$ has no infinitely large independent set, then S can be covered with a finite number of edges of $\mathcal{K}_d$.

Proof of Corollary 5. By Theorem 2, $\mathcal{K}_d$ satisfies the fractional Helly property. As $\mathcal{K}_d^{(d+1)}[S]$ has no infinitely large independent set, there exists some finite p such that $\mathcal{K}_d^{(d+1)}[S]$ has no independent set of size p by Theorem 4. We can use Theorem 1 to conclude that S can be covered with $C(p, d)$ edges of $\mathcal{K}_d$. □

[1] Note that the fractional Helly property is true for all finite hypergraph, as we can choose a small enough β. Instead, we could define the fractional Helly property for a family of finite hypergraphs, but it is easier to deal only with one infinite hypergraph.

Similarly, the fractional Helly property and the (p, q)-theorem about hyperplanes intersecting convex sets by Alon and Kalai [1] imply the following infinite variant.

Corollary 6. *Let $\mathcal{F}$ be a family of compact convex sets in $\mathbb{R}^d$. If among every $\aleph_0$ members of $\mathcal{F}$ some $d+1$ can be pierced by a hyperplane, then all the members of $\mathcal{F}$ can be pierced by finitely many hyperplanes.*

A subset $S \subset \mathbb{Z}^d$ is called a convex lattice set, if there is a convex set $C \subset \mathbb{R}^d$ with $S = C \cap \mathbb{Z}^d$. The fractional Helly and $(p, d+1)$-theorems about convex lattice sets by Bárány and Matoušek [3] imply the following infinite variant.

Corollary 7. *Let $\mathcal{F}$ be a family of compact convex sets in $\mathbb{R}^d$. If among every $\aleph_0$ members of $\mathcal{F}$ some $d+1$ contain a point in their intersection with integer coordinates, then all the members of $\mathcal{F}$ can be pierced by finitely many points with integer coordinates.*

Using the fractional Helly and $(p, k+2)$-Theorems about k-flats intersecting Euclidean balls proved in [6], Theorem 4 also provides an alternative proof of Theorem 3 of Keller and Perles [9], which initiated this whole line of research.

We could give a long list of other corollaries of Theorem 4, one for each case where a fractional Helly and a (p, q)-type result is known, but such a list would add little to the paper. Section 2 contains the proof of our main Theorem 4.

2 Proof of Theorem 4

In Section 2.1, we show a class $M_s^{(q)}(t)$ of forbidden subhypergraphs of hypergraphs satisfying the fractional Helly property. In Section 2.2, we show a lemma about finding highly homogeneous subhypergraphs in infinite hypergraphs. Finally, in Section 2.3, we prove Theorem 4 by showing how homogeneous $M_s^{(q)}(t)$-free hypergraphs with arbitrary large finite independent sets contain infinitely large independent sets.

2.1 A consequence of the fractional Helly property

Motivated by Holmsen [4], for any $s, t \geq q$ we define $M_s^{(q)}(t)$ as a class of q-uniform hypergraphs as follows. Take st vertices divided into s parts of size t such that we have a complete q-uniform s-partite hypergraph among the parts, but no edge inside any part. There is no restriction on the "mixed" edges that intersect more than one, but less than q parts. If $q = 2$, there are no mixed edges, the only graph in the family $M_s^{(2)}(t)$ is the complete s-partite graph $K_{t,\dots,t}$. For $q > 2$, however, there are several different q-uniform graphs in $M_s^{(q)}(t)$. We call a q-uniform hypergraph $M_s^{(q)}(t)$-free if it contains none of them as an induced subgraph. By monotonicity, if a hypergraph is $M_s^{(q)}(t)$-free, it is also $M_s^{(q)}(t+1)$-free and $M_{s+1}^{(q)}(t)$-free.

Holmsen [4] proved that for any $s \geq q$, the $M_s^{(q)}(q)$-free q-uniform hypergraphs satisfy the fractional Helly property, which can be interpreted as that the 'fractional Helly number' of any hypergraph is at most as large as the 'colorful Helly number' of the hypergraph (see also [5]). In the opposite direction, we observe the following.

Claim 8. *Every q-uniform hypergraph that has the fractional Helly property is $M_q^{(q)}(t)$-free for $t > \frac{q-1}{\beta}$ where β belongs to $\alpha = \frac{q!}{q^q}$.*

Proof. A graph from $M_q^{(q)}(t)$ would have at least $t^q \geq \frac{q!}{q^q}\binom{qt}{q}$ edges, so by the fractional Helly property would contain a clique of size βqt, but the largest clique in any graph from $M_q^{(q)}(t)$ has size at most $(q-1)q$. □

Thus, $M_s^{(q)}(q)$-freeness implies the fractional Helly property for any $s \geq q$ by [4], and the fractional Helly property implies $M_q^{(q)}(t)$-freeness for some $t \geq q$ by Claim 8. Although it will not be used in the proof of Theorem 4, for the sake of completeness, we observe that neither of the above implications can be reversed.

Claim 9. *There are hypergraphs with the fractional Helly property which are not $M_s^{(q)}(q)$-free for any $s \geq q$, and there are hypergraphs without the fractional Helly property which are $M_q^{(q)}(t)$-free for all $t > q$.*

Proof. For the first part, let H be the complement of an infinite matching, i.e., the complement of infinitely many pairwise disjoint edges containing q vertices each. This is clearly not $M_s^{(q)}(q)$-free for any $s \geq q$, and any n vertices contain a clique of size at least $\frac{q-1}{q}n$, so $\beta = \frac{q-1}{q}$ is a good choice for every α.

For the second part, for any n we construct a hypergraph H_n which is $M_q^{(q)}(q+1)$-free, dense, but has no clique of linear size. It is a basic result from off-diagonal (hypergraph) Ramsey-theory that for every $s > q \geq 2$ we can color the edges of $K_n^{(q)}$ with red and blue such that there is no red $K_s^{(q)}$ and no blue $K_{f_s(n)}^{(q)}$, where $f_s(n) = o(n)$ (see for example the survey [10]). If H consists of the blue edges of such a coloring, then its complement contains no $K_{q+1}^{(q)}$, so H is dense and $M_q^{(q)}(q+1)$-free, but its largest clique is $o(n)$, so H does not have the fractional Helly property. Finally, since H is $M_q^{(q)}(q+1)$-free, it is $M_q^{(q)}(t)$-free for every $t > q$. □

2.2 Homogenization

The main ingredient in the proof of Theorem 4 is a Ramsey-type statement about the existence of highly homogeneous subhypergraphs in infinite hypergraphs; we state this as an independent lemma. For $1 \leq p \leq q$, and a sequence $(V_i)_{i\in\mathbb{N}}$ of sets, we call a q-tuple $(v_1, v_2, \ldots, v_q)$ an *increasing q-tuple of $(V_i)_{i\in\mathbb{N}}$ starting with $(v_1, \ldots, v_p)$* if there exist indices $i_1 < i_2 < \ldots < i_q$ such that $v_j \in V_{i_j}$. We say that a q-uniform hypergraph H spanned by $\cup_i V_i$ is *homogeneous with respect to an increasing p-tuple* $(v_1, \ldots, v_p)$, if either all increasing q-tuples starting with $(v_1, \ldots, v_p)$ are edges, or no increasing q-tuples starting with $(v_1, \ldots, v_p)$ are edges. We say that $H[\cup_i V_i]$ is *p-homogeneous*, if it is homogeneous with respect to every growing p-tuple. Finally, for an infinite sequence $(V_i)_{i\in\mathbb{N}}$ of sets, a *subsequence of subsets* is an infinite sequence $(V_i')_{i\in\mathbb{N}}$ of sets such that there exists $i_1 < i_2 < \ldots$ with $V_j' \subset V_{i_j}$.

Lemma 10. *For every q and sequence of integers $(n_i')_{i\in\mathbb{N}}$ there exists a sequence of integers $(n_i)_{i\in\mathbb{N}}$ such that if $(V_i)_{i\in\mathbb{N}}$ is a sequence of pairwise disjoint vertex sets of a q-uniform hypergraph H with $|V_i| \geq n_i$ for all i, then we can find a subsequence of subsets $(V_i')_{i\in\mathbb{N}}$ of $(V_i)_{i\in\mathbb{N}}$ such that*

1. *We have $|V_i'| \geq n_i'$ for all i.*
2. *$H[\cup_i V_i']$ is $(q-1)$-homogeneous.*

The proof uses standard (hypergraph) Ramsey-type arguments. Lemma 10 implies the existence of p-homogeneous subhypergraphs for every $1 \leq p < q$ by an easy induction argument. For proofs see the arxiv version of the paper [7].

Corollary 11. *For every $1 \leq p < q$ and a sequence of integers $(n'_i)_{i\in\mathbb{N}}$ there exists a sequence of integers $(n_i)_{i\in\mathbb{N}}$ such that the following holds for every q-uniform hypergraph H. If $(V_i)_{i\in\mathbb{N}}$ is a sequence of pairwise disjoint vertex sets of H with $|V_i| \geq n_i$ for all i, then we can find a subsequence of subsets $(V'_i)_{i\in\mathbb{N}}$ of $(V_i)_{i\in\mathbb{N}}$ such that*

1. *We have $|V'_i| \geq n'_i$ for all i.*
2. *$H[\cup_i V'_i]$ is p-homogeneous.*

2.3 Putting it all together

Proof of Theorem 4. Let H be an infinite q-uniform hypergraph satisfying the fractional Helly property and with arbitrarily large independent sets. We know that H is $M_q^{(q)}(t)$-free with some finite $t \geq q$ by Claim 8. By applying Corollary 11 with $p = 1$ and $n'_i = t$ for all i, we obtain a sequence n_i such that from any pairwise disjoint independent sets $V_1, V_2, \ldots$ of $V(H)$ with $|V_i| \geq n_i$, we can find disjoint independent sets $V'_1, V'_2, \ldots$ with $|V'_i| = t$ such that $H[\cup_i V'_i]$ is 1-homogeneous. (Such large enough disjoint V_i necessarily exist because H contains arbitrarily large independent sets.) Since H is $M_q^{(q)}(t)$-free, and each V'_i is independent, every V'_i has to contain a vertex v_i such that no increasing q-tuple starting at v_i is an edge, otherwise because of p-homogeneity we could obtain a hypergraph from $M_q^{(q)}(t)$. Therefore, the $\{v_i | i \in \mathbb{N}\}$ form an infinite independent set, finishing the proof. □

References

[1] Noga Alon and Gil Kalai. Bounding the piercing number. *Discrete & Computational Geometry*, 13:245–256, 1995.

[2] Noga Alon and Daniel J Kleitman. Piercing convex sets and the Hadwiger-Debrunner (p, q)-problem. *Advances in Mathematics*, 96(1):103–112, 1992.

[3] Imre Bárány and Jiří Matoušek. A fractional Helly theorem for convex lattice sets. *Advances in Mathematics*, 174(2):227–235, 2003.

[4] Andreas F. Holmsen. Large cliques in hypergraphs with forbidden substructures. *Combinatorica*, 40(4):527–537, 2020.

[5] Andreas F. Holmsen and Donggyu Lee. Radon numbers and the fractional Helly theorem. *Isr. J. Math.*, 241(1):433–447, 2021.

[6] Attila Jung and Dömötör Pálvölgyi. k-dimensional transversals to fat convex sets. *arXiv preprint arXiv:2311.15646*, 2024.

[7] Attila Jung and Dömötör Pálvölgyi. A note on infinite versions of (p, q)-theorems. *arXiv preprint*, 2024.

[8] Meir Katchalski and A. Liu. A problem of geometry in $\mathbf{R}^n$. *Proc. Amer. Math. Soc.*, 75(2):284–288, 1979.

[9] Chaya Keller and Micha A Perles. An $(\aleph_0, k+2)$-theorem for k-transversals. *38th International Symposium on Computational Geometry (SoCG 2022) and arXiv preprint arXiv:2306.02181*, 2022–2023.

[10] Dhruv Mubayi and Andrew Suk. A survey of hypergraph Ramsey problems. *Discrete Mathematics and Applications*, pages 405–428, 2020.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Minimal non-EPT graphs based on minimal non-interval graphs

(Extended abstract)

Victoria Kaial* Annegret K. Wagler*

Abstract

Given a set $\mathcal{P}$ of paths over a graph, its *edge intersection graph* is a graph each of whose vertices corresponds to a path in $\mathcal{P}$ and where two vertices are connected if and only if the paths they correspond to share at least one edge. Interval and EPT are respectively the classes of edge intersection graphs over paths and trees. The complete characterization of EPT graphs by minimal forbidden subgraphs is not yet known.

Since interval graphs are in particular EPT and, therefore, every non-EPT graph contains a minimal non-interval graph as an induced subgraph, where the complete list of the latter is known (Lekkerkerker, Boland, 1962), we first address the question of which minimal non-interval graphs are also minimal non-EPT. Some of the previously known minimal non-EPT graphs can be constructed by adding a vertex to a minimal non-interval graph (Golumbic, Jamison, 1985; Alcón, et.al., 2010). We complete the list of all minimal non-EPT graphs that can be obtained by adding exactly one vertex to a minimal non-interval graph.

1 Introduction

Graphs expressing the intersection of some objects play a role in many applications and are accordingly widely studied in the literature, see e.g. [2, 3, 4, 7]. In this paper, we consider intersection graphs of paths in a host graph H: given a collection $\mathcal{P}$ of paths over H, its *edge* (resp. *vertex) intersection graph* $EI(\mathcal{P})$ (resp. $VI(\mathcal{P})$) is a graph G whose vertices are in a one-to-one correspondence with the paths in $\mathcal{P}$ where two vertices are adjacent if and only if the corresponding paths have at least one edge (resp. vertex) of H in common; $(\mathcal{P}, H)$ is said to be an EP (resp. a VP) representation of G. Any graph admitting an EP (resp. a VP) representation in a host tree is said to be an EPT (resp. a VPT) graph and the class formed by all EPT (resp. VPT) graphs is again called EPT (resp. VPT). The study of EPT and VPT graphs was initiated in 1985 by Golumbic & Jamison [5, 6] and by Monma & Wei [10] who also showed that the classes EPT and VPT are incomparable. While the class

*University Clermont Auvergne, Clermont Auvergne INP, LIMOS (UMR 6158 CNRS), Clermont-Ferrand, France. E-mail: victoria.kaial@uca.fr, annegret.wagler@uca.fr

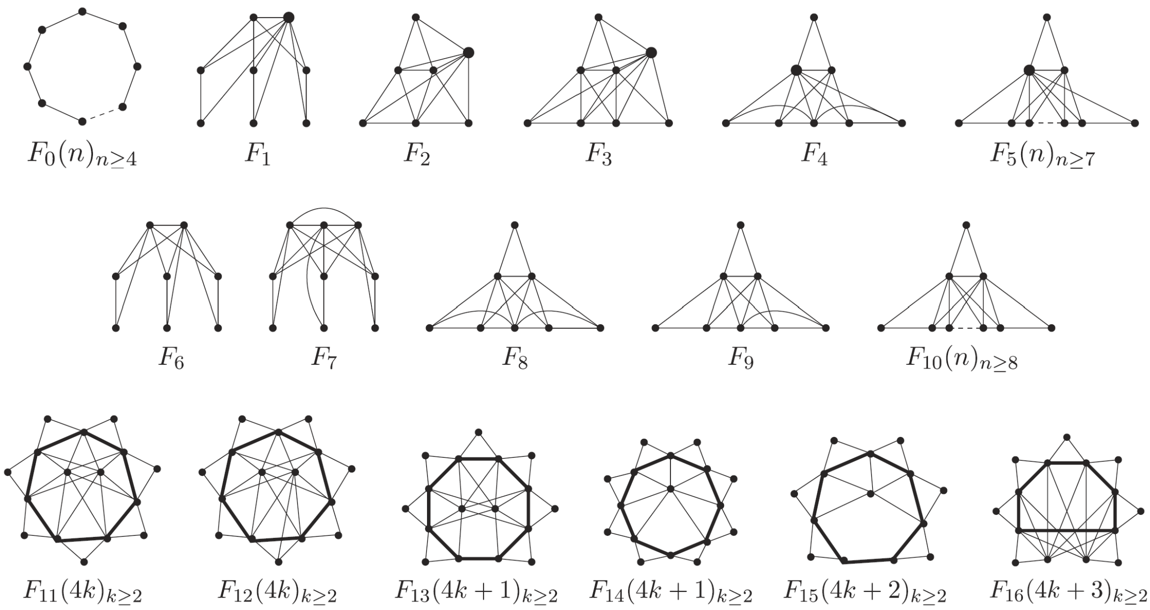


Figure 1: Complete list of minimal non-VPT graphs [9]. In $F_{11},\ldots,F_{16}$, the vertices in the bold cycle form a clique. Figure from [9].

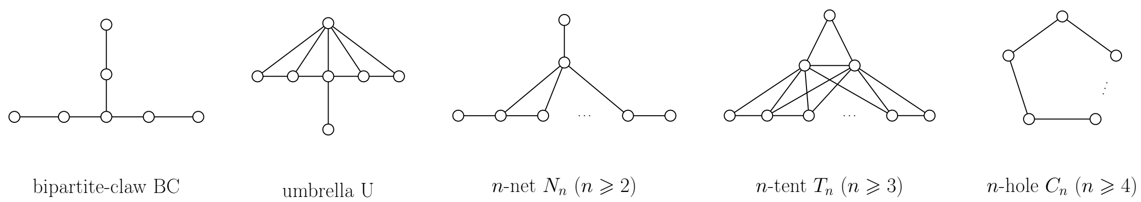


Figure 2: Complete list of all minimal non-interval graphs by [8].

VPT is characterized by a complete list of minimal non-VPT graphs presented by Lévêque et. al [9] (see Figure 1), it has been shown in [6] that recognizing EPT graphs is NP-hard. Determining the complete list of minimal non-EPT graphs remains an open problem.

The aim of this paper is to study minimal non-EPT graphs, based on a characterization of interval graphs by minimal forbidden induced subgraphs due to Lekkerkerker and Boland [8], see Figure 2. In fact, interval graphs are the edge intersection graphs of subpaths over a path and thus form a subclass of EPT graphs. We observe the following:

Observation 1.1. *Every non-EPT graph contains a minimal non-interval graph as an induced subgraph.*

In this paper, we study minimal non-EPT graphs from this perspective. A canonical question in this context is the following:

Problem 1. *Which minimal non-interval graphs are also minimal non-EPT?*

We note that Golumbic and Jamison [6] showed that every n-hole C_n for $n \geq 4$ is EPT (with a unique EPT-representation in a star) and exhaustively answered the question which EPT graphs can be obtained by adding one vertex to a hole as follows:

Theorem 1.2 ([6]). *A graph obtained from an n-hole C_n with $n \geqslant 4$ by adding one vertex v is EPT if and only if v has*

- *at most one neighbor in C_n,*
- *2 or 3 consecutive neighbors on C_n,*
- *2 pairs of 2 consecutive neighbors on C_n.*

In consequence, every graph obtained from an n-hole C_n by adding one vertex in a different way is non-EPT and minimal as all proper induced subgraphs are EPT or even interval graphs. Furthermore, three minimal non-EPT graphs are presented in [1] and analyzing their structure shows that they are obtained from the minimal non-interval graphs 2-net, 3-tent and 4-tent (see Figure 2) by adding a universal vertex (i.e. a vertex being adjacent to all other vertices). Accordingly, we pose the following problem:

Problem 2. *Characterize the minimal non-EPT graphs obtained by adding one vertex to a minimal non-interval graph.*

In Section 2, we present the resulting characterizations and close with some concluding remarks and lines of future research in Section 3.

2 New classes of minimal non-EPT graphs

To address Problem 1, we first recall that n-holes are shown to be EPT by [6], hence it is left to treat the remaining minimal non-interval graphs (see Figure 2). They are all chordal, i.e., without an n-hole C_n for $n \geq 4$ as an induced subgraph. Syslo showed in [11] that all chordal EPT graphs are VPT, and Monma & Wei [10] further proved that the class of chordal EPT graphs coincides with $VPT \cap EPT$. This implies:

Observation 2.1. *Every chordal non-VPT graph is non-EPT.*

Moreover, we can prove the following property of minimal non-EPT graphs:

Lemma 2.2. *Every minimal non-EPT graph is 2-connected.*

Combining these facts with the complete list of minimal non-VPT graphs from [9] in Figure 1 enables us to prove:

Theorem 2.3. *The only minimal non-interval minimal non-EPT graphs are n-tents for $n \geqslant 5$.*

Proof (sketch) We treat all members of the complete list of minimal non-interval graphs from [8] (see Figure 2):

- Bipartite claw BC, umbrella U and all n-nets N_n for $n \geq 2$ are not 2-connected and thus not minimal non-EPT by Lemma 2.2.
- EPT representations of 3- and 4-tent can be constructed easily.
- All n-tents for $n \geq 5$ are minimal non-VPT by [9] (see Figure 1, graphs $F_{10}(n)$, $n \geqslant 8$), thus also non-EPT by Observation 2.1, and minimal as all proper induced subgraphs are interval graphs.

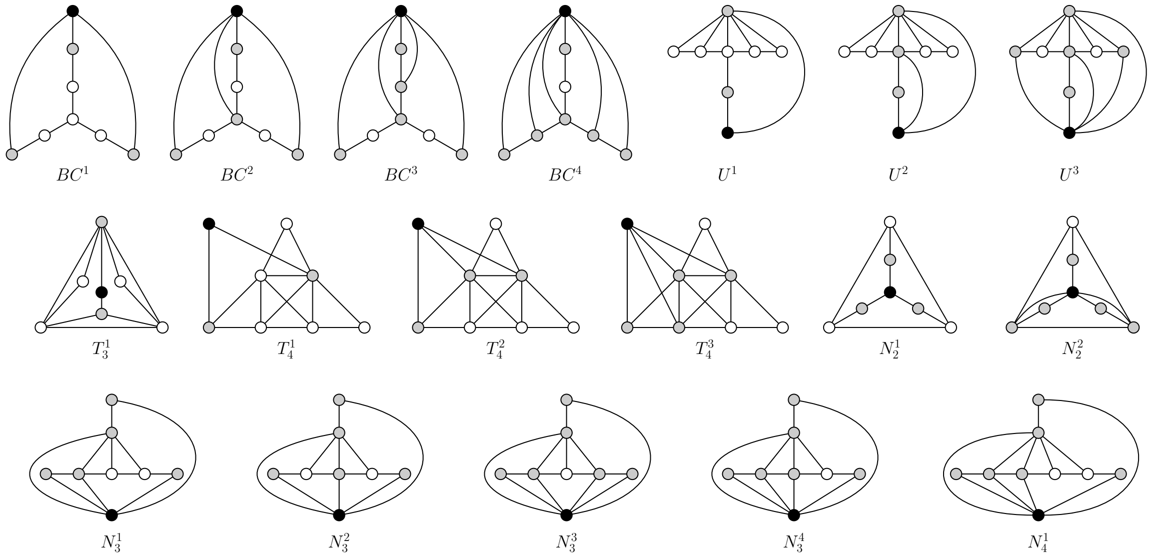


Figure 3: All minimal non-EPT graphs $G \oplus v$ (where v is not universal) for G bipartite-claw, umbrella, n-tent $n \geqslant 3$) and n-net ($n \geqslant 2$) in that order from top left to bottom right. The vertex v is indicated in black and its neighborhood in gray. Note the following pairs are isomorphic: $BC^4 \cong U^1$, $T_3^1 \cong N_2^2$, and $T_4^1 \cong N_3^4$.

- All n-holes C_n for $n \geq 4$ are shown to be EPT by [6]. □

We next address the question which minimal non-EPT graphs can be obtained by adding one vertex v to a minimal non-interval but EPT graph G, denoted by $G \oplus v$.

Theorem 2.4. *The minimal non-EPT graphs $G \oplus v$ obtained by adding one vertex v to a minimal non-interval graph G are exactly the graphs where G is*

- *an n-hole and the neighborhood of v is different from the cases in Theorem 1.2;*
- *a bipartite claw, an umbrella, a 3-tent, a 4-tent or an n-net for $n \geq 2$ and either v is universal or $G \oplus v$ is one of the graphs listed in Figure 3.*

Proof (sketch) We have to treat all minimal non-interval EPT graphs G, i.e., all minimal non-interval graphs except n-tents with $n \geq 5$ by Theorem 2.3.

If G is a hole C_n, we deduce the assertion from Theorem 1.2 by [6]. For all other minimal non-interval EPT graphs $G \neq C_n$,

- if v is a universal vertex or $G \oplus v$ is T_4^3 or U^2, the graph $G \oplus v$ is chordal and minimal non-VPT (see graphs $F_1, \ldots F_6$ and F_9 in Figure 1 by [9]), hence also non-EPT by Observation 2.1;
- otherwise, $G \oplus v$ is minimal non-EPT if and only if it is one of the non-chordal graphs listed in Figure 3.

In all cases, minimality follows from the fact that all proper induced subgraphs are EPT or even interval graphs.

All other such graphs $G \oplus v$ are either EPT (e.g. if they are not 2-connected) or contain a minimal non-EPT graph as proper induced subgraph. This can be shown for G being a bipartite claw, an umbrella, a 3-tent, a 4-tent or an n-net with $n \in \{2, 3, 4\}$ by a case analysis and for n-nets with $n \geq 5$ with the help of some technical lemmas. □

3 Concluding remarks

In this work, we studied minimal non-EPT graphs in relation to their minimal non-interval subgraphs and addressed the two problems of characterizing the minimal non-interval graphs that are also minimal non-EPT and all minimal non-EPT graphs obtained from a minimal non-interval but EPT graph by adding one vertex. We showed that the n-tents for $n \geq 5$ are exactly the minimal non-interval minimal non-EPT graphs and provided a complete list of all minimal non-EPT graphs obtained from a minimal non-interval but EPT graph, i.e., from bipartite claw, umbrella, 3-tent, 4-tent, n-nets for $n \geq 2$ and n-holes for $n \geq 4$, by adding one vertex. We note that adding a universal vertex to all of them except C_4 yields a minimal non-EPT graph, and that for any n-net with $n \geq 5$, only adding a universal vertex results in a minimal non-EPT graph.

Our lines of future research include finding and describing more minimal non-EPT graphs, e.g., to characterize all minimal non-EPT graphs containing a universal vertex, further minimal non-EPT graphs obtained from minimal non-interval graphs by adding two or more vertices or by applying some operations to known minimal non-EPT graphs.

References

[1] Liliana Alcón, Marisa Gutierrez, and María Pía Mazzoleni. A necessary condition for EPT graphs and a new family of minimal forbidden subgraphs. *Matemática Contemporânea*, 39(13), 2010.

[2] Liliana Alcón, Maria Mazzoleni, and Tanilson Santos. Relationship among B1-EPG, VPT and EPT graphs classes. *Discussiones Mathematicae Graph Theory*, 43(3):841, 2023.

[3] Guillermo Durán, Luciano N. Grippo, and Martín D. Safe. Structural results on circular-arc graphs and circle graphs: a survey and the main open problems. *Discrete Applied Mathematics*, 164:427–443, 2014.

[4] Martin Charles Golumbic. *Algorithmic graph theory and perfect graphs.* Number 57 in Annals of discrete mathematics. Elsevier, Amsterdam, 2004.

[5] Martin Charles Golumbic and Robert E. Jamison. Edge and vertex intersection of paths in a tree. *Discrete Mathematics*, 55(2):151–159, 1985.

[6] Martin Charles Golumbic and Robert E. Jamison. The edge intersection graphs of paths in a tree. *Journal of Combinatorial Theory, Series B*, 38(1):8–22, 1985.

[7] Victoria Kaial, Hervé Kerivin, and Annegret K. Wagler. On non-superperfection of edge intersection graphs of paths. *Discrete Optimization*, 54:100857, 2024.

[8] Cornelis Lekkerkerker and Johan Boland. Representation of a finite graph by a set of intervals on the real line. *Fundamenta Mathematicae*, 51(1):45–64, 1962.

[9] Benjamin Lévêque, Frédéric Maffray, and Myriam Preissmann. Characterizing path graphs by forbidden induced subgraphs. *Journal of Graph Theory*, 62(4):369–384, 2009.

[10] Clyde L. Monma and Victor K Wei. Intersection graphs of paths in a tree. *Journal of Combinatorial Theory, Series B*, 41(2):141–181, 1986.

[11] Maciej M. Sysło. Triangulated edge intersection graphs of paths in a tree. *Discrete Mathematics*, 55(2):217–220, 1985.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The α-representation for Tait coloring and sums over spanning trees

(Extended abstract)

Ilyas Kalimullin* Eduard Lerner†

Abstract

Consider a connected pseudograph H such that each edge is associated with weight x_e, $x_e \in \mathbb{F}_3$; $\mathcal{T}(H)$ is the set of spanning trees of graph H. Assume that $s(H;\mathbf{x}) = \sum_{T\in\mathcal{T}(H)} \prod_{e\in E(T)} x_e$. Let G be a maximal planar graph (arbitrary planar triangulation) such that each face F is assigned the value $\alpha(F) = \pm 1 \in \mathbb{F}_3$. Then we can associate each edge with $x_e = \alpha(F'_e) + \alpha(F''_e)$, where F'_e and F''_e are the faces containing edge e. Let us define the value $w_G(\mathbf{x})$ as $\left(\frac{s(G/W^*(\mathbf{x});\mathbf{x})}{3}\right)/(-3)^{(|V(G/W^*(\mathbf{x}))|-1)/2}$; here $\left(\frac{x}{3}\right)$ is the Legendre symbol, G/W is the graph with the contracted set of vertices W, while $W^*(\mathbf{x})$ is a set of vertices W, $W \subseteq V(G)$, with minimal cardinality such that $s(G/W;\mathbf{x})$ differs from zero. In the following, we prove that the number of Tait colorings for graph G equals the tripled sum $w_G(\mathbf{x}(\alpha))$ with respect to all possible vectors $\alpha \in \{-1,1\}^{\mathcal{F}(G)}$ such that $G/W^*(\mathbf{x}(\alpha))$ has an odd number of vertices, where $\mathcal{F}(G)$ is the set of faces of graph G.

1 Introduction

The idea of this work has a long history. Let notation $\mathcal{T}(G)$ stand for the set of spanning trees of connected graph G. Consider sums

$$s(G;\mathbf{x}) = \sum_{T\in\mathcal{T}(G)} \prod_{e\in E(T)} x_e, \tag{1}$$

where x_e are elements of finite field $\mathbb{F}_q$. In December 1997, when giving a talk at the Gelfand Seminar at Rutgers University, Maxim Kontsevich proposed the conjecture that the number of non-zero values of (1) for $\mathbf{x} \in \mathbb{F}_q^{E(G)}$ is a polynomial with respect to q. This conjecture was inspired by studying analogous sums (with real positive x_e) in quantum field theory. Although this conjecture was never published, it has aroused the interest of experts in combinatorics

*Student of Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kremlevskaya ul. 18, Kazan, 420008 Russia. E-mail: `ilyaskalimullinn@yandex.ru`.

†Institute of Mathematics and Mechanics, Kazan Federal University, Kremlevskaya ul. 18, Kazan, 420008 Russia. E-mail: `eduard.lerner@gmail.com`. Supported by RSF (project No. 24-21-00158).

(see [18, 5, 19]). Sometime later, this conjecture was refuted [4]. Note that in the refuted conjecture one actually considers the sum of weights $w_G(\mathbf{x})$ that equal one when $s(G;\mathbf{x})$ differs from zero. A proper weight $w_G(\mathbf{x})$ is related to the value of a multidimensional Gaussian sum (an analog of the Gaussian integral) over a finite field such that the Laplace–Kirchhoff matrix of quadratic form is parameterized by values $\mathbf{x}$. For example, in the case of field $\mathbb{F}_3$, an analog of Gaussian integral obeys formula (2) given below. If we apply these formulas and carefully adjust the techniques of the so-called α-representation, which are used in quantum theory in the case of a real field, to the case of a finite one, then we obtain a new representation for the flow polynomial of graph [8]. Moreover, this representation allows for a generalization for the case of an arbitrary matroid representable over field $\mathbb{F}_q$ [10]. In the case of a regular matroid, this formula is even simpler.

The goal of this paper is to obtain the α-representation for the number of Tait colorings for an arbitrary maximal planar graph. For the cubic graph dual to the considered one, this representation was recently obtained in [10]. In our case, there occur sums with respect to spanning trees, which unites this representation with the initial Kontsevich conjecture.

Let us now state the main result. We will consider not only the sums (1) for initial simple graphs G, but also sums $s(H;\mathbf{x})$ for pseudographs H obtained from graph G by contracting all the vertices that belong to set W, $W \subseteq V(G)$; denote the pseudograph H as G/W. Evidently, the value $s(H;\mathbf{x})$ is independent of loops of graph H (as distinct from multiple edges). Let notation $W^*(\mathbf{x})$ stand for an arbitrary set of vertices W with minimal cardinality such that sum $s(G/W;\mathbf{x})$ differs from zero.

In what follows, we consider only the field $\mathbb{F}_3$ of three elements $\{-1,0,1\}$. In such notation of elements, the *Legendre symbol* $\left(\frac{x}{3}\right)$, $x \in \mathbb{F}_3$, coincides with the corresponding real value x. Assume that for arbitrary $\mathbf{x} \in \mathbb{F}_3^{E(G)}$ the weight $w_G(\mathbf{x})$ obeys the formula

$$w_G(\mathbf{x}) = \left(\frac{s(G/W^*(\mathbf{x});\mathbf{x})}{3}\right) / (-3)^{(|V(G/W^*(\mathbf{x}))|-1)/2}. \qquad (2)$$

Here $\sqrt{-3} = -i\sqrt{3}$, although this fact does not affect the statement of the main theorem, because it does not contain odd powers of (-3). In the following, we prove that the weight $w_G(\mathbf{x})$ is independent of the choice of sets $W^*(\mathbf{x})$ (with equal cardinalities).

Let G be *a maximal planar graph*, i.e., a planar graph such that each face is a triangle. Recall that *a Tait coloring* is a coloring of all edges of graph G in three colors so that edges of one and the same face are colored differently. The existence of such a coloring for any G is equivalent to the assertion of the Four Color Theorem. Denote *the number of Tait colorings* for graph G as $\mathrm{Tai}(G)$. Evidently, $\mathrm{Tai}(G)$ is a multiple of 3. We need the value $\mathrm{Tai}_0(G) = \mathrm{Tai}(G)/3$.

Let notation $\mathcal{F}(G)$ stand for *the set of faces* of graph G. Let us associate each face F with the variable $\alpha(F)$ which takes on values in the set $\{1,-1\}$ of invertible elements of field $\mathbb{F}_3$ (*below we denote this set as* $\mathbb{F}_3^*$). The vector $(\alpha(F), F \in \mathcal{F}(G))$ corresponds to $\mathbf{x}(\alpha) = (x_e, e \in E(G))$, where $x_e = \alpha(F_e') + \alpha(F_e'')$; here F_e' and F_e'' are faces containing edge e.

Theorem 1. *The following formula is valid:* $\mathrm{Tai}_0(G) = \sum w(G;\mathbf{x}(\alpha))$*; the sum is calculated with respect to all vectors* $\alpha \in (\mathbb{F}_3^*)^{\mathcal{F}(G)}$ *such that* $G/W^*(\mathbf{x}(\alpha))$ *has an odd number of vertices.*

For example, in the case of K_4, there are 16 values $\alpha \in (\mathbb{F}_3^*)^{\mathcal{F}(K_4)}$ that fall into three distinct cases (see Fig. 1). If α is the same for all faces, then no vertex contraction is needed, and we obtain $K_4/W^* = K_4$ with an even number of vertices. If α differs for only one face, the vertices still do not contract. If two faces have the value $+1$ and the other two have -1, then exactly two vertices must be contracted, resulting in K_4/W^* having three vertices. In this case, we obtain $w(K_4; \mathbf{x}(\alpha)) = \left(\frac{-1}{3}\right)/(-3) = \frac{1}{3}$. Thus, $\mathrm{Tai}_0(K_4) = \binom{4}{2} \cdot \frac{1}{3} = 2$.

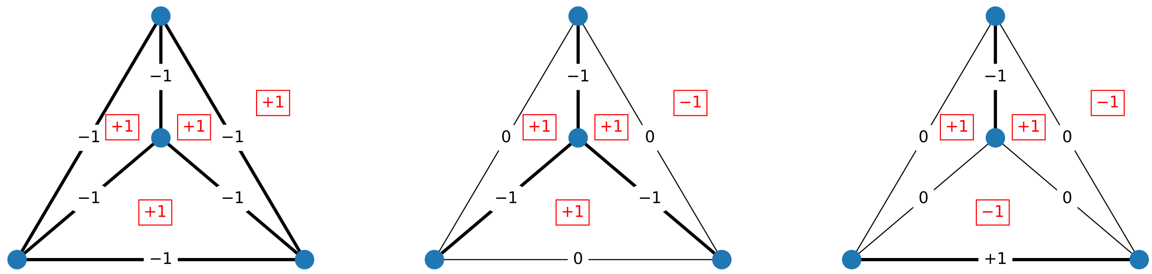


Figure 1: Three cases of the values of α for K_4

Note that the Four Color Theorem allows for an elegant algebraic statement in terms of the *graph polynomial* of G (this notion was introduced by N.Alon and M. Tarsi in [1], though the main variant of Theorem 1.1 in this paper was proposed earlier by Yu. V. Matiyasevich in [12]). In particular $\mathrm{Tai}(G)$ coincides with certain coefficients of the graph polynomial of the line graph of $\widetilde{G}$, where $\widetilde{G}$ is the dual graph to G (see [16], [1], [13]). See related bibliographic references in [14].

2 Gaussian sums and the Laplace–Kirchhoff matrix

In the case of a real field, the application of the classical α-representation implies the use of explicit formulas for the calculation of Gaussian integrals with the imaginary unit in the exponent. In the case of finite fields, we use multidimensional Gaussian sums. The statement of Theorem 1 contains explicit formulas that are valid in the case of field $\mathbb{F}_3$.

Assume that C is an arbitrary symmetric matrix $n \times n$, whose elements belong to $\mathbb{F}_3$, and $\mathbf{y}^T C \mathbf{y}$ is a quadratic form with this matrix ($\mathbf{y}$ is a vector column of the corresponding dimension). The following sum represents an analog of the multidimensional Gaussian integral:

$$\mathrm{Gau}(C) = \sum\nolimits_{\mathbf{y} \in \mathbb{F}_3^n} \exp(2\pi i\, \mathbf{y}^T C \mathbf{y}/3).$$

In the case $n = 1$, we get the so-called quadratic Gaussian sum $g(c) = \sum_{y \in \mathbb{F}_3} \exp(2\pi i\, cy^2/3)$. By elementary calculations, we make sure that $g(0) = 3$, otherwise $g(c) = \left(\frac{c}{3}\right) i\sqrt{3}$. See [7] for the historical background of the calculation of the quadratic Gaussian sum for an arbitrary field $\mathbb{F}_p$, where p is a prime number, $p > 2$; see [11, Theorem 5.15] for the general case of field $\mathbb{F}_q$, $q = p^d$.

Remark 1. *If matrices C and A are congruent, i.e., $A = P^T C P$, where P is a non-singular $n \times n$ matrix, then* $\mathrm{Gau}(C) = \mathrm{Gau}(A)$.

Remark 1 is valid because by putting $\mathbf{y}' = P\mathbf{y}$ we reduce the sum $\mathrm{Gau}(A)$ to $\mathrm{Gau}(C)$.

Lemma 1 (a particular case of Lemma 8 in [8] for the field $\mathbb{F}_3$)**.** *Consider a symmetric $n \times n$ matrix C of rank r, whose elements belong to $\mathbb{F}_3$; let $\det C_r$ be an arbitrary non-zero principal* *minor of order r of the matrix C. The following formula is valid:*

$$\frac{\mathrm{Gau}(C)}{3^n} = \left(\frac{\det C_r}{3}\right)\left[\frac{i}{\sqrt{3}}\right]^r. \tag{3}$$

In view of Remark 1 the proof of Lemma 1 is reduced to considering the diagonal case (see [17, Chapters IV] for the reduction of a quadratic form over a finite field to the diagonal form). By factorization, we reduce the diagonal case to the one-dimensional variant considered above. The multiplicative property of the Legendre symbol allows us to write the final result in the form (3).

Corollary 1. *For any symmetric matrix C of rank r, the value $\left(\frac{\det C_r}{3}\right)$ is independent of the choice of C_r.*

Corollary 2. *Assume that all elements of a symmetric matrix C are linear functions of a certain collection of variables $\alpha \in (F_3^*)^k$, while $r(C(\alpha))$ is the rank of this matrix. Then* $\sum_{\substack{\alpha:\ \alpha\in(F_3^*)^k,\\ r(C(\alpha)) \bmod 2=1}} \mathrm{Gau}(C(\alpha)) = 0.$

Proof. Let us replace α in the sum under consideration by $-\alpha$. Note that when calculating $\mathrm{Gau}(C(-\alpha))$ we replace the sign of the value $\det C_r$ with the opposite (i.e., use the term $(-1)^r \det C_r$). Therefore, the considered sum equals itself with the opposite sign. □

Let now G be an arbitrary multigraph and let $L(G;\lambda)$ be a weighted Laplace–Kirchhoff matrix of graph G, i.e., $L(G;\lambda) = B\Lambda B^T$, where B is the oriented incidence matrix, while Λ is the diagonal matrix, whose diagonal elements are equal to λ_e.

Lemma 2 (cf. Theorem 6 in [8])**.** *Let $C = L(G;\mathbf{x})$. Then the right-hand side of formula* (3) *coincides with the right-hand side of formula* (2)*.*

Proof. The principal minor of the matrix $L(G;\mathbf{x})$ equals the determinant of the submatrix obtained from $L(G;\mathbf{x})$ by deleting rows and columns, whose indices belong to the set W. According to Lemma 1, it suffices to prove that this minor coincides with $s(G/W;\mathbf{x})$. This fact is well known in the case where $W = \emptyset$ (see, for example, [2], as well as [18] and references therein). The general case is reduced to a particular one, because the matrix $L(G/W;x)$ is obtained from $L(G;x)$ by deleting the rows and columns corresponding to the vertices in W, and adding a row and column corresponding to the resulting contracted vertex. These new elements of the matrix $L(G/W;x)$ are fully determined by the remaining part of the matrix because the sum of the elements of any row or column in this matrix equals zero. □

3 The Heawood theorem and the Fourier transform

To prove Theorem 1 we need the Heawood representation for $\mathrm{Tai}_0(G)$ as the number of nowhere-zero solutions of a system of linear equations over $\mathbb{F}_3$.

Proposition 1 ([6])**.** *Let G be a maximal planar graph. Let us associate each face F of graph G with variable $\sigma(F)$, which takes on values in set $\mathbb{F}_3^*$. Then $\mathrm{Tai}_0(G)$ equals the number of all possible collections of spins $(\sigma(F), F \in \mathcal{F}(G))$, such that for any graph vertex v, the sum $\sigma(F)$ calculated with respect to all faces F that contain v, equals zero.*

See also [15, Theorem 9.3.4] for the proof of this proposition; see [3] (as well as [9]) for its other proofs.

We also make use of some simple properties of the Fourier transform over field $\mathbb{F}_3$. Consider complex-valued functions $f(k)$, whose argument k belongs to field $\mathbb{F}_3$. The inverse Fourier transform of such functions is usually understood as the function $\widehat{f}(y) = \sum\limits_{k\in\mathbb{F}_3} f(k)\frac{\exp(2\pi i\, ky/3)}{3}$.

Let $\mathbf{1}(k)$ be the function $f(k)$ that is identically equal to one; let symbol $\delta(y)$ denote the delta function (the Kronecker symbol): $\delta(0) = 1$, $\delta(y) = 0$ with all $y \in \mathbb{F}_3^*$. We can easily make sure that

$$\widehat{\mathbf{1}}(y) = \delta(y). \tag{4}$$

Relation (4) implies the following formula, which plays an important role in further calculations:

$$\sum\nolimits_{y\in\mathbb{F}_3^*} \exp(2\pi i\, ky/3) = 3\delta(k) - 1 = 3\delta(k^2) - 1 = \sum\nolimits_{a\in\mathbb{F}_3^*} \exp(2\pi i\, k^2 a/3). \tag{5}$$

Remark 2. *In the case of a finite field, we can consider the function $f(k) = 1 - \delta(k)$ as the norm of an element of the finite field. Thus, the sum $\sum_{y\in\mathbb{F}_p^*} \exp(2\pi i\, ky/p)$ is the Fourier transform of the norm raised to a certain power. An analog of this sum in the case of a real field is the Fourier transform of a (generalized) function $|k|^\gamma$. In quantum field theory, such functions often represent the so-called propagators of Feynman amplitudes. The paper [20, p. 691] has given rise to the parametric representation of the integrals of propagators of Feynman amplitudes as integrals of the exponent. In the mentioned paper, K. Symanzik uses the symbol α for the analog of variable a introduced by us. The technique based on the use of this representation in quantum field theory (in the next section we implement its simplified analog for $\mathbb{F}_3$) is called the α-representation. Following this tradition, we will use the same notation.*

4 The α-representation for $\mathrm{Tai}_0(G)$

Proof of theorem 1. According to the Heawood theorem (Proposition 1),

$$\mathrm{Tai}_0(G) = \sum_{\sigma\in\{-1,1\}^{\mathcal{F}(G)}} \prod_{v\in V(G)} \delta\Big(\sum_{F:v\in F} \sigma(F)\Big). \tag{6}$$

Let us modify the right-hand side of formula (6), using the fact that each δ-function represents the inverse Fourier transform of the function $\mathbf{1}(\cdot)$ (see (4)). Representing the product of exponents as the exponent of the sum and changing the summation order, we conclude that

$$\mathrm{Tai}_0(G) = \sum_{\mathbf{k}\in\mathbb{F}_3^{V(G)}} \sum_{\sigma\in\{-1,1\}^{\mathcal{F}(G)}} \exp\left(\frac{2\pi i}{3} \sum_{v\in V(G)} k_v \sum_{F:v\in F} \sigma(F)\right) /3^{|V(G)|}.$$

We can represent the sum in the exponent in another way, namely,

$$\sum_{v\in V(G)} k_v \sum_{F:v\in F} \sigma(F) = \sum_{F\in\mathcal{F}(G)} \sigma(F) \sum_{v\in F} k_v.$$

This allows us to use the formula (5). We obtain the relation

$$\mathrm{Tai}_0(G) = \sum_{\mathbf{k}\in\mathbb{F}_3^{V(G)}} \sum_{\alpha\in\{-1,1\}^{\mathcal{F}(G)}} \exp\left(\frac{2\pi i}{3} \sum_{F\in\mathcal{F}(G)} \alpha(F) \left(\sum_{v\in F} k_v\right)^2\right) /3^{|V(G)|}.$$

The sum in the exponent can be expressed differently, specifically,

$$\sum_{F\in\mathcal{F}(G)} \alpha(F)\left(\sum_{v\in F} k_v\right)^2 = \sum_{v_1\in V}\sum_{v_2\in V} k_{v_1}k_{v_2} \sum_{\substack{F\in\mathcal{F}(G):\\ v_1\in F, v_2\in F}} \alpha(F).$$

Taking into account the fact that $2(\alpha(F'_e)+\alpha(F''_e)) = -(\alpha(F'_e)+\alpha(F''_e))$, we get the relation

$$\mathrm{Tai}_0(G) = \sum_{\alpha\in\{-1,1\}^{\mathcal{F}(G)}} \sum_{\mathbf{k}\in\mathbb{F}_3^{V(G)}} \frac{\exp(2\pi i\,\mathbf{k}\,L(\mathbf{x}(-\alpha))\,\mathbf{k}^T\ /3)}{3^{|V(G)|}} = \sum_{\alpha\in\{-1,1\}^{\mathcal{F}(G)}} \frac{\mathrm{Gau}(L(\mathbf{x}(-\alpha)))}{3^{|V(G)|}}$$

(here $\mathbf{k} = (k_v, v\in V(G))$). Note that we iterate over all possible values of α, thus we can replace $\mathbf{x}(-\alpha)$ with $\mathbf{x}(\alpha)$. In view of Corollary 2 and Lemma 2 this equality is equivalent to the assertion of Theorem 1. □

5 Conclusion

In the case of finite fields, the base of the α-representation is the interpretation of the desired value as the number of nowhere-zero solutions to a system of linear equations. For the number of Tait colorings, this base is ensured by the Heawood theorem. The application of formula (5) allows us to reduce further calculations to the evaluation of multidimensional Gaussian sums, for which the principal minors of the matrix of the quadratic form can be expressed explicitly in terms of the Legendre symbol. In the case of Tait colorings, we can visually interpret these minors as the sum with respect to spanning trees. We plan to further develop the α-representation technique in this evident case, which is also related to the assertion of the Four Color Theorem.

References

[1] N. Alon and M. Tarsi. A note on graph colorings and graph polynomials. *Journal of combinatorial theory, Series B*, 70:197–201, 1997.

[2] R. B. Bapat. *Graphs and Matrices*. Universitext. Springer, London, second edition, 2014.

[3] E. G. Belaga. Mod 3 arithmetic on triangulated Riemann surfaces. *Theoretical Computer Science*, 263:123–137, 2001.

[4] P. Belkale and P. Brosnan. Matroids, motives, and a conjecture of Kontsevich. *Duke Mathematical Journal*, 116(1):147–188, 2003.

[5] F. Chung and C. Yang. On polynomials of spanning trees. *Annals of Combinatorics*, 4(1):13–25, 2000.

[6] P. J. Heawood. On the four-colour map theorem. *Quarterly Journal of Pure and Applied Mathematics*, 29:270–285, 1898.

[7] K. Ireland and M. I. Rosen. *A Classical Introduction to Modern Number Theory*, volume 84 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 1990.

[8] A. P. Kuptsov, E. Yu. Lerner, and S. A. Mukhamedjanova. Flow polynomials as Feynman amplitudes and their α-representation. *Electronic Journal of Combinatorics*, 24(1):Paper 11, 19 pp., 2017.

[9] E. Yu. Lerner. The Heawood approach to Tait colorings and defining vertex sets. 2024. `arXiv:2411.15992`.

[10] E. Yu. Lerner. The α-representation for the tait coloring and for the characteristic polynomial of matroid, 2025. `arXiv:2503.14285`.

[11] R. Lidl and H. Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1997.

[12] Yu. V. Matiyasevich. A criterion for vertex colorability of a graph stated in terms of edge orientations. *Diskretnyi Analiz*, 26:65–71, 1974. `arXiv:0712.1884`.

[13] Yu. V. Matiyasevich. Some algebraic methods for calculation of the number of colorings of a graph. *Zapiski Nauchnykh Seminarov POMI*, 293:193–205, 2001.

[14] Yu. V. Matiyasevich. Four color theorem from three points of view. *Illinois Journal of Mathematics*, 60:185–205, 2016.

[15] O. Ore. *The Four-Color Problem.* Academic Press, New York and London, 1967.

[16] D. E. Scheim. The number of edge 3-colorings of a planar cubic graph as a permanent. *Discrete Mathematics*, 8:377–382, 1974.

[17] J.-P. Serre. *A Course in Arithmetic.* Springer-Verlag, New York, 1973.

[18] R. P. Stanley. Spanning trees and a conjecture of Kontsevich. *Annals of Combinatorics*, 2(4):351–363, 1998.

[19] J. R. Stembridge. Counting points on varieties over finite fields related to a conjecture of Kontsevich. *Annals of Combinatorics*, 2(4):365–385, 1998.

[20] K. Symanzik. Dispersion relations and vertex properties in perturbation theory. *Progress of Theoretical Physics*, 20(5):690–702, 1958.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Distance 2 versions and relative versions for difference patterns

(Extended abstract)

Thomas Karam*

Abstract

We provide two applications of a conceptually simple covering technique to density theorems and conjectures on patterns in sets involving set differences: (i) analogues of these statements to distance 2 versions of the pattern, and (ii) reduction of these statements to relative versions.

1 Introduction

The results described below are proved and discussed further in [12] together with some more results. Throughout this paper we will use the following notations. If n is a positive integer then we will write $[n]$ for the set $\{1, \dots, n\}$ of positive integers between 1 and n. We will often use notations such as $[n]^{d_1} \cup \dots \cup [n]^{d_s}$, where $s, d_1, \dots, d_s$ are positive integers; unless stated otherwise they will always be understood as disjoint unions: for instance, $[n] \cup [n]$ will refer not to $[n]$ but to the disjoint union of two copies of $[n]$. We will also often write a subset $A \subset [n]^{d_1} \cup \dots \cup [n]^{d_s}$ as $A_1 \cup \dots \cup A_s$. In such decompositions it will be implicit that A_i is a subset of the ith part $[n]^{d_i}$ of the union for every i, unless stated otherwise.

1.1 Background on patterns in set systems

Density theorems in combinatorics (of which Szemerédi's theorem from 1975 is a celebrated representative example) are still far from completely understood even purely qualitatively. For instance, Gowers [7] describes the following conjecture (in its version with $d_1 = 1, \dots, d_s = s$) as a central open problem in Ramsey theory.

Conjecture 1.1. *Let $k, s, d_1, \dots, d_s$ be positive integers and let $\delta > 0$. If n is large enough depending on $k, s, d_1, \dots, d_s, \delta$ only then, for every subset A of the set $[k]^{[n]^{d_1}} \times \dots \times [k]^{[n]^{d_s}}$ with density at least δ there is some $\emptyset \neq S \subset [n]$ and some $y \in [k]^{[n]^{d_1} \setminus S^{d_1}} \times \dots \times [k]^{[n]^{d_s} \setminus S^{d_s}}$ such that whenever the coordinates of x are the same within each of the sets $S^{d_1}, \dots, S^{d_s}$, and coincide with those of y outside these sets, we have that $x \in A$.*

*Mathematical Institute, University of Oxford. Email: thomas.karam@maths.ox.ac.uk. Supported by ERC Advanced Grant 883810.

One motivation to consider Conjecture 1.1 is that it would simultaneously imply both the Bergelson-Leibman theorem [3] and the density Hales-Jewett theorem [5], [6] [14], two generalisations of Szemerédi's theorem which respectively replace arithmetic progressions by polynomial progressions and subsets of integers by higher-dimensional sets.

Since Conjecture 1.1 implies all three theorems, the proofs of which are difficult, it appears natural to attempt to look for new basic difficulties that do not arise in any of these three theorems and treat them in isolation before attempting a solution to Conjecture 1.1. Taking $k = 2$ in Conjecture 1.1 is such a step. First, it does away with the structure of arithmetic progression and its accompanying difficulties, and the special cases $k = 2$ of the three theorems are degenerate or at least much easier to prove. Second, the special case $k = 2$ of Conjecture 1.1, which we are about to state, appears to on its own be a difficult problem.

Conjecture 1.2. *Let $s, d_1, \dots, d_s$ be positive integers and let $\delta > 0$. Then, for n large enough depending on $s, d_1, \dots, d_s, \delta$ only, every subset of $\mathcal{P}([n]^{d_1} \cup \dots \cup [n]^{d_s})$ that has density at least δ contains all 2^s sets $A \cup \bigcup_{r \in T} S^{d_r}$ where $T \subset [s]$, the unions are disjoint unions, and the sets $\emptyset \neq S \subset [n]$ and $A \subset [n]^{d_1} \cup \dots \cup [n]^{d_s}$ are common to all 2^s sets.*

On the way to Conjecture 1.2, we may first ask for two of the required 2^s sets, and to obtain two such sets it suffices to consider the case where one of the points is obtained by taking T to be empty. Regardless of the choice of the second point, Conjecture 1.2 in turn specialises as follows (with parameters $s, d_1, \dots, d_s$ that might be lower than originally).

Conjecture 1.3. *Let $s, d_1, \dots, d_s$ be positive integers and let $\delta > 0$. Then, for n large enough depending on $s, d_1, \dots, d_s, \delta$ only, every subset of $\mathcal{P}([n]^{d_1} \cup \dots \cup [n]^{d_s})$ that has density at least δ contains a pair (A, B) of distinct subsets of $[n]^{d_1} \cup \dots \cup [n]^{d_s}$ such that $A \subset B$ and $B \setminus A = S^{d_1} \cup \dots \cup S^{d_s}$ for some $S \subset [n]$.*

1.2 Main results

With Conjecture 1.3 in sight, we may define an *oriented* graph with vertex set $\mathcal{P}([n]^{d_1} \cup [n]^{d_2} \cup \dots \cup [n]^{d_s})$, and join A to B by an edge if (A, B) is as required in the conclusion of 1.3. The statement of Conjecture 1.3 can then be reformulated as stating that every dense subset $\mathcal{A}$ of the vertex set contains (for n large enough) a pair of vertices that is joined by an edge. We shall not prove this but we shall instead prove that $\mathcal{A}$ must contain a pair (A, B) of vertices such that for some U (not necessarily in $\mathcal{A}$) both (U, A), (U, B) are edges of the oriented graph. In particular, in the nonoriented version of the presently defined graph, the vertices A and B are at distance at most 2 from one another. Our first result is a "distance at most 2" analogue of Conjecture 1.3 that we previously described.

Theorem 1.4. *Let $s, d_1, \dots, d_s, m$ be positive integers and let $\delta > 0$. Then, for n large enough depending on $s, d_1, \dots, d_s, m, \delta$ only, every subset $\mathcal{A}$ of $\mathcal{P}([n]^{d_1} \cup \dots \cup [n]^{d_s})$ that has density at least δ contains a pair (A, B) of distinct subsets such that for some $U \in \mathcal{P}([n]^{d_1} \cup \dots \cup [n]^{d_s})$ (not necessarily in $\mathcal{A}$) we have $U \subset A$, $U \subset B$ and*

$$A \setminus U = S_1^{d_1} \cup \dots \cup S_1^{d_s} \text{ and } B \setminus U = S_2^{d_1} \cup \dots \cup S_2^{d_s}$$

for some $S_1, S_2 \subset [n]$. We may furthermore require S_1, S_2 to both have size m and S_1 to be disjoint from S_2, which provides $A \Delta B = (S_1^{d_1} \cup S_2^{d_1}) \cup \dots \cup (S_1^{d_s} \cup S_2^{d_s})$. Alternatively we may require S_2 to be a strict subset of S_1, which provides $A \Delta B = (S_1^{d_1} \setminus S_2^{d_1}) \cup \dots \cup (S_1^{d_s} \setminus S_2^{d_s})$.

Our second main result is an equivalence between several versions of Conjecture 1.3. To state it we introduce a few more notations. If d is a positive integer, then we say that a subset $A \subset [n]^d$ is *symmetric* if $\mathbb{1}_A(x_1, \dots, x_d) = \mathbb{1}_A(x_{\sigma(1)}, \dots, x_{\sigma(d)})$ for every permutation σ of $[d]$, and write $\mathcal{P}([n]^d)_{\mathrm{Sym}}$ for the collection of symmetric subsets of $\mathcal{P}([n]^d)$.

Theorem 1.5. *Conjecture 1.3 is equivalent to each of the following statements.*

(i) Let s, d be positive integers and let $\delta > 0$. Then, for n large enough depending on d and δ only, every subset $\mathcal{A}$ of $\mathcal{P}([n]^d \cup \cdots \cup [n]^d)$ that has density at least δ contains a pair (A, B) of distinct subsets of $[n]^d \cup \cdots \cup [n]^d$ such that A is contained in B and $B \setminus A = S^d \cup \cdots \cup S^d$ for some $S \subset [n]$, with all disjoint unions taken over s copies.

(ii) Let d be a positive integer and let $\delta > 0$. Then, for n large enough depending on d and δ only, every subset $\mathcal{A}$ of $\mathcal{P}([n]^d)$ that has density at least δ contains a pair (A, B) of distinct subsets of $[n]^d$ such that A is contained in B and $B \setminus A = S^d$ for some $S \subset [n]$.

(iii) Let d be a positive integer. There exists a sequence $(\mathcal{A}_m)_{m \geq 1}$ of subsets $\mathcal{A}_m \subset [m]^d$ such that for every $\delta > 0$, if m is large enough (depending on d, δ and on the sequence $(\mathcal{A}_m)_{m \geq 1}$) then every $\mathcal{A} \subset \mathcal{A}_m$ with $|\mathcal{A}| \geq \delta |\mathcal{A}_m|$ contains a pair (A, B) of distinct subsets of $[m]^d$ satisfying $A \subset B$ and $B \setminus A = S^d$ for some $S \subset [n]$.

(iv) Let d be a positive integer and let $\delta > 0$. Then, for n large enough depending on d and δ only, every subset $\mathcal{A}$ of $\mathcal{P}([n]^d)_{\mathrm{Sym}}$ that has density at least δ contains a pair (A, B) of distinct subsets of $[n]^d$ such that A is contained in B and $B \setminus A = S^d$ for some $S \subset [n]$.

2 Overview of the proof methods

2.1 Proof of Theorem 1.4

Suppose that we want to show that for some sequence Ω_n of "universes" indexed by n and some pattern P on pairs of elements (both in the same Ω_n), we have that whenever $\delta > 0$ is a positive real number, n is large enough (depending on δ), and $\mathcal{A}$ is a subset of Ω_n with density at least δ in Ω_n we can always find a pair (A, B) of distinct elements of $\mathcal{A}$ satisfying P.

Then one line of argument runs as follows. Suppose that we can, for large enough n, define a non-empty collection $\mathcal{W}_n$ of subsets of Ω_n satisfying the four following properties.

(i) If (A, B) is a pair of distinct elements of Ω_n which belongs to the same $\mathcal{C} \in \mathcal{W}_n$, then (A, B) satisfies P.

(ii) Every $\mathcal{C} \in \mathcal{W}_n$ has the same size $K > 0$.

(iii) Every element of Ω_n belongs to the same number $L > 0$ of collections $\mathcal{W}_n$.

(iv) The size of each $\mathcal{C} \in \mathcal{W}_n$ tends to infinity with n.

Then a double-counting argument allows us to conclude as desired: the average density $\mathbb{E}_{\mathcal{C} \in \mathcal{W}_n} \frac{|\mathcal{A} \cap \mathcal{C}|}{|\mathcal{C}|}$ can by (ii) be rewritten as

$$K^{-1} |\Omega_n| \mathbb{E}_{\mathcal{C} \in \mathcal{W}_n} \mathbb{E}_{A \in \Omega_n} \mathbb{1}_{A \in \mathcal{A} \cap \mathcal{C}} = \delta K^{-1} |\Omega_n| \mathbb{E}_{A \in \mathcal{A}} \mathbb{E}_{\mathcal{C} \in \mathcal{W}_n} \mathbb{1}_{A \in \mathcal{C}}.$$

2.1 Proof of Theorem 1.4

Assumption (iii) then shows that for every $A \in \mathcal{A}$ the inner expectation is equal to $L/|\mathcal{W}_n|$, so we obtain

$$\mathbb{E}_{\mathcal{C}\in\mathcal{W}_n}\frac{|\mathcal{A}\cap\mathcal{C}|}{|\mathcal{C}|} = \delta(|\Omega_n|L/|\mathcal{W}_n|K) = \delta : \tag{1}$$

indeed both products $|\Omega_n|L$ and $|\mathcal{W}_n|K$ count the number of pairs $(A,\mathcal{C}) \in \Omega_n \times \mathcal{W}_n$ satisfying $A \in \mathcal{C}$, so the parenthetical term is equal to 1. The identity (1) states that the average density of $\mathcal{A}\cap\mathcal{C}$ inside a random collection $\mathcal{C} \in \mathcal{W}_n$ is equal to δ, so we can in particular find some $\mathcal{C} \in \mathcal{W}_n$ such that the density of $\mathcal{A}\cap\mathcal{C}$ inside $\mathcal{C}$ is at least δ. Provided that n is large enough, we can by (iv) find a pair (A,B) of two distinct elements of $\mathcal{A}$ that both belong to $\mathcal{C}$, and that pair hence satisfies P by (i).

In many situations, we do not have (i)-(iv) in full, but weaker versions are satisfied which nonetheless suffice, especially for (ii) and (iii). For instance, we may have that most $\mathcal{C} \in \mathcal{W}_n$ have approximately the same size, and that most elements of Ω_n belong to approximately the same number of collections $\mathcal{C} \in \mathcal{W}_n$, these two properties often being established using concentration arguments. There, rather than (1) we instead obtain the sufficient lower bound

$$\mathbb{E}_{\mathcal{C}\in\mathcal{W}_n}\frac{|\mathcal{A}\cap\mathcal{C}|}{|\mathcal{C}|} \geq \tau\delta \tag{2}$$

for some $\tau > 0$ (such as $1/2$).

An example of a setting where the present proof technique works straightforwardly, with (i)-(iv) completely fulfilled is the task of showing that if $\mathcal{A}$ is a subset of $\mathbb{Z}_2^{\mathbb{Z}_n}$ with density equal to some $\delta > 0$, and n is large enough depending on δ, then we can find distinct $A, B \in \mathcal{A}$ such that the symmetric difference $A\Delta B$ is an interval modulo n. For every $C \in \mathbb{Z}_2^{\mathbb{Z}_n}$ and every $y \in \mathbb{Z}_n$ we define the collection

$$\mathcal{C}(C,y) = \{C, C+\mathbb{1}_{\{y\}}, C+\mathbb{1}_{\{y,y+1\}}, C+\mathbb{1}_{\{y,y+1,y+2\}}, \dots, C+1\}.$$

In this case, the common size of the collections $\mathcal{C}(C,y)$ is equal to $n+1$, and (1) then shows that it suffices that $n > \delta^{-1}$ for $\mathcal{A}$ to contain a desired pair (A,B).

That (i) is satisfied in this setting follows from the fact that the symmetric difference of two intervals intervals in $C(C,y)$ is an interval, but many patterns P are not "transitive" in the sense that is not the case that if (A,B) and (A,C) satisfy P then (B,C) satisfies P: in particular, whenever $d \geq 2$ it is never the case that if A,B,C are subsets of $\mathbb{Z}_2^{[n]^d}$ such that $A\Delta B = X^d$ and $A\Delta C = Y^d$ for some non-empty distinct $X,Y \subset [n]$, then $B\Delta C = Z^d$ for some non-empty $Z \subset [n]$. Likewise if $A,B,C \in \mathcal{P}([n]^2)$ satisfy $A \subset B \subset C$ with $B \setminus A = Y^d$ and $C \setminus B = X^d$ for some non-empty $X,Y \subset [n]$ then the set difference $C \setminus A$ is never of the type Z^d with $Z \subset [n]$.

The requirement for transitivity is a key limitation of the method that we have described so far, but it nonetheless still provides a recipe for obtaining a distance 2 version of the pattern. For every $A \in \Omega_n$, we define the collection $\mathcal{C}(A)$ to be the collection of all B such that (A,B) satisfies P; if we can show, using the argument described so far, that $\mathcal{C}(A)$ contains at least two elements B,C of $\mathcal{A}$, then we have that B,C are distance 2 apart for P, in the sense that there exists A such that (A,B) and (A,C) both satisfy P. This is the basic idea behind how we will prove Theorem 1.4. (It will suffice for us to define $\mathcal{C}(A)$ for some well-chosen "independent" $A \in \Omega_n$ rather than for all $A \in \Omega_n$, and we shall do so to avoid unnecessary technical complications in the proofs of these results.)

2.2 Proof of Theorem 1.5

Until this point all that we have used from (2) is that (provided that n is large enough) one of the $\mathcal{C} \in \mathcal{W}_n$ contains at least two elements of $\mathcal{A}$. But (2) shows the much stronger claim that $\mathcal{A}$ has density at least $\tau\delta$ inside one of the $\mathcal{C} \in \mathcal{W}_n$. This shows that in the distance 2 results we can not only find pairs $(A, B), (A, C) \in \Omega_n \times \mathcal{A}$ satisfying P, but, for any integer k and n large enough depending on k,δ only, pairs $(A, B_1), \dots, (A, B_k) \in \Omega_n \times \mathcal{A}$.

However, having $\mathcal{A} \cap \mathcal{C}$ dense inside some $\mathcal{C} \in \mathcal{W}_n$ also allows us to obtain a very different type of statement besides distance 2 results. Suppose that we aim for the original distance 1 problem stated in the first paragraph of this section, and that the argument as described so far does not suffice (as we have previously discussed, this is the case whenever P does not exhibit transitivity). Then a variation of the argument may allow us to impose extra structure on the elements of $\mathcal{A}$.

We say that two sets V, V' are *isomorphic* (for a pattern P, which will always be clear given the context) if there exists a bijection $h : V \to V'$ such that for every pair (A, B) of elements of V, the pair $(A, B) \in V \times V$ satisfies P if and only if the pair $(h(A), h(B)) \in V' \times V'$ satisfies P.

Suppose that for some sequence of $\Omega'_n \subset \Omega_n$ we know the distance 1 result analogous to the one that we seek to prove: that is, we know that for every $\delta' > 0$, for n large enough depending on δ' only every subset $\mathcal{A}'$ of Ω'_n with density at least δ' contains a pair (A, B) of elements satisfying P. If we can define a sequence $\mathcal{W}_n$ of collections of subsets of Ω_n such that, for some m tending to infinity with n, each $\mathcal{C} \in \mathcal{W}_n$ is isomorphic to some Ω'_m, and the collections $\mathcal{W}_n$ satisfy (ii), (iii), (iv), then by using (2) and taking $\delta' = \tau\delta$ in the assumption above, we conclude that $\mathcal{A}$ contains a pair (A, B) of elements satisfying P. In other words, we have reduced the problem from dense subsets of Ω_n to dense subsets of Ω'_m. The basic strategy behind much of the proof of Theorem 1.5 will be modelled on this argument, even if only a relaxed version of (iii) will hold.

The role of each reference below is discussed in the full version [12] of the paper.

References

[1] N. Alon, *Graph-codes*, European J. Combin, **116**, (2024), Paper No. 103880.

[2] R. Alweiss, *On a Conjecture of Gowers on Clique Differences*, arxiv:2011.04039 (2020).

[3] V. Bergelson and A. Leibman. *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), 725–753.

[4] M. Campos, S. Griffiths, R. Morris, J. Sahasrabudhe, *An exponential improvement for diagonal Ramsey*, arXiv:2303.09521 (2024).

[5] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem for* $k = 3$, Discrete Math. **75** (1989), 227-241.

[6] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64-119.

[7] W. T. Gowers, *The first unknown case of polynomial DHJ*, `https://gowers.wordpress.com/2009/11/14/the-first-unknown-case-of-polynomial-dhj/`.

[8] W. T. Gowers and T. Karam, *Low-complexity approximations for sets defined by generalizations of affine conditions*, arXiv:2306.00747 (2023).

[9] B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), 481–547.

[10] P. Gupta, N. Ndiaye, S. Norin, L. Wei, *Optimizing the CGMS upper bound on Ramsey numbers*, arXiv:2407.19026 (2024).

[11] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222-229.

[12] T. Karam, *Three applications of coverings to difference patterns*, arXiv:2408.06812 (2024).

[13] S. Peluse, *Subsets of* $\mathbb{F}_p^n \times \mathbb{F}_p^n$ *without L-shaped configurations*, Compos. Math. **160** (2024), 176-236.

[14] D. H. J. Polymath, *A new proof of the density Hales-Jewett theorem*, Ann. Math. **175** (2012), 1283-1327.

[15] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. **30** (1930), 264–286.

[16] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.

[17] E. Sperner, *Ein Satz über Untermengen einer endlichen Menge*, Math. Z. **27** (1928), 544-548.

[18] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299-345.

[19] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Archief voor Wiskunde **15** (1927), 212-216.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Unions of intervals in codes based on powers of sets

(Extended abstract)

Thomas Karam*

Abstract

We prove that for every integer $d \geq 2$ there exists a dense collection of subsets of $[n]^d$ such that no two of them have a symmetric difference that may be written as the dth power of a union of at most $\lfloor d/2 \rfloor$ intervals. This provides a limitation on reasonable tightenings of a question of Alon from 2023 and of a conjecture of Gowers from 2009, and investigates a direction analogous to that of recent works of Conlon, Kamčev, Leader, Räty and Spiegel on intervals in the Hales-Jewett theorem.

The full proof of the results in this paper is performed in [9].

1 Introduction and main result

We will throughout use the following notations. If n is a positive integer, then $[n]$ will denote the set $\{1, \ldots, n\}$ of positive integers between 1 and n. If a,b are positive integers with $a \leq b$, then $[a, b]$ will denote the set $\{a, a+1, \ldots, b\}$ of $b - a + 1$ consecutive integers starting at a and ending at b, and we will refer to such a set as an *interval*. If $A \subset B$ are two finite sets and B is non-empty, then we will say that the ratio $|A|/|B|$ is the *density* of A inside B.

One research theme involves extending to patterns on set systems (and more generally to patterns on high-dimensional families) results on patterns in the integers. For instance, the Hales-Jewett theorem (proved in [7]) and the density Hales-Jewett theorem (proved by Furstenberg and Katznelson [4], [5], then by the Polymath1 project [14]) are such generalisations of van der Waerden's theorem and of Szemerédi's theorem respectively. Likewise, the Bergelson-Leibman theorem [2] is a corresponding generalisation of the polynomial van der Waerden theorem. One goal in this programme, discussed by Gowers [6], is to combine these generalisations.

*Mathematical Institute, University of Oxford. Email: `thomas.karam@maths.ox.ac.uk`. Supported by ERC Advanced Grant 883810.

Conjecture 1.1. *[6, Conjecture 3] Let k, d be positive integers and let $\delta > 0$. If n is large enough depending on k, d, δ only then for every subset $\mathcal{A}$ of the set*

$$\mathcal{K} = [k]^{[n]} \times \cdots \times [k]^{[n]^d}$$

with density at least δ there is a non-empty subset $S \subset [n]$ and an element

$$y \in [k]^{[n] \setminus S} \cup \cdots \cup [k]^{[n]^d \setminus S^d}$$

such that whenever the coordinates of $x \in \mathcal{K}$ are constant on each of the sets $S, \ldots, S^d$, and coincide with those of y outside these sets, we have that $x \in \mathcal{A}$.

The case $k = 2$ is an important step towards Conjecture 1.1 because it is still open while at the same time avoiding the extra difficulties coming from arithmetic progressions that arise for $k \geq 3$. In turn, for $k = 2$, the case where we have $\mathcal{A} = \mathcal{A}' \times [2]^{[n]^2} \times \cdots \times [2]^{[n]^d}$ with $\mathcal{A}' \subset [2]^n$ can immediately be seen to reduce to Sperner's theorem [15], but the "opposite" specialisation $\mathcal{A} = [2]^{[n]} \times \cdots \times [2]^{[n]^{d-1}} \times \mathcal{A}'$ with $\mathcal{A}' \subset [2]^{[n]^d}$ remains a conjecture which can be reformulated as follows.

Conjecture 1.2. *Let d be a positive integer and let $\delta > 0$. Then, for n large enough depending on d, δ only, every subset $\mathcal{A}$ of $\mathcal{P}([n]^d)$ that has density at least δ contains a pair (A, B) of distinct subsets of $[n]^d$ such that A is contained in B and $B \setminus A = S^d$ for some $S \subset [n]$.*

Meanwhile, a second direction of research has been to ask for strengthening of results on patterns in set systems by requiring intervals in these patterns. In the case of the Hales-Jewett theorem, substantial progress has been made, particularly for the alphabet [3], which is now very well understood. Say that a *combinatorial line* in $[3]^n$ is a subset $L \subset [3]^n$ such that for some strict subset S of $[n]$ and some element $y \in [3]^{[n] \setminus S}$, an element $x \in [3]^n$ belongs to L if and only if all coordinates of x that are in $[n] \setminus S$ coincide with those of y and all coordinates of x that are in S are the same. For every positive integer r, let $s(r)$ be the smallest integer such every colouring of $[3]^n$ with r colours contains a monochromatic combinatorial line for which the wildcard set - that is, the set S - may be written as a union of at most s intervals. In a sequence of works by Conlon and Kamčev [3], Leader and Räty [13], and Kamčev and Spiegel [8], the value of $s(r)$ has been determined for every r: we have $s(r) = r$ if r is odd and $s(r) = r - 1$ if r is even.

Finally, a third avenue, described by Alon [1] last year and which since then has received much interest and led to much follow-up work, consists in exploring questions of the following type. Given a set $\mathcal{H}$ of graphs with vertex set $[n]$, we may ask for the largest size of a "graph-code" with respect to $\mathcal{H}$, that is, for the largest size of a set $\mathcal{A}$ of graphs with vertex set $[n]$ such that no two distinct graphs in $\mathcal{A}$ have edge sets of which the *symmetric difference* is a graph in $\mathcal{H}$. As discussed in [1], one of the motivations for considering this class of problems is the case where $\mathcal{H}$ is the set of all cliques with vertex set contained in $[n]$, which provides a symmetric difference version of a graph variant of the case $d = 2$ of Conjecture 1.2: that graph variant had previously been suggested ([6], Conjecture 4) as a possible Polymath project.

The weaker, symmetric difference version of Conjecture 1.2 (that is, the statement that we obtain if we replace the two conditions $A \subset B$ and $B \setminus A = S^d$ by the single condition $A \Delta B = S^d$) is to our knowledge still open. Our main result will assert that even in this version (and therefore in Conjecture 1.2 as well), the set S^d cannot be required to be a union of too few intervals.

Theorem 1.3. *For all positive integers $n \geq d \geq 2$ there exists $\mathcal{A} \subset \mathcal{P}([n]^d)$ with $|\mathcal{A}| = |\mathcal{P}([n]^d)|/2$ such that there exists no pair $(A, B) \in \mathcal{A} \times \mathcal{A}$ satisfying $A \Delta B = S^d$ for some non-empty interval $S \subset [n]$ that can be written as a union of at most $\lfloor d/2 \rfloor$ intervals.*

As a last comment on the existing literature, we recall that the related theme of pairs of subsets of $[n]$ (rather than $[n]^d$) of which some combination avoids a specified class of sets has been studied in a systematic way in several works: this is for instance respectively the case in [12] by Leader and Long for forbidden set differences, in [10] by Karpas and Long for forbidden symmetric differences, and in [11] by Keevash and Long for forbidden intersections.

2 The case of one interval

In the present section we prove the following special case of Theorem 1.3, only requiring that symmetric differences avoid powers of intervals. The full proof of Theorem 1.3, building on the proof of the special case that we are about to describe, is contained in [9].

Theorem 2.1. *For all positive integers $n, d \geq 2$ there exists $\mathcal{A} \subset \mathcal{P}([n]^d)$ with $|\mathcal{A}| = |\mathcal{P}([n]^d)|/2$ such that there exists no pair $(A, B) \in \mathcal{A} \times \mathcal{A}$ satisfying $A \Delta B = I^d$ for some non-empty interval $I \subset [n]$.*

We have chosen to do so of course for the reader's convenience but also because later on the full proof of Theorem 1.3 will have the same general structure, and we will then focus on the additional complications relatively to the case of Theorem 2.1.

With respect to the graph-codes setting of Alon, essentially the same proof also provides a similar statement for clique symmetric differences.

Theorem 2.2. *Let $n \geq 2$ be an integer. There exists a subset $\mathcal{G}$ of the set of (nonoriented, loopless) graphs on the vertex set $[n]$ that has density at least $1/2$ in that set and such that no two distinct graphs $G_1, G_2 \in \mathcal{G}$ are two distinct graphs have edge sets the symmetric difference of which is a clique with vertex set indexed by an interval.*

We now perform the successive steps that will lead us to Theorem 2.1.

Proof of Theorem 2.1. We first note that it suffices to obtain $\mathcal{A}$ satisfying the inequality $|\mathcal{A}| \geq |\mathcal{P}([n]^d)|/2$, since if this inequality is strict, then by averaging, $\mathcal{A}$ must contain a pair A,B of elements that only differ by some diagonal element, that is, some element $(a, \dots, a) \in [n]^d$ with $a \in [n]$.

Next, we note that we may reduce to the case $d = 2$. To do this, we define a map $i_d : [n]^2 \to [n]^d$ by

$$i_d(x, y) = i_d(x, \dots, x, y)$$

for all $x, y \in [n]$. If $\mathcal{A} \subset \mathcal{P}([n]^2)$ then we define $\mathcal{A}' \subset \mathcal{P}([n]^d)$ to be the collection of subsets of $[n]^d$ that have a restriction to

$$i_d([n]^2) = \{(x, \dots, x, y) : (x, y) \in [n]^2\}$$

which coincides with $i_d(A)$ for some $A \in \mathcal{A}$. The densities $|\mathcal{A}'|/|\mathcal{P}([n]^d)|$ and $|\mathcal{A}|/|\mathcal{P}([n]^2)|$ are the same, and if $A', B' \in \mathcal{A}'$ satisfy $A' \Delta B' = I^d$ for some non-empty interval I of $[n]$, then the sets $A = i_d^{-1}(A' \cap i_d([n]^2))$ and $B = i_d^{-1}(B' \cap i_d([n]^2))$ satisfy $A \Delta B = I^2$.

After that, we define a (nonoriented, loopless) graph G with vertex set $\mathcal{P}([n]^2)$ and where two distinct vertices A,B are joined by an edge if and only if there exists an interval $I \subset [n]$ such that $A\Delta B = I^2$. In the remainder of the proof we will show that the graph G is bipartite; to conclude the proof we may then take $\mathcal{A}$ to be one larger of the two parts of the bipartition. (In fact, by the first observation from the present proof, they cannot have different sizes.)

In turn, to show that G is bipartite it suffices to show that G contains no cycle of odd length, and to establish this it suffices to prove that the family of functions

$$\{\mathbb{1}_{[a,b]^2} : 1 \leq a \leq b \leq n\}$$

is linearly independent over $\mathbb{F}_2$. Indeed, suppose that G contains a cycle $A_0, \dots, A_k$ with $A_k = A_0$ and such that for every $i \in [k]$ we have $\{A_{i-1}, A_i\} \in E(G)$, that is, $A_i \Delta A_{i-1} = I_i^2$ for some interval $I_i \subset [n]$. Then we have

$$I_1^2 \Delta \dots \Delta I_k^2 = \emptyset,$$

that is, the linear relation

$$\mathbb{1}_{I_1^2} + \dots + \mathbb{1}_{I_k^2} = 0$$

over $\mathbb{F}_2$. The intervals $I_1, \dots, I_k$ might not yet be pairwise distinct; we consider a maximal subfamily of pairwise distinct intervals among $I_1, \dots, I_k$, which without loss of generality we can assume to be $I_1, \dots, I_l$ for some $l \leq k$; letting (for each $h \in [l]$) $A_h \in \mathbb{N}$ be the number of appearances of I_h among $I_1, \dots, I_k$ we obtain

$$A_1 \mathbb{1}_{I_1^2} + \dots + A_l \mathbb{1}_{I_l^2} = 0$$

over $\mathbb{F}_2$. Assuming that $\mathbb{1}_{I_1^2}, \dots, \mathbb{1}_{I_l^2}$ are linearly independent over $\mathbb{F}_2$, each of the integers $A_1, \dots, A_l$ must be even and their sum k must hence be even, so the cycle $A_0, \dots, A_k$ must have even length.

Finally we establish the desired linear independence. We begin by noting that whenever $1 \leq a \leq b \leq n$ we may write the set differences

$$[a, b]^2 \setminus [a, b-1]^2 \text{ and } [a+1, b]^2 \setminus [a+1, b-1]^2$$

respectively as

$$\{(a, b), (a+1, b), \dots, (b-1, b), (b, b), (b, b-1), \dots, (b, a+1), (b, a)\}$$
$$\{(a+1, b), \dots, (b-1, b), (b, b), (b, b-1), \dots, (b, a+1)\}.$$

This provides the identity

$$\mathbb{1}_{[a,b]^2} - \mathbb{1}_{[a,b-1]^2} - \mathbb{1}_{[a+1,b]^2} + \mathbb{1}_{[a+1,b-1]^2} = \mathbb{1}_{\{(a,b),(b,a)\}},$$

which holds over $\mathbb{R}$ and hence also over $\mathbb{F}_2$. Let $\mathrm{Sym}([n]^2)$ be the linear space of functions $u : [n]^2 \to \mathbb{F}_2$ satisfying $u(x, y) = u(y, x)$ for all $(x, y) \in [n]^2$. The family

$$\{\mathbb{1}_{\{(a,a)\}} : a \in [n]\} \cup \{\mathbb{1}_{\{(a,b),(b,a)\}} : 1 \leq a < b \leq n\}$$

constitutes a basis of $\mathrm{Sym}([n]^2)$ and has size $n(n+1)/2$, and as we have now shown that it is spanned by the family

$$\{\mathbb{1}_{[a,b]^2} : 1 \leq a \leq b \leq n\}$$

which has exactly the same size, that family is also a basis of $\mathrm{Sym}([n]^2)$, and is in particular linearly independent over $\mathbb{F}_2$. □

To prove Theorem 2.2 it suffices to go through the proof of Theorem 2.1 while making three modifications: ignoring the first two paragraphs, ignoring the n diagonal elements of $[n]^2$, and ignoring the n intervals in $[n]$ with size 1.

It is worthwhile to note how the proof of Theorem 2.1 fails in the case $d = 1$, where

$$\mathbb{1}_{I_1 \cup I_2} - \mathbb{1}_{I_1} - \mathbb{1}_{I_2} = 0$$

for any disjoint intervals $I_1, I_2 \subset [n]$, and how it fails to provide an analogous negative result to the case $d = 2$ of Conjecture 1.2, where for any pairwise disjoint intervals $I_1, I_2, I_3 \subset [n]$ we have

$$\mathbb{1}_{(I_1 \cup I_2 \cup I_3)^2} - \mathbb{1}_{(I_1 \cup I_2)^2} - \mathbb{1}_{(I_1 \cup I_3)^2} - \mathbb{1}_{(I_2 \cup I_3)^2} + \mathbb{1}_{I_1^2} + \mathbb{1}_{I_2^2} + \mathbb{1}_{I_3^2} = 0.$$

This identity provides a cycle with odd length not only in the graph G used in the proof of Theorem 2.1, but also a cycle with odd length in the analogous graph for the case $d = 2$ of Conjecture 1.2, that is, the graph with vertex set $\{0,1\}^{[n]^2}$ where we join two distinct sets A, B by an edge if $A \subset B$ and $B \setminus A = S^2$ for some $S \subset [n]$. Indeed if for instance, I_1, I_2, I_3 are taken to be $\{1\}, \{2\}, \{3\}$ respectively, then the sequence of sets with respective indicator functions represented by the following matrices provides a cycle with length 7.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

The role of each reference below is discussed in the full version [9] of the paper.

References

[1] N. Alon, *Graph-codes*, European J. Combin. **116** (2024).

[2] V. Bergelson and A. Leibman. *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), 725–753.

[3] D. Conlon and N. Kamčev. *Intervals in the Hales–Jewett theorem*, to appear in European. J. Combin, arXiv:1801.08919 (2018).

[4] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem for* $k = 3$, Discrete Math. **75** (1989), 227-241.

[5] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64-119.

[6] W. T. Gowers, *The first unknown case of polynomial DHJ*, `https://gowers.wordpress.com/2009/11/14/the-first-unknown-case-of-polynomial-dhj/`.

[7] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222-229.

[8] N. Kamčev and C. Spiegel, *Another Note on Intervals in the Hales–Jewett Theorem*, Electron. J. Combin. **29** (2022).

[9] T. Karam, *Unions of intervals in codes based on powers of sets*, arxiv:2408.15144 (2024).

[10] I. Karpas and E. Long, *Set families with a forbidden pattern*, European. J. Combin. **62** (2017), 183-196.

[11] P. Keevash and E. Long, *Forbidden vector-valued intersections*, Proc. London Math. Soc. **121** (2020), 702-742.

[12] I. Leader and E. Long, *Forbidding a set difference of size 1*, Discret. Appl. Math. **169** (2014), 219-224.

[13] I. Leader and E. Räty, *A note on intervals in the Hales–Jewett theorem*, Electron. J. Combin. **25** (2018).

[14] D. H. J. Polymath, *A new proof of the density Hales-Jewett theorem*, Ann. Math. **175** (2012), 1283-1327.

[15] E. Sperner, *Ein Satz über Untermengen einer endlichen Menge*, Math. Z. **27** (1928), 544-548.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On small densities defined without pseudorandomness

(Extended abstract)

Thomas Karam*

Abstract

We identify an assumption on linear forms $\phi_1, \ldots, \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ that is much weaker than approximate joint equidistribution on the Boolean cube $\{0,1\}^n$ and is in a sense almost as weak as linear independence, but which guarantees that every subset of $\{0,1\}^n$ on which none of $\phi_1, \ldots, \phi_k$ has full image has a density which tends to 0 with k. This density is at most quasipolynomially small in k, a bound that is necessarily close to sharp.

1 Background and main results

1.1 Background on mod-p linear forms and basic difficulties

All our statements in this paper will be uniform in the integer n, which will not play any significant role. The present paper will involve restrictions of mod-p forms (that is, of linear forms $\mathbb{F}_p^n \to \mathbb{F}_p$) to subsets of $\mathbb{F}_p^n$ such as $\{0,1\}^n$ and more generally S^n for some non-empty subset S of $\mathbb{F}_p$. This topic discussed here is treated more thoroughly in the paper [10].

Although the distributions of mod-p forms and k-tuples of mod-p forms on $\{0,1\}^n$ for some integer k present additional difficulties compared to their counterparts on the whole of $\mathbb{F}_p^n$, there remains a known sufficient condition which guarantees that a k-tuple of forms is approximately jointly equidistributed.

Definition 1.1. *Let p be a prime. If $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ is a mod-p form defined by $\phi(x) = a_1x_1 + \cdots + a_nx_n$, then we say that the* support $Z(\phi)$ *is the set $\{i \in [n] : a_i \neq 0\}$ and that the* support size *of ϕ is the size of $Z(\phi)$.*

If $k \geq 1$, $r \geq 0$ are integers, then we say that mod-p forms $\phi_1, \ldots, \phi_k$ are r-separated *if the support size of the linear combination $a_1\phi_1 + \cdots + a_k\phi_k$ is at least r for every $(a_1, \ldots, a_k) \in \mathbb{F}_p^k \setminus \{0\}$.*

*Mathematical Institute, University of Oxford. Email: thomas.karam@maths.ox.ac.uk. Supported by ERC Advanced Grant 883810.

Proposition 1.2 ([6], Proposition 2.4)**.** *Let p be a prime, let S be a subset of $\mathbb{F}_p$ with size at least 2, and let $k \geq 1$, $r \geq 0$ be integers. If $\phi_1, \ldots, \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ are r-separated mod-p forms, then for every $(y_1, \ldots, y_k) \in \mathbb{F}_p^k$ we have*

$$|\mathbb{P}_{x \in \{0,1\}^n}(\phi_1(x) = y_1, \ldots, \phi_k(x) = y_k) - p^{-k}| \leq (1 - p^{-2})^r.$$

1.2 Main results

Previous works [6], [11] relied primarily on Proposition 1.2 to show that a set defined by linear mod-p conditions has small density. One quantitative way in which Proposition 1.2 may be viewed as not so strong is that in order to immediately deduce from Proposition 1.2 an upper bound of the type

$$|\mathbb{P}_{x \in \{0,1\}^n}(\phi_1(x) = y_1, \ldots, \phi_k(x) = y_k) - p^{-k}| \leq cp^{-k}$$

for some fixed $c > 0$, the lower bound r on the supports of the linear combinations $a_1\phi_1 + \cdots + a_k\phi_k$ with $a \in \mathbb{F}_p^k \setminus \{0\}$ that we must take grows linearly in k.

In particular, if we consider strict subsets $E_1, \ldots, E_k$ of $\mathbb{F}_p$ and the set Λ of $x \in \{0,1\}^n$ satisfying all conditions $\phi_1(x) \in E_1, \ldots, \phi_k(x) \in E_k$, as has been done in [6], then it suffices for $E_1, \ldots, E_k$ to have size 2 or more for the upper bound on the density of Λ that comes from applying the bounds from Proposition 1.2 to be $2^k(1 - p^{-2})^r$ and for that upper bound to be meaningful for large k we must then require r to be at least linear in k.

Our main contribution in the present paper will be to show that there exists some lower bound on r that is *uniform* in k, and which suffices to guarantee that the density of Λ tends to 0 as k tends to infinity. We will state and prove our results in the general case where the alphabet of the variables is an arbitrary subset S of $\mathbb{F}_p$ containing at least two elements, as the proofs in this setting are the same as those of the special case $S = \{0, 1\}$.

Definition 1.3. *Let p be a prime, let S be a subset of $\mathbb{F}_p$ with size at least 2, and let E be a strict subset of $\mathbb{F}_p$. Then there exists a smallest nonnegative integer*

$$L = L(S, E) \leq \lfloor \frac{|E| - 1}{|S| - 1} \rfloor + 1 \leq |E| \tag{1}$$

such that for any $a_1, \ldots, a_L \in \mathbb{F}_p^*$ the set $a_1S + \cdots + a_LS$ is not contained in E. We will write $L(S)$ for $L(S, \mathbb{F}_p)$.

The existence of L follows from the $k = 1$ case of Proposition 1.2, but it also follows immediately from the Cauchy-Davenport inequality, which furthermore provides the inequality (1). We note that if $S = \{0, 1, \ldots, |S| - 1\}$ and $E = \{0, 1, \ldots, |E| - 1\}$, then (1) becomes an equality. We are now ready to state our main result in this paper.

Theorem 1.4. *Let p be a prime, and let S be a subset of $\mathbb{F}_p$ with size at least 2. Then there exists some $a = a(p, S) > 0$ such that the following holds. If $k \geq 1$ is an integer, $E_1, \ldots, E_k$ are strict subsets of $\mathbb{F}_p$, and $\phi_1, \ldots, \phi_k$ are mod-p forms satisfying*

$$|Z(\phi_j - \phi_i)| \geq 2 \max_{1 \leq t \leq k} L(S, E_t) - 1 \tag{2}$$

for any $1 \leq i < j \leq k$, then the density of points $x \in S^n$ satisfying $\phi_1(x) \in E_1, \ldots, \phi_k(x) \in E_k$ is at most $O(k^{-a/\log\log k})$ as k tends to infinity.

In the case where the sets $E_1, \dots, E_k$ all have size bounded above by some common value strictly less than p, Theorem 1.4 specialises as follows thanks to the inequality (1).

Corollary 1.5. *Let p be a prime, let S be a subset of $\mathbb{F}_p$ with size at least 2, and let $1 \leq r \leq p-1$ be an integer. Then there exists some $a = a(p, S, r) > 0$ such that the following holds. If $k \geq 1$ is an integer, $E_1, \dots, E_k$ are strict subsets of $\mathbb{F}_p$ each with size at most r, and $\phi_1, \dots, \phi_k$ are mod-p forms satisfying $|Z(\phi_j - \phi_i)| \geq 2\lfloor \frac{r-1}{|S|-1} \rfloor + 1$ for any $1 \leq i < j \leq k$, then the density of points $x \in S^n$ satisfying $\phi_1(x) \in E_1, \dots, \phi_k(x) \in E_k$ is at most $O(k^{-a/\log\log k})$ as k tends to infinity.*

The bound in the assumption of Theorem 1.4 is optimal, at least in the case where the sets $E_1, \dots, E_k$ are the same set E, as can be seen from the definition of L and the triangle inequality on the support size.

We will use the following proposition, proved in [6], which informally states that the probability that a given mod-p form takes any prescribed value is either 0 or bounded below by some positive quantity. Proposition 1.6 will again be used in Section 2.

Proposition 1.6 ([6], Proposition 2.5)**.** *Let p be a prime, let S be a non-empty subset of $\mathbb{F}_p$, and let $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ be a linear form. Then for every $y \in \mathbb{F}_p$ we have*

$$\mathbb{P}_{x \in S^n}(\phi(x) = y) = 0 \text{ or } \mathbb{P}_{x \in S^n}(\phi(x) = y) \geq \beta(p, S)$$

where $\beta(p, S) = |S|^{-\lceil (p-1)/(|S|-1) \rceil}$.

We note that the bounds in Theorem 1.4 and in Corollary 1.5 cannot be too far from tight, because there are simple examples satisfying power lower bounds in the other direction (and furthermore, uniformly with respect to $r \geq 1$).

Example 1.7. *Let $r, t \geq 1$ be integers and let $k = p^t$. Let $\psi_1, \dots, \psi_t : \mathbb{F}_p^n \to \mathbb{F}_p$ be linear forms with pairwise disjoint supports each with size at least r. The k linear combinations $\phi_1, \dots, \phi_k$ of $\psi_1, \dots, \psi_t$ then satisfy $|Z(\phi_j - \phi_i)| \geq r$ for any $1 \leq i < j \leq k$, but the set $\{x \in \{0,1\}^n : \phi_1(x) = 0, \dots, \phi_k(x) = 0\}$ is the same as the set $\{x \in \{0,1\}^n : \psi_1(x) = 0, \dots, \psi_t(x) = 0\}$, so by Proposition 1.6 has density at least $2^{-(p-1)t} = k^{-a}$ where $a = \log(2^{p-1})/\log p$.*

Throughout, we will often write (ϕ, E) for a condition $\phi(x) \in E$ where x is an element of S^n, as was done in [6]. If ϕ, ψ are two mod-p forms, then we will say that the *support distance* (or more simply *distance*) between ϕ and ψ is the support size of $\phi - \psi$. If ϕ_0 is a mod-p form and $r \geq 0$ is an integer, then we will say that the *ball* with *radius* r centered at ϕ_0 is the set of mod-p forms ϕ such that the support size of $\phi - \phi_0$ is at most r.

2 Proof of the main result

In this section we prove Theorem 1.4. Although there are several ways of writing this proof, the choice that we ultimately made was to do so in a way that mirrors the high-level structure of that of [6, Theorem 1.2] (performed in Section 3 of that paper). In both situations the proof proceeds in three successive stages establishing the desired result, in the cases where the linear forms that we start with (i) constitute a sunflower, (ii) are contained in a ball with bounded radius, and finally (iii) are completely arbitrary. Let us begin by recalling the relevant formal definition of a sunflower.

Definition 2.1. *Let p be a prime, let I be a finite set, let $\phi_i : \mathbb{F}_p^n \to \mathbb{F}_p$ be a linear form for every $i \in I$, and let ϕ_0 be an additional such form. We say that the forms $\phi_0 + \phi_i$ with $i \in I$ constitute a* sunflower *with* center ϕ_0 *if the supports of the forms ϕ_i with $i \in I$ are pairwise disjoint.*

2.1 Exponentially small densities from the sunflower structure

The most basic structure on mod-p forms that we will use to ensure that a set of points of S^n satisfying many conditions involving these forms has small density - exponentially small in the number of forms for now, although not for long - is a sunflower.

Proposition 2.2. *Let p be a prime and let $k \geq 1$ be an integer. Suppose that $E_1, \dots, E_k$ are strict subsets of $\mathbb{F}_p$ and that $\phi_0, \phi_1, \dots, \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ are linear forms such that $\phi_0 + \phi_1, \dots, \phi_0 + \phi_k$ constitute a sunflower with center ϕ_0 and $\phi_1, \dots, \phi_k$ each have support size at least* $\max_{1 \leq i \leq k} L(S, E_i)$. *Then the set*

$$\{x \in S^n : (\phi_0 + \phi_1)(x) \in E_1, \dots, (\phi_0 + \phi_k)(x) \in E_k\}$$

has density at most $p(1 - \beta(p, S))^k$ inside S^n.

Proof. For every $y \in \mathbb{F}_p$ the probabilities

$$\begin{aligned}
&\mathbb{P}_{x \in S^n}((\phi_0 + \phi_1)(x) \in E_1, \dots, (\phi_0 + \phi_k)(x) \in E_k, \phi_0(x) = y) \\
&\mathbb{P}_{x \in S^n}(\phi_1(x) \in E_1 - \{y\}, \dots, \phi_k(x) \in E_k - \{y\}, \phi_0(x) = y) \\
&\mathbb{P}_{x \in S^n}(\phi_1(x) \in E_1 - \{y\}, \dots, \phi_k(x) \in E_k - \{y\}) \\
&\mathbb{P}_{x \in S^n}(\phi_1(x) \in E_1 - \{y\}) \dots \mathbb{P}_{x \in S^n}(\phi_k(x) \in E_k - \{y\})
\end{aligned}$$

do not decrease from one to the next. (The equality between the last two comes from the disjointness of the supports of $\phi_1, \dots, \phi_k$.) The sets $E_1 - \{y\}, \dots, E_k - \{y\}$ have the same sizes as $E_1, \dots, E_k$ respectively, and every term of the product is strictly less than 1 by definition of $L(S, E_1), \dots, L(S, E_k)$, so is at most $1 - \beta(p, S)$ by Proposition 1.6. We obtain

$$\mathbb{P}_{x \in S^n}((\phi_0 + \phi_1)(x) \in E_1, \dots, (\phi_0 + \phi_k)(x) \in E_k, \phi_0(x) = y) \leq (1 - \beta(p, S))^k$$

for every $y \in \mathbb{F}_p$ and conclude by the law of total probability. □

2.2 From sunflowers to small balls

We next tackle the case where the assumption that the mod-p forms constitute a sunflower is relaxed to the assumption that they are contained in a ball of bounded radius. First we adapt the proof of the Erdős-Rado sunflower theorem (established in [3]).

Proposition 2.3. *Let p be a prime, let $t \geq 1, k \geq 1, r \geq 0$ be integers. Assume that $\phi_0 + \phi_1, \dots, \phi_0 + \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ are linear forms and that $\phi_1, \dots, \phi_k$ have supports at most r. If $k \geq p^r r! t^r$ then we can find $I \subset [k]$ with $|I| \geq t$ such that the linear forms $\phi_0 + \phi_i$ with $i \in I$ constitute a sunflower (the center of which may be different from ϕ_0).*

Proof. We proceed by induction on r. If $r = 0$ then the result is immediate. Suppose that for some $r \geq 1$ the result holds for $r - 1$. Let $M \subset [k]$ be a maximal set such that the linear forms ϕ_i with $i \in M$ have pairwise disjoint supports. If $|M| \geq t$ then we take $I = M$ and we are done. Otherwise, for every $i \in [k]$ the support of ϕ_i intersects the union $\cup_{j \in M} Z(\phi_j)$ of supports, which has size at most rt, in at least one element. By the pigeonhole principle we can find $J \subset [k]$ with size at least $k/(p-1)rt$ such that all forms ϕ_i with $i \in J$ have a common element z in their support, and furthermore have the same coefficient $u \in \mathbb{F}_p$ at that element of the support. The forms $\phi_i - ux_z$ with $i \in J$ then all have supports with size at most $r-1$. We then conclude by the inductive hypothesis. □

Our result in the case of forms contained in balls of bounded radius then follows from applying Proposition 2.3 and Proposition 2.2 successively.

Corollary 2.4. *Let p be a prime, let $k \geq 1$ and $r \geq 0$ be integers. If $\phi_0, \phi_1, \dots, \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ are linear forms, and $E_1, \dots, E_k$ are strict subsets of $\mathbb{F}_p$ satisfying $\max_{1 \leq i \leq k} L(S, E_i) \leq |Z(\phi_i)| \leq r$ for every $i \in [k]$ then*

$$\mathbb{P}_{x \in S^n}((\phi_0 + \phi_1)(x) \in E_1, \dots, (\phi_0 + \phi_k)(x) \in E_k) \leq p(1 - \beta(p, S))^{(k/p^r r!)^{1/r}}.$$

2.3 Finishing the proof

We are now ready to finish the proof of Theorem 1.4.

Proof of Theorem 1.4. Let $\epsilon > 0$. We distinguish two cases. If we can find a set $I \subset [k]$ with size $u = \lceil p \log(2\epsilon^{-1}) \rceil$ satisfying $|Z(\sum_{i \in I} a_i \phi_i)| \geq 5p^4 \log(2\epsilon^{-1})$ for every $a \in \mathbb{F}_p^I \setminus \{0\}$ then by Proposition 1.2 we have $\mathbb{P}_{x \in S^n}(\forall i \in I, \phi_i(x) = y_i) \leq 2p^{-u}$ for every $y \in \mathbb{F}_p^I$, and hence

$$\mathbb{P}_{x \in S^n}(\forall i \in I, \phi_i(x) \in E_i) \leq 2(1 - p^{-1})^u \leq \epsilon.$$

If on the other hand we cannot find a set I as above, then we can partition $[k]$ into at most p^u sets such that all forms $\phi_0 + \phi_i$ with index i in a given set are contained in some ball of radius at most $r = 5p^4 \log(2\epsilon^{-1})$. By the pigeonhole principle we can find such a set containing at least k/p^u forms. The assumption (2) together with the triangle inequality show that all but at most one of these forms is at a distance at least $\max_{1 \leq i \leq k} L(S, E_i)$ from the center of the ball. Using that $1 - \beta(p, S) \leq \exp(-\beta(p, S))$, Corollary 2.4 then provides the upper bound

$$\mathbb{P}_{x \in S^n}(\forall i \in I, \phi_i(x) \in E_i) \leq p \exp\left(-\beta(p, S)((k - p^u)/p^u p^r r!)^{1/r}\right).$$

It suffices that $k \leq 2p^u p^r r! (\beta(p, S)^{-1} \log(p\epsilon^{-1}))^r$ for the previous right-hand side to be at most ϵ. Plugging in the values of u and r in terms of ϵ, that becomes

$$k \leq 2p^{p \log(2\epsilon^{-1}) + 1} p^{5p^4 \log(2\epsilon^{-1})} (5p^4 \log(2\epsilon^{-1}))! (\beta(p, S)^{-1} \log(p\epsilon^{-1}))^{5p^4 \log(2\epsilon^{-1})}. \tag{3}$$

As ϵ tends to 0, the right-hand side (the dominant terms of which are the last two) grows as $\Omega(\epsilon^{-c(p) \log \log \epsilon^{-1}})$ for some $c(p) > 0$. □

The role of each reference below is discussed in the full version [10] of the paper.

2.3 Finishing the proof

References

[1] R. Alweiss, S. Lovett, K. Wu, J. Zhang, *Improved bounds for the sunflower lemma*, Ann. of Math. **194** (2021), 795-815.

[2] A. Bhangale, S. Khot, D. Minzer, *Effective bounds for restricted* 3*-arithmetic progressions in* $\mathbb{F}_p^n$, arXiv:2308.06600 (2023).

[3] P. Erdős and R. Rado, *Intersection theorems for systems of sets*, Journal of the London Mathematical Society, **35** (1960), 85–90.

[4] W. T. Gowers, *The first unknown case of polynomial DHJ*, https://gowers.wordpress.com/2009/11/14/the-first-unknown-case-of-polynomial-dhj. (Blog post.)

[5] W. T. Gowers and T. Karam, *Equidistribution of high-rank polynomials with variables restricted to subsets of* $\mathbb{F}_p$, arXiv:2209.04932 (2022).

[6] W. T. Gowers and T. Karam, *Low-complexity approximations for sets defined by generalizations of affine conditions*, arXiv:2306.00747 (2023).

[7] B. Green, *100 open problems*, manuscript.

[8] B. Green and T. Tao, *The distribution of polynomials over finite fields, with applications to the Gowers norms.* Contributions to Discrete Mathematics, **4** (2009), no. 2, 1-36.

[9] Jan Hazla, Thomas Holenstein, Elchanan Mossel, *Product space models of correlation: Between noise stability and additive combinatorics*, Discrete Anal. **19** (2018).

[10] T. Karam, *On small densities defined without pseudorandomness*, arxiv:2405.04391 (2024).

[11] T. Karam, *On the expressive power of mod-p linear forms on the Boolean cube*, arXiv:2309.14229 (2023).

[12] T. Karam, *Ranges of polynomials control degree ranks of Green and Tao over finite prime fields*, arXiv:2305.11088 (2023).

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On distinguishing subsets of the Boolean cube by mod-p forms

(Extended abstract)

Thomas Karam*

Abstract

Let $(\mathcal{A}_i)_{i \in [s]}$ be a sequence of dense subsets of the Boolean cube $\{0,1\}^n$ and let p be a prime. We show that if s is assumed to be superpolynomial in n then we can find distinct i, j such that the two distributions of every mod-p linear form on $\mathcal{A}_i$ and $\mathcal{A}_j$ are almost positively correlated. We also prove that if s is merely assumed to be sufficiently large independently of n then we may require the two distributions to have overlap bounded below by a positive quantity depending on p only. We finally provide analogous results where the sets $\mathcal{A}_i$ are instead dense subsets of disjoint regions of the cube.

The results presented here are discussed further and proved in [10]. For p a prime and n a positive integer, we will throughout say that a *mod-p linear form*, or *mod-p form* for short, is a linear form $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$. Such a form can be written as $\phi(x) = a_1 x_1 + \cdots + a_n x_n$ for some $a_1, \ldots, a_n \in \mathbb{F}_p$, and we will say that the *support* of ϕ is the set $\{z \in [n] : a_z \neq 0\}$. If ϕ, ψ are two mod-p forms then we will say that the *support distance* between them is the size of the support of the difference $\phi - \psi$.

1 Restrictions of mod-p forms to the cube

One central tool in the analysis of Boolean functions and in theoretical computer science more broadly is discrete Fourier analysis. There, the characters involved in the definition of the Fourier coefficients are functions $\mathbb{F}_2^n \to \mathbb{F}_2$ of the type $x \to (-1)^{\phi(x)}$, for some linear form $\phi : \{0,1\}^n \to \mathbb{F}_2$, where $\{0,1\}$ is identified with $\mathbb{F}_2$. Restrictions to $\{0,1\}^n$ of mod-p forms for some general prime p appear to be variants of these objects that are natural to study. Because $\{0,1\}$ can be identified to the whole of $\mathbb{F}_2$, whereas $\{0,1\}$ can only be embedded into $\mathbb{F}_p$, we can expect behaviour arising from linear forms with $p \geq 3$ on the cube to be at least a little more complex, and in some ways this is indeed the case.

The following question arises in numerous contexts as a way of measuring, informally speaking, the power of expression of a set of functions or of data. Given a class $\mathcal{C}$ of objects

*Mathematical Institute, University of Oxford. Email: thomas.karam@maths.ox.ac.uk. Supported by ERC Advanced Grant 883810.

and a class $\mathcal{F}$ of functions defined on $\mathcal{C}$, how many objects in $\mathcal{C}$ can we select such that for any two of them we can find a function in $\mathcal{F}$ that clearly separates them ? In our case, the class $\mathcal{C}$ of objects will be that of sufficiently dense subsets of the cube $\{0,1\}^n$ and the class $\mathcal{F}$ of functions will be the functions sending a set in $\mathcal{C}$ to the distribution of a mod-p form on that set. Answering this question happens to furthermore be of independent Ramsey-theoretic interest, since it can be equivalently formulated as follows: how many dense subsets of the cube can we choose until we necessarily can find a pair of them for which the pair of distributions of every mod-p form are not too far apart ?

2 Notions of separation on the distributions of mod-p forms

We begin by defining a number of ways to measure proximity between two probability distributions.

Definition 2.1. *Let $m \geq 2$ be an integer. For any m-tuple $(X_1, \dots, X_m)$ of distributions each taking values in the same finite set D, we define the following quantities.*

(i) The diameter $\mathrm{Diam}(X_1, \dots, X_m)$ *is defined to be* $\max_{1 \leq i < j \leq m} \mathrm{TV}(X_i, X_j)$*, where we write* $\mathrm{TV}(X_i, X_j) = \sum_{d \in D} |\mathbb{P}[X_i = d] - \mathbb{P}[X_j = d]|$ *for the total variation distance between X_i and X_j for all $i, j \in [m]$.*

(ii) The correlation $\mathrm{cor}(X_1, \dots, X_m)$ *is defined as* $\big(\sum_{d \in D} \mathbb{P}[X_1 = d] \dots \mathbb{P}[X_m = d]\big) - 1/|D|^{m-1}$.

(iii) The overlap $\omega(X_1, \dots, X_m)$ *is defined to be* $\sum_{d \in D} \min(\mathbb{P}[X_1 = d], \dots, \mathbb{P}[X_m = d])$.

We note in particular that if all but at most one of the variables $X_1, \dots, X_m$ are uniformly distributed (resp. approximately uniformly distributed), then $\mathrm{cor}(X_1, \dots, X_m)$ is zero (resp. small in absolute value), and that if $\omega(X_1, \dots, X_m) > 0$, then the intersection of the ranges of $X_1, \dots, X_m$ is not empty.

These definitions allow us to establish a qualitative hierarchy in the extent to which distributions of random variables $X_1, \dots, X_m$ are not too far apart. Each of the following properties is qualitatively stronger than the next, in the sense that whenever $i \in \{1, 2, 3\}$, for every choice of the parameters in the $i+1$th property which is sufficiently close to 0 (in a manner that depends on $|D|$ and m), there is a choice of parameters (which may further depend on $|D|$ and m) in the ith property which implies it.

1. Close distributions for some $\eta > 0$: $\mathrm{Diam}(X_1, \dots X_m) \leq \eta$.
2. Almost positive correlation for some $\nu > 0$: $\mathrm{cor}(X_1, \dots X_m) \geq -\nu$.
3. Overlap bounded away from 0 for some $A > 0$: $\omega(X_1, \dots X_m) \geq A$.
4. Overlapping distributions: $\omega(X_1, \dots X_m) > 0$.

If E is an event and $\mathcal{A}$ is a non-empty subset of $\{0,1\}^n$ then we shall write $\mathbb{P}_{\mathcal{A}}[E]$ for the probability that x satisfies E when x is chosen uniformly at random inside $\mathcal{A}$. If F is a function defined on $\{0,1\}^n$ and $\mathcal{A}, \mathcal{B}$ are non-empty subsets of $\{0,1\}^n$, then we write $\mathrm{TV}_{\mathcal{A},\mathcal{B}}(F)$ for the total variation distance between the distributions of $F(x)$ when x is chosen uniformly at random from $\mathcal{A}$ and when it is chosen uniformly at random from $\mathcal{B}$. If $m \geq 2$ is a positive integer and $\mathcal{A}_1, \dots, \mathcal{A}_m$ are non-empty subsets of $\{0,1\}^n$, then we write $\mathrm{Diam}_{\mathcal{A}_1,\dots,\mathcal{A}_m}(F)$ for $\mathrm{Diam}(X_1, \dots, X_m)$, where for each $i \in [m]$ the distribution of the variable X_i is the distribution of $F(x)$ for x chosen uniformly at random in $\mathcal{A}_i$. We define $\mathrm{cor}_{\mathcal{A}_1,\dots,\mathcal{A}_m}(F)$ and $\omega_{\mathcal{A}_1,\dots,\mathcal{A}_m}(F)$ in a similar way.

3 Main results

Starting with a number s of non-empty subsets $\mathcal{A}_1, \dots, \mathcal{A}_s$ of $\{0,1\}^n$, we can ask whether it is possible to obtain a pair (i,j) of distinct elements of $[s]$ such that all mod-p forms have close distributions on the pair $(\mathcal{A}_i, \mathcal{A}_j)$. As we will now illustrate, this is not possible in general, even if the number s grows exponentially with n.

In the case $p = 2$, that can be seen from the fact that the $2^n - 1$ non-zero linear forms $\phi_1, \dots, \phi_{2^n-1} : \mathbb{F}_2^n \to \mathbb{F}_2$ each have a different linear hyperplane of $\mathbb{F}_2^n$ as their kernels; since these hyperplanes have pairwise intersections of size 2^{n-2}, the hyperplanes $\mathcal{A}_i = \ker \phi_i$ with $i \in [2^n - 1]$ are such that $\phi_i(\mathcal{A}_i)$ only takes the value 0 on $\mathcal{A}_i$ but takes the values 0 and 1 with equal probability on every $\mathcal{A}_j$ with $j \neq i$.

For $p \geq 3$, it is no longer true in general that mod-p forms are determined by their preimages of 0 inside $\{0,1\}^n$, even up to proportionality: for instance the mod-5 forms $x_1 + x_2$ and $2x_1 + x_2$ both have the same preimage $\{x_1 = x_2 = 0\}$ of 0 inside $\{0,1\}^n$. However, we can find a set of mod-p forms with size exponential in n and within which this is the case. We begin by partitioning $[n]$ into pairwise disjoint sets $I_1, \dots, I_{2d}$ each of size p, with $d = \lfloor n/2p \rfloor$ and a remainder set I_0 with size at most $2p - 1$. We then restrict our attention to the mod-p forms ϕ of the type $\phi(x) = a_1 x_1 + \dots + a_n x_n$ where for every $u \in [d]$, all coefficients a_z with $z \in I_{2u-1} \cup I_{2u}$ are the same. We can then find a maximal set of such mod-p forms with size at least $p^{d-1} = p^{\lfloor n/2p \rfloor - 1}$ such that no two of them are proportional. We then index the forms in this set as $\phi_1, \dots, \phi_t$.

We then take x to be the element of $\{0,1\}^n$ such that $x_i = 0$ whenever $i \in I_0$ or i in I_u for some odd u, and $x_i = 1$ whenever $i \in I_u$ for some non-zero even u. We have $\phi_i(x) = 0$ for every $i \in [t]$, since the contribution of each set I_u of coordinates is zero. Let now i,j be distinct indices in $[t]$. Because ϕ_i and ϕ_j are not proportional, there exist distinct $u_1, u_2 \in [d]$ such that the pair $(A_{1,i}, A_{2,i})$ of coefficients of ϕ_i on $I_{2u_1-1} \cup I_{2u_1}$ and on $I_{2u_2-1} \cup I_{2u_2}$ is not proportional to the pair $(A_{1,j}, A_{2,j})$ of coefficients of ϕ_j on $I_{2u_1-1} \cup I_{2u_1}$ and on $I_{2u_2-1} \cup I_{2u_2}$. Viewing $A_{1,i}, A_{2,i}, A_{1,j}, A_{2,j}$ as integers in $[0, p-1]$, we use x to define a new element y of $\{0,1\}^n$ by changing $A_{2,i}$ of the coordinates of x in I_{2u_1-1} to 1 and changing $A_{1,i}$ of the coordinates of x in I_{2u_2} to 0. We then have

$$\begin{aligned} \phi_i(y) - \phi_i(x) &= A_{2,i}A_{1,i} - A_{1,i}A_{2,i} = 0 \\ \phi_j(y) - \phi_j(x) &= A_{2,i}A_{1,j} - A_{1,i}A_{2,j} \neq 0, \end{aligned}$$

so $\phi_i(y) = 0$ but $\phi_j(y) \neq 0$. We have shown that for every pair (i,j) of distinct elements of $[t]$, the set $\phi_i^{-1}(0)$ is not contained in the set $\phi_j^{-1}(0)$.

This can be used to show (but this is not completely straightforward) that any event on $\{0,1\}^n$ defined by a bounded number k of mod-p forms has probability either equal to 0 or bounded below in a way that depends on p and k only. This implies in particular a positive lower bound that depends only on p on the probability, for x chosen uniformly in $\{0,1\}^n$, of the event $\phi_j(x) = 0, \phi_i(x) \neq 0$, and this now allows us to conclude as we did for $p = 2$.

Example 3.1. *Let $p \geq 2$ be a prime. The subsets $\mathcal{A}_1 = \phi_1^{-1}(0), \dots, \mathcal{A}_t = \phi_t^{-1}(0)$ of $\{0,1\}^n$ each have density at least $2^{-(p-1)}$ inside $\{0,1\}^n$ and satisfy*

$$\mathrm{TV}_{\mathcal{A}_i,\mathcal{A}_j}\, \phi_i \geq \mathbb{P}_{\mathcal{A}_i}[\phi_i = 0] - \mathbb{P}_{\mathcal{A}_j}[\phi_i = 0] \geq 2^{-2(p-1)}$$

for any pair (i,j) of distinct elements of $[t]$.

We will nonetheless begin by showing that close distributions can be obtained if we only consider mod-p forms with bounded support size and assume the number of dense subsets of $\{0,1\}^n$ to grow superpolynomially in n.

Proposition 3.2. *Let $\delta > 0$, let $k \geq 1$ be an integer, and let $\nu > 0$. Then there exists $c_{\mathrm{TV}}(\delta,\nu,k) > 0$ such that that the following holds. If s is a positive integer and $\mathcal{A}_1, \dots, \mathcal{A}_s$ are non-empty subsets of $\{0,1\}^n$ with density at least δ inside $\{0,1\}^n$, there exists a positive integer $N(\nu,k)$ such that for all $n \geq N(\nu,k)$, there exists a subset $S \subset [s]$ with size at least $s/(n+1)^{c_{\mathrm{TV}}(\delta,\nu,k)}$ such that $\mathrm{Diam}_{\mathcal{A}_i : i \in S}\, g \leq \nu$ for every function g defined on $\{0,1\}^n$ and determined by at most k coordinates.*

Proposition 3.2 does not involve the structure of the function g, and in particular does not require g to be a mod-p form. However, we will use some of the lemmas and ideas involved in the proof of Proposition 3.2 in our later proofs, and in the case where g is a mod-p form it seems worthwhile to compare its conclusion to those that we can or cannot obtain when g is not assumed to have bounded support size, as we will do in the final summarising table of the present section.

We note that for every $k \geq 1$ the power bound on s in terms of n which suffices for the set S in the conclusion of Proposition 3.2 to contain at least two elements cannot be replaced by a too small power of n (how small depends on k).

Example 3.3. *Let $k \geq 1$ be an integer. For every subset I of $[n]$ with size k let ϕ_I be the mod-p form such that $\phi_I(x)$ is defined to be the sum of the coordinates x_i with $i \in I$, reduced modulo p, and let $\mathcal{A}_I = \phi_I^{-1}(0)$. Then the sets $\mathcal{A}_I$ each have density at least $2^{-(p-1)}$, and whenever I, J are distinct subsets of [n] with size k we have $\mathbb{P}_{\mathcal{A}_I}[\phi_I = 0] - \mathbb{P}_{\mathcal{A}_J}[\phi_I = 0] \geq 1/2$.*

Let us now examine what we can or cannot hope for in that case, where we ask for a result on all mod-p forms rather than merely those with bounded support size. We may first aim for a weaker conclusion and ask for a pair (i,j) of distinct elements of $[s]$ such that all mod-p forms have almost positive correlation on the pair $(\mathcal{A}_i, \mathcal{A}_j)$. If the number s of sets $\mathcal{A}_i$ is only known to tend to ∞ with n then the following example shows that for $p \geq 3$ it is still not possible in general to obtain such a pair. To state it, it will be convenient to identify elements $x \in \{0,1\}^n$ with respective subsets $A \subset [n]$ defined by $A = \{i \in [n] : x_i = 1\}$. We note that this example does not apply to $p = 2$, as it then leads to a correlation of zero, rather than to a negative correlation.

Example 3.4. *Let $p \geq 3$ be a prime, let $s = \lfloor n/p \rfloor$, let $Z_1, \dots, Z_s$ be pairwise disjoint subsets of $[n]$ each with size $\frac{p-1}{2}$, and let $\mathcal{A}_i := \{A \subset [n] : |A \cap Z_i| = 0\}$ for each $i \in [s]$. Then for each pair (i,j) of distinct elements of $[s]$ the correlation $\mathrm{cor}_{\mathcal{A}_i,\mathcal{A}_j}\, \phi_{i,j}$ of the form $\phi_{i,j} : A \mapsto |A \cap Z_i| - |A \cap Z_j|$ is equal to $2^{-(p-1)} - 1/p$, which is negative.*

However, we shall show that the statement that we have just been aiming for becomes true if we modify it in either of two ways. In one direction, if s is superpolynomial in n then it is always possible to find the desired pair (i,j) such that every mod-p form has almost positive correlation on the pair $(\mathcal{A}_i, \mathcal{A}_j)$. In another direction, merely having s tend to infinity with n is enough to ensure the existence of such a pair (i,j) of distinct elements of $[s]$ such that every mod-p form has overlap bounded below on the pair $(\mathcal{A}_i, \mathcal{A}_j)$ by some function of p only. The following two theorems will be our main results: one on the correlation and the other on the overlap.

Theorem 3.5. *Let p be a prime, let $m \geq 2$ be an integer, and let $\delta > 0, \nu > 0, \epsilon > 0$. Then there exists $C_{\mathrm{cor}}(\delta, \nu, \epsilon, p, m) > 0$ such that the following holds. If s is a positive integer and $\mathcal{A}_1, \dots, \mathcal{A}_s$ are non-empty subsets of $\{0,1\}^n$ with density at least δ inside $\{0,1\}^n$, then there exists a subset R of $[s]$ with size at least $s/(n+1)^{C_{\mathrm{cor}}(\delta,\nu,\epsilon,p,m)}$ such that for at least $(1-\epsilon)|R|^m$ of m-tuples $(i_1, \dots, i_m) \in R^m$ we have $\mathrm{cor}_{\mathcal{A}_{i_1}, \dots, \mathcal{A}_{i_m}} \phi \geq -\nu$ for every mod-p form.*

Theorem 3.6. *Let p be a prime, let $m \geq 2$ be an integer, and let $\delta > 0, \nu > 0, \epsilon > 0$. Then there exists $c_\omega(\delta, \nu, \epsilon, p, m) > 0$ such that the following holds. If s is a positive integer and $\mathcal{A}_1, \dots, \mathcal{A}_s$ are non-empty subsets of $\{0,1\}^n$ with density at least δ inside $\{0,1\}^n$, then there exists a subset Q of $[s]$ with size at least $c_\omega(\delta, \nu, \epsilon, p, m)s$ such that for at least $(1-\epsilon)|Q|^m$ of the m-tuples $(i_1, \dots, i_m) \in Q^m$ we have $\omega_{\mathcal{A}_{i_1}, \dots, \mathcal{A}_{i_m}} \phi \geq 2^{-(m+1)(p-1)} - \nu$ for every mod-p form.*

Our results and counterexamples allow us to obtain the following table, which summarises for any prime $p \geq 3$ and any integer $m \geq 2$ whether from a collection of dense subsets of the cube we can always find an m-tuple of subsets guaranteeing the properties we have been discussing on all mod-p forms simultaneously, depending on how many subsets we begin with.

Number of subsets $\mathcal{A}_i$	Tending to ∞	$\Omega(n^C)$ for all $C > 0$
Overlapping distributions	Yes	Yes
Overlap bounded below away from zero	Yes	Yes
Close distributions when the support has bounded size	No	Yes
Almost positive correlation	No	Yes
Close distributions	No	No

The role of each reference below is discussed in the full version [10] of the paper.

References

[1] P. Erdős and R. Rado, *Journal of the London Mathematical Society*, Second Series **35** (1960), 85-90.

[2] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem for $k = 3$*, Discrete Math. **75** (1989), 227-241.

[3] H. Furstenberg and Y. Katznelson, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64-119.

[4] W. T. Gowers, *The first unknown case of polynomial DHJ*, https://gowers.wordpress.com/2009/11/14/the-first-unknown-case-of-polynomial-dhj. (Blog post.)

[5] W. T. Gowers and T. Karam, *Equidistribution of high-rank polynomials with variables restricted to subsets of* $\mathbb{F}_p$, arXiv:2209.04932 (2022).

[6] W. T. Gowers and T. Karam, *Low-complexity approximations for sets defined by generalizations of affine conditions*, arXiv:2306.00747 (2023).

[7] W. T. Gowers and T. Karam, *Modular obstructions to the polynomial density Hales-Jewett conjecture*, in preparation.

[8] B. Green and T. Tao, *The distribution of polynomials over finite fields, with applications to the Gowers norms.* Contributions to Discrete Mathematics, **4** (2009), no. 2, 1-36.

[9] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222-229.

[10] T. Karam, *On the expressive power of mod-p linear forms on the Boolean cube*, arxiv:2309.14229 (2023).

[11] T. Karam, *Ranges of polynomials control degree ranks of Green and Tao over finite prime fields*, arXiv:2305.11088 (2023).

[12] D. H. J. Polymath, *A new proof of the density Hales-Jewett theorem*, Ann. of Math. **175** (2012) 1283-1327.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Monochromatic trees in 2-coloured dense graphs

(Extended abstract)

Jasmin Katz* Matías Pavez-Signé† Jozef Skokan‡

Abstract

In 2012, Schelp proposed the following degree version of the Burr-Erdős conjecture on trees: is it true that for any $\varepsilon > 0$ and $\Delta \geq 2$, there is n_0 such that if $n \geq n_0$ and G is a graph on $(2+\varepsilon)n$ vertices with minimum degree $\delta(G) > 3|G|/4$, then every 2-colouring of the edges of G yields a monochromatic copy of each n-vertex tree with maximum degree at most Δ? We prove that Schelp's conjecture is true in a strong form, showing that it holds even if one removes the εn term.

1 Introduction

A graph G is Ramsey for a graph H if every 2-edge colouring of G contains a monochromatic copy of H. For any graph G with a 2-edge colouring we will write $G = R_G \cup B_G$ to indicate the red and blue graph, respectively. For a graph H, the *Ramsey number* $R(H)$ of H is the minimum number n such that K_n is Ramsey for H. In this paper, we are interested in studying Ramsey-type problems for trees. In 1976, Burr and Erdős [4] conjectured that for any n-vertex tree T, the Ramsey number of T satisfies $R(T) \leq 2n-2$ for even n and $R(T) \leq 2n-3$ for odd n. Although this conjecture is still open, we know by a result of Zhao [14] that $R(T) \leq 2n-3$ holds for any n-vertex tree T for sufficiently large odd n.

While the Ramsey number $R(T)$ gives the minimum number of vertices required for a graph to be Ramsey for T, it is interesting to consider whether it is possible to find a sparser graph on $R(T)$ vertices that is Ramsey for T. In this direction, we study how resilient the property of being Ramsey for T is under the deletion of edges. In the uncoloured setting, an influential result by Komlós, Sárközy and Szemerédi [10] states that the property of containing

*Department of Mathematics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. E-mail: j.katz3@lse.ac.uk.

†Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, Universidad de Chile, Beauchef 851, Santiago, Chile. E-mail: mpavez@dim.uchile.cl. Supported by Fondecyt Regular grant No. 1241398 and ANID Basal Grant CMM FB210005.

‡Department of Mathematics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. E-mail: j.skokan@lse.ac.uk.

all n-vertex trees with bounded maximum degree remains true even after an adversary deletes a $(1/2 - o(1))$-fraction of the neighbourhood of each vertex in K_n.

Theorem 1.1. *For all $\delta \in (0,1)$, there are n_0 and $c \in (0,1)$ such that the following is true for all graphs G and for all trees T with $|V(G)| = |V(T)| = n \geq n_0$. If $\Delta(T) \leq \frac{cn}{\log n}$ and $\delta(G) \geq (1+\delta)\frac{n}{2}$, then T is a subgraph of G.*

In 2012, Schelp [13, Question 22] asked for which graphs H does there exist a constant $0 < c < 1$ such that every graph G of order $R(H)$ with $\delta(G) > c|H|$ is Ramsey for H. Schelp made several conjectures of this nature for many classes of graphs - namely for cycles, paths and trees of bounded maximum degree.

For these classes of graphs, it was conjectured that the correct constant c is $3/4$. In the case of paths, Schelp conjectured if $|G| = R(P_n)$ and $\delta(G) > 3|G|/4$, then G is Ramsey for P_n, for which an approximate version was proved in [9]. In the case of cycles, it is known $R(C_n) \leq 2n - 1$ for all $n \geq 4$, in which case Li, Nikiforov and Schelp [11] conjectured that any 2-edge coloured n-vertex graph G with minimum degree more than $3n/4$ contained a monochromatic cycle of length ℓ for all $\ell \in [4, \lceil n/2 \rceil]$. This conjecture was proven by Benevides, Łuczak, Scott, Skokan and White [1].

Here we address Schelp's question for bounded degree trees [13, Question 12], which was given as follows. Let $\varepsilon > 0, \Delta \in \mathbb{N}$ and n be large, and let G be a graph of order $(2+\varepsilon)n$ and minimum degree $\delta(G) > 3|G|/4$. If T is an n-vertex tree with $\Delta(T) \leq \Delta$, is G Ramsey for T? We answer this question in the affirmative, and show that the εn term is not necessary.

Theorem 1.2. *For all $\Delta \geq 2$ there is some $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. If G is a 2-edge-coloured graph with $|G| = 2n-1$ and $\delta(G) \geq \lfloor 3|G|/4 \rfloor$, then G contains a monochromatic copy of every n-vertex tree with maximum degree at most Δ.*

The extremal example is given by a graph G^* on $2n-2$ vertices, which we describe now. The vertices of G^* are partitioned into four sets W_1, W_2, W_3 and W_4 of orders $n/2, n/2-1, n/2$ and $n/2-1$ respectively (assuming n is even). The edge set of this graph is as follows. Each $G^*[W_i]$ forms a complete graph (the edge colouring doesn't matter). Both $G^*[W_1, W_2]$ and $G^*[W_3, W_4]$ form complete bipartite graphs, with all edges coloured blue. Finally, $G^*[W_1, W_4]$ and $G^*[W_2, W_3]$ form complete bipartite graphs with all edges coloured red. The minimum degree of this graph is $3n/2 - 3 = \lfloor 3|G^*|/4 \rfloor - 1$ (this is the degree of the vertices in W_1 and W_3), it does not contain any tree on n vertices for the reason that each monochromatic component only covers $n-1$ vertices.

2 Sketch of Proof

Given $\Delta \geq 2$, we let n be large enough and let G be an arbitrary graph satisfying the assumptions of Theorem 1.2 with $G = R_G \cup B_G$ a 2-edge colouring of G. We use a stability approach, where we find that either G contains a monochromatic copy of every bounded-degree tree, or the structure of G is close to the extremal example G^*. As with many of the results in this area, our proof uses the regularity method. We choose some $0 < \varepsilon \ll d \ll \delta \ll 1/\Delta$, and apply a 2-edge coloured version of Szemeredi's Regularity Lemma to our graph to obtain an (ε, d)-reduced 2-edge coloured graph H. This gives a partition of $V(G)$ into a bounded number of *clusters*, whereby the bipartite subgraphs between most of the clusters behave *pseudorandomly* in each colour class. These clusters then form the vertex set of the reduced graph $H = R_H \cup B_H$. Given any set of clusters $S \subseteq V(H)$, we let $\bigcup S \subseteq V(G)$ denote the union of the vertices in the clusters of S. As the reduced graph H inherits a weakened version of the minimum degree condition of G, that is $\delta(H) \geq (3/4 - \delta)|H|$, we may apply the following lemma from [1] to H .

Lemma 2.1 (Lemma 4.1 in [1]). *Let $0 < \delta < 1/36$, and let H be a graph of sufficiently large order k with $\delta(H) \geq (3/4 - \delta)k$. Suppose we are given a 2-edge colouring $H = R_H \cup B_H$. Then one of the following holds.*

(i) There is a non-bipartite component of R_H or B_H which contains a matching on at least $(1/2 + \delta)k$ vertices.

(ii) There is a bipartite component C of R_H or B_H which contains a matching on at least $(1/2 + \delta)k$ vertices and $|V(C)| \geq (1 - 5\delta)k$.

(iii) There is a partition $V(H) = U_1 \cup \dots \cup U_4$ with $\min_i |U_i| \geq (1/4 - 3\delta)k$ such that there are no blue edges from $U_1 \cup U_2$ to $U_3 \cup U_4$ and no red edges from $U_1 \cup U_3$ to $U_2 \cup U_4$.

In case (i), we have a monochromatic, connected, and non-bipartite cluster matching in the reduced graph, which covers more than n vertices of G. A result from [2, Proposition 5.8] implies that any bounded degree tree can be embedded into such a structure.

Proposition 2.2. *Let $\Delta \geq 2, \varepsilon > 0$ and let n be sufficiently large. Let T be a tree on n vertices with $\Delta(T) \leq \Delta$. Let G be a 2-edge coloured graph with $(\varepsilon, 5\sqrt{\varepsilon})$-reduced graph H. Suppose H contains a monochromatic non-bipartite component containing a matching M with $|M| \geq (1 + 100\sqrt{\varepsilon})n|H|/|G|$ vertices. Then G contains a monochromatic copy of T.*

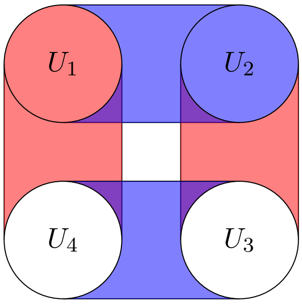
Case (ii) splits into two cases. Let C have 2-edge colouring $C = R_C \cup B_C$, and assume B_C is the bipartite component (these arguments hold if red and blue are reversed). In the first case, the reduced bipartite graph is *sufficiently unbalanced*, meaning that the largest bipartition class X of B_C satisfies $|X| > (1/2 + 30\delta)|H|$. In this case we consider the graph $R_G[\bigcup X]$. As there are few blue edges between these vertices, a simple calulation shows that this graph has minimum degree more than $(1/2 + \delta)n$, and we are able to apply Theorem 1.1 to find a monochromatic copy of every bounded degree tree.

Otherwise, we say C is *almost-balanced*. In this case, we require the following lemma.

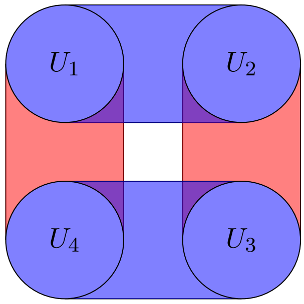
Lemma 2.3. *Given $\Delta \geq 2$, let $\Delta^{-1} \gg d \gg \varepsilon > 0$ and let n be sufficiently large. Let T be a tree on n vertices with $\Delta(T) \leq \Delta$. Let G be a 2-edge coloured graph on $2n - 1$ vertices with (ε, d)-reduced graph H. If H contains a monochromatic bipartite component C with almost balanced bipartition classes covering at least $(1 - \delta)|H|$ clusters, and minimum degree at least $(1 - \delta)|H|/4$, then G contains a monochromatic copy of T.*

The proof of this lemma uses a connected cluster-matching argument, similar to the proof of Proposition 2.2, and so we omit the details here. We apply this lemma in the following way. If there are no red edges between the bipartition classes of C, we are able to conclude that the minimum blue degree of C is at least $(1 - \delta)|H|/4$ and thus we can find a copy of every bounded degree tree. However, if there are red edges between the bipartition classes of C, the blue degree may not be high enough. We instead use a version of Dirac's theorem to analyse R_C. This allows us to either find a large monochromatic non-bipartite component in R_C, in which case we can conclude by Proposition 2.2, or we find that either R_C or B_C has minimum degree $(1 - \delta)|H|/4$, and so we can conclude by Lemma 2.3.

In case (iii), the structure of the graph is already close to the extremal example. However, it is possible that there may only be $(1-\varepsilon/2)n$ vertices contained in the clusters of the largest monochromatic component of the reduced graph– and hence it is not possible to embed using only the properties of the regular partition.



R_H and B_H are non-bipartite.



R_H is bipartite.

We again distinguish between two cases. In the first case, there is some $i \in [4]$ such that $R_H[U_i] \neq \emptyset$ and some $j \in [4]$ such that $B_H[U_j] \neq \emptyset$, implying that both R_H and B_H are non-bipartite. Let $W_i = \bigcup U_i$ be the set of vertices in G contained in the clusters of each U_i. We use the following lemma.

Lemma 2.4. *Given $\Delta \geq 2$, let $\Delta^{-1} \gg d \gg \varepsilon > 0$ and let n be sufficiently large. Let T be a tree on n vertices with $\Delta(T) \leq \Delta$. Let G be a 2-edge coloured graph on $2n-1$ vertices with (ε, d)-reduced graph H. Let U_i be partition of $V(H)$ satisfying Lemma 2.1* (iii)*, such that $R_H[U_1] \neq \emptyset$ and $B_H[U_2] \neq \emptyset$. Let $W_i = \bigcup U_i$. Then the following hold.*

1. *If there is a vertex $v \in V(G)$ such that v has at least one blue neighbour in $W_1 \cup W_2$ and at least $\Delta - 1$ blue neighbours in $W_3 \cup W_4$, then B_G contains a copy of T.*

2. *If there is a vertex $v \in V(G)$ such that v has at least one red neighbour in $W_1 \cup W_4$ and at least $\Delta - 1$ red neighbours in $W_2 \cup W_3$, then R_G contains a copy of T.*

For each $i \in [4]$,the structure of H implies that either $\delta(R_G[W_i, W_{i+1}]) \geq (1-\delta)|W_{i+1}|$ or $\delta(B_G[W_i, W_{i+1}]) \geq (1-\delta)|W_{i+1}|$ (where the indices are taken modulo four). We use Lemma 2.4 to conclude that either we can find a monochromatic copy of of every bounded degree tree, or we can assign each vertex of $G \setminus \bigcup_{i\in[4]} W_i$ to some W_i, while maintaining this high monochromatic minimum degree condition. Finally, using that $|G| = 2n-1$, we find some W_i of size at least $\lceil n/2 \rceil$. As $\delta(G) \geq \lfloor 3|G|/4 \rfloor$, each vertex of W_{i+2} must have a neighbour in W_i. Whichever colour this edge is, we can find a monochromatic copy of every bounded degree tree using Lemma 2.4.

