Mathematisches Forschungsinstitut Oberwolfach

Report No. 1/2017

DOI: 10.4171/OWR/2017/1

Combinatorics

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

1 January – 7 January 2017

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 abstracts of the presentations and a summary of the problem session.

Mathematics Subject Classification (2010): 05-XX.

Introduction by the Organisers

The workshop Combinatorics, organized by Jeff Kahn (Piscataway), Angelika Steger (Zürich) and Benny Sudakov (Zürich), was held the first week of January, 2017. Despite the early point in the year the meeting was well attended, with roughly 50 participants from the US, Canada, Brazil, UK, Israel, and various European countries. The program consisted of 11 plenary lectures and 18 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 statements of 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 two of the more spectacular developments.

The first of these is Ron Peled's resolution (with Yinon Spinka) of a more than 30-year-old conjecture of R. Kotecký on the antiferromagnetic Potts model. (Their results apply in the far more general context of *graph homomorphisms,* but here we only discuss the special case that was the subject of R. Kotecký's conjecture.)

For a fixed $q \geq 3$, large fixed d, and very large n, we consider σ chosen uniformly at random from the set of proper q-colorings of the "box"

$$\Lambda_n^d = \{x \in \mathbb{Z}^d : \|x\|_\infty \leq n\}.$$

In the language of statistical mechanics this is the *antiferromagnetic Potts model at temperature zero.* Kotecký conjectured (circa 1985) an affirmative answer to the basic ("phase transition" or, more properly, "phase coexistence") question for this model; namely, *does the model admit multiple maximal entropy Gibbs measures?* Such an answer is supplied by the following combinatorial statement, proved by Peled and Spinka, in which ∂ denotes boundary (thus $\partial\Lambda_n^d = \{x \in \mathbb{Z}^d : \|x\|_\infty = n\}$).

Theorem. *If $A \cup B$ is an equipartition of $\{1, \ldots, q\}$ (that is, $||A| - |B|| \leq 1$) and we condition on σ assigning colors from A (resp. B) to all even (resp. odd) vertices of $\partial\Lambda_n^d$, then*

$$\Pr(\sigma(\underline{0}) \in A) > 1 - \epsilon_d,$$

where $\epsilon_d \to 0$ as $d \to \infty$.

That is: in high enough (fixed) dimension, the influence of the boundary conditions on behavior at the origin persists as the boundary recedes. For $q = 3$ this had been known for about 15 years and stronger results have more recently been established by Peled and others; but the jump to $q \geq 4$ had proved wholly intractable until this breakthrough.

Our second highlight is the work discussed by Daniela Kühn (obtained jointly with Stefan Glock, Allan Lo and Deryk Osthus), giving a "second generation" proof of a vast generalization of the "Existence Conjecture" for block designs. (This famous nineteenth century problem was settled a few years ago by Peter Keevash and first announced at the January 2014 edition of "Combinatorics.") To give some flavor, we just state (skipping some specifics) a special case of the original conjecture:

For fixed $2 \leq t \leq k$ and any sufficiently large n obeying some obviously necessary congruential restrictions, it is possible to find a collection of k-subsets of $\{1, \ldots, n\}$ containing each t-element set exactly once.

For $t = 2$ this was proved in celebrated work of R.M. Wilson in the early 1970's, but there was really nothing between Wilson and Keevash that could be called serious progress on this basic problem; indeed—perhaps most strikingly—*for $t \geq 6$ not a single example of such a collection was known to exist.*

Keevash's work, based on a new notion of "randomized algebraic construction," was a complete departure from the methods that have dominated design theory for the last forty years. (Recent times have seen a number of such examples; that is, of celebrated problems that had usually been attacked by other methods succumbing to ideas more prevalent in the community at the heart of the present meeting.)

The formidable work of Kühn *et al.* goes well beyond Keevash, whose approach already encompasses much more than the Existence Conjecture (as indeed it must to succeed). Here the algebraic aspects of Keevash are replaced by a highly sophisticated ("iterative") version of the "absorbing method" of Rödl, Ruciński and Szemerédi, and the work may be considered a sort of culmination (at least to date) in the development of this powerful approach. A description of even a representative subset of the new results seems infeasible for the present summary. The new, more purely combinatorial approach provides considerable flexibility that looks sure to lead to many further developments.

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-1049268, "US Junior Oberwolfach Fellows".

Workshop: Combinatorics

Table of Contents

Abstracts

Equiangular lines and spherical codes

Igor Balla
(joint work with F. Draxler, P. Keevash, and B. Sudakov)

1. Introduction

A set of lines through the origin in $\mathbb{R}^n$ is called *equiangular* if any pair of lines defines the same angle. Equiangular sets of lines appear naturally in various areas of mathematics. In elliptic geometry, they correspond to equilateral sets of points, or, in other words, to regular simplexes [7]. In frame theory, so-called optimal Grassmannian frames form equiangular sets of lines [8]. In the theory of polytopes, the convex hull of the points of intersection of an equiangular set of lines with the unit sphere is a spherical polytope of some kind of regularity [3].

It is therefore a natural question to determine the maximum cardinality $N(n)$ of an equiangular set of lines in $\mathbb{R}^n$. This is considered to be one of the founding problems of algebraic graph theory, see e.g. [6, p. 249]. It was first studied by Haantjes [7] in 1948, who showed that $N(3) = N(4) = 6$ and that an optimal configuration in 3 (and 4) dimensions is given by the 6 diagonals of a convex regular icosahedron. In 1966, van Lint and Seidel [10] formally posed the problem of determining $N(n)$ for all positive integers n.

A general upper bound of $\binom{n+1}{2}$ on $N(n)$ was established by Gerzon, see [9]. In dimensions 2 and 3 this gives upper bounds of 3 and 6, respectively, matching the actual maxima. In $\mathbb{R}^7$, the above bound shows $N(7) \leq 28$. This can be achieved by considering the set of all 28 permutations of the vector (1,1,1,1,1,-3,-3) [10]. It is also known that there is an equiangular set of 276 lines in $\mathbb{R}^{23}$, see e.g. [9], which again matches Gerzon's bound. Strikingly, these four examples are the only known ones to match his bound. In fact, for a long time it was even an open problem to determine whether n^2 is the correct order of magnitude. In 2000, de Caen [2] constructed a set of $2(n+1)^2/9$ equiangular lines in $\mathbb{R}^n$ for all n of the form $3 \cdot 2^{2t-1} - 1$.

Interestingly, all known examples of size $\Theta(n^2)$ have a common angle on the order of $\arccos(1/\sqrt{n})$. On the other hand, all known construction of equiangular lines with a fixed common angle have much smaller size. It is therefore natural to consider the maximum number $N_\alpha(n)$ of equiangular lines in $\mathbb{R}^n$ with common angle $\arccos\alpha$, where α does not depend on dimension. This question was first raised by Lemmens and Seidel [9] in 1973, who showed that for sufficiently large n, $N_{1/3}(n) = 2n - 2$ and also conjectured that $N_{1/5}(n)$ equals $\lfloor 3(n-1)/2 \rfloor$. This conjecture was later confirmed by Neumaier [11]. Interest in the case where $1/\alpha$ is an odd integer was due to a general result of Neumann [9, p. 498], who proved that if $N_\alpha(n) \geq 2n$, then $1/\alpha$ is an odd integer.

Despite active research on this problem, for many years these were the best results known. Recently, Bukh [1] made important progress by showing that

$N_\alpha(n) \le c_\alpha n$, where $c_\alpha = 2^{O(1/\alpha^2)}$ is a large constant only depending on α. We completely resolve the question of maximizing $N_\alpha(n)$ over constant α, showing that for sufficiently large n, $N_\alpha(n)$ is maximized at $\alpha = \frac{1}{3}$.

Theorem 1. *Fix $\alpha \in (0,1)$. For n sufficiently large relative to α, the number of equiangular lines in $\mathbb{R}^n$ with angle $\arccos\alpha$ is bounded by $2n-2$ if $\alpha = \frac{1}{3}$ and by $1.93n$ otherwise.*

A more general setting than that of equiangular lines is the framework of *spherical L-codes*, introduced in a seminal paper by Delsarte, Goethals and Seidel [5] in 1977 and extensively studied since.

Definition 2. Let L be a subset of the interval $[-1,1)$. A finite non-empty set C of unit vectors in Euclidean space $\mathbb{R}^n$ is called a *spherical L-code*, or for short an *L-code*, if $\langle x, y\rangle \in L$ for any pair of distinct vectors x, y in C.

Note that if $L = \{\pm\alpha\}$ where $\alpha \in [0,1)$, then an L-code corresponds to a set of equiangular lines with common angle $\arccos\alpha$. For $L = [-1,\beta]$, finding the maximum cardinality of an L-code is equivalent to the classical problem of packing spherical caps of angular radius $\frac{1}{2}\arccos\beta$ on the sphere. Generalising Gerzon's result, Delsarte, Goethals and Seidel [4] obtained bounds on the cardinality of sets of lines having k distinct angles. That is, they proved that, for $L = \{-\alpha_1, \dots, -\alpha_k, \alpha_1, \dots, \alpha_k\}$ and $\alpha_1, \dots, \alpha_k \in [0,1)$, spherical L-codes have size at most $O(n^{2k})$. They subsequently extended this result to an upper bound of $O(n^{|L|})$ on the size of any L-code [5].

Bukh [1] observed that, in some sense, the negative values of L pose less of a constraint on the size of L-codes than the positive ones, as long as they are separated away from 0. Specifically, he proved that for $L = [-1,-\beta] \cup \{\alpha\}$, where $\beta \in (0,1)$ is fixed, the size of any L-code is at most linear in the dimension. Motivated by the above-mentioned work of Delsarte, Goethals and Seidel [4] Bukh made an analogous conjecture for $[-1,-\beta] \cup \{\alpha_1, \dots, \alpha_k\}$-codes, which we verify as follows.

Theorem 3. *Let $L = [-1,-\beta] \cup \{\alpha_1, \dots, \alpha_k\}$ for some fixed $\beta \in (0,1]$. Then there exists a constant $c_{\beta,k}$ such that any spherical L-code in $\mathbb{R}^n$ has size at most $c_{\beta,k} n^k$. Moreover, if $0 \le \alpha_1 < \ldots < \alpha_k < 1$ are also fixed then such a code has size at most*

$$2^k (k-1)! \Big(1 + \frac{\alpha_1}{\beta}\Big) n^k + o(n^k).$$

In particular, if $\alpha_1, \dots, \alpha_k$ are fixed this substantially improves the aforementioned bound of Delsarte, Goethals and Seidel [4, 5] from $O(n^{2k})$ to $O(n^k)$. We furthermore show that the second statement of Theorem 2 is tight up to a constant factor.

Theorem 4. *Let n, k, r be positive integers and $\alpha_1 \in (0,1)$ with k and α_1 being fixed and $r \le \sqrt{n}$. Then there exist $\alpha_2, \dots, \alpha_k$, $\beta = \alpha_1/r - O(\sqrt{\log(n)/n})$ and a spherical L-code of size $(1+r)\binom{n}{k}$ in $\mathbb{R}^{n+r}$ with $L = [-1,-\beta] \cup \{\alpha_1, \dots, \alpha_k\}$.*

2. Open problems

If $\alpha = 1/(2r-1)$ for some positive integer r, one can construct an equiangular set of $r\lfloor (n-1)/(r-1) \rfloor$ lines with angle $\arccos(1/(2r-1))$. For $r = 2$ [10] and $r = 3$ [11], these are the extremal constructions. This motivates the following conjecture, which was also raised by Bukh [1].

Conjecture 5. *Let $r \geq 2$ be a positive integer. Then, for sufficiently large n,*

$$N_{\frac{1}{2r-1}}(n) = \frac{r(n-1)}{r-1} + O(1).$$

If α is not of the above form, there aren't any known constructions with more than $(1 + o(1))n$ lines.

For a set of lines attaining k prescribed angles, we have the general bound of $O(n^{2k})$ by Delsarte, Goethals and Seidel [4]. The constructions of $\Omega(n^2)$ equiangular lines shows that this bound is the correct order of magnitude for $k = 1$. However, no constructions of size $\Omega(n^{2k})$ are known for $k \geq 2$.

Theorem 3 shows that spherical $[-1, 0) \cup \{\alpha\}$-codes in $\mathbb{R}^n$ with α fixed can have size at least $n^{3/2-\epsilon}$, answering a question of Bukh [1]. On the other hand, we can show that such codes have size at most $O(n^2)$. However, if we allow α to depend on n, then no upper bound is known and the best construction we have is just that of equiangular lines with size $\Omega(n^2)$. Since $[-1, 0)$-codes and $\{\alpha\}$-codes in $\mathbb{R}^n$ both have size at most $n + 1$, it is plausible to conjecture that $[-1, 0) \cup \{\alpha\}$-codes have size at most $O(n^2)$.

References

[1] Bukh, B. (2015). Bounds on equiangular lines and on related spherical codes. *SIAM J. Discrete Math., 30,* pp. 549–554.

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

[3] Coxeter, H. S. M. (1973). **Regular polytopes** (3rd ed.). Dover Publications.

[4] Delsarte, P., Goethals, J. M., & Seidel, J. J. (1975). Bounds for systems of lines, and Jacobi polynomials. *Philips Research Reports, 30,* 91–105.

[5] Delsarte, P., Goethals, J. M., & Seidel, J. J. (1977). Spherical codes and designs. *Geometriae Dedicata, 6,* 363–388.

[6] Godsil, C., & Royle, G. (2001). **Algebraic Graph Theory**. Springer-Verlag New York, Inc.

[7] Haantjes, J. (1948). Equilateral point-sets in elliptic two- and three-dimensional spaces. *Nieuw Arch. Wisk., 22,* 355–362.

[8] Heath, R. W., & Strohmer, T. (2003). Grassmannian frames with applications to coding and communication. *Applied and computational harmonic analysis, 14,* 257–275.

[9] Lemmens, P. W. H., & Seidel, J. J. (1973). Equiangular Lines. *Journal of Algebra, 24,* 494–512.

[10] van Lint, J. H., & Seidel, J. J. (1966). Equilateral point sets in elliptic geometry. *Indag. Math., 28,* 335–348.

[11] Neumaier, A. (1989). Graph Representations, Two-Distance Sets, and Equiangular Lines. *Linear Algebra and its Applications, 114,* 141–156.

On the number of points in general position in the plane

József Balogh
(joint work with J. Solymosi)

We say that a set of $n \geq d+1$ points on the Euclidean space $\mathbb{R}^d$ is in *general position* if no hyperplane contains $d+1$ points. One might expect that if in a set no hyperplane contains $d+2$ points then a large subset of the n points can be selected such that the points of this subset are in general position. Determining that in the worse case how many points in general position could be chosen - like many other similar problems - was initiated by Paul Erdős [6, 8]. Here, we consider only the planar ($d = 2$) case, the method can be easily extended for higher dimsenion. The best known upper and lower bounds are far from each other and even a sublinear upper bound was hard to achieve. Our primary goal is to prove the existence of a better construction for the upper bound.

Füredi [9] proved the lower bound $c\cdot\sqrt{n \log n}$ using a general result of Phelps and Rödl [15], who gave a lower bound on the independence number of partial Steiner Systems. Füredi [9] proved $o(n)$ upper bound using the Density Hales-Jewett Theorem (DHJ) of Katznelson and Furstenberg [10, 11]. In this application Füredi takes a random (generic) projection of the combinatorial cube of the combinatorial space $\{1, 2, 3\}^N$ to the plane. This point set has no colinear four-tuple, however for any $\delta > 0$, if N is large enough then any subset of points of size at least $\delta \cdot 3^N$ contains a colinear triple, a projection of a combinatorial line. The best known bound for DHJ follows from a recent proof by D.H.J. Polymath [16]. A subset of $\{1, 2, 3\}^N$ of density δ contains a combinatorial line if N is at least a tower of 2's of height $O(1/\delta^3)$. We improve the upper bound of Füredi [9].

Theorem 1. *As n goes to infinity, there exists a point set containing no four points in a line such that its every subset of size $n^{5/6+o(1)}$ contains three points on a line.*

Proof. (sketch) We start with $[m]^3$, the 3-dimensional integer grid with side-length m. Then we sparsen this point set, we keep each point with probability $p = \frac{1}{n \log^{1/2} n}$. From each of the colinear 4-tuples of the random point set we remove a point, yielding a point set without a colinear 4-tuple. An elementary argument gives that most of the points are not removed. With a random generic projection we map the points into the plan. With high probability we keep colinear point tuples colinear, and we will not create new colinear point tuples. The heart of the argument is to prove that these $n \approx m^3 p$ points satisfy the condition of the theorem.

We will use the method of hypergraph containers for the proof. This useful method was recently introduced independently by Balogh, Morris and Samotij [4], and by Saxton and Thomason [18]. Roughly speaking, it says that if a hypergraph $\mathcal{H}$ has a uniform edge distribution, then one can find a relatively small collection of sets, *containers*, covering all independent sets in $\mathcal{H}$. One can also require that the

container sets span few edges only. In our applications this later condition guarantees that all container sets are small. Our geometric construction determines a hypergraph where all large subsets contain an edge (e.g. a collinear triple). It is far from being clear, if 5/6 is the right constant in the exponent, however it could be that the point set that we constructed is close to be best possible, at least it should be close to be optimal, and only the limitation of our method is the reason that we cannot determine it.

We refer the readers to [2–4, 18] for more details and applications on the container method. The detailed proof of the theorem appears in [5]. Q.E.D.

The ε-nets are important concepts in computational geometry and approximation theory in computer science. On a base set A of n elements and a family of subsets $\mathcal{F} \subset 2^A$ a set $E \subset A$ is an ε-net if for every $B \in \mathcal{F}$ if $|B| \geq \varepsilon n$ then $E \cap B \neq \emptyset$. It is expected that if $\mathcal{F}$ has low complexity then one can find a small ε-net. In their seminal paper Haussler and Welzl [12] proved that for any range space with VC-dimension d, there is an ε-net of size $O((d/\varepsilon)\log(d/\varepsilon))$. Here we are concerned with a special case. Answering a question by Matoušek, Seidel, and Welzl [13], Alon [1] proved that the minimum possible size of an ε-net for point objects and line ranges in the plane is super-linear in $1/\varepsilon$. Alon used the density Hales-Jewett theorem for the construction, so his bound was just enough to break linearity. Alon conjectured that the upper bound is close to the truth. We cannot quite answer determine the correct function, however we are able to give a better lower bound, $(1/\varepsilon)\log^{1/3-o(1)}(1/\varepsilon)$, for the range space of lines. Our proof is somewhat similar in soul to the proof of Theorem 1, but different in many ways. We consider a different grid, and apply the container method somewhat different way.

Note that the container method was applied in a similar way in the graph setting [14] and [26], and in additive combinatorics [2], and there are several more recent papers, here we just listed the ones which were motivating our work. The merit of our work is that we have realized that this machinery is useful in discrete geometry. Given this method, still it is possible that better results could be obtained, in case someone was able to find more useful hypergraphs, or prove a variant of a container theorem taylored to these particular hypergraphs.

1. Acknowledgements

The authors are thankful to János Pach who suggested that our method might be applicable in the ε-nets problem. We also thank Noga Alon and Wojciech Samotij for fruitful discussions on the ε-nets problem.

References

[1] N. Alon, A non-linear lower bound for planar epsilon-nets, Proc. of the 51th IEEE FOCS (2010), 341–346. Also: *Discrete and Computational Geometry*, **47** (2012), 235–244.

[2] J. Balogh, H. Liu, and M. Sharifzadeh, The number of subsets of integers with no k-term arithmetic progression, to appear in *Int. Math. Res. Notices.*

[3] J. Balogh, H. Liu, M. Sharifzadeh and A. Treglown, Sharp bound on the number of maximal sum-free subsets of integers, submitted.
[4] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, **28** (2015), 669–709.
[5] J. Balogh and J. Solymosi, On the number of points in general position in the plane, to appear in *Discrete Analysis*.
[6] P. Erdős: On some metric and combinatorial geometric problems, *Discrete Mathematics*, **60** (1986), 147–153.
[7] P. Erdős: Problems and results on extremal problems in number theory, geometry, and combinatorics, Proceedings of the 7th Fischland Colloquium, I (Wustrow, 1988), Rostock. Math. Kolloq., No. 38 (1989), 6–14.
[8] P. Erdős: Some old and new problems in combinatorial geometry, Applications of discrete mathematics (Clemson, SC, 1986), 32–37, SIAM, Philadelphia, PA, 1988.
[9] Z. Füredi, Maximal independent subsets in Steiner systems and in planar sets, *SIAM J. Discrete Math.,* 4 (1991), 196–199.
[10] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem for $k = 3$, Graph Theory and Combinatorics (Cambridge, 1988), *Discrete Math.*, **75** (1989), 227–241.
[11] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. Anal. Math.*, **57** (1991), 64–119.
[12] D. Haussler and E. Welzl, ε-nets and simplex range queries, *Discrete and Computational Geometry*, (1987), 127–151.
[13] J. Matousek, R. Seidel and E. Welzl, How to net a lot with little: Small ε-nets for disks and halfspaces, In Proc. 6th Annu. ACM Sympos. Comput. Geom., 16–22, 1990.
[14] R. Nenadov and A. Steger, A short proof of the Random Ramsey theorem, *Combinatorics, Probability, and Computing*, **25** (2016), 130–144.
[15] K. T. Phelps and V. Rödl, Steiner triple systems with minimum independence number, *Ars Combin.*, **21** (1986), 167–172.
[16] D.H.J. Polymath, A new proof of the density Hales-Jewett theorem. *Ann. of Math.*, **175** (2012), no. 3, 1283–1327.
[17] V. Rödl, A. Ruciński and M. Schacht, An exponential-type upper bound for Folkman numbers, to appear in *Combinatorica*.
[18] D. Saxton and A. Thomason, Hypergraph containers, *Inventiones mathematicae*, (2015), 1–68.

Permanents and complex zeros

Alexander Barvinok

The permanent of an $n \times n$ complex matrix $A = (a_{ij})$ is defined as

$$\operatorname{per} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where S_n is the symmetric group of all permutations of the set $\{1, \ldots, n\}$.

Theorem 1. *If $A = (a_{ij})$ is an $n \times n$ complex matrix such that*

$$|1 - a_{ij}| \ \leq \ 0.5 \quad \textit{for all} \quad i, j$$

then $\operatorname{per} A \neq 0$.

Theorem 1 implies

Theorem 2. *If $A = (a_{ij})$ is an $n \times n$ complex matrix such that*

$$|1 - a_{ij}| \ \leq \ 0.49 \quad \textit{for all} \quad i, j$$

then, for any $0 < \epsilon < 1$, the value of $\ln \operatorname{per} A$ is approximated within an additive error ϵ in $n^{O(\ln n - \ln \epsilon)}$ time by a polynomial of degree $O(\ln n - \ln \epsilon)$ in the entries a_{ij} of A.

Theorem 3. *If $A = (a_{ij})$ is an $n \times n$ complex matrix such that*

$$\delta \ \leq \ \Re\, a_{ij} \ \leq \ 1 \quad \textit{and} \quad |\Im\, a_{ij}| \ \leq \ \frac{1}{2}\delta^2 \quad \textit{for all} \quad i, j$$

and some $0 < \delta \leq 1$ then $\operatorname{per} A \neq 0$.

Theorem 3 implies

Theorem 4. *If $A = (a_{ij})$ is an $n \times n$ real matrix such that*

$$\delta \ \leq \ a_{ij} \ \leq \ 1 \quad \textit{for all} \quad i, j$$

and some $0 < \delta \leq 1$ then, for any $0 < \epsilon < 1$, the value of $\ln \operatorname{per} A$ is approximated within an additive error ϵ in $n^{O_\delta(\ln n - \ln \epsilon)}$ time by a polynomial of degree $O_\delta(\ln n - \ln \epsilon)$ in the entries a_{ij} of A.

The implications Theorem 1 $\Longrightarrow$ Theorem 2 and Theorem 3 $\Longrightarrow$ Theorem 4 illustrate a general principle: if a combinatorially defined polynomial $p : \mathbb{C}^n \longrightarrow \mathbb{C}$ (such as the permanent or the hafnian of a matrix, the multi-dimensional permanent of a tensor, the independence polynomial of a graph, etc.) has no zeros in a domain $\Omega \subset \mathbb{C}^n$ then it can be efficiently approximated in a slightly smaller domain $\Omega' \subset \Omega$.

References

[1] A. Barvinok, *Approximating permanents and hafnians*, Discrete Analysis (2017), to appear.

One-sided epsilon-approximants

Boris Bukh
(joint work with G. Nivasch)

A common theme in mathematics is approximation of large, complicated objects by smaller, simpler objects. This paper proposes a new notion of approximation in combinatorial geometry, which we call one-sided ε-approximants. It is a notion of approximation that is in strength between ε-approximants and ε-nets. We recall these two notions first.

Let $P \subset \mathbb{R}^d$ be a finite set, and $\mathcal{F} \subset 2^{\mathbb{R}^d}$ a family of sets in $\mathbb{R}^d$. In applications, the family $\mathcal{F}$ is usually a geometrically natural family, such as the family of all halfspaces, the family of all simplices, or the family of all convex sets. A finite set $A \subset \mathbb{R}^d$ is called an *ε-approximant for P with respect to $\mathcal{F}$* if

$$\left| \frac{|C \cap P|}{|P|} - \frac{|C \cap A|}{|A|} \right| \leq \varepsilon \qquad \text{for all } C \in \mathcal{F}.$$

The notion of an ε-approximant was introduced by Vapnik and Chervonenkis [6] in the context of statistical learning theory. They associated to each family $\mathcal{F}$ a number VC-dim$(\mathcal{F}) \in \{1, 2, 3, \ldots, \infty\}$, which has become known as *VC dimension*, and proved that if VC-dim$(\mathcal{F}) < \infty$, then every set P admits an ε-approximant A of size $|A| \leq C_{\text{VC-dim}(\mathcal{F})}\varepsilon^{-2}$, a bound which does not depend on the size of P. The ε-approximants that they constructed had the additional property that $A \subset P$. Following tradition, we say that A is a *strong ε-approximant* if $A \subset P$. When we wish to emphasize that our ε-approximants are not necessarily subsets of P, we call them *weak ε-approximants*. The bound has been improved to $|A| \leq C_{\text{VC-dim}(\mathcal{F})}\varepsilon^{-2+2/(\text{VC-dim}(\mathcal{F})+1)}$ (see [4, Theorem 1.2] and [5, Exercise 5.2.7]) which is optimal [2].

In a geometric context, Haussler and Welzl [3] introduced ε-nets. With P and $\mathcal{F}$ as above, a set N is called an *ε-net for P with respect to $\mathcal{F}$* if

$$\frac{|C \cap P|}{|P|} > \varepsilon \implies C \cap N \neq \emptyset \qquad \text{for all } C \in \mathcal{F}.$$

An ε-approximant is an ε-net, but not conversely. While an ε-net is a weaker notion of approximation, its advantage over an ε-approximant is that every set P admits an ε-net of size only $C_{\text{VC-dim}(\mathcal{F})}\varepsilon^{-1}\log\varepsilon^{-1}$, which is smaller than the bound for the ε-approximants. The ε-nets constructed by Haussler and Welzl are also strong, i.e., they satisfy $N \subset P$.

