Mathematisches Forschungsinstitut Oberwolfach

Report No. 42/2025

DOI: 10.4171/OWR/2025/42

Combinatorics, Probability and Computing

Organized by
Robert Morris, Rio de Janeiro
Oliver Riordan, Oxford
Lisa Sauermann, Bonn

14 September – 19 September 2025

Abstract. The main theme of the workshop was the use of probabilistic methods in combinatorics and related fields. This area is evolving extremely quickly, with the introduction of powerful new methods and the development of increasingly sophisticated techniques, and there have been a number of very significant breakthroughs in the area in recent years. The workshop emphasized several of these recent breakthroughs, which include foundational results in the theory of random graphs and processes, and also applications of probabilistic techniques in Ramsey theory, design theory, and group theory, and of combinatorial techniques to problems in number theory, functional analysis, and high-dimensional geometry.

Mathematics Subject Classification (2020): 05-XX, 60-XX.

Introduction by the Organizers

The meeting was very well attended with 48 participants from around the globe, including from Germany, Switzerland, Austria, Belgium, Croatia, Sweden, Israel, the UK, the US, Canada, Brazil, Australia and New Zealand. In addition, many excellent mathematicians who would have loved to participate could not be invited due to space constraints. The program consisted of 8 main lectures, 18 shorter talks, and a problem session, as well as plenty of time for discussion.

A number of major recent breakthroughs were presented during the workshop, including the proof of a famous conjecture of Erdős about induced Ramsey numbers, which was announced by Marcelo Campos in his talk. More precisely, Campos (joint work with Lucas Aragão, Gabriel Dahia, Rafael Filipe and João Pedro

Marciano) showed that there exists $n = r^{O(rk)}$ such that, with high probability, every r-colouring of the edges of the random graph $G(n, 1/2)$ contains a monochromatic induced copy of every graph with k vertices. A key tool in the proof was a new kind of container lemma, which was discovered recently by Campos and Wojciech Samotij. Julian Sahasrabudhe presented another recent breakthrough in Ramsey theory (joint work with Marcelo Campos, Matthew Jenssen and Marcus Michelen): a randomized construction that improved the best-known lower bound on the off-diagonal Ramsey numbers $R(3, k)$.

Another highlight of the workshop was the talk of Richard Montgomery, who announced the resolution (joint work with Natalie Behague and Daniel Il'kovič) of the famous Kim–Vu sandwich conjecture, which states that if $d \gg \log n$, then the random d-regular graph $G(d)$ can be 'sandwiched' between two Erdős–Rényi random graphs $G(n, p)$ and $G(n, q)$ with $p, q = \big(1 + o(1)\big)d/n$, in the following sense: there exists a coupling (G_1, H, G_2) such that

$$G_1 \sim G(n,p), \quad H \sim G(d), \quad G_2 \sim G(n,q) \quad \text{and} \quad \mathbb{P}\big(G_1 \subset H \subset G_2\big) = 1 - o(1).$$

Two other beautiful talks about random graphs were given by Julia Böttcher (joint work with Peter Allen, Yoshiharu Kohayakawa and Mihir Neve), on applications of the sparse blow-up lemma to extremal problems about bounded-degree spanning subgraphs of random graphs, and by Nina Kamčev (joint work with Nicolas Broutin, Gábor Lugosi, Bruce Reed, and Liana Yepremyan) on ways of detecting planted trees in sparse random graphs, focusing in particular on the case of a uniformly-chosen random tree with $k = k(n)$ vertices. Another important breakthrough was presented by Marcus Kühn, who described his recent resolution (joint work with Felix Joos) of a notorious conjecture about the final number of edges in the hypergraph removal process. In this fundamental random process, uniformly-random copies of a k-uniform hypergraph are removed one by one from the complete hypergraph $K_n^{(k)}$ until no copies are left.

Several talks focused on design theory, including those of Peter Keevash, who presented a new proof of the existence of designs, and Noga Alon, who presented solutions to two problems of Erdős about partial designs. Furthermore, the talk of Michelle Delcourt discussed her recently introduced method of refined absorption (joint with Luke Postle), as well as several of its applications (joint work with Cicley Henderson, Thomas Lesgourges, Tom Kelly and Luke Postle), such as a construction of designs with high girth (joint with Luke Postle).

There were also several talks on additive combinatorics, including Huy Tuan Pham's stunning description of his recent work with Noga Alon on the independence number of sparse random Cayley graphs, and the number of sets with small sumset. The key step in their proof is an 'efficient covering' lemma for sets with small sumset, which also resolves a conjecture of Shachar Lovett. Marcus Michelen gave a beautiful talk on his work with Oren Yakir, in which they study a complex analogue of the classical Littlewood-Offord problem, and deduce results about the separation of the roots of random polynomials, and Mehtaab Sawhney described his extremely technically impressive proof (joint with Michael Jaber, Yang Liu, Shachar Lovett and Anthony Ostuni) of a quasipolynomial bound for the corners theorem, which is a two-dimensional generalization of Roth's theorem on 3-term arithmetic progressions.

Finally, there were a number of talks on applications of probabilistic and combinatorial techniques to resolve problems from other areas of mathematics. In particular, Marius Tiba presented his work with Alessio Figalli and Peter van Hintum on the stability of the Brunn–Minkowski and Prékopa–Leindler inequalities, and Liana Yepremyan (joint with Matija Bucić, Bryce Frederickson, Alp Müyesser and Alexey Pokrovskiy) and Matija Bucić (joint work with Benjamin Bedert, Noah Kravitz, Richard Montgomery and Alp Müyesser) both presented progress towards Graham's rearrangement conjecture in general groups (originally, Graham posed this conjecture for $\mathbb{Z}_p$, but it naturally generalizes to all groups).

As always, and on behalf of all participants, the organizers would like to thank the staff and the director of the Mathematisches Forschungsinstitut Oberwolfach for providing such a stimulating and inspiring atmosphere.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-2230648, "US Junior Oberwolfach Fellows".

Workshop: Combinatorics, Probability and Computing

Table of Contents

Noga Alon
On designs and partial designs 2249

Maria Axenovich (joint with Noga Alon, John Goldwasser)
Hypercube statistics 2251

Julia Böttcher (joint with Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, Mihir Neve)
Robustness and resilience for spanning graph of bounded degree 2253

Domagoj Bradač (joint with Zach Hunter and Benny Sudakov)
Lower bounds for Ramsey numbers of bounded degree hypergraphs 2255

Matija Bucić (joint with Benjamin Bedert, Noah Kravitz, Richard Montgomery, Alp Müyesser)
On Graham's rearrangement conjecture over $\mathbb{F}_2^n$ 2257

Marcelo Campos (joint with Lucas Aragão, Gabriel Dahia, Rafael Filipe, João Marciano)
Induced Ramsey Numbers 2260

Michelle Delcourt (joint with Cicley Henderson, Thomas Lesgourges, Tom Kelly, Luke Postle)
Applications of Refined Absorption 2262

Oliver Janzer (joint with Julien Portier)
Nearly tight bounds for MaxCut in hypergraphs 2265

Nina Kamčev (joint with Nicolas Broutin, Gábor Lugosi, Bruce Reed, and Liana Yepremyan)
Detecting planted trees in sparse random graphs 2267

Peter Keevash
A short proof of the existence of designs 2269

Marcus Kühn (joint with Felix Joos)
The Hypergraph Removal Process 2273

Matthew Kwan (joint with Kalina Petrova, Mehtaab Sawhney)
Parities in random Latin squares 2275

Jeck Lim (joint with David Conlon)
Sums of algebraic dilates 2278

Sam Mattheus (joint with Yamaan Attwa, Tibor Szabó, Jacques Verstraete)
Improved bounds for the minimum degree of minimal multicolor Ramsey graphs . . . 2281

Marcus Michelen (joint with Oren Yakir)
On the separation of roots of random polynomials . . . 2285

Richard Montgomery (joint with Natalie Behague, Daniel Il'kovič)
On the Kim–Vu sandwich conjecture . . . 2287

Konstantinos Panagiotou (joint with Umberto de Ambroggio, Tamas Makai, Annika Steibel)
Dispersion on the complete graph . . . 2289

Huy Tuan Pham (joint with Noga Alon)
Low-complexity approximations of sets with small doubling and applications . . . 2292

Alexey Pokrovskiy (joint with Alp Müyesser)
Group-harmonious labelling of trees . . . 2295

Julian Sahasrabudhe (joint with Marcelo Campos, Matthew Jenssen, Marcus Michelen)
A new lower bound for $R(3, k)$ . . . 2297

Maya Sankar
The Turán Density of Tight Cycles . . . 2300

Mehtaab Sawhney (joint with Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni)
Quasipolynomial bounds for the corners theorem . . . 2304

Benny Sudakov (joint with Zach Hunter, Aleksa Milojević, István Tomon)
Disjoint pairs in set systems and combinatorics of low rank matrices . . . 2307

Tibor Szabó (joint with Stefan Glock, Olaf Parczyk, Silas Rathke)
The maximum diameter of simplicial complexes . . . 2310

Marius Tiba (joint with Alessio Figalli, Peter van Hintum)
Stability of Functional and Geometric Inequalities . . . 2312

Liana Yepremyan (joint with Matija Bucic, Bryce Frederickson, Alp Müyesser, Alexey Pokrovskiy)
Approximate versions of Graham's conjecture . . . 2315

Abstracts

On designs and partial designs

Noga Alon

1. Results

We consider two related open problems of Erdős on block designs. Recall that a family of subsets $A_1, A_2 \dots, A_m$ of a finite set X is a (pairwise balanced incomplete) block design if every pair of distinct elements of X is contained in exactly one of the subsets A_i. It is a partial design if every pair of distinct elements of X is contained in at most one of the subsets A_i (equivalently, if $|A_i \cap A_j| \leq 1$ for all $1 \leq i < j \leq m$.)

The first problem deals with partial designs and appears in [6], see also [4], problem number 664.

Problem 1. *Is it true that for every fixed positive constant $c < 1$ there is a finite constant $C = C(c)$ so that the following holds. For every m and n and for every family of subsets $\{A_1, A_2, \dots, A_m\}$ of $[n] = \{1, 2, \dots, n\}$ that satisfies $|A_i| > c\sqrt{n}$ for all $1 \leq i \leq m$, and $|A_i \cap A_j| \leq 1$ for all $1 \leq i < j \leq m$, there is a subset $B \subset [n]$ so that $0 < |B \cap A_i| \leq C$ for all $1 \leq i \leq m$?*

The second problem appears in [5], page 35, see also [4], problem number 732.

Problem 2. *Call a sequence $n \geq x_1 \geq x_2 \geq \dots \geq x_m \geq 2$* block-compatible for n *if there is a pairwise balanced block design $A_1, A_2, \dots, A_m$ of m subsets of $[n]$ such that $|A_i| = x_i$ for $1 \leq i \leq m$. Is there an absolute constant $c > 0$ so that for all large n there are at least $e^{cn^{1/2} \log n}$ sequences that are block-compatible for n ?*

We show that the answer to the first problem is "no" and the answer to the second is "yes". The proofs are short, based on appropriate modifications of the family of lines of a projective plane which form a block design with $m = n = q^2 + q + 1$ subsets of cardinality $q + 1 = (1 + o(1))\sqrt{n}$ of a set of size $n = q^2 + q + 1$. It is well known that such a plane exists for any prime power q.

The following result settles Problem 1

Theorem 3. *Let q be a (large) prime power and put $m = n = q^2 + q + 1$. Then there is a partial design consisting of m subsets $A_1, A_2, \dots, A_m$ of an n element set P, so that $|A_i| > 0.4\sqrt{n}$ for all $1 \leq i \leq m$, $|A_i \cap A_j| \leq 1$ for all $1 \leq i < j \leq m$, and for any subset B of P that has a nonempty intersection with all sets A_j, there is some $1 \leq i \leq m$ so that $|B \cap A_i| \geq 0.1 \log n$.*

The next result settles problem 2.

Theorem 4. *Let q be a large prime power and put $n = q^2 + q + 1$. Let $S = (x_1 \geq x_2 \geq x_3 \geq \dots \geq x_m)$ be any sequence of integers satisfying*

$$q + 1 \geq x_1 \geq x_2 \geq x_3 \dots \geq x_n \geq 3,$$

$$m = n + \sum_{i=1}^{n} [\binom{q+1}{2} - \binom{x_i}{2}],$$

and $x_i = 2$ for all $n < i \leq m$. Then S is block-compatible for n. Therefore, the number of sequences that are block-compatible for n is at least

$$\binom{n+q-2}{q-2} = 2^{(0.5+o(1))n^{1/2} \log n}.$$

2. Remarks

- It is easy to see that the estimate in Theorem 3 is tight up to constant factors, for every partial design in which all blocks are of sizes $\Theta(\sqrt{n})$.
- Problem 1 for block designs (and not for partial designs) was also asked by Erdős in [5]. This remains open although we suspect that the answer here is negative as well. We suggest the following conjecture which, if true, would establish this negative answer.

  Conjecture 5. *Let q be a (large) prime power, put $n = q^2 + q + 1$, let P be the set of n points of a projective plane of order q and let $L_1, L_2, \ldots, L_n$ be the sets of points of its lines. Let R be a random subset of P obtained by picking each point of P randomly and independently to lie in R with probability $1/2$. Then with high probability the smallest cardinality of a subset B of R that intersects all the subsets $L_1 \cap R, L_2 \cap R, \ldots, L_n \cap R$ satisfies $|B|/q > f(q)$ for some function $f(q)$ tending to infinity as q tends to infinity. In fact, this may even be true with $f(q) = \Omega(\log q)$.*

  This conjecture remains open, although related results have been proved in [2], [3] using the container method. The parameters in these papers are very different and it seems that a proof here, if true, would require additional ideas.
- Call a sequence $n \geq x_1 \geq x_2 \geq \ldots \geq x_m \geq 2$ *line-compatible for n* if there is a set P of n points in the Euclidean plane R^2 so that for the family $L_1, L_2, \ldots, L_m$ of all lines in R^2 determined by the points of P, $|L_i \cap P| = x_i$ for $1 \leq i \leq m$. Note that every line-compatible sequence for n is also block-compatible for n, but the converse is not true. Erdős conjectured in [5] (see also [4], problem 733) that the number of sequences which are line compatible for n is only $2^{O(n^{1/2})}$. This upper bound was proved by Szemerédi and Trotter in [7]. Note that in view of Theorem 4 this is much smaller than the number of block-compatible sequences for n.

  Indeed, there are far more block designs on n points than designs that can be described by the lines determined by a set of points in the plane. This is demonstrated by the following result.

  Proposition 6.
  1. *The number of hypergraphs on n labelled vertices whose edges form a block design is $2^{\Theta(n^2 \log n)}$.*

(2) The number of hypergraphs whose vertices are n labelled points in R^2 and whose edges are the sets of points contained in the lines determined by the points is only $2^{\Theta(n \log n)}$.

The detailed proofs of all the results above can be found in [1].

References

[1] Noga Alon, Problems and results in Extremal combinatorics - V, to appear.
[2] József Balogh and Wojciech Samotij, An efficient container lemma, Discrete Anal. 2020, Paper No. 17, 56 pp.
[3] József Balogh and József Solymosi, On the number of points in general position in the plane, Discrete Anal. (2018), Paper No. 16, 20 pp.
[4] Thomas Bloom, Erdős' Open Problems, https://www.erdosproblems.com/all
[5] Paul Erdős, On the combinatorial problems which I would most like to see solved, Combinatorica (1981), 25–42.
[6] Paul Erdős, Some unsolved problems, in: Combinatorics, geometry and probability (Cambridge, 1993) (1997), 1–10.
[7] Endre Szemerédi and William T. Trotter, Extremal problems in discrete geometry, Combinatorica 3 (1983), no. 3-4, 381–392.

Hypercube statistics

Maria Axenovich

(joint work with Noga Alon, John Goldwasser)

Let Q_n be the hypercube of dimension n whose vertices are identified with n-component binary vectors. For a subset A of vertices of Q_n and $d \leq n$, let $\lambda(n, d, s, A)$ denote the fraction of subcubes Q_d of Q_n that contain exactly s vertices of A. Let $\lambda(n, d, s)$ denote the maximum possible value of $\lambda(n, d, s, A)$ as A ranges over all subsets of vertices of Q_n, and let $\lambda(d, s)$ denote the limit of this quantity as n tends to infinity. It is easy to see that the limit exists, and is the infimum over n of $\lambda(n, d, s)$ as for any fixed d, s the function $\lambda(n, d, s)$ is monotone non-increasing in n.

The problem of determining or estimating the quantities $\lambda(n, d, s)$ and $\lambda(d, s)$ is motivated by the questions and results of Goldwasser and Hansen on counting structural configurations in hypercubes [6], as well as by the results on edge-statistics in graphs by Alon, Hefetz, Krivelevich, and Tyomkyn [1], Kwan, Sudakov, and Tran [7], Martinsson, Mousset, Noever, and Trujic [8], and Fox and Sauermann [7].

Clearly $\lambda(d, s) = \lambda(d, 2^d - s)$ and $\lambda(d, 0) = 1$. In addition, if $s = 2^{d-1}$, then we see that $\lambda(d, s) = 1$ by taking all vertices of the hypercube with even number of ones.

To state our upper bounds on $\lambda(d, s)$, consider the generalized Johnson's graph $J(4s, 2s, s)$ whose vertex set is the set of $2s$-element subsets of a $4s$-element set, in which two vertices are adjacent if and only if the corresponding sets intersect in exactly s elements. Let $\omega(s) = \omega(J(4s, 2s, s))$ denote the clique number of $J(4s, 2s, s)$. It is known that $\omega(s) \leq 4s - 1$, see for example Godsil and Royle

[14]. Let $t(n,k)$ be the number of edges in the Turán graph $T(n,k)$, that is, the complete k-partite n-vertex graph with parts that are as equal as possible. Denote the density $t(n,k)/\binom{n}{2}$ by $\pi(n,k)$.

Theorem 1. *Let s and d be integers. Then $\lambda(d,s) = 1$ if and only if $s \in \{0, 2^d, 2^{d-1}\}$. If $1 < s < 2^{d-1}$, then*

$$\lambda(d,s) \le \lambda(d+2,d,s) = \pi(d+2,\omega(s)) \le \left(1 - \frac{1}{4s-1}\right)\left(1 + \frac{1}{d+1}\right).$$

In particular, $\lambda(d+2,d,s) = 1$ iff $d+2 \le \omega(s)$. When $s = 1$, we have $\lambda(d,1) \le \lambda(d+2,d,1) = \pi(d+2,3)$ for $d < 6$, and $\lambda(d+2,d,1) = 3/4$ otherwise.

Note that the general upper bound implies in particular that if s is not large, say, $0 < s < d/8$, then $\lambda(d,s) \le 1 - \Omega(1/s)$.

To state our lower bounds on $\lambda(d,s)$ we need to define c_d and $c(d,k)$. Let c_d denote the probability that a random d by d binary matrix whose rows are random independent non-zero vectors of $\mathbb{F}_2^d$ is nonsingular (in $\mathbb{F}_2$). It is easy and well known that

$$c_d = \prod_{i=1}^{d-1}\left(1 - \frac{2^i - 1}{2^d - 1}\right),$$

which is roughly 0.289 for large d.

For $1 \le k \le d$, let $c(d,k)$ denote the probability that a random $(d-k)$ by d binary matrix whose columns are uniform random vectors in $\mathbb{F}_2^{d-k}$ is of rank $d-k$ (over $\mathbb{F}_2$). By choosing the rows (not the columns) of the matrix one by one ensuring that each row does not lie in the span of the previous ones it is easy to see that

$$c(d,k) = \prod_{i=0}^{d-k-1}\left(1 - \frac{2^i}{2^d}\right).$$

Note that this quantity is larger than $1 - \frac{1}{2^k}$.

The first simple lower bound in the theorem below appears in the paper Goldwasser and Hansen [6], we state it here for completeness. Note that this lower bound approaches $e^{-1} \approx 0.37$ as d tends to infinity.

Theorem 2. *For any integer $d \ge 2$, $\lambda(d,1) \ge \left(1 - 2^{-d}\right)^{2^d - 1}$. For all admissible d and s, $\lambda(d,s) \ge c_d$. Moreover, for every s of the form $s = 2^k \cdot j$, where j is an odd integer, which satisfies $0 < s \le 2^{d-1}$, $\lambda(d,s) \ge c(d,k)$. In particular, for any s which is a power of 2, $\lambda(d,s) \ge 1 - \frac{1}{s}$.*

We summarise the best bounds we know on $\lambda(d,1)$ when $d = 2, 3$, or 4, as well as the known exact values of $\lambda(d,s)$. Observe that $\lambda(d,1) \ge 2/(d+1)$ by the following construction. The Hamming weight of a binary vector is its number of 1's. For a fixed d, let A be the set of all vertices in Q_n with Hamming weight divisible by $d+1$. A copy of a Q_d-cube contains precisely one vertex in A if and only if the smallest Hamming weight of any of its vertices is congruent to 0 or 1 (mod $d+1$). Together with upper bounds established by Baber [2] using the Flag Algebra method, we have the following estimates for $d = 2, 3$, and 4: $2/3 \leq \lambda(2,1) \leq 0.68572$, $0.5 \leq \lambda(3,1) \leq 0.61005$, and $0.4 \leq \lambda(4,1) \leq 0.60254$. Motivated by our paper, Bodnár and Pikhurko [3] determined $\lambda(d,s)$ for three pairs (d,s) where $\lambda(d,s) \neq 1$. Using the Flag Algebra method they proved that $\lambda(3,2) = 8/9$, $\lambda(4,2) = 264/343$ and $\lambda(4,4) = 26/27$. Rahil Baber (personal communication) was able to re-prove these results using the Flag Algebra method with his code. At the moment these are the only pairs (d,s) for which $\lambda(d,s)$ is not 1 and is known precisely. In all three cases the lower bounds follow from the proof of Theorem 2.

References

[1] N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn, Edge-statistics on large graphs, Combin. Probab. Comput. 29 (2020), no. 2, 163–189.

[2] R. Baber, Some results in extremal combinatorics, Ph.D. Thesis, 2011.

[3] L. Bodnár and O. Pikhurko, Some exact values of the inducibility and statistics constants for hypercubes, arXiv:2503.03408, 2025.

[4] J. Fox and L. Sauermann, A completion of the proof of the edge-statistics conjecture, Adv. Comb. (2020), Paper No. 4, 52.

[5] C. Godsil and G. Royle, Algebraic Graph Theory, Springer-Verlag, New York, 2001, xx+439 pp.

[6] J. Goldwasser and R. Hansen, Inducibility in the hypercube J. Graph Theory 105 (2024), no. 4, 501–522.

[7] M. Kwan, B. Sudakov, and T. Tran, Anticoncentration for subgraph statistics, J. Lond. Math. Soc. (2) 99 (2019), 757–777.

[8] A. Martinsson, F. Mousset, A. Noever, and M. Trujic, The edge-statistics conjecture for $\ell < k^{6/5}$, Israel J. Math. 234 (2019), 677–690.

Robustness and resilience for spanning graph of bounded degree

Julia Böttcher

(joint work with Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, Mihir Neve)

Much progress has been made concerning Dirac-type conditions for the containment of spanning substructures in dense graphs. One conjecture that, however, still remains open is the well-known conjecture of Bollobás, Eldridge, and Catlin [1, 2], which states that any n-vertex graph with minimum degree of $(1 - \frac{1}{\Delta+1})n$ contains any n-vertex graph with maximum degree Δ as a subgraph. The best general condition for all Δ is still given by the following classic theorem of Sauer and Spencer [3].

Theorem 1 (Sauer–Spencer Theorem). *Let $\Delta > 0$ be given. Suppose that G is an n-vertex graph with minimum degree $\delta(G) \geq (1 - \frac{1}{2\Delta})n$ and H is an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$. Then H is a spanning subgraph of G.*

For bipartite graphs, improving on [5], in recent work the following was obtained, which shows that in general one can do better when additionally the chromatic number is restricted.

Theorem 2 (Allen, Böttcher, Skokan, Sudakov). *There is a constant $c > 0$ such that for each sufficiently large n and each $\Delta \geq 1$, every n-vertex graph G with minimum degree*

$$\delta(G) \geq \Big(1 - c\frac{\log \Delta}{\Delta}\Big)n\,,$$

contains each n-vertex bipartite graph H with maximum degree at most Δ.

This is tight up to the value of c.

Moving to sparser hostgraphs, we recently obtained the following robustness and resilience results. Given a graph G we write $G(p)$ for the subgraph of G in which every edge of G is included independently with probability p. The binomial random graph $G(n,p)$ then equals $K_n(p)$. The *maximum* 1*-density* $m_1(H)$ of the graph H is defined as follows. Let $d_1(H) = e(H)/(v(H)-1)$ denote the 1*-density* of H, and let $m_1(H) = \max_{H' \subseteq H} d_1(H')$, where the maximum runs over all subgraphs of H with at least two vertices.

Theorem 3 (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 $\delta(G) \geq \big(\frac{2\Delta-1}{2\Delta} + \gamma\big)n$, then for $p \geq Cn^{-1/m_1(H)} \log n$ asymptotically almost surely the graph H is a subgraph of $G(p)$.*

For the proof of this theorem, we combine a spread version of the dense case of the sparse blow-up lemma from [4] with an extension of the Sauer-Spencer Theorem.

Theorem 4 (Resilient Sauer–Spencer Theorem). *For all $\Delta \geq 2, \gamma \in (0, 1/2\Delta)$, there exists a constant $C > 0$ such that for $p \geq (\log n/n)^{1/\Delta}$, the following holds asymptotically almost surely for $\Gamma \sim G(n,p)$. Let G be a spanning subgraph of Γ with minimum degree $\delta(G) \geq \big(1 - \frac{1}{2\Delta} + \gamma\big)pn$, and let H be an n-vertex graph with maximum degree $\Delta(H) \leq \Delta$ and with at least Cp^{-2} vertices which are not contained in any triangles of H. Then, G contains a copy of H.*

The proof of this theorem uses the sparse blow-up lemma. The restriction concerning vertices which are not in triangles is necessary here. The bound on p is unlikely to be optimal.

References

[1] 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.

[2] P. A. Catlin, *Embedding subgraphs and coloring graphs under extremal degree conditions*, Thesis (Ph.D.)–The Ohio State University, 1976.

[3] N. Sauer and J. Spencer, *Edge disjoint placement of graphs*, J. Combin. Theory Ser. B **25** (1978), no. 3, 295–302.

[4] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, Discrete Anal. (2025), Paper No. 8, 141 pp.

[5] B. Csaba, *On the Bollobás–Eldridge conjecture for bipartite graphs*, Combin. Probab. Comput. **16** (2007), no. 5, 661–691.

Lower bounds for Ramsey numbers of bounded degree hypergraphs

Domagoj Bradač
(joint work with Zach Hunter and Benny Sudakov)

1. Introduction

Given a positive integer q and a k-uniform hypergraph (or k-graph for short) H, the q-color Ramsey number of H, denoted by $r(H;q)$ is the minimum integer N such that in any q-coloring of the complete k-uniform hypergraph on N vertices, there is a monochromatic copy of H.

One of the key directions in (hyper)graph Ramsey theory is studying the Ramsey numbers of sparse (hyper)graphs. A seminal result of Chvátal, Rödl, Szemerédi and Trotter [2] states that graphs with bounded degree have linear Ramsey numbers. A hypergraph analogue of this result was proved by Cooley, Fountoulakis, Kühn and Osthus [5, 6] and for 3-uniform hypergraphs independently by Nagle, Olsen, Rödl and Schacht [9].

Our topic is understanding the above results quantitatively. Namely, given integers k, q, Δ with Δ sufficiently large, we define $C^{(k)}(\Delta;q)$ as the infimum value of C such that for all sufficiently large n, any n-vertex k-graph H with maximum degree Δ satisfies $r(H;q) \leq Cn$.

For graphs, there has been a long line of study estimating $C^{(2)}(\Delta;q)$. For two colors the current best bounds are $2^{\Omega(\Delta)} \leq C^{(2)}(\Delta;2) \leq 2^{O(\Delta \log \Delta)}$, where the lower bound is due to Graham, Rödl and Rucinski [8] and the upper bound due to Conlon, Fox and Sudakov [4]. For more than two colors, the best upper bound is $C^{(2)}(\Delta;q) \leq 2^{O_q(\Delta^2)}$ by Fox and Sudakov [7].