In the second case, either R_H or B_H is bipartite– meaning that all U_i are monochromatic. Assume $R_H(U_i) = \emptyset$ for all $i \in [4]$ (these arguments hold if red and blue are reversed). In this case, we cannot hope to have a statement as strong as in Lemma 2.4 in the colour red, as there are examples of unbalanced trees which cannot be embedded easily into bipartite structures. Instead, note that both $B_G(W_1 \cup W_2)$ and $B_G(W_3 \cup W_4)$ are almost complete, so we look to find a copy of every bounded degree tree in one of these components. For this, we still need to add the vertices of $G \setminus \bigcup_{i\in[4]} W_i$. We find that we are able to embed each bounded-degree tree if some vertex has red degree Δ to at least three W_i and then use this observation to conclude that either we can embed each bounded-degree tree in red, or we can assume that every vertex of $G \setminus \bigcup_{i\in[4]} W_i$ has reasonable blue degree to one of the blue components. In the last case, we can then find component of B_G of order at least n and density high enough to embed every bounded degree tree.

3 Remarks

The upper bound for the Ramsey number of an n-vertex tree T given by $R(T) \leq 2n-2$ is tight, with a star on an even number of vertices as the extremal example. However, stars are the most unbalanced trees possible, in the sense that one of the bipartition classes consists of a single vertex. Gerencsér and Gyárfás [8] showed that for a path P on n vertices (the most balanced tree on n vertices), the Ramsey number is given by $R(P) = \lfloor 3n/2 \rfloor - 1$, quite far from the upper bound $R(P) \leq 2n-2$.

These examples suggest a relation between the order of the bipartition classes of a tree and its Ramsey number. In 1974, Burr [3] conjectured that if T is a tree with bipartition classes of order n_1 and n_2, with $n_1 \geq n_2 \geq 2$, then $R(T) = \max\{n_1 + 2n_2, 2n_1\} - 1$. This conjecture has been disproved, as Norin, Sun and Zhao [12] showed that there are trees T with $n_1 = 2n_2$ and $R(T) \geq 4.2n_2 - o(n_2)$. Nonetheless, it remains open for bounded-degree trees. Letting $R_B(T) := \max\{n_1 + 2n_2, 2n_1\} - 1$, we believe it would be interesting to answer the following question: for any $\Delta > 2, \varepsilon > 0$ and n large enough, letting T be a tree on n vertices with $\Delta(T) \leq \Delta$, does every 2-edge coloured graph G on $(1+\varepsilon)R_B(T)$ vertices with $\delta(G) > 3|G|/4$ contain a monochromatic copy of T?

References

[1] F. S. Benevides, T. Łuczak, A. Scott, J. Skokan, and M. White. Monochromatic cycles in 2-coloured graphs. *Combin. Probab. Comput.*, 21(1-2):57–87, 2012.

[2] G. Besomi, M. Pavez-Signé, and M. Stein. Degree conditions for embedding trees. *SIAM J. Discrete Math.*, 33(3):1521–1555, 2019.

[3] S. A. Burr. Generalized Ramsey theory for graphs—a survey. In *Graphs and combinatorics (Proc. Capital Conf., George Washington Univ., Washington, D.C., 1973)*, volume Vol. 406 of *Lecture Notes in Math.*, pages 52–75. Springer, Berlin-New York, 1974.

[4] S. A. Burr and P. Erdős. Extremal Ramsey theory for graphs. *Utilitas Math.*, 9:247–258, 1976.

[5] G. A. Dirac. Some theorems on abstract graphs. *Proc. London Math. Soc. (3)*, 2:69–81, 1952.

[6] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. The size Ramsey number. *Period. Math. Hungar.*, 9(1-2):145–161, 1978.

[7] C. G. Fernandes, T. Naia, G. Santos, and M. Stein. Packing large balanced trees into bipartite graphs, 2024. `arXiv:2410.13290v1 [math.CO]`.

[8] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math*, 10:167–170, 1967.

[9] A. Gyárfás and G. N. Sárközy. Star versus two stripes Ramsey numbers and a conjecture of Schelp. *Combin. Probab. Comput.*, 21(1-2):179–186, 2012.

[10] J. Komlós, G. N. Sárközy, and E. Szemerédi. Spanning trees in dense graphs. *Combin. Probab. Comput.*, 10(5):397–416, 2001.

[11] H. Li, V. Nikiforov, and R. H. Schelp. A new class of Ramsey-Turán problems. *Discrete Math.*, 310(24):3579–3583, 2010.

[12] S. Norin, Y. R. Sun, and Y. Zhao. Asymptotics of Ramsey numbers of double stars, 2016. `arXiv:1605.03612v1 [math.CO]`.

[13] R. H. Schelp. Some Ramsey-Turán type problems and related questions. *Discrete Math.*, 312(14):2158–2161, 2012.

[14] Y. Zhao. Proof of the $(n/2 - n/2 - n/2)$ conjecture for large n. *Electron. J. Combin.*, 18(1):Paper 27, 61pp., 2011.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Coloring Grids Avoiding Bicolored Paths

(Extended abstract)

Derman Keskinkılıç* Lale Özkahya†

Abstract

The star chromatic number on a graph is the minimum number of colors in a proper vertex coloring forbidding any P_4 with two colors (bicolored). This problem was introduced by Grünbaum (1973) together with the acyclic coloring of graphs, where bicolored cycles are avoided. In this paper, we study a generalization of this problem, by considering proper vertex coloring on graphs forbidding bicolored paths of a fixed length that was initially discussed by Alon, Mcdiarmid, and Reed (1991). Here, we study this problem on products of two paths. We show that at least 4 colors are needed to properly color the product of paths, $P_m \square P_n$, avoiding a bicolored P_k, unless $n < k-2$ or $m < k-2$. With this result, the above question is settled for all k on 2-dimensional grids.

1 Introduction

The *star coloring* problem on a graph G asks to find the minimum number of colors in a proper coloring forbidding a bicolored (2-colored) P_4, called the star-chromatic number $\chi_s(G)$. This problem is introduced by Grünbaum [8], who proved that a graph with maximum degree 3 has an acyclic coloring with 4 colors. Similarly, *acyclic chromatic number* of a graph G, $a(G)$, is the minimum number of colors used in a proper coloring not having any bicolored cycle, also called acyclic coloring of G [8]. Both, the star coloring and acyclic coloring problems are shown to be NP-complete by Albertson et al. [1] and Kostochka [12], respectively.

The star coloring problem has been studied widely on many different graph families, such as product of graphs, planar, and outerplanar graphs, in [1, 10, 13] to name a few. Similarly, acyclic coloring of these graph families has been studied widely, such as [3, 4]. Alon, Mcdiarmid, and Reed [2] proved that there exist graphs G with maximum d for which $a(G) = \Omega((d^{\frac{4}{3}})/(logd)^{\frac{1}{3}})$. In [2], it is also shown that for any graph G with maximum degree d, $a(G) = O(d^{\frac{4}{3}})$. Recently, there have been some improvements in the constant factor of the

*Department of Computer Engineering at Hacettepe University, Beytepe, Ankara, Turkiye. E-mail: dermanakgol@gmail.com.

†Department of Computer Engineering at Hacettepe University, Beytepe, Ankara, Turkiye. E-mail: ozkahya@cs.hacettepe.edu.tr.

upper bound in [5, 7, 14] by using the entropy compression method. Similar results for the star chromatic number of graphs are obtained by Fertin et al. [6], showing $\chi_s(G) \leq \lceil 20d^{3/2} \rceil$ for any graph G with maximum degree d. Alon, Mcdiarmid, and Reed claim in [2] that an upper bound similar to above can be shown when a bicolored path, P_k, is not allowed in a proper vertex coloring. In [5], this chromatic number is studied for paths of even order, and later studied for all paths in [9] and [11]. This problem has been also generalized to subgraphs other than paths.

In this paper, we study this problem on products of paths, in particular on 2-dimensional grids. We call a proper vertex coloring of a graph G without a bicolored copy of P_k a *P_k-coloring* of G for $k \geq 4$. The minimum number of colors needed for a P_k-coloring of G is called *P_k-chromatic number* of G, denoted by $s_k(G)$, where the value for $k = 4$ corresponds to the star chromatic number. In [11], Kırtışoğlu and the second author show that $s_k(P_{k-3} \square P_n) = 3$ for all $k \geq 5$ and $n \geq 1$, by providing the following colorings for the case $k = 5, 6$, which can be generalized to all $k \geq 5$.

$$\begin{array}{ccccccc} 1 & 2 & 3 & 1 & 2 & 3 & \ldots \\ 2 & 3 & 1 & 2 & 3 & 1 & \ldots \end{array} \qquad \begin{array}{ccccccc} 1 & 2 & 3 & 1 & 2 & 3 & \ldots \\ 2 & 3 & 1 & 2 & 3 & 1 & \ldots \\ 1 & 2 & 3 & 1 & 2 & 3 & \ldots \end{array}$$

This coloring pattern having columns with alternating colors from $(1, 2)$, $(2, 3)$, $(3, 1)$, respectively, yields a valid 3-coloring for any $k \geq 6$ and $n \geq 1$, showing $s_k(P_{k-3} \square P_n) = 3$. Note that in such colorings, a bicolored P_k has to have at least $k - 2$ vertices in the same column, thus cannot be found in $P_{k-3} \square P_n$ colored according to the pattern above. In [11], it is also observed that for $k = 5, 6$, $s_k(P_{k-2} \square P_n) = 4$ for all $n \geq k - 2$ and this is conjectured to hold for all k. With our main theorem below, we confirm this conjecture showing that there is no proper 3-coloring of $P_m \square P_n$ avoiding a bicolored P_k, for $m, n \geq k - 2$.

Theorem 1. *For any $k \geq 5$, $s_k(P_{k-2} \square P_{k-2}) = 4$, thus giving $s_k(P_m \square P_n) = 4$ for all $n, m \geq k - 2$.*

2 Main Result

We call a maximal connected subgraph induced by vertices having only two colors a *bicolored component.* To prove Theorem 1, we analyze bicolored components containing vertices from anyone of the sides of the grid. These components belong to one of the groups below:
1) *complete bicolored component:* a component that has vertices in two opposite sides of the grid, i.e., top and bottom sides, or left and right sides.
2) *partial bicolored component:* a component that is not complete, but has vertices on at least one of the sides of the grid.
In the remaining, we assume that the sides of the grid that a partial component may intersect are top and left sides, since remaining cases are symmetric. We categorize each partial bicolored component C as:
Type 1: if, w.l.o.g., the vertices of C are only on the top side,
Type 2: if the vertices of C are on the top and left sides.
We see examples of type-1 and type-2 partial bicolored (as red-blue colored) components in Figure 1.(a), and in Figure 1.(b),(c), respectively. The following definitions are associated with a (partial or complete) bicolored component C :

- *Boundary of C:* the walk traversing the outer face of C in clockwise direction when C is considered as a planar subgraph in the grid drawing. For example, in Figure 1.(c), the boundary of C is the walk (starting at any vertex) $(s, a, b, a, d, e, t, e, d, a, s)$.

- *Partial walk* B^C*:* A maximal segment of the boundary of C such that no edge on the sides of the grid is traversed from the outside of the grid. In addition, if C is a partial bicolored component, we let the starting vertex of B^C be the rightmost vertex of C on the top side of the grid. Hence, it is possible to have more than one partial walk on the boundary of C only if C is a complete bicolored component. Some examples for the partial walks are shown in Figure 1, where B^C is described by the vertex sequence $(s, a, b, a, c, a, d, a, t)$, (s, a, b, c, d, c, t), (s, a, b, a, d, e, t), and $(s, a, b, a, c, d, e, f, t)$, respectively. In Figure 1.(a), the boundary of C happens to be the same as B^C.
- s^C, t^C: The first and last vertex on B^C, respectively. For example, in Figure 1.(c), continuing the walk B^C after t would traverse the edge te, hence traversing the left side of the grid from the outside. Similarly, the start vertex, s, in the examples in Figure 1, is chosen according to this maximality property of B^C.

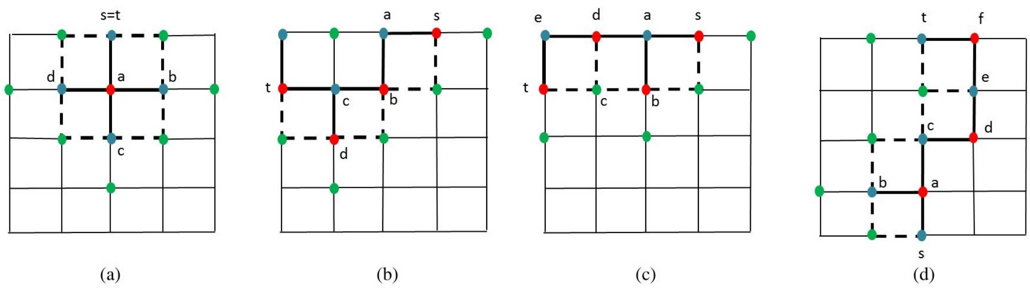


Figure 1: Examples of bicolored components showing the edges of B^C (bold edges) and the neighboring edges of another bicolored component (dashed edges). In (d), only B^C is shown, without showing the entire bicolored component C.

In the remaining, we use shortly *3-coloring* meaning a proper 3-coloring of $V(P_{k-2}\Box P_{k-2})$, $k \geq 5$, using colors red, blue and green. With Lemma 2 below, we make a generalization about the boundary structure of bicolored components.

Lemma 2. *For any $k \geq 5$, let C be a (partial or complete) bicolored component in a 3-coloring of $P_{k-2}\Box P_{k-2}$, and label the vertices along a partial walk of C, B^C, as $v_1, v_2, \ldots, v_r$, where $v_1 = s^C$ and $v_r = t^C$. Then, the following hold:*

1. *r is an odd integer and $r \geq 3$.*
2. *the angle between the edges v_iv_{i+1} and $v_{i+1}v_{i+2}$, $i \leq r-2$, along B^C is $90°$ if and only if i is odd.*
3. *If C is, w.l.o.g., red-blue colored, then there is a α-green colored connected subgraph, α being the color of v_1, induced by the vertices v_i and $u_{(i+1)/2}$, for odd i, $1 \leq i \leq r-2$, where $u_{(i+1)/2}$ is the green vertex on the C_4 containing $\{v_i, v_{i+1}, v_{i+2}\}$.*

In Figure 2, we show the only possible cases of angles (omitting symmetric cases) between consecutive edges along B^C, where w, x, y, z represent the vertices $v_iv_{i+1}, v_{i+2}, v_{i+3}$, respectively, for some i along B^C. The leftmost case in this figure shows that if two consecutive angles along B^C are $90°$, then this contradicts with the fact that x and y are on the boundary. Lemma 3 below is used to show that Theorem 1 holds in case there is a complete bicolored component in a 3-coloring of $P_{k-2}\Box P_{k-2}$, $k \geq 5$.

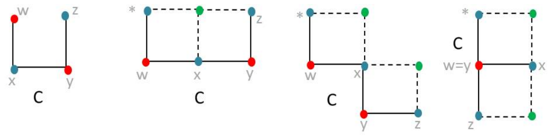


Figure 2: The only possible configurations (omitting symmetric cases) along B^C in terms of the angles between consecutive triples of vertices. The first case shows that two consecutive angles cannot be both 90°. Vertices marked with * are part of C.

Lemma 3. *For any $k \geq 5$, if a 3-coloring of $P_{k-2} \Box P_{k-2}$ has a complete bicolored component, then there is a bicolored P_k.*

Proof of Theorem 1: We know that $s_k(P_m \Box P_n) \leq 4$ as $s_5(P_m \Box P_n) = 4$, for any $m, n \geq 3$, $k \geq 5$ ([11]). To prove that $s_k(P_m \Box P_n) \geq 4$, it suffices to show that $s_k(P_{k-2} \Box P_{k-2}) \geq 4$. We consider any 3-coloring of $P_{k-2} \Box P_{k-2}$ for $k \geq 5$, and show that it has a complete bicolored component, hence a bicolored P_k by Lemma 3.

Assume that such a 3-coloring does not have a complete bicolored component. Let C be a partial, w.l.o.g., red-blue colored component containing vertices from top side and possibly left side of the grid. Let $B^C = (v_1, v_2, \ldots, v_r)$ be a partial walk described earlier, where $v_1 = s^C$and $v_r = t^C$. By Lemma 2, for each odd $1 \leq i \leq r-2$, the angle between the edges $v_i v_{i+1}$ and $v_{i+1} v_{i+2}$, along B^C is 90° and there is a α-green colored connected subgraph, α being the color of v_1, induced by v_i and $u_{(i+1)/2}$, for odd i, $1 \leq i \leq r-2$, where $u_{(i+1)/2} \notin C$ is the green vertex on the C_4 containing $\{v_i, v_{i+1}, v_{i+2}\}$, call it A^C. Let D^C be the bicolored component containing A^C. If D^C is a complete bicolored component, we are done by Lemma 3. Otherwise, D^C is a partial bicolored component satisfying the same assumptions on C. Thus, we keep replacing C with D^C and find the new D^C iteratively, until it is a complete bicolored component. As D^C has vertices that are not contained by any C used in earlier iterations, this procedure stops successfully after a finite number of iterations and we are done by Lemma 3.

References

[1] Michael O Albertson, Glenn G Chappell, Hal A Kierstead, André Kündgen, and Radhika Ramamurthi. Coloring with no 2-colored p_4's. *the electronic journal of combinatorics*, pages R26–R26, 2004.

[2] Noga Alon, Colin Mcdiarmid, and Bruce Reed. Acyclic coloring of graphs. *Random Structures & Algorithms*, 2(3):277–288, 1991.

[3] Noga Alon, Bojan Mohar, and Daniel P Sanders. On acyclic colorings of graphs on surfaces. *Israel Journal of Mathematics*, 94(1):273–283, 1996.

[4] Oleg V Borodin. On acyclic colorings of planar graphs. *Discrete Mathematics*, 25(3):211–236, 1979.

[5] Louis Esperet and Aline Parreau. Acyclic edge-coloring using entropy compression. *European Journal of Combinatorics*, 34(6):1019–1027, 2013.

[6] Guillaume Fertin, André Raspaud, and Bruce Reed. Star coloring of graphs. *Journal of Graph Theory*, 47(3):163–182, 2004.

[7] Daniel Gonçalves, Mickaël Montassier, and Alexandre Pinlou. Entropy compression method applied to graph colorings. In *ICGT: International Colloquium on Graph Theory and Combinatorics*, Grenoble, France, June 2014.

[8] Branko Grünbaum. Acyclic colorings of planar graphs. *Israel journal of mathematics*, 14(4):390–408, 1973.

[9] Jianfeng Hou and Hongguo Zhu. Coloring graphs without bichromatic cycles or paths. *Bulletin of the Malaysian Mathematical Sciences Society*, pages 1–13, 2020.

[10] Hal A Kierstead, André Kündgen, and Craig Timmons. Star coloring bipartite planar graphs. *Journal of graph theory*, 60(1):1–10, 2009.

[11] Alaittin Kırtışoğlu and Lale Özkahya. Coloring of graphs avoiding bicolored paths of a fixed length. *Graphs and Combinatorics*, 40(1):1–11, 2024.

[12] Alexandr V Kostochka. Upper bounds of chromatic functions of graphs. In *Doct. Thesis.* Novosibirsk, 1978.

[13] Alexandr V Kostochka and Leonid S Mel'nikov. Note to the paper of grünbaum on acyclic colorings. *Discrete Mathematics*, 14(4):403–406, 1976.

[14] Sokol Ndreca, Aldo Procacci, and Benedetto Scoppola. Improved bounds on coloring of graphs. *European Journal of Combinatorics*, 33(4):592–609, 2012.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Blocking Planes by Lines in PG(n, q)

(Extended abstract)

Benedek Kovács* Zoltán Lóránt Nagy† Dávid R. Szabó‡

Abstract

In this paper, we study the cardinality of the smallest set of lines of the finite projective spaces PG(n, q) such that every plane is incident with at least one line of the set. This is the first main open problem concerning the minimum size of (s, t)-blocking sets in PG(n, q), where we set $s = 2$ and $t = 1$. In PG(n, q), an (s, t)-blocking set refers to a set of t-spaces such that each s-space is incident with at least one chosen t-space. This is a notoriously difficult problem, as it is equivalent to determining the size of certain q-Turán designs and q-covering designs. We present an improvement on the upper bounds of Etzion and of Metsch via a refined scheme for a recursive construction, which in fact enables improvement in the general case as well.

1 Introduction

Let q be a prime power and let $\mathbb{F}_q$ denote the finite field of q elements. PG(n, q) denotes the n-dimensional projective space over $\mathbb{F}_q$, while a subspace of dimension d will be called a d-space.

In the history of finite geometry, one of the most crucial central problems is to describe the size and structure of *blocking sets*. A blocking set with respect to s-dimensional subspaces is a point set of PG(n, q) which meets every s-dimensional subspace in at least one point. As observed by Bose and Burton, the point set of an $(n-s)$-dimensional subspace is the smallest possible blocking set with respect to s-dimensional subspaces [4]. From this classical result, a whole theory of so-called Bose-Burton type theorems has emerged. Indeed, the above concept has some far-reaching generalisations, when one can take sets of subspaces of a given dimension t instead of point sets (which correspond to the case $t = 0$). Let $t \leq s$ be nonnegative integers.

*ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. Supported by the EKÖP-24 University Excellence Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund and the University Excellence Fund of Eötvös Loránd University. E-mail: `benoke98@student.elte.hu`

†ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. The author is supported by the Hungarian Research Grant (NKFIH) No. PD 134953. and No. K. 124950 and the University Excellence Fund of Eötvös Loránd University E-mail: `nagyzoli@cs.elte.hu`

‡ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. The author is supported by the University Excellence Fund of Eötvös Loránd University E-mail: `szabo.r.david@gmail.com`

In this general context, an (s,t)-*blocking set* B is a set of t-spaces of PG(n,q) such that every s-space of PG(n,q) contains at least one element of B.

An example for (s,t)-blocking sets can be obtained from geometric t-spreads, first introduced by André [1].

Definition 1.1 (t-spread)**.** A *t-spread* of a projective space PG(n,q) is a set of t-dimensional subspaces for some $t < n$, such that every point of the space lies in exactly one of the elements of the spread.

The necessary and sufficient condition for the existence of t-spreads in PG(n,q) is that $(t+1) \mid (n+1)$.

Beutelspacher and Ueberberg found the minimal (s,t)-blocking sets in the following cases.

Theorem 1.2 (Beutelspacher and Ueberberg, [2])**.** *Suppose that $\mathcal{B}$ is an (s,t)-blocking set. If $n \leq s + s/t - 1$ or $t = 0$, then*

$$|\mathcal{B}| \geq (q^{(t+1)(n+1-s)} - 1)/(q^{t+1} - 1),$$

where equality is attained if and only if $\mathcal{B}$ is a so-called geometric t-spread in a subspace of dimension equal to $(n+1-s)(t+1)-1$.

In the last fifty years the general problem of determining the smallest cardinality of a blocking set has been studied by several authors both in the original case when $t = 0$, and the general case ($t \in \mathbb{N}$) (see [3, 23] and references therein). The original blocking set problem has generated a considerable amount of research, eventually leading up to the Linearity Conjecture of Sziklai, see [26]. Note that the general problem can be considered as a q-analogue of Mantel's theorem about triangle-free graphs. A connection to maximum rank distance codes [15] and to q-designs [5, 13] has been investigated [15], which topics have been extensively studied later on (see [17], [9], and [16], and the references therein). In recent years, additional motivation for such results has been highlighted by the increasing interest in codes over the Grassmannian as a result of their application to error correction in network coding as was demonstrated by Koetter and Kschischang [21]. Even so, only a few instances are known where the problem is completely solved [4, 22, 25].

Known results in the case $t = 1$. Consider the problem of finding minimal $(s,1)$-blocking sets in PG(n,q). Theorem 1.2 solves the problem on the interval $s \leq n \leq 2s - 1$. In the case of the subsequent interval $2s \leq s \leq 3s - 3$, by pulling back the $(q^{2s}-1)/(q^2-1)$ lines forming a minimal $(s,1)$-blocking set in a suitable $2s-1$-dimensional quotient space, Metsch constructed an $(s,1)$-blocking set in PG(n,q). This construction is minimal and essentially is the unique minimal one.

Theorem 1.3 (Eisfeld and Metsch, [11, Theorem 1.2.], [22, Theorem 1.2])**.** *Let $2s \leq n \leq 3s-3$ for $s \geq 3$, or $n = 4$ for $s = 2$. Let q be a prime power. Write $k := n - 2s$. Assume further that $q \geq 4^{k+1} + 2^{k+1} + 1$ if $n \neq 2s$. Then every $(s,1)$-blocking set $\mathcal{B}$ in* PG(n,q) *satisfies*

$$|\mathcal{B}| \geq q^{2k+2} \cdot \frac{q^{2s}-1}{q^2-1} + \sum_{i=0}^{k} q^i \sum_{j=k}^{n-s} q^j.$$

Furthermore, there are blocking sets $\mathcal{B}$ with equality above, and the structure of such sets is known.

The case $(s,t)=(2,1)$**.** In the examples above, n is bounded. Our main intention is to determine the size of (s,t)-blocking sets in projective spaces of higher dimension, at least when $(s,t)=(2,1)$. Even in this case, the problem is solved only in low dimensions. Our main function will be as follows.

Definition 1.4. Let $f(n,q)$ denote the smallest possible size of a $(2,1)$-blocking set in PG(n,q) for $n\geq -1$.

Apart from the trivial case $n=2$, Theorem 1.2 and Theorem 1.3 gives the following known values.

Theorem 1.5. *For every prime power q, we have $f(2,q)=1$, $f(3,q)=q^2+1$, $f(4,q)=q^4+2q^2+q+1$.*

While the proof for $n=3$ was not very involved, the 4-dimensional case required a tour-de-force proof [11]. In fact, Metsch [22] notes that his refined double counting technique even leads to an exact result for $n=5$ as well with $f(5,q)=q^6+2q^4+2q^3+2q^2+q$ for large enough q, but the argument remained unpublished.

Let us point out that these results translate to results concerning q-covering designs and q-Turán designs, which will be detailed in the next section. Note that $f(n,q)$ is the minimum size of a q-Turán design $T_q[n+1,3,2]$.

In general, the q-analog Schönheim bound, formulated by Etzion and Vardy [15], gives the best known lower bound. It provides a lower bound for every (s,t)-blocking set; below we apply it to the case $(s,t)=(2,1)$.

Theorem 1.6 (Etzion, Vardy, [15])**.** $f(n,q)\geq\left\lceil\frac{q^{n+1}-1}{q^{n-1}-1}\cdot f(n-1,q)\right\rceil$.

If one applies Proposition 1.6 recursively and omits the ceiling, this leads to the following explicit lower bound.

Proposition 1.7 (Lower bound)**.** *For every $n\geq 4$ and prime power q, we have $f(n,q)\geq(q^{n+1}-1)(q^n-1)\cdot\frac{q^4+2q^2+q+1}{(q^5-1)(q^4-1)}$. In particular, estimating the rational function from below for $n\geq 5$ gives*

$$q^{2n-4}+2q^{2n-6}+q^{2n-7}+2q^{2n-8}\leq f(n,q).$$

Our contribution is an improvement on the upper bounds according to the result below, which roughly matches the bound above.

Theorem 1.8. *If $n\geq 2$ is fixed and q is a variable prime power, then*

$$f(n,q)\leq q^{2n-4}+2q^{2n-6}+2q^{2n-7}+3q^{2n-8}+3q^{2n-9}+3q^{2n-10}+3q^{2n-11}+O(q^{2n-12}).$$

Our recursive construction for $f(n,q)$ (see Proposition 2.8) uses a parameter k. The construction in the cases $k\in\{0,1\}$ specialises to that of Eisfeld's and Metsch's [11, 22], which obtains bounds on $f(n,q)$ in dimensions $n=4$ and $n=5$ relying on the case of $n=3$ when we know the exact result via Theorem 1.2. We show that one can improve upon this if we start off the recursion from dimension $n=3$, and crucially, instead of augmenting the dimension by one, we use our recursion with a dimension step depending on n, which enables us to block the planes more efficiently.

Related notions For a better overview, we introduce the concepts of q-covering designs, q-Turán designs, q-Steiner designs and subspace designs in general and their relations to each other. These can be investigated as structures on the linear space, but also as structures on the projective space. Observe that there is a difference of 1 between the dimension of a subspace of the linear space and the dimension of the corresponding subspace in the projective geometry. Indeed, (projective) $(k-1)$-subspaces of $\mathrm{PG}(n-1, q)$ correspond to k-subspaces of $\mathbb{F}_{q^n}$.

Definition 1.9. Let $r \leq k \leq n$.

- A q**-covering design** $C_q[n, k, r]$ is a set S of k-spaces of an n-dimensional vector space, such that each r-space of the vector space **is contained in at least one element of** S. Let $\mathcal{C}_q(n, k, r)$ denote the minimum number of k-spaces in a q-covering design $C_q(n, k, r)$.
- A q**-Turán design** $T_q[n, k, r]$ is a collection S of r-spaces of an n-dimensional vector space, such that each k-space of the vector space **contains at least one element of** S. Let $\mathcal{T}_q(n, k, r)$ denote the minimum number of r-spaces in a q-Turán design $T_q(n, k, r)$.
- A q**-Steiner design** $S_q[n, k, r]$ is a collection S of k-spaces of an n-dimensional vector space, such that each r-space of the vector space **is contained in exactly one element of** S.
- Finally, an $(n, k, r, \lambda)_q$ **subspace design** is a collection S of k-spaces of an n-dimensional vector space, such that each r-space of the vector space **is contained in exactly** λ **elements of** S.

We recall the following well-known observation.

Proposition 1.10 (Duality between coverings and Turán designs)**.** *An* $(n, k, r, \lambda)_q$ *subspace design is a q-Steiner design if* $\lambda = 1$*. Since the dual structure of a q–Turán design* $T_q(n, k, r)$ *is a q–covering design* $C_q(n, n-r, n-k)$*, we have* $\mathcal{T}_q(n, k, r) = \mathcal{C}_q(n, n-r, n-k)$*. Hence, a q-Steiner design is a q-covering design which corresponds also to a q-Turán design with appropriate parameters.*

Using these notations, in order to study minimal-size blocking sets $\mathcal{B}$ of lines of $\mathrm{PG}(n-1, q)$ such that every plane of $\mathrm{PG}(n-1, q)$ contains at least one element of $\mathcal{B}$, we wish to improve bounds on $\mathcal{T}_q(n, 3, 2) = \mathcal{C}_q(n, n-2, n-3)$.

2 Constructions

Our aim is to construct (s, t)-blocking sets recursively by passing to the subspace and the quotient spaces.

Definition 2.1 (Grassmannian)**.** Let X be a projective space over a field $\mathbb{F}$. For $-1 \leq d \leq \dim(X)$, let

$$\mathcal{G}_d(X) := \{Y \subseteq X : Y \text{ is a } d\text{-dimensional projective space (over } \mathbb{F})\},$$

the *Grassmannian* of projective subspaces of dimension d. In particular, $\emptyset \subseteq X$ has dimension -1. Write $\mathcal{G}_d(n, q)$ as a shorthand for $\mathcal{G}_d(\mathrm{PG}(n, q))$.

Definition 2.2 (Quotient space)**.** For $K \in \mathcal{G}_k(X)$, the *quotient space* X/K is a projective space of dimension $\dim(X/K) = \dim(X) - k - 1$ given by

$$\mathcal{G}_r(X/K) := \{Y \in \mathcal{G}_{r+k+1}(X) : K \subseteq Y\}$$

for any $-1 \leq r \leq \dim(X/K)$. Note that $\mathcal{G}_{-1}(X/K) = \{K\}$.

2.1 Recursive construction for general (s, t)

Definition 2.3 (Partial blocking sets)**.** For $t \leq s \leq \dim(X)$, we say $\mathcal{B} \subseteq \mathcal{G}_t(X)$ is a *blocking set* for $\mathcal{S} \subseteq \mathcal{G}_s(X)$ if for every $S \in \mathcal{S}$ there exists $T \in \mathcal{B}$ such that $T \subseteq S$. In this case, we also say $\mathcal{B}$ *blocks* $\mathcal{S}$.

The next statement is fundamental in the paper on which all constructions are built. Here we construct partial blocking sets from those of a subspace and the corresponding quotient space. The main idea is to pass to the underlying vector space and consider the split exact sequence $0 \to U \to V \to V/U \to 0$ and use its compatibility with taking subspace intersections and linear spans.

Lemma 2.4 (Recursive construction of partial blocking sets)**.** *Let X be an n-dimensional projective space, $K \in \mathcal{G}_k(X)$, and assume the integers s_1, s_2, t_1, t_2 satisfy $-1 \leq t_1 \leq s_1 \leq \dim(K)$ and $-1 \leq t_2 \leq s_2 \leq \dim(X/K)$. Write $t := t_1 + t_2 + 1$ and $s := s_1 + s_2 + 1$.*

If $\mathcal{B}_K \subseteq \mathcal{G}_{t_1}(K)$ blocks $\mathcal{S}_K \subseteq \mathcal{G}_{s_1}(K)$ and $\mathcal{B}_{X/K} \subseteq \mathcal{G}_{t_2}(X/K)$ blocks $\mathcal{S}_{X/K} \subseteq \mathcal{G}_{s_2}(X/K)$, then $\mathcal{B} \subseteq \mathcal{G}_t(X)$ blocks $\mathcal{S} \subseteq \mathcal{G}_s(X)$ where

$$\mathcal{B} := \left\{T \in \mathcal{G}_\bullet(X) : K \cap T \in \mathcal{B}_K, \langle K, T\rangle \in \mathcal{B}_{X/K}\right\},$$
$$\mathcal{S} := \left\{S \in \mathcal{G}_\bullet(X) : K \cap S \in \mathcal{S}_K, \langle K, S\rangle \in \mathcal{S}_{X/K}\right\}.$$

Taking the union of partial blocking sets given by Lemma 2.4 over $-1 \leq s_1 \leq \dim(K)$, we can obtain an (s, t)-blocking set in an n-dimensional space from blocking sets corresponding to smaller parameters. The next statement is one of the many ways to achieve this.

Theorem 2.5 (An explicit basic recursive construction)**.** *Let $-1 \leq t \leq s \leq n$ and $-1 \leq k \leq n-s-1$. Let X be an n-dimensional projective space and $K \in \mathcal{G}_k(X)$. For $-1 \leq d \leq \min\{k, t\}$, let $\mathcal{B}_d$ be an $(s-d-1, t-d-1)$-blocking set in X/K. Then the disjoint union*

$$\mathcal{B} := \bigcup_{d=-1}^{\min\{k,t\}} \{T \in \mathcal{G}_t(X) : \dim(K \cap T) = d, \langle K, T\rangle \in \mathcal{B}_d\}$$

in an (s, t)-blocking set in X.

2.2 Improved recursion for $(s, t) = (2, 1)$

In the case $(s, t) = (2, 1)$, Lemma 2.4 can be applied in a more complex way to improve the construction of Theorem 2.5 as follows. Let $S \in \mathcal{G}_2(X)$ be an arbitrary plane we want to block, and set $d := \dim(K \cap S)$. We treat the $d = -1, 2$ cases as in Theorem 2.5. However, for $d = 0$, we carefully modify the partial blocking sets of Lemma 2.4 so that they also handle the $d = 1$ case. The idea is to pass to the quotient space X/P for *every* $P \in K$ (instead of the usual single space X/K) and choose a suitable blocking set in every quotient *depending* on P. This extra P-dependency gives us enough freedom to have suitable control over the lines intersecting K in a single point which ultimately gives the next construction.

Proposition 2.6 (Improved recursion for $(s,t) = (2,1)$)**.** *Let X be an n-dimensional projective space and $K \in \mathcal{G}_k(X)$ for $-1 \leq k \leq n-3$. Let $\mathcal{B}_K$ be a $(2,1)$-blocking set in K, $\mathcal{B}_{X/K}$ be a $(2,1)$-blocking set in X/K. Assume that $\mathcal{B}_{X/P}$ is a $(1,0)$-blocking set in X/P for every $P \in \mathcal{G}_0(K)$ such that for every $L \in \mathcal{G}_1(K) \setminus \mathcal{B}_K$, the set $\bigcup\{\mathcal{B}_{X/P} \setminus \mathcal{G}_0(K/P) : P \in \mathcal{G}_0(L)\}$ blocks $\{S \in \mathcal{G}_2(X) : K \cap S \in \mathcal{G}_0(L)\}$. Then the disjoint union*

$$\mathcal{B} := \{T \in \mathcal{G}_1(X) : \langle K, T\rangle \in \mathcal{B}_{X/K}\} \cup \Big(\bigcup_{P \in \mathcal{G}_0(K)} (\mathcal{B}_{X/P} \setminus \mathcal{G}_0(K/P))\Big) \cup \mathcal{B}_K$$

is a $(2,1)$-blocking set in X.

Next, we apply this to the following setup. Pick a polarity ϕ of X such that $K \subseteq \phi(K)$ (i.e. the corresponding symmetric bilinear form on the underlying vector space vanishes on the subspace corresponding to K). Then we set $\mathcal{B}_{X/P} := \mathcal{G}_0(H_P)$ (a minimal $(1,0)$-blocking set in X/P), where the hyperplane $H_P \in \mathcal{G}_{n-2}(X/P)$ is given by $\phi(P) \in \mathcal{G}_{n-1}(X)$. With this choice, Proposition 2.6 produces a $(2,1)$-blocking set $\mathcal{B}$ in X from $(2,1)$-blocking sets $\mathcal{B}_K$ in K and $\mathcal{B}_{X/K}$ in X/K. The fact that $\dim(K), \dim(X/K) < \dim(X)$ (for $K \notin \{\emptyset, X\}$ enables a recursive construction of $(2,1)$-blocking sets in projective spaces of arbitrary dimension.

Remark 2.7. Over finite fields, the $k = 0$, $n = 4$ case of this construction is essentially the same as in [11] where its minimality is also proved. The $k = 1$, $n = 5$ case is the same as [22, Theorem 1.2], and Metsch conjectures its minimality. Compare Theorem 1.3 and Proposition 2.8.

This recursive construction above over $\mathbb{F}_q$ gives the following recursive upper bound for $f(n,q)$.

Proposition 2.8 (Recursive upper bound)**.** *For any integer $n \geq -1$ and $-1 \leq k \leq \frac{n-1}{2}$, we have*

$$f(n,q) \leq q^{2k+2} f(n-k-1,q) + f(k,q) + \sum_{i=0}^{k} q^i \sum_{j=k}^{n-2} q^j$$

For a fixed prime power q, we start from the known exact values for $f(3,q)$ and $f(4,q)$ (see Theorem 1.5), and use Proposition 2.8 to find upper bounds for $f(n,q)$ recursively for each $n \geq 5$. Iteratively, we would like to choose the value $0 \leq k \leq \frac{n-1}{2}$ giving the best upper bound. A technical analysis gives the following optimal choices.

Proposition 2.9. *In the recursion above, the best bound for $f(n,w)$ is obtained if at every step, we choose the unique $0 \leq k \leq \frac{n-1}{2}$ such that $n-k$ is a power of 2.*

Computing the recursion from Proposition 2.8 with the optimal choices from Proposition 2.9, we obtain Theorem 1.8, the main statement of the paper.

3 Concluding remarks and open problems

Definition 3.1 (Density of line sets)**.** Define the *density* by $\varrho(n,q) := f(n,q)/|\mathcal{G}_1(n,q)|$.

For every q, the function $n \mapsto \varrho(n,q)$ is increasing by Theorem 1.6 and is bounded by Theorem 1.8. More concretely, Theorem 1.8 yields

$$\frac{1}{q^2+q} \leq \varrho(4,q) \leq \lim_{n\to\infty} \varrho(n,q) \leq \frac{1}{q^2+q-2q^{-1}-2q^{-2}+4q^{-3}}.$$

We pose the following open problem.

Problem 3.2. *For every prime power q, determine* $\lim_{n\to\infty} \varrho(n, q)$.

Throughout the paper, our main focus was to give bounds on $(2, 1)$-blocking sets. However, the main idea of the recursive refinement presents an approach that can improve general cases as well, albeit it would be more involved. This is the reason we did not want to elaborate on it. Many statements of the paper hold in the general case which could be the starting point of a more general argument, c.f. Lemma 2.4, especially for $t = s - 1$.

The paper discussed a new bound on q-covering and q-Turán designs, which raised increasing interest due to their connections with constant-dimension codes. In particular, a q-Steiner design is an optimal constant-dimension code. Metsch [24] conjectured that no such designs $S_q[n, k, r]$ exist, provided that $k > r > 1$. However, this turned out to be false, as Braun, Etzion, Östergård, Vardy, and Wassermann showed the existence of $S_2[13, 3, 2]$ 2-Steiner designs [5].

Very recently, Keevash, Sah and Sawhney proved the existence of subspace designs [20] with any given parameters, provided that the dimension of the underlying space is sufficiently large in terms of the other parameters of the design and satisfies the obvious necessary divisibility conditions, which settled the corresponding open problem from the 1970s. In particular, for $s > t \geq 1$ fixed integers and q fixed prime power, they showed that if n attains a certain threshold, then once $\begin{bmatrix} k-i \\ r-i \end{bmatrix}_q \mid \begin{bmatrix} n-i \\ r-i \end{bmatrix}_q$ holds for all $0 \leq i \leq r - 1$, a subspace design $S_q[n, k, r]$ exists. However, this cannot be applied in our case, since from the dual of the $(2, 1)$-blocking sets in $\mathrm{PG}(n, q)$, the corresponding parameter set of $\mathcal{C}_q[n+1, n-1, n-2]$, namely $k = n - 1$, $r = n - 2$ will not be independent of n.

We mention the connection point between maximum rank distance codes (MRD codes) and covering designs. This has been discussed by Pavese [25] who used lifted MRD codes to improve bounds on $\mathcal{C}_q[2n, 4, 3]$ and on $\mathcal{C}_q[3n+\ 8, 4, 2]$.

References

[1] André, J. (1954), Über nicht-Desarguessche Ebenen mit transitiver Translationsgruppe, Mathematische Zeitschrift, 60, 156–186.

[2] Beutelspacher, A., Ueberberg, J. (1991). A Characteristic Property of Geometric t-Spreads in Finite Projective Spaces. Europ. J. Comb, **12**(4), 277-281.

[3] Blokhuis, A., Sziklai, P., Szőnyi, T. (2011). Blocking sets in projective spaces. Current research topics in Galois geometry, 61-84.

[4] Bose, R. C. Burton, R. C. (1966). A characterization of flat spaces in a finite geometry and the uniqueness of the Hamming and the MacDonald codes, J. Comb. Th. **1** 96-104.

[5] Braun, M., Etzion, T., Östergård, P. R., Vardy, A., Wassermann, A. (2016, January). Existence of q-analogs of Steiner systems. In Forum of Mathematics, Pi (Vol. 4, p. e7). Cambridge University Press.

[6] Cameron, P. J. (1992). Projective and Polar Spaces. University of London, Queen Mary and Westfield College.

[7] Chaudhury, A. Patkós, B. (2010). Shadows and Intersections in Vector Spaces, Journal of Combin. Theory, Ser. A **117** 1095-1106.

[8] Cossidente, A., Pavese, F. (2016). A Note on q-Covering Designs in PG$(5, q)$. Journal of Combinatorial Designs, 24(8), 383-388.

[9] Csajbók, B., Marino, G., Polverino, O., Zullo, F. (2017). Maximum scattered linear sets and MRD-codes. Journal of Algebraic Combinatorics, 46, 517-531.

[10] Dentice, E. F., Zanella, C. (2009). Bose–Burton type theorems for finite Grassmannians. Discrete Mathematics, 309(2), 363-370.

[11] Eisfeld, J., Metsch, K. (1997). Blocking s-dimensional subspaces by lines in PG$(2s, q)$, Combinatorica, 17(2), 151-162.

[12] Etzion, T. (2013). Problems on q-analogs in coding theory. arXiv preprint arXiv:1305.6126.

[13] Etzion, T. (2014). Covering of subspaces by subspaces. Designs, codes and cryptography, 72(2), 405-421.

[14] Etzion, T., Kurz, S., Otal, K., Özbudak, F. (2020). Subspace packings: constructions and bounds. Designs, Codes and Cryptography, 88(9), 1781-1810.

[15] Etzion, T., Vardy, A. (2011). On q-analogs of Steiner systems and covering designs. Advances in Mathematics of Communications, 5(2), 161-176.

[16] Etzion, T., Kurz, S., Otal, K., Özbudak, F. (2020). Subspace packings: constructions and bounds. Designs, Codes and Cryptography, 88(9), 1781-1810.

[17] Gabidulin, E. M. (1985). Theory of codes with maximum rank distance. Problemy peredachi informatsii, 21(1), 3-16.

[18] Hirschfeld, J. W. P. Projective geometries over finite fields, Clarendon Press, Oxford, 1999.

[19] Katona, G., Nemetz, T., Simonovits, M. (1964). On a graph problem of Turán. Mat. Lapok, 15(1–3), 228-238.

[20] Keevash, P., Sah, A., Sawhney, M. (2022). The existence of subspace designs. arXiv preprint arXiv:2212.00870.

[21] Koetter, R., Kschischang, F. R. (2008). Coding for errors and erasures in random network coding. IEEE Transactions on Information theory, 54(8), 3579-3591.

[22] Metsch, K. (2004). Blocking subspaces by lines in PG(n, q), Combinatorica **24** 459-486.

[23] Metsch, K. (2003). Blocking sets in projective spaces and polar spaces. Journal of Geometry, 76(1), 216-232.

[24] Metsch, K. (1999). Bose-Burton type theorems for finite projective, affine and polar spaces, Surveys in combinatorics, (Canterbury), London Math. Soc. Lecture Note Ser., vol. 267, Cambridge Univ. Press, Cambridge, 137-166.

[25] Pavese, F. (2020). On q-covering designs. The Electronic Journal of Combinatorics, P1-35.

[26] Sziklai, P. (2008). On small blocking sets and their linearity. Journal of Combinatorial Theory, Series A, 115(7), 1167-1182.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Resolution of the no-$(k+1)$-in-line problem when k is not too small

(Extended abstract)

Benedek Kovács * Zoltán Lóránt Nagy † Dávid R. Szabó ‡

Abstract

What is the maximum number of points that can be selected from an $n \times n$ square lattice such that no $k+1$ of them are in a line? This has been asked more than 100 years ago for $k = 2$ and it remained wide open ever since. In this paper, we prove the precise answer is kn, provided that $k > C\sqrt{n \log n}$ for an absolute constant C. The proof relies on carefully constructed bi-uniform random bipartite graphs and concentration inequalities.

1 Introduction

A set of points in the plane is said to be in *general position* if no three of the points lie on a common line. Motivated by a problem concerning the placement of chess pieces, Dudeney [4] asked how many points may be placed in an $n \times n$ grid so that the points are in general position. This No-Three-In-Line problem has received considerable attention: for history and background, we refer to the excellent book of Brass, Moser, and Pach [3] and that of Eppstein [5]. Brass, Moser and Pach called this one of the oldest and most extensively studied geometric questions concerning lattice points.

Concerning an upper bound, it is straightforward to see that at most $2n$ points can be placed in general position. For rather small n, several examples have been constructed where the theoretical bound $2n$ can be attained, see e.g. [2, 6]. However, it is still an open problem to determine the answer for general n.

The earliest lower bound is due to Erdős ([20]), and uses the *modular parabola*, consisting of the points $(i, i^2) \pmod p$. If $n = p$ is a prime number, then this yields n points in general position in $\mathbb{Z}^2 \cap [1, n]^2$. If n is not a prime, then taking p to be the largest prime not exceeding n and applying the modular parabola in a $p \times p$ subgrid yields $n - o(n)$ points in general position.

*ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. E-mail: .
Supported by the University Excellence Fund of Eötvös Loránd University
†ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. E-mail: .
Supported by the University Excellence Fund of Eötvös Loránd University
‡ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. E-mail: .
Supported by the University Excellence Fund of Eötvös Loránd University

The best general lower bound for the No-Three-In-Line problem relies on a construction due to Hall, Jackson, Sudbery, and Wild [11]. Their construction is built up from constructions in subgrids where one places points on a hyperbola $xy = t \pmod p$, where p is a prime slightly smaller than $n/2$, and yields $\frac{3}{2}n - o(n)$ points in general position. Interestingly, even what to conjecture on the *asymptotical value* of the answer for the No-Three-In-Line problem is far from clear, with prominent experts believing it could be roughly $1.5n$ via the previous construction (see Green [8]), or it could be $2n$ thus the upper bound can be attained asymptotically in general [3], or somewhere in between [5, 10]. Some related problems were studied recently in [1] and [19].

We are interested in a generalization which might be the most natural one, where a collinear point set of $k+1$ points is forbidden in the grid $[1, n] \times [1, n]$.

Definition 1.1. Let $f_k(n)$ denote the maximum size of a point set chosen from the vertices of the $n \times n$ grid $\mathcal{G} = [1, n] \times [1, n]$ in which at most k points lie on any line.

Note that in a more general context, this problem has been studied before, e.g. by Lefmann [15], but no exact value has been determined up to our best knowledge. Our result shows that the upper bound kn, which is based on a simple pigeonhole argument, is tight when the order of magnitude of k is slightly larger than $\sqrt{n}$.

Theorem 1.2. *For every $C > \frac{5}{2}\sqrt{35} \approx 14.79$, there exists N_C such that whenever $n \geq N_C$ and $k \geq C\sqrt{n \log n}$, we have*

$$f_k(n) = kn.$$

In fact, we may take any $C > 25/2 = 12.5$.
In the following sections we present the main tools and outline the proof of the main result. For the full proof we refer to [14].

2 Preliminaries

2.1 Notation

If $a, b \in \mathbb{Z}$, then the notation $[a, b]$ will refer to the integer interval $\{x \in \mathbb{Z} : a \leq x \leq b\}$. A *grid* is a set $\mathcal{G} = I_1 \times I_2$ where $I_1 = [a_1, b_1]$ and $I_2 = [a_2, b_2]$ for some integers $a_1 \leq b_1$ and $a_2 \leq b_2$. We say that $\mathcal{G}$ is an $n_1 \times n_2$ grid, where $n_j = |I_j| = b_j - a_j + 1$ for $j = 1, 2$. We call $\mathcal{G}$ a *square grid* if $n_1 = n_2$.

We will often identify a subset S of a grid $\mathcal{G}$ with a bipartite graph with vertex classes I_1 and I_2, where edge (i_1, i_2) appears in the graph if and only if $(i_1, i_2) \in S$. An *r-factor* in a bipartite graph is a spanning subgraph where every vertex has degree r. Analogously, an r-factor in a grid $\mathcal{G}$ is a subset of $\mathcal{G}$ containing exactly r points in each row and column. A *t-matching* is a matching of size t, while in a grid it corresponds to t points with no two lying on the same axis-parallel line.

A *secant* ℓ for a grid $\mathcal{G}$ is a line of the Euclidean plane that meets $\mathcal{G}$ in at least two points. We call a line ℓ *generic* if it is not horizontal or vertical. Let $\mathcal{L}(\mathcal{G})$ be the set of all generic secants of $\mathcal{G}$, and for a fixed real $\kappa \geq 0$, let $\mathcal{L}_{>\kappa}(\mathcal{G}) = \{\ell \in \mathcal{L}(\mathcal{G}) : |\ell \cap \mathcal{G}| > \kappa\}$ be defined as the generic secants admitting more than κ grid points.

A set $S \subseteq \mathcal{G}$ is a *no-$(k+1)$-in-line set* if $|S \cap \ell| \leq k$ for every Euclidean line ℓ. A no-$(k+1)$-in-line set $S \subseteq \mathcal{G}$ is said to have *reserve h* for some $h \in \mathbb{N}$, if $|S \cap \ell| \leq k - h$ for every $\ell \in \mathcal{L}(\mathcal{G})$.

For events E_n indexed with a parameter n, we say that E_n occurs *with high probability* (w.h.p.) if the probability of E_n tends to 1 as $n \to \infty$.

2.2 Tools

Concentration results The well-known Chernoff–Hoeffding bound states that the sum of independent $\{0,1\}$-valued random variables is highly concentrated around the expected value.

Theorem 2.1 (Chernoff bound for the binomial distribution). *Let* $\mathrm{Bin}\,(m, \alpha)$ *denote the binomial distribution with* m *trials and success probability* α. *If* $X \sim \mathrm{Bin}\,(m, \alpha)$, *then for any* $\alpha < \beta \leq 1$ *we have*

$$\mathbb{P}(X > \beta m) \leq e^{-D(\beta||\alpha)m},$$

where $D(\cdot||\cdot)$ *is the so-called relative entropy function satisfying* $D(\beta||\alpha) \geq 2(\beta-\alpha)^2$.

We will apply a variant where the condition on independence can be replaced with a more general condition.

Theorem 2.2 (Linial-Luria [17]). *Let* $X_1, \ldots, X_m$ *be indicator random variables,* $0 < \beta < 1$ *and* $t, m \in \mathbb{N}$ *with* $0 < t < \beta m$. *Then*

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i \geq \beta m\right) \leq \frac{1}{\binom{\beta m}{t}} \sum_{|T|=t} \mathbb{P}\left(\bigcap_{i\in T}\{X_i = 1\}\right).$$

In particular, if $\mathbb{P}\left(\bigcap_{i\in T}\{X_i = 1\}\right) \leq K\alpha^t$ *for every* T *of size* $t = \left(\frac{\beta-\alpha}{1-\alpha}\right)m$, *where* $0 < \alpha < \beta$ *and* $K > 0$ *are constants, then*

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i > \beta m\right) \leq K e^{-D(\beta||\alpha)m} \leq K e^{-2(\beta-\alpha)^2 m}.$$

Note that the upper bound $Ke^{-2(\beta-\alpha)^2 m}$ *also clearly holds in the case when* $0 \leq \alpha < 1$ *and* $\beta > \alpha$ *(not excluding* $\beta \geq 1$*).*

Counting regular subgraphs that contain a given matching In [12], Isaev and McKay presented a series of enumeration results based on the calculation of corresponding high-dimensional complex integrals. An illustrative example is the following. Let $N_H(\mathbf{d})$, resp. $\tilde{N}_H(\mathbf{d})$ denote the number of graphs on n vertices, resp. number of bipartite graphs on $n_1+(n-n_1)$ vertices which have vertex degrees $\mathbf{d} = (d_1, \ldots, d_n)$, include H^+ as a subgraph, and are edge-disjoint from another graph H^-, where H denotes the pair (H^+, H^-) of fixed edge-disjoint graphs on the n-vertex underlying set. Then, following the earlier work of Barvinok, Canfield, Gao, Greenhill, Hartigan, Isaev, McKay, Wang, Wormald (see [9, 16, 18] and the references therein), they expressed $N_H(\mathbf{d})$ and $\tilde{N}_H(\mathbf{d})$ as a complex integral and gave their order of magnitude, provided that some moderate assumptions hold on H and on the degree sequence $\mathbf{d}$. We state here a special case of their result on bipartite graphs, when H^+ is a matching and H^- is an empty graph.

Theorem 2.3 (Special case of [12, Theorem 5.8]). *There is an absolute constant* $K \geq 1$ *satisfying the following. Let* $0 \leq r \leq m$ *be integers,* V_1 *and* V_2 *be disjoint sets of cardinality* m, M *be a fixed (not necessarily perfect) matching of the complete bipartite graph* Γ *on the vertex set* $V_1 \cup V_2$. *If* $G \subseteq \Gamma$ *is a uniform random* r*-regular bipartite graph, then*

$$\mathbb{P}(M \subseteq E(G)) \leq K \cdot (r/m)^{|M|}.$$

3 Proof of Theorem 1.2

First we note that the case $k \geq \frac{2}{3}n$ is treated separately, and it can be handled by an explicit construction. Then we may assume that $k \leq \frac{5}{6}n$ holds. We show the existence of a suitable no-$(k+1)$-in-line set S of maximal size in the grid $\mathcal{G} := [1,n]^2$ using a probabilistic construction (Subsection 3.1). Such a set corresponds to a k-regular bipartite graph via assigning two sets of vertices of size n to the set if columns and rows of the grid, and joining a pair of vertices assigned to a row and a column if their intersection point is in S.

To keep the idea transparent, we first prove that the construction produces the desired no-$(k+1)$-in-line set with positive probability only in the case $4 \mid n$ and $10 \mid k$. The constructed point set will have some reserve on the generic secants. This enables us to adjust S slightly by deleting a small r-factor from S, or adding new rows and columns to $\mathcal{G}$ while keeping the property that the resulting no-$(k+1)$-in-line set has maximal size. In this way, we can handle the cases where n and k do not necessarily satisfy the divisibility conditions (Subsection 3.2).

3.1 Case $4 \mid n$ and $10 \mid k$: bi-uniform random construction

The main idea is to choose the point set of size kn from the grid (or the k-regular subgraph of $K_{n,n}$) in a random manner, in such a way that the expected number of the chosen points from ℓ is significantly less than k for every generic line ℓ.

To do so, we design a bi-uniform random graph (Definition 3.1), in which the probability that an individual edge is contained in a graph will be one of two prescribed values, depending on the edge class it belongs to. After this step, we may apply concentration inequalities to show the existence of a suitable point set.

In our approach, the existence of a suitable random point set demands a divisibility condition on the size of the grid. Thus we also show a method which provides suitable sets of kn' points in an $n' \times n'$ grid from previously constructed suitable sets of size kn in an $n \times n$ grid, where n is only slightly smaller than n' and satisfies the required divisibility condition.

First we establish the random subset $S \subseteq \mathcal{G} = [1,n]^2$ when n is a multiple of 4, k is a multiple of 10, and $C\sqrt{n \log n} \leq k \leq \frac{5}{6}n$ holds.

Let us decompose $\mathcal{G}$ into 16 subgrids of equal size $\frac{n}{4} \times \frac{n}{4}$, see Figure 1. Denote these subgrids by

$$\mathcal{G}_{i,j} := \left[(i-1)\frac{n}{4}+1, i\frac{n}{4}\right] \times \left[(j-1)\frac{n}{4}+1, j\frac{n}{4}\right]$$

for $1 \leq i,j \leq 4$. Mark the subgrid $\mathcal{G}_{i,j}$ as *sparse* if $i=j$ or $i+j=5$ (i.e. if it intersects any of the longest diagonals), and mark it as *dense* otherwise. As indicated in Figure 1, we marked eight of the 16 subgrids as sparse and the other eight as dense.

Definition 3.1 (Bi-uniform random construction)**.** In each subgrid $\mathcal{G}_{i,j}$, we define the point set $S_{i,j} \subseteq \mathcal{G}_{i,j}$ as follows.

Let $S_{i,j}$ be an $r_{i,j}$-factor in the subgrid $\mathcal{G}_{i,j}$ chosen uniformly at random, and independently for each pair (i,j), where the entries of the matrix $R := ((r_{i,j})) \in [0, n/4]^{4\times 4}$ are given in (3.1) as follows.

$$r_{i,j} := \begin{cases} 2\frac{k}{10}, & \text{if } \mathcal{G}_{i,j} \text{ is sparse} \\ 3\frac{k}{10}, & \text{if } \mathcal{G}_{i,j} \text{ is dense} \end{cases} \tag{3.1}$$

Our point set is defined as $S := \bigcup_{i,j=0}^{3} S_{i,j} \subseteq \mathcal{G}$.

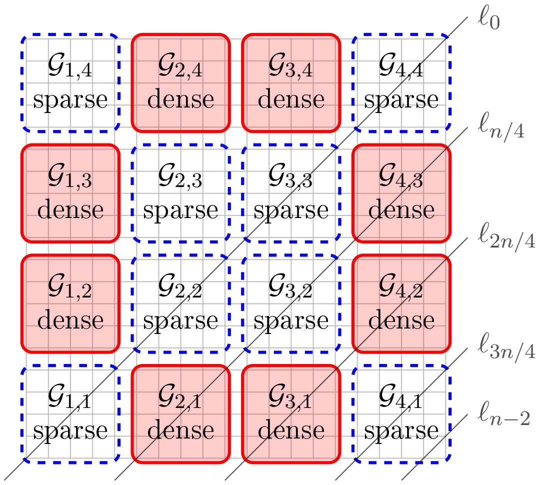


Figure 1: The arrangement of the sparse and dense subgrids and the critical secant lines

The key idea is the following. We will show that this probabilistic construction works with positive probability if $|S \cap \ell| = k$ for every vertical or horizontal line ℓ, and the expected value of $|S \cap \ell|$ is significantly less than k for any other line ℓ. Intuitively, the second condition implies that even with deviations from the expected value, with positive probability $|S \cap \ell| \leq k$ still should hold for every ℓ. To extend this section to the general case (without requiring $4 \mid n$ and $10 \mid k$), we need to construct a no-$(k+1)$-in-line set having small positive reserve.

Lemma 3.2. *Suppose that $\frac{5}{6}n \geq k \geq C\sqrt{n \log n}$ for a suitable constant $C > \frac{5}{2}\sqrt{35}$. Then*

$$\mathbb{P}(\exists \ell \in \mathcal{L} : |\ell \cap S| > k - 15) \leq \frac{1}{\log n}.$$

This lemma is proved using the Linial–Luria variant of the Chernoff bound (Theorem 2.2) by giving an upper bound on $p_{\ell,h} = \mathbb{P}(|\ell \cap S| > k-h) < \frac{1}{n^4 \log n}$ for every $\ell \in \mathcal{L}_{>k}(\mathcal{G})$ and $h > 0$ positive integer. This would imply the statement due to the union bound as $|\mathcal{L}_{>k}(\mathcal{G})| \leq n^4$ follows from the fact that two points in the grid determine at most one secant. To show $\mathbb{P}(|\ell \cap S| > k - h) < \frac{1}{n^4 \log n}$ for every generic secant line ℓ, it is enough to consider the secants of slope ± 1, as other lines contain less points in expectation. For lines with slope ± 1, one can easily bound the distribution of the size of the secant with respect to the subgrids $\mathcal{G}_{i,j}$, and have to overcome slight technical difficulties to handle the sum of these. Here we apply Theorem 2.3.

Corollary 3.3. *For every $C > \frac{5}{2}\sqrt{35}$, there exists n_C such that for every $n \geq n_C$ with $4 \mid n$, and for every k satisfying $C\sqrt{n \log n} \leq k \leq \frac{5}{6}n$ and $10 \mid k$, there exists a no-$(k+1)$-in-line set $S \subseteq [1,n]^2$ of size kn having reserve 15. In particular, $f_k(n) = kn$.*

3.2 General case

To reduce the general case to the one discussed in Subsection 3.1 (namely $4 \mid n$ and $10 \mid k$), we need the following elementary constructions to adjust the values of k and n suitably.

Lemma 3.4 (Adjusting k). *Let $0 \le h \le k \le n$ be integers. Assume $S \subseteq [1,n]^2$ is a no-$(k+1)$-in-line set of size kn having reserve h. Then there is a no-$(k'+1)$-in-line set $S' \subseteq [1,n]^2$ of size $k'n$ having reserve h' for every $0 \le h' \le k'$ with $k - k' = h - h' \ge 0$.*

Lemma 3.5 (Adjusting n). *Let $0 \le h'' \le k' \le n$ be integers with $2 \mid h''$. Assume $S' \subseteq [1,n]^2$ is a no-$(k'+1)$-in-line set of size $k'n$ having reserve h''. Then there exists a no-$(k'+1)$-in-line set $S'' \subseteq [1,n']^2$ of size $k'n'$ for $n' := n + \frac{h''}{2}$.*

4 Concluding remarks and open questions

Our paper completely solves the no-$(k+1)$-in-line problem provided that k is sufficiently large with respect to n, with a rather moderate lower bound on k. It would be very interesting to see whether a similar statement holds when k is in a lower range, $k = \Omega(n^{\varepsilon})$ for every $\varepsilon > 0$. We pose this as an open problem. We believe that the upper bound $f_n(k) \le kn$ is tight in this range as well, that is why we did not elaborate too much on optimizing the constant C for which our proof method works with $k > C\sqrt{n \log n}$.

At last, we point out a different approach to prove that our Bi-uniform construction yields a suitable point set, via the bipartite variant of the celebrated Kim-Vu sandwich conjecture. The original conjecture reads as follows.

Conjecture 4.1 (Kim-Vu sandwich conjecture). If $d \gg \log n$, then for some sequences $p_1 = p_1(n) \sim d/n$ and $p_2 = p_2(n) \sim d/n$ there is a joint distribution of a random d-regular graph $\mathbb{R}(n,d)$ and two binomial random graphs $\mathbb{G}(n,p_1)$ and $\mathbb{G}(n,p_2)$ such that with high probability, we have

$$\mathbb{G}(n,p_1) \subseteq \mathbb{R}(n,d) \subseteq \mathbb{G}(n,p_2).$$

This has been proved by Gao, Isaev and McKay [7] for $d \gg \log^4 n$. Klimošová, Reiher, Ruciński and Šileikis recently proved the following bipartite variant of the Kim-Vu sandwich conjecture. Let $\mathbb{R}(n,n,p)$ denote a random graph chosen uniformly from the set of pn-regular bipartite graphs on $n+n$ vertices, and let $\mathbb{G}(n,n,p)$ denote the binomial bipartite random graph where each edge of $K_{n,n}$ is chosen independently with probability p.

Theorem 4.2 (Sandwich theorem on regular random bipartite graphs, [13]). *If $p \gg \frac{\log n}{n}$ and $1-p \gg \left(\frac{\log n}{n}\right)^{1/4}$, then for some $p' \sim p$, there exists a joint distribution of $\mathbb{G}(n,n,p')$ and $\mathbb{R}(n,n,p)$ such that $\mathbb{G}(n,n,p') \subseteq \mathbb{R}(n,n,p)$ holds with high probability, while for $p \gg \left(\frac{\log^3 n}{n}\right)^{1/4}$, the opposite embedding holds.*

These sandwich theorems can reduce the study of any monotone graph property of the random d-regular (bipartite) graphs in the regime $d \gg \log n$ to study the same property in $\mathbb{G}(n,p)$ (or $\mathbb{G}(n,n,p)$), which can usually be handled much easier.

In order to prove that every line intersects our carefully designed random point set in at most k points, we might apply the above sandwich theorem and bound the probability that a generic line contains more k points. Here we can apply the original variant of Chernoff's bound to estimate the intersection size within a subgrid $\mathcal{G}_{i,j}$. However, the condition $p \gg \left(\frac{\log^3 n}{n}\right)^{1/4}$ means that we could only use it in the range $k \gg (n \log n)^{3/4}$. Note that it is commonly believed that the statement of Theorem 4.2 also holds when $p \gg \log^c n/n$ for some absolute constant c, just like in the generic case of the sandwich theorem of Gao, Isaev and McKay [7].

References

[1] Oswin Aichholzer, David Eppstein, and Eva-Maria Hainzl. Geometric dominating sets - a minimum version of the No-Three-In-Line problem. *Computational Geometry*, 108:Article 101913, 2023.

[2] David Brent Anderson. Update on the no-three-in-line problem. *Journal of Combinatorial Theory, Series A*, 27(3):365–366, 1979.

[3] Peter Brass, William O. J. Moser, and János Pach. *Lattice Point Problems*, pages 417–433. Springer New York, New York, NY, 2005.

[4] Henry E Dudeney. *Amusements in mathematics*, volume 473. Courier Corporation, 1917.

[5] David Eppstein. *Forbidden configurations in discrete geometry.* Cambridge University Press, 2018.

[6] Achim Flammenkamp. Progress in the no-three-in-line problem, II. *Journal of Combinatorial Theory, Series A*, 81(1):108–113, 1998.

[7] Pu Gao, Mikhail Isaev, and Brendan McKay. Kim-Vu's sandwich conjecture is true for $d \geq \log^4 n$. *arXiv preprint arXiv:2011.09449*, 2020.

[8] Ben Green. 100 open problems. Manuscript, available on request to Professor Green.

[9] Catherine Greenhill and Brendan D McKay. Random dense bipartite graphs and directed graphs with specified degrees. *Random Structures & Algorithms*, 35(2):222–249, 2009.

[10] Richard K. Guy and Patrick A. Kelly. The no-three-in-line problem. *Canadian Mathematical Bulletin*, 11:527–531, 1968.

[11] Richard R Hall, Terence H Jackson, Anthony Sudbery, and Ken Wild. Some advances in the no-three-in-line problem. *Journal of Combinatorial Theory, Series A*, 18(3):336–341, 1975.

[12] Mikhail Isaev and Brendan D McKay. Complex martingales and asymptotic enumeration. *Random Structures & Algorithms*, 52(4):617–661, 2018.

[13] Tereza Klimošová, Christian Reiher, Andrzej Ruciński, and Matas Šileikis. Sandwiching biregular random graphs. *Combinatorics, Probability and Computing*, 32(1):1–44, 2023.

[14] Benedek Kovács, Zoltán Lóránt Nagy, and Dávid R Szabó. Settling the no-$(k+1)$-in-line problem when k is not small. *arXiv preprint arXiv:2502.00176*, 2025.

[15] Hanno Lefmann. No l grid-points in spaces of small dimension. In *Algorithmic Aspects in Information and Management*, pages 259–270, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[16] Anita Liebenau and Nick Wormald. Asymptotic enumeration of digraphs and bipartite graphs by degree sequence. *Random Structures & Algorithms*, 62(2):259–286, 2023.

[17] Nathan Linial and Zur Luria. Chernoff's inequality-a very elementary proof. *arXiv preprint arXiv:1403.7739*, 2014.

[18] Brendan D McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the International Congress of Mathematicians 2010*, pages 2489–2501. World Scientific, 2010.

[19] Dániel T Nagy, Zoltán Lóránt Nagy, and Russ Woodroofe. The extensible no-three-in-line problem. *European Journal of Combinatorics*, 114:103796, 2023.

[20] Karl F. Roth. On a problem of Heilbronn. *The Journal of the London Mathematical Society*, 26:198–204, 1951.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Two-colorability of random non-uniform hypergraphs

(Extended abstract)

Jakub Kozik *† Grzegorz Ryn ‡ Toullec Streicher Ronan §

Abstract

Two-colorability of random hypergraphs has been extensively studied in the literature for over half a century. Most of the efforts were devoted to studying uniform hypergraphs (i.e. with all the edges of the same size). Random models played significant role in this research. Our main motivation was to find out how some of the well-known results on random sparse uniform hypergraphs can be translated to the non-uniform case. We have examined both algorithmic and non-constructive bounds for 2-colorability thresholds for certain natural models of random nonuniform hypergraphs. In both of these problems, we have managed to find sufficient conditions for 2-colorability that aggregate information about different sizes of edges. Additionally, in the case of algorithmic lower bounds, we discovered an interesting behaviour when the discrepancy of the sizes of edges can be used to improve the natural generalization of the uniform bounds.

1 Introduction

Hypergraph $\mathcal{H} = (V, E)$ is a pair consisting of a set of *vertices* V and a family of sets of vertices $E \subset \mathcal{P}(V)$, called *(hyper-)edges*. Hypergraph is *k-uniform* if all its edges are of size k. Such hypergraphs are also called *k-graphs*. Just like hypergraphs can be considered as generalizations of ordinary graphs, their coloring problems can be seen as straightforward analogs of graph coloring problems. For a hypergraph $\mathcal{H} = (V, E)$, vertex coloring $c : V \to \mathbb{N}$ is *proper* if every edge contains at least two vertices x, y such that $c(x) \neq c(y)$ (in other words no edge is *monochromatic*). Two colors are enough to make this notion interesting – as observed by Lovász [1] the computational problem of recognizing 2-colorable hypergraphs is NP-complete.

An interesting perspective on hypergraph 2-coloring comes from considering it as a constraint satisfaction problem. Apparent similarity to SAT allows for a translation of a number of

*Theoretical Computer Science, Jagiellonian University, Kraków, Poland. E-mail: jakub.kozik@uj.edu.pl.
†Research partially supported by Polish National Science Center (2023/49/B/ST6/01738).
‡Theoretical Computer Science, Jagiellonian University, Kraków, Poland. E-mail: grzegorz.ryn@student.uj.edu.pl.
§École normale supérieure de Lyon, Lyon, France. E-mail: ronan.toullec_streicher@ens-lyon.fr.

results to the framework of hypergraph coloring. Significant difference reveals itself when one tries to construct a small no-instance. In [2] Erdős and Hajnal asked about the minimum number of edges in a k-graph which is not 2-colorable. Shortly after that Erdős came up with asymptotic lower and upper bounds. In [3], he showed that for large enough k, any k-uniform hypergraph with at most $(1-\varepsilon)2^k \ln 2$ edges is 2-colorable. A year later, in [4] he showed with a probabilistic argument that there exist a non-2-colorable hypergraph with $k^2 \cdot 2^{k+1}$ edges. This is one of the textbook examples of the situation when randomization yields a construction which is significantly smaller than any known deterministic examples. Constraint satisfaction perspective also inspired parallel line of research where 2-colorability is studied for random hypergraphs with fixed size of edges but the number of vertices tending to infinity. Erdős' upper bound literally translate to this framework yielding Theorem 1.1 below. See Alon and Spencer [5] for the adapted proof. By a direct analogy to the well-known random graph model $G(n,p)$, we denote by $H_k(n,p)$ the probabilistic space of random k-graphs on the set V of n vertices, given by including each k-subset of V to the set of edges independently with probability p.

Theorem 1.1 (uniform upper bound [5]).
For $p(n) \geq \frac{\ln 2}{2} \cdot 2^k \cdot n$, random hypergraph $H_k(n,p(n))$ is a.a.s. not 2-colorable.

The evolution of lower bounds has been more convoluted. In [5] Alon and Spencer presented an algorithm that almost surely produces a proper 2-coloring for $H_k(n,p)$ whenever the expected number of edges in that hypergraph does not exceed $c\cdot(2^k/k^2)\cdot n$. The next significant progress came in 2001 with Achlioptas, Kim, Krivelevich, Tetali [6] publishing an algorithm that almost surely 2-colors random k-graphs $H_k(n,p)$ with at most $c \cdot (2^k/k) \cdot n$ edges on average. Finally lower bound for 2-colorability has been improved in 2002 by Achlioptas and Moore [7] to almost asymptotically match the upper one. This argument however is not constructive.

Theorem 1.2 (uniform 2-colorability lower bound [7]).
Let $\varepsilon > 0$ and $k \geq k_0(\varepsilon)$. Then for $p(n)$ such that the expected number of edges in a random hypergraph is $\left(\frac{\ln 2}{2} \cdot \left(2^k - 1\right) - \frac{1+\varepsilon}{2}\right) \cdot n$, random hypergraph $H_k(n,p(n))$ is a.a.s. 2-colorable.

Both, the algorithm from [6] and the nonconstructive lower bound from [7] are going to be our reference points for the developments on non-uniform hypergraphs.

It's worth noting that [6] provides the best currently known algorithmic lower bound for 2-colorability. Moreover, there are reasons to believe that no *significant* improvement is possible. In 2008 Achlioptas and Coja-Oghlan [8] published a paper in which they analyze how the space (which is essentially a subset of the Hamming hypercube) of proper 2-colorings changes as probability of an edge increases. They have showed that for random k-uniform hypergraphs with the expected number of edges between $(1+\varepsilon_k)\frac{2^{k-1}}{k}\ln k\cdot n$ and $(1-\varepsilon_k)2^{k-1}\ln 2\cdot n$ the space of the solutions is *shattered*. This means that the space of proper 2-colorings becomes disconnected into *many* small components that are *very far* apart of each other wrt the Hamming distance. In this and a number of related problems the occurrence of this phenomenon coincides with the point where the best known algorithms break. Moreover this kind of behaviour has been formally proven to be a barrier for certain classes of algorithms [9].

1.1 Non-uniform hypergraphs

A hypergraph is called a k^+-graph if all its edges are of size at least k. Extending the described results to the case of k^+-graphs is often a challenging endeavor. The basic question of Erdős and Hajnal (about the minimum number of edges in a non-2-colorable k-graph) has been reformulated for non-uniform hypergraphs by Erdős and Lovász in [10] as follows: what is the minimum number $f(k)$ for which any k^+-graph with the expected number of monochromatic edges[1] not exceeding $f(k)$ is 2-colorable. Erdős and Lovász conjectured that such defined function is unbounded. It has been confirmed by Beck in 1978 [11], who provided a lower bound of the order $\log^*(k)$. The best currently known lower bound for $f(k)$ from [12] is of the order $\log(k)$. Note that the corresponding bound for k-graphs from [13] is roughly $\sqrt{k/\log(k)}$. While it seems harder to color non-uniform hypergraphs, no one so far managed to exploit non-uniformity in the construction of non-2-colorable hypergraphs. In fact, the best upper bound for $f(k)$ is still $\Theta(k^2)$ which results from the Erdős' upper bound for the uniform case.

Note that $f(k)$ can be equivalently defined as the sum of weights of edges, where an edge of size k gets weight 2^{-k+1}. That gives us a clue about how to translate the statements of the theorems to the non-uniform world.

2 Main results

We are going to work with the following model of a random non-uniform hypergraphs. For a function $\mathcal{M} : \mathbb{N} \to [0,1]$ let random hypergraph $H(n;\mathcal{M})$ have set of vertices $V = \{1,\ldots,n\}$ and let the (random) set of edges be constructed by including every subset of $S \subset V$ independently with probability $\mathcal{M}(|S|)$. The support of $\mathcal{M}$, denoted by *supp*$(\mathcal{M})$, is the set of numbers k for which $\mathcal{M}(k) \neq 0$. By an analogy to k being fixed in the case of uniform hypergraphs, and in order to avoid discussing certain technical issues, we consider only functions $\mathcal{M}$ with finite support. This is implicitly assummed in the statements of the theorems. We are mainly interested in asymptotic results, when the sizes of edges are bounded and the number of vertices tends to infinity. Then, the assymptotic statements about random hypergraphs have to involve working with the sequence of random hypergraphs. For the problems of our interest, the edge probabilities in such sequences are tending to 0 with n. By the common convention of the area we do not explicitely represent the dependence of the probabilities in $\mathcal{M}$ on n. The main theorems are formulated in a way that implicitely defines the probabilities for every n by specifying the expected number of edges of given size.

2.1 Algorithmic 2-coloring of non-uniform hypergraphs

We start by presenting the algorithm of Achlioptas, Kim, Krivelevich, Tetali from [6] (see listing below). For convenience we assume that the number of vertices n is even. Given some partial coloring[2] an edge of size k is called an *l-tail* if $(k-l)$ of its vertices are colored and all of them are assigned the same color.
The main result of [6] can be now restated as follows: for p such that $p \cdot \binom{n}{k} \leq \lambda \cdot \frac{2^k}{k} \cdot n$, AKKT algorithm a.a.s. properly colors $H_k(n,p)$[3].

Observe that this algorithm essentially does not care about any edge till it becomes a 3-tail. Moreover, when an edge becomes a 3-tail, it does not matter how many of its vertices are already colored (i.e. what is the size of that edge). Hence, it is natural to consider its behaviour

[1] in a uniformly random 2-coloring of the vertices
[2] throughout this paper the colorings will be done with RED and BLUE.
[3] For $k \geq 10$, $\lambda \leq \frac{1}{50}$ is enough. For $k \geq 40$, it can be improved to $\lambda \leq \frac{1}{10}$.

Algorithm 1 AKKT algorithm

do $n/2$ times
  if there exists some 3-tail or 2-tail e **then**
    pick any two uncolored vertices $x, y \in e$ and color x to RED and y to BLUE.
  else
    pick any two uncolored vertices x, y and color x to RED and y to BLUE.
  end if
end do

for non-uniform hypergraphs.

An important property of this procedure is that every considered partial coloring is equitable (i.e. the number of RED vertices equals the number of BLUE ones). Moreover, the coloring decisions depend only on the currently observed 3-tails and 2-tails. In particular the potential edges with a larger number of not yet colored vertices are irrelevant at that point and do not even have to be inspected for being an actual edges of the underlying hypergraph. As a result, for every $i \in \{0, \ldots, n/2-1\}$, the distribution of new 3-tails produced by coloring selected vertices in the i-th step does not depend on the previously chosen vertices. Then it is easy to conclude that the new 3-tails are in fact distributed as the edges of a random 3-graph on the remaining vertices with specific probability of including an edge at the i-th step $p(i)$. (In case of k-graphs the value of $p(i)$ depends only on the number of remaining vertices and the original p.) We can therefore abstract from the underlying (random) instance and focus on the probability that the algorithm is successful as a function of the sequence of probabilities $(p(i))_{i=0,\ldots,n/2-1}$. Most of the arguments from the original analysis of this procedure in [6] are valid in such extended framework.

In the uniform case, the algorithm succeeds on graphs with $\lambda \cdot (2^k/k) \cdot n$ edges on average. That suggests assigning weights $(2^k/k)$ to edges of size k. Let $\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_l}$ be a sequence of positive real numbers and $H(n; \mathcal{M})$ be the random hypergraph model with $supp(\mathcal{M}) = \{k_1, \ldots, k_l\}$, such that the expected number of hyperedges of arity k_i is $\lambda_{k_i} \cdot (2^{k_i}/k_i) \cdot n$. The results of [6] imply that if $\lambda_{k_i} \geq 1$ for some i, then the edges of arity k_i are enough to make the algorithm fail. Natural generalization of the result for k-graphs, would be to show that whenever the average weighted sum of edges is bounded by $\lambda \cdot n$, the procedure is likely to succeed (for the same value of λ as used in [6]). Indeed, once the (random) instance has been abstracted to the sequence of probabilities $(p(i))_{i=0,\ldots,n/2-1}$, the tools of [6] can be used in a straightforward manner to deduce such result – if

$$\sum_{k \in supp(\mathcal{M})} \lambda_k < \lambda, \tag{1}$$

then the procedure a.a.s. succeeds. Interestingly, this natural condition in some cases is a serious overkill. E.g. when we consider only two sizes of edges and they differ significantly (e.g. edge sizes k and k' satisfy $k' = k^4$), there are no steps for which both λ_k and $\lambda_{k'}$ contribute significantly to $p(i)$. In such a case the success condition would be closer to the maximum of λ_k and $\lambda_{k'}$ being bounded by λ. Taking into account this kind of behaviour we have shown the following general condition for the random hypergraph to be a.a.s. 2-colorable by AKKT procedure.

Theorem 2.1 (AKKT nonuniform sufficient condition).
If in every phase of AKKT algorithm run on the random hypergraph, the expected number of new 3-tails is bounded by $1-\varepsilon$ for $\varepsilon > 0$, then a.a.s. the algorithm succeeds in finding a proper 2-coloring.

This condition, although not directly related to the specific parametrization (i.e., the expected number of hyperedges of each size), can be used to derive the key properties used in [6] to show that AKKT succeeds. That enable easy adaptation to the non-uniform setting. We present below a more technical version of this condition, by expressing the expected number of new 3-tails using the parametrization of our random hypergraph model. Fourth derivative and exponent 3 in the formula are the consequence of focusing on 3-tails.

Theorem 2.2 (AKKT nonuniform aggregating condition).
Let $H(n;\mathcal{M})$ be a random hypergraph and let $(\lambda_k)_{k\in supp(\mathcal{M})}$ be such that the expected number of edges of each size k in $H(n;\mathcal{M})$ equals $\lambda_k \cdot \frac{2^k}{k} \cdot n$. Let

$$f(x) := \sum_{k\in supp(\mathcal{M})} \frac{\lambda_k}{k} \cdot x^k$$

Then, for any $\varepsilon > 0$ and $\min(supp(\mathcal{M}))$ large enough, the following condition is sufficient for the AKKT procedure to produce a proper 2-coloring almost surely

$$\max_{x\in[0,1]} \left(\frac{16}{3}(1-x)^3 f^{(4)}(x) \right) < 1-\varepsilon,$$

where $f^{(4)}$ is the fourth derivative of f.

It is not immediately clear how to compare the technical condition above with the natural generalization (1). In order to illustrate the difference we constructed a family of the sequences of edge sizes and corresponding lambdas s.t. their sum can be arbitrarily large.

Theorem 2.3 (Unbounded sum of lambdas).
For every $c > 0$, there exists a random hypergraph $H(n;\mathcal{M})$ with corresponding sequence $(\lambda_k)_{k\in supp(\mathcal{M})}$ such that the expected number of edges of each size k in $H(n;\mathcal{M})$ equals $\lambda_k \cdot \frac{2^k}{k} \cdot n$, for which

$$\sum_{k\in supp(\mathcal{M})} \lambda_k \geq c$$

and AKKT procedure succeeds a.a.s.

As we can see, 2-colorable non-uniform random hypergraphs can have much more[4] edges on the average than the uniform ones as long as their sizes are sufficiently spread.

2.2 2-coloring threshold for non-uniform hypergraphs

It is already known, that the 2-colorability threshold for k-graphs is around the values of $p^*(n)$ for which the expected number of edges is $\lambda_k^* \cdot 2^k \cdot n$, for some λ_k^* that is close to $\ln 2/2$

[4] in the sense of the sum of lambdas

but whose precise value has not been determined yet[5]. For $\lambda_k \geq \ln 2/2$ the corresponding random hypergraph is a.a.s. not-2-colorable, and by the results of [7] and later improvements [14, 15], for slightly smaller λ_k it is eventually very likely to be 2-colorable. That justifies parametrization of the average number of edges as $\lambda_k \cdot 2^k \cdot n$ and focusing on how the values of λ_k affect 2-colorability. In a nonuniform case, for a sequence $\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_l}$ of positive real numbers, we are going to work with corresponding random hypergraph $H(n; \mathcal{M})$ in which the expected number of edges of size k_i is $\lambda_{k_i} \cdot 2^{k_i} \cdot n$. Within this setting, the proof of the upper bound for k-graphs generalizes in a straightforward way.

Theorem 2.4 (non-uniform upper bound)**.**
Let $H(n; \mathcal{M})$be a random hypergraph model with corresponding sequence $(\lambda_k)_{k \in supp(\mathcal{M})}$ for which, for every $k \in supp(\mathcal{M})$, the expected number of edges of size k equals $\lambda_k \cdot 2^k \cdot n$. Suppose that $\sum_{k \in supp(\mathcal{M})} \lambda_k > \frac{\ln 2}{2}$, then

$$\lim_{n \to \infty} \mathcal{P}(H(n; \mathcal{M}) \text{ is 2-colorable}) = 0.$$

We immediately observe that, unlike in the case of the algorithmic lower bound, the sizes of edges are irrelevant provided they are compensated by probability. In other words it is really the sum of λ_k (not the maximum) that makes the (random) hypergraph not 2-colorable. The non-uniform structure which we have exploited with AKKT appears to be completely insignificant around the threshold. As explained below, it is not only a feature of this particular argument, since we observe an analogous behavior in the analysis of nonconstructive lower bounds.

The first lower bound for the uniform case of the order $2^k \cdot n$ was obtained in [7] by the second moment method. Later is has been observed (see e.g. [14]) that this proof can be significantly simplified by focusing on equitable colorings. Note that once again, the equitable colorings announce their prominent role. Second moment argument is used to prove that, for certain range of the values of p, the probability that corresponding random k-graph is 2-colorable is bounded away from 0. Then, by the general properties of the thresholds of monotonic properties developed by Friedgut [16], it can be concluded that for every slightly smaller value of p corresponding random hypergraph is a.a.s. 2-colorable. The proof that 2-colorability of hypergraphs exhibits a sharp (non-uniform) threshold can be found in [17].

With a bit of technical effort we managed to show that non-uniform hypergraphs are 2-colorable with positive probability just before the non-2-colorability bound of Theorem 2.4. Our proof is also based on the second moment method, and just like [14] focuses on equitable 2-colorings. The obtained bound for the sum of λ_k is weaker than the known corresponding bounds for the uniform case. However, the ideas that allowed for the improvements of the uniform case are likely to be adapted to the non-uniform framework. The technical cost of doing so seems to be considerable.

Theorem 2.5 (non-uniform 2-colorability threshold)**.**
Let $\varepsilon > 0$ and $H(n; \mathcal{M})$ be a random hypergraph model with large enough $\min(supp(\mathcal{M}))$ and corresponding sequence λ_k for $k \in supp(\mathcal{M})$ such that the expected number of edges of size k equals $\lambda_k \cdot 2^k \cdot n$.

[5] To be precise it is also possible that the threshold is nonuniform, i.e. $\lambda_k^* = \lambda_k^*(n)$ and it does not converge with $n \to \infty$.

Suppose that $\sum_{k \in supp(\mathcal{M})} \frac{\lambda_k}{1-2^{1-k}} < \frac{\ln 2}{2}(1-\varepsilon)$*, then*

$$\liminf_{n\to\infty} \mathcal{P}(H(n;\mathcal{M}) \text{ is 2-colorable}) > 0.$$

The thing stopping us from finishing the proof and showing that such hypergraphs are 2-colorable almost surely is the lack of tools for handling sharp thresholds in the non-uniform case. We explored more recent generalizations of these methods from [18] by Hatami and Molloy, and tried to embed our non-uniform model in their (much more general than the original) random CSP model. We did not succeed and eventually convinced ourselves that stronger tools are necessary. We are currently working on the multi-dimensional analogue of the Friedgut's results from [16].

3 Conclusion

A number of results concerning the limits of 2-colorability of large (and sparse) random k-graphs can be carried out to non-uniform framework. While the methods for non-constructive upper and lower bounds turns out to be rather oblivious to the distribution of edge sizes, specific configurations can be exploited in the proofs of the algorithmic (but weaker) lower bounds. It is particularly interesting to study the behavior of such coloring procedures from the point of view of algorithmic barriers (as defined in [8]).

It is expected that 2-colorability for non-uniform models exhibit (possibly non-uniform) sharp-threshold behavior. That would allow for strengthening our non-constructive lower bound from Theorem 2.5 and show that for slightly smaller sums of lambdas, the limit is not merely positive but equals 1. Although nonuniformity do not seem to introduce any essentially new phenomena when the sharp-thresholds are considered, the common tools of this area need to be sharpened by a multi-dimensional analysis before they can be applied to nonuniform hypergraphs.

[1] L. Lovász. Coverings and colorings of hypergraphs. *Proc. 4th Southeastern Conference of Combinatorics, Graph Theory, and Computing*, pages 3–12, 1973.

[2] Paul Erdős and András Hajnal. On a property of families of sets. *Acta Mathematica Academiae Scientiarum Hungaricae*, 12:87–123, 1961.

[3] Paul Erdős. On a combinatorial problem. *Nordisk Matematisk Tidskrift*, 11:5–10, 40, 1963.

[4] Paul Erdős. On a combinatorial problem. II. *Acta Mathematica Academiae Scientiarum Hungaricae*, 15:445–447, 1964.

[5] Noga Alon and Joel H. Spencer. A note on coloring random k-sets.

[6] Dimitris Achlioptas, Jeong Han Kim, Michael Krivelevich, and Prasad Tetali. Two-coloring random hypergraphs. *Random Structures Algorithms*, 20(2):249–259, 2002.

[7] Dimitris Achlioptas and Cristopher Moore. On the 2-colorability of random hypergraphs. In *Randomization and approximation techniques in computer science*, volume 2483 of *Lecture Notes in Comput. Sci.*, pages 78–90. Springer, Berlin, 2002.

[8] Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '08, pages 793–802, Washington, DC, USA, 2008. IEEE Computer Society.

[9] David Gamarnik. The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences*, 118(41):e2108492118, 2021.

[10] Paul Erdős and László Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II*, volume 10 of *Colloquia Mathematica Societatis János Bolyai*, pages 609–627. North-Holland, Amsterdam, 1975.

[11] József Beck. On 3-chromatic hypergraphs. *Discrete Mathematics*, 24(2):127–137, 1978.

[12] Lech Duraj, Grzegorz Gutowski, and Jakub Kozik. A Note on Two-Colorability of Nonuniform Hypergraphs. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 46:1–46:13, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[13] Jaikumar Radhakrishnan and Aravind Srinivasan. Improved bounds and algorithms for hypergraph 2-coloring. *Random Structures & Algorithms*, 16(1):4–32, 2000.

[14] Martin Dyer, Alan Frieze, and Catherine Greenhill. On the chromatic number of a random hypergraph. *J. Combin. Theory Ser. B*, 113:68–122, 2015.

[15] Peter Ayre, Amin Coja-Oghlan, and Catherine Greenhill. Hypergraph coloring up to condensation. *Random Structures & Algorithms*, 54(4):615–652, 2019.

[16] Ehud Friedgut and Jean Bourgain. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American Mathematical Society*, 12(4):1017–1054, 1999.

[17] Ehud Friedgut. Hunting for sharp thresholds. *Random Struct. Algorithms*, 26(1–2):37–51, January 2005.

[18] Michael Molloy Hamed Hatami. Sharp thresholds for constraint satisfaction problems and homomorphisms. *Random Structures & Algorithms*, 33(3):310–332, 2008.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Uniform Turán density—palette classification*

(Extended abstract)

Daniel Kráľ† Filip Kučerák‡ Ander Lamaison§ Gábor Tardos¶

Abstract

In the 1980s, Erdős and Sós initiated the study of Turán hypergraph problems with a uniformity condition on the distribution of edges, and specifically asked to determine the uniform Turán densities of the hypergraphs $K_4^{(3)-}$ and $K_4^{(3)}$. In these two and many additional cases, the tight lower bounds are provided by a so-called palette construction.

Lamaison [arXiv:2408.09643] showed that the uniform Turán density of a 3-uniform hypergraph H is equal to the supremum of the densities of palettes that H is not colorable with. We give a necessary and sufficient condition, which is easy to verify, on the existence of a 3-uniform hypergraph colorable by a set of palettes and not colorable by another given set of palettes, and demonstrate how the condition can be used to find a 3-uniform hypergraph with a given value of uniform Turán density.

1 Introduction

Turán problems are one of the most fundamental problems in extremal combinatorics; they ask to determine the minimum density, which is called *Turán density*, that guarantees the existence of a given substructure (the name originates from the classical theorem of Turán [27] concerning complete graphs). While Turán densities of graphs are well-understood due to work

*This research started during the first author's stay at HUN-REN Alfréd Rényi Institute of Mathematics in the spring 2024, which was supported by the ERC Synergy Grant No. 810115 (DYNASNET). The first and fourth authors also acknowledge support by the National Science Foundation under Grant No. DMS-1928930 while the authors were in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Spring 2025 semester.

†Institute of Mathematics, Leipzig University, Augustusplatz 10, 04109 Leipzig, and Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany. E-mail: daniel.kral@uni-leipzig.de. Previous affiliation: Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic.

‡Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic. E-mail: kucerak@mail.muni.cz.

§Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. This author was supported by IBS-R029-C4. E-mail: ander@ibs.re.kr.

¶HUN-REN Rényi Institute, Reáltanoda utca 13–15, Budapest, Hungary. This author was also supported by the ERC Advanced Grants "ERMiD" and "GeoScape" and by the National Research, Development and Innovation Office project K-132696. E-mail: tardos@renyi.hu.

of Erdős and Stone [7], the same is not the case for hypergraphs, where even some of the most basic problems have stayed unchallenged for many decades. For example, the Tetrahedron Problem, which asks for the Turán density of the complete 3-uniform hypergraph $K_4^{(3)}$, has resisted attempts for its resolution since its formulation 80 years ago [27]. We remark that determining the Turán density of all complete k-uniform hypergraphs for $k \geq 3$ is a \$1 000 problem of Erdős [4] (and Erdős offered \$500 for determining the Turán density of any single non-trivial complete k-uniform hypergraph). We refer to [3, 8, 18] for additional results and to the surveys by Keevash [12] and Sidorenko [26] for a more comprehensive treatment of the matter.

Almost all known and conjectured extremal constructions for Turán problems in the (hyper)graph setting have large independent sets, i.e., the edges are spread in a highly non-uniform way in the host hypergraph. This led Erdős and Sós [5, 6] to studying Turán densities with the additional condition on the uniform distribution of the edges of the host hypergraph. Formally, we say that an n-vertex k-uniform hypergraph is (d, ε)-uniformly dense if every subset of $n' \geq \varepsilon n$ vertices spans at least $d\binom{n'}{k}$ edges, and the *uniform Turán density* of a k-uniform hypergraph H is the infimum over all d such that for every $\varepsilon > 0$, there exists n_ε such that every k-uniform hypergraph with $n \geq n_\varepsilon$ vertices that is (d, ε)-uniformly dense contains H as a subhypergraph. We remark that the notion is trivial for graphs as the uniform Turán density of every graph is equal to zero.

Until about a decade ago, there was little progress concerning the uniform Turán densities since Erdős and Sós introduced the notion and asked about determining the uniform Turán densities of the complete 3-uniform hypergraph $K_4^{(3)}$ and the 3-uniform hypergraph $K_4^{(3)-}$, which is $K_4^{(3)}$ with an edge removed. The latter was resolved by Glebov, Volec and the first author [11] using the flag algebra method of Razborov [17], and a direct combinatorial argument using the hypergraph regularity method was given by Reiher, Rödl and Schacht [23].

The use of hypergraph regularity method revolutionized the area. Reiher, Rödl and Schacht [21] classified 3-uniform hypergraphs with uniform Turán density equal to 0, and the uniform Turán density of every 3-uniform hypergraph is at least 1/27 unless it is equal to 0; the value of 1/27 was shown to be tight in [10]. Families of 3-uniform hypergraphs with uniform Turán density equal to 1/27, 4/27, 1/4 and 8/27 were identified in [1, 2, 9, 16], and very recently the uniform Turán densities of all generalized stars, largely extending the result concerning $K_4^{(3)-}$, have been determined by Wu and the third author [15]. For further exposition including results on stronger notions of uniform density such as e.g. [20, 22, 24], we refer the reader to the survey by Reiher [19] on the topic.

1.1 Palette constructions

In our further exposition, we restrict our attention to 3-uniform hypergraphs only. All lower bounds on the uniform Turán density come from so-called *palette constructions*, generalizing the construction of Rödl [25] in the tetrahedron case. A *palette* $\mathcal{P}$ is a pair (C, T) with $T \subseteq C^3$; the elements of C are *colors* and the elements of T are (feasible) *triples*. The *density* $d(\mathcal{P})$ of a palette $\mathcal{P} = (C, T)$ is $|T|/|C|^3$, and the *Lagrangian* $L(\mathcal{P})$ is the maximum weighted density, i.e.

$$\max_{p:C\to[0,1],\sum p=1} \sum_{(x,y,z)\in T} p(x)p(y)p(z),$$

where the maximum is taken over all probability distributions p on C. Consider the probability distribution p maximizing the expression and the following random n-vertex hypergraph H_n with vertices $v_1, \ldots, v_n$: color each pair (v_i, v_j), $1 \le i < j \le n$, with a color $c \in C$ with probability $p(c)$, and include $\{v_i, v_j, v_k\}$, $1 \le i < j < k \le n$, as an edge if $(x, y, z) \in T$ where x is the color of (v_i, v_j), y that of (v_i, v_k) and z that of (v_j, v_k). Observe that H_n is $(L(\mathcal{P}) - \varepsilon, \varepsilon)$-uniformly dense with positive probability when n is sufficiently large (and ε is fixed).

We say that a hypergraph H is $\mathcal{P}$-colorable if there exists an ordering $v_1, \ldots, v_N$ of its vertices and coloring of their pairs with the colors from C such that every edge $\{v_i, v_j, v_k\}$ of H, $1 \le i < j < k \le N$, satisfies that $(x, y, z) \in T$ where x is the color of (v_i, v_j), y is that of (v_i, v_k) and z is that of (v_j, v_k). Observe that if H is not $\mathcal{P}$-colorable, then the uniform Turán density of H is at least $L(\mathcal{P})$. The empirical evidence strongly suggested that the palette constructions always provide tight lower bounds. This has recently been proven by the third author [14] using a combination of the hypergraph regularity method and probabilistic arguments.

Theorem 1. *The uniform Turán density of any 3-uniform hypergraph H is equal to the supremum of the Lagrangian of a palette $\mathcal{P}$ such that H is not $\mathcal{P}$-colorable.*

We remark that Theorem 1 also holds when Lagrangian is replaced with density.

Theorem 1 presented a breakthrough in regard to the methodology as it replaced complex arguments based on the hypergraph regularity with much simpler arguments concerning colorability by palettes. In particular, it led to determining the uniform Turán densities of generalized stars [15], a result vastly extending the solution of the problem of Erdős-Sós on the uniform Turán density of $K_4^{(3)-}$.

2 Result

Given that the colorability of palettes with certain densities reflects the uniform Turán density of a hypergraph H, we arrive at the following question:

For which palettes $\mathcal{P}_1, \ldots, \mathcal{P}_r$ and $\mathcal{P}'_1, \ldots, \mathcal{P}'_q$ does there exist a hypergraph H that is $\mathcal{P}_i$-colorable for every $i = 1, \ldots, r$ but not $\mathcal{P}'_j$-colorable for any $j = 1, \ldots, q$?

The main motivation for this question (as we demonstrate in Section 3), is obtaining a simple tool for constructing hypergraphs with a given uniform Turán density. Indeed, the case when $r = 1$ was resolved independently of us by King, Piga, Sales and Schülke [13] in their work on feasible uniform Turán densities of families of hypergraphs, where they showed every real that is the Lagrangian of a palette is the uniform Turán density of a finite family of hypergraphs. Our main result (given as Theorem 3) completely answers this question by giving a necessary and sufficient condition, which is easy to verify.

To state our results, we need to introduce additional terminology. A *homomorphism* from a palette $\mathcal{P} = (C, T)$ to a palette $\mathcal{P}' = (C', T')$ is a mapping $f : C \to C'$ such that $(f(x), f(y), f(z)) \in T'$ for every triple $(x, y, z) \in T$. The *inverse* $\mathrm{inv}(\mathcal{P})$ of a palette $\mathcal{P} = (C, T)$ is the palette (C, T') such that $(x, y, z) \in T'$ iff $(z, y, x) \in T$. Observe that if there exists a homomorphism from a palette $\mathcal{P}$ to a palette $\mathcal{P}'$ or to $\mathrm{inv}(\mathcal{P}')$, then every $\mathcal{P}$-colorable hypergraph is also $\mathcal{P}'$-colorable. This simple condition turned out to be the sought condition when $r = 1$.

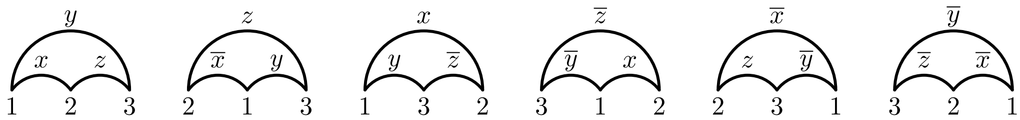


Figure 1: Visualization of the triples included in $T^{(s)}$ for a triple $(x, y, z) \in T$.

Theorem 2. *Let $\mathcal{P}$ and $\mathcal{P}_0$ be two palettes. There exists a hypergraph that is $\mathcal{P}$-colorable but not $\mathcal{P}_0$-colorable if and only if there is no homomorphism from the palette $\mathcal{P}$ to the palette $\mathcal{P}_0$ or to the palette* $\mathrm{inv}(\mathcal{P}_0)$.

The general setting when $r > 1$ is much more complex. We need two additional definitions. The product $\mathcal{P}_1 \times \mathcal{P}_2$ of palettes $\mathcal{P}_1 = (C_1, T_1)$ and $\mathcal{P}_2 = (C_2, T_2)$ is the palette (C, T) such that $C = C_1 \times C_2$ and $((x_1, x_2), (y_1, y_2), (z_1, z_2)) \in T$ iff $(x_1, y_1, z_1) \in T_1$ and $(x_2, y_2, z_2) \in T_2$. The *symmetrization* of a palette $\mathcal{P} = (C, T)$, denoted by $\mathcal{P}^{(s)}$, is the palette $(C^{(s)}, T^{(s)})$ such that

$$\begin{aligned} C^{(s)} &= \{c_1, \dots, c_k, \overline{c_1}, \dots, \overline{c_k}\} \text{ and} \\ T^{(s)} &= \{(x, y, z), (\overline{x}, z, y), (y, x, \overline{z}), (\overline{y}, \overline{z}, x), \\ &\qquad (z, \overline{x}, \overline{y}), (\overline{z}, \overline{y}, \overline{x}) \text{ for every } (x, y, z) \in T\}, \end{aligned}$$

where $c_1, \dots, c_k$ are the colors contained in C. Informally speaking, the palette $\mathcal{P}^{(s)}$ has two twin colors for each color of $\mathcal{P}$, and contains six triples for each triple (x, y, z) of $\mathcal{P}$ that are obtained as follows (see Figure 1): color the edges of the transitive triangle with colors x, y and z in a way that the order given by the edge orientation is consistent with the triple (x, y, z), and then consider all six possible orderings of the vertices and include the triple that corresponds to the coloring with respect to the considered ordering with the colors of the reversed edges replaced with their twin colors.

We are now ready to state our main result.

Theorem 3. *Let $\mathcal{P}_1, \dots, \mathcal{P}_r$ and $\mathcal{P}'_1, \dots, \mathcal{P}'_q$ be $r + q$ palettes. There exists a hypergraph H that is $\mathcal{P}_i$-colorable for every $i \in [r]$ but not $\mathcal{P}'_j$-colorable for any $j \in [q]$ if and only if for every $i \in [r]$ and $j \in [q]$, there is no homomorphism from the palette $\mathcal{P}_i \times \prod_{i' \in [r] \setminus \{i\}} \mathcal{P}^{(s)}_{i'}$ to $\mathcal{P}'_j$ or to* $\mathrm{inv}(\mathcal{P}'_j)$.

The proof of the of Theorem 3 has two steps, which we sketch on a very high level. We first argue using the hypergraph regularity method and suitable tools from Ramsey theory that if every *linearly ordered* hypergraph that is $\mathcal{P}_i$-colorable for all $i \in [r]$ is also $\mathcal{P}'_j$-colorable, then there exists a homomorphism of the form given in the statement. If no such homomorphism exists for any j, we consider the ordered hypergraphs that are $\mathcal{P}_i$-colorable for all $i \in [r]$ but not $\mathcal{P}'_j$-colorable and plant their random permutations in an almost disjoint way in a large host hypergraph. As the planted copies behave independently, we can show that the host hypergraph (without any linear order assumed) is $\mathcal{P}_i$-colorable for every $i \in [r]$ but not $\mathcal{P}'_j$-colorable for any $j \in [q]$ with high probability.

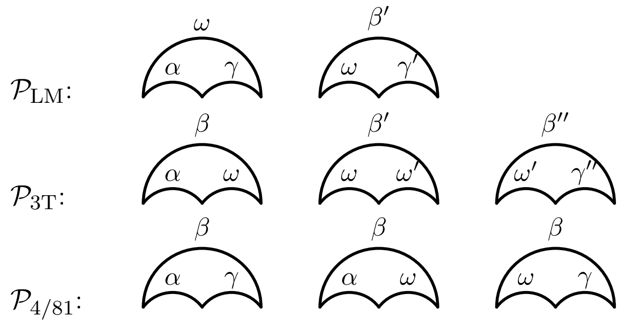


Figure 2: The triples of the palette $\mathcal{P}_{\mathrm{LM}}$ with colors $\alpha, \beta', \gamma, \gamma', \omega$, the palette $\mathcal{P}_{\mathrm{3T}}$ with colors $\alpha, \beta, \beta', \beta'', \gamma'', \omega, \omega'$, and the palette $\mathcal{P}_{4/81}$ with colors $\alpha, \beta, \gamma, \omega$.

3 Example of application

We now show how Theorem 3 can be used to derive the existence of a hypergraph with uniform Turán density equal to 4/81. This is far from coming for free, e.g. the existence of a hypergraph with uniform Turán density equal to 1/27 is the main result of [10]. Consider the palettes $\mathcal{P}_{\mathrm{LM}}$, $\mathcal{P}_{\mathrm{3T}}$ and $\mathcal{P}_{4/81}$ given in Figure 2. It can be shown that for any palette $\mathcal{P}$ with $d(\mathcal{P}) > 4/81$, there is a homomorphism from $\mathcal{P}_{\mathrm{LM}}$ or from $\mathcal{P}_{\mathrm{3T}}$ to $\mathcal{P}$ or to $\mathrm{inv}(\mathcal{P})$. Since neither the palette $\mathcal{P}_{\mathrm{LM}} \times \mathcal{P}_{\mathrm{3T}}^{(s)}$ nor the palette $\mathcal{P}_{\mathrm{3T}} \times \mathcal{P}_{\mathrm{LM}}^{(s)}$ has a homomorphism to the palette $\mathcal{P}_{4/81}$ and the Lagrangian of the palette $\mathcal{P}_{4/81}$ is 4/81, Theorems 1 and 3 imply that there exists a 3-uniform hypergraph with uniform Turán density equal to 4/81.

References

[1] M. Bucić, J. W. Cooper, D. Král', S. Mohr and D. Munhá Correia: *Uniform Turán density of cycles*, Transactions of the AMS **376** (2023), 4765–4809.

[2] A. Chen and B. Schülke: *Beyond the broken tetrahedron*, `preprint arXiv:2211.12747`.

[3] F. Chung and L. Lu: *An upper bound for the Turán number $t_3(n, 4)$*, Journal of Combinatorial Theory Series A **87** (1999), 381–389.

[4] P. Erdős: *On the combinatorial problems which I would most like to see solved*, Combinatorica **1** (1981), 25–42.

[5] P. Erdős: *Problems and results on graphs and hypergraphs: similarities and differences*, in: J. Nešetřil and V. Rödl (eds.), Mathematics of Ramsey theory (1990), 223–233.

[6] P. Erdős and V. T. Sós: *On Ramsey-Turán type theorems for hypergraphs*, Combinatorica **2** (1982), 289–295.

[7] P. Erdős and A. H. Stone: *On the structure of linear graphs*, Bulletin of the American Mathematical Society **52** (1946), 1087–1091.

[8] P. Frankl and Z. Füredi: *An exact result for 3-graphs*, Discrete Mathematics **50** (1984), 323–328.

[9] F. Garbe, D. Iľkovič, D. Kráľ and A. Lamaison: *Hypergraphs with uniform Turán density equal to 8/27* (2024), `preprint arXiv:2407.05829`.

[10] F. Garbe, D. Kráľ and A. Lamaison: *Hypergraphs with minimum positive uniform turán density*, Israel Journal of Mathematics **259** (2024), 701–726.

[11] R. Glebov, D. Kráľ and J. Volec: *A problem of Erdős and Sós on 3-graphs*, Israel J. Math. **211** (2016), 349–366.

[12] P. Keevash: *Hypergraph Turán problems*, in: R. Chapman (ed.), Surveys in Combinatorics 2011, London Mathematical Society Lecture Note Series (2011), 83–140.

[13] D. King, S. Piga, M. Sales and B. Schülke: *On possible uniform turán densities*, `preprint arXiv:2504.21220`.

[14] A. Lamaison: *Palettes determine uniform Turán density*, `preprint arXiv:2408.09643`.

[15] A. Lamaison and Z. Wu: *The uniform Turán density of large stars*, `preprint arXiv:2409.03699`.

[16] H. Li, H. Lin, G. Wang and W. Zhou: *Hypergraphs with a quarter uniform Turán density*, `preprint arXiv:2305.11749`.

[17] A. A. Razborov: *Flag algebras*, J. Symbolic Logic **72** (2007), 1239–1282.

[18] A. A. Razborov: *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discrete Math. **24** (2010), 946–963.

[19] C. Reiher: *Extremal problems in uniformly dense hypergraphs*, European Journal of Combinatorics **88** (2020), 103117.

[20] C. Reiher, V. Rödl and M. Schacht: *Embedding tetrahedra into quasirandom hypergraphs*, Journal of Combinatorial Theory Series B **121** (2016), 229–247.

[21] C. Reiher, V. Rödl and M. Schacht: *Hypergraphs with vanishing Turán density in uniformly dense hypergraphs*, Journal of the London Mathematical Society **97** (2018), 77–97.

[22] C. Reiher, V. Rödl and M. Schacht: *On a generalisation of Mantel's theorem to uniformly dense hypergraphs*, International Mathematics Research Notices **16** (2018), 4899–4941.

[23] C. Reiher, V. Rödl and M. Schacht: *On a Turán problem in weakly quasirandom 3-uniform hypergraphs*, Journal of the European Mathematical Society **20** (2018), 1139–1159.

[24] C. Reiher, V. Rödl and M. Schacht: *Some remarks on* π, in: S. Butler, J. Cooper and G. Hurlbert (eds.), Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham (2018), 214–239.

[25] V. Rödl: *On universality of graphs with uniformly distributed edges*, Discrete Mathematics **59** (1986), 125–134.

[26] A. Sidorenko: *What we know and what we do not know about Turán numbers*, Graphs and Combinatorics **11** (1995), 179–199.

[27] P. Turán: *On an extremal problem in graph theory*, Matematikai és Fizikai Lapok **48** (1941), 436–452.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Every fully graphic degree sequence region is P-stable

(Extended abstract)

Péter L. Erdős* Gábor Lippner † Lajos Soukup‡

Abstract

Let $n > c_1 \geq c_2$ be positive integers, and let Σ be an even integer with $n \cdot c_1 \geq \Sigma \geq n \cdot c_2$. Let $\mathbb{D} = \mathbb{D}(n, \Sigma, c_1, c_2)$ denote the set of all degree sequences $(d_1, \ldots, d_n)$ of length n with sum Σ and with $c_1 \geq d_i \geq c_2$. Furthermore assume that all sequences in $\mathbb{D}$ are graphic (such a $\mathbb{D}$ is called **fully graphic**). Our main result is that slightly perturbing a degree sequence in $\mathbb{D}$ never increases the number of its realizations by more than a multiplicative factor of $O(n^{13})$. This can be viewed as a strengthening of P-stability, a concept introduced by Jerrum and Sinclair in 1989. In particular, it implies fast mixing of the switch Markov-chain on all degree sequences in $\mathbb{D}$. Here we sketch the main ideas of our proof.

Keywords fully graphic degree sequences, simple degree sequence region, P-stable degree sequences, switch Markov chain

1 Introduction

A **degree sequence** is a sequence $\mathbf{d} = (d_1, \ldots, d_n)$ of positive integers with even sum, denoted by Σ. (In this paper, the parameter Σ is always even.) We say $\mathbf{d}$ is graphic if there exists a simple graph on the vertex set $\{v_1, \ldots, v_n\}$ such that vertex v_i has degree d_i for every $i \in [n]$. Let $n > c_1 \geq c_2$ and Σ be positive integers. Denote

$$\mathbb{D}(n, \Sigma, c_1, c_2) = \{(d_1, \ldots, d_n) : \quad c_1 \geq d_i \geq c_2 \;\text{ and }\; \textstyle\sum_{i=1}^{n} d_i = \Sigma\}. \tag{1}$$

We call arbitrary unions of such sets **simple degree sequence regions**. A special case arises when one does not fix the sum of the degrees:

$$\mathbb{D}(n, c_1, c_2) = \bigcup \{\mathbb{D}(n, \Sigma, c_1, c_2) : \forall \Sigma \;\text{ with }\; n \cdot c_1 \geq \Sigma \geq n \cdot c_2\} \tag{2}$$

*HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u 13–15 Budapest, 1053 Hungary. E-mail: `erdos.peter@renyi.hu`. Supported in part by NKFIH grant SNN 135643, K 132696 and he would like to thank the Institute for Computational and Experimental Research in Mathematics (ICERM) in Providence, Rhode Island, for its hospitality during the Fall research semester in 2024.

†Northeastern University, 360 Huntington Ave, Boston, MA, 02115. E-mail: `g.lippner@northeastern.edu`. Supported in part Simons Foundation Collaboration Grant No 953420

‡HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u 13–15 Budapest, 1053 Hungary. E-mail: `soukup.lajos@renyi.hu`.

We refer to an arbitrary union of such sets as a **very simple degree sequence region**. Typically, a (very) simple degree sequence region may contain both graphic and non-graphic degree sequences. If a region is such that it only contains graphic degree sequences, then it is called **fully graphic**. All of these definitions were introduced in [2].

The notion of P-stability of an infinite set of graphic degree sequences was introduced by Jerrum and Sinclair in [6]. The notion plays a crucial role in complexity theory and graph theory. Here we recall this concept and introduce a more refined variant of it that is similar to 2-k-stability used in [3]. One benefit is that the more refined version applies to finite sets of sequences as well. At the same time, all regions that are known to be P-stable satisfy this more refined version.

Given a graphic degree sequence D of length $|D| = n$, denote $\mathcal{G}(D)$ the set of all its realizations, and let

$$\partial^{++}(D) = \sum_{1\le i<j\le n} \left|\mathcal{G}\left(D + 1^i_j\right)\right| / |\mathcal{G}(D)|. \tag{3}$$

where the vector 1^i_j consists of all zeros, except at the ith and jth coordinates, where the values are $+1$. The operation $D \mapsto D + 1^i_j$ is called a **perturbation operation** on degree sequences. Its inverse, given by $D \mapsto D - 1^i_j$, is an analogous operation. Let us emphasize that we do not assume that i and j are different, so for example the operation $D \mapsto D + 1^j_j$ is also defined, and it adds 2 to the j^{th} coordinate.

A family $\mathbb{D}$ of degree sequences is called **P-stable** if there is a polynomial $p(n)$ such that $\partial^{++}(D) \le p(n)$ for each graphic element D of $\mathbb{D}$. If we want to emphasize a specific polynomial, we will say that the family is **p(n)-stable**. While P-stability is only relevant for infinite families, $p(n)$-stability is meaningful even for a single sequence. It is important that we do not require that every element of a P-stable (or $p(n)$-stable) family to be graphic.

In the years between 1989 and 2024 only four P-stable infinite degree sequence sets were found, by Jerrum, McKay, Sinclair and by Greenhill and her colleagues. All four regions are defined via a system of inequalities. Let φ a relation on $\mathbb{N}^4$ then define

$\mathbb{D}[\varphi] = \bigcup \left\{\mathbb{D}(n, \Sigma, c_1, c_2) : (n, \Sigma, c_1, c_2) \in \mathbb{N}^4 \text{ such that } \varphi(n, \Sigma, c_1, c_2) = \textbf{TRUE}\right\}.$

The known P-stable degree sequence sets are the following:

(P1) (Jerrum, McKay, Sinclair [5]) The very simple degree sequence region $\mathbb{D}[\varphi_{JMS}]$ is P-stable, where $\varphi_{JMS} \equiv (c_1 - c_2 + 1)^2 \le 4c_2(n - c_1 - 1)$.

(P2) (Jerrum, McKay, Sinclair [5]) The simple degree sequence region $\mathbb{D}[\varphi^*_{JMS}]$ is P-stable, where $\varphi^*_{JMS} \equiv (\Sigma - nc_2)(nc_1 - \Sigma) \le (c_1 - c_2)\left\{(\Sigma - nc_2)(n - c_1 - 1) + (nc_1 - \Sigma)c_2\right\}$.

(P3) (Greenhill, Sfragara [4]) The simple degree sequence region $\mathbb{D}[\varphi_{GS}]$ is P-stable, where $\varphi_{GS} \equiv (2 \le c_2 \text{ and } 3 \le c_1 \le \sqrt{\Sigma/9})$. (This result was not announced explicitly, but [4, Lemma 2.5] clearly proved this fact.)

(P4) (Gao, Greenhill [3]) $\mathbb{D}[\varphi_{GG}]$ is P-stable, where $\varphi_{GG} \equiv (\sum_{j=1}^{c_1} d_j + 6c_1 + 2 \le \sum_{j=c_1+1}^{n} d_j)$.

In [2] it was proved that the first three regions are fully graphic, while it is easy to see that the same holds for the fourth region as well. That paper also proved that every fully graphic very simple degree region is also P-stable. In this paper we will extend this result by showing that every fully graphic simple region is $3n^{13}$-stable, and hence P-stable.

2 The fully graphic simple regions are P-stable

In this Section we present our main results.

Theorem 1. *If $\mathbb{D}(n, \Sigma, c_1, c_2)$ is fully graphic, then for each $D \in \mathbb{D}(n, \Sigma, c_1, c_2)$ we have*

$$\partial^{++}(D) \leq 3 \cdot n^{13}.$$

The proof of Theorem 1 can be derived from Theorem 3 below similarly to how the proof of Theorem 4.4. follows from Lemma 4.5 in [2, Page 18] so we omit it here.

Definition 2. Let G be a graph, and let v and v' be (not necessarily distinct) vertices of G. An alternating trail of edges and non-edges between v and v' of odd length is called **edge-abundant** if its first element is an edge, (hence the trail contains more edges than non-edges). An edge-abundant trail of length at most 11 is called **witness trail**.

Theorem 3. *Assume that $D \in \mathbb{D}(n, \Sigma, c_1, c_2)$, $1 \leq p, q \leq n$ and let $D^\diamond = D + 1_q^p$. Suppose G is a graph with vertex set $V = \{v_1, \ldots v_n\}$ and degree sequence $D^\diamond$, moreover, $\Gamma_G(v_p) = \Gamma_G(v_q)$. If $\mathbb{D}(n, \Sigma, c_1, c_2)$ is fully graphic, then there exists a witness trail between v_p and v_q in G.*

Remark. *It is well-known, and easy to show, that there always exists an edge-abundant alternating trail from v_p to v_q. The key point of Theorem 3 is that guarantees the existence of a* short *one.*

Our strategy to prove Theorem 3 can be outlined as follows:

- Assuming there are no $v_p \to v_q$ witness trails in G allows us to identify certain structural properties of G, similar to what has been previously used in [5, 2].
- Still assuming the lack of witness trails, we are able to gradually change G to a new graph G' that has what we call a **dream structure** (see below in the statement of Theorem 4 and in Figure 1).
- The graph G' has a degree sequence of the form $\tilde{D} = D'' + 1_q^p$ for some $D'' \in \mathbb{D}(n, \Sigma, c_1, c_2)$. However, the dream structure is designed such that whenever a graph with degree sequence $\tilde{D}$ has this structure, then $D'' = \tilde{D} - 1_q^p$ cannot be graphic. This leads to a contradiction because $\mathbb{D}(n, \Sigma, c_1, c_2)$ was assumed to be fully graphic.

Theorem 4. *Let $D'' \in \mathbb{D}(n, \Sigma, c_1, c_2)$, $1 \leq p \leq q \leq n$, and let $G' \in \mathcal{G}(D'' + 1_q^p)$,. Assume furthermore that $V(G')$ can be partitioned into four subsets $V(G') = S \uplus K' \uplus Y' \uplus R'$ satisfying the following properties: (a) $G'[K']$ is a complete subgraph; (b) $G'[K'; R']$ is a complete bipartite subgraph; (c) $G'[Y']$ is an empty subgraph; (d) $G'[Y'; R']$ is an empty bipartite subgraph; (e) $S = \{v_p, v_q\}$ and $\Gamma(v_p) \cup \Gamma(v_q) \subseteq K'$. Then the degree sequence D'' is not graphic. (See Figure 1)*

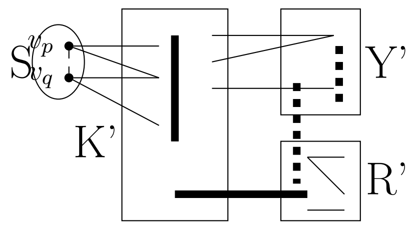


Figure 1: The Dream Structure. The thick dashed lines indicate no edge at all, the solid thick lines indicate all edges.

Proof. Assume on the contrary that we have a graph G'' with degree sequence D''. Then there is an edge-abundant alternating train in the symmetric difference of G' and G'' between v_p and v_q. However, the dream structure forces any alternating trail starting at v_p to alternate forever between K' and Y' and thus never reach v_q. ☐

Proof of Theorem ??

We prove the contrapositive of the statement of the theorem: We assume that there is no witness trail between v_p and v_q in G, and we will obtain a non-graphic element $D^\circ \in \mathbb{D}(n, \Sigma, c_1, c_2)$ by constructing a graph $G^\bullet$ having a dream structure and degree sequence $D^\circ + 1_q^p$. The proof consists of two parts. First, we show that the graph G has one of two possible specific structures. Second, we obtain $G^\bullet$ by transforming G gradually until it has the dream structure. This will be done differently based on which original structure G has.

Part 1: We start by defining pairwise disjoint subsets of the vertices: Let $S = \{x_p, x_q\}$; $X = \Gamma(x_p) = \{v_1 \ldots, v_\ell\}$; $Y = \{y \in V : |X \setminus \Gamma_G(y)| \geq 2\}$; $Z = \Gamma_G(Y) \setminus X = \{v_{\ell+1}, \ldots, v_k\}$; $K = X \cup Z$; and $R = V \setminus (S \cup X \cup Y \cup Z)$.

Lemma 5 (Jerrum, McKay Sinclair [5]). *The lack of $v_p \to v_q$ edge-abundant alternating trails of length at most 7 imply that:* (i) *$x_p x_q$ is not edge ;* (ii) *$G[Y]$ is independent; finally* (iii) *$G[K]$ is a clique.*

So far, we repeated arguments of [5]. The subsequent observations are new: based on the absence of $v_p \to v_q$ witness trails we are able to restrict the structure of G even further. Consider the following partition of R. Let $R_0 = \{r \in R : K \subset \Gamma_G(r)\}$; $R_i = \{r \in R : K \setminus \Gamma_G(r) = \{w_i\}\}$ for $1 \leq i \leq k$; and $R_\infty = \{r \in R : |K \setminus \Gamma_G(r)| \geq 2\}$. The next lemma describes properties of this partition. Its proof is routine and hence omitted.

Lemma 6. (1) *R_∞ does not contain edges.* (2) *There is no edge between R_i and R_j for $i \neq j \in (\{1, 2, \ldots, k\} \cup \{\infty\})$.* (3) *If there exists an edge (a, b) in R_i for some $1 \leq i \leq k$ then $R_j = \emptyset$ provided $1 \leq j \leq k$ and $j \neq i$.*

After this preparation we can see that we can distinguish two cases concerning the structure of G.

Case I: $R_N \coloneqq \bigcup\{R_i : 1 \leq i \leq k\} \cup R_\infty$ is an independent subset (see Figure2a),

Case II: There exists $1 \leq i \leq k$ such that R_i contains edges (see Figure2b).

Part 2: Our goal now is to transform our degree sequence D into a new degree sequence $D'' \in \mathbb{D}(n, \Sigma, c_1, c_2)$ and our graph G into a realization $G' \in \mathcal{G}(D'' + 1_q^p)$ which has a dream structure. In both cases, we will describe the sequence of transformations to achieve this, but will not provide proof of the key fact that the resulting degree sequence D'' is indeed in $\mathbb{D}(n, \Sigma, c_1, c_2)$. The proof of this (in particular that the c_1 bound is not violated) requires a careful and technical analysis that is again based on the absence of witness trails. In the second case we also need to keep track of how witness trails change during the transformation process. We omit this part of the proof due to space constraints.

Definition 7. Assume that H is a graph and x, y, z are distinct vertices such that (x, y) is an edge and (x, z) is a non-edge. The well-known **hinge-flip** operation $(x, y) \Rightarrow (x, z)$ deletes the edge (x, y) from H and adds the edge (x, z) to obtain a new graph H'.

Case I. There is no edge in R_N.

We generate a sequence of graphs $G_0, G_1, \ldots, G_m$ and their degree sequences $D_0, D_1, \ldots, D_m$ with consecutive hinge-flip operations as follows.

Let $G_0 = G$, and so $D_0 = D^\diamond = D + 1_q^p$. Assume that G_ℓ is already produced. Pick $x_\ell \neq z_\ell \in R_0$ and $y_\ell \in R_N$ such that (x_ℓ, y_ℓ) is an edge and (x_ℓ, z_ℓ) is a non-edge in G_ℓ.

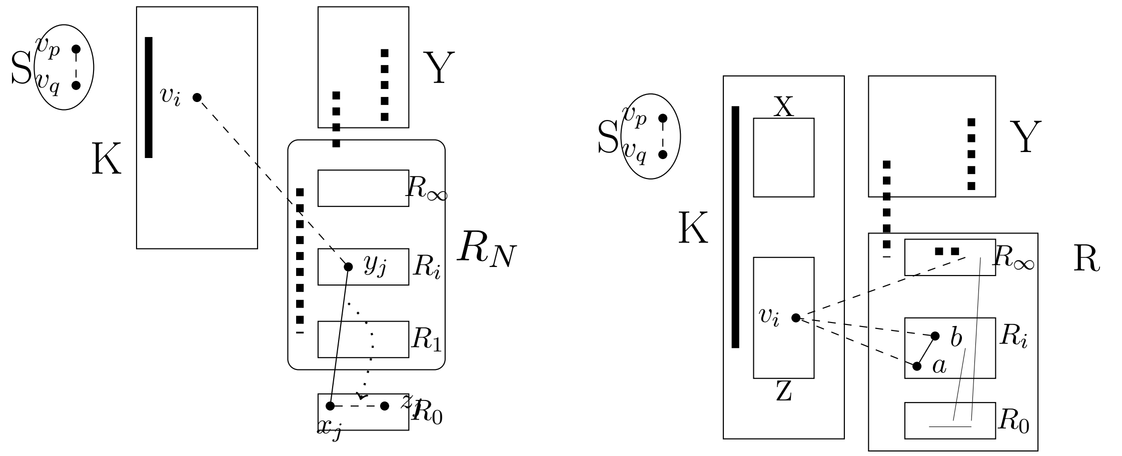
Apply the corresponding hinge-flip operation $(x_\ell, y_\ell) \Rightarrow (x_\ell, z_\ell)$ for G_ℓ to obtain $G_{\ell+1}$. (We will call this operation a **downward twist**.) We stop the process if there are no suitable x_ℓ, y_ℓ, and z_ℓ.

Now let $R_0^0 := \{c \in R_0 : \Gamma_m(c) \cap R_N \text{ is not empty}\}$ and $R_0^1 = R_0 \setminus R_0^0$.

Lemma 8. *It is easy to see that G_m has a dream structure, namely $S := \{v_p, v_q\}$, $K' := K \cup R_0^0$, $Y' := Y \cup R_N$, and $R' := R_0^1$.*

Case II. There is an edge in R_N.

In this case there exists $1 \le i \le k$ such that R_i contains an edge ab and $R = R_0 \cup R_i \cup R_\infty$. (See Figure 2b.) This case is more challenging than **Case I** was. We will do it in two steps.



(a) In R_N there is no edge. The vertices $x_j, z_j \in R_0$ and $y_j \in R_N$. The curved, loosely dotted arrow indicates a hinge-flip operation.

(b) In Y, in R_∞, and in $G[Y, R]$ there is no edge.

Figure 2: The solid lines are existing edges, the dashed line are missing edges The thick lines note all edges / non-edges in the subgraph.

Step 1: In the first step we will "uplift" edges from R_i into the bipartite graph $G[R_i, K]$. Let $H_1, \ldots, H_t$ be an enumeration of the connected components of the subgraph $G[R_i]$. In each component H_s fix a vertex r_s and choose a spanning tree T_s with root r_s such that $\Gamma_{T_s}(r_s) = \Gamma_{G[R_i]}(r_s)$.

We define the **upward twist** operation of this graph as follows. In each spanning tree T_s let orient the edges outward from the root r_s. Then, for each oriented edge $\overrightarrow{xy}$ execute the hinge-flip operation $(y, x) \Rightarrow (y, v_i)$. At the end we obtain the graph $G^\star$ with degree sequence $D^\star$. Clearly, the degree sum does not change along the operations.

Step 2: In the second step we imitate the "downward twist" construction of **Case I** in the graph $G^\star$. Denote

$$R_i^1 := \{r_\beta : 1 \le \beta \le t\}, \text{ and } R_i^0 := (R_i \setminus R_i^1),$$

so R_i^1 consists of the roots of the spanning tree in R_i. Then write

$$R_0^\star = R_0 \cup R_i^0 \text{ and } R_N^\star = R_\infty \cup R_i^1.$$

After this preparation we can proceed similarly to Case I. Generate a sequence of graphs $G_0, G_1, \ldots, G_m$ and their degree sequences $D_0, D_1, \ldots, D_m$ with consecutive hinge-flip operations as follows.

Let $G_0 = G^\star$. Assume that G_ℓ is already produced. Pick $x_\ell \neq z_\ell \in R_0^\star$ and $y_\ell \in R_N^\star$ such that (x_ℓ, y_ℓ) is an edge and (x_ℓ, z_ℓ) is a non-edge in G_ℓ. Apply the corresponding hinge-flip operation $(x_\ell, y_\ell) \Rightarrow (x_\ell, z_\ell)$ for G_ℓ to obtain $G_{\ell+1}$ (These are just the downward twist operations from Case I.) We stop the construction if there are no suitable x_ℓ, y_ℓ and z_ℓ. Finally let

$$R_K = \{r \in R_0^\star : \text{there is } G_m\text{-edge between } r \text{ and } R_N^\star\}.$$

We have the set $S = \{v_p, v_q\}$, and let

$$K' = K \cup R_K, \qquad Y' = Y \cup R_N^*, \qquad R' = R_0^* \setminus R_K.$$

Now it is easy to see that (S, K', Y', R') is a dream configuration for G_m which finishes the proof of Theorem 3.

Based on our results, as well as the known examples of non-P-stable sequences, it is reasonable to conjecture that the set of degree sequences that are ”far” from being not graphic form a P-stable family.

References

[1] P. Erdős – T. Gallai: Graphs with prescribed degree of vertices (in Hungarian), *Mat. Lapok* **11** (1960), 264–274.

[2] P.L. Erdős – I. Miklós – L. Soukup: Fully graphic degree sequences and P-stable degree sequences, *Advances in Applied Mathematics* **163** (2025), #102 805, pp. 29. DOI:10.1016/j.aam.2024.102805,

[3] P. Gao – C. Greenhill: Mixing time of the switch Markov chain and stable degree sequences, *Discrete Applied Mathematics* **291** (2021), 143–162. DOI:10.1016/j.dam.2020.12.004

[4] C. Greenhill – M. Sfragara: The switch Markov chain for sampling irregular graphs, *Theoretical Comp. Sci* **719** (2018), 1–20. DOI:10.1016/j.tcs.2017.11.010

[5] M.R. Jerrum – B.D. McKay – A. Sinclair: When is a Graphical Sequence Stable? Chapter 8 in *Random Graphs: Volume 2* Eds.: A. Frieze and T. Luczak, Wiley, (1992), pp. 304. ISBN: 978-0-471-57292-3

[6] M.R. Jerrum – A. Sinclair: Approximating the Permanent, *SIAM J. Comput.*, **18** (6) (1989), 1149–1178. DOI:10.1137/0218077

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Polystochastic matrices of order 4 with zero permanent

(Extended abstract)

Aleksei L. Perezhogin* Vladimir N. Potapov† Anna A. Taranenko‡
Sergey Yu. Vladimirov§

Abstract

A multidimensional nonnegative matrix is called polystochastic if the sum of its entries over each line is equal to 1. The permanent of a multidimensional matrix is the sum of products of entries over all diagonals. We prove that if d is even, then the permanent of a d-dimensional polystochastic matrix of order 4 is positive, and for odd d, we give a complete characterization of d-dimensional polystochastic matrices with zero permanent.

A d-dimensional matrix A of order n is an array $A = (a_\alpha)_{\alpha \in I_n^d}$, $a_\alpha \in \mathbb{R}$, where $I_n^d = \{\alpha = (\alpha_1, \dots, \alpha_d),\ \alpha_i \in \{0, \dots, n-1\}\}$ is the index set of A. A *k-dimensional plane* in a matrix A is a k-dimensional submatrix obtained by fixing $d-k$ positions of indices and letting the values in other k positions vary from 0 to $n-1$. A 1-dimensional plane of the matrix A is said to be a *line*, and $(d-1)$-dimensional planes are *hyperplanes*. Two multidimensional matrices are called *equivalent* if they are obtained from each other by permutations of positions of indices or by permutations of parallel hyperplanes.

The *support* $\mathrm{supp}(A)$ of a matrix A is the set of all indices α for which $a_\alpha \neq 0$. A d-dimensional matrix A of order n is a *polystochastic matrix* if $a_\alpha \geq 0$ for all α and the

*Sobolev Institute of Mathematics, 630090 Novosibirsk, Russia. E-mail: `pereal@math.nsc.ru`. Supported by the state contract of the Sobolev Institute of Mathematics (project no. FWNF-2022-0017).

†Independent researcher. E-mail: `quasigroup349@gmail.com`.

‡Sobolev Institute of Mathematics, 630090 Novosibirsk, Russia. E-mail: `taa@math.nsc.ru`. Supported by the state contract of the Sobolev Institute of Mathematics (project no. FWNF-2022-0017).

§Novosibirsk State University, 630090 Novosibirsk, Russia. E-mail: `s.vladimirov@g.nsu.ru`.

sum of entries over each line of A is equal to 1. 2-dimensional polystochastic matrices are known as doubly stochastic. If all entries of a multidimensional matrix A are either 0 or 1, then A is a $(0,1)$-*matrix*, and d-dimensional polystochastic $(0,1)$-matrices are said to be *d-dimensional permutations*.

A *diagonal* in a d-dimensional matrix A of order n is a set $\{\alpha^1, \ldots, \alpha^n\}$ of n indices such that each pair α^i and α^j is distinct in all components. The *permanent* of a multidimensional matrix A is given by

$$\mathrm{per} A = \sum_{p \in D(A)} \prod_{\alpha \in p} a_\alpha,$$

where $D(A)$ is the set of all diagonals of A.

It is easy to see that the number of perfect matchings in a d-partite d-uniform hypergraph with parts of size n is equal to the permanent of the d-dimensional matrix of order n representing the adjacency of parts (see, e.g., [9]). The problem of determining the positivity of the permanent of a d-dimensional $(0,1)$-matrix of order n is NP-hard because for $d = 3$ it is equivalent to the 3-dimensional matching problem (one of Karp's 21 NP-complete problems). Certain conditions on the number of 1s in lines, sufficient for the positivity of the permanent of d-dimensional $(0,1)$-matrices, were established in [1] (in the context of hypergraphs).

The permanent of multidimensional permutations is closely related to the number of transversals in latin squares and hypercubes. A d-dimensional latin hypercube Q of order n is a multidimensional matrix filled with n symbols so that each line contains all different symbols. 2-dimensional latin hypercubes are usually called latin squares. A transversal in a latin hypercube Q is a diagonal that contains all n symbols.

There is a one-to-one correspondence between d-dimensional latin hypercubes Q of order n and $(d+1)$-dimensional permutations A of order n: an entry $q_{\alpha_1,\ldots,\alpha_d}$ of Q equals α_{d+1} if and only if an entry $a_{\alpha_1,\ldots,\alpha_{d+1}}$ of A equals 1. Jurkat and Ryser [13] were the first to note that the number of transversals in a latin hypercube Q coincides with the permanent of the corresponding polystochastic matrix A.

The well-known Birkhoff theorem (see, e.g., [4]) states that every doubly stochastic matrix has a positive permanent and is a convex combination of some permutation matrices. However, for $d \geq 3$ there exist d-dimensional polystochastic matrices with zero permanent, even when the matrix is a multidimensional permutation. For example, latin squares corresponding to the Cayley table of groups $\mathbb{Z}_n$ of even order n have no transversals.

