Mathematisches Forschungsinstitut Oberwolfach

Report No. 1/2020

DOI: 10.4171/OWR/2020/1

Combinatorics

Organized by
Jeff Kahn, Piscataway
Angelika Steger, Zürich
Benny Sudakov, Zürich

5 January – 11 January 2020

Abstract. Combinatorics is a fundamental mathematical discipline that focuses on the study of discrete objects and their properties. The present workshop featured research in such diverse areas as Extremal, Probabilistic and Algebraic Combinatorics, Graph Theory, Discrete Geometry, Combinatorial Optimization, Theory of Computation and Statistical Mechanics. It provided current accounts of exciting developments and challenges in these fields and a stimulating venue for a variety of fruitful interactions. This is a report on the meeting, containing extended abstracts of the presentations and a summary of the problem session.

Mathematics Subject Classification (2010): 05-XX.

Introduction by the Organizers

The workshop Combinatorics, organized by Jeff Kahn (Piscataway), Angelika Steger (Zürich) and Benny Sudakov (Zürich), was held the first week of January, 2020. Despite the early point in the year the meeting was well attended, with roughly 50 participants from the US, Canada, UK, Israel, and various European countries. The program consisted of 11 plenary lectures and 17 shorter contributions, including the presentations by Oberwolfach Leibniz graduate students. There was also a lively problem session led by Nati Linial. The plenary lectures were chosen to provide both overviews of the state of the art in various areas and in-depth treatments of major new results. The short talks ranged over a broad range of topics, including, *for example* (far from an exhaustive list), graph theory, coding theory, probabilistic combinatorics, discrete geometry, extremal combinatorics and Ramsey theory, additive combinatorics, and theoretical computer science. As in the past, particular stress was placed on providing a platform for younger researchers to present themselves and their results.

This report contains extended abstracts of the talks and some discussion of the problems that were posed at the problem session. This was a particularly successful edition of the meeting Combinatorics, in large part because of the exceptional strength and range of both the participants and the results presented. While it is hard to do justice to a meeting at such a level in this short summary, we here highlight just three of the more spectacular developments.

The first highlight is the work discussed by Jinyoung Park (joint with Keith Frankston, Jeff Kahn, and Bhargav Narayanan), proving a fractional version of the "expectation-threshold" conjecture of Kahn and Kalai.

For any increasing family $\mathcal{F}$, the "expectation threshold" $q(\mathcal{F})$ is a trivial lower bound on the threshold $p_c(\mathcal{F})$. Kahn and Kalai conjectured that this bound is never far from the truth. Talagrand proposed a sort of LP relaxation of the Kahn-Kalai conjecture with $q(\mathcal{F})$ replaced by $q_f(\mathcal{F})$, the "fractional expectation threshold," and then a strengthening with $|X|$ replaced by $\ell(\mathcal{F})$, the maximum size of a minimal element of $\mathcal{F}$. The main result of the work under discussion is a proof of the latter conjecture.

This easily implies several heretofore difficult results and conjectures in probabilistic combinatorics, including thresholds for perfect hypergraph matchings (Johansson–Kahn–Vu), bounded degree spanning trees (Montgomery), and bounded degree graphs (new). The key idea of the proof can also be applied to resolve (and vastly extend) the "axial" version of the random multi-dimensional assignment problem, which was earlier considered by Martin–Mézard–Rivoire and Frieze–Sorkin.

The second highlight is the proof of Ringel's Conjecture for sufficiently large trees, presented by Alexey Pokrovskiy (joint work with Richard Montgomery and Benny Sudakov). Decomposition problems have a very long and illustrious history dating back to Euler's work on Latin squares from 18th century all the way to the recent proof of existence of designs by Keevash. Ringel's conjecture is one of the oldest and best known conjectures on graph decompositions. It states that for any tree T of size n the edges of the complete graph K_{2n+1} can be partitioned into edge-disjoint subgraphs isomorphic to T.

Before this breakthrough Ringel's conjecture has only been established for some very special classes of trees and certain asymptotic results were known. The proof, which is inspired by randomized decompositions and the absorption technique, manages to surpass a major obstruction of working with trees of possibly unbounded degree, which could be helpful for approaching a number of related long standing conjectures on graph decompositions and graceful labellings.

The third higlight is a new breakthrough concerning the chromatic number of random graphs by Annika Heckel and Oliver Riordan, which was presented by Heckel. Studying the chromatic number of random graphs is one of the oldest and most studied topics in random graph theory, dating back to seminal papers of Erdős and Rényi.

In a breakthrough paper from 1978, Bollobás established the asymptotic value of the chromatic number of a random graph $G_{n,\frac{1}{2}}$: w.h.p. $\chi(G_{n,\frac{1}{2}}) \sim \frac{n}{2\log_2 n}$. As a follow-up question, various researchers have asked and studied how sharp is the *concentration* of $\chi(G_{n,p})$.

On a slightly different note, in 1992 Erdős asked how accurately can we determine $\chi(G_{n,\frac{1}{2}})$—could it be that it is *not concentrated* on a series of intervals of constant length? Recently, Heckel gave a positive answer towards this question, namely that $\chi(G_{n,\frac{1}{2}})$ is not w.h.p. concentrated on fewer than $n^{\frac{1}{4}-\varepsilon}$ consecutive values.

In this work, Heckel and Riordan extend this result to an almost optimal one. They show that for any $\varepsilon > 0$, there is no sequence of intervals of length $n^{\frac{1}{2}-\varepsilon}$ which contain $\chi(G_{n,\frac{1}{2}})$ with high probability. This lower bound is tight up to the constant $\varepsilon > 0$.

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-1641185, "US Junior Oberwolfach Fellows". Moreover, the MFO and the workshop organizers would like to thank the Simons Foundation for supporting Ron Aharoni in the "Simons Visiting Professors" program at the MFO.

Workshop: Combinatorics

Table of Contents

Abstracts

Intersection theory in combinatorics and discrete geometry

Karim Adiprasito

I surveyed some applications of Hodge and intersection theory to combinatorics and discrete geometry. I started with an application of the Hodge-Riemann relations to polytope theory, answering a question of Gromov:

Theorem 1. *An infinitesimal deformation of a polytope that does not decrease dihedral angles is a combination of normal equivalence and isometry.*

On a more sophisticated level, I then presented an application of Hodge theory to matroid theory

Theorem 2 (Adiprasito-Huh-Katz '18)**.** *The characteristic polynomial of a matroid has log-concave coefficients.*

Finally, I presented a solution to the g-conjecture for simplicial spheres, and its relation to the Hall Marriage Theorem

Theorem 3 (A, arxiv:1812.10454)**.** *Consider a PL* $(d-1)$*-sphere* Σ*, or more generally a PL rational homology sphere of that dimension, and the associated graded commutative face ring* $\mathbb{R}[\Sigma]$ *(see Stanley, Birkhäuser Prog. in Math. 1996). Then there exists an open dense subset of the Artinian reductions* $\mathcal{R}$ *of* $\mathbb{R}[\Sigma]$ *and an open dense subset* $\mathcal{L} \subset A^1(\Sigma)$*, where* $A(\Sigma) \in \mathcal{R}$*, such that for every* $k \le \frac{d}{2}$*, we have:*

(1) Generic Lefschetz theorem: *For every* $A(\Sigma) \in \mathcal{R}$ *and every* $\ell \in \mathcal{L}$*, we have an isomorphism*

$$A^k(\Sigma) \xrightarrow{\ \cdot \ell^{d-2k}\ } A^{d-k}(\Sigma).$$

(2) Hall-Laman relations: *The* Hodge-Riemann bilinear form

$$\begin{array}{ccccccc} \mathrm{Q}_{\ell,k}: & A^k(\Sigma) & \times & A^k(\Sigma) & \longrightarrow & A^d(\Sigma) \cong \mathbb{R} \\ & a & & b & \longmapsto & \deg(ab\ell^{d-2k}) \end{array}$$

is nondegenerate when restricted to any squarefree monomial ideal in $A(\Sigma)$*, as well as the annihilator of any squarefree monomial ideal.*

The colorful world of rainbow sets

Ron Aharoni

(joint work with Joseph Briggs, Ron Holzman, Zilin Jiang, Jinha Kim and Minki Kim)

1. Introduction - the intersection of a matroid and a filter

Given a family $\mathcal{S} = (S_1, \ldots, S_m)$ of not necessarily distinct sets, an $\mathcal{S}$-choice function is a partial function f such that $Dom(f) \subseteq \mathcal{S}$, and $f(S_i) \in S_i$ for all $S_i \in Dom(f)$. The image of f is then called a rainbow set. The term originates in viewing the sets S_i as "colors". We keep track of the "color" of each element of the rainbow set. So, when we speak of the rainbow set we keep in mind the function defining it. If $Dom(f) = \mathcal{S}$ we say that the rainbow set is full.

Rainbow sets can be viewed in the wider context of matroids. By duplicating vertices we can assume that the sets S_i are disjoint, and then a rainbow set is an independent set in the partition matroid whose parts are the sets S_i. This point of view enables formulating rainbow sets problems in a familiar framework - asking for the existence of a set that is small in one sense, and large in another. This means belonging to the intersection of a complex (closed down hypergraph) and a filter (closed up hypergraph). To obtain meaningful results, either the complex or the complement of the filter (which is of course another complex) should carry more structure, and a suitable structure is that of a matroid.

Let $\mathcal{C}$ be a complex and $\mathcal{M}$ a matroid on the same vertex set V. The script above means asking for two possible types of sets, one that is large in $\mathcal{M}$ and small in $\mathcal{C}$, or small in $\mathcal{M}$ and large in $\mathcal{C}$. "Small" in $\mathcal{M}$ or in $\mathcal{C}$ simply means belonging to them, while "large in $\mathcal{M}$" has, besides the meaning of not belonging to the matroid, also the meaning of spanning. So, we can look for a set A satisfying:

(1) A spans $\mathcal{M}$ and belongs to $\mathcal{C}$, or
(2) $A \in \mathcal{M} \setminus \mathcal{C}$.

In fact, the two are one and the same problem. Let $D(\mathcal{C}) = \{S \subseteq V \mid S^c \notin \mathcal{C}\}$ (this is called the "combinatorial Alexander dual" of $\mathcal{C}$). Then it is easily seen that $A \in \mathcal{M} \setminus \mathcal{C}$ if and only if A^c is spanning in $\mathcal{M}^*$ (the matroidal dual of $\mathcal{M}$) and belongs to $A(\mathcal{C})$.

A main tool for tackling this type of problems is provided by topology: connectivity of simplicial complexes. Given a simplicial complex $\mathcal{C}$, the *topological connectivity* $\eta(\mathcal{C})$, is the smallest dimension of a hole in $\mathcal{C}$. This has a homotopic version - the smallest d for which there exists a continuous function $f : S^{d-1} \to ||C||$ (the geometric realization of $\mathcal{C}$) that cannot be extended to a continuous function $\tilde{f} : B^d \to ||\mathcal{C}||$. It also has a homological version - the smallest d such that the homology group $\tilde{H}_{d-1}$ is non-trivial. There is also a dual notion - $\lambda'(\mathcal{C})$ is the largest d such that $\tilde{H}_{d-1}$ is non-trivial (here the "sphere" version is not applicable. It is possible that $\tilde{H}_{d-1} \neq 0$, and there is no empty sphere of dimension larger than $d+1$). Also let $\lambda(\mathcal{C}) = \max\{\lambda'(\mathcal{C}[S]) \mid S \subseteq V(\mathcal{C})\}$. Since the Alexander dual of

$\mathcal{C}$ is a "mirror image" of $\mathcal{C}$, it is not entirely surprising (though in no way trivial) that the homology groups of $D(\mathcal{C})$ are mirror images of the homology groups of $\mathcal{C}$. A folklore theorem (if names are to be attached, they are those of Stanley and Kalai), is that

$$\tilde{H}_i(D(\mathcal{C})) = \tilde{H}_{n-i-3}(\mathcal{C})$$

where $n = |V(\mathcal{C})|$. In particular,

$$\lambda'(\mathcal{C}) \leq n - \eta(D(C)) - 1. \tag{1}$$

Two theorems were proved around the same time, on the two types of problems. Regarding (1), the following was proved:

Theorem 1. [1] *If $\eta(\mathcal{C}[S]) \geq rank(\mathcal{M}.S)$ for every $S \subseteq V$ then there exists a spanning set of $\mathcal{M}$ that belongs to $\mathcal{C}$.*

Here $\mathcal{M}.S$ is the contraction of $\mathcal{M}$ to S. As to (2), the corresponding theorem is:

Theorem 2. [10] *If $\lambda(\mathcal{C}) \leq rank(\mathcal{M}[S^c])$ for every $S \in \mathcal{C}$ then $\mathcal{M} \setminus \mathcal{C} \neq \emptyset$.*

As mentioned above, the two theorems are in fact the same (a fact that was not recognized at the time of their inception). To prove Theorem 2 from Theorem 1, we use (1), and the easy identity

$$D(K)[S] = D(lk_K(S^c))$$

whenever $S^c \in K$ (the link $lk_K(\sigma)$ of a face σ is the complex consisting of the sets complementing σ to a face of K). We also use another theorem of Kalai-Meshulam:

Theorem 3. [11] *If $\lambda(K) \leq d$ then $\lambda(lk_K(\sigma)) \leq d$ for every $\sigma \in K$.*

Theorem 2 yields also:

Theorem 4. *Let $\mathcal{N}$ be a matroid and $\mathcal{K}$ a complex on the same ground set V. Suppose that $\lambda(\mathcal{K}) \leq d$. If for every $S \in D(\mathcal{K})$ we have $rank_{\mathcal{N}}(S) \leq |S| - d$ then there exists a set belonging to $\mathcal{N} \setminus K$.*

2. Partition matroids

In most known applications of the above two theorems the matroid considered is a partition matroid. In this case Theorem 1 is the so called "Topological Hall" theorem:

Theorem 5. *If $\mathcal{C}$ is a complex, $V(C)$ is partitioned into sets $V_1, \ldots, V_m$, and $\eta(C[\bigcup_{i \in I}]) \geq |I|$ for every $I \subseteq [m]$ then there exists a full rainbow set belonging to $\mathcal{C}$.*

(The grain of the theorem appears in [4], the formulation above appears in [13], ascribed to the speaker).

Theorem 4 becomes:

Theorem 6. *If $\lambda(\mathcal{C}) \leq d$ and $\mathcal{S} = (S_1, \ldots, S_{d+1})$ is a family of sets such that $S_i \notin \mathcal{C}$ for all $i \leq m$, then there exists an $\mathcal{S}$-rainbow set not belonging to $\mathcal{C}$.*

2.1. **Colorful Caratheodory.** A famous result of type (2) is the Bárány–Lovász colorful version of Caratheodory's theorem. For a set $V = \{\vec{v}_i,\ i \in I\}$ of vectors in $\mathbb{R}^d$ let

$$cone(V) = \{\sum_{i \in I} \alpha_i \vec{v}_i \mid \forall i\ \alpha_i \geq 0\}$$

and

$$conv(V) = \{\sum_{i \in I} \alpha_i \vec{v}_i \mid \forall i\ \alpha_i \geq 0 \text{ and } \sum_i \alpha_i = 1\}$$

Theorem 7 (Bárány [6]).

(1) If $S_1, \ldots, S_d$ are sets of vectors in $\mathbb{R}^d$ satisfying $\vec{v} \in cone(S_i)$ for all $i \leq d$ then there exists a rainbow set S such that $\vec{v} \in\ (S)$.

(2) If $S_1, \ldots, S_{d+1}$ are sets of vectors in $\mathbb{R}^d$ satisfying $\vec{v} \in conv(S_i)$ for all $i \leq d+1$ then there exists a rainbow set S such that $\vec{v} \in conv(S)$.

This follows from Theorem 6, and the facts that the complex $NONCONE(v) = \{S \in \mathbb{R}^d \mid v \notin cone(S)\}$ satisfies $\lambda(NONCONE(v)) = d-1$, and the complex $NONCONV(v) = \{S \in \mathbb{R}^d \mid v \notin conv(S)\}$ satisfies $\lambda(NONCONE(v)) = d$. (These facts are proved using LP duality and the Nerve Theorem, see e.g [14].)

2.2. **Rainbow matchings.** The following theorem was proved by Drisko [9] in a special case, formulated in terms of Latin rectangles, and in [1] in its generality.

Theorem 8 ([1,9]). *$2n-1$ matchings of size n in a bipartite graph have a rainbow matching of size n.*

The case in which all matchings live in $K_{n,n}$ follows from Theorem 7, since a set F of edges in $K_{n,n}$ contains a matching of size n if and only if $\vec{1}$, the all 1's vector, belongs to $cone(\{\chi_f \mid f \in F\})$. Here, and below, χ_A denotes the characteristic function of the set A. In this case the vectors χ_f belong to a $2n-1$-dimensional vector space, since they live in $\mathbb{R}^{2n}$, and all satisfy a linear equality - the sum over the first side of the bipartite graph is equal to the sum on the other side. The general case follows from Theorem 4 and the following result:

Theorem 9. [5] *Given any bipartite graph G and a natural number n, the complex $\{F \subseteq E(G) \mid \nu(F) < n$ has $\lambda \leq 2n-2$.*

As to general graphs, Seunghun Lee proved, using discrete Morse theory:

Theorem 10. [12] *Given any bipartite graph G and a natural number n, the complex $\{F \subseteq E(G) \mid \nu(F) < n$ has $\lambda \leq 3n-3$.*

By Theorem 2 This implies that $3n-2$ matchings of size n in any graph have a rainbow matching of size n. This was first proved, combinatorially, in [2]. Lee's bound on λ is sharp, so the bound $3n-2$ cannot be improved using this method. Using combinatorial methods, it was proved in [3] that $3n-3$ matchings of size n suffice. The conjecture is that $2n$ matchings suffice. This would follow, by a trick of doubling the matchings, from the following conjecture of the speaker and Eli Berger:

Conjecture 11. *n matchings of size n in any graph have a rainbow matching of size $n-1$.*

In [5] the following was proved:

Theorem 12. *$2n$ matchings of size n in any graph have a rainbow set F with $\nu^*(F)=n$.*

2.3. **Reachability in networks.** A *network* is a triple $\mathcal{N}=(D,S,T)$, where D is a directed graph, S ("sources") and T ("targets") are two subsets of $V=V(D)$, that in this paper are assumed to be disjoint. We write V° for $V\setminus(S\cup T)$. It is assumed that no edge goes into S or leaves T. By $E(\mathcal{N})$ we denote the edge set of D. Let $NR(\mathcal{N})$ be the set of subsets of $E(\mathcal{N})$ not containing an $S-T$ path.

Theorem 13.

$$\lambda(NR(\mathcal{N}))\leq|V^\circ|.$$

Together with Theorem 2 this implies:

Theorem 14. *Any family of $|V^\circ|+1$ $S-T$ paths has a rainbow $S-T$ path.*

This has an easy combinatorial proof: take a maximal rainbow directed forest F starting in S. If it does not reach T, then it uses at most $|V^\circ|$ paths from the family, so there exists a path P from the family that is not represented. On its way to T, P leaves F, and the edge leaving F can be used to extend F, contradicting the maximality of F.

We conclude with two more rainbow results. The first is a common generalization of Theorems 8 and 14. The known proof for it is topological, but needs more tools than just Theorem 2:

Theorem 15. *Let $\mathcal{N}$ be a network with $|V^\circ(\mathcal{N})|=q$, and let p be an integer. Let $\mathcal{F}=(F_1,\ldots,F_{2p-1+q})$ be a family of sets of edges, satisfying $\nu^P(F_i)\geq p$ for all $i\leq 2p-1+q$. Then there exists an $\mathcal{F}$-rainbow set R with $\nu^P(R)\geq p$.*

The case $q=0$ is Theorem 8, and the case $p=1$ is Theorem 14.

And a last result, from [5], again proved by both topological (see [8]) and combinatorial methods:

Theorem 16. *Any family $\mathcal{A}=(A_1,\ldots,A_n)$ of edge sets of odd cycles on a set of size n has a rainbow odd cycle. If n is even then $n-1$ odd cycles suffice.*

References

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

[2] R. Aharoni, E. Berger, M. Chudnovsky, D. Howard, P. Seymour, Large rainbow matchings in general graphs, *European J. Combin.*, 79 (2019) 222–227.

[3] R. Aharoni, J. Briggs, J. Kim and M. Kim, How many matchings are needed for a large rainbow matching, *in preparation.*

[4] R. Aharoni, P. Haxell, Hall's theorem for hypergraphs, J. Graph Theory 35.2 (2000) 83-88.

[5] R. Aharoni, R. Holzman and Z. Jiang, Rainbow fractional matchings, *Combinatorica*, to appear.

[6] I. Bárány, A generalization of Carathéodory's theorem, Discrete Math. 40 (1982) 141–152.
[7] J. Barát, A. Gyárfás, G. Sárközy, Rainbow matchings in bipartite multigraphs, Period. Math. Hung. 74 (2017) 108–111.
[8] M. K. Chari, On discrete Morse functions and combinatorial decompositions, Discrete Math. **217** (2000), 101–113.
[9] A. A. Drisko, Transversals in row-latin rectangles, J. Combin. Theory, Ser. A 84 (1998) 181–195.
[10] G. Kalai, R. Meshulam, A topological colourful Helly theorem, Adv. Math. 191 (2005) 305–311.
[11] G. Kalai and R. Meshulam, Intersections of Leray complexes and regularity of monomial ideals, Journal of Combinatorial Theory, Series A Volume **113** (2006), 1586–1592.
[12] S. Lee, *private communication.*
[13] R. Meshulam, The clique complex and hypergraph matching, *Combinatorica* **21** (2001), 89–94.
[14] G. Wegner, d-Collapsing and nerves of families of convex sets, Arch. Math. (Basel) 26 (1975) 317–321.

The hat guessing number of a graph

Noga Alon
(joint work with Omri Ben-Eliezer, Chong Shangguan, Itzhak Tamo)

Consider the following hat guessing game: n players are placed on n vertices of a graph, each wearing a hat whose color is arbitrarily chosen from a set of q possible colors. Each player can see the hat colors of his neighbors, but not his own hat color. All of the players are asked to guess their own hat colors simultaneously, according to a predetermined guessing strategy and the hat colors they see, where no communication between them is allowed. Given a graph G, its hat guessing number $\mathrm{HG}(G)$ is the largest integer q such that there exists a guessing strategy guaranteeing at least one correct guess for any hat assignment of q possible colors.

In 2008, Butler et al. [2] asked whether the hat guessing number of the complete bipartite graph $K_{n,n}$ is at least some fixed positive (fractional) power of n. We answer this question affirmatively, showing that for sufficiently large n, the complete r-partite graph $K_{n,\ldots,n}$ satisfies $\mathrm{HG}(K_{n,\ldots,n}) = \Omega(n^{\frac{r-1}{r}-o(1)})$. Our guessing strategy is based on a probabilistic construction and other combinatorial ideas, and can be extended to show that $\mathrm{HG}(\vec{C}_{n,\ldots,n}) = \Omega(n^{\frac{1}{r}-o(1)})$, where $\vec{C}_{n,\ldots,n}$ is the blow-up of a directed r-cycle, and where for directed graphs each player sees only the hat colors of his outneighbors.

Additionally, we consider related problems like the relation between the hat guessing number and other graph parameters, and the linear hat guessing number, where the players are only allowed to use affine linear guessing strategies. Several nonexistence results are obtained by using well-known combinatorial tools, including the Lovász Local Lemma and the Combinatorial Nullstellensatz [1]. Among other results, it is shown that with linear guessing functions, the hat guessing number of $K_{n,n}$ is smaller than 4, exhibiting a huge gap from the $\Omega(n^{\frac{1}{2}-o(1)})$ (nonlinear) hat guessing number of this graph.

References

[1] N. Alon, *Combinatorial Nullstellensatz*, Combin. Probab. Comput. **8** (1999), 7–29.
[2] S. Butler, M. T. Hajiaghayi, R. D. Kleinberg, and T. Leighton, *Hat guessing games*, SIAM J. Discrete Math. **22** (2008), 592–605.

Erdős-Szekeres theorem for multidimensional arrays

Matija Bucić

(joint work with Benny Sudakov and T. Tran)

A classical paper of Erdős and Szekeres [3] from 1935 is one of the starting points of a very rich discipline within combinatorics: Ramsey theory. A main result of the paper, which has become known as the Erdős-Szekeres theorem, says that any sequence of $(n-1)^2+1$ distinct real numbers contains either an increasing or decreasing subsequence of length n, and this is tight. Among simple results in combinatorics, only few can compete with this one in terms of beauty and utility. See, for example, Steele [11] for a collection of six proofs and some applications.

A very natural question which arises is how does one generalise the Erdős-Szekeres theorem to higher dimensions? The main concept which does not have an obvious generalisation is that of the monotonicity of a subsequence. Several candidates have been proposed [2, 6–10, 12] but perhaps the most natural one was introduced more than 25 years ago by Fishburn and Graham [4]. A multidimensional array is said to be monotone if for each dimension all the 1-dimensional subarrays along the direction of this dimension are increasing or are all decreasing. To be more formal, a d-dimensional array f is an injective function from $A_1 \times \ldots \times A_d$ to $\mathbb{R}$ where $A_1, \ldots, A_d$ are non-empty subsets of $\mathbb{Z}$; we say f has size $|A_1| \times \ldots \times |A_d|$.

Definition (Monotone array)**.** *A d-dimensional array $f\colon A_1 \times \ldots \times A_d \to \mathbb{R}$ is monotone if for each $i \in [d]$ one of the following alternatives occurs:*

(i) $f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ *is increasing in x for all choices of a_j's.*
(ii) $f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_d)$ *is decreasing in x for all choices of a_j's.*

For example, of the following 2-dimensional arrays first and second are monotone, while the third is not (since some rows contain increasing and some rows decreasing sequences).

7	8	9
4	5	6
1	2	3

1	3	6
2	5	7
4	8	9

7	8	9
6	5	4
1	2	3

The higher dimensional version of the Erdős-Szekeres problem introduced by Fishburn and Graham [4] now becomes: given positive integers d and n, determine the smallest N such that any d-dimensional array of size $N \times \ldots \times N$ contains a monotone d-dimensional subarray of size $n \times \ldots \times n$, we denote this N by $M_d(n)$. The Erdős-Szekeres theorem can now be rephrased as $M_1(n) = (n-1)^2 + 1$.

Fishburn and Graham [4, Section 3] showed that $M_2(n) \leq \text{towr}_5(O(n))$[1], that $M_3(n)$ is bounded by a tower of height at least a tower in n and that $M_d(n)$ is bounded from above by an Ackermann-type[2] function of order at least d for $d \geq 4$. We significantly improve upon these results.

Theorem 1.
(i) $M_2(n) \leq 2^{2^{(2+o(1))n}}$,
(ii) $M_3(n) \leq 2^{2^{(2+o(1))n^2}}$,
(iii) $M_d(n) \leq 2^{2^{2^{(1+o(1))n^{d-1}}}}$ *for* $d \geq 4$,
where the terms $o(1)$ *tend to* 0 *as* $n \to \infty$.

Fishburn and Graham introduced another very natural generalisation of the notion of monotonicity of a sequence to higher dimensional arrays. A multidimensional array is said to be *lexicographic* if for any two entries the one which has the larger position in the first coordinate in which they differ is larger. For example, the following array is lexicographic:

3	6	9
2	5	8
1	4	7

An array is said to be *lex-monotone* if it is possible to permute the coordinates and reflect the array along some dimensions to obtain a lexicographic array. To be more formal, for two vectors $\boldsymbol{u} = (u_1, \ldots, u_d)$ and $\boldsymbol{v} = (v_1, \ldots, v_d)$ in $\mathbb{R}^d$, we write $\boldsymbol{u} <_{\text{lex}} \boldsymbol{v}$ if $u_i < v_i$, where i is the smallest index such that $u_i \neq v_i$.

Definition (Lex-monotone array)**.** *A d-dimensional array f is said to be lex-monotone if there exist a permutation $\sigma\colon [d] \to [d]$ and a sign vector $\boldsymbol{s} \in \{-1, 1\}^d$ such that*

$$f(\boldsymbol{x}) < f(\boldsymbol{y}) \Leftrightarrow (s_{\sigma(1)}x_{\sigma(1)}, \ldots, s_{\sigma(d)}x_{\sigma(d)}) <_{lex} (s_{\sigma(1)}y_{\sigma(1)}, \ldots, s_{\sigma(d)}y_{\sigma(d)}).$$

Note that a 1-dimensional array is lex-monotone if and only if it is a monotone sequence. The following 2-dimensional arrays are lex-monotone since for the first one the above matrix is obtained by swapping the coordinates, for the second one by reflecting along the first dimension and for the third by performing both of these operations.

7	8	9
4	5	6
1	2	3

9	6	3
8	5	2
7	4	1

9	8	7
6	5	4
3	2	1

Given positive integers d and n, let $L_d(n)$ denote the minimum N such that for any d-dimensional array of size $N \times \ldots \times N$, one can find a lex-monotone subarray

[1]We define the tower function $\text{towr}_k(x)$ by $\text{towr}_1(x) = x$ and $\text{towr}_k(x) = 2^{\text{towr}_{k-1}(x)}$ for $k \geq 2$.

[2]The Ackermann function A_k of order k is defined recursively by $A_k(1) = 2, A_1(n) = 2n$ and $A_k(n) = A_{k-1}(A_k(n-1))$. It is an incredibly fast growing function, for example $A_2(n) = 2^n$, $A_3(n) = \text{towr}_n(2)$ and $A_4(n)$ is a tower of height tower of height tower, iterated n times, of 2.

of size $n \times \ldots \times n$. Fishburn and Graham [4, Theorem 1] showed that $L_d(n)$ exists. This result has been used to prove interesting results in poset dimension theory [5] and computational complexity theory [1].