Most geometrically important families $\mathcal{F}$ have a bounded VC dimension. A notable exception is the family $\mathcal{F}_{\text{conv}}$ of all convex sets. Indeed, it is easy to see that a set of n points in convex position does not admit any strong ε-net of size smaller than $(1-\varepsilon)n$ with respect to $\mathcal{F}_{\text{conv}}$. Alon, Bárány Füredi, and Kleitman [1] showed that for every $P \subset \mathbb{R}^d$ there exists a (weak) ε-net of size bounded solely by a function of ε and d. No extension of their result to ε-approximants is possible.

Proposition 1. *If $P \subset \mathbb{R}^2$ is a set of n points in convex position, then every ε-approximant with respect to $\mathcal{F}_{\text{conv}}$ has size at least $n(\frac{1}{4} - \varepsilon/2)$.*

Proof. Let $p_1, p_2, \ldots, p_n$ be the enumeration of the vertices of P in clockwise order along the convex hull of P. For $i = 1, \ldots, \lfloor (n-1)/2 \rfloor$ write T_i for the triangle $p_{2i-1}, p_{2i}, p_{2i+1}$. Suppose $A \subset \mathbb{R}^2$ is an ε-approximant for P. Let $I \stackrel{\text{def}}{=} \{i : T_i \cap A = \emptyset\}$. Note that $|I| \geq n/2 - 2|A| - 1$ since each point of A lies in at most two triangles. Define $S \stackrel{\text{def}}{=} \{p_1, p_3, p_5, \ldots\}$ to be the odd-numbered points, and let $S' \stackrel{\text{def}}{=} S \cup \{p_{2i} : i \in I\}$. Let $C \stackrel{\text{def}}{=} \operatorname{conv} S$ and $C' \stackrel{\text{def}}{=} \operatorname{conv} S'$. Then $C \cap A = C' \cap A$, but $|C' \cap P|/|P| - |C \cap P|/|P| = |I|/|P| > \varepsilon$ if $|A| < |P|(\frac{1}{4} - \varepsilon/2)$. □

In light of Proposition 1, we introduce a new concept. A multiset $A \subset \mathbb{R}^d$ is a *one-sided ε-approximant for P with respect to the family $\mathcal{F}$* if

$$\frac{|C \cap P|}{|P|} - \frac{|C \cap A|}{|A|} \leq \varepsilon \qquad \text{for all } C \in \mathcal{F}.$$

In other words, if $C \in \mathcal{F}$, then C might contain many more points of A than expected, but never much fewer. It is clear that an ε-approximant is a one-sided ε-approximant, and that a one-sided ε-approximant is an ε-net.

Our main result shows that allowing one-sided errors is enough to sidestep the pessimistic Proposition 1.

Theorem 2. *Let $P \subset \mathbb{R}^d$ be a finite set, and let $\varepsilon \in (0,1]$ be a real number. Then P admits a one-sided ε-approximant with respect to $\mathcal{F}_{\text{conv}}$ of size at most $g(\varepsilon, d)$, for some g that depends only on ε and on d.*

Unfortunately, due to the use of a geometric Ramsey theorem, our bound on g is very weak:

$$g(\varepsilon, d) \leq \text{tw}_d\big(\varepsilon^{-c}\big)$$

for some constant $c > 1$ that depends only on d, where the tower function is given by $\text{tw}_1(x) \stackrel{\text{def}}{=} x$ and $\text{tw}_{i+1}(x) \stackrel{\text{def}}{=} 2^{\text{tw}_i(x)}$. We believe this bound to be very far from sharp.

References

[1] Noga Alon, Imre Bárány, Zoltán Füredi, and Daniel J. Kleitman. Point selections and weak ϵ-nets for convex hulls. *Combin. Probab. Comput.*, 1(3):189–200, 1992. `http://www.tau.ac.il/~nogaa/PDFS/abfk3.pdf`.

[2] R. Alexander. Geometric methods in the study of irregularities of distribution. *Combinatorica*, 10(2):115–136, 1990.

[3] David Haussler and Emo Welzl. ε-nets and simplex range queries. *Discrete Comput. Geom.*, 2(2):127–151, 1987.

[4] Jiří Matoušek. Tight upper bounds for the discrepancy of half-spaces. *Discrete Comput. Geom.*, 13(3-4):593–601, 1995.

[5] Jiří Matoušek. *Geometric discrepancy: an illustrated guide*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1999.

[6] Vladimir N. Vapnik and Alexey Ja. Červonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Teor. Verojatnost. i Primenen.*, 16:264–279, 1971. `http://mi.mathnet.ru/tvp2146`.

Induced subgraphs and chromatic number

Maria Chudnovsky

(joint work with A. Scott, P. Seymour, and S. Spirkl)

Let G be a graph. The *clique number* of G is the size of the largest set of pairwise adjacent vertices, and the *chromatic number* is the smallest number k such that the vertices of G can be colored with k colors and no two adjacent vertices receive the same color. The clique number of G is denoted by $\omega(G)$, and the chromatic number by $\chi(G)$. Clearly $\chi(G) \geq \omega(G)$, and an interesting question is what can be said about bounding the chromatic number from above by a function of the clique number.

It turns out that in general no such bound exists, because of the following result of [3].

Theorem 1. *For every $g, k > 0$ there exists a graph with chromatic number at least k and with no cycle of length less than g.*

Let $\mathcal{C}$ be a class of graphs closed under taking induced subgraphs. We say that $\mathcal{C}$ is *χ-bounded* if there exists a function $f_{\mathcal{C}} : N \to N$ such that for every graph $G \in \mathcal{C}$ $\chi(G) \leq f_{\mathcal{C}}(\omega(G))$. We call such $f_{\mathcal{C}}$ a *χ-bounding function* for $\mathcal{C}$. For a graph H, G is said to be *H-free* if G has no induced subgraph isomorphic to H. What can be said in this context about H-free graphs? Theorem 1 implies:

Theorem 2. *If the class of H-free graph is χ-bounded, then H is a forest.*

The converse of Theorem 2 is a well-known conjecture [4, 7]:

Conjecture 3. *For every forest H there exists a function f_H such that $\chi(G) \leq f_H(G)$ for every H-free graph G.*

Very little is known to-date about Conjecture 3, but the case where H is a path was proved in [4]. To solve this special case a technique was developed, to which we now refer as "building a Gyárfás path". The idea is to construct an induced path in the graph, that starts at an arbitrary vertex, chips off pieces of the graph, each of which has small chromatic number, and advances in the direction of the "unexplored" part of the graph (that still has large chromatic number).

Here is another angle on χ-boundedness: what if induced cycles of certain lengths are forbidden? This is promising because of the next result, known as The Strong Perfect Graph Theorem [1]. For a graph G, the complement G^c of G is the graph with vertex set $V(G)$, and such that $uv \in E(G^c)$ if and only if $uv \notin E(G)$. The *parity* of a cycle C is the parity of $|V(C)|$; thus a cycle can be *odd* or *even*.

Theorem 4. *Let $\mathcal{C}$ be the class of graphs G with no induced cycles of odd length at least five in G, and none in G^c. Then the function $f(x) = x$ is a χ-bounding function for $\mathcal{C}$.*

In [5] three conjecture were made, in the spirit of Theorem 4, but without mentioning complements:

Conjecture 5. *The class of graphs with no odd induced cycles of length at least five is χ-bounded.*

Conjecture 6. *For every l, the class of graphs with no induced cycles of length at least l is χ-bounded.*

Conjecture 7. *For every l, the class of graphs with no odd induced cycles of length at least l is χ-bounded.*

Conjecture 5 was proved in [10], and Conjecture 6 in [2]. Recently we were able to prove Conjecture 7. The proof has two main parts. In the first part we assume that for every vertex v of G, the subgraph of G induced by the vertices at distance at most two from v has bounded chromatic number. In this case we construct a Gyárfás path P of length l, that can be completed to induced cycles of two different parities (constructing the completions in the new insight in our

proof, that was missing until now). Having dealt with this case, we can assume that for every graph with no long odd induced cycle, and with small ω and large χ, there is a vertex v whose second neighborhood (the vertices at distance exactly two from v) induces a graph with large chromatic number. Now we follow ideas from [2], repeatedly extracting such vertices v. This process allows us to either build an induced subdivision of a large complete bipartite graph in G (which we use to construct a long induced path that can then be completed to induced cycles of different parities), or build a vast variety of induced subgraphs in G, including cycles of any length.

References

[1] M. Chudnovsky, N.Robertson, P.Seymour, and R.Thomas, *The strong perfect graph theorem*, Annals of Math **164** (2006), 51–229.

[2] M. Chudnovsky, A. Scott and P. Seymour, *Induced subgraphs of graphs with large chromatic number III. Long holes*, Combinatorica, to appear

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

[4] A. Gyárfás, *On Ramsey covering-numbers*, *Coll. Math. Soc. János Bolyai*, in *Infinite and Finite Sets*, North Holland/American Elsevier, New York (1975), 10.

[5] A. Gyárfás, *Problems from the world surrounding perfect graphs*, Proceedings of the International Conference on Combinatorial Analysis and its Applications, (Pokrzywna, 1985), *Zastos. Mat.* **19** (1987), 413–441.

[6] A. Scott and P. Seymour, *Induced subgraphs of graphs with large chromatic number I. Odd holes*, J. Combinatorial Theory, Ser. B, **121** (2016), 68–84.

[7] D.P. Sumner, *Subtrees of a graph and chromatic number* in The Theory and Applications of Graphs, (G. Chartrand, ed.), John Wiley & Sons, New York (1981), 557–576.

Universality-type results in random graphs

David Conlon

(joint work with A. Ferber, R. Nenadov, and N. Škorić)

Given a family of graphs $\mathcal{H}$, a graph G is said to be $\mathcal{H}$-universal if it contains every graph in $\mathcal{H}$ as a subgraph. For $\mathcal{H} = \mathcal{H}(n, \Delta)$, the class of graphs with at most n vertices and maximum degree at most Δ, a result of Alon and Capalbo [1] shows that there are universal graphs with $O(n^{2-2/\Delta})$, a result which is sharp up to the implied constant. In this talk, we explore the extent to which the binomial random graph $G(n,p)$ may be used as a universal graph for $\mathcal{H}(n,\Delta)$.

In particular, we discuss the proof of the following theorem, due to the speaker, Ferber, Nenadov and Škorić [3].

Theorem 1. *For any constant $\epsilon > 0$ and integer $\Delta \geq 3$, the random graph $G(n,p)$ is a.a.s. universal for the family $\mathcal{H}((1-\epsilon)n, \Delta)$, provided that $p \gg n^{-\frac{1}{\Delta-1}} \log^5 n$.*

This improves a result of Alon, Capalbo, Kohayakawa, Rödl, Ruciński and Szemerédi [2], who proved the analogous statement for $p \geq c_{\Delta,\epsilon}(\log n/n)^{1/\Delta}$. This latter probability serves as a natural boundary for problems involving the embedding of large graphs with maximum degree Δ, since it is the lowest probability at which one can expect that every collection of Δ distinct vertices will have many

common neighbours, a property that facilitates a greedy embedding strategy. Going below this probability introduces significant difficulties.

We also discuss a resilient version of this theorem, due to the speaker and Nenadov [4]. We say that a graph G is $\mathcal{H}$-Ramsey-universal if every two-colouring of the edges of G contains a colour class which is $\mathcal{H}$-universal. For this concept, our main result, concerns $\mathcal{F}(n,\Delta)$, the class of triangle-free graphs with at most n vertices and maximum degree at most Δ.

Theorem 2. *For any integer $\Delta \geq 5$, there exist constants $B \in \mathbb{N}$ and $C > 0$ such that if $N = Bn$ and $p = p(N) \geq (C \log N/N)^{\frac{1}{\Delta-1/2}}$ then $G(N,p)$ is a.a.s. $\mathcal{F}(n,\Delta)$-Ramsey-universal.*

This improves a result of Kohayakawa, Rödl, Schacht and Szemerédi [5], who again proved an analogous result for $p \geq c_\Delta(\log n/n)^{1/\Delta}$. However, the cost of bypassing their bound is that we have to restrict to $\mathcal{F}(n,\Delta)$, the class of triangle-free graphs in $\mathcal{H}(n,\Delta)$, a restriction not required in [5]. It would be very interesting to remove this caveat from our work.

References

[1] N. Alon and M. Capalbo, *Optimal universal graphs with deterministic embedding*, in Proc. of the Nineteenth Annual ACM-SIAM SODA (2008), 373–378.

[2] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński and E. Szemerédi, *Universality and Tolerance*, in Proc. 41 IEEE FOCS, IEEE (2000), 14–21.

[3] D. Conlon, A. Ferber, R. Nenadov and N. Škorić, *Almost-spanning universality in random graphs*, to appear in Random Structures Algorithms.

[4] D. Conlon and R. Nenadov, *Size Ramsey numbers of triangle-free graphs with bounded degree,* submitted.

[5] Y. Kohayakawa, V. Rödl, M. Schacht and E. Szemerédi, *Sparse partition universal graphs for graphs of bounded degree*, Adv. Math. **226**, 2011, 5041–5065.

On a resilience Littlewood-Offord problem

Asaf Ferber

(joint work with A. Bandeira and M. Kwan)

Let $a = (a_i)_{i=1}^n$ be a fixed sequence of nonzero real numbers, and for a sequence of i.i.d. (independent, identically distributed) Rademacher random variables $\xi = (\xi_i)_{i=1}^n$ (meaning $\Pr[\xi_i = 1] = \Pr[\xi_i = -1] = 1/2$), define the random sum

$$X = X_a(\xi) = \sum_{i=1}^n a_i \xi_i.$$

Sums of this form are ubiquitous in probability theory. For example, X can be interpreted as the outcome of an unbiased random walk with step sizes given by a. The central limit theorem asserts that if the a_i are all equal then X asymptotically has a normal distribution. More flexible variants of the central limit theorem allow the a_i to differ to an extent, and give quantitative control of the distribution of X. An important example is the Berry-Esseen theorem [1, 4], which gives an

estimate for the probability that X lies in a given interval, comparing it to the corresponding probability for an appropriately scaled normal distribution (we give a precise statement, adapted to our context, later in the paper). The Berry-Esseen theorem is effective when the a_i are of the same order of magnitude, in which case it can be used to easily deduce the estimate

$$\Pr[X = x] = O\left(\frac{1}{\sqrt{n}}\right)$$

for any x. Qualitatively, it guarantees that X is unlikely to be concentrated on any particular value (X is *anti-concentrated*).

Over half a century ago, in connection with their study of random polynomials, Littlewood and Offord [6] considered anti-concentration in the general setting where no assumption is made on a, other than that its entries being nonzero. The classical result of Littlewood and Offord [6] strengthened by Erdős [2] states that for any a, for all $x \in \mathbb{R}$ we have

$$\Pr[X = x] \leq \binom{n}{\lfloor n/2 \rfloor} \Big/ 2^n = O\left(\frac{1}{\sqrt{n}}\right),$$

which is sharp for the sequence $a = (1, 1, \ldots, 1)$. This result is particularly remarkable due to the fact that if one does not assume anything about a_i, then the distribution of X may be far from normal and Berry-Esseen type bounds may no longer be meaningful.

Since the Littlewood-Offord problem was first introduced, many variants of it have been addressed; one particularly interesting line of research involves the relationship between the structure of a and the resulting concentration probability $\max_x \Pr[X = x]$.

Erdős and Moser [3] and Sárközy and Szemerédi [10] considered the case where the a_i are all distinct, and showed that the stronger bound $\Pr[X = x] = O\left(n^{-3/2}\right)$ holds. Halász [5] gave even stronger bounds for sequences which are "arithmetically unstructured" in an appropriate sense. More recently, Tao and Vu [12, 13] and Nguyen and Vu [8] investigated the inverse problem of characterizing the arithmetic structure of a given the concentration probability $\max_x \Pr[X = x]$.

Many fruitful connections have been found between Littlewood-Offord-type problems and various areas of mathematics. In particular, Littlewood-Offord-type theorems are essential tools in some of the landmark results in random matrix theory (see for example [11, 12]). In particular, the Littlewood-Offord theorem gives an upper bound on the probability that a particular row of a random ± 1 matrix is orthogonal to a given vector, and can thus be used to bound the probability that a Bernoulli random matrix is singular.

In this talk we are interested in studying a "resilience" version of the Littlewood-Offord problem. Given a sequence $a \in (\mathbb{R} \setminus \{0\})^n$ and a real number $x \in \mathbb{R}$, we know that most sequences $\xi \in \{-1, 1\}^n$ do not satisfy the event $\{X(\xi) = x\}$. We are interested in understanding whether most sequences ξ are "far" from this event. In order to make this question precise we need a few definitions. Given two sequences $\xi, \xi' \in \{-1, 1\}^n$ we define $d(\xi, \xi')$ to be the *Hamming distance* between

ξ and ξ' (that is, $d(\xi,\xi')$ denotes the number of coordinates in which ξ and ξ' differ). If $S \subset \{-1,1\}^n$ is a subset of the hypercube we further define $d(\xi,S)$ as the minimum Hamming distance from ξ to a point in S. Finally, for a fixed sequence a of non-zero reals and $\xi \in \{-1,1\}^n$, let us define

$$R_x(\xi) := R_x^a(\xi) = d\left(\xi, X^{-1}(x)\right),$$

which is the minimum number of signs one needs to change in ξ in order to satisfy $X = x$. We refer to $R_x(\xi)$ as the *resilience* of ξ with respect to the event $\{X \neq x\}$, and if $R_x > k$ we say ξ is *k-resilient* (for completeness, if one cannot obtain x at all then we set $R_x(\xi) = \infty$).

We prove an optimal result (both lower and upper bounds), showing that whp $R_x(\xi) = (1+o(1))\log_3 \log n$.

References

[1] A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Transactions of the American Mathematical Society **49** (1941), no. 1, 122–136.

[2] P. Erdős, *On a lemma of Littlewood and Offord*, Bulletin of the American Mathematical Society **51** (1945), no. 12, 898–902.

[3] P. Erdős, *Extremal problems in number theory*, Proc. Sympos. Pure Math., Vol. VIII, AMS, Providence, R.I (1965), 181–189.

[4] C.-G. Esseen, *On the Liapounoff limit of error in the theory of probability*, Arkiv för Matematik, Astronomi och Fysik **A28** (1942), no. 9, 1–19.

[5] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Periodica Mathematica Hungarica **8** (1977), 197–211.

[6] J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation. III*, Math. Mat. Sbornik N.S. **12** (1943), no. 3, 277–286.

[7] R. Meka, O. Nguyen, and V. Vu, *Anti-concentration for polynomials of Rademacher random variables and applications in complexity theory*, arXiv preprint arXiv:1507.00829 (2015).

[8] H. Nguyen and V. Vu, *Optimal inverse Littlewood-Offord theorems*, Advances in Mathematics **226** (2011), no. 6, 5298–5319.

[9] H. Nguyen and V. Vu, Small probability, *inverse theorems, and applications*, Erdős Centennial, Bolyai Society Mathematical Studies, Vol. 25 (2013), 409–463.

[10] A. Sárközy and E. Szeméredi, *Über ein Problem von Erdős und Moser*, Acta Arithmetica **11** (1965), 205–208.

[11] T. Tao and V. Vu, *From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices*, Bulletin of the American Mathematical Society **46** (2009), no. 3, 377–396.

[12] T. Tao and V. H. Vu, *Inverse Littlewood-Offord theorems and the condition number of random discrete matrices*, Annals of Mathematics (2009), 595–632.

[13] T. Tao and V. H. Vu, *A sharp inverse Littlewood-Offord theorem*, Random Structures & Algorithms **37** (2010), no. 4, 525–539.

Arithmetic progressions, regularity, and removal

Jacob Fox
(joint work with L. M. Lovasz, H. T. Phan, and Y. Zhao)

Szemerédi's regularity lemma [28] and its variants are some of the most powerful tools in combinatorics. Szemerédi used an early version in the proof of his celebrated theorem extending Roth's theorem to arithmetic progressions of any given length. The regularity lemma roughly says that the vertex set of every graph can be partitioned into a small number of parts such for most of the pairs of parts, the induced bipartite subgraph between the pair is pseudorandom.

Green [16] developed an arithmetic regularity lemma and used it to prove two extensions of Roth's theorem, known as the arithmetic removal lemma and the popular differences theorem. A *triangle* in $\mathbb{F}_p^n$ is a triple a, b, c of elements with $a+b+c = 0$. The first theorem is known as the arithmetic triangle removal lemma, as it is an arithmetic analogue of the triangle removal lemma.

Theorem 1 (Green [16]). *For each $\epsilon > 0$ and prime p there is $\delta = \delta(\epsilon, p) > 0$ such that if $A, B, C \subset \mathbb{F}_p^n$ with at most δp^{2n} triangles in $A \times B \times C$, then we can remove at most ϵp^n elements from A, B, C and remove all triangles.*

A random subset A of $\mathbb{F}_p^n$ of density α almost surely satisfies that for every nonzero $d \in \mathbb{F}_p^n$, the density of three-term arithmetic progressions with common difference d that are in A is close to α^3. For $p = 3$, Edel's construction [8] shows that there are sets with density α whose density of three-term arithmetic progressions is $O(\alpha^{4.63})$, which is much smaller than the random bound, α^3. However, the following theorem of Green [16], answering a question of Bergelson, Host, and Kra [1], shows that there is a nonzero d for which the density of three-term arithmetic progressions with common difference d is at least almost the random bound, α^3.

Theorem 2 (Green [16]). *For each $\epsilon > 0$ and prime p there is a least positive integer $n_p(\epsilon)$ such that the following holds. For each $n \geq n_p(\epsilon)$ and every subset A of $\mathbb{F}_p^n$ with density α, there is a nonzero d in $\mathbb{F}_p^n$ such that the density of three-term arithmetic progressions with common difference d that are in A is at least $\alpha^3 - \epsilon$.*

A major drawback in applying the graph regularity lemma of Szemerédi or the arithmetic regularity lemma of Green is that the bound on the number of parts is enormous as a function of the approximation parameter ϵ (the bound is a tower of twos of height $\epsilon^{-O(1)}$). This gives seemingly weak bounds for the various applications. That such a huge bound is necessary in Szemerédi's regularity lemma was proved by Gowers [17] using a probabilistic construction. Later, Conlon and Fox [4], Moshkovitz and Shapira [25], and Fox and Lovász [11] gave improvements on various aspects. Similarly, Green [16] showed that a tower-type bound is neccesary in the arithmetic regularity lemma, and Hosseini et al. [20] improved the tower lower bound to height $\epsilon^{-\Omega(1)}$.

While a tower-type bound is known to be necessary for these regularity lemmas, for the various applications, it was generally believed that much better bounds

should hold, and this would be shown by developing alternative techniques to the regularity method. This turned out to be the case for many such applications. A few examples include Gowers' influential proof of Szemerédi's theorem [18] which introduced higher Fourier analysis, new bounds on Ramsey numbers of sparse graphs (see [5] and [23] and their references) using the greedy embedding method or dependent random choice, certain applications in extremal graph theory using the absorption method [26], [24], and the first author's proof of the graph removal lemma [10]. Prior to this work, no known application required the tower-type bound coming from applying a regularity lemma.

Green's proof of the arithmetic removal lemma, Theorem 1, used the arithmetic regularity lemma, giving a bound on $1/\delta$ which is a tower of twos of height polynomial in $1/\epsilon$. Král', Serra, and Vena [22] showed that the arithmetic triangle removal lemma follows from the triangle removal lemma for graphs. Green's proof shows the result further holds in any abelian group, and the Král'-Serra-Vena proof shows that an analogue holds in any group.

Green [16] posed the problem of improving the quantitative bounds on the arithmetic triangle removal lemma, and, in particular, asked whether a polynomial bound holds. Green's problem has received considerable attention by many researchers and was featured in more than a dozen papers. This is in part due to its applications and connections to several major problems in number theory, combinatorics, and computer science. The previous best known bound for the arithmetic triangle removal lemma came from the improved bound on the triangle removal lemma for graphs [10] together with the Král'-Serra-Vena reduction. It gives a bound on $1/\delta$ in the arithmetic triangle removal lemma which is a tower of twos of height logarithmic in $1/\epsilon$, see [10] for details. An alternative, Fourier-analytic proof of this bound in the case when $p = 2$ was given by Hatami, Sachdeva, and Tulsiani [19]. Prior to this work, this tower-type bound was the best known bound for the arithmetic triangle removal lemma.

In the case $p = 2$, obtaining better upper bounds for δ has also received much attention due in part to its close connection to property testing, and, in particular, testing triangle-freeness. Bhattacharyya and Xie [2] were the first to give a non-trivial upper bound on δ. Fu and Kleinberg [15] provided a simple construction showing that $\delta \leq \epsilon^{C_2 - o(1)}$, where $C_2 \approx 13.239$ is defined below. It is based on a construction from Coppersmith and Winograd's famous matrix multiplication algorithm [6]. It was widely conjectured that the exponent in the bound is not optimal (and that the bound is perhaps even superpolynomial).

We solve Green's problem by proving an essentially tight bound for the arithmetic triangle removal lemma in vector spaces over finite fields. We show that a polynomial bound holds, and further determine the best possible exponent for every prime p. It involves, for every p, constants c_p which will be between 0 and 1 and will be defined below, and $C_p = 1 + 1/c_p$. In particular, we have[1] $c_2 = 5/3 - \log 3 \approx .0817$ and $C_2 = 1 + 1/c_2 \approx 13.239$ as in Fu and Kleinberg [15],

[1]Here, and throughout, all logarithms are base 2.

and $c_3 = 1 - \frac{\log b}{\log 3}$, where $b = a^{-2/3} + a^{1/3} + a^{4/3}$, and $a = \frac{\sqrt{33}-1}{8}$. This gives $c_3 \approx .0775$, and $C_3 \approx 13.901$.

Theorem 3. *[Fox and Lovász [12]] Let $0 < \epsilon \leq 1/2$ and $\delta = (\epsilon/3)^{C_p}$. If $A, B, C \subset \mathbb{F}_p^n$ with less than δN^2 triangles in $A \times B \times C$, then we can remove ϵN elements from $A \cup B \cup C$ so that no triangle remains. Furthermore, this bound is tight in that we must have $\delta \leq \epsilon^{C_p - o(1)}$.*

Theorem 3 appears in [12] and uses recent breakthroughs on the cap set problem by Croot, Lev, and Pach [7], Ellenberg and Gijswijt [9] and its generalization to the multicolored sum-free problem by Blasiak et al. [3], Alon, and Kleinberg, Sawin, and Speyer [21] together with probabilitistic and combinatorial arguments.

With Pham [13], we recently determined tight tower-type bounds for Green's theorem on popular progression differences, Theorem 2, which is tight up to an absolute constant factor for p fixed. We prove lower and upper bounds on the necessary dimension which grow as an exponential tower of twos of height $\Theta(\log(1/\epsilon))$. Note that this is the first application of a regularity lemma for which the tower-type bound that come from using the regularity lemma is necessary.