This observation can be extended to latin hypercubes. Let $\mathcal{Q}_n^d$ be the d-dimensional latin hypercube of order n such that $q_\alpha \equiv \alpha_1 + \cdots + \alpha_d \mod n$. In [12], Wanless showed that if n and d are even, then the latin hypercube $\mathcal{Q}_n^d$ has no transversals. This implies that the permanent of the multidimensional permutation $\mathcal{M}_n^d$ corresponding to the latin hypercube $\mathcal{Q}_n^{d-1}$ is zero when n is even and d is odd. Later, in [2] Child and Wanless proved that modifications of the matrix $\mathcal{M}_n^d$ in r consequentive hyperplanes, where $r(r-1) < n$, also produce polystochastic matrices with zero permanent for even n and odd d.

There are no known examples of latin squares of odd order with no transversals, and in 1967 Ryser [7] conjectured that every latin square of odd order has a transversal. Despite

Montgomery's recent breakthrough [5], proving that every latin square of large order n has a near transversal (a partial transversal of length $n-1$), the Ryser's conjecture is still open.

In [8] Sun proved that latin hypercubes $\mathcal{Q}_n^d$ of odd dimension d have a transversal, and so all matrices $\mathcal{M}_n^d$ of even dimension d have a positive permanent. In that paper, he also conjectured that all 3-dimensional latin hypercubes have a transversal.

On the basis of these results and computer enumeration of latin hypercubes of small order and dimension [3], Wanless conjectured the following.

Conjecture 1 (Wanless, [12]). *Every latin hypercube of odd order or odd dimension has a transversal.*

This conjecture is trivial for latin hypercubes of order 2 and easy to prove for order 3 [9]. In [10], Taranenko showed that, except for the hypercube $\mathcal{Q}_4^d$ of even dimension d, all latin hypercubes of order 4 have a transversal, and Perezhogin, Potapov, and Vladimirov in [6] proved that all latin hypercubes of order 5 have a transversal. In [9], Taranenko extended Conjecture 1 from multidimensional permutations to polystochastic matrices.

Conjecture 2 (Taranenko, [9]). *The permanent of every polystochastic matrix of odd order n or even dimension d is positive.*

It is straightforward to show that all polystochastic matrices of order 2 with zero permanent are matrices $\mathcal{M}_2^d$ with odd d (see [9]). It is also proved in [9] that every polystochastic matrix of order 3 has a positive permanent. Finally, in [11], Taranenko proved that the permanent of every 4-dimensional polystochastic matrix of order 4 is positive.

We prove that for even d, every d-dimensional polystochastic of order 4 has a positive permanent, and for odd d, we show that d-dimensional polystochastic matrices of order 4 with zero permanent are either equivalent to $\mathcal{M}_4^d$ or matrices constructed by Child and Wanless in [2]. Thus, we confirm Conjecture 2 for polystochastic matrices of order $n=4$.

1 Main result and proof outline

Let $\mathcal{M}_n^d$ be the d-dimensional permutation of order n such that $m_\alpha = 1$ if $\alpha_1 + \cdots + \alpha_d \equiv 0 \mod n$. Denote by $\mathcal{L}_n^d$ the family of d-dimensional polystochastic matrices of order n obtained as a convex sum $\lambda\mathcal{M}_n^d + (1-\lambda)M$, $0<\lambda<1$, where M is the matrix equivalent to $\mathcal{M}_n^d$ such that $\operatorname{supp}(M) = \{\alpha : \alpha_1 + \cdots + \alpha_{d-1} + \pi(\alpha_d) \equiv 0 \mod n\}$, π is the transposition (01).

It is not hard to prove (see [2, 12]) that matrices $\mathcal{M}_n^d$ and $\mathcal{L}_n^d$ have zero permanent if d is odd and n is even. All known examples of d-dimensional polystochastic matrices of order 4 with zero permanent have odd dimension d and are equivalent to $\mathcal{M}_4^d$ or some matrix from $\mathcal{L}_4^d$.

The main result of the present paper is that $\mathcal{M}_4^d$ and matrices from the family $\mathcal{L}_4^d$ of odd dimension d are the unique (up to equivalence) polystochastic matrices of order 4 with zero permanent.

Theorem 1. *Let A be a d-dimensional polystochastic matrix of order 4 such that* $\mathrm{per}A = 0$. *Then d is odd and A is equivalent to $\mathcal{M}_4^d$ or some matrix from $\mathcal{L}_4^d$.*

We will say that a polystochastic matrix A is a *sesquialteral permutation* if every line of A contains no more than two nonzero entries. Note that multidimensional permutations and polystochastic matrices that has entries only from the set $\{0, 1/2\}$ (double permutations) are sesquialteral permutations.

A collection of indices U, $U \subset I_n^d$, is a *unitrade*, if every line of a d-dimensional matrix of order n contains zero or two elements from U. A unitrade U is called a *bitrade* if there is a sign function $\sigma : U \to \{\pm 1\}$ such that for every α, β from the same line it holds $\sigma(\alpha) \neq \sigma(\beta)$.

The definitions imply that for every sesquialteral permutation A the set of indices $U(A) = \{\alpha | 0 < a_\alpha < 1\}$ is a unitrade. Note that if $U(A)$ is bitrade with a sign function σ, then there is a multidimensional permutation M such that $\mathrm{supp}(M) \subseteq \mathrm{supp}(A)$, $\mathrm{supp}(M) = \{\alpha \in \mathrm{supp}(A) : a_\alpha = 1 \text{ or } \sigma(\alpha) = 1\}$. By the induction on the size of the support of A, we get that in this case the matrix A is a convex combination of multidimensional permutations.

The proof of the main theorem is by consecutive narrowing the set of matrices that can have the zero permanent. Firstly, we show that only sesquialteral permutations can have zero permanent.

Lemma 1. *Let A be a d-dimensional polystochastic matrix of order 4. If there exists a line in A that contains at least three nonzero entries, then* $\mathrm{per}A > 0$.

Next, we prove that if a sesquialteral permutation A of order 4 has zero permanent, then the support of each 3-dimensional plane of A is equivalent to a matrix from a certain list provided by total computer enumeration of possible supports. Thus we reduce our consideration to some class of sesquialteral permutations that we call suspicious. The similar reduction was previously used for transversals in latin hypercubes of order 5 in [6].

Using restrictions on planes of suspicious sesquialteral permutations A, we show that the unitrade $U(A)$ is a bitrade.

Lemma 2. *Let A be a suspicious d-dimensional sesquialteral permutation of order 4 and $U = U(A)$ be the unitrade of A. Then U is a bitrade.*

It means that if a d-dimensional polystochastic matrix A of order 4 has a zero permanent, then A is a convex sum of some multidimensional permutations with zero permanent. Such permutations were previously described in [10].

Theorem 2 ([10], Theorem 5)**.** *Let A be a d-dimensional permutation of order 4 such that* $\mathrm{per}A = 0$. *Then d is odd and A is equivalent to the matrix $\mathcal{M}_4^d$.*

Thus we obtain the following result.

Lemma 3. *Let A be a d-dimensional polystochastic of order 4 such that* $\mathrm{per}A = 0$. *Then d is odd and A is equivalent to a convex sum of matrices equivalent to $\mathcal{M}_4^d$.*

In the last part of the proof we show that if an odd-dimensional matrix A is equal to a convex sum of matrices equivalent to $\mathcal{M}_4^d$ and has a zero permanent, then A is equivalent to $\mathcal{M}_4^d$ or some matrix from $\mathcal{L}_4^d$. The technique is similar to one used for the proof of Theorem 2 in [10].

References

References

[1] R. Aharoni, A. Georgakopoulos, P. Sprüssel. Perfect matchings in r-partite r-graphs. European J. Combin. 30 (2009), P. 39–42.

[2] B. Child, I. M. Wanless. Multidimensional permanents of polystochastic matrices. Linear Algebra Appl. 586 (2020), P. 89–102.

[3] B. D. McKay, I. M. Wanless. A census of small latin hypercubes. SIAM J. Discrete Math. 22 (2008), P. 719–736.

[4] H. Minc. Permanents. Encyclopedia of Mathematics and its Applications, 6. Addison-Wesley Publishing Co., Reading, MA, 1978. xviii+205 pp. ISBN: 0-201-13505-1

[5] R. Montgomery. A proof of the Ryser–Brualdi–Stein conjecture for large even n. ArXiv:2310.19779. Available at https://arxiv.org/abs/2310.19779.

[6] A. L. Perezhogin, V. N. Potapov, S. Y. Vladimirov. Every latin hypercube of order 5 has transversals. J. Combin. Des. 32 (2024), P. 679–699.

[7] H. J. Ryser. Neuere Probleme der Kombinatorik. Vortrage über Kombinatorik Oberwolfach (Juli 24–29, 1967), P. 69–91.

[8] Z.-W. Sun. An additive and restricted sumsets. Math. Res. Lett. 15(6) (2008), P. 1263–1276.

[9] A. A. Taranenko. Permanents of multidimensional matrices: properties and applications. J. Appl. Ind. Math. 10(4) (2016), P. 567–604. (Translated from Diskretn. Anal. Issled. Oper. 23(4) (2016), P. 35–101).

[10] A. A. Taranenko. Transversals in completely reducible multiary quasigroups and in multiary quasigroups of order 4. Discrete Math., 341(2) (2018), P. 405–420.

[11] A. A. Taranenko. Positiveness of the permanent of 4-dimensional polystochastic matrices of order 4. Discrete Appl. Math. 276 (2020), P. 161–165.

[12] I. M. Wanless. Transversals in Latin squares: a survey. Surveys in combinatorics 2011, volume 392 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 2011, P. 403–437.

[13] W. B. Jurkat and H. J. Ryser, Extremal configurations and decomposition theorems. I, J. Algebra 8 (1968), P. 194–222.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Palettes determine uniform Turán density

(Extended abstract)

Ander Lamaison*

Abstract

Turán problems, which concern the minimum density threshold required for the existence of a particular substructure, are among the most fundamental problems in extremal combinatorics. We study Turán problems for hypergraphs with an additional uniformity condition on the edge distribution. This kind of Turán problems was introduced by Erdős and Sós in the 1980s but it took more than 30 years until the first non-trivial exact results were obtained when Glebov, Král' and Volec [Israel J. Math. 211 (2016), 349–366] and Reiher, Rödl and Schacht [J. Eur. Math. Soc. 20 (2018), 1139–1159] determined the uniform Turán density of $K_4^{(3)-}$.

Subsequent results exploited the powerful *hypergraph regularity method*, developed by Gowers and by Nagle, Rödl and Schacht about two decades ago. Central to the study of the uniform Turán density of hypergraphs are *palette constructions*, which were implicitly introduced by Rödl in the 1980s. We prove that palette constructions always yield tight lower bounds, unconditionally confirming present empirical evidence. This results in new and simpler approaches for determining uniform Turán densities, which completely bypass the use of the hypergraph regularity method.

1 Introduction

In extremal combinatorics, Turán problems, which vastly generalize the classical Turán's theorem from 1941, concern the threshold density for the existence of a specific substructure in a host structure; this threshold density is referred to as the *Turán density*. While Turán densities are very well-understood in the case of graphs [16, 21, 7], Turán problems concerning hypergraphs are one of the most challenging problems in extremal combinatorics. Indeed, Erdős offered \$500 for determining the Turán density of *any* complete hypergraph and \$1000 for determining the Turán densities of all complete hypergraphs. To this day, the value of the Turán density of the complete r-uniform hypergraph with t vertices has not been obtained for any $t > r > 2$. Even for the very simple hypergraph $K_4^{(3)-}$, obtained by removing an edge from $K_4^{(3)}$, the Turán density is unknown [11].

*Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. Supported by IBS-R029-C4. **Email address**: ander@ibs.re.kr

Most of the extremal constructions for Turán problems in the hypergraph setting have large independent sets, i.e., linear-sized sets of vertices with no edges. This led Erdős and Sós [5, 6] to propose studying the *uniform Turán density* of hypergraphs, which is the density threshold for the existence of a hypergraph with the additional requirement that the edges of the host hypergraph are distributed uniformly. We define this notion as follows:

Definition. *A hypergraph H is said to be* $(d, \varepsilon, \therefore)$-dense *if any subset $S \subseteq V(H)$ of size at least $\varepsilon|V(H)|$ has edge density at least $d - \varepsilon$. A sequence of hypergraphs $\{H_i\}_{i=1}^{\infty}$ is said to be* locally d-dense *if each H_i is $(d, \varepsilon_i, \therefore)$-dense with $\varepsilon_i \to 0$, and $|V(H_i)| \to \infty$.*

The uniform Turán density $\pi_{\therefore}(F)$ *of an r-uniform hypergraph F is defined as the supremum of the values of d, for which there exists a locally d-dense sequence of r-uniform hypergraphs which do not contain F as a subgraph.*

Only recently, the resistance of uniform Turán densities has been broken using approaches based on the hypergraph regularity method [19, 1, 9], starting with resolving a 30-year-old problem by Erdős and Sós on determining the uniform Turán density of the 3-uniform hypergraph $K_4^{(3)-}$. Erdős [6] conjectured that $\pi_{\therefore}(K_4^{(3)-}) = 1/4$. Glebov, Král' and Volec [10] gave a computer-assisted proof of this conjecture, which was then proved combinatorially by Reiher, Rödl and Schacht [19].

We now briefly survey recent results on exact values of the uniform Turán densities of 3-uniform hypergraphs (or 3-graphs for short). Reiher, Rödl and Schacht [18] characterized 3-graphs F with $\pi_{\therefore}(F) = 0$. As a consequence of this characterization, they deduced that every 3-graph F with non-zero uniform Turán density satisfies $\pi_{\therefore}(F) \geq 1/27$. In other words, there is a "jump" phenomenon: $\pi_{\therefore}(F)$ does not take values in $(0, 1/27)$. Garbe, Král' and the author [9] constructed 3-graphs with uniform Turán density 1/27. Other classes of 3-graphs whose uniform Turán density are known include tight cycles [1], and a specific family of 3-graphs with uniform Turán density 8/27 [8]; see [2, 15] for more constructions with the same uniform Turán density as some of the examples above.

In all these cases, palette constructions, which were introduced by Reiher [17], extending constructions by Erdős and Hajnal [4] and Rödl [20] and which we introduce in the next section, play a key role. The main result of this submission states that palette constructions are always optimal: the uniform Turán density of every 3-graph F is the supremum of the densities of the palette constructions not containing F as a subgraph.

2 Palettes

Palettes can be seen as a way to generate locally dense sequences of 3-graphs, and also to describe the structure of the subgraphs generated by it.

Definition. *A* palette $\mathcal{P}$ *is a pair $(\mathcal{C}, \mathcal{A})$, where $\mathcal{C}$ is a finite set (whose elements we call* colors*) and a set of (ordered) triples of colors $\mathcal{A} \subseteq \mathcal{C}^3$, which we call the* admissible triples. *The density of $\mathcal{P}$ is $d(\mathcal{P}) := |\mathcal{A}|/|\mathcal{C}|^3$.*

If $\mathcal{P}$ is a palette with density d, one can construct a locally d-dense sequence of 3-graphs $\{H_i\}_{i=1}^{\infty}$. To generate H_n, take $[n]$ as the vertex set. Color each pair of vertices uv with a color $\varphi(uv) \in \mathcal{C}$ chosen uniformly at random. The edges of H_n are the triples $u < v < w$ such that $(\varphi(uv), \varphi(uw), \varphi(vw)) \in \mathcal{A}$. With high probability, the resulting sequence is locally d-dense, as one can show with standard probabilistic tools.

Definition. *We say that a* 3*-graph* F admits *a palette* $\mathcal{P}$ *if there exists an order* $\preceq$ *on* $V(F)$ *and a function* $\varphi : \binom{V(F)}{2} \to \mathcal{C}$ *such that for every edge* $uvw \in E(F)$ *with* $u \prec v \prec w$ *we have* $(\varphi(uv), \varphi(uw), \varphi(vw)) \in \mathcal{A}$.

It is easy to see that any subgraph of H_n admits the palette $\mathcal{P}$. Therefore, if a 3-graph F does not admit $\mathcal{P}$, then it is not a subgraph of the locally d-dense sequence $\{H_i\}_{i=1}^{\infty}$, and so $\pi_{\therefore}(F) \geq d = d(\mathcal{P})$.

All the lower bound constructions for the tight results on uniform Turán density mentioned above are derived from palettes via this procedure. The same applies to the conjectured optimal constructions for other families of 3-graphs, including complete graphs and stars [17]. This motivated the following question, which had circulated in the community and is explicitly discussed in [17, Section 3]: can all lower bounds on uniform Turán density be obtained or approximated arbitrarily with palette constructions? Our main theorem answers this question in the affirmative.

Theorem 2.1. *For every* 3*-graph* F*,*

$$\pi_{\therefore}(F) = \sup\{d(\mathcal{P}) : \mathcal{P} \textit{ palette, } F \textit{ does not admit } \mathcal{P}\}. \tag{1}$$

The proof of the main theorem is based on a method developed by Reiher [17], using some auxiliary structures called reduced hypergraphs. This method is itself based on the hypergraph regularity method, where the properties of locally dense sequences of hypergraphs are used to simplify the structure of their regularity partitions. To prove the main theorem, we show that every large enough reduced hypergraph, with edge density d, contains large repetitive substructures (which can be associated with the regularity partitions of palette constructions) with density at least $d - \varepsilon$. The proof is based on an iterative algorithmic approach, in which randomized vertex selections are alternated with applications of the hypergraph Ramsey theorem.

3 Applications

An important reason why Theorem 2.1 is significant is that, judging from all the proofs available so far, the right hand side of (1) is considerably easier to compute than the left hand side. In fact, while proofs about the uniform Turán density always use the hypergraph regularity method or Reiher's reduced hypergraphs (with the exception of [10], which uses the flag algebra method), computing the right side of (1) does not rely on regularity at all. Therefore, Theorem 2.1 yields an easier method to study the uniform Turán density of hypergraphs.

One noteworthy example of this is the tight cycle $F = C_\ell^{(3)}$, for $\ell \geq 5$ not divisible by 3. These cycles do not admit the palette $\mathcal{P} = (\mathcal{C}, \mathcal{A})$ with $\mathcal{C} = \{1, 2, 3\}$ and $\mathcal{A} = \{(1,2,2), (1,2,3), (1,3,2), (1,3,3)\}$, so $\pi_{\therefore}(C_\ell^{(3)}) \geq 4/27$. In his Master's thesis, Cooper [3] proved that $C_\ell^{(3)}$ admits every palette with density greater than $4/27$, in a five page proof consisting mostly of simple double-counting arguments. It took three years and a considerable amount of effort to prove that $\pi_{\therefore}(C_\ell^{(3)}) = 4/27$ in a highly technical, thirty page proof [1].

In the full paper version of this extended abstract we give a one page proof of the equality $\pi_{\therefore}(K_4^{(3)-}) = 1/4$, simplifying the proof from [19]. Here we give a short proof of Reiher, Rödl and Schacht's characterization of hypergraphs with uniform Turán density equal to zero.

Theorem 3.1 ([18]). *Let $\mathcal{P}_0$ be the palette $(\mathcal{C}_0, \mathcal{A}_0)$ with $\mathcal{C}_0 = \{1,2,3\}$ and $\mathcal{A}_0 = \{(1,2,3)\}$. Then for every 3-graph F, $\pi_{\therefore}(F) = 0$ iff F admits $\mathcal{P}_0$.*

Proof. If F does not admit $\mathcal{P}_0$, then $\pi_{\therefore}(F) \geq d(\mathcal{P}_0) = 1/27 > 0$. On the other hand, suppose that F does admit $\mathcal{P}_0$. By definition, there exists an order $\preceq$ on $V(F)$ and a coloring $\varphi : \binom{V(F)}{2} \to \mathcal{C}_0$ such that, for every edge $uvw \in E(F)$ with $u \prec v \prec w$, we have $(\varphi(uv), \varphi(uw), \varphi(vw)) = (1,2,3)$.

We claim that F admits all palettes with at least one admissible triple. Indeed, let $\mathcal{P} = (\mathcal{C}, \mathcal{A})$ be a non-empty palette. Let (a,b,c) be an element of $\mathcal{A}$ (the colors a,b,c do not need to be distinct). Let ψ be the function with $\psi(1) = a$, $\psi(2) = b$ and $\psi(3) = c$. Then every edge uvw with $u \prec v \prec w$ satisfies $(\psi \circ \varphi(uv), \psi \circ \varphi(uw), \psi \circ \varphi(vw)) = (a,b,c) \in \mathcal{A}$, meaning that F admits $\mathcal{P}$. From (1), we obtain $\pi_{\therefore}(F) = 0$. $\square$

The simplicity of the palette method compared to the regularity method has allowed for more exact values of $\pi_{\therefore}$ to be found. Previous to Theorem 2.1, the only values of d for which hypergraphs F with $\pi_{\therefore}(F) = d$ were known to exist were $0, 1/27, 4/27, 1/4$ and $8/27$. Using Theorem 2.1, for each $k \geq 2$ we constructed a hypergraph F_k with $\pi_{\therefore}(F_k) = \frac{1}{2} - \frac{1}{2k}$.

The k-star S_k is a 3-graph on $k+1$ vertices $u, v_1, v_2, \ldots, v_k$, where the edges are all the triples of the form uv_iv_j with $1 \leq i < j \leq k$. Reiher, Rödl and Schacht [19] studied the uniform Turán density of stars, and proved that

$$\frac{k^2 - 5k + 7}{(k-1)^2} \leq \pi_{\therefore}(S_k) \leq \left(\frac{k-2}{k-1}\right)^2.$$

In [14], Wu and the author used Theorem 2.1 to prove that $\pi_{\therefore}(S_k) = \frac{k^2-5k+7}{(k-1)^2}$ for all $k \geq 48$, providing another infinite family of values of the uniform Turán density.

Further information on how to use Theorem 2.1 to study the uniform Turán density of hypergraphs can be found in the paper [12] by King, Sales and Schülke, as well as in upcoming work by Král', Kučerák, Tardos and the author [13].

References

[1] Matija Bucić, Jacob W. Cooper, Daniel Král', Samuel Mohr, and David Munhá Correia. Uniform Turán density of cycles. *Transactions of the American Mathematical Society*, 376(7):4765–4809, 2023.

[2] August Y. Chen and Bjarne Schülke. Beyond the broken tetrahedron. Preprint. arXiv:2211,12747, 2022.

[3] Jacob W. Cooper. Uniform Turán density. Msc thesis, University of Warwick, 2018.

[4] Paul Erdős and András Hajnal. On Ramsey like theorems, problems and results. In *Combinatorics (Proc. Conf. Combinatorial Math.)*, pages 123–140. Mathematical Institute, Oxford, 1972.

[5] Pál Erdős and Vera T. Sós. On Ramsey-Turán type theorems for hypergraphs. *Combinatorica*, 2(3):289–295, 1982.

[6] Paul Erdős. Problems and results on graphs and hypergraphs: similarities and differences. In J. Nešetřil and V. Rödl, editors, *Mathematics of Ramsey theory*, pages 223–233. Springer-Verlag, 1990.

[7] Paul Erdős and Arthur H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(12):1087–1091, 1946.

[8] Frederik Garbe, Daniel Il'kovič, Daniel Král', Filip Kučerák, and Ander Lamaison. Hypergraphs with uniform Turán density equal to 8/27. Preprint. arXiv:2407.05829, 2024.

[9] Frederik Garbe, Daniel Král', and Ander Lamaison. Hypergraphs with minimum positive uniform Turán density. *Israel Journal of Mathematics*, 259(2):701–726, 2024.

[10] Roman Glebov, Jan Volec, and Daniel Král'. A problem of Erdős and Sós on 3-graphs. *Israel Journal of Mathematics*, 211(1):349–366, 2016.

[11] Peter Keevash. Hypergraph Turán problems. In R. Chapman, editor, *Surveys in Combinatorics 2011*, London Mathematical Society Lecture Note Series, pages 83–140. Cambridge University Press, 2011.

[12] Dylan King, Marcelo Sales, and Bjarne Schülke. Lagrangians are attained as uniform turán densities. Preprint. arXiv:2412.07297, 2024.

[13] Daniel Král', Filip Kučerák, Ander Lamaison, and Gábor Tardos. Uniform turán density – classification of palettes. In preparation.

[14] Ander Lamaison and Zhuo Wu. The uniform Turán density of large stars. Preprint. arXiv:2409.03699, 2024.

[15] Hao Li, Hao Lin, Guanghui Wang, and Wenling Zhou. Hypergraphs with a quarter uniform Turán density. Preprint. arXiv:2305.11749, 2023.

[16] Willem Mantel. Problem 28. *Wiskundige Opgaven*, 10:60–61, 1907.

[17] Christian Reiher. Extremal problems in uniformly dense hypergraphs. *European Journal of Combinatorics*, 88:103117, 2020.

[18] Christian Reiher, Vojtěch Rödl, and Mathias Schacht. Hypergraphs with vanishing Turán density in uniformly dense hypergraphs. *Journal of the London Mathematical Society*, 97(1):77–97, 2018.

[19] Christian Reiher, Vojtěch Rödl, and Mathias Schacht. On a Turán problem in weakly quasirandom 3-uniform hypergraphs. *Journal of the European Mathematical Society*, 20(5):1139–1159, 2018.

[20] Vojtěch Rödl. On universality of graphs with uniformly distributed edges. *Discrete Mathematics*, 59(1-2):125–134, 1986.

[21] Paul Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

A hypergraph bandwidth theorem

(Extended abstract)

Richard Lang* Nicolás Sanhueza-Matamala †

Abstract

A cornerstone of extremal graph theory due to Erdős and Stone states that the edge density which guarantees a fixed graph F as subgraph also asymptotically guarantees a blow-up of F as subgraph. It is natural to ask whether this phenomenon generalises to vertex-spanning structures such as Hamilton cycles. This was confirmed by Böttcher, Schacht and Taraz for graphs in the form of the Bandwidth Theorem. Our main result extends the phenomenon to hypergraphs.

A graph on n vertices that robustly contains a Hamilton cycle must satisfy certain conditions on space, connectivity and aperiodicity. Conversely, we show that if these properties are robustly satisfied, then all blow-ups of cycles on n vertices with clusters of size at most $\mathrm{poly}(\log\log n)$ are guaranteed as subgraphs. This generalises to powers of cycles and to the hypergraph setting.

As an application, we recover a series of classic results and recent breakthroughs on Hamiltonicity under degree conditions, which are then immediately upgraded to blown up versions. The proofs are based on a new setup for embedding large substructures into dense hypergraphs, which is of independent interest and does not rely on the Regularity Lemma or the Absorption Method.

1 Introduction

2 Introduction

An old problem of Turán is to determine the optimal edge density $\tau(F)$ that guarantees a copy of a k-uniform hypergraph F in a host hypergraph G. In the graph setting, Erdős and Stone [7] proved that under mildly stronger assumptions, one can even guarantee a *blow-up* of F. Formally, such a blow-up is formed by replacing the vertices of F with vertex-disjoint clusters and each edge with a complete k-partite k-graph whose parts are the corresponding

*Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona, Spain. E-mail: `richard.lang@upc.edu`. Supported by the Ramón y Cajal programme (RYC2022-038372-I) and by grant PID2023-147202NB-I00 funded by MICIU/AEI/10.13039/501100011033.

†Departamento de Ingeniería Matemática, Universidad de Concepción, Concepción, Chile. E-mail: `nsanhuezam@udec.cl`. Supported by ANID-FONDECYT Regular Nº1251121.

clusters. Erdős [6] extended this phenomenon to hypergraphs, showing that the presence of many copies of F already forces a blow-up of F with clusters of size $\operatorname{polylog} n$. Combining this with a supersaturation argument [8] extends the Erdős–Stone Theorem to hypergraphs, which is insofar remarkable that the value of $\tau(F)$ is generally unknown. Random constructions show that the cluster sizes in these blow-ups cannot exceed $\operatorname{polylog} n$. Nevertheless, an ongoing branch of research is dedicated to optimising the corresponding bounds [2, 4, 19].

So far, the order of the substructures has been tiny in comparison the host structure. It is then natural to ask whether the Erdős–Stone theorem generalises to much larger, even vertex-spanning substructures. To extend Turán-type problems in this direction, one typically replaces bounds on density with some condition on the degrees. For instance, the classic theorem of Dirac [5] provides optimal minimum degree conditions for the existence of a Hamilton cycle in an (ordinary) graph. This was later generalised to powers of cycles by Komlós, Sárközy and Szemerédi [12], which can be viewed as a combination of Dirac's and Turán's theorem. Against the backdrop of these advances, Bollobás and Komlós [13] asked whether one can find a vertex-spanning blow-up of a power of a cycle under marginally stronger degree conditions. This vertex-spanning extension of the Erdős–Stone theorem was confirmed by Böttcher, Schacht and Taraz [3] by proving the Bandwidth Theorem.

Our main result presents a solution to this question for hypergraphs. To state Dirac-type results for hypergraphs, we introduce some terminology. For $1 \leqslant d < k$, the *minimum d-degree* $\delta_d(G)$ of a k-uniform hypergraph G is the maximum m such that every set of d vertices is in at least m edges. The $(t-k+1)$*st power of a Hamilton cycle* $C \subseteq G$ has a cyclical vertex ordering such that every set of t cyclically consecutive vertices forms a clique $K_t^{(k)}$. Given this, the *threshold* $\delta_d^{\mathsf{HC}}(k,t)$ is the infimum $\delta \in [0,1]$ such for every $\varepsilon > 0$, there is n_0 such every k-graph G on $n \geqslant n_0$ vertices with $\delta_d(G) \geqslant (\delta+\varepsilon)\binom{n-d}{k-d}$ contains the $(t-k+1)$st power of a Hamilton cycle. For instance, Dirac's theorem implies that $\delta_1(2,2) = 1/2$ and the above-mentioned result of Komlós, Sárközy and Szemerédi implies that $\delta_1(2,t) = 1 - 1/t$. In general, we can view $\delta_d^{\mathsf{HC}}(k,t)$ as the analogue of $\tau(K_t^{(k)})$ for Hamilton cycles. As with the latter, little is known about its value except for the cases $k - d \leqslant 3$ as summarised in Theorem 3.1. Given this, we can formalise the aforementioned result of Böttcher, Schacht and Taraz [3]:

Theorem 2.1 (Bandwidth Theorem)**.** *For every $2 \leqslant k \leqslant t$ and $\varepsilon > 0$, there exist c and n_0 with the following properties. Let G be a graph on $n \geqslant n_0$ vertices with*

$$\delta_1(G) \geqslant \big(\delta_1^{\mathsf{HC}}(2,t) + \varepsilon\big)n.$$

Let H be a subgraph of an n-vertex blow-up of the $(t-1)$st power of a cycle with clusters of size at most cn and maximum degree $\Delta(H) \leqslant \Delta$. Then $H \subseteq G$.

The conjecture of Bollobás and Komlós [13] and its solution were originally stated in terms of the bandwidth parameter, hence the name of the theorem. It is not hard to see that Theorem 2.1 implies this formulation [16, Lemma 10.1].

In the spirit of the Erdős–Stone problem, our goal is to state a hypergraph bandwidth theorem in terms of $\delta_d^{\mathsf{HC}}(k,t)$. However, in addition to the existence of Hamilton cycles, one also needs a certain guarantee of consistency with regards to their location when working in the hypergraph setting. In the full version of this paper [18], we define a threshold $\delta_d^{\mathsf{HF}}(k,t)$ for *Hamilton frameworks* that captures this phenomenon. Moreover, we prove that $\delta_d^{\mathsf{HC}}(k,t) = \delta_d^{\mathsf{HF}}(k,t)$ in all cases where $\delta_d^{\mathsf{HC}}(k,t)$ is known (see Theorem 3.4). Given this, we can formulate a simplified version of our main result.

Theorem 2.2. *For every $1 \leqslant d < k \leqslant t$ and $\varepsilon > 0$, there exist c and n_0 with the following properties. Let G be a k-uniform hypergraph on $n \geqslant n_0$ vertices with*

$$\delta_d(G) \geqslant \left(\delta_d^{\mathsf{HF}}(k,t) + \varepsilon\right)\binom{n-k}{k-d}.$$

Let H be an n-vertex blow-up of the $(t-k+1)$st power of a k-uniform cycle with clusters of (not necessarily uniform) size at most $(\log\log n)^c$. Then $H \subseteq G$.

We conclude this overview with a summary of our main outcomes:

- The threshold $\delta_d^{\mathsf{HF}}(k,t)$ is conceptually simpler than $\delta_d^{\mathsf{HC}}(k,t)$ and thus easier to determine. Plain powers of Hamilton cycles can be recovered by taking clusters of size one in Theorem 2.2. So in particular $\delta_d^{\mathsf{HC}}(k,t) \leqslant \delta_d^{\mathsf{HF}}(k,t)$.
- Theorem 2.2 is a consequence of a more general result where the requirement on the minimum degree is replaced by a more general condition based on subgraph counts. This corresponds to Erdős' theorem [6], where a blow-up of F is forced via the presence of many copies of F instead of the Turán density $\tau(F)$.
- We also prove a Hamilton-connectedness result, which allows to find Hamilton paths with some prescribed end-tuples. Together with our former work with Joos [10] this leads to counting and random-robust Dirac-type theorems.
- Our proofs do not rely on the Regularity Lemma, the Blow-up Lemma or the Absorption Method, which are the prevalent techniques in the field. Instead, we develop a new setup which uses 'blow-up covers' to tackle these type of problems.

For a gentle introduction to the proof ideas, we also refer to our related work [17], where a similar approach is used to find blow-ups with optimal cluster sizes in the more basic setting of Hamilton cycles in graphs with large minimum degree.

3 Further background and applications

The study of Hamiltonicity under degree conditions dates back to Dirac [5], who proved that a graph G on $n \geqslant 3$ vertices with minimum degree $\delta_1(G) \geqslant n/2$ contains a Hamilton cycle. In the hypergraph setting, Hamilton cycles (as defined above with $t = k$) where investigated first by Katona and Kierstead [11] and later in the seminal contribution of Rödl, Ruciński and Szemerédi [24]. Over the past two decades, significant efforts have been directed towards extending their work [9, 15, 21, 22, 23, 24], and the minimum degree threshold for Hamilton cycles has been determined in the following range:

Theorem 3.1. *We have $\delta_{k-1}^{\mathsf{HC}}(k,k) = 1/2$, $\delta_{k-2}^{\mathsf{HC}}(k,k) = 5/9$ and $\delta_{k-3}^{\mathsf{HC}}(k,k) = 5/8$.*

Powers of Hamilton cycles for hypergraphs were first studied by Bedenknecht and Reiher [1] for 3-uniform graphs. More recently, Pavez-Signé, Sanhueza-Matamala and Stein [20] generalised their work to higher uniformities. Thus far, research has focused on the case of codegrees. For $k \leqslant t$, set $f_k(t) = 1 - 1/\left(\binom{t-1}{k-1} + \binom{t-2}{k-2}\right)$.

Theorem 3.2 ([1, 20])**.** *We have $\delta_{k-1}^{\mathsf{HC}}(k,t) \leqslant f_k(t)$ for every $2 \leqslant k \leqslant t$.*

As discussed before, bandwidth theorems are relatively well-understood in the graph setting. For hypergraphs however, no equivalent (or even partial) results have been obtained, to the best of our knowledge. We believe that vertex-spanning blow-ups of powers of cycles appear under the same (asymptotic) minimum degree conditions as powers of Hamilton cycles. Given Theorem 2.2, this can be formalised as follows.

Conjecture 3.3. *For all $1 \leqslant d < k \leqslant t$, we have $\delta_d^{\mathsf{HF}}(k,t) = \delta_d^{\mathsf{HC}}(k,t)$.*

We confirm Conjecture 3.3 in all cases where $\delta_d^{\mathsf{HC}}(k,t)$ is known (see Theorem 3.1).

Theorem 3.4. *We have $\delta_{k-1}^{\mathsf{HF}}(k) = 1/2$, $\delta_{k-2}^{\mathsf{HF}}(k) = 5/9$ and $\delta_{k-3}^{\mathsf{HF}}(k) = 5/8$.*

It is worth noting that Theorem 3.4 does not rely on the technical setup of our past work [15]. Instead, we integrate the structural ideas from the proofs of Theorem 3.1 into our framework to give new much cleaner proofs. Finally, we also extend Theorem 3.2 to blow-ups as follows.

Theorem 3.5. *We have $\delta_{k-1}^{\mathsf{HF}}(k,t) \leqslant f_k(t)$ for every $2 \leqslant k \leqslant t$.*

In summary, we recover all known results on Hamiltonicity and extend them to the bandwidth setting, which allows us to find vertex-spanning blow-ups of powers of cycles.

4 Methodology

There are two prevalent approaches for the embedding for spanning structures in graphs and hypergraphs. The first approach combines Szemerédi's Regularity Lemma with the Blow-up Lemma of to Komlós, Sárközy and Szemerédi [12]. This framework was key in the proof of the Bandwidth Theorem (Theorem 2.1). The second approach is based on the Absorption Method often also in conjunction with a (hypergraph) Regularity Lemma. This setup was introduced by Rödl, Ruciński and Szemerédi [24] to find Hamilton cycles under codegree conditions (meaning $k - d = 1$) and has formed the basis of all contributions in this direction thereafter.

Both methods come with advantages and drawbacks. The combination of the Regularity and Blow-up Lemma offers a systematic way to tackle embedding problems for dense graphs. However, this framework also comes with many technical challenges, which usually leads to long and convoluted proofs [16]. This phenomenon is most dramatic for (proper) hypergraphs, where even stating the definitions requires effort and attention to detail. In this setting, we are moreover still lacking some the basic instruments such as a Blow-up Lemma with image restrictions. The Absorption Method, on the other hand, avoids many of these technicalities and has therefore become the approach of choice in the hypergraph setting. It is particularly well-suited for the embedding of spanning structures that exhibit strong symmetries. Absorption for less symmetrical structures becomes more difficult and it is not well-understood whether it works at all. For instance, there is so far no proof of the Bandwidth Theorem using the Absorption Method. Finally, we note that neither of these methods works systematically well with graphs of unbounded maximum degree.

To overcome these limitations, we propose a new framework for embedding large structures into dense hypergraphs. Our approach is best explained in the context of a Blow-up Lemma, which can be summarised as follows. Broadly speaking, the Regularity Lemma allows us to approximate a given host-graph G with a quasirandom blow-up of a *reduced graph R* of constant order. So the vertices of R are replaced by (disjoint, linear-sized) vertex clusters and the edges of R are replaced by quasirandom partite graphs. Importantly, the Regularity Lemma guarantees that R approximately inherits structural properties of G such as degree conditions. The Blow-up Lemma then tells us that one can effectively assume that these partite graphs are complete. Ignoring a number of technicalities, this reduces the problem of embedding a guest graph H into G to embedding H into a complete blow-up of R. This last step is known as the *allocation* of H, and there is a well-tested set of techniques available to facilitate this step. Therefore, the reduction to the allocation problem often presents considerable step towards the successful embedding of H into G.

In contrast to this, our method proceeds as follows. Instead of approximating the entire structure of G with a quasirandom blow-up, we only cover the vertex set of G with a family of complete blow-ups $R_1(\mathcal{V}_1), \ldots, R_\ell(\mathcal{V}_\ell)$ that are interlocked in a cycle-like fashion. As before, the *reduced graphs* $R_1, \ldots, R_\ell$ inherit structural properties of G such as degree conditions. Now suppose we are given a blow-up of a cycle H. We subdivide H into blow-ups of paths $H_1, \ldots, H_\ell$, where each H_i has approximately $|\bigcup \mathcal{V}_i|$ vertices. This effectively reduces the problem of embedding H into G to embedding each H_i into the blow-up $R_i(\mathcal{V}_i)$, which is analogous to the above detailed allocation problem.

The connectivity, space and divisibility properties guaranteed by the threshold $\delta_d^{\mathsf{HF}}(k,t)$ (or a more general set-up) allow us to carry out the allocation into each $R_i(\mathcal{V}_i)$ by combining arguments from our past work [14, 16] with several new ideas. Moreover, the consistency property of $\delta_d^{\mathsf{HF}}(k,t)$ then guarantees that the allocations for $R_i(\mathcal{V}_i)$ and $R_{i+1}(\mathcal{V}_{i+1})$ are properly synchronised.

References

[1] W. Bedenknecht and Chr. Reiher. Squares of Hamiltonian cycles in 3-uniform hypergraphs. *Random Structures Algorithms*, 56(2):339–372, 2020.

[2] B. Bollobás and P. Erdős. On the structure of edge graphs. *Bull. London Math. Soc*, 5(317-321):31, 1973.

[3] J. Böttcher, M. Schacht, and A. Taraz. Proof of the bandwidth conjecture of Bollobás and Komlós. *Math. Ann.*, 343(1):175–205, 2009.

[4] V. Chvátal and E. Szemerédi. On the Erdős–Stone theorem. *J. London Math. Soc.(2)*, 23(2):207–214, 1981.

[5] G. A. Dirac. Some theorems on abstract graphs. *Proc. Lond. Math. Soc.*, 3(1):69–81, 1952.

[6] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.*, 2:183–190, 1964.

[7] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[8] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.

[9] J. Han and Y. Zhao. Forbidding Hamilton cycles in uniform hypergraphs. *J. Combin. Theory Ser. A*, 143:107–115, 2016.

[10] F. Joos, R. Lang, and N. Sanhueza-Matamala. Robust Hamiltonicity. *arXiv:2312.15262*, 2023.

[11] G. Y. Katona and H. A. Kierstead. Hamiltonian chains in hypergraphs. *J. Graph Theory*, 30(3):205–212, 1999.

[12] J. Komlós, G. N. Sárközy, and E. Szemerédi. On the Pósa-Seymour conjecture. *J. Graph Theory*, 29(3):167–176, 1998.

[13] J. Komlós and M. Simonovits. Szemerédi's Regularity Lemma and its applications in graph theory. In D. Miklós, V. T. Sós, and T. Szőnyi, editors, *Combinatorics, Paul Erdős is Eighty*, volume 2, pages 295–352. Bolyai Society Mathematical Studies, 1996.

[14] R. Lang. Tiling dense hypergraphs. *arXiv:2308.12281*, 2023.

[15] R. Lang and N. Sanhueza-Matamala. Minimum degree conditions for tight Hamilton cycles. *J. Lond. Math. Soc.*, 105(4):2249–2323, 2022.

[16] R. Lang and N. Sanhueza-Matamala. On sufficient conditions for spanning structures in dense graphs. *Proc. Lond. Math. Soc.*, 127(3):709–791, 2023.

[17] R. Lang and N. Sanhueza-Matamala. Blowing up Dirac's theorem. *arXiv:2412.19912*, 2024.

[18] R. Lang and N. Sanhueza-Matamala. A hypergraph bandwidth theorem. *arXiv:2412.14891*, 2024.

[19] V. Nikiforov. Complete r-partite subgraphs of dense r-graphs. *Discrete Math.*, 309(13):4326–4331, 2009.

[20] M. Pavez-Signé, N. Sanhueza-Matamala, and M. Stein. Towards a hypergraph version of the Pósa–Seymour conjecture. *Adv. Comb.*, 7 2023.

[21] J. Polcyn, Chr. Reiher, V. Rödl, A. Ruciński, M. Schacht, and B. Schülke. Minimum pair-degee for tight Hamiltonian cycles in 4-uniform hypergraphs. *Acta Mathematica Universitatis Comenianae*, 88(3):1023–1027, 2019.

[22] J. Polcyn, Chr. Reiher, V. Rödl, and B. Schülke. On Hamiltonian cycles in hypergraphs with dense link graphs. *J. Combin. Theory Ser. B*, 150:17–75, 2021.

[23] Chr. Reiher, V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi. Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs. *Proc. Lond. Math. Soc.*, 119(2):409–439, 2019.

[24] V. Rödl, A. Ruciński, and E. Szemerédi. An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The rank partition of a matroid and partial symmetries on tensors

(Extended abstract)

Inês Legatheaux Martins*

Abstract

The rank partition of a matroid is an enumerative invariant which measures how close a matroid is to being a union of bases. In this paper, we introduce partial symmetry classes of tensors associated with the rook monoid and we give new applications of the rank partition in the context of the annulment of a partially symmetrized tensor.

1 Introduction

In 1990, J. A. Dias da Silva [8] considered the problem of partitioning a matroid $\mathcal{M}$ into independent sets of prescribed sizes. His approach led to the definition of a matroid invariant which measures how close $\mathcal{M}$ is to being a union of bases. The *rank partition* of $\mathcal{M}$ is the sequence $\rho_{\mathcal{M}} = (\rho_1, \rho_2, \ldots)$ defined by requiring that, for each $j \geq 1$, the partial sum $\rho_1 + \rho_2 + \ldots + \rho_j$ is the maximum size of a union of j independent sets in $\mathcal{M}$. It is a partition of the number of non-loop elements of $\mathcal{M}$ and it has numerous applications [4, 6, 7, 9, 10, 11, 18].

The work in this article originated in an attempt to generalize an interpretation of $\rho_{\mathcal{M}}$ which was already present in [8]. Henceforth, n is a fixed positive integer, $\mathbb{F}$ is a field of characteristic zero and V is a d-dimensional $\mathbb{F}$-space such that $d \geq n$.

It is well known that the symmetric group S_n acts by place permutations on the nth tensor power of V, herein denoted by $\otimes^n V$. If λ is a partition of n, let χ_λ be the corresponding irreducible character of S_n. The projection operator $\pi_\lambda : \otimes^n V \longrightarrow \otimes^n V$ onto the isotypic component of $\otimes^n V$ indexed by χ_λ is given by

$$\pi_\lambda(v_1 \otimes \ldots \otimes v_n) = \frac{\chi_\lambda(1_{\mathbf{n}})}{n!} \sum_{\sigma \in S_n} \chi_\lambda(\sigma) v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}\,, \text{ for } v_1, \ldots, v_n \in V. \tag{1}$$

*Faculty of Sciences of the University of Lisbon, Campo Grande, Edifício C6, Piso 2, 1749-016, Lisboa, Portugal. E-mail: ilegatheaux@gmail.com. This research was made within the activities of the Center for Mathematical Studies (University of Lisbon, Portugal) and was partially supported by the Portuguese Science Foundation (FCT) through the Strategic Project $UID/04561/2025$.

The image of π_λ is a *symmetry class of tensors.* If $v_1, \ldots, v_n \in V$, an element of $\otimes^n V$ of the form $\pi_\lambda(v_1 \otimes \ldots \otimes v_n)$ is called a *symmetrized decomposable tensor.*

Symmetry classes of tensors arise in a wide variety of areas [2, 22, 23, 24, 27, 31]. In this context, the problem of finding conditions for a symmetrized tensor to be zero is critical. The question was settled by C. Gamas [19] and J. A. Dias da Silva [8] showed that Gamas's result can be rephrased in terms of the rank partition.

Theorem 1.1 (Dias da Silva)**.** Let $\mathbf{v} = (v_1, \ldots, v_n)$ be a family of nonzero vectors in V which realizes the matroid $\mathcal{M}(\mathbf{v})$ over $\mathbb{F}$. Let λ be a partition of n and let λ' be its conjugate. If $\leq_d$ is the dominance order on partitions of n, then

$$\pi_\lambda(v_1 \otimes \ldots \otimes v_n) \neq 0 \quad \Leftrightarrow \quad \lambda' \leq_d \rho_{\mathcal{M}(\mathbf{v})}\,.$$

Since 2009, A. Berget [3, 4, 5] made use of the deep interactions between the representation theories of S_n and $GL(V)$ to derive concise proofs of classical results involving symmetrized tensors. His work led to another viewpoint on the rank partition. Let $GL(V)$ act diagonally on $\otimes^n V$. If $\mathbf{v} = (v_1, \ldots, v_n)$ is a family of nonzero vectors in V and $\mathbf{v}^{\otimes} = v_1 \otimes \ldots \otimes v_n \in \otimes^n V$, let $G(\mathbf{v}^{\otimes})$ be the cyclic $\mathbb{F}GL(V)$-module generated by $\mathbf{v}^{\otimes}$. If λ is a partition of n, Schur–Weyl duality says that $G(\mathbf{v}^{\otimes})$ contains an isomorphic copy of a simple $\mathbb{F}GL(V)$-module indexed by λ if and only if $\pi_\lambda(\mathbf{v}^{\otimes}) \neq 0$ and thus, by Theorem 1.1, if and only if $\lambda' \leq_d \rho_{\mathcal{M}(\mathbf{v})}$. Similarly, if $S(\mathbf{v}^{\otimes})$ is the $\mathbb{F}S_n$-module generated by $\mathbf{v}^{\otimes}$, the decomposition of $S(\mathbf{v}^{\otimes})$ into simple submodules is controlled by $\rho_{\mathcal{M}(\mathbf{v})}$.

The archetypal structure when it comes to extend the principle of symmetry is the *rook monoid* (also called the *symmetric inverse monoid* [21, 30]). Denoted by R_n, it is the monoid of all bijective partial maps from $\mathbf{n} = \{1, \ldots, n\}$ to itself. It contains S_n and it is isomorphic to the monoid under matrix multiplication of all $n \times n$ matrices with at most one entry equal to 1 in each row and column and zeros elsewhere. The latter first appeared in [28] in the context of the Iwahori–Hecke algebra for the monoid of all $n \times n$ matrices over a finite field.

In 2002, L. Solomon [29] established a Schur–Weyl duality type result between $GL(V)$ and R_n, in which R_n acts on tensors by "partial" place permutations. This influential result allowed us to lay the foundations of a natural generalization of the theory of symmetry class of tensors. As such, *partial symmetry classes of tensors* and *partially symmetrized decomposable tensors* are introduced in Section 3. The main goal of this paper is to characterise the rôle of the rank partition in the context of the annulment problem for partially symmetrized tensors. Section 3 also contains several new combinatorial results on this elegant matroid invariant.

2 Background

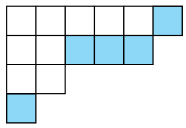
2.1 Partitions and diagrams

A *partition μ of r*, denoted by $\mu \vdash r$, is a sequence $(\mu_1, \mu_2, \ldots)$ of non-negative decreasing integers $\mu_1 \geq \mu_2 \geq \ldots$ such that $\sum_{i \geq 1} \mu_i = r$. The number of nonzero parts of μ, written $l(\mu)$, is the *length* of μ. We agree that there is a partition of 0, represented by (0).

If $\mu \vdash r$, the *shape of* μ is $[\mu] = \{(i, j) : i \geq 1 \text{ and } 1 \leq j \leq \mu_i\}$. As usual, $[\mu]$ can be depicted as an array of r square boxes arranged in $l(\mu)$ left-justified rows such that the ith row contains exactly μ_i boxes. The *conjugate* of μ is the sequence $\mu' = (\mu'_1, \mu'_2, \ldots)$, where each μ'_j corresponds to the number of boxes in the jth column of $[\mu]$. Clearly, $\mu' \vdash r$. For completeness, $[(0)]$ is identified with $\emptyset$.

Given $\lambda = (\lambda_1, \lambda_2, \ldots) \vdash n$ and $\mu = (\mu_1, \mu_2, \ldots) \vdash r$, we write $\mu \subseteq \lambda$ to indicate that $\mu_i \leq \lambda_i$, for all $i \geq 1$. Note that $\mu \subseteq \lambda$ implies that $l(\mu) \leq l(\lambda)$. We refer to the sequence $(\lambda_1 - \mu_1, \lambda_2 - \mu_2, \ldots)$ as the *skew partition* λ/μ. The *skew diagram of* λ/μ, denoted by $[\lambda/\mu]$, corresponds to the array of $n - r$ boxes that appear in $[\lambda]$ but not in $[\mu]$. For instance, if $\lambda = (6, 5, 2, 1) \vdash 14$ and $\mu = (5, 2^2) \vdash 9$, then $[\lambda/\mu]$ is given by the shaded region in

$[\lambda/\mu]$:



If $\mu \subseteq \lambda$, the skew partition λ/μ is a *horizontal strip* if $\lambda'_j - \mu'_j \leq 1$, for all $j \geq 1$. In other words, λ/μ is a horizontal strip if $[\lambda/\mu]$ has at most one box in each column. We also adopt the convention that (0) is a horizontal strip.

2.2 The rank partition of a matroid

A *matroid* on the finite set S is a pair $\mathcal{M} = (S, \mathcal{I})$, where $\mathcal{I}$ is a non-empty hereditary collection of subsets of S, called *independent sets*, which satisfies the following condition:

- for every $I, J \in \mathcal{I}$, if $|I| < |J|$, there is some $x \in I \setminus J$ such that $J \cup \{x\} \in \mathcal{I}$.

A *basis* of $\mathcal{M}$ is a maximal independent set contained in S. It is known that all bases have the same size. This number is called the *rank of* $\mathcal{M}$ and denoted by $r_{\mathcal{M}}$.

All matroids considered in this paper are loopless and thus we shall not mention this restriction in the sequel. If $\mathbf{v} = (v_1, \ldots, v_n)$ is a family of nonzero vectors in V, we write $\mathcal{M}(\mathbf{v})$ for the matroid on $\mathbf{n} = \{1, \ldots, n\}$ whose independent sets are those $I \subseteq \mathbf{n}$ such that $\{v_i : i \in I\}$ is linearly independent in V. In this context, $\mathcal{M}(\mathbf{v})$ is called a *vector matroid* and $\mathbf{v}$ is said to *realize* $\mathcal{M}(\mathbf{v})$ *over* $\mathbb{F}$.

Covering and packing problems arise naturally in matroid theory [12, 13, 14] and have numerous applications [17, 20, 25, 32, 34]. Given a matroid $\mathcal{M}$ on S, the following problem was considered in [8]: if $\mu = (\mu_1, \mu_2, \ldots) \vdash |S|$ with $l(\mu) = k$, do there exist pairwise disjoint independent sets $I_1, \ldots, I_k$ in $\mathcal{M}$ such that

$$|I_j| = \mu_j, \text{ for } j = 1, \ldots, k \text{ and } S = I_1 \cup \ldots \cup I_k\,? \tag{2}$$

If a family $(I_1, \ldots, I_k)$ of independent sets of $\mathcal{M}$ satisfies (2), it is known as a *μ-coloring of* $\mathcal{M}$. Determining if μ-colorings exist leads to the following definition.

Definition 2.1. Let $\mathcal{M} = (S, \mathcal{I})$ be a matroid. The *rank partition of* $\mathcal{M}$ is the sequence of integers $\rho_{\mathcal{M}} = (\rho_1, \rho_2, \ldots)$ defined by the condition that, for each $j \geq 1$,

$$\rho_1 + \rho_2 + \ldots + \rho_j = \max\{|I| : I = I_1 \cup \ldots \cup I_j,\ I_i \in \mathcal{I},\ \text{for } i = 1, \ldots, j\}. \tag{3}$$

The proof that $\rho_{\mathcal{M}} \vdash |S|$ is by no means obvious [8, p. 27]. The first part of $\rho_{\mathcal{M}}$ is $r_{\mathcal{M}}$ and $l(\rho_{\mathcal{M}})$ is the covering number of $\mathcal{M}$. The next result can be found in [8, Theorem 1].

Theorem 2.2 (Dias da Silva)**.** Let $\mathcal{M}$ be a matroid on S and let $\mu = (\mu_1, \mu_2, \ldots)$ be a partition of $|S|$. There exists a μ-coloring of $\mathcal{M}$ if and only if $\mu \leq_d \rho_{\mathcal{M}}$.

2.3 The rook monoid

We recall that R_n is the monoid of all bijective maps σ with domain $\mathcal{D}(\sigma) \subseteq \mathbf{n}$ and range $\mathcal{R}(\sigma) \subseteq \mathbf{n}$ endowed with the usual product of partial maps. We agree that R_n contains a map $\epsilon_\emptyset$ with empty domain and range which behaves as a zero element. With this convention, $|R_n| = \sum_{r=0}^{n} \binom{n}{r}^2 r!$.

Note that any $\sigma \in R_n$ such that $D(\sigma) = R(\sigma) = \mathbf{n}$ is a permutation of $\mathbf{n}$ and thus $S_n \subseteq R_n$. A minute's thought reveals that $S_r \subseteq R_n$, for $r = 0, 1, \ldots, n$, where $S_0 = \{\epsilon_\emptyset\}$ is a group with a single element. On the other hand, the set of idempotents of R_n is the monoid of all partial identities $\epsilon_X : X \to X$, with $X \subseteq \mathbf{n}$.

The following decomposition of the monoid algebra $\mathbb{F}R_n$ is known since the $1950's$. However, it was first written explicitly by L. Solomon [29, Lemma 2.17].

Theorem 2.3. There exists an isomorphism of $\mathbb{F}$-algebras

$$\mathbb{F}R_n \cong \bigoplus_{r=0}^{n} \mathcal{M}_{\binom{n}{r}}(\mathbb{F}S_r),$$

where, for $1 \leq r \leq n$, $\mathcal{M}_{\binom{n}{r}}(\mathbb{F}S_r)$ is the $\mathbb{F}$-algebra of all matrices with rows and columns indexed by subsets of $\mathbf{n}$ of size r and entries in $\mathbb{F}S_r$, and $\mathcal{M}_{\binom{n}{0}}(\mathbb{F}S_0) = \mathbb{F}$.

Theorem 2.3 implies that $\mathbb{F}R_n$ is a finite-dimensional (split) semisimple $\mathbb{F}$-algebra and that the isomorphisms classes of simple $\mathbb{F}R_n$-modules are in one-to-one correspondence with those of the group algebras $\mathbb{F}S_r$, for $0 \leq r \leq n$. As such, let $\mathcal{I}(R_n) = \bigcup_{0 \leq r \leq n} \{R_\mu \ : \ \mu \vdash r\}$ be a full set of simple $\mathbb{F}R_n$-modules. If $0 \leq r \leq n$ and $\mu \vdash r$, we also write $\widehat{\chi_\mu}$ for the corresponding irreducible character of R_n.

We end this section by giving a new description of the central primitive idempotents of $\mathbb{F}R_n$. If $X \subseteq \mathbf{n}$, let η_X be the element of $\mathbb{F}R_n$ defined by

$$\eta_X = \sum_{Y \subseteq X} (-1)^{|X|-|Y|} \epsilon_Y. \tag{4}$$

L. Solomon [29, Lemma 2.6] proved that the η_X are pairwise orthogonal idempotents of $\mathbb{F}R_n$. Combined with Theorem 2.3, this allowed us to establish the next result.

Theorem 2.4. The central primitive idempotent of $\mathbb{F}R_n$ corresponding to (0) is $\widehat{\epsilon_{(0)}} = \epsilon_\emptyset$. If $1 \leq r \leq n$ and $\mu \vdash r$, the central primitive idempotent indexed by μ is

$$\widehat{\epsilon_\mu} = \sum_{\substack{X \subseteq \mathbf{n}, \\ |X| = r}} \left(\frac{\widehat{\chi_\mu}(\epsilon_X)}{r!} \sum_{w \in S_X} \widehat{\chi_\mu}(w\epsilon_X) w \eta_X \right),$$

where, for each $X \subseteq \mathbf{n}$ such that $|X| = r$, $S_X = \{w \in S_n : w(t) = t, \text{ if } t \notin X\}$.

3 The rank partition and partially symmetrized tensors

We are now ready to introduce a generalization of symmetrized tensors associated with the rook monoid. As before, let V be a d-dimensional $\mathbb{F}$-space such that $d \geq n$. Let $U = U_1 \oplus U_0$ be such that $U_1 = V$ and $\dim_{\mathbb{F}}(U_0) = 1$ and, for each $i \in \{0, 1\}$, let $\mathbf{p}_i : U \to U$ be the projection of U onto U_i which annihilates U_j for $j \neq i$. If $\sigma \in R_n$, we may write $\sigma = w\epsilon_{\mathcal{D}(\sigma)}$ for some $w \in S_n$. If $u_1, \dots, u_n \in U$, define

$$(u_1 \otimes \dots \otimes u_n) \cdot \sigma = \mathbf{p}_{i_1}\left(u_{w(1)}\right) \otimes \dots \otimes \mathbf{p}_{i_n}\left(u_{w(n)}\right), \tag{5}$$

where $\mathbf{p}_{i_t}\left(u_{w(t)}\right) = \mathbf{p}_0\left(u_{w(t)}\right)$ if $t \notin \mathcal{D}(\sigma)$, and $\mathbf{p}_{i_t}\left(u_{w(t)}\right) = \mathbf{p}_1\left(u_{w(t)}\right)$ otherwise. It can be shown that (5) does not depend on the representation of $\sigma = w\epsilon_{\mathcal{D}(\sigma)}$, with $w \in S_n$, and that it turns $\otimes^n U$ into a right $\mathbb{F}R_n$-module. On the other hand, $GL(V)$ acts on U by fixing U_0 and diagonally on $\otimes^n U$ and this action commutes with that of R_n. The proof of the next result can be found in [29, Theorem 5.10].

Theorem 3.1 (Solomon)**.** Let $U = V \oplus U_0$ be the direct sum of the natural module and the trivial module for $GL(V)$. Then $GL(V)$ and $\mathbb{F}R_n$ generate full centralizers of each other on $\otimes^n U$ and, provided that $\dim_{\mathbb{F}}(V) \geq n$, $\mathrm{End}_{GL(V)}(\otimes^n U) \cong \mathbb{F}R_n$.

Note that Theorem 3.1 has been translated in the language of Schur algebras [1]. Matroids also arise naturally in this context. If $\mathbf{u} = (u_1, \dots, u_n)$ is a family of nonzero vectors in U, then $\mathbf{u}$ realizes $\mathcal{M}(\mathbf{u})$ over $\mathbb{F}$. As it will soon become apparent, the combinatorics of $\mathcal{M}(\mathbf{u})$ are linked to the decomposition of the $GL(V)$-module generated by $\mathbf{u}^{\otimes} = u_1 \otimes \dots \otimes u_n$, denoted by $G(\mathbf{u}^{\otimes})$. By Theorem 3.1, similar results apply to $R(\mathbf{u}^{\otimes})$, the $\mathbb{F}R_n$-module spanned by $\mathbf{u}^{\otimes}$. If $R(\mathbf{u}^{\otimes})$ contains an isomorphic copy of R_μ for $\mu \vdash r$, we shall say that μ *occurs in* $R(\mathbf{u}^{\otimes})$.

The central primitive idempotents of a (split) semisimple algebra determine its isotypic components. Thus, if $0 \leq r \leq n$ and $\mu \vdash r$, we define $\widehat{\pi_\mu} : \otimes^n U \to \otimes^n U$ by $\widehat{\pi_\mu}(x) = x \cdot \widehat{\epsilon_\mu}$, for $x \in \mathbb{F}R_n$. The image of $\widehat{\pi_\mu}$, denoted by $\mathcal{U}_\mu(R_n)$, is a *partial symmetry class of tensors.* If $u_1, \dots, u_n \in U$, tensors of the form $\widehat{\pi_\mu}(u_1 \otimes \dots \otimes u_n)$ are called *partially symmetrized decomposable tensors.* An explicit formula for $\widehat{\pi_\mu}(u_1 \otimes \dots \otimes u_n)$ can be deduced from Theorem 2.4 and (5). Since it will play no rôle in the sequel, we omit the details.

If $r = n$ and $\lambda \vdash n$, it can be proved that $\mathcal{U}_\lambda(R_n) = \widehat{\pi_\mu}(\otimes^n U) = \pi_\lambda(\otimes^n V)$, where π_λ is given by (1). In particular, $\mathcal{U}_{(1^n)}(R_n) = \bigwedge^n V$, the nth exterior power of V, and $\mathcal{U}_{(n)}(R_n) = Sym^n(V)$, the nth symmetric power of V, realised in $\otimes^n U$.

If $\mu \vdash r$ with $0 \leq r \leq n$, then $\widehat{\pi_\mu} : \otimes^n U \to \otimes^n U$ is the projection operator onto the isotypic component of $\otimes^n U$ indexed by μ. By Schur–Weyl duality, we also have $\widehat{\pi_\mu} \in \mathrm{End}_{GL(V)}(\otimes^n U)$. Hence, the proof of the next result is straightforward.

Proposition 3.2. Let $\mathbf{u}^{\otimes} = u_1 \otimes \dots \otimes u_n \in \otimes^n U$, where $u_1, \dots, u_n$ are nonzero vectors in U. If $1 \leq r \leq n$ and $\mu \vdash r$, then μ occurs in $R(\mathbf{u}^{\otimes})$ if and only if μ occurs in $G(\mathbf{u}^{\otimes})$ if and only if $\widehat{\pi_\mu}(\mathbf{u}^{\otimes}) \neq 0$.

From now on, $\mathbf{u} = (u_1, \dots, u_n)$ is a fixed family of nonzero vectors in U and $\mathbf{u}^{\otimes} = u_1 \otimes \dots \otimes u_n$. As before, $\mathbf{u}$ realizes the matroid $\mathcal{M}(\mathbf{u})$ over $\mathbb{F}$. Our first result is an attempt to understand if the rank partition of $\mathcal{M}(\mathbf{u})$ plays the same rôle for partially symmetrized decomposable tensors as it does in the classical case.

Theorem 3.3. Let $1 \leq r \leq n$ and let $\mu \vdash r$. If $\widehat{\pi_\mu}(\mathbf{u}^{\otimes}) \neq 0$, then there is a partition $\lambda \vdash n$ such that $\mu \subseteq \lambda$ and

(*i*) the skew partition λ/μ is a horizontal strip,

(*ii*) $\lambda' \leq_d \rho_{\mathcal{M}(\mathbf{u})}$.

As an example, let $U = V \oplus U_0$ be such that $\dim_{\mathbb{F}}(V) = 7$ and $U_0 = \mathbb{F}e_\infty$, where $e_\infty \neq 0$. If $x, y \in V$ are linearly independent in V, let $u_1 = x + e_\infty$, $u_2 = y + e_\infty$, $u_3 = y + e_\infty$, $u_4 = x$ and $u_5 = y + e_\infty$. It is easily seen that $\rho_{\mathcal{M}(\mathbf{u})} = (3, 1^2) \vdash 5$. Theorem 3.3 implies that $\widehat{\pi_\mu}(\mathbf{u}^\otimes) = 0$, for all $\mu \in \{(1^4), (2^2, 1), (2, 1^3), (1^5)\}$.

For simplicity, let $\mathbf{p} = \mathbf{p}_1$ be the projection of U onto V which annihilates U_0. For each $1 \leq j \leq n$, let $v_j = \mathbf{p}(u_j) \in V$. We adopt the convention that there is at least one $v_j \neq 0$. Thus, there is $1 \leq t \leq n$ such that $\mathbf{v} = (v_1, \ldots, v_t)$ is a family of nonzero vectors in V (after reordering) that realizes $\mathcal{M}(\mathbf{v})$ over $\mathbb{F}$. Our main result is a complete characterization of when a partially symmetrized tensor is non null.

Theorem 3.4. Let $1 \leq r \leq n$ and let $\mu \vdash r$. The following are equivalent:

(*i*) $\widehat{\pi_\mu}(\mathbf{u}^\otimes) \neq 0$.

(*ii*) There is a partition $\lambda \vdash t$ such that $\mu \subseteq \lambda$ and $\lambda' \leq_d \rho_{\mathcal{M}(\mathbf{v})}$.

Let us consider the previous example once again. In such case, $v_1 = x$, $v_2 = y$, $v_3 = y$, $v_4 = x$ and $v_5 = y$ and thus $\mathbf{v} = (v_1, v_2, v_3, v_4, v_5)$ realizes $\mathcal{M}(\mathbf{v})$ over $\mathbb{F}$. It is clear that $\rho_{\mathcal{M}(\mathbf{v})} = (2, 2, 1) \vdash 5$. By Theorem 3.4, the decomposition of $R(\mathbf{u}^\otimes)$ is

$$\begin{aligned} R(\mathbf{u}^\otimes) \quad &\cong \quad d_0 R_{(0)} \oplus d_1 R_{\ydiagram{1}} \oplus d_2 R_{\ydiagram{2}} \oplus d_3 R_{\ydiagram{1,1}} \oplus d_4 R_{\ydiagram{3}} \oplus d_5 R_{\ydiagram{2,1}} \oplus d_6 R_{\ydiagram{4}} \\ &\oplus \quad d_7 R_{\ydiagram{3,1}} \oplus d_8 R_{\ydiagram{2,2}} \oplus d_9 R_{\ydiagram{3,2}} \oplus d_{10} d_8 R_{\ydiagram{4,1}} \end{aligned}$$

where the d_i are positive integers. The partitions that occur in $G(\mathbf{u}^\otimes)$ are precisely the same (with possibly distinct multiplicities). We also remark that $\rho_{\mathcal{M}(\mathbf{v})} \leq_d \rho_{\mathcal{M}(\mathbf{u})}$.

By now, it should be clear that rank partitions play an important part in the decomposition into simple modules of several important cyclic modules. On the other hand, the proofs of theorems 3.3 and 3.4 make use of a link between rank partitions, the dominance order and weak maps [33, p. 254]. As with other aspects of this paper, this suggests that the rank partition deserves even more study.

References

[1] Carlos A. M. André and Inês Legatheaux Martins. Schur-Weyl dualities for the rook monoid: an approach via Schur algebras. *Semigroup Forum*, 109(1):38–59, 2024.

[2] Natália Bebiano, Chi-Kwong Li, and João da Providência. Generalized numerical ranges of permanental compounds arising from quantum systems of bosons. *Electron. J. Linear Algebra*, 7:73–91, 2000.

[3] Andrew Berget. A short proof of Gamas's theorem. *Linear Algebra Appl.*, 430(2-3):791–794, 2009.

[4] Andrew Berget. Two results on the rank partition of a matroid. *Port. Math.*, 68(4):411–423, 2011.

[5] Andrew Berget. Equality of symmetrized tensors and the flag variety. *Linear Algebra Appl.*, 438(2):658–662, 2013.

[6] Andrew Berget, J. A. Dias da Silva, and Amélia Fonseca. Vanishing of doubly symmetrized tensors. *Electron. J. Combin.*, 20(2):Paper 60, 9, 2013.

[7] Richard A. Brualdi and Amélia Fonseca. On the quasi-transversals of a matroid. *Discrete Math.*, 146(1-3):293–297, 1995.

[8] J. A. Dias da Silva. On the μ-colorings of a matroid. *Linear and Multilinear Algebra*, 27(1):25–32, 1990.

[9] J. A. Dias da Silva and Rosário Fernandes. The rank partition and the covering number of the elements of the dual matroid. *Ars Combin.*, 59:129–144, 2001.

[10] J. A. Dias da Silva and Amélia Fonseca. The r-depth of a matroid. *Discrete Math.*, 141(1-3):85–93, 1995.

[11] J. A. Dias da Silva and Amélia Fonseca. Computing the rank partition and the flag transversal of a matroid. *Discrete Math.*, 216(1-3):85–110, 2000.

[12] Jack Edmonds. Lehman's switching game and a theorem of Tutte and Nash-Williams. *J. Res. Nat. Bur. Standards Sect. B*, 69B:73–77, 1965.

[13] Jack Edmonds. Minimum partition of a matroid into independent subsets. *J. Res. Nat. Bur. Standards Sect. B*, 69B:67–72, 1965.

[14] Jack Edmonds and D. R. Fulkerson. Transversals and matroid partition. *J. Res. Nat. Bur. Standards Sect. B*, 69B:147–153, 1965.

[15] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:51–57, 1966.

[16] Paul Erdős and Arthur H Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(12):1087–1091, 1946.

[17] Genghua Fan, Hongbi Jiang, Ping Li, Douglas B. West, Daqing Yang, and Xuding Zhu. Extensions of matroid covering and packing. *European J. Combin.*, 76:117–122, 2019.

[18] Rosário Fernandes. Rank partitions and covering numbers under small perturbations of an element. *Linear Multilinear Algebra*, 57(4):387–398, 2009.

[19] Carlos Gamas. Conditions for a symmetrized decomposable tensor to be zero. *Linear Algebra Appl.*, 108:83–119, 1988.

[20] Bernhard Korte and Jens Vygen. *Combinatorial optimization*, volume 21 of *Algorithms and Combinatorics*. Springer, Heidelberg, fifth edition, 2012. Theory and algorithms.

[21] Mark V. Lawson. *Inverse semigroups*. World Scientific Publishing Co. Inc., River Edge, NJ, 1998. The theory of partial symmetries.

[22] Chi-Kwong Li and Alexandru Zaharia. Induced operators on symmetry classes of tensors. *Trans. Amer. Math. Soc.*, 354(2):807–836 (electronic), 2002.

[23] Marvin Marcus and Henryk Minc. Generalized matrix functions. *Trans. Amer. Math. Soc.*, 116:316–329, 1965.

[24] Russell Merris. *Multilinear algebra*, volume 8 of *Algebra, Logic and Applications*. Gordon and Breach Science Publishers, Amsterdam, 1997.

[25] C. St. J. A. Nash-Williams. Edge-disjoint spanning trees of finite graphs. *J. London Math. Soc.*, 36:445–450, 1961.

[26] C. St. J. A. Nash-Williams. Decomposition of finite graphs into forests. *J. London Math. Soc.*, 39:12, 1964.

[27] Thomas H. Pate. Partitions, irreducible characters, and inequalities for generalized matrix functions. *Trans. Amer. Math. Soc.*, 325(2):875–894, 1991.

[28] Louis Solomon. The Bruhat decomposition, Tits system and Iwahori ring for the monoid of matrices over a finite field. *Geom. Dedicata*, 36(1):15–49, 1990.

[29] Louis Solomon. Representations of the rook monoid. *J. Algebra*, 256(2):309–342, 2002.

[30] Benjamin Steinberg. *Representation theory of finite monoids*. Universitext. Springer, Cham, 2016.

[31] Maria M. Torres and Pedro C. Silva. Tensors, matchings and codes. *Linear Algebra Appl.*, 436(6):1606–1622, 2012.

[32] W. T. Tutte. On the problem of decomposing a graph into n connected factors. *J. London Math. Soc.*, 36:221–230, 1961.

[33] Neil White, editor. *Theory of matroids*, volume 26 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1986.

[34] Daqing Yang. Decomposing a graph into forests and a matching. *J. Combin. Theory Ser. B*, 131:40–54, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Tight Hamilton Cycles in Hypergraphs with Dichotomous Degree Conditions

(Extended abstract)

Shoham Letzter*† Arjun Ranganathan*‡

Abstract

A *tight Hamilton cycle* in a k-uniform hypergraph (henceforth *k-graph*) is a cyclic ordering of its vertices such that every set of k consecutive vertices forms an edge. We study a variant of typical Dirac-type problems where, instead of requiring a universal minimum co-degree, we consider k-graphs where each $(k-1)$-set either supports nearly all of the vertices or does not support any vertices at all. Specifically, Illingworth, Lang, Müyesser, Parczyk and Sgueglia recently conjectured that, for all $k \geq 2$ and sufficiently large n, if G is an n-vertex k-graph with no isolated vertices such that each $(k-1)$-set contained in an edge is contained in at least $(1-1/k)n-(k-2)$ edges, then G has a tight Hamilton cycle. It is easy to show that this bound is best possible.

Our main contribution is a complete resolution of this conjecture. We use a novel blow-up tiling technique developed by Lang which avoids traditional approaches using the regularity lemma and the blow-up lemma. Our proof involves a stability analysis to distinguish between extremal and non-extremal cases.

1 Introduction

A seminal result of Dirac [2] states that any graph on $n \geq 3$ vertices and minimum degree at least $n/2$ contains a Hamilton cycle. Several natural extensions of Dirac's theorem have been studied over the years such as in digraphs [3], in random graphs [16], and rainbow versions [6], just to state a few. We will focus on hypergraph generalisations.

Formally, a *k-uniform hypergraph* or a *k-graph* G has of a set of vertices $V(G)$ and a set of edges $E(G)$, where each edge consists of exactly k vertices. Given a set of $k-1$ vertices $S \in \binom{V(G)}{k-1}$, we will say that *co-degree* of S is $d_{k-1}(S) = |\{v \in V(G) : S \cup \{v\} \in E(G)\}|$, and define the *minimum co-degree* of G to be $\delta_{k-1}(G) = \min\{d_{k-1}(S) : S \in \binom{V(G)}{k-1}\}$.

*Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK
†Email: `s.letzter@ucl.ac.uk`.
‡Email: `arjun.ranganathan.24@ucl.ac.uk`.

For any $\ell \in [k-1]$, we say that a k-graph is an ℓ-cycle if there is a cyclic ordering of its vertices such that every edge consists of k consecutive vertices and every pair of consecutive edges (with the ordering inherited from the vertex ordering) intersects in exactly ℓ vertices. A *Hamilton ℓ-cycle* in a k-graph G is an ℓ-cycle in G that spans $V(G)$. If $\ell = k-1$, we say it is a *tight Hamilton cycle*, and if $\ell = 1$, we call it a *loose Hamilton cycle*. Since each edge of a Hamilton ℓ-cycle contains $k-\ell$ vertices that were not in the previous edge, a necessary condition for the existence of such a cycle is that $k-\ell$ divides $|V(G)|$.

Katona and Kiearstead [7] conjectured that every k-graph with $\delta_{k-1} \geq \lfloor (n-k+3)/2 \rfloor$ has a tight Hamilton cycle, and showed that this is tight. Rödl, Ruciński and Szemerédi [19,21] first proved this asymptotically for $k \geq 3$ and $\delta_{k-1}(G) \geq n/2+o(n)$, and later obtained an exact result for $k=3$ [23]. Hàn and Schacht [4] showed that for $k \geq 3$ and $1 \leq \ell < k/2$, if $\delta_{k-1}(G) \geq n/(2k-2\ell)+o(n)$, then G has a Hamilton ℓ-cycle, and this was independently proved for $\ell = 1$ by Keevash, Kühn, Mycroft and Osthus [8]. Again, this bound is tight up to the error term. If $\ell \geq k/2$ and $k-\ell \mid k$, the best possible asymptotic bound is $n/2+o(n)$, which follows from [19,21] as a tight cycle ensures an ℓ-cycle (if the necessary divisibility conditions hold), and the lower bound arises from the existence of a perfect matching (see [14,18,22]). Conversely, when $k-\ell \nmid k$, Kühn, Mycroft and Osthus [12] showed that

$$\delta_{k-1}(G) \geq \frac{n}{\left\lceil \frac{k}{k-\ell} \right\rceil (k-\ell)} + o(n)$$

suffices and is optimal up to the $o(n)$ term, yielding a much lower threshold.

We study a recently proposed variant of these questions – we look for tight Hamilton cycles in hypergraphs where, for any $S \in \binom{V(G)}{k-1}$, either $d_{k-1}(S)$ is very large or it is zero. Formally, we define the *minimum supported co-degree* $\delta^*(G)$ of a k-graph G to be the maximum integer d such that every $(k-1)$-set contained in at least one edge of G is contained in at least d edges of G. As one would expect, the minimum supported co-degree necessary to guarantee a tight Hamilton cycle is significantly higher than simply $n/2+o(n)$. Indeed, consider an n-vertex k-graph $G_{\mathcal{E}}$ with $k \mid n$ with a partition of the vertex set $V(G_{\mathcal{E}}) = A \sqcup B$ such that $|A| = n/k+1$ and $G_{\mathcal{E}}$ contains every edge that has at most one vertex from A. It is straightforward to see that $\delta^*(G_{\mathcal{E}}) = (1-1/k)n-(k-1)$. However, since any matching in $G_{\mathcal{E}}$ misses at least one vertex in A, clearly $G_{\mathcal{E}}$ cannot contain a perfect matching, and hence does not have a tight Hamilton cycle. More fundamentally, $G_{\mathcal{E}}$ cannot contain a *perfect fractional matching*, which is defined as an edge weighting function $w : E(G_{\mathcal{E}}) \to [0,1]$ such that for each $v \in V(G_{\mathcal{E}})$, the sum of the weights of edges incident to v is exactly one. Clearly, a perfect matching is just a special case where $w(e) \in \{0,1\}$ for all edges e. Illingworth, Lang, Müyesser, Parczyk and Sgueglia [5] conjectured that $G_{\mathcal{E}}$ is the extremal construction.

Conjecture 1.1. *For every $k \geq 2$, there is some sufficiently large $n_0 \in \mathbb{N}$ such that every k-graph G not containing any isolated vertices with $|V(G)| = n \geq n_0$ and $\delta^*(G) \geq (1-1/k)n-(k-2)$ contains a tight Hamilton cycle.*

Our main result is a proof of this conjecture for all uniformities $k \geq 3$. Since Dirac's theorem corresponds to the conjecture for $k=2$, this provides a complete resolution.

2 Useful Techniques and Results

One of our main tools is an embedding method used in [5], which is based on a new blow-up tiling technique by Lang [15]. In order to state these results, we first make some definitions.

We say that a k-graph F^* is a *blow-up* of a k-graph F if it can be obtained by replacing the vertices of F with pairwise disjoint non-empty vertex sets $\{V_u : u \in V(F)\}$ and replacing each edge $u_1 \dots u_k \in E(F)$ with every possible k-partite edge across $V_{u_1} \times \cdots \times V_{u_k}$. We say F^* a *(γ, m)-regular blow-up* if each part has size in $[(1-\gamma)m, (1+\gamma)m]$, and simply say it is an *m-regular blow-up* if $\gamma = 0$. A *(γ, m)-nearly-regular blow-up* means that all but at most one part have size in $[(1-\gamma)m, (1+\gamma)m]$, and the exceptional part contains exactly one vertex. We define a *cyclic chain tiling* of a k-graph G to be a sequence of subgraphs $H_1, \dots, H_\ell$ that span $V(G)$ with $V(H_i) \cap V(H_j) \neq \emptyset$ if and only if $j = i \pm 1$ (addition and subtraction modulo ℓ), in which case the intersection is equal to a common edge of H_i and H_j. Finally, for integers $s, k \geq 1$ and any $\eta > 0$, we set $\mathcal{P}(s, k, \eta)$ to be the family of s-vertex k-graphs with no isolated vertices and minimum supported co-degree at least $(1 - 1/k - \eta)s$.

We now state a slight modification of Lemma 2.1 of [5].

Lemma 2.1. *Let $1/n \ll 1/m_2 \ll 1/m_1 \ll 1/s, \gamma \ll \eta, 1/k \leq 1$. Let G be an n-vertex k-graph without isolated vertices and $\delta^*(G) \geq (1-1/k)n - (k-2)$. Then, there exists a sequence of s-vertex k-graphs $F_1, \dots, F_\ell \in \mathcal{P}(s, k, \eta)$ and a cyclic chain tiling $F_1^*, \dots, F_\ell^*$ of G such that each F_i^* is a (γ, m_i')-nearly-regular blow-up of F_i for some $m_i' \in [m_1, m_2]$. Additionally, the common edge of every pair (F_i^*, F_{i+1}^*) avoids the singleton parts if they exist.*

In other words, we can find a cyclic chain tiling of G with blow-ups of k-graphs that satisfy a slightly weaker minimum supported co-degree condition than G. The heart of Lemma 2.1 is actually Lemma 4.4 of [15]. Loosely speaking, it provides a set of regular blow-ups F_i^* as above, and they cover all but at most $o(n)$ vertices (it is actually a stronger result that allows the F_i to inherit more general properties than just a degree condition).

3 Proof Outline

Throughout the proof sketch, we will assume that all quantities are suitably divisible and will also avoid taking floors and ceilings (assuming all relevant numbers are integers). This will not affect the overall strategy. Our proof entails separate analysis for k-graphs that are "structurally close" to the extremal example of Illingworth et al. [5] and those that are not.

3.1 Non-Extremal Hypergraphs

Let us first prove the conjecture for when G is "far" from extremal, in the sense that $V(G)$ does not have a subset of size n/k containing very few supported pairs.

Theorem 3.1. *Let $1/n \ll \varepsilon \ll 1/k \leq 1/3$, and let G be an n-vertex k-graph with no isolated vertices and $\delta^*(G) \geq (1-1/k)n - (k-2)$. Suppose every $A \subseteq V(G)$ of size at least n/k contains at least εn^2 supported pairs. Then G contains a tight Hamilton cycle.*

The first step of the proof is to obtain a blow-up chain $F_1^*, \dots, F_\ell^*$ of G from Lemma 2.1. We will look for a tight Hamilton cycle inside each F_i^*, and connect these cycles via the edges that link the tiles F_i^*. Crucially, we use Lemma 4.4 of [15] to argue that the F_i inherit not only the approximate minimum supported degree, but also a slightly weaker version of the non-extremal structure of G. We show that each vertex in the graph of supported pairs either has degree at least $(1 - 1/k + \eta)n$ or its non-neighbourhood has at least $c\eta n^2$ edges for some $c > 0$.

Let us suppose that F^* is a (γ, m)-regular blow-up of F (technically we look at nearly-regular blow-ups, but the arguments are essentially identical). As discussed previously, our main barrier for a tight Hamilton cycle is the existence of a perfect fractional matching.

A standard method to find perfect fractional matchings is via Farkas' lemma [1,17]. It is a linear algebraic result that can be suitably applied to the adjacency matrix of a k-graph to determine the existence of a perfect fractional matching. The argument involves using the degree condition and the non-extremal structure to construct an edge subject to constraints based on a given set of vertex weights (see [9,20] for similar applications). We use a variant of Farkas' lemma to obtain a perfect fractional matching w in F^* with rational edge weights.

We now make a crucial observation that allows us to "blow-up" to a perfect matching. Let K be some integer such that $Kw(e)$ is an integer for all $e \in E(F^*)$, and consider the K-regular blow-up F^{**} of F^*. Each vertex $u \in V(F^*)$ is blown up to a set of K vertices V_u, and we split this into parts $\{V_{u,e} : e \in E(F^*), u \in e\}$, where $|V_{u,e}| = Kw(e)$. Delete every edge incident to $V_{u,e}$ that is not a blow-up of e, and do this for all $u \in V(F^*)$ and $e \in E(F^*)$. Since w is a perfect fractional matching, it is easy to show that the edges we are left with constitute a perfect matching in F^{**}. By construction, this matching is a union of disjoint complete balanced k-partite k-graphs corresponding to blow-ups of edges of F^*.

Consequently, we think of each tile F_i^* as being obtained via two (sufficiently large) consecutive blow-ups. The degree conditions in F yield a tight walk W that contains every edge of F (Lemma 3.1 [5]) which can be blown up to a tight path W^* by picking distinct copies for each repetition of a vertex along W. We then replace the edges of the path with a spanning path in the corresponding complete k-partite k-graphs of the perfect matching. This connects the perfect matching, forming a tight Hamilton cycle in F_i^*.

Of course, it may not be possible to get the final blow-up to be exactly the right size of the tile due to divisibility conditions, and we may have a nearly spanning matching that leaves a few vertices uncovered. The high degree in F allows us to create W such that every $v \in V(F)$ can be "inserted" into W within some subwalk $W_v \subseteq W$ to create a new tight walk. This "absorbing" feature allows us to incorporate the few uncovered vertices. Hence, overall, we reserve this connecting path in F a priori, and blow-up the rest of F in two steps to find a large matching. We then use the connecting path to join the edges of the matching and to cover any remaining vertices to form a tight cycle that spans F_i^*, and finally to move on to the next blown-up tile in the chain.

3.2 Extremal Hypergraphs

Theorem 3.2. *Let $1/n \ll \varepsilon \ll 1/k \leq 1/3$. Suppose G be an n-vertex k-graph with no isolated vertices and $\delta^*(G) \geq (1 - 1/k)n - (k - 2)$. If G admits a partition of its vertex set $V(G) = A \sqcup B$ into two sets such that $|A| = n/k$ and A contains at most εn^2 supported pairs, then G has a tight Hamilton cycle.*

Let us first introduce some notation. Given a k-graph G and any $i \in [k - 1]$, we define the *i-shadow of G*, denoted $\partial^i G$, to be the i-graph with $V(\partial^i G) = V(G)$ whose edges are the supported i-sets of G. For any $U \subseteq V(G)$, let $\partial^i_G[U]$ be the subgraph of $\partial^i G$ induced by U.

Our strategy is inspired by the construction of powers of Hamilton cycles for near-extremal graphs in the resolution of the Pósa-Seymour conjecture by Komlós, Sárközy and Szemerédi [10,11]. We find a $(k-1)$-uniform tight cycle in $\partial^{k-1}[B]$ that spans B, and then "insert" vertices of A after every set of $k-1$ vertices in the cycle to form a k-uniform tight Hamilton cycle in G.

The first step is to show that $\partial^{k-1}[B]$ is quite "dense". In particular, we use a simple counting argument to show that for all $i \in [k-2]$, every supported i-set must support almost all the vertices of B, because otherwise it would create too many supported pairs inside A. Once we have this, we make an arbitrary partition of B in $k-1$ equally sized parts $B_1, \dots, B_{k-1}$ and consider the corresponding $(k-1)$-partite subgraph of $\partial^{k-1}[B]$, say $T^{k-1}(B)$. We perform a semi-random greedy procedure to construct a perfect matching M in $T^{k-1}(B)$. Roughly speaking, we pick a random subset of δn vertices in B_1, where $\delta \gg \varepsilon$, and order them arbitrarily. We consider the first vertex of this random set and select a uniformly random neighbour in B_2. Then, for the second vertex, we pick a random neighbour in B_2 that has not yet been chosen in the previous steps, and so on. The remaining vertices in B_1 are matched to B_2 using Hall's theorem. In the next step, we consider supported pairs across $B_1 \times B_2$ and match into B_3, and so on. Next, we form an auxiliary digraph D, whose vertices are the edges of M. Given two edges $u_1 \dots u_{k-1}$ and $v_1 \dots v_{k-1}$ of M with $u_i, v_i \in B_i$ for each i, we direct an edge from $u_1 \dots u_{k-1}$ to $v_1 \dots v_{k-1}$ if $u_1 \dots u_{k-1} v_1 \dots v_{k-1}$ forms a $(k-1)$-uniform tight path in $\partial^{k-1}[B]$. Analogous to the matching procedure, the degree conditions will imply that D has large minimum in and out degrees with high probability. We then construct a directed Hamilton cycle in D, again using a random greedy algorithm. This will exactly correspond to a spanning $(k-1)$-uniform tight cycle, say C, in $T^{k-1}(B)$.

While the density of $\partial^{k-1}[B]$ would likely allow for a deterministic construction of a spanning cycle in $\partial^{k-1}[B]$, we may not be able to insert vertices of A into an arbitrary such cycle. A simple double counting argument of edges across the partition $A \sqcup B$ shows that there can be very few $(k-1)$-sets in $\partial^{k-1}[B]$ that support at most $|A| - \varepsilon' n$ vertices in A ($\varepsilon \ll \varepsilon' \ll 1/k$). Hence, we show that we can work with a modified $\partial^{k-1}[B]$ which does not have any such edges, but still has sufficiently high degree conditions to allow the random constructions to work. We then argue that, with high probability, for each $v \in A$, there are several pairs of consecutive $(k-1)$-sets $e, f \in C$ such that $e \cup \{v\} \cup f$ is a tight k-uniform path in G. Finally, an application of Hall's theorem will complete the proof.

The only problem with the strategy described is that there could be a few vertices in A that do not have a high enough degree in B, and there could be a few vertices in B that are part of several sets of $k-1$ with low degree into A, making it difficult to avoid such sets in C. Consequently, we modify A and B beforehand by transferring some "problematic" vertices across the partition. The issue now is that we may no longer have $|A| = |B|/(k-1)$. In this case, we argue that we can find a collection of small supported sets within A which, when individually contracted to vertices, give the desired relative sizes of A and B. We should mention this argument is the only part of the proof that uses the exact minimum supported co-degree condition. We build C in $T^{k-1}(B)$ so that it contains some a priori reserved $(k-1)$-edges between which we can insert these few exceptional vertices.

References

[1] V. Chvátal, *Linear programming*, Macmillan, 1983.

[2] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **3** (1952), no. 1, 69–81.

[3] A. Ghouilahouri, *A sufficient condition for the existence of a Hamiltonian circuit*, Weekly Reports of the Sessions of the Academy of Sciences **251** (1960), no. 4, 495–497.

[4] H. Hàn and M. Schacht, *Dirac-type results for loose Hamilton cycles in uniform hypergraphs*, Journal of Combinatorial Theory, Series B **100** (2010), no. 3, 332–346.

[5] F. Illingworth, R. Lang, A. Müyesser, O. Parczyk, and A. Sgueglia, *Spanning spheres in Dirac hypergraphs*, arXiv preprint arXiv:2407.06275 (2024).

[6] F. Joos and J. Kim, *On a rainbow version of Dirac's theorem*, Bulletin of the London Mathematical Society **52** (2020), no. 3, 498–504.

[7] G. Y. Katona and H. A. Kierstead, *Hamiltonian chains in hypergraphs*, Journal of Graph Theory **30** (1999), no. 3, 205–212.

[8] P. Keevash, D. Kühn, R. Mycroft, and D. Osthus, *Loose Hamilton cycles in hypergraphs*, Discrete Mathematics **311** (2011), no. 7, 544–559.

[9] P. Keevash and R. Mycroft, *A geometric theory for hypergraph matching*, vol. 233, American Mathematical Society, 2015.

[10] J. Komlós, G. N. Sárközy, and E. Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Random Structures & Algorithms **9** (1996), no. 1-2, 193–211.

[11] J. Komlós, G. N Sárközy, and E. Szemerédi, *Proof of the Seymour conjecture for large graphs*, Annals of Combinatorics **2** (1998), 43–60.

[12] D. Kühn, R. Mycroft, and D. Osthus, *Hamilton ℓ-cycles in uniform hypergraphs*, Journal of Combinatorial Theory, Series A **117** (2010), no. 7, 910–927.

[13] D. Kühn and D. Osthus, *Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree*, Journal of Combinatorial Theory, Series B **96** (2006), no. 6, 767–821.

[14] D. Kühn and D. Osthus, *Matchings in hypergraphs of large minimum degree*, Journal of Graph Theory **51** (2006), no. 4, 269–280.

[15] R. Lang, *Tiling dense hypergraphs*, arXiv preprint arXiv:2308.12281 (2023).

[16] C. Lee and B. Sudakov, *Dirac's theorem for random graphs*, Random Structures & Algorithms **41** (2012), no. 3, 293–305.

[17] L. Lovász and M. D. Plummer, *Matching theory*, vol. 367, American Mathematical Soc., 2009.

[18] K. Markström and A. Ruciński, *Perfect matchings (and Hamilton cycles) in hypergraphs with large degrees*, European Journal of Combinatorics **32** (2011), no. 5, 677–687.

[19] V. Rödl, A. Ruciński, and E. Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combinatorics, Probability and Computing **15** (2006), no. 1-2, 229–251.

[20] V. Rödl, A. Ruciński, and E. Szemerédi, *Perfect matchings in uniform hypergraphs with large minimum degree*, European Journal of Combinatorics **27** (2006), no. 8, 1333–1349.

[21] V. Rödl, A. Ruciński, and E. Szemerédi, *An approximate Dirac-type theorem for k-uniform hypergraphs*, Combinatorica **28** (2008), no. 2, 229–260.

[22] V. Rödl, A. Ruciński, and E. Szemerédi, *Perfect matchings in large uniform hypergraphs with large minimum collective degree*, Journal of Combinatorial Theory, Series A **116** (2009), no. 3, 613–636.

[23] V. Rödl, A. Ruciński, and E. Szemerédi, *Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs*, Advances in Mathematics **227** (2011), no. 3, 1225–1299.

13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25–29, 2025

Path degeneracy and applications

(Extended abstract)

Yuquan Lin* Patrice Ossona de Mendez**

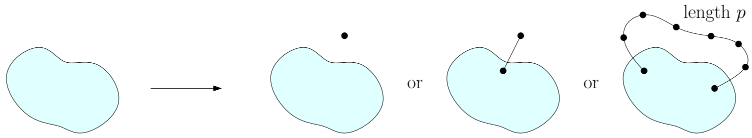
Abstract

In this work, we relate girth and path-degeneracy in classes with sub-exponential expansion, with explicit bounds for classes with polynomial expansion and proper minor-closed classes that are tight up to a constant factor (and tight up to second order terms if a classical conjecture on existence of g-cages is verified). As an application, we derive bounds on the generalized acyclic indices, on the generalized arboricities, and on the weak coloring numbers of high-girth graphs in such classes.

1 Introduction and previous work

Intuitively, a graph is *p-path degenerate* if it can be constructed by iteratively adding paths of length p. An original version of this notion was introduced in [16] in a study of the circular chromatic number of graphs with large girth excluding a minor, where it was proved that high-girth graphs excluding a minor are almost bipartite.

In this paper, we consider a slightly different, more robust definition. Formally, a graph is *p-path degenerate* if it can be constructed from the empty graph by successively adding an isolated vertex, a vertex of degree one, or a path of length p with both endpoints in the previously constructed graph:



*Southeast University, Nanjing, Jiangsu, China and CAMS (CNRS, UMR 8557), Paris, France. email: yqlin@seu.edu.cn
**CAMS (CNRS, UMR 8557), Paris, France. email: pom@ehess.fr

In this work, we improve the proof and bounds given in [16], and obtain the following results on the minimum girth g_p ensuring that a graph in a given class $\mathscr{C}$ is p-path degenerate:

- if $\mathscr{C}$ is sub-exponential, we give a short proof that g_p exists (already shown in [16]);
- if $\mathscr{C}$ has polynomial expansion, we prove $g_p = O(p \log_2 p)$;
- if $\mathscr{C}$ is a proper minor-closed class, we prove $g_p = O(p)$.

More precise bounds are given below in Section 2. These results are applied to provide new bounds for chromatic invariants $(\mathrm{a}'_r, \mathrm{arb}_r, \mathrm{wcol}_r)$ for graphs with large girth. Precisely, the invariants we study are:

- the *generalized r-acyclic chromatic index* $\mathrm{a}'_r(G)$ of a graph G, which is the minimum number of colors of a proper edge coloring such that every cycle C gets at least $\min(|C|, r)$ colors. This invariant has been introduced by Gerke, Greenhill, and Wormald [6] as a generalization of the acyclic chromatic index [1, 5].

- the *generalized r-arboricity* $\mathrm{arb}_r(G)$ of a graph G, which is the minimum number of colors of a (possibly improper) edge coloring such that every cycle C gets at least $\min(|C|, r+1)$ colors. This invariant has been introduced by Nešetřil, Zhu, and the second author [17] as a generalization of graph arboricity.

- the *weak coloring number* $\mathrm{wcol}_r(G)$ of a graph G introduced by Kierstead and Yang [12] in the context of game coloring, and intensively used in Sparisty theory [15, 18] as a generalization of the coloring number.

The full version of this work can be found in [14].

2 Path degeneracy

First, we relate the path degeneracy for graphs with large girth to (shallow) topological minors by showing that every graph with girth at least $2p-1$ (this bound is tight) is p-path degenerate if and only if it contains no topological minor of depth $(p/2-1)$ with minimum degree at least 3. Then, we deduce bounds on the girth of the graph ensuring p-path degeneracy in some nice classes. A classical conjecture on existence of g-cages will be of particular importance in this setting.

Conjecture 2.1 ([3, p. 164])**.** *For infinitely many integers g there exists a cubic graph with girth at least g and order at most $c2^{g/2}$.*

Recall that for a graph G and an integer r, $\nabla_r(G)$ denotes the maximum of $\|H\|/|H|$ over minors of G at depth r [15] and that the *expansion* of a class $\mathscr{C}$ is the function $\mathrm{Exp}_{\mathscr{C}} : \mathbb{N} \to [0, +\infty]$ defined by

$$\mathrm{Exp}_{\mathscr{C}}(r) = \sup\{\nabla_r(G) \colon G \in \mathscr{C}\}.$$

A class $\mathscr{C}$ has *bounded expansion* (resp. *sub-exponential expansion*, *polynomial expansion*) if $\mathrm{Exp}_{\mathscr{C}}(r)$ is finite for every integer r (resp. $\mathrm{Exp}_{\mathscr{C}}(r) = 2^{o(r)}$, $\mathrm{Exp}_{\mathscr{C}}(r) = r^{O(1)}$). The class $\mathscr{C}$ is *proper minor-closed* if every minor of a graph in $\mathscr{C}$ belongs to $\mathscr{C}$ and $\mathscr{C}$ excludes at least one graph. Note that a class $\mathscr{C}$ is included in a proper minor-closed class if and only if $\mathrm{Exp}_{\mathscr{C}}(r) = O(1)$.

We start with a very general statement.

Theorem 2.2. *Let $\mathscr{C}$ be a class with sub-exponential expansion. Then, for every integer p there exists an integer g_p such that every graph $G \in \mathscr{C}$ with girth at least g_p is p-path degenerate.*

Sketch. If a graph $G \in \mathscr{C}$ has large girth and is not p-path degenerate, then it contains, as a minor of depth $(p-1)/2$, a graph with large girth and minimum degree 3. This minor, in turns, contains minors of depth r with degree about 2^r [13] (for r bounded by 1/4th of the girth). If the original girth is sufficiently large, this contradicts the sub-exponential bound of the expansion of $\mathscr{C}$. □

For a class with polynomial expansion, a more precise analysis allows us to give a more precise bound on g_p in terms of the branch W_{-1} of Lambert's W function (whose branches form the converse relation of the complex function $f(z) = ze^z$).

Theorem 2.3. *Let $\mathscr{C}$ be a class with polynomial expansion, with $\mathrm{Exp}_{\mathscr{C}}(r) \leq a(r + \frac{1}{2})^b$ for some positive reals a, b.*

Then, every graph $G \in \mathscr{C}$ with

$$\mathrm{girth}(G) > \max\left(7, 2\left\lfloor -\frac{2b}{\log 2} W_{-1}\left(-\frac{\log 2}{(24\sqrt{2}a)^{1/b}\, bp}\right)\right\rfloor + 4\right)(p-1) \sim 4bp\log_2 p$$

is p-path degenerate.

We further prove that our bounds are tight up to a constant factor.

Theorem 2.4. *For every real $b > 0$, there exists a class $\mathscr{C}$ with expansion $\mathrm{Exp}_{\mathscr{C}}(r) = O(r^b)$ such that for infinitely many $p \in \mathbb{N}$ the class contains a graph G with*

$$\mathrm{girth}(G) \geq -\frac{8b}{3\log 2} W_{-1}\left(-\frac{\log 2}{pb}\right)(p-1) \sim \frac{8}{3} bp\log_2 p$$

that is not p-path degenerate.

Moreover, the above lower bound can be improved to

$$-\frac{4b}{\log 2} W_{-1}\left(-\frac{\log 2}{pb}\right)(p-1) \sim 4bp\log_2 p,$$

if we assume that Conjecture 2.1 is true. Hence, the bound provided in Theorem 2.3 is likely to be tight (up to second order terms).

Eventually, we give some bounds for proper minor-closed classes of graphs, by considering two different bounds on the density of shallow minors of high girth graphs with minimum degree 3 [4, 13].

Theorem 2.5. *Let $\mathscr{C}$ be a minor-closed class of graphs and let d be the maximum average degree of the graphs in $\mathscr{C}$. Then, for every integer $p \geq 2$, every graph $G \in \mathscr{C}$ with girth strictly greater than*

$$(4\log_2 d + 2\log_2(\min\{d, 576\}) + 3)(p-1)$$

is p-path degenerate.

Moreover, the bound cannot be improved to better than $(\frac{8}{3}\log_2(d-2) + c)(p-1)$.

In the case where the class $\mathscr{C}$ is the class of all graphs excluding K_k as a minor, we derive (according to [19] for the upper bound) that the maximum girth $g(K_k, p)$ of a K_k-free graph that is not p-path degenerate satisfies

$$\left(\frac{8}{3}\log_2 k + c_1\right)(p-1) \leq g(K_k, p) < (4\log_2 k + 2\log_2\log_2 k + c_2)(p-1),$$

for some constants c_1, c_2. We conjecture that $g(K_k, p)$ is close to the upper bound.

Conjecture 2.6. *We have*

$$g(K_k, p) = (4\log_2 k + O(1))(p-1).$$

Note that Conjecture 2.1 would imply, if true, a lower bound of the form $(4\log_2(d-2) + c)(p-1)$ for unboundedly many d (and all p) for Theorem 2.5 and, for infinitely many k and every p, a lower bound of the form $(4\log_2 k + c)(p-1)$ for $g(K_k, p)$.

3 Applications

We review here some of the applications of our study of path degeneracy.

Generalized r-arboricity

The generalized r-arboricity $\mathrm{arb}_r(G)$ is unbounded for planar graphs [2]. However, for every class with bounded expansion, $\mathrm{arb}_r(G)$ is bounded for graphs with sufficiently large girth (depending on r) [17]. In this context, we focus on the bound for the generalized r-arboricity of large girth graphs in a given class. We study the invariant $\mathrm{A}_r(\mathscr{C})$ of a class $\mathscr{C}$ with bounded expansion, which was defined in [2] as

$$\mathrm{A}_r(\mathscr{C}) := \min\{k : \exists g\ \forall G \in \mathscr{C}\ (\mathrm{girth}(G) \geq g) \Rightarrow (\mathrm{arb}_r(G) \leq k)\}.$$

We confirm the conjecture that for every integer $r \geq 1$, every proper minor-closed class $\mathscr{C}$ satisfies $\mathrm{A}_r(\mathscr{C}) \leq r+1$ [2, Conjecture 7] by proving the following theorem in a strong form.

Theorem 3.1. *Let $\mathscr{C}$ be a class with sub-exponential expansion and let r be a positive integer.*

If $\mathscr{C}$ contains some graph G with $r < \mathrm{girth}(G) < \infty$, then $\mathrm{A}_r(\mathscr{C}) = r+1$; otherwise, $\mathrm{A}_r(\mathscr{C}) = 1$.

Generalized r-acyclic chromatic index

In general, large girth graphs can have arbitrary large generalized r-acyclic chromatic indices, for instance, Greenhill and Pikhurko [7] proved that for fixed integer $r \geq 4$, there exist graphs G such that $a'_r(G) \geq c\Delta^{\lfloor r/2 \rfloor}$ for some constants $c > 0$. However, if restricted to some nice classes of graphs, large girth graphs may be forced to have small generalized r-acyclic chromatic indices, in the sence that they are close to the best possible value of $\max\{\Delta, r\}$. For example, combining [10, Theorem 3.3] and [21, Theorem 6] together, we obtain for every integer $r \geq 3$, every planar graph G with $\Delta \geq 3$ and $\mathrm{girth}(G) \geq 5r+1$ satisfies $a'_r(G) = \max\{\Delta, r\}$. In this context, we extend the existence of a linear bound for the girth ensuring $a'_r(G) \leq \max\{\Delta, r\}$ to all proper minor-closed classes.

Theorem 3.2. *Let $r \geq 3$ be an integer, let $\mathscr{C}$ be a class with sub-exponential class, and let $g(p)$ be the girth ensuring p-path degeneracy in $\mathscr{C}$.*

Then, every graph $G \in \mathscr{C}$ with maximum degree Δ and girth at least $g(r+1)$ that is not a forest satisfies $\mathrm{a}'_r(G) = \max\{\Delta, r\}$.

Weak coloring number

Given an integer $r \geq 1$, the *weak r-coloring number* $\mathrm{wcol}_r(G)$ of a graph G is a minimum integer k such that there exists a linear order π on the vertices of G, so that for every $v \in V(G)$, there are at most k vertices $u \leq_\pi v$ (possible $u = v$) that can be reached from v by a path P with length at most r, for which every internal vertex w of P satisfies $u <_\pi w$.

The growth rate of weak r-coloring number for graphs in a fixed proper minor-closed class $\mathscr{C}$ has been extensively studied, see [9, Table 1] for a summary on the lower and upper bounds on $\max_{G\in\mathscr{C}} wcol_r(G)$. In particular, for the class of all K_k-minor free graphs ($k \geq 4$), the lower and upper bounds on $\max_{G\in\mathscr{C}} wcol_r(G)$ are $\Omega(r^{t-2})$ [8] and $O(r^{t-1})$ [20], respectively. Even for the class of planar graphs $\mathscr{C}$, the exact growth rate of maximum weak r-coloring numbers is still unknown, and the known lower and upper bounds on $\max_{G\in\mathscr{C}} wcol_r(G)$ are $\Omega(r^2 \log r)$ [11] and $O(r^3)$ [20], respectively.

In this context, we prove that for large girth graphs in any proper minor-closed class $\mathscr{C}$ in Theorem 3.3, the weak r-coloring numbers are close to $r + 2$.

Theorem 3.3. *Let $r \geq 1$ be an integer, let $\mathscr{C}$ be a class with sub-exponential class, and let $g(p)$ be the girth ensuring p-path degeneracy in $\mathscr{C}$.*

Then, for every graph $G \in \mathscr{C}$, we have

$$\mathrm{wcol}_r(G) \leq \begin{cases} r + 2 + \lfloor \log_2(\frac{q-1}{q-r}) \rfloor, & \textit{if } \mathrm{girth}(G) \geq g(2q) \textit{ and } r < q < 2r, \\ r + 2, & \textit{if } \mathrm{girth}(G) \geq g(4r). \end{cases}$$

In particular, we have the following bounds for the two extreme cases of $q = r + 1$ and $q = 2r - 1$:

$$\mathrm{wcol}_r(G) \leq \begin{cases} r + 2 + \lfloor \log_2 r \rfloor, & \text{if } \mathrm{girth}(G) \geq g(2r + 2), \\ r + 3, & \text{if } \mathrm{girth}(G) \geq g(4r - 2). \end{cases}$$

The first author is supported by the China Scholarship Council (CSC) and SEU Innovation Capability Enhancement Plan for Doctoral Students (CXJH_SEU 24119).

This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 810115 – DYNASNET).

References

[1] N. Alon, C. McDiarmid, and B. Reed. Acyclic coloring of graphs. *Random Structures & Algorithms*, 2(3):277–288, 1991.

[2] T. Bartnicki, B. Bosek, S. Czerwiński, M. Farnik, J. Grytczuk, and Z. Miechowicz. Generalized arboricity of graphs with large girth. *Discrete Mathematics*, 342(5):1343–1350, 2019.

[3] B. Bollobás. *Extremal graph theory.* Academic Press, 1978.

[4] R. Diestel and C. Rempel. Dense minors in graphs of large girth. *Combinatorica*, 25:111–116, 2004.

[5] J Fiamcik. The acyclic chromatic class of a graph. *Math. Slovaca*, 28(2):139–145, 1978.

[6] S. Gerke, C. Greenhill, and N. Wormald. The generalized acyclic edge chromatic number of random regular graphs. *Journal of Graph Theory*, 53(2):101–125, 2006.

[7] C. Greenhill and O. Pikhurko. Bounds on the generalised acyclic chromatic numbers of bounded degree graphs. *Graphs and Combinatorics*, 21(4):407–419, 2005.

[8] M. Grohe, S. Kreutzer, R. Rabinovich, S. Siebertz, and K. Stavropoulos. Coloring and covering nowhere dense graphs. *SIAM Journal on Discrete Mathematics*, 32(4):2467–2481, 2018.

[9] J. Hodor, H. La, P. Micek, and C. Rambaud. Weak coloring numbers of minor-closed graph classes. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3325–3334. SIAM, 2025.

[10] J. Hou, J. Wu, G. Liu, and B. Liu. Acyclic edge colorings of planar graphs and seriesparallel graphs. *Science in China Series A: Mathematics*, 52(3):605–616, 2009.

[11] G. Joret and P. Micek. Improved bounds for weak coloring numbers. *arXiv preprint arXiv:2102.10061*, 2021.

[12] H.A. Kierstead and D. Yang. Orderings on graphs and game coloring number. *Order*, 20:255–264, 2003.

[13] D. Kühn and D. Osthus. Minors in graphs of large girth. *Random Structures & Algorithms*, 22(2):213–225, 2003.

[14] Y. Lin and P. Ossona de Mendez. Path degeneracy and applications. arXiv:2503.18614 [math.CO], 2025.

[15] J. Nešetřil and P. Ossona de Mendez. *Sparsity (Graphs, Structures, and Algorithms)*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. 465 pages.

[16] J. Nešetřil and P. Ossona de Mendez. A note on circular chromatic number of graphs with large girth and similar problems. *Journal of Graph Theory*, 80(4):268–276, 2015.

[17] J. Nešetřil, P. Ossona de Mendez, and X. Zhu. Coloring edges with many colors in cycles. *Journal of Combinatorial Theory, Series B*, 109:102–119, 2014.

[18] S. Siebertz. On the generalized coloring numbers. arXiv preprint arXiv:2501.08698, 2025.

[19] A. Thomason. The extremal function for complete minors. *Journal of Combinatorial Theory, Series B*, 81(2):318–338, 2001.

[20] J. Van Den Heuvel, P. Ossona de Mendez, D. Quiroz, R. Rabinovich, and S. Siebertz. On the generalised colouring numbers of graphs that exclude a fixed minor. *European Journal of Combinatorics*, 66:129–144, 2017.

[21] X. Zhang, G. Wang, Y. Yu, J. Li, and G. Liu. On r-acyclic edge colorings of planar graphs. *Discrete Applied Mathematics*, 160(13-14):2048–2053, 2012.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Partition subcubic planar graphs into independent sets

(Extended abstract)

Xujun Liu* Yan Wang†

Abstract

A packing $(1^\ell, 2^k)$-coloring of a graph G is a partition of $V(G)$ into ℓ independent sets and k 2-packings (whose pairwise vertex distance is at least 3). The square coloring of planar graphs was first studied by Wegner in 1977. Thomassen, and independently Hartke et al. proved one can always square color a cubic planar graph with 7 colors, i.e., every subcubic planar graph is packing (2^7)-colorable. We focus on packing $(1^\ell, 2^k)$-colorings, which lie between proper coloring and square coloring. Gastineau and Togni proved every subcubic graph is packing $(1, 2^6)$-colorable. They also asked whether every subcubic graph except the Petersen graph is packing $(1, 2^5)$-colorable.

In this paper, we prove an analogue result of Thomassen and Hartke et al. on packing coloring that every subcubic planar graph is packing $(1, 2^5)$-colorable. This also answers the question of Gastineau and Togni affirmatively for subcubic planar graphs. Moreover, we prove that there exists an infinite family of subcubic planar graphs that are not packing $(1, 2^4)$-colorable, which shows that our result is the best possible. Besides, our result is also sharp in the sense that the disjoint union of Petersen graphs is subcubic and non-planar, but not packing $(1, 2^5)$-colorable.

*Department of Foundational Mathematics, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu Province, 215123, China, xujun.liu@xjtlu.edu.cn; the research of X. Liu was supported by the National Natural Science Foundation of China under grant No. 12401466 and the Research Development Fund RDF-21-02-066 of Xi'an Jiaotong-Liverpool University.

†School of Mathematical Sciences, Shanghai Frontier Science Center of Modern Analysis (CMA-Shanghai), Shanghai Jiao Tong University, Shanghai, 200240, China, yan.w@sjtu.edu.cn; the research of Y. Wang was supported by National Key R&D Program of China under grant No. 2022YFA1006400, Shanghai Municipal Education Commission (No. 2024AIYB003), National Natural Science Foundation of China under grant No. 12201400 and Explore X project of Shanghai Jiao Tong University.

1 Introduction

An i-packing in a graph G is a set of vertices whose pairwise vertex distance is at least $i+1$. Let $S=(s_1,s_2,...,s_k)$ be a non-decreasing sequence of positive integers. A *packing S-coloring* of a graph G is a partition of $V(G)$ into sets $V_1,...,V_k$ such that each V_i is an s_i-packing. In particular, a packing $(1^\ell,2^k)$-coloring of a graph G is a partition of $V(G)$ into ℓ independent sets and k 2-packings. This concept was introduced by Goddard and Xu [16] and it has drawn much attention in graph coloring (e.g., see [4, 7, 8, 12, 14, 15, 19, 20, 22]). The *packing chromatic number* (PCN), $\chi_p(G)$, of a graph G is the minimum k such that G has a packing $(1,2,...,k)$-coloring. The notion of PCN was first studied under the name broadcast chromatic number by Goddard, Hedetniemi, Hedetniemi, Harris, and Rall [15] in 2008, which was motivated by a frequency assignment problem in broadcast networks. They asked whether the PCN of subcubic graphs is bounded by a constant. Balogh, Kostochka, and Liu [4] answered their question in the negative using the probabilistic method. Later, Brešar and Ferme [7] gave an explicit construction which shows it is unbounded. The 1-subdivision of a graph G is obtained by replacing each edge with a path of two edges. Gastineau and Togni [14] asked whether the PCN of the 1-subdivision of a subcubic graph is bounded by 5 and subsequently, Brešar, Klavžar, Rall, and Wash [8] conjectured it is true. The current best upper bound (6) was recently proved by Liu, Zhang, and Zhang [19].

The famous Four Color Theorem, which was proved by Appel and Haken [1], Appel, Haken, and Koch [2], as well as Robertson, Sanders, Seymour, and Thomas [21] states that every planar graph is 4-colorable, i.e., every planar graph is packing (1^4)-colorable. The square of a graph G, denoted by G^2, is the graph obtained from G by adding the edges joining vertices with distance exactly two. Wegner [24] conjectured in 1977 that if G is a planar graph with maximum degree Δ then

$$\chi(G^2)\le\begin{cases}7 & \Delta=3,\\ \Delta+5 & 4\le\Delta\le 7,\\ \lfloor\frac{3\Delta}{2}\rfloor & \Delta\ge 8.\end{cases}$$

Note that a coloring of G^2 using k colors is equivalent to a packing (2^k)-coloring of G. Thomassen [23] and independently Hartke, Jahanbekam, and Thomas [13] confirmed Wegner's conjecture [24] for the case when $\Delta=3$ by proving that every subcubic planar graph is packing (2^7)-colorable. Their result is also sharp due to the existence of subcubic planar graphs that are not packing (2^6)-colorable.

Theorem 1 (Thomassen [23]; Hartke et al. [13]). *Every subcubic planar graph is packing (2^7)-colorable.*

Bousquet, Deschamps, Meyer, and Pierron [6] recently showed every planar graph with maximum degree at most four is packing (2^{12})-colorable. The square coloring of planar graphs with girth conditions has also been considered by many researchers. For example, Dvořák, Král, Nejedlý, and Škrekovski [11] showed a planar graph with girth at least six and sufficiently large maximum degree Δ is packing $(2^{\Delta+2})$-colorable. The list version of square coloring is also studied (e.g., see [3, 5, 10, 17]). In particular, Cranston and Kim [10] proved the list chromatic number of the square of a subcubic graph, except the Petersen graph, is at most 8.

By Brooks' theorem [9], every subcubic graph except K_4 has a proper 3-coloring, i.e., a packing (1^3)-coloring. Gastineau and Togni [14] showed that if one can prove every subcubic graph except the Petersen graph has a packing $(1^2, 2^2)$-coloring then the conjecture of Brešar et al. [8] is true. This also motivates the study of packing $(1^\ell, 2^k)$-colorings, which lie between proper coloring and square coloring. Gastineau and Togni [14] proved that every subcubic graph is packing $(1^2, 2^3)$-colorable. Furthermore, they asked whether every subcubic graph except the Petersen graph is packing $(1, 1, 2, 3)$-colorable. This question still remains open. Very recently, Liu et al. [19] showed every subcubic graph is packing $(1, 1, 2, 2, 3)$-colorable and conjectured that every subcubic graph except the Petersen graph is packing $(1^2, 2^2)$-colorable.

For packing $(1, 2^k)$-coloring, Tarhini and Togni [22] proved that every cubic Halin graph is packing $(1, 2^5)$-colorable. Mortada and Togni [20] showed that every subcubic graph with no adjacent heavy vertices (degree three vertices with all neighbours also being 3-vertices) is packing $(1, 2^5)$-colorable. Gastineau and Togni [14] proved every subcubic graph is packing $(1, 2^6)$-colorable. Moreover, they asked the following question after performing a computer search on graphs with small order.

Question 2 (Gastineau and Togni [14])**.** *Is it true that every subcubic graph except the Petersen graph is packing* $(1, 2^5)$*-colorable?*

2 Main result and sharpness example

In this paper, weprove an analogue result of Thomassen [23] and Hartke et al. [13] that all subcubic planar graphs are packing $(1, 2^5)$-colorable. This also answers Question 2 in the affirmative for subcubic planar graphs.

Theorem 3. *Every subcubic planar graph is packing* $(1, 2^5)$*-colorable.*

Our result is the best possible since the Petersen graph is non-planar and is not packing $(1, 2^5)$-colorable. To see this, we know the independence number of the Petersen graph is four and its diameter is two. Therefore, at most four vertices can be colored by the 1-color and each of the remaining at least six vertices must receive a distinct 2-color, which is impossible.

Furthermore, we give an infinite family of non-packing-$(1, 2^4)$-colorable subcubic planar graphs in the following example. This shows the sharpness of our result in another sense.

Example 4. *Let G be the graph in Figure 1, G_1 be the subgraph of G induced by the vertices $\{v_1, \ldots, v_7\}$, and G_2 be the subgraph of G induced by the vertices $\{u_1, \ldots, u_7\}$. Obviously, the graph G is subcubic and planar. We provide a packing $(1, 2^5)$-coloring of G using the colors $1, A, B, C, D, E$, where 1 is the 1-color and each of A, B, C, D, E is a 2-color (see*

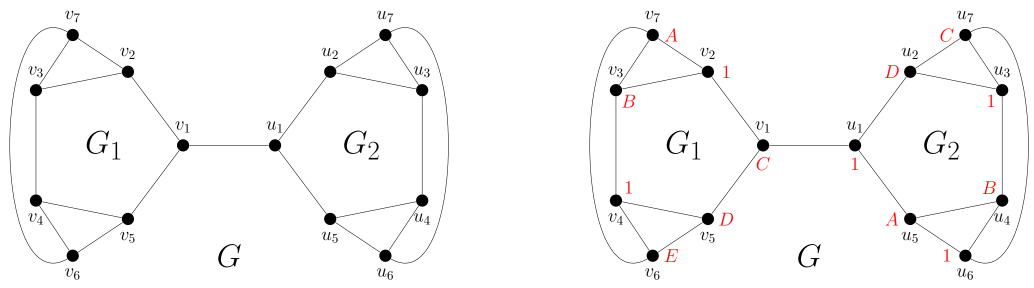


Figure 1: A non-packing-$(1, 2^4)$-colorable subcubic planar graph.

Figure 1 right picture). We show G is not packing $(1, 2^4)$-colorable. Thus, the disjoint union of G's is an infinite family of non-packing-$(1, 2^4)$-colorable subcubic planar graphs.

Proof. Suppose G has a packing $(1, 2^4)$-coloring using colors $1, A, B, C, D$, where 1 is the 1-color and each of A, B, C, D is a 2-color. We show that $v_1 \in G_1$ must be colored with 1. Suppose not, say v_1 is colored with A. Note that G_1 has diameter two and thus each of the colors A, B, C, D can be used at most once. Since $v_2v_3v_7$ and $v_4v_5v_6$ are two triangles, the color 1 can be used at most once in each of the triangles. However, we have at least four uncolored vertices and each of them requires a distinct color from A, B, C, D, which is a contradiction to the fact that G_1 has diameter two. Similarly, we can show $u_1 \in G_2$ must be colored with 1. This is a contradiction since $u_1v_1 \in E(G)$. □

3 Reducible configurations

It suffices to show Theorem 3 for connected graphs since otherwise we can apply the argument to each component. Suppose that Theorem 3 is false, i.e., there are subcubic planar graphs that are not packing $(1, 2^5)$-colorable. Let G be a counterexample with smallest $|V(G)|$. Our plan is first to show that G must be a cubic graph and cannot contain each configuration in Figure 2. We also prove that there are no separating cycles of length up to 7, as well as the non-existence of some specific separating 8-cycle (which may occur in our Configurations). Then we use the discharging method to redistribute charges to show that G is too dense to be a planar graph.

Lemma 5. *The graph G is cubic and each configuration in Figure 2 does not exist.*

We prove the non-existence of configurations by extending partial packing $(1, 2^5)$-colorings (its existence is guaranteed by the minimality of G after the deletion of each configuration) to G. The details of the proof of the non-existence of configurations can be found in the arXiv version [18].

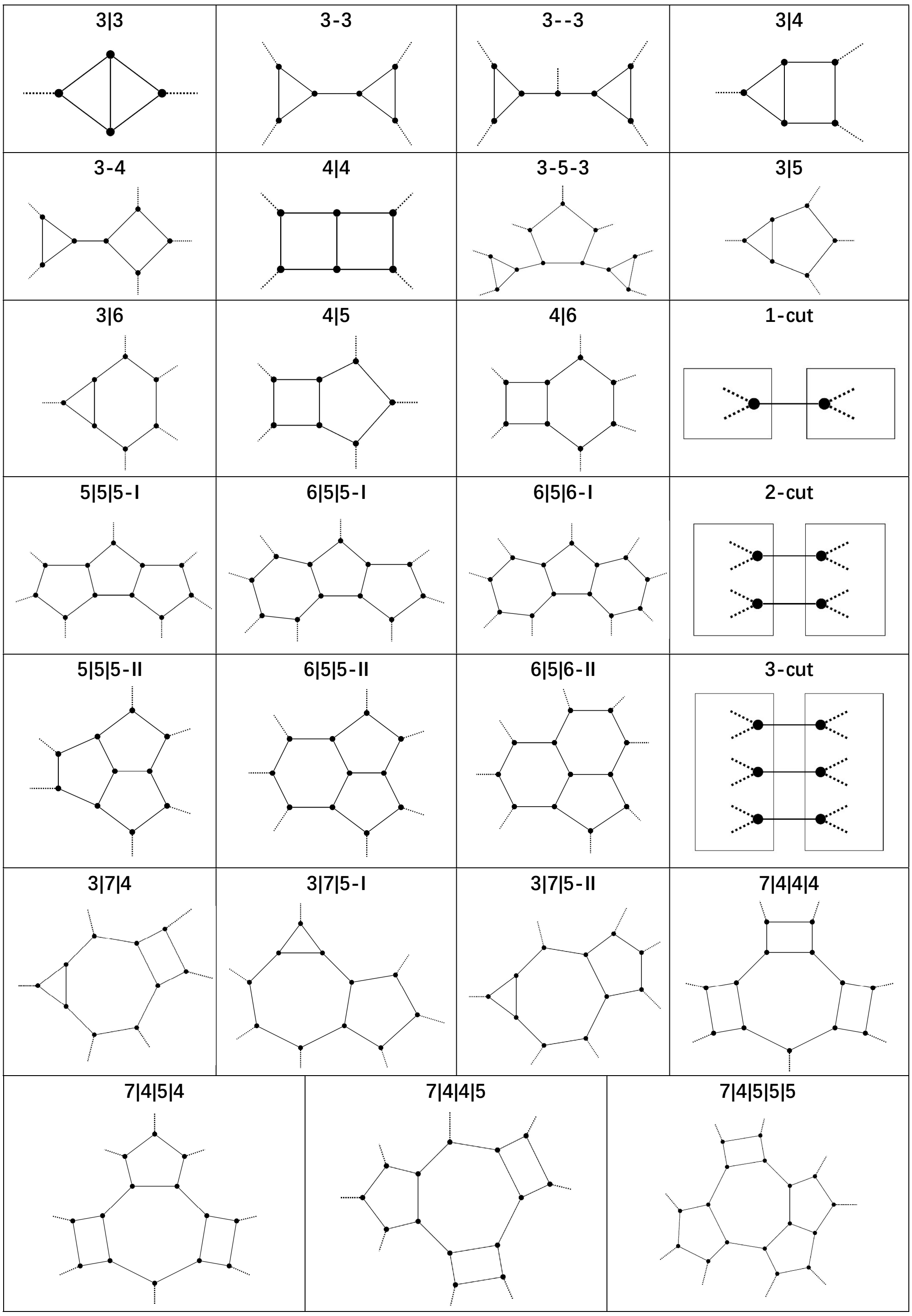


Figure 2: Configurations.

4 Proof of Theorem 3 - the Discharging Method

Suppose that Theorem 3 is false, i.e., there are subcubic planar graphs that are not packing $(1, 2^5)$-colorable. Let G be a counterexample with smallest $|V(G)|$. Note that G is a connected planar graph. According to Euler's formula for connected planar graphs, we use face charging:

$$\sum_{v \in V(G)} (2d(v) - 6) + \sum_{f \in F(G)} (\ell(f) - 6) = -12. \tag{1}$$

Let the initial charge of each vertex in G be $Ch(v) = 2d(v) - 6$ and each face in G be $Ch(f) = \ell(f) - 6$. Our goal is to redistribute the charges so that the final charge $Ch^*(v)$ for each vertex v and $Ch^*(f)$ for each face f satisfy $Ch^*(v) \geq 0$ and $Ch^*(f) \geq 0$. This gives us $\sum_{v \in V(G)} Ch^*(v) + \sum_{f \in F(G)} Ch^*(f) \geq 0$, which contradicts (1).

We call a vertex v or a face f *happy* if $Ch^*(v) \geq 0$ or $Ch^*(f) \geq 0$ respectively. Now we define the rules for redistributing the charges.

(i): Each 3-face receives charge 1 from each of its adjacent faces of length at least 7.

(ii): Each 4-face receives charge $\frac{1}{2}$ from each of its adjacent faces of length at least 7.

(iii): Each 5-face receives charge $\frac{1}{4}$ from each of its adjacent faces of length at least 7.

We show all the vertices and faces are happy by the following claims.

Claim 6. *All the vertices and faces of length at most* 6 *are happy.*

Proof. After applying rules **(i)**,**(ii)**, and **(iii)**, all the vertices are happy since G is a cubic graph by Lemma 5 and the charges on vertices are unchanged by the redistributing rules. By Lemma 5, there are no $3|3, 3|4, 3|5$, and $3|6$, a 3-face must be adjacent to three faces of length at least 7 and therefore receives charge 1 from each adjacent face. The final charge of a 3-face is $3-6+3 \cdot 1 = 0$ and thus every 3-face is happy. Since there are no $3|4, 4|4, 4|5$, and $4|6$, a 4-face must be adjacent to four faces of length at least 7 and therefore receives charge $\frac{1}{2}$ from each adjacent face. The final charge is $4-6+4 \cdot \frac{1}{2} = 0$ and thus every 4-face is happy. For a 5-face, since there are no $3|5, 4|5$, it is not adjacent to 3-faces and 4-faces. Furthermore, by Lemma 5, a 5-face is adjacent to at most one face of length at most 6, and thus it receives charge $\frac{1}{4}$ from each of at least four adjacent faces of length 7 or more. The final charge is at least $5-6+4 \cdot \frac{1}{4} = 0$ and thus every 5-face is happy. All the 6-faces are happy since a 6-face does not give any charge to adjacent faces. □

Claim 7. *All the* 7*-faces are happy.*

Proof. By Lemma 5, there are no $3|3, 3\text{-}3$, and 3- -3, a 7-face f can be adjacent to at most one 3-face.

Case 1: f is adjacent to exactly one 3-face. Furthermore, by Lemma 5, it cannot be adjacent to any other faces of length at most 5. Therefore, f is happy since the final charge for f is $7-6-1 = 0$.

Case 2: f is not adjacent to any 3-faces.

Case 2.1: f is adjacent to at least one 4-face. By Lemma 5, there are no $4|4$, $7|4|4|4$, f can be adjacent to at most two 4-faces. If f is adjacent to two 4-faces, then by Lemma 5, there are no 4|4, 4|5, 7|4|5|4, and 7|4|4|5, the other faces that f is adjacent to are of length at least 6. Thus, the final charge for f is $7-6-2\cdot\frac{1}{2}=0$ and f is happy. So f is adjacent to one 4-face. By Lemma 5, there are no 5|5|5-I and 7|4|5|5|5, f can be adjacent to at most two faces of length 5. The final charge of f is at least $7-6-\frac{1}{2}-2\cdot\frac{1}{4}=0$ and f is happy.

Case 2.2: f is not adjacent to any 3-faces or 4-faces. By Lemma 5, there is no 5|5|5-I and thus f can be adjacent to at most four 5-faces. The final charge of f is at least $7-6-4\cdot\frac{1}{4}=0$ and f is happy. □

Claim 8. *All the 8-faces are happy.*

Proof. For a 8-face f, since there are no 3|3, 3-3, 3- -3, f can be adjacent to at most two 3-faces.

Case 1: f is adjacent to exactly two 3-faces. Then they must be in opposite positions. Furthermore, since there are no 3|4, 3|5, 3-4, and 3-5-3, the other faces that f is adjacent to are of length at least 6. Thus, the final charge of f is $8-6-2\cdot 1=0$.

Case 2: f is adjacent to exactly one 3-face. Since there are no 3|4, 3-4, and 4|4, f can be adjacent to at most two 4-faces. If f is adjacent to two 4-faces, then it cannot be adjacent to any 5-faces. Thus, the final charge of f is $8-6-1-2\cdot\frac{1}{2}=0$ and f is happy. If f is adjacent to one additional 4-face, then it can be adjacent to at most two 5-faces. Thus, the final charge of f is at least $8-6-1-\frac{1}{2}-2\cdot\frac{1}{4}=0$ and f is happy. Thus, f is not adjacent to any 4-face. By Lemma 5, there is no 5|5|5-I, it can be adjacent to at most four 5-faces. Thus, the final charge of f is at least $8-6-1-4\cdot\frac{1}{4}=0$ and f is happy.

Case 3: f is not adjacent to any 3-face. Since there is no 4|4, f can be adjacent to at most four 4-faces. Let y be the number of 4-faces that f is adjacent to and z be the number of 5-faces that f is adjacent to. Since there is no 4|4, $0\le y\le 4$. Since there is no 5|5|5-I, $0\le z\le 5$. Since there is no 4|5, $2y+\frac{3}{2}z\le 8$. We conclude that f is happy since the final charge is at least $8-6-y\cdot\frac{1}{2}-z\cdot\frac{1}{4}=2-\frac{1}{4}(2y+z)\ge 2-\frac{1}{4}(2y+\frac{3}{2}z)=0$. □

Claim 9. *All the 9-faces are happy.*

Proof. For a 9-face f, since there are no 3|3, 3-3, 3- -3, f can be adjacent to at most two 3-faces.

Case 1: f is adjacent to exactly two 3-faces. The two 3-faces must have distance exactly 3. By Lemma 5, there are no 3|4, 3-4, 3|5, 3-5-3, f cannot be adjacent to any 4-face and can be adjacent to at most two more 5-faces. Therefore, the final charge is at least $9-6-2\cdot 1-\frac{1}{4}\cdot 2=0$.

Case 2: f is adjacent to exactly one 3-face. Let y be the number of 4-faces that f is adjacent to and z be the number of 5-faces that f is adjacent to. By Lemma 5, there is no 3|4, 3|5, 3-4, 4|4, 4|5, and 5|5|5-I, $0\le y\le 2$, $0\le z\le 4$, and $2y+\frac{3}{2}z\le 7$. We conclude that f is happy since the final charge is at least $9-6-1-y\cdot\frac{1}{2}-z\cdot\frac{1}{4}=2-\frac{1}{4}(2y+z)\ge 2-\frac{7}{4}=\frac{1}{4}>0$.

Case 3: f is not adjacent to any 3-face. Since there is no 4|4, f can be adjacent to at most four 4-faces. Let y be the number of 4-faces that f is adjacent to and z be the number of 5-faces that f is adjacent to. We know $0\le y\le 4$. Since there is no 5|5|5-I, $0\le z\le 6$. Since there is no 4|5, $2y+\frac{3}{2}z\le 9$. We conclude that f is happy since the final charge is at least $9-6-y\cdot\frac{1}{2}-z\cdot\frac{1}{4}=3-\frac{1}{4}(2y+z)\ge 3-\frac{9}{4}=\frac{3}{4}>0$. □

Claim 10. *All the faces of length at least* 10 *are happy.*

Proof. For a face of length $\ell \geq 10$. Let x, y, z be the number of $3, 4, 5$-faces that f is adjacent to respectively. By Lemma 5, $0 \leq 4x \leq \ell$, $0 \leq 2y \leq \ell$, $0 \leq \frac{3}{2}z \leq \ell$, $3x + 2y \leq \ell$, $2x + \frac{3}{2}z \leq \ell$, and $2y + \frac{3}{2}z \leq \ell$. Thus, the final charge is at least

$$\ell - 6 - x - y \cdot \frac{1}{2} - z \cdot \frac{1}{4} = \ell - 6 - \frac{7}{30}(3x + 2y) - \frac{3}{20}(2x + \frac{3}{2}z) - \frac{1}{60}(2y + \frac{3}{2}z)$$

$$\geq \ell - 6 - (\frac{7}{30} + \frac{3}{20} + \frac{1}{60}) \cdot \ell = \frac{3}{5}\ell - 6 \geq 0,$$

since $\ell \geq 10$. Therefore, all faces of length at least 10 are happy. □

By Claims 6, 7, 8, 9 and 10, it is a contradiction to (1) since all the vertices and faces are happy. This completes the proof of Theorem 3. □

5 Concluding Remarks

We believe the answer to Question 2 is affirmative and therefore posed the following conjecture.

Conjecture 11. *Every subcubic graph except the Petersen graph is packing* $(1, 2^5)$*-colorable.*

References

[1] K. Appel and W. Haken, Every planar map is four colorable. Part I. Discharging, *Illinois J. Math.* **21** (1977), 429–490.

[2] K. Appel, W. Haken and J. Koch, Every planar map is four colorable. Part II. Reducibility, *Illinois J. Math.* **21** (1977), 491–567.

[3] O. Amini, L. Esperet and J. Heuvel, A unified approach to distance-two colouring of graphs on surfaces, *Combinatorica* **33** (3) (2013),253–296.

[4] J. Balogh, A. Kostochka and X. Liu, Packing chromatic number of cubic graphs, *Discrete Math.* **341** (2018), 474–483.

[5] M. Bonamy, B. Lévêque and A. Pinlou, List coloring the square of sparse graphs with large degree, *European J. Combin.* **41** (2014), 128–137.

[6] N. Bousquet, Q. Deschamps, L. Meyer and T. Pierron, Square Coloring Planar Graphs with Automatic Discharging, *SIAM Journal on Discrete Mathematics* **38** (1) (2024), 504–528.

[7] B. Brešar and J. Ferme, An infinite family of subcubic graphs with unbounded packing chromatic number, *Discrete Math.* **341** (2018), 2337–2342.

[8] B. Brešar, S. Klavžar, D.F. Rall and K. Wash, Packing chromatic number, $(1, 1, 2, 2)$-colorings, and characterizing the Petersen graph, *Aequationes Math.* **91** (2017), 169–184.

[9] R.L. Brooks, On coloring the nodes of a network, *Proc. Cambridge Phil. Soc.* **37** (1941) 194–197.

[10] D. Cranston and S. Kim, List-coloring the square of a subcubic graph. *J. Graph Theory* **57** (2008), 65–87.

[11] Z. Dvořák, D. Král, P. Nejedlý and R. Škrekovski, Coloring squares of planar graphs with girth six, *European J. Combin.* **29** (2008), 838–849.

[12] J. Fiala, S. Klavžar and B. Lidický, The packing chromatic number of infinite product graphs, *European J. Combin.* **30** (2009), 1101–1113.

[13] S. Hartke, S. Jahanbekam and B. Thomas, The chromatic number of the square of subcubic planar graphs, arXiv:1604.06504.

[14] N. Gastineau and O. Togni, S-packing colorings of cubic graphs, *Discrete Math.* **339** (2016), 2461–2470.

[15] W. Goddard, S.M. Hedetniemi, S.T.Hedetniemi, J.M. Harris and D.F. Rall, Broadcast chromatic numbers of graphs, *Ars Combin.* **86** (2008), 33–49.

[16] W. Goddard and H. Xu, The S-packing chromatic number of a graph, Discuss. Math. Graph Theory **32** (2012), 795–806.

[17] S. Kim and B. Park, Counterexamples to the List Square Coloring Conjecture, *J. Graph Theory* **78** (2015), 239–247.

[18] X. Liu and Y. Wang, Partition subcubic planar graphs into independent sets, arXiv:2408.12189.

[19] X. Liu, X. Zhang and Y. Zhang, Every subcubic graph is packing $(1, 1, 2, 2, 3)$-colorable, arXiv:2404.09337.

[20] M. Mortada and O. Togni, Further results and questions on S-packing coloring of subcubic graphs, arXiv:2407.07424.

[21] N. Robertson, D. P. Sanders, P. D. Seymour and R. Thomas, The four colour theorem, *J. Comb. Theory, Ser. B* **70** (1997), 2–44.