Turning our attention to hypergraphs, the best upper bounds are $C^{(k)}(\Delta,q) \leq \mathrm{tw}_k(c\Delta)$, for $k \geq 4$, and $C^{(3)}(\Delta,q) \leq \mathrm{tw}_3(c'\Delta \log \Delta)$, for $k=3$ due to Conlon, Fox and Sudakov [3]. Some evidence that these bounds are close to the truth is given by constructions of Conlon, Fox and Sudakov for $k=3$ and of Bradač, Fox and Sudakov for $k \geq 4$, of k-graphs with n vertices, maximum degree $O_k(n)$ and 4-color hypergraph Ramsey number $\mathrm{tw}_k(\Omega(n))$. Thus it is a natural question posed in [3] and reiterated in [1] whether there exist k-graphs H with maximum degree Δ and arbitrarily many vertices for which the Ramsey number is at least $\mathrm{tw}_k(c_k\Delta) \cdot n$, that is, to determine whether $C^{(k)})(\Delta;4) \geq \mathrm{tw}_k(c_k\Delta)$, for some $c_k > 0$ depending only on k. Our main theorem is a positive answer to the above question.

Theorem 1. *For any $k \geq 2$, there is a constant $c_k > 0$ such that for any integers $\Delta \geq 1/c_k$ and $n \geq \Delta$, there exists a k-uniform n-vertex hypergraph H with maximum degree at most Δ whose 4-color Ramsey number is at least* $\mathrm{tw}_k(c_k\Delta) \cdot n$.

2. Proof ideas

Our proof borrows some ideas from the work of Graham, Rödl and Ruciński [8] who proved the lower bound for graphs. Additionally, as a building block, we use the aforementioned construction by Bradač, Fox and Sudakov [1] of k-graphs with m edges whose 4-color Ramsey number is $\mathrm{tw}_k(\Theta(\sqrt{m}))$.

The k-graph H in Theorem 1 is the edge union of two hypergraphs. The first one, which we shall call H_R, where R stands for "random", is obtained from the binomial random hypergraph with average degree $O(\Delta)$ by removing high degree vertices. The second one, which we denote by H_E, where E stands for "expander", is obtained by taking $O(n/\Delta)$ random copies of a k-graph H_0 on $O(\Delta)$ vertices with maximum degree $O(\Delta)$. The construction of our gadget H_0 is a small modification of the construction from [1] and is obtained by considering a d-regular 2-graph F with expansion properties (here d depends only on k), and putting a k-edge for any $k-1$ vertices spanning a tree in F and any distinct k-th vertex.

Informally speaking, H_0 has two crucial properties. First, any pair of vertices is contained in a k-edge. Second, given any set $W \subset V(H_0)$ of size $\varepsilon|V(H_0)|$ (where $\varepsilon > 0$ is not too small with respect to d), the auxiliary $(k-1)$-graph $H_0' = H_0'(W)$ where e is an edge if $e \cup \{w\} \in E(H_0)$ for some $w \in W$ "behaves like" the $(k-1)$-uniform gadget hypergraph we define from F. More specifically, since F is an expander, there will be $U \subset V(F)$ of size $(1-\varepsilon)|V(F)|$ so that each $u \in U$ has a neighbor in W, whence H_0' contains a $(k-1)$-uniform edge for every $k-2$ vertices in U spanning a tree in F, along with any distinct $(k-1)$-th vertex. Consequently, any two vertices in U belong to a common edge of H_0', and we can further iterate this (defining an analogous auxiliary $(k-2)$-graph H_0'', it should again behave like our gadget in uniformity $k-2$). The fact that the hypergraph 'behaves similarly' at lower uniformities is what allows us to use stepping-up techniques.

The coloring that avoids a monochromatic copy of H is obtained by blowing up a modification of the usual stepping-up coloring. More concretely, a variant of the stepping-up coloring yields a 4-coloring on $\mathrm{tw}_k(c_k\Delta)$ vertices, which, for example, has no monochromatic copy of $K_\Delta^{(k)}$. Then, we take an $\Omega_k(n/\Delta)$-blow-up of this coloring to get a coloring on $\mathrm{tw}_k(c_k\Delta) \cdot n$ vertices. There is some freedom when performing this blow-up and we can tailor it to our setting to ensure the resulting coloring avoids monochromatic copies of H.

References

[1] D. Bradač, J. Fox, and B. Sudakov. Ramsey numbers of hypergraphs with a given size. *Math. Proc. Cambridge Philos. Soc.*, 178(1):31–44, 2025.

[2] C. 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.

[3] D. Conlon, J. Fox, and B. Sudakov. Ramsey numbers of sparse hypergraphs. *Random Structures & Algorithms*, 35(1):1–14, 2009.

[4] D. Conlon, J. Fox, and B. Sudakov. On two problems in graph Ramsey theory. *Combinatorica*, 32(5):513–535, 2012.

[5] O. Cooley, N. Fountoulakis, D. Kühn, and D. Osthus. 3-uniform hypergraphs of bounded degree have linear Ramsey numbers. *J. Combin. Theory Ser. B*, 98(3):484–505, 2008.

[6] O. Cooley, N. Fountoulakis, D. Kühn, and D. Osthus. Embeddings and Ramsey numbers of sparse k-uniform hypergraphs. *Combinatorica*, 29(3):263–297, 2009.

[7] Jacob Fox and Benny Sudakov. Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica*, 29(2):153–196, 2009.

[8] R. L. Graham, V. Rödl, and A. Ruciński. On graphs with linear Ramsey numbers. *Journal of Graph Theory*, 35(3):176–192, 2000.

[9] B. Nagle, S. Olsen, V. Rödl, and M. Schacht. On the Ramsey number of sparse 3-graphs. *Graphs Combin.*, 24(3):205–228, 2008.

On Graham's rearrangement conjecture over $\mathbb{F}_2^n$

Matija Bucić
(joint work with Benjamin Bedert, Noah Kravitz, Richard Montgomery, Alp Müyesser)

A sequence $g_1, g_2, \ldots, g_n$ of elements of a (multiplicative) group G is a *valid ordering* if the partial products

$$g_1, \quad g_1g_2, \quad g_1g_2g_3, \quad \ldots, \quad g_1 \cdots g_n$$

are all distinct. Which subsets of groups admit valid orderings? Variants of this natural problem have been studied in many different cases over the years.

The first question in this direction appeared in 1961, when Gordon [9], motivated by constructions of complete Latin squares, asked for which finite groups the entire group has a valid ordering. Gordon gave a complete characterization in the abelian case: A finite (additive), nontrivial abelian group G admits a valid ordering if and only if $\sum_{g\in G} g \neq 0$, this being the obvious necessary condition for the existence of such an ordering. In 1974, Ringel [16] posed the closely related problem of characterising the groups G whose elements can be ordered as $g_1, \ldots, g_n$ in such a way that $g_1 = g_1g_2 \cdots g_n = \mathrm{id}$ but otherwise all partial products are distinct. The motivation for this question came from Ringel's solution [17] of the Heawood map colouring conjecture.

The nonabelian case of Gordon's problem is more subtle, since there are some small nonabelian groups (such as S_3) that, for no apparent reason, fail to have valid orderings. In 1981, Keedwell [12] posed the bold conjecture that every sufficiently large nonabelian group has a valid ordering. Müyesser and Pokrovskiy [13] recently proved Keedwell's conjecture as a consequence of their more general probabilistic analogue of the Hall–Paige Conjecture [7,11] concerning the existence of transversals in multiplication tables. This work also shows that large groups have an ordering, in the sense that Ringel asked for, if and only if the product of all group elements (in any order) is an element of the commutator subgroup[1].

Here, we will be concerned not only with the case when an entire group G admits a valid ordering but with the more general question of when an *arbitrary* subset S of a given group G admits a valid ordering. Notice that when S contains the identity element, every possible valid ordering of S must start with the identity, since otherwise two consecutive partial products would be equal. Thus, if G is abelian and $\sum_{g\in S} g = 0$, then there cannot be a valid ordering of S. In order to avoid this obstruction, we restrict our attention to subsets S not containing the identity, and the following is our central question.

[1] This condition is equivalent to existence of an ordering $g_1, \ldots, g_n$ such that $g_1g_2 \cdots g_n = \mathrm{id}$.

Question 1. *For which groups G does every subset $S \subseteq G \setminus \{\mathrm{id}\}$ admit a valid ordering?*

It seems feasible that the answer to this question is affirmative for every finite group G. At first glance, finding valid orderings for smaller subsets S might seem like an easier task, since there is more space to place the partial products without creating collisions. However, the potential obstructions for small S are at least as rich as for Gordon's setting $S = G \setminus \{\mathrm{id}\}$, since a small set S may itself be a subgroup of G, or could be a complicated conglomeration of approximate subgroups and random-like sets. In the graph-theoretic formulation of these problems, which we will describe below, Gordon's setting corresponds to the complete graph case (in particular, a directed variant of a well-known conjecture of Andersen [2]), whereas Theorem 1 corresponds to a sparse analogue. Such sparse analogues in extremal graph theory tend to be harder and less well understood than their dense counterparts.

The simplest instance of Question 1 is when $G = \mathbb{F}_p$, for a prime p. This problem was first posed by Graham [10] in 1971 and later reiterated in an open problems book of Erdős and Graham [8].

Conjecture 2 (Graham)**.** *Let p be prime. Then every subset of $\mathbb{F}_p \setminus \{0\}$ admits a valid ordering.*

Most previous work towards Conjecture 2 has concerned the edge cases where either S or $\mathbb{F}_p \setminus S$ is very large. The best result for small sets S is due to Bedert and Kravitz [3], who showed that every set $S \subseteq \mathbb{F}_p \setminus \{0\}$ of size at most $e^{\log^{1/4} p}$ has a valid ordering. For very large sets S, the aforementioned result of Müyesser and Pokrovskiy [13] establishes Conjecture 2 for all sets $S \subseteq \mathbb{F}_p \setminus \{0\}$ of size at least $(1 - o(1))p$ (and indeed proves an analogous result for all finite groups). The intermediate regime remains open.

Various groups of authors have considered instances of Question 1 other than $G = \mathbb{F}_p$. In particular, Alspach [5] conjectured an affirmative answer to Question 1 for all finite abelian groups G, and Alspach and Liversidge [1] confirmed this for subsets of size up to 11. For extensions of this problem to a nonabelian setting, see [6, 14] and the dynamic survey of Ollis [15].

In a different direction, Bucić, Frederickson, Müyesser, Pokrovskiy, and Yepremyan [4] have recently provided an affirmative answer to an "approximate" relaxation of Question 1. They showed that every finite subset S of any group G has an ordering in which all but $o(|S|)$ partial products are distinct.

Main results. Despite the partial progress discussed above, there is no infinite class of groups G for which we have a complete understanding of Question 1. Our main result remedies this situation for the family of groups $\mathbb{F}_2^n$.

Theorem 3. *There is an absolute constant C such that for all $n \in \mathbb{N}$, every set $S \subseteq \mathbb{F}_2^n \setminus \{0\}$ of size at least C has a valid ordering.*

Our proof of Theorem 3 treats the "sparse S" and "dense S" regimes separately. Our argument for the sparse case makes use of the specific structure of $\mathbb{F}_2^n$, but our argument for the dense case applies to general (even nonabelian) groups. In particular, we are able to provide an affirmative answer to Question 1 if one restricts attention to subsets S of size at least $|G|^{1-c}$; this significantly improves on the result of Müyesser and Pokrovskiy [13], which treats only subsets S of size $(1-o(1))|G|$.

Theorem 4. *There is an absolute constant $c > 0$ such that for any finite (possibly nonabelian) group G, every subset $S \subseteq G \setminus \{\mathrm{id}\}$ of size at least $|G|^{1-c}$ admits a valid ordering.*

References

[1] B. Alspach and G. Liversidge, *On strongly sequenceable abelian groups*, Art Discrete Appl. Math. (2020).

[2] L. D. Andersen, *Hamilton circuits with many colours in properly edge-coloured complete graphs*, Math. Scand. **64** (1989), no. 1, 5–14.

[3] B. Bedert and N. Kravitz, *Graham's rearrangement conjecture beyond the rectification barrier*, Isr. J. Math. (to appear), arXiv preprint 2409.07403.

[4] M. Bucić, B. Frederickson, A. Müyesser, A. Pokrovskiy, and L. Yepremyan, *Towards Graham's rearrangement conjecture via rainbow paths*, arXiv preprint 2503.01825 (2025).

[5] S. Costa and M. A. Pellegrini, *Some new results about a conjecture by Brian Alspach*, Arch. Math. **115** (2020), no. 5, 479–488.

[6] S. Costa, S. Della Fiore, and E. Engel, *Graham's rearrangement for dihedral groups*, arXiv preprint 2503.18101 (2025).

[7] S. Eberhard, F. Manners, and R. Mrazović, *An asymptotic for the Hall–Paige conjecture*, Adv. Math. **404** (2022), 108423.

[8] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory*, Université de Genève, L'Enseignement Mathématique, Geneva, 1980.

[9] B. Gordon, *Sequences in groups with distinct partial products*, Pacific J. Math. **11** (1961), 1309–1313.

[10] R. L. Graham, *On sums of integers taken from a fixed sequence*, Proceedings of the Washington State University Conference on Number Theory (Washington State Univ., Pullman, Wash., 1971), Washington State University, Department of Mathematics, Pi Mu Epsilon, Pullman, WA, 1971, pp. 22–40.

[11] M. Hall and L. J. Paige, *Complete mappings of finite groups*, Pacific J. Math. **5** (1955), 541–549.

[12] A. Keedwell, P. Cameron, J. Hirschfeld, and D. Hughes, *Sequenceable groups: a survey*, LMS Lecture Notes **49** (1981), 205–215.

[13] A. Müyesser and A. Pokrovskiy, *A random Hall-Paige conjecture*, Inventiones mathematicae **240** (2025), 779–867.

[14] M. A. Ollis, *Sequences in dihedral groups with distinct partial products*, arXiv preprint 1904.07646.

[15] M. A. Ollis, *Sequenceable groups and related topics*, Electron. J. Combin. **DS10** (2002), 34.

[16] G. Ringel, *Cyclic arrangements of the elements of a group.*, Notices Amer. Math. Soc. **21** (1974), no. 1, A–95.

[17] G. Ringel, *Map color theorem*, Die Grundlehren der mathematischen Wissenschaften, vol. 209, Springer-Verlag, New York-Heidelberg, 1974.

[18] B. Sudakov, *Restricted subgraphs of edge-colored graphs and applications*, arXiv preprint 2412.13945.

Induced Ramsey Numbers

Marcelo Campos
(joint work with Lucas Aragão, Gabriel Dahia, Rafael Filipe, João Marciano)

We write $G \xrightarrow{\text{ind}} H$ to denote the following property: for any red/blue colouring of the edges of G, there exists an induced monochromatic copy of H (that is, a copy of H which is induced in G, and all of its edges have the same colour). We then define

$$R_{\text{ind}}(H) = \min\big\{v(G) : G \xrightarrow{\text{ind}} H\big\}.$$

In particular, observe that we have $R_{\text{ind}}(K_k) = R(k)$, the usual Ramsey number, since every copy of K_k in a graph G is also an induced subgraph of G. For general graphs H, on the other hand, Erdős remarked that even "the existence of [the induced Ramsey number] is not at all obvious."

Deuber [3], Erdős, Hajnal and Posa [6] and Rödl [10] independently established in the 1970s that $R_{\text{ind}}(H)$ is finite for every graph H. While none of these works provide an explicit dependency on k, the number of vertices of H, [5] later remarked that the best bound that one can deduce from the proofs in [3, 6, 10] is of the form

$$R_{\text{ind}}(H) \leq 2^{2^{k^{1+o(1)}}}.$$

Nevertheless, Erdős [4, 5] conjectured, first implicitly in 1975 and then explicitly in 1984, that the function $R_{\text{ind}}(H)$ should grow at most exponentially as a function of $v(H) = k$ for every H. Note that, if true, this would be best possible, since we have $R_{\text{ind}}(K_k) = R(k) \geq 2^{k/2}$.

Using the techniques of [10], one can prove Erdős' conjecture for bipartite H. The problem is much harder for non-bipartite H, however, and the next significant advance was not obtained until almost 25 years later, by [9]. By taking G (the host graph) to be a random graph built using projective planes, they showed, among other results, that

$$R_{\text{ind}}(H) \leq k^{O(k \log k)} \qquad (1)$$

for every graph H with k vertices. [7] later provided an explicit, pseudorandom graph attaining the bound in (1).

A few years later, [1] removed a factor of $\log k$ from the exponent in (1) and showed, using an explicit graph, that

$$R_{\text{ind}}(H) \leq k^{O(k)}. \qquad (2)$$

In order to prove (2), the authors of [1] developed a general method for proving Ramsey-type theorems using pseudorandom properties of the host graph G.

The result presented in the talk confirms Erdős' conjecture for all graphs H.

Theorem 1. *There exists a constant $C > 0$ such that*

$$R_{\text{ind}}(H) \leq 2^{Ck}$$

for every graph H with k vertices.

We also consider the induced Ramsey number for r-colourings. Define $R_{\text{ind}}(H;r)$ to be the *r-colour induced Ramsey number* of a graph H; that is, the minimum number of vertices of a graph G such that every r-colouring $c\colon E(G) \to [r]$ of the edges of G contains an induced monochromatic copy of H. The techniques used in [1] and [9] do not work in this more general setting, and provide no bounds when $r \geq 3$, but [8] introduced a different approach in 2009, which can be used to show that

$$R_{\text{ind}}(H;r) \leq r^{O(rk^2)}. \tag{3}$$

The large gap between (3) and the known bounds for the $r = 2$ case motivated Problem 3.5 in [2] that asks if one could show that, for fixed $r \in \mathbb{N}$,

$$R_{\text{ind}}(H;r) \leq 2^{k^{1+o(1)}}.$$

The methods used to establish Theorem 1 can also solve this problem in a very strong form.

Theorem 2. *There exists a constant $C > 0$ such that*

$$R_{\text{ind}}(H;r) \leq r^{Crk} \tag{4}$$

for every $r \geq 2$ and every graph H with k vertices.

These methods moreover imply the stronger statement that for almost all graphs G with $N = r^{Crk}$ vertices every r-colouring of the edges of G contains an induced monochromatic copy of *every* graph H on k vertices.

References

[1] D. Conlon, J. Fox, and B. Sudakov. On two problems in graph Ramsey theory. *Combinatorica*, 32:513–535, 2012.

[2] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. *Surveys in combinatorics*, 424:49–118, 2015.

[3] W. Deuber. A generalization of Ramsey's theorem. *Infinite and Finite Sets*, 10:323–332, 1975.

[4] P. Erdős. Problems and results on finite and infinite graphs. *Recent advances in graph theory*, pages 183–192. (loose errata), 1975.

[5] P. Erdős. On some problems in graph theory, combinatorial analysis and combinatorial number theory. *Graph theory and combinatorics*, pages 1–17, 1984.

[6] P. Erdős, A. Hajnal, and L. Pósa. Strong embeddings of graphs into colored graphs. *Infinite and Finite sets*, 10:585–595, 1975.

[7] J. Fox and B. Sudakov. Induced Ramsey-type theorems. *Adv. Math.*, 219:1771–1800, 2008.

[8] J. Fox and B. Sudakov. Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica*, 29:153–196, 2009.

[9] Y. Kohayakawa, H. J. Prömel, and V. Rödl. Induced Ramsey numbers. *Combinatorica*, 18: 373–404, 1998.

[10] V. Rödl. The dimension of a graph and generalized Ramsey theorems. Master's thesis, Charles University, 1973.

Applications of Refined Absorption

Michelle Delcourt
(joint work with Cicley Henderson, Thomas Lesgourges, Tom Kelly, Luke Postle)

The study of combinatorial designs has a rich history spanning nearly two centuries. In a recent breakthrough, the notorious Existence Conjecture for Combinatorial Designs dating back to the 1800s was proved in full by Keevash [12] via the method of *randomized algebraic constructions*. Subsequently, Glock, Kühn, Lo, and Osthus [10] provided an alternate purely combinatorial proof of the Existence Conjecture via the method of *iterative absorption*. Very recently Delcourt and Postle [6] introduce a novel method of *refined absorption* for designs and use it to provide a new proof of the Existence Conjecture. Our method can also be used in a 'black box' fashion and applied to many other problems in probabilistic design theory. In this talk, we explore some of out recent results including the Existence of High Girth Designs (Delcourt and Postle [7]), the large minimum degree setting (Delcourt, Henderson, Lesgourges, and Postle [2]), and finding sufficiently spread distributions on designs (Delcourt, Kelly, and Postle [3]).

1. Introduction. One of the most classical theorems in all of design theory is a result of Kirkman [13] which classifies for which n there exists a set of triples of an n-set V such that every pair in V is in exactly one triple; such a set is called a *Steiner Triple System*. From the graph theoretic perspective, this is equivalent to a decomposition of the edges of the complete graph K_n into edge-disjoint triangles. Kirkman proved that the necessary divisibility conditions for this, namely that $3 \mid \binom{n}{2}$ and $2 \mid (n-1)$ which equates to $n \equiv 1, 3 \mod 6$, are also sufficient.

Arguably the most studied object in design theory is a natural generalization of this object as follows. A *Steiner system* with parameters (n, q, r) is a set S of q-subsets of an n-set V such that every r-subset of V belongs to exactly one element of S. More generally, a *design* with parameters (n, q, r, λ) is a set S of q-subsets of an n-set V such that every r-subset of V belongs to exactly λ elements of S. The notorious Existence Conjecture originating from the mid-1800's asserts that designs exist for large enough n provided the obvious necessary divisibility conditions are satisfied as follows: Let $q > r \geq 2$ and $\lambda \geq 1$ be integers. If n is sufficiently large and $\binom{q-i}{r-i} \mid \lambda\binom{n-i}{r-i}$ for all $0 \leq i \leq r-1$, then there exists a design with parameters (n, q, r, λ).

2. Nash–Williams' Conjecture. Focusing for a moment on triangle decompositions, for host graphs other than K_n, K_3-divisibility may not be sufficient to guarantee the existence of a K_3-decomposition; consider for instance the cycle C_6 which is certainly K_3-divisible but does not contain K_3 as a subgraph, let alone an edge decomposition into copies of K_3. A natural question is if we impose a local density condition on G such as bounding the minimum degree from below can we guarantee the existence of a K_3-decomposition?

The most famous conjecture in this setting is one by Nash–Williams from 1970:

Conjecture 1 (Nash–Williams [15]). *Let G be a K_3-divisible graph with n vertices and minimum degree $\delta(G) \geq \frac{3}{4}n$. If n is sufficiently large, then G admits a K_3-decomposition.*

In an addendum to Nash–Williams' original article [15], it is noted that if true, then this conjecture would be best possible as Graham was able to produce a construction showing that the fraction 3/4 is tight.

From the perspective of optimization, one could instead study the so-called fractional relaxation of the problem. Here $\delta^*_{K_3}$ denotes the *fractional K_3-decomposition threshold*, that is, the infimum of all real numbers c such that every graph G with minimum degree at least $c \cdot v(G)$ has a *fractional K_3-decomposition* (an assignment of non-negative weights to the triangles of G such that for each edge, the sum of the weights of triangles containing that edge is exactly 1). In a breakthrough result Barber, Kühn, Lo, and Osthus [1] show that the existence of a K_3-decomposition is fundamentally related to the existence of a fractional K_3-decomposition as follows.

Theorem 2 (Barber, Kühn, Lo, and Osthus [1]). *Let $\varepsilon > 0$. Any sufficiently large, K_3-divisible graph G on n vertices with minimum degree $\delta(G) \geq \left(\max\left\{\delta^*_{K_3}, \frac{3}{4}\right\} + \varepsilon\right) \cdot n$ admits a K_3-decomposition.*

Conjecture 3 (Fractional Nash–Williams). *If G is a graph on n vertices with minimum degree $\delta(G) \geq \frac{3}{4}n$, then G admits a fractional K_3-decomposition.*

Subsequently over the years, there has been much interest in the fractional relaxation; the best-known upper bound is due to Delcourt and Postle [4] from 2021 who proved that $\delta^*_{K_3} \leq \frac{7+\sqrt{21}}{14} \approx 0.82733$.

3. Generalizing Erdős' High Girth Existence Conjecture. We say a (j, i)-*configuration* in a set P of sets on a ground set V is a set of i elements of P spanning at most j elements of V. Thus a (j, i)-configuration in a K_3-packing is a set of i triangles spanning at most j vertices. Observe that one triangle is a $(4, 1)$-configuration and any two triangles sharing a vertex is a $(5, 2)$-configuration. Indeed, every $(n, 3, 2)$-Steiner system contains an $(i+3, i)$-configuration for every $1 \leq i \leq n-3$; this observation prompted Erdős to study $(i+2, i)$-configurations in the 1970s. The *girth* of a triangle packing is the smallest integer $i \geq 2$ such that the packing contains an $(i+2, i)$-configuration. In 1973 Erdős [8] conjectured:

Conjecture 4. *For every integer $g \geq 3$, every sufficiently large K_3-divisible complete graph admits a K_3-decomposition with girth at least g.*

This was recently proved by Kwan, Sah, Sawhney, and Simkin [14]. For Steiner systems more generally, in 2022 the approximate version of this High Girth Existence Conjecture was settled by Delcourt and Postle [5] and, independently, Glock, Joos, Kim, Kühn, and Lichev [9]; that is, they proved the existence of approximate (n, q, r)-Steiner systems of high girth with almost full size. In fact, both papers developed a general methodology that finds almost perfect matchings in hypergraphs that avoid a set of forbidden submatchings provided certain degree and codegree conditions are met. In particular, the general results then imply the approximate version of the High Girth Existence Conjecture as a corollary. Recently, in 2024 Delcourt and Postle [7] proved the High Girth Existence Conjecture in full using their newly developed refined absorption methodology.

4. Erdős meet Nash–Williams. The common generalization of the two preceding conjectures was dubbed the "Erdős meets Nash-Williams' Conjecture" by Glock, Kühn, and Osthus [11] in 2021. The main result in Delcourt, Henderson, Lesgourges, and Postle [2] is the following:

Theorem 5. *For every integer $g \geq 3$ and real $\varepsilon > 0$, any sufficiently large K_3-divisible graph G on n vertices with minimum degree $\delta(G) \geq \left(\max\left\{\delta^*_{K_3}, \frac{3}{4}\right\}+\varepsilon\right)\cdot n$ admits a K_3-decomposition with girth at least g.*

Combined with the previous work of Delcourt and Postle [4] we obtain:

Corollary 6 (Corollary to 5)**.** *For every integer $g \geq 3$ and real $\varepsilon > 0$, every sufficiently large K_3-divisible graph G on n vertices with minimum degree $\delta(G) \geq \left(\frac{7+\sqrt{21}}{14} + \varepsilon\right) \cdot n$ admits a K_3-decomposition with girth at least g.*

References

[1] B. Barber, D. Kühn, A. Lo, and D. Osthus, *Edge-decompositions of graphs with high minimum degree*, Advances in Mathematics **288** (2016), 337–385.

[2] M. Delcourt, C. Henderson, T. Lesgourges, and L. Postle, *Erdős meets Nash-Williams*, preprint arXiv:2507.23624 (2025).

[3] M. Delcourt, T. Kelly, and L. Postle, *Thresholds for $(n, q, 2)$-Steiner Systems via Refined Absorption*, preprint arXiv:2402.17858 (2024).

[4] M. Delcourt and L. Postle, *Progress towards Nash-Williams' Conjecture on triangle decompositions*, Journal of Combinatorial Theory, Series B **146** (2021), 382–416.

[5] M. Delcourt and L. Postle, *Finding an almost perfect matching in a hypergraph avoiding forbidden submatchings*, arXiv preprint arXiv:2204.08981 (2022).

[6] M. Delcourt and L. Postle, *Refined Absorption: A New Proof of the Existence Conjecture*, preprint arXiv:2402.17855 (2024).

[7] M. Delcourt and L. Postle, *Proof of the High Girth Existence Conjecture via Refined Absorption*, preprint arXiv:2402.17856 (2024).

[8] P. Erdős, *Problems and results in combinatorial analysis*, Colloq. Internat. Theor. Combin. Rome (1973), 3–17.

[9] S. Glock, F. Joos, J. Kim, M. Kühn, and L. Lichev, *Conflict-free hypergraph matchings*, Journal of the London Mathematical Society **109** 5 (2024).

[10] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption: Hypergraph F-designs for arbitrary F*, Memoirs of the American Mathematical Society **284** 1406 (2023).

[11] S. Glock, D. Kühn, and D. Osthus, *Extremal aspects of graph and hypergraph decomposition problems*, Surveys in Combinatorics, Cambridge University Press (2021), 235–266.