Note that any lex-monotone array is monotone, so a very natural strategy to bound $L_d(n)$ is to first find a monotone subarray and then within this subarray find a lex-monotone subarray. This motivates the following problem which is of independent interest. For positive integers d and n, we define $F_d(n)$ to be the minimum N such that any d-dimensional monotone array of size $N \times \ldots \times N$, contains a lex-monotone subarray of size $n \times \ldots \times n$. It is easy to see by the above reasoning that $L_d(n) \leq M_d(F_d(n))$. Fishburn and Graham [4, Lemma 1] showed $F_2(n) \leq 2n^2 - 5n + 4$ and $F_3(n) \leq 2^{2n+o(n)}$, while for $d \geq 4$ their argument gives $F_d(n) \leq \mathrm{towr}_{d-1}(O_d(n))$. We determine $F_2(n)$ completely and significantly improve the bound for all $d \geq 3$.

Theorem 2.

(i) $F_2(n) = 2n^2 - 5n + 4$,

(ii) $F_d(n) \leq 2^{(c_d+o(1))n^{d-2}}$ *for* $d \geq 3$*, where* $c_d = \frac{1}{2}(d-1)!$ *and the term* $o(1)$ *tends to* 0 *as* $n \to \infty$*.*

Part (i) of Theorem 2 answers a question of Fishburn and Graham asking whether $F_2(n) = (1 + o(1))n^2$, in negative. Combining Theorems 1 and 2 with the inequality $L_d(n) \leq M_d(F_d(n))$ gives the following upper bounds on $L_d(n)$.

Theorem 3.

(i) $L_2(n) \leq 2^{2^{(4+o(1))n^2}}$,

(ii) $L_3(n) \leq 2^{2^{2^{(2+o(1))n}}}$,

(iii) $L_d(n) \leq \mathrm{towr}_5\left(O_d(n^{d-2})\right)$ *for* $d \geq 4$,

where the terms $o(1)$ *tend to* 0 *as* $n \to \infty$*.*

References

[1] M. Bodirsky and J. Kára, *The complexity of temporal constraint satisfaction problems*, J. ACM **57** (2010), Article No. 9.

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

[3] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compos. Math. **2** (1935), 463–470.

[4] P. C. Fishburn and R. L. Graham, *Lexicographic Ramsey Theory*, J. Comb. Theory Ser. A **62** (1993), 280–298.

[5] S. Felsner, P. C. Fishburn and W. T. Trotter, *Finite three dimensional partial orders which are not sphere orders*, Discrete Math. **201** (1999), 101–132.

[6] K. Kalmanson, *On a theorem of Erdős and Szekeres*, J. Comb. Theory Ser. A **15** (1973), 343–346.

[7] J. B. Kruskal, *Monotonic subsequences*, Proc. Amer. Math. Soc. **4** (1953), 264–274.

[8] N. Linial and M. Simkin, *Monotone subsequences in high-dimensional permutations*, Comb. Prob. Comput. **27** (2018), 69–83.

[9] A. P. Morse, *Subfunction structure*, Proc. Am. Math. Soc. **21** (1969), 321–323.

[10] R. Siders, *Monotone subsequences in any dimension*, J. Comb. Theory Ser. A **85** (1999), 243–253.

[11] J. M. Steele, *Variations on the monotone subsequence theme of Erdős and Szekeres.* In Discrete probability and algoritheorems 1995 (pp. 111–131). Springer, New York, NY.
[12] T. Szabó and G. Tardos, *A multidimensional generalization of the Erdős-Szekeres lemma on monotone subsequences*, Comb. Prob. Comput. **10** (2001), 557–565.

Periodic words, common subsequences and frogs

Boris Bukh
(joint work with Christopher Cox)

The longest common subsequence problem. A *word* is a finite sequence of symbols from some alphabet. We denote by $\operatorname{len} W$ the length of the word W. A *subsequence* of a word W is a word obtained from W by deleting some symbols from W; the symbols in a subsequence are not required to appear contiguously in W. A *common subsequence* between words W and W' is a subsequence of both W and W'. We denote by $\operatorname{LCS}(W, W')$ the length of the *longest common subsequence* between W and W'. We write W_i for the i'th symbol of W, with indexing starting from 0.

Throughout the paper, we use Σ to denote the alphabet, and we write $R \sim \Sigma^n$ to indicate that R is a word chosen uniformly at random from Σ^n. A long-standing problem is to understand $\operatorname{LCS}(R, R')$ for a pair of independently chosen words $R, R' \sim \Sigma^n$. Whereas it is known that

$$\mathbb{E}\operatorname{LCS}(R, R') = \gamma n + o(n) \tag{1}$$

for some constant γ depending on $|\Sigma|$, little else is known. We mention three open problems.

(1) The rate of convergence in (1) is unknown. The original proof of (1) by Chvátal and Sankoff [2] did not supply any bound on the $o(n)$ term. Alexander [1] showed that $\mathbb{E}\operatorname{LCS}(R, R') = \gamma n + O(\sqrt{n \log n})$.
(2) The value of γ, which is often called the *Chvátal–Sankoff constant*, is unknown. The best rigorous bounds for the binary alphabet are due to Lueker [4], whereas Kiwi, Loebl and Matoušek [3] gave an asymptotic for γ as $|\Sigma| \to \infty$.
(3) It is believed that $\operatorname{LCS}(R, R')$ is approximately normal, and that its variance is linear in n. Yet it is not even known that $\operatorname{Var}\operatorname{LCS}(R, R')$ tends to infinity with n.

Periodic words. A word W is *k-periodic* if $W_{i+k} = W_i$ holds for all values of i, for which both sides are defined (that is for $i = 0, 1, \ldots, \operatorname{len} W - k - 1$). For a word W of length k, write $W^{(n)}$ for the k-periodic word of length n which is obtained by repeating W the appropriate number of times (which might be fractional if k does not divide n). For example, if $W = aba$, then $W^{(8)} = abaabaab$. Additionally, write $W^{(\infty)}$ to denote the k-periodic word obtained by repeating W ad infinitum.

In attempt to solve the problems enumerated above, in this paper we tackle a simpler random variable $\operatorname{LCS}(R, W^{(n)})$ where W is a fixed word. We give answers to the analogues of all three problems. These answers are summarized in the following theorem. For a visualization of the following theorem, see Figure 1.

Theorem 1. *Let ρ be a positive real number. Fix $W \in \Sigma^k$ and let $R \sim \Sigma^n$ be an n-letter random word. Then*

$$\mathbb{E}\,\mathrm{LCS}(R, W^{(\rho n)}) = \gamma_W n - \tau_W \sqrt{n} + O(1),$$

where

(i) $\gamma_W = \gamma_W(\rho)$ *is a non-negative piecewise linear function of* ρ.
(ii) The slope of $\gamma_W(\rho)$ *is a non-increasing function of* ρ.
(iii) $\tau_W = \tau_W(\rho)$ *is nonzero only at the points where the slope of* $\gamma_W(\rho)$ *changes, and* τ_W *is strictly positive at those points.*
(iv) The random variable $\mathrm{LCS}(R, W^{(\rho n)})$ *is asymptotically normal with linear variance if and only if* $\rho > 1/|\Sigma|$, $\tau_W(\rho) = 0$ *and either*
  (a) the slope of $\gamma_W(\rho)$ *is positive, or*
  (b) there is some symbol in Σ *which does not appear in* W.
(v) There exists an algoritheorem that computes γ_W *and* τ_W *from* W.

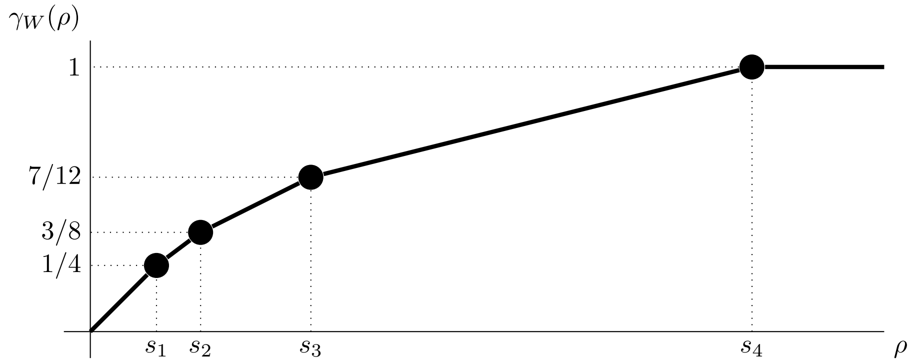

FIGURE 1. The plot of $\gamma_W(\rho)$ for $\Sigma = [4]$ and $W = 1234$. Here, $s_1 = 1/4$, $s_2 = 5/12$, $s_3 = 5/6$ and $s_4 = 5/2$. Furthermore, $\tau_W(s_1) = \sqrt{\frac{3}{512\pi}}$, $\tau_W(s_2) = \sqrt{\frac{145}{13824\pi}}$, $\tau_W(s_3) = \sqrt{\frac{79}{3456\pi}}$, $\tau_W(s_4) = \sqrt{\frac{5}{128\pi}}$ and $\tau_W(\rho) = 0$ otherwise.

From item (iii), it is clear that $\tau_W \neq 0$ happens rarely. However, it does happen for infinitely many W even in the case $\rho = 1$.

Item (iv) extends a result of Matzinger–Lember–Durringer [5], who showed that $\mathrm{Var}\,\mathrm{LCS}(R, W^{(n)})$ is linear in n when $|\Sigma| = 2$.

Frog dynamics. The key to Theorem 1 is the analysis of the following dynamical system. Let W be a fixed word, and set $k = \mathrm{len}\, W$. Imagine a circle of k lily pads, each of which is occupied by a frog. The k frogs vary from a large nasty frog to a little harmless froggie. They all face in the same (circular) direction. At each time step $t = 0, 1, \ldots$, the following happens:

(1) The monster living below pokes some of the frogs with its tentacles. Each poked frog gets agitated, and wants to jump away.
(2) In the order of descending nastiness, starting from the nastiest frog, each of the agitated frogs will leap to the next 'available' lily pad, that is either empty or occupied by a less menacing frog. Doing so causes the current occupant become agitated, and the frog that just hopped calms down.

    This process repeats until all frogs are content once more.

Below, in Figure 2, is an example of one round of this process, where here and thereafter we denote the frogs 🐸$_1$, 🐸$_2, \ldots,$ 🐸$_k$ in the order of nastiness, with 🐸$_1$ being the nastiest.

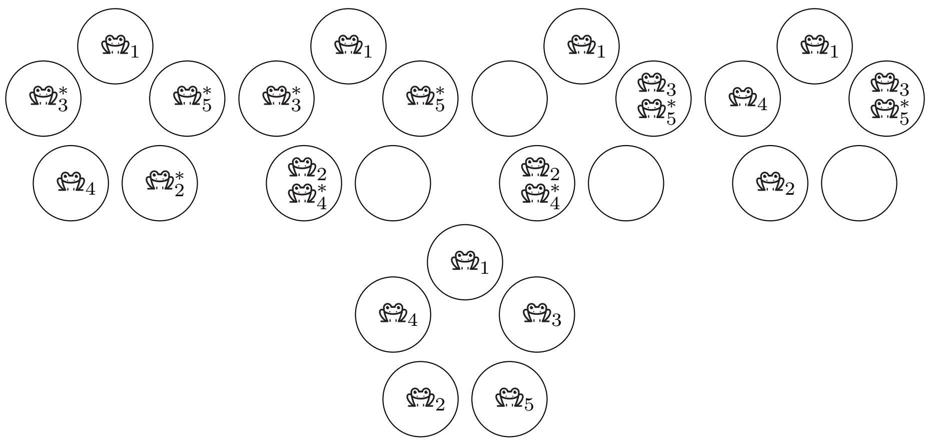

Figure 2. The sequence of frog hops resulting from poking frogs 🐸$_2$, 🐸$_3$ and 🐸$_5$. Here, frogs move in the anti-clockwise direction, and a * indicates that the frog is agitated.

References

[1] Kenneth S. Alexander. The rate of convergence of the mean length of the longest common subsequence. *Ann. Appl. Probab.*, 4(4):1074–1082, 1994.

[2] Václáv Chvátal and David Sankoff. Longest common subsequences of two random sequences. *J. Appl. Probability*, 12:306–315, 1975.

[3] Marcos Kiwi, Martin Loebl, and Jiří Matoušek. Expected length of the longest common subsequence for large alphabets. *Adv. Math.*, 197(2):480–498, 2005.

[4] George S. Lueker. Improved bounds on the average length of longest common subsequences. *Journal of the ACM*, 56(3):1–38, May 2009.

[5] Heinrich Matzinger, Jüri Lember, and Clement Durringer. Deviation from mean in sequence comparison with a periodic sequence. *ALEA Lat. Am. J. Probab. Math. Stat.*, 3:1–29, 2007.

Containers for the Maximum Weight Stable Set Problem

Maria Chudnovsky
(joint work with Tara Abrishami and Marcin Pilipczuk)

1. Introduction

An *independent set* (or *stable set*) in a simple graph G is a set $I \subseteq V(G)$ such that no edge in $E(G)$ has both endpoints in I. Given a graph G with non-negative vertex weights, the Maximum Weight Independent Set problem (MWIS) asks for an independent set of G with the greatest total weight. The MWIS problem is NP-hard in general.

For a graph G and a set F of non-adjacent vertex pairs of G, we denote by $G+F$ the graph with vertex set $V(G)$ and edge set $E(G) \cup F$. A *hole* in a graph is an induced cycle of length at least four. A graph is *chordal* if it has no holes. A set $F \subseteq \binom{V(G)}{2} \setminus E(G)$ is a *fill-in* or a *chordal completion* (of G) if $G+F$ is a chordal graph. A fill-in F is *minimal* if it is inclusion-wise minimal. Let $X \subseteq V(G)$. We say that X is a *potential maximal clique* of G if there exists a minimal chordal completion F of G such that X is a maximal clique of $G+F$.

The following theorem is the starting point of our inquiry.

Theorem 1 ([1])**.** *Given a graph G with a weight function on its vertex set and a family $\mathcal{P}$ that contains all potential maximal cliques of G, one can solve* MWIS *in G in time polynomial in $|V(G)|$ and $|\mathcal{P}|$.*

Let G be a graph and $X \subseteq V(G)$. For $s, t \in V(G) \setminus X$, we say that X is an *s,t-separator* if s and t lie in different connected components of $G - X$. An s,t-separator is a *minimal s,t-separator* if it is an inclusion-wise minimal s,t-separator. X is said to be a *minimal separator* if there exist $s, t \in V(G)$ such that X is a minimal s,t-separator in G. We say that a component D of $G \setminus X$ is a *full component for X* if every vertex of X has a neighbor in D. It is easy to see that:

Theorem 2. *Let G be a graph. $X \subseteq V(G)$ is a minimal separator if and only if at least two components of $G \setminus X$ are full for X.*

It turns out that Theorem 1 has an analogue that can be expressed using minimal separators, which is somewhat more natural from the graph-theoretic view point.

Theorem 3 ([1])**.** *Given a graph G with a weight function on its vertex set and a family $\mathcal{S}$ that contains all minimal separators of G, one can solve* MWIS *in G in time polynomial in $|V(G)|$ and $|\mathcal{S}|$.*

Let $\mathcal{C}$ be a class of graphs. We say that $\mathcal{C}$ has the *polynomial separator property* if there exists $d > 0$ such that every graph $G \in \mathcal{C}$ has at most $|V(G)|^d$ minimal separators. Thus Theorem 3 immediately implies that MWIS can be solved in polynomial time on any graph class with the polynomial separator property (assuming the list of separators can be produced in polynomial time). This fact was used in [3] and [4] to show that MWIS can be solved in polynomial time on certain subclasses of graphs with no long hole, and graphs with no even hole. Moreover, in [4] a conjecture is made characterizing the polynomial separator property by forbidden induced subgraphs.

Given a graph G and a set $I \subseteq V(G)$, we say that a chordal completion F is *I-good* if F is minimal, and each vertex pair in F is disjoint from I. In their milestone paper Lokshtanov, Vatshelle, and Villanger [5] were able to significantly strengthen Theorem 1, as follows (we rephrase the statement to fit our exposition).

Theorem 4. *Given a graph G with a weight function on its vertex set, and a family Π of subsets $V(G)$ with the property that*

- *for every maximal independent set I of G there there exists an I-good chordal completion F of G such that every maximal clique of $G+F$ belongs to Π,*

one can compute the maximum weight of an independent set in G it time polynomial in $|V(G)|$ and $|\Pi|$.

Our first result is the following significant strengthening of Theorem 4. We first introduce the notion of "safe containers". These are somewhat similar in spirit to [2] and [6]. Let G be a graph, let $I \subseteq V(G)$, and let $p > 0$ be an integer. For a set $X \subset V(G)$, we say that $C \subseteq V(G)$ is an *(I,p)-safe container* for X if $X \subseteq C$ and $|C \cap I| \leq p$. We prove:

Theorem 5. *Let $p > 0$ be an integer. Given a graph G with a weight function on its vertex set, and a family Π of subsets $V(G)$ with the property that*

- *for every maximal independent set I of G there there exists an I-good chordal completion F of G such that for every maximal clique X of $G+F$, Π contains an (I,p)-safe container for X,*

one can compute the maximum weight of an independent set in G it time polynomial in $|V(G)|$ and $|\Pi|$.

We then prove:

Theorem 6. *Let G be a graph with no hole of length at least five. Then we can construct in time polynomial in $|V(G)|$ a family Π of subsets of $V(G)$ such that*

- *for every maximal independent set I of G there there exists an I-good chordal completion F of G such that for every maximal clique X of $G+F$, Π contains an $(I,1)$-safe container for X, and*
- *$|\Pi| \leq |V(G)|^{35}$.*

Together Theorem 5 and Theorem 6 imply:

Theorem 7. *The* MWIS *Problem can be solved in polynomial time on the class of graphs with no hole of length at least five.*

One should note that Theorem 1 is not strong enough to use for the class of graphs with no hole of length at least five, because of the family of "prisms" (a prism is a graph obtained from two disjoint cliques by adding a matching between them).

To prove Theorem 6 we first observe that if I is a maximal independent set of G, F is an I-good chordal completion of G, and X is a clique of $G+F$, then $|X \cap I| \leq 1$; and by a result of [5] it is enough to handle such X that are disjoint from I. To do so we use the approach of considering minimal separators instead of maximal cliques of chordal completions. We first show:

Theorem 8. *Let G be a graph with no hole of length at least five. Then we can construct in time polynomial in $|V(G)|$ a family Σ of subsets of $V(G)$ such that*

- *for every maximal independent set I of G, and every minimal separator S of G such that $S \cap I = \emptyset$, Σ contains an $(I,0)$-safe container for X, and*
- $|\Sigma| \leq |V(G)|^{10}$.

References

[1] V. Bouchitté and I. Todinca *Treewidth and Minimum Fill-in: Grouping the Minimal Separators* SIAM J Comput. 31 (2001) 212-232.
[2] J. Balogh, R. Morris, W. Samotij, *The Method of Hypergraph Containers* Proceedings of the International Congress of Mathematicians ICM 2018, (2019) 3059-3092.
[3] M. Chudnovsky, M. Pilipczuk, M. Pilipczuk and S. theoremassé, *On the Maximum Weight Independent Set Problem in graphs without induced cycles of length at least five*, submitted for publication.
[4] M. Chudnovsky, S. Thomassé, N. Trotignon and K. Vušković, *Maximum independent sets in (pyramid, even hole)-free graphs*, submitted for publication.
[5] D. Lokshtanov, M. Vatshelle and Y. Villanger, Independent Set in P_5-Free Graphs in Polynomial Time, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algoritheorems, SODA 2014, Portland, Oregon, USA*, SIAM (2014) 570-581.
[6] D. Saxton and A. Thomason, *Hypergraph Containers* Inventiones mathematicae 201 (2015), 925–992.

Euclidean Ramsey theory

Jacob Fox
(joint work with David Conlon)

Let $\mathbb{E}^n$ denote n-dimensional Euclidean space, that is, $\mathbb{R}^n$ equipped with the Euclidean distance. Following Erdős, Graham, Montgomery, Rothschild, Spencer and Straus [5], we study the following question.

Question 1. *For which subsets $K \subset \mathbb{E}^n$ does every red/blue-coloring of $\mathbb{E}^n$ contain a red pair of points of distance one or a blue isometric copy of K?*

In what follows, we will write ℓ_m for a sequence of m points on a line with consecutive points of distance one and $\mathbb{E}^n \longrightarrow (\ell_2, K)$ if every red/blue-coloring of $\mathbb{E}^n$ contains either a red copy of ℓ_2 or a blue copy of K, where a copy of a set will always mean an isometric copy. Conversely, $\mathbb{E}^n \not\longrightarrow (\ell_2, K)$ expresses the fact that there is some red/blue-coloring of $\mathbb{E}^n$ which contains neither a red copy of ℓ_2 nor a blue copy of K.

The problem of determining which n and K satisfy the relation $\mathbb{E}^n \longrightarrow (\ell_2, K)$ has received considerable attention, with a particular focus on small values of n. For example, Erdős et al. [5] showed that $\mathbb{E}^2 \longrightarrow (\ell_2, \ell_4)$ and $\mathbb{E}^2 \longrightarrow (\ell_2, K)$ for any three-point set K. Juhász [9] later improved the latter result to cover all four-point planar sets, while just recently Tsaturian [17] improved the former result by showing that $\mathbb{E}^2 \longrightarrow (\ell_2, \ell_5)$. In three dimensions, Iván [8] showed that $\mathbb{E}^3 \longrightarrow (\ell_2, K)$ for any five-point set $K \subset \mathbb{E}^3$. The particular case where $K = \ell_5$ was recently improved by Arman and Tsaturian [1], who showed that $\mathbb{E}^3 \longrightarrow (\ell_2, \ell_6)$.

On the other hand, Csizmadia and Tóth [3] identified a set K of 8 points in the plane, namely, a regular heptagon with its center, such that $\mathbb{E}^2 \not\longrightarrow (\ell_2, K)$. This improved a result of Juhász [9], who had previously identified a set K of 12 points with the same property. Our chief concern in this paper will be with extending these results to higher dimensions by studying the smallest possible size of a set $K \subset \mathbb{E}^n$ such that $\mathbb{E}^n \not\longrightarrow (\ell_2, K)$.

In general, $|K|$ can be unbounded in terms of n and still satisfy $\mathbb{E}^n \longrightarrow (\ell_2, K)$. For example, any subset K of the unit sphere in $\mathbb{E}^n$ satisfies $\mathbb{E}^n \longrightarrow (\ell_2, K)$. Indeed, in a red/blue-coloring of $\mathbb{E}^n$, if there is no red point, then we clearly get a copy of K, while if there is a red point, then the sphere of radius one around that point must be blue, so we again get a blue copy of K.

However, our main result shows that under some mild conditions a set $K \subset \mathbb{E}^n$ such that $\mathbb{E}^n \longrightarrow (\ell_2, K)$ can have size at most exponential in n. To state the result, we say that a point set $S \subset \mathbb{E}^n$ is *t-separated* if any two points in S have distance at least t. Here and throughout, we use log to denote log base 2.

Theorem 2. *If $R > 2$ and K is a 1-separated set of points in $\mathbb{E}^n$ with diameter at most $R-1$ and $|K| > 10^{4n} \log R$, then $\mathbb{E}^n \not\longrightarrow (\ell_2, K)$.*

In particular, for $m = 10^{5n}$, we see that $\mathbb{E}^n \not\longrightarrow (\ell_2, \ell_m)$. This simple corollary is already enough to answer a problem raised by Erdős et al. [5], namely, whether, for every natural number d, there is a natural number n depending only on d such that $\mathbb{E}^n \to (\ell_2, K)$ for every $K \subset \mathbb{E}^d$. Erdős et al. state that they expect the answer to this question to be negative and our result confirms this already for $d = 1$, a special case stressed in [5], showing that n must grow logaritheoremically in the size of $|K|$.

The exponential dependence in Theorem 2, and hence the logaritheoremic dependence above, is also necessary. In fact, Szlam [15] proved the stronger result that every red/blue-coloring of $\mathbb{E}^n$ contains either a red copy of ℓ_2 or a blue *translate* of any set K of size at most $2^{c'n}$. For the sake of completeness, we include his short proof here. We will need the seminal result of Frankl and Wilson [6] that there exists a positive constant c' such that any coloring of $\mathbb{E}^n$ with at most $2^{c'n}$ colors contains a pair of points of distance one with the same color (see [14] for the current best estimate on c').

Suppose now that $K = \{k_1, \ldots, k_t\} \subset \mathbb{E}^n$ is a set of size at most $2^{c'n}$ and there is a red/blue-coloring of $\mathbb{E}^n$ with no blue copy of K. Then, for each $p \in \mathbb{E}^n$, there is at least one i such that $p + k_i$ is red, since otherwise the set $p + K$ would be a blue translate of K. We may therefore color the points of $\mathbb{E}^n$ in $t \leq 2^{c'n}$ colors, giving the point p the color i for some i such that $p + k_i$ is red, always choosing the minimum such i. By the result of Frankl and Wilson, there must then exist two points p and p' of distance one which receive the same color, say j. But then $p + k_j$ and $p' + k_j$ are two points of distance one both of which are colored red. This gives the required result. In particular, we have the following counterpart to Theorem 2, which we again stress is due to Szlam [15].

Theorem 3. *There exists a positive constant c' such that $\mathbb{E}^n \longrightarrow (\ell_2, K)$ for any set $K \subset \mathbb{E}^n$ of size at most $2^{c'n}$.*

The full version appeared in [2].

References

[1] A. Arman and S. Tsaturian, *A result in asymmetric Euclidean Ramsey theory*, Discrete Math. (2018), 1502–1508.
[2] D. Conlon and J. Fox, *Lines in Euclidean Ramsey theory*, Discrete Comput. Geom. **61** (2019), 218–225.
[3] G. Csizmadia and G. Tóth, Note on a Ramsey-type problem in geometry, *J. Combin. Theory Ser. A* **65** (1994), 302–306.
[4] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer and E. G. Straus, *Euclidean Ramsey theorems I*, J. Combin. Theory Ser. A **14** (1973), 341–363.
[5] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. H. Spencer and E. G. Straus, *Euclidean Ramsey theorems II*, in Infinite and finite sets (Colloq., Keszthely, 1973), Vol. I, 529–557, Colloq. Math. Soc. János Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.
[6] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), 357–368.
[7] P. Frankl and V. Rödl, *A partition property of simplices in Euclidean space*, J. Amer. Math. Soc. **3** (1990), 1–7.
[8] L. Iván, *Monochromatic point sets in the plane and in space*, Master's thesis, University of Szeged, 1979.
[9] R. Juhász, *Ramsey type theorems in the plane*, J. Combin. Theory Ser. A **27** (1979), 152–160.
[10] I. Kříž, *Permutation groups in Euclidean Ramsey Theory,* Proc. Amer. Math. Soc. **112** (1991), 899–907.
[11] J. Matoušek, *Lectures on discrete geometry*, Springer-Verlag, New York, 2002.
[12] J. Milnor, *On the Betti numbers of real varieties*, Proc. Amer. Math. Soc. **15** (1964), 275–280.
[13] O. A. Oleĭnik and I. B. Petrovskiĭ, *On the topology of real algebraic surfaces*, Izvestiya Akad. Nauk SSSR Ser. Mat. **13** (1949), 389–402; see also: Amer. Math. Soc. Translations **7** (1962), 399–417.
[14] A. M. Raigorodskii, *On the chromatic number of a space*, Russian Math. Surveys **55** (2000), 351–352.
[15] A. D. Szlam, *Monochromatic translates of configurations in the plane*, J. Combin. Theory Ser. A **93** (2001), 173–176.
[16] R. Thom, *Sur l'homologie des variétés algébriques réelles*, in Differential and Combinatorial Topology, 255–265, Princeton Univ. Press, Princeton, NJ, 1965.
[17] S. Tsaturian, *A Euclidean Ramsey result in the plane*, Electron. J. Combin. **24** (2017), Paper 4.35, 9 pp.

A New Bound for the Brown–Erdős–Sós Problem

Lior Gishboliner
(joint work with David Conlon, Yevgeny Levanzov, Asaf Shapira)

Extremal combinatorics, and extremal graph theory in particular, asks which global properties of a graph force the appearance of certain local substructures. Perhaps the most well-studied problems of this type are Turán-type questions, which ask for the minimum number of edges that force the appearance of a fixed subgraph F. Recall that an r-uniform hypergraph (r-graph for short) is composed of a ground set V of size n (the vertices) and a collection E of subsets of V (the edges), where each edge is of size exactly r. A (v, e)-*configuration* is a hypergraph with e edges and at most v vertices. Denote by $f_r(n, v, e)$ the largest number of edges in an r-graph on n vertices that contains no (v, e)-configuration. Estimating the asymptotic growth of this function for fixed integers r, e, v and large n is one of the most well-studied and influential problems in extremal graph theory. For example, when $e = \binom{v}{r}$ we get the well-known Turán problem of determining the maximum possible number of edges in an r-graph that contains no complete r-graph on v vertices. As another example, the case $r = 2$, $v = 2t$ and $e = t^2$ is essentially equivalent to the Zarankiewicz–Kővári–Sós–Turán problem, which asks for the maximum number of edges in a graph without a complete bipartite graph $K_{t,t}$.