[22] B. Tarhini and O. Togni, S-Packing Coloring of Cubic Halin Graphs, *Discrete Appl. Math.* **349** (31) (2024), 53–58.

[23] C. Thomassen, The square of a planar cubic graph is 7-colorable, *J. Comb. Theory, Ser. B* **128** (2018), 192–218.

[24] G. Wegner, Graphs with given diameter and a coloring problem. Technical report, University of Dortmund, 1977.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Asymptotic Enumeration of Subclasses of Level-2 Phylogenetic Networks

(Extended abstract)

Hexuan Liu* Bing-Ze Lu† Taoyang Wu‡ Guan-Ru Yu§

Abstract

This paper studies the enumeration of seven subclasses of level-2 phylogenetic networks under various planarity and structural constraints, including terminal planar, tree-child, and galled networks. We derive their exponential generating functions, recurrence relations, and asymptotic formulas. Specifically, we show that the number of networks of size n in each class follows:

$$N_n \sim c \cdot n^{n-1} \cdot \gamma^n,$$

where c is a class-specific constant and γ is the corresponding growth rate. Our results reveal that being terminal planar can significantly reduce the growth rate of general level-2 networks, but has only a minor effect on the growth rates of tree-child and galled level-2 networks. Notably, the growth rate of 3.83 for level-2 terminal planar galled tree-child networks is remarkably close to the rate of 2.94 for level-1 networks.

1 Introduction

Phylogenetic trees have been instrumental in evolutionary biology, illustrating how species diverge from common ancestors. However, their limitations in representing events like hybridization and horizontal gene transfer led to the development of phylogenetic networks. These networks extend trees by incorporating complex evolutionary events, offering a more flexible framework for modeling evolutionary histories.

A **phylogenetic network** with n leaves is an acyclic directed graph with a single source (the root) and a set of sinks (leaves) bijectively labeled by $\{1, 2, \ldots, n\}$. Throughout this paper, all phylogenetic networks are binary, satisfying the following conditions:

*Department of Pure and Applied Mathematics, Graduate School of Fundamental Science and Engineering, Waseda University, 169-8555 Tokyo, Japan. E-mail: lhx@ruri.waseda.jp

†Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan. E-mail: wind951016@gmail.com

‡School of Computing Sciences, University of East Anglia (UEA), Norwich, NR4 7TJ, UK. E-mail: taoyang.wu@uea.ac.uk

§Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan. E-mail: gryu@math.nsysu.edu.tw

(i) If $n > 1$, the root has exactly one child, and a leave has exactly one parent.

(ii) Each interior node is either a tree node (one parent and two children) or a reticulation node (two parents and one child).

(iii) For every 2-connected component B with three or more vertices, there are at least two cut arcs directed away from B.

A phylogenetic network is a **tree-child network** if every interior node has at least one child that is either a tree node or a leaf (see [2,5]). Similarly, it is a **galled network** if every reticulation node lies in a unique tree cycle (see [6,8]). A **galled tree-child network** (GTC) is a phylogenetic network that is both a galled network and a tree-child network (see [3]).

A particularly important class is **level-k networks**, where each 2-connected component contains at most k reticulation nodes (see [7]). See Figure 1 for examples of level-2 networks. This hierarchical classification controls the complexity of networks, making them computationally tractable. For instance, level-1 networks correspond to galled trees (see [9,10]), while higher levels (such as level-2) model more complex evolutionary histories (see [1,11,12]).

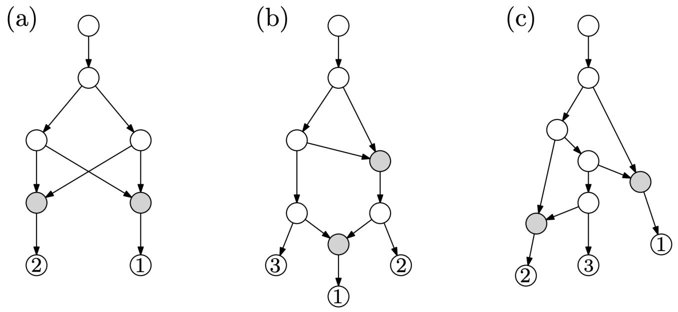


Figure 1: Examples of level-2 networks where reticulation nodes are represented by filled circles. (a) is a galled network which is not tree-child; (b) is a tree-child network which is not galled; (c) is a galled tree-child network. Both (b) and (c) are terminal and outer planar, whereas (a) is neither.

The **planarity** of phylogenetic networks is crucial for their visualization and interpretation. A phylogenetic network is *planar* if it can be embedded on an Euclidean xy-plane without edge crossings, except at their endpoints. Based on planarity, phylogenetic networks are classified into four types:

- **Planar networks** ($\mathcal{N}_p$): Admit a planar embedding with no edge crossings.
- **Upward planar networks** ($\mathcal{N}_u$): Networks with a planar drawing in which each directed edge is a curve of monotonically non-decreasing in the y-coordinate.
- **Terminal planar networks** ($\mathcal{N}_t$): Planar networks with a planar drawing in which all leaves and the root are on the outer face.
- **Outer planar networks** ($\mathcal{N}_o$): Planar networks with a planar drawing in which all vertices are on the outer face boundary.

These classes satisfy the inclusion relation: $\mathcal{N}_p \supset \mathcal{N}_u \supset \mathcal{N}_t \supset \mathcal{N}_o$.

In fact, level-1 or 2 constraints are closely tied to planarity, leading to the following properties (see [14]).

Proposition 1.1. *The following statements hold:*

(i) All level-1 networks are outer planar.

(ii) All level-2 networks are upward planar. Moreover, a level-2 network N is terminal planar if and only if N is outer planar.

Proposition 1.1 establishes that level-2 tree-child networks are not only planar but also upward planar. Under the level-2 constraint, network planarity reduces to two cases: outer planar networks and non-outer planar networks (which remain upward planar).

It is well known that there are $(2n-3)!!$ binary phylogenetic trees with n leaves. However, counting phylogenetic networks is considerably more challenging. Consequently, much of the recent progress has focused on subclasses with certain topological constraints, such as level-1 and level-2 networks [1], tree-child networks [5, 13], and galled networks [6, 8].

This paper focuses on enumerating subclasses of level-2 networks, including **level-2 tree-child networks** and **level-2 galled networks** under various planarity constraints. We derive functional equations for their exponential generating functions and asymptotic growth rates, revealing their structural properties and combinatorial behavior. Additionally, we study **galled tree-child networks**, the intersection of the two classes above. Our results deepen the mathematical understanding of level-2 networks. In Section 2, we present the main results, followed by the methodology in Section 3, where generating functions and theoretical tools are employed for enumeration.

2 Main Results

Theorem 2.1. *Let t_n denote the number of level-2 tree-child networks with n leaves. Then, the asymptotic growth of t_n is given by*

$$t_n \sim c \cdot n^{n-1} \cdot \gamma^n, \tag{1}$$

where $c \approx 0.0667418464$ and $\gamma \approx 4.6710490708$. Moreover, each of the seven subclasses of level-2 networks in the following table has the same asymptotic growth pattern as in Eq.(1) with distinct constants c and growth rates γ as summarized in the table.

	general	***tree-child***	***galled***	***GTC***
arbitrary *planar* *upward planar*	$c = 0.02931010$ $\gamma = 15.4332995$	$c = 0.06674185$ $\gamma = 4.67104907$	$c = 0.05885954$ $\gamma = 6.42241234$	$c = 0.07888067$ $\gamma = 3.98275804$
terminal planar *outer planar*	$c = 0.03486095$ $\gamma = 12.9230111$	$c = 0.07450612$ $\gamma = 4.33252428$	$c = 0.06965278$ $\gamma = 5.39994365$	$c = 0.08586449$ $\gamma = 3.83201916$

Remark 2.2. *The first cell follows from Bouvel et al. [1]: they showed that the exponential growth rate for level-1 networks is* 2.94 *and that for level-2 networks is $\gamma = 15.43$. In contrast, we found that the rate for level-2 terminal planar networks is $\gamma = 12.92$, indicating a noticeable gap. However, under the tree-child or galled constraints, this gap is significantly smaller. Moreover, the growth rate of level-2 galled tree-child networks ($\gamma = 3.98/3.83$) is rather close to that of level-1 networks. This suggests that the tree-child and galled constraints impose limitations similar to those of level-1 networks.*

3 Asymptotics via Generating Functions

In this section, we describe a general methodology for deriving asymptotic results on the number of networks from their generating functions. We illustrate this approach using level-2 tree-child networks as a primary example to outline a proof for Theorem 2.1.

Lemma 3.1. *Let t_n denote the number of level-2 tree-child networks with n leaves, and let $T(x)$ be the corresponding exponential generating function defined as*

$$T(x) = \sum_{n \geq 1} t_n \frac{x^n}{n!}. \tag{2}$$

Then, $T(x)$ satisfies the functional equation:

$$\begin{aligned} T = x + \frac{T^2}{2} &+ \frac{1}{2}\left(\left(\frac{1}{1-T}\right)^2 - 1\right)T + \frac{3}{2}\left(\frac{1}{1-T}\right)^2\left(\frac{T}{1-T}\right)\left(\left(\frac{1}{1-T}\right)^2 - 1\right)T \\ &+ \left(\frac{1}{1-T}\right)^4\left(\left(\frac{1}{1-T}\right)^2 - 1\right)T^2 + \frac{1}{4}\left(\frac{1}{1-T}\right)^2\left(\left(\frac{1}{1-T}\right)^2 - 1\right)^2 T^2, \end{aligned}$$

which is derived from the decomposition illustrated in Figure 2.

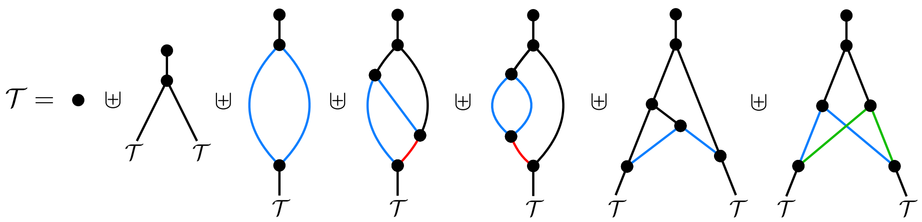


Figure 2: Decomposition of the combinatorial structure for level-2 tree-child networks. Red edges indicate that they cannot be empty (must contain some insertion). In contrast, pairs of blue or green edges represent that, within each pair of edges of the same color, both cannot be empty simultaneously.

Due to space constraints, we omit the detailed derivation of the generating function (only the decomposition is shown), as it is closely tied to the combinatorial structure of the problem and involves a lengthy process. Instead, we present the functional equation obtained from our analysis. By comparing the coefficients on both sides of the equation and iteratively substituting the results from the left-hand side into T on the right-hand side, we derive the following series expansion:

$$T(x) = x + \frac{3}{2}x^2 + 11x^3 + \frac{745}{8}x^4 + \frac{3333}{4}x^5 + \frac{127395}{16}x^6 + \frac{637635}{8}x^7 + \frac{105501013}{128}x^8 + \cdots.$$

This expansion corresponds to the EGF (2), yielding the sequence

$$\{t_n\}_{n \geq 1} = \{1, 3, 66, 2235, 99900, 5732775, 401710050, 33232819095, \cdots\}.$$

Next, we introduce the key tools used to analyze the n-th order coefficients of the generating function. We begin with the following version of singular inversion theorem.

Theorem 3.2. *(Singular inversion theorem) [4] Let $C(z)$ be a generating function with $C(0) = 0$, satisfying $C(z) = z\phi(C(z))$, where $\phi(z) = \sum_{n\geq 0}\phi_n z^n$ is a power series such that: (i) $\phi_0 \neq 0$, (ii) all ϕ_n are non-negative reals, (iii) $\phi(z) \not\equiv \phi_0 + \phi_1 z$ (i.e., ϕ is nonlinear).*

Let R be the radius of convergence of ϕ at 0. Assume (iv) ϕ is analytic at 0 (so $R > 0$), (v) the equation $\phi(z) - z\phi'(z) = 0$ has a unique solution $\tau \in (0, R)$, and (vi) ϕ is aperiodic.

Then:

- *The radius of convergence of $C(z)$ is $\rho = \frac{\tau}{\phi(\tau)}$.*
- *Near ρ, $C(z) \sim \tau - \sqrt{\frac{2\phi(\tau)}{\phi''(\tau)}}\sqrt{1-\frac{z}{\rho}}$.*
- *$[z^n]C(z) = \frac{1}{n}[z^{n-1}]\phi(z)^n$. Furthermore, as $n \to \infty$, $[z^n]C(z) \sim \sqrt{\frac{\phi(\tau)}{2\phi''(\tau)}}\frac{\rho^{-n}}{\sqrt{\pi n^3}}$.*

We now apply Theorem 3.2 to the functional equation derived from Lemma 3.1. First, we rewrite the equation into a form compatible with Theorem 3.2, namely $T(z) = z\phi(T(z))$. The function $\phi(z)$ is given by:

$$\phi(z) = \frac{-4(z-1)^6}{2z^7 - 18z^6 + 67z^5 - 126z^4 + 124z^3 - 70z^2 + 30z - 4} \quad = 1 + \frac{3}{2}z + \frac{35}{4}z^2 + \frac{403}{8}z^3 + \cdots.$$

We can verify that $\phi_0 \neq 0$ and that ϕ is nonlinear. In fact, it is straightforward to check that all coefficients ϕ_n are positive rational numbers. By solving the characteristic equation

$$\phi(z) - z\phi'(z) = 0,$$

we obtain a unique solution $\tau \approx 0.1226285445$ (computed using Maple and the Newton method). From this, we further derive

$$\rho = \frac{\tau}{\phi(\tau)} \approx 0.0787573489.$$

The asymptotic expression for t_n is then given by

$$t_n = n! \cdot [x^n]T(x) \sim \sqrt{\frac{\phi(\tau)}{\phi''(\tau)}}\, n^{n-1}(\rho e)^{-n},$$

where, after calculation, we obtain the constants in Theorem 2.1:

$$\sqrt{\frac{\phi(\tau)}{\phi''(\tau)}} \approx 0.0667418464 \quad \text{and} \quad \frac{1}{\rho e} \approx 4.6710490707.$$

References

[1] M. Bouvel, P. Gambette, and M. Mansouri (2020). Counting phylogenetic networks of level 1 and level 2, *J. Math. Biol.*, **81**(6-7), 1357–1395.

[2] G. Cardona and L. Zhang (2020). Counting and enumerating tree-child networks and their subclasses, *J. Comput. System Sci.*, **114**, 84–104.

[3] Y.-S. Chang, M. Fuchs, and G.-R. Yu (2024). Galled Tree-Child Networks, *LIPICS*, **AofA 2024**, 8:1–8:13.

[4] P. Flajolet and R. Sedgewick (2009). *Analytic Combinatorics*, Cambridge University Press.

[5] M. Fuchs, G.-R. Yu, and L. Zhang (2021). On the asymptotic growth of the number of tree-child networks, *European J. Combin.*, **93**, 103278.

[6] M. Fuchs, G.-R. Yu, and L. Zhang (2022). Asymptotic enumeration and distributional properties of galled networks, *J. Comb. Theory Ser. A.*, **189**, 105599.

[7] P. Gambette, V. Berry, and C. Paul (2009). The structure of level-k phylogenetic networks. In *Twentieth Annual Symposium on Combinatorial Pattern Matching (CPM'09)*, **volume 5577** of Lecture Notes in Computer Science, 289–300. Springer.

[8] A. D. M. Gunawan, J. Rathin, and L. Zhang (2020). Counting and enumerating galled networks, *Discrete Appl. Math.*, **283**, 644–654.

[9] K. Huber, L. Van Iersel, V. Moulton, C. Scornavacca, and T. Wu (2017). Reconstructing phylogenetic level-1 networks from nondense binet and trinet sets. *Algorithmica*, **77**(1), 173–200.

[10] K. Huber, V. Moulton, C. Semple, and T. Wu (2018). Quarnet inference rules for level-1 networks. *Bulletin of Mathematical Biology*, **80**, 2137–2153.

[11] L. Van Iersel, J. Keijsper, S. Kelk, L. Stougie, F. Hagen, and T. Boekhout (2009). Constructing level-2 phylogenetic networks from triplets. *IEEE/ACM Transactions in Computational Biology and Bioinformatics*, **6**(4), 667–681.

[12] L. Van Iersel and V. Moulton (2014). Trinets encode tree-child and level-2 phylogenetic networks. *Journal of Mathematical Biology*, **68**(7), 1707–1729.

[13] C. McDiarmid, C. Semple, and D. Welsh (2015). Counting phylogenetic networks. *Annals of Combinatorics*, **19**, 205–224.

[14] V. Moulton and T. Wu (2022). Planar rooted phylogenetic networks, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **20**(2), 1289–1297.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On degree-balanced edge partitions

(Extended abstract)

Borut Lužar* Paweł Pękała† Jakub Przybyło‡ Roman Soták§

Abstract

We investigate two notions of degree-balanced edge decompositions, which seemingly have nothing in common but surprisingly intertwine in various ways.

The first of these concerns majority edge decompositions, more commonly known as majority edge colourings. A graph is said to be majority edge c-colourable if its edges can be assigned one of c colours in such a way that each colour class induces a subgraph including at most $1/2$ the edges incident with every vertex. A natural generalization of this concept allows at most $1/k$'th fraction of monochromatic edges incident to any vertex. If one aims to use at most $k+1$ colours to achieve such an edge colouring, which is typically the least reasonable number, this usually results in nearly balanced degrees across different colours at each vertex. We particularly focus on list variants of this notion. An elegant result in this direction builds upon Galvin's well-known theorem related to Dinitz's conjecture.

The second problem concerns a conjecture that every d-regular graph G of order n contains a subgraph H which is called irregular. This subgraph is required to have a nearly balanced degree distribution, meaning that for any fixed degree, the number of vertices attaining it in H differs from $n/(d+1)$ by at most 2. We discuss a solution to this problem in the case of cubic graphs, for which an optimal result can be obtained – one even stronger than the postulated above.

This conjecture was formulated in 2023 by Alon and Wei in a paper where the authors introduced several interesting tools to take advantage of the probabilistic

*Faculty of Information Studies in Novo mesto, Slovenia, Rudolfovo Institute, Novo mesto, Slovenia. E-mail: `borut.luzar@gmail.com`.

†AGH University of Krakow, al. A. Mickiewicza 30, 30-059 Krakow, Poland. E-mail: `ppekala@agh.edu.pl`.

‡AGH University of Krakow, al. A. Mickiewicza 30, 30-059 Krakow, Poland. E-mail: `jakubprz@agh.edu.pl`.

§Pavol Jozef Šafárik University, Faculty of Science, Košice, Slovakia. E-mail: `roman.sotak@upjs.sk`.

method in obtaining an asymptotic approximation of the desired property. Interestingly, one of these tools, derived from linear algebra, played a key role in the study of the first of the two discussed problems.

1 Irregular subgraph of a regular graph

In [2] Alon and Wei formulated an intriguing problem concerning the existence of a spanning subgraph H of any given regular graph G with almost perfectly balanced distribution of all possible degrees in H. Let $m(H,k)$ denote the number of vertices of degree k in H.

Conjecture 1 (Alon & Wei [2]). *Every d-regular graph G on n vertices contains a spanning subgraph H such that for every k, $0 \le k \le d$,*

$$\left| m(H,k) - \frac{n}{d+1} \right| \le 2\,.$$

Such subgraph H is called irregular. Note that since G is regular, then the complement of an irregular H in G, i.e. the graph $(V(G), E(G) \smallsetminus E(H))$ is also irregular. Thus, in fact, we strive to edge-decompose a regular graph into its two irregular subgraphs.

Alon and Wei [2] confirmed Conjecture 1 asymptotically for sufficiently small d relative to n, proving that every n-vertex d-regular graph with $d = o((n/\log n)^{1/3})$ contains a spanning subgraph H satisfying $m(H,k) = (1+o(1))\frac{n}{d+1}$, for all $0 \le k \le d$. Later, Fox, Luo, and Pham [12] strengthened this result by significantly expanding the range of d for which it holds, extending it to all $d = o(n/(\log n)^{12})$. The proofs in both papers rely on the probabilistic method and analyze a randomized procedure for selecting H. This approach had been previously used, e.g., in [13, 28, 29, 30], to study a related concept – the so-called irregularity strength of graphs, which likely played a role in motivating research on Conjecture 1; see, for instance, [1, 7, 8, 9, 10, 11, 16, 17, 20, 24, 25, 28, 26, 27] for further results in this direction. Recently, Ma and Xie [23] introduced a novel deterministic technique based on local adjustments, leading to the first bound independent of n. Specifically, they proved that every d-regular multigraph G on n vertices contains a spanning subgraph H such that, for all k with $0 \le k \le d$, the deviation satisfies $|m(H,k) - \frac{n}{d+1}| \le 2d^2$.

We provide a proof of Conjecture 1 in the case of cubic graphs, that is for $d = 3$.

Theorem 2 (Lužar, Przybyło & Soták [22]). *Every cubic graph G on n vertices, not isomorphic to K_4, $K_{3,3}$, or $3K_4$, contains a spanning subgraph H such that for every k, $0 \le k \le 3$,*

$$m(H,k) \in \left\{ \left\lfloor \frac{n}{d+1} \right\rfloor, \left\lceil \frac{n}{d+1} \right\rceil \right\}.$$

We note that Conjecture 1 in the case of $d = 3$ was independently confirmed by Ma and Xie [23], who refined their local adjustments method to achieve this result. Our approach differs from theirs in two key aspects. First, we employ an entirely different technique – one that is strictly constructive and leads to a straightforward algorithm for directly generating the desired subgraph. More importantly, our result for cubic graphs is stronger than that in [23], which essentially establishes Conjecture 1 for this graph class by proving an upper bound of 2 on the maximum deviation of a subgraph degree frequency from $n/4$. In contrast, we provide a complete characterization of this invariant, determining the exact maximum deviation achievable for any cubic graph. Apart from a few exceptional cases, this deviation is always either 0 or 1/2.

Remarks on the proof of Theorem 2. The proof is divided into two stages.

Stage 1 addresses the case where the graph is connected. In this stage, we present a direct and efficient algorithm for constructing the desired irregular subgraph H. Starting from an empty subgraph, we iteratively add small groups of edges to increase the number of vertices of a given degree in H, typically introducing one or two such vertices at a time. The process begins with vertices of degree 3. While ensuring that H contains the required number of degree-3 vertices, we must also control the number of vertices with other degrees that arise in the process. To achieve this, we keep H as compact as possible – typically ensuring that it remains connected. We then proceed in a similar manner to introduce the desired number of vertices of degree 2 and 1, which ultimately results in an appropriate number of isolated vertices in H.

Stage 1 not only consist in proving Theorem 2 for connected cubic graphs, as for Stage 2 we need a stronger result. A d-regular graph G is said to be $(n_d, n_{d-1}, \ldots, n_0)$-*decomposable* if there exists a subgraph H of G such that $n_k = m(H, k)$, for each $k \in \{0, \ldots, d\}$. Stage 1 amounts to proving the following lemma.

Lemma 3 (Lužar, Przybyło & Soták [22]). *Let G be a connected cubic graph on n vertices. The following statements hold:*

(i) if $n = 4t$ and G is not isomorphic to K_4, then G is (t, t, t, t)-decomposable;

(ii) if $n = 4t$, then G is $(t-1, t-1, t+1, t+1)$-decomposable;

(iii) if $n = 4t+2$ and G is not isomorphic to $K_{3,3}$, then G is $(t, t+1, t, t+1)$-decomposable;

(iv) if $n = 4t+2$, then G is $(t-1, t, t+1, t+2)$-decomposable.

Stage 2 focuses on graphs G with multiple components. In this stage, we essentially prove that, regardless of the distribution of component sizes, it is always possible to construct the desired subgraph H by appropriately assembling related subgraphs obtained within each component via Lemma 3. □

We also pose in [22] a conjecture that the same optimal result as in Theorem 2, with a finite set of exceptions should be attainable for every d.

2 Majority edge colourings

Consider a graph $G = (V, E)$ and a vertex $v \in V$. For any subset $F \subseteq E$, let $F(v)$ denote the set of edges in F that are incident to v, and define $d_F(v) := |F(v)|$. Let $\omega : E \to C$ be an edge colouring of G. For each $c \in C$, denote by E_c the set of edges coloured c. The colouring ω is called a *majority edge colouring* if, for every colour c, each vertex v is incident with at most $(1/2)d(v)$ edges coloured c, i.e., $d_{E_c}(v) \leq (1/2)d(v)$. This concept was inspired by its vertex analogue, introduced in the context of directed graphs by Kreutzer, Oum, Seymour, van der Zypen and Wood [19]. For additional results on this topic, see [3, 4, 5, 15, 18], which trace back to results of Lovász [21].

Unlike vertex colourings, majority edge colourings do not always exist. In particular, any graph containing a vertex of degree 1 cannot admit such a colouring. However, it was shown in [6] that every graph with minimum degree $\delta \geq 4$ has a majority edge 3-colouring. Much of the research on generalisations of majority edge colourings focuses on determining the minimum degree thresholds that ensure the existence of such colourings, for a given number of colours. For an integer $k \geq 2$, we define a $1/k$-majority edge colouring as an edge colouring $\omega : E \to C$ such that $d_{E_c}(v) \leq (1/k)d(v)$ for every vertex v and colour c. Such a colouring cannot exist in graphs with vertices of degree smaller than k, nor if fewer than k colours are used. Moreover, even for arbitrarily large minimum degrees, there exist graphs that do not admit a $1/k$-majority edge k-colouring – for instance, any graph containing a vertex whose degree is not divisible by k. Consequently, the smallest number of colours for which a minimum degree threshold guaranteeing the existence of a $1/k$-majority edge colouring may exist is $k+1$. Note that within such a colouring with $k+1$ colours – only slightly more than k – the resulting colour degrees around each vertex are typically well balanced. Pursuing this direction of research, it was shown in [6] that for every integer $k \geq 2$, there exists a threshold $\delta^{(k)} = O(k^3 \log k)$ such that every graph with minimum degree $\delta \geq \delta^{(k)}$ admits a $1/k$-majority edge $(k+1)$-colouring. Accidentally, a major breakthrough in refining this threshold was then achieved by means of an observation of Alon and Wei [2], originally developed using a linear algebra-based approach in the context of the problem discussed in Section 1.

Lemma 4 (Alon & Wei [2]). *Let $G = (V, E)$ be a graph, and let $z : E \to [0, 1]$ be a weight function assigning to each edge $e \in E$ a real weight $z(e)$ in $[0, 1]$. Then there is a function $x : E \to \{0, 1\}$ assigning to each edge an integer value in $\{0, 1\}$ such that for every $v \in V$*

$$\sum_{e \ni v} z(e) - 1 < \sum_{e \ni v} x(e) \leq \sum_{e \ni v} z(e) + 1.$$

This lemma allowed us to prove in [31] that for every integer $k \geq 2$, if a graph G has minimum degree $\delta \geq \frac{7}{4}k^2 + \frac{1}{2}k$, then G is $1/k$-majority edge $(k+1)$-colourable. It turned out that this result established the correct order of magnitude for the optimal minimum value of $\delta^{(k)}$, as there exist graphs with minimum degree $\delta = k^2 - 1$ that are not $1/k$-majority edge $(k+1)$-colourable. This led to the formulation of the following conjecture.

Conjecture 5 (Przybyło & Pękała [31]). *For every integer $k \geq 2$, if a graph G has minimum degree $\delta \geq k^2$, then G is $1/k$-majority edge $(k+1)$-colourable.*

Let us also mention that a corresponding conjecture in the context of bipartite graphs was fully resolved in [31] using a slightly modified version of Lemma 4, originally due to Alon and Wei.

Our research has recently focused on a natural list variant of the majority edge colouring concept. In particular, we established the following result, with a bound on the minimum degree that is nearly as small as the best known bound in the non-list variant.

Theorem 6 (Przybyło & Pękała [32]). *For every integer $k \geq 2$, each graph G with minimum degree $\delta \geq 2k^2 - 2k$ has a $1/k$-majority edge colouring from lists of size $k+1$.*

This result provides a bound on δ of the correct order of magnitude. It is roughly only twice as large as the one stated in the following conjecture, which cannot be further lowered due to the existence of the aforementioned graph family.

Conjecture 7 (Przybyło & Pękała [32]). *For every integer $k \geq 2$, if a graph G has minimum degree $\delta(G) \geq k^2$, then G has a $1/k$-majority edge colouring from any lists of size $k+1$.*

An interesting aspect of Theorem 6 is that its proof is relatively simple. It relies on an application of Galvin's famous theorem, which confirms the List Colouring Conjecture for bipartite graphs.

Theorem 8 (Galvin's theorem, [14]). *For every bipartite graph G, $\chi'_l(G) = \Delta(G)$.*

Remarks on the proof of Theorem 6. We first, in a sense, split the vertices of large degree in the investigated graph G into multiple copies, each with the degree in $[k, 2k-1]$. We then argue that it suffices to prove the theorem for such near-regular graphs, i.e., graphs in which all degrees lie within this interval. Next, we apply Euler's Theorem to orient the edges of such a graph in a way that ensures the out-degree is nearly equal to the in-degree for every vertex in the resulting digraph. Subsequently, we associate this digraph with a bipartite graph by introducing two copies, say v' and v'', of each original vertex v, and we define edges $u'v''$ corresponding to each arc (u, v) of the digraph. We then apply Galvin's theorem to colour the edges of the bipartite graph. It turns out that this colouring, when transferred one-to-one to the original graph G, satisfies our requirements, provided that certain details are properly addressed. □

We also investigated several other variants of list majority colourings, allowing for varying acceptance thresholds for different colours. For further details, we refer the interested reader to [32]. There, we specifically employ the probabilistic method to study these variants.

References

[1] M. Aigner and E. Triesch. Irregular assignments of trees and forests. *SIAM J. Discrete Math.*, 3(4):439–449, 1990.

[2] N. Alon and F. Wei. Irregular subgraphs. *Combinatorics, Probability and Computing*, 32(2):269–283, 2023.

[3] M. Anastos, A. Lamaison, R. Steiner, and T. Szabó. Majority colorings of sparse digraphs. *Electron. J. Combin.*, 28(2):#P2.31, 2021.

[4] M. Anholcer, B. Bosek, and J. Grytczuk. Majority choosability of digraphs. *Electron. J. Combin.*, 24(3):#P3.57, 2017.

[5] M. Anholcer, B. Bosek, J. Grytczuk, G. Gutowski, J. Przybyło, and M. Zając. Mrs. correct and majority colorings. *arXiv preprint*, 2022. arXiv:2207.09739.

[6] F. Bock, R. Kalinowski, J. Pardey, M. Pilśniak, D. Rautenbach, and M. Woźniak. Majority edge-colorings of graphs. *Electron. J. Combin.*, 30(1):#P1.42, 2023.

[7] G. Chartrand, M.S. Jacobson, J. Lehel, O.R. Oellermann, S. Ruiz, and F. Saba. Irregular networks. *Congr. Numer.*, 64:197–210, 1988.

[8] B. Cuckler and F. Lazebnik. Irregularity strength of dense graphs. *J. Graph Theory*, 58(4):299–313, 2008.

[9] J.H. Dinitz, D.K. Garnick, and A. Gyárfás. On the irregularity strength of the m $\times$ n grid. *J. Graph Theory*, 16:355–374, 1992.

[10] R.J. Faudree and J. Lehel. Bound on the irregularity strength of regular graphs. *Colloq Math Soc János Bolyai*, 52:247–256, 1987.

[11] M. Ferrara, R.J. Gould, M. Karoński, and F. Pfender. An iterative approach to graph irregularity strength. *Discrete Appl. Math.*, 158:1189–1194, 2010.

[12] J. Fox, S. Luo, and H.T. Pham. On random irregular subgraphs. *Random Struct. Alg.*, 64:899–917, 2024.

[13] A. Frieze, R.J. Gould, M. Karoński, and F. Pfender. On graph irregularity strength. *J. Graph Theory*, 41(2):120–137, 2002.

[14] F. Galvin. The list chromatic index of a bipartite multigraph. *J. Combin. Theory Ser. B*, 63(1):153–158, 1995.

[15] A. Girão, T. Kittipassorn, and K. Popielarz. Generalized majority colourings of digraphs. *Combin. Probab. Comput.*, 26(6):850–855, 2017.

[16] A. Gyárfás. The irregularity strength of $K_{m,m}$ is 4 for odd m. *Discrete Math.*, 71:273–274, 1998.

[17] M. Kalkowski, M. Karoński, and F. Pfender. A new upper bound for the irregularity strength of graphs. *SIAM J. Discrete Math.*, 25:1319–1321, 2011.

[18] F. Knox and R. Šámal. Linear bound for majority colourings of digraphs. *Electron. J. Combin.*, 25(3):#P3.29, 2018.

[19] S. Kreutzer, S. Oum, P. Seymour, D. van der Zypen, and D.R. Wood. Majority colourings of digraphs. *Electron. J. Combin.*, 24(2):#P2.25, 2017.

[20] J. Lehel. Facts and quests on degree irregular assignments. *Graph Theory, Combinatorics and Applications, Willey*, pages 765–782, 1991.

[21] L. Lovász. On decomposition of graphs. *Studia Sci. Math. Hungar.*, 1:237–238, 1996.

[22] B. Luzar, J. Przybyło, and R. Soták. Degree-balanced decompositions of cubic graphs. *arXiv preprint*, 2024. arXiv:2408.16121.

[23] J. Ma and S. Xie. Finding irregular subgraphs via local adjustments. *arXiv preprint*, 2024. arXiv:2406.05675.

[24] P. Majerski and J. Przybyło. On the irregularity strength of dense graphs. *SIAM J. Discrete Math.*, 28(1):197–205, 2014.

[25] T. Nierhoff. A tight bound on the irregularity strength of graphs. *SIAM J. Discrete Math.*, 13(3):313–323, 2000.

[26] J. Przybyło. Irregularity strength of regular graphs. *Electron. J. Combin.*, 15(1):#R82, 2008.

[27] J. Przybyło. Linear bound on the irregularity strength and the total vertex irregularity strength of graphs. *SIAM J. Discrete Math.*, 23(1):511–516, 2009.

[28] J. Przybyło. Asymptotic confirmation of the faudree-lehel conjecture on irregularity strength for all but extreme degrees. *J. Graph Theory*, 100:189–204, 2022.

[29] J. Przybyło and F. Wei. On the asymptotic confirmation of the faudree-lehel conjecture for general graphs. *Combinatorica*, 43:791–826, 2023.

[30] J. Przybyło and F. Wei. Short proof of the asymptotic confirmation of the faudree-lehel conjecture. *Electron. J. Combin.*, 30(4):#P4.27, 2023.

[31] P. Pękała and J. Przybyło. On generalised majority edge-colourings of graphs. *Electron. J. Combin.*, 31(4):#P4.66, 2024.

[32] P. Pękała and J. Przybyło. On list extensions of the majority edge colourings. *arXiv preprint*, 2024. arXiv:2502.12688.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Perfect matchings in r-graphs

(Extended abstract)

Yulai Ma* Davide Mattiolo† Eckhard Steffen‡ Isaak H. Wolf§
Junxue Zhang¶

Abstract

An r-regular graph is an r-graph, if every odd set of vertices is connected to its complement by at least r edges. Seymour [On multicolourings of cubic graphs, and conjectures of Fulkerson and Tutte. *Proc. London Math. Soc.* (3), 38(3): 423-460, 1979] conjectured (1) that every planar r-graph is r-edge colorable and (2) that every r-graph has $2r$ perfect matchings such that every edge is contained in precisely two of them.

We give a brief overview of different research approaches and partial results on these conjectures with a focus on H-colorings of r-graphs.

We then study variants of these conjectures and introduce several equivalent statements to these conjectures.

1 Introduction

All graphs considered in this paper are finite and may have parallel edges but no loops. The vertex set of a graph G is denoted by $V(G)$ and its edge set by $E(G)$. A graph is *r-regular* if

*Department of Mathematics, Paderborn University, 33098 Paderborn, Germany. Email: yulai.ma@upb.de.
Supported by Sino-German (CSC-DAAD) Postdoc Scholarship Program 2021 (57575640).
†Department of Computer Science, KU Leuven Kulak, 8500 Kortrijk, Belgium. Email: davide.mattiolo@kuleuven.be.
Supported by a Postdoctoral Fellowship of the Research Foundation Flanders (FWO), project number 1268323N.
‡Department of Mathematics, Paderborn University, 33098 Paderborn, Germany. Email: es@upb.de.
§Department of Mathematics, Paderborn University, 33098 Paderborn, Germany. Email: isaak.wolf@upb.de.
Funded by Deutsche Forschungsgemeinschaft (DFG) - 445863039.
¶School of Mathematics and Statistics, Beijing Institute of Technology, Beijing, China. Email: jxuezhang@163.com.

every vertex has degree r. An r-regular graph is an *r-graph*, if $|\partial_G(X)| \geq r$ for every $X \subseteq V(G)$ of odd cardinality, where $\partial_G(X)$ denotes the set of edges that have precisely one vertex in X.

A (t, r)-PM of an r-graph G is a multiset of $t \cdot r$ perfect matchings $M_1, \ldots, M_{tr}$ of G such that every edge of G is contained in exactly t of them. Clearly, if an r-graph G has a $(1, r)$-PM, then G is r-edge-colorable. In this case, we say that G is class 1, otherwise it is class 2.

Our research is mainly motivated by the following two conjectures.

Conjecture 1.1 ([19])**.** *For $r \geq 1$, every r-graph has $(2, r)$-PM.*

For $r = 3$ the conjecture was attributed to Berge and Fulkerson [5], who put it into print (cf. [19]). This conjecture is open even for planar r-graphs. However, Seymour stated the following stronger conjecture for planar graphs.

Conjecture 1.2 ([19, 20])**.** *For $r \geq 1$, every planar r-graph is class 1.*

1.1 H-coloring

As a unifying approach to study some hard conjectures on cubic graphs, Jaeger [10] introduced colorings with edges of another graph. To be precise, let G and H be graphs. An *H-coloring* of G is a mapping $f : E(G) \to E(H)$ such that

- if $e_1, e_2 \in E(G)$ are adjacent, then $f(e_1) \neq f(e_2)$,
- for every $v \in V(G)$ there exists a vertex $u \in V(H)$ with $f(\partial_G(v)) = \partial_H(u)$.

If such a mapping exists, then we write $H \prec G$ and say H *colors* G. A set $\mathcal{A}$ of connected r-graphs such that for every connected r-graph G there is an $H \in \mathcal{A}$ with $H \prec G$ is said to be *r-complete*. For every $r \geq 3$, let $\mathcal{H}_r$ be an inclusion-wise minimal r-complete set. By definition, any r-graph G of class 1 can be colored with any r-graph H. Indeed, let $M_1, \ldots, M_r$ be r pairwise disjoint perfect matchings of G and v a vertex of H with $\partial_H(v) = \{e_1, \ldots, e_r\}$. Every edge of M_i of G can be mapped to e_i in H. Hence, the aforementioned questions and conjectures reduce to r-graphs of class 2.

For $r = 3$, Jaeger [10] conjectured that the Petersen graph P colors every bridgeless cubic graph. If true, this conjecture would have far reaching consequences. For instance, it would imply that the Berge-Fulkerson Conjecture and the 5-Cycle Double Cover Conjecture (see [23]) are also true. The Petersen Coloring Conjecture is a starting point for research in several directions. Different aspects of it are studied and partial results are proved, see for instance [3, 7, 9, 11, 16, 18, 22].

Analogously to the case $r = 3$, if all elements of $\mathcal{H}_r$ would satisfy the generalized Berge-Fulkerson Conjecture, then every r-graph would satisfy it. Mazzuoccolo et al. [17] asked whether there exists a connected r-graph H such that $H \prec G$ for every (simple) r-graph G, for all $r \geq 3$.

The following theorem is our main result on H-coloring of r-graphs.

Theorem 1.3. *Either $\mathcal{H}_3 = \{P\}$ or $\mathcal{H}_3$ is a unique infinite set. Moreover, if $r \geq 4$, then $\mathcal{H}_r$ is a unique infinite set.*

A similar result is obtained for simple r-graphs.

Order structure

Jaeger [8] initiated the study of the Petersen Coloring Conjecture in terms of partial ordered sets. DeVos, Nešetřil and Raspaud [3] studied cycle-continuous mappings and asked whether there is an infinite set $\mathcal{G}$ of bridgeless graphs such that every two of them are cycle-continuous incomparable, i.e. there is no cycle-continuous mapping between any two graphs in $\mathcal{G}$. Šámal [22] gave an affirmative answer to the above question by constructing such an infinite set $\mathcal{G}$ of bridgeless cubic graphs. In fact, he also mentioned that this result can be considered in view of a quasi-order induced by cycle-continuous mappings on the set of bridgeless cubic graphs. That is, this quasi-ordered set contains infinite antichains.

For every integer $r \geq 1$, H-colorings give a quasi-order on the set of r-graphs, which is denoted by $(\mathcal{G}_r, \prec)$. Thus, Theorem 1.3 can be restated as follows.

Theorem 1.3'. *For $r = 3$, either $\mathcal{H}_3 = \{P\}$ or $\mathcal{H}_3$ is a unique infinite antichain in $(\mathcal{G}_3, \prec)$. For each $r \geq 4$, $\mathcal{H}_r$ is a unique infinite antichain in $(\mathcal{G}_r, \prec)$.*

1.2 Variants and equivalent statements

In this part we focus on some variants and combinations of the aforementioned conjectures.

The following conjecture, which naturally generalizes the edge-coloring version of Tutte's 4-flow conjecture for cubic graphs to r-graphs, is attributed to Seymour in [4].

Conjecture 1.4. *Let G be an r-graph. If G has no Petersen-minor, then G is class 1.*

The following two conjectures are weak versions of Conjecture 1.1.

Conjecture 1.5. *There is a $t \geq 1$ such that for all $r \geq 1$ every r-graph has a (t, r)-PM.*

Conjecture 1.6. *For every $r \geq 1$ there is a $t_r \geq 1$ such that every r-graph has a (t_r, r)-PM.*

Clearly, Conjecture 1.1 implies Conjecture 1.5, which implies Conjecture 1.6. The reductions of these two conjectures to planar r-graphs or to Petersen-minor-free r-graphs are also open.

Conjecture 1.7. *1. There is a $t \geq 1$ such that for all $r \geq 1$ every planar r-graph has a (t, r)-PM.*
2. For every $r \geq 1$ there exists a $t_r \geq 1$ such that every planar r-graph has a (t_r, r)-PM.

Conjecture 1.8. *1. There is a $t \geq 1$ such that for all $r \geq 1$ every Petersen-minor-free r-graph has a (t, r)-PM.*
2. For every $r \geq 1$ there exists a $t_r \geq 1$ such that every Petersen-minor-free r-graph has a (t_r, r)-PM.

Clearly, if Conjecture 1.2 is true, then both statements of Conjecture 1.7 are true for every $r \geq 1$. Similarly, if Conjecture 1.4 is true, then both statements of Conjecture 1.8 are true for every $r \geq 1$.

Every K_4-minor-free graph is planar and Conjecture 1.2 is confirmed for K_4-minor-free r-graphs by Seymour [21]. Natural subsets of the set of Petersen-minor-free r-graphs are the sets of K_5-minor-free and of $K_{3,3}$-minor-free r-graphs. We will prove that the restriction of Conjecture 1.8 to these sets of r-graphs is equivalent to Conjecture 1.7. The following theorem is the main result. The underlying simple graph of an r-graph G is denoted by G_s. Further, the crossing number of a graph H is denoted by $cr(H)$.

Theorem 1.9. *For any $t \geq 1$ and $r \geq 1$, the following statements are equivalent.*

1. *Every planar r-graph has a (t, r)-PM.*
2. *Every K_5-minor-free r-graph has a (t, r)-PM.*
3. *Every $K_{3,3}$-minor-free r-graph has a (t, r)-PM.*
4. *Every r-graph G with $cr(G_s) \leq 1$ has a (t, r)-PM.*

Conjecture 1.2 is proved for $r \leq 8$ in a sequel of papers [1, 2, 4, 6]. Thus, for $r \leq 8$, every r-graph with no K_5-minor or with no $K_{3,3}$-minor or with underlying simple graph with crossing number at most 1 has a (t, r)-PM for every $t \geq 1$.

We finish with the following statement on the structure of possible minimum counterexamples to aforementioned conjectures.

Theorem 1.10. *If H is a possible minimum counterexample to one of the statements of Conjectures 1.2 - 1.8 or to any of statements 1, 2, or 3 of Theorem 1.9, then H is a 3-connected r-graph and it does not contain a non-trivial tight edge-cut.*

2 Sketch of the proof of Theorem 1.3

We first characterize $\mathcal{H}_r$ for every $r \geq 3$.

Theorem 2.1. *Let $r \geq 3$ and let G be a connected r-graph. The following statements are equivalent.*

1) $G \in \mathcal{H}_r$.

2) The only connected r-graph coloring G is G itself.

3) G cannot be colored by a smaller r-graph.

The following corollary is a direct consequence of Theorem 2.1.

Corollary 2.2. *For every $r \geq 3$, there exists only one inclusion-wise minimal r-complete set, i.e. $\mathcal{H}_r$ is unique.*

Next we show that if $\mathcal{H}_r$ has more than one element, then it has infinitely many elements. Let G and G' be two disjoint r-graphs of class 2 with $e \in E(G)$ and $e' \in E(G')$. Denote by $(G,e)|(G',e')$ the set of all graphs obtained from G by replacing the edge e of G by (G',e'), that is, deleting e from G and e' from G', and then adding two edges between $V(G)$ and $V(G')$ such that the resulting graph is regular. We use $G|(G',e')$ to denote the set of all graphs obtained from G by replacing each edge of G by (G',e').

Theorem 2.3. *Let $\mathcal{M}$ be a multiset of $r-3$ perfect matchings of P, where $r \geq 3$, and let $e_0 \in E(P^{\mathcal{M}})$, where $P^{\mathcal{M}}\colon = P + \sum_{M \in \mathcal{M}} M$. Let G be an r-graph such that $G \not\cong P^{\mathcal{M}}$. If $G \in \mathcal{H}_r$, then $G|(P^{\mathcal{M}}, e_0) \subset \mathcal{H}_r$.*

Now we are going to show that if $r \geq 4$, then $\mathcal{H}_r$ has more than one element.

One major fact that we use is that every r-graph can be decomposed into a k-graph of class 1 and an $(r-k)$-regular graph, for a suitable $k \in \{1, \dots, r\}$. For every r-graph G let $\pi(G)$ be the largest integer t such that G has t pairwise disjoint perfect matchings. Let $r \geq 3$ and $k \in \{1, \dots, r\}$ be integers. Let $\mathcal{G}(r,k) = \{G\colon G \text{ is an } r\text{-graph with } \pi(G) = k\}$. Note that $\mathcal{G}(r, r-1) = \emptyset$, since every r-graph with $r-1$ pairwise disjoint perfect matchings is a class 1 graph and thus, it has r pairwise disjoint perfect matchings. If $k \leq r-2$, then the elements of $\mathcal{G}(r,k)$ are class 2 graphs and $\mathcal{G}(r,i) \cap \mathcal{G}(r,j) = \emptyset$, if $1 \leq i \neq j \leq r-2$. The sets $\mathcal{G}(r,k)$ are studied intensively in [12, 13]. Here we are interested in the subset of $\mathcal{G}(r,k)$ consisting of all such graphs with the smallest order. This set is denoted by $\mathcal{T}(r,k)$. Let $\mathcal{T}_r$ be the set of the smallest r-graphs of class 2. By definition, $\mathcal{T}_r \subseteq \bigcup_{i=1}^{r-2} \mathcal{T}(r,i)$.

Lemma 2.4. *Let $r \geq 3$ and let G be an r-graph of class 2 such that $\pi(G') > \pi(G)$ for every r-graph G' with $|V(G')| < |V(G)|$. If H is a connected r-graph with $H \prec G$, then $H \cong G$.*

It follows from Lemma2.4 and Theorem 2.1 that $\mathcal{T}_r \subseteq \mathcal{H}_r$ for each $r \geq 3$. For two positive integers $k \leq n$, let $p'(n,k)$ be the number of partitions of n into k parts. Set $p'(0,k) = 1$.

Theorem 2.5. *If $3 \leq r \leq 8$, then $|\mathcal{T}_r| = p'(r-3,6)$, and if $r \geq 9$, then $|\mathcal{T}_r| > p'(r-3,6)$.*

Now Theorem 1.3 follows from Theorems 2.3, and 2.5, and the fact, that there are poorly matchable r-graphs for each $r \geq 4$. The proofs of the results announced in this extended abstract can be found in the full versions of the papers [14, 15].

References

[1] Maria Chudnovsky, Katherine Edwards, Ken-ichi Kawarabayashi, and Paul D. Seymour. Edge-colouring seven-regular planar graphs. *Journal of Combinatorial Theory, Series B*, 115:276–302, 2015.

[2] Maria Chudnovsky, Katherine Edwards, and Paul D. Seymour. Edge-colouring eight-regular planar graphs. *Journal of Combinatorial Theory, Series B*, 115:303–338, 2015.

[3] Matt DeVos, Jaroslav Nešetřil, and André Raspaud. On edge-maps whose inverse preserves flows or tensions. In *Graph Theory in Paris*, Trends Math., pages 109–138. Birkhäuser, Basel, 2007.

[4] Zdeněk Dvořák, Ken-ichi Kawarabayashi, and Dan Kráľ. Packing six T-joins in plane graphs. *Journal of Combinatorial Theory, Series B*, 116:287–305, 2016.

[5] Delbert Ray Fulkerson. Blocking and anti-blocking pairs of polyhedra. *Mathematical Programming*, 1(1):168–194, 1971.

[6] Bertrand Guenin. *Packing T-joins and edge colouring in planar graphs.* Faculty of Mathematics, University of Waterloo, 2003.

[7] Jonas Hägglund and Eckhard Steffen. Petersen-colorings and some families of snarks. *Ars Math. Contemp.*, 7(1):161–173, 2014.

[8] François Jaeger. On graphic-minimal spaces. *Ann. Discrete Math.*, 8:123–126, 1980.

[9] François Jaeger. On five-edge-colorings of cubic graphs and nowhere-zero flow problems. *Ars Combin.*, 20(B):229–244, 1985.

[10] François Jaeger. Nowhere-zero flow problems. In *Selected topics in graph theory, 3*, pages 71–95. Academic Press, San Diego, CA, 1988.

[11] Ligang Jin and Yingli Kang. Partially normal 5-edge-colorings of cubic graphs. *European J. Combin.*, 95:Paper No. 103327, 14, 2021.

[12] Yulai Ma, Davide Mattiolo, Eckhard Steffen, and Isaak H. Wolf. Pairwise disjoint perfect matchings in r-edge-connected r-regular graphs. *SIAM Journal on Discrete Mathematics*, 37(3):1548–1565, 2023.

[13] Yulai Ma, Davide Mattiolo, Eckhard Steffen, and Isaak H. Wolf. Edge-connectivity and pairwise disjoint perfect matchings in regular graphs. *Combinatorica*, 44(2):429–440, 2024.

[14] Yulai Ma, Davide Mattiolo, Eckhard Steffen, and Isaak H. Wolf. Sets of r-graphs that color all r-graphs. *Combinatorica*, 45(2):16, 2025.

[15] Yulai Ma, Eckhard Steffen, Isaak H. Wolf, and Junxue Zhang. Some conjectures on r-graphs and equivalences. *arXiv preprint arXiv:2411.01753*, 2024.

[16] Giuseppe Mazzuoccolo and Vahan V. Mkrtchyan. Normal edge-colorings of cubic graphs. *J. Graph Theory*, 94(1):75–91, 2020.

[17] Giuseppe Mazzuoccolo, Gloria Tabarelli, and Jean Paul Zerafa. On the existence of graphs which can colour every regular graph. *Discrete Applied Mathematics*, 337:246–256, 2023.

[18] François Pirot, Jean-Sébastien Sereni, and Riste Škrekovski. Variations on the Petersen colouring conjecture. *Electron. J. Combin.*, 27(1):Paper No. 1.8, 14, 2020.

[19] P. D. Seymour. On multicolourings of cubic graphs, and conjectures of Fulkerson and Tutte. *Proceedings of the London Mathematical Society. Third Series*, 38(3):423–460, 1979.

[20] Paul D. Seymour. Some unsolved problems on one-factorizations of graphs. In *Graph Theory and Related Topics*, pages 367–368. Academic Press, New York, 1979.

[21] Paul D. Seymour. Colouring series-parallel graphs. *Combinatorica*, 10(4):379–392, 1990.

[22] Robert Šámal. Cycle-continuous mappings—order structure. *J. Graph Theory*, 85(1):56–73, 2017.

[23] Cun-Quan Zhang. *Integer flows and cycle covers of graphs*, volume 205 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1997.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Sharp thresholds for higher powers of Hamilton cycles in random graphs

(Extended abstract)

Tamás Makai * Matija Pasch† Kalina Petrova ‡ Leon Schiller §

Abstract

For $k \geq 4$, we establish that $p = (e/n)^{1/k}$ is a sharp threshold for the existence of the k-th power H of a Hamilton cycle in the binomial random graph model. Our proof builds upon an approach by Riordan based on the second moment method, which previously established a weak threshold for H. This method expresses the second moment bound through contributions of subgraphs of H, with two key quantities: the number of copies of each subgraph in H and the subgraphs' densities. We control these two quantities more precisely by carefully restructuring Riordan's proof and treating sparse and dense subgraphs of H separately. This allows us to determine the exact constant in the threshold[1].

1 Introduction

Research devoted to problems concerning Hamilton cycles in the binomial random graph model $G(n, p)$[2] began with the introduction of random graphs by Erdős and Rényi [4] in 1960. The extensive history of inquiry into this topic is well exemplified by Frieze's survey [7]. The threshold for the appearance of the Hamilton cycle was established as $p = (1+o(1)) \ln n/n$ and in later works the hitting time and other more precise results were derived [1, 2, 4, 10, 11, 16].

Kühn and Osthus [12] were the first to investigate the threshold for the k-th power of a Hamilton cycle.[3] They proved the upper bound $n^{-1/2+\varepsilon}$ for the case $k = 2$. Since then, the

*Department of Mathematics, LMU Munich, Munich 80333 Germany. Email: makai@math.lmu.de.

†Email: matija.pasch@gmail.com

‡Institute of Science and Technology Austria (ISTA), Klosterneurburg 3400, Austria. kalina.petrova@ist.ac.at. Supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034413.

§Hasso Plattner Institute, University of Potsdam, Potsdam 14482, Germany. Email: leon.schiller@hpi.de

[1]The full proof is available at https://arxiv.org/abs/2502.14515.

[2]In the *binomial random graph* $G(n, p)$ on n vertices, each potential edge is included with probability $p = p(n)$ independently.

[3]The *k-th power* of a graph G is obtained by including an edge $\{u, v\}$ if the distance between u and v in G is at most k.

threshold for the square of a Hamilton cycle has been the focus of a great deal of research. Nenadov and Škorić [14] established the upper bound $O(\ln^4 n/n^{1/2})$, which was subsequently improved to $O(\ln^3 n/n^{1/2})$ by Fischer, Škorić, Steger, and Trujić [5], and then to $O(\ln^2 n/n^{1/2})$ by Montgomery [13], all employing the absorption method and connection techniques. The correct order of magnitude $\Theta(n^{-1/2})$ was finally shown by Kahn, Narayanan, and Park [9], who followed a different approach, based on the resolution of the fractional 'expectation threshold' conjecture [6]. They conjectured that the sharp threshold should be $\sqrt{e/n}$ — a hypothesis that has attracted significant interest in the community [7, 15].

Turning to higher powers of Hamilton cycles, Kühn and Osthus [12] observed that for $k \geq 3$, the correct order of magnitude $\Theta(n^{-1/k})$ of the threshold for the k-th power follows from the first moment bound and a result of Riordan [17], which applies to a broad class of spanning subgraphs. While [3] suggests that Riordan's theorem also implies the sharp threshold, this is not directly the case, as applying [17] requires $pn^{1/k} \to \infty$. In this paper, we extend Riordan's approach to show that the sharp threshold for the k-th power of a Hamilton cycle for $k \geq 4$ matches the lower bound given by the first moment method.

Theorem 1. *For all $k \geq 4$, we have that $p^* = (e/n)^{1/k}$ is a sharp threshold for the existence of the k-th power H of a Hamilton cycle in $G(n,p)$. That is, for all $\varepsilon > 0$ and all $p \leq (1-\varepsilon)p^*$, there is no copy of H in $G(n,p)$ whp*[4]*, and for all $p \geq (1+\varepsilon)p^*$, there is a copy of H in $G(n,p)$ whp.*

Independent of our result, Zhukovskii [18] presented a more general result that established a sharp threshold for a wider class of spanning regular subgraphs, which includes every power of the Hamilton cycle and in particular the case $k = 2$. He used a different method, namely fragmentation, which was also employed by Kahn, Narayanan, and Park in [9].

1.1 Proof strategy

Theorem 1 consists of two statements, the lower bound and the upper bound on the threshold for the containment of the k-th power of a Hamilton cycle in $G(n,p)$. The lower bound follows easily from the first moment method and has been known for a long time. In the remainder of the paper, we therefore focus on the upper bound.

Our approach is inspired by the method in [17], where Riordan gives sufficient conditions for the appearance of a large family of spanning subgraphs H in $G(n,p)$. The author makes use of the second moment method. As a consequence of that, the main objective becomes showing that the variance of the number X of copies of H in $G(n,p)$ is negligible relative to its squared expectation. The key idea then is to write the ratio between the variance of X and $(\mathbb{E}(X))^2$ as a sum of contributions of all subgraphs of H. This makes it much easier to estimate each term.

The analysis in [17] is not tight enough to yield the sharp threshold for powers H of Hamilton cycles. The first additional key idea we introduce is to treat sparse and dense subgraphs of H separately. For the contributions of sparse subgraphs, the bound from [17] is sufficient. However, for the contributions of dense subgraphs, we significantly improve the bound for our specific spanning graph, as opposed to Riordan's general result. For this, we rely on the fact that when $F \subset H$ has almost the same density as H, there are very few copies of F in H.

[4]A sequence E_n of events holds *with high probability* (whp), if the probability of E_n tends to one as n tends to infinity.

The final step is to sum up the bounds over every subgraph of H. Since we differentiate between dense and sparse subgraphs of H, it is not possible to factorize the sum of the bounds into contributions of the individual components (and then sum the bounds for connected subgraphs over the same vertex set) as in [17] (for details, see Section 2.2). Instead, we first establish the partial sums of the bounds over every subgraph with the same vertex set and component structure. We upper bound these partial sums in a different way, depending on the density of the densest graph with a fixed component structure. Summing these contributions gives a suitable upper bound on the variance of the number of copies of H, which together with the second moment method implies our result.

2 Preliminaries and proof outline

We start by introducing the relevant notions and notation in Section 2.1. Next, we give a proof outline of our main result in Section 2.2.

2.1 Notions and notation

Throughout, we consider graphs with vertices in $[n] = \{1, \dots, n\}$ for n sufficiently large. Let K_n be the complete graph on the vertex set $[n]$. The uniform Erdős–Rényi model $G(n, M)$ is the random graph on the vertex set $[n]$ in which M of the $\binom{n}{2}$ potential edges are sampled uniformly at random. We use the standard Bachmann-Landau notation $o(\cdot)$, $O(\cdot)$, $\Theta(\cdot)$, $\Omega(\cdot)$ and $\omega(\cdot)$, exclusively with respect to n. We omit rounding of real numbers to their nearest integers whenever it is not essential for the argument.

For a graph F, let $V(F)$ be the set of vertices of F, let $E(F)$ be the set of edges of F, and let $C(F)$ be the set of connected components of F. Let $|F| = |V(F)|$ be the order of F, $c(F) = |C(F)|$ be the number of components of F, and $e(F) = |E(F)|$ be the number of edges of F. For graphs F and G, we denote by $F \subset G$ the statement that F is a (not necessarily proper) subgraph of G, and similarly for a subset T of a set S.

Since our proof decomposes various subgraphs of H into different types of connected components, we need some specific definitions for that purpose. Let $r(F) = |F| - c(F)$ be the rank of F. We call an edge *isolated* if it is the only edge in its component. Let

$$F^\circ = \bigcup_{C \in C(F) : |C| > 2} C$$

be F without isolated vertices and isolated edges. If $e(F^\circ) > 0$, we define the density of F as

$$\gamma(F) = \frac{e(F^\circ)}{|F^\circ| - 2c(F^\circ)} = \frac{e(F^\circ)}{r(F^\circ) - c(F^\circ)}.$$

A graph F is *good* if $V(F) = [n]$ and for all $C \in C(F)$, we have $|C| \neq 2$, i.e. F has no isolated edges. For graphs F, G, let $X_F(G)$ be the number of copies of F in G.

We fix $k \geq 4$ and we denote by H the k-th power of the cycle with vertices $[n]$ and edges $\{v, v+1\}$ for $v \in [n]$, with sums considered modulo n. That is, the edges of H are given by $\{v, v+i\}$ for $v \in [n]$ and $i \in [k]$.

2.2 Proof outline

The proof of the upper bound in Theorem 1 partially follows the proof strategy in [17]. In particular, we also work in $G(n, M)$ instead of $G(n, p)$. As noted in the remark on page 131 in [17], this is necessary since the variance of the number of copies of H in $G(n, p)$ is too high for the second moment method to yield the desired result. Indeed, when considering the probability that two overlapping copies of H both appear, one can observe that the appearance of one copy significantly increases the probability that the other copy is also present. Another way to think of this phenomenon is that planting H in $G(n, p)$ significantly increases the expected number of edges, whereas in $G(n, M)$ that is not the case. We now state our main result for $G(n, M)$.

Theorem 2. *For all $k \geq 4$ and $\varepsilon > 0$, the following holds for the sequence $p^* = (e/n)^{1/k}$. For all $p \geq (1 + \varepsilon)p^*$, the uniform Erdős-Rényi random graph $G(n, \binom{n}{2}p)$ contains a copy of the k-th power H of a Hamilton cycle whp.*

Note that the upper bound in Theorem 1 follows from Theorem 2 using the concentration of the number of edges.

From this point on, we fix $\varepsilon > 0$ and $p = (1+\varepsilon)(e/n)^{1/k}$. We start by applying the second moment method, as in [17]. Letting X denote the number of copies of H in $G(n, \binom{n}{2}p)$, we have by Chebyshev's inequality that

$$\mathbb{P}(X = 0) \leq \mathbb{P}\left((X - \mu)^2 \geq \mu^2\right) \leq \frac{\mathrm{Var}(X)}{\mu^2}, \tag{1}$$

where $\mu = \mathbb{E}(X)$. Let $f = 1 + \frac{\mathrm{Var}(X)}{\mu^2}$, where f depends on p. Our aim is to show $f = 1 + o(1)$, which in turn implies $\mathbb{P}(X = 0) = o(1)$. We start by stating an upper bound on f given in [17], which applies to our setting. For this purpose, we define the following quantity. Note that the sum below considers only spanning subgraphs of H, but these may have isolated vertices:

$$S = \sum_{\substack{F \subset H,\\ |F| = n}} \left(\frac{1}{p} - 1\right)^{e(F)} \left(1 + n^{-1/2}\right)^{r(F)} \frac{X_F(H)}{X_F(K_n)}.$$

A key idea in Lemma 4.2 of [17] is to rewrite f as a sum of contributions over all spanning subgraphs of H, which in our case leads to

$$f \leq \left(1 + o(1)\right) \exp\left(-\frac{1-p}{p} \cdot \frac{2k^2 n}{n-1}\right) S.$$

Next, the proof in [17] bounds S in terms of a sum S' of the contributions of all good subgraphs, i.e., graphs with no isolated edges, each multiplied by a factor that accounts for adding all the possible sets of isolated edges to it. Here we make our first modification to [17], which is that we define S' so that it gives a tighter upper bound on S. With $\sum'$ denoting the summation only over good graphs, let

$$S' = \sum_{\substack{F \subset H,\\ |F| = n}}' \phi(F), \qquad \phi(F) = (p^{-1} - 1)^{e(F)} \nu^{r(F)} \frac{X_F(H)}{X_F(K_n)}, \qquad \nu = \left(1 + \frac{1}{\sqrt{n}}\right) \exp\left(\frac{14k^2}{\sqrt{np}}\right), \tag{2}$$

be our version of S' (cf. the math display above Lemma 4.3 in [17], where ν is replaced by 2). Note that $\nu = 1 + o(1)$. Similarly to [17], we show the following bound.

Lemma 3. *We have $f \leq (1 + o(1))S'$.*

The proof of the lemma mainly follows and slightly improves upon the corresponding result in [17]. It only remains to show that $S' \leq 1 + o(1)$.

The next step is to establish an upper bound for $X_F(H)/X_F(K_n)$. We now start to significantly deviate from the proof in [17], since we treat sparse and dense subgraphs $F \subset H$ differently. We say that $F \subset H$ is ζ'-dense if $e(F^\circ) > 0$ and $\gamma(F) \geq \gamma(H) - \zeta'$. Conversely, we refer to $F \subset H$ as ζ'-sparse if $e(F) > 0$, and either $e(F^\circ) = 0$, or $\gamma(F) < \gamma(H) - \zeta'$. For this distinction based on density, let $\zeta := \zeta(n)$ tend to 0 sufficiently slowly, more precisely, let ζ be of order $\omega(\ln^{-1}(n))$.

In any case, we can use (a slightly enhanced version of) Lemma 4.4 in [17]. While this bound is rather weak, it proves to be useful whenever the density of F is sufficiently small.

Lemma 4. *For good spanning subgraphs $F \subset H$, we have*

$$\frac{X_F(H)}{X_F(K_n)} \leq \left(\frac{3ke}{n}\right)^{r(F)}.$$

On the other hand, for $(1 + \gamma(H))\zeta$-dense good subgraphs, we obtain a much stronger bound on $\frac{X_F(H)}{X_F(K_n)}$. This improvement over the bound in [17] is specific to powers of Hamilton cycles, and it is vital to our proof.

Lemma 5. *If $F \subset H$ is a $(1 + \gamma(H))\zeta$-dense good spanning subgraph, we have*

$$\frac{X_F(H)}{X_F(K_n)} \leq \left(\left(1 + O\left(\sqrt{\zeta}\right)\right)\frac{e}{n}\right)^{r(F)}.$$

There are two crucial properties of dense subgraphs that facilitate the improvement in Lemma 5 compared to Lemma 4. For one, since the subgraph is dense, most vertices must have full degree $2k$ which restricts the number of possible embeddings into H. Additionally, since small subgraphs (that is, components on few vertices) are sparse, dense subgraphs cannot have too many small components and thus not too many components overall.

The last observation is no longer true when $k \leq 3$, as then some small subgraphs are dense. For example, we have $\gamma(K_3) = 3$ for the triangle K_3, while the density of the k-th power of a Hamilton cycle is $k + o(\zeta)$. This breaks the bound on the number of components used in the proof of Lemma 5.

Turning back to the case $k \geq 4$, we give an overview of the remainder of the proof to motivate what follows. The key idea is the following: Depending on the density of F, we apply Lemma 4 or 5 in order to bound $X_F(H)/X_F(K_n)$, which in turn provides the required bound on $\phi(F)$ and consequently also on S' as defined in (2). At first glance, one might think that with this in mind, following [17] would easily give our result. However, dealing with S' is non-trivial and presents some serious challenges. In [17], this is achieved using a bound similar to Lemma 4, which can be factorized over the components of F (as both $r(F)$ and $e(F)$ are additive over them) and thus every component can be considered separately. Then the partial sums over all connected subgraphs on the same vertex set are evaluated, before establishing an upper bound on S'.

In our case, there is no simple way to create a factorization over the components, where every component contributes in the same manner to every subgraph it is contained in. We require different upper bounds on $X_F(H)/X_F(K_n)$ depending on the density of F, and since the same component can appear in subgraphs with a wide range of densities, any such factorization should depend on the whole of F, not just on an individual component. One might think that this issue could be circumvented by differentiating based on the density of each component instead of the density of the entire subgraph. However, with this approach, one loses control over the number of components of a subgraph, which is crucial for the argument.

In order to overcome both challenges — factorizing the sum and bounding the number of components — we change the order of the factorization step and the step that groups the subgraphs by their vertex set. In [17] this reversal would correspond to taking the partial sums over all subgraphs on the same vertex set with the same component structure.

The densest of these subgraphs will play a vital role. For any $F \subset H$ define the *completion* $F^* \subset H$ of F as the unique graph given by

$$F^* = \bigcup_{C \in C(F)} H\big[V(C)\big],$$

where $H[V]$ denotes the subgraph induced by H on V. Our next goal is to determine upper bounds on the partial sums, which can be factorized, considering ζ-sparse and ζ-dense completions separately. If the completion is ζ-sparse, then we can use Lemma 4 as the upper bound for $X_F(H)/X_F(K_n)$ in every summand. On the other hand, for a ζ-dense completion we establish an upper bound which roughly corresponds to using Lemma 5 as an upper bound on $X_F(H)/X_F(K_n)$ in every summand. This can only be achieved because even the weak bound Lemma 4 for the sparse graphs with this ζ-dense completion gives an insignificant contribution. This then allows for factorization and the proof is completed by summing separately over the ζ-sparse and the ζ-dense completions.

We now return to the technical details of the proof. As mentioned above, we split S' into more than two parts, based not only on the density of the individual F, but also on the density of their *completions* F^*. More precisely, we split the sum S' into

$$S' = 1 + S_{\mathrm{s}} + S_{\mathrm{ds}} + S_{\mathrm{dd}}, \tag{3}$$

where 1 accounts for the empty graph. For non-empty good subgraphs F, we differentiate based on the density of F and its completion F^*. The contribution S_{s} is the sum of $\phi(F)$ over all F with ζ-sparse completions F^*. The contribution S_{ds} is the sum of $\phi(F)$ over all F with ζ-dense completions F^* such that $\gamma(F)/\gamma(F^*) < 1 - \zeta$. The last contribution S_{dd} is the sum of $\phi(F)$ over all F with ζ-dense completions F^* such that $\gamma(F)/\gamma(F^*) \geq 1 - \zeta$. We will show that the three contributions are in $o(1)$, employing a different argument in each case.

In the case of S_{s}, we consider graphs with sparse completions. The lower densities allow us to compensate for the weaker bound on $X_F(H)/X_F(K_n)$ provided by Lemma 4.

Proposition 6. *We have $S_{\mathrm{s}} = o(1)$.*

The more interesting case occurs when we are dealing with good graphs that have a dense completion. As mentioned above, this roughly corresponds to using the stronger upper bound on $X_F(H)/X_F(K_n)$ from Lemma 5 and aggregating over subgraphs. This holds for the summands of S_{dd} (as any subgraph involved is $(1 + \gamma(H))\zeta$-dense) and we show that this still holds when including S_{ds}.

Proposition 7. *Let $\mathcal{F}^*$ denote the set of all good completions in H and let*

$$\bar{S}_{\mathrm{d}} := \sum_{\substack{F \in \mathcal{F}^* \\ F \text{ is } \zeta\text{-dense}}} p^{-e(F)} \left((1 + O(\sqrt{\zeta})) \frac{e}{n} \right)^{r(F)}.$$

Then we have $S_{\mathrm{ds}} + S_{\mathrm{dd}} \leq (1 + o(1))\bar{S}_{\mathrm{d}}$.

This allows us to limit our attention to $\bar{S}_{\mathrm{d}}$. Finally, by extending the summation domain, we can factorize the sum, which in turn yields sums over the components and thereby allows us to focus on a sum over all connected subgraphs. Distinguishing between ζ-sparse and ζ-dense subgraphs then establishes that $\bar{S}_{\mathrm{d}}$ is small.

Proposition 8. *We have $\bar{S}_{\mathrm{d}} = o(1)$.*

We are now ready to give the proof of our main theorem for $G(n, \binom{n}{2}p)$.

Proof of Theorem 2. Recall that we fixed $p = (1 + \varepsilon)(e/n)^{1/k}$. It suffices to prove that $G(n, \binom{n}{2}p)$ whp contains a copy of H. Monotonicity of subgraph containment then implies the assertion for $G(n, \binom{n}{2}p')$ with $p' \geq p$.

By Lemma 3, we have $f \leq (1 + o(1))\, S'$. Propositions 6 to 8 and (3) establish that $S' = 1 + o(1)$, so by (1) and the fact that $f = \frac{\mathrm{Var}(X)}{\mu^2} + 1$, we have

$$\mathbb{P}(X = 0) \leq f - 1 = o(1),$$

where X is the number of copies of H in $G(n, \binom{n}{2}p)$. □

3 Acknowledgments

This project was initiated during the research workshop in Buchboden, June 2024. In particular, we want to thank Angelika Steger and Konstantinos Panagiotou, who paved the way for this project.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi, *First occurrence of Hamilton cycles in random graphs*, Cycles in graphs (Burnaby, B.C., 1982), vol. 115, North-Holland, Amsterdam, 1985, pp. 173–178.

[2] B. Bollobás, *The evolution of sparse graphs*, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 35–57.

[3] A. Dudek, C. Reiher, A. Ruciński, and M. Schacht, *Powers of Hamiltonian cycles in randomly augmented graphs*, Random Structures & Algorithms 56, no. 1 (2020): 122-141.

[4] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutato Int. Kozl. 5 (1960), 17–61.

[5] M. Fischer, N. Škorić, A. Steger, and M. Trujić, *Triangle resilience of the square of a Hamilton cycle in random graphs*, Journal of Combinatorial Theory, Series B 152 (2022), 171-220.

[6] K. Frankston, J. Kahn, B. Narayanan, and J. Park, *Thresholds versus fractional expectation-thresholds*, Annals of Mathematics 194.2 (2021), 475-495.

[7] A. Frieze, *Hamilton cycles in random graphs: a bibliography*, arXiv preprint, arXiv:1901.07139.

[8] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, John Wiley & Sons, 2000.

[9] J. Kahn, B. Narayanan, and J. Park, *The threshold for the square of a Hamilton cycle*, Proceedings of the American Mathematical Society 149.8 (2021): 3201-3208.

[10] J. Komlós and E. Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Math. 43 (1983), 55–63.

[11] A. D. Korshunov, *Solution of a problem of Erdős and Rényi on Hamiltonian cycles in nonoriented graphs*, Soviet Math. Doklady 17 (1976), 760-764.

[12] D. Kühn and D. Osthus, *On Pósa's conjecture for random graphs*, SIAM Journal on Discrete Mathematics 26 (2012), 1440-1457.

[13] R. Montgomery, *Topics in random graphs*, Lecture notes (2018).

[14] R. Nenadov and N. Škorić, *Powers of Hamilton cycles in random graphs and tight Hamilton cycles in random hypergraphs*, Random Structures Algorithms 54 (2019), 187–208.

[15] W. Perkins, *Searching for (sharp) thresholds in random structures: where are we now?*, arXiv preprint, arXiv:2401.01800, 2024.

[16] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Math. 14 (1976), 359–364.

[17] O. Riordan, *Spanning Subgraphs of Random Graphs*, Combinatorics, Probability and Computing, 2000.

[18] M. Zhukovskii, *Sharp thresholds for spanning regular graphs*, arXiv preprint, arXiv:2502.14794.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Effective Khovanskii, Ehrhart Polytopes, and the Erdős Multiplication Table Problem

(Extended abstract)

Anna Margarethe Limbach* Robert Scheidweiler† Eberhard Triesch‡

Abstract

Let $P(k, n)$ be the set of products of k factors from the set $\{1, \dots, n\}$. In 1955, Erdős posed the problem of determining the order of magnitude of $|P(2, n)|$ and proved that $|P(2, n)| = o(n^2)$ for $n \to \infty$. In 2015, Darda and Hujdurović asked whether, for each fixed n, $|P(k, n)|$ is a polynomial in k of degree $\pi(n)$ - the number of primes not larger than n. Recently, Granville, Smith and Walker published an effective version of Khovanskii's Theorem. We apply this new result to show, that for each integer n, there is a polynomial q_n of degree $\pi(n)$ such that $|P(k, n)| = q_n(k)$ for each $k \geq n^2 \cdot \left(\prod_{m=1}^{\pi(n)} \log_{p_m}(n)\right) - n + 1$. Moreover, we give an upper estimate of the leading coefficient of q_n.

1 Introduction

Let

$$P(k, n) = \left\{ \prod_{j=1}^{n} j^{\lambda_j} \middle| \lambda_j \in \mathbb{N}_0, \sum_{j=1}^{n} \lambda_j = k \right\}$$

be the set of products of k factors from the set $\{1, \dots, n\}$ and denote by $p(k, n)$ its cardinality. In 1955 Erdős considered the problem of estimating $p(2, n)$ and showed that $p(2, n) = o(n^2)$ (cf. [6]). His estimates for this so-called Erdős Multiplication Table Problem were considerably refined in 2008 by Ford in [7] and, for general fixed k, by Koukoulopoulos [10]. Motivated by a graph coloring problem, namely determining the product irregularity strength of complete bipartite graphs, Darda and Hujdurović asked in [3] whether, for a fixed n, the quantity $p(k, n)$ is a polynomial in k of degree $\pi(n)$ - the number of primes not larger than n. They showed

*Institute of Computer Science, Czech Academy of Science, Pod Vodárenskou věží 271/2, 182 00 Praha 8, Czech Republic. E-mail: `limbach@cs.cas.cz`. Supported by Czech Science Foundation Project 21-21762X. With institutional support RVO:67985807

†Faculty of Electrical Engineering & Information Technology, Hochschule Düsseldorf - University of Applied Sciences, Münsterstraße 156, 40476 Düsseldorf, Germany.

‡Lehrstuhl II für Mathematik, RWTH Aachen University, Pontdriesch 10-12, 52062 Aachen, Germany.

this to be true for $n = 1, ..., 10$. We apply Khovanskii's Theorem and Ehrhart's Theorem to answer this question affirmatively for each n and for every k larger than some threshold k_n, which depends on n. Moreover, we give an expression as well as an upper bound for the polynomials leading coefficient. By applying a recent result from Granville, Smith and Walker, we establish an upper bound on k_n.

2 Reformulation by Prime Factorization

Let $p = (p_1, \ldots, p_{\pi(n)})$ be the $\pi(n)$-tuple of distinct prime numbers at most n. We define

$$M_n := \left\{ \alpha \in \mathbb{N}_0^{\pi(n)} \,\middle|\, \prod_{j=1}^{\pi(n)} p_j^{\alpha_j} \leq n \right\}$$

to be the set of $\pi(n)$-tuples of p-adic evaluations of numbers not exceeding n. Given that $\{1, \ldots, n\} = \{\prod_{j=1}^{\pi(n)} p_j^{\alpha_j} \mid \alpha \in M_n\}$ and that multiplication can be expressed as addition of exponents, it follows that

$$P(k, n) = \left\{ \prod_{j=1}^{\pi(n)} p_j^{\alpha_j} \,\middle|\, \alpha \in kM_n \right\},$$

where we use the notation

$$kA := \{a_1 + \ldots + a_k : a_i \in A \text{ for all } i\}$$

for a set A and a positive integer k. This implies $p(k, n) = |P(k, n)| = |kM_n|$.

However, the size of sumsets of the form kA is well understood. A famous theorem by Khovanskii says the following:

Theorem 2.1 ([9])**.** *Let $A \subseteq \mathbb{Z}^d$ be finite. There is a polynomial $P_A \in \mathbb{Q}[X]$ of degree at most d, and a threshold $N_{Kh}(A)$, such that $|NA| = P_A(N)$ for every integer $N > N_{Kh}(A)$.*

Thus we obtain the following intermediate result:

Lemma 2.2. *For each $n \in \mathbb{N}$ there is a polynomial $q_n \in \mathbb{Q}[X]$ of degree at most $\pi(n)$ and a threshold k_n such that $p(k, n) = q_n(k)$ for each $k \geq k_n$.*

In the following, we investigate q_n as well as the threshold k_n.

3 Relation to Ehrhart Polynomials

We now consider a *full dimensional lattice polytope* Q in $\mathbb{R}^d$, i. e., a polytope whose vertices are all integral and which contains $d + 1$ affine independent points. Denote by $\mathrm{int}(Q)$ the set of integral points in Q. Expanding Q by a factor of $t \in \mathbb{N}$ in each dimension, gives the scaled-up polytope $t \cdot Q := \{tq \mid q \in Q\}$ and the number of integral points in $t \cdot Q$ is denoted by $L(Q, t) = |\,\mathrm{int}(t \cdot Q)|$. Note that as Q is a polytope, we have

$$t \cdot Q = \{tq \mid q \in Q\} = \{q_1 + \ldots + q_t \mid q_i \in Q \text{ for all } i\} = tQ.$$

As was proven by Ehrhart in 1962 [4], the function $L(Q,t)$ is a polynomial in t of degree d, the Ehrhart polynomial of Q. It is well-known that the leading coefficient of $L(Q,t)$ is the d-dimensional volume of Q. Denote by $k \star Q = k(\mathrm{int}(Q))$ the set of all vectors in $\mathbb{R}^d$ which can be obtained as a sum of k integral vectors in Q. Clearly, $k \star Q \subseteq k \cdot Q$ and the polytope Q is called *integrally closed* if $k \star Q = \mathrm{int}(k \cdot Q)$ for all k. There are very simple polytopes which are not integrally closed. The following example is well-known ([1], Example 2.56 (c)):

Example 3.1. *Let* $Q := \{(x,y,z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, 6x + 10y + 15z \leq 30\}$. *As is easily verified, the integral points in* Q *are*

$$(5,0,0),(3,1,0),(2,0,1),(1,2,0),(0,3,0),(0,1,1),(0,0,2)$$

and component-wise smaller points in $\mathbb{N}_0^3$. *The point* $(4,2,1) = 2 \cdot (2,1,1/2)$ *is in* $2 \cdot Q$, *but not in* $2 \star Q$, *hence the simplex* Q *is not integrally closed.*

Now let Q_n denote the convex hull $H(M_n)$, which is a polytope of dimension $\pi(n)$. Then $k \star Q_n = kM_n$ and thus $p(k,n) = |kM_n| \leq L(Q_n, k)$. In fact, we checked that Q_n is integrally closed and hence the polynomials $L(Q_n,t)$ and $q_n(t)$ are equal for $n \leq 20$. By adapting the example above, we can, however, show that the polytopes Q_n are not integrally closed in general:

Example 3.2. *Choose some large prime* p *and let* $n = p^5$. *By Bertrand's postulate (see [2], easier proofs can be found in [11] and [5]), we can choose primes* q *and* r *such that* $n^{1/3} \leq q \leq 2n^{1/3}$ *and* $n^{1/2} \leq r \leq 2n^{1/2}$. *We write the points in* Q_n *in the form* $(x,y,z,...)$ *where the coordinates* x,y,z *correspond to the primes* p,q,r, *respectively. Then a nonnegative integral vector* $(x,y,z,0,...,0)$ *is in* Q_n *if and only if*

$$\frac{\log(p)}{\log(n)}x + \frac{\log(q)}{\log(n)}y + \frac{\log(r)}{\log(n)}z \leq 1.$$

This is equivalent to

$$\frac{1}{5}x + (\frac{1}{3} + \epsilon)x + (\frac{1}{2} + \delta)z \leq 1$$

where $0 \leq \epsilon, \delta \leq \log(2)/\log(n) < 1/60$ *for large* n. *As in the example above,* $(4,2,1,0,...,0)$ *is in* $2 \cdot Q_n$, *but not in* $2M_n$.

This example shows that the polynomial $q_n(t)$ does not always coincide with the Ehrhart polynomial $L(Q_n,t)$. However, we can show that both have the same degree and leading coefficient. This is the topic of the next theorem:

Theorem 3.3. *(i) For all* $k, n \in \mathbb{N}$, *the inequalities*

$$p(k,n) \leq L(Q_n,k) \leq p(k+\pi(n),n)$$

hold.

(ii) The polynomial q_n *has degree* $\pi(n)$ *with leading coefficient equal to the* $\pi(n)$*-dimensional volume of the polytope* Q_n.

Proof. (i) Let $d := \pi(n)$. We show that $kM_n \subseteq \operatorname{int}(k \cdot Q_n) \subseteq (k+d)M_n$, where the first inclusion was already explained above. For the second, recall that, by Caratheodory's Theorem, each point $x \in \operatorname{int}(k \cdot Q_n)$ can be written as a linear combination of the (integral) vertices of $k \cdot Q_n$. As one of those vertices is the origin, we obtain $x = \sum_{i=1}^{d} \alpha_i v_i$ with $v_i \in M_n$, $\alpha_i \geq 0$ for all i and $\sum_{i=1}^{d} \alpha_i \leq k$. It follows that $\tilde{x} := \sum_{i=1}^{d} \lceil \alpha_i \rceil v_i$ is in $(k+d)M_n$. The set M_n is a down-set, i. e., for each $y \in M_n$ and each $z \leq y$ (component-wise), we obtain $z \in M_n$. Consequently, kM_n and $(k+d)M_n$ are down-sets as well and $x \leq \tilde{x}$ implies that x is in $(k+d)M_n$ as well.

(ii) This follows immediately from (i) and from the properties of Ehrhart polynomials.

□

4 Application of the Effective Khovanskii Theorem

In a new preprint, Granville, Smith and Walker prove an effective version of Khovanskii's theorem, i. e., an upper bound on N_{Kh}:

Theorem 4.1 ([8])**.** *Let $A \subseteq \mathbb{Z}^d$ be finite and let $H(A) \subseteq \mathbb{R}^d$ be its convex hull. Then,*

$$N_{Kh}(A) \leq d!|A|^2 \operatorname{Vol}(H(A)) - |A| + 1.$$

By the injectivity of the logarithm on $\mathbb{R}_{\geq 0}$ we obtain

$$M_n = \left\{ \alpha \in \mathbb{N}_0^{\pi(n)} \;\middle|\; \sum_{j=1}^{\pi(n)} \alpha_j \cdot \log(p_j) \leq \log(n) \right\},$$

thus, it is easy to see that

$$H(M_n) \subseteq \left\{ x \in \mathbb{R}_{\geq 0}^{\pi(n)} \;\middle|\; \sum_{j=1}^{\pi(n)} x_j \cdot \log(p_j) \leq \log(n) \right\}$$

and hence $\operatorname{Vol}(H(M_n)) \leq \frac{1}{\pi(n)!} \prod_{j=1}^{\pi(n)} \log_{p_j}(n)$.

We obtain the following main result:

Theorem 4.2. *For each $n \in \mathbb{N}$ there is a polynomial $q_n \in \mathbb{Q}[t]$ of degree $\pi(n)$ with leading coefficient $\operatorname{Vol}(H(M_n))$ such that $p(k,n) = q_n(k)$ for each $k \geq n^2 \cdot \left(\prod_{m=1}^{\pi(n)} \log_{p_m}(n)\right) - n + 1$. Furthermore, the leading coefficient fulfills $\operatorname{Vol}(H(M_n)) \leq \frac{1}{\pi(n)!} \prod_{j=1}^{\pi(n)} \log_{p_j}(n)$.*

5 Future Research

The result by Granville, Smith, and Walker is quite general, as it can be applied to every finite subset of $\mathbb{Z}^d$. We wonder, whether one can get tighter bounds by modifying their proof and using more properties of M_n, for example that M_n is downward-closed or that M_n is the set of integer vectors of a polytope.

References

[1] Winfried Bruns and Joseph Gubeladze. *Polytopes, rings, and K-theory*, volume 27. Springer, 2009.

[2] Pafnuty Lvovich Chebyshev. Mémoire sur les nombres premiers. *J. Math. Pures Appl.*, 17:366–390, 1852.

[3] Ratko Darda and Ademir Hujdurović. On bounds for the product irregularity strength of graphs. *Graphs and Combinatorics*, 31(5):1347–1357, 2015.

[4] Eugene Ehrhart. Sur les polyèdres rationnels homothétiques à n dimensions. *CR Acad. Sci. Paris*, 254:616, 1962.

[5] Paul Erdős. Beweis eines satzes von tschebyschef. *Acta Scientifica Mathematica*, 5:194–198, 1932.

[6] Paul Erdős. Some remarks on number theory. *Riveon Lematematika*, 9:45–48, 1955.

[7] Kevin Ford. The distribution of integers with a divisor in a given interval. *Annals of mathematics*, pages 367–433, 2008.

[8] Andrew Granville, Jack Smith, and Aled Walker. Improved stability for the size and structure of sumsets. *arXiv preprint arXiv:2406.03275*, 2024.

[9] Askold Georgievich Khovanskii. Newton polyhedron, hilbert polynomial, and sums of finite sets. *Functional Analysis and Its Applications*, 26(4):276–281, 1992.

[10] Dimitris Koukoulopoulos. Localized factorizations of integers. *Proceedings of the London Mathematical Society*, 101(2):392–426, 2010.

[11] Srinivasa Ramanujan. A proof of bertrand's postulate. *Journal of the Indian Mathematical Society*, 11(181-182):27, 1919.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

First-order convergence law for small permutation classes

(Extended abstract)

Klas Markström* Eero Räty†

Abstract

A permutation class $\mathcal{F}$ is said to be small if it has growth-rate less than κ, where $\kappa \approx 2.20557$. Vatter established a structural classification result on small permutation classes and conjectured that every small permutation class has a rational generating function. The conjecture was later settled by Albert, Ruškuc and Vatter.

Our main result in this extended abstract is that every small permutation class satisfies a first-order convergence law. Given a small permutation class $\mathcal{F}$, we construct a regular language $L_{\mathcal{F}}$ and a length-preserving bijection $f_{\mathcal{F}} : L_{\mathcal{F}} \to \mathcal{F}$ that is a first-order sentence homomorphism, i.e. for every first-order sentence Ψ in $\mathcal{F}$ there exists a first-order sentence Φ in $L_{\mathcal{F}}$ so that

$$f_{\mathcal{F}}^{-1}\left(\{\sigma \models \Psi : \sigma \in \mathcal{F}\}\right) = \{w \models \Phi : w \in L_{\mathcal{F}}\}.$$

Combining this with the well known first-order closure property of regular languages, i.e. the fact that for every regular language L and first-order sentence Ψ, the set $K = \{w \models \Psi : w \in L\}$ is also a regular language, our main result follows. While our proof uses Vatter's structural classification of small permutation classes, the explicit construction of the languages $L_{\mathcal{F}}$ is highly technical and requires new ideas.

1 Introduction

A permutation class $\mathcal{F}$ is a subset of permutations that is closed under taking subpermutations. Given a permutation class $\mathcal{F}$, we define its *upper growth-rate* $\overline{\mathrm{gr}}\,(\mathcal{F})$ by setting

$$\overline{\mathrm{gr}}\,(\mathcal{F}) = \limsup_{n\to\infty} |\mathcal{F}_n|^{1/n}.$$

The Stanley-Wilf conjecture, which was proved by Marcus and Tardos [9], states that whenever $\mathcal{F}$ is a proper permutation class, then the upper growth-rate $\overline{\mathrm{gr}}\,(\mathcal{F})$ is finite. We say that a

*Institutionen för matematik och matematisk statistik, Umeå Universitet, Sweden. E-mail: `klas.markstrom@umu.se`

†Institutionen för matematik och matematisk statistik, Umeå Universitet, Sweden. E-mail: `eero.raty@umu.se`. Supported by postdoctoral grant 213-0204 from the Olle Engkvist Foundation and a postdoctoral grant from Osk. Huttunen Foundation

permutation class $\mathcal{F}$ is *small* if $\overline{\mathrm{gr}}(\mathcal{F}) < \kappa$, where $\kappa \approx 2.20557$ is the unique positive root of the polynomial $1 + 2x^2 - x^3$.

Small permutation classes have been studied extensively [1, 2, 5, 7, 8, 12], and their structure is understood very well. Vatter [12] proved a classification result for small permutation classes, and he proved that there are only countably many small permutation classes. He also conjectured that every small permutation class has a rational generating function, and commented that "... and this is exactly the sort of problem one expects to be able to handle with regular languages". Albert, Ruškuc and Vatter [5] later settled this conjecture and proved, among other results, that every small permutation class $\mathcal{F}$ has a rational generating function. However, their proof does not imply that every small permutation class $\mathcal{F}$ is in bijection with a regular language.

When studying first-order sentences on permutation classes, we use the standard logical model called Theory of Two orders (abberviated as TOTO). Informally speaking, we may view permutations as sets of points in the plane with no two contained on the same horizontal or vertical line. The model is equiped with two orders $\leq_1$ and $\leq_2$, which allow us to compare the two coordinates of the points in a permutation σ. For a formal description, see e.g. [3, 4]

Very recently, the study of first-order convergence laws on permutation classes has gathered attention. Braunfeld and Kukla [6] conjectured that every proper permutation class satisfies a first-order convergence law, and the conjecture was proved for the avoidance class Av (231) by Albert, Bouvel, Féray and Noy [4], and for Av (321) by Özdemir [11].

In this extended abstract, we give an outline of our upcoming work [10] in which we prove that small permutation classes satisfy a first-order convergence law, thus providing further support for the conjecture of Braunfeld and Kukla. More precisely, we prove that every small permutation class is in a bijection with a regular language. Furthermore, we show that the bijection can be chosen to be a sufficiently structured function in the following sense. We say that a function $f : A \to B$ is a *FOS-homomorphism* if for every first-order sentence Ψ in B, there exists a first-order sentence Φ in A for which $f^{-1}(\{b \models \Psi : b \in B\}) = \{a \models \Phi : a \in A\}$.

Theorem 1. *Let $\mathcal{F}$ be a small permutation class. Then there exists a regular language L and a length-preserving bijection $f : L \to \mathcal{F}$ that is a FOS-homomorphism. In particular, $\mathcal{F}$ satisfies a first-order convergence law.*

While we chose not to give a full background on the theory of regular languages, we recall the following standard fact, which states that the restriction of a regular language by a first-order sentence is also a regular language. Together with the fact that the function f in Theorem 1 can be chosen to be a FOS-homomorphism, the convergence law for small permutation classes follows quite easily.

Theorem 2. *Let L be a regular language, and let Ψ be a first-order sentence. Then $K = \{w \models \Psi : w \in L\} \subseteq L$ is also a regular language.*

For simplicity, we only give a brief outline of the proof of Theorem 1 for a specific family of small permutation classes denoted by $\tilde{\mathcal{O}}_k^{[d]}$, which we will now define. We start with some definitions. An *interval* in the permutation σ is a contiguous set of indices $I = [a, b] \cap \mathbb{Z}$ so that the set of values $\sigma(I) = \{\sigma(c) : c \in I\}$ is also a contiguous set of integers. Given a permutation τ of length k and permutations $\sigma_1, \ldots, \sigma_k$, we define the *inflation* of τ by $\sigma_1, \ldots, \sigma_k$ to be the permutation $\tau[\sigma_1, \ldots, \sigma_k]$ obtained by replacing each entry $\tau(i)$ by an interval that is order-isomorphic to the permutation σ_i. As an example, we have $213[132, 21, 1] = 354\,21\,6$. Given two permutation classes $\mathcal{F}$ and $\mathcal{G}$, we define the *inflation* of $\mathcal{F}$ by $\mathcal{G}$ to be the permutation class given by

$$\mathcal{F}[\mathcal{G}] = \{\tau[\sigma_1, \ldots, \sigma_k] : \tau \in \mathcal{F}_k \text{ and } \sigma_1, \ldots, \sigma_k \in \mathcal{G}\}.$$

For a permutation class $\mathcal{F}$, we define the sets $\mathcal{F}^{[d]}$ inductively by setting $\mathcal{F}^{[1]} = \mathcal{F}$ and $\mathcal{F}^{[d+1]} = \mathcal{F}\left[\mathcal{F}^{[d]}\right]$.

Define the *increasing oscillating sequence* to be the sequence (o_n) given by

$$4, 1, 6, 3, 8, 5, \ldots, 2k+2, 2k-1, \ldots.$$

Let ω_k denote the permutation induced by the points $\{(i, o_i) : 1 \leq i \leq k\}$, and let $\mathcal{O}_k$ denote the permutation class consisting of subpermutations of ω_k. Finally, we define $\tilde{\mathcal{O}}_k$ to be the permutation class that is the union of $\mathcal{O}_k$ and the set containing all finite monotone permutations. We also write $\mathcal{Q}_k$ for the set of non-monotone permutations in $\mathcal{O}_k$, and we note that $\mathcal{Q}_k$ is a finite set.

2 Encoding the elements of $\tilde{\mathcal{O}}_k^{[d]}$ by elements of a regular language

Instead of working directly with $\tilde{\mathcal{O}}_k^{[d]}$, we define the sets $\mathcal{S}_{d,k}$ as follows. When $d = 1$, we simply set $\mathcal{S}_{1,k} = \tilde{\mathcal{O}}_k$. Suppose that $d \geq 2$ and that we have defined $\mathcal{S}_{d-1,k}$. We set $\mathcal{S}_{d,k}$ to be the set of tuples $\left(\tau, \overline{\sigma}_1, \ldots, \overline{\sigma}_{|\tau|}\right)$, where $\tau \in \tilde{\mathcal{O}}_k$ and $\overline{\sigma}_i \in \mathcal{S}_{d-1,k}$ for every $i \in [|\tau|]$. Note that we have natural projection maps $\pi_{d,k} : \mathcal{S}_{d,k} \to \tilde{\mathcal{O}}_k^{[d]}$, where $\pi_{1,k}$ is the identity map and for $d \geq 2$, $\pi_{d,k}$ is defined by setting $\pi_{d,k}\left(\tau, \overline{\sigma}_1, \ldots, \overline{\sigma}_{|\tau|}\right) = \tau\left[\pi_{d-1,k}\left(\overline{\sigma}_1\right), \ldots, \pi_{d-1,k}\left(\overline{\sigma}_{|\tau|}\right)\right]$.

Our first goal is to construct a set of alphabets $\Sigma_{d,k}$ and a function $f_{d,k} : \Sigma_{d,k}^* \to \mathcal{S}_{d,k} \cup \{\star\}$, where the special $\star$-entry is used as the image of certain words that we deem unsuitable to be used to decode elements in $\mathcal{S}_{d,k}$. As our goal is to show that $L_{d,k} := \Sigma_{d,k}^* \setminus f_{d,k}^{-1}(\star)$ is a regular language, we need the set $f_{d,k}^{-1}(\star)$ to be sufficiently structured. By Theorem 2, it is certainly sufficient for us to construct the function $f_{d,k}$ so that $f_{d,k}^{-1}(\star)$ can be described with a first-order sentence.

We perform the construction by induction on d. For the base case $d = 1$, we take $\Sigma_{1,k} = \{a_I, a_D\} \cup \bigcup_{\tau \in \mathcal{Q}_k} \{a_{\tau,i} : 1 \leq i \leq |\tau|\}$, and we define the function $f_{1,k}$ by setting

1. $f_{1,k}\left(a_I^n\right) = 12\ldots n$ for all n, $f_{1,k}\left(a_D^n\right) = n(n-1)\ldots 1$ for all $n \geq 2$ and $f_{1,k}\left(a_D\right) = \star$.
2. $f_{1,k}\left(a_{\tau,1} \ldots a_{\tau,|\tau|}\right) = \tau$ for every $\tau \in \mathcal{Q}_k$.
3. $f_{1,k}(w) = \star$ for any other $w \in \Sigma_{d,k}^*$.

Since $\mathcal{Q}_k$ is a finite set, it is not too hard to write down a first-order sentence $\Psi_{1,k}$ so that $w \models \Psi_{1,k}$ if and only if $f_{1,k}(w) = \star$. Furthermore, we note that the pre-image of any other element than $\star$ contains exactly one word. This is the motivation behind the choice of $f_{1,k}(a_D) = \star$.

Having constructed $\Sigma_{d-1,k}$ and $f_{d-1,k}$, we enumerate the elements of $\Sigma_{d-1,k}$ as $\{a_1, \ldots, a_N\}$. We set $\Sigma_{d,k} = \Sigma_{d,k,1} \cup \Sigma_{d,k,2}$, where $\Sigma_{d,k,1}$ and $\Sigma_{d,k,2}$ are defined by

$$\Sigma_{d,k,1} = \{a_{\tau,j,i} : \tau \in \mathcal{Q}_k, j \in [|\tau|], i \in [N]\}, \text{ and}$$
$$\Sigma_{d,k,2} = \{a_{\alpha,\beta,i} : \alpha \in \{I, D\}, \beta \subseteq \{L, R\}, i \in [N]\}.$$

We say that the *type* of the letter $b \in \Sigma_{d,k,2}$ is the value of $\tau \in \mathcal{Q}_k$, and the *type* of the letter $b \in \Sigma_{d,k,1}$ is the value of the parameter $\alpha \in \{I, D\}$. Given a word $w \in \Sigma_{d,k}^*$, we say that w has *consistent type* if each of its letters share the same type, which we will denote by $t(w) \in \{I, D\} \cup \mathcal{Q}_k$, and also refer as the type of w. We set $f_{d,k}(w) = \star$ for every w that does not have a consistent type, and we remark that checking whether a word has consistent type or not can be checked by using a first-order sentence.

We start by defining the function $f_{d,k}$ on consistent words w whose type $t(w)$ is in $\mathcal{Q}_k$. Let $\tau = t(w)$, and for each $j \in [|\tau|]$ we define u_j to be the possibly empty subword of w consisting of all letters that are of the form $a_{\tau,j,i}$ for some $i \in [N]$. Let $\rho_1 : \Sigma_{d,k,1} \to \Sigma_{d-1,k}$ denote the projection-map obtained by setting $\rho_1(a_{\tau,i,j}) = a_j$, and we extend ρ_2 naturally to words of consistent type in $\Sigma_{d,k,1}^*$. Note that each of the letters of w belongs to precisely one subword u_j, and for each j we have $\rho_1(u_j) \in \Sigma_{d-1,k}$. For each j, we set $\overline{\sigma}_j = f_{d-1,k}(\rho_1(u_j)) \in \mathcal{S}_{d-1,k} \cup \{\star\}$, with the convention that the empty word is mapped to $\star$.

Given these preparations, we say that w *decomposes properly* if w coincides with the concatenation $u_1 u_2 \dots u_{|\tau|}$ of the subwords u_j defined as above. If w does not decompose properly, or there exists i for which we have $\overline{\sigma}_i = \star$, we set $f_{d,k}(w) = \star$. Otherwise, we set $f_{d,k}(w) = \left(\tau, \overline{\sigma}_1, \dots, \overline{\sigma}_{|\tau|}\right)$.

We now move on to defining the function $f_{d,k}$ on consistent words w whose type $t(w)$ is either I or D. In this case, we say that w decomposes properly if w can be written as a concatenation $w = u_1 \dots u_m$ of subwords $u_1, \dots, u_m$ where m is an arbitrary integer, and so that for each $i \in [m]$, the subword $u_i = u_{i,1} \dots u_{i,k_i}$ satisfies the following conditions:

1. If $k_i = 1$, the letter $u_{i,1}$ is of the form $a_{t(w),\{L,R\},j}$ for some $j \in [N]$.
2. If $k_i > 1$, the letter $u_{i,1}$ is of the form $a_{t(w),\{L\},j_1}$, the letter u_{i,k_i} is of the form $a_{t(w),\{R\},j_{k_i}}$, and every other letter $u_{i,s}$ is of the form $a_{t(w),\emptyset,j_s}$for some $j_1, j_s, j_{k_i} \in [N]$.

Define the projection map $\rho_2 : \Sigma_{d,k,2} \to \Sigma_{d-1,k}$ analogously by setting $\rho_2(a_{\alpha,\beta,j}) = a_j$, and again we extend ρ_2 to words in $\Sigma_{d,k,2}^*$, and we set $\overline{\sigma}_i = f_{d-1,k}(\rho_2(u_i))$

As before, if w does not decompose properly, or there exists i for which we have $\overline{\sigma}_i = \star$, we set $f_{d,k}(w) = \star$. Otherwise, we set $f_{d,k}(w) = (\tau_m, \overline{\sigma}_1, \dots, \overline{\sigma}_m)$, where τ_m is the monotone permutation of length m whose direction is specified by the parameter $t(w) \in \{I, D\}$.

One can check that the function $f_{d,k}$ defined above indeed satisfies the property that the set $f_{d,k}^{-1}(\star)$ can be described as the truth-set of a first-order sentence $\Psi_{d,k}$. Furthermore, one may also prove by induction on d that the function $f_{d,k}$ restricted to the truth-set of $\Psi_{d,k}$ is in fact a bijection whose image is $\mathcal{S}_{d,k}$.

By using the natural projection map $\pi_{d,k} : \mathcal{S}_{d,k} \to \tilde{\mathcal{O}}_k^{[d]}$, one concludes that there exists a regular language $L_{d,k}$ and a surjection $g_{d,k} := \pi_{d,k} \circ f_{d,k} : L_{d,k} \to \tilde{\mathcal{O}}_k^{[d]}$, but the problem is that the function $g_{d,k}$ is not an injection. Indeed, the function $g_{d,k}$ is not an injection as there are permutations $\sigma \in \tilde{\mathcal{O}}_k^{[d]}$ for which there exists multiple ways to express σ as $\sigma = \tau\left[\sigma_1, \dots, \sigma_{|\tau|}\right]$ with $\tau \in \tilde{\mathcal{O}}_k$ and $\sigma_i \in \tilde{\mathcal{O}}_k^{[d-1]}$.

To overcome this difficulty, we use the notion of left-greedy representations of permutations $\sigma \in \tilde{\mathcal{O}}_k^{[d]}$, which is a unique way of choosing a representative of the form $\sigma = \tau\left[\sigma_1, \dots, \sigma_{|\tau|}\right]$ for a permutation $\sigma \in \tilde{\mathcal{O}}_k^{[d]}$, where $\tau \in \tilde{\mathcal{O}}_k$ and $\sigma_i \in \tilde{\mathcal{O}}_k^{[d-1]}$ (see e.g. Proposition 2.4 in [5]). It turns out that these conditions can be written by using first-order sentences; we omit the details.

Lemma 3. *There exists a first-order sentence* Φ *for which the restriction of the function* $g_{d,k} = \pi_{d,k} \circ f_{d,k}$ *on* $\{w \in \Sigma_{d,k} : w \models \Phi\}$ *induces a bijection whose image is* $\tilde{\mathcal{O}}_k^{[d]}$*. In particular, this restriction induces a bijection between a regular language and* $\tilde{\mathcal{O}}_k^{[d]}$*.*

Finally, we also highlight an important property of the construction. Indeed, it is easy to see that for each letter w_i, one may backtrack the iterative construction to find a unique representative point p_j in the permutation $\sigma = g_{d,k}(w)$ that naturally corresponds to w_i. Furthermore, it is not too hard to check that w_i belongs to one of the subwords u_s defined above if and only if p_j belongs to the permutation σ_s in the left-greedy representation of σ.

3 Correspondence of first-order sentences

Note that there are three types of atomic formulas in the logical model TOTO for permutations, namely $x = y$, $x \leq_1 y$ and $x \leq_2 y$. Given a first-order sentence Ψ, we may write it in the disjunctive normal form, i.e. as

$$\Psi = Q_1 x_1 \dots Q_c x_c \left(p_{11} \vee \dots \vee p_{1a_1}\right) \wedge \dots \wedge \left(p_{b1} \vee \dots \vee p_{ba_b}\right),$$

where $Q_i \in \{\forall, \exists\}$ are quantifiers, $x_1, \dots, x_c$ are variables, and each p_{ij} is an atomic formula or a negation of an atomic formula.

In order to prove that $g_{d,k}$ is a FOS-homomorphism, it turns out to be enough to find appropriate sentences that correspond to the three different atomic formulas and their negations. Indeed, if q_{ij} is such a sentence corresponding to p_{ij}, then it is not too hard to check that the sentence $\Phi = Q_1 x_1 \dots Q_c x_c \left(q_{11} \vee \dots \vee q_{1a_1}\right) \wedge \dots \wedge \left(q_{b1} \vee \dots \vee q_{ba_b}\right)$ in $\Sigma_{d,k}^*$ corresponds to Ψ.

We now sketch the idea behind constructing the appropriate sentence in $\Sigma_{d,k}^*$ which corresponds to the atomic sentences $x \leq_a y$, where $a \in \{1, 2\}$. Given points p and q in $\sigma = g_{d,k}(w)$, let w_p and w_q denote the letters corresponding to these two points, as descried at the end of the previous chapter. Let $\sigma = \tau\left[\sigma_1, \dots, \sigma_{|\tau|}\right]$ be the left-greedy representation of σ, and for simplicity let us assume that $\tau \in \mathcal{Q}_k$.

If $p \in \sigma_i$ and $q \in \sigma_j$ for $i \neq j$, the statement $p \leq_a q$ is equivalent to $i < j$ if $a = 1$, and to $\tau(i) < \tau(j)$ if $a = 2$. Since the statements $p \in \sigma_i$ and $q \in \sigma_j$ are equivalent to having $w_p \in u_i$ and $w_q \in u_j$, this condition can easily be described by using a first-order sentence.

Now suppose that $i = j$, whence both w_p and w_q belong to the same subword u_i, and both p and q belong to σ_i. This allows us to reduce the problem to asking the same question for points p and q in the sub-interval $\sigma_i \in \tilde{\mathcal{O}}_k^{[d-1]}$, and the corresponding letters w_p and w_q in the subword u_i. Since $\sigma_i = g_{d-1,k}(\rho_2(u_i))$, we may proceed by induction. Also, checking the base case $d = 1$ is not too hard to handle.

Checking the details when τ is a monotone permutation follows similar ideas, although the execution is slightly more complicated, as the length m of the monotone permutation may not be bounded.

References

[1] M. Albert, M. Atkinson, M. Bouvel, N. Ruškuc, and V. Vatter. Geometric grid classes of permutations. *Trans. Amer. Math. Soc.*, 365(11):5859–5881, 2013.

[2] M. Albert, M. Atkinson, and R. Brignall. Permutation classes of polynomial growth. *Ann. Comb.*, 11:249–264, 2007.

[3] M. Albert, M. Bouvel, and V. Féray. Two first-order logics of permutations. *J. Combin. Theory Ser. A*, 171: Paper 105158, 2020.

[4] M. Albert, M. Bouvel, V. Féray, and M. Noy. A logical limit-law for 231-avoiding permutations. *Discrete Math. Theor. Comput. Sci.*, 26(11): Paper 11751, 2024.

[5] M. Albert, N. Ruškuc, and V. Vatter. Inflations of geometric grid classes of permutations. *Israel J. Math.*, 205:73–108, 2015.

[6] S. Braunfeld and M. Kukla. Logical limit laws for layered permutations and related structures. *Enumer. Comb. Appl.*, 2(4): Article #S4PP2, 2022.

[7] S. Huczynska and V. Vatter. Grid classes and the fibonacci dichotomy for restricted permutations. *Electron. J. Combin*, 13: Article R54, 2006.

[8] T. Klaiser and M. Klazar. On growth rates of closed permutation classes. *Electron. J. Combin*, 9(2): Article R10, 2003.

[9] A. Marcus and G. Tardos. Excluded permutation matrices and the stanley-wilf conjecture. *J. Combin. Theory Ser. A*, 107(1):153–160, 2004.

[10] K. Markström and E. Räty. A first-order isomorphism between small permutation classes and regular languages, and applications to logical limit laws. *Manuscript under preparation*, 2025.

[11] A Özdemir. First-order convergence for 321-avoiding permutations. *Arxiv preprint, arXiv:2312.0174*, 2023.

[12] V. Vatter. Small permutation classes. *Proc. Lond. Math. Soc*, 103(5):879–921, 2011.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Local Shearer Bound

(Extended abstract)

Anders Martinsson* Raphael Steiner†

Abstract

We prove the following local strengthening of Shearer's classic bound on the independence number of triangle-free graphs: For every triangle-free graph G there exists a probability distribution on its independent sets such that every vertex v of G is contained in a random independent set drawn from the distribution with probability $(1-o(1))\frac{\ln d(v)}{d(v)}$. This resolves the main conjecture raised by Kelly and Postle (2018) about fractional coloring with local demands, which in turn confirms a conjecture by Cames van Batenburg et al. (2018) stating that every n-vertex triangle-free graph has fractional chromatic number at most $(\sqrt{2}+o(1))\sqrt{\frac{n}{\ln(n)}}$. Addressing another conjecture posed by Cames van Batenburg et al., we also establish an analogous upper bound in terms of the number of edges.

To obtain these results we prove a more general technical theorem that works in a weighted setting and may be of independent interest. As a further application of this more general result, we derive a new spectral upper bound on the fractional chromatic number of triangle-free graphs: We show that every triangle-free graph G satisfies $\chi_f(G) \leq (1+o(1))\frac{\rho(G)}{\ln \rho(G)}$ where $\rho(G)$ denotes the spectral radius, improving the bound implied by Wilf's classic spectral estimate for the chromatic number by a $\ln \rho(G)$ factor.

*Department of Computer Science, ETH Zürich, Zurich, Switzerland. E-mail: aomartinsson@gmail.com.

†Department of Mathematics, ETH Zürich, Zurich, Switzerland. E-mail: raphaelmario.steiner@math.ethz.ch. Supported by the SNSF Ambizione Grant No. 216071.

1 Introduction

One of the most fundamental problems in all of combinatorics concerns bounding the famous Ramsey number $R(\ell, k)$, which may be defined as the smallest number n such that every graph on n vertices contains either a clique of size ℓ or an independent set of size k. The first highly challenging instance of this general problem is the determination of the Ramsey-numbers $R(3, k)$, posed as a price-money question by Erdős already back in 1961 [25] (see also this Erdős problem page entry). Currently, the best known asymptotic bounds are

$$\left(\frac{1}{4} - o(1)\right) \frac{k^2}{\ln k} \leq R(3, k) \leq (1 + o(1))\frac{k^2}{\ln k}.$$

The lower bound was established by an analysis of the famous *triangle-free process* independently by Fiz Pontiveros, Griffiths and Morris [27] and Bohman and Keevash [9] in 2013. Proving an upper bound on $R(3, k)$ is equivalent to establishing a lower bound on the *independence number* $\alpha(G)$ (i.e., the size of a largest independent set in G) for all triangle-free graphs on n vertices. In 1980, Ajtai, Komlós and Szemerédi [1] famously proved that every triangle-free graph G on n vertices with average degree $\overline{d}$ satisfies $\alpha(G) \geq c\frac{\ln \overline{d}}{\overline{d}}$, where $c > 0$ is some small absolute constant, which is easily seen to imply that $R(3, k) \leq O\left(\frac{k^2}{\log k}\right)$. The stronger upper bound on $R(3, k)$ stated above is due to a strengthening of the result of Ajtai, Komlós and Szemerédi, established in a landmark result by Shearer in 1983. Namely, Shearer [44] improved the constant factor c in this bound significantly by showing that $\alpha(G) \geq \frac{(1-\overline{d})+\overline{d}\ln \overline{d}}{(\overline{d}-1)^2}n = (1 - o(1))\frac{\ln \overline{d}}{\overline{d}}n$. Further refining this result, Shearer [45] proved in 1991 that every triangle-free graph G satisfies $\alpha(G) \geq \sum_{v \in V(G)} g(d_G(v))$, where $d_G(v)$ denotes the degree of v in G and $g(d) = (1 - o(1))\frac{\ln d}{d}$ is a recursively defined function. This bound is slightly better than Shearer's first bound in terms of the average degree for graphs with unbalanced degree sequences.

These two classic bounds on the independence number of triangle-free graphs due to Shearer have essentially remained the state of the art on the topic for four decades, and due to their ubiquity have found widespread application as a tool across many areas of extremal and probabilistic combinatorics. By a result of Bollobás [10], it is known that Shearer's bounds are tight up to a multiplicative factor of 2. Because of the relation to the Ramsey-numbers $R(3, k)$ discussed above, any constant factor improvement of Shearer's longstanding bounds would be a major breakthrough in Ramsey theory. Due to this, a lot of research has been devoted to finding strengthenings and generalizations of Shearer's bounds: We refer to [22, 34] for recent surveys covering Shearer's bound and relations to the hard-core model in statistical mechanics as well as the theory of graph coloring and to [3, 17, 18, 20, 21, 23, 42] for some extensions and generalizations of Shearer's bounds.

The study of lower bounds on the independence number is closely connected to the theory of graph coloring. Recall that in a *proper graph coloring* vertices are assigned colors such that neighboring vertices have distinct colors, and the *chromatic number* $\chi(G)$ of a graph G is the smallest amount of colors required to properly color G. It is easily seen by the definition (by considering a largest "color class") that every graph G on n vertices has an independent set of size at least $\frac{n}{\chi(G)}$. An even stronger lower bound on the independence number is provided by the well-known *fractional chromatic number* $\chi_f(G)$ of the graph. The fractional chromatic number has many different equivalent definitions (see the standard textbook [43] on fractional coloring as a reference). Here, we shall find the following definition convenient: $\chi_f(G)$ is the minimum real number $r \geq 1$ for which there exists a probability distribution on the independent sets of G such that a random independent set I sampled from this distribution contains any given vertex $v \in V(G)$ with probability at least $\frac{1}{r}$. By considering the expected size of a random set drawn from such a distribution, one immediately verifies that $\alpha(G) \geq \frac{n}{\chi_f(G)}$ holds for every graph G. In general, the latter lower bound $\frac{n}{\chi_f(G)}$ on the independence number is stronger than the lower bound $\frac{n}{\chi(G)}$, as there are graphs (such as the Kneser graphs [6, 37]) for which $\chi_f(G)$ is much smaller than $\chi(G)$.

Given these lower bounds of the independence number in terms of the (fractional) chromatic number, it is natural to ask whether there are analogues or strengthenings of Shearer's bounds that provide corresponding *upper* bounds for the (fractional) chromatic number. A prime example of such a result is a recent breakthrough of Molloy [40], who proved that $\chi(G) \leq (1+o(1))\frac{\Delta}{\ln \Delta}$ for every triangle-free graph G with maximum degree Δ, where the $o(1)$-term vanishes as $\Delta \to \infty$. This strengthened a longstanding previous bound of the form $O\left(\frac{\Delta}{\ln \Delta}\right)$ due to Johansson [33] and recovers Shearer's independence number bound in the case of regular graphs in a stronger form. As with Shearer's bound, it is known that Molloy's bound is optimal up to a factor of 2, and improving the constant 1 to any constant below 1 would be a major advance in the field. Several interesting strengthenings and generalizations of Johansson's and Molloy's results have been proved in the literature, see e.g. [2, 4, 5, 7, 11, 12, 13, 17, 31, 32, 42] for some selected examples.

Our results. We shall be concerned with the following conjecture from 2018 posed by Kelly and Postle [35] that claims a local strengthening of Shearer's bounds that can also be seen as a degree-sequence generalization of Molloy's bound for fractional coloring.

Conjecture 1.1 (Local fractional Shearer/Molloy, cf. Conjecture 2.2 in [35]). *For every triangle-free graph there exists a probability distribution on its independent sets such that every vertex $v \in V(G)$ appears with probability at least $(1-o(1))\frac{\ln d_G(v)}{d_G(v)}$ in a random independent set sampled from the distribution. Here, the $o(1)$ term represents any function that tends to 0 as the degree grows.*

To see that this conjecture indeed strengthens Shearer's bounds, note that the expected size of a random independent set drawn from a distribution as given by the conjecture is

$$\sum_{v\in V(G)}(1-o(1))\frac{\ln d_G(v)}{d_G(v)},$$

which recovers Shearer's second (stronger) lower bound [45] on the independence number up to lower-order terms. But on top of that, and this explains the word "local" in the name of the conjecture, the distribution in Conjecture 1.1 guarantees that *every* vertex can be expected to be contained in the random independent set a good fraction of the time (and lower degree vertices are contained proportionally more frequently). This relates back to the previously discussed fractional chromatic number, and, for instance, directly implies that $\chi_f(G) \leq (1+o(1))\frac{\Delta(G)}{\ln \Delta(G)}$ for every triangle-free graph, which recovers the fractional version of Molloy's bound.

Adding to that, Conjecture 1.1 connects to several other notions of graph coloring discussed in detail by Kelly and Postle, see in particular [35, Proposition 1.4] which provides many different equivalent formulations of Conjecture 1.1. One of these involves the notion of *fractional coloring with local demands* introduced by Dvořák, Sereni and Volec [24]. Following Kelly and Postle [35], given a graph G and a so-called *demand function* $h : V(G) \to [0,1]$ that assigns to each vertex its individual "demand", an *h-coloring* of a graph G is a mapping $c : V(G) \to 2^{[0,1]}$ that assigns to every vertex $v \in V(G)$ a measurable subset $c(v) \subseteq [0,1]$ of measure at least $h(v)$, in such a way that adjacent vertices in G are assigned disjoint subsets. Since the function h does not have to be constant but can depend on local information concerning the vertex v in G, this setting extends the usual paradigm of graph coloring in a local manner. Kelly and Postle [35, Proposition 1.4] proved that Conjecture 1.1 is equivalent to the statement that every triangle-free graph has an h-coloring, where $h : V(G) \to [0,1]$ is a function depending only on the vertex-degrees such that $h(v) = (1-o(1))\frac{\ln d_G(v)}{d_G(v)}$. We refer to the extensive introduction of [35] for further applications of the conjecture.

In one of their main results, Kelly and Postle [35, Theorem 2.3] proved a relaxation of Conjecture 1.1, replacing the bound $(1-o(1))\frac{\ln d_G(v)}{d_G(v)}$ with the asymptotically weaker $\frac{(1-o(1))\ln d_G(v)}{2e\cdot d_G(v)\ln\ln d_G(v)}$. As our first main result, we fully resolve Conjecture 1.1.

Theorem 1.2. *For every triangle-free graph G there exists a probability distribution $\mathcal{D}$ on the independent sets of G such that*

$$\mathbb{P}_{I\sim D}[v \in I] \geq (1-o(1))\frac{\ln(d_G(v))}{d_G(v)}$$

for every $v \in V(G)$. Here the $o(1)$ represents a function of $d_G(v)$ that tends to 0 as the degree grows.

A pleasing consequence of Theorem 1.2 is that it can also be used to fully resolve another conjecture about fractional coloring raised in 2018 by Cames van Batenburg, de Joannis de Verclos, Kang and Pirot [14]: Motivated by the aforementioned problem of estimating the Ramsey-number $R(3,k)$, in 1967 Erdős asked the fundamental question of determining the maximum chromatic number of triangle-free graphs on n vertices. An observation of Erdős and Hajnal [26] combined with Shearer's bound implies an upper bound $(2\sqrt{2}+o(1)))\sqrt{\frac{n}{\ln n}}$ for this problem. In recent work of Davies and Illingworth [19], this upper bound was improved by a $\sqrt{2}$-factor to the current state of the art $(2+o(1))\sqrt{\frac{n}{\ln n}}$. The current best lower bound for this quantity is $(1/\sqrt{2}-o(1))\sqrt{\frac{n}{\ln n}}$, coming from the aforementioned lower bounds on $R(3,k)$ [9, 27].

Cames van Batenburg et al. [14] studied the natural analogue of this question for fractional coloring and made the following conjecture.

Conjecture 1.3 (cf. Conjecture 4.3 in [14])**.** *As $n \to \infty$, every triangle-free graph on n vertices has fractional chromatic number at most $(\sqrt{2}+o(1))\sqrt{\frac{n}{\ln n}}$.*

In one of their main results [14, Theorem 1.4], Cames van Batenburg et al. proved the fractional version of the result of Davies and Illingworth, namely an upper bound $(2+o(1))\sqrt{\frac{n}{\ln n}}$ on the fractional chromatic number. Using a connection between Conjectures 1.1 and 1.3 proved by Kelly and Postle [35, Proposition 5.2], we are able to confirm Conjecture 1.3 too.

Theorem 1.4. *The maximum fractional chromatic number among all n-vertex triangle-free graphs is at most*

$$(\sqrt{2}+o(1))\sqrt{\frac{n}{\ln(n)}}.$$

We also prove a similar upper bound on the fractional chromatic number of triangle-free graphs in terms of the number of edges, as follows.

Theorem 1.5. *The maximum fractional chromatic number among triangle-free graphs with m edges is at most*

$$(18^{1/3}+o(1))\frac{m^{1/3}}{(\ln m)^{2/3}}.$$

Theorem 1.5 comes very close to confirming another conjecture of Cames van Batenburg et al. [14, Conjecture 4.4], stating that every triangle-free graph with m edges has fractional chromatic number at most $(16^{1/3}+o(1))m^{1/3}/(\ln m)^{2/3}$. In fact, after a personal communication with the authors of [14] it turned out that the constant $16^{1/3}$ seems to be due to a miscalculation on their end. In particular, it was claimed in [14] that the conjectured bound on the fractional chromatic number can be verified in the special case of d–regular triangle–free graphs using the upper bound $\chi_f(G) \le \min\left((1+o(1))d/\ln d, n/d\right)$. However, assuming $n=(1+o(1))d^2/\ln d$ and thus $m=(1/2+o(1))d^3/\ln d$, this upper bound simplifies to $(1+o(1))d/\ln d = (1+o(1))(2m)^{1/3}/(\ln m^{1/3})^{2/3} = (1+o(1))(18m)^{1/3}/(\ln m)^{2/3}$, matching our bound in Theorem 1.5.

To prove our main result, Theorem 1.2, we establish a key technical result that goes beyond Theorem 1.2 and generalizes it to a vertex-weighted setting (see in Section 2). As a consequence of this more general result, we obtain a new spectral upper bound on the fractional chromatic number of triangle-free graphs, stated below. In the following, as is standard, $\rho(G)$ denotes the *spectral radius* of G, i.e., the spectral radius of its adjacency matrix.

Theorem 1.6. *Every triangle-free graph G satisfies*

$$\chi_f(G) \le (1+o(1))\frac{\rho(G)}{\ln \rho(G)},$$

where the $o(1)$ term represents a function of $\rho(G)$ that tends to 0 as $\rho(G)$ grows.

Theorem 1.6 lines up nicely with a rich area of research that is concerned with spectral bounds on the (fractional) chromatic number, see e.g. Chapter 6 of the textbook on spectral graph theory [15] by Chung and [8, 16, 28, 30, 36, 39, 41] for some small selection of articles on the topic. Theorem 1.6 relates to Wilf's classic spectral bound [46] on the chromatic number, which states that every connected graph G satisfies $\chi(G) \le \rho(G)+1$ with equality if and only if G is an odd cycle or a complete graph. Theorem 1.6 shows that at least for the fractional chromatic number, a $\ln \rho(G)$-factor can be saved compared to Wilf's bound when the graph is assumed to be triangle-free. In fact, we think that our upper bound stated in Theorem 1.6 should extend from fractional to ordinary coloring, as follows.

Conjecture 1.7. *Every triangle-free graph G satisfies $\chi(G) \le (1+o(1))\frac{\rho(G)}{\ln \rho(G)}$.*

Theorem 1.6 also relates to a conjecture of Harris [29], stating that every triangle-free d-degenerate graph G satisfies $\chi_f(G) \le O(\frac{d}{\ln d})$. The first author [38] recently confirmed Harris' conjecture, thus recovering Theorem 1.6 in a more general form (since every graph G is $\rho(G)$-degenerate), albeit with a worse constant factor.

2 Key technical result

In the following, given a weight function $w : V(G) \to \mathbb{R}_+$ on the vertices of a graph G and a subset $X \subseteq V(G)$, $w(X) := \sum_{v \in X} w(v)$ denotes the total weight of X.

The key technical contribution of our work is a weighted generalization of Theorem 1.2, in which the distribution can be skewed towards vertices with a high weight relative to their neighbors. From this result, which seems of independent interest, all our main results (Theorems 1.2, 1.4, 1.5, 1.6) can be deduced without much work.

Theorem 2.1. *For every triangle-free graph G and every strictly positive weight function $w : V(G) \to \mathbb{R}_+$ on the vertices there exists a probability distribution $\mathcal{D}$ on the independent sets of G such that*

$$\mathbb{P}_{I \sim \mathcal{D}}[v \in I] \ge f\left(\frac{w(N_G(v))}{w(v)}\right)$$

for every vertex $v \in V(G)$, where $f(x) = \frac{(1-x)+x\ln(x)}{(x-1)^2} = (1-o(1))\frac{\ln(x)}{x}$.

As an example, we present the short deduction of Theorem 1.6 from Theorem 2.1.

Proof of Theorem 1.6. Let G be any given triangle-free graph. W.l.o.g., we may assume that G is connected. So let G be such a graph, and let $A \in \mathbb{R}^{V(G)\times V(G)}$ be its adjacency matrix. By definition, A has non-negative entries, and hence we may apply the Perron-Frobenius theorem to find that $\rho(A) = \rho(G)$ is an eigenvalue of A and that there exists a corresponding eigenvector $\mathbf{u} \in \mathbb{R}^{V(G)}$ with non-negative entries. So we have $A\mathbf{u} = \rho(G)\mathbf{u}$, which reformulated means that

$$\sum_{x \in N_G(v)} \mathbf{u}_x = \rho(G)\mathbf{u}_v$$

for every $v \in V(G)$. This equality in particular implies that if at least one neighbor of a vertex v has a positive entry in $\mathbf{u}$, then so does v. Hence, since G is a connected graph, it follows that $\mathbf{u}_v > 0$ for every $v \in V(G)$. Now interpret the entries of the vector $\mathbf{u}$ as a strictly positive weight assignment to the vertices of G. Then, by Theorem 2.1, there exists a probability distribution $\mathcal{D}$ on the independent sets of G such that for every $v \in V(G)$, we have

$$\mathbb{P}_{I \sim D}[v \in I] \geq f\left(\frac{\sum_{x \in N_G(v)} \mathbf{u}_x}{\mathbf{u}_v}\right) = f(\rho(G)).$$

By definition of the fractional chromatic number, this implies that $\chi_f(G) \leq \frac{1}{f(\rho(G))}$, as desired. This concludes the proof. $\square$

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *Journal of Combinatorial Theory, Series A*, 29(3):354–360, 1980.

[2] N. Alon, M. Krivelevich, and B. Sudakov. Coloring graphs with sparse neighborhoods. *Journal of Combinatorial Theory, Series B*, 77(1):73–82, 1999.

[3] Noga Alon. Independence numbers of locally sparse graphs and a Ramsey type problem. *Random Structures & Algorithms*, 9(3):271–278, 1996.

[4] J. Anderson, A. Bernshteyn, and A. Dhawan. Colouring graphs with forbidden bipartite subgraphs. *Combinatorics, Probability & Computing*, 32(1):45–67, 2023.

[5] James Anderson, Abhishek Dhawan, and Aiya Kuchukova. Coloring locally sparse graphs, 2024.

[6] J Bárány. A short proof of Kneser's conjecture. *Journal of Combinatorial Theory, Series A*, 25(3):325–326, 1978.

[7] A. Bernshteyn, T. Brazelton, R. Cao, and A. Kang. Counting colorings of triangle-free graphs. *Journal of Combinatorial Theory, Series B*, 161:86–108, 2023.

[8] Y. Bilu. Tales of Hoffman: Three extensions of Hoffman's bound on the graph chromatic number. *Journal of Combinatorial Theory, Series B*, 96(4):608–613, 2006.

[9] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Random Structures & Algorithms*, 58(2):221–293, 2021.

[10] Béla Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, 83:433—436, 1981.

[11] M. Bonamy, T. Kelly, P. Nelson, and L. Postle. Bounding χ by a fraction of δ for graphs without large cliques. *Journal of Combinatorial Theory, Series B*, 157:263–282, 2022.

[12] P. Bradshaw, B. Mohar, and L. Stacho. Bipartite graphs are $(\frac{4}{5}-\varepsilon)\frac{\Delta}{\log \Delta}$-choosable, 2024.

[13] S. Cambie, W. Cames van Batenburg, E. Davies, and R. J. Kang. Packing list-colorings. *Random Structures & Algorithms*, 64(1):62–93, 2024.

[14] W. Cames van Batenburg, R. de Joannis de Verclos, R. J. Kang, and F. Pirot. Bipartite induced density in triangle-free graphs. *The Electronic Journal of Combinatorics*, 27(2):P2–34, 2020.

[15] F. R. K. Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[16] D. Cvetkovic. Chromatic number and the spectrum of a graph. *Publ. Inst. Math.(Beograd)*, 14(28):25–38, 1972.

[17] E. Davies, R. de Joannis de Verclos, R. J. Kang, and F. Pirot. Coloring triangle-free graphs with local list sizes. *Random Structures & Algorithms*, 57(3):730–744, 2020.

[18] E. Davies, R. de Joannis de Verclos, R. J. Kang, and F. Pirot. Occupancy fraction, fractional colouring, and triangle fraction. *Journal of Graph Theory*, 97(4):557–568, 2021.

[19] E. Davies and F. Illingworth. The χ-Ramsey problem for triangle-free graphs. *SIAM Journal on Discrete Mathematics*, 36(2):1124–1134, 2022.

[20] E. Davies, R. J. Kang, F. Pirot, and J. Sereni. Graph structure via local occupancy, 2020.

[21] Ewan Davies, Matthew Jenssen, Will Perkins, and Barnaby Roberts. On the average size of independent sets in triangle-free graphs. *Proceedings of the American Mathematical Society*, 146(1):111–124, 2018.

[22] Ewan Davies and Ross J. Kang. The hard-core model in graph theory, 2025.

[23] Abhishek Dhawan. Bounds for the independence and chromatic numbers of locally sparse graphs, 2024.

[24] Z. Dvořák, J. Sereni, and J. Volec. Subcubic triangle-free graphs have fractional chromatic number at most 14/5. *Journal of the London Mathematical Society*, 89(3):641–662, 2014.

[25] P. Erdős. Some unsolved problems. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, pages 221–254, 1961.

[26] P. Erdős and A. Hajnal. Chromatic number of finite and infinite graphs and hypergraphs. *Discrete Mathematics*, 53:281–285, 1985.

[27] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and the Ramsey number $R(3, k)$. *Memoirs of the American Mathematical Society*, 263:125 pp, 2020.

[28] K. Guo and S. Spiro. New eigenvalue bound for the fractional chromatic number. *Journal of Graph Theory*, 106(1):167–181, 2024.

[29] David G. Harris. Some results on chromatic number as a function of triangle count. *SIAM Journal on Discrete Mathematics*, 33(1):546–563, 2019.

[30] A. J. Hoffman. On eigenvalues and colorings of graphs. *Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., 1969)*, 1970.

[31] E. Hurley, R. de Joannis de Verclos, and R. J. Kang. An improved procedure for colouring graphs of bounded local density. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 135–148. SIAM, 2021.

[32] E. Hurley and F. Pirot. Uniformly random colourings of sparse graphs. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1357–1370, New York, NY, USA, 2023. Association for Computing Machinery.

[33] Anders Johansson. Asymptotic choice number for triangle free graphs. Technical report, Technical report 91-5, DIMACS, 1996.

[34] D. Kang, D. Kühn, A. Methuku, and D. Osthus. Graph and hypergraph colouring via nibble methods: A survey. *Proceedings of the 8th European Congress of Mathematics*, pages 771–823, 2023.

[35] T. Kelly and L. Postle. Fractional coloring with local demands and applications to degree-sequence bounds on the independence number. *Journal of Combinatorial Theory, Series B*, 169:298–337, 2024.

[36] M. Kwan and Y. Wigderson. The inertia bound is far from tight. *Bulletin of the London Mathematical Society*, 56(10):3196–3208, 2024.

[37] Laszlo Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[38] Anders Martinsson. Triangle-free d-degenerate graphs have small fractional chromatic number, 2025.

[39] B. Mohar. Eigenvalues and colorings of digraphs. *Linear Algebra and its Applications*, 432(9):2273–2277, 2010.

[40] M. Molloy. The list chromatic number of graphs with small clique number. *Journal of Combinatorial Theory, Series B*, 134:264–284, 2019.

[41] V. Nikiforov. Chromatic number and spectral radius. *Linear algebra and its applications*, 426(2-3):810–814, 2007.

[42] F. Pirot and J. Sereni. Fractional chromatic number, maximum degree, and girth. *SIAM Journal on Discrete Mathematics*, 35(4):2815–2843, 2021.

[43] E. R. Scheinerman and D. H. Ullman. *Fractional graph theory: a rational approach to the theory of graphs*. Dover publications Inc., 2011.

[44] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Mathematics*, 46(1):83–87, 1983.

[45] J. B. Shearer. A note on the independence number of triangle-free graphs, II. *Journal of Combinatorial Theory, Series B*, 53(2):300–307, 1991.

[46] H. S. Wilf. The eigenvalues of a graph and its chromatic number. *Journal of the London Mathematical Society*, s1-42(1):330–332, 1967.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Local Degree Sum Conditions for the Existence of Paths and k-trees in Graphs

(Extended abstract)

Haruhide Matsuda*

Abstract

For an integer $k \geq 2$, a k-tree is defined as a tree with maximum degree at most k. If a k-tree T spans a graph G, then T is called a spanning k-tree of G. Since a 2-tree is a path, a k-tree is an extended concept of a path. In particular, a spanning 2-tree is a Hamiltonian path. Therefore, a spanning k-tree is an extended concept of a Hamiltonian path.

This research gives local degree sum conditions for the existence of a path and a k-tree containing all the vertices of the set of vertices S in a graph.

1 Introduction

We consider only finite undirected graphs without loops or multiple edges.

A *Hamiltonian cycle* of a graph is a cycle containing all the vertices of the graph and a *Hamiltonian path* of a graph is a path containing all the vertices of the graph. For an integer $k \geq 2$, a *k-tree* of a graph is defined as a tree with maximum degree at most k. If a k-tree T spans a graph G, then T is called *a spanning k-tree* of G. Since a 2-tree is a path, a k-tree is an extended concept of a path. In particular, a spanning 2-tree is a Hamiltonian path. Therefore, a spanning k-tree is an extended concept of a Hamiltonian path.

Before we give some results, we need some notations. The order of a graph G is denoted by $|G|$. For a vertex of a graph G, we denote the degree of x in G by $\deg_G(x)$, the neighborhood of x in G by $N_G(x)$, and $N_G(x) \cup \{x\}$ by $N_G[x]$.

For the set of vertices S in a graph G, write $\bigcup_{x\in S} N_G(x)$ as $N_G(S)$ and $(\bigcup_{x\in S} N_G(x)) \cup S$ as $N_G[S]$. Let $\sigma_k(S; G)$ denote the smallest sum of degrees of k independent vertices in the set of vertices S of a graph G, i.e.,

$$\sigma_k(S;G) := \min\left\{\sum_{i=1}^{k} \deg_G(x_i) \mid \{x_1, \dots, x_k\} \text{ is a set of } k \text{ independent vertices in } S\right\}.$$

If $S = V(G)$, we write $\sigma_k(G)$ instead of $\sigma_k(V(G); G)$.

*Department of Computer and Communications Engineering, Shibaura Institute of Technology, 307 Fukasaku, Minuma-ku, Saitama 337-8577, Japan. E-mail: hmatsuda@sic.shibaura-it.ac.jp. This work was supported by JSPS KAKENHI Grant Number 20K03724.

2 Preliminaries

There are now many sufficient conditions for a graph to have a cycle, a path, and a k-tree. Many of them involve some sort of degree condition. In particular, Ore's theorem is a famous result among them.

Theorem 1 (Ore [1])**.** *If a graph G of order $|G| \geq 3$ satisfies*

$$\sigma_2(G) \geq |G|,$$

then G has a hamiltonian cycle.

The following theorem can be derived from Theorem 1.

Theorem 2. *If a connected graph G of order $|G| \geq 2$ satisfies*

$$\sigma_2(G) \geq |G| - 1,$$

then G has a hamiltonian path.

Since a hamiltonian path is a spanning 2-tree, Win extended the above result to a spanning k-tree.

Theorem 3 (Win [4])**.** *Let $k \geq 2$ be an integer. If a connected graph G of order $|G| \geq k$ satisfies*

$$\sigma_k(G) \geq |G| - 1,$$

then G has a spanning k-tree.

The degree sum condition $\sigma_k(G)$ is a condition over an entire graph G. When it does not hold for the entire graph, but only for a certain set of vertices, Shi and Ota independently obtain some results similar to Theorem 1.

Theorem 4 (Shi [3], Ota [2])**.** *Let G be a 2-connected graph and let S be a set of vertices of G. If*

$$\sigma_2(S; G) \geq |G|,$$

then G has a cycle through all vertices of S.