Theorem 4 (Fox-Pham [13])**.** *Recall $n_p(\epsilon)$ is the least positive integer such that the following holds. For each $n \geq n_p(\epsilon)$ and every subset A of $\mathbb{F}_p^n$ with density α, there is a nonzero $d \in \mathbb{F}_p^n$ such that the density of three-term arithmetic progressions with common difference d that are in A is at least $\alpha^3 - \epsilon$. For $\epsilon < 2^{-500} p^{-8}$, we have $n_p(\epsilon)$ grows as an exponential tower of p's of height $\Theta(\log(1/\epsilon))$.*

We further determine for $p \geq 19$ and $\beta < \alpha^3$ the correct tower height, up to an absolute constant factor and an additive constant depending on p, of the minimum dimension n needed to guarantee that for any subset of $\mathbb{F}_p^n$ of density α there is a nonzero d such that the density of three-term arithmetic progressions with common difference d is at least β. In particular, to obtain half the random bound (when $\beta = \alpha^3/2$), we get the dimension we need to guarantee such a d grows as a tower of twos of height proportional to $(\log p) \log\log(1/\alpha)$.

In joint work with Zhao [14], we extend Theorem 4 to general abelian groups. This required new ideas to overcome challenges related to the lack of subgroups in the general setting. For the upper bound, we extend our arguments to work for Fourier analysis on Bohr sets. For the lower bound, new construction ideas.

References

[1] V. Bergelson, B. Host and B. Kra, Multiple recurrence and nilsequences, with an appendix by Imre Ruzsa, *Invent. Math.* **160** (2005), 261–303.

[2] A. Bhattacharyya and N. Xie, Lower bounds for testing triangle-freeness in Boolean functions, *Comput. Complexity* **24** (2015), 65–101.

[3] J. Blasiak, T. Church, H. Cohn, J. Grochow, E. Naslund, W. Sawin, and C. Umans, On cap sets and the group-theoretic approach to matrix multiplication, *Discrete Analysis* 2017:3, 27pp.

[4] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas, *Geom. Funct. Anal.* **22** (2012), 1191–256.

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

[6] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progressions, *J. Symbolic Comput.* **9** (1990), 251–280.

[7] E. Croot, V. F. Lev, and P. P. Pach, Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small, *Annals of Math.* **185** (2017), 331–337.

[8] Y. Edel, Extensions of generalized product caps, *Designs, Codes and Cryptography* **31** (2004), 5–14.

[9] J. S. Ellenberg andD. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Annals of Math.* **185** (2017), 339–343.

[10] J. Fox, A new proof of the graph removal lemma, *Annals of Math.* **174** (2011), 561–579.

[11] J. Fox and L. M. Lovász, A tight bound for Szemerédi's regularity lemma, *Combinatorica*, to appear.

[12] J. Fox and L. M. Lovász, A tight bound for Green's arithmetic triangle removal lemma in vector spaces, arXiv:1606.01230.

[13] J. Fox and H. T. Pham, Popular progession differences in vector spaces, in preparation.

[14] J. Fox, H. T. Pham, and Y. Zhao, Popular progession differences in abelian groups, in preparation.

[15] H. Fu and R. Kleinberg, Improved lower bounds for testing triangle-freeness in Boolean functions via fast matrix multiplication. Approximation, randomization, and combinatorial optimization, 669–676, LIPIcs. Leibniz Int. Proc. Inform., 28, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014.

[16] B. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Anal.* **15** (2005), 340–376.

[17] W. T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, *Geom. Funct. Anal.* **7** (1997), 322–337.

[18] W. T. Gowers, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.* **11** (2001), 465–588.

[19] P. Hatami, S. Sachdeva, and M. Tulsiani, An arithmetic analogue of Fox's triangle removal argument, *Online J. Anal. Comb.* **11** (2016), Art. 1, 17 pp.

[20] K. Hosseini, S. Lovett, G. Moshkovitz and A. Shapira, An improved lower bound for arithmetic regularity, *Math. Proc. Camb. Phil. Soc.* **161** (2016), 193–197.

[21] B. Kleinberg, W. F. Sawin, D. E. Speyer, The growth rate of tri-colored sum-free sets, arXiv:1607.00047.

[22] D. Král', O. Serra, and L. Vena, A combinatorial proof of the removal lemma for groups, *J. Combin. Theory Ser. A* **116** (2009), 971–978.

[23] C. Lee, Ramsey numbers of degenerate graphs, *Annals of Math.*, to appear. arXiv:1505.04773.

[24] I. Levitt, G. Sarkozy, and E. Szemerédi, How to avoid using the Regularity Lemma: Pósa's conjecture revisited, *Discrete Math.* **310** (2010), 630–641.

[25] G. Moshkovitz and A. Shapira, A short proof of Gowers' lower bound for the regularity lemma, *Combinatorica* **36** (2016), 187–194.

[26] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs, *Combin. Probab. Comput.* **15** (2006), 229–251.

[27] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[28] E. Szemerédi, Regular partitions of graphs, Problémes combinatoires et th eorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.

Kneser ranks of random graphs and minimum difference representations

Zoltán Füredi
(joint work with I. Kantor)

1. Kneser representations

A representation of a graph G is an assignment of mathematical objects of a given kind (intervals, disks in the plane, finite sets, vectors, etc.) to the vertices of G in such a way that two vertices are adjacent if and only if the corresponding sets satisfy a certain condition (intervals intersect, vectors have different entries in each coordinate, etc.). Representations of various kinds have been studied extensively, see, e.g., [5], the monograph [11], or from information theory point of view [9].

The *Kneser graph* $\mathrm{Kn}(s,k)$ (for positive integers $s \geq 2k$) is a graph whose vertices are all the k-subsets of the set $[s] := \{1,2,\dots,s\}$, and whose edges connect two sets if they are disjoint. An assignment $v \mapsto A_v$ to the vertices $v \in V$ of a graph $G = (V,E)$ is called a *Kneser representation of rank k* if each A_v has size k, the sets are distinct, and A_u and A_v are disjoint if and only if $uv \in E$.

Every graph G on n vertices with minimum degree δ has a Kneser representation of rank $n-\delta$. For every $i \in V(G)$, let A'_i be the set of the edges adjacent to i in the complement of G. We have $A'_i \cap A'_j = 1$ if $ij \notin E(G)$, otherwise $A'_i \cap A'_j = 0$. To turn this *co-star representation* into a Kneser representation add pairwise disjoint sets of labels to the sets $A'_1,\dots,A'_n$ to increase their cardinality to exactly $n-\delta(G)$.

Let $\mathcal{G}(n)$ denote the set of $2^{\binom{n}{2}}$ (labelled) graphs on $[n]$ and let $\mathcal{G}(n,k,\mathrm{Kneser})$ denote the family of graphs on $[n]$ having a Kneser representation of rank k. $G \in \mathcal{G}(n,k,\mathrm{Kneser})$ is equivalent to the fact that G is an *induced* subgraph of some Kneser graph $\mathrm{Kn}(s,k)$. We have

$$\mathcal{G}(n,1,\mathrm{Kneser}) \subseteq \mathcal{G}(n,2,\mathrm{Kneser}) \subseteq \cdots \subseteq \mathcal{G}(n,n-1,\mathrm{Kneser}) = \mathcal{G}(n).$$

Let $f_{\mathrm{Kneser}}(G)$ denote the smallest k such that G has a Kneser representation of rank k. We have seen that $f_{\mathrm{Kneser}}(G) \leq n-\delta$. We have

Theorem 1. *There exist constants $c_2 > c_1 > 0$ such that for $G \in \mathcal{G}(n)$ with high probability*

$$c_1 \frac{n}{\log n} < f_{\mathrm{Kneser}}(G) < c_2 \frac{n}{\log n}.$$

1.1. **Thickness of clique covers.** The *clique cover number* $\theta_1(G)$ of a graph G is the minimum number of cliques required to cover the edges of graph G. Recall that in a random graph $G \in \mathcal{G}(n,p)$, each of the $\binom{n}{2}$ edges occurs independently with probability p. Similarly, $\mathcal{G}(n,n,p)$ denotes the class of graphs $\mathcal{G}(n,n)$ with the probability of a given graph $G \in \mathcal{G}(n,n)$ is $p^{e(G)}(1-p)^{n^2-e(G)}$. Frieze and Reed [7] proved that for p constant, $0 < p < 1$, there exist constants $c'_i = c'_i(p) > 0$,

$i = 1, 2$ such that for $G \in \mathcal{G}(n,p)$ with high probability

$$c_1' \frac{n^2}{(\log n)^2} < \theta_1(G) < c_2' \frac{n^2}{(\log n)^2}.$$

The *thickness* θ_0 of a clique cover $\mathcal{C} := \{C_1, \ldots, C_m\}$ of G is the maximum degree of the hypergraph $\mathcal{C}$, i.e., $\theta_0(\mathcal{C}) := \max_{v \in V(G)} \deg_{\mathcal{C}}(v)$. The minimum thickness among the clique covers of G is denoted by $\theta_0(G)$. A clique cover $\mathcal{C}$ corresponds to a set representation $v \mapsto A_v$ in a natural way $A_v := \{C_i : v \in C_i\}$, yielding

$$\theta_0(G) \leq f_{\text{Kneser}}(\overline{G}) \leq \theta_0(G) + 1.$$

Theorem 2. *For constant $0 < p < 1$ there exist constants $c_i = c_i(p) > 0$, $i = 1, 2$ such that for $G \in \mathcal{G}(n,p)$ with high probability*

$$c_1 \frac{n}{\log n} < \theta_0(G) < c_2 \frac{n}{\log n}.$$

2. Minimum difference representations

In *difference representations*, generally speaking, vertices are adjacent if the representing sets are sufficiently different.

Definition 3. Let G be a graph on the vertices $[n] = \{1, \ldots, n\}$. A *k-min-difference representation* $(A_1, \ldots, A_n)$ of G is an assignment of a sets so that

$$ij \in E(G) \;\Leftrightarrow\; \min\{|A_i \setminus A_j|, |A_j \setminus A_i|\} \geq k.$$

Let $\mathcal{G}(n, k, \min)$ be the set of graphs with $V(G) = [n]$ that have a k-min-difference representation. The smallest k such that $G \in \mathcal{G}(n, k, \min)$ is denoted by $f_{\min}(G)$.

The co-star representation shows that $f_{\min}(G)$ exists and it is at most $n - \delta(G)$. Boros, Collado, Gurvits, and Kelmans [3] showed that many n-vertex graphs belong to $\mathcal{G}(n, 2, \min)$ and they did not find any graph with $f_{\min}(G) \geq 3$. Boros, Gurvitch and Meshulam [4] asked whether the value of $f_{\min}$ over all graphs is bounded by a constant. This question was answered in the negative by Balogh and Prince [2], who proved that for every k there is an n_0 such that whenever $n > n_0$, then for a graph G on n vertices we have $f_{\min}(G) \geq k$ with high probability. They used a Ramsey-type result of Balogh and Bollobás [1], so their bound on n_0 is a tower function of k.

Our main result is a significant improvement of the Balogh-Prince result. Let $\mathcal{G}(n,n)$ denote the family of 2^{n^2} bipartite graphs G with $n + n$ vertices.

Theorem 4. *There is a constant $c > 0$ such that for almost all bipartite graphs $G \in \mathcal{G}(n,n)$ one has $f_{\min}(G) \geq cn/(\log n)$.*

Corollary 5. *There is a constant $c > 0$ such that almost all graphs G on n vertices satisfy*

$$f_{\min}(G) \geq \frac{c \log n}{\log \log n}.$$

3. Prague dimension

The *Prague dimension* (it is also called *product dimension*) $f_{\mathrm{Pra}}(G)$ of a graph G is the smallest integer k such that one can find vertex distinguishing good colorings $\varphi_1, \dots, \varphi_k : V(G) \to \mathbb{N}$. This means that $\varphi_i(u) \neq \varphi_i(v)$ for every edge $uv \in E(G)$ and $1 \leq i \leq k$ but for every non-edge $\{u, v\}$, there exists an i with $\varphi_i(u) = \varphi_i(v)$, moreover the vectors $(\varphi_1(u), \varphi_2(u), \dots, \varphi_k(u))$ and $(\varphi_1(v), \varphi_2(v), \dots, \varphi_k(v))$ are distict for $u \neq v$. Two vertices are adjacent if and only if their labels disagree in every φ_i. We have

$$f_{\min}(G) \leq f_{\mathrm{Kneser}}(G) \leq f_{\mathrm{Pra}}(G).$$

Conjecture 6. *For a constant probability $0 < p < 1$ there exists a constant $c_{2,\mathrm{Pra}} = c_{2,\mathrm{Pra}}(p) > 0$, such that for $G \in \mathcal{G}(n,p)$ and also for $G \in \mathcal{G}(n,n,p)$ with high probability*

$$f_{\mathrm{Pra}}(G) < c_{2,\mathrm{Pra}} \frac{n}{\log n}.$$

Problem 7. Is it true that for any fixed $0 < p < 1$ for $G \in \mathcal{G}(n,p)$ with high probability one has $\Omega(n/(\log n)) \leq f_{\min}(G)$?

The Kneser rank and Prague dimension can be $\Omega(n)$. E.g., $f_{\mathrm{Kneser}}(\overline{K_{1,n-1}}) = n - 1$. No such results are known for $f_{\min}$.

Problem 8. What is the maximum of $f_{\min}(G)$ over the set of n-vertex graphs G? Is it true that it is $o(n)$ for every $G \in \mathcal{G}(n) \cup \mathcal{G}(n,n)$?

References

[1] J. Balogh and B. Bollobás, *Unavoidable traces of set systems*, Combinatorica **25** (2005), 633–643.

[2] J. Balogh and N. Prince, *Minimum difference representations of graphs*, Graphs Combin. **25** (2009), 647–655.

[3] E. Boros, R. Collado, V. Gurvich, and A. Kelmans, *On 2-representable graphs*, manuscript, DIMACS REU Report, July, (2000).

[4] E. Boros, V. Gurvich, and R. Meshulam, *Difference graphs*, Discrete Math. **276** (2004), 59–64.

[5] N. Eaton and V. Rödl, *Graphs of small dimensions*, Combinatorica **16** (1996), 59–85.

[6] P. Erdős, A. W. Goodman, and L. Pósa, *The representation of a graph by set intersections*, Canad. J. Math. **18** (1966), 106–112.

[7] A. Frieze and B. Reed, *Covering the edges of a random graph by cliques*, Combinatorica **15** (1995), 489–497.

[8] P. Hamburger, A. Por, and M. Walsh, *Kneser representations of graphs*, SIAM J. Discrete Math. **23** (2009), 1071–1081.

[9] J. Körner and A. Monti, *Compact representations of the intersection structure of families of finite sets*, SIAM J. Discrete Math. **14** (2001), 181–192.

[10] L. Lovász, J. Nešetřil, and A. Pultr, *On a product dimension of graphs*, J. Combin. Theory Ser. B **29** (1980), 47–67.

[11] T. A. McKee and F. R. McMorris, *Topics in intersection graph theory*, SIAM Monographs on Discrete Mathematics and Applications, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[12] S. Poljak, A. Pultr, and V. Rödl, *On the dimension of Kneser graphs*, in Algebraic methods in graph theory, Vol. I, II (Szeged, 1978), **25** Colloq. Math. Soc. János Bolyai, North-Holland, Amsterdam-New York, 1981, 631–646.

Performance of Sequential Local Algorithms for the Random NAE-K-SAT Problem

David Gamarnik
(joint work with M. Sudan)

In this work we consider a class of algorithms which we dub "sequential local algorithms" that capture local implementations of message-passing based decimation algorithms, including the BP-guided and the SP-guided decimation algorithms. We analyze the behavior of local sequential algorithms on random instances of "Not-All-Equal-K-SAT (NAE-K-SAT)".

Our setting and results. The NAE-K-SAT problem is a Boolean constraint satisfaction problem closely related to more commonly studied K-SAT problem. An instance of the NAE-K-SAT problem consists of a collection of N K-clauses on n Boolean variables $x_1, \ldots, x_n$. Each K-clause is given by K-literals, where each literal is either one of the variables or its negation. The clause is satisfied by a Boolean assignment to the variables if at least one of the literals is satisfied (set to 1) and at least one is unsatisfied (set to 0). (This symmetry between satisfied and unsatisfied literals lends a convenient symmetry to the NAE-K-SAT problem that is not shared by the K-SAT counterpart). The collection of N clauses is satisfied by a Boolean assignment if each clause is satisfied. Given $\ell \leq N$, we say that a Boolean assignment ℓ-satisfies the collection if at most ℓ clauses are violated.

In this work we consider the ability to find satisfying and ℓ-satisfying assignments to random instances of the NAE-K-SAT problem. Here the N clauses are chosen uniformly and independently from the collection of $2^K \cdot \binom{n}{K}$ possible K-clauses. Throughout the paper we consider the regime where the number of variables n grows, but the clause size K remains constant. In particular we consider the setting where $N = d \cdot n$ for some constant $d = d(K)$ that depends on K, but not n, and consider what is the largest d for which there exists a efficient algorithm for identifying an ℓ-satisfying satisfying assignment with probability bounded away from zero as $n \to \infty$ by some function going to zero at some rate with n. We will be interested primarily in regime where ℓ is a linear function of N and therefore n as well. The parameter d is often referred to as the formula density. Of course, no algorithm can find a satisfying assignment if none exists; and the limit of when such an assignment exists is well-studied. In particular Coja-Oghlan and and Panagiotou [7] have established that random instances of the NAE-K-SAT problem are satisfiable with high probability (w.h.p.) when the density d is below $d_s = 2^{K-1} \ln 2 - \ln 2/2 - 1/4 - o_K(1)$, and is not satisfiable w.h.p. when $d > d_s$. Earlier bound was obtained by Achlioptas and Moore [4] who obtained a bound of the form $d_s = 2^{K-1} \ln 2 - O_K(1)$. Here $o_K(\cdot)$ and $O_K(\cdot)$ denote standard order of magnitude notation as K increases. Our interest is in determining how qualitatively close to this threshold an efficient algorithm can get, i.e., how does the largest density at which the algorithm manages to find a satisfying or even ℓ-satisfying assignments compare with d_s.

The class of algorithms that we explore in this work are what we call "local sequential algorithms". A sequential local algorithm can be described roughly as follows. The algorithm works by assigning Boolean values to variables sequentially, where a chosen variable is assigned its value by a potentially probabilistic choice, which depends on the local neighborhood of the variable at the time the choice is made. The local neighborhood is defined to be the graph-theoretic $B(r)$ ball of small value r radius with respect to the underlying factor graph on the set of variables and clauses, to be defined later. Once a variable is assigned a value, the formula is simplified (removing some clauses, and restricting others). This in turn may influence the local neighborhoods of other variables, and when the future variables are set to particular Boolean values, this is done with respect to thus possibly modified neighborhoods. The algorithm continues with its iterations till all variables are set.

Local sequential algorithms capture restricted versions of BP- and SP-guided decimation algorithms, specifically when the number of message passage iterations used between every decimation step is bounded by $O(r)$. In the specific context of BP-guided decimation algorithm based on r iterations, the local rule assigns value 1 to a variable x with probability equal to the fraction of assignments in which x is assigned value 1 among all assignments that satisfy all clauses in the local neighborhood $B(r)$. The SP-guided decimation algorithm uses a more complex rule for its assignments. It is based on lifting the Boolean constraint satisfaction problem to a constraint satisfaction problem involving three decisions, as opposed to two decisions, but otherwise follows the same spirit.

Our main contribution is to show that, w.h.p. as the size of the instance diverges to infinity, every "balanced" sequential local algorithm fails to produce a ℓ-satisfying assignment when the ratio d of the number of clauses to the number of variables exceeds $(1 + o_K(1))\frac{2^{K-1}}{K} \ln^2 K$, clause size K is sufficiently large (but independent from the number of variables), and $r = O\left((\ln \ln n)^{O(1)}\right)$. Specifically, we will show this when the ratio of ℓ to the total number of clauses dn is below a certain constant less than unity, (see the aforementioned theorem for details). "Balance" is a technical condition, which says that the local algorithm respects the inherent symmetry between 0 and 1. It is a condition satisfied by all known algorithms inlcuding BP- and SP-guided decimation, as we establish.

Our bound on the ratio d is reasonably close to bounds at which simple algorithms actually work. In particular, it is well-known that a very simple Unit Clause algorithm is capable of finding satisfying assignments for this problem when d is below $\rho\frac{2^{K-1}}{K}$ for some universal constant ρ, [3] for K sufficiently large. The Unit Clause algorithm is the best known algorithm for this problem. (A better algorithm is known for the random K-SAT problem that works up to clause to variables ratio $(1 - o_K(1))\frac{2^K}{K} \ln K$ [5]. It is likely that a similar idea can be applied to the NAE-K-SAT setting, but such a result is not available to the best of our knowledge.) One of the hopes was that BP- and SP-guided decimation algorithms may be able to bridge this factor of K between the unit clause algorithm

and the satisfiability threshold d_s above. Our result however implies that, short of possibly a $\ln^2 K$ multiplicative factor, the "infamous" factor $O(K)$ gap between the satisfiability threshold and the region achievable by known algorithms cannot be broken by means of sequential local algorithms, in particular by BP- and SP-guided decimation algorithms with $O(r)$ number of rounds of message passing iterations.

Previously, Coja-Oghlan [6] showed that the BP-guided decimation algorithm fails to find satisfying assignments for random K-SAT problem when $d \geq \rho \frac{2^K}{K}$ for some universal constant ρ, for an arbitrary number of iterations r, which in particular might depend on the number of variables. (Here 2^K factor is an "appropriate" substitution for 2^{K-1} when switching from NAE-K-SAT to the K-SAT problem. We maintain this distinction, even though technically it is eliminated by constant ρ). It is reasonable to expect that his result holds also for NAE-K-SAT problem using the same analysis. Thus our result partially reproduces the main result of [6] in the special case when the number of iterations is bounded by $O\left((\ln \ln n)^{O(1)}\right)$ (short of additional $\ln^2 K$ factor). At the same time, however, our result is applied in a "blanket way" to a broad class of algorithms, including most notably SP-guided decimation algorithm with the number of iterations bounded by the same value, and our analysis is insensitive to the details of the algorithm.

After the first version of our paper was posted, we have became aware of the result by Hetterich [9] who shows that the SP-guided decimation algorithm w.h.p. fails to find a satisfying assignment of a random K-SAT formula above density $(d_s/K) \ln K$ for $d_s = 2^K \ln 2$, for all sufficiently large K. His result, unlike ours, does not assume any bound on the number of iterations of the SP-guided decimation algorithm and applies to a slightly smaller formula density $(d_s/K) \ln K$ as opposed to density $(d_s/K) \ln^2 K$ appearing in our main result.

Techniques. Our main proof technique relies on the intricate geometry of the solution space of the random NAE-K-SAT problem. Specifically it relies on the so-called *m-overlap* structure of nearly satisfying assignments of random NAE-K-SAT, which relates to the space of possible pair-wise Hamming distances between m such satisfying assignments. Previously this overlap structure was studied for the case $m = 2$ for random K-SAT problem, and several other related problems, including the problem of proper coloring of sparse random graphs [2],[1]. A certain *shattering* property was established which roughly speaking says that above a certain density, the Hamming distance between every pair of satisfying assignments (overlap) normalized by the number of variables, is either smaller than a certain constant δ_1 or larger than some constant $1 \geq \delta_2 > \delta_1$. As as result the solution space can be partitioned into different subsets (clusters) based on their proximity to each other. For the case of NAE-K-SAT problem this 2-overlap property can be established for densities d approximately $d > d_s/2$. (A weaker version of this result corresponding to "almost" all pairs does hold at densities above $O\left(\frac{d_s}{K} \ln K\right)$ [11]). Unfortunately, this is not strong enough to cover the regime of $d > (d_s/K) \ln^2 K$ claimed in our main theorem, so instead we have to establish a certain property

regarding m-overlaps of satisfying assignments, for appropriately chosen $m > 2$. This is the essence of a structural theorem which we prove in this paper. Roughly speaking this theorem says that, with probability at least $1 - \exp(-\Omega(n))$, when $d \geq (1+\epsilon)\frac{d_s}{K} \ln^2 K$, and K is sufficiently large, one cannot find $m \approx \epsilon K / \ln K$ satisfying assignments such that the Hamming distance (overlap) between every pair of the assignments normalized by the number of variables is $\approx \ln K / K$. The result applies to ℓ-satisfying assignments as well for sufficiently small $\ell < dn$. Then for every $\beta \in (0,1)$, we establish the following result. If a sequential local algorithm was capable of finding an ℓ-satisfying assignment, with probability at least $n^{-(\ln \ln n)^{O(1)}}$, then by running the algorithm m times and constructing a certain interpolation scheme, one can construct m ℓ-satisfying assignments such that the pairwise normalized distance between any pair of these assignments is $\approx \beta$, hence a contradiction. Our super polynomial upper bound $n^{-(\ln \ln n)^{O(1)}}$ on the likelihood of success also rules out the possibility of running the algorithm for polynomially many independent trials in the hope of finding at least one ℓ-satisfying assignment.

The link between the overlap property and the ensuing demise of local algorithms was recently established by authors [8] in a different context of finding a largest independent set in a random regular graph. There the argument was used to show that so-called i.i.d. factor based local algorithms are incapable of finding nearly largest independent sets in random regular graphs, refuting an earlier conjecture by Hatami, Lovász and Szegedy [10]. The result was further strengthened by Rahman and Virag [13], who obtained essentially the tightest possible result, using m-overlap structures of "large" independent set. Our use of m-overlaps is inspired by this work, though the set of restrictions on the m-overlaps is much simpler than the one appearing in [13].

An important technical and conceptual difference between the present work and that of [8] and [13] is that algorithms considered in the aforementioned papers are not sequential. Instead the decision taken by each variable in those models are taken simultaneously for all variables. In the case of sequential local algorithms, since the variables are set sequentially, the decision for one variable can be non-localized for the remaining variables, thus creating potential long-range dependencies. We deal with this potential long-range impact of decisions as follows. We associate variables with random i.i.d. weights chosen from an arbitrary continuous distribution, for example a uniform distribution. The weights are used solely to determine the order of fixing the values of the variables during the progression of the sequential local algorithm. Specifically the largest weight first rule is used. The decision to fix the value of a particular variable then can only impact variables with lower weights. Specifically if the value of variable x is fixed now, the value of variable y can be impacted *only* if there exists a sequence $x_0 = x, x_1, \ldots, x_\ell = y$ such that the distance between x_i and x_{i+1} is at most r (the radius of the decision making rule) and the weight of x_i is larger than that of x_{i+1} for all i. For a given set of variables $x_0, \ldots, x_\ell$ the likelihood of this total order of variables is $1/\ell!$ which decays faster than any exponential function in ℓ. This coupled with

the fact that the growth rate of nodes at distance at most $r\ell$ from x is at most exponential in $r\ell$, will allow us to control the range of influence of the variable x when its value is set. A similar idea of controlling the range of influence is used in the analysis of local algorithms in several places, including [12]. Bounding the ranges of influences is a crucial idea in implementing the interpolation scheme and constructing m assignments with "non-existent" normalized overlaps β.