[12] P. Keevash, *The existence of designs*, preprint arXiv:1401.3665 (2014).

[13] T. Kirkman, *On a problem in combinatorics*, Cambridge Dublin Math. Journal **2** (1847), 191–204.

[14] M. Kwan, A. Sah, M. Sawhney, and M. Simkin, *High-girth Steiner triple systems*, Annals of Mathematics **200**, 3 (2024), 1059–1156.

[15] C. Nash-Williams, *An unsolved problem concerning decomposition of graphs into triangles*, Combinatorial Theory and its Applications **3** 1070 (1970), 1179–1183.

Nearly tight bounds for MaxCut in hypergraphs

Oliver Janzer
(joint work with Julien Portier)

Given a graph G, a *cut* refers to a partition of the vertex set into two subsets, with the *size* of the cut defined as the number of edges that have one vertex in each of these subsets. The MaxCut of a graph G is the maximum size of a cut over all possible bipartitions of the vertex set of G, and is denoted by $\mathrm{mc}(G)$. The MaxCut problem has been central in both combinatorics and theoretical computer science, and appears in Karp's famous list of 21 NP-complete problems. As a simple probabilistic argument shows that $\mathrm{mc}(G) \geq m/2$ for any graph G with m edges, the focus has been on showing bounds on the *surplus* $\mathrm{sp}(G)$ of a graph, defined as $\mathrm{sp}(G) = \mathrm{mc}(G) - m/2$. A celebrated result by Edwards [3] shows that $\mathrm{sp}(G) \geq (\sqrt{8m+1}-1)/8$, which is tight for a complete graph with an odd number of vertices. Very recently, an inverse theorem for MaxCut has been established by Balla, Hambardzumyan and Tomon [1], stating that if a graph G with m edges has $\mathrm{sp}(G) = O(\sqrt{m})$, then G contains a clique of order $\Omega(\sqrt{m})$. Addressing a question of Erdős and Lovász (see [4]), much work has been devoted to establishing stronger lower bounds for $\mathrm{sp}(G)$ under the assumption that G is H-free, for some fixed graph H.

A central problem in the area concerns the extension of Edwards' result to hypergraphs. An *r-cut* of a k-uniform hypergraph G is a partition of the vertex set of G into r parts, and the *size* of the cut is defined as the number of hyperedges having at least one vertex in each part of the cut. We are then interested in the maximum size of an r-cut over all possible partitions of the vertex set of G, which is denoted by $\mathrm{mc}_r(G)$. As shown by Erdős and Kleitman [5], by assigning each vertex to one of the r parts independently and uniformly at random, it follows that $\mathrm{mc}_r(G) \geq \frac{S(k,r)r!}{r^k}m$ for every graph G with m edges, where $S(k,r)$ is the Stirling number of second kind, denoting the number of unlabelled partitions of $\{1,\dots,k\}$ into r nonempty sets. The *r-surplus* (or *r-excess*) of G is then defined as $\mathrm{sp}_r(G) = \mathrm{mc}_r(G) - \frac{S(k,r)r!}{r^k}m$. Similarly, we say that the surplus of an r-cut in a k-graph with m edges is the size of the cut minus $\frac{S(k,r)r!}{r^k}m$. Conlon, Fox, Kwan and Sudakov [2] showed that for hypergraphs, Edwards' bound can be significantly improved.

Theorem 1 (Conlon–Fox–Kwan–Sudakov [2])**.** *For every $2 \leq r \leq k$ with $(r,k) \neq (2,2)$ and $(r,k) \neq (2,3)$, every k-uniform hypergraph with m edges has an r-cut of surplus $\Omega(m^{5/9})$.*

They noted that for $(r,k) = (2,3)$, Steiner triple systems show that there are 3-uniform hypergraphs on m edges with 2-surplus $\Theta(\sqrt{m})$. They furthermore observed that with high probability the binomial random k-graph $G_k(n, n^{3-k})$ has r-surplus $O(m^{2/3})$, where m is the number of its edges. This disproved a conjecture of Scott [7], which had predicted that the complete k-graph has the smallest 2-surplus.

Conlon et al. [2] conjectured that this random hypergraph is asymptotically optimal whenever $(r,k) \neq (2,2),(2,3)$.

Conjecture 2 (Conlon–Fox–Kwan–Sudakov [2]). *For any $2 \leq r \leq k$ with either $k \geq 4$ or $r \geq 3$, every k-uniform hypergraph with m edges has an r-cut of surplus $\Omega(m^{2/3})$.*

In the particular case where $r \in \{k-1,k\}$, Räty and Tomon [6] recently improved Conlon, Fox, Kwan and Sudakov's bound to $\Omega(m^{3/5-o(1)})$ by using spectral techniques.

Theorem 3 (Räty–Tomon [6]). *If $r \in \{k-1,k\}$ and $r \geq 3$, then every k-uniform hypergraph with m edges has an r-cut of surplus $\Omega(m^{3/5-o(1)})$.*

Räty and Tomon pointed out that their methods do not extend to the range $2 \leq r \leq k-2$. Obtaining strong results in the case $r=2$ is of particular interest for several reasons. Firstly, it is this case that has been the most extensively studied in Theoretical Computer Science. This problem is known as *Max E_k-Set Splitting*, or, in the case the hypegraph is not uniform, as *Max Set Splitting*. Secondly, this case has close connections to other well-studied problems such as negative (or positive) discrepancy, bisection width and hypergraph colourings.

Using a novel approach, we prove the following approximate version of the Conlon–Fox–Kwan–Sudakov conjecture for all values of r.

Theorem 4. *For any $\varepsilon > 0$, there exists some $k_0 = k_0(\varepsilon)$ such that for any $k > k_0$ and $2 \leq r \leq k$, every k-uniform hypergraph with m edges has an r-cut of surplus $\Omega(m^{2/3-\varepsilon})$.*

In the case of linear hypergraphs, still for $r \in \{k-1,k\}$, Räty and Tomon [6] proved a stronger bound $\Omega(m^{3/4-o(1)})$. They noted that this is asymptotically optimal, by constructing k-graphs with r-surplus $O(m^{3/4})$ in which each pair of vertices is in at most $O(\log m)$ edges (i.e., the hypergraphs are nearly linear). They conjectured that the $o(1)$ term can be removed and that the result can be extended to all $2 \leq r \leq k$.

We completely resolve this conjecture.

Theorem 5. *For any $2 \leq r \leq k$ with either $k \geq 4$ or $r \geq 3$, every k-uniform linear hypergraph with m edges has an r-cut of surplus $\Omega(m^{3/4})$.*

We also show that this is tight up to a constant factor by constructing linear (not just nearly linear) k-graphs with r-surplus $O(m^{3/4})$.

Proposition 6. *Fix some integers $2 \leq r \leq k$. There is a constant $\alpha = \alpha(k)$ such that, for every $m \in \mathbb{N}$, there exists an m-edge k-uniform linear hypergraph with r-surplus at most $\alpha m^{3/4}$.*

References

[1] I. Balla, L. Hambardzumyan, and I. Tomon, *Factorization norms and an inverse theorem for MaxCut*, arXiv preprint arXiv:2506.23989, 2025.
[2] D. Conlon, J. Fox, M. Kwan, and B. Sudakov, *Hypergraph cuts above the average*, Israel Journal of Mathematics, **233** (2019), 67–111.
[3] C. S. Edwards, *An improved lower bound for the number of edges in a largest bipartite subgraph*, In Proc. 2nd Czechoslovak Symposium on Graph Theory, Prague, pages 167–181, 1975.
[4] P. Erdős, *Problems and results in graph theory and combinatorial analysis*, Proc. British Combinatorial Conf., 5th, pages 169–192, 1975.
[5] P. Erdős and D. J. Kleitman, *On coloring graphs to maximize the proportion of multicolored k-edges*, Journal of Combinatorial Theory, **5** (1968), 164–169.
[6] E. Räty and I. Tomon, *Large cuts in hypergraphs via energy*, Mathematical Proceedings of the Cambridge Philosophical Society (2025), 1–17.
[7] A. D. Scott, *Judicious partitions and related problems*, In Surveys in Combinatorics 2005, pages 95–117.

Detecting planted trees in sparse random graphs

Nina Kamčev
(joint work with Nicolas Broutin, Gábor Lugosi, Bruce Reed, and Liana Yepremyan)

Detecting structured *anomalies* in large random systems is a fundamental statistical problem with applications in epidemiology, cybersecurity, and information networks. In this setting, one observes a large graph and wishes to decide whether the observed structure is consistent with a baseline random model (the *null model*) or whether it contains a small, atypical pattern. Examples include identifying an early-stage epidemic spread, locating a cyberattack, or detecting the emergence of a signalling motif in a biological network. These questions can be formalised as hypothesis testing problems on random graphs: under the *null hypothesis* the graph is drawn from a random model (in our case the Erdős–Rényi graph $G(n,p)$), while under the *alternative* a copy of a smaller template graph is embedded at a random location, and the remaining edges are generated according to $G(n,p)$.

This framework dates back to the planted clique problem introduced by Jerrum [8] and Kučera [9], and subsequently developed in works such as [1,3,4,6]. Algorithmic clique detection is a major open problem in theoretical computer science and combinatorics [1]. Related detection and recovery questions have been studied for other types of subgraphs and random models, for example in [2,5,7,10–12]. In this note we focus on the case when the planted subgraph is a tree. Our aim is to understand when it is statistically possible to detect the presence of a planted tree inside a sparse random graph.

Formally, let T be a tree on k vertices and consider the random graph $G(n,p)$ on vertex set $[n]$, where each edge is included independently with probability p. Under the null hypothesis, the observation is $G \sim G(n,p)$. Under the alternative, a uniformly random embedding of T is first selected and its edges are added to the graph, while all other potential edges appear independently with probability

p. The resulting graph is denoted $G_p(T) = G(n,p) \cup T$. A *test* observes G and must decide whether the instance was drawn from the null or the alternative. We say that *detection is possible* if there exists a test that succeeds with probability $1 - o(1)$ as $n \to \infty$. Throughout we focus on the sparse regime $p = c/n$, with $c > 0$ fixed, and in particular on the supercritical regime $c > 1$, where $G(n, c/n)$ contains a giant component of linear size, as well as large 'typical' trees.

The detectability of a planted tree depends crucially on its structure. When T is a star on k vertices, detection is immediate for all $c > 0$ once $k \gg \log n / \log\log n$, using the maximum degree in a typical random graph. At the opposite extreme, when T is a path on k vertices, detection fails for any fixed $c > 1$ as long as $k = o(\sqrt{n})$, as shown by Massoulié, Stephan, and Towsley [10]. Between these two cases lie the "typical" trees – uniformly random labelled trees on k vertices – whose detection behaviour does not seem to be governed by any single local feature of T. How small can such a planted tree T can be while still leaving a detectable property in $G(n,p) \cup T$?

Our results reveal a dichotomy depending on the mean degree c.

Theorem 1 (Possibility of detecting most known trees). *Let $p = c/n$. There is a constant C such that the following holds. Let T be a uniform random labelled tree on k vertices. For most trees T, the following holds.*

(a) If $c < 1.2$ and $k > C \log n$, then there is a test which distinguishes $G^p(T)$ from $G(n,p)$ with high probability.
(b) For $c > 500$ and $k = o(\sqrt{n})$, every test fails with probability $\frac{1}{2} - o(1)$ (even when the structure of T is revealed in advance).

Both bounds on k are optimal up to constant factors.

We also establish analogous results for the case in which the planted tree itself (denoted $\mathbf{T}$) is chosen uniformly at random and its shape is *not* revealed to the detector. In this more challenging setting, detection is still possible for $pn < 1.1$ when $k \geq C \log^2 n$. The same impossibility regime (b) applies for $c > 500$.

The surprising aspect of the positive results is that detection remains possible even in the supercritical regime, where $G(n,p)$ contains random trees of size at least $\Omega(n^{1/3})$. The reason lies in subtle changes in local expansion properties caused by the planting. A typical random tree contains many "hairy" paths – paths whose vertices have an unusually large boundary in $G(n,p) \cup T$ (or $G(n,p) \cup \mathbf{T}$)– and the presence of such paths can be used as a test distinguishing it from $G(n,p)$ with probability $1 - o(1)$.

It could also come as a surprise that the size of the smallest detectable tree transitions from the order of magnitude $\log^2 n$ to $\sqrt{n}$ as the mean degree pn increases, and understanding this transition remains an intriguing open problem. We also show a similar transition for the detection of a fixed d-ary tree.

While our emphasis here is on detection, a refinement of these ideas also yields consequences for partial reconstruction.

Open problems. We point out only a few of the many open problems remaining in the area. Rather informally, given a class of k-vertex trees $\mathcal{T}_k$, define the *detection threshold* as 'the smallest' $k = k(n, p)$ such that $G_p(T)$ can be distinguished from $G(n, p)$ for $T \in \mathcal{T}_k$, with probability $1 - o(1)$.

- For $pn = 1 + \epsilon(n)$ and $\epsilon = o(1)$ appropriately parametrised, what is the detection threshold of a path?
- What is the maximum value of pn for which the detection threshold of typical random trees is polylog(n)? What about d-ary trees?
- For $pn < 1.1$, is it possible to detect an *unknown* random tree of size $O(\log n)$?
- Can our method (the *hairy-path test*) be performed in polynomial time?

More broadly, it would be interesting to investigate alternative models in which edges are removed from $G(n, p)$ according to a specific distribution (see, e.g., [5]).

References

[1] N. Alon, M. Krivelevich, and B. Sudakov, *Finding a large hidden clique in a random graph*, *Random Structures & Algorithms* **13** (1999), 457–466.

[2] E. Arias-Castro and N. Verzelen, *Community detection in dense random networks*, *Annals of Statistics* **42** (2014), no. 3, 940–969.

[3] Y. Dekel, O. Gurel-Gurevich, and Y. Peres, *Finding hidden cliques in linear time with high probability*, *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics*, SIAM (2011), 67–75.

[4] Y. Deshpande and A. Montanari, *Finding hidden cliques of size $\sqrt{N/e}$ in nearly linear time*, *Foundations of Computational Mathematics* **15** (2015), no. 4, 1069–1128.

[5] D. Elimelech and W. Huleihel, *Detecting arbitrary planted subgraphs in random graphs*, *arXiv preprint arXiv:2503.19069*, 2025.

[6] U. Feige and D. Ron, *Finding hidden cliques in linear time*, *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*, Discrete Mathematics and Theoretical Computer Science (2010), 189–204.

[7] B. Hajek, Y. Wu, and J. Xu, *Computational lower bounds for community detection on random graphs*, *Conference on Learning Theory*, PMLR (2015), 899–928.

[8] M. Jerrum, *Large cliques elude the Metropolis process*, *Random Structures & Algorithms* **3** (1992), no. 4.

[9] L. Kučera, *Expected complexity of graph partitioning problems*, *Discrete Applied Mathematics* **57** (1995), no. 2, 193–212.

[10] L. Massoulié, L. Stephan, and D. Towsley, *Planting trees in graphs, and finding them back*, *Conference on Learning Theory*, PMLR (2019), 2341–2371.

[11] M. Moharrami, C. Moore, and J. Xu, *The planted matching problem: Phase transitions and exact results*, *Annals of Applied Probability* **31** (2021), no. 6, 2663–2720.

[12] A. Rotenberg, W. Huleihel, and O. Shayevitz, *Planted bipartite graph detection*, *IEEE Transactions on Information Theory*, 2024.

A short proof of the existence of designs

Peter Keevash

The existence of designs was one of the oldest problems in combinatorics, studied by many 19th century mathematicians, including (perhaps in chronological order) Plücker, Kirkman, Sylvester, Woolhouse, Cayley and Steiner. Steiner's name became attached to the fundamental object of interest, namely a *Steiner system* with parameters (n, q, r), which is a set S of q-subsets of an n-set X such that every r-subset of X is contained in exactly one element of S. There are some natural necessary divisibility conditions (discussed below) on n in terms of q and r for the existence of such a system S. The existence conjecture states that these conditions suffice for $n > n_0(q, r)$ large enough in terms of q and r. There are now three quite different proofs of this conjecture, chronologically [1–3]; we refer the reader to these papers for more history of the problem and references to the large associated literature (the solution of the conjecture has inspired several other dramatic breakthroughs in design theory). We give yet another proof, which has the advantages of being much shorter and giving a reasonable bound for n_0.

To state the result, we adopt the formulation of hypergraph decompositions. We denote the complete r-graph on a set S by $\binom{S}{r}$: this is a hypergraph with vertex set S and edge set consisting of all r-subsets of S. We also write $K_n^r = \binom{[n]}{r}$, where $[n] := \{1, \dots, n\}$. We identify hypergraphs with their edge sets, so $|H|$ counts edges (we let v_H count vertices). A Steiner system with parameters (n, q, r) is equivalent to a K_q^r-decomposition of K_n^r, that is, a partition D of (the edge set of) K_n^r into copies of K_q^r. We can also think of D as a perfect matching (partition of the vertex set into edges) in the design hypergraph $K_q^r(K_n^r)$, where for an r-graph G we write $K_q^r(G)$ for the $\binom{q}{r}$-graph H with $V(H) = G$ (vertices of H are edges of G) and edges consisting of all (edge sets of) copies of K_q^r in G. The necessary divisibility conditions mentioned above appear in the integral relaxation, where for $\Phi \in \mathbb{Z}^{K_q^r(G)}$ we define $\partial\Phi \in \mathbb{Z}^G$ by $(\partial\Phi)_e = \sum\{\Phi_Q : e \in Q\}$: if $\partial\Phi = G' \subseteq G$ we call Φ an integral K_q^r-decomposition of G', noting that if Φ is $\{0, 1\}$-valued then we can identify Φ with a (true) K_q^r-decomposition of G'. For $G \subseteq K_n^r$ we say that G is K_q^r-divisible if $\partial\Phi = G$ for some $\Phi \in \mathbb{Z}^{K_q^r(K_n^r)}$; thus K_q^r-divisibility is a necessary condition for having a K_q^r-decomposition.

Theorem 1. *For all $q > r \geq 1$ there is n_0 so that if $n \geq n_0$ and K_n^r is K_q^r-divisible then K_n^r has a K_q^r-decomposition.*

We fix $q > r \geq 1$, $\rho = (6\binom{q}{r})^{-2}$, $\alpha = (2q)^{-r}\rho$ and $n > n_0 = (4q)^{90q/\alpha}$.

1. Proof overview

Our strategy for proving Theorem 1 proceeds via the following five steps. In broad outline, as in all previous proofs, it is an absorption strategy, following the framework of [1], with simplifications to remove extra properties that are not needed for our current purpose.

1. Randomly reserve a sparse subgraph $R \subseteq K_n^r$, to be used in Step 4.
2. Find an 'absorber' $A \subseteq K_n^r \setminus R$, to be used in Step 5.
3. Find a 'regularity boosted' set of q-cliques $H \subseteq K_q^r(G)$, where $G := K_n^r \setminus (A \cup R)$.
4. Find edge-disjoint cliques $D \subseteq H$ with leave $L := G \setminus \bigcup D \subseteq R$.
5. Find clique decompositions D_L of $A \cup L$, so $D \cup D_L$ of G.

Here Step 1 is a simple Chernoff bound, Step 3 is a case of [2, Lemma 6.3], Step 4 is achieved by well-known random greedy arguments (the Rödl nibble) and Step 5 follows easily from the previous steps. The key to the proof is the second step, formulated in the following lemma, in which we see the meaning of 'absorber' (also called 'omni-absorber' in [1]): it is a subgraph A disjoint from R such that $A \cup L$ has a K_q^r-decomposition for any K_q^r-divisible $L \subseteq R$. The main novelty of our new approach is in finding a relatively simple proof of this lemma.

Lemma 2. *(Absorber) There is some $n^{-\alpha/4}$-bounded K_q^r-divisible $A \subseteq K_n^r \setminus R$ that is an absorber for R, that is, $A \cup L$ has a K_q^r-decomposition for any K_q^r-divisible $L \subseteq R$.*

Our basic gadget for various clique exchange operations in our construction consists of an r-graph Ω with two K_q^r decompositions Υ^+ and Υ^-. We have a designated clique $\hat{Q}^+ \in \Upsilon^+$ such that the cliques $\hat{Q}^e \in \Upsilon^-$ such that the cliques $\hat{Q}^e \in \Upsilon^-$ that share an edge with $\hat{Q}^+$ are 'maximally disjoint'.

Our first application of this gadget (in the Splitting step 3 below) is to modify some integral decomposition $\Phi \in \mathbb{Z}^{K_q^r(K_n^r)}$ without changing the value of $\partial\Phi$ so as to eliminate a copy of some clique Q. To do so, we fix any copy $\phi(\Omega)$ of Ω with $\phi(\hat{Q}^+) = Q$ and define $\Phi' \in \mathbb{Z}^{K_q^r(K_n^r)}$ by $\Phi' = \Phi + \phi(\Upsilon^-) - \phi(\Upsilon^+)$, meaning that each $\Phi'_{Q'}$ is $\Phi_{Q'} + 1$ for $Q' \in \phi(\Upsilon^-)$, or is $\Phi_{Q'} - 1$ for $Q' \in \phi(\Upsilon^+)$, or is $\Phi_{Q'}$ otherwise. Then $\partial\Phi' = \partial\Phi$, as $\phi(\Upsilon^+)$ and $\phi(\Upsilon^-)$ cover the same set of edges, so their contributions cancel. Furthermore, as $\hat{Q}^+ \in \Upsilon^+$ we have $\Phi'_Q = \Phi_Q - 1$; if $\Phi_Q > 0$ we interpret this as removing a copy of Q from Φ and replacing it by an equivalent set of signed cliques. Similarly, if $\Phi_Q < 0$ then we consider $\Phi' = \Phi - \phi(\Upsilon^-) + \phi(\Upsilon^+)$, which we interpret as removing a negative copy of Q from Φ and replacing it by an equivalent set of signed cliques.

For our second application of the gadget (in the Elimination steps 4 and 5 below) we rename some $\hat{Q}^{e_0}$ as $\hat{Q}^-$, thus focussing on two cliques $\hat{Q}^\pm \in \Upsilon^\pm$ with $\hat{Q}^+ \cap \hat{Q}^- = \{e_0\}$, which we use to modify some integral decomposition $\Phi \in \mathbb{Z}^{K_q^r(K_n^r)}$ without changing the value of $\partial\Phi$ so as to eliminate a 'cancelling pair' $Q^+ - Q^-$, consisting of two cliques of opposite sign in Φ that intersect exactly in some edge e. To do so, we fix any copy $\phi(\Omega)$ of Ω with $\phi(\hat{Q}^\pm) = Q^\pm$ and replace Φ by $\Phi' = \Phi + \phi(\Upsilon^-) - \phi(\Upsilon^+)$. Then $\partial\Phi' = \partial\Phi$ is unchanged, $\Phi'_{Q^+} = \Phi_{Q^+} - 1$ and $\Phi'_{Q^-} = \Phi_{Q^-} + 1$. If $\Phi_{Q^+} > 0$ and $\Phi_{Q^-} < 0$ then the interpretation is that we have removed a positive copy of Q^+ and a negative copy of Q^- and replaced them by an equivalent set of signed cli

Assuming the existence of an integral absorber (in step 1 below), the construction of the absorber proceeds as follows.

1. (Integral absorber) Find $\mathcal{Q}_1 \subseteq K_q^r(K_n^r)$ such that $\partial\mathcal{Q}_1$ is $n^{-\alpha/2}$-bounded, every edge of K_n^r is in at most two cliques of $\mathcal{Q}_1$, and for any K_q^r-divisible $J \in \mathbb{Z}^{K_n^r}$ supported in R there is $\Phi \in \mathbb{Z}^{\mathcal{Q}_1}$ with $\partial\Phi = J$.

2. (Local decoders) Apply a random greedy algorithm to choose $(q+r)$-sets Z_e containing e for each $e \in \bigcup \mathcal{Q}_1$, so that each r-graph $\binom{Z_e}{r} \setminus \{e\}$ is disjoint from all others and from $\bigcup \mathcal{Q}_1$. Let $\mathcal{Q}_2 = \bigcup_{e \in \bigcup \mathcal{Q}_1} K^r_q(\binom{Z_e}{r})$.

3. (Splitting) Let $Q_1, \dots, Q_t$ be a sequence in $K^r_q(K^r_n)$ consisting of $2^{q+1}r!$ copies of each q-clique in $\mathcal{Q}_1 \cup \mathcal{Q}_2$, where $2^q r!$ are labelled $+$ and $2^q r!$ are labelled $-$. Apply a random greedy algorithm to choose copies $\Omega_i = \phi_i(\Omega)$ of Ω with $\phi_i(\hat{Q}^+) = Q_i$, where each $\Omega_i \setminus Q_i$ is edge-disjoint from all others and from $A_0 := \bigcup(\mathcal{Q}_1 \cup \mathcal{Q}_2)$, and $V(\Omega_i) \setminus V(Q_i)$ is disjoint from $V(\Omega_j) \setminus V(Q_j)$ whenever Q_i and Q_j share an edge.

4. (Elimination) For each $i \in [t]$ and $Q' \in \Upsilon^- \cup \Upsilon^+ \setminus \{\hat{Q}^+\}$ we call $\phi_i(Q')$ a splitting clique, with the same sign as Q_i if $Q' \in \Upsilon^-$ or the opposite sign if $Q' \in \Upsilon^+$; if Q' shares an edge with $\hat{Q}^+$ we call $\phi_i(Q')$ near, otherwise it is far. Let $(Q^-_1, Q^+_1), \dots, (Q^-_{t'}, Q^+_{t'})$ be a sequence consisting of all pairs of oppositely signed near cliques with a common edge. Apply a random greedy algorithm to choose copies $\Omega'_i = \phi'_i(\Omega)$ with $\phi'_i(\hat{Q}^\pm) = Q^\pm_i$ where each $\Omega'_i \setminus (Q^-_i \cup Q^+_i)$ is edge-disjoint from all others and from all previously chosen cliques.

5. (Further Elimination) For each $i \in [t']$ and $Q' \in \Upsilon^\pm \setminus \{\hat{Q}^-, \hat{Q}^+\}$ we call $\phi'_i(Q')$ an elimination clique with the opposite sign to Q'. We call a negative elimination clique bad (for e) if it shares some edge e with some negative near clique. Let $(Q'^-_1, Q'^+_1), \dots, (Q'^-_{t''}, Q'^+_{t''})$ be a sequence consisting of all pairs where each Q'^-_i is a negative elimination clique that is bad for some e_i and Q'^+_i is the positive splitting clique containing e_i (which is unique and far). Apply a random greedy algorithm to choose copies $\Omega''_i = \phi''_i(\Omega)$ with $\phi''_i(\hat{Q}^\pm) = Q'^\pm_i$ where each $\Omega''_i \setminus (Q'^-_i \cup Q'^+_i)$ is edge-disjoint from all others and from all previously chosen cliques.

6. (Conclusion) For each Q'^-_i as in Step 5 and each $Q' \in \Upsilon^\pm \setminus \{\hat{Q}^-, \hat{Q}^+\}$ we call $\phi''_i(Q')$ a further elimination clique of the opposite sign to Q'. We define $\mathcal{Q} = \mathcal{Q}^+ \cup \mathcal{Q}^-$, where $\mathcal{Q}^+$ contains all positive splitting cliques, positive elimination cliques and positive further elimination cliques, and $\mathcal{Q}^-$ contains all negative far splitting cliques, negative good elimination cliques and negative further elimination cliques. We define $A := \bigcup \mathcal{Q}^-$.

This completes the construction of the absorber, modulo various ingredients that have appeared in [1–3].

References

[1] M. Delcourt and L. Postle, Refined absorption: a new proof of the existence conjecture, arXiv:2402.17855.

[2] S. Glock, D. Kühn, A. Lo and D. Osthus, The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F, arXiv:1611.06827 and arXiv:1706.01800, *Mem. Amer. Math. Soc.* 284 (2023).

[3] P. Keevash, The existence of designs, arXiv:1401.3665.

The Hypergraph Removal Process

Marcus Kühn
(joint work with Felix Joos)

Let $\mathcal{F}$ be a k-uniform hypergraph where $k \geq 2$. Starting with a complete k-uniform hypergraph on n vertices, the *$\mathcal{F}$-removal process* iteratively removes all edges of a copy of $\mathcal{F}$ chosen uniformly at random among all remaining copies of $\mathcal{F}$ until no copies are left. We use $R_n(\mathcal{F})$ to denote the (random) final number of edges present after termination. Besides generating hypergraphs without copies of $\mathcal{F}$, the $\mathcal{F}$-removal process also yields maximal packings of edge-disjoint copies of $\mathcal{F}$ and is perhaps the most natural way to approach the problem of generating hypergraph packings.

