Final report

Dániel Gerbner

During this project I have studied several loosely connected problems. First I would like to list those in the area of extremal set theory. In what follows the underlying set is always $[n] = \{1, \ldots, n\}$. I am going to use both the expressions family and hypergraph for subsets of the power set of $[n]$.

Together with Máté Vizer we studied the so-called tilted Sperner families. These were defined by Leader and Long [19]. A family $\mathcal{F}$ is called (p, q)-tilted Sperner if there are no distinct $F, G \in \mathcal{F}$ with $p|F \setminus G| = q|G \setminus F|$. They determined asymptotically the maximum size of such a family in case p and q are co-prime. Long [20] added the additional property (pattern) that a pair (F, G) is only forbidden if in addition to the previous property $f > g$ for all $f \in F$ and $G \in G$ (here we use that the elements of the underlying set are numbers). This is an unusual restriction, usually the elements of the underlying set are treated equally. He examines the maximum size of a family satisfying these properties, and gives the upper bound $O(e^{120\sqrt{\log n}}2^n/\sqrt{n})$. We have improved the bound to $O(\sqrt{\log n}2^n/\sqrt{n})$ in [16]. Our proof is based on a new approach to the permutation method together with standard probabilistic arguments. We show that for every set there is a so-called (p, q)-cut point: an element x such that p/q times the number of points of A that are smaller than x is close to the number of points not belonging to A that are larger than x. With probabilistic arguments we show that for most members of $\mathcal{F}$ the cut point is close to $pn/(p + q)$, and the permutation method gives that for any given point x there cannot be too many sets in $\mathcal{F}$ with x as cut point.

With Abhishek Methuku and Casey Tompkins we studied the following problem. We are given a poset P and we want to find the largest family $\mathcal{F}$ not containing P as a (not necessary induced) subposet, such that $\mathcal{F}$ is also intersecting. Without the intersecting property this is a well-known and well-studied question. With the intersecting property the only known results were for Sperner familes [21] and k-Sperner families [6]. We give bounds for general posets in [12], using ideas from [4]. As a corollary, we get an infinite family of posets where the bound is sharp, at least for odd n. We also give a sharp bound for the case P is the butterfly poset, using an advanced version of Katona's cycle method. The result is the same that we get for 2-Sperner families. We show that even though there might be members of the family that are in the middle of a 3-chain, we can delete them if we count the isolated members with a larger weight, where isolated means they do not appear in any chain of length more than 1. Furthermore, we prove some LYM-type inequalities for intersecting butterfly-free families.

With Fabricio Benevides, Cory Palmer and Dominik Vu we considered the problem of finding the smallest d-separating family consisting of sets of size at most k. A family $\mathcal{F}$ is called d-separating if for any two distinct sets of size d there is a member of $\mathcal{F}$ that intersects one and is disjoint from the other. This question comes from search theory (also called group testing). Suppose we are given a set containing d unknown defective elements. We can ask queries that correspond to subsets of the base set, and the answer to a query shows if it contains a defective element or not. It is easy to see that a set of queries determine all the d defective elements if and only if the corresponding subsets form a d-separating family. It is a well-studied problem without the size restriction. If we restrict the size of the queries to at most k, but only assume that there are at most d defective elements, then strong results were achieved by D'yachkov and Rykov [5]. For our problem, we gave upper and lower bounds with a small gap. Surprisingly, what we get is about half of what D'yachkov and Rykov get for their version. It shows that even though having d defectives is the hardest case, knowing that there cannot be less than d defectives helps a lot. This is different from the usual situation in search theory. We also give sharp result for $d = 2$ and asymptotically tight result for $d = 3$. For the upper bound we use a construction of linear, regular, uniform hypergraphs with large girth by Ellis and Linial. Girth in hypergraphs leads to the next topic.

With Cory Palmer we considered the definition of cycles in hypergraphs given by Berge. A set of k edges $E_1, \dots, E_k$ on k vertices $x_1, \dots, x_k$ is a k-cycle if E_i contains x_i and x_{i+1} for $1 \le i \le k$, where $x_{k+1} := x_1$. Other containment is allowed, for example k hyperedges all containing the same k vertices is a k-cycle. We have generalized this definition to every graph G in [13]. A hypergraph $\mathcal{H}$ is Berge-G if there is a bijection $f : E(G) \to E(H)$ such that for $e \in E(G)$ we have $e \subset f(e)$. A hypergraph is G-free if it does not contain a subhypergraph that is Berge-G. The maximum cardinality of

G-free hypergraphs were studied in several papers by Győri, Lemons and Katona in case G is a cycle or path. They also consider a weighted version where we add up the sizes of the hyeredges. We showed that if all hyperedges are large enough, then even this sum is at most quadratic. For $K_{s,t}$-free hypergraphs we gave the upper bound $O(n^{2-1/s})$ where $s \leq t$.