Yamashita obtained the following result which is the refinement of Theorem 4.

Theorem 5 (Yamashita [5])**.** *Let G be a 2-connected graph of order $|G| \geq 3$. If a set of vertices S of G satisfies*

$$\sigma_2(S; G) \geq |N_G[S]|,$$

then G has a cycle which contains all the vertices of S.

It is easy to have the following theorem by using the argument of the proof of Theorem 5.

Theorem 6. *Let G be a connected graph of order $|G| \geq 2$. If a set of vertices S of G satisfies*

$$\sigma_2(S; G) \geq |N_G[S]| - 1,$$

then G has a path which contains all the vertices of S.

3 Main result

Since a path can be considered a 2-tree, we extended Theorem 6 to k-trees. The following main theorem is a natural extension of Theorems 3 and 6. In particular, the main theorem coincides with Win's theorem for $S = V(G)$.

Theorem 7. *Let $k \geq 2$ be an integer and let G be a connected graph of order $|G| \geq k$. If a set of vertices S of G satisfies*

$$\sigma_k(S;G) \geq |N_G[S]| - 1,$$

then G has a k-tree which contains all the vertices of S.

4 Proof of Theorem 7

By Theorem 6, we consider the case when $k \geq 3$. Suppose that G satisfies the condition on $\sigma_k(S;G)$ but has no k-tree which contains all the vertices of S. Let T be a maximal k-tree of G satisfying the following two conditions (i) and (ii);

(i) T contains as many vertices of S as possible and

(ii) $|V(T) \setminus S|$ is as small as possible under the condition (i).

Since T doesn't contain all the vertices of S, G has a vertex $v \in S$ not contained in T. As G is connected, G has a path P which connects v and T. Let w be the vertex of T on the path P that is nearest to v.

If $\deg_T(w) < k$, then $T \cup P$ is a k-tree which contains more vertices of S than T. This contadicts the choice of T. Hence $\deg_T(w) = k$. Let $D_1, D_2, \ldots, D_k$ be the components of $T \setminus \{w\}$, and let u_i be the vertex in D_i adjacent to w and let x_i be a leaf of T contained in D_i. Note that $u_j = x_j$ for all j's with $|D_i| = 1$.

Claim 1. *$x_i \in S$.*

Proof. To the contrary, suppose that $x_i \notin S$. Then $T - \{x_i\}$ is a k-tree of G such that $|V(T) \cap S| = |(V(T) \setminus \{x_i\}) \cap S|$ and $|V(T) \setminus S| > |(V(T) \setminus \{x_i\}) \setminus S|$. This contradicts the choice (ii) of T. Consequently, $x_i \in S$. □

Claim 2. *$\{x_1, x_2, \ldots, x_k\} \subseteq S$ is an independent set of G and $N_G(x_i) \subseteq V(T) \cup (V(G) \setminus S)$ for all $1 \leq i \leq k$.*

Proof. By the choice of T, we easily see that $N_G(x_i) \subseteq V(T) \cup (V(G) \setminus S)$ for all $1 \leq i \leq k$. Claim 1 implies $\{x_1, x_2, \ldots, x_k\} \subseteq S$. If x_i and x_j are adjacent in G, then $T + x_i x_j - w u_i + P$ is a k-tree of G, a contradiction. Hence $\{x_1, x_2, \ldots, x_k\} \subseteq S$ is an independent set of G. □

Note that if $\sigma_k(S;G) = \infty$, then S contains no k independent vertices and so the theorem can be shown immediately. Thus we may assume that $\sigma_k(S;G)$ is finite.

For any D_t with $1 \leq t \leq k$, choose a vertex x_a from $\{x_1, x_2, \ldots, x_k\} \setminus \{x_t\}$ so that

$$|N_G(x_a) \cap D_t| = \max_{i \neq t} |N_G(x_i) \cap V(D_t)|.$$

Then $\deg_T(z) = k$ for all $z \in N_G(x_a) \cap D_t$; otherwise if $\deg_T(z) < k$ for some $z \in N_G(x_a) \cap D_t$, then $T + z x_a - w u_t + P$ is a k-tree of G, which is contrary to the choice of T.

We regard each D_t as a rooted tree with root x_t. For a vertex x of D_t, the set of children of x is denoted by Child(x). Let $z \in N_G(x_a) \cap D_t$. Then no child z_1 of z is adjacent to x_t in G since otherwise $T - z_1z + z_1x_t + zx_a - wu_t + P$ is a k-tree of G, a contradiction. Hence for each $r \in N_G(x_a) \cap D_t$, r, Child(x), and $N_G(x_t) \cap D_t$ are pairwise disjoint. By $\deg_T(z) = k$ for all $z \in N_G(x_a) \cap D_t$ and by the choice of x_a, we obtain

$$\begin{aligned}|N_G[S] \cap D_t| &\geq |N_G(x_t) \cap V(D_t)| + \sum_{r \in N_G(x_a) \cap V(D_t)} |\mathrm{Child}(r)| + |\{x_t\}| \\ &= |N_G(x_t) \cap V(D_t)| + (k-1)|N_G(x_a) \cap V(D_t)| + 1 \\ &\geq \sum_{j=1}^{k} |N_G(x_j) \cap V(D_t)| + 1.\end{aligned}$$

Claim 3. *Each x_i and x_j has no common neighbor in $V(G) \setminus V(T)$.*

Proof. Suppose that x_i and x_j has common neighbor $c \in V(G) \setminus V(T)$ for some $1 \leq i < j \leq k$. Then $T + cx_i + cx_j - wu_i + P$ is a k-tree of G, a contradiction. □

Hence it follows from $v \notin N_G(S) \cap (V(G) \setminus V(T))$, the above inequality, Claims 2 and 3 that

$$\begin{aligned}\sum_{j=1}^{k} \deg_G(x_j) &\leq \sum_{j=1}^{k} \left(\sum_{i=1}^{k} (|N_G(x_j) \cap V(D_i)| + |N_G(x_j) \cap (V(G) \setminus V(T))| + |\{w\}| \right) \\ &\leq \sum_{j=1}^{k} (|N_G[S] \cap D_i| - 1) + |N_G(S) \cap (V(G) \setminus V(T))| + k \\ &\leq |N_G[S] \cap T| - 1 + |N_G(S) \cap (V(G) \setminus V(T))| = |N_G[S]| - 2.\end{aligned}$$

This contradicts the assumption $\sigma_k(S; G) \geq |N_G[S]| - 1$. Therefore the theorem is proved. □

5 Remarks

Note that S is not necessarily "connected" in Theorems 4–7. When $S = V(G)$ we obtain Theorems 1–3 by Theorems 4–7.

In addition, Theorems 6 and 7 are sharp in the sense that we cannot replace the lower bounds of the degree sum conditions by $|N_G[S]| - 2$. Consider, for example, the complete bipartite graph $G(A, B)$ with partite sets A and B such that $|A| = t$ and $|B| = (k-1)t + 2$, where t is any positive integer. Let $S = B$ and choose any k independent vetices in S. Then $|N_G[S]| = kt + 2$ and $\sigma_k(S; G) = kt = |N_G[S]| - 2$. Suppose that the graph $G(A, B)$ has k-tree T which contains all the vertices of S. Then the number of edges in T joining A to B is at most $kt = |N_G[S]| - 2$ by $\deg_T(x) \leq k$ for all $x \in V(T)$. This contradicts the fact that the number of edges in T is $|T| - 1 = |N_G[S]| - 1$. Hence the lower bounds of the degree sum conditions in Theorems 6 and 7 are sharp.

The author would like to thank the referees for making many helpful suggestions and comments.

References

References

[1] O. Ore, A note on hamiltonian circuits. Am. Math. Monthly **67** (1960), 55.

[2] K. Ota, Cycles through prescribed vertices with large degree sum. Discrete Math. **145** (1995), 201–210.

[3] R. Shi, 2-neighborhoods and hamiltonian conditions. J. Graph Theory **16** (1992), 267–271.

[4] S. Win, Existenz von Gerüsten mit vorgeschriebenem Maximalgrad in Graphen. Abh. Math. Seminar Univ. Hamburg **43** (1975), 263–267.

[5] T. Yamashita, a private communication.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Low-codimensional Subvarieties Inside Dense Multilinear Varieties

(Extended abstract)

Luka Milićević*

Abstract

Let $G_1, \dots, G_k$ be finite-dimensional vector spaces over a prime field $\mathbb{F}_p$. Let V be a variety inside $G_1 \times \cdots \times G_k$ defined by a multilinear map. We show that if $|V| \geq c|G_1| \cdots |G_k|$, then V contains a subvariety defined by at most $K(\log_p c^{-1} + 1)$ multilinear forms, where K depends on k only. This result is optimal up to multiplicative constant and is relevant to the partition vs. analytic rank problem in additive combinatorics.

1 Introduction

Throughout the paper, we work with a prime field $\mathbb{F}_p$, which is fixed, and finite-dimensional vector spaces $G_1, \dots, G_k$ and H over $\mathbb{F}_p$. We say that a function $\Phi : G_1 \times G_2 \times \cdots \times G_k \to H$ is a *multilinear map* if it is linear in each of its k arguments separately. Furthermore, if the codomain is $\mathbb{F}_p$ instead of H, we say that Φ is a *multilinear form*.

An important goal in additive combinatorics is to obtain a precise understanding of multilinear forms that are not quasirandom. One way of formalizing quasirandomness in this context is to consider the distribution of values of a multilinear form $\alpha : G_1 \times \cdots \times G_k \to \mathbb{F}_p$. If we fix $x_1, \dots, x_{k-1}$, we may consider the corresponding linear form $y_k \mapsto \alpha(x_1, \dots, x_{k-1}, y_k)$. Since it is a linear form, it is either 0, or it takes all values $0, 1, \dots, p-1 \in \mathbb{F}_p$ an equal number of times. Hence, the deviation of distribution of values of α from the uniform distribution corresponds to the frequency of 0 as the value, and a particularly elegant way to express this is to consider the *bias*, defined by

$$\operatorname{bias} \alpha = \mathop{\mathbb{E}}_{x_1 \in G_1, \dots, x_k \in G_k} \omega^{\alpha(x_1, \dots, x_k)},$$

*Mathematical Institute of the Serbian Academy of Sciences and Arts, Kneza Mihaila 36, 11000 Belgrade, Serbia. E-mail: `luka.milicevic@turing.mi.sanu.ac.rs`. Supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through the Mathematical Institute of the Serbian Academy of Sciences and Arts, and by the Science Fund of the Republic of Serbia, Grant No. 11143, *Approximate Algebraic Structures of Higher Order: Theory, Quantitative Aspects and Applications* - A-PLUS.

where $\mathbb{E}_{x\in X}$ stands for the average over elements x of a set X and $\omega = \exp(2\pi i/p)$. Hence, small bias corresponds to quasirandom behaviour. A closely related quantity is the *analytic rank*, defined in [4] as $\operatorname{arank}\alpha = \log_p \operatorname{bias}\alpha^{-1}$.

Let us remark that there are algebraic obstructions to uniform distribution of values. For example, if $\alpha(x_1,\dots,x_k)$ factorizes as $\beta(x_I)\gamma(x_{[k]\setminus I})$, where β is a multilinear form on variables with indices in a set $I \subseteq [k]$ and γ is a multilinear form on the remaining variables, then it is easy to see that $\operatorname{bias}\alpha \geq 1/p$. A remarkable fact is that this is essentially the only way to have large bias. To make this formal, define the *partition rank* [14] of a multilinear form α, denoted $\operatorname{prank}\alpha$, as the smallest number r such that α can be written as a sum of r multilinear forms that factorize in the way above. It is not hard to see that $\operatorname{prank}\alpha = r$ implies $\operatorname{bias}\alpha \geq p^{-r}$. In the opposite direction we have the following theorem.

Theorem 1. *For each $k, r \in \mathbb{N}$ there exists a positive integer $K = K(k,r)$ for which the following holds. Let $\alpha : G_1 \times \dots \times G_k \to \mathbb{F}_p$ be a multilinear form such that* $\operatorname{arank}\alpha \leq r$. *Then* $\operatorname{prank}\alpha \leq K$.

This theorem was first proved in the context of polynomials rather than multilinear forms by Green and Tao [5]. Their approach was refined by Kaufman and Lovett [8] and adjusted to the multilinear setting by Bhowmick and Lovett [2]. These results gave Ackermann-type dependence on the analytic rank. An improvement was obtained by Janzer [6], which gave tower-type bounds. The dependence was improved to polynomial bounds by Janzer [7] and by the author [12]. Finally, the current bounds

$$\operatorname{prank}\alpha \leq K(k)\Big(\operatorname{arank}\alpha(\log_p(\operatorname{arank}\alpha + 1) + 1)\Big)$$

were obtained by Moshkovitz and Zhu [13]. See also the discussion in [13] for other related results.

The key open problem in this context, stated in [1, 9, 10, 11], is to show that the partition rank is linear in the analytic rank. In this paper, we make a contribution to this problem by showing linear bounds for a weaker structural result, which was proved with weaker bounds and played a key role in an earlier proof [12] of (polynomial-bounds version of) Theorem 1. We will discuss the relevance to the proof in [12] after the statement. First, we need some further definitions.

By a *multilinear variety* in $G_1 \times \dots \times G_k$ we think of a zero set of a collection of multilinear maps, which are allowed to depend on a proper subset of coordinates. For example, $U \times V \subseteq G_1 \times G_2$ for some subspaces $U \leq G_1$ and $V \leq G_2$ is also a multilinear variety. We say that a multilinear variety $V \subseteq G_1 \times \dots \times G_k$ has *codimension at most* r if it can be represented as a common zero set of at most r multilinear forms.

The next theorem is our main result. It shows that dense multilinear varieties contain low-codimensional subvarieties. We call it a weak structural result for biased multilinear forms, for the reasons explained after the statement.

Theorem 2 (Weak structural result with linear bounds)**.** *For each $k \in \mathbb{N}$ there exists a positive integer $K = K(k)$ (independent of p) for which the following holds. Let $V \subseteq G_{[k]}$ be a multilinear variety of density c. Then V contains a multilinear variety of codimension at most $K(\log_p c^{-1} + 1)$.*

Going back to Theorem 1, the proof in [12] was a complicated inductive scheme, where Theorems 1 and 2 were two of several steps that implied each other in the inductive step. In particular, Theorem 1 for lower arity was used to infer Theorem 2 (with polynomial bounds). In this paper, we circumvent this, which opens up the possibility of improving the bounds given by the proof in [12].

We remark that biased multilinear forms correspond to dense multilinear varieties. Namely, if $\alpha(x_{[k]})$ has bias c, then we may use a dot product $\cdot$ on G_k to write $\alpha(x_{[k]}) = A(x_{[k-1]})\cdot x_k$ for some multilinear map $A : G_1 \times \dots \times G_{k-1} \to G_k$. The equality $|\{x_{[k-1]} \in G_{[k-1]} : A(x_{[k-1]}) = 0\}| = \operatorname{bias}\alpha|G_{[k-1]}|$ then holds. Thus, for a biased multilinear form α, the theorem above allows us to find a multilinear variety of lower order and of essentially optimal codimension inside the zero set of α, which is a structural result on α.

Let us mention a related result of Chen and Ye [3], who recently showed that the geometric rank, which is the algebro-geometric codimension of the variety $\{A = 0\}$ above, is linear in $\operatorname{arank}\alpha$. Our result shows that algebro-geometric codimension of some irreducible component of $\{A = 0\}$ satisfies linear bounds in $\operatorname{arank}\alpha$. On the other hand, our definition of codimension is more informative than its algebro-geometric counterpart. Moreover, our proof is direct and completely combinatorial, while the proof in [3] depends on various tools in algebraic geometry, including Lang-Weil bounds.

2 Preliminaries

Given an index set $I \subseteq [k]$, we write $G_I = \prod_{i\in I} G_i$. Given a set $X \subseteq G_1 \times \dots \times G_k$, and a tuple x_I, consisting of elements $x_i \in G_i$ for $i \in I$, we define the *slice* X_{x_I} to be the set $\{y_{[k]\setminus I} \in G_{[k]\setminus I} : (x_I, y_{[k]\setminus I}) \in X\}$.

The next lemma allows us to approximate dense varieties by low-codimensional ones.

Lemma 3 (Approximating dense varieties externally, Lemma 12 in [12])**.** *Let $\Phi\colon G_{[k]} \to H$ be a multilinear map. Then, there is a multilinear map $\phi\colon G_{[k]} \to \mathbb{F}_p^s$ such that $\{\Phi = 0\} \subset \{\phi = 0\}$ and $|\{\phi = 0\} \setminus \{\Phi = 0\}| \leq p^{-s}|G_{[k]}|$.*

The next lemma shows that we may use directional convolutions to fill in very dense subsets of low-codimensional varieties. To state it, we define *convolution in direction i* as the operator that acts on functions $f : G_{[k]} \to \mathbb{R}$ as $\mathbf{C}_i f(x_{[k]}) = \mathbb{E}_{y_i\in G_i} f(x_{[i-1]}, y_i + x_i, x_{[i+1,k]}) f(x_{[i-1]}, y_i, x_{[i+1,k]})$.

Lemma 4 (Directional convolutions of varieties)**.** *Let $W \subseteq G_{[k]}$ be a multilinear variety of codimension r. Suppose that $B \subseteq W$ is a subset of size $|B| \leq 2^{-2k}p^{-kr}|G_{[k]}|$. Then*

$$\mathbf{C}_k\mathbf{C}_{k-1}\dots\mathbf{C}_1\mathbb{1}_{W\setminus B}(x_{[k]}) > 0$$

holds for all $x_{[k]} \in W$.

Proof. We prove the claim by induction on k. The base case $k = 1$ is a standard fact in additive combinatorics. Turning to the inductive step, assume the claim holds for $k - 1$. Fix arbitrary $x_{[k]} \in W$. Let $A \subseteq G_k$ be the set of all y_k such that $|B_{y_k}| \geq 2^{-2(k-1)}p^{-(k-1)r}|G_{[k-1]}|$. By averaging, $|A| \leq \frac{1}{4p^r}|G_k|$. Let $y_k \in W_{x_{[k-1]}} \setminus A$ be arbitrary. Hence, since $y_k \notin A$ and

$W_{y_k} \neq \emptyset$ is a multilinear variety in $G_{[k-1]}$ of codimension at most r, then we may apply inductive hypothesis to obtain

$$\mathbf{C}_{k-1} \ldots \mathbf{C}_1 \mathbb{1}_{W_{y_k} \setminus B_{y_k}}(x_{[k-1]}) > 0.$$

Since $|A|< \frac{1}{2}|W_{x_{[k-1]}}|$, there exists y_k such that $y_k, y_k + x_k$ both belong to $W_{x_{[k-1]}} \setminus A$. Thus

$$\mathbf{C}_k \mathbf{C}_{k-1} \ldots \mathbf{C}_1 \mathbb{1}_{W \setminus B}(x_{[k]}) \geq 0. \qquad \square$$

3 Proof of Theorem 2

We prove the theorem by induction on the arity k of the multilinear maps in question. The following lemma is of key importance for the induction step. Assuming the inductive hypothesis, it allows us to find a bounded codimension multilinear variety W such that fibers of the given variety V associated to points of W are dense. We use $O(\cdot)$ notation to hide implicit constants that depend on k only, and not on p.

Lemma 5. *Suppose that Theorem 2 for arity k holds with constant $K = K(k)$. Let $V \subset G_{[k+1]}$ be a variety of density c. Then there exists a multilinear variety $W \subseteq G_{[k]}$ of codimension at most $K(\log_p c^{-1} + 1)$ such that for each $x_{[k]} \in W$ we have $|V_{x_{[k]}}|\geq (c/p)^{O(1)}|G_{k+1}|$.*

Proof. Let $c' > 0$ be a parameter to be specified later. Let us pick a random $x_{k+1} \in G_{k+1}$ and consider the corresponding slice $U = V_{x_{k+1}}$ in $G_{[k]}$. Let $B \subseteq U$ be the set of all $x_{[k]} \in U$ such that $|V_{x_{[k]}}|\leq c'|G|$. Note that these points have sparse fibres of V, thus we think of them as bad points. By linearity of expectation, we have

$$\begin{aligned}
\mathbb{E}|U| - \frac{cc'^{-1}}{2}|B| &= \sum_{x_{[k]} \in G_{[k]}} \mathbb{P}\Big(x_{[k]} \in V_{x_{k+1}}\Big) - \frac{cc'^{-1}}{2} \sum_{\substack{x_{[k]} \in G_{[k]} \\ |V_{x_{[k]}}|\leq c'|G_{k+1}|}} \mathbb{P}\Big(x_{[k]} \in V_{x_{k+1}}\Big) \\
&= \sum_{x_{[k]} \in G_{[k]}} \mathbb{P}\Big(x_{k+1} \in V_{x_{[k]}}\Big) - \frac{cc'^{-1}}{2} \sum_{\substack{x_{[k]} \in G_{[k]} \\ |V_{x_{[k]}}|\leq c'|G_{k+1}|}} \mathbb{P}\Big(x_{k+1} \in V_{x_{[k]}}\Big) \\
&= \sum_{x_{[k]} \in G_{[k]}} \frac{|V_{x_{[k]}}|}{|G_{k+1}|} - \frac{cc'^{-1}}{2} \sum_{\substack{x_{[k]} \in G_{[k]} \\ |V_{x_{[k]}}|\leq c'|G_{k+1}|}} \frac{|V_{x_{[k]}}|}{|G_{k+1}|} \\
&\geq |G_{k+1}|^{-1}|V| - \frac{c}{2}|G_{[k]}| \geq \frac{c}{2}|G_{[k]}|.
\end{aligned}$$

Hence, there is a choice of element $x_{k+1} \in G_{k+1}$ for which $|U|\geq \frac{c}{2}|G_{[k]}|$ and $|B|\leq 2c^{-1}c'|G_{[k]}|$. Applying induction hypothesis, i.e. Theorem 2 for arity k, to U, we find a multilinear variety $W \subseteq U$ of codimension $r \leq K(\log_p c^{-1} + 2)$.

We set $c' = 2^{-2k-1}p^{-2kK}c^{kK+1}$. This ensures that $c' \leq 2^{-k}cp^{-kr}$, allowing us to use Lemma 4 to obtain $\mathbf{C}_k\mathbf{C}_{k-1}\ldots\mathbf{C}_1\mathbb{1}_{W\setminus B}(x_{[k]}) > 0$ for all $x_{[k]} \in W$. It follows that whenever $x_{[k]} \in W$, then $V_{x_{[k]}} \geq c'^{2^k}|G_{k+1}|$, since we have intersection of 2^k subspaces $V_{a_{[k]}}$ of density at least c', where $a_{[k]}$ ranges over 2^k points of a parallelepiped in $W \setminus B$ that attests to $\mathbf{C}_k\mathbf{C}_{k-1}\ldots\mathbf{C}_1\mathbb{1}_{W\setminus B}(x_{[k]}) > 0$. $\square$

We are now ready to prove Theorem 2.

Proof of Theorem 2. As mentioned before, we prove the theorem by induction on k. The base case is $k = 1$ when V becomes a subspace of density c. Then V is a subspace of codimension exactly $\log_p c^{-1}$, completing the proof of the base of induction, with $K(1) = 1$.

We now turn to the inductive step. Suppose that Theorem 2 holds for arity k with constant $K = K(k)$. For each $i \in [k+1]$ apply Lemma 5 to get a multilinear variety $W^{(i)} \subseteq G_{[k+1]\setminus\{i\}}$ of codimension $r \le K(\log_p c^{-1} + 2)$ such that $|V_{x_{[k+1]\setminus\{i\}}}| \ge c'|G_i|$ for all $x_{[k+1]\setminus\{i\}} \in W^{(i)}$, where $c' \ge (c/p)^{O(1)}$. We slightly modify varieties $W^{(i)}$ by defining further multilinear varieties $U^{(i)}$ as zero sets of all multilinear forms that do not depend on variable x_i that appear in definition of some $W^{(j)}$. Thus, we allow repeated use of multilinear forms defining varieties $W^{(j)}$ and the codimension of $U^{(i)}$ is at most $(k+1)r$. Set $\tilde{V} = V \cap \Big(\cap_{i\in[k+1]} G_i \times U^{(i)}\Big)$. Note that variety $\tilde{V}$ has the property that whenever $x_{[k+1]\setminus\{i\}} \in U^{(i)}$ then $|\tilde{V}_{x_{[k+1]\setminus\{i\}}}| \ge c''|G_i|$, where $c'' \ge (c/p)^{O(1)}$. This follows from the fact that $\tilde{V}_{x_{[k+1]\setminus\{i\}}}$ is intersection of the subspace $V_{x_{[k+1]\setminus\{i\}}}$, which is dense as $x_{[k+1]\setminus\{i\}} \in U^{(i)} \subseteq W^{(i)}$, with at most k subspaces, namely $U^{(j)}_{x_{[k+1]\setminus\{i\}}}$, of codimension $(k+1)r$. Let $\varepsilon > 0$ be a parameter to be specified later. Use Lemma 3 to approximate $\tilde{V}$ externally up to error density ε by a multilinear variety A of codimension $\log_p \varepsilon^{-1}$. Set $\tilde{A} = A \cap \Big(\cap_{i\in[k+1]} G_i \times U^{(i)}\Big)$. Note that $|\tilde{A} \setminus \tilde{V}| \le |A \setminus \tilde{V}| \le \varepsilon |G_{[k+1]}|$. We claim that $\tilde{A} = \tilde{V}$.

Claim 6. *If $\varepsilon < {c''}^{k+1}$ then varieties $\tilde{A}$ and $\tilde{V}$ are the same.*

Proof. Suppose that $\tilde{V} \subsetneq \tilde{A}$. There exists a point $x_{[k+1]} \in \tilde{A} \setminus \tilde{V}$. By induction on $i \in [0, k+1]$ we show that there exist at least ${c''}^{i}|G_{[i]}|$ of choices of $y_{[i]}$ such that $(y_{[i]}, x_{[i+1,k+1]}) \in \tilde{A}\setminus\tilde{V}$. The base case $i = 0$ is trivial. Supposing the claim holds for some i, take any $(y_{[i]}, x_{[i+1,k+1]}) \in \tilde{A}\setminus\tilde{V}$. Hence, $\tilde{V}_{y_{[i]},x_{[i+2,k+1]}}$ is a proper subspace of $\tilde{A}_{y_{[i]},x_{[i+2,k+1]}}$. Therefore, the number of y_{i+1} such that $(y_{[i+1]}, x_{[i+2,k+1]}) \in \tilde{A} \setminus \tilde{V}$ is at least $|\tilde{V}_{y_{[i]},x_{[i+2,k+1]}}|$, which is at least $c''|G_{i+1}|$, since $(y_{[i]}, x_{[i+1,k+1]}) \in \tilde{A}$, so $(y_{[i]}, x_{[i+2,k+1]}) \in U^{(i+1)}$. Having proved the inductive step, we obtain $|\tilde{A} \setminus \tilde{V}| \ge {c''}^{k+1}|G_{[k+1]}|$, which is a contradiction with the choice of A. □

Take $\varepsilon = {c''}^{k+1}/2$. Since $\tilde{A}$ has codimension $\log_p \varepsilon^{-1} + (k+1)^2 r \le O(\log_p c^{-1} + 1)$, we are done. □

References

[1] K. Adiprasito, D. Kazhdan, and T. Ziegler. On the Schmidt and analytic ranks for trilinear forms. *arXiv preprint*, `arXiv: 2102.03659`, 2021.

[2] A. Bhowmick and S. Lovett. Bias vs structure of polynomials in large fields, and applications in effective algebraic geometry and coding theory. *IEEE Trans. Inform. Theory*, 69:963–977, 2022.

[3] Q. Chen and K. Ye. Stability of ranks under field extensions. *arXiv preprint*, `arXiv:2409.04034`, 2024.

[4] W.T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011.

[5] B. Green and T. Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2):1–36, 2009.

[6] O. Janzer. Low analytic rank implies low partition rank for tensors. *arXiv preprint*, `arXiv:1809.10931`, 2018.

[7] O. Janzer. Polynomial bound for the partition rank vs the analytic rank of tensors. *Discrete Anal.*, paper no. 7:18 pp, 2020.

[8] T. Kaufman and S. Lovett. Worst case to average case reductions for polynomials. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008)*, pages 166–175, 2008.

[9] D. Kazhdan and T. Ziegler. Applications of algebraic combinatorics to algebraic geometry. *Indag. Math. new ser.*, 32:1412–1428, 2021.

[10] A. Lampert and T. Ziegler. Relative rank and regularization. *Forum Math. Sigma*, 12(e29), 2024.

[11] S. Lovett. The analytic rank of tensors and its applications. *Discrete Anal.*, paper no. 7:10 pp., 2019.

[12] L. Milićević. Polynomial bound for partition rank in terms of analytic rank. *Geom. Funct. Anal.*, 29:1503–1530, 2019.

[13] G. Moshkovitz and D. G. Zhu. Quasi-linear relation between partition and analytic rank. *arXiv preprint*, `arXiv:2211.05780`, 2022.

[14] E. Naslund. The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$. *J. Combin. Theory Ser. A*, 174(105190), 2020.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Properly colored Hamilton cycles and the Bollobás-Erdős conjecture

(Extended abstract)

Aleksa Milojević *   Richard Montgomery†   Alexey Pokrovskiy ‡
Benny Sudakov§

Abstract

We show that if the edges of the large complete graph K_n are colored such that every vertex is incident to fewer than $\lfloor n/2 \rfloor$ edges of the same color, then there exists a properly colored Hamilton cycle, i.e. a Hamilton cycle in which no two consecutive edges have the same color. This result is tight, it answers an old question of Daykin and confirms the conjecture of Bollobás and Erdős.

1 Introduction

Let K_n^c denote an edge-colored complete graph K_n with the edge coloring c, and let $\Delta(K_n^c)$ denote the maximum number of edges of the same color incident to a vertex. In 1976 [4], Daykin asked whether there is some $\mu > 0$ such that $\Delta(K_n^c) < \mu n$ implies the existence of a properly-colored Hamilton cycle in K_n^c. Shortly after, this was proved independently by Bollobás and Erdős [2] with $\mu = 1/69$, and Chen and Daykin [3] with $\mu = 1/17$. Bollobás and Erdős observed that, for each $m \in \mathbb{N}$ and $n = 4m + 1$, there is a red/blue coloring K_n^c with $\Delta(K_n^c) = \lfloor n/2 \rfloor$, and any such colouring can have no properly-colored cycle with odd length, and hence no Hamilton cycle. Motivated by this, they conjectured that as long as $\Delta(K_n^c) < \lfloor n/2 \rfloor$, then K_n^c will have a properly-colored Hamilton cycle.

After subsequent improvements by Shearer [8] (using $\mu = 1/7$) and Alon and Gutin [1] (using $\mu = 1 - 1/\sqrt{2} - o(1)$), in 2016 Lo [7] showed that we may take $\mu = 1/2 - o(1)$, thus showing that

*ETH Zürich, Switzerland. Supported by SNSF Grant 200021-228014. E-mail: aleksa.milojevic@math.ethz.ch.

†Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. Research supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978). Email: richard.montgomery@warwick.ac.uk

‡Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: dralexeypokrovskiy@gmail.com

§ETH Zürich, Switzerland. Supported by SNSF Grant 200021-228014. Email: benjamin.sudakov@math.ethz.ch

Bollobás and Erdős's conjecture holds asymptotically. Here we aim to show that the Bollobás-Erdős conjecture holds for all sufficiently large n, as follows.

Theorem 1.1. *There exists $n_0 \in \mathbb{Z}_{>0}$ such that the following holds for each $n \geq n_0$. Any edge-colored complete graph K_n^c with $\Delta(K_n^c) < \lfloor n/2 \rfloor$ contains a properly colored Hamilton cycle.*

When $n \neq 3 \mod 4$, Theorem 1.1 is tight, as shown by a range of extremal examples given later as **E1**-**E4** in Section 2. In addition to the example of Erdős and Bollobás above (**E1**) which shows that it is tight if $n = 1 \mod 4$, constructions by Lo [7] (**E2** when $k = 1$, see also **F1**) and Fujita and Magnant [5] (**E3**) show that it is tight when $n = 0, 1 \mod 4$ and $n = 0 \mod 2$, respectively. We will give a new construction (**E4**) which shows again that the result is tight if $n = 0 \mod 2$, as well as giving a wide generalisation of Lo's construction (**E2**).

In [2], Bollobás and Erdős also considered the following problem. Let $\delta_{\mathrm{col}}(K_n^c)$ be the minimum over $v \in V(K_n^c)$ of the number of different colours that appear around v. Bollobás and Erdős gave a construction showing that we may have $\delta_{\mathrm{col}}(K_n^c) = \lfloor (n-1)/3 \rfloor$ and yet no properly-colored Hamilton cycle. The construction of Fujita and Magnant [5] given later as **E3** improved this by showing that we can have $\delta_{\mathrm{col}}(K_n^c) = n/2$ if n is even yet have no properly-colored Hamilton cycle. Lo observed that, as $\Delta(K_n^c) + \delta_{\mathrm{col}}(K_n^c) \leq n$, his result quoted above immediately implies that this cannot be improved beyond $\delta_{\mathrm{col}}(K_n^c) = (1 + o(1))n/2$. Similarly, Theorem 1.1 immediately implies the following, which is tight when n is even.

Corollary 1.2. *There exists n_0 such that the following holds for each $n \geq n_0$. Any edge-colored complete graph K_n^c with $\delta_{col}(K_n^c) > \lceil n/2 \rceil$ contains a properly colored Hamilton cycle.*

Fujita and Magnant [5] conjecture that this should hold for all edge-colored (not necessarily complete) graphs G with $\delta_{\mathrm{col}}(G) \geq \lceil (n+1)/2 \rceil$.

2 Extremal examples

Here we give the five extremal examples which each show that the assumption of Theorem 1.1 cannot be relaxed to $\Delta(K_n^c) \leq \lfloor n/2 \rfloor$. These examples are depicted in Figure 1.

Example E1. Let $n = 4k + 1$ and let G be a $2k$-regular graph with vertex set $[n]$. Create K_n^c by coloring all the edges of the graph G red and all the other edges of K_n blue. No Hamilton cycle in this edge-coloring is properly colored, since the edges of every properly-colored cycle have to alternate between red and blue, and thus has an even number of vertices. In this coloring, we have $\Delta(K_n^c) = 2k = \lfloor n/2 \rfloor$.

Example E2. Let $n = 4k$ or $4k + 1$. Let $0 \leq m \leq 2k$ be even. Partition the vertices of K_n into A and B with sizes $2k + m + 1$ and $2k - m - 1$ (if $n = 4k$) or $2k + m + 1$ and $2k - m$ (if $n = 4k + 1$). Further partition A into sets $A_1, \ldots, A_\ell$, each with size at least $m + 1$, for any $\ell \geq 2$. For each $i \in [\ell]$, pick a new color c_i and color all the edges between A_i and B with that color. Pick a new color c_0. For each $i \in [\ell]$, arbitrarily pick a m-regular subgraph on A_i and color these edges with color c_i and color all other edges in A_i with color c_0. Finally, color the edges in B arbitrarily, so that each vertex in B has at most $2k$ edges of each color next to it in total, and let K_n^c be the result.

Every properly colored cycle in K_n^c contains an even number of vertices of A. Indeed, if uv is an edge of the cycle with $u \in B, v \in A_i$ for some $i \in [\ell]$, then the next edge must be vv' for some

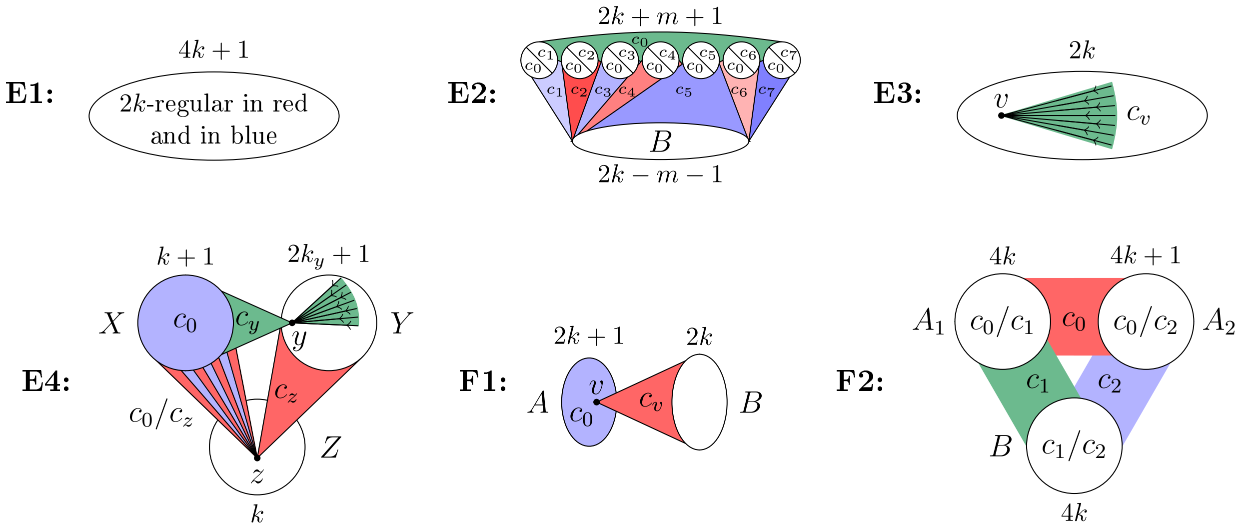


Figure 1: The extremal examples **E1**–**E4** and two special cases of **E2** (**F1** and **F2**).

$v' \in A_j$ with $j \in [\ell] \setminus \{i\}$ to avoid repeating c_i. Then, the next edge must be $v'u'$ for some $u' \in B$ (a color-c_j edge), in order to not repeat c_0, or $v'u'$ for some $u' \in A_j$ (a color-c_j edge). Continuing, we see that whenever the cycle enters A, it must use an even number of vertices from A before exiting to B. Thus, as A has an odd number of vertices, there must be no properly colored Hamilton cycle. Note that $\Delta(K_n^c) = 2k = \lfloor n/2 \rfloor$, so this is indeed an extremal example.

Example E3. Suppose $n = 2k$ and let T be a tournament with vertex set $[n]$ and maximum indegree at most k which does not contain a directed Hamilton cycle. For example, T can be constructed by taking a $(k-1)$-regular tournament on $2k-1$ vertices and adding a single vertex beating all other vertices. Then, assign a color c_v to every vertex $v \in V(T)$ and color all the edges of T directed towards v with color c_v. Let K_n^c be the coloured complete graph corresponding to T (i.e., having forgotten the directions of the edges).

Suppose K_n^c has a properly-colored Hamilton cycle. The corresponding Hamilton cycle in T cannot be directed, so must contain two edges directed towards the same vertex, v say. Both these edges were assigned color c_v, so the cycle is not proper, a contradiction. Thus, this is an extremal example.

Note that if $n = 2k+1$ then any n-vertex tournament with maximum indegree at most $k = \lfloor n/2 \rfloor$ is regular, and thus contains a directed Hamilton cycle (see, [9]), so that this construction only gives extremal examples with an even number of vertices.

Example E4. Let $k_y + 1 \le k$ and set $n = 2k + 2k_y + 2$. Partition $[n] = X \cup Y \cup Z$ into part of sizes $k+1, 2k_y+1$ and k, respectively. Color the edges in X in color c_0. Take distinct new colors c_v, $v \in Y \cup Z$. Direct the edges inside Y into a regular tournament and for each edge $\vec{uv}$ give the edge uv color c_v. For each $x \in X$ and $y \in Y$, give xy color c_y. For each $z \in Z$ and $y \in Y$, give zy color c_z. Color some edges between X and Z with c_0 so that each vertex in X is in $k_y + 1$ of these edges and each vertex in Z is in $\lfloor (k+1) \cdot (k_y+1)/k \rfloor$ or $\lceil (k+1) \cdot (k_y+1)/k \rceil$ of these edges.

Note that vertices in y are incident to at most $k + k_y + 1 = n/2$ edges of color c_y and vertices in X are adjacent to $k + k_y + 1 = n/2$ edges with color c_0. Finally, vertices in Z are adjacent to at most $(2k_y+1) + (k+1) - \lfloor (k+1) \cdot (k_y+1)/k \rfloor = n/2 - \lfloor (k_y+1)/k \rfloor \le n/2$ edges with color c_z. Thus, $\Delta(K_n^c) = n/2$.

Now, suppose that K_n^c has a properly-colored Hamilton cycle S. Suppose there are ℓ maximal intervals $I_1, \ldots, I_\ell$ on S which include at least one vertex in X and no vertices in Z. Label intervals $J_1, \ldots, J_\ell$ so that S is $I_1 J_1 \ldots I_\ell J_\ell$. For each $i \in [\ell]$, let x_i be the number of vertices in $I_i \cap X$ and let z_i be the number of vertices in $J_i \cap Z$, so that $x_i, z_i \geq 1$. As $\sum_{i\in[\ell]} x_i = k+1 > \sum_{i\in[\ell]} z_i$ we can assume[1] by rechoosing the starting interval that, for each $j \in [\ell]$, $\sum_{i\in[j]} x_i > \sum_{i\in[j]} z_i$. Observe that if S travels out of X directly into Y then the next vertex not in Y must be some $z \in Z$ (by a similar argument to that for **E3**), and the preceding edge will have color c_z. Thus, as S cannot have more than two vertices in X in a row, we have that $x_i \leq 2$ for each $i \in [\ell]$. As $x_1 > z_1$, we must then have that $x_1 = 2$ and z_1. Thus, regardless of whether the vertex, z say, in J_1 is preceded by a vertex in X (and thus by two consecutive vertices in X) or Y in the cycle, the preceding edge will have colour c_z. Thus the next vertex in the cycle must be a vertex in X, reached by a colour-c_0 edge. As the next edge in the cycle cannot again have color c_0, we must have $x_2 = 1$, and thus both $z_2 = 1$ and the edge preceding the next vertex in Z, some z' say, has colour c'_z, so that the next vertex is in X and reached by a color-c_0 edge. Repeating this shows eventually that $z_\ell = 1$ and the single vertex in $J_\ell \cap Z$, z'' say, is preceded by a color-$c_{z''}$ edge on the path. Thus, the edge after z'' in the cycle has color c_0, and I_1 cannot have two vertices in X, a contradition. Thus, K_n^c contains no properly-colored Hamilton cycle, and this is indeed an extremal example.

Example F1. Take the example **E2** in the case $m = 0$. In this case, A is partitioned into singletons, so that each vertex in v is assigned some color, c_v say, so that every edge from v to B has color c_v.

Example F2. Let $\ell = 2$, $k = 3k'$, $m = 2k'$ and $n = 4k+1$ in example **E2**. Then, $|A_1| = 4k'$, $|A_2| = 4k'+1$, and $|B| = 4k'$. Furthermore, color B using only colors c_1 and c_2, so that the graph induced on B is $(2k-1)$-regular in color-c_2 edges and $(2k)$-regular in color-c_1 edges.

3 Proof sketch

In this section, we will present a high-level overview of the proof. Our proof uses the absorption method and consists of five steps, which we now outline.

Step 1 - Constructing the reservoir

Fix parameters $\theta \ll \eta \ll \gamma \ll 1$. The first step of the proof will be to pick a random set W of θn vertices, which we will call the reservoir. The main property of this set W will be that any two edges can be connected by many properly colored path using the vertices of W.

This will be possible as long as the coloring of K_n^c is not very close to Example **E3**. We will deal with this case through a separate argument in Step 5. The main statement in this step is the following.

Lemma 3.1. *Let $\gamma \in (0,1)$ be a small parameter and let $m \geq \gamma^{-2}$ be an integer. Suppose that K_m^c is colored such that $\Delta(K_m^c) < (1+\gamma)m/2$, and suppose it is not possible to change $O(\gamma m^2)$ edges of K_m^c and obtain the coloring from Example **E3**.*

[1] This is by the lorry/fuel circuit problem. Choose j which minimises $\sum_{i\in[j]} x_i - \sum_{i\in[j]} z_i$ and relabel $i \mapsto i-j-1$, so that, now, $\sum_{i\in[j]} x_i - \sum_{i\in[j]} z_i \geq 0$ for each $j \in [\ell]$. If $\sum_{i\in[j]} x_i = \sum_{i\in[j]} z_i$ for some $j \in [\ell]$, then take the largest j such that $\sum_{i\in[j]} x_i = \sum_{i\in[j]} z_i = 0$ and relabel $i \mapsto i-j-1$.

Then, for every set $F \subseteq V(K_m)$ of at most γm forbidden vertices and any four distinct vertices $x, y, z, t \in V(K_m^c) \backslash F$, then there exists a properly colored path P of length at most γ^{-1} starting in vertices x, y and ending in vertices z, t, in that order.

Step 2 - Constructing the absorber

The goal in this step will be to construct an absorbing path P with the following property: for every subset $W' \subseteq W$, there exists a properly colored path on $V(P) \cup W'$ ending in the same edges as P. Unfortunately, we will not be able to guarantee something quite so strong - some vertices of W may be very hard to absorb. However, the following proposition is a slight weakening of our statement which will still be sufficient.

Proposition 3.2. *Suppose that the coloring of K_n^c satisfies $\Delta(K_n^c) < \lfloor n/2 \rfloor$ and that is it is not γ-close to either one of the examples **F1**, **F2** or **E3**.*

Then, there exists a set $V_{\text{bad}} \subseteq V(K_n)$ of at most $2\eta^{1/4}n$ non-absorbable vertices, such that for every reservoir set $W \subseteq V(K_n) \backslash V_{\text{bad}}$ and every set F of $\eta^5 n$ forbidden vertices, there is a properly colored path P disjoint from F, with the following properties.

We have $|W \cap P| \leq 4$, $|P| \leq O(\gamma^{-1}|W|)$ and for every subset $W' \subseteq W \backslash V_{\text{bad}}$, there exists a properly colored path on the vertex set $W' \cup V(P)$ ending in the same edges as P.

Step 3 - Constructing the linear forest

Consider now the set of vertices with W and $V(P)$ removed, and denote by n' its cardinality. Since we have removed very few vertices, we still have a bound on the maximum number of edges of the same color incident to a vertex, say $\Delta(K_{n'}^c) \leq n/2 \leq (1 + O(\gamma^{-1}\theta))n'/2$. Our next step will be to show that the set of remaining vertices can be partitioned into a bounded number of properly colored paths.

Proposition 3.3. *Let $\gamma \in (0, 1)$ be a parameter and let K_n^c be an edge-coloring of the complete graph satisfying $\Delta(K_n^c) \leq (1 + \gamma)n/2$. Then, either K_n^c has a spanning linear forest containing at most $\exp(O(\gamma^{-1}))$ properly colored paths or it is γ-close to the Example **E4**.*

Step 4 - Completing the proof

In this step, we combine the previous three results to give the proof of Theorem 1.1, under the assumption that K_n^c is not γ-close to one of the Examples **F1**, **F2**, **E3**, **E4**. Note that we have on purpose avoided giving the exact definition of γ-closeness, since it is quite technical. However, the essence of the definition is clear - a coloring is γ-close to a certain example if one can change the colors of γn^2 edges and obtain the coloring of this example.

Having constructed the properly colored paths disjoint from $W \cup V(P)$, use Lemma 3.1 to show that they can be joined with P and with each other into a single long cycle, which misses only some vertices of W. If we denote by W' the set of vertices missed by this long path, we can absorb them into P using Proposition 3.2, thus constructing a Hamilton cycle.

Step 5 - Extremal examples

The previous four steps show that Theorem 1.1 holds under the assumption that the coloring of K_n^c is not γ-close to one of the four extremal examples, namely Examples **F1**, **F2**, **E3** and **E4**. We have to analyze these four cases separately, and show that in each one of those Theorem 1.1 still holds using an independent argument.

References

[1] Noga Alon and Gregory Gutin. Properly colored Hamilton cycles in edge-colored complete graphs. *Random Structures & Algorithms*, 11(2):179–186, 1997.

[2] Béla Bollobás and Paul Erdős. Alternating Hamiltonian cycles. *Israel Journal of Mathematics*, 23:126–131, 1976.

[3] C. C. Chen and David E Daykin. Graphs with Hamiltonian cycles having adjacent lines different colors. *Journal of Combinatorial Theory, Series B*, 21(2):135–139, 1976.

[4] David E. Daykin. Graphs with cycles having adjacent lines different colors. *Journal of Combinatorial Theory, Series B*, 20(2):149–152, 1976.

[5] Shinya Fujita and Colton Magnant. Properly colored paths and cycles. *Discrete applied mathematics*, 159(14):1391–1397, 2011.

[6] Peter Keevash, Daniela Kühn, and Deryk Osthus. An exact minimum degree condition for Hamilton cycles in oriented graphs. *J. Lond. Math. Soc. (2)*, 79(1):144–166, 2009.

[7] Allan Lo. Properly coloured Hamiltonian cycles in edge-coloured complete graphs. *Combinatorica*, 36:471–492, 2016.

[8] James Shearer. A property of the colored complete graph. *Discrete Mathematics*, 25(2):175–178, 1979.

[9] Carsten Thomassen. Hamilton circuits in regular tournaments. In *North-Holland Mathematics Studies*, volume 115, pages 159–162. Elsevier, 1985.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Using induced subgraphs to express the coefficients of the chromatic polynomial

(Extended abstract)

Kerri Morgan* Lluís Vena†

Abstract

Whitney showed that the coefficients of the chromatic polynomial, $P(G;\lambda)$, could be expressed in terms of the number of (induced) subgraphs of G where the coefficient of $\lambda^{|G|-p}$ is given as a polynomial on $\binom{x_i}{k}$ with integer coefficients, and where the x_i are the number of induced copies of subgraphs.

Farrell applied this approach to find the coefficients of $\lambda^{|G|-3}$ and $\lambda^{|G|-4}$, and Morgan and Delbourgo followed with the coefficient of $\lambda^{|G|-5}$. However, computation of these expressions can be tedious.

Our main contribution is that the finding of these expressions can be systematised, and that they do not depend on the 2–connected graphs with $\leq p+1$ vertices that are formed by gluing two 2–connected graphs through a common clique. For instance, we apply our approach to find an explicit expression of the coefficient of $\lambda^{|G|-6}$ for when the graph is 4-colourable.

As an application, we have used this approach to investigate chromatic uniqueness of wheel graphs; for instance, we can give an alternative proof to the fact that the wheels with an odd number of vertices are chromatically-unique. The wheels of orders 6 and 8 are not chromatically unique, but the situation is unknown for larger order. Understanding the interplay between induced subgraphs of these wheels may shed light on these questions; one can explicitly find, using these coefficient restrictions, the graphs that are chromatically equivalent to the wheel for 6 and 8 vertices.

1 Introduction

The chromatic polynomial $P(G;\lambda)$ of a graph G gives, as its evaluations on the positive integers k, the number of proper colourings of G using k colours. In particular, the chromatic polynomial has $0,1,\ldots,\chi(G)-1$ as roots. The polynomial is multiplicative over connected

*School of Science (Department of Mathematical and Geospatial Sciences), RMIT University, Melbourne, Australia. E-mail: `kerri.morgan@rmit.edu.au`

†Serra Húnter Fellow, IMTech Fellow, Centre de Recerca Matemàtica Fellow, Department of Mathematics, Universitat Politècnica de Catalunya - BarcelonaTech (UPC), Barcelona, Spain. E-mail: `lluis.vena@upc.edu`.
Supported by the Grant PID2023-147202NB-I00 funded by MICIU/AEI/10.13039/501100011033

components, is zero if the graph has loops and can be defined as $P(G;\lambda) = P(G-e;\lambda) - P(G/e;\lambda)$ when e is a non-loop edge where $P(K_1;\lambda) = \lambda$.

The chromatic polynomial can be given [8] as:

$$P(G;\lambda) = \sum_{A \subseteq E(G)} (-1)^{|A|} \lambda^{k(A)} \tag{1}$$

where $k(A)$ is the number of components of the graph $G = (V(G), A)$ and $A \subseteq E(G)$. In particular, the coefficient of $\lambda^{|V(G)|-p}$ is given, up to a sign, by the number of subsets of edges spanning a subgraph of G with $|V(G)|-p$ components. Whitney [8] gave the following expression for the chromatic polynomial of a graph G of order n as $P(G;\lambda) = \sum_{i,j}(-1)^{i+j} m_{ij} \lambda^{n-i}$ where m_{ij} is the number of 2-connected subgraphs of G of rank i and nullity j. He [9] showed that this could be expressed as $P(G;\lambda) = \sum_i m_i \lambda^{n-i}$ where $m_i = \sum_j (-1)^{i+j} m_{ij}$ and $(-1)^i m_i$ is the number of subgraphs of G with i edges containing no broken circuits. Building on this work, Farrell [5] showed that the coefficients of the chromatic polynomial could be expressed as

$$P(G;\lambda) = \sum_i c_{n-i} \lambda^{n-i} \tag{2}$$

where each c_{n-i} is an expression in the counts of 2-connected induced subgraphs of G (we count subgraphs and induced subgraphs in terms of edge sets of G, see (3)).

In Theorem 1, which has been introduced in [6], we give a more precise description of how c_{n-i} can be written as an expression in the counts of 2-connected induced subgraphs, and also allows us to easily implement an algorithm that finds such expression: the complexity of the algorithm for $n-i$ depends on the cube of the number of connected graphs on $i+1$ vertices.

Given A, a set of edges of G, the graph $(V(A), A)$ is the graph with A as its set of edges, and where $V(A) = \{v \in V(G) \mid \exists e \in A, v \text{ adjacent to } e\}$ is its set of vertices. Given a graph H, we let

$$\mathrm{sube}(H,G) = \sum_{A \subseteq E(G)} \mathbf{1}_{(V(A),A) \text{ isomorphic to } H} \quad \Big\| \quad \mathrm{inde}(H,G) = \sum_{A \subseteq E(G)} \mathbf{1}_{G \text{ restr. } V(A) \text{ iso to } H} \tag{3}$$

be, respectively, the *number of subgraphs* isomorphic to H in G and the *number of induced subgraphs* isomorphic to H in G.[1] Regarding the previous work on some specific coefficients, the following are found in [5]:

$$\begin{aligned} c_{n-3} &= -\binom{m}{3} + (m-2)t + C_4 - 2K_4 := \\ &\quad -\binom{\mathrm{inde}(K_2,G)}{3} + (\mathrm{inde}(K_2,G)-2)\mathrm{inde}(K_3,G) + \mathrm{inde}(C_4,G) - 2\mathrm{inde}(K_4,G)\ , \end{aligned} \tag{4}$$

$$\begin{aligned} c_{n-4} &= \binom{m}{4} - \binom{m-2}{2}t + \binom{t}{2} - (m-3)\cdot C_4 + (2m-9)\cdot K_4 \\ &\quad - 6\cdot \mathrm{inde}(K_5,G) - \mathrm{inde}(C_5,G) + \mathrm{inde}(\theta_{2,2,2},G) + 3\mathrm{inde}(W_5,G) + 2\mathrm{inde}(W_5 \setminus \{\mathrm{spoke}\},G)\ . \end{aligned} \tag{5}$$

Whitney's [8] main interest was to give a general account on the expressions that appear in the general coefficient of λ^{n-i} in terms of 2-connected subgraphs, while in [4, 5, 2] the primary focus was to give an expression in terms of induced subgraphs and with the minimum number of terms as possible; the price to pay was that only the first terms could be computed (with reasonable effort) exactly.

[1] Here we consider the subgraphs (both induce and non-induced) as subsets of edges; thus $\mathrm{sube}(C_4,C_4)$ and $\mathrm{inde}(C_4,C_4)$ are 1, while the number of subgraphs of C_4 in C_4 with the usual understanding is 8.

In this paper, we apply our approach to finding chromatic coefficients to investigate chromatic uniqueness of wheel graphs; including an alternative proof to the chromatic-uniqueness of wheels of odd order. For instance, using these coefficient restrictions, one can explicitly find the graphs that are chromatically equivalent to the wheels for 6 and 8 vertices. This may mean that finding further such restrictions may provide information on the chromaticity of larger wheels.

2 Our results

Let $\mathcal{B}$ denote the set of 2–connected graphs. Consider the multiset of elements of $\mathcal{B}$, $\mathcal{T} = \{T_1, ..., T_1, ..., T_r, ..., T_r\}$, with t_i copies of T_i. Then $\Gamma(\mathcal{T}) = (T_1, \ldots, T_r)$ provides a sequence of elements of $\mathcal{T}$ without repetition, $n(\mathcal{T}) = (t_1, \ldots, t_r)$ and $v(\mathcal{T}) = (|V(T_1)|, \ldots, |V(T_r)|)$ give, respectively, the sequence of the number of copies that each T_i has in $\mathcal{T}$ and the number of vertices of each graph in $\mathcal{T}$ (these two sequences have an ordering consistent with $\Gamma(\mathcal{T})$), and for these vectors we define their dimension as the number of their components, so $\dim(n(\mathcal{T})) = \dim(t_1, \ldots, t_r) = r$. Note that a multiset of graphs, such as $\mathcal{T}$ can be viewed as a graph, denoted as $G(\mathcal{T})$, with vertex set $\sqcup_{T\in\mathcal{T}} V(T)$ and edge set $\sqcup_{T\in\mathcal{T}} E(T)$, thus having $t_1 + \ldots + t_r$ connected components.

A 2-connected graph $G = (V, E)$ is said to be *clique-separable* if there is a partition of V into three non-empty vertex sets $V = V_1 \sqcup V_2 \sqcup V_3$ such that there are no edges between V_1 and V_3, V_2 is a complete graph on $|V_2| \geq 2$ vertices, $V_1 \sqcup V_2$ and $V_3 \sqcup V_2$ induce two 2–connected graphs with $\geq |V_2| + 1$ vertices each.

Theorem 1 ([6]). *The chromatic polynomial $P(G; \lambda)$ can be computed as*

$$P(G;\lambda) = \sum_{p=0}^{|G|} \left[\sum_{\substack{\mathcal{T} \text{ multiset of } \mathcal{B} \\ (v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T}) \leq p}} c_p(\mathcal{T}) \prod_{i\in[\dim(n(\mathcal{T}))]} \binom{inde(\Gamma(\mathcal{T})_i, G)}{n(\mathcal{T})_i} \right] \lambda^{|G|-p} \tag{6}$$

$$P(G;\lambda) = \sum_{p=0}^{|G|} \left[\sum_{\substack{\mathcal{T} \text{ multiset of } \mathcal{B} \\ (v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T}) \leq p}} s_p(\mathcal{T}) \prod_{i\in[\dim(n(\mathcal{T}))]} \binom{sube(\Gamma(\mathcal{T})_i, G)}{n(\mathcal{T})_i} \right] \lambda^{|G|-p} \tag{7}$$

where: $(v(\mathcal{T}) - (1, \ldots, 1)) \cdot n(\mathcal{T})$ is the usual scalar product of two vectors, $inde(\cdot, G)$ and $sube(\cdot, G)$ are given by (3), both $c_p(\mathcal{T})$ and $s_p(\mathcal{T})$ are integers depending solely on $\mathcal{T}$ and p (not on G), and $|G| := |V(G)|$. Furthermore:

(i) $c_p(\mathcal{T}) = 0$ *if* $(v(\mathcal{T}) - (1, \ldots, 1)) \cdot n(\mathcal{T}) > p$

(ii) $c_p(\mathcal{T}) = 0$ *if a* $T \in \mathcal{T}$ *is clique-separable*

(iii) *if* $\mathcal{T} = \{T\}$ *and* $|T| = p + 1$,

$$c_p(\mathcal{T}) = \sum_{A\subseteq E(T),\ (V(T),A)\ 2\text{-connected}} (-1)^{|A|}$$

(iv) *if* $\mathcal{T} = \{T\}$ *and* $|T| = p + 1$ *and any* $i \geq 1$,

$$c_{p+i}(\mathcal{T}) = - \sum_{\substack{\mathcal{T}' \text{ multiset of } \mathcal{B},\ \mathcal{T}'\neq\mathcal{T} \\ \mathcal{T}' \text{ containing subgraphs of } T}} c_{p+i}(\mathcal{T}') \prod_{j\in[\dim(n(\mathcal{T}'))]} \binom{inde(\Gamma(\mathcal{T}')_j, T)}{n(\mathcal{T}')_j}$$

(v) when $|\mathcal{T}| = t \geq 2$ *and for each* $i \geq 0$ *we have:*

$$c_{(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})+i}(\mathcal{T}) = \sum_{k_1+\ldots+k_t=i,\ k_s\geq 0} \prod_{T_t\in\mathcal{T}} c_{|T_t|-1+k_t}(\{T_t\}) \ .$$

(vi) for each $p \geq 0$ *and* $\mathcal{T}$ *multiset of* $\mathcal{B}$*,* $c_p(\mathcal{T})$ *are determined by (i), (ii), (iii), (iv), (v).*

The proof of Theorem 1, follows the arguments of both [8, 4, 5, 2] with the aim of giving a general account of the coefficients (in the style of [8]), but in terms of induced subgraphs (as in [4, 5, 2]). The argument consists on two parts, in the first one shows that, indeed, the coefficients can be expressed using only 2-connected subgraphs; then one uses the chromatic polynomials of particular graphs (namely the graph with additional isolated vertices) to conclude the main properties. We highlight (ii) and (vi): the expressions only depend on the non-clique-separable induced subgraphs, and they can be computed recursively (albeit the recursion may involve some terms: in order to compute $c_p(K_4)$ one needs a recursion involving ≥ 100 terms).

Remark. The fact that the coefficients c_{n-i}, and, more generally, the whole chromatic polynomial of G, depends on the counts of its finite subgraphs has been extensively used in the literature, see, for instance [1, 3, 7].

3 Applications

Algorithmic approach. You need one variable for each multiset of 2-connected graphs that are not clique-joined, and such that the sum of the number of vertices minus one is equal to $p-1$. All these can be recursively found using multisets obtained in the computation of previous coefficients. Lastly, the multisets formed by a singleton graph H can be retrieved using some previously found coefficients of multisets via a linear equation with coefficients depending on the multiset (of induced subgraphs of H) and H.

On wheels. As W_{2n-1} are 3-colourable, no induced copy of K_4 can be found in a graph chromatically equivalent to them. Then Lemma 2 gives an alternative proof that W_{2n-1} are chromatically unique [10]; in its proof we combine the properties of (5) and (4), Theorem 1, and the fact that a graph with no induced cycles larger than the triangle is chordal and then it is either a clique or a clique-join.

Lemma 2. *If* G *is a graph on* $n \geq 5$ *vertices and chromatically equivalent to the wheel* W_n *(that is,* $P(G;\lambda) = P(W_n;\lambda)$*), and* $G \ncong W_n$*, then* G *has at least* 2 *induced* C_4 *and an induced* K_4*.*

On the expressions of further coefficients. We have also obtained expressions for the 6-th coefficients explicitly (c_{n-6}) for graphs of chromatic number at most 4 with the aim of better understanding the wheels with an even number of vertices; an expression for c_{n-5} was given in [4, 2].

From the known coefficients, any graph chromatically equivalent to the wheel of order $n+1$, n odd, would require

$$\mathrm{inde}(C_4) - 2\,\mathrm{inde}(K_4) = 0, \text{ and}$$
$$\mathrm{inde}(K_2) - 2\,\mathrm{inde}(K_3) = 0, \text{ and}$$
$$\mathrm{inde}(\theta_{2,2,2}) - 3\,\mathrm{inde}(K_4) - \mathrm{inde}(C_5) + 2\,\mathrm{inde}(W_4\text{-spoke}) + 3\,\mathrm{inde}(W_4) = 0\ .$$

Acknowledgements. This work started in the MATRIX Workshop on Uniqueness and Discernment in Graph Polynomials. MATRIX is Australia's international research institute for the mathematical sciences.

References

[1] M. Abért and T. Hubai. Benjamini-schramm convergence and the distribution of chromatic roots for sparse graphs. *Combinatorica*, 35:127–151, 2015.

[2] H. Bielak. A new method for counting chromatic coefficients. *Annales UMCS Informatica AI*, 3(2005):179–189, 2005.

[3] P. Csikvári and P. E. Frenkel. Benjamini–schramm continuity of root moments of graph polynomials. *European Journal of Combinatorics*, 52:302–320, 2016.

[4] D. Delbourgo and K. Morgan. On the fifth chromatic coefficient. *Australasian Journal of Combinatorics.*, 84(1):56–85, 2022.

[5] E. Farrell. On chromatic coefficient. *Discrete Mathematics*, 29:257–265, 1980.

[6] D. Garijo Royo, D. Orden Martín, and F. Santos Leal, editors. *Discrete Mathematics Days 2024*. Editorial Universidad de Alcalá, Oct. 2024.

[7] V. Patel and G. Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.

[8] H. Whitney. The coloring of graphs. *The Annals of Mathematics*, 33:688, 1932.

[9] H. Whitney. A logical expansion in mathematics. *Bulletin of the American Mathematical Society*, 8:572–579, 1932.

[10] S.-J. Xu and N.-Z. Li. The chromaticity of wheels. *Discrete Mathematics*, 51(2):207–212, 1984.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Counting subsets of integers with no arithmetic progression of a given length

(Extended abstract)

Patrick Morris [*] Miquel Ortega[†] Juanjo Rué [‡]

Abstract

Cameron and Erdős asked if the number of subsets of $[n]$ that avoid arithmetic progressions of length k is at most $2^{(1+o(1))r_k(n)}$, where $r_k(n)$ is the maximum size of a k-AP-free subset of $[n]$. Balogh, Liu, and Sharifzadeh made substantial progress towards answering this question by showing that the number of such subsets is $2^{O(r_k(n))}$ for some infinite sequence of integers n. We extend this result by proving a bound of the form $2^{O(r_k(n))}$ for all $n \in \mathbb{N}$.

1 Introduction

One fundamental question in extremal combinatorics consists of counting objects in a given family that satisfy certain properties. In the arithmetic setting, Cameron and Erdős [6] laid out several questions and conjectures on counting subsets of $[n] = \{1, \ldots, n\}$ under different arithmetic constraints, such as being sum-free, or not containing an arithmetic progression of a given length. These questions are tightly related to the foundational problem of extremal combinatorics of determining the maximum size of a set satisfying such constraints. Indeed, given a family of subsets $\mathcal{A}$ of $[n]$ closed under taking subsets and with maximum size $r_{\mathcal{A}}(n)$, it holds that

$$2^{r_{\mathcal{A}}(n)} \leq |\mathcal{A}| \leq \sum_{i=0}^{r_{\mathcal{A}}(n)} \binom{n}{i} \leq 2^{r_{\mathcal{A}}(n)\log(n)}.$$

[*]Departament de Matemàtiques and IMTECH, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. Email: pmorrismaths@gmail.com

[†]Departament de Matemàtiques and IMTECH, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. Email: miquel.ortega.sanchez-colomer@upc.edu

[‡]Departament de Matemàtiques and IMTECH, Universitat Politècnica de Catalunya (UPC) and Centre de Recerca Matemàtica (CRM), Barcelona, Spain. Email: juan.jose.rue@upc.edu

Researchers of this work are funded through the research projects: PID2023-147202NB-I00 (COCOA) and PCI2024-155080-2 (SRC-ExCo) (PM, MO, JR), MSCA-RISE-2020-101007705 (RandNet) (MO, JR), by the Spanish State Research Agency, through the Severo Ochoa and Maria de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M) (JR), by the European Union's Horizon Europe Marie Skłodowska-Curie grant RAND-COMB-DESIGN - project number 101106032 (PM), and by the FPI grant PRE2021-099120 (MO)

Determining which of the two bounds is closer to the truth usually requires a deeper understanding of the behaviour of sets in $\mathcal{A}$, since it somehow measures how much these are clustered together.

The objects we are concerned with in this work are subsets of $[n]$ that contain no arithmetic progression of length k (or k-AP for short). This property has been one of the main areas of study in additive combinatorics. In particular, one of the cornerstones of the field is the study of $r_k(n)$, the size of the largest k-AP-free set in $[n]$, with particular focus on the case $k = 3$. The classical bounds on $r_k(n)$ are Szemerédi's [18] celebrated theorem, which states that

$$r_k(n) = o(n),$$

and Behrend's [4] construction of a 3-AP-free set, which roughly gives that

$$r_k(n) \geq r_3(n) \gtrsim \frac{n}{2^{(2\sqrt{2}+o(1))\sqrt{\log_2(n)}}}, \tag{1}$$

later generalized by Rankin [15] for larger k.

In fact, very recently there have been major improvements on many of these bounds. Currently, the best known results in the case $k = 3$ are of the form

$$n \exp(-C\sqrt{\log(n)}) \lesssim r_3(n) \lesssim n \exp(-c(\log n)^{1/9})$$

where $C > 0$ in the lower bound is an explicit constant that improves Behrend's construction, due to Elsholtz, Hunter, Proske and Sauermann [7], and the upper bound is due to Bloom and Sisask [5], after the breakthrough of Kelley and Meka [11]. For general k, the current records are

$$n \exp\left(-C_k \log(n)^{1/(k-1)}\right) \lesssim r_k(n) \lesssim n \exp\left(-(\log \log n)^{c_k}\right),$$

for constants $c_k, C_k > 0$. The best lower bound is still the one achieved by Rankin's generalization of Behrend's construction, and the upper bound has been proved by Leng, Sah, and Sawhney [13]. There has been a lot of effort in the last decades dedicated to improving the bounds on $r_k(n)$ before reaching the current status; further references may be found in the cited work. In any case, the precise asymptotics for $r_k(n)$ are still far from being established, with a particularly large gap in the case $k > 3$. This is what made the following question of Cameron and Erdős [6] on the number of k-AP-free sets seem somewhat elusive.

Question 1.1. How does the number of k-AP-free subsets of $[n]$ compare with $2^{r_k(n)}$? In particular, is it $2^{r_k(n)(1+o(1))}$?

However, a surprising result of Balogh, Liu and Sharifzadeh in [1], showed that a deep understanding of the behaviour of $r_k(n)$ is not crucial to establish such a kind of bounds.

Theorem 1.1 (Balogh, Liu, and Sharifzadeh)**.** *There exists a constant $C = C(k) > 0$ such that the number of k-AP-free subsets of $[n]$ is at most $2^{Cr_k(n)}$ for infinitely many values of n.*

The main contribution of our work is a proof that the bound of Theorem 1.1 in fact holds for all values of n.

Theorem 1.2. *There exists a constant $C = C(k) > 0$ such that the number of k-AP-free subsets of $[n]$ is at most $2^{Cr_k(n)}$.*

1.1 Further context

In the setting of extremal graph theory, counting problems were initiated in the 1970s by Erdős, Kleitman, and Rothschild [9]. In particular, they proved that the number of K_r-free graphs on $[n]$ vertices is $2^{(1+o(1))\,\mathrm{ex}(n,K_r)}$, where $\mathrm{ex}(n,H)$ is the maximum size of an H-free graph (again the lower bound here simply follows from taking subsets of a K_r-free graph of maximum size). In the following years, this result was extended to all non-bipartite graphs [8] using the regularity method (see, for example, [14] for a more detailed history). However, the progress on analogous results for bipartite H was much slower. One of the main reasons for this is that $\mathrm{ex}(n,H) = o(n^2)$ for bipartite H, which means that tools such as the regularity lemma of Szemerédi cannot be applied. Another important reason is the fact that precise asymptotics for $\mathrm{ex}(n,H)$ are not known for most bipartite graphs, similarly to the case of k-APs.

The field was revitalized with the advent of the container method. In particular, a flurry of results were proved using the theory of hypergraph containers, a set of tools to count independent sets in hypergraphs simultaneously developed by Balogh, Morris, and Samotij [2] and Saxton and Thomason [17]. We highlight two applications that are particularly relevant for our work. Firstly, Morris and Saxton [14] established that the number of $C_{2\ell}$-free graphs on n vertices is of the form $2^{O(n^{1+1/\ell})} = 2^{O(\mathrm{ex}(n,C_{2\ell}))}$. In fact, they also established a much finer count on the number of $C_{2\ell}$-free graphs of a given size on n vertices. Secondly, using some of the ideas of Morris and Saxton's result and inspired by the work of Balogh, Liu and Sharifzadeh on Theorem 1.1, Ferber, McKinley and Samotij [10] practically settled the problem of counting H-free graphs for general bipartite H. In particular, they proved the following theorem.

Theorem 1.3. *Let H be an arbitrary graph containing a cycle. Suppose that there are positive constants α and A such that $\mathrm{ex}(n,H) \le An^{\alpha}$ for all n. Then there exists a constant C depending only on α, A and H such that for all n, the number of H-free graphs on n vertices is at most $2^{Cn^{\alpha}}$.*

This proves that the number of H-free graphs is $2^{O(\mathrm{ex}(n,H))}$ if $\mathrm{ex}(n,H) = \Theta(n^{\alpha})$ for some constant $\alpha > 0$ for bipartite H, a conjecture that is widely believed to be true. In particular, their result includes all previously known cases of counting H-free graphs up to a constant in the exponent.

Our proof of Theorem 1.2 is similar in spirit to that of Theorem 1.3 in [10] and our main motivation was to have a more complete picture of these counting problems in the additive context. More precisely, in the arithmetic setting, Theorem 1.1 has also spanned further work. In particular, it has been generalised first by Kim [12] for counting corner-free subsets of grids, and then by Behague, Hyde, Morrison, Noel and Wright [3], who have recently proved the following counting version of the Multidimensional Szemerédi Theorem.

Theorem 1.4. *Let d be a positive integer and let $X \subset \mathbb{N}^d$ be a finite set such that $|X| \ge 3$. For infinitely many $n \in \mathbb{N}$, the number of X-free subsets of $[n]^d$ is $2^{O(r_X(n))}$.*

In their theorem, a set is *X-free* if it does not contain a set of the form

$$\{\boldsymbol{b} + r\boldsymbol{x} \colon \boldsymbol{x} \in X\},$$

where $\boldsymbol{b} \in \mathbb{R}^d$ and $r \in \mathbb{R}_{\ge 0}$. We are currently working on extending our methods to also prove such a result for all values of n.

1.2 Proof sketch

As in most applications of the container method, the key ingredient of the proof is a supersaturation result. In other words, given a set of size somewhat greater than $r_k(n)$, one must find many k-APs contained in it. A very simple supersaturation result has the following form.

Lemma 1.5. *For every integer $k \geq 3$ and set $A \subset [n]$ of size $|A| = 2r_k(n)$, there are at least $r_k(n)$ k-APs contained in A.*

Indeed, this may be proved by greedily removing k-APs from the set and finding new ones while $|A| > r_k(n)$. This naive lemma is far from being enough to prove our desired results, but is useful to obtain a stronger version that is sufficient.

Lemma 1.6. *Given an integer $k \geq 3$, the following holds for n, m large enough. For any set $A \subset [n]$ set of size $|A| = mr_k(n)$, the number of k-APs contained in A is at least*

$$m^{\sqrt{\log(m)}} r_k(n). \tag{2}$$

In fact, any bound of the form $m^{k+\varepsilon} r_k(n)$ would suffice to prove Theorem 1.2, applying the container method with an argument due to Morris and Saxton [14]. Our proof of (2) consists of taking a random arithmetic progression L of size roughly $n/m^{\sqrt{\log(m)}}$, intersecting it with A, and applying Lemma 1.5. The key observation is that one only needs $|A \cap L|$ to be larger than $2r_k(|L|)$ in order to apply the lemma, since L is an interval up to translation and scaling, and k-APs are invariant under such transformations.

In order to actually implement such a sketch, some more care is needed. For example, it is useful to embed the set in the modular setting and take arithmetic progressions modulo a prime p, and, besides that, some control on the growth rate of $r_k(n)$ is needed, which we achieve with the following inequalities.

Lemma 1.7. *The extremal number $r_k(n)$ satisfies the inequalities*

$$r_k(mn) \leq mr_k(n) \tag{3}$$

and

$$r_k(m)r_k(n) \leq r_k(2mn) \tag{4}$$

for all $m, n \in \mathbb{Z}_{>0}$.

Inequality (3) is straightforward, and inequality (4) can be proved using an argument given by Ruzsa in [16, Theorem 6.3].

References

[1] József Balogh, Hong Liu, and Maryam Sharifzadeh. The number of subsets of integers with no k-term arithmetic progression. *Int. Math. Res. Not. IMRN*, (20):6168–6186, 2017.

[2] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28(3):669–709, 2015.

[3] Natalie Behague, Joseph Hyde, Natasha Morrison, Jonathan A. Noel, and Ashna Wright. An approximate counting version of the multidimensional szemerédi theorem, 2023.

[4] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U.S.A.*, 32:331–332, 1946.

[5] Thomas F. Bloom and Olof Sisask. An improvement to the Kelley-Meka bounds on three-term arithmetic progressions, 2023. arXiv:2309.02353.

[6] P. J. Cameron and P. Erdős. On the number of sets of integers with various properties. In *Number theory (Banff, AB, 1988)*, pages 61–79. de Gruyter, Berlin, 1990.

[7] Christian Elsholtz, Zach Hunter, Laura Proske, and Lisa Sauermann. Improving Behrend's construction: Sets without arithmetic progressions in integers and over finite fields, 2024. arXiv:2406.12290.

[8] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2(2):113–121, 1986.

[9] P. Erdős, D. J. Kleitman, and B. L. Rothschild. Asymptotic enumeration of K_n-free graphs. In *Colloquio Internazionale sulle Teorie Combinatorie (Roma, 1973), Tomo II*, pages 19–27. Accad. Naz. Lincei, Rome, 1976.

[10] Asaf Ferber, Gweneth McKinley, and Wojciech Samotij. Supersaturated sparse graphs and hypergraphs. *Int. Math. Res. Not. IMRN*, 2020(2):378–402, 03 2018.

[11] Zander Kelley and Raghu Meka. Strong bounds for 3-progressions, 2024. arXiv:2302.05537.

[12] Younjin Kim. The number of k-dimensional corner-free subsets of grids. *Electron. J. Combin.*, 29(2):Paper No. 2.53, 19, 2022.

[13] James Leng, Ashwin Sah, and Mehtaab Sawhney. Improved bounds for Szemerédi's theorem, 2024. arXiv:2402.17995.

[14] Robert Morris and David Saxton. The number of $C_{2\ell}$-free graphs. *Adv. Math.*, 298:534–580, 2016.

[15] R. A. Rankin. Sets of integers containing not more than a given number of terms in arithmetical progression. *Proc. Royal Soc. Edinburgh, Sec. A*, 65:332–344, 1962.

[16] Imre Z. Ruzsa. Solving a linear equation in a set of integers. I. *Acta Arith.*, 65(3):259–282, 1993.

[17] David Saxton and Andrew Thomason. Hypergraph containers. *Invent. Math.*, 201(3):925–992, 2015.

[18] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Signed Planar Graphs with Fractional Balanced Chromatic Number Greater than 2

(Extended abstract)

Reza Naserasr* Lan Anh Pham* Cyril Pujol* Huan Zhou*†

Abstract

A fractional balanced coloring of a signed graph (G, σ) is an assignment of nonnegative weights to balanced sets (i.e., sets that do not induce a negative cycle) such that the total weight assigned to each vertex is at least 1. The minimum possible total weight among all such colorings is called the fractional balanced chromatic number, denoted by $\chi_{fb}(G, \sigma)$. We present an example of a planar signed simple graph with fractional balanced chromatic number greater than 2, thereby disproving a conjecture by Bonamy, Kardoš, Kelly, and Postle. Their conjecture suggested that the fractional arboricity of any planar graph is at most 2, where fractional arboricity assigns weights to acyclic vertex sets rather than to sets that induce no negative cycles. Clearly, for any signature σ of G, the fractional arboricity of G, denoted by $a_f(G)$, serves as an upper bound for $\chi_{fb}(G, \sigma)$.

We show that the supremum of the fractional balanced chromatic number of planar signed simple graphs is at least $2 + \frac{1}{41}$. Additionally, we provide an example of a graph with fractional arboricity $a_f(G) = 2 + \frac{2}{25}$.

1 Introduction

The Four-Color Theorem has played a central role in driving many of the developments in graph theory. Its simplicity and elegance allow it to be expressed in various forms, each

*Université Paris Cité, CNRS, IRIF, F-75013, Paris, France. Supported by ANR-18-IdEx-0001. Emails: reza@irif.fr, lananh2734@gmail.com, cpujol@irif.fr, zhou@irif.fr.

†Affiliated also with Zhejiang Normal University, Jinhua, China.

providing unique perspectives and new directions for exploration. One of the earliest reformulation, proposed by Tait in the 1884 [8], gave rise to the theory of (nowhere-zero) flows and established a connection with the Hamiltonicity of cubic bridgeless planar graphs. Specifically, Tait conjectured that every 3-connected cubic planar graph is Hamiltonian, pointing out that such graphs are also 3-edge-colorable. Before proposing this conjecture, Tait had introduced the first reformulation of the Four-Color Conjecture, stating that every cubic bridgeless planar graph is 3-edge-colorable.

In 1946, Tutte [9] disproved this strong conjecture by building a cubic bridgeless planar graph which is not Hamiltonian. Tutte's example was composed of three identical pieces, which are now called Tutte's fragment.

The dual of Tutte's fragment is known as Wenger's graph and is used by various authors to disprove various potential extensions or relaxations of the four-color theorem. Wenger, in [10], used this gadget to build a planar graph which cannot be partitioned into a bipartite graph and a forest. In this work we similarly use Wenger's gadget to build some examples of planar (signed) graphs whose fractional balanced chromatic number and whose fractional arboricity is strictly larger than 2.

A *signed graph* (G, σ) is a graph $G = (V, E)$ together with an assignment $\sigma : E(G) \to \{+, -\}$. An edge with sign $-$ is a *negative edge* and an edge with $+$ is a *positive edge*. Sign of a substructure in (G, σ) is the product of the signs of edge in the structure. A set of vertices is said to be *balanced* if it induces no negative cycle.

Signed planar simple graphs have also attracted significant attention and research. A reformulation of a conjecture of Máčajová, Raspaud, Škoveira from 2016, [6], is to claim that every signed planar (simple) graph can be partitioned into two balanced sets.

This conjecture, was refuted by Kardoš and Narboni in [4]. They translated the problem to a dual version and used Tutte's fragment to build a counterexample to the dual notion. A direct proof of the claim, using Wenger graph, is given in [7].

Assume (G, σ) is a signed graph without negative loops. A *balanced* (p, q)*-coloring* of (G, σ) is an assignment of q colors to each vertex of G from a platter of p colors, such that each colors class induces a balanced set.

The *fractional balanced chromatic number* of a signed graph (G, σ) is defined as

$$\chi_{fb}(G, \sigma) = \min\{\frac{p}{q} : (G, \sigma) \text{ admits a } (p, q)\text{-coloring } \}.$$

This notion is introduced only recently in [3] and [5].

Fractional balanced chromatic number of a signed graph (G, σ) is upper-bounded by the fractional arboricity of G, which was firstly introduced by Yu and Zuo in [11].

Fractional arboricity of G, denoted $a_f(G)$, is the minimum of $\frac{p}{q}$ such that one can assign (at least) q colors to each vertex using only a total of (at most) p colors such that each color class is acyclic.

Albertson and Berman [1] conjectured that every planar graph G has an acyclic set of order at least $\frac{|V(G)|}{2}$. With the size of maximum acyclic set being limited by $\frac{|V(G)|}{a_f(G)}$ and that this is normally a fair enough of an estimate, Bonamy, Kardoš, Kelly and Postle conjectured in [2] that:

Conjecture 1. *Every planar graph satisfies $a_f(G) \leq 2$.*

Using the signed version of Wenger graph as a key gadget, we build a planar signed (simple) graph for which $\chi_{fb}(G, \sigma) > 2$. The underlying graph of this construction then is a counterexample to Conjecture 1.

Theorem 2. *The supremum of the fractional balanced chromatic number of planar signed simple graphs is at least $2 + \frac{1}{41}$.*

Theorem 3. *There exists a planar graph G with $a_f(G) = 2 + \frac{2}{25}$.*

2 Main result

We need to build a planar signed simple graph (G, σ) which admits no $(2k, k)$-coloring for any integer k. Based on the example (G, σ), we can construct a sequence of planar graphs whose fractional balanced chromatic number approaches $2 + \frac{1}{41}$ as a limit, but we do not have a concrete example of a planar signed simple graph with exactly this value. On the other hand, with similar operations, we can built graphs whose fractional arboricity is exactly $2 + \frac{2}{25}$.

In the following, we will show the fractional balanced chromatic number of (G, σ) is bigger than 2.

Triangle property. Given a (planar) signed graph (G, σ), a negative triangle of (G, σ) is said to have the triangle property if in any balanced $(2k, k)$-coloring, each color appears at least in one of its vertices. A positive triangle is said to have the triangle property if each color appears at most on two of its vertices.

Lemma 4. *Given a signed plane graph (G, σ), there exists a plane extension (G', σ') such that in every balanced $(2k, k)$-coloring of (G', σ'), each facial triangle of G has the triangle property.*

Proof. If T is a facial negative triangle, then we may add a new vertex u_T to the face joining it to all three vertices and choose a signature such that the induced signed K_4 is switching equivalent to $(K_4, -)$.

As each triple of vertices of $(K_4, -)$ induces a negative triangle, in any $(2k, k)$-coloring of $(K_4, -)$ each color appears on at most on two vertices. But then it follows from a basic counting that each color appears exactly twice. Hence, the colors that appear on u_T appear only once on T and the others appear twice on T.

Thus we may assume the property is validated for any facial negative triangle, and use it to enforce the property on facial positive triangles. To this end, we consider a positive facial triangle and we assume it is switched so that all edges are positive. Then we complete this face as in Figure 1, and assume each negative facial triangle admits the triangle property. We now observe that if a color c appears in all three of the u_i, then it can appear in none of the u'_j which is not possible. $\square$

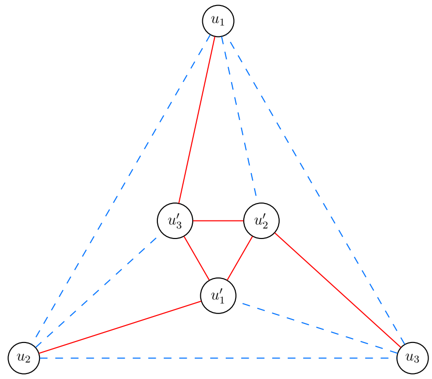


Figure 1: Completing the positive triangle

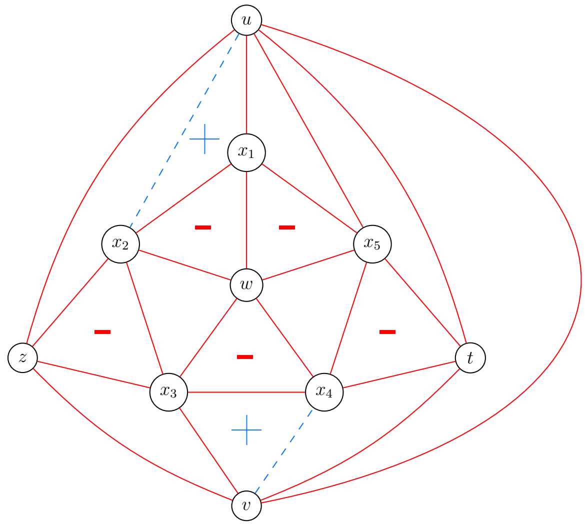


Figure 2: The signed graph $\widehat{W}$

Assuming every facial triangle must satisfy the triangle property next we build the key gadget.

Lemma 5. *In every balanced $(2k, k)$-coloring of the signed graph $\widehat{W}$ of Figure 2 satisfying the triangle property for every facial triangle, u and v have no common color.*

Proof. Toward a contradiction, assume ϕ is such a coloring and that $1 \in \phi(u) \cap \phi(v)$. We partition vertices into two parts, part A consisting of vertices that have 1 in their color set, and $\bar{A}$, vertices that do not have the color 1. Observe that A must be a balanced set, and since uvz and uvt are negative triangles, we have $z, t \in \bar{A}$.

We claim that $x_1 \in \bar{A}$. Otherwise, $x_1 \in A$ and first of all by the triangle property on ux_1x_2 we have $x_2 \in \bar{A}$ and, secondly, since ux_1x_5 is a negative cycle, we have $x_5 \in \bar{A}$. But then by the triangle property on zx_2x_3 and tx_4x_5 we have $x_3, x_4 \in A$, however then the triangle vx_3x_4 fails the triangle property.

Next we claim that $w \in A$. Otherwise, because of the triangle property on negative triangles x_1x_2w and x_5x_1w we have $x_2, x_5 \in A$. Observe that each of ux_2x_3v and ux_5x_4v induces a negative 4-cycle, thus we have $x_3, x_4 \in \bar{A}$. But then the triangle x_3x_4w fails the triangle property.

For the four vertices x_2, x_3, x_4, x_5 we then have the following conditions. Of x_2, x_5 at least one is in $\bar{A}$ because of the negative 4-cycle ux_2wx_5. Of x_3, x_4, because of the triangle property on vx_3x_4, at least one is in $\bar{A}$. Of x_2, x_3 one has to be in A because of the triangle property on zx_2x_3 and similarly, because of the triangle tx_4x_5 at least one of x_4, x_5 should be in A. That leaves us with two possibilities: i. $x_2, x_4 \in A$ and $x_3, x_5 \in \bar{A}$; ii. $x_2, x_4 \in \bar{A}$ and $x_3, x_5 \in A$. In case *i.* A induces negative 5-cycle ux_2wx_4v and in case *ii.* A induces negative 5-cycle ux_5wx_3v. This contradicts the fact that A is a balanced set, thus proving the claim. ☐

Remark 6. *To prove the lemma 5 we have used the triangle property on the following seven triangles:* ux_1x_2,vx_3x_4, wx_1x_2, wx_1x_5, wx_3x_4, zx_2x_3 *and* tx_4x_5, *where the first two are positive and five others are negative.*

Theorem 7. *There exists a planar signed simple graph whose fractional balanced chromatic number is strictly larger than* 2.

Proof. First, applying Lemma 4, add necessary pieces inside seven faces in Remark 6 of $\widehat{W}$ so that in the resulting plane signed graph $\widehat{W}'$ each facial triangle of $\widehat{W}$ has the triangle property. Then consider three vertices and between each pair put a copy of $\widehat{W}'$ where u and v are the two selected vertices. Let (G, σ) be the resulting signed graph. By the construction G is a simple graph. We have $\chi_{fb}(G, \sigma) > 2$ because otherwise it admits a $(2k, k)$-coloring for some k but then, by Lemma 5, we must use distinct colors for each of the original three vertices, but that requires $3k$ colors. □

3 Remarks

Since we use three copies of the gadget $\widehat{W}'$ and identify three pairs of vertices, the final graph (G, σ) in proof of Theorem 7 has 66 vertices. Furthermore, by identifying the 3 copies of z in the central 6-face, we get an example on 64 vertices.

Extending the construction on K_4, we will have a graph on 130 vertices whose fractional balanced chromatic number is $2+2/85$, and whose fractional arboricity is $2+2/31$. Similarly, by identifying copies of z, we may reduce the number of vertices to 127.

By iterating the construction, we can show that the supremum of the fractional balanced chromatic number of planar signed simple graphs is at least $2 + 1/41$. Unlike the case of balanced coloring, we can give a concrete example G of a planar graph satisfying $a_f(G) = 2 + 2/25$, where this value remains unchanged regardless of whether the construction is iterated.

References

[1] Michael O Albertson and David M Berman. A conjecture on planar graphs. *Graph theory and related topics*, 357:357, 1979.

[2] Marthe Bonamy, František Kardoš, Tom Kelly, and Luke Postle. Fractional vertex-arboricity of planar graphs, 2020.

[3] Andrea Jimenez, Jessica McDonald, Reza Naserasr, Kathryn Nurse, and Daniel Quiroz. Balanced-chromatic number and hadwiger-like conjectures. *Priprint*, 2024+.

[4] František Kardoš and Jonathan Narboni. On the 4-color theorem for signed graphs. *European J. Combin.*, 91:Paper No. 103215, 8, 2021.

[5] Luis Kuffner, Reza Naserasr, Lujia Wang, Xiaowei Yu, Huan Zhou, and Xuding Zhu. Fractional balanced colouring of signed graphs. *Priprint*, 2025+.

[6] Edita Máčajová, André Raspaud, and Martin Škoviera. The chromatic number of a signed graph. *Electron. J. Combin.*, 23(1):Paper 1.14, 10, 2016.

[7] Reza Naserasr and Lan Anh Pham. Complex and homomorphic chromatic number of signed planar simple graphs. *Graphs Combin.*, 38(3):Paper No. 58, 22, 2022.

[8] P. G. Tait. Listing‘s topology. *Phil. Mag. (5th Ser.)*, 17:30–46, 1884.

[9] W. T. Tutte. On Hamiltonian circuits. *J. London Math. Soc.*, 21:98–101, 1946.

[10] Gerd Wegner. Note on a paper of B. Grünbaum on acyclic colorings. *Israel J. Math.*, 14:409–412, 1973.

[11] Qinglin Yu and Liancui Zuo. Fractional vertex arboricity of graphs. In *Discrete geometry, combinatorics and graph theory*, volume 4381 of *Lecture Notes in Comput. Sci.*, pages 245–252. Springer, Berlin, 2007.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Multipartite nearly orthogonal sets over finite fields

(Extended abstract)

Rajko Nenadov * Lander Verlinde †

Abstract

For a field $\mathbb{F}$ and integers d, k and ℓ, a set $A \subseteq \mathbb{F}^d$ is called (k, ℓ)-nearly orthogonal if all vectors in A are non-self-orthogonal and every $k+1$ vectors in A contain $\ell+1$ pairwise orthogonal vectors. Recently, Haviv, Mattheus, Milojević and Wigderson have improved the lower bound on nearly orthogonal sets over finite fields, using counting arguments and a hypergraph container lemma. They showed that for every prime p and an integer ℓ, there is a constant $\delta(p, \ell)$ such that for every field $\mathbb{F}$ of characteristic p and for all integers $d \geq k \geq \ell + 1$, $\mathbb{F}^d$ contains a (k, ℓ)-nearly orthogonal set of size $d^{\delta k / \log k}$. This nearly matches an upper bound $\binom{d+k}{k}$ coming from Ramsey theory. Moreover, they proved the same lower bound for the size of a largest set A where for any two subsets of A of size $k+1$ each, there is a vector in one of the subsets orthogonal to a vector in the other one. We prove a common generalisation of this result, showing that essentially the same lower bound holds for the size of a largest set $A \subseteq \mathbb{F}^d$ with the stronger property that given any family of subsets $A_1, \ldots, A_{\ell+1} \subseteq A$, each of size $k+1$, we can find a vector in each A_i such that they are all pairwise orthogonal. Rather than combining both counting and container arguments, we make use of a multipartite asymmetric container lemma that allows for non-uniform co-degree conditions. This lemma was first discovered by Campos, Coulson, Serra and Wötzel, and we provide a new and short proof for this lemma.

1 Introduction

Consider an arbitrary field $\mathbb{F}$ and integers d, k, ℓ such that $k \geq \ell$. A set $A \subseteq \mathbb{F}^d$ of non-self-orthogonal vectors such that for any $k+1$ vectors there is a subset of size $\ell+1$ of pairwise orthogonal vectors is called a (k, ℓ)-nearly orthogonal set. The largest size of such a set is denoted by $\alpha(d, k, \ell, \mathbb{F})$. The question of determining this number for $\mathbb{F} = \mathbb{R}$ was posed by Erdős in 1988 (see [10]).

*School of Computer Science, University of Auckland, New Zealand. Email: rajko.nenadov@auckland.ac.nz.
Research supported by the New Zealand Marsden Fund.
†School of Computer Science, University of Auckland, New Zealand. Email: lver263@aucklanduni.ac.nz.

An easy lower bound when $\ell = 1$ can be found by considering the vectors of k pairwise disjoint orthogonal bases of $\mathbb{R}^d$, showing that the minimal size is at least kd. Erdős conjectured this to be tight for $k = 2$, which was proven in 1991 by Rosenfeld [18]. However, this lower bound is not tight for larger k, as was shown by Füredi and Stanley [14], proving the existence of a (5,1)-nearly-orthogonal set in $\mathbb{R}^4$ of size 24. A more general lower bound was given by Alon and Szegedy [3] (extending a result of Frankl and Rödl [13]), showing that for every integer ℓ, there exists a constant $\delta(\ell) > 0$ such that whenever $k \geq \max\{\ell, 3\}$, $\alpha(d, k, \ell, \mathbb{R}) \geq d^{\delta \log k / \log\log k}$.

More recently, Balla [4] considered a bipartite version of this problem when $\ell = 1$. For integers d, k and a field $\mathbb{F}$, let $\beta(d, k, \mathbb{F})$ denote the maximum size of a set $A \subseteq \mathbb{F}^d$ of non-self-orthogonal vectors such that for any two subsets $A_1, A_2 \subseteq A$ of size $k + 1$ each, there exist $v_1 \in A_1$ and $v_2 \in A_2$ that are orthogonal. It clearly follows that $\alpha(d, k, 1, \mathbb{F}) \geq \beta(d, k, \mathbb{F})$, as one can consider $A_1 = A_2$. For this extended notion Balla proved that for all integers d and $k \geq 3$, there exists a constant $\delta > 0$ such that $\beta(d, k, \mathbb{R}) \geq d^{\delta \log k / \log\log k}$, matching the lower bound on $\alpha(d, k, 1, \mathbb{R})$ by Alon and Szegedy. Nevertheless, this lower bound is still rather far away from the best current upper bound $\alpha(d, k, 1, \mathbb{R}) = O(d^{(k+1)/3)})$ for fixed k, proven by Balla, Letzter and Sudakov [5].

Besides $\mathbb{R}$, the maximal size of nearly orthogonal sets has also been studied in finite fields. This was first motivated by Codenotti, Pudlák and Resta [11], who showed the link between nearly-orthogonal sets and computational complexity of arithmetic circuits. Golovnev and Haviv [15] proved the existence of a constant $\delta > 0$ such that, for infinitely many integers d, $\alpha(d, 2, 1, \mathbb{F}_2) \geq d^{1+\delta}$, showing a clear difference between nearly orthogonal sets in finite fields compared to the reals, where the upper bound by Balla et al. implied $\alpha(d, 2, 1, \mathbb{R}) = O(d)$. This result was improved by Bamberg, Bishnoi, Ihringer and Ravi [8], giving an explicit construction of a (2,1)-nearly orthogonal set in $\mathbb{F}_2^d$ of size $d^{1.2895}$. Chawin and Haviv [10] considered the bipartite version and proved that for every prime p, there exists a positive constant $\delta(p)$ such that for every field $\mathbb{F}$ of characteristic p and for all integers $k \geq 2$ and $d \geq k^{1/(p-1)}$, it holds that $\beta(d, k, \mathbb{F}) \geq d^{\delta k^{1/(p-1)} / \log k}$. Hence, for $p = 2$, it follows that

$$\alpha(d, k, 1, \mathbb{F}_2) \geq \beta(d, k, \mathbb{F}_2) \geq d^{\Omega(k / \log k)}. \tag{1}$$

This almost matches an upper bound that follows easily from Ramsey theory, up to the logarithmic term in the exponent: consider an arbitrary $(k, 1)$-almost orthogonal set in $\mathbb{F}^d$ and construct its orthogonality graph, where the vertices correspond to the vectors and there is an edge if the vectors are orthogonal. Clearly this graph does not contain a clique of size $d + 1$, nor does it have an independent set of size $k + 1$, because that would contradict that the set is $(k, 1)$-almost orthogonal. Hence, the size of the set is upper bounded by the Ramsey number $R(d + 1, k + 1)$. The upper bound on $R(d + 1, k + 1)$ by Erdős and Szekeres [12] gives $\alpha(d, k, 1, \mathbb{F}_2) < \binom{d+k}{k}$. For $d \gg k$, this implies $\alpha(d, k, 1, \mathbb{F}_2) \leq d^{\Theta(k)}$.

In a recent paper, Haviv, Mattheus, Milojević and Wigderson [16] obtained the bound matching (1) for fields of characteristic larger than 2, and also extended it to the non-bipartite setup for arbitrary $\ell > 1$. Specifically, they showed that for every prime p and integer $\ell \geq 1$, there exist constants $\delta_1(p), \delta_2(p, \ell) > 0$, such that for every field $\mathbb{F}$ of characteristic p and for all integers $d \geq k \geq \ell + 1$, it holds that

$$\beta(d, k, \mathbb{F}) \geq d^{\delta_1 k / \log k}$$
$$\alpha(d, k, \ell, \mathbb{F}) \geq d^{\delta_2 k / \log k},$$

proving that $\alpha(d,k,\ell,\mathbb{F}) \geq \beta(d,k,\mathbb{F})$ also holds. We propose a generalisation of this result. Instead of considering $\alpha(d,k,\ell,\mathbb{F})$ and $\beta(d,k,\mathbb{F})$ separately, we consider a common generalisation $\beta(d,k,\ell,\mathbb{F})$ defined as the maximum size of a set $A \subseteq \mathbb{F}$ of non-self-orthogonal vectors such that for any $\ell+1$ subsets $A_1,\dots,A_{\ell+1} \subseteq A$ of size $k+1$ each, there exist $v_1 \in A_1,\dots,v_{\ell+1} \in A_{\ell+1}$ that are all pairwise orthogonal. We prove analogous bounds on $\beta(d,k,\ell,\mathbb{F})$.

Theorem 1. *For every prime p and integer $\ell \geq 1$, there exists a constant $\delta = \delta(p,\ell) > 0$ such that for every field $\mathbb{F}$ of characteristic p and for all integers $d \geq k \geq \ell+1$, it holds that*

$$\beta(d,k,\ell,\mathbb{F}) \geq d^{\delta k/\log k}.$$

Note that Theorem 1 implies the result by Haviv et al. By taking $\ell = 1$, we retrieve the bound on $\beta(d,k,\mathbb{F})$ and by setting $A_1 = \cdots = A_{\ell+1}$, we retrieve the bound on $\alpha(d,k,\ell,\mathbb{F})$.

The main distinction between the proofs of the results is that we unify their methods to find the nearly-orthogonal sets and their bipartite variants. To be more precise, their proof relies on the fact that the orthogonality graph G arising from finite fields has strong pseudorandom properties, and uses two different counting results:

1. Using an argument based on ideas of Alon and Rödl [2], Haviv et al. derived an upper bound on the number of pairs of sets in pseudorandom graphs with no edge across (used to bound β).

2. Using hypergraph containers [6, 19], they upper bounded the number of sets of size k without a copy of K_ℓ (used to bound α).

Having these counting results at hand, the proof proceeds by showing that a random induced subgraph of G avoids sets of certain size which do not contain a copy of K_ℓ, and it avoids pairs of subsets without an edge across. To make this actually work, a tensor-product trick along the lines of Alon and Szegedy [3] is applied.

We consolidate these two counting results by developing a container-type result for ℓ-tuples of subsets without any clique K_ℓ that contains a vertex in each subset of the tuple (Lemma 3). We will call such a tuple a 'bad' ℓ-tuple, as this is exactly the structure that we want to avoid. In particular, rather than counting the number of ℓ-tuples of k-subsets without such a copy of K_ℓ, we construct a small family $\mathcal{C}$ of small subsets with the property that for any bad ℓ-tuple $(U_1,\dots,U_\ell)$, there exists $S \in \mathcal{C}$ such that $U_i \subseteq S$ for some $i \in [\ell]$.

The proof of Lemma 3 uses an asymmetric, multipartite hypergraph container lemma. This lemma was first developed by Campos, Coulson, Serra and Wötzel [9, Theorem 2.1]. They came up with this lemma in the study of sets with bounded doubling factor. An iterated application of this lemma on the right hypergrapgh yields Lemma 3, as we will show in Section 3. In Section 2, we first review this asymmetric container lemma. We provide a different formulation to the lemma, which is practically equivalent to the one by Campos et al. The advantage of this formulation is that allows us to give a considerably shorter proof of the lemma, hopefully further improving accessibility of asymmetric, multipartite containers.

2 Containers with varying co-degree conditions

Hypergraph containers, developed by Balogh, Morris and Samotij [6] and Saxton and Thomason [19], provide a framework for studying independent sets in hypergraphs. On a high level, the main idea of hypergraph containers is to show that given any independent set I in a hypergraph $\mathcal{H}$, one can carefully choose a small subset $F \subseteq I$, called a *fingerprint*, such that just based on F one can confine I to some $C \subseteq V(\mathcal{H})$, called a *container*. Importantly, the size of C is bounded away from $|V(\mathcal{H})|$. This innocent looking statement has far reaching consequences. For a thorough introduction, see [7].

To make use of the standard hypergraph container framework, one typically needs that the given hypergraph has fairly uniform degree and co-degree distribution. This will not be the case in our application. In particular, our hypergraph is ℓ-partite and degrees depend on which part the vertex belongs to. More generally, the number of edges sitting on a given set of vertices depends on which parts these vertices lie in. Therefore, we require an asymmetric multipartite container lemma. The asymmetric lemma of Campos, Coulson, Serra and Wötzel [9] is developed for exactly this case. In this section, we present and prove a slight reformulation of the lemma, which allows for a more concise and straightforward analysis of the container algorithm. The statement is inspired by the one used by Nenadov and Pham [17].

Let V be an arbitrary set. For a subset $T \subseteq V$, we let $\langle T \rangle$ denote the upset of T, i.e. $\langle T \rangle = \{S \subseteq V : T \subseteq S\}$.

Lemma 2. *Let $\ell \in \mathbb{N}$ and let $\mathcal{H}$ be an ℓ-partite ℓ-uniform hypergraph on the vertex set $(V_1, \ldots, V_\ell)$. Suppose there exist $p_1, \ldots, p_\ell \in (0, 1]$, $K > 1$ and a probability measure ν over 2^V, $V = \bigcup_i V_i$, supported on $E(\mathcal{H})$ such that for all sets $S \subseteq [\ell]$, $T \in \prod_{i \in S} V_i$ and $i \in S$, we have*

$$\nu(\langle T \rangle) \leq \frac{K}{|V_i|} \prod_{j \in S \setminus \{i\}} p_j. \tag{2}$$

Let $U = (U_1, \ldots, U_\ell)$, $U_i \subseteq V_i$, be an arbitrary independent set in $\mathcal{H}$ such that $|U_j| \geq p_j |V_j|$ for all $j \in [\ell]$. Then,

- *(i) there is a unique ℓ-tuple $\mathbf{F} = (\emptyset, F_2, \ldots, F_\ell)$ associated with $(U_1, \ldots, U_\ell)$, where $F_j \subseteq U_j$ and $|F_j| = |V_j| p_j$ for $2 \leq j \leq \ell$;*
- *(ii) there is an index $i = i(\mathbf{F}) \in [\ell]$, a corresponding set $C_i = C_i(\mathbf{F}) \subseteq V_i$ and a constant $\zeta = \zeta(K, \ell) > 0$ such that $|C_i| \leq (1 - \zeta)|V_i|$ and $U_i \subseteq F_i \cup C_i$.*
- *(iii) Furthermore, given any $U'_1, \ldots, U'_\ell$ such that $F_i \subseteq U'_i \subseteq U_i$, the ℓ-tuple $\mathbf{F} = (\emptyset, F_2, \ldots, F_\ell)$ associated with $(U_1, \ldots, U_\ell)$ is also associated with $(U'_1, \ldots, U'_\ell)$.*

Whenever ν is the uniform probability measure over $\mathcal{H}$, i.e. $\nu(X) = 1/e(\mathcal{H})$ if $X \in E(\mathcal{H})$ and $\nu(X) = 0$ otherwise, this condition is just a co-degree condition analogous to the one in [6, 19]. Indeed, we have

$$\nu(\langle T \rangle) = \sum_{S \subseteq V : T \subseteq S} \nu(S) = \Delta_T \cdot \frac{1}{e(\mathcal{H})},$$

where Δ_T denotes the codegree of the set T, that is, the number of hyperedges in $\mathcal{H}$ which contain T. Moreover, it is exactly when considering the uniform measure, one retrieves the asymmetric container lemma of Campos et al. Our proof follows the recent proof of the original hypergraph containers by Nenadov and Pham closely.

3 Nearly Orthogonal Sets

The following result has the role of Theorem 2.4 and Theorem 2.6 in [16]. Unlike these two theorems, here we do not count the number of 'bad' tuples of sets. Rather, we give an upper bound on the size of a family which underpins such tuples. To this end, we will need the notion of (n, D, λ)-graphs. Recall that a D-regular graph on n-vertices is an (n, D, λ)-graph if the absolute values of all but the largest eigenvalue of its adjacency matrix are at most λ. Given a graph G, we denote with $\mathfrak{I}_\ell(G)$ the family of all ℓ-tuples $(U_1, \ldots, U_\ell)$, where $U_i \subseteq V(G)$ for each $i \in [\ell]$, such that there is no copy $(v_1, \ldots, v_\ell)$ of K_ℓ in G with $v_i \in U_i$ for $i \in [\ell]$. Note that there is no condition on disjointness of the U_i.

Lemma 3. *Let $c \in (0, 1/2]$ be a constant and n a sufficiently large integer. Suppose G is an (n, D, λ)-graph with $D \geq cn$. There exists $C = C(\ell, c) > 0$ and a family $\mathcal{C}$ of subsets of $V(G)$ with the following properties:*

- $|\mathcal{C}| \leq n^{C \log n}$;
- *each $S \in \mathcal{C}$ is of size $|S| \leq C\lambda$;*
- *for every $(U_1, \ldots, U_\ell) \in \mathfrak{I}_\ell(G)$, there exists $S \in \mathcal{C}$ and $i \in [\ell]$ such that $U_i \subseteq S$.*

Note that we make no assumptions on λ in Lemma 3. However, as $D \geq cn$, for $\lambda > |D|/(c \cdot C)$ we have $C\lambda > n$, thus the upper bound on $S \in \mathcal{C}$ becomes trivial. The proof of this lemma is obtained by constructing a hypergraph $\mathcal{H}$ by taking ℓ disjoint copies of $V(G)$, named $V_1, \ldots, V_\ell$, and putting a hyperedge on top of $(v_1, \ldots, v_\ell)$, $v_i \in V_i$ for $i \in [\ell]$, if the corresponding vertices in G form a copy of K_ℓ. We iteratively apply Lemma 2 until we obtain the desired family $\mathcal{C}$.

With Lemma 3 at hand, the proof of Theorem 1 follows the structure of Haviv et al. We consider the orthogonality graph $G(p, t)$, which has vertex set equal to the vectors in $\mathbb{F}_p^t$ and $u \sim v$ if $\langle u, v \rangle = 0$. It is well known that this graph is an (n, D, λ)-graph (e.g. see [1, Page 220]). Furthermore, we sample enough m-tuples of non-self-orthogonal vectors, where m and t are carefully chosen integers so that $t^m \leq d$, along with some other conditions to make the computations work. Since the orthogonality graph is an (n, D, λ)-graph, we find a family $\mathcal{C}$ of containers using Lemma 3. We show that with non-zero probability, no k vectors contained in an m-tuple of our random set are all contained in some $S \in \mathcal{C}$. Consider the set of m-tuples that satisfies this property. Setting W to be the set of vectors obtained by the m-fold tensor product of every tuple, and adding $d - t^m$ zeroes at the of each vector in W, yields the desired multipartite nearly orthogonal set in $\mathbb{F}_p^d$.

References

[1] N. Alon and M. Krivelevich. Constructive bounds for a Ramsey-type problem. *Graphs and Combinatorics*, 13:217–225, 1997.

[2] N. Alon and V. Rödl. Sharp bounds for some multicolor Ramsey numbers. *Combinatorica*, 25:125–141, 2005.

[3] N. Alon and M. Szegedy. Large sets of nearly orthogonal vectors. *Graphs and Combinatorics*, 15(1):1–4, 1999.

[4] I. Balla. Orthonormal representations, vector chromatic number, and extension complexity. *Bulletin of the London Mathematical Society*, 2023.

[5] I. Balla, S. Letzter, and B. Sudakov. Orthonormal representations of H-free graphs. *Discrete & Computational Geometry*, 64:654–670, 2020.

[6] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[7] J. Balogh, R. Morris, and W. Samotij. The method of hypergraph containers. In *Proceedings of the international congress of mathematicians 2018, ICM 2018, Rio de Janeiro, Brazil, August 1–9, 2018. Volume IV. Invited lectures*, pages 3059–3092. Hackensack, NJ: World Scientific; Rio de Janeiro: Sociedade Brasileira de Matemática (SBM), 2018.

[8] J. Bamberg, A. Bishnoi, F. Ihringer, and A. Ravi. Ramsey numbers and extremal structures in polar spaces. *arXiv preprint arXiv:2406.03043*, 2024.

[9] M. Campos, M. Coulson, O. Serra, and M. Wötzel. The typical approximate structure of sets with bounded sumset. *SIAM Journal on Discrete Mathematics*, 37(3):1386–1418, 2023.

[10] D. Chawin and I. Haviv. Nearly orthogonal sets over finite fields. In *40th International Symposium on Computational Geometry (SoCG 2024)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[11] B. Codenotti, P. Pudlák, and G. Resta. Some structural properties of low-rank matrices related to computational complexity. *Theoretical Computer Science*, 235(1):89–107, 2000.

[12] P. Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[13] P. Frankl and V. Rödl. Forbidden intersections. *Transactions of the American Mathematical Society*, 300:259–286, 1987.

[14] Z. Füredi and R. Stanley. Sets of vectors with many orthogonal pairs. *Graphs and Combinatorics*, 8:391–394, 1992.

[15] A. Golovnev and I. Haviv. The (generalized) orthogonality dimension of (generalized) Kneser graphs: Bounds and applications. *Theory of Computing*, 18(1):1–22, 2022.

[16] I. Haviv, S. Mattheus, A. Milojević, and Y. Wigderson. Larger nearly orthogonal sets over finite fields. *Discrete Mathematics*, 348(4):114373, 2025.

[17] R. Nenadov and H. T. Pham. Short proof of the hypergraph container theorem. *arXiv preprint arXiv:2408.08514*, 2024.

[18] M. Rosenfeld. Almost orthogonal lines in E^d. *DIMACS Series in Discrete Mathematics*, 1991.

[19] D. Saxton and A. Thomason. Online containers for hypergraphs, with applications to linear equations. *Journal of Combinatorial Theory, Series B*, 121:248–283, 2016.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Sharpenings of the Erdős-Ko-Rado theorem

(Extended abstract)

Gyula O.H. Katona*

Abstract

Let $\mathcal{F}$ be a family of k-element subsets of an n-element set where $0 < k \leq n/2$ are integers. The theorem of Erdős, Ko and Rado claims that if $\mathcal{F}$ is intersecting then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. Let $\ell \geq 3$ be an integer and take the sum of the sizes of the pairwise intersections of an arbitrarily chosen subfamily of ℓ members of $\mathcal{F}$. If $\mathcal{F}$ is intersecting then this sum is at least $\binom{\ell}{2}$. Is this weaker condition sufficient to ensure that $|\mathcal{F}| \leq \binom{n-1}{k-1}$ still holds? We will show that even a smaller lower bound for the sum of the sizes of the pairwise intersection is sufficient.

Joint work with Kartal Nagy.

1 Introduction

Let $[n] = \{1, 2, \ldots, n\}$ be our underlying set. $\binom{[n]}{k}$ will denote the family of all k-element subsets of $[n]$. A family $\mathcal{F} \subset \binom{[n]}{k}$ is called intersecting if any pair of its members have a non-empty intersection. The celebrated theorem of Erdős, Ko and Rado [1] determines the maximum size of an intersecting family of k-element subsets.

Theorem 1 ([1] Erdős-Ko-Rado, 1961)**.** *If $2k \leq n, \mathcal{F} \subset \binom{[n]}{k}$ is intersecting then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}. \tag{1}$$

This estimate is sharp: choosing all k-element sets containing the element 1 gives this number of subsets.

Choose an integer $\ell \geq 2$ and take the following sum.

$$\sum_{1 \leq i < j \leq \ell} |F_i \cap F_j|. \tag{2}$$

*Rényi Institute, Hungarian Academy of Sciences,Budapest, Reáltanoda u. 13–15, 1053 Hungary. E-mail: ohkatona@renyi.hu. Research was supported by the Hungarian National Research, Development and Innovation Office NKFIH under grant numbers SSN135643 and K132696.

If $\mathcal{F}$ is intersecting then every term in (2) is at least 1, therefore (2) is at least $\binom{\ell}{2}$. Does this weaker condition

$$\binom{\ell}{2} \leq \sum_{1\leq i<j\leq \ell} |F_i \cap F_j|$$

imply (1)? Much more is true for large n.

Paper [1] contains a more general form of Theorem 1, but it holds only for large n. A family $\mathcal{F} \subset \binom{[n]}{k}$ is called t-intersecting $(1 \leq t < k)$ if $|F_1 \cap F_2| \geq t$ holds for every pair of members of $\mathcal{F}$.

Theorem 2 ([1] Erdős-Ko-Rado, 1961)**.** *If $\mathcal{F} \subset \binom{[n]}{k}$ is t-intersecting and n is large enough then*

$$|\mathcal{F}| \leq \binom{n-t}{k-t}. \tag{3}$$

We found an improvement of this statement, too. If the family is t-intersecting then (2) is at least $\binom{\ell}{2}t$. This value is also decreased .

2 Results

Theorem 3. *Let $2 \leq k, \ell$ be integers and suppose that $\mathcal{F} \subset \binom{[n]}{k}$. If*

$$\binom{\ell-1}{2} + 1 \leq \sum_{1\leq i<j\leq \ell} |F_i \cap F_j| \tag{4}$$

holds for any subfamily $\{F_1, F_2, \ldots F_\ell\}$ of distinct members of $\mathcal{F}$ and n is large enough then

$$|\mathcal{F}| \leq \binom{n-1}{k-1}.$$

Theorem 4. *Let $1 \leq t < k, 3 \leq \ell$ be integers and suppose that $\mathcal{F} \subset \binom{[n]}{k}$. If*

$$\binom{\ell}{2}t - 1 \leq \sum_{1\leq i<j\leq \ell} |F_i \cap F_j|$$

holds for any subfamily $\{F_1, F_2, \ldots F_\ell\}$ of distinct members of $\mathcal{F}$ and n is large enough then

$$|\mathcal{F}| \leq \binom{n-t}{k-t}.$$

3 An open problem

Let us try to decrease the left hand side of (4). The following construction shows that the bound $\binom{\ell-1}{2}$ is not sufficient to ensure the statement of Theorem 3. Let $\mathcal{F}$ be the family of all k-element subsets containing the element 1, and $\mathcal{G}$ be the family of those k-element subsets which do not contain 1, but contain the set $\{2, \ldots, \ell-1\}$ as a subset. It is easy to see that

$\mathcal{F} \cup \mathcal{G}$ satisfies the condition of Theorem 3 with $\binom{\ell-1}{2}$ on the left hand side of (4). Namely, if we choose x members of $\mathcal{F}$ and $\ell - x$ members of $\mathcal{G}$ then the sum (2) becomes

$$\binom{x}{3} + \binom{\ell - x}{2}(\ell - 2).$$

We need to prove that this is at least $\binom{\ell-1}{2}$. However, the equation

$$\binom{x}{2} + \binom{\ell - x}{2}(\ell - 2) = \binom{\ell - 1}{2}$$

has the roots $\ell - 1$ and $\ell - 2$, therefore the desired inequality holds for every integer $0 \leq x \leq \ell$. We believe that this is the best construction.

Conjecture 5. *Let $2 \leq k, 3 \leq \ell$ be integers and suppose that $\mathcal{F} \subset \binom{[n]}{k}$. If*

$$\binom{\ell - 1}{2} \leq \sum_{1 \leq i < j \leq \ell} |F_i \cap F_j|$$

holds for any subfamily $\{F_1, F_2, \ldots F_\ell\}$ of distinct members of $\mathcal{F}$ and n is large enough then

$$|\mathcal{F}| \leq \binom{n-1}{k-1} + \binom{n - \ell + 1}{k - \ell + 2}.$$

The case $\ell = 3$ gives back a special case of the Erdős Matching Conjecture for large n. The state of art of this conjecture can be found in [2].

References

[1] P. Erdős, Chao Ko and R. Rado, Intersection theorems for systems of finite sets, *Quarterly Journal of Mathematics, Second Series* (1961) **12** 313–320.

[2] Peter Frankl and Andrey Kupavskii The Erdős Matching Conjecture and concentration inequalities, *Journal of Combinatorial Theory, Series B* (2002) **157** 366-400.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Enumeration of 3-regular one-face maps on orientable or non-orientable surface up to all symmetries

(Extended abstract)

Alexander Omelchenko* Igor Labutin†

Abstract

We obtain explicit analitical formulas for enumerating 3-regular one-face maps on orientable and non-orientable surfaces of a given genus g up to all symmetries.

Keywords: map; surface; orbifold; enumeration; 3-regular maps; sensed maps; unsensed maps

1 Introduction

A *one-face* (or *unicellular*) topological *map* M on a compact surface X of genus g is a 2-cell embedding of a connected graph G (possibly with loops or multiple edges) such that the only connected component of $X - G$ is homeomorphic to a disk. Vertices and edges of M are the 0- and 1-dimensional cells of G [8]. Here, we consider both orientable and non-orientable surfaces without boundary. An orientable surface of genus g is a sphere with g handles, whereas a non-orientable surface of genus g is a sphere with g crosscaps (Möbius bands).

Two maps M_1 and M_2 on X are *isomorphic* if there is a homeomorphism of X that induces an isomorphism of the underlying graphs. Equivalence under such isomorphisms splits all maps on X into unlabelled classes. On an orientable surface X^+, homeomorphisms can be orientation-preserving or orientation-reversing. Hence, *sensed maps* are classified up to orientation-preserving homeomorphisms, whereas *unsensed maps* are classified up to all homeomorphisms (including orientation-reversing in the orientable case). On non-orientable surfaces X^-, all homeomorphisms are counted, so there is only one notion of unlabelled maps, often also termed unsensed.

*Constructor University Bremen, Campus Ring 1, 28759 Bremen, Germany. E-mail: aomelchenko@constructor.university
†Constructor University Bremen, Campus Ring 1, 28759 Bremen, Germany. E-mail: igor.labutin@gmail.com

The general problem of counting such maps by genus, edges, and various constraints has a rich history. Starting from planar enumeration by Tutte [14, 15], major advances have been made for rooted or sensed maps on higher-genus surfaces. Liskovets [11] pioneered enumeration of planar maps under symmetries. Later, Mednykh and Nedela [2] obtained formulas for sensed maps on orientable surfaces of genus g, using a technique based on quotient orbifolds. This orbit-counting perspective was further extended to hypermaps [3] and regular maps [5, 6, 9]. More recently, enumerating *unsensed* maps on surfaces of arbitrary genus attracted considerable attention: in [1], unsensed maps on orientable surfaces were fully enumerated, and in [10], a general framework was established for counting unsensed maps (including both orientable and non-orientable cases) by expressing these numbers in terms of quotient maps on suitable orbifolds.

Despite these developments, obtaining *closed-form* expressions for unsensed maps in specific regularities or face structures has remained challenging. Even on the torus, where [7] derived formulas for unsensed r-regular maps, the arguments become intricate. In the present paper, we focus on the special but important case of 3*-regular one-face maps*, i.e. cubic unicellular maps, and succeed in deriving explicit analytic formulas for these objects on both orientable and non-orientable surfaces of arbitrary genus g. To our knowledge, such formulas for unlabelled cubic one-face maps have not appeared before, although enumerations of related families (e.g. rooted cubic maps) can be found in classical works by Walsh and Lehman [13] and, more recently, Bernardi and Chapuy [4, 12].

Our approach leverages the general orbifold framework of [10] but capitalizes on a key simplification: for cubic one-face maps, the underlying quotient orbifolds turn out to be particularly straightforward (either closed with a few branch points, or a single boundary component possibly with index-2 branch points). This allows us to reduce the enumeration to counting *precubic* maps (with vertices of degrees 1 and 3) on simpler surfaces, for which exact formulas already exist. In this way, we derive explicit expressions for the total number of 3-regular one-face maps on any genus-g surface, counted up to all symmetries. We provide the final formulas, illustrate them with numerical tables, and confirm their correctness via an independent map-generation procedure [16].

Overall, our results offer a new instance where unsensed map enumeration, previously seen as intractable for higher-genus surfaces in closed form, can be completely resolved for a substantial class of maps. We hope this will stimulate further research on enumerating other classes of unsensed maps by combining orbifold methods with known enumerations of special map families.

2 Enumeration of unsensed 3-regular one-face maps on orientable surfaces

Let X_g^+ be a closed orientable surface of genus g. In [10], the following important formula for determining the numbers $\tilde{\tau}_{X_g^+}(n)$ of *unsensed* orientable maps with n edges was derived:

$$\bar{\tau}_{X_\chi^+}(n) = \frac{1}{2}\Big(\tilde{\tau}_{X_g^+}(n) + \frac{1}{2n}\sum_{\substack{m|2n\\ lm=n}}\ \sum_{O\in \mathrm{Orb}^-(X_\chi^+/\mathbb{Z}_{2l})} \tau_O(2m)\cdot \mathrm{Epi}_o^+(\pi_1(O),\mathbb{Z}_{2l})\Big). \tag{1}$$

Here, $\tau_O(2m)$ is the number of quotient maps with $2m$ flags on the orbifold O, and $\mathrm{Orb}^-(X_\chi^+/\mathbb{Z}_{2l})$ is the set of orbifolds O arising from orientation-reversing homeomorphisms of the surface X_χ^+.

Meanwhile,

$$\tilde{\tau}_{X_g^+}(n) = \frac{1}{2n} \sum_{\substack{l|2n \\ l\cdot m=2n}} \sum_{O\in \mathrm{Orb}(X_g^+/\mathbb{Z}_l)} \mathrm{Epi}_o(\pi_1(O),\mathbb{Z}_l)\cdot \tau_O(m) \tag{2}$$

is the number of sensed maps on X_g^+, where $O = X_g^+/\mathbb{Z}_l$ is a quotient of the surface X_g^+ under the action of a cyclic subgroup $\mathbb{Z}_l$ of the group of automorphisms of X_g^+, and $\mathrm{Epi}_o(\pi_1(O),\mathbb{Z}_l)$ and $\mathrm{Epi}_o^+(\pi_1(O),\mathbb{Z}_{2l})$ are some integer coefficients (the numbers of order-preserving epimorphisms from the fundamental group $\pi_1(O)$ of the orbifold O onto the cyclic group $\mathbb{Z}_l$). The term $\tau_O(m)$ denotes the number of rooted quotient maps with m darts on the orbifold O.

Using the technique described in detail in [5], one can obtain the following formula for counting 3-regular one-face maps on an orientable surface X_g^+ of genus g (see formula (20) in [6]):

$$\begin{aligned}\tilde{\tau}^{(3)}(g) = {} & \frac{\tau_+^{(3)}(g)}{2\,(6g-3)} + \sum_{\mathfrak{g}=0}^{\lfloor g/2\rfloor} \frac{(4g-2-2\mathfrak{g})!}{2\cdot 3^{\mathfrak{g}}\,\mathfrak{g}!\,(2g-1-\mathfrak{g})!\,(2g-4\mathfrak{g}+1)!} \\ & + \frac{(2g-2)!}{6\cdot(g-1)!} \sum_{\mathfrak{g}=0}^{\lfloor (g+1)/3\rfloor} \left(\frac{3}{4}\right)^{\mathfrak{g}-1} \frac{2^{g+1-3\mathfrak{g}} + (-1)^{g-\mathfrak{g}}}{\mathfrak{g}!\,(g+1-3\mathfrak{g})!} \\ & + \sum_{k=\lfloor g/2\rfloor}^{\lfloor (2g-2)/3\rfloor} \sum_{\mathfrak{g}=0}^{k-\lfloor g/2\rfloor} \frac{3^{\mathfrak{g}-2}\left[2^{2g-1-3k} + (-1)^k\right](2k-2\mathfrak{g})!}{\mathfrak{g}!\,(k-\mathfrak{g})!\,(4k+3-2g-4\mathfrak{g})!\,(2g-1-3k)!}.\end{aligned} \tag{3}$$

Here, $\tau_+^{(3)}(g)$ is the number of *rooted* 3-regular one-face maps on the orientable surface X_g^+, equal to

$$\tau_+^{(3)}(g) = \frac{2\,(6g-3)!}{12^g\,g!\,(3g-2)!}. \tag{4}$$

In order to count maps by formula (1), it remains to enumerate maps on orbifolds for orientation-reversing homeomorphisms h. It turns out that in this case the orbifolds admit a simple description, allowing one to derive a concise result:

Theorem 1. *The numbers $\bar{\tau}_+^{(3)}(g)$ of unsensed 3-regular one-face maps on orientable surfaces X_g^+ are equal to*

$$\bar{\tau}_+^{(3)}(g) = \frac{1}{2}\left(\tilde{\tau}_+^{(3)}(g) + \tau_+^{(3)}(g/2) + \tau_-^{(3)}(g)\right), \tag{5}$$

where $\tilde{\tau}_+^{(3)}(g)$ is the number of sensed 3-regular one-face maps on an orientable surface X_g^+ calculated by the formula (3), $\tau_+^{(3)}(g/2)$ is the number of rooted 3-regular one-face maps on an orientable surface $X_{g/2}^+$ calculated by the formula (4) in the case of even g and equal to 0 in the case of odd g, and $\tau_-^{(3)}(g)$ is the number of rooted 3-regular one-face maps on a non-orientable surface X_g^- calculated by the formula (see [12])

$$\tau_-^{(3)}(g) = \begin{cases} \dfrac{2^{2h-2}\,h!\,(6h-2)!}{3^{h-1}\,(2h)!\,(3h-1)!} \displaystyle\sum_{i=0}^{h-1} \binom{2i}{i} 16^{-i}, & h = g/2,\ g \text{ is even}, \\ \dfrac{2^{6h}\,(3h)!}{3^h\,h!}, & h = (g-1)/2,\ g \text{ is odd}. \end{cases} \tag{6}$$

g	$\tau_+^{(3)}(g)$	$\tilde{\tau}_+^{(3)}(g)$	$\bar{\tau}_+^{(3)}(g)$
1	1	1	1
2	105	9	8
3	50050	1726	927
4	56581525	1349005	676445
5	117123756750	2169056374	1084610107
6	386078943500250	5849686966988	2924847922929
7	1857039718236202500	23808202021448662	11904101304325611
8	12277353837189093778125	136415042681045401661	68207521363461659373
9	106815706684397824557193750	1047212810636411989605202	523606405320272947813801
10	1183197582943074702620035168750	10378926166167927379808819918	5189463083084174721816125584

Table 1: The numbers of 3-regular one-face maps on orientable surfaces X_g^+

3 Enumeration of unsensed 3-regular one-face maps on non-orientable surfaces

In the case of non-orientable surfaces, the corresponding formula for the number $\bar{\tau}_{X_\chi^-}(n)$ of unsensed maps on a non-orientable surface X_χ^- of Euler characteristic χ is [10]

$$\bar{\tau}_{X_\chi^-}(n) = \frac{1}{4n} \sum_{\substack{m|2n \\ lm=2n}} \sum_{O \in \mathrm{Orb}(X_\chi^-/\mathbb{Z}_l)} \tau_O(2m) \cdot \big(\mathrm{Epi}_o(\pi_1(O), \mathbb{Z}_l) - \mathrm{Epi}_o^+(\pi_1(O), \mathbb{Z}_l)\big). \tag{7}$$

Here, $\mathrm{Orb}(X_\chi^-/\mathbb{Z}_l)$ is a set of orbifolds arising from homeomorphisms of X_χ^-. An approach similar to the one described above allows us to derive an exact formula in the non-orientable case. Because of its length, we omit the explicit form. However, the final result entirely parallels the orientable case, exploiting orbifolds with boundary or with certain branch indices.

The results presented in this article allowed us to enumerate unsensed 3-regular one-face maps on orientable and non-orientable surfaces of a given genus g. In Table 1, we provide the results for rooted, sensed, and unsensed maps on orientable surfaces of genus $g \in [1, 10]$. In Table 2, we provide the results for rooted and unsensed maps on non-orientable surfaces of genus $g \in [2, 20]$.

To verify the obtained analytical results, we also implemented an algorithm for generating maps on orientable or non-orientable surfaces, based on the ideas formulated in [16]. The numerical results obtained by generating such maps coincided with the first terms given by the analytical formulas obtained in this paper.

g	$\tau_-^{(3)}(g)$	$\bar{\tau}_-^{(3)}(g)$
2	6	2
3	128	11
4	3780	144
5	163840	3627
6	8828820	149288
7	587202560	8170800
8	45821335560	545671762
9	4133906022400	43063046307
10	421946699674500	3906934079662
11	48151737348915200	401264673924438
12	6070544859205827000	45988979036528440
13	838225443769915801600	5821010056777072838
14	125787689149526729325000	806331341176441101980
15	20385642792484352294912000	121343111865634574938768
16	3548258423062128985899690000	19712546794881999409462482
17	660168656191813264718430208000	3438378417666873290074260643
18	130746565669943973430227429382500	640914537597785062325259175158
19	27463016097579431812286696652800000	127143593044349500804170430994988
20	6098023559259606741021710317037175000	26745717365173718867249062116990380

Table 2: The numbers of 3-regular one-face maps on non-orientable surfaces X_g^-

References

[1] A. Mednykh A. Breda d'Azevedo and R. Nedela. Enumeration of maps regardless of genus: Geometric approach. *Discrete Mathematics, 310:1184–1203*, 2010.

[2] R. Nedela A. Mednykh. Enumeration of unrooted maps of a given genus. *J. Combin. Theory Ser. B, 96(5):709–729*, 2006.

[3] R. Nedela A. Mednykh. Enumeration of unrooted hypermaps of a given genus. *Discrete Mathematics, 310:518–526*, 2010.

[4] G. Chapuy. A new combinatorial identity for unicellular maps, via a direct bijective approach. *Advances in Applied Mathematics, Volume 47, Issue 4, pp. 874–893, 2011*, 2011.

[5] A. Omelchenko E. Krasko. Enumeration of 4-regular one-face maps. *European Journal of Combinatorics, 62(5):167–177*, 2017.

[6] A. Omelchenko E. Krasko. Enumeration of regular maps on surfaces of a given genus. *Journal of Mathematical Sciences, 232(1):44–60*, 2018.

[7] A. Omelchenko E. Krasko. Enumeration of r-regular maps on the torus. part ii: Unsensed maps. *Discrete Mathematics. Vol. 342. No. 2. P. 600–614*, 2019.

[8] Ping Zhang Jonathan L. Gross, Jay Yellen. *Handbook of Graph Theory. 2nd Edition.* Chapman and Hall/CRC, 2013.

[9] E. Krasko and A. Omelchenko. Enumeration of r-regular maps on the torus. part i: Rooted maps on the torus, the projective plane and the klein bottle. sensed maps on the torus. *Discrete Mathematics*, 342(2):584–599, 2019.

[10] E. Krasko and A. Omelchenko. Enumeration of unsensed orientable and non-orientable maps. *European Journal of Combinatorics, 86:103093*, 2020.

[11] V. Liskovets. Enumeration of nonisomorphic planar maps. *Selecta Math. Sovietica, 4:303–323*, 1985.

[12] G. Chapuy O. Bernardi. Counting unicellular maps on non-orientable surfaces. *Advances in Applied Mathematics, 48(2):259-275*, 2011.

[13] A. Lehman T. Walsh. Counting rooted maps by genus, i. *J. Combin. Theory Ser. B, 13:192–218*, 1972.

[14] W. Tutte. A census of planar triangulations. *Canad. J. Math., 14:21–38*, 1962.

[15] W. Tutte. A census of planar maps. *Canad. J. Math., 15:249–271*, 1963.

[16] T. Walsh. Space-efficient generation of nonisomorphic maps and hypermaps. *Journal of Integer Sequences, 18*, 2015.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Lattice Path Interpretations of the Pearson Distribution Systems

(Extended abstract)

Alexander Omelchenko*

Abstract

We propose a combinatorial framework that models the seven classical Pearson distribution types via weighted Motzkin paths with linearly dependent step weights. The generating functions of these path ensembles satisfy Pearson-type differential equations, and their asymptotic behavior matches the corresponding continuous distributions. This provides a novel discrete interpretation of Pearson's system and reveals a new connection between enumerative combinatorics and classical probability.

Keywords: Pearson distributions; Motzkin paths; generating functions; weighted lattice paths; asymptotic analysis; enumerative combinatorics; discrete models of continuous distributions.

The Pearson distribution system is a family of seven classical continuous probability distributions (Types I–VII) introduced by Karl Pearson in 1895 [3, 8, 10]. This system was developed to model data with adjustable skewness and kurtosis, and it includes many well-known distribution families (for example, Type I includes the beta distributions on a bounded interval, Type III includes the gamma/chi-square distributions on a half-line, and Type VII includes Student's t-distributions on the whole real line). Each Pearson type is characterized by a differential equation (the Pearson equation) satisfied by its probability density function. In particular, the density $f(x)$ of a Pearson distribution satisfies a first-order linear differential equation of the form

$$\frac{f'_x}{f} = -\frac{\alpha_0 x + \gamma_0}{ax^2 + bx + c}, \tag{1}$$

for some constants $\alpha_0, \gamma_0, a, b, c$ (with $a = 0$ in some cases). This equation encodes how the shape of $f(x)$ changes across its support and underpins Pearson's classification. It is a classical tool in propability and statistics, but so far it has had little presence in enumerative combinatorics. The motivation of this work is to bridge this gap: we ask whether Pearson's continuous distributions can be understood as limits of natural combinatorial models, thereby linking classical distribution theory with lattice path enumeration.

*Constructor University Bremen, Campus Ring 1, 28759 Bremen, Germany. E-mail: aomelchenko@constructor.university.

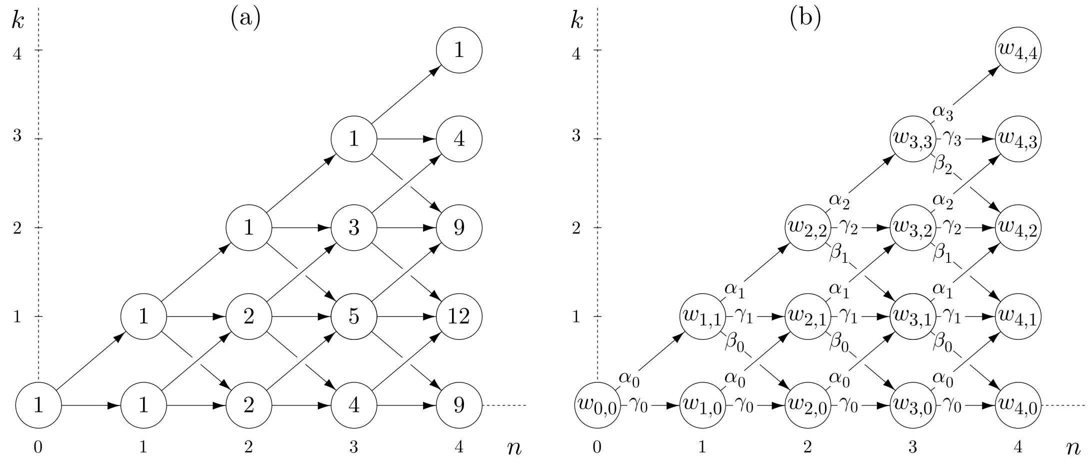


Figure 1: The classical Motzkin triangle (a), a triangle with multiplicities (b)

Our paper establishes a novel connection between the Pearson distribution system and weighted lattice paths, specifically Motzkin paths with linearly varying step weights. A Motzkin path is a lattice path consisting of up-steps $(1,1)$, down-steps $(1,-1)$, and level steps $(1,0)$ that never fall below the horizontal axis. A Motzkin triangle (1(a)) enumerates paths in the positive quadrant of the plane (t,x) issuing from the origin and consisting of the vectors $(1,1)$, $(1,0)$, and $(1,-1)$. Paths that end on the x-axis are exactly Motzkin paths [9].