It turns out that merely getting close to the order of magnitude of $R_n(K_3)$ is challenging. As a first step, Spencer [6] as well as Rödl and Thoma [5] proved that typically, that is with high probability (with probability tending to 1 as $n \to \infty$), $R_n(K_3) = o(n^2)$ holds. Grable [4] improved this to $R_n(K_3) \leq n^{11/6+o(1)}$ and described how to obtain $R_n(K_3) \leq n^{7/4+o(1)}$. Following these attempts to determine $R_n(K_3)$, Spencer conjectured that typically $R_n(K_3) = n^{3/2\pm o(1)}$ holds and offered \$200 for a resolution [4, 7]. Eventually in 2015, Bohman, Frieze and Lubetzky [2] famously confirmed this conjecture. However, since then there have been neither results improving on their estimates nor any results that establish the correct order of magitude of $R_n(\mathcal{F})$ for any nontrivial case beyond $\mathcal{F} = K_3$. In fact, obtaining asymptotic estimates for $R_n(K_4)$ was considered a central open problem in the area.

Relying on the same heuristic as for the triangle, Bennett and Bohman [1] state the following more general "folklore" conjecture predicting $R_n(\mathcal{F})$ whenever $\mathcal{F}$ is the k-uniform complete hypergraph $K_\ell^{(k)}$ on ℓ vertices.

Conjecture 1 ([1, Conjecture 1.2])**.** *Let $2 \leq k < \ell$. Then, with high probability,*

$$n^{k-\frac{\ell-k}{\binom{\ell}{k}-1}-o(1)} \leq R_n(K_\ell^{(k)}) \leq n^{k-\frac{\ell-k}{\binom{\ell}{k}-1}+o(1)}.$$

Our main result confirms this conjecture. In fact, we prove a significantly stronger result. Using $v(\mathcal{F})$ to denote the number of vertices and $e(\mathcal{F})$ to denote the number of edges of $\mathcal{F}$, the *k-density* of $\mathcal{F}$ is $\rho_{\mathcal{F}} := (e(\mathcal{F}) - 1)/(v(\mathcal{F}) - k)$ if $v(\mathcal{F}) \geq k + 1$. As in [3], we say that $\mathcal{F}$ is *strictly k-balanced* if $\mathcal{F}$ has at least three edges and satisfies $\rho_{\mathcal{G}} < \rho_{\mathcal{F}}$ for all proper subgraphs $\mathcal{G}$ of $\mathcal{F}$ that have at least two edges. Note that $K_\ell^{(k)}$ is strictly k-balanced for all $2 \leq k < \ell$. The following is a corollary of our main result.

Theorem 2. *Let $k \geq 2$ and consider a strictly k-balanced k-uniform hypergraph $\mathcal{F}$ with k-density ρ. Then, for all $\varepsilon > 0$, there exists $n_0 \geq 0$ such that for all $n \geq n_0$, with probability at least $1 - \exp(-(\log n)^{5/4})$, we have*

$$n^{k-1/\rho-\varepsilon} \leq R_n(\mathcal{F}) \leq n^{k-1/\rho+\varepsilon}.$$

As part of our proof, we rely on a difference equation argument based on supermartingale concentration. Such arguments often require very carefully chosen collections of random variables and the approach of Bohman, Frieze and Lubetzky is no exception. While they find an ingenious way to explicitly describe the collection in the triangle case, the complexity of the underlying structures very quickly makes such explicit descriptions practically infeasible, which limits such an approach to this special case.

Our new approach crucially relies on circumventing the explicit choice of such a collection. We take what usually constitutes the heart of such an argument, namely explicitly finding a suitable collection of random variables, and replace it with an implicit definition which in many aspects works by design but consists of otherwise unknown structures. We are then left with proving abstract properties of these implicitly given structures that we then subsequently rely on in the proof. To give an example, for the analysis of the triangle case in [2], substructures called *fans* that essentially correspond to graphs that for some $\ell \geq 1$ consist of vertices $u, v_1, \ldots, v_\ell$ and the edges $\{u, v_i\}$ and $\{v_j, v_{j+1}\}$ where $1 \leq i \leq \ell$ and $1 \leq j \leq \ell - 1$ play a key role. In contrast, in our more general analysis, we instead work with maximizers of density-based optimization problems that we consider without concrete knowledge of their structure.

Observe that complete (hyper)graphs exhibit a very high degree of symmetry while most strictly k-balanced hypergraphs have locally and globally essentially no symmetries. This complicates the analysis further and requires us to dedicate substantial parts of the proof to dealing with the extension from cliques to general strictly k-balanced hypergraphs.

References

[1] P. Bennett and T. Bohman, *A natural barrier in random greedy hypergraph matching*, Combin. Probab. Comput. **28** (2019), 816–825.

[2] T. Bohman, A. Frieze, and E. Lubetzky, *Random triangle removal*, Adv. Math. **280** (2015), 379–438.

[3] T. Bohman and P. Keevash, *The early evolution of the H-free process*, Invent. Math. **181** (2010), 291–336.

[4] D. A. Grable, *On random greedy triangle packing*, Electron. J. Combin. **4** (1997), Research Paper 11, 19.

[5] V. Rödl and L. Thoma, *Asymptotic packing and the random greedy algorithm*, Random Structures Algorithms **8** (1996), 161–177.

[6] J. Spencer, *Asymptotic packing via a branching process*, Random Structures Algorithms **7** (1995), 167–172.

[7] N. C. Wormald, *The differential equation method for random graph processes and greedy algorithms*, In M. Karoński and H. Prömel, Eds., Lectures on Approximation and Randomized Algorithms (1999), 73–155.

Parities in random Latin squares

Matthew Kwan
(joint work with Kalina Petrova, Mehtaab Sawhney)

An $n \times n$ *Latin square* is an $n \times n$ array filled with n different "symbols" (usually taken to be the integers $1, \ldots, n$), with the property that each symbol appears exactly once in each row and each column. For example, the multiplication table of a group is always a Latin square; in general, Latin squares can be interpreted as multiplication tables of a class of algebraic structures called *quasigroups*. See for example [14] for an introduction to this vast subject.

Each row or column of an $n \times n$ Latin square L can be interpreted as a permutation of order n, which can be either even or odd. Let $N_{\mathrm{row}}(L)$ be the number of odd row permutations, and let $N_{\mathrm{col}}(L)$ be the number of odd column permutations. If L is the multiplication table of a group, then either $N_{\mathrm{row}}(L) = N_{\mathrm{col}}(L) = 0$ or $N_{\mathrm{row}}(L) = N_{\mathrm{col}}(L) = n/2$. However, for general Latin squares, the row and column parities can have much richer behaviour, and much is still unknown. For example, one of the most important conjectures in this direction is the *Alon–Tarsi conjecture*, which (in probabilistic language) says that if n is even, and $\mathbf{L}$ is a uniformly random $n \times n$ Latin square, then

$$\Pr\left[N_{\mathrm{row}}(\mathbf{L}) \text{ is even}\right] = \Pr\left[N_{\mathrm{row}}(\mathbf{L}) + N_{\mathrm{col}}(\mathbf{L}) \text{ is even}\right] \neq \frac{1}{2}.$$

(The first equality is not part of the original conjecture; it was observed by Huang and Rota [11], in a paper where they also observed that the Alon–Tarsi conjecture has a number of surprising consequences in seemingly unrelated areas of mathematics; see [10] for a modern survey). On the other hand, Alpöge [2] and independently Cavenagh and Wanless [8] proved that

$$\Pr\left[N_{\mathrm{row}}(\mathbf{L}) \text{ is even}\right] = \frac{1}{2} + o(1). \tag{1}$$

as $n \to \infty$. In other words, if the Alon–Tarsi conjecture is true, then it is true "just barely".

Going far beyond (1), it has been suggested by Peter Cameron (in a variety of different sources; see for example [4–7]) that the row parities of a Latin square might be statistically completely unconstrained, in the sense that one can model the n row parities of a random Latin square by simply making n independent coin flips.

Conjecture 1. *Let* $\mathbf{L}$ *be a uniformly random* $n \times n$ *Latin square. Then the distribution of* $N_{\mathrm{row}}(\mathbf{L})$ *is approximately the binomial distribution* $\mathrm{Bin}(n, 1/2)$*, as* $n \to \infty$.

(Note that exchanging rows does not affect the distribution of $\mathbf{L}$, so the sequence of row parities $\vec{\xi}_{\mathrm{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ has a permutation-invariant distribution. This means that if we condition on $N_{\mathrm{row}}(\mathbf{L})$, then $\vec{\xi}_{\mathrm{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ is a uniformly random sequence in $(\mathbb{Z}/2\mathbb{Z})^n$, constrained to have exactly $N_{\mathrm{row}}(\mathbf{L})$ "1"s. That is to say, in order to understand the distribution of $\vec{\xi}_{\mathrm{row}}(\mathbf{L})$, it is enough to understand the distribution of $N_{\mathrm{row}}(\mathbf{L})$.)

To elaborate on the attribution/history of Conjecture 1: the starting point seems to have been the problem session at the *British Combinatorial Conference* in 1993, where Cameron asked a related question that was extended by Jeannette Janssen. However, at that time they did not seem to be very confident that the statement of Conjecture 1 was actually true (in the BCC problem list [5] they phrase the question as "is it true that..."). Cameron later posed the problem more assertively in a 2002 survey on permutations and permutation groups [4], and it seems he first referred to it as a "conjecture" in a 2003 lecture on random Latin squares [6]. However, as far as we can tell, he never stated the problem in a fully precise form (always using language like "approximately").

There are many different ways to compare distributions, to make rigorous sense of the word "approximately". We are able to confirm Conjecture 1 in many different senses. For example, one strong way to compare distributions is in terms of *total variation distance*: for two probability distributions μ, ν on the same space, write $\mathrm{d_{TV}}(\mu,\nu) = \sup_A |\mu(A) - \nu(A)|$ (so for any event A, the probabilities of A with respect to μ and ν differ by at most $\mathrm{d_{TV}}(\mu,\nu)$). The following theorem is a consequence of our main (technical) result.

Theorem 2. *Let* $\mathbf{L}$ *be a uniformly random* $n \times n$ *Latin square. Then*

$$\lim_{n\to\infty} \mathrm{d_{TV}}\big(N_{\mathrm{row}}(\mathbf{L}),\, \mathrm{Bin}(n, 1/2)\big) = 0.$$

Equivalently, writing $\vec{\xi}_{\mathrm{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ *for the sequence of parities of rows of* $\mathbf{L}$*, and writing* $\mathrm{Unif}((\mathbb{Z}/2\mathbb{Z})^n)$ *for the uniform distribution on* $(\mathbb{Z}/2\mathbb{Z})^n$*, we have*

$$\lim_{n\to\infty} \mathrm{d_{TV}}\big(\vec{\xi}_{\mathrm{row}}(\mathbf{L}),\, \mathrm{Unif}((\mathbb{Z}/2\mathbb{Z})^n)\big) = 0.$$

To briefly discuss why this conjecture took so long to be resolved (and, in our opinion, why it is so interesting): generally speaking, it is quite easy to make plausible predictions about uniformly random Latin squares, by making various kinds of approximate independence assumptions (for example, Cameron's conjecture can be justified from the point of view that there is "no obvious reason" for the parities of different rows to be correlated). However, it is surprisingly difficult to rigorously prove anything nontrivial about uniformly random Latin squares, or even to study them empirically.

The main issues are that Latin squares do not enjoy any neat recursive structure, and they are very "rigid" objects, in the sense that there are only very limited ways to make a "local perturbation" to change a Latin square into another one. To highlight the difficulties here, we remark that (despite some very ambitious conjectures; see [1, Section 4.1]) we still have a rather poor understanding of the total number of $n \times n$ Latin squares (the best known upper and lower bounds differ by exponential factors; see [21, Section 17]), and there is no rigorously justified way to efficiently sample a uniformly random Latin square (there are certain ergodic Markov chains on the space of $n \times n$ Latin squares [12, 20], but these are not known to be rapidly mixing). By now, there are quite a few known results about random

Latin squares; the proofs of many of these results have required fundamental new additions to a very limited toolbox of techniques, and Theorem 2 is no exception. Two of the most important new ingredients are a new re-randomisation technique via "stable intercalate switchings" (building on ideas from [17–19]), and a new approximation theorem comparing random Latin squares with a certain independent model (building on ideas from [9, 15]).

References

[1] Y. Alimohammadi, P. Diaconis, M. Roghani, and A. Saberi, *Sequential importance sampling for estimating expectations over the space of perfect matchings*, Ann. Appl. Probab. **33** (2023), no. 2, 799–833.

[2] L. Alpoge, *Square-root cancellation for the signs of Latin squares*, Combinatorica **37** (2017), no. 2, 137–142.

[3] C. Bowtell and R. Montgomery, *Almost every Latin square has a decomposition into transversals*, arXiv:2501.05438.

[4] P. J. Cameron, *Permutations*, Paul Erdős and his mathematics, II (Budapest, 1999), Bolyai Soc. Math. Stud., vol. 11, János Bolyai Math. Soc., Budapest, 2002, pp. 205–239.

[5] P. J. Cameron, *BCC problem list*, `https://webspace.maths.qmul.ac.uk/p.j.cameron/bcc/allprobs.pdf`, 2001.

[6] P. J. Cameron, *Random Latin squares*, slides for a presentation at the RAND-APX meeting at Oxford in 2003, `https://webspace.maths.qmul.ac.uk/p.j.cameron/slides/rand-apx.pdf`, 2003.

[7] P. J. Cameron, *A niggling problem*, posted on Peter Cameron's blog, `https://cameroncounts.wordpress.com/2015/01/24/a-niggling-problem/`, 2015.

[8] N. J. Cavenagh and I. M. Wanless, *There are asymptotically the same number of Latin squares of each parity*, Bull. Aust. Math. Soc. **94** (2016), no. 2, 187–194.

[9] A. Ferber and M. Kwan, *Almost all Steiner triple systems are almost resolvable*, Forum of Mathematics, Sigma, vol. 8, Cambridge University Press, 2020, p. e39.

[10] B. Friedman and S. McGuinness, *The Alon-Tarsi conjecture: a perspective on the main results*, Discrete Math. **342** (2019), no. 8, 2234–2253.

[11] R. Huang and G.-C. Rota, *On the relations of various conjectures on Latin squares and straightening coefficients*, Discrete Math. **128** (1994), no. 1-3, 225–236.

[12] M. T. Jacobson and P. Matthews, *Generating uniformly distributed random Latin squares*, J. Combin. Des. **4** (1996), no. 6, 405–437.

[13] J. C. M. Janssen, *On even and odd Latin squares*, J. Combin. Theory Ser. A **69** (1995), no. 1, 173–181.

[14] A. D. Keedwell and J. Dénes, *Latin squares and their applications*, second ed., Elsevier/North-Holland, Amsterdam, 2015, With a foreword to the previous edition by Paul Erdös.

[15] M. Kwan, *Almost all Steiner triple systems have perfect matchings*, Proc. Lond. Math. Soc. (3) **121** (2020), no. 6, 1468–1495.

[16] M. Kwan, A. Sah, and M. Sawhney, *Note on random latin squares and the triangle removal process*, arXiv:2109.15201.

[17] M. Kwan, A. Sah, and M. Sawhney, *Large deviations in random Latin squares*, Bull. Lond. Math. Soc. **54** (2022), no. 4, 1420–1438.

[18] M. Kwan, A. Sah, M. Sawhney, and M. Simkin, *Substructures in Latin squares*, Israel J. Math. **256** (2023), no. 2, 363–416.

[19] M. Kwan and B. Sudakov, *Intercalates and discrepancy in random Latin squares*, Random Structures Algorithms **52** (2018), no. 2, 181–196.

[20] A. O. Pittenger, *Mappings of Latin squares*, Linear Algebra Appl. **261** (1997), 251–268.

[21] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, 2nd ed., Cambridge University Press, Cambridge, 2001.

Sums of algebraic dilates

Jeck Lim
(joint work with David Conlon)

1. Introduction

For any subset A of $\mathbb{C}$ and $\lambda_1, \dots, \lambda_k \in \mathbb{C}$, the sum of dilates $A+\lambda_1 \cdot A+\cdots+\lambda_k \cdot A$ is given by

$$A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A := \{a_0 + \lambda_1 a_1 + \cdots + \lambda_k a_k : a_0, a_1, \dots, a_k \in A\}.$$

Our concern here will be with estimating the minimum size of $|A+\lambda_1 \cdot A+\cdots+\lambda_k \cdot A|$ in terms of $|A|$. For $\lambda_1, \dots, \lambda_k \in \mathbb{Q}$, this problem was essentially solved by Bukh [3], from whose results it follows that if $\lambda_i = p_i/q$ for q as small as possible for such a common denominator, then

$$|A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A| \geq (|p_1| + \cdots + |p_k| + |q|)|A| - o(|A|)$$

for all finite subsets A of $\mathbb{C}$, which is best possible up to the lower-order term. This result was later sharpened by Balog and Shakan [1] when $k = 1$ and then Shakan [13] in the general case, improving the $o(|A|)$ term to a constant depending only on $\lambda_1, \dots, \lambda_k$.

When at least one of the λ_i is transcendental, it was shown by Konyagin and Łaba [7] that

$$|A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A| = \omega(|A|).$$

The problem of giving more precise lower bounds for $|A + \lambda \cdot A|$ when λ is transcendental was studied in some depth by Sanders [10, 11] and Schoen [12], with progress tied to advances in quantitative estimates for Freiman's theorem on sets of small doubling. Using quite different techniques, Conlon and Lim [5] recently resolved this problem, showing that there is a constant c such that

$$|A + \lambda \cdot A| \geq e^{c\sqrt{\log |A|}}|A|,$$

which, by a construction of Konyagin and Łaba, is best possible up to the value of c.

Our focus here will be on the complementary case, where each of $\lambda_1, \dots, \lambda_k$ is algebraic. Early results in this direction were proved by Breuillard and Green [2] and Chen and Fang [4], with the latter showing that, for any fixed $\lambda \geq 1$, $|A + \lambda \cdot A| \geq (1 + \lambda)|A| - o(|A|)$ for all finite subsets A of $\mathbb{R}$. The problem of giving more precise lower bounds for $|A + \lambda \cdot A|$ when λ is algebraic was raised explicitly by Shakan [13] and by Krachun and Petrov [8], with the latter authors conducting the first systematic study and making the first concrete conjectures.

To state their conjecture, suppose that $f(x) \in \mathbb{Z}[x]$ is the minimal polynomial of λ, assumed to have coprime coefficients, and $f(x) = \prod_{i=1}^{d}(a_i x + b_i)$ is a full complex factorisation of f. If we set $H(\lambda) := \prod_{i=1}^{d}(|a_i| + |b_i|)$, the conjecture of Krachun and Petrov [8] states that for any algebraic number λ,

$$|A + \lambda \cdot A| \geq H(\lambda)|A| - o(|A|)$$

for all finite subsets A of $\mathbb{C}$.

Krachun and Petrov [8] gave some evidence for their conjecture by proving it in the special case where $\lambda = \sqrt{2}$. Subsequently, as a consequence of their work [6] on a conjecture of Bukh regarding sums of linear transformations, Conlon and Lim verified the conjecture for all λ of the form $(p/q)^{1/d}$ with $p, q, d \in \mathbb{N}$. More recently, Krachun and Petrov [9] have revisited the problem, proving their conjecture in full whenever λ is an algebraic integer.

We prove their conjecture in full for all algebraic numbers. Our method also extends to longer sums of algebraic dilates, so we will state our results in that level of generality. Given algebraic numbers $\lambda_1, \ldots, \lambda_k$, recall that if the field extension $K := \mathbb{Q}(\lambda_1, \ldots, \lambda_k)$ of $\mathbb{Q}$ is of degree $d = \deg(K/\mathbb{Q})$, then there are exactly d different complex embeddings $\sigma_1, \ldots, \sigma_d : K \to \mathbb{C}$. We also need to define the *denominator ideal* (see, for example, [14]), which is the ideal in the ring of integers $\mathcal{O}_K$ given by

$$\mathfrak{D} := \{x \in \mathcal{O}_K \mid x\lambda_l \in \mathcal{O}_K \text{ for } l = 1, \ldots, k\}.$$

The key quantity $H(\lambda_1, \ldots, \lambda_k)$ that plays the role of $H(\lambda)$ for sums of many algebraic dilates is then

$$H(\lambda_1, \ldots, \lambda_k) := N_{K/\mathbb{Q}}(\mathfrak{D}) \prod_{i=1}^{d}(1 + |\sigma_i(\lambda_1)| + |\sigma_i(\lambda_2)| + \cdots + |\sigma_i(\lambda_k)|),$$

where $N_{K/\mathbb{Q}}(\mathfrak{D})$ is the ideal norm of $\mathfrak{D}$.

With this definition in place, our main result, which is best possible up to the behaviour of the lower-order term, is as follows.

Theorem 1. *For any algebraic numbers $\lambda_1, \ldots, \lambda_k$,*

$$|A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A| \geq H(\lambda_1, \ldots, \lambda_k)|A| - o(|A|)$$

for all finite subsets A of $\mathbb{C}$.

2. Freiman-type structure theorem

One of the main ingredients in the proof of Theorem 1 is a Freiman-type structure theorem on sets with small sums of dilates, that is, sets A with

$$|A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A| \leq C|A|.$$

Any such A also has small doubling, so, by the usual version of Freiman's theorem, it must be contained in a small generalised arithmetic progression (GAP). However, a GAP does not necessarily have a small sum of dilates. Our variant of Freiman's theorem, stated below, says instead that A is contained in what we call an $\mathcal{O}_K$-GAP, which shares with A the property that it has a small sum of dilates.

Theorem 2. *For every $C > 0$ and $p \in \mathbb{N}$, there are constants n and F such that for any $A \subset K$ satisfying*

$$|A + \lambda_1 \cdot A + \cdots + \lambda_k \cdot A| \leq C|A|,$$

there exists a p-proper $\mathcal{O}_K$-GAP $P \subset K$ containing A of dimension at most n and size at most $F|A|$.

3. Lattice densities

Another crucial ingredient is a high-dimensional generalisation of the notion of local densities, which we call "lattice densities", that allows one to represent a discrete set $A \subset \mathbb{Z}^d$ with a continuous set $\overline{A}$. More precisely, for a (periodic) set $A \subseteq \mathbb{Z}^d$ and a flag of lattices $\mathcal{F} = \{L_0 \subseteq L_1 \subseteq \cdots \subseteq L_k\}$, the *lattice density* $\mathrm{LD}(A; \mathcal{F})$ is a compact downset in $[0,1]^{k+1}$ which encodes information about the density of A relative to the lattices L_l. Our continuous representation $\overline{A}$ is then a compact subset of $\mathbb{R}^{d+k+1}$ with a base in $\mathbb{R}^d$ resembling A and fibres in $\mathbb{R}^{k+1}$ equal to the local lattice density at each point of A. This allows us to convert the discrete sums of dilates problem into an analogous continuous one, which is much easier to solve.

References

[1] A. Balog and G. Shakan, On the sum of dilations of a set, *Acta Arith.* **164** (2014), 153–162.

[2] E. Breuillard and B. Green, Contractions and expansion, *Eur. J. Combin.* **34** (2013), 1293–1296.

[3] B. Bukh, Sums of dilates, *Combin. Probab. Comput.* **17** (2008), 627–639.

[4] Y.-G. Chen and J.-H. Fang, Sums of dilates in the real numbers, *Acta Arith.* **182** (2018), 231–241.

[5] D. Conlon and J. Lim, Sums of transcendental dilates, *Bull. London Math. Soc.* **55** (2023), 2400–2406.

[6] D. Conlon and J. Lim, Sums of linear transformations, to appear in *Trans. Amer. Math. Soc.*

[7] S. Konyagin and I. Łaba, Distance sets of well-distributed planar sets for polygonal norms, *Israel J. Math.* **152** (2006), 157–179.

[8] D. Krachun and F. Petrov, On the size of $A + \lambda A$ for algebraic λ, *Mosc. J. Comb. Number Theory* **12** (2023), 117–126.

[9] D. Krachun and F. Petrov, Tight lower bound on $|A + \lambda A|$ for algebraic integer λ, preprint available at arXiv:2311.09399 [math.CO].

[10] T. Sanders, Appendix to "Roth's theorem on progressions revisited" by J. Bourgain, *J. Anal. Math.* **104** (2008), 193–206.

[11] T. Sanders, On the Bogolyubov–Ruzsa lemma, *Anal. PDE* **5** (2012), 627–655.

[12] T. Schoen, Near optimal bounds in Freiman's theorem, *Duke Math. J.* **158** (2011), 1–12.

[13] G. Shakan, Sum of many dilates, *Combin. Probab. Comput.* **25** (2016), 460–469.

[14] D. Singhal and Y. Lin, Primes in denominators of algebraic numbers, *Int. J. Number Theory* **20** (2024), 327–348.

Improved bounds for the minimum degree of minimal multicolor Ramsey graphs

Sam Mattheus
(joint work with Yamaan Attwa, Tibor Szabó, Jacques Verstraete)

We say that a graph G is *r-Ramsey* for a graph H, denoted by $G \to (H)_r$ if every r-coloring of the edges of G contains a monochromatic copy of H. A graph G is called *r-Ramsey-minimal for H* if it is r-Ramsey for H, but no proper subgraph of it is. The set of all r-Ramsey-minimal graphs for H is denoted by $\mathcal{M}_r(H)$. The classical Ramsey number $R_r(H)$, one of the most well-studied parameters in Combinatorics, is then the smallest number of vertices of a graph in $\mathcal{M}_r(H)$. Following the pioneering work of Folkman [5] on the smallest clique number of Ramsey graphs for the clique, Burr, Erdős and Lovász [3] in 1976 initiated the systematic study of the extremal behaviour of several other graph parameters. In their seminal paper they investigated the chromatic number, the maximum and the minimum degree, and the connectivity.

In this paper we will be particularly interested in the minimum degree of minimal Ramsey graphs. For a graph H and number r of colors we define

$$s_r(H) := \min\{\delta(G) \mid G \in \mathcal{M}_r(H)\},$$

to be the smallest possible minimum degree that could occur among minimal r-Ramsey graphs for H. For cliques in the classical two-color case, Burr et al. [3] established the following precise result:

$$s_2(K_k) = (k-1)^2. \tag{1}$$

Upon first glance, this result looks extremely surprising. First, it determines the *exact* value of the smallest possible minimum degree in a minimal 2-Ramsey graph for K_k. This is in stark contrast with our knowledge about the smallest *number of vertices* in such graphs, which is hopelessly out of reach. Furthermore, the value of the smallest minimum degree turns out to be just a quadratic function of k. This is incredibly small considering that we know that even the smallest of Ramsey graphs will have exponentially many vertices. How could it then be possible to create a (necessarily enormous) 2-Ramsey graph for K_k, that has a vertex with just $(k-1)^2$ neighbors, such that the presence of this vertex, and in fact any of its incident edges are *crucial* in guaranteeing the 2-Ramsey-ness of said enormous graph?

Fox, Grinshpun, Liebenau, Person and Szabó [6] investigated the behaviour $s_r(K_k)$ for more than two colors. They found that for any fixed clique order $k \geq 3$ there exist positive constants c_k, C_k, such that

$$c_k r^2 \frac{\log r}{\log \log r} \leq s_r(K_k) \leq C_k r^2 (\ln r)^{8k^2}. \tag{2}$$

For the triangle K_3, a slightly stronger lower bound was given in [6], which was proved to be tight up to a constant factor by Guo and Warnke [7].

$$s_r(K_3) = \Theta(r^2 \log r). \tag{3}$$

These results establish that for any fixed clique order k the value of the smallest minimum degree $s_r(K_k)$ is quadratic in the number r of colors, up to some logarithmic factor. The power of the logarithm in the gap between the upper and lower bounds however depends significantly on k.

On the other end of the spectrum, when the number r of colors is constant, Hàn, Rödl, and Szabó [8] determined the order of magnitude of the smallest minimum degree $s_r(k)$, up to a logarithmic factor. More generally, they have shown that there exists a constant C such that for every $k^2 > r$

$$s_r(K_k) \leq Cr^3k^2 \log^3 r \log^2 k. \tag{4}$$

Considering that we know from [6] and [3] that $s_r(K_k) \geq s_2(K_k) = (k-1)^2$, the bound (4) establishes that $s_r(K_k)$ is quadratic up to a $\log^2$-factor for any fixed number r of colors.