With Balázs Keszegh and Balázs Patkós we considered a new variant of forbidden subposet problems. As the paper is in preparation and cannot be found even on my homepage or arXiv, I want to give a bit more details. Given a poset P, instead of counting the members of a P-free family, we count the copies of another poset Q (and, as usual, we want to maximize this number). The similar question for graphs is also a new area, and I am going to describe my results regarding that later. Here the results and the methods are very different from the old question of determining the cardinality, i.e. counting the copies of the 1-element poset. We determine the maximum for a couple pairs P, Q in [10]. In fact, if both are chains then the result follows from an old result of mine with Balázs Patkós [15].

Some of our results follow from known results or can be proved applying basic methods. But the most surprising development is that we can use profile polytopes in this area. For a family $\mathcal{F}$, let $f(\mathcal{F}) = (f_0, f_1, \dots, f_n)$ denote the *profile vector* of $\mathcal{F}$, where $f_i = |\{F \in \mathcal{F} : |F| = i\}|$. Many problems in extremal finite set theory ask for the largest size of a family in a class $\mathbb{A} \subseteq 2^{2^{[n]}}$. This question is equivalent to determining $\max_{\mathcal{F}\in\mathbb{A}} f(\mathcal{F}) \cdot \mathbf{1}$, where $\mathbf{1}$ is the vector with all entries being 1, and $\cdot$ denotes the scalar product. More generally, if for a given weight function $w : 2^{[n]} \to \mathbb{R}$ we want to maximize $w(\mathcal{F}) := \sum_{F\in\mathcal{F}} w(F)$, where $w(H)$ depends only on the size of F (i.e. $w(H) = w(H')$ whenever $|H| = |H'|$ holds), then this is equivalent to maximizing $f(\mathcal{F}) \cdot \mathbf{w}$, where $\mathbf{w} = (w_0, w_1, \dots, w_n)$ with w_i being the value of $w(H)$ for every i-element set H. As $\mathbb{A} \subseteq 2^{2^{[n]}}$ holds, we have $\{f(\mathcal{F}) : \mathcal{F} \in \mathbb{A}\} \subset \mathbb{R}^{n+1}$ and therefore we can consider its convex hull $\mu(\mathbb{A})$ that we call the *profile polytope* of $\mathbb{A}$. It is well known that any weight function with the above property is maximized by an *extreme point* of $\mu(\mathbb{A})$ (a point that is not a convex combination of any other points of $\mu(\mathbb{A})$). Determining the extreme points of the profile polytope of a class of families is the ultimate answer to the weighted extremal questions. It is known for several families, however I am not aware of any earlier results where it was used in the proof.

Let me illustrate how profile polytopes can be applied to our generalized forbidden subposet problems by sketching the proof of the following statement. The number of 2-chains in families without 3-chains is at most $\binom{n}{\lceil \frac{2n}{3} \rceil}\binom{\lceil \frac{2n}{3} \rceil}{\lfloor \frac{n}{3} \rfloor}$. If $\mathcal{F}$ is family without 3-chains, then it can be partitioned into $\mathcal{F}_1 \cup \mathcal{F}_2$ such that $\mathcal{F}_1 = \{F \in \mathcal{F} : \nexists F' \in \mathcal{F}\ F' \subset F\}$ and $\mathcal{F}_2 = \mathcal{F} \setminus \mathcal{F}_1$. Obviously, any copy of a 2-chain in $\mathcal{F}$ (i.e., a pair of sets in containment) contains one element from $\mathcal{F}_1$ and $\mathcal{F}_2$ each. Therefore one can count these copies by summing over all sets $F \in \mathcal{F}_2$ the number $s(F)$ of sets in $\mathcal{F}_1$ contained by F. As $\mathcal{F}_1$ is an antichain, by a theorem of Sperner we obtain $s(F) \leq \binom{|F|}{\lfloor \frac{|F|}{2} \rfloor} =: w(F)$. Clearly, $w(F)$ is a weight function that depends only on the size of F. Therefore we can apply Katona's result [18] that states that essential extreme points of the profile polytope of Sperner families are the profile vectors of full levels of Q_n. It follows that we only have to maximize $\binom{n}{i}\binom{i}{\lfloor i/2 \rfloor}$ over i. An easy calculation shows that the maximum is attained at $i = 2n/3$ (here we omit floor and ceiling signs as uniqueness depends on the residue of n modulo 3). This yield the upper bound and the corresponding construction (the family consisting of subsets of size $n/3$ and $2n/3$) shows that this is tight.