References

[1] Dimitris Achlioptas and Amin Coja-Oghlan, *Algorithmic barriers from phase transitions*, Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on, IEEE, 2008, pp. 793–802.

[2] D. Achlioptas, A. Coja-Oghlan, and F. Ricci-Tersenghi, *On the solution space geometry of random formulas*, Random Structures and Algorithms **38** (2011), 251–268.

[3] Dimitris Achlioptas, Jeong Han Kim, Michael Krivelevich, and Prasad Tetali, *Two-coloring random hypergraphs*, Random Structures & Algorithms **20** (2002), no. 2, 249–259.

[4] Dimitris Achlioptas and Cristopher Moore, *Random k-sat: Two moments suffice to cross a sharp threshold*, SIAM Journal on Computing **36** (2006), no. 3, 740–762.

[5] Amin Coja-Oghlan, *A better algorithm for random k-sat*, SIAM Journal on Computing **39** (2010), no. 7, 2823–2864.

[6] A. Coja-Oghlan, *On belief propagation guided decimation for random k-SAT*, Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2011, pp. 957–966.

[7] A. Coja-Oglan and K. Panagiotou, *Catching the k-NAESAT threshold*, Proceedings of the 44th symposium on Theory of Computing, ACM, 2012, pp. 899–908.

[8] David Gamarnik and Madhu Sudan, *Limits of local algorithms over sparse random graphs*, Annals of Probability. To appear.

[9] Samuel Hetterich, *Analysing survey propagation guided decimation on random formulas*, arXiv preprint arXiv:1602.08519 (2016).

[10] H. Hatami, L. Lovász, and B. Szegedy, *Limits of local-global convergent graph sequences*, Preprint at http://arxiv.org/abs/1205.4356.

[11] Andrea Montanari, Ricardo Restrepo, and Prasad Tetali, *Reconstruction and clustering in random constraint satisfaction problems*, SIAM Journal on Discrete Mathematics **25** (2011), no. 2, 771–808.

[12] H.N. Nguyen and K. Onak, *Constant-time approximation algorithms via local improvements*, Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on, IEEE, 2008, pp. 327–336.

[13] Mustazee Rahman and Balint Virag, *Local algorithms for independent sets are half-optimal*, arXiv preprint arXiv:1402.0485 (2014).

Forbidden vector-valued intersections

Peter Keevash
(joint work with E. Long)

The Frankl-Rödl forbidden intersection theorem is a fundamental result of Extremal Set Theory, which has had a wide range of applications to different areas of mathematics, including discrete geometry [3], communication complexity [9] and quantum computing [1].

Let $[n] = \{1, \dots, n\}$ and let $\binom{[n]}{k} = \{A \subset [n] : |A| = k\}$. For $\mathcal{A} \subset \binom{[n]}{k}$ and $t \in [n]$ let $\mathcal{A} \times_t \mathcal{A}$ be the set of all $(A, B) \in \mathcal{A} \times \mathcal{A}$ with $|A \cap B| = t$. Note that $\binom{[n]}{k} \times_t \binom{[n]}{k}$ is non-empty if and only if $\max(2k - n, 0) \le t \le k$. Frankl and Rödl proved the following 'supersaturation theorem', showing that if t is bounded away from these extremes and $\mathcal{A}$ is 'exponentially dense' in $\binom{[n]}{k}$ then $\mathcal{A} \times_t \mathcal{A}$ is 'exponentially dense' in $\binom{[n]}{k} \times_t \binom{[n]}{k}$.

Theorem 1 (Frankl–Rödl [2])**.** *Let $0 < n^{-1} \ll \delta \ll \epsilon < 1$ and $\max(2k - n, 0) + \epsilon n \le t \le k - \epsilon n$. Suppose $\mathcal{A} \subset \binom{[n]}{k}$ with $|\mathcal{A}| \ge (1 - \delta)^n \binom{n}{k}$. Then $|\mathcal{A} \times_t \mathcal{A}| \ge (1 - \epsilon)^n \left| \binom{[n]}{k} \times_t \binom{[n]}{k} \right|$.*

In a recent survey on the Borsuk problem, Kalai [7] remarked that the Frankl–Rödl theorem can be used to give a counterexample to the Borsuk conjecture (the Frankl–Wilson intersection theorem [4] was used in Kahn and Kalai's celebrated counterexample [6]), and suggested that improved bounds might follow from a suitably generalised Frankl–Rödl theorem. He proposed the following supersaturation conjecture as a possible step in this direction, in which one measures a set by its size $|A| = \sum_{i \in A} 1$ and its sum $\sum A = \sum_{i \in A} i$. Let $[n]_{k,s}$ be the set of $A \subset [n]$ with $|A| = k$ and $\sum A = s$. For $\mathcal{A} \subset [n]_{k,s}$ write $\mathcal{A} \times_{(t,w)} \mathcal{A} = \{(A, B) \in \mathcal{A} \times \mathcal{A} : A \ne B \text{ with } |A \cap B| = t \text{ and } \sum(A \cap B) = w\}$.

Conjecture 2 (Kalai)**.** *Let $0 < n^{-1} \ll \delta \ll \epsilon, \alpha_1, \alpha_2, \beta_1, \beta_2 < 1$, $k = \lfloor \alpha_1 n \rfloor$, $s = \lfloor \alpha_2 \binom{n}{2} \rfloor$, $t = \lfloor \beta_1 n \rfloor$ and $w = \lfloor \beta_2 \binom{n}{2} \rfloor$. Suppose $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \ge (1 - \delta)^n |[n]_{k,s}|$. Then $|\mathcal{A} \times_{(t,w)} \mathcal{A}| \ge \lfloor (1 - \epsilon)^n \left| [n]_{k,s} \times_{(t,w)} [n]_{k,s} \right| \rfloor$.*

Somewhat surprisingly, this conjecture is false in general! In fact, although the conjecture holds in a number of natural special cases, it fails quite dramatically in general; for most pairs (α_1, α_2) there is exactly one choice of (β_1, β_2) for which Conjecture 2 holds.

Our first result exactly describes those values of $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ for which Conjecture 2 holds. Note that the set $[n]_{k,s}$ is exponentially large in n if and only if $(\alpha_1, \alpha_2) \in \Lambda := \{(x, y) : 0 < x < 1, x^2 < y < 2x - x^2\}$. As Conjecture 2 is only interesting in this setting, we will restrict attention to this region. We say that $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0, 1]^4$ is (δ, ϵ, n)-good if with $k, s, t, w \in \mathbb{N}$ as in Conjecture 2, all $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \ge (1 - \delta)^n |[n]_{k,s}|$ satisfy $|\mathcal{A} \times_{(t,w)} \mathcal{A}| \ge \lfloor (1 - \epsilon)^n \left| [n]_{k,s} \times_{(t,w)} [n]_{k,s} \right| \rfloor$.

Theorem 3. *There are functions $\beta_1, \beta_2 : \Lambda \to \mathbb{R}$ such that the following holds. Consider the sets:*

$$\Gamma_1 = \big\{(\alpha_1, \alpha_2, \beta_1, \beta_2) : (\alpha_1, \alpha_2) \in \Lambda,\ \alpha_1 \ne \alpha_2 \text{ and } \beta_i = \beta_i(\alpha_1, \alpha_2)\big\};$$
$$\Gamma_2 = \big\{(\alpha, \alpha, \beta, \beta) : 0 < \alpha < 1 \text{ and } \max(2\alpha - 1, 0) < \beta < \alpha\big\};$$
$$\Gamma_3 = \big\{(1/2, 1/2, \beta_1, \beta_2) : (2\beta_1, 2\beta_2) \in \Lambda\big\}.$$

Then given $\mathbf{g}' \in \Gamma = \bigcup_{i \in [3]} \Gamma_i$ and $\epsilon \gg \delta \gg n^{-1}$, any $\mathbf{g} \in [0, 1]^4$ with $\|\mathbf{g} - \mathbf{g}'\|_1 \le \delta$ is (δ, ϵ, n)-good. On the other hand, given $\mathbf{g} \in [0, 1]^4 \setminus \Gamma$ there is $\epsilon > 0$ such that $\mathbf{g}$ is not (δ, ϵ, n)-good for any $\epsilon \gg \delta \gg n^{-1}$.

We have the following interpretation of elements of $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in \Gamma$:

- **Popular intersections:** For $(\alpha_1, \alpha_2) \in \Lambda$ with $\alpha_1 \neq \alpha_2$ there is exactly one $(\alpha_1, \alpha_2, \beta_1, \beta_2) \in \Gamma_1$. For n large, the value $(\beta_1 n, \beta_2 \binom{n}{2})$ is essentially the most popular intersection between sets in $[n]_{k,s}$, where $(k, s) = (\alpha_1 n, \alpha_2 \binom{n}{2})$.
- **Doubly random:** If $A \subset [n]$ is a uniformly random set of size $k = \alpha n$, the expected size of $\sum A$ is $\alpha \binom{n}{2} + o(n^2)$. Similarly, if two sets A and B in $[n]_{k,s}$ with $|A \cap B| = \beta n$ are randomly selected then the expected value of $\sum (A \cap B)$ is $\beta \binom{n}{2} + o(n^2)$. Theorem 3 for Γ_2 shows Conjecture 2 holds for 'random βn intersections' between 'random αn sets', provided α and β satisfy the Frankl-Rödl conditions.
- **Uniformly random sets:** Lastly, most sets $A \in [n]_{k,s}$ with $(k, s) = (\frac{1}{2}n, \frac{1}{2}\binom{n}{2})$ have $|A \cap [2L]| = L \pm o(n) = |A \cap [n - 2L, n]|$ for all $L \leq n/2$. Intersections of type $(t, w) = (\beta_1 n, \beta_2 \binom{n}{2})$ can only occur between such sets if $\frac{1}{2}\binom{2t}{2} + o(n^2) \leq w \leq \frac{1}{2}\left(\binom{n+1}{2} - \binom{n-2t+1}{2}\right) + o(n^2)$. Equivalently $(2\beta_1 + o(1), 2\beta_2 + o(1)) \in \Lambda$. Theorem 3 for Γ_3 shows that Conjecture 2 is true for $(\alpha_1, \alpha_2) = (1/2, 1/2)$ provided this necessary condition is fulfilled.

Although the bounds from Conjecture 2 in general do not hold, it is still natural to ask if we can find any (t, w)-intersections in $\mathcal{A}$, and if so, how many. We obtain essentially best possible answers to these questions (although for simplicity we will not state these results in this abstract). Our results apply to vector-valued intersections defined by more general vectors in higher dimensions, and also to cross-intersection theorems for two or more families over vectors over any finite alphabet. However, for simplicity in this abstract we stick with the setting of Kalai's conjecture. For cleaner notation and compatibility with the general setting we introduce the notation $\mathcal{X} = [n]_{k,s}$ and $I(\mathcal{A}) = \mathcal{A} \times_{(t,w)} \mathcal{A}$.

Our next theorem shows that, there are only two obstructions to a supersaturation result as in Kalai's conjecture; either there exists a small set $\mathcal{B}_{full}$ which artificially inflates the number of (t, w)-intersections in $\mathcal{X}$, or there is a large set $\mathcal{B}_{empty}$ which contains no (t, w)-intersections. Furthermore, in the latter case, our result still gives supersaturation relative to $\mathcal{B}_{full}$.

Theorem 4. *Let $n^{-1} \ll \delta \ll \epsilon, \alpha_1, \alpha_2, \beta_1, \beta_2$. Then one of the following holds:*

(1) *All $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| \geq (1-\delta)^n |\mathcal{X}|$ satisfy $|I(\mathcal{A})| \geq (1-\epsilon)^n |I(\mathcal{X})|$;*
(2) *There exists $\mathcal{B}_{full} \subset \mathcal{X}$ with $|\mathcal{B}_{full}| \leq (1-\delta)^n |\mathcal{X}|$ and*
$$|I(\mathcal{X}) \setminus I(\mathcal{B}_{full})| \leq (1-\delta)^n |I(\mathcal{X})|.$$
Furthermore, for $\mathcal{B} \subset \mathcal{B}_{full}$ with $|\mathcal{B}| \geq (1-\delta)^n |\mathcal{B}_{full}|$ we have $|I(\mathcal{B})| \geq (1-\epsilon)^n |I(\mathcal{X})|$;
(3) *There is $\mathcal{B}_{empty} \subset \mathcal{X}$ with $|\mathcal{B}_{empty}| \geq \lfloor (1-\epsilon)^n |\mathcal{X}| \rfloor$ and $I(\mathcal{B}_{empty}) = \emptyset$.*

A key paradigm of our approach is that $\mathcal{V}$-intersection theorems often have equivalent formulations in terms of certain maximum entropy measures, and that the necessary condition for these theorems appears naturally as a condition on the measures.

To illustrate this point, we recast the Frankl-Rödl theorem in such terms. Again we identify subsets of $[n]$ with their characteristic vectors in $\{0,1\}^n$, on which we introduce the product measure $\mu_p(\mathbf{x}) = \prod_{i\in[n]} p_{x_i}$, where $p_1 = k/n$ and $p_0 = 1-p_1$. Pairs of subsets are identified with $\{0,1\}^n \times \{0,1\}^n$, which we can identify with $(\{0,1\}\times\{0,1\})^n$, on which we introduce the product measure $\mu_q(\mathbf{x},\mathbf{x}') = \prod_{i\in[n]} q_{x_i,x_i'}$, where $q_{1,1} = t/n$, $q_{0,1} = q_{1,0} = (k-t)/n$ and $q_{0,0} = (n-2k+t)/n$. It is not hard to see that the hypothesis of Theorem 1 is essentially equivalent to $\mu_p(\mathcal{A}) > (1-\delta)^n$ and the conclusion to $\mu_q(\mathcal{A}\times_t\mathcal{A}) > (1-\epsilon)^n$. Furthermore, the assumption on t can be rephrased as $q_{j,j'} \geq \epsilon$ for all $j,j' \in \{0,1\}$, and this indicates the condition that we need in general.

The next result gives our probabilistic forbidden intersection theorem. We say that the product measure $\mu_{\mathbf{q}}$ is κ-bounded if all $q^i_{j,j'} \in [\kappa, 1-\kappa]$.

Theorem 5. *Let $0 < n^{-1}, \delta \ll \kappa, \gamma, \epsilon, \alpha_1, \alpha_2, \beta_1, \beta_2$. Suppose that $\mu_{\mathbf{q}}$ is a κ-bounded product measure on $(\{0,1\}\times\{0,1\})^n$ with $\sum_i (1,i)q^i_{1,1} = (t,w)$ and both marginals $\mu_{\mathbf{p}}$. Let $\mathcal{A} \subseteq \{0,1\}^n$ with $\mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta)^n$. Then $\mu_{\mathbf{q}}(I(\mathcal{A})) > (1-\epsilon)^n$.*

We briefly indicate how Theorem 5 can be used to prove Theorem 4. First we prove a large deviation principle, which can be roughly stated as saying that the uniform measure on $\mathcal{X}$ has the same exponential tail events in $\mathcal{X}$ as the measure $\mu_{\mathbf{p}}$ on $\{0,1\}^n$ that has maximum entropy subject to $\sum_i p_i(1,i) = (k,s)$. Thus we can apply Theorem 5 to any $\mathcal{A}$ as in Theorem 4 and any κ-bounded $\mu_{\mathbf{q}}$ with marginals $\mu_{\mathbf{p}}$. Similarly, there is a maximum entropy measure $\mu_{\mathbf{q}}$ on $(\{0,1\}\times\{0,1\})^n$ with the same exponential tail events in $I(\mathcal{X})$ as the uniform measure.

The first question to consider is whether $\mu_{\mathbf{q}}$ is κ-bounded. If not then we cannot guarantee any (t,w)-intersections. The key parameter in understanding this case is the well-known [10] Vapnik-Chervonenkis dimension of $I(\mathcal{X}) \subseteq (\{0,1\}\times\{0,1\})^n$. To see why it is natural to consider the VC-dimension, consider the problem of finding an intersection of size $n/3$ among subsets of $[n]$ of size $2n/3$. The conditions of the Frankl-Rödl theorem are not satisfied, and indeed the conclusion is not true: take $\mathcal{A} = \{A \in \binom{[n]}{2n/3} : 1 \notin A\}$. Considering $\binom{[n]}{2n/3} \times_{n/3} \binom{[n]}{2n/3}$ as a subset of $(\{0,1\}\times\{0,1\})^n$, we see that no coordinate can take the value $(0,0)$, so there is not even a shattered set of size 1! Modifying this example in the obvious way we see that it is natural to assume a bound that is linear in n. We also note that this example shows that the 'Frankl-Rödl analogue' of Conjecture 2 is not true, and hints towards a counterexample for Kalai's conjecture.

If $\mu_{\mathbf{q}}$ is κ-bounded then the next question to consider is whether its marginals $\mu_{\widetilde{\mathbf{p}}}$ are close to the maximum entropy measure $\mu_{\mathbf{p}}$ that describes $\mathcal{X}$. If so, then for $\mathcal{A}) \subseteq \mathcal{X}$ we have the implications $|\mathcal{A}|/|\mathcal{X}| > (1-\delta)^n \Rightarrow \mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta')^n \Rightarrow \mu_{\mathbf{q}}(I(\mathcal{A})) > (1-\delta'')^n \Rightarrow I(\mathcal{A})/I(\mathcal{X}) > (1-\epsilon)^n$, i.e. we have supersaturation as in Kalai's Conjecture. On the other hand, if $\mu_{\widetilde{\mathbf{p}}}$ is far from $\mu_{\mathbf{p}}$ then it concentrates on some small subset $\mathcal{B}_{full}$ of $\mathcal{X}$, which accounts for most of the (t,w)-intersections in $\mathcal{X}$.

This sketch also gives some indication of how the values in Theorem 3 arise. We can explicitly describes the maximum entropy measures $\mu_{\mathbf{q}}$ and $\mu_{\mathbf{p}}$ considered

above as Boltzmann distributions. It turns out to be quite rare for a Boltzmann distribution to be a marginal of another, although occasionally it does occur. One simple instance is given by taking the measure $\mu_{\mathbf{q}}$ formed by two independent copies of $\mu_{\mathbf{p}}$, which leads to the 'popular intersection' case of Theorem 3.

References

[1] H. Buhrman, R. Cleve and A. Wigderson, Quantum vs. classical communication and computation, *Proceedings of 30th STOC* (1998), 63–68.

[2] P. Frankl and V. Rödl, Forbidden intersections, *Trans. Amer. Math. Soc.* **300** (1987), 259–286.

[3] P. Frankl and V. Rödl, A partition property of simplices in Euclidean space, *J. Amer. Math. Soc.* **3** (1990), 1–7.

[4] P. Frankl and R.M. Wilson, Intersection theorems with geometric consequences, *Combinatorica* **1** (1981), 357–368.

[5] E.T. Jaynes, Information Theory and Statistical Mechanics, *Physical Review Series II* **106** (1957), 620–630.

[6] J. Kahn and G. Kalai, A counterexample to Borsuk's conjecture, *Bull. Amer. Math. Soc.* **29** (1993), 60–62.

[7] G. Kalai, Some old and new problems in combinatorial geometry I: Around Borsuk's problem, in *Surveys in Combinatorics* 2005, 147–174, London Math. Soc. Lecture Note Ser., 424, Cambridge Univ. Press, 2015.

[8] P. Keevash and E. Long, Frankl–Rödl type theorems for codes and permutations, *Trans. Amer. Math. Soc.* **369** (2017), 1147–1162.

[9] J. Sgall, Bounds on pairs of families with restricted intersections, *Combinatorica* **19**(4), (1999), 555-566.

[10] V.N. Vapnik and A.Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** (1971), 264–280.

The distribution of the number of triangles in a random graph

Swastik Kopparty

(joint work with J. Gilmer)

We will work with the Erdos-Rényi random graph $G(n,p)$. Recall that $G(n,p)$ is the random undirected graph G on n vertices sampled by including each of the $\binom{n}{2}$ possible edges into G independently with probability p. Let S_n be the random variable equal to the number of triangles in $G(n,p)$. Let $\mu_n = \mathbf{E}[S_n] = p^3\binom{n}{3}$ and $\sigma_n = \sqrt{\mathbf{Var}[S_n]} = \Theta(n^2)$ (see [2] for an exact expression of σ_n). Our main result [2] states that if p is a fixed constant in $(0,1)$, then the distribution of S_n is *pointwise* approximated by a discrete Gaussian distribution:

$$\Pr[S_n = k] = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-((k-\mu_n)/\sigma_n)^2/2} \pm o(1/n^2). \tag{1}$$

Thus, for every $k \in \mu_n \pm O(\sigma_n)$, we determine the probability that $G(n,p)$ has exactly k triangles, up to a $(1+o(1))$ multiplicative factor.

The proofs are based on bounding the characteristic function $\psi(t)$ of S_n, and involves a number of different conditioning arguments for different ranges of t.

More recently, Berkowitz [1] greatly enhanced the power of these methods with some elegant conceptual and technical ideas, and gave a significantly improved quantitative version of this result. His main result states (with the same setup) that for each fixed $\epsilon > 0$,

$$\Pr[S_n = k] = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-((k-\mu_n)/\sigma_n)^2/2} \pm o(\frac{1}{n^{2.5-\epsilon}}). \tag{2}$$

This improved quantitative bound then enables Berkowitz to deduce an ℓ_1-distance bound between the distribution of S_n and the discrete Gaussian distribution.

$$\sum_{k=-\infty}^{\infty} \left| \Pr[S_n = k] - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-((k-\mu_n)/\sigma_n)^2/2} \right| \leq o(\frac{1}{n^{0.5-\epsilon}}). \tag{3}$$

References

[1] R. Berkowitz, *A Quantitative Local Limit Theorem for Triangles in Random Graphs*, arXiv preprint arXiv:1610.01281 (2016).

[2] J. Gilmer, S. Kopparty, *A local central limit theorem for triangles in a random graph*, Random Structures and Algorithms (2016).

Local 3-profiles of graphs

Daniel Král'

(joint work with R. Glebov, A. Grzesik, P. Hu, T. Hubai, J. Volec)

Many problems in graph theory relate to understanding possible combinations of densities of subgraphs in graphs. One of the recent breakthroughs in extremal graph theory was the description of possible densities of complete graphs in graphs with a given edge density, which is given in the exciting work of Razborov [4], Nikiforov [3] and Reiher [5]. While problems related to possible densities of small graphs may look innocent at the first sight, they can become incredibly challenging. For example, determining the minimum possible sum of densities of $\overline{K_4}$ and K_4 is a well-known problem open for more than five decades.

Huang, Linial, Naves, Peled and Sudakov [1], building on their results from [2], determined possible densities of $\overline{K_3}$ and K_3. We contribute to this line of research by completing the description of possible densities of all pairs of 3-vertex graphs. More precisely, let $d_i(G)$ be the density of the 3-vertex i-edge graph in a graph G, i.e., the probability that three random vertices induce a subgraph with i edges, and let $S \subseteq \mathbb{R}^4$ be the set of all quadruples (d_0, d_1, d_2, d_3) that are arbitrary close to 3-vertex graph densities in arbitrary large graphs. Huang et al. [1] determined the projection of the set S to the (d_0, d_3) plane, and we determine the projection of the set S to all the remaining planes.

References

[1] H. Huang, N. Linial, H. Naves, Y. Peled and B. Sudakov, *On the 3-local profiles of graphs*, J. Graph Theory **76** (2014), 236–248.

[2] H. Huang, N. Linial, H. Naves, Y. Peled and B. Sudakov, *On the densities of cliques and independent sets in graphs*, Combinatorica, to appear.

[3] V. Nikiforov, *The number of cliques in graphs of given order and size*, Trans. Amer. Math. Soc. **363** (2011), 1599–1618.

[4] A. Razborov, *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), 603–618.

[5] C. Reiher, *The clique density theorem*, Ann. of Math. **184** (2016), 683–707.

The existence of designs – beyond quasirandomness

Daniela Kühn

(joint work with S. Glock, A. Lo, and D. Osthus)

1. Block designs and Steiner systems

An (n, q, r, λ)-*design* is a set X of q-subsets (often called 'blocks') of some n-set V, such that every r-subset of V belongs to exactly λ elements of X. (Note that this makes only sense if $q > r$, which we assume throughout.) In the case when $r = 2$, this coincides with the notion of balanced incomplete block designs. An $(n, q, r, 1)$-design is also called an (n, q, r)-*Steiner system*. There are some obviously necessary 'divisibility conditions' for the existence of a design: consider some subset S of V of size $i < r$ and assume that X is an (n, q, r, λ)-design. Then the number of elements of X which contain S is $\lambda\binom{n-i}{r-i}/\binom{q-i}{r-i}$. We say that the necessary divisibility conditions are satisfied if $\binom{q-i}{r-i}$ divides $\lambda\binom{n-i}{r-i}$ for all $0 \leq i < r$.

The 'Existence conjecture' states that for given q, r, λ, the necessary divisibility conditions are also sufficient for the existence of an (n, q, r, λ)-design, except for a finite number of exceptional n. Its roots can be traced back to work of e.g. Plücker, Kirkman and Steiner in the 19th century. Over a century later, a breakthrough result of Wilson [17–19] resolved the graph case $r = 2$.

For $r \geq 3$, much less was known until relatively recently. In 1963, Erdős and Hanani proposed an approximate version of the Existence conjecture for the case of Steiner systems. More precisely, they asked whether one can find blocks which cover every r-set at most once and cover all but $o(n^r)$ of the r-sets, as n tends to infinity. This was proved in 1985 by Rödl [15] via his celebrated 'nibble' method.

Teirlinck [16] was the first to prove the existence of designs for arbitrary $r \geq 6$ (this however requires $q = r + 1$ and λ large compared to q). Kuperberg, Lovett and Peled [13] proved a 'localized central limit theorem' for rigid combinatorial structures, which implies the existence of designs for arbitrary q and r, but again for large λ. There are many constructions resulting in sporadic and infinite families of designs (see e.g. the handbook [4]). However, the set of parameters they cover is

very restricted. In particular, even the existence of infinitely many Steiner systems with $r \geq 4$ was open until recently.

In a recent breakthrough, Keevash [10] proved the Existence conjecture in general, based on the method of 'Randomised algebraic constructions'.

In [7], we provide a new proof of the Existence conjecture based on the method of iterative absorption. In fact, our main theorem is considerably more general than this (as well as the results in [10]). Roughly speaking, our main result concerns $K_q^{(r)}$-decompositions of hypergraphs whose clique distribution fulfils certain regularity requirements. A 'regularity boosting' process frequently enables us to meet these requirements even if the clique distribution of the original hypergraph does not satisfy them. This enables us to strengthen the results in [10] as well as to derive a number of new results which go beyond the setting of quasirandom hypergraphs. In particular, in combination with this regularity boosting technique, our main result gives us a minimum degree version (see Theorem 1), a resilience version and effective bounds for a typicality version.

The method of iterative absorption was initially introduced in [10, 11] to find Hamilton decompositions of graphs. In the meantime it has been successfully applied to verify the Gyárfás-Lehel tree packing conjecture for bounded degree trees [9], as well as to find decompositions of dense graphs into a given graph F [2, 3, 6].