When both k and r are increasing, say $r = r(k)$ is a decent increasing function of k, the best known upper bounds vary depending on how fast r grows. In the range $r(k) < k^2$, the upper bound of (4) is the best we know.

In the complementary range of $r(k) \geq k^2$ another construction of Fox et al. [6] gives[1]

$$s_r(K_k) = O(r^3k^3 \log^3 k). \tag{5}$$

Bamberg, Bishnoi, and Lesgourgues [1] developed a generalization of this construction and used that to obtain

$$s_r(K_k) = O(r^{5/2}k^5). \tag{6}$$

This represents the best known upper bound when $r(k) \geq \Omega\left(\frac{k^4}{\log^6 k}\right)$ and k tends to infinity.

1. Our results

Summarizing the above: (1) When either the order k of the clique or the number r of colors is constant, the smallest minimum degree $s_r(K_k)$ is quadratic, up to poly-logarithmic factors, in terms of r and k, respectively; (2) when k and r both tend to infinity, the best known upper bounds are polynomial, with the degree of r (and sometimes also of k) being more than two. Bamberg et al. [1] in fact conjectured that an upper bound r^2k^2, up to logarithmic factors, should hold in all ranges of the parameters.

In this paper we give new constructions which establish this for a large range of the parameters and improve the best known upper bounds for every large enough k and $r \geq c\frac{k}{\log^2 k}$.

Our first main theorem improves the bounds (6), (5) and (4) whenever k tends to infinity and r is large enough.

[1] In [6] only the weaker upper bound $s_r(K_k) \leq q^3 = O(r^3k^6)$ is stated. However, Bishnoi and Lesgourgues [2] recently observed that the choice $q \sim rk^2$ used in [6] for the parameter q is suboptimal and the calculation there also works with $q \sim rk \log k$, resulting in (5).

Theorem 1. *For all sufficiently large k, r satisfying $k \leq r \log^2 r$, we have*

$$s_r(K_k) \leq 2^{400} k^2 r^{2+\frac{30}{k}} \log^{20} r \log^{20} k.$$

Note that this upper bound is of the form $(rk)^{2+o(1)}$ and the error term becomes logarithmic when $r(k) = e^{O(k \log k)}$.

For constant k, our second main result reduces the power of the log factor in the upper bound of (2) from $8k^2$ to 2.

Theorem 2. *For all $k \geq 3$ there exists a constant C_k such that for all $r \geq 2$*

$$s_r(K_k) \leq C_k (r \log r)^2.$$

Combined with the lower bound of (2), Theorem 2 determines the value of $s_r(K_k)$ up to a factor $O(\log r \log\log r)$, for every fixed $k \geq 4$.

2. On the proof

Instead of dealing with minimal Ramsey graphs, the proofs of all the known bounds on $s_r(K_k)$ work with an alternative function, distilled by Fox et al. [6] from the original argument of Burr et al. [3] for $s_2(K_k) = (k-1)^2$.

Definition 3 (Color Pattern)**.** *A sequence of pairwise edge-disjoint graphs $G_1, \ldots, G_r$ on the same vertex set V is called an r-color pattern on V (where the edges of G_i are said to have color i). The color pattern is K_{k+1}-free if G_i is K_{k+1}-free for every $i = 1, \ldots r$.*

Given a color pattern $G_1, \ldots, G_r$ on the vertex set V and an r-coloring $c : V \to [r]$ of the vertices, a strongly monochromatic *copy of a graph H according to c is a copy of H whose edges and vertices all have the same color.*

Definition 4. *Let $r, k \geq 2$ be positive integers, we define $P_r(k)$ to be the smallest positive integer n such that there exists a K_{k+1}-free color pattern $G_1, \ldots, G_r$ on the vertex set $[n]$ such that every r-coloring of $[n]$ induces a strongly monochromatic K_k.*

The connection between s_r and P_r is summarized in the following lemma.

Lemma 5. [6, Theorem 1.5] *For all integers $r, k \geq 2$ we have $s_r(K_{k+1}) = P_r(k)$.*

To prove an upper bound on $P_r(k)$, one needs to construct a K_{k+1}-free r-color pattern $G_1, \ldots, G_r$ with the specific property about strongly monochromatic K_k. As it happens, at the moment we have no other idea of guaranteeing the existence of a strongly monochromatic K_k in an arbitrary r-coloring of the vertices, but requiring that *each* of the graphs G_i has a K_k in *each* subset of size at least n/r and then use this for the largest color class in $[n]$. To this end, we define for a graph G and a positive integer k the parameter $\alpha_k(G)$ to be the order of the largest K_k*-free* induced subgraph of G.

Observation 6. [6, Lemma 4.1] *If there exists a K_{k+1}-free color pattern $G_1, \ldots, G_r$ on $[n]$ such that $\alpha_k(G_i) < \frac{n}{r}$ for every $i = 1, \ldots, r$, then $s_r(K_{k+1}) = P_k(r) \leq n$.*

We will also follow this road and construct K_{k+1}-free r-color patterns on $[n]$ with α_k-values less than n/r. Before constructing r edge-disjoint K_{k+1}-free graphs with $\alpha_k < n/r$ however, one better deals with the "simpler" problem of constructing just one. This is exactly the task of the well-studied Erdős-Rogers function $f_{k,k+1}(n)$ which asks for the smallest value of $\alpha_k(G)$ of K_{k+1}-free graphs G on n vertices. Given a good Erdős-Rogers graph, one then "only" has to pack as many of them as possible on n vertices. Indeed, Fox et al. [6, Conjecture 5.2] even predicted that for every fixed $k \geq 3$ we will have $P_r(k) = \Theta(r \cdot (f_{k,k+1}(r))^2)$. Considering the recent improvements of Mubayi and Verstraete [9] on the Erdős-Rogers function, Theorem 2 comes within a log-factor of resolving this conjecture. Good constructions for the Erdős-Rogers function will be extremely useful for us as well, but creating color patterns using them requires additional ideas.

The general approach to construct the desired color patterns is to start with an "appropriate" u-uniform linear hypergraph on $[n]$ with essentially as many hyperedges as possible, that is, in the order $\frac{n^2}{u^2}$. Then one assigns one of r colors to each hyperedge "appropriately", to indicate which color pairs of vertices inside the hyperedge will receive should they be chosen to be an edge at all. Here the linearity of the hypergraph plays a crucial role: every pair of vertices belongs to (at most) one hyperedge. Finally, one constructs the graphs G_i by dropping an "appropriately" random k-partite graph within each hyperedge of assigned color i, where the choices for different hyperedges are usually independent. This random choice must balance that no K_{k+1} is created from the edges coming from within different hyperedges, yet there are enough edges so that any n/r-subset of the vertices contains a K_k. The crux of the matter is how to define the various occurrences of "appropriate" above so that they complement each other well.

In the construction of Fox et al. [6] for (2) (crucially making use of the Erdős-Rogers construction of Dudek, Retter, and Rödl [4]) and that of Hàn et al. [8] for (4) the linear hypergraph is essentially given by the lines of an (affine or projective) plane of order q and the "appropriate" color assignment chosen randomly. In the constructions of Fox et al. [6] for (5) and of Bamberg et al. [1] for (6) the linear hypergraph is given by some (pseudo)lines in a higher dimensional space and the color-assignment is defined algebraically. The point of these assignments is to ensure that the hypergraph of each color class is *triangle-free*, hence the K_{k+1}-freeness of each G_i will be automatic once the graphs inside the hyperedges are K_{k+1}-free.

In our construction we also start with the projective plane, working in the dual setup, so the lines will correspond to vertices and the vertices correspond to the hyperedges. We choose the order to be q^2, so we are able to make use of Hermitian unitals and its beneficial algebraic and combinatorial properties. One of the main ideas of our construction is to combine the probabilistic color assignment to the hyperedges with an appropriate algebraic one. Unlike in [6] and [1], our algebraic color assignment will not guarantee immediate K_{k+1}-freeness, but will however ensure that the analysis of K_{k+1}-freeness will only have to consider very limited types of forbidden events. The random part of the color assignment then helps to limit the number of bad events within those types.

References

[1] John Bamberg, Anurag Bishnoi, and Thomas Lesgourgues. The minimum degree of minimal Ramsey graphs for cliques. 54(5):1827–1838.
[2] Anurag Bishnoi and Thomas Lesgourgues. A new upper bound on the minimum degree of minimal Ramsey graphs.
[3] Stefan A Burr, Paul Erdős, and Laszlo Lov´ász. On graphs of Ramsey type. 1(1):167–190.
[4] Andrzej Dudek, Troy Retter, and Vojtěch Rödl. On generalized Ramsey numbers of Erdős and rogers. 109:213–227.
[5] Jon Folkman. Graphs with monochromatic complete subgraphs in every edge coloring. 18(1):19–24.
[6] Jacob Fox, Andrey Grinshpun, Anita Liebenau, Yury Person, and Tibor Szabó. On the minimum degree of minimal Ramsey graphs for multiple colours. 120:64–82.
[7] He Guo and Lutz Warnke. Packing nearly optimal Ramsey $r(3,t)$ graphs. 40(1):63–103.
[8] Hiêp Hàn, Vojtěch Rödl, and Tibor Szabó. Vertex folkman numbers and the minimum degree of minimal Ramsey graphs. 32(2):826–838.
[9] Dhruv Mubayi and Jacques Verstraete. On the order of Erdős-rogers functions.

On the separation of roots of random polynomials

Marcus Michelen

(joint work with Oren Yakir)

Consider a random polynomial of the form

$$f_n(z) = \sum_{j=0}^{n} \xi_j z^j$$

where ξ_j are independent and identically distributed variables uniformly chosen from $\{-1, 1\}$. What can be said about the distribution of roots of f_n? Classical results assert that with high probability the vast majority of roots are near the unit circle and they are approximately equidistributed. This means that for each $\varepsilon > 0$, the proportion of roots in the annulus $\{z \in \mathbb{C} : 1-\varepsilon \leq |z| \leq 1+\varepsilon\}$ tends to 1 as $n \to \infty$ with high probability. Similarly, the proportion of roots with argument in $[a, b]$ tends to $(b-a)/(2\pi)$ as $n \to \infty$ with high probability.

A more qualitative feature of roots of random functions is that they tend to experience *repulsion*. This means that for a small ball B near the unit circle, one expects that the probability there are two roots in B is much less than the square of the probability that there is at least one root in B.

One natural way to quantify repulsion is to examine the smallest distance between roots. In particular, define

$$m_n = \min_{\alpha \neq \beta} |\alpha - \beta|$$

where the minimum is over all pairs of distinct roots of f_n. We prove a limit theorem for m_n, meaning that for each $t > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}(m_n \geq tn^{-5/4}) = \exp(-c_* t^4)$$

where $c_* > 0$ is a universal constant.

Our theorem in fact holds in a more general setting: we need only have that the coefficients are independent and identically distributed with subgaussian tails and that $\mathbb{P}(\xi_1 = 0) = 0$. As a corollary, this shows that with high probability a random polynomial with independent and subgaussian coefficients has no repeated roots with high probability (except perhaps at the origin).

One may think of the scale $n^{-5/4}$ as witnessing the repulsion of the roots. A toy model for what the separation would look like without repulsion is to consider placing n independent points in the annulus $\{z : 1 - 1/n \leq |z| \leq 1 + 1/n\}$ and setting σ_n to be the minimal separation. There, one can calculate that $\sigma_n = \Theta(n^{-3/2})$ with high probability. In particular, the typical gap is larger for random polynomials than for independent points.

The proof technique amounts to understanding and quantifying the local repulsion between roots. In particular, for a point z near the unit circle, we prove that having a pair of close roots near z can be approximated by a certain event depending only on the triple $(f(z), f'(z), f''(z))$. We then may approximate the probability that this event holds by proving a Gaussian comparison theorem for vectors $((f(z_j), f'(z_j), f''(z_j))_{j=1}^M$ for a finite set of points $\{z_j\}_{j=1}^M$.

At the technical core of this problem is that the strength of the Gaussian comparison theorem inherently must depend on the arithmetic properties of the points involved. This is a complex analogue of the classical Littlewood-Offord problem—which in fact originates in the study of random polynomials—and so we eliminate candidate double roots near arithmetically structured points by proving a weaker Gaussian comparison theorem but showing that there are not too many of them.

Finally, Gaussian comparison arguments may only hold close to the unit circle. Macroscopically far from the unit circle, no such comparison will hold. For example, if the coefficients of f_n lie in $\{-1, 1\}$ then there are deterministically no roots in the disk $\{z : |z| < 1/2\}$; if the coefficients are standard gaussian random variables, then for each ε, the probability there is a root in $\{z : |z| > \varepsilon\}$ is uniformly bounded below. To eliminate potential close roots far from the unit circle, we compare the random polynomial f_n to the infinite power series

$$f_\infty(z) = \sum_{j=0}^{\infty} \xi_j z^j$$

which is almost-surely analytic in the open unit disk. By a perturbative argument, we show that f_∞ almost-surely has no repeated roots in the open unit disk. By a compactness and monotonicity argument, this implies that f_n does not have a pair of roots that are close to each other in a disk of radius, say, $1 - \varepsilon$.

On the Kim–Vu sandwich conjecture

Richard Montgomery
(joint work with Natalie Behague, Daniel Il'kovič)

Random graphs have been a central object of study in Combinatorics since the foundational work of Erdős and Rényi [9] in 1959. The binomial random graph, or Erdős-Rényi random graph, $G(n,p)$ has vertex set $[n] = \{1,\dots,n\}$ and each potential edge included independently at random with probability p. Along with the closely-related uniformly random graph with n vertices and $\lceil p\binom{n}{2}\rceil$ edges, $G(n,p)$ is the most studied random graph model. The next most studied model is probably that of the random regular graph $G_d(n)$, which is chosen uniformly at random from all d-regular graphs with vertex set $[n]$. Throughout this paper, and as is common, we will implicitly assume that dn is even, so that the set of such graphs is non-empty.

The study of random regular graphs began in earnest in the late 1970's, with early work including that by Bender and Canfield [1, 2], by Bollobás [3], and by Wormald [24, 25]. Compared to $G(n,p)$, where the independence of the edges allows the use of a wide variety of techniques, studying $G_d(n)$ is often much more difficult. For example, from developments [5, 16] of the breakthrough work of Pósa [20] in 1976, it has been long understood when we may expect $G(n,p)$ to be Hamiltonian. It was widely anticipated that, for each $3 \le d \le n-1$, $G_d(n)$ should be Hamiltonian with high probability (i.e., with probability $1-o(1)$), but proving this took the combined work of many authors [4,6,7,17,21,22] over the course of 20 years (see the surveys [10,26] for more details), using a variety of tools in different regimes for d, from the configuration model [4], through switching methods [18], to estimates on the number of regular graphs [19, 23].

This increased difficulty and technicality gives great appeal to finding links from $G_d(n)$ to $G(n,p)$, so that we may hopefully deduce properties of $G_d(n)$ from those known for $G(n,p)$. In 2004, Kim and Vu [14] formalised this desire in their famous 'sandwich conjecture'. They conjectured that, if $d = \omega(\log n)$, then there are $p_*, p^* = (1+o(1))d/n$ and a coupling (G_*, G, G^*) such that $G_* \sim G(n,p_*)$, $G \sim G_d(n)$, $G^* \sim G(n,p^*)$, and $\mathbb{P}(G_* \subset G \subset G^*) = 1-o(1)$. If true, the requirement $d = \omega(\log n)$ cannot be removed; as is well-known, for each $C > 0$ there is some $\varepsilon > 0$ such that if $d = C\log n$ then, with probability $1-o(n^{-1})$, $G(n,(1+\varepsilon)d/n)$ has minimum degree less than d, and hence contains no d-regular subgraph.

Kim and Vu [14] proved the lower part of their conjecture when $\log n \ll d \ll n^{1/3}/\log^2 n$, and a weakened upper part for the same range of d. More specifically, with high probability their coupling (G_*, G, G^*) satisfied $G_* \subset G$ and $\Delta(G\setminus G^*) \le (1+o(1))\log n$. In 2017, Dudek, Frieze, Ruciński and Šileikis [8] extended this by showing that the lower part holds when $d = o(n)$, and that it can be generalised to hypergraphs. Subsequently, a major breakthrough was made by Gao, Isaev and McKay [12, 13], who gave the first coupling (G_*, G, G^*) without a weakened upper part, that is, in which $\mathbb{P}(G_* \subset G \subset G^*) = 1-o(1)$. This allowed them to show that the sandwich conjecture is true if $\min\{d, n-d\} \gg n/\sqrt{\log n}$. Klimošová, Reiher, Ruciński and Šileikis [15] then extended this to show that it holds for $\min\{d, n-d\} \gg (n \log n)^{3/4}$ (while generalising it to biregular graphs). Very significant progress was then made again by Gao, Isaev and McKay [11], who showed that the conjecture is true provided that $d \gg \log^4 n$.

This talk discussed work of the speaker with Natalie Behague and Daniel Il'kovič proving the sandwich conjecture in full, in the following very slightly stronger form.

Theorem 1. *For each $\varepsilon > 0$ there is some $C > 0$ such that the following holds for each $d \geq C \log n$. There is a coupling (G_*, G, G^*) of random graphs such that $G_* \sim G(n, (1-\varepsilon)d/n)$, $G \sim G_d(n)$, $G^* \sim G(n, (1+\varepsilon)d/n)$, and $\mathbb{P}(G_* \subset G \subset G^*) = 1 - o(1)$.*

In their breakthrough initial work, Gao, Isaev and McKay [12, 13] introduced a beautiful natural coupling process, and analysed it in the regime $\min\{d, n-d\} \gg n/\sqrt{\log n}$. In order to prove Theorem 1, we analyse this process throughout the whole regime $d = \omega(\log n)$. As highlighted by Gao, Isaev and McKay [13, Question 2.4], key to this analysis is showing that in certain random graphs F a uniformly chosen d-regular subgraph is likely to include any given edge in F with roughly equal probability, and our success in doing so is of some independent interest.

References

[1] E. A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Mathematics*, 10(2):217–223, 1974.

[2] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[3] 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.

[4] B. Bollobás. Almost all regular graphs are Hamiltonian. *European Journal of Combinatorics*, 4(2):97–106, 1983.

[5] B. Bollobás. The evolution of sparse graphs. *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57, 1984.

[6] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combinatorics, Probability and Computing*, 11(3):249–261, 2002.

[7] C. Cooper, A. Frieze, B. Reed, and O. Riordan. Random regular graphs of non-constant degree: Independence and chromatic number. *Combinatorics, Probability and Computing*, 11(4):323–341, 2002.

[8] A. Dudek, A. Frieze, A. Ruciński, and M. Šileikis. Embedding the Erdős–Rényi hypergraph into the random regular hypergraph and Hamiltonicity. *Journal of Combinatorial Theory, Series B*, 122:719–740, 2017.

[9] P. Erdős and A. Rényi. On random graphs I. *Publ Math Debrecen*, 6:290–297, 1959.

[10] A. Frieze. Hamilton cycles in random graphs: a bibliography. *arXiv preprint arXiv:1901.07139*, 2019.

[11] P. Gao, M. Isaev, and B. McKay. Kim-Vu's sandwich conjecture is true for $d \geq \log^4 n$. *arXiv preprint arXiv:2011.09449*, 2023.

[12] P. Gao, M. Isaev, and B. D. McKay. Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 690–701. SIAM, 2020.

[13] P. Gao, M. Isaev, and B. D. McKay. Sandwiching dense random regular graphs between binomial random graphs. *Probability Theory and Related Fields*, 184(1-2):115–158, 2022.

[14] J. H. Kim and V. H. Vu. Sandwiching random graphs: universality between random graph models. *Advances in Mathematics*, 188(2):444–469, 2004.
[15] T. Klimošová, C. Reiher, A. Ruciński, and M. Šileikis. Sandwiching biregular random graphs. *Combinatorics, Probability and Computing*, 32(1):1–44, 2023.
[16] J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete mathematics*, 43(1):55–63, 1983.
[17] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363, 2001.
[18] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223, 1981.
[19] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs of high degree. *European Journal of Combinatorics*, 11(6):565–580, 1990.
[20] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.
[21] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are Hamiltonian. *Random Structures & Algorithms*, 3(2):117–125, 1992.
[22] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Structures & Algorithms*, 5(2):363–374, 1994.
[23] E. Shamir and E. Upfal. Large regular factors in random graphs. In *North-Holland Mathematics Studies*, volume 87, pages 271–282. Elsevier, 1984.
[24] N. C. Wormald. The asymptotic connectivity of labelled regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):156–167, 1981.
[25] N. C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):168–182, 1981.
[26] N. C. Wormald. Models of random regular graphs. *London Mathematical Society lecture note series*, pages 239–298, 1999.

Dispersion on the complete graph

Konstantinos Panagiotou
(joint work with Umberto de Ambroggio, Tamas Makai, Annika Steibel)

The *dispersion process* introduced by Cooper, McDowell, Radzik, Rivera and Shiraga [3] consists of particles moving on the vertices of a given graph G. A particle is said to be *happy* if there are no other particles occupying the same vertex and *unhappy* otherwise. Initially, $M \geq 2$ (unhappy) particles are placed on some vertex of G. Subsequently, at discrete time steps, all unhappy particles move *simultaneously* and *independently* to a neighbouring vertex selected uniformly at random, while the happy particles remain in place. The process terminates at the first time step at which all particles are happy; we call this (random) time step the *dispersion time*.

It is a simple observation that increasing the number of particles makes it more and more difficult for the particles to disperse quickly. This effect is quite well-understood and causes a sharp transition when the underlying graph is the complete graph on n vertices with loops, in which case we write $T_{n,M}$ for the dispersion time started with M particles at an arbitrary vertex. The typical order of $T_{n,M}$ changes rather abruptly around $M = n/2$. Indeed, if $M = M(n) = (1+\varepsilon)n/2 \in \mathbb{N}$ for some sequence $\varepsilon = \varepsilon(n) \in [-1, 1]$, then in [3] it was established that $T_{n,M}$ is typically

- at most logarithmic in n when $\limsup_{n\to\infty} \varepsilon < 0$ and
- at least exponential in n when $\liminf_{n\to\infty} \varepsilon > 0$.

Our first main contribution is a thorough analysis of this phase transition, in particular, we study how the dispersion time transits from a mere $O(\ln n)$ to an enormous $e^{\Theta(n)}$. To this end, we write (again) $M = (1+\varepsilon)n/2$, but this time $|\varepsilon| = o(1)$. We establish that the process exhibits three qualitatively different behaviours based on the asymptotics of ε, where, informally speaking, $T_{n,M}$ smoothly changes from $|\varepsilon|^{-1}\ln(\varepsilon^2 n)$ to $n^{1/2}$ and then to $\varepsilon^{-1}e^{\Theta(\varepsilon^2 n)}$; in particular, $T_{n,M} = \Theta(n^{1/2})$ whenever $M = n/2 + O(\sqrt{n})$ within a *scaling window* of size $O(\sqrt{n})$. Our first main results, contained in the paper [1], are the following theorems. The first one addresses the upper tail of $T_{n,M}$.

Theorem 1. *There is a $C > 0$ such that the following is true for sufficiently large n and all $A \geq 1$. Let $M = (1+\varepsilon)n/2 \in \mathbb{N}$, where $\varepsilon = o(1)$ and $|\varepsilon| < 1/9$. If $\varepsilon < -en^{-1/2}$, then*

$$\mathbb{P}\big(T_{n,M} > AC|\varepsilon|^{-1}\ln(\varepsilon^2 n)\big) \leq e^{-A}.$$

Moreover, if $|\varepsilon| \leq en^{-1/2}$, then

$$\mathbb{P}\big(T_{n,M} > ACn^{1/2}\big) \leq e^{-A}.$$

Finally, if $\varepsilon > en^{-1/2}$, then

$$\mathbb{P}\big(T_{n,M} > A\varepsilon^{-1}e^{C\varepsilon^2 n}\big) \leq e^{-A}.$$

Our second main result provides corresponding bounds for the lower tail.

Theorem 2. *There is a $c > 0$ such that the following is true for sufficiently large n and all $A \geq 1$. Let $M = (1+\varepsilon)n/2 \in \mathbb{N}$, where $|\varepsilon| = o(1)$ and $|\varepsilon| < 1/9$. If $\varepsilon < -en^{-1/2}$, then*

$$\mathbb{P}\big(T_{n,M} \leq c|\varepsilon|^{-1}\ln(\varepsilon^2 n)/A\big) \leq e^{-A}.$$

Moreover, if $|\varepsilon| \leq en^{-1/2}$, then

$$\mathbb{P}\big(T_{n,M} \leq cn^{1/2}/A\big) \leq e^{-A}.$$

Finally, if $\varepsilon > en^{-1/2}$, set $k_0 := e^{c\varepsilon^2 n}$. Then

$$\mathbb{P}\big(T_{n,M} < \varepsilon^{-1}k_0/A\big) \leq \begin{cases} \exp\left(-\frac{A\varepsilon^2 n}{k_0}\right) & , \text{if } A > k_0 \\ A^{-1} & , \text{if } A \leq k_0 \end{cases}.$$

Let us briefly discuss some consequences of our results. First of all, the two theorems combined imply that in probability

$$T_{n,M} = \Theta(|\varepsilon|^{-1}\ln(\varepsilon^2 n)) \quad \text{if} \quad \varepsilon < -en^{-1/2},$$

and

$$T_{n,M} = \Theta(n^{1/2}) \quad \text{if} \quad |\varepsilon| = O(n^{-1/2}).$$

For larger ε we obtain the slightly weaker uniform estimate that in probability

$$\ln(T_{n,M}) = \Theta(\varepsilon^2 n + \ln(\varepsilon^{-1})) \quad \text{if} \quad \varepsilon = \omega(n^{-1/2}).$$

These results imply that there is critical window around $n/2$ of order $n^{1/2}$, where the dispersion time is of order $n^{1/2}$ as well. Moreover, the dispersion time increases smoothly from logarithmic to $n^{1/2}$ and then to exponential when we gradually increase ε. Apart from these estimates we can also use our main theorems to obtain information about, for example, the expectation of $T_{n,M}$. In particular, Theorem 1 guarantees that $T_{n,M}$ has an exponential(-ly thin) upper tail and so $T_{n,M}$ is, after appropriate normalization, integrable; we readily obtain that the aforementioned bounds hold for $\mathbb{E}[T_{n,M}]$ as well.

Our second main result, see the paper [2], is concerned with a fine analysis of the dispersion process within the critical window, that is, when $M = n/2 + O(n^{1/2})$. We establish that that the dispersion time, scaled by $n^{-1/2}$, converges in distribution to some continuous and almost surely positive random variable.

Theorem 3. *Let $\alpha \in \mathbb{R}$ and $M = M(n) = n/2 + \alpha n^{1/2} + o(n^{1/2}) \in \mathbb{N}$. Then there is a continuous and almost surely positive random variable $T^{(\alpha)}$ such that, as $n \to \infty$,*

$$n^{-1/2} T_{n,M} \xrightarrow{d} T^{(\alpha)} .$$

Within the proof of Theorem 3 we derive several useful properties of the distribution of $T^{(\alpha)}$. In order to give some description at this point we need to step back a bit and introduce some notation. Let us write U_t for the (random) number of unhappy particles at the end of step t, so that $U_0 = M$, and let us fix some $\delta > 0$. It is not difficult to establish that U_t drops rather quickly to $\Theta(n^{1/2})$ particles. In particular, with probability at least $1 - \delta$, after $t^* \sim \frac{4}{7}\delta n^{1/2}$ steps we have that $U_{t^*} \sim n^{1/2}/\delta$. After t^* steps the process $(U_t)_{t \geq t^*}$ of unhappy particles starts fluctuating significantly. In order to get a grip on it, we scale time and space by a factor of $n^{1/2}$ and establish that $(n^{-1/2} U_{t^* + \lfloor s n^{1/2} \rfloor})_{s \geq 0}$ converges weakly to a diffusion process.

Lemma 4. *Let $\alpha \in \mathbb{R}$ and $M = n/2 + \alpha n^{1/2} + o(n^{1/2}) \in \mathbb{N}$. Let $\delta > 0$ and*

$$T_{n,M,\delta} := \inf\{t > 0 : U_t \leq n^{1/2}/\delta\}$$

be the first step at which there are at most $n^{1/2}/\delta$ unhappy particles. Then, as $n \to \infty$, weakly

$$\left(n^{-1/2}\, U_{T_{n,M,\delta} + \lfloor s n^{1/2} \rfloor}\right)_{s \geq 0} \to X,$$

where X is a logistic branching process starting from $X_0 = \delta^{-1}$. In particular, if we denote by B a standard Brownian motion, then X uniquely satisfies the stochastic differential equation

$$dX_s = \left(2\alpha X_s - \frac{7}{4} X_s^2\right) ds + \sqrt{X_s}\, dB_s,\ \ s > 0, \qquad \text{and} \qquad X_0 = \delta^{-1}. \tag{1}$$

A main tool that we use in the proof of Theorem 3 is the concept of *diffusion approximation*, which allows us to approximate a sequence $(\mathbf{Y}^{(n)})_{n \in \mathbb{N}}$ of Markov chains with values in $\mathbb{R}^d$, where $d \in \mathbb{N}$, by a continuous-time stochastic process; the classic book [4, Ch. 8] contains an extensive treatment.