Let us slightly modify the Motzkin triangles by assigning positive integers α_k, β_k, and γ_k to the arrows (1(b)). The number assigned to the arrow will be interpreted as its multiplicity, i.e., as the number of various arrows going in the given direction. Triangles of this kind are called *Motzkin triangles with multiplicities*. It is evident that the conventional Motzkin triangle provides special case of the corresponding triangles with multiplicities, with $\alpha_k = \beta_k = \gamma_k = 1$.

To determine the number of weighted paths in a Motzkin triangle, we introduce the generating function

$$w(x,t) = \sum_{n=0}^{\infty} P_n(x)\frac{t^n}{n!}, \tag{2}$$

where the coefficients of the polynomial

$$P_n(x) = \sum_{k=0}^{n} w_{n,k}\, x^k$$

give the weight of the path from the origin to the point (n,k). The numbers $w_{n,k}$ satisfy the recurrence relation

$$w_{n+1,k} = \alpha_{k-1}w_{n,k-1} + \gamma_k w_{n,k} + \beta_k w_{n,k+1}, \tag{3}$$

which leads to the following recurrence relation for the polynomials $P_n(x)$:

$$P_{n+1}(x) = (ax^2+bx+c)P_n'(x) + \left(\alpha_0 x + \gamma_0 + \frac{\beta_0 - c}{x}\right)P_n(x) \tag{4}$$

In order for $P_n(x)$ to remain polynomials, the equality $\beta_0 = c$ must hold. As a result, for the generating function $w(x,t)$, we obtain the following Cauchy problem for a first-order partial differential equation:

$$\frac{\partial w(x,t)}{\partial t} = (ax^2 + bx + c)\frac{\partial w(x,t)}{\partial x} + (\alpha_0 x + \gamma_0)\, w(x,t), \quad w(x,t)|_{t=0} = 1. \tag{5}$$

Its solution reduces to solving a system of ordinary differential equations of the form:

$$\frac{dw}{w(\alpha_0 x + \gamma_0)} = -\frac{dx}{ax^2 + bx + c} = dt \tag{6}$$

The first of these equations, rewritten as

$$\frac{w'_x}{w} = -\frac{\alpha_0 x + \gamma_0}{ax^2 + bx + c}, \tag{7}$$

is known as the *Pearson equation.* The defining parameter here is the so-called "kappa-Pearson":

$$\kappa = \frac{b^2}{4ac}$$

This remarkable fact—that the generating function of weighted Motzkin paths satisfies the same differential equation as a Pearson density—was first observed in a special case in [7]. In that work, asymptotic formulas for the number of paths $w_{n,k}$ were derived, suggesting a deep connection between discrete path counts and classical Pearson densities. However, a comprehensive combinatorial interpretation of all seven Pearson types was not given in that initial study. Our approach generalizes and systematizes this connection, providing a unified framework to interpret each Pearson distribution type as a discrete limit of a suitable weighted lattice path ensemble.

In this framework, a single parameter κ governs the linear step weights, and by tuning κ we recover the different regimes corresponding to Pearson's Types I–VII. Combinatorially, varying κ changes the dominant behavior of the path-counting distribution $w_{n,k}$, much like altering the coefficients in Pearson's differential equation yields different families of continuous densities. For example, one choice of κ produces path counts whose normalized distribution approaches a Beta distribution (Type I) in the limit of large n, while another choice yields counts approaching a Gamma distribution (Type III), and yet another yields heavy-tailed counts analogous to Student's t (Type VII). More precisely, each Pearson type corresponds to a distinct "phase" determined by the sign or factorization of the quadratic $ax^2 + bx + c$. The crucial insight is that the distribution given by $w_{n,k}$ (after suitable normalization and scaling) converges to the corresponding Pearson law, thereby providing a new combinatorial interpretation of these classical families.

We illustrate our approach with several explicit examples of weighted Motzkin path models, deriving closed-form generating functions and outlining how their coefficient asymptotics align with known continuous limits. Techniques such as singularity analysis and the saddle-point method [1, 2, 4, 5, 6] confirm that, in each regime, the discrete path ensemble behaves like the corresponding Pearson density in the large-n limit. Full generality and rigorous proofs for all seven types are deferred to a longer version of the paper; here, we focus on presenting the overarching framework and highlighting instructive examples. By doing so, we build a bridge between enumerative combinatorics and classical distribution theory: the Pearson families, historically a purely analytic construct, gain a foothold in lattice path enumeration, and the rich methods of combinatorial generating functions can be applied to classical probability in novel ways.

References

[1] Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theoretical Computer Science*, 281(1–2):37–80, 2002.

[2] Edward A. Bender. Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A*, 15(1):91–111, 1973.

[3] Harald Cramér. *Mathematical Methods of Statistics.* Princeton University Press, Princeton, NJ, 1966.

[4] Philippe Flajolet. Combinatorial aspects of continued fractions. *Discrete Mathematics*, 32(2):125–161, 1980.

[5] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge University Press, Cambridge, 2009.

[6] W. K. Hayman. A generalisation of Stirling's formula. *Journal für die reine und angewandte Mathematik*, 196:67–95, 1956.

[7] V. R. Meshkov, A. V. Omelchenko, M. I. Petrov, and E. A. Tropp. Dyck and motzkin triangles with multiplicities. *Moscow Mathematical Journal*, 10(3):611–628, 2010.

[8] Karl Pearson. Contributions to the mathematical theory of evolution. ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London. Series A*, 186:343–414, 1895.

[9] Richard P. Stanley. *Enumerative Combinatorics, Vol. 2.* Cambridge University Press, Cambridge, 2001.

[10] B. L. Van der Waerden. *Mathematische Statistik.* Springer-Verlag, Berlin, 1957.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On the isomorphism problem of Cayley digraphs

(Extended abstract)

Péter P. Pálfy*

Abstract

A finite group G is called a DCI-group provided any two Cayley digraphs $Cay(G, S)$ and $Cay(G, T)$ are isomorphic iff the two connection sets S and T are images of each other via automorphisms of the group G. Building on the vast literature on CI- and DCI-groups we present an updated list of possible DCI-groups. We also point out the most important open problems related to the classification of DCI-groups.

1 Results and open problems

If α is an automorphism of the finite group G and $S \subset G$, then the Cayley digraphs $Cay(G, S)$ and $Cay(G, S^\alpha)$ are obviously isomorphic. If the converse is also true, i.e., if for any two Cayley digraphs of the group G, $Cay(G, S)$ and $Cay(G, T)$ can be isomorphic only when T is the image of S under some group automorphism, then we say that G is a DCI-group. If we require this property only for undirected graphs (that is, whenever $S^{-1} = S$ and $T^{-1} = T$) then G is called a CI-group. The definition goes back to László Babai [2], who was inspired by a research problem published by András Ádám [1] in 1967. Determining which groups are CI- or DCI-groups is still an open problem, despite the vast literature devoted to studying Cayley (di)graphs from this point of view. Some recent progress, especially a breakthrough result of Ted Dobson, Mikhail Muzychuk, and Pablo Spiga [6] makes it possible to restrict further the class of candidates satisfying these elusive properties. In [6] a list of putative CI-groups is given. However, the subtle differences between CI-groups and DCI-groups make it more promising to deal with DCI-groups first. For example, there exist exactly three cyclic CI-groups that are not DCI-groups, as it is established by the most prominent result in this area of research, the determination of cyclic CI- or DCI-groups due to Mikhail Muzychuk [13], that gives the solution to Ádám's original problem.

Theorem 1. *(Muzychuk [13]) A cyclic group of order n is a DCI-group iff $n = 2^k m$, where $k \in \{0, 1, 2\}$ and m is an odd square-free integer. In addition, the cyclic groups of order 8, 9, and 18 are CI-groups (but not DCI-groups).*

*HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: ppp@renyi.hu. Supported by the National Research, Development and Innovation Office of Hungary (NKFIH), grant no. 138596.

If the following conjecture is true, then it will be enough to consider coprime indecomposable groups, i.e., such groups that cannot be written as a direct product of two proper subgroups of coprime orders.

Conjecture 1. *(Kovács and Muzychuk [8]; Dobson [4, Conjecture 43]) If G_1 and G_2 are DCI-groups, such that $(|G_1|, |G_2|) = 1$, then $G_1 \times G_2$ is a DCI-group as well.*

This is clearly not true for CI-groups, as the examples of $C_8 \times C_3$ and $C_9 \times C_4$ show. There are several cases when the validity of this conjecture has been verified: $C_p^4 \times C_q$ ($p \neq q$ primes — Kovács and Ryabov [10]), $C_p^2 \times C_m$ and $C_p^2 \times C_q^2$ ($p \neq q$ primes, m square-free — Kovács et al. [9]), $Q_8 \times C_p$ (p prime — Somlai [16]).

The main result of the present paper shows that the structure of coprime indecomposable DCI-groups is severely restricted.

Theorem 2. *If G is a coprime indecomposable DCI-group, then one of the following occurs:*

(a) $G = C_p^r$ *is an elementary abelian p-group for some prime p and $r \geq 1$;*

(b1) $G = D_{2m} = \langle a, b \mid a^m = 1, b^2 = 1, b^{-1}ab = a^{-1} \rangle$ *is a dihedral group of order $2m$, where m is an odd square-free integer;*

(b2) $G = \hat{D}_{4m} = \langle a, b \mid a^m = 1, b^4 = 1, b^{-1}ab = a^{-1} \rangle$ *of order $4m$ is a central extension of a dihedral group, where m is an odd square-free integer;*

(c) *G is one of the groups C_4, Q_8 (the quaternion group of order 8), $Alt(4)$ (the alternating group of degree 4 and order 12).*

In Section 2 we will sketch the proof of this result, here we just comment on the various cases in the theorem.

Let $G = C_p^r$ be an elementary abelian p-group of rank r. In a major work Yang-Quan Feng and István Kovács [7] proved that if $r \leq 5$ then C_p^r is a DCI-group. For $p = 2$ this is best possible, as C_2^6 is not a CI-group, see Lewis Nowitz [14]. For $p = 3$ Pablo Spiga [17] gave an example showing that C_3^8 is not a CI-group, while the cases $r = 6$ and $r = 7$ are still open. For arbitrary primes Gábor Somlai [15] gave a very smart construction yielding that C_p^{2p+3} is not a CI-group.

Problem 1. *For a given prime p, what is the largest r such that C_p^r is a DCI-group? Is it bounded for all primes p or the rank increases with p?*

The groups D_{2m} and $\hat{D}_{4m}$ in cases (b1) and (b2) of Theorem 2 have the same order as the cyclic groups in Muzychuk's celebrated result (Theorem 1), and contain a cyclic subgroup of index two. So it seems plausible that these groups are DCI-groups indeed, although it is known only when m is an odd prime number [12]. We hope that the following conjecture holds true.

Conjecture 2. *For every odd square-free number m, the dihedral group D_{2m} and its central extension $\hat{D}_{4m}$ are DCI-groups.*

Finally, note that the groups C_4, Q_8 and $Alt(4)$ are indeed DCI-groups [3].

2 Sketch of proof of Theorem 2

We start by summarizing some well-known arguments. It is easy to see that subgroups of DCI-groups are DCI-groups as well. Since the order of an element in a DCI-group cannot be divisible by 8 or by a square of an odd prime, the exponent of a Sylow p-subgroup in a DCI-group is either p or 4. An obvious necessary condition for the DCI property of G is that any two subgroups of the same order must be images of each other via automorphisms of G. Applying this criterion to the Sylow p-subgroup, one obtains that it is either elementary abelian, or C_4, or Q_8.

Now we consider groups whose order has two distinct prime divisors. Among these groups a complete classification of coprime indecomposable DCI-groups is available.

Lemma 1. *Let G be a finite group such that its order has exactly two distinct prime divisors, and suppose that G is coprime indecomposable (i.e., not nilpotent). Then G is a DCI-group iff G is one of the groups D_{2p}, $\hat{D}_{4p}$ (with an odd prime p), or $Alt(4)$.*

Proof. These groups are indeed DCI-groups: for D_{2p} see [2]; for $\hat{D}_{4p}$ see [12] (the proof given there for the CI property also works for the DCI property); for $Alt(4)$ see [3]. Conversely, checking the list of possible CI-groups given in [12], it remains to show that the following groups are not DCI-groups: (1) semidirect products $M \rtimes \langle z \rangle$, where M is an elementary abelian p-group of order at least p^2, z has order 2 or 4, and z acts by inverting elements of M; (2) semidirect products $M \rtimes \langle z \rangle$, where M is an elementary abelian p-group, z has order 3, and z acts nontrivially on M by raising every element to the same power. In case (1) the recent discovery by Dobson, Muzychuk and Spiga [6] shows that these groups are not DCI-groups. Case (2) was settled in [5]. $\square$

Turning to the general case let G be a coprime indecomposable DCI-group, and let the prime divisors of $|G|$ be $p_1, \ldots, p_k$, where $k \geq 2$. By a crucial result of Cai Heng Li [11] every DCI-group is solvable. For solvable groups Philip Hall's classical theorem yields a system of Sylow subgroups $P_1, \ldots, P_k$ such that for each prime divisor p_i ($i = 1, \ldots, k$) a Sylow p_i-subgroup is chosen and $P_iP_j = P_jP_i$ holds for each pair $i \neq j$. Let us draw a graph Γ with vertices $p_1, \ldots, p_k$, where $\{p_i, p_j\}$ is an edge if the subgroup P_iP_j is not nilpotent. This graph is connected because G is coprime indecomposable. Moreover, Lemma 1 gives that the prime 2 is incident to every edge, hence Γ is a star. Note that in the three cases of Lemma 1 the Sylow 2-subgroups are different: it is C_2 in D_{2p}; C_4 in $\hat{D}_{4p}$; and C_2^2 in $Alt(4)$. So if the Sylow 2-subgroup of G is $C_2 = \langle z \rangle$, then for every odd prime divisor p of $|G|$ the Sylow p-subgroup has order p and is inverted by conjugation with z. Since there is no edge between odd primes, the Sylow subgroups corresponding to the odd primes commute with each other, and their direct product is a cyclic normal subgroup of odd square-free order, inverted by z. Thus we obtain $G = D_{2m}$ as described in case (2a) of Theorem 2. Similarly, if the Sylow 2-subgroup of G is C_4, then we obtain $G = \hat{D}_{4m}$ as in case (2b). Finally, if the Sylow 2-subgroup of G is C_2^2, then Γ has just one edge, and $G = Alt(4)$.

References

[1] A. Ádám, Research problem 2-20, J. Combin. Theory 2 (1967), 309.

[2] L. Babai, Isomorphism problem for a class of point-symmetric structures, Acta Math. Acad. Sci. Hungar. 29 (1977), 329–336.

[3] M. Conder and C. H. Li, On isomorphisms for finite Cayley graphs, European J. Combin. 19 (1998), 911–919.

[4] T. Dobson, On the isomorphism problem for Cayley graphs of alelian groups whose Sylow subgroups are elementary abelian or cyclic, Electron. J. Comb. 25 (2018), #P2.49

[5] T. Dobson, J. Morris, P. Spiga, Further restrictions on the structure of finite DCI-groups: an addendum, J. Algebraic Combin. 42 (2015), 959–969.

[6] T. Dobson, M. Muzychuk, P. Spiga, Generalised dihedral CI-groups, Ars Math. Contemp. 22 (2022), paper 2.07.

[7] Y.-Q. Feng, I. Kovács, Elemetary abelian groups of rank 5 are DCI-groups, J. Combin. Theory, Ser. A 157 (2018), 162–204.

[8] I. Kovács, M. Muzychuk, The group $\mathbf{Z}_p^2 \times \mathbf{Z}_q$ is a CI-group, Comm. Algebra 37 (2009), 3500–3515.

[9] I. Kovács, M. Muzychuk, P. P. Pálfy, G. Ryabov, G. Somlai, CI-property of $C_p^2 \times C_n$ and $C_p^2 \times C_q^2$ for digraphs, J. Combin. Theory, Ser. A 196 (2023), article 105738.

[10] I. Kovács, G. Ryabov, The group $C_p^4 \times C_q$ is a DCI-group, Discrete Math. 345 (2022), article 112705.

[11] C. H. Li, Finite CI-groups are soluble, Bull. London Math. Soc. 31 (1999), 419–423.

[12] C. H. Li, Z. P. Lu, P. P. Pálfy, Further restrictions on the structure of finite CI-groups, J. Algebraic Combin. 26 (2007), 161–181.

[13] M. Muzychuk, On Ádám's conjecture for circulant graphs, Discrete Math. 176 (1997), 285–298.

[14] L. A. Nowitz, A non-Cayley-invariant Cayley graph of the elementary Abelian group of order 64, Discrete Math. 110 (1992), 223–228.

[15] G. Somlai, Elementary abelian p-groups of rank $2p+3$ are not CI-groups, J. Algebraic Combin. 34 (2011), 323–335.

[16] G. Somlai, The Cayley isomorphism property for groups of order $8p$, Ars Math. Contemp. 8 (2015), 433–444.

[17] P. Spiga, CI-property of elementary abelian 3-groups, Discrete Math. 309 (2009), 3393–3398.

Stabbing non-piercing sets and face lengths in large girth plane graphs

(Extended abstract)*

Dömötör Pálvölgyi † Kristóf Zólomy ‡

Abstract

We show that a non-piercing family of connected planar sets with bounded independence number can be stabbed with a constant number of points. As a consequence, we answer a question of Axenovich, Kießle and Sagdeev about the largest possible face length of an edge-maximal plane graph with girth at least g.

1 Introduction

Define a region as a connected planar compact set whose boundary consists of a finite number of disjoint Jordan curves. One of these curves is the outer boundary of the region, while the rest cut out "holes". A family $\mathcal{F}$ of regions is in general position if for any two regions from $\mathcal{F}$ their boundaries intersect in finitely many points.[1] Such a family $\mathcal{F}$ is *non-piercing* if $F \setminus G$ is connected for any two regions $F, G \in \mathcal{F}$. For example, the family of all disks is non-piercing and, more generally, so is a pseudo-disk family, defined as a family of simply connected regions whose boundaries intersect pairwise at most twice; these include families formed by homothetic[2] copies of a fixed convex set. However, a family of axis-parallel rectangles in general position is not necessarily non-piercing, as two rectangles can cross each other without any of them containing a vertex of the other.

For a family $\mathcal{F}$, the independence number $\nu(\mathcal{F})$ is the size of the largest subfamily of pairwise disjoint sets, that is, the smallest number such that among any $\nu(\mathcal{F})+1$ sets there are two that intersect. The piercing number $\tau(\mathcal{F})$ is the least number of points that pierce $\mathcal{F}$, that is, the size of the smallest point set that intersects every set from $\mathcal{F}$. Our main result is the following.

*The full version of the paper can be found at `https://arxiv.org/abs/2504.10618`.

†ELTE Eötvös Loránd University and HUN-REN Alfréd Rényi Institute of Mathematics, Budapest. E-mail: `domotor.palvolgyi@ttk.elte.hu`.

‡ELTE Eötvös Loránd University, Budapest. E-mail: `zkr@student.elte.hu`.

[1]This is a technical condition that was introduced in [13] and makes many of the arguments simpler, though most often could be omitted. We need to assume it as we will use a result from [13] which was proved under this assumption.

[2]A homothetic copy is a copy that is translated and scaled by a positive scalar factor.

Theorem 1. *There is a function f such that if $\mathcal{F}$ is a family of non-piercing regions, then $\tau(\mathcal{F}) \leq f(\nu(\mathcal{F}))$.*

Note that this implies that the disjointness graph of a family of non-piercing regions is χ-bounded.[3] For results similar to Theorem 1 about homothetic copies of a fixed convex set with explicit bounds, see [9].

In particular, if $\mathcal{F}$ is a family of pairwise intersecting, non-piercing regions, then there exists an absolute constant T such that $\mathcal{F}$ can be pierced with at most T points. This result was proved earlier for pseudo-disk families [1] with different methods, using sweepings, which do not generalize to non-piercing families. Instead, our proof, which can be found in Section 2, uses the standard machinery developed to prove (p, q)-theorems. We do not know the best possible value for T. In case of pairwise intersecting disks, this is a well-studied problem where we know that the optimal value is four, for which by now there are several different proofs, see [7, 15, 5]. It is entirely possible that the answer in case of pairwise intersecting non-piercing regions is also four but at the moment this is not known even for pseudo-disks.

As an application of Theorem 1, we answer a recent question of Axenovich, Kießle and Sagdeev [3] about the longest possible face length of an edge-maximal plane graph with girth at least g, which was our main motivation.[4]

Theorem 2. *Suppose that G is a plane graph with girth at least g, and that G is edge-maximal with regards to these two properties. Then the length of any facial cycle of G is at most Kg for some absolute constant K.*

The exact statement of the problem and the proof can be found in Section 3.

2 Proof of Theorem 1

In this section we present the proof of Theorem 1. We start with some definitions, related to (p, q)-theorems; for a complete survey of such results, see [11].

Following Hadwiger and Debrunner [10], we say that a family $\mathcal{G}$ has the (p, q)-property, if for every subfamily $\mathcal{G} \subset \mathcal{F}$ with $|\mathcal{G}| = p$, there exists a subsubfamily $\mathcal{H} \subset \mathcal{G}$ of size q with a non-empty intersection $\cap\mathcal{H} \neq \emptyset$. In other words, from every p sets from $\mathcal{F}$, some q intersect. It was shown by Alon and Kleitman [2] that if a family $\mathcal{F}$ of compact convex sets in $\mathbb{R}^d$ satisfies the (p, q)-property for any $p \geq q \geq d + 1$, then $\mathcal{F}$ can be pierced with $T(p, q)$ points. Later, this result was extended from convex sets to several other families. For the version that we need, we need to define the Vapnik-Chervonenkis dimension. The VC-dimension of a family $\mathcal{F}$ is the largest d for which exists a set of d elements, X, such that for every subset $Y \subset X$ there exists a set $F \subset \mathcal{F}$ such that $X \cap F = Y$. The dual

[3] For the definition and a survey of χ-boundedness, see [14].

[4] Note that since the submission of our Extended Abstract to EuroComb, we learned that they have also proved this result, with very different methods, and that their result will also appear in the proceedings of EuroComb.

VC-dimension d^* of $\mathcal{F}$ is the VC-dimension of the dual family $\mathcal{F}^*$, in which the roles of elements and sets are swapped, with the containment relation reversed. It is well-known and easy to see that $d^* \leq 2^d$. Matoušek [12] showed that bounded VC-dimension and an appropriate (p, q)-property imply the existence of a small hitting set.

Theorem 3 (Matoušek [12]). *If the dual VC-dimension of $\mathcal{F}$ is at most $q-1$, and $\mathcal{F}$ satisfies the (p, q)-property for some $p \geq q$, then the sets of $\mathcal{F}$ can be hit with at most T points, where T is a constant depending on p and q.*

Therefore, in order to prove Theorem 1, it would be sufficient to show that non-piercing regions have bounded (dual) VC-dimension, and if a family $\mathcal{F}$ of non-piercing regions has bounded independence number $\nu(\mathcal{F})$, then they also satisfy the (p, q)-property for some large enough q.

Lemma 4. *If $\mathcal{F}$ is a non-piercing family of regions, then the VC-dimension and the dual VC-dimension of $\mathcal{F}$ are at most 4.*

Due to space constraints, we omit the proof.

Now we only need to show that a (p, q)-property holds for some $p \geq q \geq 5$ in every family $\mathcal{F}$ of non-piercing regions with bounded independence number $\nu(\mathcal{F})$. We first make some definitions.

For a family $\mathcal{F}$ and collection of elements P, define the dual intersection hypergraph $\mathcal{H}(\mathcal{F}, P)$ such that the vertices correspond to members of $\mathcal{F}$, while hyperedges correspond to elements of P such that for every $p \in P$ the vertex set $H_p = \{F \in \mathcal{F} : p \in F\}$ forms a hyperedge. The Delaunay graph $\mathcal{D}(\mathcal{F})$ is the subgraph of $\mathcal{H}(\mathcal{F}, P)$ that contains only the hyperedges with exactly two vertices, that is, a pair of vertices corresponding to the sets $F, G \in \mathcal{F}$ are connected by an edge if there is an element $p \in F \cap G$ which is not contained in any other member of $\mathcal{F}$. Raman and Ray proved (in a much more general form) that if $\mathcal{F}$ is a family of non-piercing regions, then $\mathcal{D}(\mathcal{F})$ is planar.

Corollary 5 (of Raman and Ray [13]). *If $\mathcal{F}$ is a family of non-piercing regions, then the Delaunay graph $\mathcal{D}(\mathcal{F})$ is planar, therefore, it can have at most $3|\mathcal{F}|$ edges.*

Now we are ready to state the last lemma needed to complete the proof.

Lemma 6. *If $\mathcal{F}$ is a family of non-piercing regions with the $(\nu + 1, 2)$-property, then $\mathcal{F}$ also has the (p, q)-property for any $p > 3e\nu(\nu + 1)q + 1$ for any $q \geq 2$, where $e = 2.71...$ is Euler's number.*

We omit the proof due to space constraints and only mention that it is based on the Clarkson-Shor method [6].

Lemmas 4 and 6 imply that $\mathcal{F}$ satisfies the assumptions of Theorem 3 with $q = 5$ and some large enough p, which implies that $\mathcal{F}$ can be stabbed with a constant number of points. This finishes the proof of Theorem 1. $\square$

3 Proof of Theorem 2

In this section, we present the exact statement and the proof of Theorem 2.

First, we introduce the definitions and notation following Axenovich, Kießle and Sagdeev [3]. A plane graph is a graph that is embedded in the plane without crossing edges. A 2-connected plane graph G is *$C_{<g}$-free* if it contains no cycle of length smaller than g. G is a *maximal $C_{<g}$-free plane graph* if adding any new edge would either create a crossing or a cycle of length less than g. Define $f_{max}(g)$ as the largest possible face length of a 2-connected maximal $C_{<g}$-free plane graph.

Axenovich, Kießle and Sagdeev [3] showed that $f_{max}(g) = 2g - 3$ for $3 \leq g \leq 6$ using a former result of Axenovich, Ueckerdt, and Weiner [4]. For larger values of g, they showed a lower bound of $3g - 9$ for $7 \leq g \leq 9$, and $3g - 12$ for $g \geq 10$.

They also showed an upper bound of $2(g-2)^2 + 1$ for any $g \geq 7$, and asked whether it could be improved to a linear upper bound. We give an affirmative answer to this question.

Theorem 7 (Theorem 2, restated)**.** *$f_{max}(g) \leq Kg$ for some absolute constant K.*

We start with a simple observation made by Axenovich, Kießle and Sagdeev [3].

Observation 8 ([3])**.** *For any two vertices u, v of a facial cycle of a maximal $C_{<g}$-free graph, their distance $d(u, v) \leq g - 2$.*

Proof. If $d(u, v) \geq g - 1$, then u and v are non-adjacent, and the edge uv could be added inside the cycle, preserving planarity, contradicting the maximality of our graph. □

Subdividing each edge with a vertex shows that it is sufficient to prove Theorem 7 when g is even—due to space constraints, we omit the details here.

Fix a maximal $\mathcal{C}_g$-free plane graph G for some even g and a facial cycle C of G, we will bound the length m of C in terms of g. By Fáry's theorem, we can assume that G is geometric, that is, each of its edges is a segment, without changing the topology of the embedding.

For two vertices u, v, let $d(u, v)$ denote their distance in G, and define the distance of a vertex u and an edge $e = vw$ as $d(u, e) = \max\{d(u, v), d(u, w)\}$. Let $\overline{B}(x, r)$ denote the closed (euclidean) disk of radius r around a point x in the plane, and for a simple curve γ, let $\overline{B}(\gamma, r) = \{x \in \mathbb{R}^2 : \exists p \in \gamma \ |x - p| \leq r\}$, the set of points at euclidean distance at most r from (a point of) γ.

For each $v_i \in C$, we define a blow-up of its $(g/2 - 1)$-neighborhood in G. In order to do this, we need to choose some small enough $\rho, \delta > 0$, and $0 < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_m < \frac{\delta}{g^2}$.

Now we can define the neighborhood regions N_i for each vertex v_i (see Figure 1).

$$N_i = \bigcup_{\substack{u \in V(G) \\ d(v_i, u) \leq g/2-1}} \overline{B}\left(u, \frac{\rho}{d(v_i, u)} + \varepsilon_i\right) \quad \cup \bigcup_{\substack{e \in E(G) \\ d(v_i, e) \leq g/2-1}} \overline{B}\left(e, \frac{\delta}{d(v_i, e)} + \varepsilon_i\right).$$

Note that the boundary of any two members of $\mathcal{N} = \{N_1, N_2, \ldots, N_m\}$ intersect in finitely many points as the radii of their blow-ups are perturbed with small amounts, so

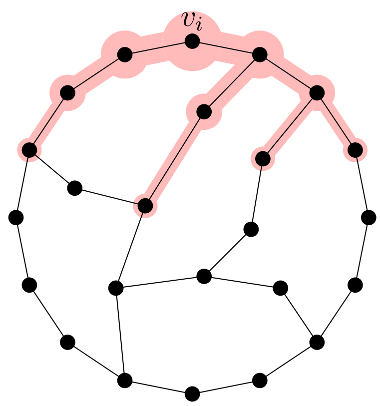


Figure 1: Illustration for a neighborhood N_i.

$\mathcal{N}$ is a family of regions in general position. In order to apply Theorem 1 to $\mathcal{N}$, we need to verify that it satisfies the two properties required by the theorem.

Lemma 9. *The set-system $\mathcal{N}$ is non-piercing.*

Due to space constraints, we omit the proof.

Lemma 10. *The set-system $\mathcal{N}$ has the $(2,2)$-property, i.e., it is pairwise intersecting.*

Proof. Take two arbitrary vertices $v_i, v_j \in C$. By Observation 8, $d(v_i, v_j) \leq g - 2$. Since g is even, this implies that there is a vertex w at distance at most $g/2 - 1$ from both v_i and v_j. By the definition of N_i and N_j, $w \in N_i \cap N_j$, which proves our statement. □

Now, Theorem 1 implies that there is a absolute constant T such that there exists a point set $Z = \{z_1, z_2, \dots, z_T\}$ hitting each member of $\mathcal{N}$.

Note that we may take the points of Z to be vertices of G, as the sets in $\mathcal{N}$ are unions of neighborhoods of vertices and edges of G, and if a set N_i contains a point in the neighborhood of an edge $e = uv$, then N_i contains both u and v. As a consequence, we have a set of T vertices such that for each vertex $v_i \in C$, there is an element of Z at distance at most $g/2 - 1$ from v_i in G.

Inspired by Axenovich, Kießle and Sagdeev [3], we define a partitioning of C into sets $C_1, C_2, \dots, C_T$ such that $v_i \in C_j$ if and only if $d(v_i, z_j) \leq d(v_i, z_{j'})$ for each $j' \neq j$, and $d(v_i, z_j) < d(v_i, z_{j'})$ for each $j' < j$. In other words, we assign each v_i to the minimum index vertex in Z which is closest to it.

Next, we present two statements (due to space constraints, without proofs) about the distributions of the sets C_i, which help us bound the number of vertices on C.

Lemma 11 ([3])**.** *If $v_{i+1}, \dots, v_{i+j} \in C_k$ are consecutive vertices of C in the same partition class, then $j \leq g - 1$.*

Lemma 12. *There cannot exist indices $h_1 < h_2 < h_3 < h_4$, for which $v_{h_1}, v_{h_3} \in C_i$ and $v_{h_2}, v_{h_4} \in C_j$ for some $i \neq j$.*

Therefore, by the Davenport–Schinzel theorem [8], we get that the length of C must satisfy $|C| \leq (2T-1)(g-1)$, which completes the proof of Theorem 7 and thus of Theorem 2. □

References

[1] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *Journal of the ACM*, 51(2):139–186, 2004.

[2] N. Alon and D. Kleitman. Piercing convex sets and the Hadwiger–Debrunner (p, q)-problem. *Advances in Mathematics*, 96:103–112, 1992.

[3] M. Axenovich, L. Kießle, and A. Sagdeev. Faces of maximal plane graphs without short cycles, 2024. arXiv:2410.13481.

[4] M. Axenovich, T. Ueckerdt, and P. Weiner. Splitting planar graphs of girth 6 into two linear forests with short paths. *Journal of Graph Theory*, 85(3):601–618, 2017.

[5] P. Carmi, M. J. Katz, and P. Morin. Stabbing Pairwise Intersecting Disks by Four Points. *Discrete and Computational Geometry*, 70:1751–1784, 2023.

[6] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete and Computational Geometry*, 4:387–421, 1989.

[7] L. Danzer. Zur Lösung des Gallaischen Problems über Kreisscheiben in der Euklidischen Ebene. *Studia Scientiarum Mathematicarum Hungarica*, 21(1–2):111–134, 1986.

[8] H. Davenport and A. Schinzel. A Combinatorial Problem Connected with Differential Equations. *American Journal of Mathematics*, 87(3):684–694, 1965.

[9] A. Dumitrescu and M. Jiang. Coloring translates and homothets of a convex body. *Beitrage zur Algebra und Geometrie*, 53(2):365–377, 2012.

[10] H. Hadwiger and H. Debrunner. Über eine Variante zum Helly'schen Satz. *Archiv der Mathematik*, 8:309–313, 1957.

[11] A. Holmsen and R. Wenger. Helly-type theorems and geometric transversals. In *Handbook of Discrete and Computational Geometry.* CRC Press, Boca Raton, FL, 2017.

[12] J. Matoušek. Bounded VC-Dimension Implies a Fractional Helly Theorem. *Discrete and Computational Geometry*, 31(2):251–255, 2004.

[13] R. Raman and S. Ray. Planar Support for Non-piercing Regions and Applications. In *26th Annual European Symposium on Algorithms*, volume 112 of *LIPIcs*, pages 69:1–69:14, 2018.

[14] A. Scott and P. Seymour. A survey of χ-boundedness. *Journal of Graph Theory*, 95(3):473–504, 2020.

[15] L. Stachó. A Gallai-féle körletüzési probléma megoldása. *Matematikai Lapok*, 32:19–47, 1981–84.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Complexity of finding maximal color-avoiding connected subgraphs

(Extended abstract)

Panna Tímea Fekete* Roland Molontay† Kitti Varga‡

Abstract

Two vertices of a (not necessarily properly) vertex-colored digraph are called vertex-ℓ-color-avoiding strongly k-arc-connected if, after the removal of the vertices of any at most ℓ colors, either at least one of the two vertices is removed, or they remain strongly k-arc-connected. The vertex-ℓ-color-avoiding strongly k-arc-connected components are maximal sets of vertices such that any two vertex belonging to the set is vertex-ℓ-color-avoiding strongly k-arc-connected in the graph.

We show that finding these components is NP-hard in general, but this problem becomes solvable in polynomial time when each vertex is single-colored. We also investigate the problem of finding the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs. We prove that this problem is likewise NP-hard in general but becomes solvable in polynomial time when each vertex is single-colored. In addition, we explore possible generalizations to k-vertex-connectivity and to arc-colored digraphs.

1 Introduction

Various theoretical frameworks have been developed to analyze network robustness. One of the most widely used approaches is percolation theory, which studies the behavior of connected components when vertices or edges fail with a given probability [3,14]. A significant limitation of standard network percolation models is that they treat failures as independent events, whereas in real-world systems, certain parts of the network may share common vulnerabilities.

A notable example is transportation networks, where different transportation modes (rail, road, air, etc.) may be affected by distinct external factors. Another example is communication networks, where routers belonging to the same country may be compromised together, requiring an alternative routing strategy that avoids any single country's infrastructure.

*HUN-REN Alfréd Rényi Institute of Mathematics,1053 Budapest, Hungary. E-mail: fekete.panna.timea@renyi.hu. Supported by the ERC Synergy Grant No. 810115 – DYNASNET.

†Department of Stochastics, Institute of Mathematics, Budapest University of Technology and Economics, 1111 Budapest, Hungary. E-mail: molontay@math.bme.hu.

‡Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, 1111 Budapest, Hungary. E-mail: vkitti@cs.bme.hu.

While in the example of transportation networks attacks target the edges, in the case of routers, they affect the vertices. In the latter scenario, we can consider two natural variations depending on how we define being compromised: whether the affected routers are entirely removed from the network and the question is how the remaining routers can still communicate; or whether routers registered to the same country can eavesdrop collectively on messages and the challenge is to ensure secure communication between two fixed routers despite potential interception.

To address such scenarios, a novel framework called color-avoiding percolation has been introduced by Krause et al. [10, 12, 13]. In this model, each vertex or edge is assigned a color representing a shared vulnerability. Two vertices are color-avoiding connected if they can reach each other via at least one path that avoids any single color.

Earlier, edge-color-avoiding colorings were studied under the name courteous edge-colorings by DeVos et al. [4]. In particular, graphs that admit a 1-courteous edge-coloring are precisely the edge-color-avoiding connected graphs. Molontay and Varga [20] investigated the computational complexity of finding the so-called color-avoiding connected components of a graph. This concept was defined by Krause et al. [12] as maximal sets of vertices such that any two vertex belonging to the set is color-avoiding connected in the graph. Pintér and Varga [21] studied the problem of finding color-avoiding connected spanning subgraphs with the minimum number of edges in edge- or vertex-colored graphs.

Typically, k-edge- or k-vertex-connected components are defined as maximal k-edge- or k-vertex-connected subgraphs. However, this definition does not necessarily coincide with the concept of maximal sets of vertices such that any two vertex belonging to the set are k-edge- or k-vertex-connected in the graph, for arbitrary $k \in \mathbb{Z}_+$. The computational and algorithmic aspects of finding maximal k-edge- or k-vertex-connected subgraphs have been studied extensively, see for example [9, 15]. The problem of finding maximal sets of vertices such that any two vertex belonging to the set are k-edge-connected in the graph has been also investigated [24].

In this paper, we address these problems within the color-avoiding framework, with a primary focus on finding maximal color-avoiding connected subgraphs. We explore the concept of color-avoidance both for arc- and vertex-colored digraphs, including two variants for the latter, motivated by the real-world scenarios described above. Our approach also considers the case when multiple color classes are attacked simultaneously, and we require higher strong arc- or vertex-connectivity. In this extended abstract, we build upon the results of Molontay and Varga [20], but with revised terminology.

We focus on the setting motivated by unidirectional communication networks, where routers registered to the same country may become simultaneously inoperative. Our goal is to ensure robustness even in scenarios where routers from multiple countries can fail at the same time, by requiring the remaining network to maintain high arc-connectivity.

2 Vertex-color-avoiding strongly arc-connected components

Let G be a digraph, let C be a finite color set, and let f be a function assigning some (possibly multiple or no) colors to every vertex. Given a set $C' \subseteq C$ of colors, by the *removal of* C' from G, we mean the removal of all vertices whose set of colors contain at least one color of C'.

Let $k, \ell \in \mathbb{Z}_+$. Two (not necessarily distinct) vertices $u, v \in V(G)$ are called *vertex-ℓ-*

color-avoiding strongly k-arc-connected if after the removal of any at most ℓ colors, either at least one of the vertices u and v is removed, or they remain strongly k-arc-connected[1]. For an example, see Figure 1.

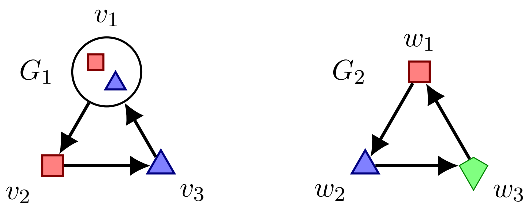


Figure 1: Examples for vertex-ℓ-color-avoiding strongly k-arc-connected vertices. The colors red, blue, and green are denoted by (red) squares, (blue) triangles, and (green) deltoids, respectively. The vertex v_1 is assigned two colors, namely red and blue. In the digraph G_1, any two vertices are vertex-2-color-avoiding strongly 1-arc-connected; thus, G_1 is vertex-2-color-avoiding strongly 1-arc-connected. In the digraph G_2, no pair of vertices are vertex-1-color-avoiding strongly 1-arc-connected.

The *vertex-ℓ-color-avoiding strongly k-arc-connected components* of a vertex-colored digraph G are maximal sets of vertices such that any two of these vertices are vertex-ℓ-color-avoiding strongly k-arc-connected. If a vertex $v \in V(G)$ is not vertex-ℓ-color-avoiding strongly k-arc-connected to any other vertex of the digraph G for some $k, \ell \in \mathbb{Z}_+$, then v is considered to form a vertex-ℓ-color-avoiding strongly k-arc-connected component of G in itself, respectively.

We say that a vertex-colored digraph on at least two vertices is *vertex-ℓ-color-avoiding strongly k-arc-connected* if it has exactly one vertex-ℓ-color-avoiding strongly k-arc-connected component. For an example, see Figure 1.

It can be proved that the problem of finding vertex-ℓ-color-avoiding strongly k-arc-connected components is NP-hard. Let us point out that the special case $k = \ell = 1$ for undirected graphs was already proved by Molontay and Varga [20].

Theorem 2.1. For any $k, \ell \in \mathbb{Z}_+$, the following problem is NP-complete even for undirected graphs.

[V$\boldsymbol{\ell}$ | $\boldsymbol{k}$A]-ConnectedComponents

Instance: a digraph G, a finite color set C, a function $f\colon V(G) \to 2^C$, and $m \in \mathbb{Z}_+$.

Question: is it true that G has a vertex-ℓ-color-avoiding strongly k-arc-connected component of size at least m?

For finding the vertex-ℓ-color-avoiding strongly k-arc-connected components of a vertex-colored digraph G for any fixed $k, \ell \in \mathbb{Z}_+$, we consider the following undirected auxiliary graph. The vertex set of the auxiliary graph is the same as that of G, and two vertices of the auxiliary graph are connected by an edge if and only if these vertices are vertex-ℓ-color-avoiding strongly k-arc-connected in G.

It is not difficult to see that a set $W \subseteq V(G)$ is the vertex set of a vertex-ℓ-color-avoiding strongly k-arc-connected component of G if and only if W spans a clique in the auxiliary graph. Also, observe that we can construct this auxiliary graph in polynomial time.

[1] Two vertices u and v of a digraph G are called *strongly k-arc-connected* if there exist at least k pairwise arc-disjoint u-v dipath, and at least k pairwise arc-disjoint v-u dipath in G.

The key idea in the proof of Theorem 2.1 is to show that the auxiliary graph can be arbitrary. Then, the NP-complete problem Clique [11] trivially reduces to the problem of finding the vertex-ℓ-color-avoiding strongly k-arc-connected components.

In the proof of Theorem 2.1, the size of the sets of colors on the vertices is a polynomial of n. An interesting question is the case when the sets of colors of the vertices have constant sizes. In the following, we consider the case when every vertex has exactly one color in its set of colors. The case when the sets of colors of the vertices have arbitrary constant sizes remains open.

Theorem 2.2. For any $k, \ell \in \mathbb{Z}_+$, the following problem can be solved in polynomial time.

[Vℓ | kA]-ConnectedComponents–EachVertexSingleColored

Instance: a digraph G, a finite color set C, a function $f\colon V(G) \to C$, and $m \in \mathbb{Z}_+$.

Question: is it true that G has a vertex-ℓ-color-avoiding strongly k-arc-connected component of size at least m?

Moreover, in this single-colored case all the vertex-ℓ-color-avoiding strongly k-arc-connected components of G can be found in polynomial time.

The proof of Theorem 2.2 is based on the following lemma, which states that the auxiliary graph defined above has some special structure.

Lemma 2.2.1. Let $k, \ell \in \mathbb{Z}_+$, let G be a digraph, let C be a finite color set, and let $f : V(G) \to C$. Then the auxiliary graph is $\overline{(\ell+2)K_2}$-free, i.e., it does not contain $\overline{(\ell+2)K_2}$ as a (not necessarily induced) subgraph.

From a theorem of Balas and Yu [2, Theorem 4], it follows that for any $p \in \mathbb{Z}_+$ a $\overline{pK_2}$-free graph has $O(n^{2p-2})$ maximal cliques. Combining this with a theorem of Tsukiyama, Ide, Ariyoshi, and Shirakawa [23, Theorem 3], these maximal cliques can be found in polynomial time. Thereby, we can conclude the proof of Theorem 2.2.

3 Maximal vertex-color-avoiding strongly arc-connected subgraphs

It is not difficult to see that even though some set $W \subseteq V(G)$ of vertices is the ground set of a vertex-ℓ-color-avoiding strongly k-arc-connected component of a vertex-colored digraph G, the induced subgraph $G[W]$ on the vertex set W is not necessarily vertex-ℓ-color-avoiding strongly k-arc-connected. Now we study such maximal vertex sets that span a vertex-color-avoiding connected subgraph. A set $W \subseteq V(G)$ of vertices is called a *maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraph* of a vertex-colored digraph G if it is inclusion-wise maximal for the property that $G[W]$ is vertex-ℓ-color-avoiding strongly k-arc-connected.

Similarly to the problem of finding the vertex-ℓ-color-avoiding strongly k-arc-connected component of a vertex-colored digraph, the NP-complete problem Clique [11] can be reduced to the problem of finding the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs.

Theorem 3.1. For any $k, \ell \in \mathbb{Z}_+$, the following problem is NP-complete even for undirected graphs.

Maximal-[Vℓ | kA]-ConnectedSubgraphs

Instance: a digraph G, a finite color set C, a function $f\colon V(G) \to 2^C$, and $m \in \mathbb{Z}_+$.

Question: is it true that G has a maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraph on at least m vertices?

Again, we consider the setting when every vertex is single-colored. In this special case, we present a polynomial-time algorithm for finding the maximal vertex-color-avoiding-connected subgraphs of a given digraph for any fixed $k, \ell \in \mathbb{Z}_+$.

In order to do so, we use the algorithm that can find the vertex-color-avoiding-connected components when every vertex is single-colored as a subroutine. Let **[V$\boldsymbol{\ell}$ | $\boldsymbol{k}$A]-Connected-Components**(G, f) denote a polynomial-time algorithm which returns the vertex sets of the vertex-ℓ-color-avoiding strongly k-arc-connected components of a digraph G with vertex-coloring $f\colon V(G) \to C$, where C is a finite color set. For ease of notation, let f_W denote the restriction of f to W for any $W \subseteq V(G)$.

The proposed algorithm, presented as Algorithm 1, iteratively takes the vertex-ℓ-color-avoiding strongly k-arc-connected components of the vertex-ℓ-color-avoiding strongly k-arc-connected components. At the end of this procedure, we get the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs.

Algorithm 1: Algorithm for finding the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs when every vertex is single-colored

Input: a digraph G, a finite color set C, and a function $f\colon V(G) \to C$.
Output: the vertex sets of the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs of G.

$\mathcal{V}_{-1} \leftarrow \{\emptyset\}$
$\mathcal{V}_0 \leftarrow \big\{V(G)\big\}$
$i \leftarrow 0$
while $\mathcal{V}_i \neq \mathcal{V}_{i-1}$ **do**
  for $W \in \mathcal{V}_i$ **do**
    $\mathcal{V}_{i+1} \leftarrow \mathcal{V}_{i+1} \cup \Big\{[\mathrm{V}\ell \,|\, k\mathrm{A}]\text{-ConnectedComponents}\big(G[W], f_W\big)\Big\}$
  $i \leftarrow i + 1$
return $\mathcal{V}_i$

In order to analyze the algorithm, let us first study the sizes of the pairwise intersections of the elements of $\mathcal{V}_i$ for any $i \in \mathbb{Z}_+$. Here, we omit the proof of the following technical lemma.

Lemma 3.1.1. Let $k, \ell \in \mathbb{Z}_+$, let G be a digraph, let C be a finite color set, and let $f\colon V(G) \to C$. Then, for any $i \in \mathbb{Z}_+$ and for any $W, W' \in \mathcal{V}_i$, the following hold.

(i) At most ℓ colors appear in $W \cap W'$.
(ii) There exists at least one color appearing in $W \cup W'$ but not in $W \cap W'$.

Now we show the correctness of Algorithm 1 and analyze its running time.

Theorem 3.2. Let $k, \ell \in \mathbb{Z}_+$ be fixed. Given a digraph G, a finite color set C, a function $f\colon V(G) \to C$,Algorithm 1 outputs the vertex sets of the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs of G and it runs in polynomial time.

Proof. We discuss the time complexity and the correctness of the algorithm separately. Due to space constraints, we omit the technical details of the proof of correctness here.

Time complexity. Let $i \in \mathbb{Z}_+$. The elements of $\mathcal{V}_i$ are of $\ell + 1$ types: those in which exactly $\ell' \in [\ell]$ colors appear, and those in which at least $\ell+1$ colors appear. Now we prove that there are at most $\binom{n}{\ell'}$ elements of $\mathcal{V}_i$ in which exactly ℓ' colors appear for any $\ell' \in [\ell]$, and there are at most $\binom{n}{\ell+1}$ elements of $\mathcal{V}_i$ in which at least $\ell + 1$ colors appear. Hence, $|\mathcal{V}_i| \in O(n^{\ell+1})$.

Let $\ell' \in [\ell]$ and let $W \in \mathcal{V}_i$ be an element in which exactly ℓ' colors appear. Then there exist ℓ' distinct vertices $w_1, \ldots, w_{\ell'} \in W$ that all have distinct colors. By Lemma 3.1.1.(ii), there is at least one color in the union of any two distinct elements of $\mathcal{V}_i$ which does not appear in their intersection. Hence for any $W' \in \mathcal{V}_i$ in which exactly ℓ' colors appear and for which $\{w_1, \ldots, w_{\ell'}\} \subseteq W'$ holds, we have $W = W'$. Therefore, there are indeed at most $\binom{n}{\ell'}$ elements of $\mathcal{V}_i$ in which exactly ℓ' colors appear.

Now let $W \in \mathcal{V}_i$ be an element in which at least $\ell+1$ colors appear. Then there exist $\ell+1$ distinct vertices $w_1, \ldots, w_{\ell+1} \in W$ that all have distinct colors. Then by Lemma 3.1.1.(i), for any $W' \in \mathcal{V}_i$ for which $\{w_1, \ldots, w_{\ell+1}\} \subseteq W'$ holds, we have $W = W'$. Therefore, there are at most $\binom{n}{\ell+1}$ elements of $\mathcal{V}_i$ in which at least $\ell + 1$ colors appear.

Now we show that the algorithm terminates after at most n iterations. To see this, let $i \in \mathbb{Z}_+$ and $W \in \mathcal{V}_i$ arbitrary. Clearly, there exists $U \in \mathcal{V}_{i-1}$ such that $W \subseteq U$. We distinguish two cases: $W = U$ or $W \subsetneq U$. In the first case, $G[U]$ is vertex-ℓ-color-avoiding strongly k-arc-connected. Thus, U is, but none of its proper subsets are contained in $\mathcal{V}_j$ for any $j \geq i$. In the latter case, clearly $|W| < |U|$ holds. Furthermore, the only element of $\mathcal{V}_0$ has size n and since $W \notin \mathcal{V}_{i-1}$, we obtain $|W| \leq n - i$ by the previous observation. Note that $\emptyset \notin \mathcal{V}_j$ for all $j \in \mathbb{Z}_+$, so each element of $\mathcal{V}_j$ has size at least one. Hence, the while loop indeed terminates after at most n iterations.

Therefore, the subroutine of finding the vertex-ℓ-color-avoiding strongly k-arc-connected components of a subgraph of G is invoked $O(n^{\ell+2})$ times, and thus by Theorem 2.2, the algorithm runs in polynomial time. □

By Theorem 3.2, we conclude that for any fixed $k, \ell \in \mathbb{Z}_+$, the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs can be found in polynomial time if each vertex is single-colored.

Corollary 3.2.1. For any $k, \ell \in \mathbb{Z}_+$, the following problem can be solved in polynomial time.
Maximal-[Vℓ | kA]-ConnectedSubgraphs-EachVertexSingleColored
Instance: a digraph G, a finite color set C, a function $f\colon V(G) \to C$, and $m \in \mathbb{Z}_+$.
Question: is it true that G has a maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraph on at least m vertices?
Moreover, in this single-colored case all the maximal vertex-ℓ-color-avoiding strongly k-arc-connected subgraphs of G can be found in polynomial time.

4 Further results

Let us summarize our results for the additional settings mentioned in the Introduction. In the case of ℓ-vertex-color-avoiding k-vertex-connectivity, the problems of finding the color-avoiding connected components and the maximal color-avoiding connected subgraphs have the same complexity as in the case of ℓ-vertex-color-avoiding k-arc-connectivity. Specifically, these problems are NP-complete when arbitrary vertex colorings are considered, but they become solvable in polynomial time when each vertex is assigned a single color. The proofs rely on similar ideas, but they are more technical than those described in Section 3.

Note that Theorem 2.2 shows that the problem [Vℓ | kA]-ConnectedComponents–EachVertexSingleColored is in the complexity class XP parameterized by k and ℓ. It remains open whether this problem is fixed-parameter tractable parameterized by these parameters.

For the other variant of color-avoidance in vertex-colored graphs, both problems become NP-hard, even when all the vertices are colored with the same single color, in both cases where higher strong arc- or vertex-connectivity is required. However, in the case of arc-colored graphs, both problems can be solved in polynomial time, even with arbitrary arc colorings, in both cases where higher strong arc- or vertex-connectivity is required.

References

[1] R. Albert, H. Jeong, and A.-L. Barabási, *Error and attack tolerance of complex networks*, Nature **406** (2000), no. 6794, 378–382.

[2] E. Balas and C. S. Yu, *On graphs with polynomially solvable maximum-weight clique problem*, Networks **19** (1989), no. 2, 247–253.

[3] R. Cohen and S. Havlin, *Percolation in complex networks*, Complex Media and Percolation Theory, 2021, pp. 419–431.

[4] M. DeVos, T. Johnson, and P. Seymour, *Cut coloring and circuit covering* (2006), available at `\url{https://web.math.princeton.edu/~pds/papers/cutcolouring/paper.pdf}`.

[5] P. T. Fekete, R. Molontay, B. Ráth, and K. Varga, *Color-avoiding percolation and branching processes*, Journal of Applied Probability **61** (2024), no. 3, 942–966.

[6] M. Giusfredi and F. Bagnoli, *A self-organized criticality method for the study of color-avoiding percolation*, Internet Science, 2019, pp. 217–226.

[7] M. Giusfredi and F. Bagnoli, *From color-avoiding to color-favored percolation in diluted lattices*, Future Internet **12** (2020), no. 8, article no. 139.

[8] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), no. 4, 357–368.

[9] W. Holberg, *The decomposition of graphs into k-connected components*, Discrete Mathematics **109** (1992), no. 1–3, 133–145.

[10] A. Kadović, S. M. Krause, G. Caldarelli, and V. Zlatić, *Bond and site color-avoiding percolation in scale-free networks*, Physical Review E **98** (2018), no. 6, article no. 062308.

[11] R. Karp, *Reducibility Among Combinatorial Problems*, Complexity of Computer Computations **40** (1972), 85–103.

[12] S. M. Krause, M. M. Danziger, and V. Zlatić, *Hidden connectivity in networks with vulnerable classes of nodes*, Physical Review X **6** (2016), no. 4, article no. 041022.

[13] S. M. Krause, M. M. Danziger, and V. Zlatić, *Color-avoiding percolation*, Physical Review E **96** (2017), no. 2, article no. 022313.

[14] M. Li, R. R. Liu, L. Lü, M.-B. Hu, S. Xu, and Y.-C. Zhang, *Percolation on complex networks: Theory and application*, Physics Reports **907** (2021), 1–68.

[15] Y. Li, Y. Zhao, G. Wang, F. Zhu, Y. Wu, and S. Shi, *Effective k-Vertex Connected Component Detection in Large-Scale Networks*, Database Systems for Advanced Applications: 22nd International Conference, DASFAA 2017, Suzhou, China, March 27-30: Proceedings, 2017, pp. 404–421.

[16] L. Lichev, *Color-avoiding percolation of random graphs: between the subcritical and the intermediate regime*, Discrete Mathematics **347** (2024), no. 1, article no. 113713.

[17] L. Lichev and B. Schapira, *Color-avoiding percolation on the Erdős-Rényi random graph* (2022), available at `\href{https://arxiv.org/abs/2211.16086}{\texttt{arXiv:\allowbreak2211.16086[math.PR]}}`.

[18] E. Malaguti and P. Toth, *A survey on vertex coloring problems*, International Transactions in Operational Research **17** (2010), no. 1, 1–34.

[19] K. Menger, *Zur allgemeinen Kurventheorie*, Fundamenta Mathematicae **10** (1927), no. 1, 96–115.

[20] R. Molontay and K. Varga, *On the complexity of color-avoiding site and bond percolation*, SOFSEM 2019: Theory and Practice of Computer Science, 2019, pp. 354–367.

[21] J. Pintér and K. Varga, *Color-avoiding connected spanning subgraphs with minimum number of edges*, Discrete Applied Mathematics **349** (2024), 25–43.

[22] L. M. Shekhtman, M. M. Danziger, I. Bonamassa, S. V. Buldyrev, G. Caldarelli, V. Zlatić, and S. Havlin, *Critical field-exponents for secure message-passing in modular networks*, New Journal of Physics **20** (2018), no. 5, article no. 053001.

[23] S. Tsukiyama, M. Ide, H. Ariyoshi, and I. Shirakawa, *A new algorithm for generating all the maximal independent sets*, SIAM Journal on Computing **6** (1977), 3.

[24] T. Wang, Y. Zhang, F.Y.L. Chin, H. Ting, Y.H. Tsin, and S. Poon, *A simple algorithm for finding all k-edge-connected components*, PLoS ONE **10** (2015), no. 9.

[25] X. Yuan, Y. Dai, H. E. Stanley, and S. Havlin, *k-core percolation on complex networks: Comparing random, localized, and targeted attacks*, Physical Review E **93** (2016), no. 6, article no. 062302.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Isometric path partition: an upper bound and a characterization of some extremal graphs

(Extended abstract)

Irena Penev* R. B. Sandeep† D. K. Supraja‡ S Taruni§

Abstract

An *isometric path* is a shortest path between two vertices. An *isometric path partition* (IPP) of a graph G is a set $\mathcal{I}$ of vertex-disjoint isometric paths in G that partition the vertices of G. The *isometric path partition number* of G, denoted by $\mathrm{ipp}(G)$, is the minimum cardinality of an IPP of G. In this article, we prove that every graph G satisfies $\mathrm{ipp}(G) \leq |V(G)| - \nu(G)$, where $\nu(G)$ is matching number of G, and we characterize all connected graphs on an odd number of vertices that attain this upper bound. Additionally, we determine all block graphs that achieve this bound.

1 Introduction

Graph partitioning and covering problems are central topics in graph theory and algorithms, including problems such as dominating set (covering by stars), clique covering (covering by cliques), coloring (partitioning by independent sets), and covering/partitioning by paths or cycles. In particular, the problems of partitioning and covering by paths have their connections to important graph-theoretic results such as the Gallai-Milgram theorem [11] and Berge's path partition conjecture [2]. These problems also have applications in various domains, including code optimization [3], bioinformatics [13], program testing [15], to name a few. In recent years, various types of path partitions have been studied, such as unrestricted path partition [7, 12, 17], induced path partition [6, 16], and isometric (shortest) path partition [5, 8]. We note that all isometric paths are induced, but not all induced paths are isometric. The focus of this article is on the problem of partitioning the vertex set of a graph into isometric paths.

In what follows, all graphs are assumed to be finite, simple, and nonnull. Unless stated otherwise, we use standard graph-theoretic terminology and notation according to West [18]. For a graph G, $V(G)$ denotes the vertex set of G and $E(G)$ denotes the edge set of G. For a nonempty set $X \subseteq V(G)$, we denote by $G[X]$ the subgraph of G induced by X. For a set $X \subsetneq V(G)$, we define $G \setminus X := G[V(G) \setminus X]$. The *length* of a path is the number of edges that it contains, and the

*Computer Science Institute (IÚUK, MFF), Charles University, Prague, Czech Republic. E-mail: `ipenev@iuuk.mff.cuni.cz`. Supported by GAČR grant 25-17377S.

†Department of Computer Science and Engineering, Indian Institute of Technology Dharwad, India. E-mail: `sandeeprb@iitdh.ac.in`. Supported by ANRF (SERB) grant MTR/2022/000692.

‡Indian Institute of Technology Dharwad, India. E-mail: `dksupraja95@gmail.com`.

§Centro de Modelamiento Matemático (CNRS IRL2807), Universidad de Chile, Santiago, Chile. E-mail: `tsridhar@cmm.uchile.cl`. Supported by Centro de Modelamiento Matemático (CMM) BASAL fund FB210005 for center of excellence from ANID-Chile and by MSCA-RISE-2020-101007705 project *RandNET*.

distance between vertices u and v in a graph G, denoted by $d_G(u,v)$, is the length of a shortest path in G between u and v. An *isometric path* between vertices x and y in G is a path of length $d_G(x,y)$ in G, with endpoints x and y. Note that any one- or two-vertex path is an isometric path, as is any induced three-vertex path. An *isometric path partition* (or *IPP* for short) of a graph G is a set of isometric paths in G such that the vertex sets of those paths form a partition of $V(G)$. The *isometric path partition number* of G, denoted by $\text{ipp}(G)$, is the minimum cardinality of an IPP of G.

Every connected graph G trivially satisfies $\text{ipp}(G) \geq \left\lceil \frac{|V(G)|}{\text{diam}(G)+1} \right\rceil$, where $\text{diam}(G)$ is the diameter of G (see [9]). This bound has been used to determine the exact isometric path partition number of hypercubes Q_n when $n+1 = 2^t$ [10], as well as of $r \times r$ tori [14]. There are polynomial-time algorithms for computing an optimal IPP in certain graph classes, including trees [17], block graphs [16], cographs, and chain graphs [5]. On the other hand, the isometric path partition problem is NP-hard for bipartite graphs with diameter 4 [8], as well as for split graphs [5] and chordal graphs [4]. As there are no significant upper bounds known for this parameter, we propose an upper bound for the isometric path partition number of a graph G based on the matching number $\nu(G)$ of G (i.e. the number of edges in a maximum matching of G), and we characterize all connected graphs of odd order (connected graphs on an odd number of vertices) that achieve this bound. We also characterize block graphs that achieve this bound.

1.1 Our contribution

For a matching M of a graph G, a vertex of G is *M-saturated* if it is incident with some edge of M, and it is *M-unsaturated* otherwise. When we write that "$p_1 \dots p_k$ is a path in G," we mean that $p_1, \dots, p_k$ are pairwise distinct vertices of G such that $p_ip_{i+1} \in E(G)$ for all $i \in \{1, \dots, k-1\}$. For notational convenience, we will identify any one-vertex path with its unique vertex, and we will identify any two-vertex path with its unique edge. Clearly, any one- or two-vertex path is isometric. So, if M is a matching of a graph G, and U is the set of all M-unsaturated vertices of G, then $\mathcal{I} := M \cup U$ is an IPP of G, and we see that $\text{ipp}(G) \leq |\mathcal{I}| = |M| + |U| = |M| + \big(|V(G)| - 2|M|\big) = |V(G)| - |M|$. By choosing M to be a maximum matching of G, so that $|M| = \nu(G)$, we obtain the following proposition.

Proposition 1.1. *Every graph G satisfies $\text{ipp}(G) \leq |V(G)| - \nu(G)$.*

In what follows, we will be interested in certain types of graphs G for which the inequality from Proposition 1.1 becomes an equality, i.e., for which $\text{ipp}(G) = |V(G)| - \nu(G)$. We begin with some definitions. A *block* of a graph G is a maximal biconnected subgraph of G. (Whenever convenient, we will identify blocks with their vertex sets.) An *odd* (resp., *even*) *block* is a block with an odd (resp., even) number of vertices. A *block graph* is a graph in which every block is a clique. Note that a connected block graph is either complete or contains a cut-vertex. Moreover, block graphs are precisely the diamond-free chordal graphs [1]. (A *chordal graph* is a graph in which all induced cycles are triangles. The *diamond* is the graph obtained from the complete graph on four vertices by deleting one edge, and a graph is *diamond-free* if it contains no induced diamond.) Our results are the following two theorems.

Theorem 1.2. *Let G be a block graph. Then $\text{ipp}(G) = |V(G)| - \nu(G)$ if and only if every component of G contains at most one even block.*

Theorem 1.3. *Let G be a connected graph on an odd number of vertices. Then, $\text{ipp}(G) = |V(G)| - \nu(G)$ if and only if G is a block graph with no even blocks.*

2 Proof outlines

2.1 Proof sketch of Theorem 1.2

In this section, we sketch the proof of Theorem 1.2. Some propositions are stated without proof; their proofs are deferred to the full manuscript associated to this extended abstract.

Proposition 2.1. *Let G be a graph. Then ipp$(G) = |V(G)| - \nu(G)$ if and only if every connected component C of G satisfies ipp$(C) = |V(C)| - \nu(C)$.*

A *leaf* in a graph G is a vertex of G that has exactly one neighbor in G. A *leaf-clique* of a block graph G is a block of G that contains at most one cut-vertex of G. Note that any connected block graph is either complete or has at least two leaf-cliques.

Proposition 2.2. *Let G be a graph on at least three vertices. Let u be a leaf of G, and let v be the unique neighbor of u in G. Then the following hold:*

(a) $|V(G \setminus \{u, v\})| - \nu(G \setminus \{u, v\}) = |V(G)| - \nu(G) - 1$;

(b) ipp$(G) - 1 \leq$ ipp$(G \setminus \{u, v\}) \leq$ ipp(G);

(c) if ipp$(G) = |V(G)| - \nu(G)$, *then* ipp$(G \setminus \{u, v\}) =$ ipp$(G) - 1$.

Proposition 2.3. *Let C be a leaf-clique of a block graph G with $|C| = k \geq 3$, and let $v \in C$ be a cut-vertex of G. Let k' be the largest even number such that $k' \leq k - 1$. Consider any $C' \subseteq C \setminus \{v\}$ such that $|C'| = k'$. Let $G' := G \setminus C'$. Then* ipp$(G) = |V(G)| - \nu(G)$ *if and only if* ipp$(G') = |V(G')| - \nu(G')$.

Proposition 2.4. *Let G be a block graph such that every component of G contains at most one even block. Then* ipp$(G) = |V(G)| - \nu(G)$.

Proof. We may assume inductively that for every block graph G' such that $|V(G')| < |V(G)|$, if every component of G' contains at most one even block, then ipp$(G') = |V(G')| - \nu(G')$. If G is disconnected, then Proposition 2.1 and the induction hypothesis immediately imply that ipp$(G) = |V(G)| - \nu(G)$, and we are done. Further, if G is a complete graph, then it is clear that ipp$(G) = |V(G)| - \nu(G)$, and again we are done. We may therefore assume that G is connected and not complete. Since G is a block graph, this implies that G has at least two leaf-cliques. Since all leaf-cliques are blocks of G, it follows that at least one leaf-clique of G, call it C, has an odd number of vertices. (Since G is connected and not complete, we see that $|C| \geq 3$.) Let $v \in C$ be a cut-vertex of G, and set $C' := C \setminus \{v\}$. Then $G \setminus C'$ is a connected block graph, and it exactly the same number of even blocks as G. So, by the induction hypothesis, we have that ipp$(G') = |V(G')| - \nu(G')$. But now Proposition 2.3 guarantees that ipp$(G) = |V(G)| - \nu(G)$. □

Proposition 2.5. *Let G be a graph, let P be an isometric path in G of length at least two, and assume that the endpoints of P are leaves of G. Then there exists an IPP $\mathcal{I}$ of G such that $P \in \mathcal{I}$ and $|\mathcal{I}| \leq |V(G)| - \nu(G) - 1$.*

Proposition 2.6. *Let G be a graph such that some component of G contains two nonadjacent leaves. Then* ipp$(G) \leq |V(G)| - \nu(G) - 1$.

Proposition 2.7. *Let G be a block graph that satisfies* ipp$(G) = |V(G)| - \nu(G)$. *Then every component of G contains at most one even block.*

Proof. We may assume inductively that for every block graph G' such that $|V(G')| < |V(G)|$, if ipp$(G') = |V(G')| - \nu(G')$, then every component of G' has at most one even block. Now, let us assume toward a contradiction that G has a component that has at least two even blocks. Then Proposition 2.1 and the induction hypothesis guarantee that G is connected. Moreover, since G contains at least two blocks, we know that G is not complete. Consequently, G contains at least two leaf-cliques. Clearly, each leaf-clique of G has at least two vertices, and moreover, each leaf-clique of G contains exactly one cut-vertex of G.

Claim 2.8. *All leaf-cliques of G contain exactly two vertices.*

Proof of Claim 2.8. Suppose otherwise, and fix a leaf-clique C of G such that $k := |C| \geq 3$. Fix a cut-vertex $v \in C$ of G. Fix the largest even integer $k' \leq k-1$ (since $k \geq 3$, we have that $k' \geq 2$). Let C' be any subset of $C \setminus \{v\}$ of size k'. Set $G' := G \setminus C'$. Then Proposition 2.3 guarantees that $\text{ipp}(G') = |V(G')| - \nu(G')$. But this contradicts the induction hypothesis, since G' is a connected block graph on fewer than $|V(G)|$ vertices, and it contains at least two even blocks (because it has exactly the same number of even blocks as G does). ♦

Since G contains at least two leaf-cliques, and since (by Claim 2.8) all leaf-cliques contain exactly two vertices, we see that G contains at least two leaves. Fix distinct leaves x and y in G. If x and y are adjacent, then $G[\{x,y\}]$ is a component of G that is isomorphic to K_2, which is impossible since G is connected and not complete. So, x and y are nonadjacent. Since G is connected, Proposition 2.6 implies that $\text{ipp}(G) \leq |V(G)| - \nu(G) - 1$, a contradiction. □

We now observe that Theorem 1.2 follows immediately from Propositions 2.4 and 2.7.

2.2 Proof of Theorem 1.3

In this section, we prove Theorem 1.3. In view of Proposition 1.1 and Theorem 1.2, it is enough to prove Proposition 2.10 (see below), which states that if G is a connected graph of odd order, satisfying $\text{ipp}(G) = |V(G)| - \nu(G)$, then G is chordal and diamond-free. Let us explain why this is indeed enough. Suppose that G is a connected graph on an odd number of vertices, as in the statement of Theorem 1.3. If G is a block graph with no even blocks, then Theorem 1.2 guarantees that $\text{ipp}(G) = |V(G)| - \nu(G)$. Conversely, suppose that $\text{ipp}(G) = |V(G)| - \nu(G)$. Then Proposition 2.10 guarantees that that G is chordal and diamond-free. So, by [1], G is a block graph. By Theorem 1.2, G has at most one even block. But note that any connected block graph with exactly one even block has an even number of vertices. Since $|V(G)|$ is odd, it follows that G has no even blocks. This shows that it is indeed enough to prove Proposition 2.10. We begin by proving the following technical proposition.

Proposition 2.9. *Let G be a connected graph on an odd number of vertices and satisfying $ipp(G) = |V(G)| - \nu(G)$. Then the following hold.*

(a) For every maximum matching M of G, and every matching edge $aa' \in M$ and M-unsaturated vertex u of G, if u has a neighbor in $\{a, a'\}$, then u is adjacent to both a and a'.

(b) For every maximum matching M of G, exactly one vertex of G is M-unsaturated.

(c) $|V(G)| = 2\nu(G) + 1$ and $ipp(G) = \nu(G) + 1$.

(d) For every $u \in V(G)$, there exists a maximum matching M of G such that u is the unique M-unsaturated vertex of G.

Proof. We first prove (a). Let M be a maximum matching of G and let U be the set of M-unsaturated vertices of G. Clearly, $|U| = |V(G)| - 2\nu(G)$, and $M \cup U$ is an IPP of G of size $|V(G)| - \nu(G)$. Now, fix a matching edge $aa' \in M$ and an M-unsaturated vertex u of G. Assume toward a contradiction that u is adjacent to exactly one vertex in $\{a, a'\}$. By symmetry, we may assume that $ua \in E(G)$ and $ua' \notin E(G)$. But now $\big(M \setminus \{aa'\}\big) \cup \big(U \setminus \{u\}\big) \cup \{uaa'\}$ is an IPP of G of size $\big(|M| - 1\big) + \big(|U| - 1\big) + 1 = |V(G)| - \nu(G) - 1$, contrary to the assumption that $\text{ipp}(G) = |V(G)| - \nu(G)$.

Next, we prove (b). Let M be a maximum matching of G, and let U be the set of all M-unsaturated vertices of G. Once again, $\mathcal{I} := M \cup U$ is an IPP of G of size $|V(G)| - \nu(G)$; since $\text{ipp}(G) = |V(G)| - \nu(G)$, $\mathcal{I}$ is a minimum IPP of G. Since $|V(G)|$ is odd, it is clear that $U \neq \emptyset$. We must show that $|U| = 1$. Suppose otherwise, and fix distinct $x, y \in U$. Let P' be an isometric path between x and y in G. But now

$$\mathcal{I}' := \{P'\} \cup \big\{P \in \mathcal{I} \mid V(P') \cap V(P) = \emptyset\big\} \cup \big\{\{v\} \mid v \in V(G) \setminus V(P') \text{ and } \exists u \in V(P') \text{ s.t. } uv \in M\big\}$$

is an IPP of G size at most $|\mathcal{I}| - 1$, contrary to the minimality of $\mathcal{I}$. This proves (b). Moreover, since $\text{ipp}(G) = |V(G)| - \nu(G)$, (c) follows immediately from (b).

It remains to prove (d). In view of (b), it is enough to show that the set $W := \big\{u \in V(G) \mid u \text{ is } M\text{-saturated for every maximum matching } M \text{ of } G\big\}$ is empty. Suppose toward a contradiction that $W \neq \emptyset$. By (b), we have that $V(G) \setminus W \neq \emptyset$. Since G is connected, there exist vertices $u \in W$ and $v \in V(G) \setminus W$ such that $uv \in E(G)$. Since $v \notin W$, there exists a maximum matching M_v of G such that v is M_v-unsaturated. Then, since $u \in W$, there exists a vertex $u' \in V(G) \setminus \{u, v\}$ such that $uu' \in M_v$. But then by (a), we have that $u'v \in E(G)$, and it follows that $M_u := \big(M_v \setminus \{uu'\}\big) \cup \{u'v\}$ is a maximum matching of G such that u is M_u-unsaturated, contrary to the fact that $u \in W$. □

Proposition 2.10. *Let G be a connected graph on an odd number of vertices and satisfying $ipp(G) = |V(G)| - \nu(G)$. Then G is a chordal and diamond-free.*

Proof. First of all, we note that Proposition 2.9(b) guarantees that $\text{ipp}(G) = \nu(G) + 1$.

Claim 2.11. *G is diamond-free.*

Proof of Claim 2.11. Suppose otherwise, and fix an induced diamond $\mathcal{D}$ in G, where $V(\mathcal{D}) = \{a, b, c, d\}$ and $E(\mathcal{D}) = \{ab, bc, cd, da, bd\}$. Using Proposition 2.9(d), we fix a maximum matching M in G such that d is the unique M-unsaturated vertex of G. (So, a, b, and c are all M-saturated.) Clearly, $M \cup \{d\}$ is an IPP of G. Since d is M-unsaturated, we see that $E(\mathcal{D}) \cap M \subseteq \{ab, bc\}$. But since edges ab and bc share an endpoint, the matching M contains at most one of them. So, by symmetry, we may assume that either $E(\mathcal{D}) \cap M = \{ab\}$ or $E(\mathcal{D}) \cap M = \emptyset$. In each case, we will derive a contradiction by exhibiting an IPP of G of size $\nu(G)$, contrary to the fact that $\text{ipp}(G) = \nu(G) + 1$.

Suppose first that $E(\mathcal{D}) \cap M = \{ab\}$. Then, since c is M-saturated, there exists some $c' \in V(G) \setminus V(\mathcal{D})$ such that $cc' \in M$. Since d is M-unsaturated and adjacent to c, Proposition 2.9(a) guarantees that $dc' \in E(G)$. But then $\big(M \setminus \{ab, cc'\}\big) \cup \{abc, dc'\}$ is an IPP of G of size $|M| = \nu(G)$, a contradiction.

Suppose now that $E(\mathcal{D}) \cap M = \emptyset$. Then, since a, b, and c are M-saturated, there exist pairwise distinct vertices $a', b', c' \in V(G) \setminus V(\mathcal{D})$ such that $aa', bb', cc' \in M$. Since d is M-unsaturated and adjacent to a, b, and c, Proposition 2.9(a) guarantees that $da', db', dc' \in E(G)$. If $a'b' \notin E(G)$, then $\big(M \setminus \{aa', bb'\}\big) \cup \{a'db', ab\}$ is an IPP of G of size $|M| = \nu(G)$, a contradiction. So, $a'b' \in E(G)$. But now $\big(M \setminus \{aa', bb', cc'\}\big) \cup \{a'b', abc, dc'\}$ is an IPP of G of size $|M| = \nu(G)$, again a contradiction. ♦

It now remains to show that G is chordal. Suppose otherwise, and fix an induced cycle $C = c_0c_1 \dots c_{k-1}c_0$ (with $k \geq 4$ and with indices taken modulo k) of G.

Claim 2.12. *For every maximum matching M of G, either all vertices of C are M-saturated or $E(C) \cap M = \emptyset$.*

Proof of Claim 2.12. Suppose otherwise, and fix a maximum matching M of G such that for some $i, j \in \{0, 1, \dots, k-1\}$, c_i is M-unsaturated and $c_jc_{j+1} \in M$. (Clearly, $i \notin \{j, j+1\}$). We may assume that M, c_i, and c_jc_{j+1} were chosen so that the path $P := c_ic_{i+1} \dots c_j$ is as short as possible. By symmetry, we may assume that $i = 0$ (and consequently, $j \in \{1, \dots, k-2\}$). Proposition 2.9(b) now guarantees that c_0 is the unique M-unsaturated vertex of G, and in particular, c_1 is M-saturated. So, there exists some $c_1' \in V(G) \setminus \{c_1\}$ such that $c_1c_1' \in M$. Since c_0 is M-unsaturated and adjacent to c_1, Proposition 2.9(a) guarantees that $c_0c_1' \in E(G)$. Now $\{c_0, c_1, c_1'\}$ induces a triangle in G; since $C = c_0c_1 \dots c_{k-1}c_0$ is an induced cycle of G of length $k \geq 4$, we see that $c_1' \notin V(C)$, and in particular, $j \neq 1$. But now $M_1 := \big(M \setminus \{c_1c_1'\}\big) \cup \{c_0c_1'\}$ is a maximum matching of G, and c_1 is M_1-unsaturated. Meanwhile, we have that $c_jc_{j+1} \in M_1$, and the path $P_1 := c_1c_2 \dots c_j$ is shorter than the path $P = c_0c_1 \dots c_j$, contrary to the minimality of P. ♦

Using Proposition 2.9(d), we fix a maximum matching M of G such that c_0 is the unique M-unsaturated vertex of G. By Claim 2.12, we have that $E(C) \cap M = \emptyset$. We now deduce that there exist pairwise distinct vertices $c_1', \dots, c_{k-1}' \in V(G) \setminus V(C)$ such that $c_1c_1', \dots, c_{k-1}c_{k-1}' \in M$.

Claim 2.13. *$c_0c'_1, c_0c'_{k-1} \in E(G)$ and $c_{k-2}c'_{k-1} \notin E(G)$.*

Proof of Claim 2.13. Since the M-unsaturated vertex c_0 is adjacent to the endpoint c_1 of the matching edge $c_1c'_1 \in M$, Proposition 2.9(a) guarantees that $c_0c'_1 \in E(G)$. Analogously, $c_0c'_{k-1} \in E(G)$. Finally, we have that $c_{k-2}c'_{k-1} \notin E(G)$, for otherwise, vertices $c_0, c_{k-2}, c_{k-1}, c'_{k-1}$ would induce a diamond in G, contrary to the fact that, by Claim 2.11, G is diamond-free. ♦

By Claim 2.13, we have that $c_{k-2}c'_{k-1} \notin E(G)$. Now, fix the smallest index $i \in \{1, \ldots, k-2\}$ such that $c_ic'_{i+1} \notin M$. By the minimality of i, and by Claim 2.13, we have that $c_jc'_{j+1} \in E(G)$ for all $j \in \{0, \ldots, i-1\}$. But now $\mathcal{I} := \{c_jc'_{j+1} \mid 0 \leq j \leq i-1\} \cup \big(M \setminus \{c_jc'_j \mid 1 \leq j \leq i+1\}\big) \cup \{c_ic_{i+1}c'_{i+1}\}$ is an IPP of G of size $\nu(G)$, contrary to the fact that $\mathrm{ipp}(G) = \nu(G)+1$. □

3 Concluding remarks

Theorem 1.3 gives a characterization of graphs on an odd number of vertices whose isometric path partition number achieves the upper bound proved in Proposition 1.1. Proving a similar characterization for even order graphs is challenging, as we can find several diverse graphs on an even number of vertices whose isometric path partition number achieves the same upper bound. Some examples of such graphs are given in Figure 1.

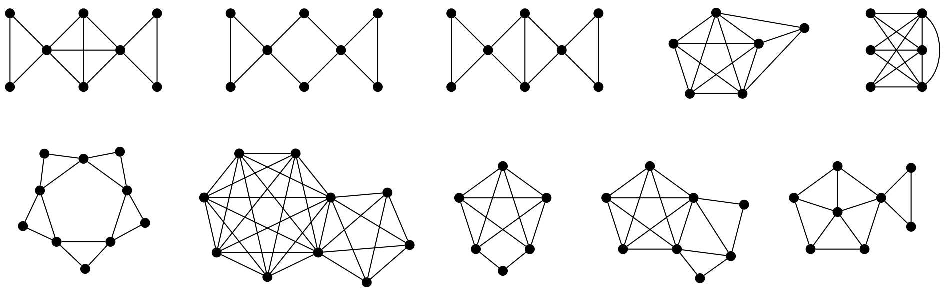


Figure 1: Examples of graphs G that have an even number of vertices and satisfy $ipp(G) = |V(G)| - \nu(G)$.

Acknowledgment

We would like to thank Florent Foucaud for the initial discussions on this problem.

References

[1] H.-J. Bandelt and H. M. Mulder. Distance-hereditary graphs. *Journal of Combinatorial Theory, Series B*, 41(2):182–208, 1986.

[2] C. Berge. Path partitions in directed graphs. In *North-Holland Mathematics Studies*, volume 75, pages 59–63. Elsevier, 1983.

[3] F. T. Boesch and J. F. Gimpel. Covering points of a digraph with point-disjoint paths and its application to code optimization. *Journal of the ACM (JACM)*, 24(2):192–198, 1977.

[4] D. Chakraborty, A. Dailly, S. Das, F. Foucaud, H. Gahlawat, and S. K. Ghosh. Complexity and Algorithms for isometric path cover on Chordal Graphs and Beyond. In *33rd International Symposium on Algorithms and Computation (ISAAC 2022)*, volume 248 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 12:1–12:17, 2022.

[5] D. Chakraborty, H. Muller, S. Ordyniak, F. Panolan, and M. Rychlicki. Covering and partitioning of split, chain and cographs with isometric paths. *Leibniz International Proceedings in Informatics. 49th International Symposium on Mathematical Foundations of Computer Science*, 2024.

[6] G. Chartrand, J. McCanna, N. Sherwani, M. Hossain, and J. Hashmi. The induced path number of bipartite graphs. *Ars combinatoria*, 37:191–208, 1994.

[7] D. G. Corneil, B. Dalton, and M. Habib. Ldfs-based certifying algorithm for the minimum path cover problem on cocomparability graphs. *SIAM Journal on Computing*, 42(3):792–807, 2013.

[8] H. Fernau, F. Foucaud, K. Mann, U. Padariya, and R. R. KN. Parameterizing path partitions. *Theoretical Computer Science*, 1028:115029, 2025.

[9] D. C. Fisher and S. L. Fitzpatrick. The isometric path number of a graph. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 38:97–110, 2001.

[10] S. L. Fitzpatrick, R. J. Nowakowski, D. Holton, and I. Caines. Covering hypercubes by isometric paths. *Discrete Mathematics*, 240(1):253–260, 2001.

[11] T. Gallai and A. N. Milgram. Verallgemeinerung eines graphentheoretischen satzes von rédei. *Acta Sc. Math*, 21:181–186, 1960.

[12] M. R. Garey, D. S. Johnson, and R. E. Tarjan. The planar hamiltonian circuit problem is NP-complete. *SIAM Journal on Computing*, 5(4):704–714, 1976.

[13] M. Lafond and V. Moulton. Path partitions of phylogenetic networks. *Theoretical Computer Science*, 1024:114907, 2025.

[14] P. D. Manuel. On the isometric path partition problem. *Discussiones Mathematicae: Graph Theory*, 41(4):1077–1089, 2021.

[15] S. C. Ntafos and S. L. Hakimi. On path cover problems in digraphs and applications to program testing. *IEEE Transactions on Software Engineering*, (5):520–529, 1979.

[16] J.-J. Pan and G. J. Chang. Induced-path partition on graphs with special blocks. *Theoretical Computer Science*, 370(1):121–130, 2007.

[17] Z. Skupién. Path partitions of vertices and hamiltonicity of graphs. In *Proceedings of the Second Czechoslovakian Symposium on Graph Theory, Prague*, 1974.

[18] D. B. West. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Robustness of the Sauer–Spencer Theorem

(Extended abstract)

Peter Allen* Julia Böttcher† Yoshiharu Kohayakawa‡ Mihir Neve§

Abstract

In this paper, we prove a robust version of a graph embedding theorem of Sauer and Spencer. To state this sparser analogue, we define $G(p)$ to be a random subgraph of G obtained by retaining each edge of G independently with probability $p \in [0,1]$, and let $m_1(H)$ be the maximum 1-density of a graph H. We show that for any constant Δ and $\gamma > 0$, if G is an n-vertex host graph with minimum degree $\delta(G) \geq (1 - 1/2\Delta + \gamma)n$ and H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$, then for $p \geq Cn^{-1/m_1(H)} \log n$, the random subgraph $G(p)$ contains a copy of H with high probability. Our value for p is optimal up to a log-factor.

In fact, we prove this result for a more general minimum degree condition on the host graph G by introducing an *extension threshold* $\delta_{\mathrm{e}}(\Delta)$, such that the above result holds for graphs G with $\delta(G) \geq (\delta_{\mathrm{e}}(\Delta) + \gamma)n$. We show that $\delta_{\mathrm{e}}(\Delta) \leq (2\Delta - 1)/2\Delta$, and further conjecture that $\delta_{\mathrm{e}}(\Delta) = \Delta/(\Delta + 1)$, which matches the minimum degree condition on G in the Bollobás–Eldridge–Catlin Conjecture. Our main tool is a vertex-spread version of the blow-up lemma of Allen, Böttcher, Hàn, Kohayakawa, and Person, which we believe to be of independent interest.

1 Introduction

Embedding problems have been of significant interest in graph theory. A graph H is said to have an *embedding* into a host graph G if there exists an injective map from $V(H)$ to $V(G)$ which maps every edge of H to some edge in G. For such problems, one is usually interested

*Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK. E-mail: p.d.allen@lse.ac.uk.

†Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK. E-mail: j.boettcher@lse.ac.uk.

‡Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, Brazil. E-mail: yoshi@ime.usp.br. Partially supported by FAPESP (2023/03167-5) and CNPq (407970/2023-1, 315258/2023-3).

§Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK. E-mail: m.s.neve@lse.ac.uk.

in finding sufficiency conditions on the host graph, such as bounds on its minimum degree, that allow for embedding a given family of graphs.

A prominent result of this flavour is Dirac's theorem [12], which states that any graph on $n \geq 3$ vertices with minimum degree at least $n/2$ contains a Hamilton cycle. Moving to different spanning subgraphs than cycles, Bollobás in 1978 [7] conjectured that if G is an n-vertex graph with minimum degree at least $(1/2+\varepsilon)n$, then G contains any spanning tree of bounded constant degree Δ for any $\varepsilon > 0$ and $n \geq n_0(\varepsilon)$. This was shown to be true by Komlós, Sárközy, and Szemerédi in 1995 [18]. In fact, they later improved their result to show that it is possible to find any spanning tree of degree $\mathcal{O}(n/\log n)$ under a similar minimum degree condition [22].

Such embedding results have been studied for several other families of subgraphs, such as clique-factors [14], F-factors [21], powers of Hamilton cycles [20], and for subgraphs with bounded degree and sublinear bandwidth [8]. A common link in all these embedding theorems is that the subgraph has sublinear bandwidth. However, the question of which minimum degree enforces the appearance of all spanning subgraphs of a given maximum degree still remains open. Observe that this class of graphs also contains expanders, which do not have sublinear bandwidth. Bollobás, Eldridge, and Catlin [6, 9] made the following conjecture.

Conjecture 1.1 (Bollobás–Eldridge–Catlin Conjecture)**.** *Let $\Delta \in \mathbb{N}$ be a positive integer. Suppose G is an n-vertex graph with minimum degree $\delta(G) \geq \frac{\Delta}{\Delta+1}n$ and H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$. Then H is a spanning subgraph of G.*

This conjecture is known to be tight, as is shown by a slightly unbalanced complete $(\Delta+1)$-partite graph, which does not contain a $K_{\Delta+1}$-factor. There have been several partial results towards the resolution of this conjecture. For instance, this conjecture is known to be true for $\Delta \leq 2$ [1, 4]; for the case where $\Delta = 3$ and n is large [11]; and when H is a bipartite graph [10]. To this date, however, the best known bound on the minimum degree that enforces all spanning subgraphs with maximum degree Δ is given by the following old and well-known theorem of Sauer and Spencer [26].

Theorem 1.2 (Sauer–Spencer Theorem)**.** *Let $\Delta > 0$ be given. Suppose that G is an n-vertex graph with minimum degree $\delta(G) \geq \frac{2\Delta-1}{2\Delta}n$ and H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$. Then H is a spanning subgraph of G.*

In general, embedding theorems such as the ones mentioned so far allow for embedding subgraphs into dense graphs, that is, having $\Theta(n^2)$ edges. A popular line of research in the past few decades has been to establish sparse analogues of such results. In such a situation, it is typical to use the high minimum degree host graph from these Dirac-type theorems as a template, and ask whether a sparse spanning subgraph of this host graph will inherit its subgraph embedding properties.

This is captured by the notion of robustness, which was proposed by Krivelevich, Lee, and Sudakov [23] in their study of Hamiltonicity. A robustness result typically examines how strongly a graph satisfies a property. To make this more precise, given a graph G and $p \in [0,1]$, let $G(p)$ be the random subgraph of G obtained by retaining edges of G independently with probability p. Further, we say that a graph property holds *with high probability* if the probability of this event tends to 1 as the size of the graph, $|G| = n$, grows to infinity.

Now suppose the graph G satisfies some property Π. Then this property Π is said to hold *robustly* for G for some $p < 1$, if the random subgraph $G(p)$ satisfies the property Π with high probability . For instance, Krivelevich, Lee, and Sudakov [23] proved that Hamiltonicity is robust for Dirac graphs. They showed that if G is an n-vertex graph with minimum degree $n/2$ and $p \geq Cn^{-1} \log n$, then $G(p)$ is Hamiltonian with high probability.

One motivation for establishing robustness results is that they immediately imply counting results on the number of embeddings in the dense host graph. In the case of robust embedding theorems with tight minimum degree conditions, these counting results reveal a discontinuous behaviour of the number of copies of the subgraph in question. For indeed, below this minimum degree threshold (say $n/2$ for Hamiltonicity), the host graph might not have any copy of the fixed subgraph, but as soon as the minimum degree condition is met, there is a sudden emergence of many copies of this subgraph in the host graph.

Robustness results have been previously studied for some graph and hypergraph properties. For instance, robust versions are known for the Hajnal–Szemerédi condition for containing clique factors [2, 25]; for the Dirac threshold for containing bounded degree trees [5, 25]; for containment of powers of Hamilton cycles [15]; for embedding subgraphs with bounded maximum degree and sublinear bandwidth [3]; and for containment of Hamilton ℓ-cycles [17] and perfect matchings [16, 25] in hypergraphs. While only a few results on robustness, as defined above, are known, we refer the interested reader to a survey by Sudakov [27] on related questions of a similar flavour.

2 Our Results

Our main result is the following robust version of the Sauer–Spencer theorem (Theorem 1.2). The value of p that we shall be working with is given by the *maximum* 1*-density* $m_1(H)$ of the graph H we are embedding, which is defined as follows. Let $d_1(H) := e(H)/(v(H) - 1)$ denote the 1*-density* of H, and let $m_1(H) := \max_{H' \subset H} d_1(H')$, where the maximum runs over all subgraphs of H with at least two vertices.

Theorem 2.1 (Robust Sauer–Spencer Theorem)**.** *For all $\gamma > 0$ and $\Delta \in \mathbb{N}$, there is a constant $C > 0$ such that if H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$ and G is an n-vertex graph with minimum degree $\delta(G) \geq (\frac{2\Delta-1}{2\Delta} + \gamma)n$, then for $p \geq Cn^{-1/m_1(H)} \log n$, with high probability, the graph H is a subgraph of $G(p)$.*

Our value for p in Theorem 2.1 is optimal up to the log factor. For the proof of Theorem 2.1, we prove and use a stronger version of Theorem 1.2, where we show that, under similar conditions, it is possible to obtain an embedding of H into G by *extending* a given partial embedding of a few vertices of H into G. In fact, as this ability to extend is the only role of Theorem 1.2 in our proof, we are able to prove a more general version of Theorem 2.1. For this, we need the following definition, which captures the above extension property.

Definition 2.2 (Extension Threshold)**.** For all $\Delta \in \mathbb{N}$, let the *extension threshold* $\delta_{\mathrm{e}}(\Delta)$ be the smallest real number δ_{e} such that the following holds. For every $\gamma > 0$ there exists $\eta > 0$ such that for all sufficiently large n, we have that if G and H are n-vertex graphs with $\delta(G) \geq (\delta_{\mathrm{e}} + \gamma)n$ and with $\Delta(H) \leq \Delta$, and if $S \subset V(H)$ is a subset of size $|S| \leq \eta n$ given with a partial embedding $\varphi_S : S \to V(G)$ of $H[S]$ into G, then there is an embedding $\varphi : V(H) \to V(G)$ of H into G that extends φ_S.

The tightness of Conjecture 1.1 implies that $\delta_{\mathrm{e}}(\Delta) \geq \Delta/(\Delta+1)$. On the other hand, our stronger version of the Sauer–Spencer theorem, as mentioned above, establishes the upper bound $\delta_{\mathrm{e}}(\Delta) \leq (2\Delta-1)/2\Delta$. We prove the following result, which implies Theorem 2.1.

Theorem 2.3. *For all $\gamma > 0$ and $\Delta \in \mathbb{N}$, there exists a constant $C > 0$ such that if H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$ and G is an n-vertex graph with minimum degree $\delta(G) \geq (\delta_{\mathrm{e}}(\Delta)+\gamma)n$, then for $p > Cn^{-1/m_1(H)} \log n$, the graph H is a subgraph of $G(p)$ with high probability.*

2.1 A Note on Optimality

It might be possible to improve Theorem 2.3 in two directions. First, we believe that if Conjecture 1.1 holds true, then it should do so robustly. Hence, we conjecture that the value of $\delta_{\mathrm{e}}(\Delta)$ should match the minimum degree condition in Conjecture 1.1 and equal $\Delta/(\Delta+1)$, in which case our bound on the minimum degree for G is asymptotically optimal.

The second direction concerns our range of values for p. A simple lower bound on the values of p which work in Theorem 2.3 is given by the containment threshold for the graph H in the Erdős–Rényi random graph $G(n,p) \coloneqq K_n(p)$. This lower bound is often used as a measure for the degree of robustness of a property Π: we say Π is *extremely robust* for G if the threshold for $G(p)$ to satisfy Π asymptotically matches the threshold for $G(n,p)$ to satisfy Π (see [23, 25]). For instance, the result in [23] says that Hamiltonicity is extremely robust for Dirac graphs, as the threshold for Hamiltonicity in $G(n,p)$ is $\Theta(\log n/n)$.

In the case of Theorem 2.3, the exact threshold for embedding a subgraph H is not entirely known. However, for any given value of $m_1(H)$, there exist graphs H for which the threshold for the containment of H in $G(n,p)$ is bounded below by $p = Cn^{-1/m_1(H)}$. Thus, over the class of graphs with $\Delta(H) \leq \Delta$ and with a fixed maximum 1-density $m_1(H)$, our value for p, and hence the strength of our robustness result, is optimal up to the log factor.

Further, it is not hard to check that among graphs H with maximum degree $\Delta \geq 2$, the quantity $m_1(H)$ is maximised by $K_{\Delta+1}$ (albeit not uniquely), which has $m_1(H) = (\Delta+1)/2$. In fact, for all H with $m_1(H) = (\Delta+1)/2$, we are able to work with the slightly better probability $Cn^{-2/(\Delta+1)}(\log n)^{1/\binom{\Delta+1}{2}}$, which is known to be optimal up to the constant [13].

3 Some Ideas in the Proof of Theorem 2.3

In order to obtain a near-optimal value of p in Theorem 2.3, we use the notions of spreadness and spread measures, as were introduced by Talagrand [28]. In their proof of the fractional version of the Kahn–Kalai conjecture, Frankston, Kahn, Narayanan, and Park [13] established a relation between the existence of such a spread measure and the threshold function for a graph property (which in this case is the best value of p that works for Theorem 2.3).

Thus, as a first step in our proof, we show the existence of a so-called vertex-spread measure, which implies the required spread measure that is sufficient for proving Theorem 2.3 for the said value of p. Such applications of spread measures have been common in proving robustness results [17, 24]. In fact, in 2023, Pham, Sah, Sawhney, and Simkin [25] proved a series of robustness results highlighting the use of spreadness in such proofs and various methods for constructing spread measures.

One of our main tools, which we use to construct this vertex-spread measure, is a spread version of the blow-up lemma of Allen, Böttcher, Hàn, Kohayakawa, and Person [3]. We believe this result will be useful to approach other problems as well. The proof of our spread blow-up lemma results from a careful probabilistic analysis of the key steps used in the proof of the sparse blow-up lemma in [3].

It should be mentioned that a spread version of the blow-up lemma of Komlós, Sárközy, and Szemerédi [19] was proved recently by Nenadov and Pham [24], but their result is not strong enough for our proof of Theorem 2.1. The main difference is that their spread blow-up lemma applies only to a bounded-size subgraph of the reduced graph of a regular partition for G, while our version applies to the entire reduced graph.

The final step of our proof involves pre-processing and partitioning the given graphs G and H, so as to satisfy the sufficient conditions for applying our spread blow-up lemma. These conditions are identical to those required for the dense blow-up lemma of [3], and we refer the interested reader to [3, Section 7.1] for these conditions and the relevant definitions. While most of the methods used in this step are standard, we need to use the extension property, as specified in Definition 2.2, to be able to partition the graph H in agreement with the requirements of our spread blow-up lemma. This completes our outline for the proof of Theorem 2.3.

References

[1] M. Aigner and S. Brandt, *Embedding arbitrary graphs of maximum degree two*, J. London Math. Soc. (2) **48** (1993), no. 1, 39–51.

[2] P. Allen, J. Böttcher, J. Corsten, E. Davies, M. Jenssen, P. Morris, B. Roberts, and J. Skokan, *A robust Corrádi-Hajnal theorem*, Random Structures Algorithms **65** (2024), no. 1, 61–130.

[3] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, 2019. Available as arXiv:1612.00622.

[4] N. Alon and E. Fischer, 2*-factors in dense graphs*, Discrete Math. **152** (1996), no. 1-3, 13–23.

[5] P. Bastide, C. Legrand-Duchesne, and A. Müyesser, *Random embeddings of bounded degree trees with optimal spread*, 2024. Available as arXiv:2409.06640.

[6] B. Bollobás and S. E. Eldridge, *Packings of graphs and applications to computational complexity*, J. Combin. Theory Ser. B **25** (1978), no. 2, 105–124.

[7] B. Bollobás, *Extremal graph theory*, London Mathematical Society Monographs, vol. 11, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978.

[8] J. Böttcher, M. Schacht, and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Math. Ann. **343** (2009), no. 1, 175–205.

[9] P. A. Catlin, *Embedding subgraphs and coloring graphs under extremal degree conditions*, ProQuest LLC, Ann Arbor, MI, 1976, Thesis (Ph.D.)–The Ohio State University.

[10] B. Csaba, *On the Bollobás–Eldridge conjecture for bipartite graphs*, Combin. Probab. Comput. **16** (2007), no. 5, 661–691.

[11] B. Csaba, A. Shokoufandeh, and E. Szemerédi, *Proof of a conjecture of Bollobás and Eldridge for graphs of maximum degree three*, Combinatorica **23** (2003), no. 1, 35–72, Paul Erdős and his mathematics (Budapest, 1999).

[12] G. A. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc. (3) **2** (1952), 69–81.

[13] K. Frankston, J. Kahn, B. Narayanan, and J. Park, *Thresholds versus fractional expectation-thresholds*, Ann. of Math. (2) **194** (2021), no. 2, 475–495.

[14] A. Hajnal and E. Szemerédi, *Proof of a conjecture of P. Erdős*, Combinatorial theory and its applications, I-III (Proc. Colloq., Balatonfüred, 1969), Colloq. Math. Soc. János Bolyai, vol. 4, North-Holland, Amsterdam-London, 1970, pp. 601–623.

[15] F. Joos, R. Lang, and N. Sanhueza-Matamala, *Robust hamiltonicity*, 2024. Available as arXiv:2312.15262.

[16] D. Y. Kang, T. Kelly, D. Kühn, D. Osthus, and V. Pfenninger, *Perfect matchings in random sparsifications of Dirac hypergraphs*, Combinatorica **44** (2024), no. 6, 1233–1266.

[17] T. Kelly, A. Müyesser, and A. Pokrovskiy, *Optimal spread for spanning subgraphs of Dirac hypergraphs*, J. Combin. Theory Ser. B **169** (2024), 507–541. MR 4790883

[18] J. Komlós, G. N. Sárközy, and E. Szemerédi, *Proof of a packing conjecture of Bollobás*, Combin. Probab. Comput. **4** (1995), no. 3, 241–255.

[19] ———, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123.

[20] ———, *Proof of the Seymour conjecture for large graphs*, Ann. Comb. **2** (1998), no. 1, 43–60.

[21] ———, *Proof of the Alon–Yuster conjecture*, Discrete Math. **235** (2001), no. 1-3, 255–269, Combinatorics (Prague, 1998).

[22] ———, *Spanning trees in dense graphs*, Combin. Probab. Comput. **10** (2001), no. 5, 397–416. MR 1869048

[23] M. Krivelevich, C. Lee, and B. Sudakov, *Robust Hamiltonicity of Dirac graphs*, Trans. Amer. Math. Soc. **366** (2014), no. 6, 3095–3130.

[24] R. Nenadov and H. T. Pham, *Spread blow-up lemma with an application to perturbed random graphs*, 2024. Available as arXiv:2410.06132.

[25] H. T. Pham, A. Sah, M. Sawhney, and M. Simkin, *A toolkit for robust thresholds*, 2023. Available as arXiv:2210.03064.

[26] N. Sauer and J. Spencer, *Edge disjoint placement of graphs*, J. Combin. Theory Ser. B **25** (1978), no. 3, 295–302.

[27] B. Sudakov, *Robustness of graph properties*, Surveys in combinatorics 2017, London Math. Soc. Lecture Note Ser., vol. 440, Cambridge Univ. Press, Cambridge, 2017, pp. 372–408.

[28] M. Talagrand, *Are many small sets explicitly small?*, STOC '10—Proceedings of the 2010 ACM International Symposium on Theory of Computing, ACM, New York, 2010, pp. 13–35.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Embedding trees using minimum and maximum degree conditions

(Extended abstract)

Alexey Pokrovskiy* Leo Versteegen† Ella Williams‡

Abstract

A variant of the notable Erdős-Sós conjecture, posed by Havet, Reed, Stein and Wood, states that every graph with minimum degree at least $2k/3$ and maximum degree at least k contains a copy of every tree with k edges. For trees of bounded maximum degree, we prove an approximate version of this, building on results already known in the dense setting due to Besomi, Pavez-Signé and Stein. We also prove similar results for related conjectures, where alternative degree combinations are considered.

1 Introduction

Which degree conditions can be imposed on a host graph in order to guarantee that it contains a copy of every tree of a fixed size? A standard observation is that a graph of minimum degree at least k contains a copy of every k-edge tree, seen by greedily embedding the vertices of the tree one-by-one, using a degeneracy ordering. This bound cannot be lowered since there must exist a vertex of degree at least k in order to embed the k-edge star, or alternatively, to avoid the components of the graph having fewer than $k+1$ vertices, in which case no k-edge tree can be embedded. Along similar lines to this observation, one

*Department of Mathematics, University College London, UK. E-mail: a.pokrovskiy@ucl.ac.uk.

†Department of Mathematics, London School of Economics and Political Science, UK. E-mail: lversteegen.math@gmail.com. Supported by the London Mathematical Society and the Heilbronn Institute for Mathematical Research through an Early Career Fellowship.

‡Department of Mathematics, University College London, UK. E-mail: ella.williams.23@ucl.ac.uk. Supported by the Martingale Foundation.

can instead ask what happens when we consider alternative degree conditions within the host graph. A notable conjecture of Erdős and Sós from the 1960s (see [6]) is a central point for much research in the area. It states that every graph with average degree strictly larger than $k-1$ contains every k-edge tree. A few decades ago, Ajtai, Komlós, Simonovits and Szemerédi announced a proof of this conjecture for large k. Although a publication of their result has not yet appeared, some of the main ideas have been communicated [1], and in the meantime several partial results have been obtained (see e.g. [4, 5, 10, 14]). As a natural adaptation of the Erdős–Sós conjecture, Havet, Reed, Stein and Wood [7] considered a combination of minimum and maximum degree requirements for the host graph, and conjectured the following.

Conjecture 1.1 (Havet, Reed, Stein and Wood [7]). *Every graph with minimum degree at least $\lfloor 2k/3 \rfloor$ and maximum degree at least k contains a copy of every k-edge tree.*

Both the minimum and maximum degree conditions in the conjecture are best possible, where the latter is necessary to again ensure there is a vertex of degree at least k. To see that the minimum degree condition cannot be lowered, consider the following construction. Suppose k is divisible by 3 and G is obtained by taking the union of two disjoint cliques, each with $2k/3-1$ vertices, and adding a universal vertex x. Then $\Delta(G) = 4k/3-2$ and $\delta(G) = 2k/3-1$. Let T be a tree containing a vertex v with $d_T(v) = 3$ such that each of the three components of $T - v$ contain exactly $k/3$ vertices (many trees of this type exist by taking any three disjoint trees on $k/3$ vertices and connecting v to one vertex in each of them). No matter which vertex of T we embed at x, all vertices in two of the components of $T - v$ must be embedded into the same clique in G, constituting at least $2k/3$ vertices, but neither clique is big enough.

Reed and Stein [12, 13] proved an exact version of Conjecture 1.1 for large spanning trees. Havet, Reed, Stein and Wood [7] proved an approximate version where the maximum degree of the host graph is at least a given function of k. This was later improved by Besomi, Pavez-Signé and Stein [2] for trees with bounded maximum degree and dense graphs.

Theorem 1.2 (Besomi, Pavez-Signé and Stein [2]). *For all $\varepsilon > 0$ there exists k_0 such that for all $n, k > k_0$ with $n \geq k \geq \varepsilon n$ the following holds. Every graph on n vertices with minimum degree at least $(2/3+\varepsilon)k$ and maximum degree at least $(1+\varepsilon)k$ contains a copy of every k-edge tree T with $\Delta(T) \leq k^{1/49}$.*

The same three authors also considered more combinations of minimum and maximum degree conditions, posing the following conjecture.

Conjecture 1.3 (Besomi, Pavez-Signé and Stein [2]). *Let $k \in \mathbb{N}$ and let $\alpha \in (0, 1/2]$. Every graph with minimum degree at least $(1+\alpha)k/2$ and maximum degree at least $2(1-\alpha)k$ contains a copy of every k-edge tree.*

We remark that for the case where $\alpha \geq 1/3$, this conjecture is implied by Conjecture 1.1, and thus the bound on maximum degree could be lowered if Conjecture 1.1 were true. On the other hand, for every odd integer $\ell \geq 5$, Besomi, Pavez-Signé and Stein [3] showed that the bounds are asymptotically best possible when $\alpha = 1/\ell$. In favour of Conjecture 1.3, an approximate version has been proven, again in the dense setting and for trees with bounded maximum degree.

Theorem 1.4 (Besomi, Pavez-Signé and Stein [3]). *For all $\varepsilon > 0$ there exists k_0 such that for all $n, k > k_0$ with $n \geq k \geq \varepsilon n$ and for each $\alpha \in (0, 1/2]$ the following holds. Every graph on n vertices with minimum degree at least $(1+\varepsilon)(1+\alpha)k/2$ and maximum degree at least $2(1+\varepsilon)(1-\alpha)k$ contains a copy of every k-edge tree T with $\Delta(T) \leq k^{1/67}$.*

Conjecture 1.3 was originally also posed for the case where $\alpha = 0$, that is, when the host graph has minimum degree at least $k/2$ and maximum degree at least $2k$. However, a counterexample to this was given by Hyde and Reed [8], who found a k-edge tree with maximum degree $(k-3)/2$ that cannot be embedded into a host graph with the given degree conditions. The host graph is constructed using a random graph. Conversely, in the same paper they also proved that minimum degree $k/2$ is sufficient for embedding all k-edge trees provided that the host graph has maximum degree at least $f(k)$ where f is some large function of k. It was commented there that taking Conjecture 1.3 with $\alpha = 0$ may be true if the inequalities are made strict in the statement. Furthermore, since the tree in the counterexample has high maximum degree, one could ask what happens if we restrict the setting only to bounded degree trees. In this direction, Besomi, Pavez-Signé and Stein suggested a stronger statement could hold for trees with maximum degree bounded by a constant.

Conjecture 1.5 (Besomi, Pavez-Signé and Stein [2]). *Let $k, \Delta \in \mathbb{N}$. Every graph with minimum degree at least $k/2$ and maximum degree at least $2(1-1/\Delta)k$ contains a copy of every k-edge tree T with $\Delta(T) \leq \Delta$.*

In [2] an example is provided showing that the bounds in Conjecture 1.5 are close to best possible. In the same paper, as evidence towards the conjecture holding, an approximate version is given again in the dense setting.

Theorem 1.6 (Besomi, Pavez-Signé and Stein [2]). *For all $\varepsilon > 0$ and $\Delta \geq 2$ there exists k_0 such that for all $n, k > k_0$ with $n \geq k \geq \varepsilon n$ the following holds. Every graph on n vertices with minimum degree at least $(1+\varepsilon)k/2$ and maximum degree at least $2(1-1/\Delta+\varepsilon)k$ contains a copy of every k-edge tree T with $\Delta(T) \leq \Delta$.*

Let us now state our main result.

Theorem 1.7. *Let $\alpha \in (0, 1/3)$. For all $\varepsilon > 0$ and $\Delta \in \mathbb{N}$, there exists k_0 such that for all $k > k_0$ the following holds. If G is a graph satisfying one of the following:*

(i) *$\delta(G) \geq (1+\varepsilon)2k/3$ and $\Delta(G) \geq (1+\varepsilon)k$;*

(ii) *$\delta(G) \geq (1+\varepsilon)(1+\alpha)k/2$ and $\Delta(G) \geq 2(1+\varepsilon)(1-\alpha)k$;*

(iii) *$\delta(G) \geq (1+\varepsilon)k/2$ and $\Delta(G) \geq 2(1+\varepsilon)(1-1/\Delta)k$;*

then G contains a copy of every tree k-edge T with $\Delta(T) \leq \Delta$.

For large trees with maximum degree bounded by a constant, property (i) of the theorem asymptotically proves Conjecture 1.1. Similarly, for this family of trees, property (ii) asymptotically proves Conjecture 1.3, noting that the case where $\alpha \geq 1/3$ is covered in a stronger sense by (i). Finally property (iii) gives an asymptotic proof of Conjecture 1.5.

2 Extending to the sparse setting

We remark that everything previously known about Conjectures 1.1, 1.3 and 1.5 holds only for dense graphs, that is, when $|G| = O(k)$. Thus in order to prove Theorem 1.7, it would be helpful if for a sparse graph G, we could reduce to the dense setting. More specifically, we would like to find a small subgraph in G that maintains both a high minimum and high maximum degree. We could then apply the corresponding dense result to this subgraph, in order to find a copy of every bounded degree tree.

The heart of our proof is the following structural lemma, which says that for every $v \in V(G)$, there are two vertex-disjoint subgraphs in G, that are small and have high minimum degree, to which v sends almost all of its neighbours.

Lemma 2.1. *For all $\varepsilon, \eta > 0$ and $\Delta \in \mathbb{N}$, there exists $k_0 \in \mathbb{N}$ such that for all $k > k_0$ the following holds. Let $a \geq 1/2$ and let G be a graph with $\delta(G) \geq (1+\varepsilon)ak$. If there exists a k-edge tree T with $\Delta(T) \leq \Delta$ such that T does not embed in G, then G contains vertex-disjoint subgraphs $C_1, \dots, C_m$ such that*

(I) *$\delta(C_i) \geq (1+\varepsilon/4)ak$ for all $i \in [m]$;*

(II) *$|C_i| < 100k$ for all $i \in [m]$; and*

(III) *for all $v \in V(G)$ there exist $i, j \in [m]$ such that $|N(v) \setminus (V(C_i) \cup V(C_j))| < \eta k$.*

It is not too hard to deduce Theorem 1.7 from Lemma 2.1.

Proof of Theorem 1.7 from Lemma 2.1. Let $\alpha \in (0, 1/3)$, $\varepsilon > 0$, $\Delta \in \mathbb{N}$ and consider a pair $(a, b) \in \{(2/3, 1), ((1+\alpha)/2, 2(1-\alpha)), (1/2, 2(1-1/\Delta))\}$. Let k_0 be sufficiently large, $k > k_0$, and T be a k-edge tree with $\Delta(T) \leq \Delta$. Suppose that G is a graph with $\delta(G) \geq (1+\varepsilon)ak$ and $\Delta(G) \geq (1+\varepsilon)bk$ such that T does not embed in G, noting that the cases (i)–(iii) of Theorem 1.7 for G are covered by the three choices for the pair (a, b). Apply Lemma 2.1 to G with ε and $\varepsilon/2$ playing the roles of ε and η respectively to obtain vertex-disjoint subgraphs $C_1, \dots, C_m$ satisfying the properties of the statement. Let $v \in V(G)$ be chosen such that $d_G(v) = \Delta(G)$, and let $i, j \in [m]$ denote the indices for which $|N(v) \setminus (V(C_i) \cup V(C_j))| < \varepsilon k/2$. Consider the subgraph $G' = G[V(C_i) \cup V(C_j) \cup \{v\}]$. We have $\Delta(G') \geq |N(v) \cap (V(C_i) \cup V(C_j))| \geq \Delta(G) - \varepsilon k/2 > (1+\varepsilon/4)bk$. Furthermore by (I), both $\delta(C_i) \geq (1+\varepsilon/4)ak$ and $\delta(C_j) \geq (1+\varepsilon/4)ak$ hold, so that $\delta(G') \geq (1+\varepsilon/4)ak$. Property (II) implies $|G'| = |C_i| + |C_j| + 1 < 200k < 4\varepsilon^{-1}k$. Observe that the choice of (a, b) as $(2/3, 1), ((1 + \alpha)/2, 2(1 - \alpha))$ or $(1/2, 2(1 - 1/\Delta))$ imply that G' satisfies the conditions of Theorem 1.2, Theorem 1.4 or Theorem 1.6 respectively, with $\varepsilon/4$ playing the role of ε. Applying the corresponding result, we deduce that T embeds in G', and thus in G, a contradiction. $\square$

In order to prove Lemma 2.1, we first aim to show that a collection of vertex-disjoint subgraphs each with minimum degree at least $(1 + \varepsilon/2)ak$ and size less than $100k$ can be found in G, so that additionally their union covers $(1-o(1))|G|$ vertices of G. We note that the minimum degree condition here is slightly larger than that of property (I), which will be useful later in the proof. We use machinery developed by Pokrovskiy [10, 11], providing a helpful decomposition result for all graphs not containing some bounded degree tree of a fixed size. A cover of a graph H is a vertex set $U \subseteq V(H)$ such that every edge in H has at least one endpoint in U. The result roughly states that if G does not contain some k-edge tree with $\Delta(T) \leq \Delta$, then a small number of edges can be deleted from G so that every component in the resulting graph has a cover of size at most $3k$. Having applied this result with suitable parameters, and viewing the obtained components as a collection of vertex-disjoint subgraphs, we make some further refinements to ensure that each subgraph has high minimum degree, deleting few vertices from their union in the process. The cover property is also used to show that each subgraph is small, as otherwise we can find high degree vertices and reduce the problem to the dense setting to embed T via previous results.

Our second aim in proving Lemma 2.1 is to show that property (III) holds for a suitable collection of subgraphs. Note that, it suffices to prove the lemma assuming that $\eta \ll \varepsilon$. We use a maximality argument, whereby the collection $(C_i)_{i\in[m]}$ of vertex-disjoint subgraphs satisfying (I) and (II) is chosen such that the number of vertices covered by their union is as large as possible. We fix a constant $\gamma = \eta/4\Delta$, and for each $i \in [m]$ define the periphery $L_i := \{v \in V(G) : |N(v) \cap V(C_i)| \geq \gamma k\}$, noting that $V(C_i) \subseteq L_i$ by the minimum degree condition. We consider those L_i sets that contain many vertices outside of $V(C_i)$ separately to those which do not expand much, and essentially prove that $\bigcup_{i\in[m]} L_i$ contains all vertices of G, and that $|L_i \setminus V(C_i)| \leq \gamma k$ for every $i \in [m]$.

The main idea here is to show that if there are vertices outside of $\bigcup_{i\in[m]} L_i$ or in the periphery sets that expand a lot, then we can either control where their neighbours lie and embed T, or we can find a subgraph amongst these vertices that has minimum degree at least $(1+\varepsilon/2)ak$. In the latter case we can apply the same logic as earlier (in our first aim) to find a collection of subgraphs satisfying properties (I) and (II), where (I) holds since we had extra space in our minimum degree condition in the argument. When combined appropriately with $(C_i)_{i\in[m]}$, we will obtain a collection covering a larger number of vertices, contradicting the maximality assumption.

Now assuming that $V(G) = \bigcup_{i\in[m]} L_i$, we gain insight into the permissible neighbourhoods of vertices in G by considering several standard arguments about splitting up the tree T into smaller subforests, and using the minimum degree condition within each subgraph (property (I)) to embed these parts greedily. For instance, for every $v \in V(G)$ we show that there are at most 2Δ distinct $\ell \in [m]$ such that v has a neighbour in L_ℓ, as otherwise T embeds in G. Since each set L_ℓ contains at most γk vertices outside of $V(C_\ell)$, then v has at most $2\Delta\gamma k = \eta k/2$ neighbours in the set $\bigcup_{\ell\in[m]}(L_\ell \setminus V(C_\ell))$. Similarly we show that if v has at least Δ neighbours in two distinct subgraphs C_i and C_j, then it cannot have another neighbour in any other C_ℓ from the collection, as otherwise we again have a greedy argument to embed T. So, for a vertex $v \in V(G)$, we pick distinct $i, j \in [m]$ such that $|N(v) \cap V(C_i)|$ and $|N(v) \cap V(C_j)|$ are maximal. If both of these sets have size at least Δ, then v cannot have a neighbour in any other subgraph C_ℓ from the collection. If on the other hand one of these sets contains fewer than Δ vertices, then by choice of i and j, v also sends fewer than Δ vertices into every other C_ℓ from the collection. Since we know there are at most 2Δ subgraphs to which v can send a neighbour, then in either case v has at most $2\Delta^2 < \eta k/2$ vertices in $\bigcup_{\ell\in[m]\setminus\{i,j\}} V(C_\ell)$.

References

[1] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. Exact solution of the Erdős–Sós conjecture. `https://imada.sdu.dk/Research/GT2015/Talks/Slides/simonovits.pdf`, 2015.

[2] G. Besomi, M. Pavez-Signé, and M. Stein. Degree conditions for embedding trees. *SIAM Journal on Discrete Mathematics*, 33(3):1521–1555, 2019.

[3] G. Besomi, M. Pavez-Signé, and M. Stein. Maximum and minimum degree conditions for embedding trees. *SIAM Journal on Discrete Mathematics*, 34(4):2108–2123, 2020.

[4] G. Besomi, M. Pavez-Signé, and M. Stein. On the Erdős–Sós conjecture for trees with bounded degree. *Combinatorics, Probability and Computing*, 30:1–21, 02 2021.

[5] A. Davoodi, D. Piguet, H. Řada, and N. Sanhueza-Matamala. Beyond the Erdős–Sós conjecture. In *European Conference on Combinatorics, Graph Theory and Applications*, pages 328–335, 2023.

[6] P. Erdős. Extremal problems in graph theory. In *Theory of Graphs and Its Applications, Proc. Sympos. Smolenice*, pages 29–36, 1964.

[7] F. Havet, B. Reed, M. Stein, and D. Wood. A variant of the Erdős-Sós conjecture. *Journal of Graph Theory*, 94(1):131–158, 2020.

[8] J. Hyde and B. Reed. Graphs of minimum degree at least $\lfloor d/2 \rfloor$ and large enough maximum degree embed every tree with d vertices. *Procedia Computer Science*, 223:217–222, 2023.

[9] R. Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.

[10] A. Pokrovskiy. Hyperstability in the Erdős-Sós Conjecture. *arXiv preprint, arXiv:2409.15191*, 2024.

[11] A. Pokrovskiy. Notes on embedding trees in graphs with $O(|T|)$-sized covers. *arXiv preprint arXiv:2409.15189*, 2024.

[12] B. Reed and M. Stein. Spanning trees in graphs of high minimum degree with a universal vertex I: An asymptotic result. *Journal of Graph Theory*, 102(4):737–783, 2023.

[13] B. Reed and M. Stein. Spanning trees in graphs of high minimum degree with a universal vertex II: A tight result. *Journal of Graph Theory*, 102(4):797–821, 2023.

[14] V. Rozhoň. A local approach to the Erdös–Sós conjecture. *SIAM Journal on Discrete Mathematics*, 33:643–664, 01 2019.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Excluding a rectangular grid

(Extended abstract)

Clément Rambaud*

Abstract

For every positive integer k, we define the k-treedepth as the largest graph parameter td_k satisfying (i) $\mathrm{td}_k(\emptyset) = 0$; (ii) $\mathrm{td}_k(G) \leq 1 + \mathrm{td}_k(G - u)$ for every graph G and every vertex $u \in V(G)$; and (iii) if G is a $(< k)$-clique-sum of G_1 and G_2, then $\mathrm{td}_k(G) \leq \max\{\mathrm{td}_k(G_1), \mathrm{td}_k(G_2)\}$, for all graphs G_1, G_2. This parameter coincides with treedepth if $k = 1$, and with treewidth plus 1 if $k \geq |V(G)|$. We prove that for every positive integer k, a class of graphs $\mathcal{C}$ has bounded k-treedepth if and only if there is a positive integer ℓ such that for every tree T on k vertices, no graph in $\mathcal{C}$ contains $T \square P_\ell$ as a minor.

This implies for $k = 1$ that a minor-closed class of graphs has bounded treedepth if and only if it excludes a path, for $k = 2$ that a minor-closed class of graphs has bounded 2-treedepth if and only if it excludes as a minor a ladder (Huynh, Joret, Micek, Seweryn, and Wollan; Combinatorica, 2021), and for large values of k that a minor-closed class of graphs has bounded treewidth if and only if it excludes a grid (Grid-Minor Theorem, Robertson and Seymour; JCTB, 1986).

As a corollary, we obtain the following qualitative strengthening of the Grid-Minor Theorem in the case of bounded-height grids. For all positive integers k, ℓ, every graph that does not contain the $k \times \ell$ grid as a minor has $(2k-1)$-treedepth at most a function of (k, ℓ).

The Grid-Minor Theorem, proved by Robertson and Seymour [10], characterizes classes of graphs having bounded treewidth in terms of excluded minors. More precisely, it asserts that a class of graphs $\mathcal{C}$ has bounded treewidth if and only if there is a positive integer ℓ such that, for every $G \in \mathcal{C}$, the $\ell \times \ell$ grid is not a minor of G. The "only if" part can be easily obtained from the fact that treewidth is monotone for the graph minor relation,

*Université Côte d'Azur, CNRS, Inria, I3S, Sophia-Antipolis, France. E-mail: clement.rambaud@inria.fr. Supported by the ANR under contract Digraphs ANR-19-CE48-0013. A full version of this article is available at https://arxiv.org/abs/2501.11617

and that grids have unbounded treewidth. Hence the core of this theorem is the fact that graphs excluding the $\ell \times \ell$ grid as a minor have treewidth at most $f(\ell)$, for some function f. A lot of research has been carried out to obtain good bounds on the smallest such function f. While the original proof of Robertson and Seymour give a superpolynomial upper bound on f, Chekuri and Chuzhoy proved a first polynomial upper bound in [2], which was later lowered down to $\ell^9 \log(\ell)^{\mathcal{O}(1)}$ by Chuzhoy and Tan [3].

The treedepth is a graph parameter larger than treewidth which can be characterized as the largest graph parameter td satisfying (i) $\mathrm{td}(\emptyset) = 0$[1], (ii) $\mathrm{td}(G) \leq \mathrm{td}(G-u)+1$ for every graph G and for every $u \in V(G)$, and (iii) $\mathrm{td}(G) \leq \max\{\mathrm{td}(G_1), \mathrm{td}(G_2)\}$ if G is the disjoint union of G_1 and G_2, for all graphs G_1, G_2. A class of graphs $\mathcal{C}$ has bounded treedepth if and only if there is a positive integer ℓ such that for every $G \in \mathcal{C}$, P_ℓ, the path on ℓ vertices, is not a minor of G. This analog of the Grid-Minor Theorem for treedepth can be easily obtained by elementary arguments, see [9, Proposition 6.1].

Huynh, Joret, Micek, Seweryn, and Wollan [8] recently introduced a variant of treedepth, called 2-treedepth and denoted by td_2, which can be characterized as the largest graph parameter satisfying (i)-(iii) of the definition of treedepth, together with the property (iv) $\mathrm{td}_2(G) \leq \max\{\mathrm{td}_2(G_1), \mathrm{td}_2(G_2)\}$ if G can be obtained from the disjoint union of G_1 and G_2 by identifying one vertex in G_1 with one vertex in G_2, for all graphs G_1, G_2. Huynh, Joret, Micek, Seweryn, and Wollan proved that a class of graphs has bounded 2-treedepth if and only if for some positive integer ℓ, every graph in this class excludes the $2 \times \ell$ grid, as a minor. While initially defined to describe the structure of graphs excluding a $2 \times \ell$ grid as a minor, it turns out that this parameter plays a important role in the study of weak coloring numbers in minor-closed classes of graphs, see [7, 6].

It is noteworthy that treewidth has a similar characterization: the treewidth is the largest graph parameter tw satisfying (i) $\mathrm{tw}(\emptyset) = -1$, (ii) $\mathrm{tw}(G) \leq \mathrm{tw}(G-u)+1$ for every graph G and for every $u \in V(G)$, and (iii) $\mathrm{tw}(G) \leq \max\{\mathrm{tw}(G_1), \mathrm{tw}(G_2)\}$ if G is a clique-sum of G_1 and G_2, for all graphs G_1, G_2. Here, a clique-sum of two graphs G_1 and G_2 is any graph G obtained from the disjoint union of G_1 and G_2 by identifying a clique in G_1 with a clique in G_2 of the same size. Furthermore, for every positive integer k, if these cliques have size less than k, then we say that G is a $(<k)$-clique-sum of G_1 and G_2. In particular, a (<1)-clique-sum of G_1 and G_2 is always the disjoint union of G_1 and G_2. This notion leads to the following generalization of treewidth, which is the subject of this paper.

For every $k \in \mathbb{N}_{>0} \cup \{+\infty\}$, we define the *k-treedepth* as the largest graph parameter td_k satisfying

(i) $\mathrm{td}_k(\emptyset) = 0$,

(ii) $\mathrm{td}_k(G) \leq 1 + \mathrm{td}_k(G-u)$ for every graph G and every vertex $u \in V(G)$, and

(iii) $\mathrm{td}_k(G) \leq \max\{\mathrm{td}_k(G_1), \mathrm{td}_k(G_2)\}$ if G is a $(<k)$-clique-sum of G_1 and G_2, for all graphs G_1, G_2.

[1] We denote by $\emptyset$ both the empty set and the graph with no vertices.

This gives a well-defined parameter because if $\mathcal{P}$ is the family of all the graph parameters satisfying (i)-(iii), then $\mathrm{td}_k\colon G \mapsto \max_{\mathrm{p}\in\mathcal{P}} \mathrm{p}(G)$ also satisfies (i)-(iii).

Note that the 1-treedepth is the usual treedepth, the 2-treedepth coincides with the homonymous parameter td_2 introduced by Huynh, Joret, Micek, Seweryn, and Wollan in [8], and that the $+\infty$-treedepth coincides with the treewidth plus 1. Hence, for every graph G,

$$\mathrm{td}(G) = \mathrm{td}_1(G) \geq \mathrm{td}_2(G) \geq \cdots \geq \mathrm{td}_{+\infty}(G) = \mathrm{tw}(G) + 1.$$

Our main result is a characterization of classes of graphs having bounded k-treedepth in terms of excluded minors, which generalizes the aforementioned results on treedepth, 2-treedepth, and treewidth.

Theorem 1. *Let k be a positive integer. A class $\mathcal{C}$ of graphs has bounded k-treedepth if and only if there exists an integer ℓ such that for every tree T on k vertices, no graph in $\mathcal{C}$ contains $T\square P_\ell$ as a minor.*

Here, $\square$ denotes the Cartesian product: for all graphs G_1, G_2, $G_1\square G_2 = \big(V(G_1) \times V(G_2), \{(u,v)(u',v') \mid (u = v \text{ and } u'v' \in E(G_2)) \text{ or } (uv \in E(G_1) \text{ and } u' = v')\}\big)$; and P_ℓ denotes the path on ℓ vertices. See Figure 1 for some examples of graphs of the form $T\square P_\ell$.

Theorem 1 gives for every positive integer k a characterization of classes of graphs having bounded k-treedepth in terms of excluded minors, which was previously known only for $k = 1$ and $k = 2$ [8]. For example, when $k = 5$, Theorem 1 applied to $k = 5$ implies that for every fixed positive integer ℓ, graphs excluding $T_1\square P_\ell, T_2\square P_\ell$, and $T_3\square P_\ell$ as minors have bounded 5-treedepth, where T_1, T_2, T_3 are the three trees on 5 vertices up to isomorphism (see Figure 1). This is optimal since each of the three families $\{T_1\square P_\ell\}_{\ell\geq 1}, \{T_2\square P_\ell\}_{\ell\geq 1}, \{T_3\square P_\ell\}_{\ell\geq 1}$ has unbounded 5-treedepth.

A similar result was obtained by Geelen and Joeris [5] for the parameter called k-treewidth, which is the minimum width of a tree decomposition with adhesion smaller than k. However, their proof relies on very different techniques. Indeed, the argument of [5] is a generalization of a proof of the Grid-Minor Theorem, while here, the structure we are looking for is more complex, but we use the Grid-Minor Theorem as a black box.

As a corollary of Theorem 1, we obtain the following structural property for the graphs excluding the $k \times \ell$ grid as a minor, when k is small compared to ℓ, which follows from the fact that the $k \times \ell$ grid is a minor of any cartesian product of a tree on $2k - 1$ vertices and a long enough path. This is a qualitative strengthening of the celebrated Grid-Minor Theorem by Robertson and Seymour [10].

Corollary 2. *There is a function $f\colon \mathbb{N}^2 \to \mathbb{N}$ such that for all positive integers k, ℓ, for every graph G, if the $k \times \ell$ grid is not a minor of G, then*

$$\mathrm{td}_{2k-1}(G) \leq f(k, \ell).$$

In [1], Biedl, Chambers, Eppstein, De Mesmay, and Ophelders defined the *grid-major height* of a planar graph H, denoted by $\mathrm{GMh}(H)$, as the smallest integer k such that H is a minor of the $k \times \ell$ grid for some integer ℓ. With this definition, Corollary 2 implies the following.

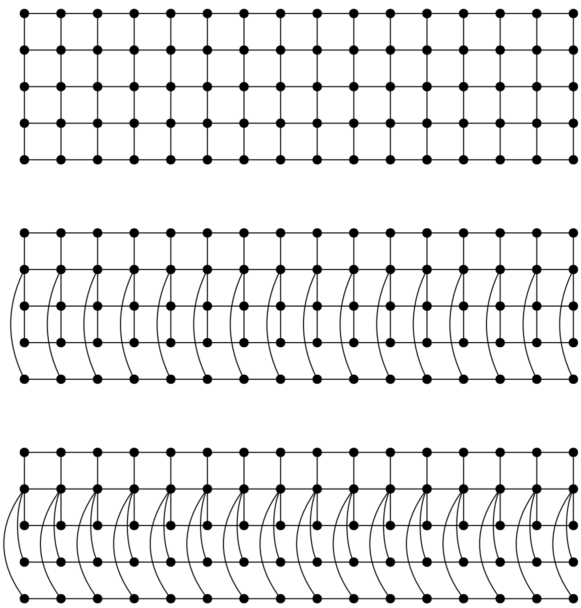


Figure 1: The three families of obstructions for 5-treedepth given by Theorem 1.

Corollary 3. *There is a function $f\colon \mathbb{N} \to \mathbb{N}$ such that the following holds. For every planar graph H with at least one vertex and for every H-minor-free graph G,*

$$\mathrm{td}_{2\,\mathrm{GMh}(H)-1}(G) \leq f(|V(H)|).$$

Moreover, GMh is the right parameter in such a result since H-minor-free graphs have unbounded $\mathrm{td}_{\mathrm{GMh}(H)-1}$, as witnessed by the family of the $(\mathrm{GMh}(H)-1)\times \ell$ grids for $\ell > 0$. This gives a better understanding of H-minor-free graphs when H is planar and has a specific structure. In the same spirit, Dujmović, Hickingbotham, Hodor, Joret, La, Micek, Morin, Rambaud, and Wood [4] recently gave a product structure for H-minor-free graphs when H is planar and has bounded treedepth.

We also investigate the variant of k-treedepth for path decompositions, instead of tree decompositions, that we call k-pathdepth and denoted by pd_k. Informally, the *k-pathdepth* pd_k is the largest parameter satisfying (i) and (ii) of the definition of k-treedepth, and an analog of (iii) ensuring that the clique-sums are made in the way of a path. This gives a family of parameters satisfying for every graph G

$$\mathrm{td}(G) = \mathrm{pd}_1(G) \geq \mathrm{pd}_2(G) \geq \cdots \geq \mathrm{pd}_{+\infty}(G) = \mathrm{pw}(G) + 1$$

and

$$\mathrm{pd}_k(G) \geq \mathrm{td}_k(G)$$

for every positive integer k. Hence, if a class of graphs has bounded k-pathdepth, then it has bounded k-treedepth and bounded pathwidth. Quite surprisingly, these two necessary conditions are also sufficient.

Theorem 4. *Let k be a positive integer and let $\mathcal{C}$ be a class of graphs. The following are equivalent.*

(1) $\mathcal{C}$ has bounded k-pathdepth.

(2) $\mathcal{C}$ has bounded pathwidth and bounded k-treedepth.

(3) There is an integer ℓ such that for every $G \in \mathcal{C}$, $\mathrm{pw}(G) \leq \ell$, and for every tree T on k vertices, $T \square P_\ell$ is not a minor of G.

Our proof actually shows the equivalence between Items (1) and (3), while the equivalence between Items (2) and (3) follows from Theorem 1.

References

[1] Therese Biedl, Erin Wolf Chambers, David Eppstein, Arnaud De Mesmay, and Tim Ophelders. Homotopy height, grid-major height and graph-drawing height. In *Graph Drawing and Network Visualization*, page 468–481. Springer International Publishing, 2019. arXiv:1908.05706.

[2] Chandra Chekuri and Julia Chuzhoy. Polynomial bounds for the Grid-Minor Theorem. *Journal of the ACM*, 63(5):1–65, 2016. arXiv:1305.6577.

[3] Julia Chuzhoy and Zihan Tan. Towards tight(er) bounds for the excluded grid theorem. *Journal of Combinatorial Theory, Series B*, 146:219–265, 2021. arXiv:1901.07944.

[4] Vida Dujmović, Robert Hickingbotham, Jędrzej Hodor, Gwenaël Joret, Hoang La, Piotr Micek, Pat Morin, Clément Rambaud, and David R. Wood. The Grid-Minor Theorem revisited. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, page 1241–1245. Society for Industrial and Applied Mathematics, 2024. arXiv:2307.02816.

[5] Jim Geelen and Benson Joeris. A generalization of the Grid Theorem. *arXiv preprint*, 2016. arXiv:1609.09098.

[6] Jędrzej Hodor, Hoang La, Piotr Micek, and Clément Rambaud. Weak coloring numbers of minor-closed graph classes. *arXiv preprint*, 2024. arXiv:2407.04588.

[7] Jędrzej Hodor, Hoang La, Piotr Micek, and Clément Rambaud. *Weak coloring numbers of minor-closed graph classes*, page 3325–3334. Society for Industrial and Applied Mathematics, 2025.

[8] Tony Huynh, Gwenaël Joret, Piotr Micek, Michał T. Seweryn, and Paul Wollan. Excluding a ladder. *Combinatorica*, 42(3):405–432, 2022. arXiv:2002.00496.

[9] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*. Springer Berlin, Heidelberg, 2012.

[10] Neil Robertson and Paul Seymour. Graph minors. V. Excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, 1986.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - August 29, 2025

An improvement on the K_2-irregular index of bipartite graphs*

(Extended abstract)

Pedro S. M. Arraes[†] Carla N. Lintzmayer[‡] Guilherme O. Mota[†]
Maycon Sambinelli[‡]

Abstract

A *locally irregular decomposition* of a graph G is a collection $\{G_1, \ldots, G_k\}$ of subgraphs of G such that $\{E(G_1), \ldots, E(G_k)\}$ is a partition of $E(G)$ and every G_i is *locally irregular*, i.e., $d_{G_i}(u) \neq d_{G_i}(v)$ for every $uv \in E(G_i)$. A *K_2-irregular decomposition* of G is a collection as the previous one with the difference that, for every G_i, all of its components are either locally irregular or isomorphic to K_2. If a graph G has a locally irregular (resp. K_2-irregular) decomposition, $\chi'_{\text{irr}}(G)$ (resp. $\chi'_{K_2\text{-irr}}(G)$) denotes the size of a smallest such decomposition. It has been conjectured that $\chi'_{K_2\text{-irr}}(G) \leq 3$ for any graph G. We prove that this holds for bipartite graphs. Our result also implies an improvement on the best-known upper bound of $\chi'_{K_2\text{-irr}}(G)$ for any graph G, from 220 to 112.

1 Introduction

A graph G is called *locally irregular* if no two adjacent vertices have the same degree. A *locally irregular decomposition* is a collection $\{G_1, \ldots, G_k\}$ of locally irregular subgraphs

*This research has been partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil – CAPES – Finance Code 001. P. S. M. Arraes was partially supported by FAPESP (2020/03336-3). C. N. Lintzmayer was partially supported by CNPq (312026/2021-8 and 404315/2023-2) and by L'ORÉAL-UNESCO-ABC For Women In Science. G. O. Mota was partially supported by CNPq (315916/2023-0, 406248/2021-4, and 407970/2023-1) and by FAPESP (2018/04876-1). M. Sambinelli was partially supported by CNPq (407970/2023-1). FAPESP is the São Paulo Research Foundation. CNPq is the National Council for Scientific and Technological Development of Brazil.

†Instituto de Matemática e Estatística, Universidade de São Paulo, SP, Brazil. E-mails: {pedro.arraes, mota}@ime.usp.br.