2. Designs in hypergraphs

We will study designs in a hypergraph setting. Here a *hypergraph* H is a pair (V, E), where V is the vertex set and the edge set E is a set of subsets of V. We identify H with E. In particular, we let $|H| := |E|$. We say that H is an *r-graph* if every edge has size r. We let $K_n^{(r)}$ denote the complete r-graph on n vertices.

Let H be some r-graph. A *$K_q^{(r)}$-decomposition* of H is a collection $\mathcal{K}$ of copies of $K_q^{(r)}$ in H such that every edge of H is contained in exactly one of these copies. More generally, a *(q, r, λ)-design* of H is a collection $\mathcal{K}$ of distinct copies of $K_q^{(r)}$ in H such that every edge of H is contained in exactly λ of these copies. Note that a (q, r, λ)-design of $K_n^{(r)}$ is equivalent to an (n, q, r, λ)-design.

For a set $S \subseteq V$ with $0 \leq |S| \leq r$, the $(r-|S|)$-graph $H(S)$ has vertex set $V \setminus S$ and contains all $(r-|S|)$-subsets of $V \setminus S$ that together with S form an edge in H. ($H(S)$ is often called the *link graph of S*.) We say that H is *(q, r, λ)-divisible* if for every $S \subseteq V$ with $0 \leq |S| \leq r-1$, we have that $\binom{q-|S|}{r-|S|}$ divides $\lambda|H(S)|$. Similarly to Section 1, this is a necessary condition for the existence of a (q, r, λ)-design of H. We say that H is *$K_q^{(r)}$-divisible* if H is $(q, r, 1)$-divisible.

We let $\delta(H)$ denote the minimum $(r-1)$-degree of an r-graph H, that is, the minimum value of $|H(S)|$ over all $S \subseteq V(H)$ of size $r-1$. The following result guarantees designs not just for $K_n^{(r)}$, but also for r-graphs which are allowed to be far from complete in the sense that they only have large minimum degree.

Theorem 1 (Minimum degree version, [7]). *For all $q > r \geq 2$ and $\lambda \in \mathbb{N}$, there exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Let*

$$c^{\diamond}_{q,r} := \frac{r!}{3 \cdot 14^r q^{2r}}.$$

Suppose that G is an r-graph on n vertices with $\delta(G) \geq (1 - c^{\diamond}_{q,r})n$. Then G has a (q, r, λ)-design if it is (q, r, λ)-divisible.

The main result of [10] implies a weaker version where $c^{\diamond}_{q,r}$ is replaced by some non-explicit $\varepsilon \ll 1/q$. Theorem 1 motivates the following very challenging problem regarding the decomposition threshold $c_{q,r}$ of $K^{(r)}_q$.

Problem 2 ([7]). *Determine the supremum $c_{q,r}$ of all $c \in [0, 1]$ with the following property: There exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, every $K^{(r)}_q$-divisible r-graph on n vertices with $\delta(G) \geq (1 - c)n$ has a $K^{(r)}_q$-decomposition.*

Theorem 1 implies that $c_{q,r} \geq c^{\diamond}_{q,r}$. It is not clear what the correct value should be. We note that for all $r, q, n_0 \in \mathbb{N}$, there exists an r-graph G_n on $n \geq n_0$ vertices with $\delta(G_n) \geq (1 - b_r \frac{\log q}{q^{r-1}})n$ such that G_n does not contain a single copy of $K^{(r)}_q$, where $b_r > 0$ only depends on r.

We now consider the graph case $r = 2$. A famous conjecture by Nash-Williams [14] on the decomposition threshold of a triangle would imply that $c_{3,2} = 3/4$. Until recently, the best bound for the problem was by Gustavsson [8], who claimed that $c_{q,2} \geq 10^{-37} q^{-94}$. Iterated absorption methods have led to significant progress in this area. For instance, the results in [6] imply that $c_{q,2} = c^*_{q,2}$, where $c^*_{q,r}$ denotes the fractional version of the decomposition threshold (the triangle case $q = 3$ was already obtained in [2]). This in turn has resulted in significantly improved explicit bounds on $c_{q,2}$, via results on fractional decompositions obtained in [1, 5]. In particular, the results from [1, 6] imply that $\frac{1}{10^4 q^{3/2}} \leq c_{q,2} \leq \frac{1}{q+1}$, where the upper bound is conjectured to be the correct value. The results in [2, 6] make (implicit) use of Szemerédi's regularity lemma, whereas our proof avoids this, resulting in much more moderate requirements on n.

References

[1] B. Barber, D. Kühn, A. Lo, R. Montgomery and D. Osthus, Fractional clique decompositions of dense graphs and hypergraphs, *arXiv:1507.04985*, 2015.

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

[3] B. Barber, D. Kühn, A. Lo, D. Osthus and A. Taylor, Clique decompositions of multipartite graphs and completion of Latin squares, *arXiv:1603.01043*, 2016.

[4] C.J. Colbourn and J.H. Dinitz, *Handbook of Combinatorial Designs*, 2nd edition, CRC Press, 2006.

[5] F. Dross, Fractional triangle decompositions in graphs with large minimum degree, *SIAM J. Discr. Math.* **30** (2016), 36–42.

[6] S. Glock, D. Kühn, A. Lo, R. Montgomery and D. Osthus, On the decomposition threshold of a given graph, *arXiv:1603.04724*, 2016.

[7] S. Glock, D. Kühn, A. Lo and D. Osthus, The existence of designs via iterative absorption, *arXiv:1611.06827*, 2016.

[8] T. Gustavsson, *Decompositions of large graphs and digraphs with high minimum degree*, PhD thesis, Univ. of Stockholm, 1991.
[9] F. Joos, J. Kim, D. Kühn and D. Osthus, Optimal packings of bounded degree trees, *arXiv:1606.03953*, 2016.
[10] P. Keevash, The existence of designs, *arXiv:1401.3665*, 2014.
[11] F. Knox, D. Kühn and D. Osthus, Edge-disjoint Hamilton cycles in random graphs, *Random Structures Algorithms* **46** (2015), 397–445.
[12] D. Kühn and D. Osthus, Hamilton decompositions of regular expanders: A proof of Kelly's conjecture for large tournaments, *Adv. Math.* **237** (2013), 62–146.
[13] G. Kuperberg, S. Lovett and R. Peled, Probabilistic existence of regular combinatorial objects, *arXiv:1302.4295*, 2013. (Extended abstract version appeared in STOC 2012.)
[14] C.St.J.A. Nash-Williams, An unsolved problem concerning decomposition of graphs into triangles, In: *Combinatorial Theory and its Applications III*, pages 1179–1183, North Holland, 1970.
[15] V. Rödl, On a packing and covering problem, *European J. Comb.* **5** (1985), 69–78.
[16] L. Teirlinck, Non-trivial t-designs without repeated blocks exist for all t, *Discrete Math.* **65** (1987), 301–311.
[17] R.M. Wilson, An existence theory for pairwise balanced designs I. Composition theorems and morphisms, *J. Combinatorial Theory Ser. A* **13** (1972), 220–245.
[18] R.M. Wilson, An existence theory for pairwise balanced designs II. The structure of PBD-closed sets and the existence conjectures, *J. Combinatorial Theory Ser. A* **13** (1972), 246–273.
[19] R.M. Wilson, An existence theory for pairwise balanced designs III. Proof of the existence conjectures, *J. Combinatorial Theory Ser. A* **18** (1975), 71–79.

Some stories of high-dimensional permutations

Nati Linial

What is the higher-dimensional analog of a permutation? If we think of a permutation as given by a permutation matrix, then the following definition suggests itself: A d-dimensional permutation of order n is an $n \times n \times \ldots n = [n]^{d+1}$ array of zeros and ones in which every *line* contains a unique 1 entry. A line here is a set of entries of the form $\{(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_{d+1}) | n \geq y \geq 1\}$ for some index $d+1 \geq i \geq 1$ and some choice of $x_j \in [n]$ for all $j \neq i$. It is easy to observe that a one-dimensional permutation is simply a permutation matrix and that a two-dimensional permutation is synonymous with an order-n Latin square.

In this lecture I surveyed three aspects of this subject:

- How many d-dimensional permutations are there? Interpolating from Stirling's formula and the estimate for the number of Latin squares, it is natural to conjecture that the number is:

$$\left((1+o(1))\frac{n}{e^d}\right)^{n^d}.$$

  In joint work with Luria [2] we proved this as an upper bound. A matching lower bound for dimensions 3 and above is still missing, though there is hope that recent advances in the study of combinatorial designs can help accomplish this.
- In joint work with Michael Simkin [3] we proved the following facts:

  - Every d-dimensional permutation of order n contains a strongly monotone subsequence of length $\Omega_d(\sqrt{n})$. The dependence on n is sharp.
  - In almost every d-dimensional permutation the length of the longest monotone subsequence is $\Theta_d(n^{\frac{d-1}{d}})$.

  The analogous statements for weakly monotone subsequences remain unknown. Also the exact dependence of these bounds on the dimension has not been determined yet.
- In recent work with Luria [4] we seek high-dimensional permutations of low discrepancy. A box $B \subseteq [n]^{d+1}$ is the cartesian product $B = \prod_1^{d+1} A_i$ and its volume is defined as $\prod |A_i|$. We conjecture that permutations X exist such that for every box B

$$||X \cap B| - vol(B)/n| \leq O(\sqrt{vol(B)}).$$

  Moreover it is conjectured that similar and even stronger bounds hold for almost every d-dimensional permutation. We say that box B is empty if $X \cap B = \emptyset$. A special case of the above conjecture which is established in [4] is that permutations exist (or even for almost all permutations) where every empty box has volume $O(n^2)$. It is interesting to note that a theorem of Kedlaya [1] shows that the Latin square corresponding to every finite group has a substantially larger empty box of volume $\Omega(n^{\frac{33}{14}})$. This clearly suggests that completely new ideas will be needed in order to find good explicit constructions in this domain.

References

[1] Kedlaya, Kiran S. Large product-free subsets of finite groups. J. Combin. Theory Ser. A 77 (1997), 339–343.
[2] N. Linial and Z. Luria, An upper bound on the number of high-dimensional permutations. Combinatorica 34(2014), 471–486.
[3] Nathan Linial, Michael Simkin, Monotone Subsequences in High-Dimensional Permutations, Combinatorics Probability and Computing, to appear arXiv:1602.02719 .
[4] Linial, Nathan; Luria, Zur Discrepancy of high-dimensional permutations. Discrete Anal. 2016, Paper No. 11, 8 pp.

Integral homology of random simplicial complexes

Tomasz Luczak
(joint work with Y. Peled)

The random 2-dimensional simplicial complex process $\{Y_2(n, M) : 0 \leq M \leq \binom{n}{3}\}$ is a Markov chain which starts with a complete graph with vertex set $[n] = \{1, 2, \ldots, n\}$ and in each step a new 2-dimensional face, chosen uniformly at random, is added. In a similar way one can define the Markov process $\{Y_2(n, p) : 0 \leq p \leq 1\}$, where the stage of this process, $Y(n, p)$, is the simplicial complex with full 1-skeleton in which each 2-dimensional edge is present independently with probability p. This model of topological structures was introduced by Linial nad Meshulam [4] who showed that the threshold for vanishing of the first homology

of $Y_2(n,p)$ over $\mathbb{F}_2$ is $(2+o(1))\ln n/n$ and it is the same as the threshold for the property that each pair 1-dimensional face is contained in some 2-dimensional face of $Y_2(n,p)$. This result was generalized by Meshulam and Wallach [6] to higher dimensions and arbitrary fixed abelian groups. Recently Kahle and Pittel [2] proved the hitting time result of Linial-Meshulam's result, showing that whp in the random simplicial complex process the first homology over $\mathbb{F}_2$ vanishes precisely in the moment when for the first time all pairs are covered by 2-dimensional faces.

For the case of $\mathbb{Z}$, which is not the field, so its homology group can have a non-trivial torsion factor, the threshold function was determined only up to a contant factor by Hoffman, Kahle and Paquette [1] who showed that whp $H_1(Y_2(n,p))$ vanishes provided $p \geq 160\ln n/n$. We determine the threshold for vanishing of the first homology over $\mathbb{Z}$ proving the following hitting time statment.

Theorem 1. *In the random 2-dimensional complex process whp the hitting time for the property that $H_1(Y_2(n,M);\mathbb{Z}) = 0$ is the same as for the property that each pair of vertices belong to some 2-dim face of $Y_2(n,M)$.*

Our argument has two main ingredients, algebraic and combinatorial. For a 2-dim simplicial complex Y with vertex set $[n]$ and the full 1-dim skeleton we define the shadow of Y over a field $\mathbb{F}$ as

$$\mathrm{SH}(Y;\mathbb{F}) = \{f \in \binom{[n]}{3} : H_1(Y;\mathbb{F}) = H_1(Y;\mathbb{F}\cup f)\}.$$

Thus, the shadow is the set of all potential 2-dim faces which, added to Y, would not affect its first homology group. It is easy to see that this group vanishes, i.e. $H_1(Y;\mathbb{F}) = 0$, if and only if the shadow is full, i.e. $|\mathrm{SH}(Y;F)| = \binom{[n]}{3}$. Moreover, the following result of Kalai [3] states that the first homology over $\mathbb{Z}$ vanishes when it vanishes over all fields which are not too large.

Lemma 2. *Let Y be a 2-dim simplicial complex on n vertices. If $H_1(Y;\mathbb{F}) = 0$ for every field $\mathbb{F}$ such that*

$$|\mathbb{F}| \leq \sqrt{3}^{\binom{n-1}{2}} \leq \exp(-n^2)\,,$$

then also $H_1(Y;\mathbb{Z}) = 0$.

It is not true that for a given field $\mathbb{F}$ the probability that $H_1(Y_2(n,p);\mathbb{F}) = 0$, which is equivalent to $|\mathrm{SH}(Y_2(n,p);\mathbb{F})| = \binom{[n]}{3}$, tends to 0 faster than $\exp(-n^2)$ for $p \sim 2\log n/n$ which is close to the threshold we are interested in. However, the following, slightly weaker, result holds.

Lemma 3. *If $p \geq \log n/n$ then*

$$\Pr\left(|SH(Y_2(n,p);\mathbb{F})| \geq \binom{[n]}{3} - \frac{n^3}{\ln\ln n}\right) = \exp(-2n^2)\,.$$

Moreover, we can supplement the above algebraic result by a combinatorial one. We say that a 2-dim simplicial complex Y on n vertices has property $\mathcal{B}$ if for each field $\mathbb{F}$ such that

$$|\mathrm{SH}(Y_2(n,p);\mathbb{F})| \geq \binom{[n]}{3} - \frac{n^3}{\ln\ln n}\,,$$

we have $|\mathrm{SH}(Y_2(n,p);\mathbb{F})| = \binom{[n]}{3}$. Then, clearly, Theorem 1 is an immediate cnsequence of Lemmata 2, 3, and the following result.

Lemma 4. *Whp the hitting time for the property $\mathcal{B}$ is the same as the hitting time for the property that each pair of vertices is covered by at least one 2-dim face.*

References

[1] C. Hoffman, M. Kahle, and E Paquette, *The threshold for integer homology in random d-complexes*, `arXiv:1308.6232`

[2] M. Kahle and B. Pittel, *Inside the critical window for cohomology of random k-complexes*, Random Structures & Algorithms **48** (2016), 102–124.

[3] G. Kalai, *Enumeration of $\mathbb{Q}$-acyclic simplicial complexes*, Israel Journal of Mathematics, **45** (1983) 337–351.

[4] N. Linial, R. Meshulam, *Homological connectivity of random 2-complexes*, Combinatorica **26** (2006), 475–487.

[5] T. Łuczak, Y. Peled, *Integral homology of random simplicial complexes*, `arXiv:1607.06985`

[6] R. Meshulam and N. Wallach, *Homological connectivity of random k-dimensional complexes*, Random Structures & Algorithms **34** (2009) 408–417.

High Dimensional Expansion

Roy Meshulam

(joint work with A. Lubotzky)

Expander graphs have been a focus of intensive research in the last four decades, with numerous applications throughout mathematics and theoretical computer science (see [2, 4]). In view of the ubiquity of expander graphs, there a recent growing interest in high dimensional notions of expansion. The k-dimensional version of the graphical Cheeger constant, called "coboundary expansion", came up independently in the work of Linial, Meshulam and Wallach [3, 6] on homological connectivity of random complexes and in Gromov's remarkable work [1] on the topological overlap property.

We proceed with the formal definitions. Let X be a finite n-dimensional pure simplicial complex. For $k \geq 0$, let $X^{(k)}$ denote the k-dimensional skeleton of X and let $X(k)$ be the family of k-dimensional faces of X, $f_k(X) = |X(k)|$. Define a positive weight function $w = w_X$ on the simplices of X as follows. For $\sigma \in X(k)$, let $c(\sigma) = |\{\eta \in X(n) : \sigma \subset \eta\}|$ and let

$$w(\sigma) = \frac{c(\sigma)}{\binom{n+1}{k+1} f_n(X)}.$$

Note that $\sum_{\sigma \in X(k)} w(\sigma) = 1$. Let $C^k(X)$ denote the space of $\mathbb{F}_2$-valued k-cochains of X with the coboundary map $d_k : C^k \to C^{k+1}$. The space of k-coboundaries of X is $B_k(X) = d_{k-1}C^{k-1}(X)$. For $\phi \in C^k(X)$, let $[\phi]$ denote the image of ϕ in $C^k(X)/B^k(X)$. Let

$$\|\phi\| = \sum_{\{\sigma \in X(k) : \phi(\sigma) \neq 0\}} w(\sigma)$$

and

$$\|[\phi]\| = \min\{\|\phi + d_{k-1}\psi\| : \psi \in C^{k-1}(X)\}.$$

Definition 1. The *k-th coboundary expansion constant* of X is

$$h_k(X) = \min\left\{\frac{\|d_k\phi\|}{\|[\phi]\|} : \phi \in C^k(X) - B^k(X)\right\}.$$

Clearly, $h_k(X) > 0$ if and only if $\tilde{H}_k(X;\mathbb{F}_2) = 0$. Indeed, $h_k(X)$ can be thought of as a measure of the resiliency of the property of having trivial k-dimensional $\mathbb{F}_2$-cohomology, under deletion of $(k+1)$-simplices from X.

As in the classical 1-dimensional case, one basic question concerns the existence of bounded degree higher dimensional expanders. Let $D_k(X)$ be the maximum number of $(k+1)$-dimensional faces of X containing a common k-dimensional face. A complex X is a (k,d,ϵ)-*expander* if

$$D_{k-1}(X) \le d \text{ and } h_{k-1}(X) \ge \epsilon.$$

In joint work with Lubotzky [5] we establish the existence of an infinite family of $(2,d,\epsilon)$-expanders for some fixed d and $\epsilon > 0$. Our proof is probabilistic and depends on the following new model, based on Latin squares, of random 2-dimensional simplicial complexes with bounded edge degrees. Let $\mathbb{S}_n$ be the symmetric group on $[n] = \{1,\ldots,n\}$. A k-tuple $(\pi_1,\ldots,\pi_k) \in \mathbb{S}_n^k$ is *legal* if $\pi_i\pi_j^{-1}$ is fixed point free for all $1 \le i < j \le k$. A *Latin Square* of order n is a legal n-tuple of permutations $L = (\pi_1,\ldots,\pi_n) \in \mathbb{S}_n^n$. Let $\mathcal{L}_n$ denote the uniform probability space of all Latin squares of order n. Let $V_1 = \{a_i\}_{i=1}^n, V_2 = \{b_i\}_{i=1}^n, V_3 = \{c_i\}_{i=1}^n$ be three disjoint sets. The complete 3-partite complex $T_n = V_1 * V_2 * V_3$ consists of all $\sigma \subset V = V_1 \cup V_2 \cup V_3$ such that $|\sigma \cap V_i| \le 1$ for $1 \le i \le 3$. An $L = (\pi_1,\ldots,\pi_n) \in \mathcal{L}_n$ determines a subcomplex $T_n^{(1)} \subset Y(L) \subset T_n$ whose 2-simplices are $[a_i, b_j, c_{\pi_i(j)}]$ where $1 \le i,j \le n$. In particular $Y(L)$ has $3n^2$ edges and every edge lies in a unique 2-simplex, i.e. $D_1(Y(L)) = 1$. Fix d and regard $\mathcal{L}_n^d$ as a uniform probability space. For $\underline{L}^d = (L_1,\ldots,L_d) \in \mathcal{L}_n^d$, let $Y(\underline{L}^d) = \cup_{i=1}^d Y(L_i)$. Note that $D_1(Y(\underline{L}^d)) \le d$. Let $\mathcal{Y}(n,d)$ denote the probability space of all complexes $Y(\underline{L}^d)$ with measure induced from $\mathcal{L}_n^d$.

Theorem 2 ([5])**.** *There exist $\epsilon > 0, d < \infty$ such that*

$$\lim_{n\to\infty} Pr[Y \in \mathcal{Y}(n,d) : h_1(Y) > \epsilon] = 1.$$

In [5] it is shown that one may take $\epsilon = 10^{-11}$ and $d = 10^{11}$. It would be of some interest to improve these constants. In particular, we conjecture that Theorem 2 holds with $d = 4$ and an appropriate $\epsilon > 0$.

References

[1] M. Gromov, *Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry*, Geom. Funct. Anal. **20** (2010), 416–526.

[2] S. Hoory, N. Linial and A. Wigderson, *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.) **43** (2006), 439–561.

[3] N. Linial and R. Meshulam, *Homological connectivity of random 2-complexes*, Combinatorica **26** (2006), 475–487.
[4] A. Lubotzky, *Expander graphs in pure and applied mathematics*, Bull. Amer. Math. Soc. (N.S.) **49** (2012), 113–162.
[5] A. Lubotzky and R. Meshulam, *Random Latin squares and* 2*-dimensional expanders*, Adv. Math. **272** (2015), 743–760.
[6] R. Meshulam and N. Wallach, *Homological connectivity of random* k*-dimensional complexes*, Random Struct. Algorithms **34** (2009), 408–417.

The sharp threshold for making squares

Rob Morris
(joint work with P. Balister and B. Bollobás)

Many of the fastest known algorithms for factoring large integers rely on finding subsequences of randomly generated sequences of integers whose product is a perfect square. Examples include Dixon's algorithm [2], the quadratic sieve [5], and the number field sieve (see, e.g., [4]); an excellent elementary introduction to the area is given by Pomerance [7]. In each of these algorithms one generates a sequence of congruences of the form

$$a_i \equiv b_i^2 \pmod{n}, \qquad i = 1, 2, \dots$$

and then one aims to find subsets of the a_i whose product is a perfect square, say $\prod_{i\in I} a_i = X^2$, so then one has $X^2 \equiv Y^2 \pmod{n}$ with $Y = \prod_{i\in I} b_i$. If one is lucky then $X \not\equiv \pm Y \pmod{n}$, in which case one can generate non-trivial factors of n as $\gcd(X \pm Y, n)$.

Motivated by this, Pomerance [6] posed in 1994 the problem of determining the threshold for the event that a sequence of random numbers $a_1, a_2, \dots$, each chosen independently and uniformly at random from the set $\{1, \dots, x\}$, contains a subset whose product is a square. To be precise, given $x \in \mathbb{N}$, let us define a random variable $T(x)$ by setting

$$T(x) := \min\left\{ N \in \mathbb{N} : \prod_{i\in I} a_i \text{ is a perfect square for some } I \subseteq \{1, \dots, N\},\ I \neq \emptyset \right\}.$$

Pomerance [8] proved that

$$T(x) = \exp\Big(\big(1 + o(1)\big) \sqrt{2 \log x \log\log x} \Big) \tag{1}$$

with high probability, and conjectured that $T(x)$ in fact exhibits a sharp threshold. Pomerance's conjecture remained wide open for over ten years, until a fundamental breakthrough was obtained by Croot, Granville, Pemantle and Tetali [1], who used a combination of techniques from number theory, probability theory and combinatorics to dramatically improve both the upper bound of Schroeppel and the lower bound of Pomerance [8], determining the location of the threshold to within a factor of $4/\pi$. To be precise, they proved that

$$\frac{\pi}{4}\big(e^{-\gamma} - \varepsilon\big) J(x) \leq T(x) \leq \big(e^{-\gamma} + \varepsilon\big) J(x) \tag{2}$$

with high probability, where[1]

$$J(x) = \min_{2 \le y \le x} \frac{\pi(y)x}{\Psi(x,y)},$$

and $\gamma \approx 0.5772$ is the Euler–Mascheroni constant.

Croot, Granville, Pemantle and Tetali [1] conjectured that the upper bound in (2) is sharp. Our main theorem confirms their conjecture.

Theorem 1. *For all $\varepsilon > 0$ we have with high probability*

$$\big(e^{-\gamma} - \varepsilon\big)J(x) \le T(x) \le \big(e^{-\gamma} + \varepsilon\big)J(x).$$

We remark that, in addition to proving the lower bound, we also obtain a new proof of the upper bound, quite different from that given in [1], as a simple consequence of our method. (We thank Jonathan Lee for pointing out to us a particularly simple and natural way of deducing this from our proof.)

Croot, Granville, Pemantle and Tetali [1] proved the lower bound in (2) via the first moment method, by counting the expected number of non-empty subsets $I \subseteq \{1, \dots, N\}$ such that $\prod_{i \in I} a_i$ is a square. Unfortunately, there exists a constant $c > 0$ such that this expected number blows up when $N \ge (e^{-\gamma} - c)J(x)$, which implies that a sharp lower bound cannot be obtained by this method.

Instead, we use the method of self-correcting martingales to follow a random process which removes numbers from the set $\{a_1, \dots, a_N\}$ as soon as we can guarantee that they are not contained in a subset whose product is a square. This is in one sense very simple: a number a_i can be discarded if there exists a prime for which a_i is the only remaining number that it divides an odd number of times. However, this apparent simplicity is deceiving, and the technical challenges involved in tracking the process are substantial. For example, we need to reveal the random numbers $\{a_1, \dots, a_N\}$ gradually (roughly speaking, prime by prime, in decreasing order), and the amount of information we are allowed to reveal at each step is rather delicate. Moreover, the removal of a number can trigger an avalanche, causing many other numbers to be removed in the same step. Fortunately, however, self-correction (which is partly a result of these avalanches) allows us to show that the process remains subcritical (in a certain natural sense), which in turn allows us to control the upper tail of the size of the avalanches. In order to do so, we need good control over the dependence between the prime factors of the numbers $\{a_1, \dots, a_N\}$ conditioned on the information we have observed so far; to obtain such control, we need some number-theoretic estimates, most of which follow from the fundamental work of Hildebrand and Tenenbaum [3] on smooth numbers.