Our focus in this paper is on a notorious question of this type, which emerged from work of Brown, Erdős and Sós [2,3] in the early 70's and came to be named after them. A special case of this so-called Brown–Erdős–Sós conjecture (see [5,6]) states the following:

Conjecture 1 (Brown–Erdős–Sós Conjecture)**.** *For every $e \geq 3$,*

$$f_3(n, e + 3, e) = o(n^2).$$

Despite much effort by many researchers, Conjecture 1 is wide open, having only been settled for $e = 3$ by Ruzsa and Szemerédi [12] in what is known as the $(6, 3)$-theorem. To get some perspective on the significance of this special case of Conjecture 1, suffice it to say that the famous *triangle removal lemma* (see [4] for a survey) was devised in order to prove the $(6, 3)$-theorem; that [12] was one of the first applications of Szemerédi's regularity lemma [16]; and that the $(6, 3)$-theorem implies Roth's theorem [11] on 3-term aritheoremetic progressions in dense sets of integers. As another indication of the importance of this problem, we note that one of the main driving forces for proving the celebrated hypergraph removal lemma, obtained by Gowers [7] and Rödl et al. [8–10] was the hope that it would lead to a proof of Conjecture 1.

Since we seem to be quite far[1] from proving Conjecture 1, it is natural to look for approximate versions. Namely, given $e \geq 3$, find the smallest $d = d(e)$ such

[1]As an indication of the difficulty of Conjecture 1, let us mention that the case $e = 4$ (i.e., the statement $f_3(n, 7, 4) = o(n^2)$) implies the notoriously difficult Szemerédi theorem [17,18] for 4-term aritheoremetic progressions, see [6].

that $f_3(n, e+d, e) = o(n^2)$. The best result of this type was obtained 15 years ago by Sárközy and Selkow [13], who proved that

$$f_3\left(n, e + 2 + \lfloor \log_2 e \rfloor, e\right) = o(n^2). \tag{1}$$

Since the result of [13], the only advance was obtained by Solymosi and Solymosi [15], who improved the bound $f_3(n, 15, 10) = o(n^2)$ that follows from (1) to $f_3(n, 14, 10) = o(n^2)$.

The main result of this paper, Theorem 2, gives the first general improvement over (1). Moreover, it shows that one can replace the $\lfloor \log_2 e \rfloor$ "error term" in (1) by a much smaller, sub-logaritheoremic, term.

Theorem 2. *For every $e \geq 3$,*

$$f_3\left(n, e + 18 \log e / \log\log e, e\right) = o(n^2).$$

By using asymptotic estimates for the factorial (in place of cruder bounds), one can replace the multiplicative constant 18 in the above theorem by $4 + o(1)$.

Although Theorem 2 deals with 3-graphs, its proof relies on an application of the r-graph removal lemma (see [7–10]) for *all* values of r. Employing the removal lemma for arbitrary r allows us to overcome a natural barrier which stood in the way of improving the result of [13].

As we mentioned above, Conjecture 1 has a more general form (see [1, 14]), which states that for every $2 \leq k < r$ and $e \geq 3$ we have $f_r(n, (r-k)e + k + 1, e) = o(n^k)$. However, it is a folklore observation that this more general version is in fact equivalent to the special case stated as Conjecture 1 (corresponding to $k = 2$ and $r = 3$). More precisely, it is known that

$$f_r(n, (r-k)e + k + d, e) \leq \binom{r}{3} e n^{k-2} \cdot f_3(n, e + 2 + d, e). \tag{2}$$

for every $2 \leq k < r$, $e \geq 3$ and $d \geq 1$. Setting $d = 1$ in (2) readily implies that Conjecture 1 is indeed equivalent to the general form of the Brown–Erdős–Sós conjecture stated above. By combining Theorem 2 with (2), we immediately obtain the following corollary.

Corollary 3. *For every $2 \leq k < r$ and $e \geq 3$,*

$$f_r\left(n, (r-k)e + k - 2 + 18 \log e / \log\log e, e\right) = o(n^k).$$

References

[1] N. Alon and A. Shapira, *On an extremal hypergraph problem of Brown, Erdős and Sós*, Combinatorica **26** (2006), 627–645.

[2] W. G. Brown, P. Erdős and V. T. Sós, *Some extremal problems on r-graphs*, in: New Directions in the Theory of Graphs, Proc. 3rd Ann Arbor Conference on Graph Theory, Academic Press, New York, 1973, 55–63.

[3] W. G. Brown, P. Erdős and V. T. Sós, *On the existence of triangulated spheres in 3-graphs and related problems*, Period. Math. Hungar. **3** (1973), 221–228.

[4] D. Conlon and J. Fox, *Graph removal lemmas*, in: Surveys in combinatorics 2013, 1–49, London Math. Soc. Lecture Note Ser., 409, Cambridge Univ. Press, Cambridge, 2013.

[5] P. Erdős, P. Frankl and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), 113–121.
[6] P. Erdős, *Problems and results in combinatorial number theory*, Journées aritheoremétiques de Bordeaux, Astérisque, 24–25 (1975), 295–310.
[7] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2007), 897–946.
[8] B. Nagle, V. Rödl and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures Algoritheorems **28** (2006), 113–179.
[9] V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures Algoritheorems **25** (2004), 1–42.
[10] V. Rödl and J. Skokan, *Applications of the regularity lemma for uniform hypergraphs*, Random Structures Algoritheorems **28** (2006), 180–194.
[11] K. F. Roth, *On certain sets of integers II*, J. Lond. Math. Soc. **29** (1954), 20–26.
[12] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, in Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, pp. 939–945, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam-New York, 1978.
[13] G. N. Sárközy and S. Selkow, *An extension of the Ruzsa-Szemerédi theorem*, Combinatorica **25** (2004), 77–84.
[14] G. N. Sárközy and S. Selkow, *On a Turán-type hypergraph problem of Brown, Erdős and T. Sós*, Discrete mathematics **297** (2005), 190–195.
[15] D. Solymosi and J. Solymosi, *Small cores in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **122** (2017), 897–910.
[16] E. Szemerédi, *Regular partitions of graphs.* In: Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), pp. 399–401, Colloq. Internat. CNRS, 260, CNRS, Paris, 1978.
[17] E. Szemerédi, *On sets of integers containing no four elements in aritheoremetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104.
[18] E. Szemerédi, *On sets of integers containing no k elements in aritheoremetic progression*, Acta Arith. **27** (1975), 199–245.

Non-concentration of the chromatic number

Annika Heckel
(joint work with Oliver Riordan)

1. Introduction

The study of the chromatic number of random graphs goes back to the foundational papers by Erdős and Rényi [9, 10] and includes some of the most celebrated results in random graph theory.

A lot of the past work on this topic has been focussed on finding the *likely value* of the chromatic number. Grimmett and McDiarmid [12] first established the order of magnitude of $\chi(G_{n,\frac{1}{2}})$ in 1975, and in a breakthrough paper in 1987, Bollobás [4] found the asymptotic value.

Theorem 1 ([4]). *With high probability,* $\chi(G_{n,\frac{1}{2}}) \sim \frac{n}{2\log_2 n}$.

Several improvements to these bounds were made by McDiarmid [16], Panagiotou and Steger [17] and Fountoulakis, Kang and McDiarmid [11]. The currently best known bounds for $\chi(G_{n,\frac{1}{2}})$ were obtained in [13].

Theorem 2. *Whp,*

$$\chi(G_{n,\frac{1}{2}}) = \frac{n}{2\log_2 n - 2\log_2\log_2 n - 2 + o(1)}. \tag{1}$$

A separate direction of research asked: how sharp is the *concentration* of $\chi(G_{n,p})$? While the bounds above give an explicit interval of length $o\left(\frac{n}{\log^2 n}\right)$ which contains $\chi(G_{n,\frac{1}{2}})$ whp, much narrower concentration is known to hold. A remarkable result of Shamir and Spencer [19] states that for *any* sequence $p = p(n)$, $\chi(G_{n,p})$ is whp contained in a (non-explicit) sequence of intervals of length about $\sqrt{n}$. For $p = \frac{1}{2}$, Alon improved this slightly to about $\frac{\sqrt{n}}{\log n}$ (see [18]).

For sparse random graphs, much more is known: Shamir and Spencer [19] also showed that for $p < n^{-\frac{5}{6}-\varepsilon}$, $\chi(G_{n,p})$ is whp concentrated on only five consecutive values; Łuczak [15] improved this to two consecutive values and finally Alon and Krivelevich [2] showed that two point concentration holds for $p < n^{-\frac{1}{2}-\varepsilon}$. In a landmark contribution, Achlioptas and Naor [1] found two *explicit* such values for $p = d/n$ where d is constant, and Coja-Oghlan, Panagiotou and Steger [8] extended this to three explicit values for $p < n^{-\frac{3}{4}-\epsilon}$.

However, while there is a wealth of results asserting sharp concentration of the chromatic number, until recently there were no non-trivial cases where $\chi(G_{n,p})$ was known *not* to be extremely narrowly concentrated.

In 1992, Erdős [3] asked the following question: How accurately can $\chi(G_{n,\frac{1}{2}})$ be estimated? Can it be shown *not* to be concentrated on a series of intervals of constant length? Bollobás [6] highlighted the question in 2004 and asked for any non-trivial examples of non-concentration of the chromatic number of random graphs, noting that "even the weakest results claiming lack of concentration would be of interest."

Recently [14], the author showed that $\chi(G_{n,\frac{1}{2}})$ is not whp concentrated on fewer than $n^{\frac{1}{4}-\varepsilon}$ consecutive values. We have now extended this to an almost optimal result.

Theorem 3. *For any constant $\varepsilon > 0$, there is no sequence of intervals of length $n^{\frac{1}{2}-\varepsilon}$ which contain $\chi(G_{n,\frac{1}{2}})$ with high probability.*

Up to the arbitrary constant $\varepsilon > 0$ in the exponent, this lower bound matches the upper bound from the classical result of Shamir and Spencer [19].

2. Proof sketch

The proof of Theorem 3 is based on the close relationship between the chromatic number $\chi\left(G_{n,\frac{1}{2}}\right)$ and the independence number $\alpha\left(G_{n,\frac{1}{2}}\right)$. The latter graph invariant is very well understood (see [7]): let

$$\alpha_0 = \alpha_0(n) = 2\log_2 n - 2\log_2\log_2 n + 2\log_2(e/2) + 1 \quad \text{and} \quad a = a(n) = \lfloor \alpha_0 \rfloor,$$

then whp $\alpha(G_{n,\frac{1}{2}}) = \lfloor \alpha_0 + o(1) \rfloor$, pinning down $\alpha(G_{n,\frac{1}{2}})$ to at most two consecutive values. In fact, for most n, whp $\alpha(G_{n,\frac{1}{2}}) = a$. Furthermore, if we let X_a denote the number of independent sets of size a in $G_{n,\frac{1}{2}}$, then the distribution of X_a is known to be approximately Poisson. In particular, X_a is not whp contained in any sequence of intervals shorter than $\sqrt{\mu_a}$, where $\mu_a = \mathbb{E}[X_a]$. It is not hard to show that there are some values n where μ_a is at least $n^{1-\varepsilon}$.

It is plausible that an optimal colouring of $G_{n,\frac{1}{2}}$ contains all or almost all independent a-sets, because such colourings maximise the expectation for a fixed number of colours. This intuition indicates that $\chi(G_{n,\frac{1}{2}})$ should vary at least as much as X_a (up to a log-factor). We show that this is indeed the case for *some* values n, where n is chosen so that μ_a is at least $n^{1-\varepsilon}$.

Starting with such an n, the main ingredient of the proof is a coupling of two random graphs $G \sim G_{n,\frac{1}{2}}$ and $G' \sim G_{n',\frac{1}{2}}$, where n' is slightly larger than n, so that G is an induced subgraph of G' and the vertex difference of G' and G can be partitioned into independent sets of size a. A key observation in the construction is the following: as long as $r = o(\sqrt{\mu_a})$, we can "plant" r random copies of an independent a-set in $G_{n',\frac{1}{2}}$ without changing the distribution of $G_{n',\frac{1}{2}}$ significantly.

The coupling then allows us to compare the chromatic numbers of $G_{n,\frac{1}{2}}$ and $G_{n',\frac{1}{2}}$, showing that the "typical intervals" which contain $\chi(G_{n,\frac{1}{2}})$ and $\chi(G_{n',\frac{1}{2}})$ whp cannot be too far apart. Comparing this with the growth rate of the known estimate (1) for $\chi(G_{n,\frac{1}{2}})$, and taking averages for a sequence of values n, yields that at least *some* of the "typical intervals" have to be long.

References

[1] D. Achlioptas and A. Naor. The two possible values of the chromatic number of a random graph. *Annals of Mathematics*, 162:1335–1351, 2005.

[2] N. Alon and M. Krivelevich. The concentration of the chromatic number of random graphs. *Combinatorica*, 17(3):303–313, 1997.

[3] N. Alon and J. Spencer. *The Probabilistic Method (With an Open Problems Appendix by Paul Erdős)*. Wiley, New York, first edition, 1992.

[4] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988.

[5] B. Bollobás. *Random Graphs*. Cambridge University Press, second edition, 2001.

[6] B. Bollobás. How sharp is the concentration of the chromatic number? *Combinatorics, Probability and Computing*, 13(01):115–117, 2004.

[7] B. Bollobás and P. Erdős. Cliques in random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 80, pages 419–427. Cambridge University Press, 1976.

[8] A. Coja-Oghlan, K. Panagiotou, and A. Steger. On the chromatic number of random graphs. *Journal of Combinatorial Theory, Series B*, 98(5):980–993, 2008.

[9] P. Erdős and A. Rényi. On random graphs, I. *Publicationes Mathematicae Debrecen*, 6: 290–297, 1959.
[10] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
[11] N. Fountoulakis, R. Kang, and C. McDiarmid. The t-stability number of a random graph. *The Electronic Journal of Combinatorics*, 17(1):R59, 2010.
[12] G. R. Grimmett and C. McDiarmid. On colouring random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 77, pages 313–324. Cambridge University Press, 1975.
[13] A. Heckel. The chromatic number of dense random graphs. *Random Structures & Algorithe-orems*, 53(1):140–182, 2018.
[14] Annika Heckel. Non-concentration of the chromatic number of a random graph. *arXiv preprint arXiv:1906.11808*, 2019.
[15] T. Łuczak. A note on the sharp concentration of the chromatic number of random graphs. *Combinatorica*, 11(3):295–297, 1991.
[16] C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics*, 141(1): 148–188, 1989.
[17] K. Panagiotou and A. Steger. A note on the chromatic number of a dense random graph. *Discrete Mathematics*, 309(10):3420–3423, 2009.
[18] A. Scott. On the concentration of the chromatic number of random graphs. *Available at* `arxiv.org/abs/0806.0178`, 2008.
[19] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica*, 7(1):121–129, 1987.

Homomorphisms from the torus

Matthew Jenssen
(joint work with Peter Keevash)

A central notion at the intersection of combinatorics and statistical physics is that of a *graph homomorphism*. From a combinatorial perspective, graph homomorphisms provide a unifying framework for a number of important graph theory notions. For the statistical physicist, homomorphisms arise in the analysis of *spin models* and their critical phenomena. Given graphs G and H we write $\mathrm{Hom}(G,H)$ for the set of homomorphisms from G into H. By considering different weight functions $\lambda : V(H) \to \mathbb{R}_{>0}$ we furnish ourselves with a rich set of probability distributions on the space $\mathrm{Hom}(G,H)$. Indeed, for each such λ we may define a probability measure $\mu_{H,\lambda}$ on $\mathrm{Hom}(G,H)$ given by

$$\mu_{H,\lambda}(f) = \frac{\prod_{v\in V(G)} \lambda(f(v))}{Z_G^H(\lambda)},$$

for $f \in \mathrm{Hom}(G,H)$ where

$$Z_G^H(\lambda) = \sum_{f\in \mathrm{Hom}(G,H)} \prod_{v\in V(G)} \lambda(f(v)).$$

We call this type of probability distribution a *spin model with hard constraints* and the normalising factor $Z_G^H(\lambda)$ is the *partition function* of the model.

Two of our main motivating examples will be the following:

Example 1. The hard-core model *: H is an edge $\{v_{in}, v_{out}\}$, with a loop at vertex v_{out}. We assign weights $\lambda(v_{out}) = 1$, $\lambda(v_{in}) = x$ for some fixed $x > 0$, known as the* fugacity *of the model. Here $Z_G^H(\lambda)$ is the hard-core model partition function (also known as the independence polynomial).*

Example 2. The zero-temperature q-state Potts model*: $H = K_q$, $\lambda \equiv 1$. In this case $\mu_{H,\lambda}$ is the uniform measure over proper q-colourings of G and $Z_G^H(\lambda)$ is the number of such colourings.*

In the physics literature, spin models are traditionally studied on the integer lattice $\mathbb{Z}^n$. This is a setting where the phenomena of *phase coexistence* and *phase transition* can be rigorously studied. Two landmark results in this field are due to Galvin and Kahn [4] and Peled and Spinka [6]. Galvin and Kahn establish phase coexistence for the hard-core model (Example 1) on $\mathbb{Z}^n$ at fugacity x where $x = x(n) \to 0$. Peled and Spinka exhibit phase coexistence for the zero-temperature q-state Potts model (Example 2) on $\mathbb{Z}^n$. Informally, these theorems show that both typical independent sets and typical proper q-colourings of $\mathbb{Z}^n$ exhibit large scale order by 'correlating' with some 'dominant phase'.

Here we present a detailed analysis for an arbitrary spin measure $\mu_{H,\lambda}$ on the discrete torus $\mathbb{Z}_m^n$, where m is even and n is large. The case $m = 2$ returns the n-dimensional discrete hypercube which we denote by Q_n. Our main result establishes the phase coexistence phenomenon in a strong form: it shows that $\mu_{H,\lambda}$ is close to a distribution defined constructively as a certain random perturbation of some dominant phase. This has several consequences:

(1) For any fixed (H, λ), we obtain a detailed structural decomposition of the set $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ and obtain sharp asymptotics for the partition function $Z_G^H(\lambda)$ where $G = \mathbb{Z}_m^n$. Special cases include asymptotics for the number of independent sets and the number of proper q-colourings of $\mathbb{Z}_m^n$ (so in particular, the discrete hypercube). We thereby resolve (in a strong form) conjectures of Engbers and Galvin [2, Conjectures 6.1 and 6.3] and a conjecture of Kahn and Park [5]Conjecture 5.1.
(2) We obtain central limit theorems which precisely describe the typical structure of a sample from $\mu_{H,\lambda}$.
(3) We obtain a slow mixing result for a natural class of Markov chain algoritheorems with stationary distribution $\mu_{H,\lambda}$.
(4) We describe in detail the behaviour of a typical *height function* on $\mathbb{Z}_m^n$ (that is, a homomorphism $f : \mathbb{Z}_m^n \to \mathbb{Z}$) and extend results of Benjamini, Häggström and Mossel [1] and Galvin [3].

Formally we show that the measure $\mu_{H,\lambda}$ can be well-approximated by a mixture of *polymer models* with *convergent cluster expansion* (we refer the reader to Scott and Sokal [7] for an excellent treatment of these notions and much more). Via the cluster expansion, we are able to gain an essentially complete probabilistic description of the measure $\mu_{H,\lambda}$ and hence a precise structural description of the set $\mathrm{Hom}(\mathbb{Z}_m^n, H)$.

Establishing convergence of the cluster expansion is a non-trivial task and requires a careful combination of entropy tools, the container method and algebraic and isoperimetric properties of the torus.

Rigorously stating our main result would require a detour into the theory of polymer models. Here we present some of the enumerative consequences of our main theorem which require less introduction.

Given a weighted graph (H, λ) and $A, B \subseteq V(H)$, we write $A \sim B$ if $\{a, b\} \in E(H)$ for all $a \in A$ and $b \in B$. We call such a pair (A, B) a *phase*. Letting $\lambda_X := \sum_{v \in X} \lambda_v$ for $X \subseteq V(H)$ we define

$$\eta_\lambda(H) := \max\{\lambda_A \lambda_B : A, B \subseteq V(H), A \sim B\}.$$

We call a phase (A, B) *dominant* if $\lambda_A \lambda_B = \eta_\lambda(H)$, and we let $\mathcal{D}_\lambda(H)$ denote the collection of all dominant phases. In Example 2, the dominant phases are all pairs (A, B) where $V(K_q) = A \cup B$ and $\{|A|, |B|\} = \{\lfloor q/2 \rfloor, \lceil q/2 \rceil\}$.

Theorem 1. *Let m be an even integer, and (H, λ) a weighted graph. Let $G = \mathbb{Z}_m^n$. For each $(A, B) \in \mathcal{D}_\lambda(H)$, there exists a sequence $(L_{A,B}(j))_{j \in \mathbb{N}}$ such that for any fixed $k \in \mathbb{N}$,*

$$Z_G^H(\lambda) = \eta_\lambda(H)^{\frac{m^n}{2}} \sum_{(A,B) \in \mathcal{D}_\lambda(H)} \exp\left\{ \sum_{j=1}^{k} L_{A,B}(j) + \epsilon_k \right\}$$

and $\epsilon_k = O(m^n n^{2(k-1)} \delta^{nk})$ for some $0 < \delta < 1$ depending only on (H, λ). In particular there exists a constant K depending only on (H, λ) such that $\epsilon_K = o(1)$. Moreover each term $L_{A,B}(j)$ can be computed in time $e^{O(j \ln j)}$.

The terms $L_{A,B}(j)$ are the terms of the cluster expansion (a type of multivariate Taylor series) mentioned above. We emphasise that since each term $L_{A,B}(j)$ can be computed in $e^{O(j \ln j)}$ time and $\epsilon_K = o(1)$ for some constant K (with ϵ_K as in the above theorem), computing an explicit asymptotic expression for $Z_G^H(\lambda)$ is a finite task for any fixed (H, λ). As a consequence Theorem 1 yields a plethora of asymptotic formulae for combinatorial quantities such as the number of proper q-colourings and the number of independent sets in $\mathbb{Z}_m^n$. We list some examples that can easily be obtained by hand. For a graph G, we let $i(G)$ and $c_q(G)$ denote the number of independent sets and proper q-colourings of G respectively.

Corollary 2.

$$c_5(Q_n) \sim 20\sqrt[3]{e} \cdot 6^{2^{n-1}} \exp\left\{ \left(\frac{4}{3}\right)^n \right\}$$

$$c_6(Q_n) \sim 20 \cdot 3^{2^n} \exp\left\{ \left(\frac{4}{3}\right)^n \right\}$$

$$c_7(Q_n) \sim 70 \cdot e^{n/2} 12^{2^{n-1}} \exp\left\{ \left(\frac{3}{2}\right)^{n-1} + \frac{1}{2}\left(\frac{4}{3}\right)^{n-1} + \frac{n^2 - n - 54}{108}\left(\frac{9}{8}\right)^{n-1} \right\}$$

$$c_8(Q_n) \sim 70 \cdot 4^{2^n} \exp\left\{ \left(\frac{3}{2}\right)^n + \frac{n^2 + 41n - 54}{108}\left(\frac{9}{8}\right)^n \right\}$$

$$i(\mathbb{Z}_m^n) \sim 2^{m^n/2+1} \exp\left\{\frac{1}{2}\left(\frac{m}{4}\right)^n\right\} \textit{ for } m = 4, 6, 8, 10, 12, 14$$

$$i(\mathbb{Z}_m^n) \sim 2^{m^n/2+1} \exp\left\{\frac{1}{2}\left(\frac{m}{4}\right)^n + 2n(2n-1)\left(\frac{m}{16}\right)^n\right\} \textit{ for } m = 16, 18, \ldots, 62\,.$$

We find it rather remarkable how well the two tools from statistical physics, polymer models and the cluster expansion, work with the graph container method, and we expect many further applications of this combination of methods.

References

[1] I. Benjamini, O. Häggström and E. Mossel *On random graph homomorphisms into* $\mathbb{Z}$, J. Comb. Theory B **78** (2000), 86–114.

[2] J. Engbers and D. Galvin *On the number of connected subsets with given cardinality of the boundary in bipartite graphs*, J. Comb. Theory B **102** (2012), 1110–1133.

[3] D. Galvin *On homomorphisms from the Hamming cube to* $\mathbb{Z}$, Israel J. Math. **138** (2003), 189–213.

[4] D. Galvin and J. Kahn *On Phase Transition in the Hard-Core Model on* $\mathbb{Z}^d$, Comb. Probab. Comput. **13** (2004), 137–164.

[5] J. Kahn and J. Park, *The number of 4-colorings of the Hamming cube*, arXiv:1808.01152 (2018)

[6] R. Peled and Y. Spinka, *Rigidity of proper colorings of* $\mathbb{Z}^d$, arXiv:1808.03597 (2018)

[7] A. Scott and A.D. Sokal, *The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma*, J. Stat. Phys. **118** (2005), 1151–1261.

Forbidden intersections for codes

Peter Keevash

(joint work with Noam Lifshitz, Eoin Long, Dor Minzer)

Many intersection problems for finite sets have natural generalisations to a setting variously described as codes, vectors or integer sequences. For example, any intersecting family of subsets of $[n]$ has size at most 2^{n-1}, and more generally any intersecting code in $[m]^n$ has size at most m^{n-1}, where we say a code $\mathcal{F} \subseteq [m]^n$ is intersecting if for any x, y in $\mathcal{F}$ there is some i with $x_i = y_i$. However, these settings are quite different, in that there are many maximum intersecting families of sets, including very symmetric examples such as the family of all sets of size $> n/2$, whereas in $[m]^n$ for $m > 2$ the only example is obtained by fixing one coordinate to have a fixed value. A more substantial difference was recently demonstrated by Eberhard, Kahn, Narayanan and Spirkl [4], who showed that adding a symmetry assumption reduces the maximum size to $o(m^n)$.

A longstanding open problem of Frankl and Füredi [7] posed the corresponding question for codes $\mathcal{F} \subseteq [m]^n$ that are *t-intersecting*, in that any x, y in $\mathcal{F}$ have *agreement* $\mathsf{agr}(x, y) = |\{i : x_i = y_i\}| \geq t$. From the perspective of coding theory, one may think of such $\mathcal{F}$ as an 'anti-code', in that we are imposing an upper bound on the Hamming distance between any two of its vectors. From a combinatorial perspective, the natural analogy is with t-intersecting k-graphs (k-uniform hypergraphs), for which the extremal question was also a longstanding open problem, posed by Erdős, Ko and Rado [6] and finally resolved by the Complete Intersection Theorem of Ahlswede and Khachatrian [1]. The analogous result for codes, resolving the problem of Frankl and Füredi, was also obtained by Ahlswede and Khachatrian [2], and independently by Frankl and Tokushige [8]. They showed that the maximum size of a t-intersecting code in $[m]^n$ is achieved by one of the following natural examples, which can thought of as Hamming balls on a subset of the co-ordinates, and which we simply call 'balls': let

$$\mathcal{S}_{t,r}[m]^n = \{x \in [m]^n : |\{j \in [1, t+2r] : x_j = 1\}| \geq t + r\}.$$

We show for any $m > 2$ and n large compared with t (but not necessarily m) that the same conclusion holds under the weaker assumption that $\mathcal{F}$ is $(t-1)$-*avoiding*, i.e. no x, y in $\mathcal{F}$ have agreement $t-1$.

Theorem 1. *For all $t \in \mathbb{N}$ there is $n_0 \in \mathbb{N}$ such that if $\mathcal{F} \subseteq [m]^n$ is a $(t-1)$-avoiding code with $m \geq 3$ and $n \geq n_0$ then $|\mathcal{F}| \leq \max_{r \geq 0} |\mathcal{S}_{t,r}[m]^n|$ with equality only when $\mathcal{F}$ is a ball.*

Theorem 1 can be viewed as an analogue for codes of the classical forbidden intersection problem for set systems, which has a substantial literature. Our proof (discussed in the next subsection) proceeds via a junta approximation result of independent interest, showing that any $(t-1)$-avoiding code is approximately contained in a t-intersecting junta (a code where membership is determined by a constant number of co-ordinates). In particular, when $t = 1$ this gives an alternative proof of the result of [4], as a family that essentially depends on few co-ordinates is very far from being symmetric.

The proof of Theorem 1 has three steps, each of which has elements of independent interest.

(1) Junta approximation: any $(t-1)$-avoiding code is approximately contained in a t-intersecting junta.
(2) Anticode Stability: a stability version of the Ahlswede-Khachatrian theorem on anticodes determines the structure of the junta from (1) – it must be a certain ball $\mathcal{F}$.
(3) Bootstrapping: given that the code of maximum size is close to $\mathcal{F}$, it must in fact be equal to $\mathcal{F}$.

The methods required to implement these three steps depend considerably on the size of m, and we need a variety of ideas in Combinatorics and Analysis, some of which are new. The most significant new idea in this paper is a random gluing operation, under which the alphabet size is reduced with the effect of either exhibiting junta-like structure in a code or significantly boosting its measure, and the analysis of this gluing operation via noise stability and a new hypercontractive inequality in general product spaces, which further extends our recent theory of global hypercontractivity introduced in [9]. This part of the argument can be viewed as a development of the Junta Method (see [3, 10, 11].)

The following is a precise statement of our junta approximation theorem, which is a stability theorem of independent interest, describing the approximate structure of any $(t-1)$-avoiding code with size that is within a constant factor of the maximum possible.

