Low Dimensional Combinatorics

KKP-133864 – Project Final Report

One of the main objectives of our project was to study hard combinatorial problems for large classes of graphs and hypergraphs arising in geometric, algebraic, and practical applications. In particular, for structures escaping the "curse of dimensionality": if they can be embedded in a bounded-dimensional space, or they have small Vapnik-Chervonenkis dimension (VC-dimension) or a short algebraic description. We made substantial progress in this direction.

Support by the KKP grant was acknowledged in 20 papers on a number of different topics. Below we summarize the most significant results, grouped according to the relevant combinatorial and geometric subfields.

1 Classical areas

In this section, we collect results in graph drawing, geometric hypergraphs, Ramsey theory, and computational geometry. These results appeared in the following publications: [7, 8, 10, 11, 16, 17, 18, 19]

1.1 Chromatic number vs. clique number

Due to recent seminal work by Chudnovsky, Scott, Seymour, and others, there has been a lot of progress concerning the χ-*boundedness* of various important graph classes. The notion was first introduced by Gyárfás and Lehel, and several conjectures made by them have been proved. We contributed to this theme by studying various classes of geometrically defined graphs.

The disjointness graph $G = G(\mathcal{S})$ of a set of segments $\mathcal{S}$ in $\mathbb{R}^d$, $d \geq 2$ is a graph whose vertex set is $\mathcal{S}$ and two vertices are connected by an edge if and only if the corresponding segments are disjoint. We proved [7] that the chromatic number of such a graph G satisfies $\chi(G) \leq (\omega(G)^4 + \omega(G))^3$, where $\omega(G)$ denotes the clique number of G. It follows that $\mathcal{S}$ has $\Omega(n^{1/5})$ pairwise intersecting or pairwise disjoint elements. Stronger bounds were established for lines in space, instead of segments. We showed [7] that computing $\omega(G)$ and $\chi(G)$ for disjointness graphs of lines in space are NP-hard tasks. However, we managed to design efficient algorithms to compute proper colorings of G in which the number of colors satisfies the above upper bounds. One cannot expect similar results for sets of continuous arcs instead of segments or lines, even in the plane. We constructed families of arcs, whose disjointness graphs are triangle-free ($\omega(G) = 2$), but whose chromatic numbers are arbitrarily large.

1.2 Crossing number

The *crossing number* of a graph G is the minimum number of edge crossings over all drawings of G in the plane. A graph G is k-crossing-critical if its crossing number is at least k, but if we remove any edge of G, its crossing number drops below k. There are examples of k-crossing-critical graphs that do not have drawings with exactly k crossings. Richter and Thomassen proved in 1993 that if G is k-crossing-critical, then its crossing number is at most $2.5k + 16$. In spite of many efforts to improve this bound, it remained the best known result until we [16] managed to break the ice and proved that the crossing number of k-crossing-critical graphs is at most $2k + 6\sqrt{k} + 44$.

Let G be a multigraph with n vertices and $e > 4n$ edges drawn in the plane such that any two parallel edges form a simple closed curve with at least one vertex in its interior and at least one vertex in its exterior. Pach and Tóth (2018) extended the Crossing Lemma of Ajtai et al. (1982) and Leighton (1983) by showing that if no two adjacent edges cross and every pair of nonadjacent edges cross at most once, then the number of edge crossings in G is at least ce^3/n^2, for a suitable constant $c > 0$. The situation turns out to be quite different if nonparallel edges are allowed to cross any number of times. We proved [18] that in this case the number of crossings in G is at least $ce^{2.5}/n^{1.5}$. The order of magnitude of this bound cannot be improved.

We call a multigraph *non-homotopic* if it can be drawn in the plane such that no two edges connecting the same pair of vertices can be continuously transformed into each other without passing through a vertex, and no loop can be shrunk to its end-vertex in the same way. It is easy to see that a non-homotopic multigraph on $n > 1$ vertices can have arbitrarily many edges. We proved in [19] that the number of crossings between the edges of a non-homotopic multigraph with n vertices and $m > 4n$ edges is larger than cm^2/n for a suitable constant $c > 0$, and that this bound is tight up to a polylogarithmic factor. We also showed the lower bound is asymptotically not sharp when n is fixed and m tends to infinity.