Using the method described above, we are able to show that, with high probability, the number $m(z)$ of 'active' numbers (i.e., elements of $\{a_1, \dots, a_N\}$ that

[1] As usual, $\pi(y)$ denotes the number of primes less than or equal to y, and $\Psi(x,y)$ denotes the number of *y-smooth* integers in $\{1, \dots, x\}$, that is, the number of integers with no prime factor strictly greater than y.

we have not yet discarded) when we reveal the zth smallest prime, tracks the (deterministic) solution of the equation

$$m(z)\exp\Big(-\int_0^{m(z)/z}\frac{1-e^{-t}}{t}\,dt\Big)=\frac{\Psi(x,q_z)}{x}\cdot N \tag{3}$$

until there are very few numbers remaining (roughly $e^{-C\sqrt{\log y_0}}y_0$ for some large constant C), at which point we can apply the first moment calculation from [1]. The heuristic reason for the appearance of the formula (3) is that the number of y-smooth numbers is concentrated (e.g., by Chernoff's inequality) for all reasonably large values of y, since the a_i are chosen independently, and is equal to the number of isolated vertices in a certain natural (random) hypergraph. We show that the number $s_k(z)$ of edges of size k is self-correcting, and satisfies[2]

$$s_k(z)\in\left(1\pm\frac{\varepsilon^k\cdot k!}{\Lambda(z)}\right)\frac{m(z)}{k(k-1)}e^{-m(z)/z}\sum_{\ell=k-1}^{\infty}\frac{1}{\ell!}\left(\frac{m(z)}{z}\right)^{\ell}$$

in the 'critical range' $e^{O(\sqrt{\log y_0})}y_0$, and that these edges are chosen *almost* independently, so in particular the degree distribution of the random hypergraph is close to Poisson. Equating these two estimates for the number of isolated vertices gives (3).

Finally, in order to prove the upper bound in Theorem 1, we observe that the ratio of the number of active numbers and active primes (that is, primes which could still appear in some square) approaches 1 when $z=\pi(y_0)$ and $N/J(x)$ approaches $e^{-\gamma}$. Thus, by adding just a few extra y_0-smooth numbers, we can apply a simple linear algebra approach due to Schroeppel in order to obtain a subset whose product is a square, as required.

References

[1] E. Croot, A. Granville, R. Pemantle and P. Tetali, Sharp transitions in making squares, *Ann. Math,* **175** (2012), 1507–1550.

[2] J.D. Dixon, Asymptotically fast factorization of integers, *Math. Comp.* **36** (1981), 255–260.

[3] A. Hildebrand and G. Tenenbaum, On integers free of large prime factors. *Trans. Amer. Math. Soc.,* **296** (1986), 265–290.

[4] A.K. Lenstra and H.W. Lenstra Jr. (eds.), The development of the number field sieve, Lecture Notes in Math., **1554**, Springer–Verlag, Berlin and Heidelberg, 1993.

[5] C. Pomerance, Analysis and comparison of some integer factoring algorithms, *Computational Methods in Number Theory,* H.W. Lenstra, Jr. and R. Tijdeman (eds.), Math. Centre Tracts **154–155**, Mathematisch Centrum, Amsterdam, 1982, pp. 89–139.

[6] C. Pomerance, The role of smooth numbers in number theoretic algorithms, *Proc. Intern. Congr. Math.,* (Zurich, 1994), Birkhaüser, Basel, 1995, pp. 411–422.

[7] C. Pomerance, A Tale of Two Sieves, *Notices Amer. Math. Soc.* **43** (1996), 1473–1485.

[8] C. Pomerance, Multiplicative independence for random integers, *Analytic Number Theory: Proceedings of a Conference in Honor of Heini Halberstam,* B. Berndt, H. Diamond and A. Hildebrand (eds.), Birkhäuser, Boston, 1996, 703–711.

[2]Here $\Lambda(z)=J(x)\Psi(x,q_z)/xz$, so $\Lambda(z)\le 1$. It follows from the work of Hildebrand and Tenenbaum [3] that moreover $\Lambda(z)=e^{-c^2}+o(1)$ if $z=\pi(y_0)\exp\big(c\sqrt{\log y_0}\big)$.

Higher-order terms in Janson's inequality

Frank Mousset
(joint work with A. Noever, K. Panagiotou, and W. Samotij)

By an *increasing family* we will mean a triple (Ω, Γ, p) where Ω is a finite set, $\Gamma = \{\gamma_1, \dots, \gamma_N\}$ is a collection of non-empty subsets of Ω, and p is a map $\Omega \to (0,1)$. Given such a triple (Ω, Γ, p), we write Ω_p for the random subset of Ω that contains each $\omega \in \Omega$ with probability $p(\omega)$, independently of the others, and for $i \in [N]$ we set $X_i = \mathbf{1}_{\gamma_i \subseteq \Omega_p}$ and $X = X_1 + \cdots + X_N$. Our goal is to prove bounds on the probability that $X = 0$. Probabilities of this type arise commonly in probabilistic combinatorics. The classical result here is Janson's inequality (where the lower bound comes from the Harris inequality):

Theorem 1. *For every increasing family (Ω, Γ, p), we have*

$$\prod_{i \in [N]} (1 - \mathbf{E}[X_i]) \leq \Pr[X = 0] \leq e^{-\mu + \Delta}, \tag{1}$$

where

$$\mu = \mathbf{E}[X] \quad \textit{and} \quad \Delta = \sum_{i<j, \gamma_i \cap \gamma_j \neq \emptyset} \mathbf{E}[X_i X_j].$$

Typically, we will apply Theorem 1 in a situation where $\sum_{i \in [n]} \mathbf{E}[X_i]^2 = o(1)$. In this case, the lower bound is asymptotic to $e^{-\mu}$. Hence, unless $\Delta = o(1)$, the upper and lower bounds in (1) may be apart by an arbitrarily large factor. The reason is that the lower bound only looks at the first-order contribution μ, whereas the upper bound also includes the second-order contribution Δ, which may be non-negligible. To tighten (1), it is natural to look at upper and lower bounds that include higher-order contributions.

Before we can state such a result, we have to make some definitions. Let $A = \{X_1, \dots, X_m\}$ be a set of real-valued random variables on the same finite probability space. The *joint moment* of the elements in A is defined as

$$\Delta(A) = \mathbf{E}[X_1 \cdots X_m]. \tag{2}$$

The *joint cumulant* may be defined as

$$\kappa(A) = \sum_{\pi \in \Pi(A)} (|\pi| - 1)!(-1)^{|\pi|-1} \prod_{P \in \pi} \Delta(P), \tag{3}$$

where we write $\Pi(A)$ for the set of all partitions of A into non-empty sets. We also define the quantity

$$\delta(A) = \Delta(A) \cdot \max\{\mathbf{E}[X_i] \mid i \in [m]\} \tag{4}$$

that will play an important role below. We define the *dependency graph* G_Γ on the vertex set $[N]$, where $\{i, j\}$ is an edge if and only if $\gamma_i \cap \gamma_j \neq \emptyset$. Moreover, we write $\mathcal{C}_i$ for the collection of all i-element subsets $A \subseteq [N]$ such that the induced

subgraph $G_\Gamma[A]$ is connected. The elements of $\mathcal{C}_i$ are called *clusters of size* i. For $i \in \mathbf{N}$ we then define

$$\Delta_i = \sum_{A \in \mathcal{C}_i} \Delta(X_i \mid i \in A), \quad \delta_i = \sum_{A \in \mathcal{C}_i} \delta(X_i \mid i \in A), \quad \kappa_i = \sum_{A \in \mathcal{C}_i} \kappa(X_i \mid i \in A).$$

With these definitions, our main result reads as follows.

Theorem 2. *Let (Ω, Γ, p) be an increasing family. For every set $A \subseteq [N]$, denote by $N(A)$ the proper neighbourhood of A in the dependency graph G_Γ, and let*

$$\rho_A = \sum_{i \in N(A)} \mathbf{E}[X_i].$$

Let $k \in \mathbf{N}$. Assume there is some $\varepsilon > 0$ such that $\mathbf{E}[X_i] \leq 1 - \varepsilon$ for all $i \in [N]$ and $\rho_A \leq 1 - \varepsilon$ for all $A \subseteq [N]$ of size at most $k + 1$. Then there exists $K = K(k, \varepsilon)$ such that

$$\big|\log \Pr[X = 0] + \kappa_1 - \kappa_2 + \kappa_3 - \cdots \pm \kappa_k\big| \leq K(\delta_{1,K} + \Delta_{k+1,K}), \tag{5}$$

where

$$\delta_{1,K} = \sum_{i=1}^{K} \delta_i \quad \text{and} \quad \Delta_{k+1,K} = \sum_{i=k+1}^{K} \Delta_i.$$

For a better understanding of the error terms, we note that if

$$\max\{\mathbf{E}[X_i] \mid i \in [N]\} \to 0,$$

which is commonly the case, then we have $\Delta_i \sim \kappa_i$ for all i. The error term $\Delta_{k+1,K}$ morally corresponds to the error term Δ in Janson's inequality, and often its magnitude is proportional to κ_{k+1}. In these situations, this error of course cannot be avoided. The error $\delta_{1,K}$ comes from replacing products by exponentials; an error of this type is also necessary. While Theorem 2 does not formally imply Janson's inequality, in many situations it more powerful, as shown by the following examples.

Example 3. Let $\Omega = \{xy \mid 1 \leq x < y \leq n\}$ be the edge set of the complete graph with n vertices, $p(e) = p$ for all $e \in \Omega$, and Γ contain all triplets of edges that form a triangle, that is,

$$\Gamma = \big\{\{xy, yz, xz\} \mid 1 \leq x < y < z \leq n\big\}.$$

Then X is the number of triangles in the random graph $G_{n,p}$. From Janson's inequality (1) it easily follows that for $p = o(n^{-1/2})$,

$$\mathbf{P}[X = 0] \sim \exp\Big(-\frac{n^3 p^3}{6} + O(n^4 p^5)\Big) = \exp(-\kappa_1 + O(\kappa_2)).$$

Applying Theorem 2 we can show much better bounds, namely, that for all $p = o(n^{-1/2})$ and all $k \in \mathbf{N}$, we have

$$\mathbf{P}[X = 0] \sim \exp(-\kappa_1 + \kappa_2 - \cdots \pm \kappa_k + O(\kappa_{k+1})). \tag{6}$$

Here $p = o(n^{-1/2})$ is essentially the range where the sequence $\kappa_1, \kappa_2, \ldots$ is decreasing, i.e., the same statement holds trivially if $p = \omega(n^{-1/2})$. For $i \leq 4$ the

values of κ_i are not too hard to compute, and from (6) it follows for example that for $p = o(n^{-7/11})$, we have

$$\mathbf{P}[X = 0] \sim \exp\Big(- \frac{n^3p^3}{6} + \frac{n^4p^5}{4} - \frac{7n^5p^7}{12} + \frac{n^2p^3}{2} - \frac{3n^4p^6}{8} + \frac{27n^6p^9}{16}\Big).$$

This result was previously obtained by Stark and Wormald [1], extending an earlier result of Wormald [2].

Example 4. Let $\Omega = [n]$, let $p(\omega) = p$ for all $\omega \in \Omega$, and let Γ be the set of all r-term arithmetic progressions in Ω. Note that X counts the number of such arithmetic progressions contained in the random binomial subset $[n]_p$. Applying Janson's inequality (1) we can obtain that for $p = o(n^{-1/r})$, we have

$$\mathbf{P}[X = 0] \sim \exp(-\kappa_1 + O(\kappa_2)),$$

where one has $\kappa_1 \asymp n^2p^r$ and $\kappa_2 \asymp n^3p^{2r-1}$. Thus we see that this statement is meaningless already for $p = \omega(n^{-1/(r-1)})$, since then $\kappa_2 \gg \kappa_1$. Using Theorem 2 we can show that for all $p = o(n^{-1/(r-1)})$ and all $k \in \mathbf{N}$ we have

$$\mathbf{P}[X = 0] \sim \exp(-\kappa_1 + \kappa_2 - \cdots \pm \kappa_k + O(\kappa_{k+1})), \tag{7}$$

which is strictly better than the bounds given by Janson's inequality in this range. Again, $p = o(n^{-1/(r-1)})$ is the range where the sequence $\kappa_1, \kappa_2, \dots$ is decreasing. For example, performing the calculations for $r = 3$ and $k = 2$ one obtains that for $p = o(n^{-4/7})$,

$$\mathbf{P}[[n]_p \text{ is 3-AP-free}] \sim \exp\Big(- \frac{n^2p^3}{4} + \frac{17n^3p^5}{24}\Big).$$

References

[1] D. Stark and N. Wormald. The probability of nonexistence of a subgraph in a moderately sparse random graph. arXiv: `1608.05193 [math.CO]`.

[2] N. Wormald. The perturbation method and triangle-free random graphs. *Random Structures Algorithms* **9**(1–2) (1996), 253–270.

Estimating the rectilinear crossing number of dense graphs

János Pach

(joint work with J. Fox and A. Suk)

A *straight-line* drawing of a graph G is a mapping which assigns to each vertex a point in the plane and to each edge a straight-line segment connecting the corresponding two points. The *rectilinear crossing number* of a graph G, $\overline{\mathrm{cr}}(G)$, is the minimum number of crossing edges in any straight-line drawing of G. Determining or estimating $\overline{\mathrm{cr}}(G)$ appears to be a difficult problem, and deciding if $\overline{\mathrm{cr}}(G) \le k$ is known to be NP-hard. In fact, the asymptotic behavior of $\overline{\mathrm{cr}}(K_n)$ is still unknown.

We present a deterministic $n^{2+o(1)}$-time algorithm that finds a straight-line drawing of any n-vertex graph G with $\overline{\mathrm{cr}}(G) + o(n^4)$ crossing edges. Together with the well-known Crossing Lemma due to Ajtai et al. [1] and Leighton, this result

implies that for any dense n-vertex graph G, one can efficiently find a straight-line drawing of G with $(1 + o(1))\,\overline{\mathrm{cr}}(G)$ crossing edges. The proof is based on a geometric partition result [3] and on an algorithmic version of the Frieze-Kannan regularity lemma [4], [2].

References

[1] M. Ajtai, V. Chvátal, M. Newborn, and E. Szemerédi, *Crossing-free subgraphs*, in: Theory and practice of combinatorics, North Holland Math. Stud. **60** (1982), 9–12.

[2] D. Dellamonica, S. Kalyariasundaram, D. Martin, V. Rödl, and A. Shapria, *An optimal algorithm for finding Frieze-Kannan regular partitions,* Comb. Probab. Comp. **24** (2015), 407–437.

[3] J. Fox, J. Pach, and A. Suk, *Density and regularity theorems for semi-algebraic hypergraphs*, in: Proc. 26th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA '15). San Diego CA, 2015, 1517–1530.

[4] A. Frieze and R. Kannan, *Quick approximation to matrices and applications,* Combinatorica **19** (1999), 175–220.

Long-range order in random colorings and graph homomorphisms

Ron Peled

(joint work with Y. Spinka)

A (graph) homomorphism from a graph G to a graph H is a function f from the vertex set of G to that of H, which preserves edges, i.e., $f(u)$ and $f(v)$ are adjacent in H whenever u and v are adjacent in G. We investigate the typical properties of a uniformly chosen homomorphism from an induced finite subgraph of $\mathbb{Z}^d$ to an arbitrary finite graph H. We do not require H to be a simple graph, i.e., loops and multiple edges are allowed (though multiple edges make no difference to the model). We show that for a large class of graphs H, when the dimension d is sufficiently high, the random homomorphism exhibits a type of long-range order. The class of graphs for which the results apply includes several well-known models as special cases, amongst which are: proper q-colorings (obtained when H is a clique on q vertices), independent sets (also called the hard-core model; obtained when H is an edge with a loop at one endpoint) and the Widom-Rowlinson model (obtained when H is a path on three looped vertices). The results extend to non-uniform distributions, in which H is equipped with vertex and edge weights, $(\lambda_a)_{a\in V(H)}, (\lambda_{\{a,b\}})_{\{a,b\}\in E(H)}$, and the probability of choosing a particular homomorphism f is proportional to

$$\prod_{u\in V(G)} \lambda_{f(u)} \prod_{\{u,v\}\in E(G)} \lambda_{\{f(u),f(v)\}},$$

but we do not formulate these extensions here.

In order to formulate our results, we require some definitions. A *phase* of the graph H is a pair (A, B) of subsets of $V(H)$ such that a and b are adjacent for all $a \in A$ and $b \in B$. A phase is called *dominant* if it maximizes the quantity $|A| \cdot |B|$. We say that two phases (A, B) and (A', B') are *direct-equivalent* if there is an

automorphism of the graph H taking A to A' and B to B', and we say that they are *equivalent* if (A, B) is direct-equivalent to either (A', B') or (B', A'). Our results apply whenever all dominant phases are equivalent. Let us fix a bipartition of $\mathbb{Z}^d$ by saying that a vertex is *even* (*odd*) if its graph-distance from the origin is even (odd), that is, a vertex has the parity of the sum of its coordinates. For a positive integer n, let Λ_n^d denote the subgraph of $\mathbb{Z}^d$ induced by the cube $[-n, n]^d$ and let $\partial\Lambda_n^d := \Lambda_n^d \setminus \Lambda_{n-1}^d$ denote the vertex boundary of Λ_n^d. Given a homomorphism f from Λ_n^d to H, we say that a vertex $v \in \Lambda_n^d$ is *in the* (A, B)*-phase* if

$$v \text{ is even and } f(v) \in A \quad \text{or} \quad v \text{ is odd and } f(v) \in B.$$

We also say that a set of vertices is in the (A, B)-phase if all its elements are such.

Theorem 1. *Let H be a finite graph in which all dominant phases are equivalent. For any $\epsilon > 0$ there exists d_0 such that for any dimension $d \geq d_0$ and any $n \geq 1$, the following holds. Suppose that f is a uniformly sampled homomorphism from Λ_n^d to H. Then, for any dominant phase P and any vertex $v \in \Lambda_n^d$, we have*

$$\Pr\big(v \textit{ is in the } P\textit{-phase} \mid \partial\Lambda_n^d \textit{ is in the } P\textit{-phase}\big) \geq 1 - \epsilon.$$

The theorem states that a random homomorphism in high dimensions exhibits a strong type of long-range order in which, subject to (A, B)-boundary conditions, a chessboard-like (A, B)-pattern emerges within the box Λ_n^d, where any even vertex is likely to take a value in A and any odd vertex a value in B. Hence, any dominant phase defines "persistent" boundary conditions.

An interesting consequence from the viewpoint of statistical mechanics is the existence of multiple Gibbs states in such models (whenever (A, A) is not the unique dominant phase). Theorem 1 admits several extensions beyond the weighted version mentioned above. These include a quantitative dependence of ϵ and d_0 on H, and a version in which $\mathbb{Z}^d$ is replaced by a 'thickened' version of $\mathbb{Z}^2$, namely $\mathbb{Z}^2 \times \{0, 1\}^d$ (in a sense, the 'global dimension' of this graph is 2 and its 'local dimension' is $d + 2$). We do not discuss these extensions further here.

A result with similar features on the hypercube graph $G = \{0, 1\}^d$ was proved earlier by Engbers–Galvin [4]. Related counting results (in more general settings) consistent with Theorem 1 include a calculation by Meyerovitch–Pavlov [12] of the limiting entropy constant as the dimension tends to infinity and an estimate by Galvin-Tetali [8] on the number of homomorphisms from the the torus graph to an arbitrary finite graph (extending the entropy method of Kahn [9] who used it to count the number of independent sets in bipartite regular graphs, proving a conjecture of Alon [1]).

Let us briefly describe the result in the context of the three well-known models mentioned above.

Proper colorings. Let $H = K_q$ be the complete graph on $q \geq 2$ vertices. Then a homomorphism from G to H is precisely a *proper q-coloring* of G, i.e., a function $f\colon V(G) \to \{1, \ldots, q\}$ such that $f(u) \neq f(v)$ when u and v are adjacent in G. In this case, the dominant phases are all equipartitions of $[q] := \{1, \ldots, q\}$ into two sets, i.e., pairs (A, B) of disjoint subsets of $[q]$ such that $|A|, |B| \geq \lfloor q/2 \rfloor$.

In particular, Theorem 1 shows that when the dimension is sufficiently high as a function of q, any such equipartition defines persistent boundary conditions. This is a trivial fact for $q = 2$. It was previously known for $q = 3$ where it is due, independently, to Peled [13] and to Galvin–Kahn–Randall–Sorkin [7], with an extension to positive temperature (the weighted case) by Feldheim–Spinka [5] who proved the Kotecký conjecture [10]. The result here is novel for $q \geq 4$. We remark that the Dobrushin uniqueness condition [2] implies that when q is sufficiently large as a function of the dimension ($q > 4d$ suffices), the model ceases to exhibit long-range order.

Independent sets. Let H be the graph with vertex set $\{0, 1\}$ and edge set $\{\{0\}, \{0, 1\}\}$, i.e., a single edge with a loop at one endpoint. Then a homomorphism from G to H can be identified with an *independent set* in G, i.e., a subset of vertices I which contains no two adjacent vertices. In this setting it is common to consider a more general distribution, called the hard-core distribution, in which an independent set I is chosen with probability proportional to $\lambda^{|I|}$, where $\lambda > 0$ is a parameter called the fugacity. In this case, there are two dominant phases, $(\{0\}, \{0, 1\})$ and $(\{0, 1\}, \{0\})$, and an extension of Theorem 1 to the weighted case shows that when the dimension is sufficiently high as a function of the fugacity λ, there are two persistent boundary conditions, one in which I consists mostly of even vertices and one in which it consists mostly of odd vertices. The fact that this is the case for sufficiently large fugacity as a function of the dimension $d \geq 2$ goes back to Dobrushin [3]. In fact, our results show that the minimal fugacity for which this behavior occurs tends to zero as the dimension tends to infinity. This result was previously known and is due to Galvin–Kahn [6], with a quantitative improvement by Peled-Samotij [14]. As in the case of proper colorings, the Dobrushin uniqueness condition [2] implies that when λ is sufficiently small as a function of the dimension ($\lambda \leq 1/2d$ suffices), the model ceases to exhibit long-range order.

The Widom-Rowlinson model. Let H be the graph with vertex set $\{-1, 0, 1\}$ and edge set $\{\{0\}, \{\pm 1\}, \{0, \pm 1\}\}$, i.e., a path on three looped vertices. Then one may regard a homomorphism from G to H as describing the territory occupied by two competing species (one represented by "1" and the other by "-1") which cannot be adjacent to one another. In this case, there are two dominant phases, $(\{0, 1\}, \{0, 1\})$ and $(\{-1, 0\}, \{-1, 0\})$, and Theorem 1 shows that when the dimension is sufficiently high, there are two persistent boundary conditions, one in which the "1" species is predominant and the other in which "-1" is predominant. Such a result was previously known only in the weighted case, by Lebowitz-Gallavotti [11], when the presence of "1" and "-1" in the configuration is encouraged by sampling configurations with probability proportional to λ raised to the number of "1" and "-1", and λ is taken sufficiently high as a function of the dimension.

References

[1] Noga Alon, *Independent sets in regular graphs and sum-free subsets of finite groups*, Israel Journal of Mathematics **73** (1991), no. 2, 247–256.

[2] Roland L. Dobrushin, *Description of a random field by means of conditional probabilities and conditions for its regularity* (Russian), Teor. Verojatnost. i Primenen **13** (1968) 201–229.

[3] Roland L. Dobrushin, *The problem of uniqueness of a Gibbsian random field and the problem of phase transitions*, Functional analysis and its applications **2** (1968), no. 4, 302–312.

[4] John Engbers and David Galvin, *H-coloring tori*, Journal of Combinatorial Theory, Series B **102** (2012), no. 5, 1110–1133.

[5] Ohad Feldheim and Yinon Spinka, *Long-range order in the 3-state antiferromagnetic Potts model in high dimensions*, arXiv preprint arXiv:1511.07877.

[6] David Galvin and Jeff Kahn, *On phase transition in the hard-core model on* $\mathbf{Z}^d$, Combinatorics, Probability and Computing **13** (2004), no. 02, 137–164.

[7] David Galvin, Jeff Kahn, Dana Randall, and Gregory Sorkin, *Phase coexistence and torpid mixing in the 3-coloring model on* $\mathbf{Z}^d$, SIAM J. Discrete Math. 29 (2015), no. 3, 1223–1244.

[8] David Galvin and Prasad Tetali, *On weighted graph homomorphisms*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **63** (2004), 97–104.

[9] Jeff Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combinatorics, Probability and Computing **10** (2001), no. 03, 219–237.

[10] Roman Kotecký, *Long-range order for antiferromagnetic Potts models*, Physical Review B **31** (1985), no. 5, 3088–3092.

[11] Joel L. Lebowitz and Giovanni Gallavotti, *Phase transitions in binary lattice gases*, Journal of Mathematical Physics **12** (1971), no. 7, 1129–1133.

[12] Tom Meyerovitch and Ronnie Pavlov, *On independence and entropy for high-dimensional isotropic subshifts*, Proc. Lond. Math. Soc. (3) **109** (2014) no. 4.

[13] Ron Peled, *High-dimensional Lipschitz functions are typically flat*, arXiv preprint arXiv:1005.4636, To appear in Annals of Probability.

[14] Ron Peled and Wojciech Samotij, *Odd cutsets and the hard-core model on* $\mathbb{Z}^d$, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 50, Institut Henri Poincaré, 2014, pp. 975–998.

An occupancy approach to bounding graph polynomials

Will Perkins

(joint work with E. Davies, M. Jenssen, and B. Roberts)

Many graph polynomials (the independence polynomial, matching polynomial, and indirectly, the Tutte and chromatic polynomials) can be viewed as *partition functions* of statistical physics models. We present a new method for proving extremal results for graph polynomials over classes of bounded degree graphs using the perspective of statistical physics.

The hard-core model at fugacity λ on a graph G is a random independent set I drawn according to the distribution

$$\mathbb{P}(I) = \frac{\lambda^{|I|}}{Z_G(\lambda)}.$$

The function $Z_G(\lambda) = \sum_I \lambda^{|I|}$ is the partition function of the hard-core model, or in graph theory, the independence polynomial. Evaluating $Z_G(1)$ counts the total number of independent sets of G.

Our approach is to maximize (or minimize) the *free energy*

$$F_G(\lambda) = |V(G)|^{-1} \log Z_G(\lambda)$$

of the hard-core model and other models from statistical physics by maximizing or minimizing the *occupancy fraction*,

$$\alpha_G(\lambda) := \lambda \cdot F_G'(\lambda) = |V(G)|^{-1}\mathbb{E}[|I|],$$

which is both the (scaled) derivative of the free energy and a probabilistic *observable* of the model. By integrating $\alpha_G(t)/t$ from 0 to λ, a bound on α_G yields a bound on the free energy. For instance, in [2] we show

Theorem 1. *For all d-regular G and all $\lambda > 0$,*

$$\alpha_G(\lambda) \leq \alpha_{K_{d,d}}(\lambda).$$

This is a strengthening of results of Kahn [6], Galvin and Tetali [5], and Zhao [9] proving via the entropy method that the free energy $F_G(\lambda)$ is maximized over d-regular graphs by $K_{d,d}$ (or to put it differently, the number of independent sets, and the independence polynomial are maximized over d-regular graphs by disjoint unions of $K_{d,d}$'s).

Our general method for proving such theorems can be summarized as follows.