References

[1] U. De Ambroggio, T. Makai, and K. Panagiotou. *Dispersion on the complete graph.* Random Structures Algorithms, 66(1):Paper No. e21266, 13, 2025. An extended abstract appeared in the Proceedings of EUROCOMB '23.

[2] U. De Ambroggio, T. Makai, K. Panagiotou and A.Steibel. *Limit Laws for Critical Dispersion on Complete Graphs.* arXiv:2403.05372, 2025+.

[3] C. Cooper, A. McDowell,T. Radzik, N. Rivera, and T. Shiraga. *Dispersion processes.* Random Structures Algorithms, 53(4):561–585, 2018.

[4] R. Durrett. *Stochastic calculus – a practical introduction.* Probability and Stochastics Series. CRC Press, Boca Raton, FL, 1996.

Low-complexity approximations of sets with small doubling and applications

Huy Tuan Pham

(joint work with Noga Alon)

Given an abelian group G and a finite subset A of G, define the sumset of A

$$A + A = \{a + b : a, b \in A\},$$

and the associated doubling constant $K = \frac{|A+A|}{|A|}$. Sumsets and sets with small doubling K are of fundamental interest in additive combinatorics. Over the years, multiple aspects of the structure of sumsets and sets with small doubling have been studied, notably the influential Freiman-Ruzsa theorem [8, 9, 18], generalized by Ruzsa [17] and Green and Ruzsa [14].doubling K must be dense subsets of certain structured objects known as coset progressions.

While structural results typically provide information about the set A given its doubling K, they provide relatively weak information about the sumset $A + A$. Motivated by fundamental applications to the investigation of sparse random Cayley graphs [1, 2, 11, 13], we will discuss new perspectives on the structure of sumsets $A + A$ of sets with small doubling, revolving around the following basic question: Does every sumset $A + A$ of a set A with small doubling K fully contain a *dense* structured subset F?

Sparse random Cayley graphs. Given an abelian group G and a symmetric subset of it S, the Cayley graph $\Gamma(G; S)$ of G with generating set S has vertex set G, and two group elements x and y are connected if and only if $y - x \in S$. A random Cayley graph $G(p)$ is obtained by selecting each equivalence class $\{x, -x\}$ to be in the generating set S independently at random with probability p.

We also consider the Cayley sum graph $\Gamma^+(G; S)$ which, given a generating set S (not necessarily symmetric), has vertex set G and two group elements x and y are connected if and only if $x + y \in S$. A random Cayley sum graph $G^+(p)$ is obtained by selecting each element x to be in S independently at random with probability p.

The independence number $\alpha(G(p))$ of random Cayley graphs has been extensively studied. In the dense case $p = 1/2$ (and more generally $p = \Theta(1)$), Alon and Orlitsky [4] showed that $\alpha(G(1/2)) = O((\log |G|)^2)$ with high probability. Green

[11] showed that whp $\alpha(G(1/2)) = \Theta(\log|G|)$ for cyclic groups $G = \mathbb{Z}_n$ and that $\alpha(G(1/2)) = \Theta(\log|G| \log\log|G|)$ for finite field vector spaces $G = \mathbb{F}_2^d$ (a similar result holds for any finite field vector space $G = \mathbb{F}_p^d$ with p fixed). Green and Morris [13] later sharpened Green's result to show that whp $\alpha(G(1/2)) = (2+o(1))\log_2|G|$ for $G = \mathbb{Z}_n$. The problem is significantly harder in the sparse case $p = o_{|G|}(1)$. The best general result in this direction is the following theorem of Alon [1, 2].

Theorem 1. *Let G be a group of size n. The independence number of the random Cayley graph $G(p)$ is at most $O(\min(p^{-2}(\log n)^2, \sqrt{n(\log n)/p}))$ whp.*

Alon conjectured that, in terms of the independence number, random Cayley graphs behave similarly to random regular graphs of the same degree.

Conjecture 2 ([2])**.** *Let G be a group of size n. The independence number of the random Cayley graph $G(p)$ is at most $\tilde{O}(p^{-1})$ whp.*

Unlike random regular graphs, random Cayley graphs, and especially sparse random Cayley graphs, are significantly harder to analyze due to the limited randomness in their definition which causes significant dependencies.

In a recent work [6], motivated by Ramsey-theoretic applications, Conlon, Fox, Pham and Yepremyan gave an improvement of Theorem 1 for general groups G.

Theorem 3. *Let G be a group of size n. The independence number of the random Cayley graph $G(p)$ is at most $O\left(p^{-2}\log n \max\left(\log p^{-1}, p^{-1}\log\left(\frac{\log n}{\log p^{-1}}\right)\right)\right)$ whp.*

This result is generally tight for the dense case $p = \Theta(1)$. On the other hand, for sparse p, such as $p = n^{-c}$ for some $c > 0$, the result only gives lower order improvements. No improvement over the exponent p^{-2} in Theorem 1 has been obtained so far. As one application of our key result, we obtain the first improvement in the exponent of p.

Theorem 4. *Let G be an abelian group of size n and let $p \leq 1/2$. Then the independence number of the random Cayley graph $G(p)$ and the random Cayley sum graph $G^+(p)$ is at most $\tilde{O}(p^{-3/2})$ whp.*

By combining our key result, the efficient covering lemma, together with input from the combinatorial approach of [6], we also obtain the sharp asymptotic for the independence number of $G^+(p)$ for $p = \tilde{\Omega}((\log n)^{-2})$, improving earlier results of Green and Morris [13], and Campos, Dahia and Marciano [5] and Nenadov [16].

Largest non-sumsets in $\mathbb{Z}_n$. One of the applications of the independence number of polynomially sparse random Cayley sum graphs is toward the following question in additive combinatorics first considered by Green [10]. Let $f(n)$ be the largest integer such that every subset of $\mathbb{Z}_n$ with size larger than $n - f(n)$ can be represented as a sumset $A + A$. Green asked to determine or estimate $f(n)$ and showed that $f(n) \geq \Omega(\log n)$.

Alon proved that $f(n) \geq \tilde{\Omega}(n^{1/2})$ and, via the upper bound $\alpha(G^+(p)) = O(p^{-2}(\log n)^2)$, that $f(n) \leq \tilde{O}(n^{2/3})$. Using Theorem 4, we obtain an improvement in the exponent of the upper bound to $f(n)$.

Theorem 5. *In the notation above, $f(n) \le \tilde{O}(n^{3/5})$. Specifically, let G be an abelian group of size n. Then there exists a subset of G of size at least $n - \tilde{O}(n^{3/5})$ which cannot be represented as a sumset $A + A$ for $A \subset G$.*

The key result - Efficient covering lemma. Our key input to Theorem 4 is a structural result showing that every sumset $A+A$ of a set A with small doubling must fully contain a *dense structured* subset F.

Theorem 6. *There exist $C, c > 0$ such that the following holds. Let G be an abelian group of order n and let $s \le n$. There exist collections $\mathcal{F}_\ell$ of subsets of G such that*

$$|\mathcal{F}_\ell| \le \exp\left(C\left(\min\left(2^{2\ell}(\log n)^2, \sqrt{2^\ell s (\log n)^{3/2}}\right)\right)\right),$$

and

$$\min_{F \in \mathcal{F}_\ell} |F| \ge c 2^\ell s/\ell^2,$$

so that the following property holds.

Let $A \subseteq G$ be such that $|A + A| \le K|A|$ and $|A| = s$. Then there exists $\ell \le \log_2 K$ and $F \in \mathcal{F}_\ell$ such that

$$A + A \supseteq F.$$

Informally speaking, the theorem yields that for every $s \le n$ and every K, there exists a dyadic scale $h = 2^\ell \le K$ and a collection of sets $\mathcal{F}$ with complexity $\log |\mathcal{F}| \le \tilde{O}(\min(h^2, \sqrt{hs}))$, such that for every subset A of size s and doubling at most K, $A + A$ fully contains a set from $\mathcal{F}$ of size at least $\tilde{\Omega}(h|A|)$.

Lovett [15] asked if it is possible to find a small collection $\mathcal{F}$ of dense sets such that, for every dense subset $A \subseteq G$, the sumset $A + A$ contains a member of $\mathcal{F}$. For every dense A, the doubling of A is clearly bounded. As a special case of Theorem 6 (when $K = O(1)$), we thus resolve Lovett's question.

Theorem 7. *Let G be an abelian group of order n. For any $\delta > 0$, there are $\epsilon > 0$ and $C > 0$ such that the following holds. There exists a collection $\mathcal{F}_\delta \subseteq 2^G$ consisting of sets of size at least ϵn with $|\mathcal{F}_\delta| \le \exp(C(\log n)^2)$ so that for every $|A| \ge \delta n$, $A + A$ fully contains a set $F \in \mathcal{F}_\delta$.*

Theorem 6 connects directly with sparse random Cayley graphs via providing an efficient *union bound obstruction* or *small cover* to the existence of large independent sets in random Cayley graphs. In probabilistic combinatorics language, a cover for a collection $\mathcal{H}$ of sets is a collection $\mathcal{G}$ such that every set in $\mathcal{H}$ contains a set in $\mathcal{G}$ as a subset. They play a critical role in the study of thresholds, particularly in the Kahn-Kalai conjecture. In our case the collection $\mathcal{H}$ is the collection of sumsets $A + A$ over sets A of suitable size. In this language, the collection $\bigcup_\ell \mathcal{F}_\ell$ describes a cover for the collection $\mathcal{H}$. We expect that the independence number of random Cayley graphs can be accurately determined via an optimally efficient cover for the collection $\mathcal{H}$ of sumsets.

The proof of Theorem 7 uses random sampling arguments in physical and Fourier spaces to produce low-complexity approximations.

Efficient approximation lemma - Optimal complexity and enumeration of sets with small doubling. Instead of efficient coverings, if we instead weaken the condition of the family $\mathcal{F}$, so that for every set in A can be *well-approximated* by at least one set in $\mathcal{F}$, then we show in future work that such $\mathcal{F}$ can be constructed with optimal complexity. This is a key tool in our work addressing a conjecture of Alon, Balogh, Morris and Samotij [3] about enumeration of sets with small doubling K in a nearly optimal range of parameters.

References

[1] N. Alon, Large sets in finite fields are sumsets, *J. Number Theory* **126** (2007), 110–118.

[2] N. Alon, The chromatic number of random Cayley graphs, *European J. Combin.* **34** (2013), 1232–1243.

[3] N. Alon, J. Balogh, R. Morris, and W. Samotij, *A refinement of the Cameron–Erdős conjecture,* Proc. Lond. Math. Soc. **108** (2014), 44–72.

[4] N. Alon and A. Orlitsky, Repeated communication and Ramsey graphs, *IEEE Trans. Inform. Theory* **41** (1995), 1276–1289.

[5] M. Campos, G. Dahia, and J. P. Marciano, *On the independence number of sparser random Cayley graphs,* J. Lond. Math. Soc. **110**(6) (2024), e70041.

[6] D. Conlon, J. Fox, H. T. Pham and L. Yepremyan, *On the clique number of random Cayley graphs and related topics*, arXiv:2412.21194.

[7] E. Croot, I. Laba, and O. Sisask, *Arithmetic progressions in sumsets and L_p almost-periodicity,* Combin. Probab. Comput. **22**(3) (2013), 351–365.

[8] G. A. Freiman, *Foundations of a Structural Theory of Set Addition,* Translation of Math. Monographs vol. 37, Amer. Math. Soc, Providence, R. I., USA, 1973.

[9] G. A. Freiman, *What is the structure of K if $K+K$ is small?,* in: Lecture Notes in Mathematics 1240, Springer-Verlag, New York-Berlin, 1987, 109–134.

[10] B. J. Green, *Essay submitted for the Smith's Prize, Cambridge University,* 2001.

[11] B. Green, *Counting sets with small sumset, and the clique number of random Cayley graphs,* Combinatorica **25** (2005), 307–326.

[12] B. Green, 100 open problems, Manuscript.

[13] B. Green and R. Morris, *Counting sets with small sumset and applications,* Combinatorica **36** (2016), 129–159.

[14] B. Green and I. Ruzsa, *Freiman's theorem in an arbitrary abelian group,* Jour. London Math. Soc. **75**(1) (2007), 163–175.

[15] S. Lovett, personal communication.

[16] R. Nenadov, *A remark on the independence number of sparse random Cayley sum graphs,* arXiv preprint (2025).

[17] I. Z. Ruzsa, *An analog of Freiman's theorem in groups,* Astérisque **258** (1999), 323–326.

[18] I. Z. Ruzsa, *Sumsets and structure,* Adv. Courses Math. CRM Barcelona, Birkhäuser Verlag, Basel, 2009, 87–210.

Group-harmonious labelling of trees

Alexey Pokrovskiy

(joint work with Alp Müyesser)

For an abelian group G, a graph T and a labelling $\phi : V(T) \to G$, define a corresponding edge labelling $\phi_e : E(T) \to G$ by $\phi_e(xy) = \phi(x) + \phi(y)$. Say that ϕ is harmonious if both ϕ and ϕ_e are injective. We will study the following question

- Let G be an order n abelian group G and T a tree on n vertices. When does T have a harmonious labelling by G?

Here we are ultimately asking for a full characterization of pairs G, T when a harmonious labelling exists. There are a variety of motivations for this question, such as a conjecture of Chang, Hsu, Rogers [3] which says that every order n tree T has a $\mathbb{Z}_n$ labelling in which the missing edge label is 0. Another related conjecture is by Graham and Sloane [2] which says that every order n tree T can be labelled by $\mathbb{Z}_{n-1}$ so that there is precisely one repetition of vertex-labels, and no repeated edge-labels. We were able to obtain a full answer to the above question in the special case when T has bounded degree.

Theorem 1 ([1])**.** *For any Δ, there exists a n_0 sufficiently large so that the following holds for any $n \geq n_0$. Let T be an n-vertex tree with $\Delta(T) \leq \Delta$ and G an abelian group of order n. There is a rainbow copy of T in K_G if, and only if, we have none of the following:*

(1) *$G = \mathbb{Z}_2^k$, $k \geq 2$ and T is a path or has precisely two vertices of even degree.*
(2) *G has characteristic m, T has adjacent vertices u and v such that $\deg(u) \equiv \deg(v) \equiv 0 \pmod{m}$ and furthermore for all $v \in V(T) \setminus \{u, v\}$, $\deg(v) \equiv 1 \pmod{m}$.*
(3) *$G = \mathbb{Z}_2^k$, $k \geq 2$, T contains precisely 4 vertices of even degree and has a perfect matching when restricted to these 4 vertices.*

Here "characteristic of G" means the least common multiple of the orders of the cyclic factors in the abelian group G. The key tool for proving this is an intermediate result about finding a harmonious labelling of a tree that uses prescribed vertex and edge labels. To state it we need the following definition.

Definition 2. *Let T be a tree. We say that an induced subforest T_{core} is a core of T if for every $d \leq \Delta(T)$, we have at least one of*

(1) *T_{core} contains all vertices v with $d_T(v) = d$.*
(2) *T_{core} and $T \setminus V(T_{core})$ both contain at least 6 vertices v with $d_T(v) = d$.*

The main intermediate result that we use in the proof of Theorem 1 is the following.

Theorem 3. *Let $\Delta^{-1} \gg \mu \gg n^{-1}$. Let G be an abelian group and T a bounded degree tree with $\Delta(T) = \Delta$. Let $V_{target}, C_{target} \subseteq G$ with $|T| = |V_{target}| = |C_{target}| + 1 \geq (1 - n^{-\mu})n$. In the case $G = \mathbb{Z}_2^k$, assume $0 \notin C_{target}$. Let T_{core} be a core of T of size $\leq n^{1-\mu}$. The following are equivalent.*

- *T has a harmonious labelling using vertex-labels V_{target} and edge-labels C_{target}.*

- *There is a harmonious labelling ϕ of T_{core} with vertex-labels contained in V_{target}, edge-labels contained in C_{target} and also*
$$\sum C_{target} = \sum_{v \in V(T_{core})} d_T(v)\phi(v)$$
$$\sum V_{target} = \sum_{v \in V(T_{core})} \phi(v).$$

There are a variety of consequences of the above theorems — for example they imply the conjectures of Chang-Hsu-Rogers and Graham-Sloane in the special cases of large bounded degree trees.

References

[1] A. Müyesser, & A. Pokrovskiy, *On the Graham–Sloane harmonious labelling conjecture*, arXiv:2509.05280 (2025).
[2] R. Graham, & N. Sloane, *On additive bases and harmonious graphs*, SIAM Journal on Algebraic Discrete Methods **1** (1980) 382–404.
[3] G. Chang, D. Hsu, & D. Rogers, *Additive variations on a graceful theme: some results on harmonious and other related graphs*, Congressus Numeratium **32** (1981) 181–197.

A new lower bound for $R(3, k)$

Julian Sahasrabudhe
(joint work with Marcelo Campos, Matthew Jenssen, Marcus Michelen)

The Ramsey number $R(\ell, k)$, is defined to be the minimum n such that every red/blue colouring of the edges of the complete graph K_n contains either a blue K_ℓ or a red K_k. In this talk we focus on the fascinating off-diagonal Ramsey numbers $R(3, k)$. The study of $R(3, k)$ has a rich history, inspiring the development of a number of important combinatorial tools. Today the Ramsey numbers $R(3, k)$ represents one of the successes of the field, and have been determined up to a factor of $4 + o(1)$, the best known bounds being

$$\left(\frac{1}{4} + o(1)\right)\frac{k^2}{\log k} \leqslant R(3, k) \leqslant \big(1 + o(1)\big)\frac{k^2}{\log k}. \tag{1}$$

Here the upper bound is due to Shearer [12], from 1983, who built on the seminal papers of Ajtai, Komlós and Szemerédi [1], which introduced the extremely influential "semi-random method". The lower bound is due, independently, to the celebrated works of Bohman and Keevash [3] and Fiz Pontiveros, Griffiths and Morris [8]. In both of these papers the authors establish this lower bound by studying the *triangle-free process*, a random process introduced by Bollobás and Erdős (see [8]) for generating a random triangle-free graph.

In their paper [8], Fiz Pontiveros, Griffiths and Morris conjectured that their lower bound (1) is sharp. In this talk I will discuss a new lower bound on $R(3, k)$, disproving their conjecture, and narrowing the gap between the upper and lower bound to a factor of $3 + o(1)$.

Theorem 1.

$$R(3,k) \geqslant \left(\frac{1}{3} + o(1)\right)\frac{k^2}{\log k}\,.$$

Perhaps the most interesting consequence of Theorem 1 is that it strongly suggests that the natural triangle-free process is *not* optimal for Ramsey and rather suggests that the optimal graphs mix randomness with some underlying structure.

Philosophy behind the construction. The main intuition behind Theorem 1 is the feeling that there should exist triangle-free graphs that are denser than the triangle-free process but still sufficiently pseudo-random so that the independence number is the same as the binomial random graph at the same density. If one thinks of the triangle-free process as the "most random" graph at its density, the idea behind our construction is to "trade-off" some of this "randomness" to obtain a slightly denser graph.

In fact we believe that a even better trade-off is possible and that one should be able to obtain a constant of $1/2 + o(1)$ in the lower-bound, but we were unable to provide such a construction. Perhaps somewhat boldly, we conjecture this would be optimal.

Conjecture 2.

$$R(3,k) = \left(\frac{1}{2} + o(1)\right)\frac{k^2}{\log k}.$$

In fact, we have a *conjectural* construction, based on the Cayley sum graph generated by the sum-free process on $\mathbb{F}_2^d$, that would imply the lower bound in this conjecture along an infinite sequence.

The construction. Our construction is actually quite easy to describe: we first sample a random graph on $n/(\log n)^2$ vertices at density $p_0 = \alpha_0(\log n/n)^{1/2}$, clean out triangles, blow it up and then run a *variant* of the triangle-free process with the resulting graph. Interestingly, this supplies the desired graph: the blow up allows for more edges and the triangle-free process "counters" the large independent sets introduced by the blow up.

A modified triangle-free process. The reason we work with a variant of the triangle-free process is that the triangle-free process exhibits a complicated dynamics that requires an extremely subtle analysis. For example, in both [3] and [8], the authors need to show the highly non-trivial interaction and "self correction" properties of certain martingales associated with the process. By contrast, in our process, we are able to avoid these subtleties entirely, by manually "steering" the process to ensure it follows a simpler trajectory. This results in a greatly simplified analysis and more straightforward proof. In particular, if we apply our version of the triangle-free process without our initial "seed" step where we take a blow up of a random graph, our analysis gives a much simpler proof of the lower bound (1) on $R(3,k)$ proved in [3, 8].

A history of the problem Following the seminal paper of Ramsey [11] in 1930, the off-diagonal Ramsey numbers $R(\ell, k)$ were introduced by Erdős and Szekeres [7], who showed $R(3,k) \leqslant k^2$. In 1980 this was improved by Ajtai, Komlós and Szemerédi [1], to

$$R(3,k) \leqslant \frac{ck^2}{\log k},$$

for some $c > 0$, in two highly influential papers that introduced the "semi-random method". Soon after, this was improved by Shearer [12] to

$$R(3,k) \leqslant \big(1 + o(1)\big)\frac{k^2}{\log k},$$

which still stands as the best known upper bound for $R(3,k)$.

The first progress on the *lower* bound is due to Erdős in the 1950s who showed $R(3,k) = \Omega(k^{1+c})$, for some $c > 0$, by giving an explicit construction [5]. In 1959, Erdős proved that $R(3,k) = \Omega(k^{3/2})$ by using a probabilistic argument before proving, in 1961, that

$$R(3,k) \geqslant \frac{ck^2}{(\log k)^2},$$

for some $c > 0$, by ingeniously modifying a sample of the binomial random graph [6]. In 1995, the breakthrough work of Kim [9] established

$$R(3,k) \geqslant \frac{ck^2}{\log k}, \tag{2}$$

for some $c > 0$, thereby determining $R(3,k)$ up to constant factors.

In 2008, a different and influential proof of this result was given by Bohman [2], who showed that the graph produced by the "triangle-free process" could also be used to prove Kim's lower bound on $R(3,k)$. The triangle-free process is a random graph process first defined by Bollobás and Erdős in 1990 (see [8]). To define the triangle-free process on $[n] = \{1, \ldots, n\}$, one starts with the empty graph G_0 on vertex set $[n]$ and then inductively defines G_{i+1} to be $G_i + e_{i+1}$ where e_{i+1} is chosen uniformly among all edges e for which $G_i + e$ is triangle free, until there is no such edge, in which case the process stops.

In the difficult and influential papers [3,8], Bohman and Keevash and, independently, Fiz Pontiveros, Griffiths and Morris, studied the trajectory of this process all the way to its (asymptotic) end, determining the the order and independence number of the terminating graph, thereby implying the bound

$$R(3,k) \geqslant \left(\frac{1}{4} + o(1)\right)\frac{k^2}{\log k}.$$

In the work we present in this talk, we discuss a new lower bound to for the Ramsey numbers $R(3,k)$. We do this by running a triangle-free-like process from a carefully chosen "seed graph". Interestingly, our modified triangle-free process is significantly easier to analyze than the original process and we can follow it to the end of its trajectory with out much difficulty.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A dense infinite Sidon sequence. *European Journal of Combinatorics*, 2(1):1–11, 1981.

[2] T. Bohman. The triangle-free process. *Advances in Mathematics*, 221(5):1653–1677, 2009.

[3] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Random Structures & Algorithms*, 58(2):221–293, 2021.

[4] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(12):292–294, 1947.

[5] P. Erdős. Remarks on a theorem of Ramsey. *Bulletin of the Research Council of Israel, Section F*, 7:21–24, 1957.

[6] P. Erdős. Graph theory and probability. II. *Canadian Journal of Mathematics*, 13:346–352, 1961.

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

[8] 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(1274):v+125, 2020.

[9] J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures & Algorithms*, 7(3):173–207, 1995.

[10] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Random Structures & Algorithms*, 7(2):145–155, 1995.

[11] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 1930.

[12] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Mathematics*, 46(1):83–87, 1983.

[13] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20(1):69–76, 1977/78.

The Turán Density of Tight Cycles

Maya Sankar

One of the most fundamental problems in extremal combinatorics lies in understanding those graphs avoiding a fixed forbidden graph F. The first results in this area were for graphs forbidding a clique K_s: Turán [12] in 1941 showed that the densest K_s-free graph on n vertices is a balanced blowup of K_{s-1}, i.e., a complete $(s-1)$-partite graph. In general, given a graph F, its *Turán number* $\mathrm{ex}(n,F)$ is defined to be the maximum number of edges in any F-free graph on n vertices. In many cases, it is impractical to determine the Turán number exactly — instead, we usually study the *Turán density* $\pi(F) = \lim_{n\to\infty} \mathrm{ex}(n,F)/\binom{n}{2}$, which limit is known to exist.

A celebrated theorem of Erdős, Stone, and Simonovits [4,5] shows that $\pi(F) = 1 - \frac{1}{\chi(F)-1}$ for any graph F. Note that this is the density of a balanced blowup of the clique $K_{\chi(F)-1}$. Indeed, we have a stronger structural result: any construction attaining $\mathrm{ex}(n,F)$ differs from a complete $(\chi(F)-1)$-partite graph in at most $o(n^2)$ edges as $n\to\infty$.

We may define Turán numbers of r-uniform hypergraphs, henceforth called *r-graphs*, analogously. If F is an r-graph, then $\mathrm{ex}(n,F)$ is defined to be the maximum number of edges in an F-free r-graph on n vertices, and its Turán density is $\pi(F) = \lim_{n\to\infty} \mathrm{ex}(n,F)/\binom{n}{r}$. The F-free r-graphs attaining $\mathrm{ex}(n,F)$ edges are called the *extremal* constructions. In contrast to the graph case, determining the Turán density of hypergraphs remains a major research area. There is no unique parameter, like the chromatic number, that determines the Turán density. In some cases, the extremal F-free construction is the densest r-graph with chromatic or rainbow chromatic number smaller than F [7, 10, 11]. However, the extremal constructions for other F are mystifying, including fractal-like iterated blowups [2, 8], algebraic constructions [9], and exponentially numerous families [1].

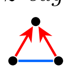
Even determining the Turán densities of seemingly simple hypergraphs remains widely open. In 1961, Turán [13] conjectured that the Turán density of the tetrahedron $K_4^{(3)}$ is 5/9. The lower bound is attained by a large family of constructions [6], but proving a matching upper bound remains open. Also open is the Turán density of the tetrahedron minus an edge $K_4^{(3)-}$, conjectured to be 2/7.

This work introduces a new algebraic parameter that should control the Turán density of various "cycle-like" hypergraphs, together with a general framework to show that these correctly determine the Turán density of sufficiently long "cycle"s. In an application of this framework, we establish the Turán density of 4-uniform tight cycles, defined as follows. The *tight r-uniform cycle of length ℓ*, denoted $C_\ell^{(r)}$, is an r-graph with $\ell > r$ vertices $v_1, \ldots, v_\ell$ such that $v_i \cdots v_{i+r-1}$ is an edge for each $1 \le i \le \ell$, with subscripts taken modulo ℓ. When ℓ is a multiple of r, the tight cycle $C_\ell^{(r)}$ is an r-partite r-graph, and an old result of Erdős [3] implies that $\pi(C_\ell^{(r)}) = 0$.

In 2024, Kamčev, Letzter, and Pokrovskiy [8] determined that $\pi(C_L^{(3)}) = 2\sqrt{3}-3$ for sufficiently large L not divisible by 3. Their result is best phrased in the language of homomorphic avoidance. A *homomorphic copy* of a tight cycle $C_L^{(r)}$ in an r-graph G, also referred to as a *tight closed walk*, is a sequence of L vertices $v_1, \ldots, v_L$ such that $v_{i+1} \cdots v_{i+r}$ is an edge for each $0 \le i < L$ with subscripts taken modulo L. If the vertices $v_1, \ldots, v_L$ are distinct, then the tight path forms a true copy of $C_L^{(r)}$ in G.