1.3 Ramsey-type results

We considered m-colorings of the edges of a complete graph, where each color class is defined semi-algebraically with bounded complexity [8]. The case $m = 2$ was first studied by Alon et al., who applied this framework to obtain surprisingly strong Ramsey-type results for intersection graphs of geometric objects and for other graphs arising in computational geometry. Considering larger values of m is relevant to problems concerning the number of distinct distances determined by a point set. For $p \geq 3$ and $m \geq 2$, the classical Ramsey number $R(p; m)$ is the smallest positive integer n such that any m-coloring of the edges of K_n, the complete graph on n vertices, contains a monochromatic K_p. It is a longstanding open problem that goes back to Schur (1916) to decide whether

$R(p;m) \leq 2^{cm}$, where $c = c(p)$. We proved that this is true if each color class is defined semi-algebraically with bounded complexity, and that the order of magnitude of this bound is tight. Our proof is based on the Cutting Lemma of Chazelle et al., and on a Szemerédi-type regularity lemma for multicolored semi-algebraic graphs, which is of independent interest. The same technique is used to address the semi-algebraic variant of a more general Ramsey-type problem of Erdős and Shelah.

1.4 Geometric arrangements

Another important objective of the project was to seek new structural information concerning various geometric arrangements, including arrangements of strings (i.e., continuous curves) and half-planes in the plane. Our goal was to apply a mix of combinatorial and algebraic techniques. It turned out to be fruitful to study extremal problems concerning forbidden homogeneous submatrices in a matrix.

A matrix is called homogeneous if all of its entries are equal. Let P be a 2×2 zero-one matrix that is not homogeneous. We proved that if an $n \times n$ zero-one matrix A does not contain P as a submatrix, then A has a $cn \times cn$ homogeneous submatrix, for a suitable constant $c > 0$ [10]. We further provided an almost complete characterization of the matrices P (missing only finitely many cases) such that forbidding P in a matrix A guarantees the existence of an $n^{1-o(1)} \times n^{1-o(1)}$ homogeneous submatrix. We applied our results to chordal bipartite graphs, totally balanced matrices, half-plane arrangements, and string graphs.

1.5 Set systems

A "cornerstone" of the proposed research was to approach notoriously difficult unsolved problems for bounded-dimensional hypergraphs (set systems). We successfully used this idea to prove a slightly weaker version the famous *Sunflower Conjecture* of Erdős and Rado.

Given a family $\mathcal{F}$ of k-element sets, some of its members, $S_1, \ldots, S_r \in \mathcal{F}$, form an *r-sunflower* if $S_i \cap S_j = S_{i'} \cap S_{j'}$ for all $i \neq j$ and $i' \neq j'$. According to the Erdős-Rado conjecture (1960), there is a constant $c = c(r)$ such that if $|\mathcal{F}| \geq c^k$, then $\mathcal{F}$ contains an r-sunflower. We came close [11] to proving this conjecture for families of bounded *Vapnik-Chervonenkis dimension*, i.e., when VC-dim$(\mathcal{F}) \leq d$. In this case, we showed that r-sunflowers exist under the slightly stronger assumption $|\mathcal{F}| \geq 2^{10k(dr)^{2\log^* k}}$. Here, $\log^*$ denotes the iterated logarithm function. We also verified the Erdős-Rado conjecture for families $\mathcal{F}$ of bounded *Littlestone dimension* and for some geometrically defined set systems.

For $n \leq d$, a family $\mathcal{F} = \{C_0, C_1, \ldots, C_n\}$ of compact convex sets in R^d is called an n-critical family provided any n members of $\mathcal{F}$ have a non-empty intersection, but $\bigcap_{i=0}^{n} C_i = \varnothing$. If $n = d$, then a lemma on the intersection of convex sets due to Klee implies that the $d+1$ members of the d-critical family enclose a “hollow” in R^d, a bounded connected component of $R^d \setminus \bigcup_{i=0}^n C_i$. We proved [17] that the closure of the convex hull of a hollow in R^d is a d-simplex.

2 Euclidean, Spherical, and Hyperbolic Geometries

We have also contributed to some purely geometric fields; see [5, 6, 9].

2.1 Triangles

We showed in [9] that every minimum area isosceles triangle containing a given triangle T shares a side and an angle with T. This proves a conjecture of Nandakumar motivated by a computational problem. We used this result to deduce that for every triangle T,

(1) there are at most 3 minimum area isosceles triangles that contain T, and

(2) there exists an isosceles triangle containing T, whose area is smaller than $\sqrt{2}$ times the area of T.