Theorem 2. *For every $t \in \mathbb{N}$ and $\eta > 0$ there are n_0 and J in $\mathbb{N}$ such that if $\mathcal{F} \subseteq [m]^n$ is a $(t-1)$-avoiding code with $m \geq 3$ and $n \geq n_0$ then there is a t-intersecting J-junta $\mathcal{J} \subseteq [m]^n$ such that $|\mathcal{F} \setminus \mathcal{J}| \leq \eta|\mathcal{J}|$.*

As mentioned above, Theorem 2 implies the result of [4], as a junta is far from being symmetric. The assumption $m \geq 3$ is necessary, as when $m = 2$ we have symmetric examples as mentioned above. When $m > m_0(t)$ is large we in fact obtain a more precise statement ($\mathcal{J}$ is a subcube of codimension t) and give effective estimates for the approximation parameter η.

Our first ingredient in the proof of Theorem 2 is a regularity lemma, showing that any code can be approximately decomposed into a constant number of pieces, each of which is pseudorandom, in a certain sense that depends on the size of m. When $m < m_0(t)$ is fixed and $n > n_0(t, m)$ is large, each piece is such that constant size restrictions cannot significantly affect the measure. This is a strong pseudorandomness condition, from which the proof can be completed fairly easily using a result of Mossel on Markov chains hitting pseudorandom sets. The idea is that, if two restrictions defining the regularity decomposition agree in fewer than t coordinates, then we can impose a further restriction to make them agree in exactly $t-1$ coordinates, with no significant loss in measure by pseudorandomness. If our code is $(t-1)$-avoiding these restrictions must be cross intersecting, but Mossel's result implies that this is impossible for pseudorandom codes of non-negligible measure.

When m is large, one cannot obtain such a strong pseudorandomness condition in a regularity lemma, so we settle for the weaker property of uncapturability: each piece is such that constant size restrictions cannot make the measure negligible. This makes it significantly harder to establish the t-intersection property as outlined above in the case that m is fixed, as uncapturability may not be preserved by further restrictions. Furthermore, if m is 'huge' (by which we mean exponential in n) then the cross-agreement statement used for fixed m is false. To see this, consider the codes $\mathcal{E}$ having all vectors with all coordinates even, and $\mathcal{O}$ having all vectors with all coordinates odd. There is no non-zero agreement between $\mathcal{E}$ and $\mathcal{O}$, yet they are both highly uncapturable, and have measure 2^{-n} (which is non-negligible when m is huge).

The above example naturally suggests a further case: we say m is 'moderate' if it is large but not huge. In this case, the high-level proof strategy is the same as for fixed m, although the required cross-agreement statement for uncapturable codes is difficult to prove, and this is where we need the most significant new ideas of the paper (gluing and global hypercontractivity). On the other hand, when m is huge, the above example shows that we need a different proof strategy. Here we draw inspiration from more combinatorial arguments of Keller and Lifshitz [11] which we adapt to the setting of codes by thinking of $\mathcal{F} \subseteq [m]^n$ as an n-partite n-graph (n-uniform hypergraph) with parts of size m. While the high-level strategy is similar to that in [11], the implementation is quite different; for example, the key to bootstrapping in this case turns out to be a subtle application of Shearer's entropy inequality.

We write $\mathcal{S}_{n,m,t}$ for a largest family among $\{\mathcal{S}_{t,r}[m]^n : r \geq 0\}$. From Theorem 2, we see that if a $(t-1)$-avoiding code $\mathcal{F} \subseteq [m]^n$ is at least as large as $\mathcal{S}_{n,m,t}$ then it is close to a t-intersecting junta. This raises the stability question for t-intersecting codes, which is the second ingredient in our proof of Theorem 1: must this junta be close to an extremal result? When m is large compared with t, it is not hard to show that such a junta must be close to a subcube of co-dimension t, i.e. the ball $\mathcal{S}_{t,0}[m]^n$. For fixed m, the picture is more complex, and the full range of balls can occur; nevertheless, we are able to establish the required stability version of the Ahlswede-Khachatrian anticode theorem.

Theorem 3. *For every $t \in \mathbb{N}$ and $\varepsilon > 0$ there is $\delta > 0$ such that if $\mathcal{F} \subseteq [m]^n$ is t-intersecting with $m \geq 3$ and $|\mathcal{F}| \geq (1-\delta)|\mathcal{S}_{n,m,t}|$ then $|\mathcal{F} \setminus \mathcal{S}| \leq \varepsilon|\mathcal{S}|$ for some copy $\mathcal{S}$ of $\mathcal{S}_{n,m,t} = \mathcal{S}_{t,r}[m]^n$, where $0 \leq r \leq t$, and $r = 0$ if $m > t+1$.*

The proof of Theorem 3 uses a local stability analysis of the compression operator of Ahlswede and Khachatrian [2], and also the corresponding stability result for t-intersecting families in the p-biased hypercube obtained by Ellis, Keller and Lifshitz [5].

References

[1] R. Ahlswede and L.H. Khachatrian, The complete intersection theorem for systems of finite sets, European Journal of Combinatorics 18:125–136, 1997.

[2] R. Ahlswede and L.H. Khachatrian, The Diametric Theorem in Hamming Spaces - Optimal Anticodes, Advances in Applied Mathematics 20:429–499, 1998.

[3] I. Dinur and E. Friedgut, Intersecting families are essentially contained in juntas, Combinatorics, Probability and Computing, 18:107–122, 2009.

[4] S. Eberhard, J. Kahn, B. Narayanan and S. Spirkl, On symmetric intersecting families of vectors, arXiv:1909.11578.

[5] D. Ellis, N. Keller and N. Lifshitz, Stability for the Complete Intersection Theorem, and the Forbidden Intersection Problem of Erdős and Sós, arXiv:1604.06135.

[6] P. Erdős, C. Ko and R. Rado, Intersection theorems for systems of finite sets, The Quarterly Journal of Mathematics 12:313–320, 1961.

[7] P. Frankl and Z. Füredi, The Erdős-Ko-Rado theorem for integer sequences, SIAM J. Algebraic Discrete Methods 1:376–381, 1980.

[8] P. Frankl and N. Tokushige, The Erdős–Ko–Rado theorem for integer sequences, Combinatorica 19:55–63, 1999.

[9] P. Keevash, N. Lifshitz, E. Long and D. Minzer, Hypercontractivity for global functions and sharp thresholds, arXiv:1906.05568.

[10] P. Keevash, N. Lifshitz, E. Long and D. Minzer, Hypergraph Turán via sharp thresholds, in preparation.

[11] N. Keller and N. Lifshitz, The Junta Method for Hypergraphs and the Erdős-Chvátal Simplex Conjecture, arXiv:1707.02643.

Matroid branch-depth and integer programming

Dan Král'

(joint work with Timothy F. N. Chan, Jacob W. Cooper, Martin Koutecký, Kristýna Pekárková)

Integer programming is a fundamental problem of importance in both theory and practice. It is well-known that integer programming in fixed dimension, i.e., with a bounded number of variables, is polynomially solvable since the work of Lenstra and Kannan [1,2] from the 1980's. Much subsequent research has focused on studying extensions and speed-ups of the results of Kannan and Lenstra. However, on the side of integer programs with many variables, research has been sparser. Until relatively recently, the most prominent tractable case is that of totally unimodular constraint matrices, i.e., matrices with all subdeterminants equal to 0 and ±1; in this case, all vertices of the feasible region are integral and algoritheorems for linear programming can be applied. Besides total unimodularity, many recent results on algoritheorems for integer programming exploited various structural properties of the constraint matrix yielding efficient algoritheorems for n-fold IPs, tree-fold IPs, multi-stage stochastic IPs, and IPs with bounded fracture number and bounded tree-width. This research culminated with an algoritheorem by Koutecký, Levin and Onn [3] who constructed a fixed parameter algoritheorem for integer programs with bounded (primal or dual) tree-depth and bounded coefficients.

The tree-depth of a constraint matrix depends on the position of its non-zero entries and thus does not properly reflect the true geometric structure of the integer program. In particular, a matrix with a large (dual) tree-depth may be row-equivalent to another matrix with small (dual) tree-depth that is susceptible to efficient algoritheorems. We will overcome this drawback with tools from matroid theory. To do so, we consider the branch-depth of the matroid defined by the columns of the constraint matrix and refer to this parameter as to the *branch-depth* of the matrix. Since this matroid is invariant under row operations, the branch-depth of a matrix is *row-invariant*, i.e., preserved by row operations.

We next give necessary definitions to state our results. To avoid our presentation becoming cumbersome through adding or subtracting one at various places, we the *depth* of a rooted tree to be the maximum number of edges on a path from the root to a leaf, and define the *height* of a rooted tree to be the maximum number of vertices on a path from the root to a leaf, i.e., the height of a rooted tree is always equal to its depth increased by one. The height of a rooted forest F is the maximum height of a rooted tree in F. The *closure* $\mathrm{cl}(F)$ of a rooted forest is the graph obtained by adding edges from each vertex to all its descendants. Finally, the *tree-depth* $\mathrm{td}(G)$ of a graph G is the minimum height of a rooted forest F such that the closure $\mathrm{cl}(F)$ of the rooted forest F contains G as a subgraph. It can be shown that the path-width of a graph G is at most its tree-depth $\mathrm{td}(G)$ decreased by one, and in particular, the tree-width of G is at most its tree-depth decreased by one.

A *depth-decomposition* of a matroid $M = (X, \mathcal{I})$ is a pair (T, f), where T is a rooted tree and f is a mapping from X to the leaves of T such that the number of edges of T is the rank of M and the following holds for every subset $X' \subseteq X$: the rank of X' is at most the number of edges contained in paths from the root to the vertices $f(x)$, $x \in X'$. The *branch-depth* $\mathrm{bd}(M)$ of a matroid M is the smallest depth of a tree T that forms a depth-decomposition of M. For example, if $M = (X, \mathcal{I})$ is a matroid of rank r, T is a path with r edges rooted at one of its end vertices, and f is a mapping such that $f(x)$ is equal to the non-root end vertex of T for all $x \in X$, then the pair (T, f) is a depth-decomposition of M. In particular, the branch-depth of any matroid M is well-defined and is at most the rank of M.

The *primal graph* of an $m \times n$ matrix A is the graph $G_P(A)$ with vertices $\{1, \ldots, n\}$, i.e., its vertices correspond to the columns of A, where vertices i and j are connected if the matrix A contains a row whose i-th and j-th entries are non-zero. Analogously, the *dual graph* of A is the graph $G_D(A)$ with vertices $\{1, \ldots, m\}$, i.e., its vertices correspond to the rows of A, where vertices i and j are connected if A contains a column whose i-th and j-th entries are non-zero, i.e., the dual graph $G_D(A)$ is isomorphic to the primal graph of the matrix A^T. The *primal tree-depth* $\mathrm{td}_P(A)$ of a matrix A is the tree-depth of its primal graph, the *dual tree-depth* $\mathrm{td}_D(A)$ is the tree-depth of its dual graph, and the *branch-depth* $\mathrm{bd}(A)$ of a matrix A is the branch-depth of the vector matroid formed by the columns of A; we also write $\mathrm{ec}(A)$ for the entry complexity of A. Since the vector matroid formed by the columns of A and the vector matroid formed by the columns of any matrix row-equivalent to A are the same, the branch-depth of A is invariant under row operations.

Our main results are the following.

Theorem 1. *Let A be a matrix over a field $\mathbb{F}$. The branch-depth of A is equal to the minimum dual tree-depth of a matrix A' that is row-equivalent to A, i.e., that can be obtained from A by row operations.*

Theorem 2. *For the parameterization by a positive integer d and a prime power q, there exists a fixed parameter algoritheorem that for a vector matroid M over the q-element field either outputs that* $\mathrm{bd}(M)$ *is larger than d, or outputs a depth-decomposition of M with depth d.*

Theorem 3. *For the parameterization by positive integers d and K, there exists a fixed parameter algoritheorem that for a vector matroid M over $\mathbb{Q}$ such that the entries of all vectors in M have complexity at most K either outputs that* $\mathrm{bd}(M)$ *is larger than d, or computes* $\mathrm{bd}(M)$ *and outputs a depth-decomposition of M with depth* $\mathrm{bd}(M)$.

Theorem 3 yields the following corollary, which generalize the fixed parameter algoritheorem of Koutecký, Levin and Onn [3] for the parameterization by the dual tree-depth.

Corollary 4. *There exists a computable function $g' : \mathbb{N}^2 \to \mathbb{N}$ such that integer programs with n variables and a constraint matrix A can be solved in time* *polynomial in* $g'(\mathrm{bd}(A), \mathrm{ec}(A))$ *and* n*, i.e., integer programming is fixed parameter tractable when parameterized by branch-depth and entry complexity.*

References

[1] R. Kannan, *Minkowski's convex body theorem and integer programming*, Mathematics of Operations Research **12** (1987), 415–440.

[2] H. W. Lenstra, Jr., *Integer programming with a fixed number of variables*, Mathematics of Operations Research **8** (1983), 538–548.

[3] M. Koutecký, A. Levin and S. Onn, *A parameterized strongly polynomial algoritheorem for block structured integer programs*, in: Proc. ICALP 2018, LIPIcs vol. 107, 85:1–85:14.

Universality of random permutations

Matthew Kwan

(joint work with Xiaoyu He)

A mathematical structure is said to be *universal* if it contains all possible substructures, in some specified sense. This notion may have been first considered in a 1964 paper by Rado [9], in which he found examples of graphs, simplicial complexes and functions which are universal in various ways. Another famous universal structure is a *de Bruijn sequence* (with parameters k and q, say), which is a string over a size-q alphabet in which every possible length-k string appears exactly once as a substring.

One topic that has received particular attention over the years is the case of universality for finite graphs. We say that a graph is k-universal (or k*-induced-universal*) if it contains every graph on k vertices as an induced subgraph. The problems that have received the most interest in this area are (1) to find a k-universal graph with as few vertices as possible, and (2) to understand for which n a "typical" n-vertex graph is k-universal. These problems are related to the problem of finding optimal *adjacency labeling schemes* in theoretical computer science; for more details we refer the reader to [3] and the references therein.

In an exciting recent paper by Alon [2], both of these problems were effectively resolved. He showed with a probabilistic proof that there exists a k-universal graph with $(1+o(1))2^{(k-1)/2}$ vertices, asymptotically matching a lower bound due to Moon [8]. Alon also showed that as soon as n is large enough that a random n-vertex graph typically contains a k-vertex clique and a k-vertex independent set, then such a random graph is typically also k-universal. His proofs involved a classification of graphs according to their numbers of automorphisms, taking advantage of the fact that graphs with few automorphisms are easier to embed into random graphs.

Alon's work essentially closes the book on the study of k-universal graphs, but substantial challenges remain in many other settings. One important example is the case of permutations, where there is no natural notion of an automorphism, and no natural scheme to embed sub-permutations using "quasirandomness" conditions. Let $\S_n$ be the set of all permutations of the n-element set $[n] := \{1, \dots, n\}$.

Say that a permutation $\sigma \in \S_n$ *contains* a *pattern* $\pi \in \S_k$ if there are indices $1 \le x_1 < \cdots < x_k \le n$ such that for $1 \le i, j \le k$ we have $\sigma(x_i) < \sigma(x_j)$ if and only if $\pi(i) < \pi(j)$. Say that σ is *k-universal* or a *k-superpattern* if it contains every $\pi \in \S_k$. As before, there are two main directions to consider: (1) finding the shortest possible k-universal permutation and (2) understanding for which n a typical length n permutation is k-universal.

As a simple lower bound for both problems, note that if $\sigma \in \S_n$ is k-universal, then we must have $\binom{n}{k} \ge k!$, since σ contains $k!$ distinct patterns. Using Stirling's approximation and the fact $\binom{n}{k} \le n^k/k!$, we deduce the lower bound

$$n \ge \left(\frac{1}{e^2} - o(1)\right) k^2.$$

For the first problem (of finding short k-universal permutations), this lower bound is not too far from best-possible: Miller [7] constructed a k-universal permutation with length $n \le (1/2 + o(1))k^2$, and the $o(1)$-term was recently improved by Engen and Vatter [5]. This constant 1/2 was conjectured to be tight by Eriksson, Eriksson, Linusson and Wästlund [6], while the constant $1/e^2$ from the lower bound was conjectured to be tight by Arratia [4].

Regarding universality of *random* permutations, much less is known. Note that containing the identity permutation $1_k \in \S_k$ is equivalent to containing an increasing sequence of length k, and the longest increasing subsequence of a typical $\sigma \in \S_n$ is known to be of length $(2 + o(1))\sqrt{n}$. It follows that we cannot hope for a typical $\sigma \in \S_n$ to be k-universal unless $n \ge (1/4 + o(1))k^2$. In 1999, Alon made the following striking conjecture (see [1, 4]).

Conjecture 1. *For a fixed $\varepsilon > 0$, a random permutation of length $(1+\varepsilon)k^2/4$ is w.h.p.*[1] *k-universal.*

Intuitively, Conjecture 1 can be justified by comparison to universality in graphs: in much the same way that cliques and independent sets are the "hardest" subgraphs to find in a random graph, it is believed that monotonically increasing and decreasing patterns are the hardest patterns to find in a random permutation. We also remark that Conjecture 1 contradicts the aforementioned conjecture by Eriksson, Eriksson, Linusson and Wästlund.

In the "ordered" setting of random permutations, most of the standard tools used in the unordered setting of graphs are not applicable, and Conjecture 1 seems rather challenging to prove. Indeed, in a recent discussion of the problem, Alon [1] highlighted the more modest problem of simply showing that for $n = 1000k^2$ a typical $\sigma \in \S_k$ is k-universal. He also observed a simple upper bound of the form $n = O(k^2 \log k)$. Our main result is the following substantial improvement.

Theorem 2. *A random permutation of length $2000k^2 \log\log k$ is w.h.p. k-universal.*

[1] We say that an event holds "with high probability", or "w.h.p." for short if it holds with probability $1 - o(1)$. Here and for the rest of the paper, asymptotics are as $k \to \infty$ and/or $n \to \infty$.

Since there is no natural notion of symmetry for permutations, we were not able to take quite the same approach as Alon took for the graph case. However, the proof of Theorem 2 still proceeds via a "structure-vs-randomness" dichotomy (see [10] for a discussion of this phenomenon in general). In our proof of Theorem 2 we show that every $\pi \in \mathcal{S}_k$ can be decomposed into a "structured part" and a "quasirandom part". The "structured part" of π is likely to appear in σ for one reason, and the "quasirandom part" is likely to appear for a different reason. It is worth mentioning here that because *most* permutations are entirely quasirandom in our sense, the following theorem also follows from our proof approach.

Theorem 3. *For any $k \geq 1$, there is a set $\mathcal{Q}_k \subseteq \mathcal{S}_k$ of $(1 - o(1))\, k!$ length-k permutations such that w.h.p. a random permutation of length $20k^2$ contains every $\pi \in \mathcal{Q}_k$.*

References

[1] N. Alon, *When are random permutations k-universal?, Combinatorics and probability*, Oberwolfach Rep. **13** (2016), no. 2, 1189–1257, Abstracts from the workshop held April 17–23, 2016, Organized by Béla Bollobás, Michael Krivelevich, Oliver Riordan and Emo Welzl.

[2] N. Alon, *Asymptotically optimal induced universal graphs*, Geom. Funct. Anal. **27** (2017), no. 1, 1–32.

[3] S. Alstrup, H. Kaplan, M. Thorup, and U. Zwick, *Adjacency labeling schemes and induced-universal graphs*, Proceedings of the forty-seventh annual ACM symposium on Theory of computing, ACM, 2015, pp. 625–634.

[4] R. Arratia, *On the Stanley-Wilf conjecture for the number of permutations avoiding a given pattern*, Electron. J. Combin. **6** (1999), Note, N1, 4.

[5] M. Engen and V. Vatter, *Containing all permutations*, arXiv preprint arXiv:1810.08252 (2018).

[6] H. Eriksson, K. Eriksson, S. Linusson, and J. Wästlund, *Dense packing of patterns in a permutation*, Ann. Comb. **11** (2007), no. 3-4, 459–470.

[7] A. Miller, *Asymptotic bounds for permutations containing many different patterns*, J. Combin. Theory Ser. A **116** (2009), no. 1, 92–108.

[8] J. Moon, *On minimal n-universal graphs*, Glasgow Mathematical Journal **7** (1965), no. 1, 32–33.

[9] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.

[10] T. Tao, *Structure and randomness in combinatorics*, 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), IEEE, 2007, pp. 3–15.

What do we know about the large-scale geometry of graphs?

Nati Linial

In my talk I have briefly addressed three topics concerning the global geometry of graphs, and its relationship with differential geometry on manifolds.

I: Whereas we know quite a bit about eigenvalues of graphs, much less is known about graphs' *eigenfunctions*. There is a considerable body of work concerning the *nodal domains* of Laplacians of smooth manifolds. This notion has a very natural graph theoretic counterpart. Let $G = (V, E)$ be a graphs and $f : V \to \mathbb{R}$ a real function. We denote by $C_f \subseteq E$ the set of those edges $xy \in E$ on which f *crosses* *zero* i.e., $f(x)f(y) < 0$. The nodal domains of f are the connected components of $G \setminus C_f$. We proved [1]:

Theorem 1. *For every $1 > p > 0$ and asymptotically almost every $G \in G(n,p)$, every eigenfunction of G's adjacency matrix has at most K nodal domains.*

The original proof gave $K \approx 40$ and combined with later work of Van Vu we know that the theorem holds with $K = 2$. I called attention to intriguing numerical results in [1] concerning these problems for random *regular* graphs, for which we presently have no theoretical explanation.

II: What is the right notion of geodetic paths in graphs? A *consistent path system* in a graph G has, for every two vertices $x, y \in V(G)$ a path P_{xy} between x and y so that the following condition holds: Whenever vertices u, v belong to the path P_{xy}, the subpath of P_{xy} between u and v coincides with P_{uv}. If every edge $e \in E$ has weight $w_e > 0$ and every P_{xy} is a w-shortest path between x and y, this clearly yields a consistent path system in G. A path system which cannot be derived from any set of weights w is called *non-additive*. In ongoing work with Daniel Cizma we proved:

Theorem 2. *Asymptotically almost every graph has a non-additive path system.*

III: The trade-off between *girth* and *diameter* in graphs is fascinating and mysterious. I recall the following old problem of mine:

Problem 3. *What is*

$$\limsup \frac{girth(G)}{diam(G)}$$

over all graphs in which every vertex has degree ≥ 3. To the best of my knowledge we only know at present that this number is in $[1, 2]$.

Even more remarkably, we do not even know

Problem 4. *Can*

$$girth(G) - diam(G)$$

be arbitrarily large for a graph G in which every vertex has degree ≥ 3? I know of (specific) graphs in the literature for which this difference is 6, but nothing larger.

I have mentioned very recent work with Michael Simkin [2] where we give a randomized construction of d-regular n-vertex graphs G with $\mathrm{girth}(G) > c \log_{d-1} n$ for any $1 > c > 0$.

References

[1] Y. Dekel, J. R. Lee, and N. Linial, Eigenvectors of random graphs: nodal domains, Random Structures Algoritheorems 39 (2011), no. 1, 39–58,.

[2] N. Linial and M. Simkin, A randomized construction of high girth regular graphs, arXiv:1911.09640

A loose short talk on loose short paths

Tomasz Łuczak
(joint work with Joanna Polcyn and Christian Reiher)

Let P_2^4 denote the 4-uniform hypergraph which consists of two edges sharing exactly one vertex. We shall study the asymptotic behaviour of the function $f_2^4(n,m)$, defined as the minimum of the maximum degrees of all P_2^4-free 4-uniform hypergraphs with n vertices and m edges. Moreover, we shall investigate the structure of extremal hypergraphs $\mathcal{F}_2^4(n,m)$, i.e. the P_2^4-free 4-uniform hypergraphs with n vertices, m edges, and maximum degree $f_2^4(n,m)$.

Let us recall that the well-known result of Frank and Füredi [1] (see also Keevash, Mubayi, and Wilson [3]) implies that, for n large enough, each P_2^4-free 4-uniform hypergraph on n vertices has at most $\binom{n-2}{2}$ edges, and this maximum is achieved only for (complete) 2-stars, i.e. hypergraphs whose edges are 4-element sets containing a given 2-element set of vertices. Thus, when m is close to its largest possible value, say when $m = \Omega(n^2)$, it is natural to expect that all hypergraphs from $\mathcal{F}_2^4(n,m)$ consist, basically, of a number of almost vertex-disjoint 2-stars, i.e. that for $m = \Omega(n^2)$ we have $f_2^4(n,m) = \Theta(n^2)$. However, after a moment of reflection, one realizes that 2-star forests are not the only P_2^4-free hypergraphs with $\Omega(n^2)$ edges. Another examples are provided by 4-uniform hypergraphs which are '2-blow-ups' of graphs in the following sense. Given a graph G, the 4-uniform hypergraph $B_2[G]$ is obtained by replacing each vertex of a graph G by a pair of vertices, called twins, so each edge of G becomes a hyperedge of $B_2[G]$. Note that the 'thick clique' $B_2[K_{\lfloor n/2\rfloor}]$ has $\binom{\lfloor n/2\rfloor}{2}$ edges but its maximum degree grows only linearly. It turns out that $B_2[K_{\lfloor n/2\rfloor}]$ is indeed the densest P_2^4-free hypergraph with linear maximum degree.

Theorem 1 (Łuczak and Polcyn [4])**.** *There exists $\bar{n}_1$ such that, for every $n \geq \bar{n}_1$ and*

$$\binom{\lfloor n/2\rfloor}{2} - \frac{n}{5} \leq m \leq \binom{\lfloor n/2\rfloor}{2},$$

we have $f_2^4(n,m) = \lfloor n/2\rfloor - 1$ and each hypergraph from $\mathcal{F}_2^4(n,m)$ is a subhypergraph of $B_2[K_{\lfloor n/2\rfloor}]$.

Moreover, there exists $\tilde{n}_1$ such that, for every $n \geq \tilde{n}_1$ and all $m \geq \binom{\lfloor n/2\rfloor}{2} + 1$, each hypergraph from $\mathcal{F}_2^4(n,m)$ has maximum degree at least $n^2/32 - n/8$, and one can delete from it at most 128 edges to obtain a union of at most four vertex-disjoint 2-stars and some number of isolated vertices.

Thus, if we rescale the function $f_2^4(n,m)$ setting for $x \in (0,1)$

$$\hat{f}_2^4(x) = \lim_{n\to\infty} \frac{f_2^4(n, x\binom{n-2}{2})}{\binom{n-2}{2}},$$

then $\hat{f}_2^4(x)$ is discontinuous at the point $x = 1/4$.

The key ingredient of the proof of Theorem 1 is the following, rather surprising, structural result, which states that each P_2^4-free 4-hypergraph, no matter how dense it is, consists of a blow-up $B_2[G]$ of some graph G, a family of 2-stars centered at pairs corresponding to the vertices of G, and at most $4n$ other edges.

Theorem 2 (Łuczak and Polcyn [4]). *For any P^4_2-free 4-uniform hypergraph $H = (V, E)$ there exists a partition of its set of vertices $V = R \cup S \cup T$, such that subhypergraphs of H defined as $H_R = \{h \in H : h \cap R \neq \emptyset\}$, $H_S = H[S]$ and $H_T = H \setminus (H_R \cup H_S)$ satisfy the following conditions:*

(i) $|H_R| \leq 4|R|$,
(ii) *H_S is a subgraph of a thick clique,*
(iii) *H_T is a family of vertex-disjoint 2-stars such that their centers are twins of H_S, whereas all other vertices are in T.*

Note that some small P^4_2-free hypergraphs, such as the one which consists of the faces of a cube, are neither 2-stars no subsets of a thick clique, and so in the above statement one cannot get rid of the 'unrestrained' edges from H_R. Nonetheless, in each P^4_2-free hypergraph H there exists a subset of vertices R which is intersected only by at most $4|R|$ edges, such that the hypergraph obtained from H by removing all vertices from R consists only of the 2-blow-up of a graph and 2-stars centered on its twins.

Similar results can be proved also for 3-uniform graphs not containing a loose path P^3_3, which is the only connected linear 3-uniform hypergraph on seven vertices of maximum degree two. In this case the P^3_3-free 3-uniform hypergraph which maximizes the number of edges is a star with $\binom{n-1}{2}$ edges, provided $n \geq 8$ (see Jackowska, Polcyn, Ruciński [2]). On the other hand, for each bipartite graph G with bipartition $V_1 \cup V_2$, the 3-uniform graph $B_{3/2}[G]$, obtained from G by replacing vertices of V_1 by twins, contains no copy of P^3_3. Note that $B_{3/2}[K_{n/4,n/2}]$ has $n^2/8$ edges and maximum degree $n/2$. In [4] we proved theorem which is an exact analogue of Theorem 1 and states, in particular, that each P^3_3-free 3-uniform hypergraph with more than $n^2/8$ edges has maximum degree $\Omega(n^2)$. Furthermore, Theorem 2 has its counterpart also in this case, since each P^3_3-free 3-uniform hypergraph consists of $B_{3/2}[G]$ for some bipartite graph G, some number of disjoint stars centered on 'singleton' vertices of $B_{3/2}[G]$, and at most $4n$ other edges.

In the paper of Łuczak, Polcyn, and Reiher [5] Theorems 1 and 2 were generalized to certain class of k-uniform hypergraphs in which some intersections of edges are prohibited. Since the statements of these results are quite technical, we only illustrate them by two examples, both concerning 16-uniform hypergraphs. Consider first the family $\mathcal{H}$ of 16-uniform hypergraphs, where for each pair of edges e and f

$$|e \cap f| \neq 1, 3, 5, 7.$$

Then there are two natural families of graphs of $\Theta(n^8)$ edges in $\mathcal{H}$. One consists of 8-stars, i.e. hypergraphs in which each edge contains a given subset of 8 vertices. The other contains naturally defined '2-blow-ups' of 8-uniform hypergraphs. It turns out that these two types of 16-uniform hypergraphs are the main building blocks of every dense hypergraph from the family $\mathcal{H}$.