Using the technique of supersaturation, it is well-known that the Turán density of $C_L^{(r)}$ is exactly the maximum possible density, asymptotically, of r-graphs avoiding tight closed walks of length L. Kamčev, Letzter, and Pokrovskiy exactly characterized the extremal hypergraphs avoiding the latter family up to $O(1)$ edges, showing that they exhibit a fractal-like structure. A key step in their argument characterizes 3-graphs avoiding tight closed walks of length 1 or 2 modulo 3, yielding a 3-uniform analogue of the statement that a graph avoids odd cycles if and only if it is bipartite.

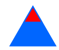
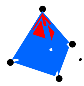
Theorem 1 (Kamčev–Letzter–Pokrovskiy [8]). *Let G be a 3-graph. The following are equivalent.*

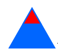
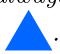
(1) G contains no tight closed walks of length 1 modulo 3.
(2) G contains no tight closed walks of length 2 modulo 3.
(3) It is possible to "orientedly color" each 2-edge of $K_{V(G)}^{(2)}$ as  *or* *so that each 3-edge of G is colored as*  *in some orientation.*

Our results further develop the machinery of Kamčev, Letzter, and Pokrovskiy, yielding a general framework to establish the Turán density of long cycle-like hypergraphs. Key to this framework is a common generalization of Theorem 1 and the equivalence between odd-cycle-free and bipartite graphs. We characterize r-graphs avoiding tight closed walks of length k modulo r as exactly those admitting certain types of "oriented colorings" of $(r-1)$-tuples of vertices. As the general statement is fairly involved, we state here the new coloring results for different residues k modulo 4. The notation $\binom{V}{3}$ is used to denote the edge set of the complete 3-graph $K_V^{(3)}$ on vertex set V.

Theorem 2. *Let G be a 4-graph. The following are equivalent.*

(1) G contains no tight closed walks of length 2 modulo 4.

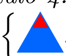
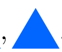
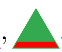
(2) There is an oriented coloring of $\binom{V(G)}{3}$ by the two pictograms  *and*  *so that, for each edge $x_1x_2x_3x_4 \in E(G)$, the coloring restricted to those four vertices is isomorphic to* *.*

In particular, three of the four triples $x_ix_jx_k$ will be colored with *, with the red portion always located at the same vertex, and the fourth triple will be colored with* *.*

Theorem 3. *Let G be a 4-graph. The following are equivalent.*

(1) G contains no tight closed walks of length 1 modulo 4.

(2) G contains no tight closed walks of length 3 modulo 4.

(3) There is an oriented coloring of $\binom{V(G)}{3}$ by the set { , , , ,  } *so that, for each edge $wxyz \in E(G)$, the coloring of $\binom{wxyz}{3}$ is isomorphic to one of the three configurations in Figure 1.*

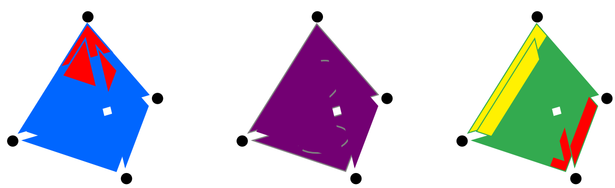


FIGURE 1. The three colorings of a 4-edge permissible in Theorem 3.

To derive our characterization, we introduce the *tight connectivity parameter*, which is a collection of subgroups of S_r measuring the ways in which an edge of an r-graph could connect to permutations of itself via tight walks. It seems likely that this algebraic parameter controls the extremal $C_L^{(r)}$-homomorprhically-free r-graphs for sufficiently large L, and indeed the same approach seems to generalize to a far larger family of "twisted" tight cycles.

Using Theorems 2 and 3, we completely determine the Turán density of all sufficiently long 4-uniform tight cycles of any other residue modulo 4.

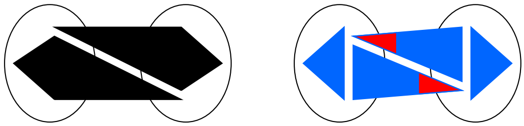


FIGURE 2. At left, a complete oddly bipartite 4-graph G_{odd}. At right, a coloring of $\binom{V(G_{\mathrm{odd}})}{3}$ that proves that G is $C_L^{(4)}$-free for $4 \nmid L$ via Theorems 2 and 3.

Theorem 4. *There exists an integer L_0 such that the following holds. For any $L > L_0$ not divisible by 4, we have $\pi(C_L^{(4)}) = \frac{1}{2}$.*

The lower bound is given by the *complete oddly bipartite* 4-graph, which is constructed as follows. Let V be a set of vertices partitioned as $A \cup B$. The complete oddly bipartite 4-graph $G_{\mathrm{odd}}(A, B)$ has vertex set V whose edge set comprises those 4-tuples intersecting A in an odd number of vertices. This hypergraph is $C_L^{(4)}$-free for all $4 \nmid L$ (see Figure 2) and has edge density approaching $\frac{1}{2}$ as $|V| \to \infty$ if the two parts have roughly equal sizes. (Actually, $e(G_{\mathrm{odd}}(A, B))$ is maximized if $|A|$ and $|B|$ differ from $|V|/2$ by $\Theta(\sqrt{|V|})$ vertices, but it is unknown which explicit choices of $|A|$ and $|B|$ maximize the part sizes [9].) In fact, our full result shows that the extremal 4-graph(s) avoiding tight walks of length L is complete oddly bipartite. As a consequence of our techniques, the extremal $C_L^{(4)}$-free 4-graph differs from a complete oddly bipartite 4-graph in at most $o(n^4)$ edges, although it seems likely that this difference is much smaller.

References

[1] J. Balogh, F. C. Clemen, and H. Luo, Non-degenerate hypergraphs with exponentially many extremal constructions, 2022. `arXiv:2208.00652`.

[2] J. Balogh and H. Luo, Turán density of long tight cycle minus one hyperedge, 2022. `arXiv:2303.10530`.

[3] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[4] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.

[5] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.

[6] A. Frohmader, More constructions for Turán's $(3, 4)$-conjecture, *Electron. J. Combin.* **15** (2008), Research Paper 137, 23.

[7] Z. Füredi and M. Simonovits, Triple systems not containing a Fano configuration, *Combin. Probab. Comput.* **14** (2005), 467–484.

[8] N. Kamčev, S. Letzter, and A. Pokrovskiy, The Turán density of tight cycles in three-uniform hypergraphs, *Int. Math. Res. Not. IMRN* (2024), 4804–4841.

[9] P. Keevash and B. Sudakov, On a hypergraph Turán problem of Frankl, *Combinatorica* **25** (2005), 673–706.

[10] P. Keevash and B. Sudakov, The Turán number of the Fano plane, *Combinatorica* **25** (2005), 561–574.

[11] D. Mubayi and O. Pikhurko, A new generalization of Mantel's theorem to k-graphs, *J. Combin. Theory Ser. B* **97** (2007), 669–678.

[12] P. Turán, On an extremal problem in graph theory, *Mat. Fiz. Lapok (in Hungarian)* **48** (1941), 436–452.
[13] P. Turán, Research problems, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** (1961), 417–423.

Quasipolynomial bounds for the corners theorem

Mehtaab Sawhney
(joint work with Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni)

The key theorem reported on is quasipolynomial bounds for the corners theorem, proved in joint work with Jaber, Liu, Lovett, and Ostuni [1].

Theorem 1. *Let $[N] = \{1, \ldots, N\}$ and $A \subseteq [N] \times [N]$. Suppose that A does not contain $(x, y), (x + d, y), (x, y + d)$ with $d \neq 0$. Then*

$$|A| \ll N^2 \exp\big(-(\log N)^{1/600}\big).$$

Via a projection argument, the above theorem recovers that a subset $B \subseteq [N]$ free of 3-term arithmetic progressions has

$$|B| \ll N \exp\big(-(\log N)^{1/600}\big).$$

This (qualitatively) recovers quasipolynomial bounds for Roth's theorem which were recently proved in breakthrough work of Kelley and Meka [2].

The corners theorem was originally proved by Ajtai and Szemerédi [3]; however, due to the use of Szemerédi's theorem on progressions of growing length, the savings over the trivial bound was exceedingly weak. The corners theorem can additionally be derived from the triangle removal lemma (by an argument of Solymosi [4]); this similarly gives weak bounds. The first reasonable bounds for the corners theorem were given by Shkredov [5], who eventually proved that

$$|A| \ll N^2 (\log\log N)^{-\Omega(1)}.$$

Therefore our results improve on the results of Shkredov by two exponential factors.

For the remainder of our discussion, we adopt the perspective of the finite-field model $\mathbb{F}_2^n \times \mathbb{F}_2^n$. Note that, unlike the case of Roth's theorem where results of Ellenberg and Gijswijt [6] (building on work of Croot, Lev, and Pach [7]) give strong bounds in the finite-field setting, for the *corners* problem there are no known power-saving bounds in the finite-field model. Let $A \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$ with $|A| = \alpha \cdot 4^n$ and no corners and let $f(x, y) = 1_A(x, y) - \alpha$. Then, in particular, we have

$$\mathbb{E}_{x,y,z}\, 1_A(x, y)\, 1_A(x + z, y)\, 1_A(x, y + z) \leq \frac{\alpha^3}{2}.$$

Note that

$$\mathbb{E}_{x,y,z}\, 1_A(x, y) \cdot \alpha \cdot 1_A(x, y + z) = \alpha\, \mathbb{E}_x \Big(\mathbb{E}_y 1_A(x, y)\Big)\Big(\mathbb{E}_z 1_A(x, z)\Big) \geq \alpha^3,$$

and therefore

$$\Big|\mathbb{E}_{x,y,z}\, 1_A(x, y)\, f(x + z, y)\, 1_A(x, y + z)\Big| \geq \frac{\alpha^3}{2}.$$

Making the change of variables $z \leftarrow x+y+z$ (using that $\mathbb{F}_2$ has characteristic 2), we obtain

$$\left|\mathbb{E}_{x,y,z}\, 1_A(x,y)\, f(y+z,y)\, 1_A(x,x+z)\right| \geq \frac{\alpha^3}{2}.$$

By iterated applications of Cauchy–Schwarz and the triangle inequality,

$$\begin{aligned}
\frac{\alpha^3}{2} &\leq \left|\mathbb{E}_{x,y,z}\, 1_A(x,y)\, f(y+z,y)\, 1_A(x,x+z)\right| \leq \mathbb{E}_{x,y}\left|\mathbb{E}_z\, f(y+z,y)\, 1_A(x,x+z)\right| \\
&\leq \left(\mathbb{E}_{x,y}\left|\mathbb{E}_z\, f(y+z,y)\, 1_A(x,x+z)\right|^2\right)^{1/2} \\
&= \left(\mathbb{E}_{x,y}\mathbb{E}_{z,z'}\, f(y+z,y)\, f(y+z',y)\, 1_A(x,x+z)\, 1_A(x,x+z')\right)^{1/2} \\
&\leq \left(\mathbb{E}_{z,z'}\left|\mathbb{E}_y\, f(y+z,y)\, f(y+z',y)\right|\right)^{1/2} \leq \left(\mathbb{E}_{z,z'}\left|\mathbb{E}_y\, f(y+z,y)\, f(y+z',y)\right|^2\right)^{1/4} \\
&= \left(\mathbb{E}_{y,y',z,z'}\, f(y+z,y)\, f(y+z',y)\, f(y'+z,y')\, f(y'+z',y')\right)^{1/4},
\end{aligned}$$

and hence

$$\mathbb{E}_{y,y',z,z'}\, f(y+z,y)\, f(y+z',y)\, f(y'+z,y')\, f(y'+z',y') \;\geq\; \frac{\alpha^{12}}{16}.$$

Note that this corresponds the counts of four-cycles in the weighted bipartite graph given by f. Via the theory of quasi-random graphs, this implies that there exist $X, Y : \mathbb{F}_2^n \to \{0,1\}$ with $\mathbb{E}[X(x)Y(y)] \geq \alpha^{O(1)}$ such that

$$\mathbb{E}[1_A(x,y)\, X(x)\, Y(y)] \geq (\alpha + \alpha^{O(1)}) \cdot \mathbb{E}[X(x)Y(y)].$$

Therefore, relative to $X \times Y$, A has increased density from α to $\alpha+\alpha^{O(1)}$. However, at this stage in the argument, X and Y may be completely arbitrary subsets of $\mathbb{F}_2^n$, and therefore repeating the density increment may not be plausible. A key innovation of Shkredov was realizing that one may replace X and Y with "quasirandom" subsets of a subgroup by splitting X and Y into Fourier-pseudorandom pieces. These two tools were sufficient to prove $|A| \ll 4^n \cdot (\log n)^{-\Omega(1)}$. (Shkredov in fact only operates in the integer setting; Green [8, 9] beautifully exposits the argument in the finite-field case.) However, both the density increment $\alpha \to \alpha + \alpha^{O(1)}$ and the pseudorandomization step, which requires passing to subspaces of codimension $\alpha^{-O(1)}$, cost a logarithm in the final bound. The key difficulty in our work is finding suitable replacements for both of these steps.

To proceed, we must handle sets A that are contained inside container sets of the form $X(x)Y(y)D(x+y)$. Let $\mathbb{E}[X] = \delta_X$, and define δ_Y and δ_D analogously. Replacing the triangle-inequality and Cauchy–Schwarz manipulations above with Hölder's inequality and *spectral positivity* (as developed by Kelley–Meka and by Kelley, Lovett and Meka [10]), we obtain

$$\mathbb{E}_{x_i,y_j\in\mathbb{F}_2^n}\left[\prod_{\substack{i\in[2]\\ j\in[k]}} 1_A(x_i,y_j)\right] \;\geq\; (1+\Omega(1))^k\, \alpha^{2k}\, \delta_X^2\, \delta_Y^k\, \delta_D^{2k}.$$

Equivalently,

$$\mathbb{E}_{x_i\in X,\, y_j\in Y}\Bigg[\prod_{\substack{i\in[2]\\ j\in[k]}} 1_A(x_i,y_j)\Bigg] \;\geq\; (1+\Omega(1))^k\,\alpha^{2k}\,\delta_D^{2k}.$$

A key feature of our result is that, given this condition, we may prove a version of "sifting" relative to pseudorandom majorants. This is closely related to counting lemmas in the combinatorics literature relative to pseudorandom majorants.

The key technical statement is the following; we defer the technical pseudorandomness condition of being (τ,ε)-spread to the main paper.

Theorem 2. *Let $\alpha,\varepsilon,\gamma,\tau\in(0,1)$ be parameters and k a positive integer, satisfying*

$$\gamma \leq (\alpha\tau)^{O\left(\varepsilon^{-2}k\log(1/\alpha)^2+\varepsilon^{-1}k\log(1/\tau)\right)}.$$

Then the following holds.

Let $T\subseteq\Omega_1\times\Omega_2$ be (τ,γ)-combinatorially spread, and let $f:\Omega_1\times\Omega_2\to[0,1]$ be supported on T. Suppose that

$$\mathbb{E}_{\substack{x_1,x_2\in\Omega_1\\ y_1,\ldots,y_k\in\Omega_2}}\left[\prod_{i=1}^{k} f(x_1,y_i)f(x_2,y_i)\right]\geq\alpha^{2k}\tau^{2k}.$$

Then there exist functions $g_1:\Omega_1\to[0,1]$ and $g_2:\Omega_2\to[0,1]$ such that

$$\mathbb{E}_{x\in\Omega_1,\,y\in\Omega_2}[f(x,y)g_1(x)g_2(y)]\geq(1-\varepsilon)\,\alpha\tau\,\mathbb{E}_{x\in\Omega_1}[g_1(x)]\,\mathbb{E}_{y\in\Omega_2}[g_2(y)]$$

and

$$\mathbb{E}_{x\in\Omega_1}[g_1(x)]\geq(\varepsilon\alpha/2)^{O(\varepsilon^{-1}k^2\log(1/\alpha))}\quad\textit{and}\quad\mathbb{E}_{y\in\Omega_2}[g_2(y)]\geq(\varepsilon\alpha/2)^{O(\varepsilon^{-1}\log(1/\alpha))}.$$

The key feature is that $\mathbb{E}[g_i]$ is independent of the size of τ, provided the majorant T is sufficiently quasirandom and that we maintain a multiplicative density increment. This theorem is ultimately proved via a combination of Hölder's inequality and densification, as developed in work of Conlon, Fox, and Zhao [11].

References

[1] Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni, and Mehtaab Sawhney. *Quasipolynomial bounds for the corners theorem.* arXiv:2504.07006, 2025.

[2] Zander Kelley and Raghu Meka. *Strong Bounds for 3-Progressions.* In *Proceedings of the 64th IEEE Symposium on Foundations of Computer Science (FOCS 2023)*, IEEE, 2023, pp. 933–973.

[3] Miklós Ajtai and Endre Szemerédi. *Sets of lattice points that form no squares.* Studia Sci. Math. Hungar. **9** (1974), 9–11.

[4] József Solymosi. *Note on a generalization of Roth's theorem.* In *Discrete and Computational Geometry: The Goodman–Pollack Festschrift*, Algorithms and Combinatorics, vol. 25, Springer, Berlin, 2003, pp. 825–827.

[5] Ilya D. Shkredov. *On a generalization of Szemerédi's theorem.* Proc. London Math. Soc. **93** (2006), no. 3, 723–760.

[6] Jordan S. Ellenberg and Dion Gijswijt. *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression.* Ann. of Math. **185** (2017), no. 1, 339–343.

[7] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach. *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small.* Ann. of Math. **185** (2017), no. 1, 331–337.

[8] Ben Green. *An argument of Shkredov in the finite field setting.* Manuscript, 2012. Available at `https://people.maths.ox.ac.uk/greenbj/papers/corners.pdf`.
[9] Ben Green, *Finite field models in additive combinatorics*, in *Surveys in Combinatorics 2005*, LMS Lecture Note Ser., vol. 327, Cambridge Univ. Press, 2005, pp. 1–27. doi:10.1017/CBO9780511734885.002.
[10] Zander Kelley, Shachar Lovett, and Raghu Meka. *Explicit separations between randomized and deterministic Number-on-Forehead communication.* In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC 2024)*, 2024. DOI: 10.1145/3618260.3649721.
[11] David Conlon, Jacob Fox, and Yufei Zhao. *A relative Szemerédi theorem.* Geom. Funct. Anal. **25** (2015), 733–762.

Disjoint pairs in set systems and combinatorics of low rank matrices

Benny Sudakov
(joint work with Zach Hunter, Aleksa Milojević, István Tomon)

1. Introduction

A central theme of Extremal Set Theory is the study of set systems satisfying certain properties about set intersections. A classical result in the area is the Erdős-Ko-Rado theorem [6], which determines precisely the maximum size of a set system in which no two sets are disjoint. This naturally motivates further questions about the maximum/minimum number of disjoint pairs of sets in set systems of given size. In particular, Daykin and Erdős [8, 14] proposed the following problem in 1981, see also Alon and Frankl [3].

Problem 1 (Daykin, Erdős)**.** *Determine/estimate the maximum number of pairs of disjoint sets in a set system of size m on a universe of size n.*

This problem remained mostly open, however, there has been considerable progress. Observe that if $\mathcal{F} \subset 2^{[n]}$ is the family containing all subsets of $\{1, ..., n/2\}$ and all subsets of $\{n/2+1, ..., n\}$, then $|\mathcal{F}| = 2^{n/2+1}$ and $\mathcal{F}$ contains $\frac{1}{4}|\mathcal{F}|^2$ (unordered) pairs of disjoint sets. Daykin and Erdős conjectured that this is optimal in the weak sense that if a family $\mathcal{F} \subseteq 2^{[n]}$ contains $\varepsilon|\mathcal{F}|^2$ pairs of sets $A, B \in \mathcal{F}$ which are disjoint for some fixed $\varepsilon > 0$, then one must have $|\mathcal{F}| \leq 2^{(1+o(1))n/2}$.

This conjecture was resolved by Alon and Frankl in 1985 [3] by an elegant probabilistic argument. They show that if $\mathcal{F}$ is of size $m = 2^{(1/2+\delta)n}$, then $\mathcal{F}$ contains at most $m^{2-\delta^2/2}$ disjoint pairs. However, as they note in their paper, this bound does not appear to be the best possible. In our first theorem, we improve upon the bound of Alon and Frankl, and obtain an asymptotically correct bound for the problem of Daykin and Erdős in the most interesting range $m \gg 2^{n/2}$.

Theorem 2. *There exists $c > 0$ such that for every positive integer n and $\delta \in (0, 1/2)$, the following holds. Let $\mathcal{F} \subseteq 2^{[n]}$ be a set family of size $m \geq 2^{(1/2+\delta)n}$. Then there are at most $m^{2-\frac{c\delta}{\log 1/\delta}}$ disjoint pairs in $\mathcal{F}$.*

In fact, Alon and Frankl [3] actually made the following more precise conjecture. Before we state it, note that it is often more natural to consider this problem in the bipartite setting. That is, instead of a single family $\mathcal{F}$, we look at disjoint pairs $(A, B) \in \mathcal{A} \times \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ might be different set systems. If $\mathcal{A}, \mathcal{B} \subset 2^{[n]}$ are of size $n^{\omega(1)}2^{n/2}$, then Alon and Frankl conjecture that there cannot be $\Omega(|\mathcal{A}||\mathcal{B}|)$ disjoint pairs $(A, B) \in \mathcal{A} \times \mathcal{B}$.

In 2015, Alon, Das, Glebov and Sudakov [2] proved this conjecture and showed that for any two families $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ with $|\mathcal{A}||\mathcal{B}| = \Theta_d(n^{2d}2^n)$, one has at most $2^{-d/150}|\mathcal{A}||\mathcal{B}|$ disjoint pairs. However, the problem of determining the maximal density of disjoint pairs between two families $\mathcal{A}, \mathcal{B}$ of size $m = \Omega_d(n^d 2^{n/2})$ was not settled. We resolve this problem completely in the following theorem, which is optimal up to the $o(1)$ term.

Theorem 3. *Let $d > 0$ be fixed and let $c_d > 0$ any constant depending on d. If $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ are set families of size $|\mathcal{A}||\mathcal{B}| \geq c_d 2^n n^{2d}$, then the number of disjoint pairs in $\mathcal{A} \times \mathcal{B}$ is at most $(1 + o(1))2^{-2d}|\mathcal{A}||\mathcal{B}|$, where $o(1) \to 0$ as $n \to \infty$.*

1.1. **Complete bipartite graphs in disjointness graphs.** We now turn our attention to a closely related extremal problem about dense disjointness graphs of set families, namely characterizing the size of the largest biclique they contain. Given two set families $\mathcal{A}$ and $\mathcal{B}$ with N disjoint pairs $(A, B) \in \mathcal{A} \times \mathcal{B}$, let $d(\mathcal{A}, \mathcal{B}) = \frac{N}{|\mathcal{A}||\mathcal{B}|}$ denote the density of disjoint pairs. We are interested in the following problem, proposed by Singer and Sudan [12].

Problem 4. *If $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ satisfy $d(\mathcal{A}, \mathcal{B}) \geq \delta$, then how large families $\mathcal{A}' \subseteq \mathcal{A}, \mathcal{B}' \subseteq \mathcal{B}$ can we find such that each pair of sets $(A, B) \in \mathcal{A}' \times \mathcal{B}'$ is disjoint?*

There are two main motivations for studying this problem. Zarankiewicz's problem is one of the long-standing open problems in Extremal Graph Theory asking to find the maximum number of edges in a bipartite graph G with parts of size m and n, which does not contain a complete bipartite graph with parts of size s and t. In the past two decades, the study of Zarankiewicz type problems in which the host graph G is assumed to satisfy certain structural, algebraic, or geometric conditions, gained a lot of attention. Considering graphs which arise as the disjointness graph of set systems fits well into this area.

On the other hand, Problem 4 is connected to the problem of finding monochromatic rectangles in low-rank matrices. The main question in this area is the well-known log-rank conjecture of Lovász and Saks [10], which is one of the fundamental open problems in communication complexity. Its combinatorial formulation asks to show that each binary matrix M of size $n \times n$ and rank r contains an all-zero or all-one submatrix of size at least $2^{-(\log r)^{O(1)}}n$. Such submatrices are called *monochromatic* or *constant*, and they do not have to consist of consecutive rows/columns. While the log-rank conjecture is still outside the reach of known techniques, Lovett [11] showed that one can always find a monochromatic submatrix of size $2^{-O(\sqrt{r}\log r)}n$. Recently, this was slightly improved by Sudakov and Tomon [13], who removed the logarithmic factor from the exponent.

In the same paper [11], Lovett proposed to study the size of all-zero rectangles in low-rank sparse real matrices, stating the following conjecture.

Conjecture 5 ([11]). *Let M be an $n \times n$ real matrix with* $\operatorname{rank}(M) = r$ *such that at most εn^2 entries of M are not zero, where $\varepsilon \in (0, 1/2)$. Then M contains an all-zero square submatrix of size at least $n \cdot \exp(-O(\sqrt{\varepsilon r}))$.*

A construction achieving the bound $n \cdot \exp(-O(\sqrt{\varepsilon r}))$ can be constructed with the help of set systems. Given a pair of set systems $\mathcal{A}, \mathcal{B} \subset 2^{[r]}$, the *intersection matrix* of $(\mathcal{A}, \mathcal{B})$ is the matrix M, whose rows and columns are indexed by the elements of $\mathcal{A}$ and $\mathcal{B}$, respectively, and $M(A, B) = |A \cap B|$. Note that $\operatorname{rank}(M) \leq r$.

Construction. Let $\mathcal{F} = \binom{[r]}{k}$ be the family of all k-element subsets of $[r]$, where $k = \sqrt{\varepsilon r}$, and let M_2 be the $n \times n$ intersection matrix of $(\mathcal{F}, \mathcal{F})$, where $n = \binom{r}{k}$.

Then all but $\Theta(\varepsilon n^2)$ entries of M_2 are 0, and M_2 contains no all-zero submatrix of size $n \cdot 2^{-\sqrt{\varepsilon r}}$. This motivates the problem of studying disjointness graphs of set systems. Specifically, Singer and Sudan [12] proposed the following conjecture. For any two set families $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ with $d(\mathcal{A}, \mathcal{B}) \geq 1 - \varepsilon$, one can find subfamilies $\mathcal{R} \subseteq \mathcal{A}, \mathcal{S} \subseteq \mathcal{B}$ such that $d(\mathcal{R}, \mathcal{S}) = 1$ and $|\mathcal{R}||\mathcal{S}| \geq |\mathcal{A}||\mathcal{B}| \cdot 2^{-O_\varepsilon(\sqrt{n})}$. In the next theorem, we show not only that this is true when $\varepsilon < 1/2$, but also find the optimal behavior in the case of sparse disjointness graphs. We say the two families $\mathcal{R}$ and $\mathcal{S}$ are *cross-disjoint*, if $d(\mathcal{R}, \mathcal{S}) = 1$, or in other words, R and S are disjoint for every $R \in \mathcal{R}$ and $S \in \mathcal{S}$.

Theorem 6. *Let $\mathcal{A}, \mathcal{B} \subset 2^{[n]}$ such that $d(\mathcal{A}, \mathcal{B}) \geq \delta$ for some $\delta \in (0, 1 - \frac{1}{n})$. Then there exist cross-disjoint subfamilies $\mathcal{R} \subseteq \mathcal{A}, \mathcal{S} \subseteq \mathcal{B}$ which satisfy*

$$|\mathcal{R}||\mathcal{S}| \geq 2^{-O(\sqrt{n \log 1/\delta})}|\mathcal{A}||\mathcal{B}|.$$

1.2. **Cover-free families and connections to coding theory.** The key insight in our proof of Theorem 6 is a covering result, which states that for a family $\mathcal{A} \subseteq 2^{[n]}$ and uniformly random sets $A_0, \dots, A_r \in \mathcal{A}$, the probability that $A_0 \subseteq A_1 \cup \dots \cup A_r$ is at least $2^{-O(n/r)}$. This is closely related to a concept in coding theory called r-cover-free families.

A set family $\mathcal{F} \subseteq 2^{[n]}$ is called *r-cover-free* if no set of $\mathcal{F}$ is covered by the union of r others. This notion was originally introduced in 1964 by Kautz and Singleton [9] in the context of coding theory, where they called such families *disjunctive codes.* The r-cover-free families were later introduced to the combinatorics community by Erdős, Frankl and Füredi [5] in 1985. Determining the maximum size of an r-cover-free family has been a long-standing open problem, of interest both in combinatorics and in coding theory. The best known upper bounds are due to D'yachkov and Rykov [4] who show $|\mathcal{F}| \leq r^{O(n/r^2)}$ (see also Füredi [7] for an elegant argument). However, the best known constructions of r-cover-free families only satisfy $|\mathcal{F}| \geq 2^{\Omega(n/r^2)}$ ([9], [5]).