Both bounds are best possible.

2.2 Cubes

It is a 300-year-old – rather counter-intuitive – observation of Prince Rupert of the Rhine that one can cut a straight tunnel in a 3-dimensional unit cube, through which another unit cube can be passed. A hundred years later, P. Nieuwland strengthened Rupert's problem and asked for the largest aspect ratio so that a larger homothetic copy of the same body can be passed through it. We showed [5] that cubes and, in fact all, rectangular boxes have Rupert-type passages in every direction that is not parallel to its faces. In the case of the cube, it was assumed without proof that the solution of the Nieuwland's problem is a tunnel perpendicular to the largest square contained by the cube. We proved that this assumption is unnecessary not only for the cube, but also for all other rectangular boxes.

2.3 Circles

We proved in [6] that for any arrangement of finitely many circles on the sphere or in the hyperbolic plane, there is a polygonal tiling, each face of which contains exactly one circle. The corresponding statement was known in the Euclidean plane. We also showed that in the Euclidean plane there is no convex disc C other than the circular one, with the property that any packing with similar copies of C can be separated by a polygonal tiling. The analogous statement is false on the sphere and it is not known in the hyperbolic plane. This study led to another characterization of the Euclidean circle: It is the only convex disc all caps of which are isosceles. It can be conjectured that restricting this condition to caps of a fixed angle would also characterize Euclidean circles.

3 Geometry and analysis

Here we describe some results whose proofs involve a mix of geometric and analytical methods.

3.1 Riemannian manifolds

The total scalar curvature of a compact submanifold, possibly with boundary, of a Riemannian manifold is the integral of the scalar curvature function of the submanifold over the submanifold with respect to the volume measure induced by the Riemannian metric. Csikós and Horváth proved in an earlier paper [20] that if a connected Riemannian manifold of dimension at least 4 is harmonic, then the total scalar curvatures of tubes of small radius about a regular curve depend only on the length of the curve and the radius of the tube, and conversely, if the latter condition holds for cylinders, i.e. for tubes about *geodesic* segments, then the manifold is harmonic.

In [21], we showed that, in contrast to the higher dimensional case, a connected 3-dimensional Riemannian manifold has the above mentioned property of tubes if and only if the manifold is a D'Atri space. (A Riemannian manifold is said to be a D'Atri space if the local geodesic reflection with respect to an arbitrary point is volume-preserving). Furthermore, if the space has bounded sectional curvature, then it is enough to require the total scalar curvature condition only for cylinders to imply that the space is D'Atri. This result led to a negative answer to the following question of Gheysens and Vanhecke [22]: Does vanishing of the total scalar curvature of tubes about curves in a 3-dimensional Riemannian manifold imply that the manifold is harmonic? To prove these statements, we gave the following characterization of D'Atri spaces in any dimension: the total scalar curvatures of any two geodesic hemispheres lying on a given geodesic sphere are equal.

3.2 Log-concave functions and computational mathematics

We gave [1] a definition of the John ellipsoid of logarithmically concave functions in an analogous manner to that used in the context of convex sets. As an application, we showed a bound on the integral of the point-wise minimum of log-concave functions. We studied a related problem in [2], a quantitative Helly type question, and obtained a general bound which can be used to answer a number of related problems.

We designed [3] an efficient algorithm for computing the covering radius of rational polytopes in $\mathbb{R}^n$. The algorithm is faster in relevant cases than the best previously known algorithm of Kannan from 1993. Applying this algorithm, we managed to solve the first open case of a variant of the Lonely Runner Conjecture. We studied [4] the Closest Vector Problem, and using a new geometric approach which relies on set coverings, we obtained a fast algorithm to approximate the optimum.

4 Applications

Apart from the algorithmic applications described in the last subsection, we found further applications of our combinatorial techniques to geometry [14]. In other cases, we used geometric intuition and concepts to tackle "abstract" (non-geometric) questions [12, 15].