For the second example let us define $\mathcal{H}'$ as the family of all 16-uniform hypergraphs such that for each pair of edges e and f we have

$$|e \cap f| \neq 1, 2, 3, 5, 6, 7, 9, 10, 11.$$

Our structural result states that every dense hypergraph from $\mathcal{H}'$ consists of the 4-blow-up $B_4[F]$ of a 4-uniform hypergraph F, some 12-stars centered at the vertices of $B_4[F]$, and a small number of unrestrained edges.

Finally, we remark that we do not know which prohibited families of subhypergraphs $\mathcal{G}$ force a nice structure in $\mathcal{G}$-free hypergraphs, similar to that described in Theorem 2. For instance, if we consider the family $\mathcal{H}''$ of all 16-uniform hypergraphs in which for every pair of edges e and f we have

$$|e \cap f| \neq 5, 6, 7, 9, 10, 11,$$

then 12-stars and 4-blow-ups of 4-uniform hypergraphs are not the only types of members of $\mathcal{H}''$ with density $\Theta(n^4)$. We can also take any $(4, i, n)$ Steiner system $\mathcal{S}_i$ with $i = 16, 17, 18, 19, 20$, and consider the hypergraph whose edges are 16-element subsets of blocks of $\mathcal{S}_i$. Thus, since there are at least seven essentially different kinds of dense hypergraphs from $\mathcal{H}''$, it is hard to expect a simple structural characterization of all hypergraphs from this family.

References

[1] P. Frankl and Z. Füredi, *Forbidding just one intersection*, J. Combinat. Th. Ser. A **36** (1985), 160–176.

[2] E. Jackowska, J. Polcyn, A. Ruciński, *Turán numbers for 3-uniform linear paths of length 3*, Electron. J. Combin. **23** (2016), #P2.30.

[3] P. Keevash, D. Mubayi, R. M. Wilson, *Set systems with no singleton intersection*, SIAM J. Discrete Math. **20** (2006), 1031–1041.

[4] T. Łuczak, J. Polcyn, *Paths in hypergraphs: a rescaling phenomenon*, SIAM J. Discrete Math., **33** (2019), 2251–2266.

[5] T. Łuczak, J. Polcyn, C. Reiher, *A tale of stars and cliques*, J. Combinat. Th. Ser. A **160** (2018), 111–135.

Symmetric intersecting families of vectors

Bhargav Narayanan

(joint work with Sean Eberhard, Jeff Kahn and Sophie Spirkl)

A family $\mathcal{A} \subset [k]^n$ is said to be *intersecting* if any two of its elements agree on at least one coordinate. Consideration of the orbits of the natural $\mathbb{Z}/k\mathbb{Z}$ action on $[k]^n$, i.e., the orbits of the map that shifts each coordinate cyclically by one, shows that any intersecting subfamily of $[k]^n$ has size at most k^{n-1}; furthermore, this bound is tight for the trivial family obtained by specifying the value of some fixed coordinate. These observations go back to Berge and Livingtson in the 1970's, and many (more substantial) generalisations are now known.

What changes when we forbid such extremal examples that are highly asymmetric, membership in which is determined by a single coordinate of outsized influence? This general line of questioning was introduced by Babai in the late

1970's, and we provide an answer in this particular instance: for fixed $k \geq 3$, if $\mathcal{A} \subset [k]^n$ is intersecting and admits a transitive automorphism group, then $|\mathcal{A}| = o(k^n)$. The requirement that $k \geq 3$ is necessary: in the Boolean hypercube $[2]^n$ with n odd, the family of vectors with more 1's than 2's is intersecting, of the maximum possible size 2^{n-1}, and invariant under the whole symmetric group.

Perhaps surprisingly, this simple and natural theorem seems resistant to elementary proof, and it may be that the more important point is the contribution to methodology. Ellis and I, in resolving an old conjecture of Frankl on symmetric 3-wise intersecting families, introduced the use of spectral machinery for tackling problems in extremal set theory involving symmetry; this framework has since been successfully adapted to resolve a few other rather old problems. However, this approach depends crucially on the interplay between up-sets, biased product measures, and 'sharp threshold' behaviour, all features absent from the problem under consideration here; for example, all of the previous work starts with the elementary observation that the p-biased measure of an up-set in $[2]^n$ is monotone increasing in p, but even this fact that has no useful analogue in $[k]^n$ for $k \geq 3$. One could, for example, try working in $[k]^n$ with the natural product order, but one is then confronted with the following: compressing an intersecting family 'upwards' preserves the intersection condition but not symmetries, while replacing a family by its 'up-closure' preserves symmetries but not the intersection condition.

Our way around these obstacles is to embed $[k]^n$ in a larger 'covering space', a suitable product of posets, in which up-closure avoids the above difficulties and appropriate analogues of biased product measures still provide the leverage we need. Having transferred our problem to this larger space, we deduce our result from a suitable variant of the sharp threshold theorem tailored to this covering space.

Singularity and universality of random integral matrices

Hoi H. Nguyen
(joint work with Melanie M. Wood)

The singularity problem in combinatorial random matrix theory states that if a square matrix $M_{n\times n}$ of size n is "sufficiently random", then $M_{n\times n}$ is non-singular asymptotically almost surely as n tends to infinity. In other words $p(M_{n\times n})$, the probability of $M_{n\times n}$ being singular, tends to 0.

This problem has a rich history. In the early 60s Komlós [8] showed that if the entries of $M_{n\times n}$ take values $\{0, 1\}$ independently with probability 1/2 then $p = O(n^{-1/2})$ [1]. This bound was significantly improved to $(1-\varepsilon)^n$ (for an explicit $0 < \varepsilon < 1/2$) by Kahn, Komlós and Szemerédi [7] in the 90's, and subsequently by Tao and Vu [15], by Rudelson and Vershynin [12], and by Bourgain, Vu and Wood [1] in the last decade. More recently Tikhomirov [16] showed that $p(M_{n\times n}) = (\frac{1}{2} + o(1))^n$, which is asymptotically optimal.

[1] Most of the results here hold for far more general matrix models.

As $M_{n\times n}$ has integral entries, these results imply that with very high probability the linear map $M_{n\times n} : \mathbb{Z}^n \to \mathbb{Z}^n$ is injective. Another important property of interest is *surjectivity*, it seems natural to wonder if with high probability $M_{n\times n} : \mathbb{Z}^n \to \mathbb{Z}^n$ is surjective. However, it is not very hard to show that the surjectivity probability for $M_{n\times n}$ goes to 0 with n. The main result of this report shows that when the matrix has more columns than rows, e.g. $M_{n\times(n+1)} : \mathbb{Z}^{n+1} \to \mathbb{Z}^n$, we have surjectivity with positive probability strictly smaller than one.

We make the following definition to restrict the types of entries our random matrices will have. We say a random integer ξ_n is α_n*-balanced* if for every prime p we have

$$\max_{r\in\mathbb{Z}/p\mathbb{Z}} \mathbf{P}(\xi_n \equiv r \pmod{p}) \leq 1 - \alpha_n. \tag{1}$$

Our main result below is not only about whether $M_{n\times(n+u)}$ is surjective, but also about the cokernel group $\mathbf{Cok}(M_{n\times(n+u)})$, which is the quotient group $\mathbb{Z}^n/M_{n\times(n+u)}(\mathbb{Z}^{n+u})$.

Theorem 1. *For integers $n, u \geq 0$, let $M_{n\times(n+u)}$ be an integral $n\times(n+u)$ matrix with entries i.i.d copies of an α_n-balanced random integer ξ_n, with $\alpha_n \geq n^{-1+\varepsilon}$ and $|\xi_n| \leq n^T$ for any fixed parameters $0 < \varepsilon < 1$ and $T > 0$ not depending on n. For any fixed finite abelian group B and $u \geq 0$,*

$$\lim_{n\to\infty} \mathbf{P}\Big(\mathbf{Cok}(M_{n\times(n+u)}) \simeq B\Big) = \frac{1}{|B|^u|\mathbf{Aut}(B)|} \prod_{k=u+1}^{\infty} \zeta(k)^{-1}$$

and

$$\lim_{n\to\infty} \mathbf{P}\Big(\mathbf{Cok}(M_{n\times(n+u)}) \text{ is cyclic}\Big) = \prod_{p \text{ prime}} (1 + p^{-(u+1)}(p-1)^{-1}) \prod_{k=u+2}^{\infty} \zeta(k)^{-1}.$$

Here $\zeta(s)$ is the Riemann zeta function. In particular, as $n \to \infty$, the map $M_{n\times(n+1)} : \mathbb{Z}^{n+1} \to \mathbb{Z}^n$ is surjective with probability approaching $\prod_{k=2}^{\infty} \zeta(k)^{-1} \approx 0.4358$. By this result, the one extra dimension mapping to $\mathbb{Z}^n$ brought the surjectivity probability from 0 to ≈ 0.4358.

For $u = 0$, the cyclicity probability has been seen in several papers studying the probability that a random lattice in $\mathbb{Z}^n$ is co-cyclic (gives cyclic quotient), in cases when these lattices are drawn from the nicest, most uniform distributions, e.g. uniform on lattices up to index X with $X \to \infty$ [2, 10, 11], or with basis with uniform entries in $[-X, X]$ with $X \to \infty$ [14]. Stanley and Wang have asked whether the probability of having cyclic cokernel is universal (see [14, Remark 4.11 (2)] and [13, Section 4]). Theorem 1 proves this universality, showing that the same probability of cocylicity occurs when the lattice is given by n random generators from a rather large class of distributions, including ones that are rather distorted mod p for each prime p.

To give a heuristic for why inverse zeta values arise in these probabilities, note that $M_{n\times(n+1)}$ is surjective if and only if its reduction to modulo p is surjective for all primes p. We then make two idealized heuristic assumptions on $M_{n\times(n+1)}$.

(i) (uniformity assumption) Assume that for each prime p the entries of $M_{n\times(n+1)}$ are uniformly distributed modulo p. In this case, a simple calculation gives the probability for $M_{n\times(n+1)}$ being surjective modulo p is $\prod_{j=2}^{n}(1-p^{-j})(1-p^{-n-1})$.
(ii) (independence assumption) We next assume that the statistics of $M_{n\times(n+1)}$ reduced to modulo p are asymptotically mutually independent for all primes p. Under these assumptions, as $n\to\infty$, the probability that $M_{n\times(n+1)}$ is surjective would be asymptotically the product of all of the surjectivity probability modulo p, which leads to the number $\prod_{k=2}^{\infty}\zeta(k)^{-1}$ as seen.

The matrices in this report do not have to satisfy either assumption, and indeed they can violate them dramatically. Also, our main results work for $\alpha_n \geq n^{-1+\varepsilon}$, which is asymptotically best possible, in terms of the exponent of n.

Moreover, we show the same results hold if we replace $\mathbf{Cok}(M_{n\times(n+1)})$ with the total sandpile group of an Erdős-Rényi simple random digraph, proving a conjecture of Koplewitz [9, Conjecture 1]. Let $M = M_{n\times n} = (x_{ij})_{1\leq i,j\leq n}$ be a random matrix where $x_{ii} = 0$ and its off-diagonal entries are i.i.d. copies of an integral random variable ξ_n satisfying (1). Let $L_M = (L_{ij})$ be the Laplacian of M, that is $L_{ij} = -x_{ij}$ if $i\neq j$ and $L_{ii} = \sum_{k=1}^{n} x_{ki}$. We then denote S_M to be the cokernel of L with respect to the group $\mathbb{Z}_0^n$ of integral vectors of zero entry-sum $S_M = \mathbb{Z}_0^n/L_M\mathbb{Z}^n$. When M is the adjacency matrix of a directed graph, this group has been called the *sandpile group without sink* [5] and the *total sandpile group* [9] of the graph.

Theorem 2. *Let $0<\varepsilon<1$ and $T>0$ be given. Let $M_{n\times n}$ be a integral $n\times n$ matrix with entries i.i.d copies of an α_n-balanced random integer ξ_n, with $\alpha_n \geq n^{-1+\varepsilon}$ and $|\xi_n|\leq n^T$. Then for any finite abelian group B,*

$$\lim_{n\to\infty}\mathbf{P}\Big(S_{M_{n\times n}}\simeq B\Big) = \frac{1}{|B||\mathbf{Aut}(B)|}\prod_{k=2}^{\infty}\zeta(k)^{-1}$$

and

$$\lim_{n\to\infty}\mathbf{P}\Big(S_{M_{n\times n}} \text{ is cyclic}\Big) = \prod_{p \text{ prime}}(1+(p^2(p-1))^{-1})\prod_{k=3}^{\infty}\zeta(k)^{-1}.$$

Applied to the Erdős-Rényi random directed graph model where each directed edge is chosen independently with probability q satisfying $\alpha_n\leq q\leq 1-\alpha_n$, via [6, Theorem 5] our result shows that with probability approximately .4358 the graph has the property that any legal chip configuration σ (that is $|\sigma|\leq|E|-|V|$) stabilizes after a finite number of legal firings.

References

[1] J. Bourgain, Van H. Vu and P. Wood, On the singularity probability of discrete random matrices. *J. Funct. Anal.*, 258(2):559–603, 2010.

[2] G. Chinta, N. Kaplan, and S. Koplewitz, The cotype zeta function of $\mathbb{Z}^d$, preprint, *arxiv.org/abs/1708.08547*.

[3] J. Clancy, N. Kaplan, T. Leake, Sam Payne and M. M. Wood, On a Cohen–Lenstra heuristic for Jacobians of random graphs, *Journal of Algebraic Combinatorics*, pages 1–23, May 2015.

[4] H. Cohen and H. W. Lenstra, Jr., Heuristics on class groups of number fields. In Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), volume 1068 of Lecture Notes in Math., pages 331–762. Springer, Berlin, 1984.

[5] M. Farrell and L. Levine, Co-Eulerian graphs, *Proceedings of the American Mathematical Society*, 144 (2016), 2847–2860.

[6] M. Farrell and L. Levine, Multi-Eulerian tours of directed graphs, *Electron. J. Combin.*, 23 (2016), no. 2, Paper 2.21, 7 pp.

[7] J. Kahn, J. Komlós and E. Szemerédi, On the probability that a random ±1-matrix is singular, *J. Amer. Math. Soc.*, 8(1):223–240, 1995.

[8] J. Komlós, On the determinant of (0, 1) matrices, *Studia Sci. Math. Hungar*, 2:7–21, 1967.

[9] S. Koplewitz, Sandpile groups and the coeulerian property for random directed graphs, *Advances in Applied Mathematics*, Volume 90, September 2017, Pages 145–159.

[10] P. Q. Nguyen and I. E. Shparlinski, Counting co-cyclic lattices, *SIAM J. Discrete Math.*, 30 (2016), no. 3, 1358–1370.

[11] V. M. Petrogradsky, Multiple zeta functions and asymptotic structure of free abelian groups of finite rank, *J. Pure Appl. Algebra*, 208 (2007), no. 3, 1137–1158.

[12] M. Rudelson and R. Vershynin, The Littlewood-Offord problem and invertibility of random matrices, *Adv. Math.*, 218(2):600–633, 2008.

[13] R. Stanley, Smith normal form in Combinatorics, *Journal of Combinatorial Theory, Series A* Volume 144, November 2016, Pages 476–495.

[14] R. Stanley and Y. Wang, The Smith normal form distribution of a random integer matrix, *SIAM J. Discrete Math.*, 31 (3), 2017, 2247–2268.

[15] T. Tao and V. Vu On the singularity probability of random Bernoulli matrices, *J. Amer. Math. Soc.*, 20(3):603–628, 2007.

[16] K. Tikhomirov, Singularity of random Bernoulli matrices, *arxiv.org/abs/1812.09016*.

[17] M. M. Wood, The distribution of sandpile groups of random graphs, *Journal of the A. M. S*, 30 (2017), pp. 915–958.

[18] M. M. Wood, Random integral matrices and the Cohen-Lenstra Heuristics, *American Journal of Mathematics*, Volume 141, Number 2, 2019 pp. 383–398.

Intersection patterns of curves

János Pach

(joint work with István Tomon)

Given a family of sets, $\mathcal{C}$, the *intersection graph of* $\mathcal{C}$ is the graph, whose vertices correspond to the elements of $\mathcal{C}$, and two vertices are joined by an edge if the corresponding sets have a nonempty intersection. Also, the *disjointness graph of* $\mathcal{C}$ is the complement of the intersection graph of $\mathcal{C}$, that is, two vertices are joined by an edge if the corresponding sets are disjoint. As usual, we denote the *clique number,* the *independence number*, and the *chromatic number* of a graph G by $\omega(G), \alpha(G)$ and $\chi(G)$, respectively.

Clique number vs. chromatic number. There are many interesting results connecting the clique number and the chromatic number of geometric intersection graphs, starting with a beautiful theorem of Asplund and Grünbaum [1], which states that every intersection graph G of axis-parallel rectangles in the plane satisfies $\chi(G) \le 4(\omega(G))^2$.

A family $\mathcal{G}$ of graphs is *χ-bounded* if there exists a function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that every $G \in \mathcal{G}$ satisfies $\chi(G) \leq f(\omega(G))$. In this case, say that the function f is *χ-bounding for $\mathcal{G}$*. Using this terminology, the result of Asplund and Grünbaum [1] mentioned above can be rephrased as follows: The family of intersection graphs of axis-parallel rectangles in the plane is χ-bounded with bounding function $f(k) = 4k^2$. (It is conjectured that the same is true with bounding function $f(k) = O(k)$.) However, an ingenious construction of Burling [3] shows that the family of intersection graphs of axis-parallel boxes in $\mathbb{R}^3$ is *not* χ-bounded. The χ-boundedness of intersection graphs of chords of a circle was established by Gyárfás [5], and Kostochka *et al.* [6, 7] proved that the best χ-bounding function $f(k)$ is between $\Omega(k \log k)$ and $2^{O(k)}$. Recently, the upper bound was improved to $O(k^2)$ by Davies and McCarty [4].

Families of curves. A *curve* or *string* in $\mathbb{R}^2$ is the image of a continuous function $\phi : [0, 1] \to \mathbb{R}^d$. A curve $C \subset \mathbb{R}^2$ is called *x-monotone* if every vertical line intersects C in at most one point. We say that C is *grounded at the curve L* if one of the endpoints of C is in L, and this is the only intersection point of C and L. A *grounded x-monotone curve* is an x-monotone curve that is contained in the half-plane $\{x \geq 0\}$, and whose left endpoint lies on the vertical line $\{x = 0\}$.

It was first suggested by Erdős in the 1970s, and remained the prevailing conjecture for 40 years, that the family of intersection graphs of curves (the family of so-called "string graphs") is χ-bounded [2, 8]. There were many promising facts pointing in this direction. Extending earlier results of McGuinness [10] and Suk [15], Rok and Walczak [14] proved the conjecture for *grounded* families of curves. Nevertheless, in 2014, Pawlik *et al.* [13] disproved Erdős's conjecture. They managed to modify Burling's above mentioned construction to obtain a sequence of finite families of *segments* in the plane whose intersection graphs, G_n, are triangle-free (that is, $\omega(G_n) = 2$), but their chromatic numbers tend to infinity, as $n \to \infty$.

Recently, Pach, Tardos and Tóth [11] proved that the family of *disjointness graphs* of curves in the plane is not χ-bounded either. However, the situation is different if we restrict our attention to *x-monotone* curves. It was shown in [9] that the family of disjointness graphs of x-monotone curves in the plane is χ-bounded with a bounding function $f(k) = k^4$. For *grounded x-monotone* curves, the same proof provides a better bounding function: $f(k) = k^2$. These results proved 25 years ago were not likely to be tight. However, in spite of many efforts, no-one has managed to improve them or to show that they are optimal.

Our results. We proved, much to our surprise, that the order of magnitude of the last two bounds cannot be improved. In fact, in the case of grounded x-monotone curves, we determined the exact value of the best bounding function for every $k \geq 2$. To the best of our knowledge, this is the first large family of non-perfect geometric disjointness graphs, for which one can precisely determine the best bounding function.

Theorem 1. *Let G be the disjointness graph of a family of grounded x-monotone curves. If $\omega(G) = k$, then $\chi(G) \leq \binom{k+1}{2}$.*

Moreover, for every positive integer $k \geq 2$, there exists a family $\mathcal{C}$ of grounded x-monotone curves such that if G is the disjointness graph of $\mathcal{C}$, then $\omega(G) = k$ and $\chi(G) = \binom{k+1}{2}$.

It turns out that disjointness graphs of grounded x-monotone curves can be completely characterized by graphs with two total orders defined on their vertex sets that satisfy some special properties. We call such graphs *magical*, and we prove Theorem 1 by studying combinatorial properties of these graphs. This novel characterization may also be useful for the solution of some other problems concerning x-monotone curves.

The disjointness graph of any collection of x-monotone curves, each of which intersects a given vertical line (the y-axis, say), is the intersection of two disjointness graphs of grounded x-monotone curves. The methods used for proving Theorem 1 can be extended to such disjointness graphs and yield sharp bounds.

Theorem 2. *Let G be the disjointness graph of a family $\mathcal{C}$ of x-monotone curves such that all elements of $\mathcal{C}$ have nonempty intersection with a vertical line l. If $\omega(G) = k$, then $\chi(G) \leq \frac{k+1}{2}\binom{k+2}{3}$.*

Moreover, for every positive integer $k \geq 2$, there exists a family $\mathcal{C}$ of x-monotone curves such that all elements of $\mathcal{C}$ have nonempty intersection with a vertical line l, the disjointness graph G of $\mathcal{C}$ satisfies $\omega(G) = k$, and $\chi(G) = \frac{k+1}{2}\binom{k+2}{3}$.

As we have mentioned before, according to [9,12], k^4 is a bounding function for disjointness graphs of *any* family of x-monotone curves. Theorem 2 implies that the order of magnitude of this bounding function is best possible. Actually, we can prove a little more.

Theorem 3. *For any positive integer k, let $f(k)$ denote the smallest m such that any K_{k+1}-free disjointness graph of x-monotone curves can be properly colored with m colors. Then we have*

$$\frac{k+1}{2}\binom{k+2}{3} \leq f(k) \leq k^2\binom{k+1}{2}.$$

Here the lower and upper bounds differ by a factor of less than 6, and there is some hope that one can determine the exact value of $f(k)$. The lower bound follows directly from Theorem 2.

References

[1] E. Asplund, B. Grünbaum. "On a coloring problem." Math. Scand. 8 (1960), 181–188.

[2] P. Brass, W. Moser, J. Pach. "Research Problems in Discrete Geometry." Springer, New York, 2005.

[3] J. P. Burling. "On Coloring Problems of Families of Prototypes (PhD thesis)." University of Colorado, Boulder, 1965.

[4] J. Davies, R. McCarty. "Circle graphs are quadratically χ-bounded." arXiv preprint (2019), arXiv:1905.11578.

[5] A. Gyárfás. "On the chromatic number of multiple interval graphs and overlap graphs." Discrete Math. 55 (1985), 161–166.

[6] A. Kostochka. "On upper bounds on the chromatic numbers of graphs." Transactions Inst. Math., Vol. 10, Siberian Branch of the Acad. Sci. USSR (1988), 204–226 (in Russian).
[7] A. Kostochka, J. Kratochvíl. "Covering and coloring polygon-circle graphs." Discrete Math 163 (1-3) (1997): 299–305.
[8] A. Kostochka, J. Nešetřil. "Chromatic number of geometric intersection graphs." in: M. Klazar (Ed.), 1995 Prague Midsummer Combinatorial Workshop, KAM Series 95–309, Charles University, Prague (1995), 43–45.
[9] D. Larman, J. Matoušek, J. Pach, J. Törőcsik. "A Ramsey-type result for convex sets." Bull. Lond. Math. Soc., 26 (1994), 132–136.
[10] S. McGuinness. "Colouring arcwise connected sets in the plane I." Graphs and Combinatorics 16 (4) (2000), 429–439.
[11] J. Pach, G. Tardos, G. Tóth. "Disjointness graphs of segments." in: 33rd Internat. Symp. Comput. Geom. (SoCG 2017), vol. **77** Leibniz Internat. Proc. Informatics (LIPIcs), 59:1–15, Leibniz-Zentrum für Informatik, Dagstuhl, 2017.
[12] J. Pach, J. Törőcsik. "Some geometric applications of Dilworth's theorem." Discrete Comput. Geom. 12 (1) (1994), 1–7.
[13] A. Pawlik, J. Kozik, T. Krawczyk, M. Lasoń, P. Micek, W. T. Trotter, B. Walczak. "Triangle-free intersection graphs of line segments with large chromatic number." J. Combin. Theory Ser. B 105 (2014), 6–10.
[14] A. Rok, B. Walczak. "Outerstring graphs are χ-bounded." in: 30th Internat. Symp. Comput. Geom. (SoCG'14), 136–143, ACM, New York, 2014.
[15] A. Suk. "Coloring intersection graphs of x-monotone curves in the plane." Combinatorica 34 (4) (2014), 487–505.

Phase transition for the number random contingency tables with non-uniform margins

Igor Pak
(joint work with Sam Dittmer and Hanbaek Lyu)

1. Introductions

Contingency tables are fundamental objects in statistics for studying dependence structure between two or more variables. They also correspond to bipartite multigraphs with given degrees and play an important role in combinatorics and graph theory. Random contingency tables have been intensely studied in a variety of regimes, yet remain largely out of reach in many interesting special cases. In this paper we extend our previous paper [2] to show phase transition for the number of contingency tables.

Let $\mathbf{r} = (r_1, \ldots, r_m) \in \mathbb{N}^m, \mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{N}^n$ be two nonnegative integer vectors with the same sum of entries. Denote by $\mathcal{M}(\mathbf{r}, \mathbf{c})$ the set of all $(n \times m)$ contingency tables with row sums r_i and column sums c_j, i.e.

(1)
$$\mathcal{M}(\mathbf{r}, \mathbf{c}) := \left\{ (a_{ij}) \in \mathbb{N}^{mn} \,\middle|\, \sum_{k=1}^{n} a_{ik} = r_i, \sum_{k=1}^{m} a_{kj} = c_j \quad \text{for all } 1 \le i \le n,\, 1 \le j \le m \right\}.$$

Let $X = (X_{ij})$ be the contingency table chosen uniformly at random from $\mathcal{M}(\mathbf{r}, \mathbf{c})$. The asymptotic properties of the entries of X as $m, n \to \infty$ is the subject of this paper. When the margins are uniform, i.e. $r_1 = \ldots = r_m$ and $c_1 = \ldots = c_n$, the exact asymptotics for $\big|\mathcal{M}(\mathbf{r},\mathbf{c})\big|$ are known, see [1].

In this paper we analyze the number of square contingency table with the first $\lfloor n^\delta \rfloor$ row and column margins $\lfloor BCn \rfloor$, and the last n row and column margins $\lfloor Cn \rfloor$. Viewing such X as block matrices, see Figure 1, it is natural to assume that the entries are again nearly independent and identically distributed within each block. However, there is still one degree of freedom remaining: the distribution of mass of each block. We establish a sharp phase transition for this distribution in our previous paper [2], where we also discuss the history of the problem, give references, details, etc.

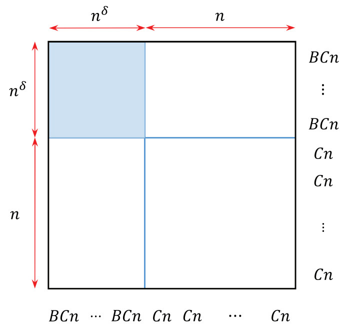

Figure 1. Contingency table with parameters n, δ, B and C. First $\lfloor n^\delta \rfloor$ rows and columns have margins $\lfloor Cn \rfloor$, the last n rows and columns have margins $\lfloor BCn \rfloor$.

This paper is not intended for publication. Due to its followup nature, we are including these results in the appendix of the `arXiv` version of [2].

2. Main results

For parameters $n \geq 1$, $0 \leq \delta \leq 1$, and $B, C \geq 0$, let $\mathcal{M}_{n,\delta}(B,C) = \mathcal{M}(\mathbf{r},\mathbf{c})$, where

$$\mathbf{r} = \mathbf{c} := \big(\lfloor BCn \rfloor, \ldots, \lfloor BCn \rfloor, \lfloor Cn \rfloor, \ldots, \lfloor Cn \rfloor\big) \in \mathbb{N}^{\lfloor n^\delta \rfloor + n}. \tag{2}$$

In other words, $\mathcal{M}_{n,\delta}(B,C)$ is the set of contingency tables whose first $\lfloor n^\delta \rfloor$ rows and columns have margin $\lfloor BCn \rfloor$ and the other n rows and columns have margin $\lfloor Cn \rfloor$, see Figure 1. Let $X = (X_{ij})$ be the random contingency table sampled uniformly from $\mathcal{M}_{n,\delta}(B,C)$. We are interested in the asymptotic behavior of $\big|\mathcal{M}_{n,\delta}(B,C)\big|$ as $n \to \infty$ for various choice of parameters δ, B and C.

Formally, we derive a phase transition for the number $|\mathcal{M}_{n,\delta}(B,C)|$ of contingency tables as B increases for large but fixed n.