‡Centro de Matemática Computação e Cognição, Universidade Federal do ABC, SP, Brazil. E-mails: {carla.negri, m.sambinelli}@ufabc.edu.br.

of G such that $\{E(G_1), \ldots, E(G_k)\}$ is a partition of $E(G)$. Notice that we can consider this partition as the set of color classes of an edge coloring, so we can also refer to a locally irregular decomposition as a *locally irregular edge-coloring*. If G has such a decomposition, we call it *decomposable* and define the *locally irregular index* $\chi'_{\text{irr}}(G)$ as the size of the smallest locally irregular decomposition of G.

These concepts were introduced by Baudon, Bensmail, Przybyło, and Woźniak [2] in 2015. They also characterized the graphs that do not have locally irregular decompositions:

- paths and cycles of odd length; and
- the class $\mathfrak{T}$ consisting of C_3 and every graph G obtained in the following way. Take any $H \in \mathfrak{T}$ with a vertex v of degree 2 belonging to a triangle; create G by identifying v with an endvertex of either an even length path or an odd length path attached to a triangle on the other endvertex.

Moreover, Baudon *et al.* [2] proved that $\chi'_{\text{irr}}(G) \leq 2$ when G is an even length path, a complete bipartite graph, a d-regular bipartite graph with $d \geq 3$, a hypercube, and a complete graph minus one or two edges; and that $\chi'_{\text{irr}}(G) \leq 3$ when G is an even length cycle, a complete graph, a decomposable tree, and a d-regular graph with $d \geq 10^7$. Based on these results, they proposed the following.

Conjecture 1.1 ([2]). *For every decomposable graph G, we have $\chi'_{\text{irr}}(G) \leq 3$.*

So far, only one graph is known to break this conjecture. Sedlar and Ŝkrekovski [7] identified a small cactus, drawn in Figure 1, whose irregular chromatic index is 4. Moreover, in a subsequent article, Sedlar and Ŝkrekovski [8] proved that all other decomposable cacti satisfy Conjecture 1.1.

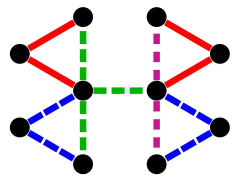


Figure 1: Graph with a locally irregular 4-edge-coloring.

Regarding decomposable graphs in general, the first upper bound for the locally irregular index was given by Bensmail, Merker, and Thomassen [3]: 328 for all decomposable graphs, and 327 if the number of edges is even. Later on, Lužar, Przybyło and Soták [6] improved it to the best known so far.

Theorem 1.2 ([6]). *Let G be a decomposable graph. Then $\chi'_{\text{irr}}(G) \leq 219$ if $|E(G)|$ is even, and $\chi'_{\text{irr}}(G) \leq 220$ otherwise.*

A relaxation of this decomposition, introduced by Bensmail and Stevens [4] in 2016, is the *regular-irregular decomposition* of G, which is a collection $\{G_1, \ldots, G_k\}$ of subgraphs of G such that $\{E(G_1), \ldots, E(G_k)\}$ is a partition of $E(G)$ and each component of every G_i is either regular or locally irregular. In their article, Bensmail and Stevens showed a result on bipartite graphs and observed that all regular components of the subgraphs in the decomposition are isomorphic to K_2. This inspired the definition of *K_2-irregular decomposition*, a regular-irregular decomposition where all regular components of its subgraphs are isomorphic to K_2. As we did for the locally irregular decomposition, we can also interpret this as a *K_2-irregular edge-coloring*. We then denote the size of a smallest K_2-irregular decomposition of a graph G, ita *K_2-irregular chromatic index*, by $\chi'_{K_2\text{-irr}}(G)$.

Notice that a locally irregular decomposition is a K_2-irregular decomposition such that none of its subgraphs have regular components. Consequently, any decomposable graph G has a K_2-irregular decomposition, implying $\chi'_{K_2\text{-irr}}(G) \leq \chi'_{\text{irr}}(G)$. Moreover, since each non-decomposable graph G has an odd number of edges, we can take an edge $e \in E(G)$ that does not disconnect the graph, and a locally irregular decomposition $\{G_1, \ldots, G_3\}$ of $G - e$[1] to obtain a K_2-irregular decomposition $\{G_1, \ldots, G_3, G[e]\}$ of G. Therefore, $\chi'_{K_2\text{-irr}}(G)$ is well-defined for every graph G.

The previous discussion together with Theorem 1.2 directly gives us $\chi'_{K_2\text{-irr}}(G) \leq 220$ for any graph G. Baudon, Bensmail, Hocquard, Senhaji, and Sopena [1] formalized the concepts regarding K_2-irregular decompositions and proposed an analog of Conjecture 1.1.

Conjecture 1.3 ([1]). *For every graph G, we have $\chi'_{K_2\text{-irr}}(G) \leq 3$.*

In the next section, we briefly show what has been done on bipartite graphs regarding the locally irregular and K_2-irregular decompositions, and then we prove our main result, which is $\chi'_{K_2\text{-irr}}(G) \leq 3$ for any bipartite graph G. Applying this bound to the proof of Theorem 1.2 shows that $\chi'_{K_2\text{-irr}}(G) \leq 112$ for any graph G.

2 Bipartite Graphs

Bipartite graphs have an important role in these problems. To give the first general upper bound for the irregular chromatic index, Bensmail *et al.* [3] showed the following. First, every graph G with $|E(G)|$ odd can be decomposed into a locally irregular graph G^* and a graph G' whose components all have an even number of edges. Next, any decomposable graph G with $|E(G)|$ even can be decomposed into a graph H with $\chi'_{\text{irr}}(H) \leq 3$ and a $(2 \cdot 10^{10} + 2)$-degenerate graph D with all components having an even number of edges, which in turn can be decomposed into at most 36 bipartite graphs, each having all components with an odd number of edges. In their article, they proved that $\chi'_{\text{irr}}(G) \leq 10$ for every decomposable bipartite graph G; and if G has an even number of edges, this bound decreases by 1. Afterward, Lužar, Przybyło, and Soták [6] reduced it to the following, which in turn implies Theorem 1.2.

Theorem 2.1 ([6]). *Let G be a bipartite graph. If $|E(G)|$ is even, then $\chi'_{\text{irr}}(G) \leq 6$; otherwise, $\chi'_{\text{irr}}(G) \leq 7$.*

As for the K_2-irregular decomposition, we mentioned that Bensmail and Stevens [4] showed a result for regular-irregular decompositions on bipartite graphs, and that their decomposition is, in particular, K_2-irregular. They proved the following.

Theorem 2.2 ([4]). *For every bipartite graph G, we have $\chi'_{K_2\text{-irr}}(G) \leq 6$.*

[1] One can show that $\chi'_{\text{irr}}(G - e) \leq 3$ for any non-decomposable graph G and non-cut edge e.

Bensmail and Stevens decompose G into an Eulerian bipartite graph G_1 and a forest G_2, and prove that these graphs have K_2-irregular decompositions of sizes, respectively, 2 and 4. The decomposition of the forest is simple. Let $T_1, \dots, T_k$ be the components of G_2. For each $i \in [k]$, choose an arbitrary root r_i of T_i and, for each vertex $v \in V(T_i)$, let $T_{i,v}$ be the subgraph induced by v and its children. Let $\{T_{i,\mathrm{e}}, T_{i,\mathrm{o}}\}$ be the decomposition of T_i such that, for each $v \in V(T_i)$ with at least one child, $T_{i,v} \subseteq T_{i,\mathrm{e}}$ if $\mathrm{dist}_{T_i}(r_i, v)$ is even, and $T_{i,v} \subseteq T_{i,\mathrm{o}}$ if $\mathrm{dist}_{T_i}(r_i, v)$ is odd. We call this the *standard tree K_2-irregular decomposition*. The K_2-irregular decomposition of G_2 is $\{G_{2,e}, G_{2,o}\}$, where $G_{2,e} = \bigcup_{i=1}^k T_{i,\mathrm{e}}$ and $G_{2,o} = \bigcup_{i=1}^k T_{i,\mathrm{o}}$.

For the Eulerian bipartite graph $G_1 = (A_1 \cup B_1, E_1)$, Bensmail and Stevens construct an *S-join*, a spanning subgraph H of G_1 such that $d_H(v)$ is odd if and only if $v \in S$; notice that $|S|$ needs to be even. When $|A_1|$ is even, take $S = A_1$ and note that H and $G - E(H)$ are locally irregular; if $|A_1|$ is odd and $|B_1|$ is even, take $S = B_1$. When both $|A_1|$ and $|B_1|$ are odd, choose $v \in A_1$ and take $S = A_1 \setminus \{v\}$; $G - E(H)$ has a path P of length at least 1 such that $G - E(H) - E(P)$ is K_2-irregular.

Inspired by this idea, we obtained the following.

Theorem 2.3. *For every bipartite graph G, we have $\chi'_{K_2\text{-irr}}(G) \leq 3$.*

Sketch of the proof. Let $G = (A \cup B, E)$ be a connected bipartite graph and consider first that $|A|$ is even. We let $A = \{v_1, \dots, v_{2a}\}$ and construct an A-join H as follows. We initially define H as an empty spanning subgraph of G. For each $i \in [a]$, we take a $v_{2i-1}v_{2i}$-path $P_i \subset G$ and update $E(H)$ as the symmetric difference of $E(H)$ and $E(P_i)$. Once this process is complete, we take $G' \coloneqq G - E(H)$. If G' is acyclic, we are done; otherwise, while it has a cycle C, we transfer C's edges from G' to H. At the end, $d_H(v_i)$ is odd for every $v_i \in A$, $d_H(u)$ is even for every $u \in B \cap V(H)$, and G' is acyclic, i.e., it is a forest. This means that H is locally irregular and G' has a K_2-irregular decomposition of size at most 2, so the result follows.

Now consider that both $|A|$ and $|B|$ are odd. We take $v \in V(G)$ with smallest degree such that $G - v$ is still connected; without loss of generality, assume $v \in A$. Then, we apply the above procedure to decompose $G - v$ into an A'-join, with $A' \coloneqq A \setminus \{v\}$, and a forest F. Our goal is to show that $\chi'_{K_2\text{-irr}}(F') \leq 2$ for $F' \coloneqq F + \{uv \colon u \in N_G(v)\}$. Since all components of F are trees, we can decompose those without neighbors of v in F' using the standard tree K_2-irregular decomposition. Thus, we assume that every component of F has a vertex $u \in N_G(v)$.

When each neighbor of v belongs to a different component of F, F' is also a forest and we are finished. Suppose now that there is a component of F with two neighbors of v. When $d_G(v) = 2$, F' has only one cycle, which necessarily has even length and contains v; let $C = (v = v_1, \dots, v_k, v_1)$ be this cycle. Let F_1, F_2 initially be empty subgraphs of F'. We distribute the edges of C equally between F_1 and F_2, where $v_iv_{i+1} \in E(F_1)$ if and only if $i \in [k]$ is odd. Let $T_1, T_2, \dots, T_k$ be the components of $F' - E(C)$ such that $v_i \in T_i$ for each $i \in [k]$. Then, for each $i \in [2, k]$, we enroot T_i on v_i and use the standard tree K_2-irregular decomposition to obtain a decomposition $\{T_{i,\mathrm{e}}, T_{i,\mathrm{o}}\}$ of T_i. If $v_{i-1}v_i \in E(F_1)$, we add the edges of $T_{i,\mathrm{e}}$ and $T_{i,\mathrm{o}}$, respectively, to F_1 and to F_2; otherwise, we do the opposite. in the end, $\{F_1, F_2\}$ is a K_2-irregular decomposition of F'.

When $d_G(v) \geq 3$, we take the largest set $\{u_1, v_1, u_2, v_2, \ldots, u_k, v_k\}$ such that for every $i \in [k]$, there is a component of F containing both u_i, v_i and $d_{F'}(u_i) = d_{F'}(v_i) = d_G(v)+2-i$. Our goal decomposition $\{F_1, F_2\}$ of F' is such that v has k neighbors in F_2, which is possible since $d_G(v) \geq 2k$. We take $F_1, F_2 \subset F'$ initially empty. Let $T_1, \ldots, T_q, T_{q+1}, \ldots, T_\ell$ be the components of F such that for every $i \in [q]$, T_i contains a vertex with degree in $[d_G(v) + 1 - k, d_G(v) + 1]$. For every $i \in [\ell]$ we enroot T_i in r_i such that, if $i \in [q]$, r_i has the smallest degree possible in $[d_G(v) + 1 - k, d_G(v) + 1]$; if $i \in (q, \min\{k, \ell\}]$, r_i is adjacent to v; and if $i \in (\min\{k, \ell\}, \ell]$, r_i is not adjacent to v. We decompose each T_i using the standard tree K_2-irregular decomposition, obtaining $\{T_{i,\mathrm{e}}, T_{i,\mathrm{o}}\}$. If $vr_i \in E(F')$, we add $T_{i,\mathrm{e}}$ and $T_{i,\mathrm{o}}$, respectively, to F_1 and F_2; otherwise, we do the opposite. As for each edge vw, we add it to F_2 if there is $i \in [k]$ for which $w = r_i$, and to F_1 if there is no such i. This is sufficient for $k \leq \ell$.

Now suppose that $\ell < k$. Proceeding as above only gives us ℓ edges incident to v in F_2, thus we need other $k - \ell$ neighbors of v. If $k = 2$, then $\ell = 1$ and F is a tree. Since $u_1, u_2 \in V(T_1)$ and $d_{F'}(u_1) > d_{F'}(u_2) \geq 4$, $u_1 \neq r_1$ and so we transfer vu_1 from F_1 to F_2 and if necessary, make the same transfer for some edges between u_1 and its descendants in F. If $k \geq 3$, for every $i \in [k-1]$ we know that $d_{F'}(u_i) = d_{F'}(v_i) = d(v)+2-i \geq 2k+2-(k-1) \geq 6$ and u_i, v_i are in the same component of F, so we can choose our $k - \ell$ vertices from $\{u_1, \ldots, u_{k-1}\}$, without loss of generality, knowing that each has degree in F' at least 6 and is not one of the roots. Let the chosen vertices be $w_1, \ldots, w_{k-\ell}$. Knowing that there may be u_i and u_j in the same component of F such that the parent of u_j is a child of u_i, we assume our vertices are ordered so that if the parent of w_p is a child of w_q for some $p \neq q \in [k - \ell]$, then $p > q$. Moreover, since all our chosen vertices have degree at least 6 in F', this means that each one has at least four children in F. We analyze our chosen vertices iteratively as follows. For each $i \in [k - \ell]$, we transfer vw_i from F_1 to F_2, which can cause w_i to have the same degree in F_2 as its parent or v. To avoid this, we transfer some edges between w_i and its children form F_1 to F_2; in order to counter any problem, w_i needs to have at least four children, which it does. $\square$

3 Concluding Remarks

We have proved that $\chi'_{K_2\text{-irr}}(G) \leq 3$ for every bipartite graph G. Applying this bound to the process of Theorem 1.2, we deduce that $\chi'_{K_2\text{-irr}}(G) \leq 112$ for any graph G. Currently, we are trying to find examples of bipartite graphs that need three colors for a K_2-irregular edge-coloring. If there are no such graphs, we believe that another method is required to prove the optimal bound. Furthermore, there is room for improving the upper bound of the irregular chromatic index on bipartite graphs.

References

[1] Olivier Baudon, Julien Bensmail, Hervé Hocquard, Mohammed Senhaji, and Eric Sopena. On locally irregular decompositions of subcubic graphs. *Opuscula Mathematica*, 38(6):795–817, 2018.

[2] Olivier Baudon, Julien Bensmail, Jakub Przybyło, and Mariusz Woźniak. On decomposing regular graphs into locally irregular subgraphs. *European Journal of Combinatorics*, 49:90–104, 2015.

[3] Julien Bensmail, Martin Merker, and Carsten Thomassen. Decomposing graphs into a constant number of locally irregular subgraphs. *European Journal of Combinatorics*, 60:124–134, 2017.

[4] Julien Bensmail and Brett Stevens. Edge-partitioning graphs into regular and locally irregular components. *Discrete Mathematics & Theoretical Computer Science*, 17(Graph Theory), 2016.

[5] Carla Negri Lintzmayer, Guilherme Oliveira Mota, and Maycon Sambinelli. Decomposing split graphs into locally irregular graphs. *Discrete Applied Mathematics*, 292:33–44, 2021.

[6] Borut Lužar, Jakub Przybyło, and Roman Soták. New bounds for locally irregular chromatic index of bipartite and subcubic graphs. *Journal of Combinatorial Optimization*, 36:1425–1438, 2018.

[7] Jelena Sedlar and Riste Škrekovski. Remarks on the local irregularity conjecture. *Mathematics*, 9(24):3209, 2021.

[8] Jelena Sedlar and Riste Škrekovski. Local irregularity conjecture vs. cacti. *Discrete Applied Mathematics*, 343:115–133, 2024.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On Counting Graph Homomorphisms by Entropy Arguments

(Extended abstract)

Igal Sason*

Abstract

This extended abstract applies properties of Shannon entropy to derive a lower bound on the number of homomorphisms from a complete bipartite graph to any bipartite graph. Further upper and lower bounds on homomorphism counts between arbitrary bipartite graphs, proofs and observations, are provided in the full version of this work, available as the arXiv preprint *Counting Graph Homomorphisms in Bipartite Settings* (*https://arxiv.org/abs/2508.06977*).

1 Introduction

Combinatorial techniques serve a vital role in addressing problems in information theory and coding theory. Several examples where tools from combinatorics and graph theory are used to study fundamental problems in information theory were briefly surveyed in [1]. Many classical and modern results in information theory can be also derived through a combinatorial perspective, particularly using the method of types [2, 3]. The reverse direction, applying information-theoretic tools to obtain combinatorial results, has proven to be equally fruitful. Notably, Shannon entropy has significantly deepened the understanding of the structural and quantitative properties of combinatorial objects by enabling concise and often elegant proofs of classical results in combinatorics (see, e.g., [4, Chapter 37], [5], [6, Chapter 22], and [7–18]).

This extended abstract takes the latter direction, employing the Shannon entropy to derive a lower bound on the number of graph homomorphisms from a complete bipartite graph to an arbitrary bipartite graph. A full version of this work, which derives both upper and lower bounds on homomorphism counts between *arbitrary* bipartite graphs, together with complete proofs and additional observations and numerical results, is available as an arXiv preprint [18].

Graph homomorphisms provide a powerful framework for the study of graph mappings, revealing insights into structural properties, colorings, and symmetries. Their applications

*Viterbi Faculty of Electrical and Computer Engineering and the Department of Mathematics, Technion—Israel Institute of Technology, Haifa 3200003, Israel. E-mail: `sason@ee.technion.ac.il`.

span multiple disciplines, including statistical physics, where they model spin systems [19], and computational complexity, where they underpin constraint satisfaction problems [20]. Research work has led to significant progress in understanding the problem of counting graph homomorphisms, a subject of both theoretical and practical relevance (see [7, 20–35]).

2 Preliminaries

In the sequel, let $\mathsf{V}(\mathsf{G})$ and $\mathsf{E}(\mathsf{G})$ denote the vertex set and edge set of a graph G, respectively. For adjacent vertices $u, v \in \mathsf{V}(\mathsf{G})$, let $e = \{u, v\} \in \mathsf{E}(\mathsf{G})$ denote the edge connecting them.

Let F and G be finite, simple, and undirected graphs. A *homomorphism* from a source graph F to a target graph G, denoted by $\mathsf{F} \to \mathsf{G}$, is a mapping $\psi\colon \mathsf{V}(\mathsf{F}) \to \mathsf{V}(\mathsf{G})$ such that every edge in F is mapped to an edge in G:

$$\{u, v\} \in \mathsf{E}(\mathsf{F}) \implies \{\psi(u), \psi(v)\} \in \mathsf{E}(\mathsf{G}). \tag{1}$$

The following establishes connections between graph homomorphisms and classical graph invariants. Let $\omega(\mathsf{G})$ and $\chi(\mathsf{G})$ denote the clique number and chromatic number of a graph G, respectively. Then, $\omega(\mathsf{G})$ is the largest $k \in \mathbb{N}$ such that there exists a homomorphism $\mathsf{K}_k \to \mathsf{G}$, and $\chi(\mathsf{G})$ is the smallest $k \in \mathbb{N}$ such that a homomorphism $\mathsf{G} \to \mathsf{K}_k$ exists. Consequently, such graph homomorphisms characterize the independence number, clique number, and chromatic number of a graph, problems known to be NP-hard [36].

Let $\mathrm{Hom}(\mathsf{F}, \mathsf{G})$ denote the set of all the homomorphisms $\mathsf{F} \to \mathsf{G}$, and define

$$\mathrm{hom}(\mathsf{F}, \mathsf{G}) \triangleq \left|\mathrm{Hom}(\mathsf{F}, \mathsf{G})\right| \tag{2}$$

as the number of such graph homomorphisms. These are called *homomorphism numbers.*

In addition to homomorphism numbers, we now introduce *homomorphism densities*, which are closely related.

Definition 2.1 (Homomorphism densities)**.** Let F and G be graphs. Let $v(\mathsf{F}) \triangleq |\,\mathsf{V}(\mathsf{F})|$ and $v(\mathsf{G}) \triangleq |\,\mathsf{V}(\mathsf{G})|$. The F-*homomorphism density in* G (or simply F-density in G) is the probability that a uniformly random mapping from $\mathsf{V}(\mathsf{F})$ to $\mathsf{V}(\mathsf{G})$ induces a graph homomorphism from F to G, i.e., it is given by

$$t(\mathsf{F}, \mathsf{G}) \triangleq \frac{\mathrm{hom}(\mathsf{F}, \mathsf{G})}{v(\mathsf{G})^{\,v(\mathsf{F})}}. \tag{3}$$

By Definition 2.1, we have $t(\mathsf{K}_1, \mathsf{G}) = 1$, and

$$t(\mathsf{K}_2, \mathsf{G}) = \frac{2\,e(\mathsf{G})}{v(\mathsf{G})^2}, \tag{4}$$

where $e(\mathsf{G}) \triangleq |\mathsf{E}(\mathsf{G})|$.

Before going into technical details, we highlight why counting graph homomorphisms is an important problem. In extremal graph theory and in the study of graph limits, homomorphism counts and homomorphism densities are basic building blocks; this viewpoint goes back to work of Lovász and collaborators [25]. In computer science, they appear naturally in the context of constraint satisfaction problems, where deciding whether a homomorphism exists, or counting how many there are, is a central computational task. And in statistical physics, partition functions of spin models can be expressed as weighted homomorphism counts, so methods and insights from this area translate directly. For background, there are excellent references by Lovász [25], by Borgs–Chayes–Lovász–Sós–Vesztergombi [22], and more recently by Yufei Zhao's textbook [29], which connect these perspectives very nicely.

3 Lower bound

In the following, we rely on properties of the Shannon entropy to derive a lower bound on the number of homomorphisms from the complete bipartite graph to any bipartite graph, and examine its tightness by comparing it to the specialized lower bound that holds by the satisfiability of Sidorenko's conjecture in the examined setting. Familiarity with Shannon entropy and its basic properties is assumed, following standard notation (see, e.g., [37, Chapter 3]).

Proposition 3.1 (Number of graph homomorphisms)**.** Let G be bipartite with partite sizes n_1, n_2 and with an edge density

$$\delta \triangleq \frac{|\mathsf{E}(\mathsf{G})|}{n_1 n_2} \in [0,1].$$

Then, for all $s, t \in \mathbb{N}$,

$$\begin{aligned} \hom(\mathsf{K}_{s,t}, \mathsf{G}) &\geq \delta^{st} (n_1^s n_2^t + n_1^t n_2^s) &(5)\\ &= \delta^{st} \hom(\mathsf{K}_{s,t}, \mathsf{K}_{n_1,n_2}) &(6)\end{aligned}$$

Proof. Let $\mathcal{U}$ and $\mathcal{V}$ denote the partite vertex sets of the simple bipartite graph G, where $|\mathcal{U}| = n_1$ and $|\mathcal{V}| = n_2$. Let (U, V) be a random vector taking values in $\mathcal{U} \times \mathcal{V}$, and suppose that $\{U, V\}$ is distributed uniformly at random on the edges of G. Then, the joint entropy of (U, V) is given by

$$\mathrm{H}(U, V) = \log \bigl|\mathsf{E}(\mathsf{G})\bigr| = \log(\delta n_1 n_2). \tag{7}$$

The random vector (U, V) can be sampled by first sampling the value $U = u$ from the marginal probability mass function (PMF) of U, denoted by P_U, and then sampling V from the conditional PMF $\mathsf{P}_{V|U}(\cdot|u)$. Construct a random vector $(U_1, \ldots, U_s, V_1, \ldots, V_t)$ as follows:

- Let $V_1, \ldots, V_t$ be conditionally independent and identically distributed (i.i.d.) given U, having the conditional PMF

$$\mathsf{P}_{V_1,\ldots,V_t|U}(v_1, \ldots, v_t|u) = \prod_{j=1}^{t} \mathsf{P}_{V|U}(v_j|u), \quad \forall\, u \in \mathcal{U}, \;\; (v_1, \ldots, v_t) \in \mathcal{V}^t. \tag{8}$$

- Let $U_1, \ldots, U_s$ be conditionally i.i.d. given $(V_1, \ldots, V_t)$, having the conditional PMF

$$\begin{aligned} &\mathsf{P}_{U_1,\ldots,U_s|V_1,\ldots,V_t}(u_1, \ldots, u_s|v_1, \ldots, v_t)\\ &\quad = \prod_{i=1}^{s} \mathsf{P}_{U_i|V_1,\ldots,V_t}(u_i|v_1, \ldots, v_t), \quad \forall\, (u_1, \ldots, u_s) \in \mathcal{U}^s, \;\; (v_1, \ldots, v_t) \in \mathcal{V}^t, \qquad (9)\end{aligned}$$

where the conditional PMFs on the right-hand side of (9) are given by

$$\begin{aligned}&\mathsf{P}_{U_i|V_1,\ldots,V_t}(u|v_1,\ldots,v_t)\\&=\frac{\mathsf{P}_U(u)\prod_{j=1}^{t}\mathsf{P}_{V|U}(v_j|u)}{\sum_{u'\in\mathcal{U}}\left\{\mathsf{P}_U(u')\prod_{j=1}^{t}\mathsf{P}_{V|U}(v_j|u')\right\}},\quad\forall\, u\in\mathcal{U},\ \ (v_1,\ldots,v_t)\in\mathcal{V}^t,\ \ i\in[s].\end{aligned}\tag{10}$$

By the construction of the random vector $(U_1,\ldots,U_s,V_1,\ldots,V_t)$ in (8)–(10), the following holds (see [18]):

1) $U_1,\ldots,U_s$ are identically distributed random variables, and $U_i \sim U$ (i.e., $\mathsf{P}_{U_i}=\mathsf{P}_U$) for all $i\in[s]$.

2) For all $i\in[s]$ and $j\in[t]$, $(U_i,V_j)\sim(U,V)$, and $(U_i,V_1,\ldots,V_t)\sim(U,V_1,\ldots,V_t)$.

Recall that, by assumption, the bipartite graph G has no isolated vertices, thus making the above construction feasible.

The joint entropy of the random subvector $(U_1,V_1,\ldots,V_t)$ then satisfies

$$\begin{aligned}\mathrm{H}(U_1,V_1,\ldots,V_t)&=\mathrm{H}(U_1)+\sum_{j=1}^{t}\mathrm{H}(V_j|U_1) &(11)\\&=\mathrm{H}(U)+t\,\mathrm{H}(V|U) &(12)\\&=t\,\mathrm{H}(U,V)-(t-1)\,\mathrm{H}(U) &(13)\\&=t\log(\delta n_1 n_2)-(t-1)\,\mathrm{H}(U) &(14)\\&\geq t\log(\delta n_1 n_2)-(t-1)\log n_1 &(15)\\&=\log(\delta^t n_1 n_2^t), &(16)\end{aligned}$$

where (11) holds by the chain rule of the Shannon entropy, since (by construction) $V_1,\ldots,V_t$ are conditionally independent given U (see (8)) and since $(U_1,V_1,\ldots,V_t)\sim(U,V_1,\ldots,V_t)$; (12) relies on the property $(U_i,V_j)\sim(U,V)$; (13) holds by a second application of the chain rule; (14) holds by (7), and finally (15) follows from the uniform bound, which states that if X is a discrete random variable supported on a finite set $\mathcal{S}$, then $\mathrm{H}(X)\leq\log|\mathcal{S}|$. In this case, $\mathrm{H}(U)\leq\log|\mathcal{U}|=\log n_1$. Consequently, the joint entropy of the random vector $(U_1,\ldots,U_s,V_1,\ldots,V_t)$ satisfies

$$\begin{aligned}\mathrm{H}(U_1,\ldots,U_s,V_1,\ldots,V_t)&=\mathrm{H}(V_1,\ldots,V_t)+\sum_{i=1}^{s}\mathrm{H}(U_i|V_1,\ldots,V_t) &(17)\\&=\mathrm{H}(V_1,\ldots,V_t)+s\,\mathrm{H}(U_1|V_1,\ldots,V_t) &(18)\\&=s\big[\mathrm{H}(V_1,\ldots,V_t)+\mathrm{H}(U_1|V_1,\ldots,V_t)\big]-(s-1)\,\mathrm{H}(V_1,\ldots,V_t) &(19)\\&=s\,\mathrm{H}(U_1,V_1,\ldots,V_t)-(s-1)\,\mathrm{H}(V_1,\ldots,V_t) &(20)\\&\geq s\log(\delta^t n_1 n_2^t)-(s-1)\,\mathrm{H}(V_1,\ldots,V_t) &(21)\\&\geq s\log(\delta^t n_1 n_2^t)-(s-1)\log(n_2^t) &(22)\\&=\log(\delta^{st} n_1^s n_2^t), &(23)\end{aligned}$$

where (17) holds by the chain rule and since (by construction) the random variables $U_1, \ldots, U_s$ are conditionally independent given $V_1, \ldots, V_t$ (see (9)); (18) holds since, by construction, all the U_i's ($i \in [s]$) are identically conditionally distributed given $(V_1, \ldots, V_t)$ (see (10)); (19) is simple algebra; (20) holds by another use of the chain rule; (21) holds by (16), and finally (22) holds by the uniform bound which implies that $\mathrm{H}(V_1, \ldots, V_t) \leq \log(|\mathcal{V}|^t) = \log(n_2^t)$.

Each vector $(U_1, \ldots, U_s, V_1, \ldots, V_t)$ can be mapped to a homomorphism from $\mathsf{K}_{s,t}$ to G via an injective mapping, where each vertex in the partite set of size s in $\mathsf{K}_{s,t}$ is mapped to a vertex from the partite set of size n_1 in G, and each vertex in the partite set of size t in $\mathsf{K}_{s,t}$ is mapped to a vertex in the second partite set of size n_2 in G. Denote the subset of such homomorphisms by $\mathcal{H}_1 \subseteq \mathrm{Hom}(\mathsf{K}_{s,t}, \mathsf{G})$. To define this mapping explicitly, label the vertices of the complete bipartite graph $\mathsf{K}_{s,t}$ by the elements of $[s+t]$, assigning the labels $1, \ldots, s$ to the vertices in the partite set of size s, and the labels $s+1, \ldots, s+t$ to those in the second partite set of size t. For every $i \in [s]$, map vertex $i \in \mathsf{V}(\mathsf{K}_{s,t})$ to vertex $U_i \in \mathsf{V}(\mathsf{G})$, and for every $j \in [t]$, map vertex $s+j \in \mathsf{V}(\mathsf{K}_{s,t})$ to vertex $V_j \in \mathsf{V}(\mathsf{G})$. Under this mapping, each edge $\{i, s+j\} \in \mathsf{E}(\mathsf{K}_{s,t})$ is mapped to the edge $\{U_i, V_j\} \in \mathsf{E}(\mathsf{G})$, thereby defining a homomorphism $\mathsf{K}_{s,t} \to \mathsf{G}$ in $\mathcal{H}_1$, since $\{U_i, V_j\} \in \mathsf{E}(\mathsf{G})$ holds by construction (see (10)). Recall that in (10), $\{U, V\}$ is uniformly distributed over the edges of the graph G, where $U \in \mathcal{U}$ and $V \in \mathcal{V}$ (by construction), P_U denotes the marginal PMF of U, and $\mathsf{P}_{V|U}$ denotes the conditional PMF of V given U. The suggested mapping is injective since it maps distinct such vectors to distinct homomorphisms in $\mathcal{H}_1 \subseteq \mathrm{Hom}(\mathsf{K}_{s,t}, \mathsf{G})$. By (2) and the uniform bound, it then follows that

$$\mathrm{H}(U_1, \ldots, U_s, V_1, \ldots, V_t) \leq \log |\mathcal{H}_1|. \tag{24}$$

Combining (23) and (24) yields

$$|\mathcal{H}_1| \geq \delta^{st} n_1^s n_2^t. \tag{25}$$

Likewise, denote by $\mathcal{H}_2$ the subset of homomorphisms $\mathsf{K}_{s,t} \to \mathsf{G}$, where each vertex in the partite set of size s in $\mathsf{K}_{s,t}$ is mapped to a vertex in the partite set of size n_2 in G, and each vertex in the partite set of size t in $\mathsf{K}_{s,t}$ is mapped to a vertex in the other partite set of size n_1 in G. Similarly to (25), we get

$$|\mathcal{H}_2| \geq \delta^{st} n_1^t n_2^s. \tag{26}$$

Since the subsets $\mathcal{H}_1$ and $\mathcal{H}_2$ form a partition of the set $\mathrm{Hom}(\mathsf{K}_{s,t}, \mathsf{G})$ (as G is a nonempty bipartite graph), it follows from (25) and (26) that

$$\mathrm{hom}(\mathsf{K}_{s,t}, \mathsf{G}) = |\mathcal{H}_1| + |\mathcal{H}_2| \geq \delta^{st} (n_1^s n_2^t + n_1^t n_2^s), \tag{27}$$

which proves the leftmost inequality in (5). □

An equivalent form of the leftmost inequality in (5) is next obtained, using the identity

$$\mathrm{hom}(\mathsf{K}_{s,t}, \mathsf{K}_{n_1,n_2}) = n_1^s n_2^t + n_1^t n_2^s. \tag{28}$$

Corollary 3.1. Let G be a simple bipartite graph with partite sets of sizes n_1 and n_2, no isolated vertices, and $\delta n_1 n_2$ edges for some $\delta \in (0, 1]$. Then, for all $s, t \in \mathbb{N}$,

$$\mathrm{hom}(\mathsf{K}_{s,t}, \mathsf{G}) \geq \delta^{st} \, \mathrm{hom}(\mathsf{K}_{s,t}, \mathsf{K}_{n_1,n_2}). \tag{29}$$

In particular, if $\mathsf{G} = \mathsf{K}_{n_1,n_2}$, then inequality (29) holds with equality (as $\delta = 1$).

Definition 3.1 (Sidorenko graph)**.** A graph H is said to be Sidorenko if it has the property that for every graph G

$$t(\mathsf{H},\mathsf{G}) \geq t(\mathsf{K}_2,\mathsf{G})^{\,e(\mathsf{H})}, \tag{30}$$

where $e(\mathsf{H}) \triangleq |\mathsf{E}(\mathsf{H})|$, or, by (3) and (4),

$$\frac{\hom(\mathsf{H},\mathsf{G})}{v(\mathsf{G})^{\,v(\mathsf{H})}} \geq \left(\frac{2\,e(\mathsf{G})}{v(\mathsf{G})^2}\right)^{e(\mathsf{H})}. \tag{31}$$

Therefore, a graph H is said to be Sidorenko if the probability that a random uniform mapping from its vertex set $\mathsf{V}(\mathsf{H})$ to the vertex set of any graph G forms a homomorphism is at least the product over all edges in H of the probability that the edge is mapped to an edge in G.

Sidorenko's conjecture states that every bipartite graph is Sidorenko. Although this conjecture remains an open problem in its full generality, it is known that every bipartite graph containing a vertex adjacent to all vertices in its other part is Sidorenko (see, e.g., [29, Theorem 5.5.14], originally proved in [30], and simplified in [31]). Additional classes of bipartite graphs that are Sidorenko have been established in [32].

Discussion 3.1 (Comparison to Sidorenko's lower bound)**.** Every complete bipartite graph is known to be Sidorenko (see [29, Theorem 5.5.12]). Specializing (31) to a complete bipartite graph $\mathsf{H} = \mathsf{K}_{s,t}$, where $s,t \in \mathbb{N}$, yields inequality (31) with $v(\mathsf{H}) = s+t$ and $e(\mathsf{H}) = st$. Let us now further specialize it to the case where G is a simple bipartite graph with partite sets of sizes n_1 and n_2, has no isolated vertices, and contains $\delta n_1 n_2$ edges for some $\delta \in (0,1]$. In this specialized setting, (31) gives

$$\hom(\mathsf{K}_{s,t},\mathsf{G}) \geq (2\delta)^{st}(n_1+n_2)^{s+t-2st}(n_1n_2)^{st} \triangleq \mathrm{LB}_1. \tag{32}$$

This lower bound on $\hom(\mathsf{K}_{s,t},\mathsf{G})$ is compared to the bound $\mathrm{LB}_2 \triangleq \delta^{st}\left(n_1^s n_2^t + n_1^t n_2^s\right)$, which appears as the leftmost inequality in (5). To compare these two lower bounds, which are symmetric in n_1 and n_2 and also in s and t, we examine the ratio $\frac{\mathrm{LB}_2}{\mathrm{LB}_1}$. Without loss of generality, assume that $p \geq q$, and let $r \triangleq \frac{\max\{n_1,n_2\}}{\min\{n_1,n_2\}} \geq 1$. By straightforward algebra, we get

$$\begin{aligned}\frac{\mathrm{LB}_2}{\mathrm{LB}_1} &= 2^{-s}\left(\frac{(1+r)^2}{2r}\right)^{s(t-1)}(1+r)^{s-t}\left(1+r^{-(s-t)}\right) && (33)\\ &\geq 2^{st-(s+t)}\left(1+r^{-|s-t|}\right). && (34)\end{aligned}$$

By the symmetry of the right-hand side of (34) in s and t, the earlier assumption that $s \geq t$ can be dropped. Consequently, the following cases hold:

(1) If $s = t$, then it follows from (34) that $\mathrm{LB}_2 \geq 2^{(s-1)^2}\,\mathrm{LB}_1$, and in particular, $\mathrm{LB}_2 \geq \mathrm{LB}_1$.

(2) Else, if $s > 1$ and $t = 1$ (i.e., $\mathsf{K}_{s,t}$ is a star graph), then by Jensen's inequality

$$\frac{\mathrm{LB}_2}{\mathrm{LB}_1} = \frac{\frac{1}{2}\left(n_1^{1-s}+n_2^{1-s}\right)}{\left(\frac{n_1+n_2}{2}\right)^{1-s}} \geq 1, \tag{35}$$

so $\mathrm{LB}_2 \geq \mathrm{LB}_1$. Due to symmetry in s and t, it also holds if $s = 1$ and $t > 1$.

(3) Otherwise (i.e., if $s,t \geq 2$ and $s \neq t$), we get from (34) that $\mathrm{LB}_2 > 2^{st-(s+t)}\,\mathrm{LB}_1 \geq 2\,\mathrm{LB}_1$.

To conclude, the lower bound on $\hom(\mathsf{K}_{s,t},\mathsf{G})$ in (5) compares favorably to Sidorenko's lower bound given in (32).

4 Concluding Remarks

We highlight several results from the full paper version on arXiv [18], which investigates homomorphism counts from arbitrary bipartite source graphs to bipartite target graphs. In this extended abstract, we focus only on the second and fourth items, choosing to present two results with proofs rather than listing all findings without any proof. The full version [18] also contains observations and numerical results that are omitted here due to space limitations.

- Combinatorial and two entropy-based lower bounds are derived for complete bipartite source graphs.
- The first entropy-based bound was also introduced here, which depends only on the sizes of the partite sets in the source and target graphs, along with the edge density of the target graph.
- The second entropy-based lower bound further incorporates the degree profiles within the target's partite sets, yielding a strengthening of the first.
- Both entropy-based bounds improve upon the inequality that is implied by Sidorenko's conjecture for complete bipartite graph sources.
- These lower bounds, combined with new auxiliary results, yield general bounds on homomorphism counts between arbitrary bipartite graphs.
- A known reverse Sidorenko inequality (by Sah, Sawhney, Stoner, and Zhao, 2020) is used to derive corresponding upper bounds.
- Numerical comparisons with exact counts in tractable cases support the effectiveness of the proposed computable bounds.

References

[1] N. Alon, "Combinatorial reasoning in information theory," 2009. here

[2] I. Csiszár, "The method of types," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2505–2523, October 1998. here

[3] I. Csiszár and J. Körner, *Information Theory, Coding Theorems for Discrete Memoryless Systems*, Second Edition, Cambridge University Press, 2011. here

[4] M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, Sixth Edition, Springer, Berlin, 2018. here

[5] F. R. K. Chung, L. R. Graham, P. Frankl and J. B. Shearer, "Some intersection theorems for ordered sets and graphs," *Journal of Combinatorial Theory, Series A*, vol. 43, no. 1, pp. 23–37, 1986. here

[6] S. Jukna, *Extremal Combinatorics with Applications in Computer Science*, second edition, Springer, 2011. here

[7] D. Galvin, "Three tutorial lectures on entropy and counting," *Proceedings of the 1st Lake Michigan Workshop on Combinatorics and Graph Theory*, Kalamazoo, MI, USA, March 2014. here

[8] N. Pippenger, "An information-theoretic method in combinatorial theory," *Journal of Combinatorial Theory, Series A*, vol. 23, no. 1, pp. 99–104, July 1977. here

[9] N. Pippenger, "Entropy and enumeration of Boolean functions," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2096–2100, September 1999. here

[10] J. Radhakrishnan, "An entropy proof of Bregman's theorem," *Journal of Combinatorial Theory, Series A*, Elsevier Science, vol. 77, no. 1, pp. 161–164, January 1997. here

[11] J. Radhakrishnan, "Entropy and counting," *Proceedings of the IIT Kharagpur, Golden Jubilee Volume on Computational Mathematics, Modelling and Algorithms*, Narosa Publishers, New Delhi, India, pp. 1–25, 2001. here

[12] E. Friedgut, "Hypergraphs, entropy, and inequalities," *The American Mathematical Monthly*, vol. 111, no. 9, pp. 749–760, November 2004. here

[13] J. Kahn, "An entropy approach to the hard-core model on bipartite graphs," *Combinatorics, Probability and Computing*, vol. 10, no. 3, pp. 219–237, May 2001. here

[14] J. Kahn, "Entropy, independent sets and antichains: a new approach to Dedekind's problem," *Proceedings of the American Mathematical Society*, vol. 130, no. 2, pp. 371–378, June 2001. here

[15] M. Madiman and P. Tetali, "Information inequalities for joint distributions, interpretations and applications," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2699–2713, June 2010. here

[16] I. Sason, "A generalized information-theoretic approach for bounding the number of independent sets in bipartite graphs," *Entropy*, vol. 23, no. 3, paper 270, pp. 1–14, 2021. here

[17] I. Sason, "Information inequalities via submodularity, and a problem in extremal graph theory," *Entropy*, vol. 24, no. 5, paper 597, pp. 1–31, 2022. here

[18] I. Sason, "Counting graph homomorphisms in bipartite settings," *priprint*, August 2025. here

[19] G. Brightwell and P. Winkler, "Graph homomorphisms and phase transitions," *Journal of Combinatorial Theory, Series B*, vol. 77, no. 2, pp. 221–262, November 1999. here

[20] P. Hell and J. Neštril, *Graphs and Homomorphisms*, Oxford Lecture Series in Mathematics and Its Applications, Oxford University Press, 2004. here

[21] P. Hell and J. Neštril, "Graphs and homomorphisms," *Topics in Algorithmic Graph Theory*, *Cambridge University Press* (edited by L. W. Beineke, M. C. Golumbic and R. J. Wilson), pp. 262–293, 2021. here

[22] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, "Counting graph homomorphisms," *Topics in Discrete Mathematics*, pp. 315–371, Springer, 2006. here

[23] P. Csikvári, N. Ruozzi, and S. Shams, "Markov random fields, homomorphism counting, and Sidorenko's conjecture," *IEEE Transactions on Information Theory*, vol. 68, no. 9, pp. 6052–6062, September 2022. here

[24] E. Friedgut and J. Kahn, "On the number of copies of one hypergraph in another," *Israel Journal of Mathematics*, vol. 105, pp. 251–256, 1998. here

[25] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society, 2012. here

[26] Z. Wang, J. Tu, and R. Lang, "Entropy, graph homomorphisms, and dissociation sets," *Entropy*, vol. 25, paper. 163, pp. 1–11, January 2023. here

[27] A. Sidorenko, "A correlation inequality for bipartite graphs," *Graphs and Combinatorics*, vol. 9, pp. 201–204, June 1993. here

[28] M. Simonovits, "Extremal graph problems, degenerate extremal problems and supersaturated graphs," Progress in graph theory, pp. 419–437, Academic Press, Toronto, ON, 1984. here

[29] Y. Zhao, *Graph Theory and Additive Combinatorics: Exploring Structures and Randomness*, Cambridge University Press, 2023. here

[30] D. Conlon, J. Fox, and B. Sudakov, "An approximate version of Sidorenko's conjecture," *Geometric and Functional Analysis*, vol. 20, pp. 1354–1366, October 2010. here

[31] D. Conlon, J. Fox, and B. Sudakov, "Sidorenko's conjecture for a class of graphs: an exposition," *unpublished note*, 2010. here

[32] D. Conlon, J. H. Kim, C. Lee, and J. Lee, "Some advances on Sidorenko's conjecture," *Journal of the London Mathematical Society*, vol. 98, no. 3, pp. 593–608, December 2018. here

[33] B. Szegedy, "An information theoretic approach to Sidorenko's conjecture," preprint, January 2015. here

[34] B. Szegedy, "Sparse graph limits, entropy maximization and transitive graphs," preprint, April 2015. here

[35] A. Sah, M. Sawhney, D. Stoner, and Y. Zhao, "A reverse Sidorenko inequality," *Inventiones Mathematicae*, vol. 221, pp. 665–711, August 2020. here

[36] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, San Francisco, 1979.

[37] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, second edition, John Wiley & Sons, 2006. here

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25–29, 2025

On H-Intersecting Graph Families

(Extended abstract)

Igal Sason*

Abstract

This paper applies the combinatorial version of Shearer's inequalities to derive a new upper bound on the maximum cardinality of a family of graphs on a fixed number of vertices, in which the intersection of every two graphs in that family contains a subgraph that is isomorphic to a specified graph H. Such families are referred to as H-intersecting graph families. The derived bound is expressed in terms of the chromatic number of H, extending the bound by Chung, Graham, Frankl, and Shearer (1986) with H specialized to a triangle.

1 Introduction

An H-intersecting family of graphs is a collection of finite, undirected, and simple graphs (i.e., graphs with no self-loops or parallel edges), whose vertices are labelled, and the intersection of every two graphs in the family contains a subgraph isomorphic to H. For instance, if H is an edge or a triangle, then every pair of graphs in the family shares at least one edge or triangle, respectively. These intersecting families of graphs play a central role in extremal combinatorics and graph theory, where determining their maximum possible size remains a longstanding challenge. Different choices of H lead to distinct combinatorial problems and structural constraints.

A pivotal conjecture, proposed in 1976 by Simonovits and Sós, concerned the maximum size of triangle-intersecting graph families—those in which the intersection of any two graphs contains a triangle. Their foundational work, initially presented in [1], along with other results on intersection theorems for families of graphs where the shared subgraphs are cycles or paths, was surveyed in [2]. The first major progress on this conjecture was made by Chung, Graham, Frankl, and Shearer [3], who utilized Shearer's inequality to establish a non-trivial bound on the largest possible cardinality of a family of triangle-intersecting graphs with a fixed number of vertices. This bound lay between the trivial bound and the conjectured bound.

The conjecture by Simonovits and Sós was ultimately resolved by Ellis, Filmus, and Friedgut [4], who proved that the largest triangle-intersecting family comprises all graphs

*Viterbi Faculty of Electrical and Computer Engineering and the Department of Mathematics, Technion—Israel Institute of Technology, Haifa 3200003, Israel. E-mail: sason@ee.technion.ac.il.

containing a fixed triangle. Building on the spectral approach in [4] (see also [5]), a recent work by Berger and Zhao [6] extended the investigation to K_4-intersecting graph families, addressing analogous questions for graph families where every pair of graphs intersects in a complete subgraph of size four. Additionally, Keller and Lifshitz [7] provided high-probability results for constructing, for every graph H, families of large random graphs with a common vertex set such that every pair of graphs contains a subgraph isomorphic to H. These are referred to as families of H-intersecting graphs.

The paper employs the combinatorial version of Shearer's lemma for upper bounding the size of H-intersecting families of graphs. An extended version of this work is available in [8].

2 Preliminaries

Definition 2.1 (Triangle-Intersecting Families of Graphs)**.** Let $\mathcal{G}$ be a family of graphs on the vertex set $[n] \triangleq \{1, \dots, n\}$, with the property that for every $\mathsf{G}_1, \mathsf{G}_2 \in \mathcal{G}$, the intersection $\mathsf{G}_1 \cap \mathsf{G}_2$ contains a triangle (i.e, there are three vertices $i, j, k \in [n]$ such that each of $\{i, j\}$, $\{i, k\}$, $\{j, k\}$ is in the edge sets of both G_1 and G_2). The family $\mathcal{G}$ is referred to as a *triangle-intersecting* family of graphs on n vertices.

The question that was posed by Simonovits and Sós [1] was how large can $\mathcal{G}$, (a family of triangle-intersecting graphs, be?

The family $\mathcal{G}$ can be as large as $2^{\binom{n}{2}-3}$. To that end, consider the family $\mathcal{G}$ of all graphs on n vertices that include a particular triangle. On the other hand, $|\mathcal{G}|$ cannot exceed $2^{\binom{n}{2}-1}$. The latter upper bound holds since, in general, a family of distinct subsets of a set of size m, where any two of these subsets have a non-empty intersection, can have a cardinality of at most 2^{m-1} ($\mathcal{A}$ and $\mathcal{A}^{\mathrm{c}}$ cannot be members of this family). The edge sets of the graphs in $\mathcal{G}$ satisfy this property, with $m = \binom{n}{2}$.

Theorem 2.1 (Ellis, Filmus, and Friedgut, [4])**.** The size of a family $\mathcal{G}$ of triangle-intersecting graphs on n vertices satisfies $|\mathcal{G}| \leq 2^{\binom{n}{2}-3}$, and it is attained by the family of all graphs with a common vertex set of n vertices, and with a fixed common triangle.

This result was proved by using discrete Fourier analysis to obtain the sharp bound in Theorem 2.1, as conjectured by Simonovits and Sós [1].

The graph K_t, with $t \in \mathbb{N}$, denotes the complete graph on t vertices, e.g., K_3 is a triangle. All results in this paper apply to finite, undirected, and simple graphs.

The first significant progress towards proving the Simonovits–Sós conjecture came from an information-theoretic approach [3]. Using the combinatorial Shearer lemma, a simple and elegant upper bound on the size of $\mathcal{G}$ was derived in [3]. That bound is equal to $2^{\binom{n}{2}-2}$, falling short of the Simonovits–Sós conjecture by a factor of 2.

Proposition 2.1 (Chung, Graham, Frankl, and Shearer, [3])**.** Let $\mathcal{G}$ be a family of K_3-intersecting graphs on a common vertex set $[n]$. Then, $|\mathcal{G}| \leq 2^{\binom{n}{2}-2}$.

We next consider more general intersecting families of graphs.

Definition 2.2 (H-intersecting Families of Graphs)**.** Let $\mathcal{G}$ be a family of graphs on a common vertex set. Then, it is said that $\mathcal{G}$ is H-intersecting if for every two graphs $\mathsf{G}_1, \mathsf{G}_2 \in \mathcal{G}$, the graph $\mathsf{G}_1 \cap \mathsf{G}_2$ contains a subgraph isomorphic to H.

The combinatorial version of Shearer's lemma, presented next, was essential in [3] for deriving Proposition 2.1. It is also used later in this work to establish a nontrivial extension of that result, providing a new upper bound on the maximum cardinality of a family of graphs with a fixed number of vertices that is H-intersecting for an arbitrary nonempty graph H.

Theorem 2.2 (Combinatorial version of Shearer's lemma, [3])**.** Let $\mathscr{F}$ be a finite multiset of subsets of $[n]$ (allowing repetitions of some subsets), where each element $i \in [n]$ is included in at least $k \geq 1$ sets of $\mathscr{F}$, and let $\mathscr{M}$ be a set of subsets of $[n]$. For every set $\mathcal{S} \in \mathscr{F}$, let the trace of $\mathscr{M}$ on $\mathcal{S}$, denoted by $\text{trace}_{\mathcal{S}}(\mathscr{M})$, be the set of all possible intersections of elements of $\mathscr{M}$ with $\mathcal{S}$, i.e.,

$$\text{trace}_{\mathcal{S}}(\mathscr{M}) \triangleq \{\mathcal{A} \cap \mathcal{S} : \mathcal{A} \in \mathscr{M}\}, \quad \forall \mathcal{S} \in \mathscr{F}. \tag{1}$$

Then,

$$|\mathscr{M}| \leq \prod_{\mathcal{S} \in \mathscr{F}} \left|\text{trace}_{\mathcal{S}}(\mathscr{M})\right|^{\frac{1}{k}}. \tag{2}$$

An open problem in extremal combinatorics is, given H and n, what is the maximum size of an H-intersecting family of graphs on n labeled vertices? It was conjectured by Ellis, Filmus, and Friedgut in [4] that every K_t-intersecting family of graphs on a common vertex set $[n]$ has size at most $2^{\binom{n}{2}-\binom{t}{2}}$, with equality for the family of all graphs containing a fixed clique on t vertices. This conjecture was also proved in [4], and was recently proved by Berger and Zhao [6], while this problem is left open for $t \geq 5$.

3 Intersecting Families of Graphs

The following result generalizes Proposition 2.1 and it extends the concept of proof in [3] to hold for every family of H-intersecting graphs on a common vertex set.

Proposition 3.1 (An upper bound on the cardinality of H-intersecting graphs, [8])**.** Let H be a non-empty graph, and let $\mathcal{G}$ be a family of H-intersecting graphs on a common vertex set $[n]$. Then,

$$|\mathcal{G}| \leq 2^{\binom{n}{2}-(\chi(\mathsf{H})-1)}. \tag{3}$$

Proof.

- Identify every graph $\mathsf{G} \in \mathcal{G}$ with its edge set $\mathsf{E}(\mathsf{G})$, and let $\mathscr{M} = \{\mathsf{E}(\mathsf{G}) : \mathsf{G} \in \mathcal{G}\}$ (all these graphs have the common vertex set $[n]$).
- Let $\mathcal{U} = \mathsf{E}(\mathsf{K}_n)$. For every $\mathsf{G} \in \mathcal{G}$, we have $\mathsf{E}(\mathsf{G}) \subseteq \mathcal{U}$, and $|\mathcal{U}| = \binom{n}{2}$.
- Let $t \triangleq \chi(\mathsf{H})$. For every unordered equipartition of $[n]$ into $t-1$ disjoint subsets, i.e., $\bigcup_{j=1}^{t-1} \mathcal{A}_j = [n]$, which satisfies $\big||\mathcal{A}_i| - |\mathcal{A}_j|\big| \leq 1$ for all $1 \leq i < j \leq t-1$, let $\mathcal{U}(\{\mathcal{A}_j\}_{j=1}^{t-1})$ be the subset of $\mathcal{U}$ consisting of all those edges that lie entirely inside one of the subsets $\{\mathcal{A}_j\}_{j=1}^{t-1}$.

- We apply the combinatorial version of Shearer's lemma (Theorem 2.2) with

$$\mathscr{F} = \{\mathcal{U}(\{\mathcal{A}_j\}_{j=1}^{t-1})\}, \tag{4}$$

taken over all unordered equipartitions of $[n]$, $\{\mathcal{A}_j\}_{j=1}^{t-1}$, as described above.

- Let $m = |\,\mathcal{U}(\{\mathcal{A}_j\}_{j=1}^{t-1})\,|$, which is independent of the equipartition since

$$m = \begin{cases} (t-1)\binom{n/(t-1)}{2} & \text{if } (t-1)|n, \\ (t-2)\binom{\lfloor n/(t-1)\rfloor}{2} + \binom{\lceil n/(t-1)\rceil}{2} & \text{if } (t-1)|(n-1), \\ \qquad\vdots & \\ \binom{\lfloor n/(t-1)\rfloor}{2} + (t-2)\binom{\lceil n/(t-1)\rceil}{2} & \text{if } (t-1)|\big(n-(t-2)\big). \end{cases} \tag{5}$$

- By (5) with $t \triangleq \chi(\mathsf{H})$, it can be verified that

$$m \le \frac{1}{\chi(\mathsf{H})-1}\binom{n}{2}. \tag{6}$$

The details of that derivation are omitted and can be found in [8].

- By a simple double-counting argument in regard to the edges of the complete graph K_n (the set $\mathcal{U}$), if k is the number of elements of $\mathscr{F}$ in which each element of $\mathcal{U}$ occurs, then

$$m\,|\mathscr{F}| = \binom{n}{2}k. \tag{7}$$

- Let $\mathcal{S} \in \mathscr{F}$. Observe that $\mathrm{trace}_{\mathcal{S}}(\mathscr{M})$, as defined in (1), forms an intersecting family of subsets of $\mathcal{S}$. Indeed,

  1. Assign to each vertex in $[n]$ the index j of the subset $\mathcal{A}_j$ $(1 \le j \le \chi(\mathsf{H})-1)$ in the partition of $[n]$ corresponding to $\mathcal{S}$. Let these assignments be associated with $\chi(\mathsf{H})-1$ color classes of the vertices.
  2. For any $\mathsf{G}, \mathsf{G}' \in \mathcal{G}$, the graph $\mathsf{G} \cap \mathsf{G}'$ contains a subgraph H (by assumption).
  3. By the definition of the chromatic number of H as the smallest number of colors that are required such that any two adjacent vertices in H are assigned different colors, it follows that there exists an edge in H whose two vertices are assigned the same index (color). Hence, that edge belongs to the set $\mathcal{A}_j$, for some $j \in [\chi(\mathsf{H})-1]$, so it belongs to $\mathcal{S}$.
  4. The complement of $\mathcal{S}$ (in $\mathcal{U}$) is therefore H-free (viewed as a graph with the vertex set $[n]$).

  Consequently, since $|\mathcal{S}| = m$, we get

$$|\mathrm{trace}_{\mathcal{S}}(\mathscr{M})| \le 2^{m-1}. \tag{8}$$

- By Theorem 2.2 (and the one-to-one correspondence between $\mathcal{G}$ and $\mathscr{M}$),

$$\begin{align} |\mathcal{G}| &= |\mathscr{M}| \\ &\le \left(2^{m-1}\right)^{\frac{|\mathscr{F}|}{k}} \tag{9} \\ &= 2^{\binom{n}{2}\left(1-\frac{1}{m}\right)} \tag{10} \\ &\le 2^{\binom{n}{2}-(\chi(\mathsf{H})-1)}, \tag{11} \end{align}$$

where (9) relies on (2) and (8), then (10) relies on (7), and (11) is due to (6).

□

The family $\mathcal{G}$ of H-intersecting graphs on n vertices can be as large as $2^{\binom{n}{2}-|\mathsf{E}(\mathsf{H})|}$. To that end, consider the family $\mathcal{G}$ of all graphs on n vertices that include a particular H subgraph. Combining this lower bound on $|\mathcal{G}|$ with its upper bound in Theorem 3 gives that the largest family $\mathcal{G}$ of H-intersecting graphs on n vertices satisfies

$$2^{\binom{n}{2}-|\mathsf{E}(\mathsf{H})|} \le |\mathcal{G}| \le 2^{\binom{n}{2}-(\chi(\mathsf{H})-1)}. \tag{12}$$

Specialization of Proposition 3.1 to a family $\mathcal{G}$ that is K_t-intersecting graphs, with $t \ge 2$, on a common vertex set $[n]$, gives that $|\mathcal{G}| \le 2^{\binom{n}{2}-(t-1)}$.

The computational complexity of the chromatic number of a graph is in general NP-hard [9]. This poses a problem in calculating the upper bound in Proposition 3.1 on the cardinality of H-intersecting families of graphs on a fixed number of vertices. This bound can be loosened, expressing it in terms of the Lovász ϑ-function of the complement graph $\overline{\mathsf{H}}$.

Corollary 3.1. Let H be a graph, and let $\mathcal{G}$ be a family of H-intersecting graphs on a common vertex set $[n]$. Then,

$$|\mathcal{G}| \le 2^{\binom{n}{2}-(\lceil\vartheta(\overline{\mathsf{H}})\rceil-1)}. \tag{13}$$

Proof. The Lovász ϑ-function of the complement graph $\overline{\mathsf{H}}$ satisfies (see Corollary 3 of [10])

$$\omega(\mathsf{H}) \le \vartheta(\overline{\mathsf{H}}) \le \chi(\mathsf{H}), \tag{14}$$

so it is bounded between the clique and chromatic numbers of H, which are both NP-hard to compute [9]. Since the chromatic number $\chi(\mathsf{H})$ is an integer, we have $\chi(\mathsf{H}) \ge \lceil\vartheta(\overline{\mathsf{H}})\rceil$. Combining (3) and the latter inequality yields (13). □

The Lovász ϑ-function of the complement graph $\overline{\mathsf{H}}$, as presented in Corollary 3.1, can be efficiently computed with a precision of r decimal digits, having a computational complexity that is polynomial in $p \triangleq |\mathsf{V}(\mathsf{H})|$ and r. It is obtained by solving the following semidefinite programming (SDP) problem [11]:

$$\boxed{\begin{array}{l} \text{maximize } \operatorname{Tr}(\mathbf{B}\,\mathbf{J}_p) \\ \text{subject to} \\ \begin{cases} \mathbf{B} \in \mathcal{S}_+^p, \ \ \operatorname{Tr}(\mathbf{B}) = 1, \\ A_{i,j} = 0 \ \Rightarrow\ B_{i,j} = 0, \quad i,j \in [p],\ i \ne j, \end{cases} \end{array}} \tag{15}$$

where the following notation is used: $\mathbf{A} = \mathbf{A}(\mathsf{H})$ is the $p \times p$ adjacency matrix of H; $\mathbf{J}_p$ is the all-ones $p \times p$ matrix, and $\mathcal{S}_+^p$ is the set of all $p \times p$ positive semidefinite matrices. The reader is referred to an account of interesting properties of the Lovász ϑ-function in [12], Chapter 11 of [13], and more recently in Section 2.5 of [14].

References

[1] M. Simonovits and V. T. Sós, "Intersection theorems for graphs," *Problemes combinatoires et théorie des graphes (Colloques internationaux du Centre National de la Recherche Scientifique (CNRS)*, University of Orsay, Orsay, France, July 1976), pp. 389–391, *Colloques internationaux du CNRS*, no. 260, Paris, France, 1978. https://users.renyi.hu/~miki/OrsayB.pdf

[2] M. Simonovits and V. T. Sós, "Intersection theorems on structures," *Annals of Discrete Mathematics*, vol. 6, pp. 301–313, 1980. https://doi.org/10.1016/S0167-5060(08)70715-5

[3] F. R. K. Chung, L. R. Graham, P. Frankl and J. B. Shearer, "Some intersection theorems for ordered sets and graphs," *Journal of Combinatorial Theory, Series A*, vol. 43, no. 1, pp. 23–37, 1986. https://doi.org/10.1016/0097-3165(86)90019-1

[4] D. Ellis, Y. Filmus, and E. Friedgut, "Triangle-intersecting families of graphs," *Journal of the European Mathematical Society*, vol. 14, no. 3, pp. 841–855, 2012. https://doi.org/10.4171/JEMS/320

[5] D. Ellis, "Intersection problems in extremal combinatorics: theorems, techniques and questions - old and new," *Surveys in Combinatorics 2022*, pp. 115–173. London Mathematical Society Lecture Note Series, Cambridge University Press, 2022. https://doi.org/10.1017/9781009093927.005

[6] A. Berger and Y. Zhao, "K_4-intersecting families of graphs," *Journal of Combinatorial Theory, Series B*, vol. 163, pp. 112–132, November 2023. https://doi.org/10.1016/j.jctb.2023.07.005

[7] N. Keller and N. Lifshitz, "A note on large H-intersecting families," *SIAM Journal on Discrete Mathematics*, vol. 33, no. 1, pp. 398–401, 2019. https://doi.org/10.1137/18M1220765

[8] I. Sason, "On H-intersecting graph families and counting of homomorphisms," *AIMS Mathematics*, vol. 10, no. 3, paper 290, pp. 6355–6378, March 2025. https://doi.org/10.3934/math.2025290

[9] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, San Francisco, 1979.

[10] L. Lovász, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1–7, January 1979. https://doi.org/10.1109/TIT.1979.1055985

[11] M. Grötschel, L. Lovász, A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica*, **1** (1981), 168–197. https://doi.org/10.1007/BF02579273

[12] D. E. Knuth, The sandwich theorem, *Electron. J. Combin.*, **1** (1994), 1–48. https://doi.org/10.37236/1193

[13] L. Lovász, *Graphs and geometry*, American Mathematical Society, **65** (2019). `https://doi.org/10.1090/coll/065`

[14] I. Sason, Observations on graph invariants with the Lovász ϑ-function, *AIMS Math.*, **9** (2024), 15385–15468. `https://doi.org/10.3934/math.2024747`

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Loops in random duals

(Extended abstract)

Robert Šámal*

Abstract

We study random graph embeddings, as defined by Stahl (JCTB 1991). In particular, we study the number of loops in the dual graph given by the random embedding. The motivation for this comes from Cycle Double Cover conjecture, which can be equivalently expressed as a search for an embedding of a given bridgeless graph that yields a dual with no loops. We provide results for some graph classes and pose several open questions.

1 Introduction

A cycle double cover (CDC for short) is a collection of cycles (2-regular connected subgraphs) that cover every edge exactly twice. The famous CDC conjecture claims existence of CDC for every bridgeless graph. Many partial results are known but the full solution remains out of reach. We refer to [4] for more detailed exposition to the problem and related results. For our purposes we mention that it is sufficient to prove the CDC conjecture for 3-regular graphs and that for such graphs the conjecture is equivalent to the existence of embedding on a surface, where every face boundary is a cycle (so-called circular 2-cell embedding).

To explain why a face boundary may not be a cycle: every edge of a graph is incident to a face on both of its "sides". However, both sides may be incident to the same face – in which case the boundary of such face is a closed walk, that is not a cycle. We call this type of edge *singular*, while the well-behaved edge (incident to two different faces) is called *regular*. So, the CDC conjecture asks us to get an embedding with all edges regular. Every 2-cell embedding (where each face has interior isomorphic to an open disk) yields a dual graph; it this dual singular edges turn into loops, while regular edges to normal edges. This explains half of the title. Before we get to the other half, we pause to specify two types of singular edges: a *negative singular* edge is such that is traversed on both sides by the same face in the same direction, while *positive singular* is traversed by the same edge, but in opposite directions. Note that the classification of edges as regular, positive singular, and negative singular is dependent on the embedding.

*Computer Science Institute of Charles University, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. Email: `samal@iuuk.mff.cuni.cz`. Supported by grant 25-16627S of the Czech Science Foundation.

There is a large body of work on embedded graphs; see, e.g. the book by Mohar and Thomassen [2]. We need only the concept of combinatorial embedding that allows us to describe the embedding of a graph on a surface without explicitly mentioning the surface: we specify a tuple (π, λ), where $\pi = (\pi_v)_{v \in V(G)}$ is a collection of *local rotations*: π_v is a cyclic ordering (permutation with one cycle) of edges incident with v (or, when speaking of simple graphs, we will consider π_v a permutation of neighbors of v). Further, $\lambda = (\lambda_e)_{e \in E(G)}$ is a collection of signs on edges: $\lambda_e = 1$ is a typical "planar" edge, while $\lambda_e = -1$ is an edge going through a crosscap, that makes a "twist" between left and right side of the face. See [2] for details. Here we only remind, that in orientable embeddings we have $\lambda_e \equiv 1$.

Next, we turn to the concept of random embeddings. Introduced by Stahl [3], an *orientable random embedding* of a graph G is given by choosing uniformly at random π_v from all unicyclic permutations, with all choices independent. We let $\lambda_e = 1$ for all edges in this case. In an *nonorientable random embedding* we add to this a random choice of $\lambda_e \in \{\pm 1\}$ for each $e \in E(G)$.

So far, the results about random embedding can be split into two categories. First of them is an explicit description of the probability distribution of genus (equivalently, number of faces) in a random embedding. This was motivated by the conjecture that this distribution is unimodal and log-concave (the latter was recently disproved by Mohar). The second part of literature deals with the expectation of genus, that sheds light on the extremal results in the theory of embedded graphs – putting more flavor to results about a maximal and minimal genus of a graph.

In this paper we want to expand beyond that. Motivated by the CDC conjecture, we study the distribution of the number of singular edges, that is the number of loops in the dual. While our results don't come close to getting a CDC (in fact many loops in the dual are typical), let us mention here a conjecture of the author and Hušek [1]: a bridgeless cubic graph with n vertices has at least $2^{n/2-1}$ cycle double covers. This bold strengthening of CDC conjecture is in fact supported by numerical evidence, and suggests that clever probabilistic techniques can be relevant in a search for CDC.

2 Results

The main results of our paper are the following:

Theorem 1. *In a random orientable embedding of K_n every edge is singular (loop of the dual) with probability $\frac{1}{2} + o(1)$. Consequently, the expected number of loops of the random dual is $\frac{1}{2}\binom{n}{2} + o(n^2)$.*

Theorem 2. *In a random nonorientable embedding of K_n every edge is positive singular, negative singular, or regular, each with probability $\frac{1}{3} + o(1)$. Consequently, the expected number of loops of the random dual is $\frac{2}{3}\binom{n}{2} + o(n^2)$.*

As K_n is edge-transitive, it is no surprise that all edges have the same probabilities. For the circular ladder CL_n – the Cartesian product of C_n and K_2 – there is some difference between probabilities of different edges.

Theorem 3. *In a random orientable embedding of CL_n each of its outer and inner edge is singular (loop of the dual) with probability $\frac{5}{9} + o(1)$, while each spoke edge is singular with probability $\frac{1}{3} + o(1)$ Consequently, the expected number of loops of the random dual is $\frac{13}{27}n + o(n)$.*

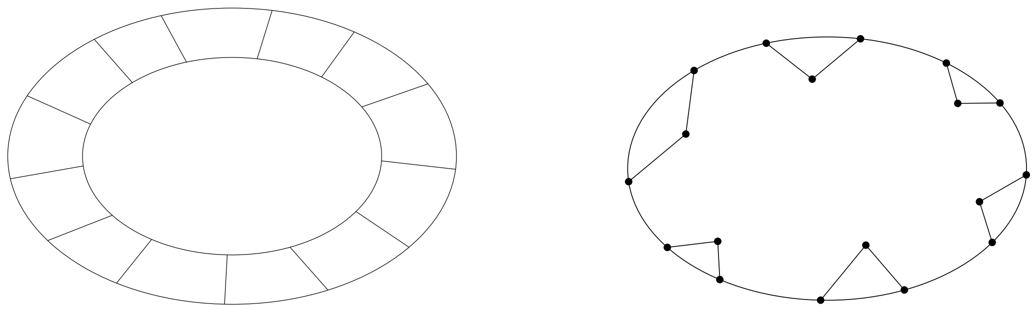


Figure 1: On the left, the circular ladder CL_{13}, on the right the graph H_{12}.

On the other hand, there are graphs with very different behavior of different edges. Let H_n be the graph C_{2n} with "every other edge doubled" (see Figure 1). Readers who don't like parallel edges can subdivide one edge in each parallel pair. For ease of formulation, we call edge of H_n bridge-like if it is in a 2-edge-cut and cyclic otherwise.

Theorem 4. *In a random orientable embedding of H_n a bridge-like e is singular with probability $1 - o(1)$. A cyclic edge is singular with probability $o(1)$.*

Theorem 5. *In a random nonorientable embedding of H_n a bridge-like e is singular with probability $1 - o(1)$. A cyclic edge is negative singular with probability $\frac{1}{2} + o(1)$ and regular with probability $\frac{1}{2} + o(1)$.*

3 Proofs

In this extended abstract we present just sketches of some of the proofs.

Proof sketch of Theorem 1. We first provide an informal idea, then we make it more precise. Let $V(G) = \{1, \dots, n\}$, we consider the face starting with edge $(1, 2)$. Let $x = \pi_2^{-1}(1)$ ($x \neq 1$ as we only consider $n > 3$). We imagine the boundary of face starting with $(1, 2)$ as a type of random walk, when we do a partial revealing of the random embedding. If this random walk meets edge $(x, 2)$ before $(1, 2)$ then the edge $\{1, 2\}$ will be visited also in the other direction, thus it will be a singular edge. And vice versa, if we meet $(1, 2)$ (in this direction) before $(x, 2)$, we close the face and $\{1, 2\}$ will be regular. Thus, we need to prove that we have (approximately) the same probability of meeting these two edges.

To make this formal, let Ω be our probabilistic space, that is the set of all tuples of cyclic rotations. We provide a mapping $f : \Omega \to \Omega$. This mapping is illustrated on Figure 2 below. Formally, to define $f(\pi)$ we let $x := \pi_2^{-1}(1)$. Then we "switch x with 1 almost everywhere" – we exchange π_x with π_1 and we exchange x with 1 at π_v with $v \neq 2$.

It is easy to show that when restricted to embeddings where $\{2, x\}$ is a regular edge, then mapping f exchanges embeddings where $\{1, 2\}$ is regular with those where this edge is singular. We leave the rest of the argument for the full paper. Note that f is only a.a.s. bijection.

□

Proof sketch of Theorem 2. Follows the same idea as the proof above, except we now have two possibilities to make edge $\{1, 2\}$ singular: we can enter it from 2 or from 1 on the "other side" (that is, from $\pi_1^{-1}(2)$). □

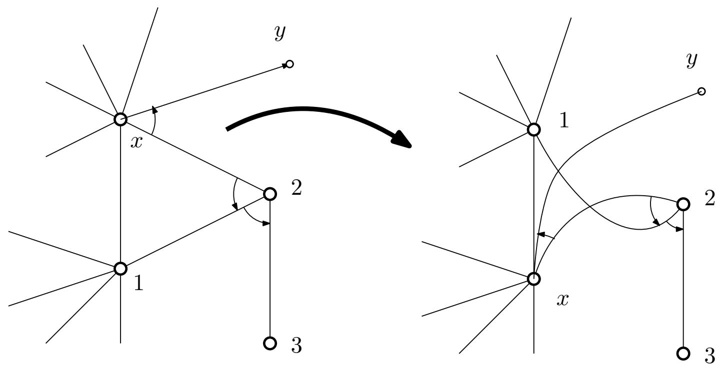


Figure 2: Illustration of the proof of Theorem 1.

Proof sketch of Theorem 4. Label vertices $v_1, v_2, \ldots, v_{2n}$. Consider a bridge-like edge $e = v_1v_2$ and the facial walks on each of its sides. At v_3 these two walks connect to a single face with probability $1/2$, and continue "in the same way" with probability $1/2$. Thus, for large enough n the probability of e being singular is close to 1. ☐

Proof sketch of Theorem 3. We proceed similarly as in the last proof sketch, except now the number of possibilities at each spoke is larger. Still we can write it explicitly. In fact, any graph sequence of this structure ("cyclically repeated gadget") can be analyzed using techniques of [1]: we describe the effect of the "gadget" by a linear mapping and analyze its eigenvalues. ☐

4 Future work

The most pressing question is a qualitative explanation of differences in the behavior of H_n and CL_n. We tentatively propose, that the size of the smallest cut containing an edge is the distinguishing factor.

Further, we would like to extend the study of loops to more properties of the dual, such as the number of short cycles. In other words, we want to study the representativity of random graph embeddings.

References

[1] Radek Hušek and Robert Šámal. Counting circuit double covers. *J. Graph Theory*, 108(2):374–395, 2025.

[2] Bojan Mohar and Carsten Thomassen. *Graphs on Surfaces.* Johns Hopkins series in the mathematical sciences. Johns Hopkins University Press, 2001.

[3] Saul Stahl. An upper bound for the average number of regions. *J. Combin. Theory Ser. B*, 52(2):219–221, 1991.

[4] Cun-Quan Zhang. *Circuit Double Cover of Graphs.* London Mathematical Society Lecture Note Series. Cambridge University Press, 2012.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Typical properties of countable graphs: flows and bridges

(Extended abstract)

Robert Šámal[*†] Hadi Zamani[*‡]

Abstract

In this paper, we have two aims: first, we study nowhere-zero flows in countably infinite graphs, observing the different roles of two types of bridges. Then, we turn to studying how typical graphs with these types of bridges are. As there is no natural probability measure to use, we turn to an analytic approach: we study which sets are nowhere-dense and which are meagre.

In studying flows in finite graphs, bridges play a crucial role – it is easy to see that every flow assigns a zero value to every bridge. It is also known that graphs with a bridge are not typical. Both dense random graphs and random regular graphs are bridgeless w.h.p. [3, 1]. In this paper, we study how this transfers to the setting of countable graphs. In Section 1, we explain the differences in the relationship between bridges and flows. We apply compactness to prove a version of Seymour's 6-flow theorem for countable graphs.

In Section 2, we explore the concepts of nowhere-dense and meagre sets. We start right here explaining why this is called for: When we extend the Erdős-Rényi random graphs to countably infinite graphs, we find that a.s. we get the Rado graph, i.e., we are only studying one graph, up to isomorphism. As another attempt, we can try to extend random regular graphs, perhaps via the configuration model [1]. However, for this, we would need a uniform probability distribution on perfect matchings on a countable set. In particular, we would need a uniform probability distribution on a countable set (namely, the distribution of neighbors of the first vertex). As this is not possible, this approach also fails. Consequently, we turn to another way to describe typical sets, common in real analysis. We study meagre/co-meagre sets (also called sets of first and of second category).

[*]Computer Science Institute of Charles University, Prague, Czech Republic. Supported by grant 25-16627S of the Czech Science Foundation. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810115).

[†]E-mail: samal@iuuk.mff.cuni.cz.

[‡]E-mail: hadiz@iuuk.mff.cuni.cz.

1 Nowhere-zero flows for infinite graphs

Definition 1 (weak flow) *A weak A-flow in a graph G given an orientation D over an abelian group A is a function $\varphi : E(G) \to A$ such that the following condition holds.*
(Kirchhoff's law for each vertex cut):

$$\forall v \in V(G) : \sum_{e \in \delta^+(v)} \varphi(e) = \sum_{e \in \delta^-(v)} \varphi(e).$$

Definition 2 (strong flow) *A strong A-flow in a graph G given an orientation D over an abelian group A is a function $\varphi : E(G) \to A$ such that the following condition holds.*
(Kirchhoff's law for each cut):

$$\forall S \subseteq V(G) : \sum_{e \in \delta^+(S)} \varphi(e) = \sum_{e \in \delta^-(S)} \varphi(e).$$

In both cases, we call φ a nowhere-zero weak/strong A-flow if $\varphi^{-1}(0) = \emptyset$.

Definition 3 *An edge e of a graph G is a bridge if deleting e increases the number of connected components of G. A bridge is called a 2-way infinite bridge if both new components are infinite and a 1-way infinite bridge otherwise.*

Every strong flow is also a weak flow. While these two concepts are equivalent in finite graphs, they differ in infinite graphs. For instance, in a 2-way infinite path, a nowhere-zero weak A-flow always exists, whereas a strong nowhere-zero flow does not. In [6], the authors constructed a 3-connected graph with two ends that admits a nowhere-zero weak A-flow but not a nowhere-zero strong A-flow. They also extended nearly all theorems of finite flow theory to infinite strong A-flows using topological compactness, including the following theorem.

Here, we use logical compactness to establish the following results concerning the existence of nowhere-zero weak 6-flows in countable locally finite graphs.

Theorem 4 *A countably infinite, locally finite graph without a 1-way infinite bridge has a weak nowhere-zero 6-flow.*

2 Metric space of rooted connected graphs

For countable graphs, we lack a good probability measure. To study what properties of graphs are typical, we follow the approach that is typical in analysis: we use a metric space of graphs and study nowhere-dense and meagre sets, which we now define properly.

We start by recalling the basic notation that we will use.

Definition 5 (ε-neighborhood) *The ε-neighborhood of a point x in a metric space (X, d) is defined as:*

$$\mathcal{U}_\varepsilon(x) = \{y \in X \mid d(x, y) < \varepsilon\}.$$

Similarly, the closed ε-neighborhood $(\overline{\mathcal{U}}_\varepsilon(x))$ includes all points satisfying $d(x, y) \leq \varepsilon$.

Definition 6 (open set) *A subset $U \subset X$ of a metric space (X, d) is said to be open if for every point $x \in U$, there exists $\varepsilon > 0$ such that $\mathcal{U}_\varepsilon(x) \subset U$.*

Definition 7 (nowhere-dense) *A set S in a metric space (X, d) is nowhere-dense if, for every $\mathcal{U}_\varepsilon(x)$, there exists a $\mathcal{U}_\delta(y) \subset \mathcal{U}_\varepsilon(x)$ such that $\mathcal{U}_\delta(y) \cap S = \emptyset$.*

Definition 8 (meagre and comeagre) *A set is meagre if it can be written as a countable union of nowhere-dense sets. A set is a comeagre if its complement is meagre.*

Next, we define the metric space under study and present the main lemma, which we use as a translation between graph theory and analysis/topology.

Definition 9 (space of rooted graphs) *Let $\mathfrak{G}^\bullet$ denote the set of all connected rooted graphs. For a natural number $D > 2$, let $\mathfrak{G}^\bullet_{\leq D}$ denote the set of* ***infinite*** *rooted connected graphs with all degrees bounded by D and $\mathfrak{G}^\bullet_{=D}$ when all degrees are equal to D. We address both of these two spaces by $\mathfrak{G}^\bullet_D$. Two graphs in these sets are considered identical if there exists a root-preserving isomorphism between them.*

Definition 10 (r-ball) *In a rooted graph G with v as the root which we write as (G, v), we define r-**ball of** (G, v) as an induced graph of G by the vertices that have distance at most r from the vertex v.*

$$B_r(G, x) = (G[\{y \in V(G) \mid d(x, y) \leq r\}], x)$$

Definition 11 (metric space $(\mathfrak{G}^\bullet, d^\bullet)$) *Define the metric on $\mathfrak{G}^\bullet$ such that for $(G_1, v_1), (G_2, v_2) \in \mathfrak{G}^\bullet$,*

$$d^\bullet((G_1, v_1), (G_2, v_2)) := \inf\left\{2^{-r} \;|B_r(G_1, v_1) \cong B_r(G_2, v_2)\right\}.$$

$(\mathfrak{G}^\bullet, d^\bullet)$ is a metric space [5].

Definition 12 ("extensions" of F) *For an r-ball $F = B_r(G, v)$ define the set of "extensions" of F,*

$$\mathfrak{G}^\bullet_F := \{(G', v') \in \mathfrak{G}^\bullet | B_{r,G'}(v') \cong (F, v)\}.$$

Lemma 13 *In the metric space $(\mathfrak{G}^\bullet, d^\bullet)$, the following properties hold:*

- *For each rooted graph (G, x), natural number r, and $0 < \varepsilon < 1$, we have*

$$\mathcal{U}_{2^{-r+\varepsilon}}((G, x)) = \overline{\mathcal{U}}_{2^{-r}}((G, x)) = \mathfrak{G}^\bullet_{B_r(G,x)}.$$

- *$\mathfrak{G}^\bullet_F$ is non-empty if F is an r-ball in $\mathfrak{G}^\bullet_D$ that contains at least one vertex of degree less than D at distance r from the root.*

- *A subset $S \subset \mathfrak{G}^\bullet$ is nowhere dense if*

$$\forall\ r\text{-ball } F,\ \exists\ r'\text{-ball } F' \text{ such that } \mathfrak{G}^\bullet_{F'} \subset \mathfrak{G}^\bullet_F \setminus S.$$

- *For all $D > 2$, $\mathfrak{G}^\bullet_D$ is not meager.*

Here we state the results of our study:

Theorem 14 *The set of bridgeless graphs is nowhere-dense,*

- *in* $\mathfrak{G}^{\bullet}_{\leq D}$ *for all* $D > 2$.
- *and in* $\mathfrak{G}^{\bullet}_{=D}$ *for all odd* $D > 2$.

In D-regular graphs for even D, 1-way infinite bridges cannot exist. Instead, we have the following:

Theorem 15 *The set of graphs without cut vertex is nowhere-dense, in* $\mathfrak{G}^{\bullet}_{D}$ *for all* $D > 2$.

Theorem 16 *The set of graphs with a 2-way infinite bridge is meagre (but not nowhere-dense), in* $\mathfrak{G}^{\bullet}_{D}$ *for all* $D > 2$.

We also have investigated some other (a)typical graph properties:

Theorem 17 *The set of graphs with edge chromatic number D is nowhere-dense in* $\mathfrak{G}^{\bullet}_{D}$ *for all* $D > 2$.

3 Proofs

Proof of Theorem 4: Let G be a graph with vertices $\{v_1, v_2, \dots\}$, where edges $\{v_i, v_j\}$ are oriented from v_i to v_j if $i < j$, and v_1 is the root.

Define G_r on $V(B_r(G, v_1)) \cup \{v_\infty\}$ as $B_r(G, v_1)$ plus edges $E_{r,\infty}$: for each $\{v_i, v_j\} \in E(G)$ with $d(v_1, v_i) = r$ and $d(v_1, v_j) = r + 1$, add $\{v_i, v_\infty\}$ to $E_{r,\infty}$ (allowing multiple edges).

For each $e \in E(G)$, introduce Boolean variables $F_{(e,1)}, \dots, F_{(e,6)}$ and enforce:

$$F_e := \bigvee_{i=1}^{6} \left(F_{(e,i)} \wedge \bigwedge_{j \neq i} \neg F_{(e,j)} \right).$$

This induces a bijection between $E(G)$ and $\bigcup_{r=0}^{\infty} E_{r,\infty}$, extending variables to G_r.

For $v_k \in V(G_r)$ with incident edges $\{\{v_k, v_{i_1}\}, \dots, \{v_k, v_{i_{\deg(v_k)}}\}\}$, define:

$$F_{v_k} := \bigvee_{\substack{(c_1,\dots,c_{\deg(v_k)}) \in \{1,\dots,6\}^{\deg(v_k)} \\ \sum_{e=\{v_k,v_{i_j}\}} c_j \cdot (-1)^{(i_j > k)} = 0}} \left(\bigwedge_{e=\{v_k,v_{i_j}\}} F_{(e,c_j)} \right),$$

enforcing Kirchhoff's law at v_k. For G_r, take $F_{G_r} := \bigwedge_{v_k \in V(G_r)} F_{v_k}$, which includes v_∞ by double counting.

By compactness [4], since each finite subset of $F_G := \{F_{G_r}\}_{r \in \mathbb{N}} \cup \{F_e\}_{e \in E(G)}$ is satisfiable via [7], F_G is satisfiable, yielding a nowhere-zero 6-flow on G.

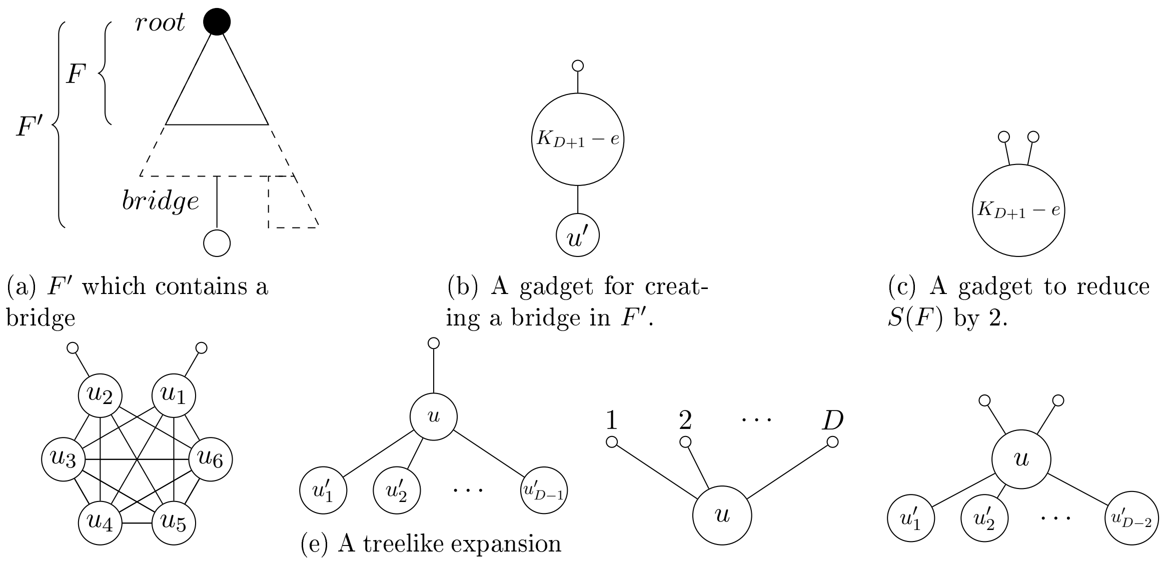
Proof of Theorem 14: To prove that the set of graphs with no bridges is nowhere-dense in $\mathfrak{G}^{\bullet}_{D}$, we must show that for each $\mathfrak{G}^{\bullet}_{F}$, there exists a non-empty open set $\mathfrak{G}^{\bullet}_{F'}$ such that every graph in $\mathfrak{G}^\bullet_{F'}$ contains a bridge. Let L_r denote the set of vertices at distance r from the root of F with degree less than D.

For $\mathfrak{G}^\bullet_{\leq D}$, we construct an r'-ball F' as follows: we add one vertex and connect it to one vertex in L_r(Figure 1a). The added edge is a bridge in F', and every graph in $\mathfrak{G}^\bullet_{F'}$ contains this edge as a bridge.

For $\mathfrak{G}^\bullet_{=D}$, we must also handle regularity. We define $S(F) := \sum_{v \in L_r}(D - \deg(v))$.

- If $S(F)$ is odd, we add $K_{D+1} - e$ gadgets (Figure 1c) to vertices in L_r until $S(F) = 1$. We then add another $K_{D+1} - e$ gadget, connecting one of its vertices to L_r and the other to a new vertex (Figure 1b). This process constructs an r'-ball F' that contains a bridge.
- If $S(F)$ is even, we add S tree-like layers (Figure 1e), resulting in an $r+2$-ball $\tilde{F}$. For this new graph, $S(\tilde{F}) = (D-1) \times S(F)$, which is even and contains more than D vertices in $L_{\tilde{r}}$. We then add a new vertex connected to D of these vertices (Figure 1f). At this point, $S(\tilde{F})$ becomes odd, allowing us to proceed as in the previous case.

Thus, every graph in the non-empty set $\mathfrak{G}^\bullet_{F'}$ contains the same bridge.



(a) F' which contains a bridge

(b) A gadget for creating a bridge in F'.

(c) A gadget to reduce $S(F)$ by 2.

(d) An example for $D = 5$.

(e) A treelike expansion to increase the size of L_r.

(f) A gadget to change the parity of $S(F)$.

(g) A gadget for creating a vertex cut.

Figure 1: Gadgets used in the construction of F'.

Proof of Theorem 15: For $\mathfrak{G}^\bullet_{\leq D}$ with $D > 2$ and $\mathfrak{G}^\bullet_{=D}$ with odd $D > 2$, the result follows directly from Theorem 14.

For $\mathfrak{G}^\bullet_{=D}$ with even $D > 2$, we proceed similarly to the previous proof. We show that for each $\mathfrak{G}^\bullet_{F}$, there exists a non-empty open set $\mathfrak{G}^\bullet_{F'}$ such that every graph in $\mathfrak{G}^\bullet_{F'}$ contains a vertex cut.

By parity, $S(F)$ is always even. If L_r contains only one vertex, that vertex is a cut vertex. If L_r contains more than one vertex, we add a new vertex and connect it to two vertices in L_r, creating a gadget as shown in Figure 1g. Since $S(F)$ remains even, we add enough gadgets (as in Figure 1c) to reduce $S(F)$ to zero. Thus, every graph in $\mathfrak{G}^{\bullet}_{F'}$ contains a cut vertex.

Proof of Theorem 16: The set of graphs with a 2-way infinite bridge is a countable union of graphs that have a 2-way infinite bridge at distance k. Therefore, it suffices to prove that having a 2-way infinite bridge at distance k is nowhere-dense.

For an r-ball F, we construct an r'-ball F' such that no 2-way infinite bridge exists at level k. If $k \leq r$, we apply the proof of Theorem 15 to construct F', which contains F and has a 1-way infinite bridge at a distance greater than k. Since one of the components must lie in a finite part of the vertex cut, this ensures that no 2-way infinite bridge exists at distance k in $\mathfrak{G}^{\bullet}_{F'}$.

If $r < k$, we extend F using treelike expansions until the size of the ball exceeds k and proceed similarly to the previous case. Thus, the set of graphs with a 2-way infinite bridge at distance k is nowhere-dense.

Proof of Theorem 17: For each D, there exists a D-regular graph G with $\chi'(G) = D + 1$. For even D, a complete graph or any regular graph of odd size has an edge chromatic number of $D + 1$. For odd D, such graphs can be constructed using Tutte's theorem.

If an arbitrary edge is removed from such a graph, its edge chromatic number remains $D+1$ due to parity constraints. Since the edge chromatic number of a graph is always greater than or equal to that of its subgraphs, we proceed as follows: In each r-ball F, we introduce $S(F)$ copies of the described gadgets and connect them to L_r, as shown in Figure 1b. This constructs F', and the resulting nonempty extensions $\mathfrak{G}^{\bullet}_{F'}$ in $\mathfrak{G}^{\bullet}_{D}$ all have an edge chromatic number equal to $D + 1$, as established in [2].

4 Future work

We suggest some future questions to continue our effort to define "probability-free" notion of typical graph property:

Question 18 *Is the fact that a property is nowhere-dense or meagre independent of the root?*

A change of root gives one type of change of metric. We can, however, define many metrics on the space of countable graphs.

Question 19 *Is there some other metric on $\mathfrak{G}^{\bullet}_{D}$ where typical properties are different? For instance, is there a metric in which bridgeless graphs are typical?*

We also want to investigate more (a)typical graph properties in the spirit of Theorem 17.

References

[1] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[2] N. G. De Bruijn and P. Erdős. A colour problem for infinite graphs and a problem in the theory of relations. *Indigationes Mathematicae*, 13:371–373, 1951.

[3] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1):261–267, 1961.

[4] K. Gödel. Die vollständigkeit der axiome des logischen funktionenkalküls. *Monatshefte für Mathematik und Physik*, 37:349–360, 1930.

[5] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[6] B. Miraftab and M.J. Moghadamzadehb. Algebraic flow theory of infinite graphs. *European Journal of Combinatorics*, 62:58–69, 2017.

[7] P.D. Seymour. Nowhere-zero 6-flows. *Journal of combinatorial theory, series B*, 30(2):130–135, 1981.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Prime graphical parking functions and strongly recurrent configurations of the Abelian sandpile model

(Extended abstract)

Thomas Selig* Haoyue Zhu†

Abstract

Parking functions were originally introduced by Konheim and Weiss (1966) in their study of hashing functions. Since then, they have become a central object in combinatorics research, with a number of variations and generalisations. Prime parking functions are parking functions that cannot be decomposed into smaller parking functions. Building on recent work by Armon *et al.* (2024), we extend the notion of primeness to a generalisation known as *graphical* parking functions, or G-parking functions. Using the classical duality betwen G-parking functions and the Abelian sandpile model (ASM), we exhibit a bijection between prime G-parking functions and what we call *strongly recurrent* configurations of the ASM. We apply this to obtain various enumerative results for prime G-parking functions on graph families.

1 Introduction

1.1 Parking functions and their variations

This extended abstract investigates the duality between parking functions and the Abelian sandpile model, from the perspective of *primeness*. More detailed examples and proofs can be found in [32]. We denote by $\mathbb{N}$ the set of strictly positive integers, and let $\mathbb{Z}_+ := \mathbb{N} \cup \{0\}$. Throughout this paper, n denotes a positive integer, and we let $[n] := \{1, \ldots, n\}$. A *parking preference* is a vector $p = (p_1, \ldots, p_n) \in [n]^n$, and we write $\mathrm{inc}\,(p) = \big(p_{(1)}, \ldots, p_{(n)}\big)$ for its non-decreasing re-arrangement.

Definition 1.1. We say that a parking preference $p \in [n]^n$ is a *parking function* if $p_{(i)} \leq i$ for all $i \in [n]$. We denote by PF_n the set of parking functions of size n.

*Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Livepool University, Suzhou, China. E-mail: Thomas.Selig@xjtlu.edu.cn. Supported by RDF-22-01-089. Corresponding author.

†Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Livepool University, Suzhou, China. E-mail: Haoyue.Zhu18@student.xjtlu.edu.cn. Supported by PGRS2012026. Corresponding author.

Remark 1.2. The name *parking function* comes from the following observation. Consider a one-way street with n parking spots and n cars, both labelled 1 to n. In this context, each car i has a preferred parking spot p_i in the street. Cars enter in order $1, \ldots, n$, and for each $i \in [n]$, car i parks in the first available spot $k \geq p_i$ (if such a spot exists). Then p is a parking function if and only if all cars are able to park through this process. Figure 1 shows an example of this process for $p = (3, 1, 3, 1)$. Here, all cars are able to park, so $p \in \mathrm{PF}_4$. However, if the last car's preferred spot were 3 (or 4), it could not find a parking place, so $(3, 1, 3, 3) \notin \mathrm{PF}_4$.

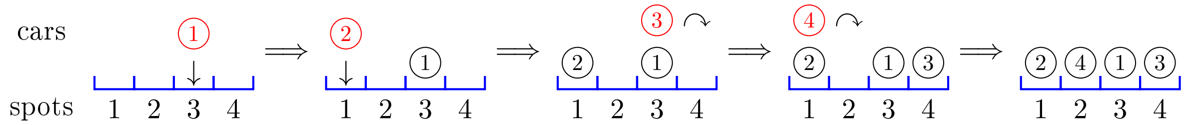


Figure 1: Illustrating the parking process for the parking function $p = (3, 1, 3, 1)$.

Parking functions were originally introduced by Konheim and Weiss [23] to study hashing functions. Since their introduction, they have been a rich research topic in Mathematics and Computer Science, with connections to various fields such as graph theory, hyperplane arrangements, discrete geometry, and the Abelian sandpile model [7, 10, 11, 33, 35]. We refer interested readers to the excellent survey by Yan [36]. In recent years, a fruitful research direction has focused on variations and extensions of the classical parking functions introduced above. Some variations modify the parking process from Remark 1.2, such as **Naples parking functions** [8, 9], **MVP parking functions** [21, 30], or **friendship parking functions** [22]. Others, such as **vector parking functions** [5, 37] or **rational parking functions** [2, 20, 24], extend the inequality condition in Definition 1.1. The paper [6] contains a fun survey of some of these variants.

1.2 Graphical parking functions

Graphical parking functions, or G-parking functions, are another extension of parking functions, introduced by Postnikov and Shapiro [25]. Here, we consider graphs $G = (V, E)$ which are undirected, finite, connected, with possible multiple edges (but no loops), and rooted at a special vertex $s \in V$, called the *sink*. We sometimes write $G = (\Gamma, s)$ for clarity (Γ is the un-rooted graph), and will simply call G a *graph*. The set of non-sink vertices will be denoted $\tilde{V} := V \setminus \{s\}$. For two vertices $v, w \in V$, we denote $\mathrm{mult}\,(vw)$ the number of edges between v and w (this can be 0). For $v \in V$ and $A \subseteq V$, we denote $N^A(v)$ the multi-set of neighbours of v in A, and $\deg^A(v) := |N^A(v)| = \sum\limits_{w \in A} \mathrm{mult}\,(vw)$. For simplicity, we usually just write $\deg(v) = \deg^V(v)$ for the degree of the vertex v. Finally, we denote $\mathbf{deg} : V \to \mathbb{N}, v \mapsto \deg(v)$ the degree function of G.

Definition 1.3. Let $G = (\Gamma, s)$ be a graph, with vertex set (resp. non-sink vertex set) V (resp. $\tilde{V}$). For $S \subseteq V$ (here, $V = \tilde{V} \cup \{s\}$ includes the sink), we denote $S^c := V \setminus S$ its complement in V. We say that a funciton $p : \tilde{V} \to \mathbb{N}$ is a *G-parking function* if, for every non-empty subset $S \subseteq \tilde{V}$, there exists $v \in S$ such that $p(v) \leq \deg^{S^c}(v)$. We denote by $\mathrm{PF}\,(G)$ the set of G-parking functions.

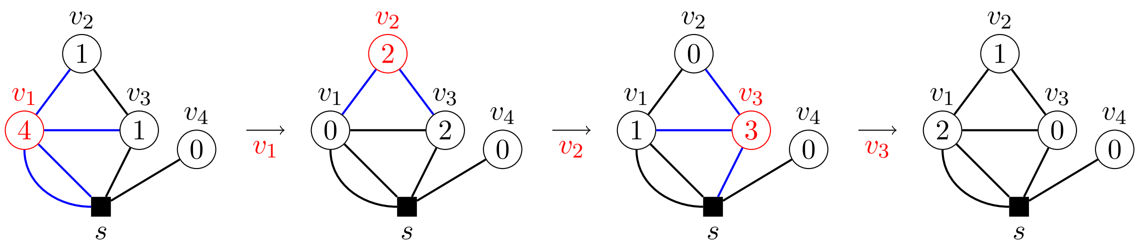


Figure 2: Illustrating the stabilisation for $c = (c(v_1), c(v_2), c(v_3), c(v_4)) = (4, 1, 1, 0) \in \mathrm{Config}\,(G)$. Vertices being toppled are represented with arrows in each phase.

Proposition 1.4. *1. We have* $\mathbf{1} \in \mathrm{PF}\,(G)$*, where* $\mathbf{1} : \tilde{V} \to \mathbb{N}, v \mapsto 1$ *is the constant function.*

2. If $p \in \mathrm{PF}\,(G)$ *is a G-parking function, then for all* $v \in \tilde{V}$*, we have* $p(v) \le \deg(v)$*.*

3. If K_n^0 *is the complete graph with vertex set* $[n] \cup \{0\}$*, rooted at* 0*, then we have* $\mathrm{PF}_n = \mathrm{PF}\left(K_n^0\right)$*.*

1.3 The Abelian sandpile model

In this section, we formally define the Abelian sandpile model, which is the second process we are interested in in this abstract. The sandpile model was initially introduced by Bak, Tang and Wiesenfeld [3, 4] as a model to study *self-organised criticality*. Shortly after, Dhar [13] formalised and generalised it, naming it the *Abelian sandpile model* (ASM).

Let $G = (\Gamma, s)$ be a graph with vertex set (resp. non-sink vertex set) V (resp. $\tilde{V}$). A (sandpile) *configuration* is a function $c : \tilde{V} \to \mathbb{Z}_+$, where we think of $c(v)$ as representing the number of "grains of sand" at the vertex v. We define $\mathrm{Config}\,(G)$ the set of all sandpile configurations on G. For $v \in \tilde{V}$ and $c \in \mathrm{Config}\,(G)$, a vertex v is said to be *stable* in the configuration c if $c(v) < \deg(v)$. The configuration c is *stable* if all its vertices are stable, and we denote $\mathrm{Stable}\,(G)$ the set of stable configurations on G. The maximal stable configuration is $\mathbf{max} := \mathbf{deg} - \mathbf{1}$.

If a vertex v is unstable, it *topples*, sending one grain along each of its incident edges. Each neighbour w of v therefore receives $\mathrm{mult}\,(wv)$ grains, while v loses $\deg(v)$ grains. Grains that are sent to the sink exit the system. Toppling v may cause some of its neighbours to become unstable, and we topple these in turn. The sink ensures that, starting from an unstable configuration c and toppling unstable vertices successively, we eventually reach a stable configuration c'. Dhar [13] showed that c' does not depend on the order in which unstable vertices topple, and we write $c' = \mathrm{Stab}(c)$. Figure 2 shows the stabilisation process for the configuration $c = (c(v_1), c(v_2), c(v_3), c(v_4)) = (4, 1, 1, 0)$.

A configuration c is called *recurrent* if there exists a configuration $d \in \mathrm{Config}\,(G)$ such that $c = \mathrm{Stab}(\mathbf{max} + d)$. We denote by $\mathrm{Rec}\,(G)$ the set of recurrent configurations on G. Recent combinatorial ASM research has seen significant progress in calculating $\mathrm{Rec}\,(G)$ on specific highly-symmetrical graph families, such as complete graphs [11] (see also [28]), complete multipartite graphs with a dominating sink [10], complete bipartite graphs where the sink is in one of the two components [17, 31], wheel and fan graphs [27], complete split graphs [12, 16], Ferrers graphs [19, 29], permutation graphs [18], and so on. We will use some of these studies in Section 3 in this paper to characterise prime parking functions for such families. We end this section by recalling the following result, which establishes the key duality between G-parking functions and recurrent configurations for the ASM on G.

Theorem 1.5 ([25, Corollary 13.7]). *For any graph G, the map $c \mapsto \mathbf{deg} - c$ is a bijection from* $\mathrm{Rec}\,(G)$ *to* $\mathrm{PF}\,(G)$.

2 Prime G-parking functions and strong recurrence

2.1 Prime (G-)parking functions

We say that a parking function $p = (p_1, \ldots, p_n) \in \mathrm{PF}_n$ has a *breakpoint* at an index $j \in [n]$ if we have $|\{i \in [n];\, p_i \leq j\}| = j$. For example, the parking function $p = (3, 1, 3, 1)$ from Figure 1 has a breakpoint at 2. If $j < n$, we can think of p as *decomposable*. Write $s_1 < \cdots < s_j$ for elements of the set $\{i \in [n];\, p_i \leq j\}$, and $s_{j+1} < \cdots < s_n$ for elements of its complement. Then by construction we have $p^{\leq j} := \left(p_{s_1}, \ldots, p_{s_j}\right) \in \mathrm{PF}_j$ and $p^{>j} := \left(p_{s_{j+1}} - j, \ldots, p_{s_n} - j\right) \in \mathrm{PF}_{n-j}$, thus decomposing p into two smaller parking functions. A *prime* parking function is a parking function whose only breakpoint is at index n, and we denote by PPF_n the set of prime parking functions of size n. Prime parking functions were introduced by Gessel (see e.g. [34, Exercise 5.49]), who showed that $|\mathrm{PPF}_n| = (n-1)^{n-1}$ (see also [15] for a bijective proof).

The concept of primeness for parking functions was recently extended in [1], with definitions for vector parking functions, (p, q)-parking functions, and two-dimensional vector parking functions. We propose here to extend the concept to G-parking functions. If $G = (\Gamma, s)$ is a graph with vertex set $V = \tilde{V} \cup \{s\}$, and $A \subsetneq \tilde{V}$, we define $G^A := G[A \cup \{s\}]$ to be the *induced* subgraph on $A \cup \{s\}$.

Definition 2.1. Let (A, B) be an ordered set partition of $\tilde{V}$, with $A, B \neq \emptyset$. For $p \in \mathrm{PF}\,(G)$, we define $p^A : A \to \mathbb{N}$ by $p^A(v) := p(v)$ for all $v \in A$, and $p^B : B \to \mathbb{Z}$ by $p^B(v) := p(v) - \deg^A(v)$ for all $v \in B$. We say that p is *decomposable* with respect to the partition (A, B) if we have $p^A \in \mathrm{PF}\left(G^A\right)$ and $p^B \in \mathrm{PF}\left(G^B\right)$. We say that p is *prime* if there exists no ordered set partition (A, B) with respect to which p is decomposable, and denote by $\mathrm{PPF}\,(G)$ the set of all prime G-parking functions.

Proposition 2.2. *If s is* not *a cut vertex of G, then $\mathbf{1} \in \mathrm{PPF}\,(G)$. Otherwise,* $\mathrm{PPF}\,(G) = \emptyset$.

Example 2.3. The G-parking function p on Figure 3 is decomposable with respect to the partitions $(\{v_1\}, \{v_2, v_3, v_4\})$ and $(\{v_3, v_4\}, \{v_1, v_2\})$ (in both cases p^A and p^B are the all-1 function $\mathbf{1}$).

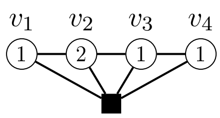


Figure 3: A graphical parking function which admits two different prime decompositions.

For $p \in \mathrm{PF}\,(G)$, a *prime decomposition* of p is an ordered set partition $(A_1, \ldots, A_k)$ of $\tilde{V}$ such that for all $i \in [k]$, we have $p^{A_i} \in \mathrm{PPF}\left(G^{A_i}\right)$, where p^{A_i} is defined by $p^{A_i}(v) := p(v) - \deg^{\bigcup_{j=1}^{i-1} A_j}(v)$ for all $v \in A_i$. It is immediate to see that every G-parking function admits a prime decomposition. However, the example from Figure 3 shows that prime decompositions are not always unique.

2.2 Strong recurrence for the ASM

Recurrent configurations of the ASM are closed under grain addition. That is, if $c \in \mathrm{Rec}(G)$, then for any $v \in \tilde{V}(G)$ such that $c(v) < \deg(v) - 1$, we have $c + \mathbb{1}_v \in \mathrm{Rec}(G)$, where $\mathbb{1}_v$ is the indicator function at v (i.e., $\mathbb{1}_v(w) = 1$ if $w = v$ and 0 otherwise). In this section, we are interested in the reverse property: which recurrent configurations are closed under grain removal?

Our motivation stems from so-called *minimal recurrent* configurations. A configuration $c \in \mathrm{Config}(G)$ is minimal recurrent if it is recurrent, and for all $v \in \tilde{V}(G)$, the configuration $c - \mathbb{1}_v$ (obtained from c by removing one grain at v) is **not** recurrent. Minimal recurrent configurations are in bijection with certain acyclic orientations of G (see e.g. [26]), and have provided a key tool in various combinatorial studies of the recurrent configurations on graph families, see for example [19, 29, 27]. Our notion of *strong recurrence* is in some sense orthogonal to that of minimal recurrence: strongly recurrent configurations remain recurrent after we remove many grains from them.

Definition 2.4. Let $G = (\Gamma, s)$ be a graph with vertex set $V = \tilde{V} \cup \{s\}$. For a recurrent configuration $c \in \mathrm{Rec}(G)$, we set $V_M(c) := \{v \in N^G(s);\, c(v) \geq \deg(v) - \mathrm{mult}(vs)\}$. For $v \in V_M(c)$, we define the configuration $c^{v-} := c - \sum\limits_{w \in \tilde{V} \setminus \{v\}} \mathrm{mult}(ws) \cdot \mathbb{1}_w$. We say that c is *strongly recurrent* (SR) if for all $v \in V_M(c)$, we have $c^{v-} \in \mathrm{Rec}(G)$. We denote by $\mathrm{SR}(G)$ the set of SR configurations on G.

In words, the transformation $c \rightsquigarrow c^{v-}$ removes grains from all neighbours of the sink except v according to their multiplicity. For many graph families, such as complete graphs, wheel graphs, fan graphs, complete multi-partite graphs with a dominating sink, and so on, this just means removing exactly one grain from all vertices except v. Leaving c unchanged at v is dictated by Dhar's burning algorithm (see [14, Section 6.2]): $V_M(c)$ is simply the set of vertices which can be burnt in the algorithm's first stage, and if we were to simultaneously remove grains from all vertices in $V_M(c)$ the configuration could no longer be recurrent. We now state the main result of this paper.

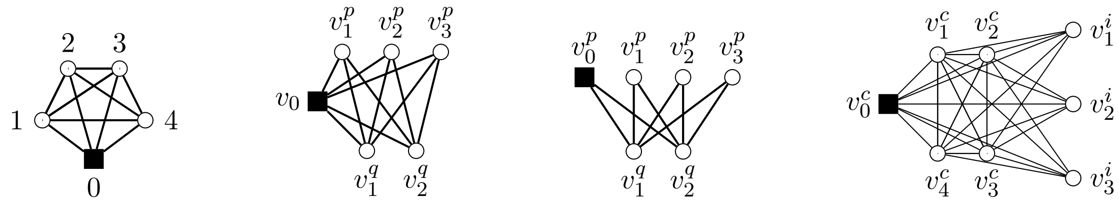
Theorem 2.5. *For any graph G, the map $c \mapsto \mathbf{deg} - c$ is a bijection from $\mathrm{SR}(G)$ to $\mathrm{PPF}(G)$.*

Corollary 2.6. *Let $p \in \mathrm{PF}(G)$, and define $V_M(p) := \{v \in \tilde{V};\, p(v) \leq \mathrm{mult}(vs)\}$. Then p is prime if and only if, for all $v \in V_M(p)$, we have $p^{v+} := p + \sum\limits_{w \in \tilde{V} \setminus \{v\}} \mathrm{mult}(ws) \cdot \mathbb{1}_w \in \mathrm{PF}(G)$.*

3 Application to graph families

In this section, we consider graph families with high degrees of symmetry (see Figure 4). We will first see that our Definition 2.1 extends some existing definitions of prime parking functions.

Proposition 3.1.
1. *If K_n^0 is the complete graph with vertex set $[n] \cup \{0\}$, rooted at 0, then we have $\mathrm{PPF}_n = \mathrm{PPF}\left(K_n^0\right)$, where PPF_n is the set of prime classical parking functions.*
2. *If $K_{p,q}^0$ is the complete tripartite graph with p vertices in one component, q vertices in the other, and a dominating sink (see Figure 4b), then $\mathrm{PPF}\left(K_{p,q}^0\right)$ is the set of prime (p, q)-parking functions, as defined in [1, Definition 4.1].*



(a) The complete graph K_4^0.

(b) The complete tripartite graph $K_{3,2}^0$.

(c) The complete bipartite graph $K_{3*,2}$.

(d) The complete split graph $S_{4*,3}$.

Figure 4: Graph families with high degrees of symmetry.

3.1 The complete bipartite case

In this section, we consider the complete bipartite graph $K_{p*,q}$. This graph consists of two independent sets $P = \{v_0^p, v_1^p, \ldots, v_p^p\}$ (including the sink v_0^p) and $Q = \{v_1^q, \ldots, v_q^q\}$, and edges between each pair $(v_a^p, v_b^q) \in P \times Q$ (see Figure 4c). A parking function $p \in \mathrm{PF}\,(K_{p*,q})$ is called *increasing* if $p\,(v_1^p) \leq \cdots \leq p\,(v_p^p)$ and $p\,(v_1^q) \leq \cdots \leq p\,(v_q^q)$. We denote by $\mathrm{PF}^{\mathrm{inc}}\,(K_{p*,q})$, resp. $\mathrm{PPF}^{\mathrm{inc}}\,(K_{p*,q})$ the set of increasing, resp. increasing prime, $K_{p*,q}$-parking functions.

Theorem 3.2. *There is a bijection from* $\mathrm{PPF}^{\mathrm{inc}}\,(K_{p*,q})$ *to* $\mathrm{PF}^{\mathrm{inc}}\left(K_{(p-1)*,q}\right)$. *In particular, we have* $\left|\mathrm{PPF}^{\mathrm{inc}}\,(K_{p*,q})\right| = \left|\mathrm{PF}^{\mathrm{inc}}\left(K_{(p-1)*,q}\right)\right| = \frac{1}{p+q-1}\binom{p+q-1}{p}\binom{p+q-1}{p-1}$.

3.2 The complete split case

The complete split graph $S_{m*,n}$ consists of a clique with vertices $C = \{v_0^c, \ldots, v_m^c\}$ (including the sink v_0^c), an independent set $I = \left\{v_1^i, \ldots, v_n^i\right\}$, and edges between each pair $(v_a^c, c_b^i) \in C \times I$ (see Figure 4d). As in the complete bipartite case, we are interested in *increasing* (prime) $S_{m*,n}$-parking functions, with analogous definitions and notation.

Theorem 3.3. *There is a bijection from* $\mathrm{PPF}^{\mathrm{inc}}\,(S_{m*,n})$ *to* $\mathrm{PF}^{\mathrm{inc}}\left(S_{(m-1)*,n}\right)$. *In particular, we have* $\left|\mathrm{PPF}^{\mathrm{inc}}\,(S_{m*,n})\right| = \left|\mathrm{PF}^{\mathrm{inc}}\left(S_{(m-1)*,n}\right)\right| = \frac{1}{m}\binom{2m-2}{m-1}\binom{2m+n-2}{n}$.

The proofs of Theorems 3.2 and 3.3 are similar. We show that if p is a prime parking function on a complete bipartite or split graph, then there is a vertex v in the sink component with $p(v) = 1$. The bijections then simply remove that vertex. This reflects the cases of classical and (p, q)-parking functions (see [34, Exercise 5.49] for the classical case, and [1, Proposition 4.2] for the (p, q) case). The enumerative formulae exploit the connection to the ASM (Theorem 1.5) and existing combinatorial characterisations on complete bipartite [17] and complete split [16] graphs.

References

[1] Sam Armon, Joanne Beckford, Dillon Hanson, Naomi Krawzik, Olya Mandelshtam, Lucy Martinez, and Catherine Yan. Primeness of generalized parking functions. *ArXiv preprint*, 2024.

[2] Jean-Christophe Aval and François Bergeron. Interlaced rectangular parking functions. *Sémin. Lothar. Comb.*, 81:article b81h, 15, 2020.

[3] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the $1/f$ noise. *Phys. Rev. Lett.*, 59:381–384, 1987.

[4] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Phys. Rev. A (3)*, 38(1):364–374, 1988.

[5] Yue Cai and Catherine H. Yan. Vector parking functions with periodic boundaries and rational parking functions. *J. Comb. Theory, Ser. A*, 173:30, 2020. Id/No 105239.

[6] Joshua Carlson, Alex Christensen, Pamela E. Harris, Zakiya Jones, and Andrés Ramos Rodríguez. Parking functions: choose your own adventure. *Coll. Math. J.*, 52(4):254–264, 2021.

[7] Denis Chebikin and Pavlo Pylyavskyy. A family of bijections between G-parking functions and spanning trees. *J. Comb. Theory, Ser. A*, 110(1):31–41, 2005.

[8] Alex Christensen, Pamela E. Harris, Zakiya Jones, Marissa Loving, Andrés Ramos Rodríguez, Joseph Rennie, and Gordon Rojas Kirby. A generalization of parking functions allowing backward movement. *Electron. J. Comb.*, 27(1):research paper p1.33, 18, 2020.

[9] Laura Colmenarejo, Pamela E. Harris, Zakiya Jones, and Christo Keller. Counting k-Naples parking functions through permutations and the k-Naples area statistic. *Enumer. Comb. Appl.*, 1(2):Article #S2R11, 2021.

[10] Robert Cori and Dominique Poulalhon. Enumeration of (p, q)-parking functions. *Discrete Math.*, 256(3):609–623, 2002.

[11] Robert Cori and Dominique Rossin. On the sandpile group of dual graphs. *Eur. J. Comb.*, 21(4):447–459, 2000.

[12] Henri Derycke, Mark Dukes, and Yvan Le Borgne. The sandpile model on the complete split graph: q,t-Schröder polynomials, sawtooth polyominoes, and a cyclic lemma. *ArXiv preprint*, 2024.

[13] Deepak Dhar. Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64(14):1613–1616, 1990.

[14] Deepak Dhar. Theoretical studies of self-organized criticality. *Phys. A Stat. Mech. Appl.*, 369(1):29–70, 2006. Fundamental Problems in Statistical Physics.

[15] Rui Duarte and António Guedes de Oliveira. The number of prime parking functions. *Math. Intell.*, 46(3):222–224, 2024.

[16] Mark Dukes. The sandpile model on the complete split graph, Motzkin words, and tiered parking functions. *J. Comb. Theory, Ser. A*, 180:15, 2021. Id/No 105418.

[17] Mark Dukes and Yvan Le Borgne. Parallelogram polyominoes, the sandpile model on a complete bipartite graph, and a q,t-Narayana polynomial. *J. Comb. Theory, Ser. A*, 120(4):816–842, 2013.

[18] Mark Dukes, Thomas Selig, Jason P. Smith, and Einar Steingrímsson. Permutation graphs and the abelian sandpile model, tiered trees and non-ambiguous binary trees. *Electron. J. Comb.*, 26(3):research paper P3.29, 25pp., 2019.

[19] Mark Dukes, Thomas Selig, Jason P. Smith, and Einar Steingrímsson. The abelian sandpile model on Ferrers graphs – a classification of recurrent configurations. *Eur. J. Comb.*, 81:221–241, 2019.

[20] Eugene Gorsky, Mikhail Mazin, and Monica Vazirani. Affine permutations and rational slope parking functions. *T. Am. Math. Soc.*, 368(12):8403–8445, 2016.

[21] Pamela E. Harris, Brian M. Kamau, J. Carlos Martínez Mori, and Roger Tian. On the outcome map of MVP parking functions: Permutations avoiding 321 and 3412, and Motzkin paths. *Enumer. Comb. Appl.*, 3(2):Article #S2R11, 2023.

[22] Yujia Kang, Thomas Selig, Guanyi Yang, Yanting Zhang, and Haoyue Zhu. On friendship and cyclic parking functions. *Enumer. Comb. Appl.*, 4(3):Article #S2R21, 2024.

[23] Alan G. Konheim and Benjamin Weiss. An occupancy discipline and applications. *SIAM J. Appl. Math.*, 14(6):1266–1274, 1966.

[24] Garrett Nelson. Rational parking functions and (m, n)-invariant sets. *ArXiv preprint*, 2025.

[25] Alexander Postnikov and Boris Shapiro. Trees, parking functions, syzygies, and deformations of monomial ideals. *T. Am. Math. Soc.*, 356:3109–3142, 2004.

[26] Matthias Schulz. Minimal recurrent configurations of chip firing games and directed acyclic graphs. In *Automata 2010. Selected papers based on the presentations at the 16th international workshop on cellular automata (CA) and discrete complex systems (DCS), Nancy, France, June 14–16, 2010.*, pages 111–124. Nancy: The Association. Discrete Mathematics & Theoretical Computer Science (DMTCS), 2010.

[27] Thomas Selig. Combinatorial aspects of sandpile models on wheel and fan graphs. *Eur. J. Comb.*, 110:23, 2023. Id/No 103663.

[28] Thomas Selig. The stochastic sandpile model on complete graphs. *Electron. J. Comb.*, 31(3):research paper P3.26, 29pp., 2024.

[29] Thomas Selig, Jason P. Smith, and Einar Steingrímsson. EW-tableaux, Le-tableaux, tree-like tableaux and the abelian sandpile model. *Electron. J. Comb.*, 25(3):research paper P3.14, 32pp., 2018.

[30] Thomas Selig and Haoyue Zhu. New combinatorial perspectives on MVP parking functions and their outcome map. *ArXiv preprint*, 2023.

[31] Thomas Selig and Haoyue Zhu. Abelian and stochastic sandpile models on complete bipartite graphs. In Shin-ichi Nakano and Mingyu Xiao, editors, *WALCOM: Algorithms and Computation*, pages 326–345, Singapore, 2025. Springer Nature Singapore.

[32] Thomas Selig and Haoyue Zhu. Prime graphical parking functions and strongly recurrent configurations of the Abelian sandpile model, 2025. In preparation, preprint to appear on *ArXiv* shortly.

[33] Richard P. Stanley. Hyperplane arrangements, interval orders, and trees. *Proc. Natl. Acad. Sci. USA*, 93(6):2620–2625, 1996.

[34] Richard P. Stanley. *Enumerative combinatorics. Volume 2*, volume 62 of *Camb. Stud. Adv. Math.* Cambridge: Cambridge University Press, 1999.

[35] Richard P. Stanley and Jim Pitman. A polytope related to empirical distributions, plane trees, parking functions, and the associahedron. *Discrete Comput. Geom.*, 27(4):603–634, 2002.

[36] Catherine H. Yan. *Parking Functions*, chapter 13. CRC Press, 2015.

[37] Mei Yin. Parking functions, multi-shuffle, and asymptotic phenomena. *La Matematica*, pages 1–25, 2023.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Bound on shortest cycle covers of graphs

(Extended abstract)

Deping Song* Xuding Zhu†

Abstract

Assume G is a bridgeless graph. A cycle cover of G is a collection of cycles of G such that each edge of G is contained in at least one of the cycles. The length of a cycle cover of G is the sum of the lengths of the cycles in the cover. The minimum length of a cycle cover of G is denoted by $scc(G)$. The minimum length of a cycle cover of G consisting of k cycles is denoted by $scc_k(G)$. It was proved independently by Alon and Tarsi and by Bermond, Jackson, and Jaeger that $scc(G) \leq \frac{5}{3}m$ for every bridgeless graph G with m edges. This remained the best-known upper bound for $scc(G)$ for 40 years. In this paper, we prove that if G is a bridgeless graph with m edges and n_2 vertices of degree 2, then $scc(G) \leq scc_3(G) < \frac{29}{18}m + \frac{1}{18}n_2$. As a consequence, we show that $scc(G) \leq scc_3(G) \leq \frac{5}{3}m - \frac{1}{42}\log m$. The upper bound $scc(G) \leq scc_3(G) < \frac{29}{18}m \approx 1.6111m$ for bridgeless graphs G of minimum degree at least 3 improves the previous known upper bound $scc(G) \leq 1.6258m$. A key lemma used in the proof confirms Fan's conjecture that if C is a circuit of G and G/C admits a nowhere zero 4-flow, then G admits a 4-flow f such that $E(G) - E(C) \subseteq \mathrm{supp}(f)$ and $|\mathrm{supp}(f) \cap E(C)| > \frac{3}{4}|E(C)|$.

1 Introduction

Graphs considered in this paper may have loops and multiple edges. For terminology and notations not defined here, we follow [3, 19]. *Contracting an edge* means deleting the edge and then identifying its ends. For a subgraph H of a graph G, let G/H be the graph obtained from G by contracting all edges of H.

*School of Mathematical Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China. E-mail: songdeping@zjnu.edu.cn

†School of Mathematical Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China. E-mail: xdzhu@zjnu.edu.cn. Supported by NSFC.

A *circuit* is a 2-regular connected graph, and a *cycle* is a graph such that the degree of each vertex is even. The *length* of a cycle or a circuit is the number of its edges. A collection $\mathcal{C}$ of cycles of a graph G *covers* G if each edge of G is contained in at least one cycle in $\mathcal{C}$. The *length* of a cycle cover $\mathcal{C}$, denoted by $\ell(\mathcal{C})$, is the sum of the lengths of its cycles. The length of the shortest cycle cover of G is denoted by $scc(G)$. The minimum length of a cycle cover of G consisting of k cycles is denoted by $scc_k(G)$.

Alon and Tarsi [1] conjectured that $scc(G) \leq \frac{7}{5}m$ for every bridgeless graph G with m edges. This conjecture is sharp, as the Petersen graph G has $scc(G) = \frac{7}{5}m = 21$. It was proved by Jamshy and Tarsi [10] that this conjecture implies the well-known cycle double cover conjecture of Seymour [14] and Szekeres [16], that every bridgeless graph has a collection of cycles that covers each edge exactly twice.

For 3 cycle covers, it was conjectured by Tarsi (page 171, [20]) that $scc_3(G) \leq \frac{22}{15}m$ for any bridgeless graph G with m edges. This conjecture is also sharp for the Petersen graph, and infinitely many graphs G with $scc_3(G) = \frac{22}{15}m$ were given by Fan and Raspaud in [6]. On the other hand, it was proved in [6] that this conjecture is implied by Fulkerson's conjecture, which says that every bridgeless graph has six cycles that cover each edge exactly 4 times.

Alon and Tarsi [1] and Bermond, Jackson, and Jaeger [2] independently proved that $scc(G) \leq \frac{5}{3}m$ for every bridgeless graph G with m edges. This upper bound on $scc(G)$ remained the best known upper bound for 40 years. For special families of graphs, there are some better upper bounds proved in the literature. Jamshy, Raspaud, and Tarsi [9] proved that if G admits a nowhere-zero 5-flow, then $scc(G) \leq \frac{8}{5}m$. Fan and Raspaud [6] proved that if the Fulkerson Conjecture is true, then $scc(G) \leq \frac{22}{15}m$. Kaiser, Král, Lidický, Nejedlý and Šámal [11] proved that if G has minimum degree at least 3 and is loopless, then $scc(G) \leq \frac{44}{27}m$. Fan [5] proved that if G has minimum degree at least 3, then $scc(G) < \frac{218}{135}m$ in the loopless case and $scc(G) < \frac{278}{171}m$ if loops are allowed. Kompišová and Lukot'ka [13] proved that $scc(G) < \frac{278}{171}m + \frac{2}{27}n_2$ for an arbitrary bridgeless graph G, where n_2 is the number of degree 2 vertices in G.

This paper proves that $scc_3(G) < \frac{29}{18}m + \frac{1}{18}n_2$. As a consequence, we show that $scc_3(G) \leq \frac{5}{3}m - \frac{1}{42}\log m$, which improves the general upper bound for $scc(G)$ by a fraction of $\log m$.

Let G be a graph and Γ be an (additive) abelian group. For an orientation D of $E(G)$, denote $E^+(v)$ (resp., $E^-(v)$) by the set of all oriented edges of $D(G)$ with their tails (resp., heads) at vertex $v \in V(G)$. For a function $f : E(G) \to \Gamma$ and $v \in V(G)$, let $f^+(v) = \sum_{e \in E^+(v)} f(e)$ and $f^-(v) = \sum_{e \in E^-(v)} f(e)$. A Γ-*flow* in G is an ordered pair (D, f) such that $f^+(v) - f^-(v) = 0$ for every vertex $v \in V(G)$. For a subgraph H of G and $a \in \Gamma$, define $E_{f=a}(H) = \{e \in E(H) : f(e) = a\}$. The *support* of f is the set of all edges of G with $f(e) \neq 0$ and is denoted by $supp(f)$. A flow (D, f) is called *nowhere-zero* if $supp(f) = E(G)$.

We shall consider $\mathbb{Z}_2 \times \mathbb{Z}_2$-flows in this paper. As $-\alpha = \alpha$ for all $\alpha \in \mathbb{Z}_2 \times \mathbb{Z}_2$, the orientation of G is irrelevant and can be omitted. To be precise, a $Z_2 \times Z_2$-flow on G is a function $f : E(G) \to Z_2 \times Z_2$ such that for each vertex v, $\sum_{e \in E(v)} f(e) = 0$, where $E(v)$ is the set of edges of G incident to v.

2 A proof sketch of the main result

A maximal path P in a graph H with all interior vertices of degree 2 (in H) is called a *thread* in H. First we prove the following lemma, which confirms a conjecture of Fan in [5].

Lemma 2.1. *Assume G is a graph and C is a circuit in G. Assume each edge in C is contained in a thread of G of length at least q. If f is a $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow on G, then G admits a $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow g such that* $\mathrm{supp}(f) - C = \mathrm{supp}(g) - C$ *and* $|E_{g=(0,0)}(C)| \leq \frac{1}{4}(|E(C)| - q)$.

It was proved by Edmonds [4] that for each bridgless cubic graph G, there is a positive integer p such that G has $3p$ perfect matchings and each edge is contained in exactly p of these matchings. The following lemma follows easily from this result.

Lemma 2.2. *Assume G is a bridgeless graph and not a circuit. Then there is a positive integer p such that G has a family $\mathcal{F}$ of $3p$ cycles that covers each edge exactly $2p$ times.*

Then we use these two lemmas to prove the following main result.

Theorem 2.3. *Let G be a bridgeless graph with m edges and n_2 vertices of degree 2, then $scc_3(G) < \frac{29}{18}m + \frac{1}{18}n_2$ and consequently $scc_3(G) \leq \frac{5}{3}m - \frac{1}{42}\log m$, where the base for* log *is* 2.

Proof Sketch. Let $\mathcal{F}$ be a family $3p$ cycles F such that each edge e of G is contained in $2p$ cycles F in $\mathcal{F}$. The average size of F is $2m/3$.

It was shown in [19] that G/F has odd-edge-connectivity at least 5. Hence G/F have a nowhere-zero $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow [8], which extends to a $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow f of G with $E(G) - F \subseteq \mathrm{supp}(f)$.

By Lemma 2.1, we may assume that for each component B of F, $|E_{f=(0,0)}, B)| < \frac{1}{4}|B|$. For $i \geq 2$, let d_i be the number of components in F with i edges. We choose the $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow f so that $|E_{f=(0,1)}(G-F)|$ is minimum, and subject to this, $|E_{f=(1,0)}(G-F)|$ is minimum. This implies that $E_{f=(0,1)}(G-F)$ is acyclic, because if C is a cycle in $E_{f=(0,1)}(G-F)$, then $f + f_{C,(1,0)}$ is a $\mathbb{Z}_2 \times \mathbb{Z}_2$-flow g with $\mathrm{supp}(g) = \mathrm{supp}(f)$ and $|E_{g=(0,1)}(G)| < |E_{f=(0,1)}(G)|$. Similarly, $E_{f=(1,0)}(G-F)$ is acyclic.

Let $C_1 = E_{f=(0,1)}(G) \cup E_{f=(1,0)}(G), C_2 = E_{f=(0,1)}(G) \cup E_{f=(1,1)}(G), C_3 = E_{f=(1,1)}(G) \cup E_{f=(1,0)}(G)$. Then C_1, C_2, C_3 is a double cover of $supp(f)$. Therefore

$$\mathcal{C}_1 = \{C_1 \Delta F, C_2 \Delta F, C_3 \Delta F\}$$

is a cycle cover of G, where each edge not in F is covered twice, and each edge $e \in F \cap \mathrm{supp}(f)$ is covered once and each edge in $E_{f=0}(F)$ is covered three times. Thus $\mathcal{C}_1$ is a cycle cover of G of length $2m - |F| + 2|E_{f=0}(G)|$. This implies that

$$\ell(\mathcal{C}_1) \leq 2m - |F| + 2|E_{f=0}(G)|.$$

We may assume that for each component B of F, $|C_2 \cap B| \leq \frac{1}{2}|B|$, for otherwise we replace C_2 with the symmetric difference $C_2 \Delta B$. Thus $|C_2 \cap F| \leq \frac{1}{2}|F| - \frac{1}{2}\sum_{i=1}^{\infty} d_{2i+1}$. Let d_1 be the number of degree 2 vertices of G contained in G/F and $\mathcal{C}_2 = \{C_1, C_2, C_1 \Delta F\}$. Then $\mathcal{C}_2$ is also a cycle cover of G. Calculation shows that the length of this cycle cover is

$$\ell(\mathcal{C}_2) \leq m + \frac{1}{2}|F| + \frac{5}{2}\sum_{i=1}^{\infty} d_{4i+1} + 3\sum_{i=0}^{\infty} d_{4i+2} + \frac{5}{2}\sum_{i=0}^{\infty} d_{4i+3} + 3\sum_{i=1}^{\infty} d_{4i} + d_1 - 3.$$

Hence

$$scc_3(G) \leq \frac{5}{6}\ell(\mathcal{C}_1) + \frac{1}{6}\ell(\mathcal{C}_2) \leq \frac{11}{6}m - \frac{1}{3}|F| + \frac{1}{6}d_1 - \frac{1}{2}.$$

By randomly choosing F from $\mathcal{F}$, we have $E[|F|] = \frac{2}{3}m$ and $E[d_1] = \frac{1}{3}n_2$. So there is a choice of F to ensure that

$$scc_3(G) \leq \frac{29}{18}m + \frac{1}{18}n_2 - \frac{1}{2}.$$

It remains to show that $scc_3(G) \leq \frac{5}{3}m - \frac{1}{42}\log m$.

As $scc_3(G) \leq \frac{5}{3}m - \frac{1}{18}(m - n_2)$, we are done if $m - n_2 > \frac{3}{7}\log m$.

Assume $m - n_2 \leq \frac{3}{7}\log m$.

Let $\mathcal{C}_1$ be the cycle cover of G constructed above. We have shown that

$$scc_3(G) \leq \ell(\mathcal{C}_1) \leq 2m - |F| + 2|E_{f=(0,0)}(G)|,$$

where $|F| \geq \frac{2}{3}m$.

Let $\epsilon = \frac{\log m}{56m}$. If $|E_{f=(0,0)}(G)| \leq (\frac{1}{4} - \epsilon)|F|$, then

$$scc_3(G) \leq 2m - (\frac{1}{2} + 2\epsilon)|F| \leq 2m - (\frac{1}{2} + 2\epsilon)\frac{2}{3}m = \frac{5}{3}m - \frac{1}{42}\log m.$$

Thus it suffices to show that for any connected component C of F, $|E_{f=(0,0)}(C)| \leq (\frac{1}{4} - \epsilon)|C|$.

If $|C| < \frac{14m}{\log m}$, then it follows from Lemma 2.1 that $|E_{f=(0,0)}(C)| \leq \frac{1}{4}|C| - \frac{1}{4} < (\frac{1}{4} - \epsilon)|C|$

Assume $|C| \geq \frac{14m}{\log m}$. Let $e_1, e_2, \ldots, e_k$ be the threads in C, with $\ell(e_i) \leq \ell(e_{i+1})$ for all i. Thus C has k vertices of degree at least 3. Hence $k \leq \frac{2}{3}(m - n_2) \leq \frac{2}{7}\log m$.

If $\ell(e_i) \leq 5^i \log m$ for all i, then $|C| \leq \sum_{i=1}^{k} 5^i \log m \leq \frac{5^{k+1}}{4}\log m$. As $k \leq \frac{2}{7}\log m$, $|C| \leq \frac{5}{4}m^{\frac{2log5}{7}}\log m$. This implies that $\frac{14m}{\log m} \leq \frac{5}{4}m^{\frac{2\log 5}{7}}\log m$, hence $56m^{1-\frac{2\log 5}{7}} \leq 5(\log m)^2$, which fails for any positive integer m.

Thus there exists i such that $\ell(e_i) \geq 5^i \log m$.

Let i_0 be the minimum index such that $\ell(e_{i_0}) \geq 5^{i_0}\log m$.

Let S be the union of all the threads $e_1, e_2, \ldots, e_{i_0-1}$. Let $G' = G/S$ and $C' = C/S$. Let $q = 5^{i_0}\log m$. Then each thread of C' has length at least q. By Lemma 2.1, we may assume that $|E_{f=(0,0)}(C')| \leq \frac{1}{4}(|E(C')| - q)$.

Then $|E_{f=(0,0)}(C)| \leq \frac{1}{4}(|E(C')| - q) + \sum_{i=1}^{i_0-1} 5^i \log m \leq \frac{1}{4}|E(C)| - \frac{5}{4}\log m < (\frac{1}{4} - \epsilon)|C|$. This completes the proof of Theorem 2.3.

References

[1] N. Alon, M. Tarsi, Covering multigraphs by simple circuits. *SIAM J. Algebraic Discrete Methods.* 6, (1985), 345–350.

[2] J.C. Bermond , B. Jackson, F. Jaeger, Shortest coverings of graphs with cycles.*J. Combin. Theory Ser. B* 35, (1983), 297–308.

[3] J. A. Bondy, U. S. R. Murty, Graph theory. Springer, 2008.

[4] J. Edmonds, Maximum matching and a polyhedron with (0,1)-vertices *J. Res. Nat. Bur. Standards Sect. B* 69, (1983), 125–130.

[5] G. Fan, Integer 4-flows and cycle covers. *Combinatorica* 37(**6**), (2017), 1097–1112.

[6] G. Fan and A. Raspaud. Fulkerson's conjecture and circuit covers. *J. Combin. Theory Ser. B* 61 (1994), no. 1, 133¨C-138.

[7] H. Fleischner, Eine gemeinsame Basis f¡§?r die Theorie der Eulerschen Graphen und den Satz von Petersen. *Monatsh. Math.* 81(**4**), (1976), 267–278.

[8] F. Jaeger, Flows and generalized coloring theorems in graphs. *J. Combin. Theory Ser. B,* 26, (1979), 205–216.