4.1 Compression schemes

Let B be a fixed finite domain and $\mathcal{F}$ a fixed family of binary functions $f : B \to \{0,1\}$. For learning functions $f \in \mathcal{F}$, several compression schemes were suggested. Intuitively, the learner is given a sample, that is, the function f restricted to a subset A of the domain B. But the learner has no space to store the entire sample. Instead, she stores only a well chosen fragment of it and should be able to restore the value $f(x)$ for any $x \in A$ from the fragment. Kuzmin and Warmuth [13] suggested the following definition for an *unlabelled compression scheme* for the family $\mathcal{F}$: it consists of two functions, the compressing function α takes an arbitrary function on a subset A of the domain B and outputs a subset C of A. The reconstruction function β takes a subset of B and an element $x \in B$ and outputs 0 or 1. Such a pair is an unlabeled compression scheme if it correctly guesses $f(x)$ for any $x \in A$, that is, if for any function $f \in \mathcal{F}$, for any subset A of the domain B, and for any element $x \in A$, we have $\beta(\alpha(f|_A), x) = f(x)$. The *size* of such a compression scheme is the size of the largest set C output by the compression function.

Kuzmin and Warmuth observed that the size of any unlabeled compression scheme for $\mathcal{F}$ must be at least the VC-dimension of $\mathcal{F}$ and proved that unlabelled compression schemes of size equal to the VC-dimension exists for any maximal size family $\mathcal{F}$ for any given VC-dimension. (Their proof contained a minor error that has since been fixed.) They conjectured that for any family $\mathcal{F}$, there exists an unlabeled compression scheme whose size is equal to the VC-dimension. In [12], we gave a counterexample to this conjecture. The proof is interesting for its novel lower bound techniques. We also explored bounds for the size of optimal unlabeled compression schemes for various function families.

4.2 Extremal edge-ordered graphs with applications in discrete geometry

One of the objectives of the project was to develop the foundations of an extremal theory of *ordered* graphs.

The basic question of Turán-type extremal graph theory is the following: what is the maximal number of edges that an n-vertex graph can have if it does not contain a fixed so-called *forbidden subgraph.* Results of this kind have found many applications in various fields of mathematics and computer science. Extensions of this theory study the same question about graphs equipped with some additional structure such as a drawing in the plane (geometric graphs, topological graphs), a linear order among the vertices (vertex-ordered graphs) among others. We introduced *edge-ordered* graphs in [14]: these are simple graphs with a linear order on their edge-set. The corresponding extremal problem is as follows: how many edges can an edge-ordered graph on n vertices have if a it does not contain a given *forbidden edge-ordered graph.* In other words, the forbidden graph comes with an edge-order and we allow our graph to contain isomorphic copies of the underlying simple graph if they have a different edge-order.

We started to build the extremal theory of edge-ordered graphs [14]. We found an interesting version of the well known Erdős-Stone-Simonovits theorem that applies to edge-ordered graphs. Each forbidden edge-ordered graph has an *order-chromatic number*, an integer $c \geq 2$ (or infinity), and the asymptotic behavior of the corresponding extremal function is determined by this number, unless it is 2. The order-chromatic number behaves differently from the classic chromatic number in the following respect. It can be exponentially (possibly twice exponentially) larger than the size of the edge-ordered graph. We showed that when one forbids more than one edge-ordered graphs, the family often behaves as an edge-ordered graph with considerably lesser order chromatic number than any of the edge-ordered graphs in the family.

We studied the extremal function of a few well-chosen forbidden patterns in [14]. They include all edge-ordered paths and cycles of up to four edges, forests consisting of star components and some other graphs. One of our results (about the extremal function of an edge-ordered path with four edges) has already found an application in discrete geometry. Keszegh and Pálvölgyi was able to improve a bound on the number of possible touchings between two restricted families of plane curves using our result.

4.3 Limit theory of finite trees

A limit theory of finite graphs was developed by Lovász, Szegedy and many others. It proved useful in extremal graph theory and, in general, it has led to a better understanding the structure of large graphs. The limit theory is based on random sampling: if we take a uniform random sample of k vertices of a much larger graph, and consider the subgraph induced by them, we obtain a distribution on the k-vertex graphs. In the limit theory of large graphs, a graph sequence is considered convergent if these distributions converge for any fixed k. Its limit object is a two-variable real function.

We developed a similar theory for the convergence and limits of finite trees in [15]. We also worked with random samples of vertices, but trees are sparse, so the induced subgraphs are not informative (they are almost always empty).