Theorem 1. *Let $\mathcal{M}_{n,\delta}(B,C)$ be as above, where $0 \leq \delta < 1$. Let $B_c = 1 + \sqrt{1 + 1/C}$. Let $f(x) = (x+1)\log(1+x) - x\log(x)$. Define*

$$\mathcal{F}_{n,\delta}(B,C) := \log|\mathcal{M}_{n,\delta}(B,C)| - \big[n^2 f(C) + B_c C \log(1 + 2C^{-1}) n^{1+\delta}\big]. \tag{3}$$

Then there exists constants $\alpha, \beta > 0$ independent of B such that the followings hold:

(i): *If $B < B_c$, then for all $n \geq 1$,*

$$\left[2n^{1+\delta} f(BC) + \frac{\alpha}{B_c - B} n^{(2\delta)\vee(3\delta-1)}\right] - \beta n \log Bn \tag{4}$$

$$\leq \mathcal{F}_{n,\delta}(B, C) \leq \left[2n^{1+\delta} f(BC) + \frac{\alpha}{B_c - B} n^{(2\delta)\vee(3\delta-1)}\right]. \tag{5}$$

(ii): *If $B > B_c$, then for all $n \geq 1$,*

$$\left[(1-\delta)n^{2\delta} \log(C(B - B_c)n) + \frac{\alpha}{B - B_c} n^{2\delta}\right] - \beta n \log Bn \tag{6}$$

$$\leq \mathcal{F}_{n,\delta}(B, C) \leq \left[(1-\delta)n^{2\delta} \log(C(B - B_c)n) + \frac{\alpha}{B - B_c} n^{2\delta}\right]. \tag{7}$$

Remark. The most interesting implication of the above result is the upper bound when $B > B_c$. There, if we fix large $n \geq 1$ and increase B past B_c, the growth rate of $\log |\mathcal{M}_{n,\delta}(B, C)|$ in B is only in the order of $n^{2\delta} \log Bn$. On the contrary, the lower bound when $B < B_c$ says that $\log |\mathcal{M}_{n,\delta}(B, C)|$ should increase at least in the order of $n^{1+\delta} f(BC)$. In the special case of $\delta = 0$, the theorem implies that $|\mathcal{M}_{n,0}(B, C)|$ grows exponentially in n as we increase B toward B_c, but it grows only polynomially in n once $B > B_c$.

References

[1] E. R. Canfield and B. D. McKay, Asymptotic enumeration of integer matrices with large equal row and column sums, *Combinatorica* **30** (2010), 655–680.

[2] S. Dittmer, H. Lyu, and I. Pak, Phase transition in random contingency tables with non-uniform margins, to appear in *Trans. AMS*, 23 pp.; `https://arxiv.org/1903.08743`.

Thresholds versus fractional expectation-thresholds

Jinyoung Park

(joint work with K. Frankston, J. Kahn and B. Narayanan)

Our most important contribution here is the proof of a conjecture of Talagrand [8] that is a fractional version of the "expectation-threshold" conjecture of Kahn and Kalai [4]. For an increasing family $\mathcal{F}$ on a finite set X, we write (with definitions below) $p_c(\mathcal{F})$, $q_f(\mathcal{F})$ and $\ell(\mathcal{F})$ for the threshold, fractional expectation-threshold, and size of a largest minimal element of $\mathcal{F}$. In this language, our main result is the following.

Theorem 1. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q_f(\mathcal{F}) \log \ell(\mathcal{F}).$$

As observed below, $q_f(\mathcal{F})$ is a more or less trivial lower bound on $p_c(\mathcal{F})$, and Theorem 1 says this bound is never far from the truth. (Apart from the constant K, the upper bound is tight in many of the most interesting cases.)

Thresholds. For a given X and $p \in [0,1]$, μ_p is the product measure on 2^X given by $\mu_p(S) = p^{|S|}(1-p)^{|X \setminus S|}$. An $\mathcal{F} \subseteq 2^X$ is *increasing* if $B \supseteq A \in \mathcal{F} \Rightarrow B \in \mathcal{F}$. If this is true (and $\mathcal{F} \neq 2^X, \emptyset$), then $\mu_p(\mathcal{F})(:= \sum\{\mu_p(S) : S \in \mathcal{F}\})$ is strictly increasing in p, and the *threshold*, $p_c(\mathcal{F})$, is the unique p for which $\mu_p(\mathcal{F}) = 1/2$. This is finer than the original Erdős–Rényi notion, according to which $p^* = p^*(n)$ is ***a*** threshold for $\mathcal{F} = \mathcal{F}_n$ if $\mu_p(\mathcal{F}) \to 0$ when $p \ll p^*$ and $\mu_p(\mathcal{F}) \to 1$ when $p \gg p^*$. (That $p_c(\mathcal{F})$ *is* always an Erdős–Rényi threshold follows from [2].)

Following [6–8], we say $\mathcal{F}$ is *p-small* if there is a $\mathcal{G} \subseteq 2^X$ such that $\mathcal{F} \subseteq \langle \mathcal{G} \rangle := \{T : \exists S \in \mathcal{G}, S \subseteq T\}$ and

$$\sum_{S \in \mathcal{G}} p^{|S|} \leq 1/2. \tag{1}$$

Then $q(\mathcal{F}) := \max\{p : \mathcal{F} \text{ is } p\text{-small}\}$, which we call the *expectation-threshold* of $\mathcal{F}$ (note the term is used slightly differently in [4]), is a trivial lower bound on $p_c(\mathcal{F})$, since for $\mathcal{G}$ as above and T drawn from μ_p,

$$\mu_p(\mathcal{F}) \leq \mu_p(\langle \mathcal{G} \rangle) \leq \sum_{S \in \mathcal{G}} \mu_p(T \supseteq S) = \sum_{S \in \mathcal{G}} p^{|S|} \;\; (= \mathbb{E}[|\{S \in \mathcal{G} : S \subseteq T\}|]). \tag{2}$$

The following statement, the main conjecture (Conjecture 1) of [4], says that for *any* $\mathcal{F}$, this trivial lower bound on $p_c(\mathcal{F})$ is close to the truth.

Conjecture 2. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q(\mathcal{F}) \log |X|.$$

We should emphasize how strong this is (from [4]: "It would probably be more sensible to conjecture that it is *not* true"). For example, it easily implies—and was largely motivated by—Erdős–Rényi thresholds for (a) perfect matchings in random r-uniform hypergraphs, and (b) appearance of a given bounded degree spanning tree in a random graph. These have since been resolved: the first—*Shamir's Problem*, circa 1980—in [3], and the second—a mid-90's suggestion of the second author—in [5]. Both arguments are difficult and specific to the problems they address (e.g. they are utterly unrelated either to each other or to what we do here).

Talagrand [6, 8] suggests relaxing "p-small" by replacing the set system $\mathcal{G}$ above by what we may think of as a *fractional* set system, g: say $\mathcal{F}$ is *weakly p-small* if there is a $g : 2^X \to \mathbb{R}^+$ such that

$$\sum_{S \subseteq T} g(S) \geq 1 \;\forall T \in \mathcal{F} \;\text{ and }\; \sum_{S \subseteq X} g(S) p^{|S|} \leq 1/2.$$

Then $q_f(\mathcal{F}) := \max\{p : \mathcal{F} \text{ is weakly } p\text{-small}\}$, the *fractional expectation-threshold* of $\mathcal{F}$, satisfies

$$q(\mathcal{F}) \leq q_f(\mathcal{F}) \leq p_c(\mathcal{F}) \tag{3}$$

(the first inequality is trivial and the second is similar to (2)), and Talagrand [8, Conjectures 8.3 and 8.5] proposes a sort of LP relaxation of Conjecture 2, and then a strengthening thereof. The first of these, the following, replaces q by q_f in Conjecture 2; the second, which adds replacement of $|X|$ by the smaller $\ell(\mathcal{F})$, is our Theorem 1.

Conjecture 3. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \le K q_f(\mathcal{F}) \log |X|.$$

Talagrand further suggests the following "very nice problem of combinatorics," which implies *equivalence of* Conjectures 2 and 3, as well as of Theorem 1 and the corresponding strengthening of Conjecture 2.

Conjecture 4. *There is a universal K such that, for any increasing $\mathcal{F}$ on a finite set X, $q(\mathcal{F}) \ge q_f(\mathcal{F})/K$.*

(That is, weakly p-small implies (p/K)-small.)

Note the interest here is in Conjecture 4 for its own sake and as the most likely route to Conjecture 2; all applications of the latter that we're aware of follow just as easily from Theorem 1.

Spread hypergraphs and spread measures. We say a *hypergraph* on the (*vertex*) set X is a collection $\mathcal{H}$ of subsets of X (*edges* of $\mathcal{H}$), with *repeats allowed*. For $S \subseteq X$, we use $\langle S \rangle$ for $\{T \subseteq X : T \supseteq S\}$, and for a hypergraph $\mathcal{H}$ on X, we write $\langle \mathcal{H} \rangle$ for $\cup_{S \in \mathcal{H}} \langle S \rangle$. We say $\mathcal{H}$ is *ℓ-bounded* (resp. *ℓ-uniform* or an *ℓ-graph*) if each of its members has size at most (resp. exactly) ℓ, and *κ-spread* if

$$|\mathcal{H} \cap \langle S \rangle| \le \kappa^{-|S|} |\mathcal{H}| \; \forall S \subseteq X. \tag{4}$$

(Note that edges are counted with multiplicities on both sides of (4).)

A major advantage of the fractional versions (Conjecture 3 and Theorem 1) over Conjecture 2—and the source of the present relevance of [1]—is that they admit, via linear programming duality, reformulations in which the specification of $q_f(\mathcal{F})$ gives a usable starting point. Following [8], we say a probability measure ν on 2^X is *q-spread* if

$$\nu(\langle S \rangle) \le q^{|S|} \; \forall S \subseteq X.$$

Thus a hypergraph $\mathcal{H}$ is κ-spread iff uniform measure on $\mathcal{H}$ is q-spread with $q = \kappa^{-1}$.

As observed by Talagrand [8], the following is an easy consequence of duality.

Proposition 5. *For an increasing family $\mathcal{F}$ on X, if $q_f(\mathcal{F}) \le q$, then there is a $(2q)$-spread probability measure on 2^X supported on $\mathcal{F}$.* □

This allows us to reduce Theorem 1 to the following alternate (actually, equivalent) statement.

Theorem 6. *There is a universal K such that for any ℓ-bounded, κ-spread hypergraph $\mathcal{H}$ on X, a uniformly random $((K\kappa^{-1} \log \ell)|X|)$-element subset of X belongs to $\langle \mathcal{H} \rangle$ w.h.p. as $\ell \to \infty$.*

References

[1] R. Alweiss, S. Lovett, K. Wu, and J. Zhang, *Improved bounds for the sunflower lemma*, Preprint, arXiv:1908.08483v1.
[2] B. Bollobás and A. Thomason, *Thresholds functions*, Combinatorica **7** (1987), 35-38.
[3] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs*, Random Structures Algoritheorems **33** (2008), 1–28.
[4] J. Kahn and G. Kalai, *Thresholds and expectation thresholds,* Combin. Probab. Comput. **16** (2007), 495–502.
[5] R. Montgomery, *Spanning trees in random graphs*, Adv. Math. **356** (2019), 106793, 92.
[6] M. Talagrand, *Are all sets of positive measure essentially convex?*, Geometric aspects of functional analysis (Israel, 1992–1994), Oper. Theory Adv. Appl., vol. 77, Birkhäuser, Basel, 1995, pp. 295–310.
[7] M. Talagrand, *The generic chaining,* Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005.
[8] M. Talagrand, *Are many small sets explicitly small?,* Proceedings of the 2010 ACM International Symposium on Theory of Computing, (2010), 13–35.

Efficiently sampling colorings with less than 11Δ/6 colours

Guillem Perarnau
(joint work with Michelle Delcourt, Luke Postle)

Counting the number of proper k-colorings of a graph is a computationally hard problem [8]. Jerrum, Valiant, and Vazirani [5] showed that one can approximate this number using an almost uniform sampler, motivating the question of finding an efficient algoritheorem to generate uniformly random colorings of a graph. This question is a central topic in computer science and statistical physics.

To this end, we study the following Markov chain Monte Carlo algoritheorem known as Glauber dynamics (e.g. see [3]). The *Glauber dynamics for k-colorings* is a Markov chain with state space the set of proper k-colorings of G and transitions given by recoloring one vertex at a time.

It is easy to check that the chain is ergodic provided that $k \geq \Delta + 2$. A well-known conjecture in the area is that Glauber dynamics mixes in polynomial time (*rapid mixing*) for every $k \geq \Delta+2$ (and hence that there exists a polynomial-time approximation scheme for k-colorings for every $k \geq \Delta+2$). Jerrum [4] showed that for $k \geq 2\Delta+1$ the mixing time is $O(n \log n)$. Salas and Sokal [7] used Dobrushin's uniqueness criterion to obtain similar results. In 2000, Vigoda [9] obtained an important breakthrough in the area by showing that flip dynamics with certain flip parameters has mixing time $O(n \log n)$ for $k > \frac{11}{6}\Delta$, implying that Glauber dynamics has mixing time $O(n^2)$ in the same regime. In the last 20 years there have been numerous improvements for particular classes of graphs, but as to the original conjecture, no improvement over Vigoda's bound had appeared. In this talk we present the first result in this direction.

Theorem 1. *The Glauber dynamics for k-colorings on a graph on n vertices with maximum degree Δ and $k \geq (11/6 - \eta)\Delta$ with $\eta = \frac{1}{84000}$ has mixing time $O\left(\log k \cdot n^2\right)$.*

A similar result has been obtained independently by Chen and Moitra with a different proof technique. See [1] for a joint extended abstract including both proofs of the result.

As in [9], Theorem 1 will follow as a corollary of a similar result for the flip dynamics. The *flip dynamics for k-colorings with flip parameters* $\mathbf{p} = (p_1, p_2, \dots)$ is a Markov chain with space state the set of proper k-colornings and transitions between states given by swapping the colors of a maximum bicolored connected set of vertices S with probability proportional to $p_{|S|}$. Using the comparison theorem of Diaconis and Saloff-Coste [2] and provided that no large flips are allowed, one can bound the mixing time of Glauber dynamics using the mixing time of the flip dynamics.

In addition, we also study the problem of sampling list colorings. Frieze and Vigoda [3] asked if the results obtained for sampling colorings could be transferred to list coloring. Jerrum's proof for $k \geq 2\Delta + 1$ [4] carries over immediately for list coloring; however, rapid mixing of Glauber dynamics for list coloring was not previously known for $k \leq 2\Delta$.

Theorem 2. *The Glauber dynamics for k-list-colorings on a graph on n vertices with maximum degree Δ and $k \geq \left(\frac{11}{6} - \eta\right)\Delta$ has mixing time $O(\log k \cdot n^2)$.*

To prove Theorems 1 and 2, we introduce two main ingredients in Vigoda's analysis of the flip dynamics. Firstly, we use a novel metric for path coupling. While most of the previous approaches to the study of the mixing time of Glauber dynamics use the Hamming metric, alternative metrics have been introduced to deal with particular instances. For example, Luby and Vigoda [6] used a metric for independent sets that modifies the Hamming distance, decreasing it for "good" pairs of independent sets. Following the same spirit, we define a new metric d for colorings that consists of the Hamming metric d_H minus a small factor d_B counting the number of "non-extremal configurations" around a vertex. We show that in a single transition of the chain, either d_H decreases in expectation or d_B increases in expectation. In both cases, this leads to an expected decrease in the metric d. In this way, it suffices to study a single-step coupling with the new metric, avoiding to do a multi-step coupling argument. Secondly, we use linear programming to obtain a set of flip parameters $\mathbf{p}$ minimizing the number of extremal configurations. The existence of a small number of extremal configurations makes it possible to analyze the expected change of the metric.

References

[1] S. Chen, M. Delcourt, A. Moitra, G. Perarnau, and L. Postle. Improved bounds for randomly sampling colorings via linear programming. *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algoritheorems*, 2019, 2216–2234.

[2] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Prob.*, **3**(3), 1993, 696–730.

[3] A. Frieze and E. Vigoda. A survey on the use of Markov chains to randomly sample colourings. *Combinatorics, complexity, and chance*, Oxford Lecture Ser. Math. Appl., Vol. 34. Oxford Univ. Press, Oxford, UK, 2007, 53–71.

[4] M.R. Jerrum. A very simple algoritheorem for estimating the number of k-colourings of a low-degree graph. *Random Structures and Algoritheorems* 7(2), 1995, 157–165.
[5] M.R. Jerrum, L.G. Valiant, and V.V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science* 43, 1986, 169–188.
[6] M. Luby and E. Vigoda, Fast convergence of the Glauber dynamics for sampling independent sets, *Random Structures & Algoritheorems*, **15**, 1999, 229–241.
[7] J. Salas and A. Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86(3-4), 1997, 551–579.
[8] L. G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8(3), 1979, 410–421.
[9] E. Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3), 2000, 1555–1569.

A proof of Ringel's Conjecture

Alexey Pokrovskiy
(joint work with Richard Montgomery, Benny Sudakov)

Ringel's Conjecture is a problem about *graph decompositions*. We say that a graph G has a *decomposition into copies of a graph H* if the edges of G can be partitioned into edge-disjoint subgraphs isomorphic to H. Many well known graph theoretic problems can be phrased in these terms. For example, in 1847, Kirkman studied decompositions of complete graphs K_n and showed that they can be decomposed into copies of a triangle if, and only if, $n \equiv 1$ or $3 \pmod 6$. Ringel's conjecture is about decompositions of complete graphs into large trees, (where a tree is a connected graph with no cycles).

Conjecture 1 (Ringel, [8])**.** *The complete graph K_{2n+1} can be decomposed into copies of any tree with n edges.*

Ringel's conjecture is one of the oldest and best known open conjectures on graph decompositions. It has been established for many very special classes of trees such as caterpillars, trees with ≤ 4 leaves, firecrackers, diameter ≤ 5 trees, symmetrical trees, trees with ≤ 35 vertices, and olive trees (see Chapter 2 of [5] and the references therein). There have also been some partial general results in the direction of Ringel's conjecture. Typically, for these results, an extensive technical method is developed which is capable of almost-packing any appropriately-sized collection of certain sparse graphs, see, e.g., [1–4, 6]. Despite the variety of techniques, these latter papers all have the same limitation, requiring that the maximum degree of the tree should be much smaller than n.

At Oberwolfach, a proof of Ringel's conjecture for sufficiently large trees was presented.

Theorem 2 ([7])**.** *For every sufficiently large n the complete graph K_{2n+1} can be decomposed into copies of any tree with n edges.*

The proof of this theorem is based on showing that a rainbow copy of every n-edge tree exists in every *near-distance*-coloured K_{2n+1}. Using a standard transformation, this implies that K_{2n+1} contains a *cyclic decomposition* by copies of any n-edge tree (a decomposition is cyclic if it is invariant under some cyclic permutation of the vertices of K_{2n+1}). The existence of such a cyclic decomposition was separately conjectured by Kotzig (see [9]).

Our approach builds on ideas from the previous research on graph decompositions and graceful labellings. From the work on graph decompositions, our approach is inspired by randomized decompositions and the absorption technique. The rough idea of absorption is as follows. Before the embedding of T we prepare a template which has some useful properties. Next we find a partial embedding of the tree T with some vertices removed such that we did not use the edges of the template. Finally we use the template to embed the remaining vertices. This idea was introduced as a general method by Rödl, Ruciński and Szemerédi and has been used extensively since then. For example, the proof of Ringel's Conjecture for bounded degree trees is based on this technique [6].

References

[1] A. Adamaszek, P. Allen, C. Grosu, and J. Hladký. Almost all trees are almost graceful. *arXiv preprint arXiv:1608.01577*, 2016.

[2] P. Allen, J. Böttcher, D. Clemens, and A. Taraz. Perfectly packing graphs with bounded degeneracy and many leaves. *arXiv preprint arXiv:1906.11558*, 2019.

[3] J. Böttcher, J. Hladký, D. Piguet, and A. Taraz. An approximate version of the tree packing conjecture. *Israel Journal of Mathematics*, 211(1):391–446, 2016.

[4] A. Ferber and W. Samotij. Packing trees of unbounded degrees in random graphs. *J. Lond. Math. Soc.*, 2016.

[5] J. A. Gallian. A dynamic survey of graph labeling. *The electronic journal of combinatorics*, 16(6):1–219, 2009.

[6] F. Joos, J. Kim, D. Kühn, and D. Osthus. Optimal packings of bounded degree trees. *J. European Math. Soc., to appear*, 2018.

[7] R. Montgomery, A. Pokrovskiy, and B. Sudakov. A proof of Ringel's Conjecture. *arXiv 2001.02665*, 2019.

[8] G. Ringel. Theory of graphs and its applications. In *Proceedings of the Symposium Smolenice*, 1963.

[9] A. Rosa. On certain valuations of the vertices of a graph. In *Theory of Graphs (Internat. Symposium, Rome*, pages 349–355, 1966.

Girth in Ramsey theory

Christian Reiher and Vojtěch Rödl

1. Introduction – Colouring vertices

We commence with an unusual formulation of the well known fact that there exist graphs whose girth and chromatic number are simultaneously arbitrarily large [2, 6, 8, 13, 16]. Recall that the *chromatic number* of a graph H is the least natural number $\chi(H)$ such that there exists a colouring of the vertices of H using $\chi(H)$ colours with the property that no edge of H is monochromatic. This is a Ramsey theoretic invariant of H, for a lower bound estimate of the form $\chi(H) > r$ can equivalently be expressed by the partition relation

$$H \longrightarrow (e)^v_r\,, \tag{1}$$

where the superscripted v on the right side means that the objects receiving colours are *vertices* and the letter e enclosed in parentheses indicates that the object we want to find monochromatically is an *edge*. Describing the absence of cycles in terms of forests one can formulate the aforementioned result as follows.

Theorem 1. *Given $r, n \in \mathbb{N}$ there exists a graph H with $H \longrightarrow (e)^v_r$ having the property that any set consisting of at most n edges of H forms a forest.*

From a Ramsey theoretic perspective, it is equally natural to investigate the problem that instead of a monochromatic edge one wishes to enforce a monochromatic copy of a given graph F. For any two graphs F and H we write $\binom{H}{F}$ for the set of all *induced* subgraphs of H isomorphic to F. For $\mathcal{H} \subseteq \binom{H}{F}$ and $r \in \mathbb{N}$ the partition relation

$$\mathcal{H} \longrightarrow (F)^v_r \tag{2}$$

is defined to hold if for every colouring of the vertices of H with r colours there exists a monochromatic copy $F_\star \in \mathcal{H}$. The existence of a system $\mathcal{H}$ having this property for given F and r is easily established by starting with a linear $v(F)$-uniform hypergraph whose chromatic number is larger than r, and inserting copies of F into its edges. Pursuing this argument further one arrives at the following result (see also [14, 22, 23]).

Theorem 2. *For every graph F and all $r, n \in \mathbb{N}$ there exists a graph H together with a system of copies $\mathcal{H} \subseteq \binom{H}{F}$ such that*

- (I) *$\mathcal{H} \longrightarrow (F)^v_r$ and*
- (II) *every $\mathcal{N} \subseteq \mathcal{H}$ with $|\mathcal{N}| \le n$ admits an enumeration $\mathcal{N} = \{F_1, \dots, F_{|\mathcal{N}|}\}$ with the property that for every $j \in [2, |\mathcal{N}|]$ the sets $\bigcup_{i<j} V(F_i)$ and $V(F_j)$ have at most one vertex in common.*

2. Colouring edges – Results

An entirely new level of difficulty emerges when edges rather than vertices are the entities subject to colouration. Henson asked more than forty years ago whether for every graph F and every number of colours r there exists a graph H with

$$H \longrightarrow (F)_r\,,$$

where now we are colouring edges[1] and the desired monochromatic occurrence of F is supposed to be induced. This problem was first solved in the 2-uniform case [5, 9, 19, 20] and later for more general structures, such as hypergraphs [1, 15]. As a matter of fact, these articles show much stronger results, asserting that one can demand H to have certain additional properties, provided that F has these

[1] Henceforth all colourings are colourings of edges and attempting to keep the notation simple we refrain from adding a superscripted "e" on the right side of our partition relations.

properties as well. For instance, these results allow to preserve the clique number, but not the girth (length of the shortest induced cycle). We may now state the simplest form of the girth Ramsey theorem, which is among the main results of this talk.

Theorem 3. *Given an integer $g \geq 2$, a graph F with* $\mathrm{girth}(F) > g$ *and a natural number r, there exists a graph H with* $\mathrm{girth}(H) > g$ *and*

$$H \longrightarrow (F)_r\,. \tag{3}$$

What can we say about longer cycles in H? Clearly, if F contains a C_m, then H needs to contain C_m as well. Moreover, as soon as both r and $e(F)$ are at least 2, there must be overlapping copies of F in H, which give rise to plenty of cycles in H that pass through several copies of F. One can insist, however, that up to some fixed length all *induced* cycles in H are inherited from the copies of F that guarantee the Ramsey property. This is made precise by the following version of the girth Ramsey theorem.

Theorem 4. *Given a graph F and positive integers r and n, there exists a graph H with $H \longrightarrow (F)_r$ such that for $m \in [2, n]$ every induced m-cycle of H is contained in an induced subgraph of H which is isomorphic to F.*

For a result in the spirit of Theorem 2 we need to define forests of copies.

Definition 5. *Suppose that F and H are graphs and that $\mathcal{N} \subseteq \binom{H}{F}$. We say that $\mathcal{N}$ is a forest of copies if there exists an enumeration $\mathcal{N} = \{F_1, \dots, F_{|\mathcal{N}|}\}$ such that for every $j \in [2, |\mathcal{N}|]$ the set $z_j = \big(\bigcup_{i<j} V(F_i)\big) \cap V(F_j)$ satisfies*

(I) *either $|z_j| \leq 1$*
(II) *or $z_j \in \big(\bigcup_{i<j} E(F_i)\big) \cap E(F_j)$.*

We would like to draw attention to a bizarre difference between this notion and the standard edge forests considered in Section 1. It is well known that if a set of edges forms a standard forest, then so does each of its subsets. As the following counterexample demonstrates, the analogous statement for forests of copies fails. Let F be a graph containing a triangle $x_0x_1x_2$. For every index $i \in \mathbb{Z}/3\mathbb{Z}$ let F_i be a graph isomorphic to F having with F the edge $x_{i+1}x_{i+2}$ but nothing else in common. Suppose that except for these intersections the copies in $\mathcal{N} = \{F, F_0, F_1, F_2\}$ are mutually disjoint. This enumeration exemplifies that $\mathcal{N}$ is a forest of copies. However its subset $\mathcal{N}^- = \mathcal{N} \smallsetminus \{F\}$ fails to be such a forest. For instance, for the enumeration $\mathcal{N}^- = \{F_0, F_1, F_2\}$ the set $z_2 = \big(V(F_0) \cup V(F_1)\big) \cap V(F_2) = \{x_0, x_1\}$ is certainly not in case (I) and, as it fails to be an edge of F_0 or F_1, it does not satisfy (II) either. By symmetry a similar problem arises when one enumerates $\mathcal{N}^-$ in any other way.

Summarising this discussion, we have seen that being a forest of copies is not preserved under taking subsets. This phenomenon explains the rôle of $\mathcal{X}$ in the third version of the girth Ramsey theorem that follows.

Theorem 6. *Given a graph F and $r, n \in \mathrm{N}$ there exists a graph H together with a system of copies $\mathcal{H} \subseteq \binom{H}{F}$ satisfying not only $\mathcal{H} \longrightarrow (F)_r$ but also the following* *statement: For every* $\mathcal{N} \subseteq \mathcal{H}$ *with* $|\mathcal{N}| \in [2, n]$ *there exists a set* $\mathcal{X} \subseteq \mathcal{H}$ *such that* $|\mathcal{X}| \le |\mathcal{N}| - 2$ *and* $\mathcal{N} \cup \mathcal{X}$ *is a forest of copies.*

If F contains triangles the upper bound on $|\mathcal{X}|$ is optimal and relates to the fact that when triangulating an $|\mathcal{N}|$-gon by means of diagonals one creates $|\mathcal{N}| - 2$ triangles. If $\mathrm{girth}(F) > g$ is known to hold for some $g \ge 2$, the upper bound on $|\mathcal{X}|$ can be improved to $|\mathcal{X}| \le (|\mathcal{N}| - 2)/(g - 1)$.

Finally, we also proved generalisations of these theorems to linear hypergraphs. All of our results are proved by means of explicit constructions.

3. Discussion of earlier results

We conclude by discussing some partial results towards the girth Ramsey theorem that have been obtained over the years. First, for general k-uniform hypergraphs Theorem 6 has been proved by Nešetřil and Rödl [18] for $n = 2$ and their approach yields Theorem 3 for $g = 3$ as well.

Most other partial results deal with the case of graphs. The main result of [18] asserts that Theorem 3 holds for $g = 4$ and, as Nešetřil and Rödl point out, by means of a more elaborate version of their argument one can treat every $g \le 7$. However, the new difficulties arising for $g = 8$ are fairly overwhelming. In general, it seems that even cycles cause more difficulties than odd cycles and, in fact, an *odd-girth* version of Theorem 3 was obtained in [17].

Rödl and Ruciński [21] proved probabilistically that Theorem 3 holds for $F = C_{g+1}$, thus answering a question of Erdős [7]. The smallest number of edges that a Ramsey graph for C_{g+1} can have was subsequently investigated by Haxell, Kohayakawa, and Łuczak [12]. The study of Ramsey properties of sparse random graphs and hypergraphs culminated in contributions by Friedgut, Rödl, and Schacht [10], and by Conlon and Gowers [4]. Recently Hàn, Rödl, Retter, and Schacht [11] studied numerical aspects associated with the case $F = C_{g+1}$ of Theorem 3 utilising the container method [3, 24].