With Balázs Keszegh, Gábor Mészáros, Balázs Patkós and Máté Vizer we considered a generalized bootstrap percolation model. The usual bootstrap percolation is defined on graphs. A set A of infected vertices is given, and another vertex gets infected if it has at least r infected neighbors. This way we have defined a process; A is called percolating if every vertex gets infected by the end of this process. For practical applications this question is very interesting on the d-dimensional grid. Balister, Bollobás, Lee and Narayanan [2] defined a line percolation model, where the vertices of a line get infected if the line already contains at least r infected vertices, thus changing the rule of infection from a local property to one, where the infection can spread much faster. We generalize it to any hypergraph, and study the case of projective planes in [7]. There are numerous parameters to consider. In the probabilistic setting we calculate the critical probability that a random set of vertices percolate. We also prove a bottleneck phenomenon: if we pick random infected vertices one by one, when we pick enough vertices to find a line with r infected vertices, the set percolates with probability tending to infinity. In the combinatorial setting we study the size of the smallest and the largest percolating set, and also the number of rounds it takes to infect all the vertices.

With Cory Palmer we consider a new variant of Turán-type problems. During studying Berge-type problems, that I have described above, we have realized that counting hyperedges in a k-uniform hypergraph that is G-free (in the Berge sense) is closely related to counting copies of K_r in G-free graphs. That motivated the general question: what is the maximum number of copies of a graph H in an F-free graph on n vertices? (This in turn motivated the similar questions in posets that I have already discussed and the following question about the number of cycles). There were sporadic results in this area, and Alon and Shikhelman published a paper [1] about this while we did our research. They determined which graphs F have the property that every F-free graph contains only $O(n)$ triangles. In [14] We extend this to every cycle showing which forbidden graphs imply only a linear number of copies of C_k and showing that for every other forbidden graph we can find quadratic many copies of C_k. We also show that for $t \geq 2$ we have that the maximum number of copies of C_k is $(\frac{1}{2} + o(1))(t-1)^{k/2}n^{k/2}/k$, and the maximum number of copies of P_k is $(\frac{1}{2} + o(1))(t-1)^{(k-1)/2}n^{(k+1)/2}$ in $K_{2,t}$-free graphs on n vertices. We also prove an Erdős-Stone-Simonovits-type result, and explore the already mentioned connection to Berge-hypergraphs. The maximum number of hyperedges in r-uniform Berge-F-free hypergraphs is at least the maximum number of copies of K_r in G-free graphs, and at most the mayimum number of copies of K_r plus number of edges in G-free graphs. Moreover, if we consider non-uniform hypergraphs, then the maximum number of hyperedges is equal to the maximum number of complete subgraphs in G-free graphs.

With Balázs Keszegh, Cory Palmer and Balázs Patkós we considered a variation of the above problem. Instead of forbidding only one cycle, we forbid almost all of them. We still wanted to study extremal properties of such graph; the interesting question turned out to be the number of cycles in such graphs. Let L be a set of positive integers. We call a graph G an *L-cycle graph* if all cycle lengths in G belong to L. Let $c(L, n)$ be the maximum number of cycles possible in an n-vertex L-cycle graph (we use $\vec{c}(L, n)$ for the number of cycles in directed graphs). In the undirected case we show that for any fixed set L, we have $c(L, n) = \Theta_L(n^{\lfloor k/\ell \rfloor})$ where k is the largest element of L and 2ℓ is the smallest even element of L (if L contains only odd elements, then $c(L, n) = \Theta_L(n)$ holds.) We also give a characterization of L-cycle graphs when L is a single element. In the directed case we prove that for any fixed set L we have $\vec{c}(L, n) = (1 + o(1))(\frac{n-1}{k-1})^{k-1}$, where k is the largest element of L. We determine the exact value of $\vec{c}(\{k\}, n)$ for every k and characterize all graphs attaining this maximum.

