Final report on the project "Extremal vector sets and limit shapes"

Gergely Ambrus

NKFIH Project no. PD-125502

In this report we summarize the results of the NKFIH project no. PD-125502 achieved by the Principal Investigator (PI), Gergely Ambrus.

The results of the project are collected in the following seven publications:

- G. Ambrus, S. Nietert, *Polarization, sign sequences and isotropic vector systems.* Pacific J. Math. **303** (2019), no. 2, 385-399.
- G. Ambrus and M. Matolcsi, *On the density of planar sets without unit distances.* Accepted for publication, Discrete Comput. Geom. (2020). `arXiv:1809.05453`.
- G. Ambrus, *On a frame energy problem.* Submitted (2020). `arXiv:2002.03974`.
- G. Ambrus, P. Nielsen, and C. Wilson, *New estimates for convex layer numbers.* Submitted (2020). `arXiv:2006.03799`.
- G. Ambrus, A. Hsu, B. Peng, and S. Jan, *The layer number of grids.* Submitted (2020). `arXiv:2009.13130`.
- G. Ambrus, *Longest k-monotone chains.* Submitted (2020). `arXiv:2009.13887`
- G. Ambrus, B. González Merino, *Large signed subset sums.* Manuscript, 2020.

The project proposal concentrated on two main topics: (1) Extremal vector systems and sharp inequalities; and (2) Limit shapes and typical behaviour of random points. During the 3-year-period of the project, the PI managed to prove substantial results in both of these topics, as well as answering questions from other fields. We list the results below.

Extremal vector systems. Given a vector set $v_1, \ldots, v_N \subset S^{d-1}$ and $p > 0$, the *ℓ_p-potential* of the system at the unit vector w as

$$U^p((v_i)_1^N, w) = \sum_{i=1}^{n} |\langle w, v_i \rangle|^p.$$

This is an analogue of the classical Riesz potential for inner products. The *ℓ_p-polarization of* $(v_i)_1^N$ is given by

$$M^p((v_i)_1^N) = \max_{v \in S^{d-1}} U^p((v_i)_1^N, v).$$

The quantity

$$M_N^p(S^{d-1}) = \min_{u_1,\ldots,u_n \in S^{d-1}} \max_{w \in S^{d-1}} \sum_{i=1}^{n} |\langle w, u_i \rangle|^p$$

is called the *Nth ℓ_p-polarization (or Chebyshev) constant of S^{d-1}*.

Polarization problems have been subject to very active research in the last 15 years, although their study dates back to at least 1967 [20]. In the planar case, $M_n^p(S^1)$ has a direct connection to the classical notions of Riesz potentials and Chebyshev constants.

In a recent article jointly written with S. Nietert [3], the PI succesfully determined the order of magnitude of $M_n^p(S^{d-1})$ for $p > 0$. Moreover, they showed that for $p = 2$, the ℓ_2-polarization optimal vector sets are exactly the isotropic vector sets, whereas for $p = 1$, they showed that the polarization problem is equivalent to that of maximizing the norm of signed vector sums. Finally, for $d = 2$, they proved the optimality of equally spaced configurations on the unit circle.

In a closely related joint work with B. González Merino [10], the PI considered the problem of large signed subset sums. The question is the following. For every $d \geqslant 2$, $n \geqslant d$ and $k \leqslant n$ positive numbers, determine the largest number $c(d, n, k)$ so that for any system of n unit vectors $u_1, \ldots, u_n \subset S^{d-1}$ there is a subset $u_{i_1}, \ldots, u_{i_k}$ of cardinality k of them with corresponding signs $\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}$, so that

$$\left| \sum_{j=1}^{k} \varepsilon_j u_{i_j} \right| \geqslant c(d, n, k).$$

This question is, in some sense, dual to the small subset sums problem studied in [12]. In [10], the PI and his coauthor completely solved the question in the case $d = 2$, and they determined the sharp asymptotics for the higher dimensional quantities. They also proved sharp results in the cases $k = 2$, and $n = d + 1$.