Recently, Alon, Gilboa and Gueron [1] asked what probability distribution μ on $2^{[n]}$ minimizes the probability $\mathbb{P}[A_0 \subseteq A_1 \cup \dots \cup A_r]$ if $A_0, \dots, A_r$ are sampled independently with respect to μ. A natural candidate for a distribution minimizing this probability is the uniform distribution on a maximal r-cover-free family $\mathcal{F}$. Surprisingly, they show that this is not the case.

One of the key observations in Alon, Gilboa and Gueron [1] is that for any distribution μ, $\mathbb{P}[A_0 \subseteq A_1 \cup \cdots \cup A_r] \geq \frac{1}{r^{O(n/r)}}$. This follows from combining the best known upper bound on r-cover-free families with a supersaturation result. We strengthen this lower bound, obtaining a tight result up to the constant in the exponent.

Lemma 7. *Let $n \geq r \geq 1$ be positive integers and let μ be a probability distribution on $2^{[n]}$. If $A_0, \ldots, A_r$ are randomly and independently drawn elements of $2^{[n]}$ with respect to μ, then*

$$\mathbb{P}[A_0 \subseteq A_1 \cup \cdots \cup A_r] \geq 2^{-n/r-2}.$$

References

[1] N. Alon, S. Gilboa, and S. Gueron, *A probabilistic variant of Sperner's theorem and of maximal r-cover free families*, Discrete Math. **343** (2020), 112027.

[2] N. Alon, S. Das, R. Glebov, and B. Sudakov, *Comparable pairs in families of sets*, J. Combin. Theory Ser. B **115** (2015), 164–185.

[3] N. Alon and P. Frankl, *The maximum number of disjoint pairs in a family of subsets*, Graphs Combin. **1** (1985), 13–21.

[4] A. G. Dỳachkov and V. V. Rykov, *Bounds on the length of disjunctive codes*, Problems Inform. Transmission **18** (1982), no. 3, 166–171.

[5] P. Erdős, P. Frankl, and Z. Füredi, *Families of finite sets in which no set is covered by the union of r others*, Israel J. Math. **51** (1985), 79–89.

[6] P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. (2) **12** (1961), 313–320.

[7] Z. Füredi, *On r-cover-free families*, J. Combin. Theory Ser. A **73** (1996), 172–173.

[8] R. Guy, *Unsolved Problems: A Miscellany of Erdős Problems*, Amer. Math. Monthly **90** (1983), 118–120.

[9] W. Kautz and R. Singleton, *Nonrandom binary superimposed codes*, IEEE Trans. Inform. Theory **10** (1964), 363–377.

[10] L. Lovász and M. Saks, *Communication complexity and combinatorial lattice theory*, J. Comput. Syst. Sci. **47** (1993), 322–349. (Conference version: 29th FOCS, 1988.)

[11] S. Lovett, *Communication is bounded by root of rank*, J. ACM **63** (2016), Art. 1, 9 pp.

[12] N. Singer and M. Sudan, *Point-hyperplane incidence geometry and the log-rank conjecture*, ACM Trans. Comput. Theory **14** (2022), no. 2, Art. 7, 16 pp.

[13] B. Sudakov and I. Tomon, *Matrix discrepancy and the log-rank conjecture*, Math. Program. (2024), to appear.

[14] I. Rival (ed.), *Ordered Sets: Proceedings of the NATO Advanced Study Institute, Banff (1981)*, NATO ASI Series C **83**, D. Reidel, Dordrecht, 1982.

The maximum diameter of simplicial complexes

Tibor Szabó

(joint work with Stefan Glock, Olaf Parczyk, Silas Rathke)

The Polynomial Hirsch Conjecture states that the diameter of the edge-vertex graph of polytopes is bounded by a polynomial of the number of their facets. This natural geometric problem is strongly connected to the central algorithmic question on the worst case running time of the simplex algorithm and its study motivated a number of abstractions and generalizations. Santos in particular asked for the maximum diameter of d-dimensional abstract simplicial complexes on n vertices.

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 $G(\mathcal{C})$ (the maximum, taken over all pairs $u, v \in V(G(\mathcal{C}))$ of vertices, of the length of a shortest u,v-path). If $G(\mathcal{C})$ is connected, $\mathcal{C}$ is called *strongly connected*.

Santos [6] 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 [3] 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. Matching the order of magnitude for all n was done by Criado and Newman [2] using a probabilistic construction with the Lovász Local Lemma for any fixed $d \geq 3$. Their result also significantly reduced the gap between upper and lower 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.

At the end of their paper, Bohman and Newman remarked that any improvement of their lower bound would be interesting. In [5] we 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 $n = 6$.

Theorem 1. [5, Parczyk, Rathke, Szabó, 2024+]

$$H_s(n,2) = \begin{cases} \lfloor \frac{1}{2}\binom{n}{2} - \frac{3}{2} \rfloor & n \neq 6 \\ 5 = \lfloor \frac{1}{2}\binom{6}{2} - \frac{3}{2} \rfloor - 1 & n = 6. \end{cases} \tag{4}$$

In [5] it is also conjectured that the simple upper bound (2) can also be achieved for all other d as long as n is large enough. In our most recent manuscript we prove this conjecture.

Theorem 2. [4, Glock, Parczyk, Rathke, Szabó, 2025+] *For every positive integer $d \geq 2$, there exists a positive integer n_0 such that for all $n > n_0$,*

$$H_s(n,d) = \left\lfloor \frac{1}{d}\binom{n}{d} - \frac{d+1}{d} \right\rfloor.$$

References

[1] T. Bohman, A. Newman, Complexes of nearly maximum diameter, (2022), arXiv:2204.11932.
[2] F. Criado, A. Newman, Randomized construction of complexes with large diameter, *Discrete Comput. Geom.*, **66**:2, (2021), 687–700.
[3] F. Criado, F. Santos, The maximum diameter of pure simplicial complexes and pseudo-manifolds, *Discrete Comput. Geom.*, **58**:3, (2017), 643–649.
[4] S. Glock, O. Parczyk, S. Rathke, T. Szabó, The maximum diamater of d-dimensional simplicial complexes, (2025), *manuscript.*
[5] O. Parczyk, S. Rathke, T. Szabó, The maximum diamater of 2-dimensional simplicial complexes, (2024), *submitted.*
[6] F. Santos, Francisco, Recent progress on the combinatorial diameter of polytopes and simplicial complex, *Transactions in Operations Research*, **21**:3, (2013), 482–484.

Stability of Functional and Geometric Inequalities

Marius Tiba

(joint work with Alessio Figalli, Peter van Hintum)

Geometric and functional inequalities play a crucial role in several problems arising in the calculus of variations, partial differential equations, geometry and combinatorics.

The classical Euclidean isoperimetric inequality states that, for any bounded open smooth set $E \subset \mathbb{R}^d$, the perimeter $P(E)$ controls the volume $|E|$:

$$P(E) \geq d|B_1|^{1/d}|E|^{(d-1)/d},$$

where B_1 is the unit ball in $\mathbb{R}^d$. Moreover, equality holds if and only if E is a ball.

A central problem in the study of geometric and functional inequalities is the stability problem. In the case of the isoperimetric inequality, this is based on the principle that if we are close to equality, then the set E is close to being a ball.

This problem has been thoroughly studied and results seek to quantify the two notions of closeness (we refer to the survey [[11], Section 3] for the history of the problem). In 2008, Fusco, Maggi and Pratelli [10] proved the following sharp stability result conjectured by Hall.

Theorem 1. *There exists $C(d) > 0$ such that the following holds. If*

$$P(E) = (1+\delta)d|B_1|^{1/d}|E|^{(d-1)/d},$$

then there exists a ball B with $|B| = |E|$ for which

$$|E\Delta B| \leq C(d)\delta^{1/2}|E|.$$

The Brunn-Minkowski inequality is a fundamental tool in convex geometry and analysis that relates the size of the sumset to the sizes of the individual sets. It states that for compact sets with positive volume $A, B \subset \mathbb{R}^d$, $|A+B|^{1/d} \geq |A|^{1/d} + |B|^{1/d}$. Here, $A + B = \{x + y\colon x \in A, y \in B\}$. Moreover, equality holds if and only if A and B are convex and homothetic.

An equivalent and natural reformulation of the Brunn-Minkowski inequality is that if $t \in (0, 1/2]$ and $A, B \subset \mathbb{R}^d$ are compact sets with equal positive volume, then

$$|tA + (1-t)B| \geq |A|.$$

Equality holds, if and only if A and B are convex and equal up to translation.

It is worth remarking that the Brunn-Minkowski inequality implies the isoperimetric inequality.

The stability for the Brunn-Minkowski inequality is based on the principle that if we are close to equality then the sets A and B are close to being convex and equal up to translation.

The problem of quantifying the two notions of closeness has been intensely investigated. To mention a few important works, a sharp stability result was obtained by Figalli, Maggi and Pratelli [8] when the sets are convex, by Figalli, Maggi and Mooney [7] when one of the sets is a ball, by Barchiesi and Julin [2] when one of the sets is convex and by van Hintum, Spink and Tiba when the sets are two dimensional. For general sets in any dimension, Figalli and Jerison [4] obtained a quantitative sub-optimal stability result.

In this talk we conclude a long line of research on this problem by proving a sharp stability result for the Brunn-Minkowski inequality for general sets in any dimension.

Theorem 2. [5] *There exists $\delta_{d,t} > 0$ s. t. the following holds. If $|A| = |B|$ and*

$$|tA + (1-t)B| = (1+\delta)|A|,$$

for some $\delta < \delta_{d,t}$, then there exists $v \in \mathbb{R}^d$ such that

$$|A\Delta(B+v)| = O_d(1)t^{-1/2}\delta^{1/2}|A|$$

and

$$|co(A) \setminus A| \leq O_{d,t}(1)\delta|A| \quad \text{and} \quad |co(B) \setminus B| \leq O_{d,t}(1)\delta|B|.$$

The Prékopa-Leindler inequality is a functional generalization of the Brunn-Minkowski inequality with applications to high dimensional probability theory. It states that if $t \in (0, 1)$ and $f, g : \mathbb{R}^d \to \mathbb{R}_+$ are continuous functions with bounded support such that

$$\int f\, dx = \int g\, dx > 0$$

and $h : \mathbb{R}^d \to \mathbb{R}_+$ is defined by

$$h(z) = \sup_{z=tx+(1-t)y} f^t(x)g^{1-t}(y),$$

then

$$\int h\ dx \geq \int f\ dx.$$

Equality holds if and only if f and g are log-concave i.e. $\log(f)$ and $\log(g)$ are concave and equal up to translation.

The Borell-Brascamp-Lieb inequality extends Prékopa-Leindler inequality to other means. It states that if $d \in \mathbb{N}$, $t \in (0,1)$, $p \in (-1/d, \infty)$ and $f, g : \mathbb{R}^d \to \mathbb{R}_+$ are continuous functions with bounded support such that

$$\int f\ dx = \int g\ dx > 0$$

and $h : \mathbb{R}^d \to \mathbb{R}_+$ is defined by

$$h(z) = \sup_{z=tx+(1-t)y} \Big(tf^p(x) + (1-t)g^p(y) \Big)^{1/p}$$

in case $p \neq 0$ and by

$$h(z) = \sup_{z=tx+(1-t)y} f^t(x)g^{1-t}(y)$$

in case $p = 0$, then

$$\int h\ dx \geq \int f\ dx.$$

Equality holds if and only if f and g are equal up to translation and p-concave i.e. $\mathrm{sign}(p)f^p$ and $\mathrm{sign}(p)g^p$ are concave in case $p \neq 0$ and $\log(f)$ and $\log(g)$ are concave in case $p = 0$.

The stability of the Prékopa-Leindler and Borell-Brascamp-Lieb inequalities is based on the principle that if we are close to equality, then the functions f and g are close to being p-concave and equal up to translation.

Böröczky, Figalli and Ramos [3] and Figalli and Ramos [9] obtained quantitative sub-optimal stability results for the Prékopa-Leindler inequality and also conjectured a sharp stability result. Balogh and Kristaly [1] established the equality case in Borell-Brascamp-Lieb inequality and Rossi and Salani proved a quantitative sub-optimal stability result when $p > 0$.

In this talk we prove sharp stability results for Prékopa–Leindler and Borell–Brascamp–Lieb inequalities, resolving the conjecture of Böröczky, Figalli and Ramos.

Theorem 3. [6] *There exists $\delta_{d,t,p} > 0$ such that the following holds. If*

$$\int h\ dx = (1+\delta) \int f\ dx$$

for some $\delta < \delta_{d,t,p}$ then there exist $u, v \in \mathbb{R}^d$ and p-concave $c\colon \mathbb{R}^d \to \mathbb{R}_+$ s.t.

$$\int |f(x) - c(x+u)| + |g(x) - c(x+v)|\ dx = O_{d,t,p}(\delta^{1/2}) \int f(x) dx.$$

References

[1] Z. Balogh and A. Kristaly, *Equality in Borell–Brascamp–Lieb inequalities on curved spaces*, Advances in Mathematics, **339** (2018), 453–494.

[2] M. Barchiesi and V. Julin, *Robustness of the Gaussian concentration inequality and the Brunn-Minkowski inequality.*, Calc. Var. Partial Differential Equations, **56** (2017).

[3] K. Böröczky, A. Figalli and J. Ramos, *A quantitative stability result for the Prékopa–Leindler inequality for arbitrary measurable functions*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **41** (2024), 565–614.

[4] A. Figalli and D. Jerison, *Quantitative stability for the Brunn–Minkowski inequality.*, Adv. Math., **314** (2017), 1–47.

[5] A. Figalli, P. van Hintum and M. Tiba, *Sharp quantitative stability of the Brunn-Minkowski inequality*, Arxiv (2023).

[6] A. Figalli, P. van Hintum and M. Tiba, *Sharp Quantitative Stability for the Prékopa-Leindler and Borell-Brascamp-Lieb Inequalities*, Arxiv (2025).

[7] A. Figalli, F. Maggi and C. Mooney, *The sharp quantitative Euclidean concentration inequality.*, Camb. J. Math., **6** (2018), 59–87.

[8] A. Figalli, F. Maggi and A. Pratelli, *A refined Brunn-Minkowski inequality for convex sets*, Ann. Inst. H. Poincare Anal. Non Lineaire, **26** (2009), 2511–2519.

[9] A. Figalli and J. Ramos, *Improved stability versions of the Prékopa-Leindler inequality*, Arxiv (2024).

[10] N. Fusco, F. Maggi and A. Pratelli, *The sharp quantitative isoperimetric inequality.*, Ann. of Math., **168** (2008), 941–980.

[11] F. Maggi, *Symmetrization, optimal transport and quantitative isoperimetric inequalities.*, Optimal transportation, geometry and functional inequalities, CRM Series, 11, Ed. Norm., Pisa, (2010) 73–120.

Approximate versions of Graham's conjecture

Liana Yepremyan

(joint work with Matija Bucic, Bryce Frederickson, Alp Müyesser, Alexey Pokrovskiy)

This talk was on a question in combinatorial number theory, first raised by Graham in 1971, and reiterated by Erdős and Graham in 1980. The conjecture says that for any p prime and any $S \subseteq \mathbb{Z}_p \setminus \{0\}$ with $|S| = d$ there exists a *rearrangement* of S as $a_1, a_2, \ldots, a_d$ such that all partial sums $\sum_{j=1}^{t} a_i$, $1 \leq t \leq d$ are distinct. Such a set S we call *rearrangeable.* Prior to our work [4], most of the progress towards Graham's conjecture has been in the cases when the *generator set* $S = \{a_1, a_2, \ldots, a_d\}$ is very small - $|S| \leq 12$ using Combinatorial Nullstellensatz or very large - $|S| \geq p - 3$ using direct, constructive arguments. Recently, Kravitz [6] and independently Sawin [8] showed that Graham's conjecture holds when $|S| \leq \log p / \log\log p$. In a subsequent paper, Bedert and Kravitz [3] showed that the conjecture holds for $|S| \leq e^{(\log p)^{1/4}}$ which has been an important milestone as this bound overcomes a natural barrier for the rectification techniques used in [6, 8].

In our work, we gave the first approximate solution to this problem for all groups Γ; for any finite group Γ and any subset $S \subseteq \Gamma$ there exists an ordering of elements of S in which at least $(1 - o(1))|S|$ many partial products are distinct. This guarantees that all subsets are approximately rearrangeable in a very general setup, and the proof uses the translation invariance in Cayley graphs. However, we also prove a stronger approximate result for certain cases of groups and generator sets. We can find a subset $S' \subseteq S$ of size $(1-o(1))|S|$ such that S' is rearrangeable in the following cases:

(a) For any Γ and S containing only *involutions*, that is, elements of order two. This in particular applies to an interesting case of the hypercube - $\Gamma = \mathbb{F}_2^n$.
(b) $|S| = \Omega(|\Gamma|)$.
(c) $d \geq p^{2/3+o(1)}$ and $\Gamma = \mathbb{Z}_p$ for some prime p.

Graham's rearrangement conjecture has a natural translation to a graph theoretic question, and that is the setup we use to prove our results. Given a subset S of a group Γ, we define the *colored directed Cayley graph* on Γ with *generator set* S to be the edge-colored directed graph $\mathrm{Cay}(\Gamma, S)$ on the vertex set Γ with an edge from a to ag of color g for every $a \in \Gamma$ and $g \in S$. It is not hard to see that the set S is rearrangeable if and only if the Cayley graph $\mathrm{Cay}(\Gamma, S)$ has a *rainbow* directed path of length $|S|-1$, that is, all edges of the path have distinct colours. Here the *length* of a path refers to the number of edges it contains. For example, if $S = \{a_{i_1}, \ldots, a_{i_d}\}$ is a rearrangement of $S \subseteq \mathbb{Z}_p \setminus \{0\}$ with distinct partial sums, we see that $(a_{i_1}, a_{i_1} + a_{i_2}, \ldots, a_{i_1} + \cdots + a_{i_d})$ is a rainbow directed path in $\mathrm{Cay}(\mathbb{Z}_p, S)$ with $d-1$ edges. Conversely, any rainbow directed path in $\mathrm{Cay}(\mathbb{Z}_p, S)$ with $d-1$ edges is of the form $(x + a_{i_1}, x + a_{i_1} + a_{i_2}, \ldots, x + a_{i_1} + \cdots + a_{i_d})$ for some $x \in \mathbb{Z}_p$ and rearrangement $S = \{a_{i_1}, \ldots, a_{i_d}\}$ with distinct partial sums.

One may wonder if the underlying structure of the groups in Cayley graphs has any importance in Graham's rearrangement conjecture and its variants. Recall an edge colouring of an undirected graph is *proper* if no two edges sharing a vertex have the same colour. In the directed setting, no pair of edges with a common start-vertex and no pair of edges with a common end point may be monochromatic. Indeed, one can ask if any d-regular digraph properly edge-colored with d colors contains a directed rainbow path with $d-1$ edges. For undirected graphs, this question has already been studied extensively. Indeed, Andersen's conjecture [1] from 1989 states that there exists a rainbow path of length $n-2$ in any properly colored K_n. Note that $n-2$ is tight here. In this direction, Hahn [5] conjectured in 1980 that in any proper edge colouring of K_n there exists a rainbow path of length $n-1$. This was refuted by Maamoun and Meyniel [7], by considering $\mathrm{Cay}(\mathbb{F}_2^k, \mathbb{F}_2^k \setminus \{0\})$ which has no rainbow Hamilton path. Andersen's conjecture has been proven to be true asymptotically by Alon, Pokrovskiy, and Sudakov [2] in 2017 who showed the existence of rainbow paths of length $n - o(n)$.

Schrijver [9] asked for a far reaching generalization of Andersen's conjecture by posing the question of the existence of a rainbow path of length $d-1$ in any properly d-edge-colored d-regular graph G, and he verified this conjecture whenever $d \leq 10$. The best general bound on Schrijver's problem prior to our work was a rainbow path of length at most $2d/3$. In our work, we settled Schrijver's question asymptotically, generalizing the approximate solution of Andersen's conjecture in this setting; any properly edge-colored d-regular graph contains a rainbow path of length $(1 - o(1))d$. While a result of this form is perhaps not very surprising for $d = \Omega(|G|)$, the significance of this result lies in covering the entire range of d, in particular, when d is a sufficiently large constant. For directed graphs, we could only get a rainbow directed path of length $d - o(d)$ in the dense regime when $d = \Omega(|G|)$. In both cases, we can relax the condition on the number of colors being arbitrary many and the graph or digraph can be almost-regular.

References

[1] Andersen, L. D., Hamilton circuits with many colours in properly edge-coloured complete graphs, Mathematica Scandinavica, pages 5–14, 19

[2] Alon, N., Pokrovskiy, A. and Sudakov, B.. Random subgraphs of properly edge-coloured complete graphs and long rainbow cycles. Israel Journal of Mathematics, 222(1):317–331, 201

[3] Bederd, B. and Kravitz, N., Graham's rearrangement conjecture beyond the rectification barrier, sr. J. Math. (to appear), arXiv:2409.07403, 2024

[4] Bucic, M., Frederickson, B., Muyesser, M., Pokrovskiy, A., Yepremyan, L., Towards Graham's rearrangement conjecture via rainbow paths, arXiv:2503.01825, 2025.

[5] Hahn, G., Un jeu de coloration, Regards sur la theorie des graphes, Actes du Colloque de Cerisy, 12, page 18, 1980.

[6] Kravitz, N., Rearranging small sets for distinct partial sums, arXiv:2409.07403, 2024

[7] Maamoun, M., Meyniel, H., On a problem of G. Hahn about coloured hamiltonian paths in K_{2t}, Discrete Mathematics, volume 51, issue 2, 1984, pages 213–214.

[8] Sawin, W., comment on the post "Ordering subsets of the cyclic group to give distinct partial sums"', MathOverflow, https://mathoverflow.net/q/202857, 2015

[9] Schrijver, L., Rainbow paths in edge-coloured regular graphs, manuscript.

Reporter: Adva Mond

Participants

Prof. Dr. Noga Alon
Department of Mathematics
Princeton University
Office: 706 Fine Hall
Washington Street
Princeton NJ 08544
UNITED STATES

Prof. Maria Axenovich
Karlsruher Institut f. Technologie (KIT)
Inst. f. Algebra und Geometrie
Englerstr. 2
76131 Karlsruhe
GERMANY

Prof. Dr. Julia Böttcher
Department of Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
UNITED KINGDOM

Domagoj Bradač
EPFL Lausanne
BM 4.134 (Bâtiment B)
Station 17
1015 Lausanne
SWITZERLAND

Dr. Matija Bucić
Princeton University
Washington Road
Princeton NJ 08544
UNITED STATES

Marcelo Campos
Instituto de Matematica Pura e
Aplicada – IMPA
Jardim Botanico
Estrada Dona Castorina, 110
Rio de Janeiro 22460-320
BRAZIL

Prof. Dr. David Conlon
Department of Mathematics
California Institute of Technology
1200 E California Boulevard
Pasadena CA 91125
UNITED STATES

Dr. Michelle Delcourt
Department of Mathematics
Toronto Metropolitan University
350 Victoria Street
Toronto, ON M5B 2K3
CANADA

Prof. Dr. Jacob Fox
Department of Mathematics
Stanford University
Building 380
Stanford CA 94305
UNITED STATES

Prof. Dr. Stefan Glock
Fakultät für Informatik und Mathematik
Universität Passau
94032 Passau
GERMANY

Dr. Annika Heckel
Matematiska Institutionen
Uppsala Universitet
Box 480
751 06 Uppsala
SWEDEN

Zach Hunter
Forschungsinstitut für Mathematik
ETH Zürich
8092 Zürich
SWITZERLAND

Oliver Janzer
Department of Pure Mathematics
and Mathematical Statistics
University of Cambridge
Wilberforce Road
Cambridge CB3 0WB
UNITED KINGDOM

Dr. Matthew Jenssen
Department of Mathematics
King's College London
London WC2R 2LS
UNITED KINGDOM

Felix Joos
Institut für Informatik
Universität Heidelberg
Im Neuenheimer Feld 205
69120 Heidelberg
GERMANY

Dr. Nina Kamcev
Department of Mathematics
University of Zagreb
Bijenicka cesta 30
10000 Zagreb 10000
CROATIA

Prof. Dr. Peter Keevash
Mathematical Institute
University of Oxford
Andrew Wiles Building
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Michael Krivelevich
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Marcus Kühn
Mathematisches Institut
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Dr. Matthew Kwan
IST Austria
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Dr. Shoham Letzter
Department of Mathematics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM

Dr. Anita Liebenau
School of Mathematics and Statistics
UNSW Sydney
Sydney NSW 2052
AUSTRALIA

Jeck Lim
CALTECH
The Division of Physics, Mathematics
and Astronomy
1200 East California Boulevard
Pasadena CA 91125
UNITED STATES

Dr. Sam Mattheus
Dept. of Mathematics and Data Science
Vrije Universiteit Brussel
Pleinlaan 2
1050 Bruxelles
BELGIUM

Leticia Mattos
Universität Heidelberg
Im Neuenheimer Feld 205
69120 Heidelberg
GERMANY

Marcus Michelen
Department of Computer Science
University of Illinois at Chicago
M/C 249, 322 SEO
851 S. Morgan Street
Chicago IL 60607-7045
UNITED STATES

Dr. Adva Mond
Department of Mathematics
King's College London
Strand
London WC2R 2LS
UNITED KINGDOM

Prof. Dr. Richard Montgomery
Mathematics Institute
University of Warwick
Gibbet Hill Road
Coventry CV4 7AL
UNITED KINGDOM

Prof. Dr. Robert Morris
IMPA
Instituto Matematica Pura e Aplicada
Estrada Dona Castorina, 110
Jardim Botanico
Rio de Janeiro 22460-320
BRAZIL

Prof. Dr. Dhruv Mubayi
Dept. of Mathematics, Statistics
and Computer Science
University of Illinois Chicago
851 S. Morgan Street
Chicago IL 60607-7045
UNITED STATES

Dr. Rajko Nenadov
School of Computer Science
The University of Auckland
38 Princes Street
1010 Auckland
NEW ZEALAND

Prof. Dr. Konstantinos Panagiotou
Mathematisches Institut
Universität München
Theresienstraße 39
80333 München
GERMANY

Dr. Jinyoung Park
Courant Institute of
Mathematical Sciences
New York University
251, Mercer Street
New York NY 10012-1110
UNITED STATES

Dr. Will Perkins
School of Computer Science
Georgia Institute of Technology
Atlanta GA 30305
UNITED STATES

Huy Tuan Pham
The Division of Physics, Mathematics
and Astronomy
California Institute of Technology
1200 E California Blvd
Pasadena CA, 91125
UNITED STATES

Dr. Alexey Pokrovskiy
Department of Mathematics,
University College London
Gordon Street
London WC1H 0AY
UNITED KINGDOM

Prof. Dr. Oliver Riordan
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Julian Sahasrabudhe
Dept. of Math. and Math. Stats.
University of Cambridge
Wilberforce road
Cambridge CB3 0WA
UNITED KINGDOM

Prof. Dr. Wojciech Samotij
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Maya Sankar
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
UNITED STATES

Prof. Dr. Lisa Sauermann
Universität Bonn
Institut für Angewandte Mathematik
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Mehtaab Sawhney
Department of Mathematics
Columbia University
Mathematics Hall, 2990 Broadway
New York City NY 10027
UNITED STATES

Prof. Dr. Mathias Schacht
Fachbereich Mathematik
Universität Hamburg
Bundesstraße 55
20146 Hamburg
GERMANY

Prof. Dr. Alex Scott
Mathematical Institute
University of Oxford
Radcliffe Observatory Quarter
Andrew Wiles Building
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Benjamin Sudakov
Department of Mathematics
ETH Zürich
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Tibor Szabó
Institut für Mathematik
Freie Universität Berlin
Arnimallee 6
14195 Berlin
GERMANY

Dr. Marius Tiba
King's College London
London WC2R 2LS
UNITED KINGDOM

Prof. Dr. Jacques Verstraete
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES

Yuval Wigderson
Institute for Theoretical Studies
ETH Zürich
Scheuchzerstrasse 70
8006 Zürich
SWITZERLAND

Dr. Liana Yepremyan
Department of Mathematics
Emory University
400 Dowman Drive
Atlanta GA 30307
UNITED STATES