References

[1] Fred G. Abramson and Leo A. Harrington, *Models without indiscernibles*, J. Symbolic Logic **43** (1978), no. 3, 572–600.

[2] Noga Alon, Alexandr Kostochka, Benjamin Reiniger, Douglas B. West, and Xuding Zhu, *Coloring, sparseness and girth*, Israel J. Math. **214** (2016), no. 1, 315–331.

[3] József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709.

[4] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), no. 2, 367–454.

[5] W. Deuber, *Generalizations of Ramsey's theorem*, Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I, North-Holland, Amsterdam, 1975, pp. 323–332. Colloq. Math. Soc. János Bolyai, Vol. 10. MR0369127 (51 #5363)

[6] P. Erdős, *Graph theory and probability*, Canadian J. Math. **11** (1959), 34–38.

[7] Paul Erdős, *Problems and results on finite and infinite graphs*, Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974), Academia, Prague, 1975, pp. 183–192. (loose errata).

[8] P. Erdős and A. Hajnal, *On chromatic number of graphs and set-systems*, Acta Math. Acad. Sci. Hungar. **17** (1966), 61–99.
[9] P. Erdős, A. Hajnal, and L. Pósa, *Strong embeddings of graphs into colored graphs*, Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I, North-Holland, Amsterdam, 1975, pp. 585–595. Colloq. Math. Soc. János Bolyai, Vol. 10.
[10] Ehud Friedgut, Vojtěch Rödl, and Mathias Schacht, *Ramsey properties of random discrete structures*, Random Structures Algoritheorems **37** (2010), no. 4, 407–436.
[11] Hiêp Hàn, Troy Retter, Vojtěch Rödl, and Mathias Schacht, *Ramsey-type numbers involving graphs and hypergraphs with large girth*. Combin. Probab. Comput. To Appear.
[12] P. E. Haxell, Y. Kohayakawa, and T. Łuczak, *The induced size-Ramsey number of cycles*, Combin. Probab. Comput. **4** (1995), no. 3, 217–239.
[13] L. Lovász, *On chromatic number of finite set-systems*, Acta Math. Acad. Sci. Hungar. **19** (1968), 59–67.
[14] Jaroslav Nešetřil and Vojtěch Rödl, *Partitions of vertices*, Comment. Math. Univ. Carolinae **17** (1976), no. 1, 85–95.
[15] Jaroslav Nešetřil and Vojtěch Rödl, *Partitions of finite relational and set systems*, J. Combinatorial Theory Ser. A **22** (1977), no. 3, 289–312.
[16] Jaroslav Nešetřil and Vojtěch Rödl, *A short proof of the existence of highly chromatic hypergraphs without short cycles*, J. Combin. Theory Ser. B **27** (1979), no. 2, 225–227.
[17] J. Nešetřil and V. Rödl, *On Ramsey graphs without cycles of short odd lengths*, Comment. Math. Univ. Carolin. **20** (1979), no. 3, 565–582.
[18] Jaroslav Nešetřil and Vojtěch Rödl, *Strong Ramsey theorems for Steiner systems*, Trans. Amer. Math. Soc. **303** (1987), no. 1, 183–192.
[19] Vojtěch Rödl, *The dimension of a graph and generalized Ramsey numbers*, 1973. Master's Thesis, Charles University, Praha, Czechoslovakia.
[20] ———, *A generalization of the Ramsey theorem*, Graphs, Hypergraphs, Block Syst. (Proc. Symp. comb. Anal., Zielona Gora, 1976), 1976, pp. 211–219.
[21] Vojtěch Rödl and Andrzej Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942.
[22] V. Rödl and N. Sauer, *The Ramsey property for families of graphs which exclude a given graph*, Canad. J. Math. **44** (1992), no. 5, 1050–1060.
[23] V. Rödl, N. Sauer, and X. Zhu, *Ramsey families which exclude a graph*, Combinatorica **15** (1995), no. 4, 589–596.
[24] David Saxton and Andrew Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992.

The upper tail for triangles in sparse random graphs

Wojciech Samotij
(joint work with Matan Harel, Frank Mousset)

Given an integer n and a real $p \in [0,1]$, let $X = X_{n,p}$ denote the number of triangles in the binomial random graph $G_{n,p}$. It is natural to ask for quantitative estimates of the tail probabilities $\Pr(X \le (1-\delta)\mathbb{E}[X])$ and $\Pr(X \ge (1+\delta)\mathbb{E}[X])$. The main motivation for studying this particular random variable is that the classical theory of large deviations applies only to linear functions of independent random variables, whereas X, defined above, is naturally expressed as a degree-three polynomial. As it turns out, these two tail probabilities are governed by very different phenomena. On the one hand, using a combination of Harris's inequality [9] and Janson's inequality [10], one can show that

$$-\log \Pr\big(X \le (1-\delta)\mathbb{E}[X]\big) = \Theta_\delta(\min\{n^2p, n^3p^3\}), \tag{1}$$

as long as $p \le 1 - \Omega(1)$. On the other hand, there are no comparably simple tools that allow one to easily obtain similar estimates on the logaritheorem of the upper tail probability.

An important reason for this difficulty is that, unlike the lower tail, the upper tail is influenced by 'local events' – the appearance of small subgraphs that increases the value X atypically. For example, there are graphs with n vertices and merely $O_\delta(n^2p^2)$ edges that contain $(1+\delta)\mathbb{E}[X]$ triangles (consider a complete graph with $C_\delta np$ vertices, for a sufficiently large constant C_δ); the upper tail probability is thus at least the probability that $G_{n,p}$ contains such a graph:

$$-\log \Pr\big(X \ge (1+\delta)\mathbb{E}[X]\big) \le O_\delta(n^2p^2\log(1/p)). \tag{2}$$

Establishing an upper bound on the upper tail probability that would match (2) proved to be a difficult challenge. This was achieved only less than ten years ago, following a sequence of papers [11–13, 15], by Chatterjee [2] and DeMarco–Kahn [6], who proved that

$$-\log \Pr\big(X \ge (1+\delta)\mathbb{E}[X]\big) = \Theta_\delta(\min\{\mathbb{E}[X], n^2p^2\log(1/p)\}),$$

which matches the upper bound (2) in the range $p \ge n^{-1}\log n$.

This still leaves open the problem of determining the dependence of the upper tail probability on the constant δ, which would be the first step towards obtaining a meaningful description of the upper tail event. In order to quantify this dependence, define

$$\Psi_X(\delta) = \min\Big\{e_G : G \subseteq K_n \text{ and } \mathbb{E}[X \mid G \subseteq G_{n,p}] \ge (1+\delta)\mathbb{E}[X]\Big\}. \tag{3}$$

It is not difficult to show that, if $\Psi_X(\delta) \to \infty$,

$$-\log \Pr\big(X \ge (1+\delta)\mathbb{E}[X]\big) \le (1+o(1)) \cdot \Psi_X(\delta)\log(1/p). \tag{4}$$

In the last several years, a powerful theory of large deviations for nonlinear functions of independent random variables has been developed in [1, 3–5, 7]. This theory allows to express the logaritheorem of the upper tail probability for X, for a certain range of p, in terms of the solution to a variational problem that generalises the definition (3). In particular, it was shown in [1, 5, 14] that the upper bound (4) is asymptotically optimal when $n^{-1/2}(\log n)^{O(1)} \ll p \ll 1$. Our main results extends this to the optimal range of densities.

Theorem 1. *Let X denote the number of triangles in $G_{n,p}$. Then, for every fixed positive constant δ and all $p = p(n)$ satisfying $n^{-1}\log n \ll p \ll 1$,*

$$-\log \Pr\big(X \ge (1+\delta)\mathbb{E}[X]\big) = (1 \pm o(1)) \cdot \Psi_X(\delta)\log(1/p).$$

In the complementary range of densities, we show that the upper tail of X resembles that of a Poisson random variable (with the same expectation).

Theorem 2. *Let X denote the number of triangles in $G_{n,p}$. Then, for every fixed positive constant δ and all $p = p(n)$ satisfying $n^{-1} \ll p \ll n^{-1}\log n$,*

$$\lim_{n\to\infty} \frac{-\log \Pr\big(X \geq (1+\delta)\mathbb{E}[X]\big)}{\mathbb{E}[X]} = (1+\delta)\log(1+\delta) - \delta.$$

Together, these theorems resolve the upper tail problem for X nearly completely. At the heart of the proof of Theorem 1 lies a general method for proving bounds on the upper tail probability of random variables that may be expressed as low-degree polynomials in i.i.d. Bernoulli variables. The proof uses an adaptation of the classical moment argument of Janson, Oleszkiewicz, and Ruciński [11]; this argument is used to formalise the intuition that the upper tail event is dominated by the appearance of near-minimisers of the combinatorial optimisation problem (3).

In the particular context of triangles in $G_{n,p}$, we say that a graph $G \subseteq K_n$ is a *core* if it is a feasible set for the above optimisation problem, it has at most $O\big(\Psi_X(\delta)\log(1/p)\big)$ edges, and it satisfies a certain natural rigidity condition. Our general method implies that the upper tail probability is approximately equal to the probability of the appearance of a core. In particular, when the number of cores of size m is $(1/p)^{o(m)}$, a property we term *entropic stability*, then a naive union bound implies that $-\log \Pr\big(X \geq (1+\delta)\mathbb{E}[X]\big) \approx \Psi_X(\delta)\log(1/p)$; fairly simple combinatorial arguments can be used to show that this is the case precisely when $n^{-1}\log n \ll p \ll 1$. The proof of Theorem 2 involves a change of measure and factorial moment estimates.

The asymptotics of $\Psi_X(\delta)$ are given by the following theorem, which extends an earlier result due to Lubetzky–Zhao [14]. For positive reals δ and c, we define

$$\varphi(\delta, c) = \min\left\{\frac{\delta^{2/3}}{2}, \frac{\lfloor \delta c/3\rfloor + \{\delta c/3\}^{1/2}}{c}, \frac{\lfloor \delta c/3\rfloor}{c} + \frac{(r\{\delta c/3\}/c)^{2/r}}{2}\right\}.$$

Theorem 3. *Let X denote the number of triangles in $G_{n,p}$. Then, for every fixed positive constant δ and all $p = p(n)$ satisfying $n^{-1} \ll p \ll 1$,*

$$\lim_{n\to\infty} \frac{\Psi_X(\delta)}{n^2p^2} = \begin{cases} \delta^{2/3}/2 & \text{if } np^2 \to 0, \\ \varphi(\delta, c) & \text{if } np^2 \to c \in (0,\infty), \\ \min\left\{\delta^{2/3}/2, \delta/3\right\} & \text{if } np^2 \to \infty. \end{cases}$$

The different possible values for this limit correspond to different types of subgraphs achieving the minimum in the definition of $\Psi_X(\delta)$. Our methods also allow us to show that if one conditions $G_{n,p}$ on the upper tail event, then with high probability, $G_{n,p}$ contains a graph closely resembling one of these minimisers. The results mentioned above extend to the case where one replaces triangles by larger cliques and, to some extent, by regular graphs.

References

[1] F. Augeri. Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős–Rényi graphs. `arXiv:1810.01558`.
[2] S. Chatterjee. The missing log in large deviations for triangle counts. *Random Structures Algoritheorems*, 40(4):437–451, 2012.
[3] S. Chatterjee and A. Dembo. Nonlinear large deviations. *Adv. Math.*, 299:396–450, 2016.
[4] S. Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdős–Rényi random graph. *European J. Combin.*, 32(7):1000–1017, 2011.
[5] N. A. Cook and A. Dembo. Large deviations of subgraph counts for sparse Erdős–Rényi graphs. `arXiv:1809.11148`.
[6] R. DeMarco and J. Kahn. Tight upper tail bounds for cliques. *Random Structures Algoritheorems*, 41(4):469–487, 2012.
[7] R. Eldan. Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations. `arXiv:1612.04346`.
[8] M. Harel, F. Mousset, and W. Samotij. Upper tails via high moments and entropic stability. `arXiv:1904.08212`.
[9] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.
[10] S. Janson. Poisson approximation for large deviations. *Random Structures Algoritheorems*, 1(2):221–229, 1990.
[11] S. Janson, K. Oleszkiewicz, and A. Ruciński. Upper tails for subgraph counts in random graphs. *Israel J. Math.*, 142:61–92, 2004.
[12] S. Janson and A. Ruciński. The deletion method for upper tail estimates. *Combinatorica*, 24(4):615–640, 2004.
[13] J. H. Kim and V. H. Vu. Divide and conquer martingales and the number of triangles in a random graph. *Random Structures Algoritheorems*, 24(2):166–174, 2004.
[14] E. Lubetzky and Y. Zhao. On the variational problem for upper tails in sparse random graphs. *Random Structures Algoritheorems*, 50(3):420–436, 2017.
[15] V. H. Vu. Concentration of non-Lipschitz functions and applications. *Random Structures Algoritheorems*, 20(3):262–316, 2002.

On counting algebraically defined graphs

Lisa Sauermann

1. Introduction

Many natural classes of graphs arising in discrete and computational geometry have been studied intensively both because of their structural properties and due to their relevance in practical applications. In this talk, we discuss an essentially tight lower bound on the number of graphs with vertex set $\{1, \dots, n\}$ in various graph classes obtained from discrete geometry. In fact, the graphs in these classes can be defined algebraically by polynomial conditions. Therefore, following an approach of Alon and Scheinerman [1], Warren's theorem [7] implies an upper bound on the number of graphs with vertex set $\{1, \dots, n\}$ in these graph classes (Warren's theorem is a variant of a theorem of Milnor [4] and Thom [6]). We show that this upper bound is essentially tight for any such class of algebraically defined graphs, assuming that the corresponding polynomials satisfy some reasonable conditions.

Intersection graphs are particularly natural classes of graphs obtained from discrete geometry. Given n geometric objects from some family $\mathcal{F}$ (for example the family of all segments in the plane) numbered from 1 to n, their intersection graph is the graph on the vertex set $\{1, \dots, n\}$ where two vertices are joined by an edge if and only if the corresponding objects intersect.

As mentioned above, for many families $\mathcal{F}$ of geometric objects, Warren's theorem [7] can be used to bound the number of graphs occuring as intersection graphs of a collection of n numbered objects in $\mathcal{F}$ (see for example [5] for segments in the plane and [3] for disks in the plane). In contrast, all known lower bounds for the number of intersection graphs of n numbered objects in a given family $\mathcal{F}$ were obtained by (sometimes fairly involved) ad hoc constructions for some specific families $\mathcal{F}$. Specifically, McDiarmid and Müller [3] proved lower bounds for disks and unit disks (in the plane), and Fox [2] provided a lower bound construction for segments (in the plane). All of these lower bounds essentially match the upper bounds that Warren's theorem [7] gives in the respective cases. However, these lower bound constructions are specific to the particular family $\mathcal{F}$ and do not easily generalize to other families $\mathcal{F}$ of geometric objects.

In this talk, we discuss a result giving an essentially tight lower bound for the number of graphs whose edges are defined using the signs of a given finite list of polynomials, assuming these polynomials satisfy some reasonable conditions. Our theorem in particular implies essentially tight lower bounds for the number of intersection graphs of segments, disks and many other geometric objects in the plane (or in higher dimension). It also implies an essentially tight lower bound for the number of graphs obtained by considering the pairwise linking or non-linking relations of n numbered disjoint circles in $\mathbb{R}^3$.

From discrete geometry, one can not only obtain graphs of interest, but also partial orders, so-called containment orders. A collection of n geometric objects from some family $\mathcal{F}$ numbered from 1 to n defines a partial order on the set $\{1, \dots, n\}$ obtained from the containment relations between the objects. Alon and Scheinerman [1] gave an upper bound for the number of containment orders obtained from a collection of n numbered objects in some family $\mathcal{F}$ using Warren's theorem [7]. For many geometric families $\mathcal{F}$, our general result implies an essentially matching lower bound for this number of containment orders. In particular, this essentially determines the number of circle orders, angle orders and containment orders obtained from polygons with a fixed number of vertices in the plane.

2. Statement of the result

In order for our result to apply straightforwardly not only to algebraically defined graphs, but also to partial orders, we work in the framework of algebraically defined edge-labelings of complete graphs. Given a finite set Λ of labels, a list of polynomials $P_1, \dots, P_k \in \mathbb{R}[x_1, \dots, x_d, y_1, \dots, y_d]$, a function $\phi : \{+, -, 0\}^k \to \Lambda$ and points $a_1, \dots, a_n$ in some open subset $U \subseteq \mathbb{R}^d$, one can define an edge-labeling of the complete graph on the vertex set $\{1, \dots, n\}$ as follows: For any $1 \le i < j \le n$ the label of the edge ij is defined to be $\phi\big(\operatorname{sgn} P_1(a_i, a_j), \dots, \operatorname{sgn} P_k(a_i, a_j)\big) \in \Lambda$. Fixing Λ, the polynomials $P_1, \dots, P_k$, the function ϕ and the subset $U \subseteq \mathbb{R}^d$, we are then concerned with the number of edge-labelings which can be obtained in this way for some points $a_1, \dots, a_n \in U$. We call all such edge-labelings $(P_1, \dots, P_k, \phi, U, \Lambda)$-*representable*

Taking the set of labels to be $\Lambda = \{$"edge", "non-edge"$\}$, edge-labelings of the complete graph on the vertex set $\{1, \dots, n\}$ correspond precisely to ordinary graphs on the vertex set $\{1, \dots, n\}$.

To motivate the set-up of $(P_1, \dots, P_k, \phi, U, \Lambda)$-representable edge-labelings, let us check that intersection graphs of open disks in the plane can be interpreted in this way: Each disk in the plane is given by specifying its center (x, y) and its radius $r > 0$. Thus, the family of open disks in the plane corresponds to the open set U of points $(x, y, r) \in \mathbb{R}^3$ with $r > 0$. Two disks corresponding to the points $(x, y, r), (x', y', r') \in U$ intersect if and only if $(x-x')^2+(y-y')^2 < (r+r')^2$. Thus, a graph on the vertex set $\{1, \dots, n\}$ is an intersection graph of n numbered open disks in the plane if and only if there are points $(x_1, y_1, r_1), \dots, (x_n, y_n, r_n) \in U$ such for all $1 \le i < j \le n$ we have $(x_i-x_j)^2+(y_i-y_j)^2-(r_i+r_j)^2 < 0$ if and only if ij is an edge of the graph. Consider the set of labels $\Lambda = \{$"edge", "non-edge"$\}$, the polynomial $P(x, y, r, x', y', r') = (x-x')^2+(y-y')^2-(r+r')^2$ and the function ϕ given by $\phi(-) =$ "edge" and $\phi(+) = \phi(0) =$ "non-edge". Then intersection graphs of n numbered open disks in the plane then correspond to (P, ϕ, U, Λ)-representable edge-labelings of the complete graph on the vertex $\{1, \dots, n\}$ with labels in Λ.

For any fixed $P_1, \dots, P_k$, ϕ, U and Λ, the following theorem gives an upper bound for the number of $(P_1, \dots, P_k, \phi, U, \Lambda)$-representable edge-labelings . It follows from a theorem of Warren [7] with exactly the same method as in [1, 3, 5].

Theorem 1. *Let us fix a finite set Λ, an integer $d \ge 1$, polynomials $P_1, \dots, P_k \in \mathbb{R}[x_1, \dots, x_d, y_1, \dots, y_d]$, a function $\phi : \{+, -, 0\}^k \to \Lambda$, and a non-empty open subset $U \subseteq \mathbb{R}^d$. Then the number of $(P_1, \dots, P_k, \phi, U, \Lambda)$-representable edge-labelings of the complete graph on the vertex set $\{1, \dots, n\}$ is at most $n^{(1+o(1))dn}$.*

Our main result, Theorem 3 below, states that under some reasonable assumptions, the upper bound in Theorem 1 is tight. In particular, we need an assumption that the open set U is reasonable shaped. This will be made precise by the following definition.

Definition 2. *Let us call a subset $U \subseteq \mathbb{R}^d$* definable by polynomials *if there exists a finite list of real polynomials $Q_1, \dots, Q_\ell$ and a subset $S \subseteq \{+, -, 0\}^\ell$ such that*

$$U = \{x \in \mathbb{R}^d \mid (\operatorname{sgn} Q_1(x), \dots, \operatorname{sgn} Q_\ell(x)) \in S\}.$$

Our main result is the following theorem.

Theorem 3. *Let us fix a finite set Λ, an integer $d \ge 1$, polynomials $P_1, \dots, P_k \in \mathbb{R}[x_1, \dots, x_d, y_1, \dots, y_d]$, a function $\phi : \{+, -, 0\}^k \to \Lambda$, and a non-empty open subset $U \subseteq \mathbb{R}^d$ which is definable by polynomials. Suppose that for any two distinct points $a, a' \in U$ there exists a point $b \in U$ with $P_s(a, b) \ne 0$ and $P_s(a', b) \ne 0$ for all $1 \le s \le k$ and such that*

$$\phi\big(\operatorname{sgn} P_1(a, b), \dots, \operatorname{sgn} P_k(a, b)\big) \ne \phi\big(\operatorname{sgn} P_1(a', b), \dots, \operatorname{sgn} P_k(a', b)\big).$$

Then there are at least $n^{(1-o(1))dn}$ *different* $(P_1, \ldots, P_k, \phi, U, \Lambda)$*-representable edge-labelings of the complete graph on the vertex set* $\{1, \ldots, n\}$*.*

Let us comment on the assumption in Theorem 3 concerning the existence of the desired point $b \in U$ for any two distinct points $a, a' \in U$. Roughly speaking, this assumption is saying that for any two distinct points $a, a' \in U$ there exists a point $b \in U$ such that for $i < j$ the pairs $(a_i, a_j) = (a, b)$ and $(a_i, a_j) = (a', b)$ would lead to different outcomes for the label of the edge ij. An assumption of such a form is necessary in Theorem 3. In applications, this assumption is usually very easy to check.

In our proof proof of Theorem 3 we use some tools from algebraic geometry and differential topology.

References

[1] N. Alon and E. R. Scheinerman, *Degrees of freedom versus dimension for containment orders*, Order **5** (1988), 11–16.

[2] J. Fox, personal communication.

[3] C. McDiarmid and T. Müller, *The number of disk graphs*, European J. Combin. **35** (2014), 413–431.

[4] J. Milnor, *On the Betti numbers of real varieties*, Proc. Amer. Math. Soc. **15** (1964), 275–280.

[5] J. Pach and J. Solymosi, *Crossing patterns of segments*, J. Combin. Theory Ser. A **96** (2001), 316–325.

[6] R. Thom, *Sur l'homologie des variétés algébriques réelles*, In Differential and Combinatorial Topology, pages 255–265, Princeton Univ. Press, 1965.

[7] H. E. Warren, *Lower bounds for approximation by nonlinear manifolds*, Trans. Amer. Math. Soc. **133** (1968), 167–178.

Fast uniform generation of regular graphs and contingency tables

Nick Wormald

(joint work with Andrii Arman and Pu Gao)

1. Introduction

Sampling discrete objects from a specified probability distribution is a classical problem in computer science, both in theory and for practical applications. Uniform generation of random graphs with a specified degree sequence is one such problem that has frequently been studied. When considering the adjacency matrices of graphs, this can be seen to be related to sampling of contingency tables with given marginals, as these are equivalent to bipartite multigraphs with specified vertex degrees.

Many results have been obtained on sampling such objects from an approximately uniform distribution; see for example [10], [9], [1], [5] and [6]. Most of these use a Monte Carlo Markov Chain approach, in which case the main results show that some appropriate Markov chain has polynomially bounded mixing time. The bounds proved are often too large to give practical algoritheorems, though in many cases it is believed that the true mixing time is much smaller than the theoretical results can demonstrate.

On the other hand, a simple rejection-based uniform generation algoritheorem is implicit in the asymptotic enumeration of graphs by degree sequence presented by Békéssy, Békéssy and Komlós [2], Bender and Canfield [3] and Bollobás [4]. The run time of this algoritheorem is linear in n but exponential in the square of the average degree. Hence it only works in practice when degrees are small. McKay and Wormald [11] used the model of random graphs given explicitly in [4] to generate a random multigraph with a given degree sequence. These multigraphs are uniform conditional upon the numbers of loops and multiple edges. Instead of repeatedly rejecting until finding a simple graph, McKay and Wormald used a switching operation to switch away the loops and multiple edges, producing (hopefully) a simple graph in the end. The graphs produced occur uniformly at random, with the given degree sequence. The algoritheorem is rather efficient when the degrees are not too large. In particular, for d-regular graphs it runs in expected time $O(d^3 n)$ when $d = O(n^{1/3})$. (Here and in the following we assume n is the number of vertices.) More switching-based algoritheorems for exactly uniform generation were given which deal with new degree sequences permitting vertices of higher degrees. The regular case was treated by Gao and Wormald [7], again using switchings, for $d = o(\sqrt{n})$ with time complexity again $O(d^3 n)$. Very non-regular but still quite sparse degree sequences (such as power law) were considered by the same authors [8], still using switchings.

We present a new technique for use in switching-based random generation of graphs with given degrees, which we call *incremental relaxation*. For graphs with m edges and maximum degree $\Delta = O(m^4)$, the "best" existing uniform sampler, by McKay and Wormald, runs in time $O(m^2\Delta^2)$. Our new one runs in expected time $O(m)$, which is effectively optimal. (We ignore logaritheoremic factors that are essentially negligible for practical purposes.) For d-regular graphs with $d^2 = o(n)$, the best existing ones run in time $O(nd^3)$. This is now improved to $O(nd + d^4)$, which is of course optimal for $d^3 = O(n)$. For random graphs with power-law degree sequences, with parameter $\gamma > 2.88$, the best known uniform algoritheorems run in time $n^{4.081}$. The new one runs in time $O(n)$.

Making use of incremental relaxation, as well as other ideas, we have a new algoritheorem for uniformly generating random contingency tables with given marginals (equivalently, bipartite multigraphs with given degree sequence) in the sparse case. If the maximum marginal is at most $(5m)^{1/4}$, where m is the sum of the margials, the expected running time is $O(m)$. Essentially, no uniform generator was previously known for this problem, and the algoritheorem covers some cases of parameters for which even no FPAUS (a commonly sought after type of approximately uniform generator) was previously known.

The basic idea of incremental relaxation can be described as follows. Let H be a (small) graph with each edge designated as positive or negative. We say that an H-*anchoring* of a graph G is an injection $Q : V(H) \to V(G)$ that maps every positive edge of H to an edge of G, and every negative edge to a non-edge of G.

(This is a generalisation of rooting at a subgraph, which usually corresponds to the case that H has positive edges only.)

Suppose the aim is to generate a graph G uniformly at random in some set $\mathcal{O}$ of graphs. However, assume that we are given an H-anchored graph (G, Q) uniformly at random. That is, each such ordered pair with $G \in \mathcal{O}$, and with Q being an H-anchoring of G, is equally likely to be presented. We can convert this to a random graph $G \in \mathcal{O}$ by finding the number $b(G)$ of H-anchorings of G, and accepting G with probability $\underline{b}(\mathcal{O})/b(G)$ where $\underline{b}(\mathcal{O})$ is a lower bound on the number of H-anchorings of any element $G' \in \mathcal{O}$. If not accepted, it is *rejected*.

However, computing $b(G)$ can be time-consuming. Our key new idea is to *incrementally* relax the constraints imposed on G by Q, so that the rejection step is replaced by a sequence of "smaller" rejection steps. Set $\emptyset = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_k = V(H)$ and let Q_i denote the restriction of Q to V_i. With this definition, for each i, Q_i is an $H[V_i]$-anchoring of G. Thus Q_i determines some subset (increasing with i) of the constraints on G corresponding to the edges of H. Given that (G, Q_i) is uniformly random, we can obtain a uniformly random anchored graph (G, Q_{i-1}) by applying a rejection strategy similar to that described above, but using only the number $b(G, Q_{i-1})$ of ways that Q_{i-1} can be extended to an $H[V_i]$-anchoring of G. This procedure of incremental relaxation of constraints can be highly advantageous if for each i, $b(G, Q_{i-1})$ can be computed much faster than $b(G)$. In this way, a sequence of uniformly random objects is obtained, involving anchorings at ever-smaller subgraphs of H, until the empty subgraph is reached, corresponding to obtaining G u.a.r.

This general scheme applies to algoritheorems such as the one in [11] for the following reason. The slowness of that algoritheorem is determined by a repeated step which essentially computes the number of choices of two 2-paths $u_1u_2u_3$ and $v_1v_2v_3$ in a graph, for which all six vertices are distinct and none of the edges u_iv_i are present ($i = 1, 2, 3$). This is the quantity $b(G)$ described above, where H is the graph with six vertices, having two 2-paths of positive edges and three negative edges corresponding to the constraints on $u_1u_2u_3$ and $v_1v_2v_3$. Upon further examination of the algoritheorem, it is easy to see that a uniformly random (G, Q) has been obtained, where $\mathcal{O}$ is a set of d-regular multigraphs with a specified number of double edges, and what is desired is a uniformly random $G \in \mathcal{O}$.

To apply incremental relaxation, we let $V_1 = \{u_1, u_2, u_3\}$ and first relax the anchoring to this set of vertices. Thus, in the first incremental step, it is required to compute the number of ways to select the second 2-path, $v_1v_2v_3$, to satisfy the constraints. This can be achieved very quickly by inclusion-exclusion, since the number of "bad" 2-paths can be found by searching G only locally, looking at the vertices of distance at most 2 from V_1. On the other hand, the algoritheorem in [11] fundamentally requires global properties of G. The second incremental step is even easier.