I also had some results that are less closely related to the main topic.

With Viola Mészáros, Dömötör Pálvölgyi, Alexey Pokrovskiy and Günter Rote [11] we studied the discrete Voronoi game. It is defined in graphs, two players alternatingly claim vertices in t rounds. At the end of the game, the remaining vertices are divided between the players, with each player receiving the vertices that are closer to his or her claimed vertices. The question is what percent of the vertices the first player can control at the end. We showed that this number can be arbitrarily small and proved bounds for general graphs depending on the one-round game. We also determined the minimum the first player can get for trees in case there are only one or two rounds. With Balázs Keszegh, Dömötör Pálvölgyi, Balázs Patkós, Máté Vizer and Wiener Gábor [9] we considered a variant of the majority problem. In the original version we are given n blue or red ball. As one step we can compare two balls if they are of the same color (there is no way to actually find out the color of the ball). The goal is to find a ball of the majority color using as few steps as possible. In our model we check k balls as a step and find a majority ball among them. The goal is to find a majority (more precisely a non-minority) ball among all the balls. With Máté Vizer [17] we considered several other variant of the same problem. Finally, I would like to mention that together with Balázs Patkós I started writing a book with the title Extremal Finite Set Theory.

References

[1] N. Alon, C. Shikhelman, Many T-copies in H-free graphs, J. Combinatorial Theory, Ser. B, to appear.

[2] P. N. Balister, B. Bollobás, J. D. Lee, B. P. Narayanan, Line percolation. arXiv:1403.6851 (2014).

[3] F. Benevides, D. Gerbner, C. Palmer, D. Vu, Generalising separated families of fixed size, submitted

[4] P. Burcsi, D. T. Nagy, The method of double chains for largest families with excluded subposets. Electronic Journal of Graph Theory and Applications (EJGTA), 1(1), 40–49, 2013.

[5] A. G. D'yachkov, V. V. Rykov, Optimal superimposed codes and designs for Renyi's search model. J. Statist. Plann. Inference, 100(2):281–302, 2002. International Conference (of the Forum for Interdisciplinary

Mathematics) on Combinatorics, Information Theory and Statistics (Portland, ME, 1997).

[6] D. Gerbner, Profile polytopes of some classes of families, Combinatorica, 33(2), 199–216, 2013.

[7] D. Gerbner, B. Keszegh, G. Mészáros, B. Patkós, M. Vizer, Line Percolation in Finite Projective Planes, submitted

[8] D. Gerbner, B. Keszegh, C. Palmer, B. Patkós, On the number of cycles when few cycle lengths are allowed, in preparation

[9] D. Gerbner, B. Keszegh, D. Pálvölgyi, B. Patkós, M. Vizer, G. Wiener, Finding a majority ball with majority answers, Electr. Notes in Disc. Math. 49,345–351.

[10] D. Gerbner, B. Keszegh, B. Patkós, Generalized forbidden subposet problems, in preparation

[11] D. Gerbner, V. Mészáros, D. Pálvölgyi, A. Pokrovskiy, and G. Rote, Advantage in the discrete Voronoi game Vol. 18, no. 3, 439–457, 2014.

[12] D. Gerbner, A. Methuku, C. Tompkins, Intersecting P-free families, submitted

[13] D. Gerbner, C. Palmer, Extremal results for Berge-hypergraphs, submitted

[14] D. Gerbner, C. Palmer, Generalized Turán problems, in preparation

[15] D. Gerbner, B. Patkós, l-chain profile vectors, SIAM J. Discrete Math. 22,185–193, 2008.

[16] D. Gerbner, M. Vizer, A note on tilted Sperner families with patterns, Discrete Mathematics 339, 2737–2741, 2016.

[17] D. Gerbner, M. Vizer, Majority problems of large query size, in preparation

[18] G.O.H. Katona, Two applications (for search theory and truth functions) of Sperner type theorems, Period. Math. Hungar. 3, 19–26, 1973.

[19] I. Leader, E. Long, Tilted Sperner families, Disc. Appl. Math., 163, part 2, 194–198, 2014.

[20] E. Long, Forbidding intersection patterns between layers of the cube, J. Comb. Theory A 134, 103–120, 2015.

[21] EC Milner, A combinatorial theorem on systems of sets, Journal of the London Mathematical Society, 1(1), 204–206, 1968.