- Choose a Gibbs measure (e.g. the hard-core model), an observable (the occupancy fraction), and a class of bounded-degree graphs (d-regular graphs).
- For a fixed depth t, consider the two-part experiment of choosing a configuration (independent set) σ from the Gibbs measure on G, choosing a vertex v uniformly at random from $V(G)$ and recording both the depth-t neighborhood of v as well as the boundary condition of the Gibbs measure induced by σ. Call the neighborhood plus boundary condition the random *local view* from the experiment.
- Relax the extremal problem on graphs to an optimization problem on probability distributions on local views, adding carefully chosen constraints that must hold for all distributions induced by graphs in the chosen class.
- Pose the relaxation as a linear program, solve it, and integrate the bound on the observable to obtain a bound on the free energy.

By restricting to triangle-free graphs and minimizing instead of maximizing, the method yields the following result [3].

Theorem 2. *For any triangle-free graph G of maximum degree at most d,*

$$\alpha_G(1) \geq (1 + o_d(1))\log d/d.$$

Theorem 2 provides an alternative proof to Shearer's classic result [8] that gives the best known upper bound on the Ramsey number $R(3,k)$. Recall that $\alpha_G(1) \cdot |V(G)|$ is the average size of a uniformly random independent set from G. The theorem begs the question of whether the maximum size of an independent set in a triangle-free graph is significantly larger than the average size.

Conjecture 3. *In any triangle-free graph the ratio of the maximum independent set size to the average independent set size is at least* $4/3$*. In any triangle-free graph of minimum degree* d*, the ratio is at least* $(2 - o_d(1))$*.*

A proof of Conjecture 3 would yield a respective factor $4/3$ or 2 improvement on the current upper bound on $R(3, k)$.

We can also apply the method to matchings in regular graphs. Let M be a random matching from a graph G chosen with probability $\mathbb{P}(M) = \frac{\lambda^{|M|}}{M_G(\lambda)}$ where $M_G(\lambda) = \sum_M \lambda^{|M|}$ is the Monomer-Dimer partition function, or the matching generating function. Bregman's theorem [1] tells us that the Monomer-Dimer free energy is maximized in the $\lambda \to \infty$ limit by $K_{d,d}$, but the result for finite λ (and the counting case $\lambda = 1$) was not known previously. In [2] we prove:

Theorem 4. *For all* d*-regular* G *and all* $\lambda > 0$*,*

$$\alpha_G^M(\lambda) \leq \alpha_{K_{d,d}}^M(\lambda),$$

and in particular,

$$|V(G)|^{-1} \log M_G(\lambda) \leq (2d)^{-1} \log M_{K_{d,d}}(\lambda).$$

The method can also be applied to the anti-ferromagnetic Potts model on cubic graphs which yields in the zero-temperature limit a tight bound on the number of proper q-colorings of a cubic graph [4].

Finally, we can obtain a richer class of extremal regular graphs if we forbid certain local structures. Examples of such results include the following from [7].

Theorem 5. *Let* $P_{5,2}$ *denote the Petersen graph. Then for every* 3*-regular, triangle-free graph* G*, and every* $\lambda \in (0, 1]$*,*

$$\alpha_G(\lambda) \geq \alpha_{P_{5,2}}(\lambda).$$

Let $H_{3,6}$ *denote the Heawood graph, the unique* $(3, 6)$*-cage graph. Then for every* 3*-regular graph* G *of girth at least* 5*, and all* $\lambda > 0$*,*

$$\alpha_G(\lambda) \leq \alpha_{H_{3,6}}(\lambda).$$

Integrating the occupancy fraction yields a tight bound on the independence polynomial and the normalized number of independent sets.

Further discussion of extremal problems for regular graphs can be found in the recent survey of Zhao [10]

References

[1] L. Bregman, *Some properties of nonnegative matrices and their permanents*, Soviet Math. Dokl **14** (1973), 945–949.

[2] E. Davies, M. Jenssen, W. Perkins, B. Roberts, *Independent sets, matchings, and occupancy fractions*, arXiv preprint arXiv:1508.04675 (2015).

[3] E. Davies, M. Jenssen, W. Perkins, B. Roberts, *On the average size of independent sets in triangle-free graphs*, arXiv preprint arXiv:1606.01043 (2016).

[4] E. Davies, M. Jenssen, W. Perkins, B. Roberts, *Extremes of the internal energy of the Potts model on cubic graphs*, arXiv preprint arXiv:1610.08496 (2016).

[5] D. Galvin, P. Tetali, *On weighted graph homomorphisms*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **63** (2004), 97–104.
[6] J. Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combinatorics, Probability and Computing, **10** (2001), 219–237. 2001.
[7] G. Perarnau, W. Perkins, *Counting independent sets in cubic graphs of given girth*, arXiv preprint arXiv:1611.01474 (2016).
[8] J. Shearer, *A note on the independence number of triangle-free graphs*, Discrete Mathematics **46**(1983), 83–87.
[9] Y. Zhao, *The number of independent sets in a regular graph*, Combinatorics, Probability and Computing **19** (2010), 315–320.
[10] Y. Zhao, *Extremal regular graphs: independent sets and graph homomorphisms*, arXiv preprint arXiv:1610.09210 (2016).

Hamiltonian cycles in 3-uniform hypergraphs

Christian Reiher
(joint work with V. Rödl, A. Ruciński, M. Schacht, E. Szemerédi)

G. A. Dirac [6] proved that every graph $G = (V, E)$ on at least 3 vertices and with minimum vertex degree $\delta(G) \geq |V|/2$ contains a Hamiltonian cycle. This result is best possible, as there are graphs G with minimum degree $\delta(G) = \lceil |V|/2 \rceil - 1$ not containing a Hamiltonian cycle.

We continue the study to which extent Dirac's theorem can be generalised to hypergraphs. Here we shall restrict to 3-uniform hypergraphs and if not mentioned otherwise by a hypergraph we will mean a 3-uniform hypergraph. Note that in this case there are at least two natural concepts of a minimum degree condition and several notions of cycle, and we briefly introduce some of them below.

For a hypergraph $H = (V, E)$ and a vertex $v \in V$ we denote by $d_H(v)$ *the degree of* v defined as

$$d_H(v) = \big|\{e \in E \colon v \in e\}\big|,$$

and by $\delta(H) = \min d_H(v)$ the *minimum vertex degree* of H taken over all $v \in V$. Similarly, for any two vertices u, $v \in V$ we denote by $d_H(u, v)$ their *pair degree*, which is the number of edges containing u and v, i.e.,

$$d_H(u, v) = \big|N_H(u, v)\big| = \big|\{e \in E \colon u, v \in e\}\big|,$$

and let $\delta_2(H) = \min d_H(u, v)$ be the *minimum pair degree* over all pairs of vertices of H.

An early notion of cycles in hypergraphs appeared in the work of Berge [1] (see, also [2]) more than 40 years ago. More recently, Katona and Kierstead [9] considered the following types of paths and cycles.

A hypergraph P is a *tight path of length* ℓ, if $|V(P)| = \ell + 2$ and there is an ordering of the vertices $V(P) = \{x_1, \ldots, x_{\ell+2}\}$ such that a triple e forms a hyperedge of P if and only if $e = \{x_i, x_{i+1}, x_{i+2}\}$ for some $i \in [\ell]$. The ordered pairs (x_1, x_2) and $(x_{\ell+1}, x_{\ell+2})$ are the *end-pairs* of P and we say that P is a tight (x_1, x_2)-$(x_{\ell+1}, x_{\ell+2})$ path. This definition of end-pairs is not symmetric and implicitly fixes a direction on P and the order of the end-pairs. Hence, we may refer to (x_1, x_2) as the *starting pair* and to $(x_{\ell+1}, x_{\ell+2})$ as the *ending pair*. All

other vertices of P are called *internal*. We sometimes identify such a path P with the sequence of its vertices $x_1 \dots x_{\ell+2}$. Moreover, a *tight cycle* C *of length* $\ell \geq 4$ consists of a path $x_1 \dots x_\ell$ of length $\ell - 2$ and the two additional hyperedges $\{x_{\ell-1}, x_\ell, x_1\}$ and $\{x_\ell, x_1, x_2\}$. In both cases the *length* of a tight cycle and of a tight path is measured by the number of hyperedges and we will use the same convention for the length of cycles, paths, and walks in graphs. For simplicity we denote edges and hyperedges by xy and xyz instead of $\{x, y\}$ and $\{x, y, z\}$.

Roughly speaking, one may think of tight paths and cycles as ordered hypergraphs such that "consecutive" edges overlap in exactly two vertices. Similarly, one may consider so-called *loose* paths and cycles, where the overlap is restricted to one vertex only. The optimal approximate minimum pair and vertex degree conditions for the existence of loose Hamiltonian cycles were obtained in [3, 10] and precise versions for large hypergraphs appeared in [5, 8].

Results on pair degree conditions implying tight Hamiltonian cycles were obtained in [15, 16]. For minimum vertex degrees, $(5/9 - o(1))n^2/2$ provides a lower bound (see Examples 2 1 – 3 below), which was conjectured to be optimal. So far only suboptimal upper bounds were obtained in [7, 13, 14]. We close this gap here, as the following result provides an asymptotically optimal minimum vertex degree condition for tight Hamiltonian cycles.

Theorem 1. *For every $\alpha > 0$ there exists an integer n_0 such that every 3-uniform hypergraph H with $n \geq n_0$ vertices and with minimum vertex degree*

$$\delta(H) \geq \left(\tfrac{5}{9} + \alpha\right)\tfrac{n^2}{2}$$

contains a tight Hamiltonian cycle.

A recent result of Cooley and Mycroft [4] establishes the existence of an almost spanning tight cycle under the same degree condition as in Theorem 1. Moreover, both these results are asymptotically best possible, as the following well known examples show.

Example 2.

(1) Consider a partition $X \dot\cup Y = V$ of a vertex set V of size n with $|X| = \lceil (n+1)/3 \rceil$ and let H be the hypergraph containing all triples $e \in V^{(3)}$ such that $|e \cap X| \neq 2$. It is not hard to show that H contains no tight Hamiltonian cycle, since two consecutive vertices in X cannot be connected to Y (see, e.g., [13]). Moreover, we have $\delta(H) \geq (5/9 - o(1))n^2/2$.

(2) Similarly, one may consider a partition $X \dot\cup Y = V$ with $|X| = \lceil 2n/3 \rceil$ and let H be the hypergraph consisting of all triples $e \in V^{(3)}$ such that $|e \cap X| \neq 2$. Again H has $\delta(H) \geq (5/9 - o(1))n^2/2$ and it contains no tight Hamiltonian cycle.

(3) The last example utilises the fact that every tight Hamiltonian cycle contains a matching of size $\lfloor n/3 \rfloor$. Again we consider a partition $X \dot\cup Y = V$ this time with $|X| = \lfloor n/3 \rfloor - 1$ and let H consist of all triples having at least one vertex in X. Consequently, H contains no matching of

size $\lfloor n/3 \rfloor$ and, hence, no tight Hamiltonian cycle. On the other hand, $\delta(H) \geq (5/9 - o(1))n^2/2$.

The proof of Theorem 1 may be found in [11].

References

[1] C. Berge, *Graphs and hypergraphs*, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. Translated from the French by Edward Minieka; North-Holland Mathematical Library, Vol. 6. MR0357172 (50 #9640)

[2] J.-C. Bermond, A. Germa, M.-C. Heydemann, and D. Sotteau, *Hypergraphes hamiltoniens*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 39–43 (French, with English summary). MR539937 (80j:05093)

[3] E. Buß, H. Hàn, and M. Schacht, *Minimum vertex degree conditions for loose Hamilton cycles in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **103** (2013), no. 6, 658–678, DOI 10.1016/j.jctb.2013.07.004. MR3127586

[4] O. Cooley and R. Mycroft, *The minimum vertex degree for an almost-spanning tight cycle in a 3-uniform hypergraph*, available at `1606.05616`. Submitted.

[5] A. Czygrinow and T. Molla, *Tight codegree condition for the existence of loose Hamilton cycles in 3-graphs*, SIAM J. Discrete Math. **28** (2014), no. 1, 67–76, DOI 10.1137/120890417. MR3150175

[6] G. A. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc. (3) **2** (1952), 69–81. MR0047308 (13,856e)

[7] R. Glebov, Y. Person, and W. Weps, *On extremal hypergraphs for Hamiltonian cycles*, European J. Combin. **33** (2012), no. 4, 544–555, DOI 10.1016/j.ejc.2011.10.003. MR2864440

[8] J. Han and Y. Zhao, *Minimum vertex degree threshold for loose Hamilton cycles in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **114** (2015), 70–96, DOI 10.1016/j.jctb.2015.03.007. MR3354291

[9] Gy. Y. Katona and H. A. Kierstead, *Hamiltonian chains in hypergraphs*, J. Graph Theory **30** (1999), no. 3, 205–212, DOI 10.1002/(SICI)1097-0118(199903)30:3<205::AID-JGT5>3.3.CO;2-F. MR1671170 (99k:05124)

[10] D. Kühn and D. Osthus, *Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree*, J. Combin. Theory Ser. B **96** (2006), no. 6, 767–821, DOI 10.1016/j.jctb.2006.02.004. MR2274077 (2007h:05115)

[11] C. Reiher, V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi, *Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs*, available at `1611.03118`. Submitted.

[12] V. Rödl and A. Ruciński, *Dirac-type questions for hypergraphs—a survey (or more problems for Endre to solve)*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 561–590, DOI 10.1007/978-3-642-14444-8_16. MR2815614 (2012j:05008)

[13] ———, *Families of triples with high minimum degree are Hamiltonian*, Discuss. Math. Graph Theory **34** (2014), no. 2, 361–381, DOI 10.7151/dmgt.1743. MR3194042

[14] V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi, *On the Hamiltonicity of triple systems with high minimum degree*, Ann. Comb. To appear.

[15] V. Rödl, A. Ruciński, and E. Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combin. Probab. Comput. **15** (2006), no. 1-2, 229–251, DOI 10.1017/S0963548305007042. MR2195584 (2006j:05144)

[16] ———, *Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs*, Adv. Math. **227** (2011), no. 3, 1225–1299, DOI 10.1016/j.aim.2011.03.007. MR2799606 (2012d:05213)

[17] Y. Zhao, *Recent advances on Dirac-type problems for hypergraphs*, available at `1508.06170`.

Incidence geometry, rank bounds for design matrices, and applications

Shubhangi Saraf
(joint work with A. Basit, Z. Dvir, and C. Wolf)

Given any set P of n points in $\mathbb{R}^d$ and suppose we know that many small subsets of the points are linearly dependent. Can we deduce a global bound on the dimension of the entire configuration? The oldest example of such a question dates back to 1893 when Sylvester, motivated by certain complex configurations arising from Elliptic curves, asked "If every line through two points in P contains a third point from P, must they all line on the same line?" [11]. The positive answer to this question, proved independently by Melchior [9] and by Gallai (answering a question of Erdos [6]), is now known as the Sylvester-Gallai (SG) theorem. This is a strong form of the local-to-global principle. Loosely speaking, it says that the existence of 'many' local dependencies (collinearity of triples is captured by affine dependence) imposes a global structure - here all points must lie in 1-dimensional affine space.

There have been several variations of this question that have been studied such as high dimensional, colorful, approximate and quantitative versions of the Sylvester-Gallai theorem, as well as their analogs over finite fields. In recent years, theorems of this nature have been useful in studying certain structural questions arising in theoretical computer science. Variants of the SG theorem were useful in understanding the structure of low-depth arithmetic circuits and the polynomial identity testing problem (PIT) [4, 8, 10]. Quantitative versions of the SG theorem were shown to be closely linked to the structure of linear Locally Correctable Codes [1, 3, 5].

This talk surveys some of these extensions, with particular emphasis on a proof technique that has recently proven to be quite powerful for proving strong bounds for several of the incidence questions, which is via rank bounds for design matrices. We will demonstrate the power of this technique by focusing on the proof of the following recent result [2].

Consider a set of n points in $\mathbb{R}^d$. The classical theorem of Sylvester-Gallai says that, if the points are not all collinear then there must be a line through exactly two of the points. Let us call such a line an "ordinary line". In a recent result, Green and Tao were able to give optimal linear lower bounds (roughly n/2) on the number of ordinary lines determined n non-collinear points in $\mathbb{R}^d$.

We will consider the analog over the complex numbers. While the Sylvester-Gallai theorem as stated above is known to be false over the field of complex numbers, it was shown by Kelly [7] that for a set of n points in $\mathbb{C}^d$, if the points don't all lie on a 2-dimensional plane then the points must determine an ordinary line. Using techniques developed for bounding the rank of design matrices, we will show that such a point set must determine at least 3n/2 ordinary lines, except in the trivial case of n-1 of the points being contained in a 2 dimensional plane.

References

[1] A. Bhattacharyya, Z. Dvir, A. Shpilka, and S. Saraf. Tight lower bounds for 2-query lccs over finite fields. In *Proc. of FOCS 2011*, pages 638–647, 2011.

[2] Abdul Basit, Zeev Dvir, Shubhangi Saraf, and Charles Wolf. On the number of ordinary lines determined by sets in complex space. In *manuscript*.

[3] B. Barak, Z. Dvir, A. Wigderson, and A. Yehudayoff. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 519–528, New York, NY, USA, 2011. ACM.

[4] Zeev Dvir and Amir Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2006.

[5] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Improved rank bounds for design matrices and a new proof of kelly's theorem. In *Forum of Mathematics, Sigma*, volume 2, page e4. Cambridge Univ Press, 2014.

[6] P. Erdos. Problems for solution: 4065–4069. *American Mathematical Monthly*, 1943.

[7] L. M. Kelly. A resolution of the sylvester - gallai problem of j. -p. serre. *Discrete & Computational Geometry*, 1:101–104, 1986.

[8] Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *FOCS '09: Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 198–207, Washington, DC, USA, 2009. IEEE Computer Society.

[9] E. Melchior. Uber vielseite der projektive ebene. *Deutsche Math.*, 5:461–475, 1940.

[10] Nitin Saxena and C. Seshadhri. From sylvester-gallai configurations to rank bounds: Improved black-box identity test for depth-3 circuits. *Foundations of Computer Science, Annual IEEE Symposium on*, 0:21–29, 2010.

[11] J. J. Sylvester. Mathematical question 11851. *Educational Times*, 59:98, 1893.

Removal Lemmas with Polynomial Bounds

Asaf Shapira
(joint work with L. Gishboliner)

A hereditary graph property $\mathcal{P}$ is said to be *testable* if there is a function $f_{\mathcal{P}}(\varepsilon)$ such that for every $\varepsilon > 0$ and for every graph G, if G is ε-far from satisfying $\mathcal{P}$ then a random subset of $f_{\mathcal{P}}(\varepsilon)$ vertices of G does not satisfy $\mathcal{P}$ with probability at least $\frac{2}{3}$. $\mathcal{P}$ is *easily testable* if one can take $f_{\mathcal{P}}(\varepsilon)$ to be polynomial in ε^{-1}. For a fixed graph H, the testability if H-freeness is equivalent to the celebrated graph removal lemma of Ruzsa and Szemerédi [9]. It was shown by Alon and Shapira [6] that every hereditary graph property is testable. This general result relied on a version of Szemerédi's regularity lemma [10] and thus gave tower-type bounds on $f_{\mathcal{P}}(\varepsilon)$. Consequently, several authors [4, 8] raised the problem of characterizing the easily testable hereditary graph properties.

We prove new sufficient and necessary conditions for a hereditary graph property to be easily testable. These general combinatorial conditios imply almost all previously known results of this type, as well as many new ones. As a corollary, we prove a conjecture of Alon, stating that every semi-algebraic graph property is easily testable.

For a hereditary graph property $\mathcal{P}$, let $\mathcal{F}(\mathcal{P})$ denote the family of minimal forbidden induced subgraphs for $\mathcal{P}$. Our first result gives a sufficient condition that guarantees that $\mathcal{P}$ is easily testable.

Theorem 1. *If $\mathcal{F}(\mathcal{P})$ is finite and contains a bipartite graph, a co-bipartite graph and a split graph, then $\mathcal{P}$ is easily testable.*

Theorem 1 implies the previously known results that H-freeness is easily testable for a bipartite H [1] and that induced H-freness is easily testable if H is a path with at most four vertices [4, 5], as well as the new results that the properties of being a line graph, being a threshold graph and being a trivially perfect graph are easily testable.

Our next result is a necessary condition for a hereditary property to be easily testable. It only applies when $\mathcal{F}(\mathcal{P})$ is finite.

Theorem 2. *If $\mathcal{F}(\mathcal{P})$ is finite and $\mathcal{P}$ is easily testable then $\mathcal{F}(\mathcal{P})$ contains a bipartite graph and a co-bipartite graph.*

Theorem 2 implies as corollaries the negative results of [1] and [5]. Our final result for the case that $\mathcal{F}(\mathcal{P})$ is finite is that the necessary condition in Theorem 2 is not sufficient.

Theorem 3. *There is a bipartite graph F_1 and a co-bipartite graph F_2 for which induced $\{F_1, F_2\}$-freeness is not easily testable.*

We now turn to consider the case where $\mathcal{F}(\mathcal{P})$ is infinite. It is natural to ask if the condition of Theorem 1 is still sufficient. Our next theorem states that it is not.

Theorem 4. *There is a graph family $\mathcal{F}$ which contains a bipartite graph, a co-bipartite graph and a split graph, for which induced $\mathcal{F}$-freeness is not easily testable.*

The family $\mathcal{F}$ in Theorem 4 is an ad hoc construction and does not arise from a familiar graph property. We prove that for many "natural" properties, the condition in Theorem 1 is sufficient. To state the exact result we need the following definition. Let H be a graph with vertices $V(H) = \{1, \ldots, r\}$ and let $g : V(H) \to \{0, 1\}$. A graph G is a *g-blowup* of H if G admits a vertex partition $V(G) = V_1 \cup \cdots \cup V_r$ such that (V_i, V_j) is a complete (resp. empty) bipartite graph whenever $(i, j) \in E(H)$ (resp. $(i, j) \notin E(H)$) and V_i is a clique (resp. an independent set) whenever $g(i) = 1$ (resp. $g(i) = 0$).

Theorem 5. *Let $\mathcal{P}$ be a hereditary graph property satisfying the following.*

(1) *$\mathcal{F}(\mathcal{P})$ contains a bipartite graph, a co-bipartite graph and a split graph.*
(2) *For every $H \in \mathcal{P}$ there is a function $g : V(H) \to \{0, 1\}$ such that every g-blowup of H satisfies $\mathcal{P}$.*

Then $\mathcal{P}$ is easily testable.

In order to describe the main application of Theorem 5, we recall the definition of *semi-algebraic graph properties*. A semi-algebraic property is given by a set

of real $2k$-variate polynomials $f_1, \ldots, f_t \in \mathbb{R}[x_1, \ldots, x_{2k}]$ and a Boolean function $\Phi : \{\text{true}, \text{false}\}^t \to \{\text{true}, \text{false}\}$. A graph G satisfies the property if there is an assignment of points $\{p_v : v \in V(G)\}$ to the vertices of G, where $p_v \in \mathbb{R}^k$, such that $(u, v) \in E(G)$ if and only if

$$\Phi\Big(f_1(p_u, p_v) \geq 0; \ldots; f_t(p_u, p_v) \geq 0\Big) = \text{true}.$$

In the expression $f_i(p_u, p_v)$, we substitute p_u into the first k variables of f_i and p_v into the last k variables of f_i. Many well studied classes of graphs, such as interval graphs, unit disc graphs and circular arc graphs, are semi-algebraic. We proved that every semi-algebraic property satisfies conditions 1-2 of Theorem 5. Thus, we got the following corollary, which confirms a conjecture of Alon [2].

Corollary 6. *Every semi-algebraic graph property is easily testable.*

The proofs of Theorems 1 and 5 rely on a lemma stating that $\mathcal{F}(\mathcal{P})$ contains a bipartite graph, a co-bipartite graph and a split graph if and only if there is $k = k(\mathcal{P})$ such that every $G \in \mathcal{P}$ has VC-dimension at most k. Next we use a theorem by Alon, Fishcer and Newman [3], stating that if only a small fraction of the induced subgraphs of G of order $\text{poly}(1/\varepsilon)$ have VC-dimension larger than k, then G admits a highly-structured (more structured than the one produced by the regularity lemma) partition with only $\text{poly}(1/\varepsilon)$ parts. In the proof of Theorem 5 there is an additional difficulty that arises from having to handle infinitely many forbidden subgraphs. Indeed, there can be a situation in which G is ε-far from satisfying $\mathcal{P}$ but the smallest $F \in \mathcal{F}(\mathcal{P})$ which is an induced subgraph of G is of size much larger than $\text{poly}(1/\varepsilon)$. In fact, such a construction is used in the proof of Theorem 4. Condition 2 of Theorem 5 allows us to overcome this difficulty, as well as other technical difficulties that arise in proofs of this type.

In order to prove 2 and 3 we prove the following new hardness result.

Theorem 7. *For every non-bipartite H, $\varepsilon > 0$ and $n \geq n_0(\varepsilon)$ there is an n-vertex graph which is ε-far from being induced H-free but contains only $\varepsilon^{c \log(1/\varepsilon)} n^{v(H)}$ (not necessarily induced) copies of H.*

The proof of Theorem 7 relies on a new variant of the Ruzsa-Szemerédi construction. Instead of using a set of integers with no solution to $x_1 + \cdots + x_k = kx_{k+1}$, as was done in previous works, we use a set with no solution to *any* equation of the form $a_1x_1 + \ldots + a_kx_k = (a_1 + \ldots + a_k)x_{k+1}$ with positive bounded $a_1, \ldots, a_k$. Furthermore, we carefully label the vertices of H in order to guarantee that every copy of H in the construction will contain a cycle which corresponds to a solution to one of the above equations.

References

[1] N. Alon, Testing subgraphs in large graphs, Random Structures and Algorithms 21 (2002), 359–370.
[2] N. Alon, private communication, 2013.
[3] N. Alon, E. Fischer, and I. Newman. Testing of bipartite graph properties. SIAM Journal on Computing 37 (2007), 959–976.

[4] N. Alon and J. Fox, Easily testable graph properties, Combin. Probab. Comput. 24 (2015), 646–657.
[5] N. Alon and A. Shapira, A characterization of easily testable induced subgraphs, Combin. Probab.Comput. 15 (2006), 791–805.
[6] N. Alon and A. Shapira, A characterization of the (natural) graph properties testable with one-sided error, SIAM Journal on Computing 37 (2008), 1703–1727.
[7] N. Alon and A. Shapira, Every monotone graph property is testable, SIAM Journal on Computing 38 (2008), 505–522.
[8] O. Goldreich, Contemplations on testing graph properties, ECCC Technical Report 2005. Also, Studies in Complexity and Cryptography 6650 (2011), 547–554.
[9] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai 18, Volume II, 939–945.
[10] E. Szemerédi, Regular partitions of graphs, In: *Proc. Colloque Inter. CNRS.*

Core Forging by Warning Propagation

Kathrin Skubch
(joint work with A. Coja-Oghlan, O. Cooley, and M. Kang)

The *k-core* of a graph is the largest subgraph of minmum degree k. It can be determined algorithmically by the *peeling process* that removes an arbitrary vertex of degree less than k while there is one. In one of the most influential contributions to the theory of random graphs Pittel, Spencer and Wormald analysed the peeling process on the Erdős-Rényi random graph via the method of differential equations [6]. They determined the precise threshold d_k from where on the k-core is non-empty w.h.p as well as the asymptotic order (i.e., number of vertices) and size (number of edges) of the k-core for $d > d_k$ and $k \geq 3$. The case $k \geq 3$ is very different from the case $k = 2$, as the 2-core simply emerges continuously (mostly) inside the giant component. By contrast, a most remarkable feature of the k-core problem for $k \geq 3$, first observed by Łuczak [3, 4], is that the order of the k-core leaps from 0 to a linear the number of vertices the very moment the k-core becomes non-empty.