In another publication related to extremal vectors sets [6], the PI studied the extremum problem

$$\max_{\substack{v_1,\ldots,v_N \in \mathbb{R}^d \\ c_1 \leqslant |v_i|^2 \leqslant c_2 \ \forall i}} \ \min_{1 \leqslant k \leqslant N} \frac{|v_k|^2}{\sigma^2 + \sum_{l \neq k} \langle v_k, v_l \rangle^2} .$$

He showed that in the case $\sigma = 0$, *uniform tight frames* are the only optimal configurations: these are systems of vectors $v_1, \ldots, v_N \in \mathbb{R}^d$ for which $|v_i| = R$ for every i and some constant R, and

$$\sum_{i=1}^{N} v_i \otimes v_i = \lambda \, I_d$$

holds with some positive constant λ. He also gave quantitative bounds on how far tight frames are being from optimal for small values of σ^2. The motivation for studying this question is the practical application in telecommunication systems, in particular by Huawei Technologies, Ltd. This result has led to an ongoing joint research project between the PI and the company.

Limit shapes and typical behaviour of random points. In the article [9], the PI studied higher order convexity properties of n random points chosen independently, uniformly in the unit square $[0, 1]^2$, denoted by X_n. The framework constructed by him is a common generalization of the famous question of the longest increasing subsequences in random permutations (see e.g. [1]) and the problem of longest convex chains [11]. Based on the work of Eliás and Matoušek, [17], he introduced the concept of k-monotone chains. He proved that if L_n^k denotes the maximal cardinality of a k-monotone chain in X_n, then for every $k \geqslant 1$,

$$\lim_{n \to \infty} n^{-\frac{1}{k+1}} \, \mathbb{E} L_n^k = \alpha_k$$

with some positive constant α_k. Furthermore, $n^{-\frac{1}{k+1}} \, L^k(X_n) \to \alpha_k$ almost surely, as $n \to \infty$. The proof is based on the subadditive ergodic theorem of Kingman [19].

The PI also established concentration properties of the random variable L_n^k by showing that for every $k \geqslant 1$, and for every $\varepsilon > 0$,

$$\mathbb{P}\left(|L_n^k - \mathbb{E} L_n^k| > \varepsilon n^{\frac{1}{2(k+1)}} \right) \leqslant 5 e^{-\varepsilon^2 / 5\alpha_k}$$

holds for every sufficiently large n. This is proved by utilizing Talagrand's concentration estimates [22].

In addition to the topics specified in the original project proposal, the PI also conducted research in the following two topics.

Density of planar sets without unit distances. In his joint work with M. Matolcsi [10], the PI studied the following question: What is the maximal upper density of a measurable planar set A with no two points at distance 1? This 40-year-old problem has attracted special attention recently, with a sequence of progressively improving estimates, the strongest of which currently being that of Bellitto, Pêcher, and Sédillot [13], who gave the upper estimate 0.25646. Erdős conjectured that the density of such a set may not exceed 0.25. In their work, M. Matolcsi and the PI show that any Lebesgue measurable, 1-avoiding planar set has upper density at most 0.25442, getting tantalizingly close to the conjecture of Erdős. In order to do so, they revert to the Fourier analytic method described by Oliveira Filho and Vallentin [21]. They applied the novel way of estimating densities of triple correlations, which translates to density estimates of prescribed triangles in the set A. These theoretical constraints are then used to formulate a linear programming problem, which may be solved with the aid of numerical methods.

Estimates for layer numbers. In the articles [7] and [8], the PI studies the following computational geometric problem. Let $X \subset \mathbb{R}^d$ be a finite point set. Define the *peeling process* as follows: in every step, we remove the set of vertices of the convex hull of the previous iteration. The sets of points removed in each step form the *convex layers* of X, while the total number of steps needed to completely delete X is the *layer number* of X, which we denote by $\tau(X)$. The convex layer decomposition of planar sets was first studied by Eddy [16] and Chazelle [14] from the algorithmic point of view. The latter article gave an $O(n \log n)$ running time algorithm for computing the convex layers of an n-element planar point set. Therefore, layer numbers may be computed quickly and efficiently.

In the article [7], the authors studied evenly distributed families of point sets contained in B^d, the d-dimensional unit ball. A family of sets $X_1, X_2, X_3, \dots \subset B^d$ is said to be *evenly distributed* if $|X_i| \to \infty$ and the minimal distance between pairs of points of X_i is of order $\Theta(|X_i|^{-1/d})$. This is a simplification of the definition appearing in [15], which studies the planar case of the problem.