[9] U. Jamshy, A. Raspaud and M. Tarsi, Short circuit covers for regular matroids with a nowhere zero 5 -flow. *J. Combin. Theory Ser. B* 43 (1987), no. 3, 354–357.

[10] U. Jamshy, M. Tarsi, Short cycle covers and the cycle double cover conjecture. *J. Combin. Theory Ser. B,* 56, (1992), 197–204.

[11] T. Kaiser, D. Král, B. Lidický, P. Nejedlý, R. Šámal, Short cycle covers of graphs with minimum degree three. *SIAM J. Discrete Math.* 24(**1**), (2010), 330–355.

[12] M. Kochol, Smallest counterexample to the 5-flow conjecture has girth at least eleven. *J. Combin. Theory Ser. B.* 100 (**4**), (2004), 381–389.

[13] A. Kompišová, R. Lukot'ka, Short cycle covers of graphs with at most 77% vertices of degree two. *Electron. J. Combin.* 27, (2020), No. 4.31, 7 pp.

[14] P. D. Seymour, Sums of circuits, in: *Graph Theory and Related Topics (J. A. Bondy and U. S. R. Murty, Eds.),* 341–355, *Academic Press, New York,* 1979.

[15] D. Song, S. Li, X. Wang, On Fan's conjecture about 4-flow. *Graphs Combin.* 1 (2025), Paper No. 7.

[16] G. Szekeres, Polyhedral decompositions of cubic graphs, *Bull. Austral. Math. Soc.* 8, (1973), 367–387.

[17] W. T. Tutte, A contribution on the theory of chromatic polynomial, *Canad. J. Math.,* 6, (1954), 80–90.

[18] X. Wang, Y. Lu, S.-G. Zhang, Note on integer 4-flows in graphs. *Acta Math. Sin. (Engl. Ser.)* 38(**9**) (2022), 1653–1664.

[19] C.-Q. Zhang, Integer flows and cycle covers of graphs. *Marcel Dekker, Inc.,* New York, 1997.

[20] C.-Q Zhang, Circuit double cover of graphs. *London Mathematical Society Lecture Note Series,* 399. *Cambridge University Press,* Cambridge, 2012.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Edge Isoperimetry of Lattices

(Extended abstract)

Cameron Strachan* Konrad Swanepoel†

Abstract

We present two results related to an edge-isoperimetric question for Cayley graphs on the integer lattice asked by Ben Barber and Joshua Erde. For any (undirected) graph G, the edge boundary of a subset of vertices S is the number of edges between S and its complement in G. Barber and Erde asked whether for any Cayley graph on $\mathbb{Z}^d$, there is always an ordering of $\mathbb{Z}^d$ such that for each n, the first n terms minimize the edge boundary among all subsets of size n. Our first result answers this question in the negative by presenting an example of a Cayley graph on $\mathbb{Z}^d$ (for all $d \geq 2$) for which there is no such ordering. Our second result is a positive example of a Cayley graph on $\mathbb{Z}^2$ that has such an ordering. This is the most complicated example known to us of a two-dimensional Cayley graph for which such an ordering exists.

1 Introduction

Definition 1. Given a graph G, the *edge boundary* of $S \subseteq V(G)$ is

$$\partial(S) := |\{uv \in E(G) : u \in S, v \notin S\}|.$$

The *edge isoperimetric problem* (EIP) of a graph G is, for a given n, to minimize $\partial(S)$ over all $S \subseteq V(G)$ where $|S| = n$. We call such minimizing sets *solutions to the EIP of G*. This classical problem has been extensively studied since the 1960s (see [12]). Although it is NP-hard in general, some special cases are known. One aspect that has received particular attention is whether nested solutions exist. A *nested solution* for the EIP of G is an ordering $v_1, v_2, \dots$ of the vertex set $V(G)$ such that for each n, the set $\{v_1, v_2, \dots, v_n\}$ is a solution to the EIP of G.

One of the first cases of the EIP that has been solved is the d-dimensional cube graph, which has nested solutions, and where the optimal shapes include subcubes [3, 10, 13, 14].

*Department of Mathematics, London School of Economics and Political Science, WC2A 2AE London, UK. E-mail: c.r.strachan@lse.ac.uk.
†Department of Mathematics, London School of Economics and Political Science, WC2A 2AE London, UK. E-mail: k.swanepoel@lse.ac.uk.

For $p = 1, \infty$, denote by $(\mathbb{Z}^d, \ell_p)$ the graph with vertex set $\mathbb{Z}^d$ where pairs of vertices have an edge if their ℓ_p distance is 1. Bollobás and Leader [4] solved the EIP for $(\mathbb{Z}^d, \ell_1)$, and proved that the solutions include cubes. Moreover, they showed that this graph has nested solutions.

Bollobás and Leader [4] also considered the EIP on finite grids $\{1, 2, \dots, k\}^d$, considered as induced subgraphs of $(\mathbb{Z}^d, \ell_1)$. It turned out that there are two types of solutions: cubes if n is small relative to the size of the grid and half-grids for large n. Furthermore, they showed the transition between these two types of solutions is not smooth, giving the first example of a graph without nested solutions.

If G is an undirected k-regular graph, for any $S \subseteq V(G)$ we have $|E(G[S])| = \frac{k|S| - \partial(S)}{2}$. If G is a directed k-regular graph, then for any $S \subseteq V(G)$ we have $|E(G[S])| = k|S| - \partial(S)$. Thus, the problem of minimizing $\partial(S)$ over all subsets with size n is the same as maximizing $|E(G[S])|$ over all subsets of size n.

In terms of this formulation, Brass [5] solved the EIP of $(\mathbb{Z}^2, \ell_\infty)$, where the optimal shapes include certain octagons. Additionally, he showed $(\mathbb{Z}^2, \ell_\infty)$ has nested solutions. For $d \geq 3$, the EIP of $(\mathbb{Z}^d, \ell_\infty)$ remains open.

Let $g_1 = (1, 0)$ and $g_2 = (1/2, \sqrt{3}/2)$. The *triangular lattice* is the set

$$\Lambda := \{mg_1 + ng_2 : m, n \in \mathbb{Z}\}.$$

We can turn Λ into a graph by joining a pair of vertices if their Euclidean distance is 1. For this graph, the EIP is solved [9] (see also [8, 11]) with solutions that include regular hexagons. Again, the graph has nested solutions. This graph is isomorphic to $\mathbb{Z}^2$, where two vertices are joined if their difference is in $\{(\pm 1, 0), (0, \pm 1), \pm(1, 1)\}$, hence it can be thought of as a graph between $(\mathbb{Z}^2, \ell_1)$ and $(\mathbb{Z}^2, \ell_\infty)$.

The above examples are all special cases of Cayley graphs on the group $\mathbb{Z}^d$ [1].

Definition 2. Let U be a finite set that generates $\mathbb{Z}^d$ as a group and does not contain the identity. The (directed) Cayley graph $\mathbb{Z}^d_U$ is the graph on the vertex set $\mathbb{Z}^d$ where (u, v) is an edge whenever $v - u \in U$. When U is symmetric (that is, $-u \in U$ for all $u \in U$), we consider $\mathbb{Z}^d_U$ to be undirected.

Given a generating set U of $\mathbb{Z}^d$, let $Z \subseteq \mathbb{R}^d$ be the zonotope $\sum_{u \in U}[0, u]$ generated by the line segments $[0, u]$, $u \in U$. Barber and Erde [1] showed that the edge boundary of $tZ \cap \mathbb{Z}^d$ for large t, asymptotically approximates the edge boundary of solutions to the EIP of $\mathbb{Z}^d_U$. Barber, Erde, Keevash and Roberts [2] showed that additionally, $tZ \cap \mathbb{Z}^d$ asymptotically approximates the shape of solutions to the EIP of $\mathbb{Z}^d_U$.

Barber and Erde [1] asked if every Cayley graph of $\mathbb{Z}^d$ has nested solutions. Despite the positive examples already given, Briggs and Wells [6] gave counterexamples for the case $d = 1$. On the other hand, they also gave a partial positive answer: for any Cayley graph of $\mathbb{Z}$, there exists an $m \in \mathbb{N}$ and an ordering $v_1, v_2 \dots$ of $\mathbb{Z}$ such that for any $n \geq m$, the set $\{v_1, v_2, \dots, v_n\}$ is a solution to the EIP. In other words, they showed that any Cayley graph on $\mathbb{Z}$ has nested solutions starting at a sufficiently large size.

We give a negative answer to the question of Barber and Erde for all $d \geq 2$ by giving an explicit example of a Cayley graph without nested solutions. Furthermore, we show this example does not have nested solutions regardless of any starting point. Thus, in dimensions 2 and higher there are stronger counterexamples than in $\mathbb{Z}^1$.

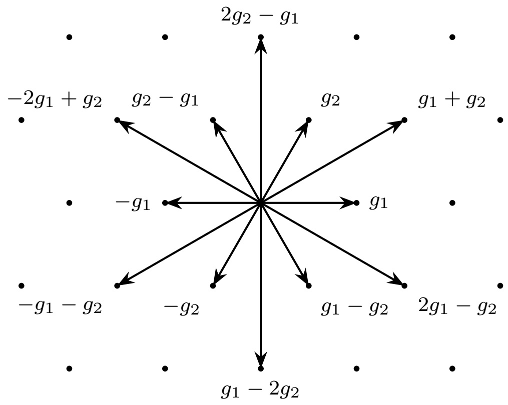


Figure 1: Generating set of Theorem 2

Theorem 1. *The EIP for $\mathbb{Z}^d_U$, where U is the generating set $\{\pm e_i : i = 1, \ldots, d\} \cup \{\pm 2e_1\}$ of $\mathbb{Z}^d$, does not have nested solutions starting at any size. In other words, for all n and each n-element subset S_n of $\mathbb{Z}^d$ for which $\partial(S_n)$ is the minimum among all n-element subsets of $\mathbb{Z}^d_U$, there does not exist a sequence $S_n \subset S_{n+1} \subset S_{n+2} \subset \cdots$ of i-element subsets S_i of $\mathbb{Z}^d$, such that for each $i \geq n$, $\partial(S_i)$ is the minimum among all i-element subsets of $\mathbb{Z}^d_U$.*

Our second result is a solution of the EIP for another Cayley graph on $\mathbb{Z}^2$ with nested solutions. The generating set for this graph is $U = \{(\pm 1, 0), (0, \pm 1),$ $\pm(1,1), \pm(1,-1), \pm(-1,2), \pm(-2,1)\}$. Thus, it contains $(\mathbb{Z}^2, \ell_\infty)$ as a subgraph. In fact, it is more suitable to consider this to be the graph on the triangular lattice with edges for all pairs at distance 1 or $\sqrt{3}$. As a Cayley graph on Λ, the generating set is depicted in Figure 1. We denote it by Λ_U.

Theorem 2. *Let Λ_U be the undirected Cayley graph with vertex set Λ and symmetric generating set $U = \{\pm g_1, \pm(g_1 + g_2), \pm g_2, \pm(2g_2 - g_1), \pm(g_2 - g_1), \pm(g_2 - 2g_1)\}$, where $g_1 = (1, 0)$ and $g_2 = (1/2, \sqrt{3}/2)$. The maximum number of edges of a subgraph of Λ_U with $n \geq 3$ vertices is*

$$e(n) := \begin{cases} 6n - 4\sqrt{6n-6} & \text{if } n = 24k^2 - 24k + 7 \text{ for some } k \in \mathbb{N} \\ \lfloor 6n - \sqrt{96n - 63} \rfloor & \text{otherwise.} \end{cases}$$

Additionally, Λ_U has nested solutions.

The first few values of n where $e(n) = 6n - 4\sqrt{6n-6}$ are $n = 7, 55, 151, 295,$ 487 and 727. In Figure 2 we depict the unique (up to translation) extremal subgraph of Λ_U with 55 vertices.

The subgraphs of Λ_U with n vertices and $e(n)$ edges are candidate extremal graphs for a problem of Erdős and Vesztergombi [7] on the maximum number of occurrences of the smallest and second smallest distances in a set of n points in the plane. Let S be a set of n points in the plane, and denote by $m_1(S)$ and $m_2(S)$ the number of occurrences of the smallest and second smallest distance in S. Let $f(n)$ be the maximum value of $m_1(S) + m_2(S)$, where the maximum is taken over all sets S of n points. It is known that $f(n) \leq 6n$ [16]. (See also [7] for further results.) Theorem 2 implies that $f(n) \geq e(n)$, with the lower bound being given by subsets of the triangular lattice, with smallest distance 1 and second smallest distance $\sqrt{3}$.

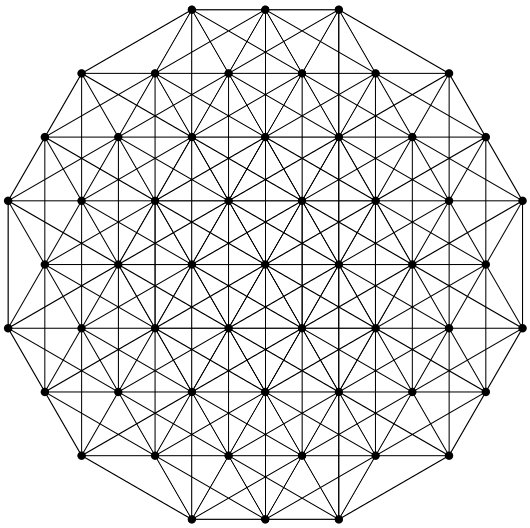


Figure 2: The extremal subgraph of Λ_U with $24k^2 - 24k + 7$ vertices ($k = 2$)

Conjecture 3. *For any sufficiently large n, $f(n) = e(n)$, with the only sets S of n points attaining $f(n) = m_1(S) + m_2(S)$ being similar to the extremal sets on the triangular lattice.*

2 Proof outline of Theorem 1

Let $\mathbb{Z}_U^d$ be the Cayley graph stated in Theorem 1 with generating set $U = \{\pm e_i : i = 1, \ldots, d\} \cup \{\pm 2e_1\}$ where $e_1, \ldots, e_d$ denote the standard unit basis of $\mathbb{Z}^d$. Let $n \in \mathbb{N}$ be arbitrary, and let S be an n-element subset of $\mathbb{Z}^d$ for which the number of edges of the subgraph $\mathbb{Z}_U^d[S]$ induced by S is the maximum among all n-element subsets of $\mathbb{Z}^d$. To prove Theorem 1 we show there is no sequence S_i, $i \geq n$, of subsets of $\mathbb{Z}^d$ such that $S = S_n$, $|S_i| = i$ for all $i \geq n$, and each S_i maximizes the number of edges in the subgraph $\mathbb{Z}_U^d[S_i]$.

We prove this by contradiction and suppose there is such a sequence. Let $m_j = \min\{x_j : (x_1, x_2, \ldots, x_d) \in S\}$ and $M_j = \max\{x_j : (x_1, x_2, \ldots, x_d) \in S\}$ for each $j = 2, \ldots, d$. Let $C = \mathbb{Z} \times \prod_{j=2}^{d} \{m_j, \ldots, M_j\}$. We derive a contradiction by first proving by induction that $S_i \subseteq C$ for all $i \geq n$. We then use the Loomis-Whitney inequality to show for a large enough $m \in \mathbb{N}$, $S_m \subseteq C$ implies S_m does not maximize the number of edges over all m-element induced subgraphs of $\mathbb{Z}_U^d$.

3 Proof outline of Theorem 2

Let Λ_U be the Cayley graph of Λ with generating set U stated in Theorem 2. Since Λ_U is a 12-regular graph, we have for any $S \subseteq \Lambda$ with n vertices

$$|E(\Lambda_U[S])| = 6n - \frac{\partial(S)}{2}. \tag{1}$$

To prove Theorem 2 we first show that any subgraph of Λ_U with $n \geq 3$ vertices has at most $e(n)$ edges. We then give an ordering $v_1, v_2, \ldots$ of Λ such that for each n, the graph

$\Lambda_U[\{v_1, v_2, \ldots, v_n\}]$ has $e(n)$ many edges. This ordering, together with the upper bound, proves that Λ_U has nested solutions.

To show that $e(n)$ is an upper bound for the number of edges of an n-vertex subgraph of Λ_U, we first show this upper bound for a particular class of subgraphs of Λ_U that can be thought of as (possibly degenerate) completely filled-up lattice 12-gons. We define these as follows.

Definition 3. For any finite $S \subseteq \Lambda$, the *hull* of S, denoted as $\mathrm{hull}(S) \subseteq \Lambda$, is the intersection of the 12 supporting half-planes of S parallel to an element of U. Denote $\mathscr{P} = \{\Lambda_U[\mathrm{hull}(S)] : S \subseteq \Lambda, |S| < \infty\}$.

We use induction on n to show that any $P \in \mathscr{P}$ with n vertices has at most $e(n)$ edges. The base case of this induction goes up to $n = 30$ and involves a Python computation. In the inductive step, we rearrange the points of P to make P "rounder". This potentially leads to constructing a new $P^* \in \mathscr{P}$ with more vertices, where we can either apply the inductive hypothesis after removing its boundary, or for a certain range of polygons very close to optimal, we need to do some exact symbolic computations using sympy.

Once we have shown the upper bound for all sets in $\mathscr{P}$, the general case is straightforward. We again use induction on n to show that any subgraph of Λ_U with n vertices has at most $e(n)$ edges. During this process we will prove additionally that when $n = 24k^2 - 24k + 7$ for some $k \in \mathbb{N}$, the n-vertex subgraph of Λ_U with $e(n)$ edges is unique up to a translation. This enables us to define an ordering of Λ by interpolating between the unique extremal subgraph P_k when $n = 24k^2 - 24k + 7$ for some $k \in \mathbb{N}$, and the unique extremal subgraph P_{k+1} when $n = 24(k+1)^2 - 24(k+1) + 7$. This interpolation entails finding a specific sequence of length 48 of the 12 orientations of sides that have to be added to P_k to get to P_{k+1}. This sequence is found by using Breadth-First Search in an auxiliary directed graph that represents all ways of adding sides.

References

[1] Ben Barber and Joshua Erde, *Isoperimetry in integer lattices*, Discrete Anal. (2018), Paper No. 7, 16. MR3819052

[2] Ben Barber, Joshua Erde, Peter Keevash, and Alexander Roberts, *Isoperimetric stability in lattices*, Proc. Amer. Math. Soc. **151** (2023), no. 12, 5021–5029. MR4648905

[3] A. J. Bernstein, *Maximally connected arrays on the n-cube*, SIAM J. Appl. Math. **15** (1967), 1485–1489. MR223260

[4] Béla Bollobás and Imre Leader, *Edge-isoperimetric inequalities in the grid*, Combinatorica **11** (1991), no. 4, 299–314. MR1137765

[5] Peter Brass, *Erdős distance problems in normed spaces*, Comput. Geom. **6** (1996), no. 4, 195–214. MR1392310

[6] Joseph Briggs and Chris Wells, *Phase transitions in isoperimetric problems on the integers*, 2024.

[7] György Csizmadia, *The multiplicity of the two smallest distances among points*, Discrete Math. **194** (1999), no. 1-3, 67–86. MR1654972

[8] Frank Harary and Heiko Harborth, *Extremal animals*, J. Combin. Inform. System Sci. **1** (1976), no. 1, 1–8. MR457263

[9] Heiko Harborth, *Solution to problem 664 A*, Elem. Math **29** (1974), 14–15.

[10] L. H. Harper, *Optimal assignments of numbers to vertices*, J. Soc. Indust. Appl. Math. **12** (1964), 131–135. MR162737

[11] ———, *The edge-isoperimetric problem for regular planar tesselations*, Ars Combin. **61** (2001), 47–63. MR1863367

[12] ———, *Global methods for combinatorial isoperimetric problems*, Cambridge Studies in Advanced Mathematics, vol. 90, Cambridge University Press, Cambridge, 2004. MR2035509

[13] Sergiu Hart, *A note on the edges of the n-cube*, Discrete Math. **14** (1976), no. 2, 157–163. MR396293

[14] John H. Lindsey II, *Assignment of numbers to vertices*, Amer. Math. Monthly **71** (1964), 508–516. MR168489

[15] L. H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. **55** (1949), 961–962. MR31538

[16] K. Vesztergombi, *Bounds on the number of small distances in a finite planar set*, Studia Sci. Math. Hungar. **22** (1987), no. 1-4, 95–101. MR913897

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Three hardness results for graph similarity problems

(Extended abstract)

He Sun* Danny Vagnozzi†

Abstract

Notions of graph similarity provide alternative perspective on the graph isomorphism problem and vice-versa. In this paper, we consider measures of similarity arising from mismatch norms as studied in Gervens and Grohe: the edit distance $\delta_{\mathcal{E}}$, and the metrics arising from ℓ_p-operator norms, which we denote by δ_p and $\delta_{|p|}$. We address the following question: can these measures of similarity be used to design polynomial-time approximation algorithms for graph isomorphism? We show that computing an optimal value of $\delta_{\mathcal{E}}$ is NP-hard on pairs of graphs with the same number of edges. In addition, we show that computing optimal values of δ_p and $\delta_{|p|}$ is NP-hard even on pairs of 1-planar graphs with the same degree sequence and bounded degree. We also study similarity problems on strongly regular graphs and prove some near optimal inequalities with interesting consequences on the computational complexity of graph and group isomorphism. The results presented first appeared in [15], where the reader may find the full proofs.

Main results

The *graph isomorphism problem* is a classical problem in theoretical computer science. It consists in determining whether there is an edge preserving bijection between the vertices of two input graphs. While there are several applications of graph isomorphism algorithms in real-world settings such as pattern recognition and computer vision, most studies on the graph isomorphism problem have been motivated by purely theoretical intricacies. On the one hand, while we do not know whether the problem is in P, it is believed not to be

*University of Edinburgh
†Durham University

NP-complete. On the other hand, following a breakthrough by Babai [2], we know that graph isomorphism is decidable in quasipolynomial time. Adding this to the fact that many practical instances are solvable in polynomial time, one can safely say that on average the graph isomorphism problem can be solved efficiently.

In comparison to graph isomorphism, *graph similarity* is a more general concept and has received plentiful attention in practical applications [6, 10, 11]; however its studies from a theoretical point of view have remained limited. Informally speaking, graph similarity is a family of decision problems, and every such problem is associated with some measure of similarity, for example a *graph metric* δ mapping pairs of graphs into non-negative real numbers. For a fixed metric δ, the input of the problem consists in a pair of graphs G, H and a real constant c, and it is asked to decide whether $\delta(G, H) \leq c$ holds. As an example, consider a metric ι assigning 0 to pairs of isomorphic graphs and 1 otherwise. Evaluating ι is equivalent to deciding graph isomorphism, which is therefore in itself a similarity problem. Another interesting metric that has appeared in the literature under different guises is the *edit distance* [13], which we denote by $\delta_{\mathcal{E}}$. Given two graphs G and H with the same number of vertices, $\delta_{\mathcal{E}}(G, H)$ is defined to be the minimum number of edges to be deleted from and added to H in order to obtain a graph that is isomorphic to G. We call $\mathtt{DIST}_{\mathcal{E}}$ the similarity problem associated to the metric $\delta_{\mathcal{E}}$. This problem is known to be NP-hard even on some very restricted classes of graphs, and is closely related to several computationally hard problems such as the maximum common edge subgraph problem or the quadratic assignment problem, both of which are well-known to be notoriously hard optimisation problems.

Very recent work by Gervens et al. [12] studies the computational complexity of graph similarity problems, and builds a theoretical framework for studying such problems. This includes some precise definition of graph metrics as well as analysing the metrics from matrix norms that are commonly used in graph theory. These metrics are all based on minimising some measure μ of 'mismatch' between two graphs, which can be seen as the discrepancy between the graphs for a given alignment. The edit distance can be classed as such a metric. Indeed, for a bijection $\pi : V(G) \to V(H)$, let G^π be the image of G under π, and $\mu_{\mathcal{E}}(G^\pi, H)$ the number of edges (u, v) of G for which π is not a local isomorphism; that is, $(\pi(u), \pi(v))$ is an edge if, and only if, (u, v) is not an edge. Then the edit distance $\delta_{\mathcal{E}}(G, H)$ is given by the minimum value of $\mu_{\mathcal{E}}(G^\pi, H)$ over all bijections π. In a more abstract sense, this can be seen as minimising the number of edges in the *mismatch graph* $G^\pi - H$, a signed graph whose adjacency matrix is given by the difference between the adjacency matrix of G^π and that of H. The other graph metrics analysed in [12] also consist in minimising some property of a mismatch graph. In particular, they consider the ℓ_p-*operator norm* of the signed and unsigned versions of the mismatch graph as a measure for mismatch. The main difference between these metrics and the edit distance is that the value of the former can be estimated by the maximum degree of the mismatch graph, hence a *local* property, whereas the latter seems to rely mostly on *global* properties of the mismatch graph. Nonetheless, these metrics all give rise to NP-hard similarity problems, and the strategies to prove the hardness result are strikingly similar; that is, by reducing the NP-hard Hamiltonian cycle problem on 3-regular graphs to the similarity problem between a 3-regular graph and a cycle of the same order. A natural question arising from these proofs is whether NP-hardness still holds when considering the similarity problem on pairs of graphs harder to distinguish isomorphism-wise than, say, a 3-regular graph and a cycle. In fact, the main question motivating this paper is the following:

> *Are there classes of graphs where some notion of similarity can be used as an efficient approximation of isomorphism?*

Put otherwise, we ask whether there are classes of graphs whose isomorphism is hard to decide but one can decide in polynomial time whether $\delta(G,H) \leq t$ for some metric δ and threshold t (which could be constant or, for example, a function of the number of vertices), so we can deduce that G and H are not isomorphic if $\delta(G,H) > t$, and deem the test inconclusive otherwise. Since 3-regular graphs are distinguished from cycles using simple properties such as the number of edges or their degree sequences, the known hardness proofs for the metrics studied in [12] do not answer the above question.

We address this issue in light of the graph metrics studied in [12]. Let $\delta_{\mathcal{E}}, \delta_p$ and $\delta_{|p|}$ be the graph metrics arising from the edit distance and the ℓ_p-operator norms of the signed and unsigned mismatch graph respectively. Denote the similarity problem associated with each of these metrics by $\texttt{DIST}_{\mathcal{E}}, \texttt{DIST}_p$ and $\texttt{DIST}_{|p|}$. The following are our main results.

Theorem 0.1. *For any constant $c > 0$ and $\epsilon \in (0, 1/2)$, the following holds: for any pair of graphs G and H with n vertices and the same number of edges, it is NP-hard to decide whether $\delta_{\mathcal{E}}(G,H) \leq cn^{\epsilon}$. Consequently, $\texttt{DIST}_{\mathcal{E}}$ is NP-hard, even when the two input graphs have the same number of edges.*

Theorem 0.2. *For any pair of 1-planar graphs G and H with n vertices, the same degree sequence and maximum degree at most 15, it is NP-hard to decide whether $\delta_p(G,H) \leq 2$ and $\delta_{|p|}(G,H) \leq 2$ for any $p \geq 1$. Consequently, $\texttt{DIST}_p$ and $\texttt{DIST}_{|p|}$ are NP-hard, even when the two input graphs have bounded degree and the same degree sequences.*

The construction employed in the proof of Theorem 0.2 is interesting in a twofold manner. Firstly, we remark that a priori, it is not clear if these graphs can be used to show the hardness of $\texttt{DIST}_{\mathcal{E}}$ for pairs of graphs with same degree sequence. This is in contrast to the more general problem, where considering as input a 3-regular graph and a cycle graph is sufficient to show the hardness of $\texttt{DIST}_{\mathcal{E}}$, $\texttt{DIST}_p$ and $\texttt{DIST}_{|p|}$. Secondly, the proof for $p = 2$ differs significantly from the proof for values of $p \neq 2$, and requires a deeper analysis of the combinatorial properties of signed graphs. These anomalies seem to point at an interesting layer of discrepancy between these similarity problems.

While the statement of Theorem 0.1 seems to diminish the hopes of a threshold function t asymptotically smaller than $\sqrt{n}$ such that $\delta_{\mathcal{E}}(G,H) \leq t(n)$ is an efficiently decidable approximation of isomorphism, neither Theorem 0.1 nor Theorem 0.2 provide a satisfactory answer to our motivating question. In fact, the pairs of graphs used in their proofs are still distinguishable by simple properties such as their degree sequences (Theorem 0.1) or spectra (Theorem 0.2). As an attempt to study the complexity of $\texttt{DIST}_p$ and $\texttt{DIST}_{|p|}$ over pairs of cospectral graphs, we prove another result with a similar effect but in a different style.

Theorem 0.3. *Let G and H be cospectral graphs on n vertices each with maximum degree $d_{\max}$, and consider some function $t(n) = o\left(n^{1/2}\right)$. The following hold:*

1. *If there is a polynomial-time algorithm deciding whether $\delta_{\mathcal{E}}(G, H) \leq t(n)$, then there is a polynomial-time algorithm deciding the isomorphism of groups as Cayley tables.*

2. *If there is a polynomial-time algorithm deciding whether $\delta_1(G, H) \leq d_{\max}/3 - 4$, then there is a polynomial-time algorithm deciding the isomorphism of groups as Cayley tables.*

This follows from a simple combinatorial analysis for bounds on alignments between *strongly regular graphs*, which may be of independent theoretical interest. A graph is said to be *strongly regular* with parameters (n, d, λ, ν) if it has n vertices, it is d-regular and every pair of adjacent (rsp. non-adjacent) vertices has λ (rsp. μ) common neighbours.

Proposition 0.4. *Let G and H be strongly regular graphs with parameters (n, d, λ, ν), where $\lambda \geq \nu$. Then, G and H are not isomorphic if and only if it $\delta_1(G, H) \geq \lambda - \nu + 1$.*

Whilst the analysis for proving the above is rather naïve at a first glance, we show that our obtained bound is almost tight up to some constant additive term.

Proposition 0.5. *There exist strongly regular graphs G and H with same parameters (n, d, λ, ν) such that $\delta_1(G, H) \geq \sqrt{n}$ and $\lambda - \nu + 1 = \sqrt{n} - 5$.*

We remark that the graphs used to prove the above are Latin square graphs arising from the Cayley tables of Abelian groups, whose isomorphism problem is well known to be decidable in polynomial time.

Finally, we comment on Theorem 0.3 in light of current knowledge on group isomorphism. The fastest known algorithm for group isomorphism runs in quasipolynomial time. However, it is believed that an efficient algorithm for group isomorphism could help overcome some of the current bottlenecks in Babai's quasipolynomial time algorithm for graph isomorphism [2]. As such, Theorem 0.3 suggests that the existence of a small threshold function t so that $\delta_{\mathcal{E}}(G, H) \leq t(n)$, or a constant c such that $\delta_1(G, H) \leq c$ provide efficiently decidable approximations of isomorphism is unlikely.

Related work

Our results have close ties to those in [1, 12, 13]. Arvind et al. [1] study the $\texttt{DIST}_{\mathcal{E}}$ problem in the context of approximating graph isomorphism. The authors provide both hardness results and efficient algorithms for different variants of the problem of approximating some optimisation version of the $\texttt{DIST}_{\mathcal{E}}$ problem. In [12, 13], the authors show the hardness of the problems $\texttt{DIST}_{\mathcal{E}}$, $\texttt{DIST}_p$ and $\texttt{DIST}_{|p|}$ for all natural numbers p, even when restricted to forests of bounded degree. However, the classes of graphs resulting from these reductions include pairs of graphs with differing number of edges. Thus, whilst the proofs of Theorems 0.1 and 0.2 are inspired from the results in [12, 13], they provide a different perspective on the nature of graph similarity problems.

Linear algebraic approaches to similarity prior to the ℓ_p-operator norms in [12] have appeared in [14, 16]. For instance, Kolla et al. [14] consider the *spectrally robust graph isomorphism problem* - a similarity problem based on the eigenvalues of Laplacian matrices.

Related to similarity problems on Latin square graphs are the works on the Hamming distance between multiplication tables of finite groups [7, 8, 9]. More recent work by Buchheim et al. [4] provides hardness results for variants of the *subgroup distance problem*.

Open problems

Our work motivates the following open problems.

> *Are* $\mathtt{DIST}_{\mathcal{E}}$*,* $\mathtt{DIST}_p$ *and* $\mathtt{DIST}_{|p|}$ *NP-hard on pairs of cospectral graphs?*

This question arises naturally from the proof of Theorem 0.2, for the graphs used in the hardness reduction are distinguishable via their spectra.

A natural follow-up question from Theorem 0.1 is to ask for which functions $t(n)$ is $\delta_{\mathcal{E}}(G,H) \leq t(n)$ decidable in polynomial time. Our result addresses such question for polynomials $t(n)$ asymptotically smaller than $\sqrt{n}$. Since any simple graph on n vertices has at most $n(n-1)/2$ edges, it is trivial to show that $\delta_{\mathcal{E}}(G,H) \leq n(n-1)/2$ always holds.

> *For what values of* $c > 0$ *and* $\epsilon > 0$ *is* $\delta_{\mathcal{E}}(G,H) \leq cn^{\epsilon}$ *decidable in polynomial time for all pairs of* n *vertex graphs* G *and* H *with equal volumes?*

Theorem 0.1 implies that there are no such values for $\epsilon < 1/2$, unless $\mathsf{P} = \mathsf{NP}$.

Finally, our results on strongly regular graphs is centred around a rather niche area of combinatorics and graph theory. Based on the fact that the bound in Proposition 0.5 might not be tight, we ask the following question.

> *Given non-isomorphic Latin square graphs* G *and* H *on* n *vertices each, does it always hold that* $\delta_1(G,H) \geq \sqrt{n}$*?*

Given the importance of Latin square graphs in applied areas of mathematics such as experiment design [3] and error correcting codes [5], we expect that the answer to the above question could be of interest beyond the pure theory.

References

[1] Vikraman Arvind, Johannes Köbler, Sebastian Kuhnert, and Yadu Vasudev. Approximate graph isomorphism. In *Mathematical Foundations of Computer Science 2012*, pages 100–111, 2012.

[2] László Babai. Graph isomorphism in quasipolynomial time. *ArXiv*, abs/1512.03547, 2015.

[3] Rosemary A. Bailey. *Design of comparative experiments*, volume 25 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2008.

[4] Christoph Buchheim, Peter J. Cameron, and Taoyang Wu. On the subgroup distance problem. *Discrete Mathematics*, 309(4):962–968, 2009.

[5] Charles J. Colbourn, Torleiv Kløve, and Alan C. H. Ling. Permutation arrays for powerline communication and mutually orthogonal Latin squares. *IEEE Transactions on Information Theory*, 50:1289–1291, 2004.

[6] Donatello Conte, Pasquale Foggia, Carlo Sansone, and Mario Vento. Thirty years of graph matching in pattern recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 18:265–298, 2004.

[7] J Dénes. On a problem of L. Fuchs. *Acta Sci. Math.(Szeged)*, 23:237–241, 1962.

[8] Diane Donovan, Sheila Oates-Williams, and Cheryl E. Praeger. On the distance between distinct group Latin squares. *Journal of Combinatorial Designs*, 5(4):235–248, 1997.

[9] Aleš Drápal. How far apart can the group multiplication tables be? *European Journal of Combinatorics*, 13(5):335–343, 1992.

[10] Frank Emmert-Streib, Matthias Dehmer, and Yongtang Shi. Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, 346-347:180–197, 2016.

[11] Pasquale Foggia, Gennaro Percannella, and Mario Vento. Graph matching and learning in pattern recognition in the last 10 years. *International Journal of Pattern Recognition and Artificial Intelligence*, 28, 2014.

[12] Timo Gervens and Martin Grohe. Graph similarity based on matrix norms. In *MFCS47th22*, pages 52:1–52:15, 2022.

[13] Martin Grohe, Gaurav Rattan, and Gerhard J. Woeginger. Graph similarity and approximate isomorphism. In *International Symposium on Mathematical Foundations of Computer Science*, 2018.

[14] Alexandra Kolla, Ioannis Koutis, Vivek Madan, and Ali Kemal Sinop. Spectrally Robust Graph Isomorphism. In *ICALP45th18*, pages 84:1–84:13, 2018.

[15] He Sun and Danny Vagnozzi. Three hardness results for graph similarity problems. *arXiv preprint arXiv:2309.03810*, 2023.

[16] Shinji Umeyama. An eigendecomposition approach to weighted graph matching problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):695–703, 1988.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

On canonical sandpile actions of embedded graphs

(Extended abstract)

Lilla Tóthmérész*

Abstract

The sandpile group of a connected graph is a group whose cardinality is the number of spanning trees. The group is known to have a canonical action on spanning trees if the graph is embedded into the plane. Recently, Merino, Moffatt and Noble defined a sandpile group variant (called Jacobian) for embedded graphs, whose cardinality is the number of quasi-trees. Baker, Ding and Kim showed that this group acts canonically on the quasitrees. We show that for any embedded Eulerian digraph, one can define a canonical action of the sandpile group on compatible Eulerian tours. Furthermore, we show that the Jacobian of an embedded graph is the usual sandpile group of the medial digraph, and the action by Baker at al. agrees with the action of the medial digraph on Eulerian tours (which fact is made possible by the existence of a canonical bijection between Eulerian tours of the medial digraph and quasi-trees due to Bouchet).

1 Introduction

The sandpile group of a (connected) graph is an Abelian group, whose order equals to the number of spanning trees. More generally, the sandpile group of an Eulerian digraph is a group whose order equals the number of in-arborescences rooted at an arbitrary vertex. Ellenberg [7] asked if the sandpile group of a graph has a "canonical" simply transitive action on the spanning trees if the graph is embedded into an orientable surface. Holroyd et al. [8] defined the rotor routing action, which depends on an embedding and a fixed vertex (root) of the graph. Chan at al. [4] showed that it is independent of the root if and only if the embedding is into the plane. Analogous results were proved about a different sandpile action [2], the Bernardi action, and the relationship of the two group actions also has a rich literature [2, 9, 11, 6]. The existence of a canonical action in the non-plane case remained open. Recently, Merino et al. [10] defined an "embedded sandpile group" called the Jacobian of the embedded graph. This group has cardinality equal to the number of quasi-trees. Baker, Ding and Kim [1] gave

*ELTE Eötvös Loránd University, Pázmány Péter sétány 1/C, Budapest, Hungary E-mail: `lilla.tothmeresz@ttk.elte.hu` Supported by the Counting in Sparse Graphs Lendület Research Group of the Alfréd Rényi Institute of Mathematics.

an alternative definition for this group, and they proved that the Jacobian of an embedded graph has a canonical action on the quasi-trees.

We have two main results: We show that for embedded Eulerian digraphs, with any root vertex, the rotor-routing action of the sandpile group can be lifted to a canonical action on Eulerian tours compatible with the embedding. We call this the tour-rotor action.

Then, we show that the Jacobian of an embedded graph is canonically isomorphic to the sandpile group of the medial digraph. It is known [3] that the Eulerian tours of the medial digraph are canonically in bijection with the quasi-trees. Furthermore, in this case all Eulerian tours are compatible. We show that the action defined by Baker, Ding and Kim [1] agrees with the tour-rotor action of the sandpile group of the medial digraph.

2 Preliminaries

We denote undirected graphs by G, and directed graphs by D. Our graphs will always be connected and our directed graphs weakly connected. A digraph D is Eulerian if the in-degree agrees with the out-degree for each vertex. For an undirected graph, a spanning tree is a connected, cycle-free subgraph containing each vertex. In a directed graph D, a spanning in-arborescence rooted at v is a weakly connected spanning subgraph where the out-degree of each vertex is 1 except for v, where the out-degree is 0. We denote the set of in-arborescences rooted at v by $\mathrm{Arb}(D, v)$.

A *ribbon digraph* D is a digraph together with a cyclic ordering of the (in- and out-) edges incident to v for each vertex v. For each ribbon digraph D, one can construct a closed orientable surface Σ, such that D is cellularly embedded in Σ, and the positive orientation around each vertex of D induces the ribbon structure (see for example [12]). We will alternatively use the abtract or the embeding viewpoint.

An *Eulerian tour* of a digraph D is a closed walk that uses each edge of D exactly once. It is well-known that a digraph has an Eulerian tour if and only if it is Eulerian. We think of an Eulerian tour without specifying its beginning and endpoint. We call an Eulerian tour of D *compatible with the ribbon structure*, if around each vertex v, the cyclic order in which the Eulerian tour traverses the out-edges of v agrees with the cyclic order induced by the ribbon structure. Note that we do not assume anything about the order in which the tour traverses the in-edges of v. For an example, see Figure 1, that shows an Eulerian ribbon digraph embedded into the plane. The red curves on the second and third panels indicate compatible Eulerian tours.

By the BEST theorem [5], for any given vertex v, the number of compatible Eulerian tours of D agrees with the number of in-arborescences of D rooted at v.

Definition 2.1 ([5], bijection between compatible Eulerian tours and $\mathrm{Arb}(D, v)$)**.** To define the bijection, we need to fix an edge $e = \overrightarrow{vw}$.

For a compatible Eulerian tour $\mathcal{E}$, think of $\overrightarrow{vw}$ as its first edge. Then, for each vertex $u \neq v$, take the out-edge of u that is last used by the tour. By [5], this subgraph is an in-arborescence rooted at v. We denote it by $A_{\overrightarrow{vw}}(\mathcal{E})$. See Figure 1 for an example.

2.1 Sandpile group

Let $D = (V, E)$ be an Eulerian digraph. We refer to a vector $x \in \mathbb{Z}^V$ as a chip configuration. For $v \in V$, we think of $x(v)$ as the number of chips on v (which might also be negative). We denote by $\mathbb{Z}_0^V$ the subgroup of $\mathbb{Z}^V$ where the sum of the coordinates is 0.

The sandpile group is defined via the so-called chip-firing moves. The firing of a vertex v modifies chip configuration x by decreasing the number of chips on v by the outdegree of v, and increasing the number of chips on each out-neighbor w of v by the multiplicity of the edge $\overrightarrow{vw}$. We define two chip configurations x and y to be *linearly equivalent* if there is a sequence of firings that transforms x to y. We denote $x \sim_D y$. Linear equivalence is indeed an equivalence relation, and it preserves the sum of the chips.

Definition 2.2 (Sandpile group)**.** For an Eulerian digraph $D = (V, E)$, the sandpile group is defined as $S(D) = \mathbb{Z}_0^V/_{\sim_D}$.

The following statement is a version of the matrix-tree theorem.

Fact 2.3. *[8] If D is Eulerian and $v \in V(D)$, then $|S(D)| = |\mathrm{Arb}(D, v)|$.*

2.2 Rotor-routing

Rotor routing is a relaxed version of chip-firing that enables one to define a simply transitive action of $S(D)$ on $\mathrm{Arb}(D, v)$ for an Eulerian digraph D.

Rotor-routing is defined on a ribbon digraph. For a directed edge $\overrightarrow{vu}$, let us introduce the notation $\mathrm{nextout}(v, \overrightarrow{vu})$ for the next out-edge after $\overrightarrow{vu}$ in the cyclic order around v.

A *rotor configuration* on D is a function ϱ that assigns to each vertex v an out-edge with tail v. We call $\varrho(v)$ the *rotor* at v.

A configuration of rotor-routing is a pair (x, ϱ), where x is a chip-configuration, and ϱ is a rotor configuration on D. We also call such pairs *chip-and-rotor configuration.*

Given a configuration (x, ϱ), a *routing* at vertex v results in the configuration (x', ϱ'), where ϱ' is the rotor configuration with

$$\varrho'(u) = \begin{cases} \varrho(u) & \text{if } u \neq v, \\ \mathrm{nextout}(v, \varrho(v)) & \text{if } u = v, \end{cases}$$

and $x' = x - \mathbf{1}_v + \mathbf{1}_{v'}$ where v' is the head of $\varrho'(v)$.

Similarly to chip-firing, we call two chip-and-rotor configurations (x_1, ϱ_1) and (x_2, ϱ_2) *linearly equivalent*, if (x_2, ϱ_2) can be reached from (x_1, ϱ_1) by a sequence of routings. We denote this by $(x_1, \varrho_1) \sim (x_2, \varrho_2)$. If D is strongly connected, this is indeed an equivalence relation [14, Proposition 3.4] (symmetry is the only nontrivial property).

It is easy to see that if x and y are chip configurations, and ϱ is an arbitrary rotor configuration, then $(x, \varrho) \sim (y, \varrho)$ if and only if $x \sim_D y$.

Let D be an Eulerian ribbon digraph. Holroyd et al. [8] defined the rotor routing action, which is a simply transitive action of $S(D)$ on $\mathrm{Arb}(D, v)$. Due to lack of space, we do not give the original definition, but the following equivalent definition from [14].

Definition 2.4 (rotor routing action)**.** Let D be an Eulerian ribbon digraph, $x \in S(D)$, and $T, T' \in \mathrm{Arb}(D, v)$. Fix an arbitrary out-edge $\overrightarrow{vw} \in E$ of v. We define $r_v(x, T) = T'$ if and only if $(x, T \cup \overrightarrow{vw}) \sim (\mathbf{0}, T' \cup \overrightarrow{vw})$.

By [14, Proposition 3.16], r_v is a simply transitive action of $S(D)$ on $\mathrm{Arb}(D, v)$, it does not depend on $\overrightarrow{vw}$ (only on v) and it agrees with the action of Holroyd et al. [8].

For undirected graphs, $\mathrm{Arb}(G^{\leftrightarrows}, v)$ is in bijection with the spanning trees of G via forgetting the orientations. Hence one can think of r_v as an action on the spanning trees.

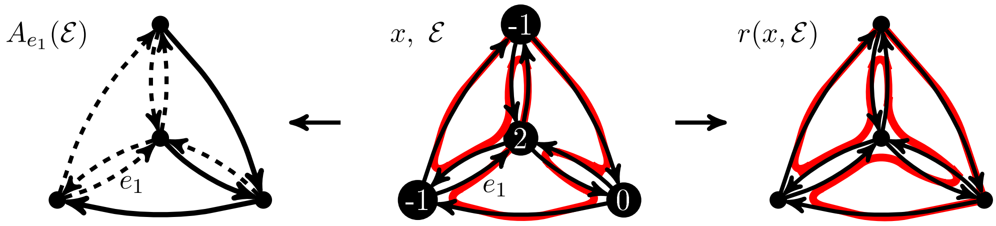


Figure 1: A compatible Eulerian tour, the corresponding arborescence, and an example for the tour-rotor action. See Example 3.4 for the details.

3 A canonical action of Eulerian ribbon digraphs on compatible Eulerian tours

For non-plane graphs, the rotor-routing action depends on the root vertex. However, we show that one can obtain a canonical action by lifting the rotor-routing action to compatible Eulerian tours via Definition 2.1. Our main theorem is that this lifted action is canonical. That is, it does not depend on the root v, and neither on the edge $\overrightarrow{vb}$ used to define the bijection.

Theorem 3.1. *Let D be an Eulerian ribbon digraph, and $\overrightarrow{ub}$ and $\overrightarrow{wz}$ be arbitrary edges. Let $x \in S(D)$, and let $\mathcal{E}$ and $\mathcal{E}'$ be compatible Eulerian tours. If we have*

$$r_u(x, A_{\overrightarrow{ub}}(\mathcal{E})) = A_{\overrightarrow{ub}}(\mathcal{E}'), \qquad \textit{then} \qquad r_w(x, A_{\overrightarrow{wz}}(\mathcal{E})) = A_{\overrightarrow{wz}}(\mathcal{E}').$$

The proof can be found in the full version [13]. As a corollary, the following action is well-defined.

Definition 3.2 (Tour-rotor action)**.** Let D be an Eulerian ribbon digraph. For $x \in S(G)$ and a compatible Eulerian tour $\mathcal{E}$, we define $r(x, \mathcal{E}) = \mathcal{E}'$, where $\mathcal{E}'$ is the unique compatible Eulerian tour such that for an arbitrary edge $\overrightarrow{vb}$, we have $r_v(x, A_{\overrightarrow{vb}}(\mathcal{E})) = A_{\overrightarrow{vb}}(\mathcal{E}')$. We call this the tour-rotor action.

Corollary 3.3. *The tour-rotor action is a well-defined, canonical simply transitive action of the sandpile group of an Eulerian ribbon digraph on the set of compatible Eulerian tours.*

Example 3.4. Figure 1 shows a plane embedded Eulerian digraph. The middle panel shows a chip configuration (white numbers on the vertices) and a compatible Eulerian tour (red curve). The tour-rotor action of the chip configuration on the tour gives the compatible Eulerian tour on the right panel (red curve).

One feels that the tour-rotor action needs to have a nice, canonical definition, that does not need fixing an arbitrary edge.

Problem 3.5. Give a canonical definition for the tour-rotor action.

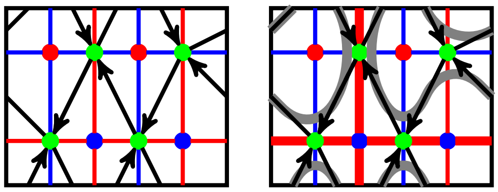
4 Jacobians of embedded graphs

Let G be a graph cellularly embedded into the closed orientable surface Σ. A dual G^* can be defined with respect to Σ: Put a vertex of G^* inside each component of $\Sigma \setminus G$, and connect dual vertices of neighboring components through each edge of G. This way, the edge sets of G and of G^* are in bijection: for an edge $e \in E(G)$, the corresponding edge $e^* \in E(G^*)$ is the one that intersects e. Call the intersection v_e.

The medial digraph of G is an Eulerian digraph $G^{\bowtie}$. The node set of $G^{\bowtie}$ is $\{v_e \mid e \in E(G)\}$. An edge of $G^{\bowtie}$ leads from v_e to v_f, if the edges $e \in E(G)$ and $f \in E(G)$ are incident to a common vertex $v \in V(G)$, and at that vertex, f is the edge following e in the ribbon structure. Below is an example for a G embedded into the torus. By Bouchet [3, Corollary 3.4], the Eulerian tours of $G^{\bowtie}$ are in canonical bijection with the quasitrees of G. See an example for this bijection on the right panel of the following figure. Since the out-degrees are all 2, all the Eulerian tours of the medial digraph are compatible with the embedding. Hence we have the following:

Theorem 4.1. *For an embedded graph G, via the tour-rotor action, $S(G^{\bowtie})$ canonically acts on the quasitrees of G.*

G: vertex, edge
G^*:vertex, edge
$G^{\bowtie}$: vertex, edge



quasitree:
thick red edges
Eulerian tour:
gray curve

Baker at al. [1] defined the Jacobian $Jac(G, \Sigma)$ of an embedded graph. Due to lack of space we are unable to give the definition, but refer to [1]. We have the following results.

Proposition 4.2. *$Jac(G, \Sigma)$ is canonically isomorphic to $S(G^{\bowtie})$.*

Theorem 4.3. *The canonical action of $Jac(G, \Sigma)$ by Baker, Ding and Kim [1] agrees with the tour-rotor action of $S(G^{\bowtie})$.*

The proofs of these two results can also be found in the full version [13].

References

[1] Matthew Baker, Changxin Ding, and Donggyu Kim. The Jacobian of a regular orthogonal matroid and torsor structures on spanning quasi-trees of ribbon graphs. *arXiv:2501.08796*, 2025.

[2] Matthew Baker and Yao Wang. The Bernardi process and torsor structures on spanning trees. *Int. Math. Res. Not. IMRN*, (16):5120–5147, 2018.

[3] André Bouchet. Maps and △-matroids. *Discrete Math.*, 78(1-2):59–71, 1989.

[4] Melody Chan, Thomas Church, and Joshua A. Grochow. Rotor-routing and spanning trees on planar graphs. *Int. Math. Res. Not. IMRN*, (11):3225–3244, 2015.

[5] Nicolaas Govert de Bruijn and Tanja van Aardenne-Ehrenfest. Circuits and trees in oriented linear graphs. *Simon Stevin*, 28:203–217, 1951.

[6] Changxin Ding. The rotor-routing torsor and the bernardi torsor disagree for every non-planar ribbon graph. *arXiv preprint arXiv:2103.01137*, 2021.

[7] J. Ellenberg. What is the sandpile torsor? *Math Overflow, http://mathoverflow.net/questions/83552.*

[8] Alexander E. Holroyd, Lionel Levine, Karola Mészáros, Yuval Peres, James Propp, and David B. Wilson. Chip-firing and rotor-routing on directed graphs. In Vladas Sidoravicius and Maria Eulália Vares, editors, *In and Out of Equilibrium 2*, volume 60 of *Progress in Probability*, pages 331–364. Birkhäuser Basel, 2008.

[9] Tamás Kálmán, Seunghun Lee, and Lilla Tóthmérész. The sandpile group of a trinity and a canonical definition for the planar bernardi action. *Combinatorica*, 42(Suppl 2):1283–1316, 2022.

[10] Criel Merino, Iain Moffatt, and Steven Noble. The critical group of a combinatorial map. *arXiv preprint arXiv:2308.13342*, 2023.

[11] Farbod Shokrieh and Cameron Wright. On torsor structures on spanning trees. *SIAM J. Discrete Math.*, 37(3):2126–2147, 2023.

[12] Carsten Thomassen. Embeddings and minors. In *Handbook of combinatorics, Vol. 1, 2*, pages 301–349. Elsevier Sci. B. V., Amsterdam, 1995.

[13] Lilla Tóthmérész. On canonical sandpile actions of embedded graphs. *arXiv preprint arXiv:2504.05275*, 2025.

[14] Lilla Tóthmérész. Algorithmic aspects of rotor-routing and the notion of linear equivalence. *Discrete Applied Mathematics*, 236:428 – 437, 2018.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The linear Turán number of the 3-graph P_5

(Extended abstract)

Chaoliang Tang[*] Hehui Wu[†] Junchi Zhang[‡]

Abstract

We prove that for any linear 3-graph on n vertices without path of length 5, the number of edges is at most $\frac{15}{11}n$, and the equality holds if and only if the graph is the disjoint union of G_0, a graph with 11 vertices and 15 edges. Thus, $ex_L(n, P_5) \leq \frac{15}{11}n$, and the equality holds if and only if $11|n$.

Keyword: linear Turán number, disjoint structure, link graph, discharging

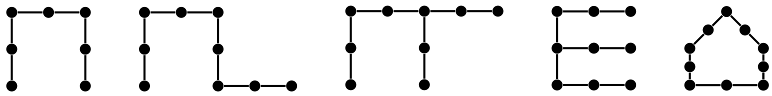
1 Introduction

A k-uniform hypergraph, or a k-graph, is a pair (V, E) such that V is the set of vertices, and E is a family of k-vertex set of V. A linear 3-graph $G(V, E)$ is a 3-uniform hypergraph such that any two vertices are contained in at most one edge. Given a linear 3 graph F, the linear Turán number $ex_L(n, F)$ is the maximal number of edges in any F-free linear 3-graph on n vertices.

Determining the Turán number is the fundamental problem in extremal graph theory. For hypergraphs, determining the linear Turán number can have quite different flavors. A celebrated result is the famous (6,3)-theorem , which is proved by Ruzsa and Szemerédi [?] that $ex_L(n, C_3) = o(n^2)$, where the cycle C_3 is the only linear 3-graph with 3 vertices and 6 edges. Determining the linear Turán number of graph seems surprisingly hard, even for the acyclic cases. Recently Gyárfás et. al. initiated the linear Turán number of some trees in [?], where they proved that $ex_L(n, P_k) \leq 1.5kn$ and determined the linear Turán number of P_3, P_4, B_4 and matchings. In [?], Tang et. al. proved that the linear Turán number of crown E_4 is $1.5n$ and completed the determination of linear Turán number for acyclic 3-graphs with at most 4 edges. A Berge path of length k in a hypergraph consists of $k+1$ distinct vertices $v_1, v_2, ..., v_{k+1}$ and k hyperedges $h_1, h_2, ..., h_k$ such that $\{v_i, v_{i+1}\} \subseteq h_i$ for all $i \in [k]$. In [?]

[*]Shanghai Center for Mathematical Statistics, Fudan Univertsity, 220 Handan Road, Shanghai 200433, China. E-mail: `cltang22@m.fudan.edu.cn`.
[†]Shanghai Center for Mathematical Statistics, Fudan Univertsity, 220 Handan Road, Shanghai 200433, China. E-mail: `hhwu@fudan.edu.cn`.
[‡]Shanghai Center for Mathematical Statistics, Fudan Univertsity, 220 Handan Road, Shanghai 200433, China. E-mail: `jczhang24@m.fudan.edu.cn`.

Győri et. al. showed that number of hyberedges in an n-vertex linear 3-graph without a Berge path of length $k \geq 4$ as a subgraph is at most $(k-1)n/6$. Further, they ask the linear Turán number of path of length larger than 4, see Conjectue 1.1, or Conjecture 4 in [?]. Note that the bound in Conjectue 1.1 is asymptotically sharp for infinitely many pairs of n and k, reached by disjoint union of Steiner triple systems on $2k$ vertices. The cases of linear cycles are more difficult than linear trees and we refer a result of asymptotic Turán number for linear 5-cycle C_5 by Gao et. al. in [?].

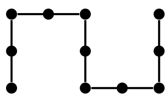


P_3 P_4 B_4 E_4 C_5

Conjecture 1.1. *Let G be an n vertex linear 3-graph, containing no linear path of length $k \geq 5$. Then the number of edges in G is at most $\frac{k}{3}n + cn$, for some universal constant c.*

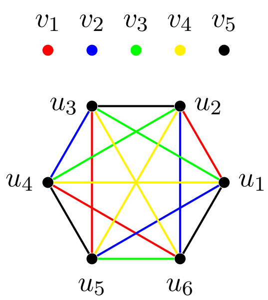
Motivated by the conjecture, we answer for the path of length 5. This is the first known path whose extremal graph is not disjoint union of Steiner triple systems. The methods we use in structural analysis is to assume disjoint subgraph in the minimal counter example, then find some particular 'weak' vertex subsets, which means that they have limited incident edges. Then we could delete these locally sparse set to obtain a contradiction to minimality. We also use the basic method of link graph and degree list on edges which is first introduced by Gyárfás et. al. in [?]. We highly believe these ideas could be used in proving the linear Turán number of longer paths.

In this paper, we determine $ex_L(n, P_5)$, in which P_5, the path of length 5, is the graph with 11 vertices and 5 edges as shown below.



P_5

Theorem 1.2. *Let G be an n vertex linear 3-graph, containing no P_5. Then the number of edges in G is at most $\frac{15}{11}n$, and the equality holds if and only if the graph is the disjoint union of G_0, a graph with 11 vertices and 15 edges as shown below.*



Vertices	$v_1, v_2, \ldots, v_5, u_1, u_2, \ldots, u_6$
Edges	$\{v_1, u_1, u_2\}, \{v_1, u_3, u_5\}, \{v_1, u_4, u_6\}$, $\{v_2, u_1, u_5\}, \{v_2, u_2, u_6\}, \{v_2, u_3, u_4\}$, $\{v_3, u_1, u_3\}, \{v_3, u_2, u_4\}, \{v_3, u_5, u_6\}$, $\{v_4, u_1, u_4\}, \{v_4, u_2, u_5\}, \{v_4, u_3, u_6\}$, $\{v_5, u_1, u_6\}, \{v_5, u_2, u_3\}, \{v_5, u_4, u_5\}$.

Figure 1: Extremal graph G_0

2 Preliminaries

Given a graph G, the neighbours of v in G are denoted by $N_G(v)$, the degree of v in G is denoted by $d_G(v)$, the minimal degree of G is denoted by $\delta(G)$, and the subgraph induced by $S \subseteq V(G)$ is denoted by $G[S]$. For any integer n, we denote by $[n]$ the set of integers $\{1, 2, \ldots, n\}$. For simplicity, if no ambiguity, we identify a graph G with its vertex set, a single vertex set $\{v\}$ with the element v, and we omit a 3-hyperedge $\{u, v, w\}$ as uvw.

The path of length k, denoted by P_k, is the graph with $2k+1$ vertices $\{u_i \mid 1 \leq i \leq 2k+1\}$ and k edges $\{u_{2j-1}u_{2j}u_{2j+1} \mid 1 \leq j \leq k\}$. Similarly, the cycle of length k, denoted by C_k, is the linear cycle with k edges.

Given a vertex subset $S \subseteq G$, let $E_i(S) = \{e \in E(G) || e \cap S| = i\}$, for $1 \leq i \leq 3$, and $E_{icd}(S) = E_1(S) \cup E_2(S) \cup E_3(S)$, the set of all edges incident to S. Note that $E_{icd}(v) = E_1(v) = d_G(v)$ for a single vertex v and one should be careful with the difference between $E(G[S])$ and $E_{icd}(S)$.

3 Main result

Now we give a sketch of the proof of Theorem 1.2. Assume G is a minimal (first with respect to number of vertices and then with respect to number of edges) linear 3-graph, not isomorphic to disjoint union of G_0, with edge density at least $\frac{15}{11}$ such that P_5 is not a subgraph of G. Let $|V(G)| = n$, then $|E(G)| \geq \frac{15}{11}n$. The edge density of a linear 3-graph with less than 11 vertices is always less than $\frac{15}{11}$ (see Steiner triple systems in [?]), so we can assume $n \geq 11$. Now the critical lemma is to give a lower bound of $E_{icd}(S)$ for every vertex set S, which is proved by contradiction to minimality of G after deleting S from G.

Lemma 3.1. *For any vertex set $S \subseteq G$, $|E_{icd}(S)| \geq \frac{15}{11}|S|$.*

Corollary 3.2. *G is connected and $\delta(G) \geq 2$.*

The rest of the proof is divided into 2 parts. Firstly, we assume there are two disjoint P_2 in G, and get a contradiction. Secondly, we use the structural result we just got to give an upper bound of the edge density.

The first part is to prove that there are no two disjoint P_2 in G. We will need some new tools, which are called 2-centers and center pairs, to analyse disjoint structure in G.

Lemma 3.3. *There are no two disjoint P_2 in G.*

Definition 3.4. *Let U be a subgraph of G.*

1. *A vertex u_0 is called a **k-center** of U if there is no path in U starting from u_0 with length larger than k.*

2. *A pair of vertices $\{u_1, u_2\} \subseteq U$ is called a **center pair** of U if they are not adjacent and every edge in U is incident to one of them.*

Example 3.5. *For example, a single edge is a graph with three 2-centers and no center pairs; S_3 is a graph with no center pairs and seven 2-centers; C_4 is a graph with one center pair and no 2-centers; P_4 is a graph with one 2-center and one center pair (see Figure 2, the red vertices are 2-centers and the pairs connected by a red line are center pairs).*

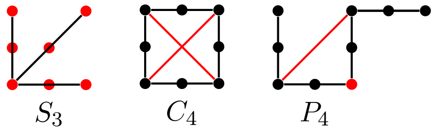


Figure 2: S_3 and P_4

One of the key idea of our proof is to consider the 2-centers and center pairs in the disjoint structures. Assume there is a connecting path between different disjoint structure, then we will prove that we can always extend the path with exception of using 2-centers and center pairs. This is shown by the following two observation. Their proofs are easily derived by contradiction, leading by the existence of P_5.

Lemma 3.6. *Let U be a non-star subgraph, V be a star disjoint from U in G and e be an edge connecting U and leaves of V. Then e consists of a 2-center from one of them and a center pair from the other.*

Lemma 3.7. *Let U, V be two non-star subgraphs of G and P be a path connecting them by W. If $|W| = 0$, then $l(P) = 1$. Moreover, the only edge e in P consists of a 2-center from one of them and a center pair from the other.*

Now consider the remaining graph after deleing vertices of any P_2 from G. Using Lemma 3.6 and Lemma 3.7, we first show that there is no P_4 or C_4 in the remaining graph. Futhermore, we show that any P_3 in $G - P_2$ has to be a component. Moreover, we deduce one step more that there is no P_3 or C_3 in $G - P_2$. Finally, we prove there is no P_2 in the remaining graph.

In these proofs we always assume there is such structure, then by Lemma 3.6 and Lemma 3.7 we can conclude the number of edges is limited, which implies the graph is locally sparse and contradicts to Lemma 3.1.

For the second part, link graph is the main tool to do structure analysis. Specifically speaking, we give a bound on the degree list of any edge of G. We say a triple of numbers $(x_1, x_2, x_3) \geq (y_1, y_2, y_3)$, if $x_i \geq y_i, i = 1, 2, 3$.

Definition 3.8. *Given a linear 3-graph G and a vertex set $S = \{x_1, x_2, \dots, x_k\}$ of G. The edge-labelled link graph (H, l) of S, denoted by $(H, l) = L(G, S)$, is the simple graph H with an edge labelling l, such that $V(H) = V(G) - S, E(H) = \{xy : x, y \in V(H)$ and $(x, y, z) \in E(G)$ for some $z \in S\}$, and $l(xy) = z$ if $\{x, y, z\} \in E(G)$ for some $z \in S$. Note that l is well-defined by linearity, of G.*

Lemma 3.9. *There's no edge $\{a, b, c\}$ in G such that $(d_G(a), d_G(b), d_G(c)) \geq (5, 5, 4)$.*

The proof of Lemma 3.9 is by taking 4 edges labelled by a, 4 edges labelled by b and 3 edges labelled by c in the link graph $G(e)$, then force the structure of $G(e)$ by Lemma 3.3.

Lemma 3.10. *For any vertex v in G, if $d_G(v) \geq 6$, then for any neighbour u of v, $d_G(u) \leq 3$. If $d_G(v) \geq 5$, then for any non-neighbour u of v, $d_G(u) \leq 2$.*

Proof. If not, then in the edge-labelled link graph $G(v, u)$, denoted by H, we can find 5 edges labelled by v together with 3 edges labelled by 2. Since by Lemma 3.3 there is no dispoint P_2, every two edge labelled by u must intersect four edges labelled by v.

It impies that each edge labelled by u must intersect two edges labelled by v by linearity. However there are only 5 edges labelled by v thus by the pigeonhole principle there must be one edge labelled by v intersect two edges labelled by u, say e_1 and e_2. Now e_1 and e_2 can intersect at most 3 edges labelled by v, contradiction. ∎

Now we can prove our main result.

Proof of Theorem 1.2:

If there's no vertex of deree larger than 4 in G, then the edge density is $\frac{4}{3}$ and we are done. Now suppose there's some vertex v in G such that $d_G(v) = k \geq 5$, let $\{v, v_{2i-1}, v_{2i} \mid i \in [k]\}$ be the edges incident to v.

When $k \geq 6$, we first claim that every other vertex in G is a neighbour of v. Assume there is a non-neighbour u, by Corollary 3.2, $d_G(u) \geq 2$. Take any two edges e_1, e_2 incident to u. There are at least $k-4 \geq 2$ edges adjacent to v disjoint with e_1, e_2 by linearity, which is a P_2 disjoint from e_1 and e_2, contradiction to Lemma 3.3. Then by Lemma 3.10, $d_G(u) \leq 3, \forall u \neq v$. So $d_G(v) = \frac{n-1}{2}$, $d_G(u) \leq 3$, $u \neq v$, and $|E(G)| \leq \frac{1}{3}(\frac{n-1}{2} + 3(n-1)) < \frac{15}{11}n$, contradicting Lemma 3.1.

When $k = 5$, any non-neighbour u of v has degree less than 3 by Lemma 3.10. Meanwhile, by Lemma 3.9, $d_G(v_{2i-1}) + d_G(v_{2i}) \leq 8, i \in [5]$. So $|E(G)| \leq \frac{1}{3}(5 + 2(n-11) + 8 \times 5) \leq \frac{15}{11}n$. Note that equality holds only when $n = 11$ and $d_G(v_{2i-1}) + d_G(v_{2i}) = 8, i \in [5]$. Thus we can see that for any P_5-free 3-graph G, its edge density is no more than $\dfrac{15}{11}$. ∎

Acknowledgements: The authors want to thank Jiayi Lu and Zhicheng Liu for their useful discussions in the graph theory group of Fudan University.

References

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Lower bounds on the minimal dispersion of point sets via cover-free families

(Extended abstract)

Matěj Trödler* Jan Volec† Jan Vybíral‡

Abstract

We elaborate on the intimate connection between the largest volume of an empty axis-parallel box in a set of n points from $[0,1]^d$ and cover-free families from the extremal set theory. This connection was discovered in a recent paper of the authors. In this work, we apply a very recent result of Michel and Scott to obtain a whole range of new lower bounds on the number of points needed so that the largest volume of such a box is bounded by a given ε. Surprisingly, it turns out that for each of the new bounds, there is a choice of the parameters d and ε such that the bound outperforms the others.

1 Introduction

Let $X \subset [0,1]^d$ be a (finite) set of points. There are several ways of how to measure whether the points of X are well spread. One way, which has recently attracted the attention of many researchers, is the so-called *dispersion*. The dispersion of X is the volume of the largest axis-parallel box in $[0,1]^d$ that contains no point of X, i.e.,

$$\operatorname{disp}(X) := \sup_{B: B\cap X=\emptyset} |B|.$$

Here, the supremum is taken over all the boxes $B = \prod_{i=1}^d (a_i, b_i)$, where $0 \le a_i < b_i \le 1$ for all $i \in [d]$, and $|B|$ stands for the (Lebesgue) volume of B. The study of this notion goes back to [4, 9] and [10].

*Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Trojanova 13, 12000 Praha, Czech Republic. E-mail: trodlmat@fjfi.cvut.cz. Supported by the grant P202/23/04720S of the Grant Agency of the Czech Republic.

†Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, Prague, 160 00, Czech Republic. E-mail: jan@ucw.cz. Supported by the grant 23-06815M of the Grant Agency of the Czech Republic.

‡Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Trojanova 13, 12000 Praha, Czech Republic. E-mail: jan.vybiral@fjfi.cvut.cz. Supported by the grant P202/23/04720S of the Grant Agency of the Czech Republic.

A very natural extremal problem is to determine the smallest possible dispersion for a given number of points in $[0,1]^d$. For fixed integers d and n, we denote the minimum dispersion of an n-point set $X \subseteq [0,1]^d$ by

$$\operatorname{disp}^*(n,d) := \inf_{\substack{X \subset [0,1]^d \\ |X|=n}} \operatorname{disp}(X).$$

It is sometimes convenient to study this problem in the inverse setting: given an integer d and $\varepsilon \in (0,1)$, how many points do we need to place into the d-dimensional unit cube (and how?) so that their dispersion can be as small as ε? The extremal problem is reflected by the quantity

$$\begin{aligned} N(\varepsilon,d) &:= \min\{n \in \mathbb{N} : \operatorname{disp}^*(n,d) \le \varepsilon\} \\ &= \min\{n \in \mathbb{N} : \exists X \subset [0,1]^d \text{ with } |X| = n \text{ and } \operatorname{disp}(X) \le \varepsilon\}. \end{aligned}$$

As mentioned above, a number of upper and lower bounds on the quantities $\operatorname{disp}^*(n,d)$ and $N(\varepsilon,d)$ have been established. Below, we mention only a selection of those results that are relevant for our work, and refer to [7] for a detailed overview. The elementary lower bound $\operatorname{disp}^*(n,d) \ge \frac{1}{n+1}$ was improved in [3] to $\operatorname{disp}^*(n,d) \ge \frac{5}{4(n+5)}$. As the next step, for $d \ge 2$ and $\varepsilon \in (0,1/4)$, it was shown in [1] that $N(\varepsilon,d) \ge \frac{\log_2 d}{8\varepsilon}$, which seems to be the first lower bound on $N(\varepsilon,d)$ that grows with d.

A further improvement was obtained by Bukh and Chao [2], who proved

$$\operatorname{disp}^*(n,d) \ge \frac{1}{e} \cdot \frac{2d}{n}\left(1 - \frac{4d}{n^{1/d}}\right). \tag{1}$$

This can be translated into the following lower bound on the inverse problem:

$$N(\varepsilon,d) \ge \frac{1}{e} \cdot \frac{d}{\varepsilon} \quad \text{for every } \varepsilon \le (8d)^{-d}. \tag{2}$$

The bounds (1) and (2) are nearly optimal when d is fixed, and we study the limiting behavior for n (or $1/\varepsilon$) tending to infinity. However, note that (1) yields (2) only for very small values of ε.

Regarding upper bounds on $N(\varepsilon,d)$, a series of recent papers [6, 11, 13] have shown that

$$N(\varepsilon,d) \le \frac{C \log d \cdot \log \frac{1}{\varepsilon}}{\varepsilon^2}. \tag{3}$$

In [12] we have shown that this bound is nearly optimal for ε being large:

$$N(\varepsilon,d) > \frac{c \log d}{\varepsilon^2 \cdot \log \frac{1}{\varepsilon}} \quad \text{for every } \varepsilon \in \left(\frac{1}{4\sqrt{d}}, \frac{1}{4}\right). \tag{4}$$

On one hand, (4) matches (3) up to a polylogarithmic factor in $1/\varepsilon$. On the other hand, (4) holds only for ε of the order at least $1/\sqrt{d}$.

The aim of this work is to establish lower bounds on $N(\varepsilon,d)$ that are valid also when $\varepsilon < 1/\sqrt{d}$. Our main result is the following.

Theorem 1. *Fix a positive integer k. There exists a constant c_k such that if d is a positive integer satisfying $d \geq d^{\frac{k}{k+1}} + k$ and $\varepsilon \in \left(0, 2^{-k-2}\right)$, then the following is true.*

(i) If $\varepsilon \geq d^{-\frac{k^2}{k+1}}$ then $N(\varepsilon, d) \geq c_k \cdot \varepsilon^{-\frac{k+1}{k}}$.

(ii) If $\varepsilon < d^{-\frac{k^2}{k+1}}$ then $N(\varepsilon, d) \geq c_k \cdot d^{\frac{k}{k+1}}\, \varepsilon^{-1}$.

Note that if $k = 1$ and ε is a constant, then the bound in Part (i) is inferior to (4) by a factor of $\log d$.

2 Part (i) of Theorem 1 and cover-free families

We utilize a strong connection between lower bounds on $N(\varepsilon, d)$ and a certain problem in extremal set theory. The following is a generalization of the notion of an *r-cover-free family*, a crucial notion in the proof of (4) in [12].

Definition 1. *Let $\mathcal{F}$ be a family of subsets of a ground set X. We say that $\mathcal{F}$ is* (k, r)-cover-free *if for all $A_1, \ldots, A_k \in \mathcal{F}$ and all $B_1, \ldots, B_r \in \mathcal{F} \setminus \{A_1, \ldots, A_k\}$ it holds*

$$\bigcap_{i=1}^{k} A_i \not\subseteq \bigcup_{j=1}^{r} B_j.$$

Let $C(k, r, d)$ be the smallest size of the ground set such that a d-element (k, r)-cover-free family exists, i.e.,

$$C(k, r, d) := \min\{n \in \mathbb{N} \colon \exists (k, r)\text{-cover-free family } \mathcal{F}_n \text{ on } [n] \text{ with } |\mathcal{F}_n| = d\}.$$

Note that the case $k = 1$ corresponds to r-cover-free families introduced in 1964 by Kautz and Singleton [5]. The following very recent result of Michel and Scott [8] plays a crucial role for us.

Theorem 2 ([8, Lemma 2.2])**.** *If k, s, t and d are positive integers satisfying $d \geq k + t$, then $C(k, s + t, d) \geq \frac{1}{2k^k} \cdot \min\{d^k,\ s(k + t)^k\}$.*

The connection between lower bounds on $N(\varepsilon, d)$ and cover-free families was first discovered in [12]. Our approach here is similar, but more elaborate and more flexible. We define a very specific set of axis-parallel boxes of volume at least ε, which allows us to translate the bounds on $N(\varepsilon, d)$ from below to the existence of (k, r)-cover-free families. The role of k in this reduction will be such that we obtain a whole range of new lower bounds, each valid only for ε that is appropriately bounded from below.

Fix the dimension d. For all positive integers k and ℓ with $k + \ell \leq d$ and every $u \in (0, 1)$, we define a collection of boxes that have some k sides equal to $(0, u)$ and some other ℓ sides equal to $(u, 1)$. Specifically, for a given set $K \subseteq [d]$ with $|K| = k$ and $L \subseteq [d] \setminus K$ with $|L| = \ell$, let $B^{K,L} \subseteq [0, 1]^d$ be defined as

$$B_u^{K,L} := I_1 \times I_2 \times \cdots \times I_d, \quad \text{where} \begin{cases} I_i = (0, u) & \text{for } i \in K, \\ I_i = (u, 1) & \text{for } i \in L, \\ I_i = (0, 1) & \text{for } i \in [d] \setminus (K \cup L). \end{cases}$$

Let $\mathcal{B}(d,k,\ell,u) := \left\{ B_u^{K,L} \subseteq [0,1]^d : K \in \binom{[d]}{k}, L \in \binom{[d]\setminus K}{\ell} \right\}$. We choose ℓ and u based on the values of k and ε to ensure that the boxes in $\mathcal{B}(d,k,\ell,u)$ have volume at least ε.

Lemma 1. *Let k and d be positive integers, let $\varepsilon \in \left(0, 2^{-k-2}\right)$, and set $u := (4\varepsilon)^{1/k}$. If $d \geq k + \left\lfloor \frac{1}{u} \right\rfloor$, then $|B| > \varepsilon$ for every $B \in \mathcal{B}\left(d, k, \left\lfloor \frac{1}{u} \right\rfloor, u\right)$.*

Let us now describe how we assign to a point set $X \subset [0,1]^d$ a certain family of subsets of $[d]$. Given $u \in (0,1)$ and $j \in [d]$, we define

$$F_j^u := \{x \in X : (x)_j < u\}.$$

The connection between minimal dispersion and extremal set theory, which plays the central role in our proof of Part (i) of Theorem 1, is the following.

Lemma 2. *Let d be a positive integer, $u \in (0,1)$, k and ℓ positive integers with $d \geq k + \ell$, and $X \subset [0,1]^d$. If X intersects every box in $\mathcal{B}(d,k,\ell,u)$ then the family $\{F_1^u, F_2^u, \ldots, F_d^u\}$ is (k,ℓ)-cover-free.*

Lemmas 1 and 2 yield a lower bound on $N(\varepsilon,d)$ in terms of $C(k,r,d)$.

Corollary 3. *If k is a positive integer and $\varepsilon \in \left(0, 2^{-k-2}\right)$ such that they satisfy $d \geq k + \left\lfloor (4\varepsilon)^{-1/k} \right\rfloor$, then $N(\varepsilon, d) \geq C\left(k, \left\lfloor (4\varepsilon)^{-1/k} \right\rfloor, d\right)$.*

3 Part (ii) of Theorem 1

The following is a rescaling-type observation analogous to [1, Lemma 1], which was stated for $\operatorname{disp}^*(n,d)$.

Lemma 3. *If d and b are positive integers and $\varepsilon \in \left(0, \frac{1}{b}\right)$, then*

$$N(\varepsilon, d) \geq b \cdot N(b \cdot \varepsilon, d).$$

Lemma 3 serves as a tool for extending the validity of the lower bound in Part (i) of Theorem 1 to the regime of ε in Part (ii). Indeed, for the k-th bound of Part (ii), we apply the lemma with $b := d^{-\frac{k^2}{k+1}} \big/ \varepsilon^{-1}$ to move in the range of parameters, where the k-th bound of Part (i) applies.

4 Conclusion

Theorem 1 yields a whole series of lower bounds on $N(\varepsilon, d)$, conveniently indexed by k. Moreover, for a fixed k, we get a lower bound that changes its nature at $\varepsilon = d^{-\frac{k^2}{k+1}}$. The bounds are relevant when d is sufficiently large. Specifically, if $d \geq 2^{k+1}$ then $k + d^{\frac{k}{k+1}} \leq k + \frac{d}{2} \leq d$, thus Theorem 1 applies.

Also note that Part (ii) of Theorem 1 approaches (2) as k tends to infinity. However, the interval where ε must lie in gets smaller when k increases.

Finally, let us discuss the interplay between the lower bounds of Theorem 1 for different values of k. To simplify the discussion, we neglect the factors depending only on k, and we also assume that d is sufficiently large. The two bounds for a fixed $k = k_0$ are better than the bounds for other values of k if $d^{-k_0} \le \varepsilon < d^{-(k_0-1)}$, i.e., when the value $d^{-\frac{k^2}{k+1}}$ lies in this interval. Therefore, for d large enough, we get the best lower bound from Theorem 1 in the following way:

1. Choose the integer k such that $\varepsilon \in \left[d^{-k}, d^{-(k-1)}\right)$.
2. If $\varepsilon \ge d^{-\frac{k^2}{k+1}}$ then $N(\varepsilon, d) \ge c_k \cdot \varepsilon^{-\frac{k}{k+1}}$ by Part (i) for k.
3. Otherwise $\varepsilon < d^{-\frac{k^2}{k+1}}$, thus $N(\varepsilon, d) \ge c_k \cdot d^{\frac{k}{k+1}} \varepsilon^{-1}$ by Part (ii) for k.

References

[1] C. Aistleitner, A. Hinrichs and D. Rudolf, *On the size of the largest empty box amidst a point set*, Discrete Appl. Math. 230 (2017), 146–150.

[2] B. Bukh and T. Chao, *Empty axis-parallel boxes*, Int. Math. Res. Notices 18 (2022), 13811–13828.

[3] A. Dumitrescu and M. Jiang, *On the largest empty axis-parallel box amidst n points*, Algorithmica 66 (2013), 225–248.

[4] E. Hlawka, *Abschätzung von trigonometrischen Summen mittels diophantischer Approximationen*, Österreich. Akad. Wiss. Math.- Naturwiss. Kl. S.-B. II, 185 (1976), 43–50.

[5] W. H. Kautz and R. C. Singleton, *Nonrandom binary superimposed codes*, IEEE Trans. Inform. Theory 10 (1964), 363–377.

[6] A. E. Litvak, *A remark on the minimal dispersion*, Commun. Contemp. Math. 23 (2021), 2050060.

[7] A. E. Litvak and G. V. Livshyts, *New bounds on the minimal dispersion*, J. Compl. 72 (2022), 101648.

[8] L. Michel and A. Scott, *Lower bounds for graph reconstruction with maximal independent set queries*, Theor. Comput. Sci. 1034 (2025), 115121

[9] H. Niederreiter, *A quasi-Monte Carlo method for the approximate computation of the extreme values of a function*, In: P. Erdös, L. Alpár, G. Halász, A. Sárközy, (eds.), Birkhäuser Basel, Basel (1983), 523-529.

[10] G. Rote and R. F. Tichy, *Quasi-Monte Carlo methods and the dispersion of point sequences*, Math. Comput. Modelling 23 (1996), 9–23.

[11] J. Sosnovec, *A note on the minimal dispersion of point sets in the unit cube*, Eur. J. Comb. 69 (2018), 255–259.

[12] M. Trödler, J. Volec, and J. Vybíral, *A tight lower bound on the minimal dispersion.* European J. Combin. 120 (2024), 103945.

[13] M. Ullrich and J. Vybíral, *An upper bound on the minimal dispersion*, J. Compl. 45 (2018), 120–126.

The Hierarchy of Saturating Matching Numbers

Hans U. Simon[*] Jan Arne Telle[†]

June 7, 2025

Abstract

In this paper, we study three matching problems all of which came up quite recently in the field of machine teaching. The cost of a matching is defined in such a way that, for some formal model of teaching, it equals (or bounds) the number of labeled examples needed to solve a given teaching task. We show how the cost parameters associated with these problems depend on each other and how they are related to other combinatorial parameters (like, for instance, the VC-dimension).[1]

1 Introduction

A binary concept class C contains classification rules (called concepts) which split the instances in its domain $X = \mathrm{dom}(C)$ into positive examples (labeled 1) and negative examples (labeled 0). Formally, a concept can be identified with a function $f : X \rightarrow \{0, 1\}$.

This paper is inspired by recent work in the field of machine teaching. In machine teaching, the goal is to find a collection S of helpful examples, called a *teaching set for the unknown target concept f in the class C*, which enables a learner (or a learning algorithm) to "infer" f from S. The three matching problems that we discuss in this paper are related to teaching models that came up quite recently [4, 14, 5, 9]. We will look at these problems from a purely combinatorial perspective but, as a service to the interested reader, we will indicate the connection to machine teaching by giving a reference to the relevant literature at the appropriate time.

2 Preliminaries

Let us first recall some definitions from learning theory. Let X be a finite set and let 2^X denote the set of all functions $f : X \rightarrow \{0, 1\}$. A family $C \subseteq 2^X$ is called a *concept class over the domain* $X = \mathrm{dom}(C)$. A set $S = \{(x_1, b_1), \ldots, (x_m, b_m)\}$ with m distinct elements

[*]Faculty of Mathematics, Ruhr-University Bochum, Germany. E-mail: hans.simon@rub.de.

[†]Department of Informatics, University of Bergen, Norway. E-mail: jan.arne.telle@uib.no.

[1]This is a short version of a longer arXiv paper [15].

$(x_i, b_i) \in X \times \{0,1\}$ is called a *labeled sample of size* $|S| = m$. A concept $f \in C$ is said to be *consistent with* S if $f(x_i) = b_i$ for $i = 1, \ldots, m$. A labeled sample S is said to be *realizable by* C if C contains a concept that is consistent with S. A C*-saturating matching* is a mapping M that assigns to each $f \in C$ a labeled sample $M(f)$ such that the following holds: first f is consistent with $M(f)$; second $f \neq g \in C$ implies that $M(f) \neq M(g)$. The *cost of* M is defined as the size of the largest labeled sample that is assigned to one of the concepts in C, i.e., $\mathrm{cost}(M) = \max_{f \in C} |M(f)|$.

In this paper, we compare 8 combinatorial parameters related to saturating matchings, called SMN, SMN′, AMN, AMN′, GMN, GMN′, VCD and RTD. The first 6 of them we now proceed to define. The parameter $\mathrm{SMN}(C)$, whose relevance for machine teaching is explained in [14], is defined as the cost of a cheapest C-saturating matching. $\mathrm{SMN}'(C)$ is defined as the smallest number d such that the number of C-realizable labeled samples of size at most d is greater than or equal to $|C|$. A C-saturating matching M is said to have the *antichain property* if the sets $M(f)$ with $f \in C$ form an antichain. The parameter $\mathrm{AMN}(C)$, whose relevance for machine teaching is explained in [9], is defined as the cost of a cheapest C-saturating matching with the antichain property. $\mathrm{AMN}'(C)$ is defined as the smallest d such that the C-realizable labeled samples of size at most d contain an antichain of size $|C|$. A C-saturating matching M' is called a *direct improvement* of another C-saturating matching M if M' differs from M only on a single concept f and the size of $M'(f)$ is smaller than the size of $M(f)$. A C-saturating matching M is called *greedy* if it does not admit for direct improvements. The parameter $\mathrm{GMN}(C)$, whose relevance for machine teaching is explained in [5], is defined as the largest possible cost of a greedy C-saturating matching. It is not hard to see that $\mathrm{SMN}(C)$ is the smallest possible cost of a greedy C-saturating matching. We define $\mathrm{GMN}'(C) := \min\left\{d : \sum_{i=0}^{d} \binom{|X|}{i} \geq |C|\right\}$. By $\mathrm{VCD}(C)$, we denote the well known VC-dimemsion of the concept class C. By $\mathrm{RTD}(C)$, we denote the so-called *recursive teaching dimension of* C. The interested reader may consult [17, 2] for a definition of $\mathrm{VCD}(C)$ and [18, 3] for a definition of $\mathrm{RTD}(C)$.

Fig. 1 below visualizes some $\leq$-relations among the combinatorial parameters that we associate with each concept class. Most of these $\leq$-relations are obvious. Here we briefly argue for the less obvious ones:

1. $\mathrm{GMN}'(C) \leq \min\{\mathrm{VCD}(C), \mathrm{RTD}(C)\}$.
   This easily follows from Sauer's Lemma [11, 12], which states that $|C| \leq \sum_{i=0}^{\mathrm{VCD}(C)} \binom{|X|}{i}$. Sauer's Lemma is known to hold also with the RTD in place of the VCD [10].

2. $\mathrm{AMN}(C) \leq \min\{\mathrm{VCD}(C), \mathrm{RTD}(C)\}$.
   This inequality is from [9] where it was shown that a variant of Balbach's subset-teaching dimension [1] is lower-bounded by $\mathrm{AMN}(C)$ and upper-bounded by $\mathrm{VCD}(C)$ and by $\mathrm{RTD}(C)$.

3. $\mathrm{AMN}'(C) \leq \mathrm{GMN}'(C)$.
   The proof of this inequality is more challenging and will be given in Section 3 below.

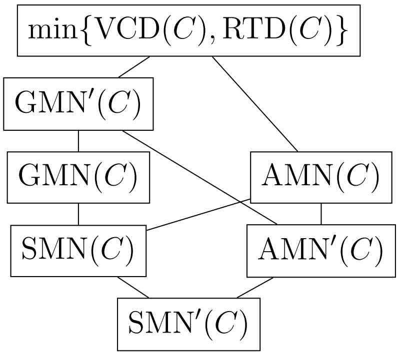


Figure 1: The parameter hierarchy: each edge represents a $\leq$-relation among two of the combinatorial parameters where higher parameter values are found higher up in the hierarchy.

3 Bounding AN′ by GMN′

The (k, n)-*family* is defined as the family of concept classes C with $|C| = k$ and $|\mathrm{dom}(C)| = n$. We denote this family by $\mathcal{C}_{k,n}$.

Example 3.1. *For $k \geq 2$ and $n \geq \log k$, we denote by $C_{k,n}$ the concept class in $\mathcal{C}_{k,n}$ given by*

$$\forall i = 0, 1, \ldots, k-1, j = 0, 1, \ldots, n-1 : i = \sum_{j=0}^{n-1} c_i(x_j)2^j \ .$$

In other words, $C_{k,n} = \{c_0, c_1, \ldots, c_{k-1}\}$, $\mathrm{dom}(C_{k,n}) = \{x_0, x_1, \ldots, x_{n-1}\}$ and $c_i(x_j)$ is the j-th bit in the binary representation of i.

It is easy to characterize the $C_{k,n}$-realizable samples:

Remark 3.2. *A labeled sample S over $\mathrm{dom}(C_{k,n})$ is $C_{k,n}$-realizable iff $\sum_{j:(x_j,1)\in S} 2^j \leq k-1$.*

According to the following quite recent result, the class $C_{k,n} \in \mathcal{C}_{k,n}$ has the smallest number of realizable labeled samples of a given size:

Theorem 3.3 (Main Theorem in [16])**.** *For each $1 \leq d \leq n$, the number of C-realizable labeled samples of size d, with C ranging over all concept classes in $\mathcal{C}_{k,n}$, is minimized by setting $C = C_{k,n}$.*[2]

Since the minimizer $C_{k,n}$ of the number of realizable labeled samples of size d is the same for all possible choices of d, we immediately obtain the following result:

[2]The main theorem in [16] is not stated in this form, but it is equivalent to what is claimed in Theorem 3.3.

Corollary 3.4. $\mathrm{SMN}'(C)$ *with* C *ranging over all concept classes in* $\mathcal{C}_{k,n}$*, is maximized by setting* $C = C_{k,n}$.

Here is another (less immediate than Corollary 3.4) application of Theorem 3.3:

Theorem 3.5. *For each concept class* C*, we have that* $\mathrm{AMN}'(C) \le \mathrm{GMN}'(C)$.

Proof. For each concept class C, let $\mathrm{AMN}''(C)$ be the smallest number d such that there exist $|C|$ many C-realizable labeled samples of size d. Clearly $\mathrm{AMN}'(C) \le \mathrm{AMN}''(C)$. It suffices therefore to show that, for each concept class C, we have that $\mathrm{AMN}''(C) \le \mathrm{GMN}'(C)$. Let C_0 be an arbitrary but fixed concept class. We set

$$k := |C_0| \ , \ \ell := \lfloor \log k \rfloor \ \text{ and } \ n := |\mathrm{dom}(C_0)| \ .$$

Note that $k \le 2^n$ because there can be at most 2^n distinct concepts over a domain of size n. By definition, $\mathrm{GMN}'(C_0)$ is the smallest number d such that $\sum_{i=0}^{d} \binom{n}{i} \ge k$. Note that $\mathrm{GMN}'(C)$ is the same number for all concept classes from the family $\mathcal{C}_{k,n}$. Specifically $\mathrm{GMN}'(C_0) = \mathrm{GMN}'(C_{k,n})$. On the other hand, it is immediate from Theorem 3.3 (and the fact that labeled samples of the same size always form an antichain) that

$$\mathrm{AMN}''(C_{k,n}) = \max\{\mathrm{AMN}''(C) : C \in \mathcal{C}_{k,n}\} \ .$$

Thus, if we knew that $\mathrm{AMN}''(C_{k,n}) \le \mathrm{GMN}'(C_{k,n})$, we could conclude that

$$\mathrm{AMN}''(C_0) \le \mathrm{AMN}''(C_{k,n}) \le \mathrm{GMN}'(C_{k,n}) = \mathrm{GMN}'(C_0) \ .$$

It suffices therefore to verify the inequality $\mathrm{AMN}''(C_{k,n}) \le \mathrm{GMN}'(C_{k,n})$. We will make use of the following auxiliary result.

Claim 1: $\mathrm{GMN}'(C_{k,n}) \le \ell$.

Proof of Claim 1: Since $k \le 2^n$, it follows that $n \ge \lceil \log k \rceil$. The function $n \mapsto \sum_{i=0}^{\ell} \binom{n}{i}$ is monotonically increasing with n. But even for $n = \lceil \log k \rceil$, we have that

$$S := \sum_{i=0}^{\ell} \binom{\lceil \log k \rceil}{i} \ge k$$

for the following reason:

- If k is a power of 2, then $\ell = \log k = \lceil \log k \rceil$ and, therefore, $S = 2^\ell = k$.
- If k is not a power of 2, then $\ell = \lfloor \log k \rfloor$ and $\lceil \log k \rceil = \ell + 1$. This implies that $S = 2^{\ell+1} - 1 > 2^{\log k} - 1 = k - 1$ and, because S is an integer, it implies that $S \ge k$.

It follows from this discussion that $\mathrm{GMN}'(C_{k,n}) \le \ell$.

Claim 2: For each $d \in [\ell]$, there exists an injective mapping f which transforms an unlabeled sample of size at most d over domain[3] $X = \{x_0, x_1, \ldots, x_{n-1}\}$ into a $C_{k,n}$-realizable labeled sample of size exactly d.

Proof of Claim 2: Let $U \subseteq X$ be an unlabeled sample of size at most d. We define $f(U)$ as the (initially empty) labeled sample that is obtained from U as follows:

1. For each $x_j \in U$ with $\ell \leq j \leq n-1$, insert $(x_j, 0)$ into $f(U)$.
2. For each $x_j \in U$ with $0 \leq j \leq \ell - 1$, insert $(x_j, 1)$ into $f(U)$.
3. While $|f(U)| < d$, pick some instance x_j from $\{x_0, \ldots, x_{\ell-1}\} \setminus f(U)$ and insert $(x_j, 0)$ into $f(U)$.

It is obvious that, after Step 3, the set $f(U)$ is of size d. Moreover

$$\sum_{j:(x_j,1)\in U} 2^j \leq \sum_{j=0}^{\ell-1} 2^j = 2^\ell - 1 \leq k-1 \ ,$$

which, according to Remark 3.2, implies that $f(U)$ is $C_{k,n}$-realizable. Finally observe that U can be reconstructed from $f(U)$:

$$U = \{x_j : (\ell \leq j \leq n-1 \wedge (x_j, 0) \in f(U)) \vee (0 \leq j \leq \ell - 1 \wedge (x_j, 1) \in f(U)\} \ .$$

Hence f is injective, which completes the proof of Claim 2.

The proof of the theorem can now be accomplished as follows. Set $d := \mathrm{GMN}'(C_{k,n})$. Then there exist at least k distinct unlabeled samples of size at most d over domain X. Thanks to Claim 1, we know that $d \leq \ell$. Thanks to Claim 2, we may now conclude that there exist at least k distinct labeled $C_{k,n}$-realizable samples over domain X, each of size exactly d. This shows that $\mathrm{AMN}''(C_{k,n}) \leq d$. □

4 Final Remarks

The full paper [15] contains some additional results that we could not discuss in this brief abstract:

- We show that each of the $\leq$-relations visualized in Fig. 1 can be made strict by choosing C properly. Hence this parameter hierarchy is proper.
- Class- and domain-monotonicity are desirable properties for parameters related to teaching. In the full paper, we determine which of our combinatorial parameters enjoy these properties.
- We determine which of our combinatorial parameters are additive (resp. sub- or super-additive) on the free combination of two concept classes.
- We evaluate our parameters on the class 2^X of all binary concepts over the domain X.

[3] We talk here about the domain of $C_{k,n}$.

References

[1] Frank Balbach. Measuring teachability using variants of the teaching dimension. *Theoretical Computer Science*, 397(1–3):94–113, 2008.

[2] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *Journal of the Association on Computing Machinery*, 36(4):929–965, 1989.

[3] Thorsten Doliwa, Gaojian Fan, Hans U. Simon, and Sandra Zilles. Recursive teaching dimension, VC-dimension and sample compression. *Journal of Machine Learning Research*, 15:3107–3131, 2014.

[4] César Ferri, José Hernández-Orallo, and Jan Arne Telle. Non-cheating teaching revisited: New probabilistic machine teaching model. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*, pages 2973–2979, 2022.

[5] Cèsar Ferri, Dario Garigliotti, Brigt Arve Toppe Håvardstun, Josè Hernández-Orallo, and Jan Arne Telle. When redundancy matters: Machine teaching of representations, 2024. arXiv2401.12711.

[6] Ziyuan Gao, Christoph Ries, Hans U. Simon, and Sandra Zilles. Preference-based teaching. *Journal of Machine Learning Research*, 18(31):1–32, 2017.

[7] Sally A. Goldman and Michael J. Kearns. On the complexity of teaching. *Journal of Computer and System Sciences*, 50(1):20–31, 1995.

[8] David G. Kirkpatrick, Hans U. Simon, and Sandra Zilles. Optimal collusion-free teaching. In *Proceedings of ALT 2019*, pages 506–528, 2019.

[9] Farnam Mansouri, Hans U. Simon, Adish Singla, and Sandra Zilles. On batch teaching with sample complexity bounded by VCD. In *Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, 2022.

[10] Rahim Samei, Pavel Semukhin, Boting Yang, and Sandra Zilles. Algebraic methods proving Sauer's bound for teaching. *Theoretical Computer Science*, 558:35–50, 2014.

[11] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

[12] S. Shelah. A combinatorial problem: Stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:247–261, 1972.

[13] Ayumi Shinohara and Satoru Miyano. Teachability in computational learning. *New Generation Computing*, 8(4):337–348, 1991.

[14] Hans U. Simon and Jan Arne Telle. MAP- and MLE-based teaching. *Journal of Machine Learning Research*, 25(96):1–34, 2024.

[15] Hans U. Simon and Jan Arne Telle. The hierarchy of saturating matching numbers. arXiv:2503.14061, 2025.

[16] Joakim Sunde, Brigt Håvardstun, Jan Kratochvíl, and Jan Arne Telle. On a combinatorial problem arising in machine teaching. In *Proceedings of the 41st International Conference on Machine Learning*, pages 47305–47313, 2024.

[17] Vladimir N. Vapnik and Alexey Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & its Applications*, 16(2):264–280, 1971.

[18] Sandra Zilles, Steffen Lange, Robert Holte, and Martin Zinkevich. Models of cooperative teaching and learning. *Journal of Machine Learning Research*, 12:349–384, 2011.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

The decompressed tree size of k-ary chains

(Extended abstract)

Michael Wallner*

Abstract

A chain is defined as a directed acyclic graph (DAG) with one source and one sink, where the children are ordered and the spanning tree computed using a depth-first search is a path. Such DAGs emerge in the context of tree compression and are therefore uniquely associated with a tree. The tree size of a DAG is defined as the size of the associated tree. For fixed out-degree $k \geq 2$, we compute the asymptotic expected decompressed tree size of a chain of size n chosen uniformly at random, and we show that it contains a stretched exponential term of the form $e^{c\sqrt{n}}$. This result also has implications for the limit distribution of Brauer chains of fixed length.

1 Introduction

Trees are a fundamental data structure in computer science. To save memory, many effective applications store them as directed acyclic graphs (DAGs) such that repeated occurrences of the same subtrees are only stored once; see Figure 1. This process is known as *sharing* or *DAG compression* and is used in XML documents [3], binary decision diagrams [16], compilers [6], and various programming languages [1]. This gives a bijection between a tree class and a DAG class. The compression is very efficient and can be implemented in worst-case time $O(n)$ [6]. The gain in memory has been analyzed previously [3,10]: A uniformly chosen simply generated tree (e.g., a binary tree) of size n has on average a compressed size (i.e., the size of the DAG) that is asymptotically equivalent to $\frac{C' n}{\sqrt{\log n}}$, with a model-dependent constant C'.

So far, nothing was known about the reverse question: What is the average *decompressed size* of a compressed tree chosen uniformly at random? In our main result we answer this question for *chains*, which are the simplest non-trivial class of compressed trees defined in Section 3. In Theorem 3.1 we prove that the expected decompressed size of chains of size n chosen uniformly at random is asymptotically equivalent to $\frac{C\, e^{c\sqrt{n}}}{n^{1/4}}$ for explicit constants $C, c > 0$. The main difficulty is that it is much more difficult to enumerate DAGs than trees.

In [12] we considered the reverse question of enumerating all compressed[1] binary trees of compressed size n. We showed that after bounding one parameter related to the height, the associated generating functions are D-finite[2]. Later, in [8] we solved the counting problem of unconstrained

*Institute of Discrete Mathematics, TU Graz, Austria. Supported by the Austrian Science Fund (FWF) P 34142.

[1]Previously, the term *compacted tree* was used.

[2]A function is *D-finite* if it satisfies a linear differential equations with polynomial coefficients.

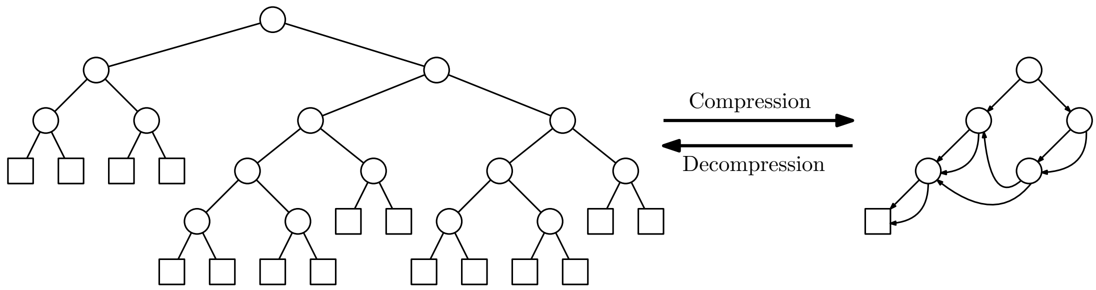


Figure 1: DAG compression: A binary tree T of size 15 on the left and its compressed representation as DAG G of compressed size 5 on the right, in which repeated occurrences of the same subtrees are factored. Thus, the *decompressed tree size* of G is $|G|_{\mathcal{T}} = 15$; see Definitions 2.1, 2.2, and 2.4

compressed binary trees using an entirely different approach on bivariate recurrence relations. We showed that their number grows for $n \to \infty$ like $\Theta\left(n!\, 4^n\, e^{3a_1 n^{1/3}}\, n^{3/4}\right)$, where $a_1 \approx -2.331$ is the largest root of the Airy function $\mathrm{Ai}(x)$ of the first kind[3]. Such a stretched exponential term did not appear in the bounded case, proving a phase transition between the bounded and unbounded model. Similar phenomena have recently also been shown in the number of relaxed k-ary trees [13], minimal deterministic finite automata accepting a finite binary language [7], Young tableaux with walls [2], and phylogenetic tree-child networks [5, 11].

2 The decompressed tree size of rooted DAGs

All trees and DAGs we consider have a distinguished root or source.

Definition 2.1 (Parameters of rooted DAGs)**.** *The* spine *of a rooted DAG G is its unique spanning tree computed in a depth-first search. The edges of the spine are referred to as* internal edges *and the other edges of G as* external edges *or* pointers*. The* fringe subgraph *of a node $v \in G$, is the induced subgraph of all nodes that can be reached from v.*

Now we are ready to define the decompression operator D that assigns to each rooted DAG a rooted tree that we call its decompressed tree.

Definition 2.2 (Decompressed tree)**.** *Let G be a rooted DAG. Its* decompressed tree $D(G)$ *is defined as follows:*

1. *Replace the sinks by leaves, i.e., nodes with out-degree* 0.
2. *Traverse the DAG along internal edges in postorder.*
3. *Iteratively, once all children of a node v have been processed, replace all pointers emanating from v by edges to copies of the fringe subtrees they are pointing to.*

Remark 2.3. Note that different DAGs $G_1 \neq G_2$ might have the same decompressed trees, i.e., $D(G_1) = D(G_2)$. However, for each tree T, there is a unique rooted DAG G with minimal number of nodes, such that $D(G) = T$. In the minimal DAG all fringe subgraphs are unique.

Definition 2.4 (Decompressed tree size)**.** *For a given size function $|\cdot| : \mathcal{T} \mapsto \mathbb{N}$ on trees $\mathcal{T}$, the* decompressed tree size $|G|_{\mathcal{T}}$ *of a DAG G is defined as*

$$|G|_{\mathcal{T}} := |D(G)|.$$

[3]The Airy function $\mathrm{Ai}(x)$ of the first kind is characterized by $\mathrm{Ai}''(x) = x\mathrm{Ai}(x)$ and $\lim_{x\to\infty} \mathrm{Ai}(x) = 0$.

The type of the decompressed tree depends on the chosen family of DAGs. Since our motivation comes from the compression of ordered trees, we study rooted ordered DAGs, in which children have a distinguished order. As the size function $|\cdot|$ we use the number of internal nodes.

Remark 2.5 (Decompressed tree statistics for DAGs). The decompressed tree size is just one example of a tree statistic that can be used to define a DAG statistic using the decompression operator. Other interesting choices are, e.g., the height or the width of the associated tree.

3 The expected decompressed size of chains

We will now introduce the family of DAGs that we consider: k-ary chains. Let k be a positive integer. We define *k-ary DAGs* as rooted ordered DAGs in which each node has exactly k outgoing edges. The corresponding decompressed trees are classical rooted ordered k-ary trees, enumerated by the Fuss–Catalan numbers $\frac{1}{(k-1)n+1}\binom{kn}{n}$. The most famous subclass are binary trees enumerated by the ubiquitous Catalan numbers $\frac{1}{n+1}\binom{2n}{n}$; see, e.g., [17] for links to hundreds of combinatorial objects.

In addition to the out-degrees, one may also fix the allowed shapes. We consider fixed shapes of the corresponding spines. Such families were (asymptotically) enumerated in [12] for binary DAGs with spines of fixed right height. The simplest non-trivial example are *chains* that are DAGs conditioned to have spines of right height 0, i.e., the spine is a sequence of nodes; see Figure 2. All fringe subgraphs of these DAGs are unique and thus all its decompressed trees are distinct.

Figure 2: Left: A binary chain (i.e., $k=2$) with 6 internal nodes. The unique root is on the very right and the unique sink on the very left. Right: The shape of ternary chains.

Let us briefly describe, why this class is interesting. First, since pointers can point to any node to the left, the number of k-ary chains with n internal nodes is $(n!)^{k-1}$. Hence, the cardinality of this subclass grows super-exponentially in the number n of internal nodes. Second, the study of this subclass was the key in all previous papers to derive asymptotic results. Third, among all compressed trees of size n, this class includes the smallest and the largest possible decompressed tree: If all pointers end in the sink, the decompressed tree has size n, while if all pointers connect with the node just before in postorder the decompressed tree has size $k^n - 1$.

Formally, let $k \geq 2$ be an integer and let $\mathcal{C}$ be the set of k-ary chains. We denote by $\mathcal{C}_n = \{C \in \mathcal{C} \ : \ |C| = n\}$ the set of k-ary chains conditioned to have n internal nodes and by $c_n := |\mathcal{C}_n| = (n!)^{k-1}$ its cardinality. Let X_n be the random variable associated with the decompressed size of chains $\mathcal{C}_n$ chosen uniformly at random, i.e.,

$$\mathbb{P}(X_n = k) = \frac{|\{C \in \mathcal{C}_n \ : \ |C|_{\mathcal{T}} = k\}|}{c_n}.$$

We are interested in the distribution of X_n. Our main result is the following asymptotics of $\mathbb{E}(X_n)$, which includes a stretched exponential. For comparison, the minimal size is $n = e^{\log(n)}$ and the maximal size is $k^n - 1 \approx e^{n\log(k)}$.

Theorem 3.1. *The expected decompressed tree size of k-ary chains with n internal nodes is for $n \to \infty$ asymptotically given by*

$$\mathbb{E}(X_n) = \frac{1}{2\sqrt{e^{k-1}\pi}\,(k-1)^{5/4}} \frac{e^{2\sqrt{(k-1)n}}}{n^{1/4}} \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right).$$

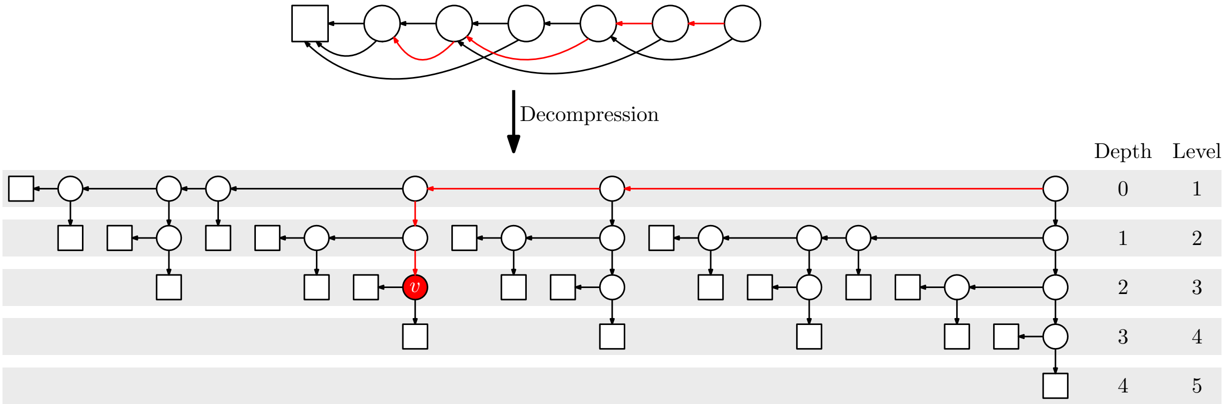


Figure 3: A binary chain C (top) and its decompressed tree $D(C)$ (bottom). The compression depths and levels from Definition 3.2 are stated on the right. The red path P_v in C is associated with the node $v \in D(C)$.

Proof (Sketch). The key idea is to relate the nodes in the decompressed trees with a statistic in the corresponding DAGs; see Figure 3. It can be shown that each node $v \in D(G)$ is in bijection with a path P_v in G from the root to the compressed instance of v. Note that in G several paths might end in the same node, but different paths correspond to different nodes in the decompressed tree $D(G)$.

Now, we use these paths to define the compression depth and the compression level of a node using the concept of pointers in DAGs from Definition 2.1.

Definition 3.2 (Compression depth and level)**.** *Let G be a rooted DAG. The* compression depth *of a node $v \in D(G)$ is defined as the number of pointers in P_v. The* compression level *of a node v is defined as its compression depth plus one; see Figure 3.*

Our strategy is to count the number of decompressed nodes per compression level in all chains of size n using a generalized exponential generating function. For this purpose, let $d_{\ell,n}$ be the number of nodes on compression level $\ell \geq 1$ in $D(\mathcal{C}_n)$, i.e., all decompressed k-ary chains of size n. We will use the follwoing specially rescaled generating function to study these numbers.

$$D_\ell(z) = \sum_{n\geq 0} d_{\ell,n} \frac{z^n}{(n!)^{k-1}}.$$

The main reason for the rescaling by $(n!)^{k-1}$ is that it allows us to extend the symbolic calculus on exponential generating functions introduced in [12] to k-ary DAGs. In particular, we can generalize [12, Lemmas 6.1 and 6.3] and prove that the operation $zF(z)$ corresponds to the symbolic construction of appending a new root node and the pointing operation $zF'(z)$ to adding a new pointer to the root. This allows us to derive a recursion for $D_\ell(z)$ and to prove the following result:

Theorem 3.3. *The bivariate generating function $D(z,q) = \sum_{\ell\geq 1} D_\ell(z) q^\ell$ is given by*

$$D(z,q) = \frac{1}{(k-1)(1-z)} \left(e^{\frac{(k-1)qz}{1-z}} - 1 \right).$$

Now, observe that the specific rescaling by $(n!)^{k-1}$ has also the advantage that $\mathbb{E}(X_n) = [z^n]D(z,1)$, since $c_n = (n!)^{k-1}$. Finally, the asymptotics of the coefficients of $D(z,1)$ can be deduced by the saddle-point method [9, Chapter VIII]. For this purpose, we use Cauchy's integral formula, which expresses the coefficients of $D(z,1)$ in terms of a complex contour integral. Then, we transform this contour such that the main contribution can be approximated by a Gaussian integral. $\square$

Remark 3.4 (Laguerre polynomials). In the binary case ($k = 2$) Theorem 3.3 directly implies the closed form $d_{\ell,n} = \frac{n!}{\ell!}\binom{n}{\ell}$. These are the coefficients of Laguerre polynomials; see A000142, A001563, A001809, A001810, A001811, A001812, for $\ell = 1, \dots, 6$ in the OEIS. The coefficients of the exponential generating function $D(z, 1)$ correspond to A070779, given by strictly partial permutations of $\{1, \dots, n\}$.

Remark 3.5 (Increasing subsequences in permutations). Let $s_{\ell,n}$ be the number of increasing subsequences of length ℓ in all permutations of n. First, choose ℓ out of n positions. Then choose ℓ out of n numbers $\{1, \dots, n\}$ and fill the chosen positions with these in increasing order. Finally, fill the remaining $n - \ell$ positions, with an arbitrary permutation of the remaining $n - \ell$ elements. Therefore, we have $s_{\ell,n} = (n-\ell)!\binom{n}{\ell}^2 = \frac{n!}{\ell!}\binom{n}{\ell} = d_{\ell,n}$. We leave it as an open problem to find a bijection between these two quantities.

4 Applications to Brauer chains

Let m, n be positive integers. A *Brauer chain*[4] *of length n for m* is a sequence of numbers $(a_0, a_1, \dots, a_n)$ such that $a_0 = 1$, $a_n = m$, and for $i > 0$ it holds that $a_i = a_{i-1} + a_j$, where $j < i$; see [4, 14]. Such addition chains play an important role in the efficient evaluation of powers x^m, in the sense of minimizing the number of multiplications needed [15, Section 4.6.3].

Proposition 4.1. *Brauer chains of length n are in bijection with binary chains with n internal nodes. In particular, the value a_n is equal to the number of leaves in the decompressed binary tree.*

Our main result Theorem 3.1 in the case of binary trees admits then the following interpretation: the expected value of a_n for an addition chain of length n chosen uniformly at random is equal to

$$\frac{1}{2\sqrt{e\pi}} \frac{e^{2\sqrt{n}}}{n^{1/4}} \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right).$$

Note that also our most general result for k-ary chains can be interpreted in terms of addition chains, in which addition is a k-ary operation.

5 Conclusion and further research directions

The motivation of this work was to show that the progress on the enumerative questions of the last years makes now statistics of the decompressed size distribution of DAGs accessible. The analysis of chains should be seen as the first step, towards understanding the decompressed size distribution of arbitrary compressed trees. A next step is to study the expected decompressed size of other DAG classes, such as compressed trees of right height less than 1, 2, etc. Another interesting question is to further analyze the limit distribution of decompressed k-ary chains. We have derived the asymptotic mean, however, we have no information on higher moments, like the variance. Finally, now that the counting problem is solved, we can analyze parameters, such as the typical shape of large decompressed trees. For example, let L_n be the random variable associated with the level of a random node in all decompressed binary chains $D(\mathcal{C}_n)$ with n internal nodes:

$$\mathbb{P}(L_n = k) = \frac{[z^n q^k]D(z, q)}{[z^n]D(z, 1)}.$$

Then, by standard methods of analytic combinatorics [9] one can show that L_n satisfies a Gaussian limit law with mean $\mathbb{E}(L_n) = \sqrt{n} + \mathcal{O}(1)$ and variance $\mathbb{V}(L_n) = \frac{\sqrt{n}}{2} + \mathcal{O}(1)$.

[4] In the literature of addition chains it is custom to use n instead of m. However, in this paper, n is always used for the size of the compressed tree, which corresponds to the length of the addition chain.

References

[1] Alfred V. Aho, Ravi Sethi, and Jeffrey D. Ullman. *Compilers, Principles, Techniques.* Addison-Wesley, Boston, 1986.

[2] Cyril Banderier and Michael Wallner. Young Tableaux with Periodic Walls: Counting with the Density Method. *Sém. Lothar. Combin.*, 85B, Article 85B.47:12 pp., 2021.

[3] Mireille Bousquet-Mélou, Markus Lohrey, Sebastian Maneth, and Eric Noeth. XML compression via directed acyclic graphs. *Theory of Computing Systems*, 57(4):1322–1371, 2015.

[4] Alfred Brauer. On addition chains. *Bull. Amer. Math. Soc.*, 45:736–739, 1939.

[5] Yu-Sheng Chang, Michael Fuchs, Hexuan Liu, Michael Wallner, and Guan-Ru Yu. Enumerative and distributional results for d-combining tree-child networks. *Adv. in Appl. Math.*, 157:Paper No. 102704, 58, 2024.

[6] Peter J. Downey, Ravi Sethi, and Robert Endre Tarjan. Variations on the common subexpression problem. *J. Assoc. Comput. Mach.*, 27(4):758–771, 1980.

[7] Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Asymptotics of Minimal Deterministic Finite Automata Recognizing a Finite Binary Language. In *AofA 2020*, volume 159 of *LIPIcs*, pages 11:1–11:13. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020.

[8] Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Compacted binary trees admit a stretched exponential. *J. Combin. Theory Ser. A*, 177:Paper No. 105306, 40, 2021.

[9] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[10] Philippe Flajolet, Paolo Sipala, and Jean-Marc Steyaert. Analytic variations on the common subexpression problem. In *Automata, languages and programming*, pages 220–234. Springer, 1990.

[11] Michael Fuchs, Guan-Ru Yu, and Louxin Zhang. On the asymptotic growth of the number of tree-child networks. *European J. Combin.*, 93:Paper No. 103278, 2021.

[12] Antoine Genitrini, Bernhard Gittenberger, Manuel Kauers, and Michael Wallner. Asymptotic enumeration of compacted binary trees of bounded right height. *J. Combin. Theory Ser. A*, 172:105177, 2020.

[13] Manosij Ghosh Dastidar and Michael Wallner. Asymptotics of Relaxed k-Ary Trees. In *AofA 2024*, LIPIcs, pages 15:1–15:13. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[14] Richard K. Guy. *Unsolved problems in number theory.* Problem Books in Mathematics. Springer-Verlag, New York, third edition, 2004.

[15] Donald E. Knuth. *The Art of Computer Programming. Vol. 2.* Addison-Wesley, Reading, MA, 1998. Seminumerical algorithms, Third edition.

[16] Christoph Meinel and Thorsten Theobald. *Algorithms and Data Structures in VLSI Design.* Springer Berlin Heidelberg, 1998.

[17] Richard P. Stanley. *Catalan Numbers.* Cambridge University Press, 2015.

Proceedings of the 13th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'25
Budapest, August 25 - 29, 2025

Arbitrary orientations of cycles in oriented graphs

(Extended abstract)

Guanghui Wang* Yun Wang† Zhiwei Zhang‡

Abstract

We show that every sufficiently large oriented graph G with both minimum indegree and outdegree at least $(3|V(G)| - 1)/8$ contains a cycle of every possible orientation and of every possible length unless G is isomorphic to an exceptional oriented graph. In particular, the oriented graph G contains a Hamilton cycle of every possible orientation. The degree condition is best possible and the result solves a problem of Debiasio, Kühn, Molla, Osthus and Taylor from 2015.

1 Introduction

The *minimum semidegree* $\delta^0(G)$ of a digraph G is the minimum of all the indegrees and outdegrees of the vertices in G. For a path or cycle in digraphs, we say that it is *directed* if all its edges are oriented in the same direction and, it is *antidirected* if it contains no directed 2-path. Throughout the paper, we will write n for the number of vertices of the graph or digraph G we are considering.

A fundamental result of Dirac [3] states that every graph G with minimum degree $n/2$ contains a Hamilton cycle. A directed version of this result was obtained by Ghouila-Houri [4], who showed that any digraph G with minimum semidegree $\delta^0(G) \geq n/2$ contains a directed Hamilton cycle. In [7] Keevash, Kühn and Osthus proved that the minimum semidegree threshold for an oriented graph G containing a directed Hamilton cycle turns out to be $(3n-4)/8$.

Instead of asking for a directed Hamilton cycle in digraphs, one may ask under certain conditions, if it is possible to show a digraph containing every possible orientation of a Hamilton cycle. For tournaments, Rosenfeld [10] conjectured that the directed Hamilton cycle is the only orientation of a Hamilton cycle that can be avoided by tournaments on arbitrarily many vertices. This conjecture was settled by Thomason [12] who proved that every tournament on $n \geq 2^{128}$ vertices contains every oriented n-cycle except possibly the directed Hamilton cycle. In [6] Havet showed that the number of vertices above can be lowered to 68. The question for digraphs was answered by Häggkvist and Thomason [5] in 1995, who proved that a minimum semidegree $\delta^0(G) \geq n/2 + n^{5/6}$ ensures the

*School of Mathematics, Shandong University, Jinan 250100, China. E-mail: ghwang@sdu.edu.cn. Supported by National Natural Science Foundation of China (12231018).
†Data Science Institute, Shandong University, Jinan 250100, China. E-mail: yunwang@sdu.edu.cn.
‡School of Mathematics, Shandong University, Jinan 250100, China. E-mail: zhiweizh@mail.sdu.edu.cn.

existence of every possible orientation of a Hamilton cycle in digraphs. In [2] Debiasio and Molla got the exact minimum semidegree threshold for an antidirected Hamilton cycle in digraphs, they showed that $\delta^0(G) \geq n/2+1$ will suffice. Later, Debiasio, Kühn, Molla, Osthus and Taylor [1] settled the problem by completely determining the exact threshold for every possible orientation of a Hamilton cycle in digraphs. They showed that $\delta^0(G) \geq n/2$ suffices if the orientation is not antidirected. For oriented graphs, Häggkvist and Thomason [5] proved that for all $\varepsilon > 0$ and all sufficiently large oriented graphs G a minimum semidegree $\delta^0(G) \geq (5/12+\varepsilon)n$ is sufficient to guarantee every possible orientation of a Hamilton cycle. They conjectured that the semidegree threshold should be $(3/8+\varepsilon)n$. This conjecture was confirmed by Kelly in [8]. Debiasio, Kühn, Molla, Osthus and Taylor [1] believe that it would be interesting to obtain an exact version of this result. In this paper, we solve this problem by proving the following.

Theorem 1.1. *There exists an integer n_0 such that every oriented graph G on $n \geq n_0$ vertices with minimum semidegree $\delta^0(G) \geq (3n-1)/8$ contains every possible orientation of a Hamilton cycle.*

Now we show that the above result is actually the best bound. For oriented graphs, recall that a minimum semidegree $\delta^0(G) \geq (3n-4)/8$ ensures the existence of a directed Hamilton cycle [7]. Not surprisingly, similar to the case in digraphs, the semidegree for the existence of antidirected Hamilton cycles slightly larger than the degree for directed Hamilton cycles. In fact, the following proposition shows that if the given Hamilton cycle has a long antidirected subpath, then the degree bound in Theorem 1.1 is the best possible.

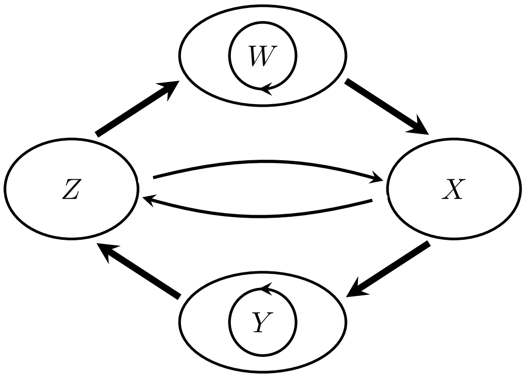


Figure 1: The oriented graph G in Proposition 1.2. The bold edges indicate that all possible edges are present and have the directed shown. The set W has size $\lfloor n/4 \rfloor$ and $|X| = |Z| = \lceil n/4 \rceil$. Each of W and Y spans an almost regular tournament, that is, the indegree and outdegree of every vertex differ by at most one. The oriented graph induced by X and Z is an almost regular bipartite tournament.

Proposition 1.2. *For any $n \geq 3$, there is an oriented graph G on n vertices with minimum semidegree $\delta^0(G) = \lceil (3n-1)/8 \rceil - 1$ which does not contain a Hamilton cycle having an antidirected subpath of order larger than $(3n+2)/4$. In particular, it has no antidirected Hamilton cycle.*

It seems to be an interesting problem to determine the exact semidegree threshold for a Hamilton cycle without antidirected subpaths of order larger than $(3n+2)/4$. The authors believe that $\delta^0(G) \geq (3n-4)/8$ will suffice.

In this paper we extend Theorem 1.1 further to a pancyclicity result for arbitrary orientations. Before state it we introduce some results on this topic. In [9], Kelly, Kühn and Osthus showed that in oriented graph the minimum semidegree condition $\delta^0(G) \geq (3n-4)/8$ gives not only a directed Hamilton cycle but also a directed cycle of every possible length.

Theorem 1.3. *[9] There exists an integer n_0 such that every oriented graph G on $n \geq n_0$ vertices with $\delta^0(G) \geq (3n-4)/8$ contains a directed t-cycle for all $3 \leq t \leq n$.*

Later, Kelly [8] proved that the above result with the addition of an error term in the degree condition can be extended to arbitrary orientations of cycles.

Theorem 1.4. *[8] Let $\alpha > 0$. Then there exists $n_0 = n_0(\alpha)$ such that if G is an oriented graph on $n \geq n_0$ vertices with minimum semidegree $\delta^0(G) \geq (3/8 + \alpha)n$ then G contains a cycle of every possible orientation and of every possible length.*

In this paper, we show that the minimum semidegree in Theorem 1.1 for arbitrary orientations of a Hamilton cycle implies pancyclicity unless the oriented graph is isomorphic to an exception.

Theorem 1.5. *There exists an integer n_0 such that if G is an oriented graph on $n \geq n_0$ vertices with minimum semidegree $\delta^0(G) \geq (3n-1)/8$, then one of the following holds.*

(i) G contains a cycle of every orientation and of every possible length.

(ii) Let G^* be the oriented graph depicted in Figure 1 which has order $8s+2$ for some s. Then G can be obtained from G^* by adding an extra vertex v and adding edges between v and G^* so that G satisfying the condition $\delta^0(G) = (3n-1)/8$.

In particular, if $\delta^0(G) \geq 3n/8$, then G contains a cycle of every orientation and of every possible length.

2 Proof sketch of Theorem 1.1

We first observe that any oriented graph satisfying the degree condition $\delta^0(G) \geq (3n-1)/8$ must be a robust outexpander or it is 'close to' the oriented graph in Figure 1. Roughly speaking, a digraph is a robust outexpander if every vertex set S of reasonable size has an outneighborhood which is at least a little large than S itself. Taylor [11] proved that every sufficiently large robust outexpander of linear minimum semidegree contains every possible orientation of a Hamilton cycle. This allows us to restrict our attention to the 'close to' case. To formally say what we mean by 'close to' the extremal oriented graph we need the definition of ε-extremal, where an oriented graph G on n vertices is ε-extremal if there exists a partition W, X, Y, Z of $V(G)$ such that $|W|, |X|, |Y|, |Z| = n/4 \pm O(\varepsilon n)$ and $e(W \cup Z, Y \cup Z) < \varepsilon n^2$.

We next divide the proof of Theorem 1.1 into two cases depending on the number of sink vertices of the cycle C, where sink and source vertices are vertices with indegree 2 and outdegree 2 in C, respectively.

2.1 C has few sink vertices

In this case the cycle C is close to being directed, which means that there are many directed subpaths of C can be used to balance the sizes of W, X, Y, Z and to cover 'bad' vertices. Here, bad vertices are the vertices with low indegree or outdegree in the vertex classes where we would expect most of their neighbors to lie.

First we will pick few short subpaths $P_1, \dots, P_l$ of the given cycle C which together contain all sink and source vertices of C and $dist_C(P_i, P_{i+1}) \equiv 0 \pmod 4$ for all i. Moreover, there are sufficiently many directed 12-paths among those l paths. After this we embed several directed 12-paths P_i to balance the number of remaining vertices in X, Z and cover all bad vertices of G, which is the main difficulty in this case. Then we move on to greedily embed the paths containing sink or source vertices in $G[W]$. Note that this is possible as the number of vertices on those paths are quite small compared to the size of W.

Finally, removing all inner vertices of the paths embedded above, say (W', X', Y', Z') is obtained from (W, X, Y, Z) by removing operations, we have $|W'| - l = |X'| = |Y'| = |Z'|$ and then we embed the left subpaths of C by applying the Blow-up Lemma. It should be noted that since all sink and source vertices are contained in P_i and $dist_C(P_i, P_{i+1}) \equiv 0 \pmod 4$, all left subpaths of C are directed and each of them has length 0 mod 4. Further, in the above process we can embed all paths P_i such that their endvertices have large indegree and outdegree in W. Then the neighbors of these endvertices in $X' \cup Z'$ can be embedded properly when we apply the Blow-up Lemma as the number of embedded paths is quite small.

2.2 C has many sink vertices

Now we discuss the case that C has at least linear sink vertices. Recall that there exists a partition (W, X, Y, Z) of $V(G)$ such that $e(W\cup Z, Y\cup Z) < \varepsilon n^2$. Then the degree condition $\delta^0(G) \geq (3n-1)/8$ shows that almost all vertices of Z have many inneighbors and outneighbors in X. However, possibly every vertex of X has either zero inneighbor or zero outneighbor in Z. For example, let X_1, X_2 be a partition of X with $|X_1| = |X_2|$ such that ZX_1X_2 is a blow-up of the directed 3-cycle. Then each vertex of X_1 has outdegree 0 in Z and each vertex of X_2 has indegree 0 in Z. To overcome this main difficulty, we first cover most of the vertices in $X \cup Z$ by using randomised embedding method. During the above embedding process, the sets W and Y are not used too many vertices so that the remaining $G[W]$ and $G[Y]$ still have large degrees. Then we use short subpaths of C containing many sink vertices to cover bad vertices and the remaining vertices in $X \cup Z$. After this, the semidegree of the remaining $G[W]$ (resp., $G[Y]$) is still large enough so that one may find a Hamilton cycle of any orientation in the remaining $G[W]$ (resp., $G[Y]$) by Theorem 1.4.

Now we briefly mention how to overcome the main difficulty, i.e., how to cover most of the vertices in $X \cup Z$. Figure 2 illustrates this procedure. For the further embedding and connecting, we first store some vertices and paths with nice properties in $\mathcal{RC}$ and $\mathcal{SE}$. Deleting all bad vertices in $\mathcal{B}$, we apply the Diregularity Lemma to the remaining $G[X \cup Z]$. Since the reduced graph inherits the minimum semidegree of $G[X \cup Z]$, there is an almost perfect matching between the clusters in X and in Z, say $Z_1X_1, \ldots, Z_{k_1}X_{k_1}$ and $X_{k_1+1}Z_{k_1+1}, \ldots, X_kZ_k$ are matching from Z to X and from X to Z, respectively. Then we randomly choose k sets $W_1, \ldots, W_k$ and $Y_1, \ldots, Y_k$ of W and Y, respectively, such that the sets W_i, X_i, Y_i, Z_i with $i \in [k]$ have the same size.

The second step is to partition the given cycle C into short paths $P_1, Q_1, \ldots, P_t, Q_t, \ldots$ such that the paths has a specific property. Based on the directions of Q_{i-1} and Q_i, the paths $P_1, \ldots, P_t$ can be divided into three types. Roughly speaking, type (**I**) consists of short paths which we can not describe its exact orientation; type (**II**) consists of short paths which is antidirected and type (**III**) consists of paths with constant length which can be embedded into $G[W]$ greedily. Inspired by the techniques used in [8] and [11], the paths of type (**I**) will randomly embedded around a directed 3-cycle $Z_iX_iY_i$ or $X_iZ_iW_i$ and the paths of type (**II**) will randomly embedded around an edge Z_iX_i, Z_iW_i or W_iX_i. Here, 'a path P is randomly embedded around F' means that we start from the specified initial vertex $V_i \in V(F)$ and assign the next vertex of P to either the successor or the predecessor of V_i in F according to the orientation of the edge. It is worthwhile mentioning that the way of partition the cycle C into subpaths $P_1, Q_1, P_2, Q_2, \ldots, P_t, Q_t, \ldots$ is totally different from the way used in [8] and [11]. Moreover, in this paper, the short paths not only randomly embedded around a directed 3-cycle but also around an edge.

An additional difficulty is that it is impossible to embed a long antidirected path in the oriented graph depicted in Figure 1. Indeed, let $P = v_1v_2 \cdots v_l$ be an antidirected path with $v_1v_2 \in E(P)$. If v_1 is embedded in $W \cup Z$, then v_2 must belong to $W \cup X$ as $v_1v_2 \in E(P)$ and then $v_3 \in W \cup Z$ as $v_3v_2 \in E(P)$. Continue this procedure until all vertices are embedded, we get that all vertices with odd indices must belong to $W \cup Z$ and vertices with even indices are in $W \cup X$. The case that v_1 is embedded in $X \cup Y$ can be verified similarly. This shows that every antidirected path of G only uses the vertices in $W \cup X \cup Z$ or the vertices in $Y \cup X \cup Z$ and thus it is impossible to embed an antidirected path of length more than $3n/4$. In the paper, special edges are useful to overcome this difficulty. Here, an edge e is *special* for the partition (W, X, Y, Z) if it belongs to $E(W \cup Z, Y \cup Z) \cup E(X \cup Y, W \cup X)$, i.e., the edges that do not shown in Figure 1. It is not difficult to check that if there is a special edge from Y to X, then one may embed a long antidirected path in $G[W \cup X \cup Z]$ first and then the special edge can help us to embed the remaining part of the path in $G[Y]$. In the paper we observe that the degree condition $\delta^0(G) \geq (3n-1)/8$ ensures two special edges unless G is isomorphic to an exceptional oriented graph.

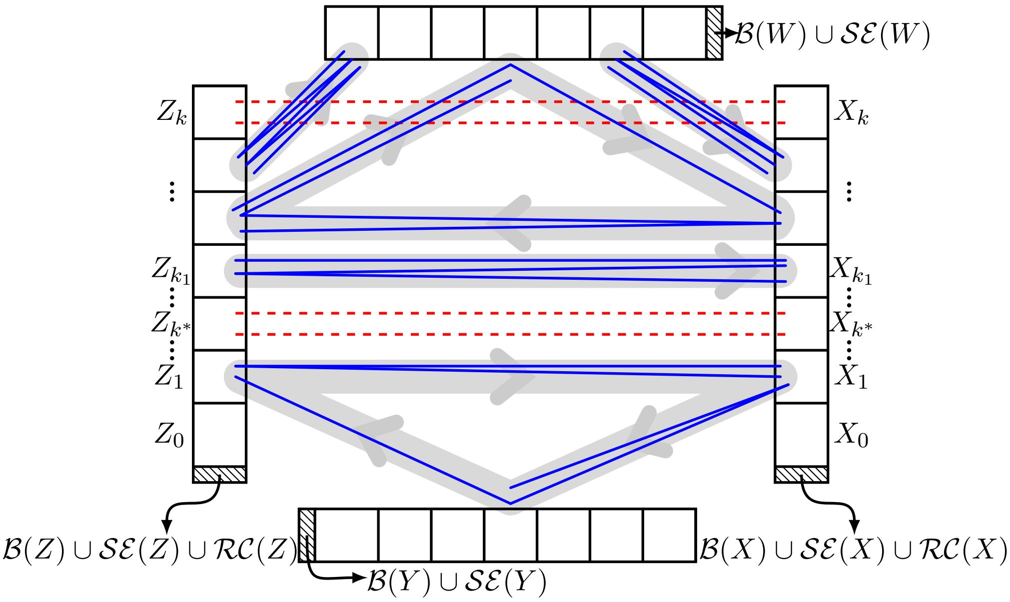


Figure 2: An illustration of how the Diregularity Lemma and the randomised embedding method may be applied.

We now end this section by showing the differences between the two cases. When C has few sink vertices, the cycle C has many directed subpaths of length 12 which can be used to cover bad vertices. However, if C has many sink vertices, then it is no longer guaranteed that C has many directed or antidirected paths. To overcome this, we will use both sink vertices and short directed paths to cover bad vertices. Another difference is that when C has few sink vertices we need to balance X and Z first, but we do not proceed this when C has many sink vertices. This is because when C has few sink vertices the cycle is close to being directed and every directed path uses the same number of vertices in X and Z. However, for the case that C has many sink vertices, we have enough sink and source vertices to cover the remaining vertices in X and Z and thus we do not need to balance X and Z at first.

3 Some ideas in the proof of Theorem 1.5

Similar with the proof of Theorem 1.1, we prove Theorem 1.5 by using the stability technique where we split the proof into two cases depending on whether G is ε-extremal to the graph depicted in Figure 1. When G is ε-extremal there is a partition (W, X, Y, Z) of $V(G)$ with very nice properties. So if $|C| < n/5$, then it seems to be not difficult to embed C into $G[W]$ as $G[W]$ is close to a regular tournament of order $|W| \sim n/4$. For the case $|C| \geq n/5$, the proof of Theorem 1.1 holds for the case that C is not Hamilton as well. If the given oriented graph G has expansion property, by applying several simple results on expansion we will show that a minimum semidegree $\delta^0(G) \geq 0.346n$ is sufficient to guarantee the cycle C.

References

[1] L. DeBiasio, D. Kühn, T. Molla, D. Osthus, and A. Taylor. Arbitrary orientations of Hamilton cycles in digraphs. *SIAM J. Discrete Math.*, 29(3):1553–1584, 2015.

[2] L. Debiasio and T. Molla. Semi-degree threshold for anti-directed Hamiltonian cycles. *Elec. J. Combin.*, 22(4), 2015.

[3] G. A. Dirac. Some theorems on abstract graphs. *Proc. London Math. Soc.*, 2(3):69–81, 1952.

[4] A. Ghouila-Houri. Une condition suffisante d'existence d'un circuit hamiltonien. *C.R. Acad. Sci. Paris*, 25:495–497, 1960.

[5] R. Häggkvist and A. Thomason. *Oriented Hamilton cycles in oriented graphs*, pages 339–354. Combinatorics, Geometry and Probability. Cambridge University Press, United Kingdom, 1997.

[6] F. Havet. Oriented Hamiltonian cycles in tournaments. *J. Combin. Theory Ser. B*, 80:1–31, 2000.

[7] P. Keevash, D. Kühn, and D. Osthus. An exact minimum degree condition for Hamilton cycles in oriented graphs. *J. London Math. Soc.*, 79(1):144–166, 2009.

[8] L. Kelly. Arbitrary orientations of Hamilton cycles in oriented graphs. *Elec. J. Combin.*, 18(1):1–25, 2011.

[9] L. Kelly, D. Kühn, and D. Osthus. Cycles of given length in oriented graphs. *J. Combin. Theory Ser. B*, 100(3):251–264, 2010.

[10] M. Rosenfeld. Antidirected Hamiltonian cycles in tournaments. *J. Combin. Theory Ser. B*, 16:234–242,1974.

[11] A. Taylor. The regularity method for graphs and digraphs. MSci thesis, School of Mathematics, University of Birmingham, UK, 2013; also available from arXiv:1406.6531.

[12] A. Thomason. Paths and cycles in tournaments. *Trans. Amer. Math. Soc.*, 296(1):167–180, 1986.