References

[1] M. Bayati, J. H. Kim and A. Saberi. A sequential algoritheorem for generating random graphs. *Algoritheoremica*, 58(4):860–910, 2010.

[2] A. Békéssy, P. Békéssy and J. Komlós. Asymptotic enumeration of regular matrices. *Stud. Sci. Math. Hungar.*, 7:343–353, 1972.

[3] 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.

[4] 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.

[5] P. Diaconis and L. Saloff-Coste. Random walk on contingency tables with fixed row and column sums. Department of Mathematics, Harvard University, 1995.

[6] S. Dittmer and I. Pak. Sampling sparse contingency tables, preprint.

[7] P. Gao and N. Wormald. Uniform generation of random regular graphs. *SIAM Journal on Computing*, 46(4):1395–1427, 2017.

[8] P. Gao and N. Wormald. Uniform generation of random graphs with power-law degree sequences. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algoritheorems*, pages 1741–1758. SIAM, 2018.

[9] C. Greenhill. The switch markov chain for sampling irregular graphs. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algoritheorems*, pages 1564–1572. SIAM, 2014.

[10] M. Jerrum and A. Sinclair. Fast uniform generation of regular graphs. *Theoretical Computer Science*, 73(1):91–100, 1990.

[11] B.D. McKay and N.C. Wormald. Uniform generation of random regular graphs of moderate degree. *Journal of Algoritheorems*, 11(1):52–67, 1990.

Equiangular lines with a fixed angle

Yufei Zhao

(joint work with Zilin Jiang, Jonathan Tidor, Yuan Yao, and Shengtong Zhang)

We say that a set of lines all passing through the origin in $\mathbb{R}^d$ is *equiangular* if their pairwise angles are all equal. It is known that the maximum number of equiangular lines in $\mathbb{R}^d$ is $\Theta(d^2)$ [3, 5], although determining the optimal leading constant is an open problem and the exact maximum number is known in only finitely many dimensions. An interesting feature of all known constructions of $\Theta(d^2)$ equiangular lines in $\mathbb{R}^d$ is that the pairwise angles approach 90° as $d \to \infty$.

What happens if we fix the angle in a configuration of equiangular lines? Let $N_\alpha(d)$ denote the maximum number of lines in $\mathbb{R}^d$ all passing through the origin with pairwise angles $\arccos \alpha$. We are interested in determining $N_\alpha(d)$ for fixed α and large d. This problem was posed by Lemmens and Seidel [5] in 1973. They completely determined the values of $N_{1/3}(d)$ for all d, and in particular proved that $N_{1/3}(d) = 2(d-1)$ for all $d \geq 15$. Neumann (see [5]) showed that $N_\alpha(d) \leq 2d$ unless $1/\alpha$ is an odd integer. It was conjectured by Lemmens and Seidel [5] and subsequently proved by Neumaier [6] that $N_{1/5}(d) = \lfloor 3(d-1)/2 \rfloor$ for all sufficiently large d. Neumaier [6] writes that "the next interesting case [$\alpha = 1/7$] will require substantially stronger techniques."

Progress was stalled on this problem until recently, when Bukh [2] showed that $N_\alpha(d) \leq C_\alpha d$ for some constant C_α, in contrast to the quadratic growth when the angle is not fixed. Then came a surprising breakthrough of Balla, Dräxler, Keevash, and Sudakov [1], who showed that $\limsup_{d\to\infty} N_\alpha(d)/d$, as a function of $\alpha \in (0,1)$, is maximized at $\alpha = 1/3$, and in fact this limit is at most 1.93 unless $\alpha = 1/3$, in which case the limit is 2.

Our main result determines the exact value of $N_{1/(2k-1)}(d)$ for all sufficiently large d, thereby confirming conjectures suggested by [5,6] and stated explicitly in [2]. We also determine $N_\alpha(d)$ for almost all other values of α. To state our result in full generality, we require the following definition.

Definition 1 (Spectral radius order)**.** *Define the* spectral radius order*, denoted $k(\lambda)$, of a real $\lambda > 0$ to be the smallest integer k so that there exists a k-vertex graph G whose adjacency matrix has spectral radius exactly λ. Set $k(\lambda) = \infty$ if no such graph exists.*

Theorem 2 (Main theorem)**.** *Fix $\alpha \in (0,1)$. Let $\lambda = (1-\alpha)/(2\alpha)$ and $k = k(\lambda)$ be its spectral radius order. The maximum number $N_\alpha(d)$ of equiangular lines in $\mathbb{R}^d$ with common angle* $\arccos\alpha$ *satisfies*

(a) $N_\alpha(d) = \lfloor k(d-1)/(k-1)\rfloor$ for all sufficiently large $d > d_0(\alpha)$ if $k < \infty$.
(b) $N_\alpha(d) = d + o(d)$ as $d \to \infty$ if $k = \infty$.

The form of the answer was conjectured by Jiang and Polyanskii [4] (building on [1]). They were the first to observe the role of $k(\lambda)$ in this problem.

If $k \geq 2$ is an integer and $\alpha = 1/(2k-1)$, then $\lambda = k-1$ and $k(\lambda) = k$ (the complete graph K_k is the graph on fewest vertices with spectral radius $k-1$), so the following corollary extends the only two previously known cases of $k = 2$ [5] and $k = 3$ [6] to all values of k.

Corollary 3. *For every fixed integer $k \geq 2$, one has $N_{1/(2k-1)}(d) = \lfloor k(d-1)/(k-1)\rfloor$ for all sufficiently large $d > d_0(k)$.*

The equiangular lines problem can be recast in graph theory terms. A set of N equiangular lines can be represented by unit vectors $v_1, \ldots, v_N \in \mathbb{R}^d$ with $\langle v_i, v_j\rangle = \pm\alpha$ for all $i \neq j$. Construct the graph G on N vertices with edge relations $i \sim j$ if and only if $\langle v_i, v_j\rangle = -\alpha$. The Gram matrix of pairwise inner products of these vectors is equal to $(\langle v_i, v_j\rangle)_{1\leq i,j\leq N} = (1-\alpha)I + \alpha(J - 2A_G)$. The Gram matrix is positive semidefinite, which implies that the following two statements are equivalent, where $\lambda = (1-\alpha)/(2\alpha)$ as earlier:

- There exists a family of N equiangular lines in $\mathbb{R}^d$ with common angle $\arccos\alpha$.
- There exists an N-vertex graph G such that the matrix $\lambda I - A_G + \frac{1}{2}J$ is positive semidefinite and has rank at most d (A_G is the adjacency matrix of $G$and J is the all-ones matrix).

Using rank-nullity relations, we can bound N in terms of the multiplicity of λ as an eigenvalue of A_G. Note that λ must be one of the two largest eigenvalues of A_G due to the positive semidefiniteness of the Gram matrix.

The top eigenvalue of A_G for a connected graph G has multiplicity one, due to the Perron–Frobenius theorem. While some graphs have high second eigenvalue multiplicity, not all graphs can arise from an equiangular lines configuration. An important step towards the solution is the following result from [1], showing that one only needs to consider graphs of bounded degree.

Theorem 4 ([1])**.** *For every $\alpha \in (0,1)$, there exists some $\Delta = \Delta(\alpha)$ so that for every set of equiangular lines in $\mathbb{R}^d$ with common angle $\arccos\alpha$, one can choose a set S of unit vectors, with one unit vector in the direction of each line in the equiangular set, so that each unit vector in S has inner product $-\alpha$ with at most Δ other vectors in S.*

We prove a new result in spectral graph theory showing that the second eigenvalue multiplicity of a connected bounded degree graph is sublinear, which is enough to deduce the main theorem concerning equiangular lines.

Theorem 5. *For every Δ there is a constant $C = C(\Delta)$ so that the multiplicity of the second largest eigenvalue of the adjacency matrix of a connected n-vertex graph with maximum degree at most Δ is at most $Cn/\log\log n$.*

To further illustrate Theorem 5, let us give some near-miss examples:

- A disjoint union of triangles has linear multiplicity of all eigenvalues, but it is not connected.
- A strongly regular graph can have linear second eigenvalue multiplicity, but it does not have bounded degree.
- The n-vertex graph obtained from a cycle of length $n/3$ by attaching two pendent edges to every vertex of the cycle has the eigenvalue 0 with linear multiplicity, but 0 is not one of the top eigenvalues.

References

[1] I. Balla, F. Dräxler, P. Keevash, and B. Sudakov, *Equiangular lines and spherical codes in Euclidean space*, Invent. Math. **211** (2018), 179–212.

[2] B. Bukh, *Bounds on equiangular lines and on related spherical codes*, SIAM J. Discrete Math. **30** (2016), 549–554.

[3] D. de Caen, *Large equiangular sets of lines in Euclidean space*, Electron. J. Combin. **7** (2000), Research Paper 55, 3 pp.

[4] Z. Jiang and A. Polyanskii, *Forbidden subgraphs for graphs of bounded spectral radius, with applications to equiangular lines*, Israel J. Math., to appear.

[5] P. W. H. Lemmens and J. J. Seidel, *Equiangular lines*, J. Algebra **24** (1973), 494–512.

[6] A. Neumaier, *Graph representations, two-distance sets, and equiangular lines*, Linear Algebra Appl. **114/115** (1989), 141–156.

Problem Session

Nathan Linial (chair)

Noga Alon

Induced subgraphs of Cayley graphs of Z_2^n

For a subset S of the elementary abelian 2-group Z_2^n let $G = G(Z_2^n, S)$ denote the Cayley graph of Z_2^n with respect to S. Is it true that for every such S and G and for every subset A of at least $2^{n-1}+1$ vertices of G, the maximum degree of the induced subgraph of G on A is at least $\sqrt{|S|}$?

The fact that this is true for the set S of the n vectors of Hamming weight 1 was proved by Huang in [1]. Together with Kai Zheng we proved that this is also true for every set obtained from the above S by adding to it another vector, and in several other cases (sometimes with a lot of room to spare). The proof for some cases applies the technique of Huang but is based on complex signing of the adjacency matrix of the graph, and it can be shown that in several cases no real signing exhibiting the result exists.

Note added in proof: We have recently proved that the answer to the question is "yes". A proof appears in: N. Alon and K. Zheng, Unitary signings and induced subgraphs of Cayley graphs of Z_2^n, manuscript, 2020.

References

[1] Hao Huang, *Induced subgraphs of hypercubes and a proof of the sensitivity conjecture* Ann. of Math. (2) 190 (2019), no. 3, 949–955.

József Balogh

Maximum number of independent sets in uniform hypergraphs

The question that we would like to draw attention concerns the hypergraph analogue of an old (and now resolved) graph-theoretic problem of Granville (see [1]). Granville raised the following problem: which d-regular graphs on n vertices have the maximum number of independent sets? This problem was also considered by Kahn [4] in the context of the hard-core model, and a complete answer is now available owing to the work of Kahn [4] and Zhao [5]: the extremal graphs are precisely those consisting of disjoint copies of the complete bipartite graph $K_{d,d}$.

Here, we shall focus on the analogous problem for r-uniform hypergraphs (or r-graphs, for short). While this is a natural problem, we emphasise that it is not even apparent what the correct conjectural analogue of the complete bipartite graph is; our aim is to remedy this situation. Of course, there are multiple notions of 'independent sets' and 'degrees' in hypergraphs, but we shall restrict ourselves to the most common ones: a subset of the vertex set of an r-graph is *independent* if it induces no edges, the *degree* of a vertex is the number of edges containing it, and an r-graph is *d-regular* if each of its vertices has degree d.

For $r \geq 2$ and $d \in \mathbb{N}$, let $\mathcal{H}_d^r$ be the d-regular r-graph on rd vertices whose d^2 edges are as follows: we mark a subset of the vertex set of order d, fix an $(r-1)$-uniform matching of size d on the remaining $(r-1)d$ unmarked vertices, and then include in the edge set of $\mathcal{H}_d^r$ each r-set consisting of a marked vertex and a matching edge. For example, $\mathcal{H}_d^2$ is the complete bipartite graph $K_{d,d}$, and $\mathcal{H}_d^3$ is the set of triangles in the graph on $3d$ vertices where d vertices are each joined to both ends of all the edges of a matching covering the other $2d$ vertices.

Writing $\mathrm{ind}(\mathcal{G})$ for the number of independent sets in an r-graph $\mathcal{G}$, an easy computation reveals that

$$\mathrm{ind}(\mathcal{H}_d^r) = (2^d - 1)(2^{r-1} - 1)^d + 2^{(r-1)d}.$$

Our main conjecture is the following.

Conjecture 1. *For all $r \geq 2$ and $d \in \mathbb{N}$, if $\mathcal{G}$ is a d-regular r-graph on n vertices, then*

$$\mathrm{ind}(\mathcal{G}) \leq \mathrm{ind}(\mathcal{H}_d^r)^{n/rd}.$$

By way of orientation, let us mention that when $r \geq 3$, a disjoint union of copies of $\mathcal{H}_d^r$ has strictly more independent sets than a comparable disjoint union of copies of the complete r-partite r-graph (when the numerics allow it, i.e., when $d = t^{r-1}$ for some $m \in \mathbb{N}$). Conjecture 1 for $r = 2$ is the aforementioned Kahn–Zhao theorem, but we are unable to verify it even when $r = 3$. Nevertheless, in the spirit of Kahn [4], we can verify our conjecture for some class of r-graphs, see [2], this class appears to capture the extent of the entropic approach of Kahn [4], reduces to the class of bipartite graphs when $r = 2$.

There has been some recent (independent) interest around finding a statement in the spirit of Conjecture 1; for example, Cohen, Perkins, Sarantis and Tetali [3] study an analogue of the problem treated here for regular *linear* r-graphs, and raise the question of what one can say about regular r-graphs in general.

References

[1] N. Alon, Independent sets in regular graphs and sum-free subsets of finite groups, *Israel J. Math.*, 73, 1991, 247–256.

[2] J. Balogh, B. Bollobás, and B. Narayaran, Counting problems in regular hypergraphs, manuscript.

[3] E. Cohen, W. Perkins, M. Sarantis, and P. Tetali, On the Number of Independent Sets in Uniform, Regular, Linear Hypergraphs, preprint.

[4] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combinatorics, Probability and Computing*, 10, 2001, 219–237.

[5] Y. Zhao, The number of independent sets in a regular graph, *Combinatorics, Probability and Computing*, 19, 2010, 315–320.

Maria Chudnovsky

Poly-time algoritheorem for stable sets in perfect graphs

Let G be a perfect graph and assume that G^c contains no induced cycle of length at least six. Let $K_1, \dots, K_t$ be a list of cliques of G such that

- $|K_i| = \omega(G)$ for every i , and
- $\bigcup_{i=1}^{t} K_i = V(G)$.

Design a combinatorial polynomial time algoritheorem that finds a stable set S of G such that $K_i \cap S \neq \emptyset$ for every i.

Gil Kalai

Avi's 1-2-3 conjecture

Find an *explicit construction* of a bipartite graph $G = (A \cup B, E)$ with colour classes A and B such that $|A| = n^3$, $|B| = n^2$, and the vertices in A have a bounded degree, where the following holds: for every $S \subseteq A$ of size n, there is a perfect matching in the subgraph $G[S, B]$.

Super-linear Gaussian elimination

Let M be an $n \times n$ matrix and let $\mathrm{gc}(M)$ be defined as

$$\mathrm{gc}(M) := \min_S \{M = E_1, \dots, E_S\},$$

where each E_i is an elementary matrix. In other words, $\mathrm{gc}(M)$ is the minimum number of Gaussian elimination operations needed to make M diagonal.

Find an explicit M such that $\mathrm{gc}(M)$ is super-liear.

Dan Král'

Cycles of length divisible by four in tournaments

Let $d(\ell, n)$ be the ratio of the maximum number of cycles of length ℓ that can be contained in an n-vertex tournament and the expected number of cycles of length ℓ in the random n-vertex tournament (note that $d(\ell, n) \geq 1$ for trivial reasons). Show that the following holds for every ℓ that is divisible by four.

$$\lim_{n\to\infty} = 1 + 2 \cdot \sum_{i=0}^{\infty} \left(\frac{2}{(2i+1)\pi} \right)^{\ell}.$$

Day [2]Conjecture 45 provided a construction giving this bound and conjectured the construction to be optimal (though he did not compute the limit), i.e., the limit is at least the right hand side. A classical result of Beineke and Harary [1] yields the conjecture $\ell = 4$ when the limit value is 4/3. Grzesik, Lovász Jr., Volec [3] and I have recently proved the conjecture for $\ell = 8$ when the limit value is 332/315. We were also able to give the following asymptotic result: for every $\varepsilon > 0$, there exists ℓ_0 such that the limit is at most

$$1 + \left(\frac{2}{\pi} + \varepsilon \right)^{\ell}$$

for every $\ell \geq \ell_0$ that divisible by four. Finally, we remark that if ℓ is not divisible by four, then the limit is equal to one [3].

References

[1] L. Beineke and F. Harary: *The maximum number of strongly connected subtournaments*, Canadian Mathematical Bulletin **8** (1965), 491–498.

[2] A. N. Day: A collection of problems in extremal combibatorics, PhD Thesis, Queen Mary University of London, 2017.

[3] A. Grzesik, D. Král', L. M. Lovász, J. Volec: *Cycles of a given length in tournaments*, in preparation.

Nati Linial

Internal partitions of graphs

An *external partition* of a graph $G = (V, E)$ is a partition of V into two sets in which every vertex has more neighbours in the other part than in its own. It is trivial to see that every graph has such a partition—namely, just take the max cut.

On the other hand, not much is known about an *internal partition* which is a partition of V into two (non-empty!) sets such that every vertex has at least as many neighbours in its own part as in the other. Not all graphs have an internal partition, for example if $d \geq 3$ is odd then $K_{d,d}$ has no such partition.

It has been conjectured (see [1, 2]) that for every d, there is $n_0(d)$ such that for all $n \geq n_0$ every n-vertex d-regular graph has an internal partition. Even more

Conjecture 2. *For every d there are only finitely many d-regular graphs with no internal partition.*

This is known for $d \in \{3, 4, 6\}$ [1]. The first unknown example is $d = 5$. On the other hand, it is known that asymptotically almost every $2d$-regular graph has an internal partition [2].

References

[1] A. Ban, N. Linial, *Internal partitions of regular graphs*, Journal of Graph Theory, 2016

[2] S. Louis, N. Linial, *Asymptotically almost every $2r$-regular graph has an internal partition*, Master Thesis, 2017

Bhargav Narayanan

Existence of geometric thresholds

Let M be a compact metric space (think unit 2-sphere) and let $G(n, r)$ denote the random graph on n vertices generated by independently sampling n points from some fixed probability distribution on M (usually, uniform), and joining any pair of points at distance at most r from each other. We say that $r^* = r^*(n)$ is a *geometric threshold* for an up-set F of graphs on $[n]$ if $\tau_r(F)$ tends to either 1 or 0 depending on whether $r \gg r^*$ or $r \ll r^*$; here, we write τ_r for the law of the random geometric graph $G(n, r)$. For what metric spaces M do all up-sets of graphs admit geometric thresholds? There are compact examples where thresholds are not guaranteed, but these examples are not 'well-behaved'. In particular, one expects a positive answer for the 2-sphere or the unit square, but very little seems to be known.

Alexey Pokrovskiy

Rainbow trees in k-factorised complete graphs

We say a colouring of a graph G is a k-factorisation if every colour class is k-regular. A coloured graph is said to be rainbow if all its edges have distinct colours.

Conjecture. *Given $k \geq 2$ in any k-factorisation of the complete graph K_{nk+1} one can find a rainbow copy of any tree with n edges.*

The conjecture is not true for $k = 1$.

Tibor Szabó

3-coloring 6-regular 6-uniform hypergraphs

Problem 1. *Is it true that every 6-regular 6-uniform hypergraph has a 3-coloring of its vertices such that no hyperedge has four vertices of the same color?*
More generally: for what values of k does it hold that every $(2k)$-regular $(2k)$-uniform hypergraph has a 3-coloring of its vertices, such that no hyperedge contains more than k vertices of the same color?

For $k = 1$ the appropriate 3-coloring always exists (of course), while for $k = 2$ there is a counterexample. For large enough constant k, the appropriate 3-coloring always exists by the Local Lemma.

The motivation for this problem comes from the concept of majority coloring of digraphs. A *majority coloring* of a digraph D with k colors is an assignment $c : V(D) \to \{1, \ldots, k\}$ such that for every $v \in V(D)$, we have $c(w) = c(v)$ for at most half of all out-neighbors $w \in N^+(v)$. This notion of coloring was first introduced and studied by Kreutzer, Oum, Seymour, van der Zypen, and Wood [2], who showed that every digraph has a majority 4-coloring. Odd directed cycles provide examples of digraphs that are not majority 2-colorable. However, so far no example of a digraph is known that requires the use of four colors. Kreutzer et al. conjectured that there are none.

Conjecture 2 ([2]). *Every digraph is majority 3-colorable.*

A natural benchmark for the study of this conjecture are r-regular digraphs. It is easy to check the validity of the conjecture for 1- and 2-regular digraphs. The Local Lemma implies that the uniform random 3-coloring works for r-regular digraphs when r is a large enough constant (Kreutzer et al [2] mention $r \geq 144$). Digraphs with small minimum out-degree seem to be outside the realm of probabilistic methods. Together with Michael Anastos, Ander Lamaison, and Raphael Steiner [1] we show that the conjecture also holds for $r = 3$ and 4. The methods used in our paper are unlikely to resolve Conjecture 2 for the open cases of 5- and 6-regular digraphs. One possible approach could be via an extension to hypergraphs: Given a 5-regular digraph D, consider the hypergraph $\mathcal{H}(D)$ with vertex set $V(D)$ and whose edges are $\{v\} \cup N^+(v)$, $v \in V(D)$. This hypergraph is 6-regular and 6-uniform. If we could now find a vertex-3-coloring of $\mathcal{H}(D)$ such that no hyperedge contains four vertices of the same color, this coloring would certainly be a majority coloring of D.

For more intriguing open probelms on majority colorings, see [1, 2].

References

[1] M. Anastos, A. Lamaison, R. Steiner, T. Szabó, Majority Colorings of Sparse Digraphs, *submitted*.

[2] S. Kreutzer, S.-i. Oum, P. Seymour, D. van der Zypen, D. Wood, *Majority Colorings of Digraphs*, The Electronic Journal of Combinatorics, **24** (2017), P2.25.

Reporters: Matija Bucić, Jinyoung Park and Miloš Trujić

Participants

Prof. Dr. Karim Adiprasito
Einstein Institute for Mathematics
The Hebrew University of Jerusalem
Edmond J. Safra Campus
Givat Ram
Jerusalem 9190401
ISRAEL

Prof. Dr. Ron Aharoni
Department of Mathematics
TECHNION-Israel Institute of
Technology
Amado Building, Rm. 600
Haifa 3200003
ISRAEL

Dr. Peter D. Allen
Department of Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
UNITED KINGDOM

Prof. Dr. Noga Alon
Department of Mathematics
Princeton University
Office: 706 Fine Hall
Princeton NJ 08544
UNITED STATES

Prof. Dr. József Balogh
Department of Mathematical Sciences
University of Illinois
Office: 233 B Illini Hall
1409 West Green Street
Urbana IL 61801
UNITED STATES

Prof. Dr. Alexander Barvinok
Department of Mathematics
University of Michigan
1858 East Hall
Ann Arbor, Michigan 48109-1043
UNITED STATES

Matija Bucić
Department of Mathematics
ETH Zürich
Rämistrasse 101
8092 Zürich
SWITZERLAND

Dr. Boris Bukh
Department of Mathematical Sciences
Carnegie Mellon University
6202 Wean Hall
5000 Forbes Avenue
Pittsburgh PA 15213
UNITED STATES

Prof. Dr. Maria Chudnovsky
Department of Mathematics
Princeton University
211 Fine Hall
68 Washington Road
Princeton, NJ 08544
UNITED STATES

Prof. Dr. David Conlon
Department of Mathematics
California Institute of Technology
1200 E California Boulevard
Pasadena CA 91125
UNITED STATES

Prof. Dr. Reinhard Diestel
Mathematisches Seminar
Universität Hamburg
Bundesstrasse 55
20146 Hamburg
GERMANY

Prof. Dr. Jacob Fox
Department of Mathematics
Stanford University
Building 380
Stanford CA 94305
UNITED STATES

Prof. Dr. Ehud Friedgut
Department of Mathematics
The Weizmann Institute of Science
234 Herzl Street
P.O. Box 26
Rehovot 7610001
ISRAEL

Lior Gishboliner
Department of Mathematics
School of Mathematical Sciences
Tel Aviv University
P.O. Box 39040
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Prof. Dr. Penny E. Haxell
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON N2L 3G1
CANADA

Dr. Annika Heckel
Mathematisches Institut
Universität München
Theresienstrasse 39
80333 München
GERMANY

Dr. Matthew Jenssen
Mathematical Institute
Oxford University
24-29 St. Giles
Oxford OX1 3LB
UNITED KINGDOM

Prof. Dr. Jeff Kahn
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway NJ 08854-8019
UNITED STATES

Prof. Dr. Gil Kalai
Einstein Institute for Mathematics
The Hebrew University of Jerusalem
Givat-Ram
Jerusalem 9190401
ISRAEL

Prof. Dr. Peter Keevash
Mathematical Institute
University of Oxford
Andrew Wiles Building
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Daniel Král
Faculty of Informatics
Masaryk University
Office: C 517
Botanická 68 A
602 00 Brno
CZECH REPUBLIC

Prof. Dr. Michael Krivelevich
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
P.O. Box 39040
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Dr. Matthew A. Kwan
Department of Mathematics
Stanford University
Building 380
Stanford CA 94305-2125
UNITED STATES

Dr. Shoham Letzter
Institute for Theoretical Studies
ETH Zürich, CLV Building
HG G 61.2
Clausiusstrasse 47
8092 Zürich
SWITZERLAND

Prof. Dr. Nathan Linial
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Givat Ram
Jerusalem 9190401
ISRAEL

Prof. Eyal Lubetzky
Courant Institute of Mathematical Sciences
New York University
Warren Weaver Hall 813
251 Mercer Street
New York NY 10012
UNITED STATES

Prof. Dr. Tomasz Łuczak
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 4
61-614 Poznań
POLAND

Richard H. Montgomery
School of Mathematics
University of Birmingham
Edgbaston
Birmingham B15 2TT
UNITED KINGDOM

Dr. Hoi H. Nguyen
Department of Mathematics
The Ohio State University
100 Mathematics Building
231 West 18th Avenue
Columbus, OH 43210-1174
UNITED STATES

Prof. Dr. János Pach
SB MATHGEOM DCG
E P F L, MA C1 577
Station 8
1015 Lausanne
SWITZERLAND

Prof. Dr. Igor Pak
Department of Mathematics
University of California, Los Angeles
Math. Sciences 6125
405 Hilgard Avenue
Los Angeles CA 90095
UNITED STATES

Prof. Dr. Konstantinos Panagiotou
Mathematisches Institut
Universität München
Theresienstrasse 39
80333 München
GERMANY

Jinyoung Park
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway NJ 08854-8019
UNITED STATES

Dr. Guillem Perarnau
Departament de Matemàtica Aplicada
Universitat Politècnica de Catalunya
Despatx 13, Edifica C 3
Avinguada del Canal Olimpic, 15
08860 Castelldefels
SPAIN

Prof. Dr. Bhargav Peruvemba Narayanan
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway NJ 08854-8019
UNITED STATES

Dr. Alexey Pokrovskiy
Department of Mathematics and
Statistics
Birkbeck College
University of London
Malet Street, Bloomsbury
London WC1E 7HX
UNITED KINGDOM

Dr. Christian Reiher
Department Mathematik
Universität Hamburg
Bundesstrasse 55
20146 Hamburg
GERMANY

Prof. Dr. Oliver M. Riordan
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Vojtěch Rödl
Department of Mathematics and
Computer Science
Emory University
400 Dowman Drive
Atlanta GA 30322
UNITED STATES

Prof. Wojciech Samotij
School of Mathematical Sciences
Tel Aviv University
P.O. Box 39040
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Lisa Sauermann
Department of Mathematics
Stanford University
Building 380
450 Serra Mall
Stanford CA 94305
UNITED STATES

Prof. Dr. Alexander Scott
Mathematical Institute
University of Oxford
Radcliffe Observatory Quarter
Andrew Wiles Building
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Asaf Shapira
Department of Mathematics
School of Mathematical Sciences
Tel Aviv University
P.O. Box 39040
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Prof. Dr. József Solymosi
Department of Mathematics
University of British Columbia
1984 Mathematics Road
Vancouver V6T 1Z2
CANADA

Prof. Dr. Angelika Steger
Institut für Theoretische Informatik
ETH Zürich, CAB G 37.2
Universitätstrasse 6
8092 Zürich
SWITZERLAND

Prof. Dr. Benjamin Sudakov
Department of Mathematics
ETH Zürich, HG G 65.1
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Tibor Szabó
Institut für Mathematik
Freie Universität Berlin
Arnimallee 6
14195 Berlin
GERMANY

Dr. Istvan Tomon
Departement Mathematik
ETH - Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Miloš Trujić
Institut für Theoretische Informatik
ETH Zürich, CAB G 32.2
Universitätstrasse 6
8092 Zürich
SWITZERLAND

Prof. Nicholas Wormald
School of Mathematical Sciences
Monash University
Room 455
9 Rainforest Walk
Clayton, VIC 3800
AUSTRALIA

Prof. Yufei Zhao
Department of Mathematics
Massachusetts Institute of Technology
Room 2-271
77 Massachusetts Avenue
Cambridge MA 02139
UNITED STATES