Instead, we considered trees as metric spaces with a normalized version of the graph distance. This way one can apply the well developed theory of metric measures. It is quite surprising that the limit objects we find (which we call them *dendrons*) are typically not metric measure spaces. The relevant metric spaces (so-called *real trees*) are needed for the description of the limit objects. However, the probability distributions are defined not directly on the real trees T, but on the product set $T \times [0,1]$. This product set is then equipped with a strange "distance function" d that does not satisfy $d(x,x) = 0$ for most of the points x. Nevertheless, sampling random points from the product, these "distances" give the limit distribution of the normalized graph distances whenever finite graphs are sampled.

We proved the existence and uniqueness of the limit dendrons for a convergent sequence of finite trees using ultralimit theory. This requires the extension of earlier results on ultralimits of metric and measure spaces.

References

[1] G. Ivanov, M. Naszódi, Functional John Ellipsoids. `https://arxiv.org/abs/2006.09934`

[2] G. Ivanov, M. Naszódi, A Quantitative Helly-type Theorem: Containment in a Homothet. `https://arxiv.org/abs/2103.04122`

[3] J. Cslovjecsek, R.D. Malikiosis, M. Naszódi and M. Schymura, Computing the covering radius of a polytope with an application to lonely runners. `https://arxiv.org/abs/2009.12080`

[4] M. Naszódi and M. Venzin, Covering convex bodies and the Closest Vector Problem. `https://arxiv.org/abs/1908.08384`

[5] A. Bezdek, Z. Guan, M. Hujter and A. Joos, Cubes and Boxes Have Rupert's Passages in Every Nontrivial Direction. *Amer. Math. Monthly* 128(6):(2021), 534–542.

[6] A. Bezdek, Separating circles on the sphere by polygonal tilling. *Beiträge zur Algebra und Geometrie / Contributions to Algebra and Geometry*, Springer, accepted. (2020), pp 1–11.

[7] J. Pach, G. Tardos and G. Tóth, Disjointness graphs of segments in the space. *Combinatorics, Probability and Computing* Volume 30 , Issue 4 , (2021), 498–512.

[8] J. Fox, J. Pach and A. Suk, The Schur-Erdős problem for semi-algebraic colorings. *Israel Journal of Mathematics* volume 239, (2020), 39–57.

[9] G. Kiss, J. Pach and G. Somlai, Minimum Area Isosceles Containers. *Journal of Information Processing* Vol.28, (2020) 759–765.

[10] D. Korándi, J. Pach and I. Tomon, Large Homogeneous Submatrices. *SIAM J. Discret. Math.* to appear.

[11] J. Fox, J. Pach and A. Suk, Sunflowers in set systems of bounded dimension. `https://arxiv.org/pdf/2103.10497.pdf`

[12] D. Pálvölgyi and G. Tardos, Unlabeled compression schemes exceeding the VC-dimension. *Discrete Applied Mathematics*, **276** (2020), 102–107.

[13] D. Kuzmin and M.K. Warmuth, Unlabeled Compression Schemes for Maximum Classes. In Proceedings of the 18th Annual Conference on Computational Learning Theory (COLT 05), Bertinoro, Italy, 591–605, June 2005.

[14] D. Gerbner, A. Methuku, D. Nagy, D. Pálvölgyi, G. Tardos and M. Vizer, Edge ordered Turán problems. submitted to JCTB.

[15] G. Elek and G. Tardos, Convergence and limits of finite trees. *Combinatorica*, to appear.

[16] J. Barát and G. Tóth, Improvement on the crossing number of crossing-critical graphs. *Discrete and Computational Geometry* (2020).

[17] J. Lehel and G. Tóth, On the hollow enclosed by convex sets. *Geombinatorics* XXX (3) (2021), 113–122.

[18] M. Kaufmann, J. Pach, G. Tóth and T. Ueckerdt, The number of crossings in multigraphs with no empty lens. *Journal of Graph Algorithms and Applications* to appear.

[19] J. Pach, G. Tardos and G. Tóth, Crossings between non-homotopic edges. Graph Drawing 2020, Vancouver, 16-18 September 2020 (best paper) *Lecture Notes in Computer Science* **12590** (2021) 359–371.

[20] B. Csikós and M. Horváth, Harmonic manifolds and tubes. *J. Geom. Anal.* 28 (2018).

[21] B. Csikós, A. Elnashar and M. Horváth, D'Atri spaces and the total scalar curvature of hemispheres, tubes and cylinders. `arXiv:2110.03678` (2021).

[22] L. Gheysens and L. Vanhecke, Total scalar curvature of tubes about curves. *Math. Nachr.* 103 (1981), 177–197.