Together with his coauthors, the PI showed that $\tau(X_i) \geqslant \Omega(|X_i|^{1/d})$ for every evenly distributed family $\{X_i\}_{i=1}^{\infty} \subset B^d$. Moreover, this bound is shown to be sharp. On the other hand, they proved that for such families of sets, $\tau(X_i) \leqslant O(|X_i|^{2/d})$ holds. This improves on the previous bounds for every $d \geqslant 3$. That this upper bound is nearly sharp is illustrated by their construction: for every $d \geqslant 1$, there exists an evenly distributed family $\{X_i^d\}_{i=1}^{\infty}$ in $\mathbb{R}^d$ with

$$\tau(X_i^d) = \Theta\left(|X_i^d|^{\frac{2}{d} - \frac{1}{d\,2^{d-1}}}\right) .$$

Another well studied question is the layer number of the grids $[n]^d = \{1, \dots, n\}^d$. Har-Peled and Lidický [18] proved that $\tau([n]^2) = \Theta(n^{4/3})$. The PI and his coauthors showed that for every $d \geqslant 1$, $\tau([n]^d) \geqslant \Omega\left(n^{\frac{2d}{d+1}}\right)$. This estimate is conjectured to be tight, since this is proved to be the layer number of the random point sets. On the other hand, they showed that for every $d \geqslant 3$, the layer number of the d-dimensional grid satisfies $\tau([n]^d) \leqslant O(n^{d-9/11})$. These are the first nontrivial estimates for the layer number of high dimensional grids. The applied methods include number theoretical and combinatorial tools.

References

[1] D. Aldous, P. Diaconis, *Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem.* Bull. Amer. Math. Soc. **36** (1999), 413–432.

[2] G. Ambrus, I. Bárány, *Longest convex chains.* Random Structures Algorithms **35** (2009), no. 2, 137–162.

[3] G. Ambrus, S. Nietert, *Polarization, sign sequences and isotropic vector systems.* Pacific J. Math. **303** (2019), no. 2, 385-399.

[4] G. Ambrus, *On a frame energy problem.* Submitted. `arXiv:2002.03974`, 1-9.

[5] G. Ambrus, M. Matolcsi, *On the density of planar sets without unit distances.* Accepted for publication, Discrete Comput. Geom. (2020). `arXiv:1809.05453`, 1–15.

[6] G. Ambrus, *On a frame energy problem.* Submitted (2020). `arXiv:2002.03974`.

[7] G. Ambrus, P. Nielsen, and C. Wilson, *New estimates for convex layer numbers.* Submitted (2020). `arXiv:2006.03799`.

[8] G. Ambrus, A. Hsu, B. Peng, and S. Jan, *The layer number of grids.* Submitted (2020). `arXiv:2009.13130`.

[9] G. Ambrus, *Longest k-monotone chains.* arXiv preprint, 2020.

[10] G. Ambrus, B. González Merino, *Large signed subset sums.* Manuscript, 2020.

[11] G. Ambrus, I. Bárány, *Longest convex chains.* Random Structures Algorithms **35** (2009), no. 2, 137–162.

[12] G. Ambrus, I. Bárány, V. Grinberg, *Short subset sums.* Linear Alg. Appl. **499** (2016), 66–78.

[13] T. Bellitto, A. Pêcher, and A. Sédillot, *On the density of sets of the Euclidean plane avoiding distance 1.* arXiv:1810.00960 (2018), 1–11.

[14] B. Chazelle, *On the convex layers of a planar set.* IEEE Trans. Inform. Theory **31** (1985), no. 4, 509–517.

[15] I. Choi, W. Joo, and M. Kim, *The layer number of α-evenly distributed point sets.* To appear in Discrete Mathematics. arXiv preprint 2006.02822 (2020), 1–14.

[16] W.F. Eddy, *Convex Hull Peeling.* In: Caussinus H., Ettinger P., Tomassone R. (eds), COMPSTAT 1982, 5th Symposium held at Toulouse 1982. Physica, Heidelberg, 1982.

[17] M. Eliáš and J. Matoušek, *Higher order Erdős-Szekeres theorems.* Adv. Math. **244** (2013), 1–15.

[18] S. Har-Peled and B. Lidický, *Peeling the Grid.* SIAM J. Discrete Math. **27** (2013), no. 2.,

[19] J.F.C. Kingman, *Subadditive ergodic theory*, Ann. Prob. **1** (1973), no. 6., 883–909.

[20] M. Ohtsuka, *On various definitions of capacity and related notions.* Nagoya Math. **30** (1967), 121–127.

[21] F. M. de Oliveira Filho and F. Vallentin, *Fourier analysis, linear programming, and densities of distance-avoiding sets in $\mathbb{R}^n$.* Journal of the European Mathematical Society **12** (2010), 1417–1428.

[22] M. Talagrand, *A new look at independence.* Ann. Probab. **24** (1996), 1–34.