Since the seminal work of Pittel, Spencer and Wormald several alternative proofs of the k-core threshold have been developed. Some of these proofs extend to hypergraphs and/or given degree sequences. Additionally, establishing a bivariate central limit theorem, Janson and Luczak [2] also studied the joint limiting distribution of the order and size of the k-core.

While one way or another all the previous work on the k-core problem has been based on tracing the peeling process in the random graph, in the present paper we develop a very different approach. Instead of analysing the peeling process, we devise a probabilistic scheme for constructing graphs with a k-core of a given order and size directly. Formally, we develop a randomised sampling algorithm `Forge` that produces a graph with a core of a given desired order and size. The output distribution of `Forge` converges in total variation to the distribution of an Erdős-Rényi random graph given the order and size of the k-core. Consequently, because the randomised construction employed by `Forge` is surprisingly simple, we can immediately read off the asymptotic number of graphs with a k-core of a

given order and size. As an application, we obtain a bivariate local limit theorem for the distribution of these parameters in the Erdős-Rényi random graph. This result sharpens the central limit theorem of Janson and Luczak [2]. Additionally, the sampling algorithm completely elucidates the way the k-core is embedded into the random graph, a question on which we obtained partial results in an earlier paper via the formalism of local weak convergence [1].

Our randomised algorithm `Forge` is based on an enhanced "configuration model" that explicitly designates which vertices will wind up in the core. Setting up such a model is anything but straightforward because the core interacts with the rest of the graph, sometimes called the "mantle", in a rather complicated way.

The key idea of the present paper is to investigate not merely the set of vertices and edges in the k-core but to also look at the "surrounding structure" of the graph from the right angle. As it turns out, the necessary additional structure can be set out concisely by way of *Warning Propagation*, a message passing scheme that plays an important role in physics work on random constraint satisfaction problems [5]. Remarkably, the Warning Propagation messages contain additional information that help to unlock the k-core problem.

References

[1] A. Coja-Oghlan, O. Cooley, M. Kang, K. Skubch: How does the core sit inside the mantle? Random Structures and Algorithms, in press.

[2] S. Janson, M. Luczak: Asymptotic normality of the k-core in random graphs Ann. Appl. Probab. **18** (2008) 1085–1137.

[3] T. Łuczak: Size and connectivity of the k-core of a random graph. Discrete Math. **91** (1991) 61–68.

[4] T. Łuczak: Sparse random graphs with a given degree sequence. in A. Frieze and T. Łuczak (eds.): Random Graphs 2, John Wiley & Sons (1992) 165–182.

[5] M. Mézard, A. Montanari: Information, physics and computation. Oxford University Press 2009.

[6] B. Pittel, J. Spencer, N. Wormald, *Sudden emergence of a giant k-core in a random graph*, Journal of Combinatorial Theory, Series B **67** (1996), 111–151

An application of the polynomial method

József Solymosi

(joint work with J. Ellenberg and J. Zahl)

The polynomial method is an important and powerful method with various applications in number theory, discrete geometry, and combinatorics. In this talk we gave a brief introduction and then sketched the proof of the theorems below. For the interested reader we recommend the excellent lecture note by Larry Guth [1].

As an application of the method we established new bounds on the number of tangencies and orthogonal intersections determined by an arrangement of curves. First, given a set of n algebraic plane curves, we showed that there are $O(n^{3/2})$ points where two or more curves are tangent. In particular, if no three curves are mutually tangent at a common point, then there are $O(n^{3/2})$ curve-curve

tangencies. Second, given a family of algebraic plane curves and a set of n curves from this family, we showed that either there are $O(n^{3/2})$ points where two or more curves are orthogonal, or the family of curves has certain special properties.

We obtained these bounds by transforming the arrangement of plane curves into an arrangement of space curves so that tangency (or orthogonality) of the original plane curves corresponds to intersection of space curves. We then bounded the number of intersections of the corresponding space curves. For the case of curve-curve tangency, we used a polynomial method technique that is reminiscent of Guth and Katz's proof of the joints theorem [2]. For the case of orthogonal curve intersections, we employed a bound of Guth and Zahl to control the number of two-rich points in space curve arrangements [3].

References

[1] L. Guth. Polynomial methods in combinatorics. *AMS University Lecture Series* Volume: 64; 2016; 273 pp

[2] L. Guth and N. Katz. On the Erdős distinct distance problem in the plane. *Ann. of Math.*, 181:155–190, 2015.

[3] L. Guth and J. Zahl. Algebraic curves, rich points, and doubly-ruled surfaces. *arXiv:1503.02173*, 2015.

Generalized Turán Numbers

Jacques Verstraëte

(joint work with M. Molloy and B. Sudakov)

The *generalized Turán number* $\mathrm{ex}(G,F)$ of two graphs G and F is the maximum number of edges in an F-free subgraph of G. Define $\mathrm{dx}(G,F)$ to be the largest minimum degree amongst all F-free spanning subgraphs of G. The determination of the Turán numbers $\mathrm{ex}(K_n,F)$ is a central problem in combinatorics. In the case that F has chromatic number $r+1\geq 2$, then for any graph G, $\mathrm{ex}(G,F)\geq (1-\frac{1}{r})e(G)$, by taking a random r-partition of the vertices of G. By Turán's Theorem, $\mathrm{ex}(K_n,F)\sim (1-\frac{1}{r})\binom{n}{2}$, and so for non-bipartite F, this determines the asymptotic minimum value of $\mathrm{ex}(G,F)$ given the number of edges in G. The focus in this paper is on the case that F is bipartite and contains a cycle. The generalized Turán numbers $\mathrm{ex}(G,F)$ have been studied at length for specific graphs G, in particular when $G=Q_n$, the boolean cube graph (see [2] and the references therein), when G is a C_{2k}-free graph (see Györi [5] and Kühn and Osthus [7]), and in random graphs (see Bohman and Keevash [1], Schacht [8] and Conlon and Gowers [3] and the references therein) and pseudorandom graphs [6].

We focus in this paper on the case that G has prescribed minimum and maximum degree. If G is a $\triangle$-regular graph, a recurring theme is whether $\mathrm{ex}(G,F)$ and $\mathrm{dx}(G,F)$ achieve their minimum values (exactly or asymptotically or in order of magnitude as $\triangle\to\infty$) when G is a disjoint union of cliques of order $\triangle+1$. Along these lines, the following conjecture was made in [4]:

Conjecture 1. *If F and G are graphs and $\triangle(G) = \triangle \to \infty$, then*

$$\mathrm{dx}(G, F) = \Omega\Big(\frac{\mathrm{dx}(K_\triangle, F)}{\triangle}\Big) \cdot \delta(G).$$

Partial results were obtained when $F = C_{2k}$ for $k \in \{2, 3, 5\}$ by Foucaud, Krivelevich and Perarnau [4]. To state our main theorem, we require the following definitions. If F and H are graphs and $f : V(F) \to V(H)$ is a homomorphism, then f is called a *local isomorphism* if $|\{f(u)f(v) : uv \in E(F)\}| = |E(F)|$. In other words, the map $\phi : E(F) \to E(H)$ defined by $\phi(uv) = f(u)f(v)$ for $uv \in E(F)$ is a bijection. For a graph F, let $\mathcal{H}_F$ denote the family of graphs H such that there is a local isomorphism $f : V(F) \to V(H)$. Our main theorem is as follows:

Theorem 2. *Let F be a graph containing a cycle. Then for any graph G of maximum degree $\triangle \geq 1$ such that $\mathrm{dx}(K_\triangle, \mathcal{H}_F) \cdot \delta(G) \geq 10^{50}\triangle \log \triangle$,*

$$\mathrm{dx}(G, F) \geq \frac{\mathrm{dx}(K_\triangle, \mathcal{H}_F)}{10^5 \triangle} \cdot \delta(G).$$

Furthermore, for all $\triangle \geq 1$,

$$\mathrm{ex}(G, F) \geq \frac{\mathrm{ex}(2\triangle, \mathcal{H}_F)}{2e\triangle^2} \cdot e(G).$$

Theorem 2 solves Conjecture 1 whenever $\mathrm{dx}(K_n, \mathcal{H}_F) = \Theta(\mathrm{dx}(K_n, F))$, which may even be true for all graphs F.

Conjecture 3. *For any graph F, show that $\mathrm{dx}(K_n, F) = \Theta(\mathrm{dx}(K_n, \mathcal{H}_F))$.*

Focusing on the case of regular graphs, we do not have an example of a family of $\triangle$-regular graphs $G_\triangle$ for which $\mathrm{dx}(G_\triangle, F) \sim \mathrm{dx}(K_\triangle, F)$ as $\triangle \to \infty$, so we pose the following problem: determine those graphs F for which there exists an $\epsilon > 0$ such that for all large enough $\triangle$ and there exists a $\triangle$-regular graph $G_\triangle$ such that

$$\mathrm{dx}(G_\triangle, F) < (1 - \epsilon)\mathrm{dx}(K_\triangle, F).$$

This is open even in the case $F = C_4$.

One of the interesting ingredients of the proof of Theorem 2 involves injective proper colorings: an *injective proper coloring* of a graph G is a proper coloring of the square graph G^2 i.e. the graph obtained by joining any pair of vertices at distance at most two in G. We prove the following result:

Theorem 4. *Let G be a graph of maximum degree $\triangle$, and let J be any graph on $5\triangle$ vertices such that $2\delta(J)\delta(G) \geq 10^{50}\triangle \log \triangle$. Then there is a spanning subgraph H of G with an injective proper J-coloring such that for all $v \in V(H)$,*

$$d_H(v) \geq \frac{d_J(\chi(v))\, d_G(v)}{10000\triangle}.$$

In particular, if G is $\triangle$-regular, this shows that even though G^2 may have chromatic number of order $\triangle^2$, one can find a spanning subgraph of G of minimum degree $\Theta(\triangle)$ whose square graph has chromatic number $\Theta(\triangle)$. In other words, G^2 may be moderately thinned out to reduce the chromatic number to $\Theta(\triangle)$.

References

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

[2] D. Conlon, An extremal theorem in the hypercube, Electron. J. Combin. 17 (2010), R111 (7 pages).

[3] D. Conlon, W. T. Gowers, Combinatorial theorems in sparse random sets, Ann. of Math. 184 (2016), 367-454.

[4] F. Foucaud, M. Krivelevich and G. Perarnau, Large subgraphs without short cycles. SIAM Journal on Discrete Mathematics 29 (2015), 65–78.

[5] E. Győri, C_6-free bipartite graphs and product representation of squares. Discrete Mathematics 165/166, 1997, 371–375.

[6] Y. Kohayakawa, V. Rödl, M. Schacht, S. Sisshoko, Turán's theorem for pseudo-random graphs, Journal of Combinatorial Theory (A) 114(4), 2007, 631–657.

[7] D. Kühn and D. Osthus, 4-cycles in graphs without a given even cycle, J. Graph Theory 48 (2005), 147–156.

[8] M. Schacht, Extremal results for random discrete structures, Ann. of Math. 184 (2), (2016), 333–365.

[9] C. Thomassen, Girth in graphs. J. Combin. Theory Ser. B. v35 i2. 129–141.

Upper tails for arithmetic progressions

Yufei Zhao

(joint work with B. Bhattacharya, S. Ganguly, and X. Shao)

Let X_k denote the number of k-term arithmetic progressions in a random subset of $\{1, 2, \dots, n\}$ where every element is included with probability p. We are interested in estimating the upper tail probability $\mathbb{P}(X_k \geq (1+\delta)\mathbb{E}X_k)$.

It may be helpful to compare with the analogous problem in random graphs. For a fixed graph H, let X_H denote the number of copies of H in an Erdős–Rényi random graph $G(n,p)$.

The problem of estimating the upper tail of these random variables has a long history, and it has been called the "infamous upper tail" [8]. Even the order of the log-probability had not been known until relatively recently. In the case of triangle counts in a random graph, the order of the log-probability was determined by DeMarco and Kahn [5], and independently, by Chatterjee [2].

Theorem 1 (Chatterjee/DeMarco–Kahn)**.** *Fix $\delta > 0$. If $p \gtrsim \log n/n$, then*

$$\mathbb{P}(X_{K_3} \geq (1+\delta)\mathbb{E}X_{K_3}) = p^{\Theta_\delta(n^2p^2)}.$$

DeMarco and Kahn [6] subsequently extended this result to cliques. The corresponding result for arithmetic progressions was proved recently by Warnke [10].

Theorem 2 (Warnke)**.** *Fix $\delta > 0$ and $k \geq 3$. If $p \gtrsim (\frac{\log n}{n})^{1/(k-1)}$, Then*

$$\mathbb{P}(X_k \geq (1+\delta)\mathbb{E}X_k) = p^{\Theta_k(\sqrt{\delta n^2 p^k})}.$$

In our work, we are interested in determining the missing constant in the exponent. See Chatterjee's recent survey [3] for a discussion of recent developments in this direction.

Chatterjee and Dembo [4] devised a new *nonlinear large deviation principle* that reduces such certain problems (random graphs, 3-term progressions) to natural variational problems, though its applicability is restricted to $p \geq n^{-\alpha}$ for some fixed $\alpha > 0$ in each case. Very recently, Eldan [7] proved a new nonlinear large deviation principle, using completely different methods, that appears to be somewhat quantitatively more efficient than [4].

The corresponding variational problem was solved in [9] for triangle counts, and more generally whenever H is a clique. This result was subsequently extended to all H in [1]. Together with the nonlinear large deviation principles, we now know the following.

Theorem 3. *Fix $\delta > 0$ and graph H with maximum degree Δ. There is some constant $\alpha_H > 0$ such that whenever $p = p_n \to 0$ with $p \geq n^{-\alpha_H}$, one has*

$$\mathbb{P}(X_H \geq (1+\delta)\mathbb{E}X_H) = p^{(1+o(1))c_H(\delta)n^2p^\Delta},$$

where $c_H(\delta) > 0$ is some explicit constant.

For example, when $H = K_3$, we have

$$c_{K_3}(\delta) = \min\{\tfrac{1}{2}\delta^{2/3}, \tfrac{1}{3}\delta\}.$$

See [1] for the formula for $c_H(\delta)$ for general H.

Our main new result is a corresponding theorem for X_k, the number of k-term arithmetic progressions in a random subset of $\{1, \dots, n\}$.

Theorem 4. *Fix $\delta > 0$ and integer $k \geq 3$. There is some constant $\alpha_k > 0$ such that whenever $p = p_n \to 0$ with $p \geq n^{-\alpha_k}$, one has. With $p \to 0$, one has*

$$\mathbb{P}(X_k \geq (1+\delta)\mathbb{E}X_k) = p^{(1+o(1))\sqrt{\delta n^2 p^k}}.$$

The best known α_k is currently being improved in work in progress, and we do not state it here.

References

[1] B. B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao, *Upper tails and independence polynomials in random graphs*, arXiv:1507.04074.

[2] S. Chatterjee, *The missing log in large deviations for triangle counts*, Random Structures Algorithms **40** (2012), 437–451.

[3] S. Chatterjee, *An introduction to large deviations for random graphs*, Bull. Amer. Math. Soc. **53** (2016), 617–642.

[4] S. Chatterjee and A. Dembo, *Nonlinear large deviations*, Adv. Math. **299** (2016), 396–450.

[5] B. DeMarco and J. Kahn, *Upper tails for triangles*, Random Structures Algorithms **40** (2012), 452–459.

[6] B. DeMarco and J. Kahn, *Tight upper tail bounds for cliques*, Random Structures Algorithms **41** (2012), 469–487.

[7] R. Eldan, Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations, arXiv:1612.04346.

[8] S. Janson and A. Ruciński, *The infamous upper tail*, Random Structures Algorithms **20** (2002), 317–342.

[9] E. Lubetzky and Y. Zhao, *On the variational problem for upper tails in sparse random graphs*, Random Structures Algorithms, to appear.
[10] L. Warnke, *Upper tails for arithmetic progressions in random subsets*, Israel J. Math., to appear.

Problem Session

(1) (Günter M. Ziegler) Describe the set of integers $m \geq 2$ for which there exists a matrix $A(t) \in \{0, 1, t\}^{(n+1)\times(n+1)}$ such that $\det(A(t)) \not\equiv 0$ and $\det(A(m)) = 0$, where "describe" could mean find its size and its maximum value.

(2) (Anders Björner) For any function $w\colon \binom{[n]}{2} \to \mathbb{R}$ and for any $G \subseteq \binom{[n]}{2}$, define $w(G) = \sum_{e\in G} w(e)$.

Conjecture. *If $n \geq 8$ then for any $w\colon \binom{[n]}{2} \to \mathbb{R}$ with $w(\binom{[n]}{2}) = 0$, the number of spanning trees $T \subset \binom{[n]}{2}$ such that $w(T) \geq 0$ is at least $2n^{n-3}$.*

Note that the above conjecture is equivalent to the statement that the w which minimizes the number of spanning trees T with $w(T) \geq 0$ is given by assigning a huge negative weight to one edge, and all the rest positive (so that nonnegative weight spanning trees are precisely those that avoid the negative edge).

(3) (Noga Alon)

Conjecture. *For all $p_1, \ldots, p_n \in \mathbb{R}^n$ with $\|p_1\| = \cdots = \|p_n\| = 1$, there exist $q_1, \ldots, q_n \in \mathbb{R}^{n/2}$ such that for all i, j*

$$|\langle p_i, p_j\rangle - \langle q_i, q_j\rangle| \leq \frac{1000}{\sqrt{n}}.$$

(4) (Boris Bukh) *Degeneracy of random subgraph.*
Let G be a graph of minimum degree D. Let $G_{1/2}$ be a random subgraph of G obtained by keeping each edge of G with probability $1/2$ independently of the other edges. Let d be the largest integer such that $G_{1/2}$ contains a non-empty subgraph of minimum degree d. How large must d be in term of D?
Trivialities:
(a) If $G = K_{D+1}$, then $d = D/2 - o(D)$ by Chernoff's bound.
(b) Since G has average degree $\geq D$, the average degree of $G_{1/2}$ is $\geq D/2 - o(D)$ implying that $d \geq D/4 - o(D)$, for any G.

Surprisingly, there is a better construction than the complete graph. Namely, if G is a $(d/3)$-blowup of a large 3-regular expander, then $d \leq D/3 + o(D)$.

Conjecture. *$d \geq D/3 - o(D)$ for any G.*

(5) (Ron Peled & Jeff Kahn) *Positive association for colorings.*
Let G be a finite bipartite graph with parts A and B. Let $q \geq 3$ and f be a uniformly sampled proper q-coloring of G. Does the set $f^{-1}(0) \cap A$ have positive association? Precisely, if $X, Y : \{0,1\}^A \to \mathbb{R}$ are increasing, is

$$\mathrm{E}[X(f|_A^{-1}(0))Y(f|_A^{-1}(0))] \geq \mathrm{E}[X(f|_A^{-1}(0))] \cdot \mathrm{E}[Y(f|_A^{-1}(0))]?$$

Even in the case of 4 vertices $u, v, x, y \in A$, we don't know if

$$\mathrm{P}[f(u) = f(v) = f(x) = f(y) = 0] \geq \mathrm{P}[f(u) = f(v) = 0]\mathrm{P}[f(x) = f(y) = 0].$$

$\Delta + 1$ colors.
Suppose that G is a finite graph with max degree Δ and $q = \Delta + 1$. Let f be a uniformly sampled proper q-coloring of G. For two vertices $x, y \in G$, does the correlation of the events $f(x) = 0$ and $f(y) = 0$ go to zero as the distance between x and y increases, uniformly in the choice of graph G and vertices x and y?

(6) (József Solymosi) Let A be an $n \times n$. $R_i = \sum_{j:j\neq i} |A_{i,j}|$ is the radius of the Gershgorin disk centered at $A_{i,i}$. The radius of the half Gershgorin disk r_i is defined as the sum of the $\lfloor n/2 \rfloor$ largest terms of $\{|A_{i,j}| : j \neq i\}$.

Theorem (Bárány, Solymosi)**.** *If A is a real non-negative valued matrix, and λ is an eigenvalue with geometric multiplicity ≥ 2, then λ lies in a half Gershgorin disk, $D(A_{i,i}, r_i)$ for some i.*

Q1: What about for eigenvalues with higher multiplicity?
Q2: What about for positive semidefinite or general (complex) matrices?

(7) (János Pach) Are pseudosegment disjointness graphs χ-bounded?

(8) (Nati Linial) For any matrix $A \in \{0,1\}^{n\times n}$ with no $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ as a submatrix, the maximium number of 1s in A is $\Theta(n^{3/2})$. Analogously, define $f(n)$ to be the maximum number of 1s, taken over all 3D matrices $A \in \{0,1\}^{n\times n\times n}$ having no $2\times 2\times 2$ subcube consisting of all 1s. It is known that $\Omega(n^{8/3}) \leq f(n) \leq O(n^{11/4})$, but the correct order of magnitude isn't known.

(9) (Swastik Kopparty) What is $\mathrm{P}[\mathrm{per}(A) = 0]$ if A is an $n \times n$ random matrix over $\mathbb{F}_3$?

Reporters: Igor Balla & Frank Mousset

Participants

Prof. Dr. Karim Adiprasito
Einstein Institute of Mathematics
The Hebrew University of Jerusalem
Edmond J. Safra Campus
Givat Ram
Jerusalem 91904
ISRAEL

Prof. Dr. Noga Alon
Department of Mathematics
Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv 69978
ISRAEL

Igor Balla
Department of Mathematics
ETH Zürich
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Jozsef Balogh
Department of Mathematics
University of Illinois at
Urbana-Champaign
1409 West Green Street
Urbana IL 61801
UNITED STATES

Prof. Dr. Alexander Barvinok
Department of Mathematics
University of Michigan
2074 East Hall
530 Church Street
Ann Arbor, MI 48109-1043
UNITED STATES

Prof. Dr. Anders Björner
Department of Mathematics
Royal Institute of Technology
100 44 Stockholm
SWEDEN

Dr. Boris Bukh
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Maria Chudnovsky
Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000
UNITED STATES

Prof. Dr. Amin Coja-Oghlan
Institut für Mathematik
Goethe-Universität Frankfurt
Robert-Mayer-Straße 6-10
60325 Frankfurt am Main
GERMANY

Dr. David Conlon
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Dr. Asaf Ferber
Department of Mathematics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307
UNITED STATES

Prof. Dr. Jacob Fox
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
UNITED STATES

Prof. Dr. Ehud Friedgut
Department of Mathematics
The Weizmann Institute of Science
P. O. Box 26
Rehovot 76100
ISRAEL

Prof. Dr. Zoltan Furedi
Alfred Renyi Institute of Mathematics
Hungarian Academy of Sciences
P.O.Box 127
1364 Budapest
HUNGARY

Prof. Dr. David Gamarnik
Massachusetts Institute of Technology
Sloan School of Management
E 53 - BS 7
77 Massachusetts Avenue
Cambridge MA 02139-4307
UNITED STATES

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

Prof. Dr. Jeff Kahn
Department of Mathematics
Rutgers University
Piscataway NJ 08854-8019
UNITED STATES

Prof. Dr. Gil Kalai
Institute of Mathematics
The Hebrew University
Givat-Ram
Jerusalem 91904
ISRAEL

Prof. Dr. Mihyun Kang
Institut für Diskrete Mathematik
Technische Universität Graz
Steyrergasse 30
8010 Graz
AUSTRIA

Prof. Dr. Peter Keevash
Mathematical Institute
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Swastik Kopparty
Department of Mathematics
Rutgers University
Busch Campus, Hill Center
New Brunswick, NJ 08854-8019
UNITED STATES

Prof. Dr. Daniel Kral
Department of Mathematics
University of Warwick
Coventry CV4 7AL
UNITED KINGDOM

Prof. Dr. Michael Krivelevich
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv 69978
ISRAEL

Dr. Daniela Kühn
School of Mathematics and Statistics
The University of Birmingham
Edgbaston
Birmingham B15 2TT
UNITED KINGDOM

Prof. Dr. Nathan Linial
School of Computer Science & Engineering
The Hebrew University
Givat Ram
Jerusalem 91904
ISRAEL

Dr. Po-Shen Loh
Department of Mathematical Sciences
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Tomasz Luczak
Wydzial Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza
Umultowska 87
61-614 Poznań
POLAND

Dr. Roy Meshulam
Department of Mathematics
Technion -
Israel Institute of Technology
Haifa 32000
ISRAEL

Richard H. Montgomery
Trinity College
Cambridge CB2 1TQ
UNITED KINGDOM

Prof. Dr. Robert Morris
IMPA
Instituto Matematica Pura e Aplicada
Estrada Dona Castorina, 110
Jardim Botanico
Rio de Janeiro 22460-320
BRAZIL

Frank Mousset
Institut für Theoretische Informatik
ETH Zürich
CAB
8092 Zürich
SWITZERLAND

Dr. Deryk Osthus
School of Mathematics and Statistics
The University of Birmingham
Edgbaston
Birmingham B15 2TT
UNITED KINGDOM

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

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

Dr. Ron Peled
Department of Mathematics
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 69978
ISRAEL

Dr. Will Perkins
School of Mathematics and Statistics
The University of Birmingham
Birmingham B15 2TT
UNITED KINGDOM

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

Dr. Wojciech Samotij
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv
Tel Aviv 69978
ISRAEL

Dr. Shubhangi Saraf
Department of Mathematics
and Computer Science
Rutgers University
Busch Campus, Hill Center
New Brunswick, NJ 08854-8019
UNITED STATES

Lisa Sauermann
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
UNITED STATES

Prof. Dr. Mathias Schacht
Fachbereich Mathematik
Universität Hamburg
Bundesstrasse 55
20146 Hamburg
GERMANY

Prof. Dr. Alexander Schrijver
Centrum voor Wiskunde en Informatica
University of Amsterdam
Science Park 123
1098 XG Amsterdam
NETHERLANDS

Prof. Dr. Asaf Shapira
Department of Mathematics
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 69978
ISRAEL

Kathrin Skubch
Institut für Mathematik
Goethe-Universität Frankfurt
Robert-Mayer-Strasse 6-10
60325 Frankfurt am Main
GERMANY

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

Prof. Dr. Angelika Steger
Institut für Theoretische Informatik
ETH Zürich
CAB G 38
Universitätsstrasse 6
8092 Zürich
SWITZERLAND

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

Dr. Yufei Zhao
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

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

Prof. Dr. Günter M. Ziegler
Institut für Mathematik
Freie Universität Berlin
Arnimallee 2
14195 Berlin
GERMANY

Dr. Jacques A. Verstraete
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES