Joint upper Banach density, VC dimensions and Euclidean point configurations

Bruno Predojević

Abstract

We study two related quantities which generalize the concept of upper Banach density of a set, to two measurable subsets of the plane. The first of them allows us to generalize a classic result on sufficiently large distances realized in a set of positive upper density, to distances between points of two sets satisfying an appropriate density condition. The second one allows us to show that for all sufficiently large scales $t > 0$ and for a smooth, closed, centrally symmetric, planar curve Γ which bounds a convex and compact region in the plane and is of non-vanishing curvature, the family consisting of portions of translates of $t\Gamma$ has the maximal possible Vapnik–Chervonenkis dimension.

1 Introduction

1.1 A brief overview of previous results

In areas as diverse as geometric measure theory, ergodic theory and combinatorics, one often strives to capture the intuitive notion of the proportion of space occupied by a given subset. One of the ways to do this rigorously is by considering the so-called upper Banach density, which is defined for a measurable set $A \subseteq \mathbb{R}^d$ as

$$\overline{\delta}(A) = \limsup_{R\to\infty} \sup_{z\in\mathbb{R}^d} \frac{|A \cap (z + [0, R]^d)|}{R^d}.$$

It is possible to show that the limit superior in the above definition can be replaced by an actual limit. Székely conjectured that all planar sets with strictly positive upper Banach density, necessarily realize all sufficiently large distances between pairs of their points. The conjecture was further popularized by Erdős in [14], sparking developments across a broad range of fields. Eventually, three distinct proofs were found: one by Furstenberg, Katznelson and Weiss [17]; one by Falconer and Marstrand [15]; and one by Bourgain [3]. In fact, Bourgain's proof solved a generalized version of the conjecture, which states that all measurable sets $A \subseteq \mathbb{R}^{d+1}$ of strictly positive upper Banach density necessarily contain the vertices of all sufficiently large dilates of an isometric copy of any non-degenerate d-dimensional simplex. This result sparked a whole series of investigations, determining which finite configurations have the property that isometric copies of all their sufficiently large dilations can be found inside sets of positive upper Banach density. Graham in [19] gave the currently most general negative answer known, namely that configurations which cannot be inscribed inside a unit sphere in $\mathbb{R}^d$, cannot satisfy the analogue of Bourgain's simplex theorem. The most general positive result is due to Lyall and Magyar [27] and it concerns configurations which are Cartesian products of simplices. A fruitful path to further generalizations opens up if one is satisfied with dropping the condition of finding isometric copies of the studied configuration and instead considers "flexible configurations", namely distance graphs, which can be thought of as graphs which carry information about the lengths of their edges. The study of Bourgain-type results for distance graphs was pioneered by Lyall and Magyar in [28]. In [25], among results for rigid configurations, namely simplices and boxes, Kovač obtained results for anisotropic scaling of a special class of distance graphs, namely distance trees, while in [26] Kovač and the present author obtained dimensionally sharp embedding results for yet another special class of distance graphs, namely hypercube graphs. In this paper, we propose a new direction for generalization of these configuration results. Instead of having a single ambient set in which we are to look for given patterns, we study two fixed ambient sets and configurations which in some prescribed way have some of their vertices lie in one and the remaining ones in the other set.

1.2 Statement of the first main result

The simplest configuration which one might consider finding between two sets is the configuration consisting of only two points, i.e., we ask under which conditions on measurable subsets $A, B \subseteq \mathbb{R}^2$ can we say that all sufficiently large distances between them are realized. One might think that asking that both sets have strictly positive upper Banach density might be enough to guarantee that an analogue of the aforementioned result (answering Székely's question) holds, however it is rather straightforward to construct a counterexample. We let A be the union of all balls centered around $(2^n, 0)$ of radius n and we let B be the union of balls centered at $(-2^n, 0)$ of radius n, where n ranges over the positive integers. Numerous other naive attempts of defining a certain "averaged" upper density of A and B fail by simply taking the first set to be the union of annuli $10^n < |x| < 2 \cdot 10^n$ and the second set to be the union of the annuli $6 \cdot 10^n < |x| < 7 \cdot 10^n$ when n ranges over the positive integers. It turns out that it is crucial to have control over "how close" the points can get to each other.

Definition 1.1. *For measurable sets $A, B \subseteq \mathbb{R}^2$, we define their joint upper Banach density as*

$$\overline{\delta}(A,B) = \sup_{M\geqslant 1} \limsup_{R\to+\infty} \sup_{z\in\mathbb{R}^2} \inf_{M\leqslant r\leqslant \frac{R}{2}} \frac{1}{(2r)^2R^2} \int_{z+[0,R]^2} \int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(y)\mathbb{1}_{[-r,r]^2}(x-y)\,\mathrm{d}y\,\mathrm{d}x.$$

A consequence of Proposition 3.2 below tells us that, for a measurable set $A \subseteq \mathbb{R}^2$, it follows that

$$\overline{\delta}(A,A) > 0 \quad \Longleftrightarrow \quad \overline{\delta}(A) > 0,$$

thus the qualitative condition of a set having positive upper Banach density generalizes to a set having strictly positive joint upper Banach density, in the sense of Definition 1.1. Thus Theorem 1.2 below is indeed a generalization of the previously mentioned classic result.

Theorem 1.2 (Two-set Székely's problem)**.** *Let $A, B \subseteq \mathbb{R}^2$ be measurable sets such that $\overline{\delta}(A,B) > 0$. Then, there exists a large enough $\lambda_0 = \lambda_0(A,B) > 0$ such that for all $\lambda \geqslant \lambda_0$, there exist $x \in A$ and $y \in B$ such that*

$$|x-y| = \lambda.$$

A prototypical example of sets A and B for which $\overline{\delta}(A,B) > 0$ are the squares of a chessboard pattern in the plane.

1.3 VC dimension and statement of second main result

The concept of VC dimension was first introduced by Vapnik and Chervonenkis in [34] and it is used to formalize the idea of complexity of a family of sets. Here, we briefly restate it. For an excellent brief explanation of the connections with statistical learning theory, we advise the reader to look at the introduction of [16].

Definition 1.3. *Let X be some ambient set and let $\mathcal{T} \subseteq \mathcal{P}(X)$ be some family of subsets of X. We say that a finite set C of cardinality $n \in \mathbb{N}$ is shattered by $\mathcal{T}$ if for every $D \subseteq C$, there exists a $T \in \mathcal{T}$ such that*

$$T \cap C = D.$$

Definition 1.4. *Let X be some ambient set and let $\mathcal{T} \subseteq \mathcal{P}(X)$ be some family of subsets of X. We say that $\mathcal{T}$ has Vapnik—Chervonenkis (VC) dimension equal to n if n is the largest positive integer for which there exists a finite set $C \subseteq X$ of cardinality n which is shattered by $\mathcal{T}$; we denote this by*

$$VC(\mathcal{T}) = n.$$

If no such positive integer n exists, then we set

$$VC(\mathcal{T}) = 0.$$

If arbitrarily large such positive integers n exist, then we set

$$VC(\mathcal{T}) = \infty.$$

Euclidean geometry provides us with the following elementary example of a family for which we can explicitly determine the VC dimension. Let $E \subseteq \mathbb{R}^d$ and $t > 0$ be arbitrary. Consider the family of E-centered spheres of radius t intersected with E,

$$\mathcal{T}_t(E) = \{S_t(e) \cap E : e \in E\},$$

where

$$S_t(e) = \{y \in \mathbb{R}^d : |e - y| = t\}.$$

From elementary geometry, chiefly the fact that $d + 1$ distinct points uniquely determine a $(d-1)$-sphere in d dimensions, one can easily deduce that

$$VC(\mathcal{T}_t(\mathbb{R}^d)) = d + 1$$

for every $t > 0$. The interested reader might also consult Definition 1.5.5. and the comments thereafter in [30].

In [16], Fitzpatrick, Iosevich, McDonald and Wyman made the connection between VC dimension of families of circles and Ramsey theory-type questions, namely they asked how large, where "large" should be understood in a suitable sense, must the set E be, before it becomes unavoidable for the family $\mathcal{T}_t(E)$ to have the maximal possible VC dimension. They showed that in the setting of 2-dimensional vector spaces over finite fields $\mathbb{F}_q^2$, with the appropriate modifications of the above definitions, if one has $|E| \geqslant Cq^{\frac{15}{8}}$, then

$$VC(\mathcal{T}_t(E)) = 3,$$

which can be seen to be the maximal dimension possible in their setting. An analogous version of the above claim in higher dimensional vector spaces over finite fields was studied [1]. In fact, Ramsey theory-type questions relating to VC dimensions have, in recent years, sparked a whole series of investigations, such as [2], [21], [23], [31]. For further results and a nice overview of problems in the context of finite fields, we advise the reader to look at McDonald's doctoral dissertation [30], where the reader can also find results in the Euclidean setting, with the notion of largeness measured by the Hausdorff dimension. In fact, measuring largeness via Hausdorff dimension seems to be a very fruitful approach in the Euclidean setting; the interested reader can consult [22].

In this paper, we will focus on the two dimensional Euclidean case where $VC(\mathcal{T}_t(\mathbb{R}^2)) = 3$. One of the easy consequences of Theorem 1.6 below is the fact that if $E \subseteq \mathbb{R}^2$ satisfies $\overline{\delta}(E) > 0$, then $VC(\mathcal{T}_t(E)) = 3$, for all sufficiently large t. Therefore, having strictly positive upper Banach density seems to be another suitable measure of largeness in the Euclidean setting, if one is interested in the aforementioned Ramsey theory-type questions. As with the generalization of Székely's problem, we are interested in generalizing this claim to two sets. Our results could very well be stated with circles, but it turns out that we can be even more general and can replace translates of the circle by translates of the curve Γ which is essentially circle-like. More precisely we let Γ be a smooth, closed, and centrally symmetric planar curve which also has non-vanishing curvature and is the boundary of a convex and compact region in the plane with a non-empty interior. So, we study translates of $t\Gamma$ by vectors from a measurable set $B \subseteq \mathbb{R}^2$, with the additional caveat that we do not consider the entirety of the curve $b + t\Gamma$, but only a portion of its arc, which lies inside a fixed measurable set A. Thus, for a fixed $t > 0$, we consider the family

$$\mathcal{T}_t(A, B) = \{(b + t\Gamma) \cap A : b \in B\}.$$

It is quite easy to see that our family can be empty for certain choices of A and B, so we are yet again interested in finding a condition on A and B which will guarantee that for all sufficiently large $t > 0$, not only are the families $\mathcal{T}_t(A, B)$ non-trivial, but they have the maximal possible VC dimension, which as another consequence of Theorem 1.6 below turns out to be 3. We now present a quantity $\overline{\delta}_{VC}(A, B)$, positivity of which is sufficient if one wishes to obtain the desired Ramsey theory-type result. We start by introducing some helpful notation. For measurable sets $A, B \subseteq \mathbb{R}^2$ and for $x, v_1, v_2, v_3, s_1, s_2, s_3 \in \mathbb{R}^2$, we define $\mathcal{F}_{A,B}(x; v_1, v_2, v_3; s_1, s_2, s_3)$ as the product

$$\begin{aligned}
&\mathbb{1}_B(x)\mathbb{1}_A(x + v_1)\mathbb{1}_A(x + v_2)\mathbb{1}_A(x + v_3)\\
&\mathbb{1}_B(x + v_1 + v_2)\mathbb{1}_B(x + v_1 + v_3)\mathbb{1}_B(x + v_2 + v_3)\\
&\mathbb{1}_B(x + v_1 + s_1)\mathbb{1}_B(x + v_2 + s_2)\mathbb{1}_B(x + v_3 + s_3).
\end{aligned}$$

This product corresponds to a particular placement of vertices of a flexible configuration from Theorem 1.7 below between sets A and B; we advise the reader to look at Figure 1.

Definition 1.5. *For two measurable sets $A, B \subseteq \mathbb{R}^2$, we define*

$$\overline{\delta}_{VC}(A, B) := \sup_{M\geqslant 1} \limsup_{R\to+\infty} \sup_{z\in\mathbb{R}^2} \inf_{M\leqslant r\leqslant 2^{-6}R} \frac{1}{(2r)^{12}R^2} \int_{z+[0,R]^2} \int_{([-r,r]^2)^6} \mathcal{F}_{A,B}(x; v_1, v_2, v_3; s_1, s_2, s_3) \,\mathrm{d}v_1\,\mathrm{d}v_2\,\mathrm{d}v_3\,\mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3\,\mathrm{d}x.$$

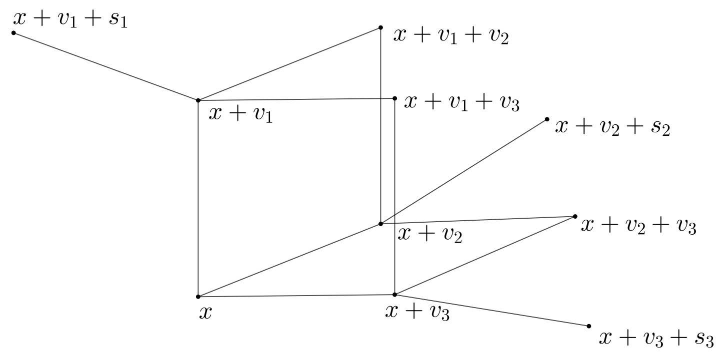

Figure 1: Flexible configuration from Theorem 1.7

Another consequence of Proposition 3.2 below tells us that for a measurable set $E \subseteq \mathbb{R}^2$, we have the chain of equivalences

$$\overline{\delta}(E) > 0 \Longleftrightarrow \overline{\delta}_{VC}(E, E) > 0 \Longleftrightarrow \overline{\delta}_{VC}(\mathbb{R}^2, E) > 0 \Longleftrightarrow \overline{\delta}_{VC}(E, \mathbb{R}^2) > 0,$$

which tells us that the qualitative assumption of quantity from Definition 1.5 being strictly positive for $A = B = E$ is equivalent to the condition of E being a set of strictly positive upper Banach density, so our new quantity is in a certain sense also a generalization of upper Banach density. Our second main results is

Theorem 1.6. *Let $A, B \subseteq \mathbb{R}^2$ be measurable and such that $\overline{\delta}_{VC}(A, B) > 0$. Let Γ be a planar curve which is closed, smooth, centrally symmetric and which is the boundary of some convex region in the plane, with a non-empty interior. There exists a critical scale $t_0 > 0$ such that, for all $t \geqslant t_0$, it follows that $\mathcal{T}_t(A, B)$ has the Vapnik–Chervonenkis dimension equal to* 3.

With the help of elementary results from differential geometry, which we briefly recall in Section 2.2, we show that Theorem 1.6 reduces to the following theorem about flexible configurations in $\mathbb{R}^2$, which is analogous to the distance graph result from [26] and which would, by minor modifications, already imply a version of Theorem 1.6 in case $A = B$ and Γ being the unit circle.

Theorem 1.7 (VC dimension, configuration variant)**.** *Let $A, B \subseteq \mathbb{R}^2$ be measurable sets which satisfy $\overline{\delta}_{VC}(A, B) > 0$. Let Γ be a planar curve which is closed, smooth, centrally symmetric and which is the boundary of some convex region in the plane, with a non-empty interior. There exists a critical scale $t_0 > 0$ such that, for all $t \geqslant t_0$, there exists a point $x \in B$ and vectors $v_1, v_2, v_3, s_1, s_2, s_3 \in t\Gamma$, such that*

$$x, x + v_1 + v_2, x + v_1 + v_3, x + v_2 + v_3, x + v_1 + s_1, x + v_2 + s_2, x + v_3 + s_3 \in B \tag{1.1}$$

and

$$x + v_1, x + v_2, x + v_3 \in A \tag{1.2}$$

and that all 10 *points listed above are mutually distinct.*

To prove Theorem 1.7, we follow the approach which has its roots in the work of Bourgain from [3] and which was further developed by Cook, Magyar and Pramanik in [4]; namely we shall introduce a certain integral counting form, which we strive to bound from below by approximating it with its suitably smoothed out versions. This approach was further polished and applied in a series of papers by Durcik, Kovač and Rimanić, such as [10], [8], [9], [25] and in a series of papers by Lyall and Magyar such as [27], [29], [28]. We emphasize that the key part of the aforementioned approximation is a bound for a particular singular Brascamp–Lieb form. The study of such forms has seen significant development in the last two decades, starting from the pioneering work of Kovač in [24] and ranging through further improvements by Durcik, Kovač, Thiele and Slavikova in a series of papers such as [6], [7], [12], [11], where [11] is currently the most general known result. In principle, the estimate we shall require follows from Lemma 2 from [12] by considering certain skew projections, which would again lead to a somewhat elaborate argument. Therefore, we prefer to give a short self-contained proof, while fully acknowledging that the approach is by no means new or original.

1.4 Organization of the paper

We briefly comment on how the rest of the paper is organized. Section 2 focuses on preliminary results and is subdivided into Subsection 2.1, in which we recall basic facts about the Fourier transform and Gaussian function, and Subsection 2.2, in which we recall basic facts from differential geometry of smooth curves in the plane, as well as prove Lemma 2.6 which, as mentioned, will be the key link in showing that Theorem 1.7 implies Theorem 1.6. Section 3 is devoted to the condition which is a simultaneous generalization of Definitions 1.1 and 1.5 and which allows us to prove Proposition 3.2, which shows that our new conditions indeed generalize the condition of a set having positive upper Banach density. Section 4 is devoted to the proof of Theorem 1.2. Section 5 is concerned with the proof of Theorem 1.6.

2 Preliminaries and notation

We shall write $A \lesssim_P B$ if the inequality $A \leqslant C_P B$ holds for some unimportant finite constant $C_P > 0$, which is allowed to depend on a set of parameters P. Similarly, we write $A \gtrsim_P B$ if $B \lesssim_P A$ holds and we also write $A \sim_P B$ if both $A \lesssim_P B$ and $A \gtrsim_P B$ hold. We shall write $\mathbb{1}_A$ for the indicator function and $\mathcal{P}(A)$ for the power set of a set A and write $\operatorname{Int}(A)$ for the interior of A. Since both problems we study concern the Euclidean plane $\mathbb{R}^2$, unless specified otherwise, the reader can assume that all our functions have $\mathbb{R}^2$ as their domain. For $r \in \{0,1\}^n$ and $v \in (\mathbb{R}^2)^n$ we often write

$$r \cdot v = r_1 v_1 + r_2 v_2 + \cdots + r_n v_n,$$

where, even if not explicitly, we assume that all tuples, either of numbers or of vectors, have coordinates denoted with the same letter but with indices, such as for instance $r = (r_1, r_2, \ldots, r_n)$ and $v = (v_1, v_2, \ldots, v_n)$. We also use the same symbol to denote both the above product and the standard Euclidean inner product on $\mathbb{R}^2$ and we trust that it will be clear from context which meaning we assign to it. We shall make use of the following "freeze functions". For positive integers $k \leqslant n$ and $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$ mutually distinct and for $a_1, a_2, \ldots, a_k \in \{0,1\}$, we consider the function $C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)} : \{0,1\}^n \to \{0,1\}^n$

defined in such a way that if

$$C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r_1, r_2, \ldots, r_n) = (s_1, s_2, \ldots, s_n),$$

then for $j \in \{1, 2, \ldots, n\}$, it follows that

$$s_j = \begin{cases} a_{i_l}, & j = i_l \text{ for some } l \in \{1, 2, \ldots, k\}, \\ r_j, & \text{otherwise.} \end{cases}$$

By $\mathrm{co}(v_1, v_2, \ldots, v_n)$ we denote the convex hull of points $v_1, v_2, \ldots, v_n \in \mathbb{R}^2$.

2.1 Gaussian functions and Fourier analysis

We use the following normalization of the Fourier transform for an integrable function f:

$$\widehat{f}(\xi) = \int_{\mathbb{R}^2} f(x) e^{-2\pi i x\cdot\xi} \,\mathrm{d}x.$$

It is a well-known fact that under this normalization, if we assume that f belongs to the Schwartz class $\mathcal{S}(\mathbb{R}^2)$, then the Fourier reconstruction formula holds pointwise, that is to say

$$f(x) = \int_{\mathbb{R}^2} \widehat{f}(\xi) e^{2\pi i x\cdot\xi} \,\mathrm{d}\xi,$$

for all $x \in \mathbb{R}^2$. Similarly, for a Borel measure μ, we use the following normalization of the Fourier transform

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i x\cdot\xi} \,\mathrm{d}\mu(x).$$

For $\lambda > 0$, an integrable function f and a Borel measure μ, we denote their respective dilations by

$$f_\lambda(x) = \lambda^{-2} f(\lambda^{-1}x), \quad \mu_\lambda(E) := \mu(\lambda^{-1}E)$$

and we define their convolution as

$$(\mu * f)(x) = \int_{\mathbb{R}^2} f(x-y) \,\mathrm{d}\mu(y).$$

By $\mathbb{g}$ we denote the standard Gaussian on $\mathbb{R}^2$ and by $\mathbb{k}$ we denote its Laplacian, while for $i = 1, 2$, by $\mathbb{h}^{(i)}$ we denote its i-th partial derivative. That is to say

$$\mathbb{g}(x) := e^{-2\pi|x|^2}, \quad \mathbb{h}^{(i)}(x) := \partial_i \mathbb{g}(x), \quad \mathbb{k}(x) := \Delta \mathbb{g}(x).$$

The key properties of the Gaussian functions used in the proofs below are the following Gaussian convolution identities, which appear in [9, Section 5]. Let $\alpha, \beta > 0$, then

$$\mathbb{g}_\alpha * \mathbb{g}_\beta = \mathbb{g}_{\sqrt{\alpha^2+\beta^2}}, \tag{2.1}$$

$$\sum_{l=1}^{2} \mathbb{h}^{(l)}_\alpha * \mathbb{h}^{(l)}_\beta = \frac{\alpha\beta}{\alpha^2+\beta^2} \mathbb{k}_{\sqrt{\alpha^2+\beta^2}}, \tag{2.2}$$

$$\mathbb{k}_\alpha * \mathbb{g}_\beta = \frac{\alpha^2}{\alpha^2+\beta^2} \mathbb{k}_{\sqrt{\alpha^2+\beta^2}}. \tag{2.3}$$

The identities follow by passing to the Fourier side and performing a simple calculation. We also make use of the following inequalities, which can be found in [9, Section 5], as well. For $t, s, r, \lambda > 0$ that satisfy $t\lambda \sim r \sim s$, it follows that

$$(\mu_t * \mathrm{g}_{t\lambda})(x) \lesssim \varepsilon^{-3} \int_1^\infty \mathrm{g}_{s\gamma}(x) \,\frac{\mathrm{d}\gamma}{\gamma^2}, \tag{2.4}$$

$$(\mu_t * \mathrm{g}_{r})(x) \lesssim \varepsilon^{-3} \int_1^\infty \mathrm{g}_{s\gamma}(x) \,\frac{\mathrm{d}\gamma}{\gamma^2}, \tag{2.5}$$

where μ denotes a Borel probability measure. By direct calculation, we check that the reparametrized heat equation is satisfied by the Gaussian functions, i.e., it holds that

$$\frac{\partial}{\partial t}\mathrm{g}_t(x) = \frac{1}{2\pi t}\mathbb{k}_t(x). \tag{2.6}$$

The following elementary lemma will play a crucial role in the proof of Theorem 1.2. The proof follows by combining an elementary dyadic decomposition of the positive real numbers and the trivial estimate for tails of Schwartz functions.

Lemma 2.1 (Essential disjointness of Gaussian functions)**.** *Let $J \in \mathbb{N}$ and let $\eta_1, \eta_2, \ldots, \eta_J \in \mathbb{R}^2$ be such that for all $j \in \{1, 2, \ldots, J-1\}$*

$$2|\eta_j| \leqslant |\eta_{j+1}|.$$

Then, the following bound, which we stress is uniform in the number of variables, holds

$$\sum_{j=1}^{J} |\widehat{\mathbb{k}}(\eta_j)| \lesssim 1.$$

Lemma 2.2 (Fourier dimension estimate)**.** *Let μ be a probability measure supported on a closed, smooth curve of non-vanishing curvature, which is the boundary of some compact and convex set in the plane. Then, the Fourier transform of μ satisfies*

(a) $\sup_{\xi\in\mathbb{R}^2} |\xi|^{\frac{1}{2}} |\widehat{\mu}(\xi)| < +\infty$.

(b) $\big|\widehat{\mu}(\lambda\xi)\widehat{\mathbb{k}}(t\lambda\xi)\big| \lesssim t^{\frac{1}{2}}$*, for all $\xi \in \mathbb{R}^2$, and all $t, \lambda > 0$.*

The proof of (a) is quite standard and can be found as Theorem 1 in Section 8.3 in [32], while (b) follows quite easily from (a); the interested reader may consult [25] or [26] for details. One measure which satisfies the conditions of the above lemma, and which we make frequent use of, is the uniform probability measure supported on the unit circle centered at the origin. We denote this measure by σ.

Lastly, we recall some basic tools from additive combinatorics. Let $f \colon \mathbb{R}^2 \to [0, 1]$ be a compactly supported measurable function. For $n \in \mathbb{N}$, we denote

$$\mathcal{F}_n(x; v_1, \ldots, v_n) := \prod_{r\in\{0,1\}^n} f(x + r_1 v_1 + r_2 v_2 + \cdots + r_n v_n),$$

for $x, v_1, \ldots, v_n \in \mathbb{R}^2$. These products are the basic building blocks of the so-called *Gowers uniformity norms* $\|\cdot\|_{\mathrm{U}^n}$, which are defined as

$$\|f\|_{\mathrm{U}^n(\mathbb{R}^2)} := \Big(\int_{(\mathbb{R}^2)^{n+1}} \mathcal{F}_n(x; v_1, \ldots, v_n) \,\mathrm{d}x \,\mathrm{d}v_1 \cdots \mathrm{d}v_n \Big)^{\frac{1}{2^n}}. \tag{2.7}$$

Uniformity norms are usually studied in the context of locally-compact Abelian groups and in this general setting, one can prove the so called Gowers–Cauchy–Schwarz inequality; the interested reader might look at [18, 20] or [13]. For our purposes, the following special case, the proof of which follows from inductively applying the Cauchy–Schwarz inequality in $\mathbb{R}^2$, will suffice:

$$\left| \int_{(\mathbb{R}^2)^{n+1}} \prod_{r\in\{0,1\}^n} f_r(x + r\cdot v)\,\mathrm{d}x\,\mathrm{d}v_1 \cdots \mathrm{d}v_n \right| \leqslant \prod_{r\in\{0,1\}^n} \|f_r\|_{\mathrm{U}^n(\mathbb{R}^2)}, \tag{2.8}$$

where $(f_r)_{r\in\{0,1\}^n}$ are bounded, real-valued and compactly supported measurable functions. By testing the above inequality with suitably chosen indicator functions of cubes (see [26] for more elaboration), we can conclude that

$$\|f\|_{\mathrm{U}^n(\mathbb{R}^2)} \gtrsim_n |Q|^{-1+(n+1)/2^n} \int_Q f(x)\,\mathrm{d}x, \tag{2.9}$$

where f is a positive function, supported on the cube Q.

2.2 Geometry of curves in the plane

Throughout this section, let Γ be a closed, smooth, centrally symmetric curve with non-vanishing curvature, which is the boundary of some compact and convex planar set with non-empty interior. Let μ be the probabilistically normalized (uniform) arclength measure supported on Γ. The following proposition is geometrically quite self-evident, however a rigorous proof requires some theory. The standard reference is a book by do Carmo [5], and the interested reader might consult Sections 1–7; chiefly Exercise 9 and 5–7; chiefly Proposition 1.

Proposition 2.3. *Let Γ be a closed, smooth planar curve which is the boundary of some compact set C with a non-empty interior. Then, C is convex if and only if, the so-called, "supporting line" property holds, namely for every point $a \in \Gamma$, there exists a unique vector $n_a \in \mathbb{R}^2$ such that for all $b \in \Gamma \setminus \{a\}$, it follows that*

$$n_a \cdot (b - a) < 0,$$

or geometrically, the entirety of the curve Γ, except for point a, is entirely contained in one of the two open half-planes determined by the line $\ell_a := a + \{n_a\}^{\perp}$.

Lemma 2.4. *Every line l in $\mathbb{R}^2$ intersects Γ in at most two points.*

Proof. Assume that l intersects Γ in three mutually distinct points x, y and z. Without loss of generality, we can assume that there exists some $t \in (0,1)$ such that $z = tx + (1-t)y$. Since z belongs to Γ, from Proposition 2.3 we conclude that there exists some $n \in \mathbb{R}^2$ such that for all $b \in \Gamma \setminus \{z\}$, it follows that $n \cdot (b - z) < 0$. In particular, this leads to a contradiction:

$$0 = n \cdot ((tx + (1-t)y) - z) = t(n \cdot (x - z)) + (1-t)(n \cdot (y - z)) < 0. \qquad \square$$

Lemma 2.5. *Let $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{R}^2$ be mutually distinct points which satisfy*

$$b_1 - a_1 = b_2 - a_2 = b_3 - a_3.$$

Then one of these points lies in the convex hull spanned by some three of the remaining points.

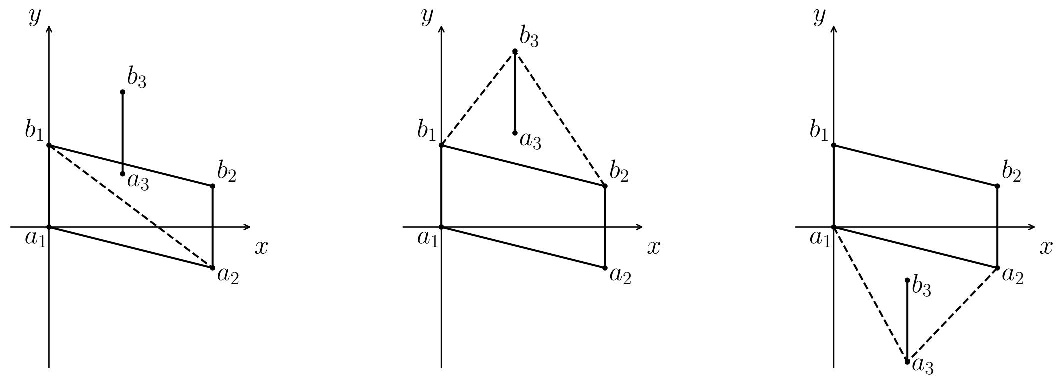

Figure 2: Sketch of proof of Lemma 2.5

Proof. By making an appropriate change of coordinates, the claim reduces to checking 3 distinct cases which correspond to different arrangements of the given points and which are illustrated in Figure 2. □

As it has been mentioned in the introduction, the following lemma is crucial for the reduction of Theorem 1.6 to Theorem 1.7.

Lemma 2.6. *Let $\Gamma \subseteq \mathbb{R}^2$ be as above.*

(a) For any mutually distinct $v_1, v_2 \in \mathbb{R}^2$, the intersection $(v_1 + \Gamma) \cap (v_2 + \Gamma)$ consists of at most two points.

(b) For two mutually distinct points $x, y \in \mathbb{R}^2$, there exist at most two translates of the curve Γ which contain both x and y.

Proof of (a). Assume that there exist at least three distinct intersections $x, y, z \in (v_1 + \Gamma) \cap (v_2 + \Gamma)$. Then, we can write

$$x = v_1 + \gamma_{1,x} = v_2 + \gamma_{2,x}, \qquad y = v_1 + \gamma_{1,y} = v_2 + \gamma_{2,y}, \qquad z = v_1 + \gamma_{1,z} = v_2 + \gamma_{2,z}, \tag{2.10}$$

for some $\gamma_{1,x}, \gamma_{1,y}, \gamma_{1,z}, \gamma_{2,x}, \gamma_{2,y}, \gamma_{2,z} \in \Gamma$. We claim that at least two of these points are equal. Assume on the contrary that all 6 points are mutually distinct. By Lemma 2.5, it follows that there exist $a, b, c, d \in \{\gamma_{1,x}, \gamma_{1,y}, \gamma_{1,z}, \gamma_{2,x}, \gamma_{2,y}, \gamma_{2,z}\}$ all mutually distinct and such that $a \in \mathrm{co}(b, c, d)$.

Firstly, notice that it necessarily follows that $a \in \mathrm{Int}(\mathrm{co}(b, c, d))$, because if a was an element from the boundary of the convex hull, this would imply the existence of the line which intersects Γ in three points, which we know can not exists because of Lemma 2.4. For instance if $a \in [b, d]$, then it lies on the line spanned by b and d and we obtain our contradiction.

Let ℓ_a be the supporting line to Γ at a, which we know exists because of Proposition 2.3, and let $\mathcal{H}$ be an open half-plane which contains $\Gamma \setminus \{a\}$. However, since $\mathcal{H}$ is a convex set, it follows that

$$a \in \mathrm{Int}(\mathrm{co}(b, c, d)) \subseteq \mathrm{co}(b, c, d) \subseteq \mathcal{H},$$

which is in contradiction with Proposition 2.3.

Therefore, we know that the set $\{\gamma_{1,x}, \gamma_{1,y}, \gamma_{1,z}, \gamma_{2,x}, \gamma_{2,y}, \gamma_{2,z}\}$ consists of at most 5 points. Thus, two of the elements listed above actually equal each other. We consider several possibilities how this might occur and show that they all lead to a contradiction. If for $a \in \{x, y, z\}$ it follows that $\gamma_{1,a} = \gamma_{2,a}$, then by (2.10), we conclude that $v_1 = v_2$, which contradicts the assumption that v_1 and v_2 are mutually distinct. If for mutually distinct $a, b \in \{x, y, z\}$ it follows that $\gamma_{1,a} = \gamma_{2,b}$, again by (2.10), we would get that

$$\gamma_{1,a} = \frac{1}{2}(\gamma_{1,b} + \gamma_{2,a}),$$

which would imply that the line determined by $\gamma_{1,b}$ and $\gamma_{2,a}$ also contains $\gamma_{1,a}$, which contradicts Lemma 2.4, since all three points $\gamma_{1,a}, \gamma_{1,b}$ and $\gamma_{2,a}$ lie on the curve Γ. Finally, if for mutually distinct $a, b \in \{x, y, z\}$ it either follows that$\gamma_{1,a} = \gamma_{1,b}$ or $\gamma_{2,a} = \gamma_{2,b}$, then, once more, (2.10) implies that either $\gamma_{2,a} = \gamma_{2,b}$ or $\gamma_{1,a} = \gamma_{1,b}$ and hence in both cases $a = b$, contradicting our assumption. Since we have exhausted all the possible cases of equality, this concludes the proof. □

Proof of (b). Assume the contrary, that is that there exist mutually distinct vectors $v_1, v_2, v_3 \in \mathbb{R}^2$ such that

$$x, y \in v_1 + \Gamma, v_2 + \Gamma, v_3 + \Gamma.$$

By central symmetry of Γ, we conclude that

$$v_1, v_2, v_3 \in (x + \Gamma) \cap (y + \Gamma).$$

However, this is a contradiction with (a), because we have assumed that v_1, v_2, v_3 are mutually distinct. □

3 Common generalization of $\overline{\delta}(A, B)$ and $\overline{\delta}_{VC}(A, B)$

In this section we explore the common thread between Definitions 1.1 and 1.5. In fact, we will study a slightly more general quantity which we now define.

Definition 3.1 (Joint upper Banach density)**.** *For $n \in \mathbb{N}$ and 2^n locally integrable functions $(f_r)_{r\in\{0,1\}^n}$, we consider the quantity*

$$\overline{\delta}((f_r)_{r\in\{0,1\}^n}) = \sup_{M\geqslant 1} \limsup_{R\to+\infty} \sup_{z\in\mathbb{R}^2} \inf_{M\leqslant r\leqslant 2^{-n}R} \frac{1}{(2r)^{2n}R^2} \int_{z+[0,R]^2} \int_{([-r,r]^2)^n} \mathcal{F}_n(x; v_1, \ldots, v_n) \, \mathrm{d}v_1 \ldots \mathrm{d}v_n \, \mathrm{d}x, \quad (3.1)$$

where for $x, v_1, \ldots, v_n \in \mathbb{R}^2$, we have slightly abused notation introduced in Section 2.1, by defining

$$\mathcal{F}_n(x; v_1, \ldots, v_n) := \prod_{(r_1,r_2,\ldots,r_n)\in\{0,1\}^n} f_{(r_1,r_2,\ldots,r_n)}(x + r_1v_1 + \cdots + r_nv_n).$$

We note that in the definition of upper Banach density we are essentially calculating limits of averages of an indicator function over a given cube. The above quantity generalizes this idea by looking at averages of multiple functions, whose variables are tied up in a hypercubic structure. It is easy to see that by picking various different indicator functions, we are essentially considering versions of upper Banach density for multiple sets. Indeed, for measurable sets $A, B \subseteq \mathbb{R}^2$ and for $n = 1$, we recover Definition 1.1, i.e.,

$$\overline{\delta}(A, B) = \overline{\delta}(\mathbb{1}_A, \mathbb{1}_B).$$

We also briefly comment that Definition 3.1 generalizes Definition 1.5 in the case when $n = 6$. More precisely, if we write v_4, v_5, v_6 instead of s_1, s_2, s_3 respectively and if we set

$$\begin{aligned} f_{(1,0,0,0,0,0)} &= f_{(0,1,0,0,0,0)} = f_{(0,0,1,0,0,0)} = \mathbb{1}_A, \\ f_{(0,0,0,0,0,0)} &= f_{(0,0,0,1,0,0)} = f_{(0,0,0,0,1,0)} = f_{(0,0,0,0,0,1)} \\ &= f_{(1,1,0,0,0,0)} = f_{(1,0,1,0,0,0)} = f_{(0,1,1,0,0,0)} = \mathbb{1}_B \end{aligned} \tag{3.2}$$

and for all other tuples $(r_1, r_2, r_3, r_4, r_5, r_6)$ not listed here, we set $f_{(r_1,r_2,\ldots,r_6)} = \mathbb{1}_{\mathbb{R}^2}$, then we can notice that

$$\overline{\delta}((f_r)_{r\in\{0,1\}^6}) = \overline{\delta}_{VC}(A, B).$$

We prefer to use both the full hypercube notation and the notation which highlights the vertices of interest to us, depending on the context. We turn our attention to showing the proposition which helped us conclude that the qualitative conditions of our new quantities being positive truly generalizes the qualitative condition of a set having strictly positive upper Banach density.

Proposition 3.2. *Consider a measurable set $A \subseteq \mathbb{R}^2$. For $n \in \mathbb{N}$, also consider 2^n locally integrable functions $(f_r)_{r\in\{0,1\}^n}$, where for each $r \in \{0,1\}^n$, $f_r \in \{\mathbb{1}_A, \mathbb{1}_{\mathbb{R}^2}\}$. Assume that at least one f_r equals $\mathbb{1}_A$. We claim that*

$$\overline{\delta}(A)^{\frac{n+1}{2^n}} \gtrsim_n \overline{\delta}((f_r)_{r\in\{0,1\}^n}) \gtrsim_n \overline{\delta}(A)^{2^n}.$$

Proof. From the definition of the upper Banach density, we conclude that there exists a sufficiently large $R_0 > 0$ such that for all $R \geqslant R_0$ and all $z \in \mathbb{R}^2$, we have that

$$\frac{|A \cap (z + [0, R]^2)|}{R^2} < 2\overline{\delta}(A), \tag{3.3}$$

Fix $M \geqslant 1$ and $z \in \mathbb{R}^2$ and denote

$$I(R, z) := \inf_{M\leqslant r\leqslant 2^{-n}R} \frac{1}{(2r)^{2n}R^2} \int_{z+[0,R]^2} \int_{([-r,r]^2)^n} \mathcal{F}_n(x; v_1, v_2, \ldots, v_n)\, \mathrm{d}v_1 \ldots \mathrm{d}v_n\, \mathrm{d}x.$$

By considering $r = 2^{-n}R$, we easily conclude that

$$I(R, z) \lesssim_n \frac{1}{(R^2)^{n+1}} \int_{(\mathbb{R}^2)^{n+1}} \prod_{r\in\{0,1\}^n} (\mathbb{1}_{(z-(R,R))+[0,2R]^2} \cdot f_r)(x + r \cdot v)\, \mathrm{d}v_1 \ldots \mathrm{d}v_n\, \mathrm{d}x.$$

Since at least one function in the above product is equal to $\mathbb{1}_{A\cap((z-(R,R))+[0,2R]^2)}$, we can bound all the functions on the remaining vertices of the hypercube by $\mathbb{1}_{(z-(R,R))+[0,2R]^2}$. After performing this bound and applying the Gowers–Cauchy–Schwarz inequality (2.8), followed by the assumption (3.3), we conclude that

$$\begin{aligned}
I(R,z) &\lesssim_n R^{-2n-2}\|\mathbb{1}_{A\cap((z-(R,R))+[0,2R]^2)}\|_{\mathrm{U}^n(\mathbb{R}^2)}\|\mathbb{1}_{(z-(R,R))+[0,2R]^2}\|_{\mathrm{U}^n(\mathbb{R}^2)}^{2^n-1} \\
&\lesssim R^{-2n-2}R^{\frac{n+1}{2^{n-1}}}\left(\frac{|A\cap((z-(R,R))+[0,2R]^2)|}{(2R)^2}\right)^{\frac{n+1}{2^n}}|((z-(R,R))+[0,2R]^2)|^{\frac{n+1}{2^n}(2^n-1)} \\
&\lesssim \overline{\delta}(A)^{\frac{n+1}{2^n}},
\end{aligned}$$

where we have used the fact that the uniformity norms of indicator functions of measurable sets can be bounded by appropriate powers of the measure of the set. Since this was true for all $z\in\mathbb{R}^2$ and $M\geqslant 1$, by taking a supremum in z, followed by a limit superior in R and finally a supremum in M, we conclude that

$$\overline{\delta}((f_r)_{r\in\{0,1\}^n})\lesssim_n \overline{\delta}(A)^{\frac{n+1}{2^n}}.$$

Next, we prove the lower bound. Unpacking the definition of the usual upper Banach density, we can find an increasing sequence of positive real numbers $(R_m)_{m\in\mathbb{N}}$ converging to infinity and a sequence $(z_m)_{m\in\mathbb{N}}$ in $\mathbb{R}^2$ such that for all $m\in\mathbb{N}$, we have

$$\int_{z_m+[0,R_m]^2}\mathbb{1}_A(x)\mathrm{d}x>\frac{\overline{\delta(A)}R_m^2}{2}. \tag{3.4}$$

Now fix some $M\geqslant 1$ and consider $m\in\mathbb{N}$ large enough so that $M<2^{-n}R_m$. For $M\leqslant r\leqslant 2^{-n}R_m$, we strive to bound

$$I(r,R_m,z_m):=\frac{1}{(2r)^{2n}R_m^2}\int_{(z_m+[0,R_m]^2)^n}\int_{([-r,r]^2)^n}\mathcal{F}_n(x;v_1,v_2,\ldots,v_n)\,\mathrm{d}v_1\ldots\mathrm{d}v_n\,\mathrm{d}x$$

from below by a positive constant multiple of $\overline{\delta}(A)^{2^n}$. To start with, we partition the cube $z_m+[0,R_m]^2=(z_1^m,z_2^m)+[0,R_m]^2$ into cubes of side-length r', where

$$r'=\frac{1}{2}\frac{R_m}{\left\lfloor\frac{R_m}{r}\right\rfloor}.$$

From elementary properties of the floor function, it follows that $\frac{r}{2}\leqslant r'\leqslant r$. Explicitly, we consider the family

$$\mathcal{Q}=\{[z_1^m+kr'',z_1^m+(k+1)r'')\times[z_2^m+lr'',z_2^m+(l+1)r''):k,l\in\{1,2\ldots,2\left\lfloor\frac{R_m}{r}\right\rfloor\}\},$$

noting that $\mathrm{card}(\mathcal{Q})=4\left\lfloor\frac{R_m}{r}\right\rfloor^2$. We begin by estimating all the functions in the product $\mathcal{F}_n(x;v_1,v_2,\ldots,v_n)$ from below by $\mathbb{1}_A$ and we also note that if we know that all 2^n points $x+r\cdot v$ lie inside the same cube Q of side-length r'', then the edges v_i lie inside the cube $[-r'',r'']^2$ and so we can actually estimate

$$I(r,R_m,z_m)\gtrsim_n\frac{1}{r^{2n}R_m^2}\sum_{Q\in\mathcal{Q}}\|\mathbb{1}_{A\cap Q}\|_{\mathrm{U}^n(\mathbb{R}^2)}^{2^n}.$$

Combining our inequality derived from Gowers–Cauchy–Schwarz (2.9) with Jensen's inequality and then applying 3.4, we obtain

$$I(r, R_m, z_m) \gtrsim_n \frac{r^{-2^{n+1}+2}|\mathcal{Q}|}{R_m^2} \left(\frac{1}{|\mathcal{Q}|} \sum_{Q \in \mathcal{Q}} \int_Q \mathbb{1}_A(x)\,\mathrm{d}x \right)^{2^n} \gtrsim_n \overline{\delta}(A)^{2^n}.$$

Since this was true for all $M \geqslant 1$, $m \in \mathbb{N}$ large enough and $M \leqslant r \leqslant 2^{-n} R_m$, we can first take the infimum in r and then take the limit superior in R and finally a supremum in M to conclude that

$$\overline{\delta}((f_r)_{r\in\{0,1\}^n}) \gtrsim_n \overline{\delta}(A)^{2^n}.$$

□

4 Proof of Theorem 1.2

We now turn to proving Theorem 1.2. In the spirit of Bourgain, we reduce this qualitative theorem to the following more quantitative formulation.

Theorem 4.1 (Two-set Szekély, quantitative version)**.** *Let $\delta \in (0,1)$ be arbitrary. Then, there exists $J = J(\delta)$ such that for all $\lambda_1, \lambda_2, \dots, \lambda_J > 1$ which satisfy $2\lambda_k \leqslant \lambda_{k+1}$ for all $k = 1, 2, \dots, J-1$, for all $R > 0$ such that $R \geqslant 2\lambda_J$ and for all $M \geqslant 1$ so that $\lambda_1, \lambda_2, \dots, \lambda_J \in [M, 2^{-1}R]$, the following holds: For an arbitrary $z \in \mathbb{R}^2$ and for all $A, B \subseteq z + [0,R]^2$ which satisfy*

$$\inf_{M \leqslant r \leqslant 2^{-1}R} \frac{1}{(2r)^2} \int_{z+[0,R]^2} \int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(y)\mathbb{1}_{[-r,r]^2}(x-y)\,\mathrm{d}y\,\mathrm{d}x > \delta R^2, \tag{4.1}$$

there exists $j \in \{1, 2, \dots, J\}$ such that for $\lambda = \lambda_j$, it follows that

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(x+v)\,\mathrm{d}\sigma_\lambda(v)\,\mathrm{d}x \gtrsim_\delta R^2. \tag{4.2}$$

Theorem 4.1 implies Theorem 1.2. Assume that Theorem 4.1 is true and that Theorem 1.2 is not. Then, there exist measurable sets $A, B \subseteq \mathbb{R}^2$ such that $\overline{\delta}(A,B) > 0$ and there exists a sequence of scales $(\lambda_m)_{m\in\mathbb{N}}$ such that for all $x \in A$ and $y \in B$, $|x-y| \neq \lambda_m$ for all $m \in \mathbb{N}$. By passing to a subsequence, we can assume that $\lambda_1 > 1$ and $2\lambda_m \leqslant \lambda_{m+1}$ for all $m \in \mathbb{N}$. We now apply Theorem 4.1 with $\delta = \frac{\overline{\delta}(A,B)}{4}$ to obtain $J \in \mathbb{N}$, for which the conclusion of Theorem 4.1 holds. Unpacking the definition of $\overline{\delta}(A,B)$, we find $M \geqslant 1$, a sequences $(R_m)_{m\in\mathbb{N}}$ of positive real numbers converging to infinity, and a sequence $(z_m)_{m\in\mathbb{N}}$ in $\mathbb{R}^2$ such that for all $m \in \mathbb{N}$, it follows that

$$\inf_{M \leqslant r \leqslant 2^{-1}R_m} \frac{1}{4r^2} \int_{z_m+[0,R_m]^2} \int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(y)\mathbb{1}_{[-r,r]^2}(x-y)\,\mathrm{d}y\,\mathrm{d}x > \delta R_m^2.$$

Since $(\lambda_m)_{m\in\mathbb{N}}$ and $(R_m)_{m\in\mathbb{N}}$ are sequences converging to infinity, there exist some $i, m \in \mathbb{N}$ such that $\lambda_i > M$ and such that $2^{-1}R_m > 2\lambda_{i+J-1}$. Clearly, all the hypotheses of Theorem 4.1 are now satisfied and so, we can find some $j \in \{0, 1, \dots, J-1\}$ such that (4.2) holds for $\lambda = \lambda_{i+j}$. This however immediately implies the existence of some $x \in A$ and $y \in B$ such that

$$|x-y| = \lambda_{i+j},$$

contradicting the very definition of scale λ_{i+j}. □

Proof of Theorem 4.1. For logical rigor's sake, we immediately choose

$$\varepsilon = \left(\frac{\delta C_s}{3C_u}\right)^2 \qquad \text{and then} \qquad J = \left\lceil \frac{3\delta C_e \log(\varepsilon^{-1})}{C_s} \right\rceil, \tag{4.3}$$

where $C_s = \inf_{z\in B(0,3)} \mathrm{g}(z)$, C_u is the implicit constant from Lemma 2.2 (b) and C_e is the implicit constant from Lemma 2.1. Let $\lambda_1, \lambda_2, \ldots, \lambda_J > 1$ be such that $2\lambda_k \leqslant \lambda_{k+1}$ for all $k = 1, 2, \ldots, J-1$ and also let $R, M > 1$ be such that $\lambda_1, \lambda_2, \ldots, \lambda_J \in [M, 2^{-1}R]$ and $R > 2\lambda_J$. Consider $z \in \mathbb{R}^2$ and $A, B \subseteq z + [0, R]^2$ which satisfy (4.1). For clarity's sake, we allow ourself the flexibility to use the symbol λ to stand for any one of the scales $\lambda_1, \lambda_2, \ldots, \lambda_J$ when the estimate in question does not hinge on precisely which one of them we are working with. To simplify notation, we consider the counting form

$$\mathcal{N}_\lambda^0 := \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(x+v)\,\mathrm{d}\sigma_\lambda(v)\,\mathrm{d}x.$$

We briefly mention that Bourgain's original proof of Szekély's theorem in [3] proceeds by taking the Fourier transform and splitting the integral analogous to the one above into "low", "high" and "mid" frequency parts. The interested reader may also wish to consult [33]. However, we rather choose to follow a variant of the scheme introduced by Cook, Magyar and Pramanik in [4], which introduces an additional parameter, which we call the "smoothness scale". More precisely, for $\epsilon > 0$, we consider the smoothed-out version of the counting form:

$$\mathcal{N}_\lambda^\epsilon := \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(x+v)(\sigma_\lambda * \mathrm{g}_{\epsilon\lambda})(v)\,\mathrm{d}v\,\mathrm{d}x.$$

We split the counting form into three parts:

$$\mathcal{N}_\lambda^0 = I_s + I_e + I_u,$$

where

$$I_s = \mathcal{N}^1, \quad I_e = \mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^1, \quad I_u = \mathcal{N}_\lambda^0 - \mathcal{N}_\lambda^\varepsilon$$

and call them by their established literature names as, "the structured part", "the error part" and the "uniform part" of $\mathcal{N}_\lambda^0$ respectively. We begin by bounding the structured part from below. We start by recalling an elementary convolution inequality

$$\sigma_\lambda * \mathrm{g}_\lambda \gtrsim \lambda^{-2}\mathbb{1}_{[-\lambda,\lambda]^2}, \tag{4.4}$$

where the implied constant has the value of $C_s = \inf_{z\in B(0,3)} \mathrm{g}(z)$. Using this bound, we immediately obtain that

$$\mathcal{N}_\lambda^1 \gtrsim R^2(R\lambda)^{-2}\int_{z+[0,R]^2}\int_{z+[0,R]^2} \mathbb{1}_A(x)\mathbb{1}_B(x+y)\mathbb{1}_{[-\lambda,\lambda]^2}(y)\,\mathrm{d}y\,\mathrm{d}x$$

and since $\lambda \in [M, 2^{-1}R]$, from (4.1), we can conclude

$$\mathcal{N}_\lambda^1 \geqslant C_s\delta R^2. \tag{4.5}$$

Our next goal is to bound $|I_u|$ from above. A standard dominated convergence argument gives that

$$\lim_{\theta\to 0}\mathcal{N}_\lambda^\theta = \mathcal{N}_\lambda^0,$$

so, we actually focus on bounding

$$|\mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^\theta|,$$

for $0<\theta<\varepsilon$. Using the re-parametrized heat equation (2.6) and the fundamental theorem of calculus, we obtain

$$|\mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^\theta| \lesssim \int_\theta^\varepsilon \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbb{1}_A(x)\mathbb{1}_B(x+y)(\sigma_\lambda * \Bbbk_{t\lambda})(y)\,\mathrm{d}y\,\mathrm{d}x\,\frac{\mathrm{d}t}{t}$$

By exploiting the central symmetry of the circle, we can actually recognize that in the integral above, we have the triple convolution $\mathbb{1}_B * \sigma_\lambda * \Bbbk_{t\lambda}$ and so, by passing to the Fourier side, using the standard formulae, the fact that A and B are subsets of the cube of side-length R and Lemma 2.2 (b), we get

$$\begin{aligned}|\mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^\theta| &\lesssim \int_\theta^\varepsilon \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\xi)||\widehat{\mathbb{1}_B}(\xi)||\widehat{\sigma_\lambda}(\xi)||\widehat{\Bbbk_{t\lambda}}(\xi)|\,\mathrm{d}\xi\,\frac{\mathrm{d}t}{t} \\ &\lesssim R^2\int_\theta^\varepsilon t^{\frac{-1}{2}}\,\mathrm{d}t = R^2(\varepsilon^{\frac12}-\theta^{\frac12}).\end{aligned}$$

Taking the limit as θ goes to 0, we conclude that

$$|\mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^0| \leqslant C_u R^2 \varepsilon^{\frac12} \tag{4.6}$$

Similarly to what we have done for the uniform part, we conclude

$$\begin{aligned}\sum_{j=1}^J \left|\mathcal{N}_{\lambda_j}^1 - \mathcal{N}_{\lambda_j}^\varepsilon\right| &\lesssim \sum_{j=1}^J \int_\varepsilon^1 \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\xi)||\widehat{\mathbb{1}_B}(\xi)||\widehat{\sigma_{\lambda_j}}(\xi)||\widehat{\Bbbk_{t\lambda_j}}(\xi)|\,\mathrm{d}\xi\,\frac{\mathrm{d}t}{t} \\ &\lesssim \int_{\mathbb{R}^2}\int_\varepsilon^1 |\widehat{\mathbb{1}_A}(\xi)||\widehat{\mathbb{1}_B}(\xi)|\sum_{j=1}^J |\widehat{\Bbbk}(\lambda_j(t\xi))|\frac{\mathrm{d}t}{t}\mathrm{d}\xi \lesssim \log(\varepsilon^{-1})R^2,\end{aligned}$$

where the last inequality follows from Lemma 2.1 applied with $\eta_j = \lambda_j t\xi$, followed by integration in t and then an application of the Cauchy-Schwarz inequality and Plancherel's theorem. A standard pigeonholing argument now allows us to conclude that there exists a scale $\lambda = \lambda_i \in \{\lambda_1, \lambda_2, \ldots, \lambda_J\}$ such that

$$\left|\mathcal{N}_{\lambda_i}^\varepsilon(\mathbb{1}_B) - \mathcal{N}_{\lambda_i}^1(\mathbb{1}_B)\right| \leqslant C_e J^{-1}\log(\varepsilon^{-1})R^2. \tag{4.7}$$

Combining the estimates (4.5), (4.7), (4.6) together with the choice of parameters (4.3) and $\lambda = \lambda_i$, as well as our decomposition of $\mathcal{N}_\lambda^0$ into the structured, error and uniform parts, we conclude

$$I \geqslant I_s - |I_e| - |I_u| \geqslant \delta R^2 - \frac{\delta R^2}{3} - \frac{\delta R^2}{3} = \frac{\delta R^2}{3}. \qquad \square$$

5 Proof of Theorem 1.6

We begin this section by proving the aforementioned reduction of Theorem 1.6 to 1.7. We then prove Theorem 1.7 employing techniques analogous to the ones used in the proof of Theorem 1.2; we first reduce Theorem 1.7 to a "quantitative" version, namely Theorem 5.1 below, which we prove by decomposing a suitable counting form into its structured, error and uniform parts, which we bound separately and devote a subsection to.

Theorem 1.7 implies Theorem 1.6. Assume that $A, B \subseteq \mathbb{R}^2$ are measurable sets for which $\overline{\delta}_{VC}(A,B) > 0$. Applying Theorem 1.7, we find a critical scale $t_0 > 0$ such that for an arbitrary $t \geqslant t_0$, we can find $x \in \mathbb{R}^2$ and $v_1, v_2, v_3, s_1, s_2, s_3 \in t\Gamma$ which satisfy conditions (1.1) and (1.2). We claim that the set

$$C := \{x + v_1, x + v_2, x + v_3\}$$

shatters the family

$$\mathcal{T}_t(A,B) = \{(b + t\Gamma) \cap A : b \in B\}.$$

Indeed, first note that $C \subseteq A$, which is certainly a necessary condition for our claim. Further, we check that for every $S \in \mathcal{P}(C)$, we can find $b \in B$ such that both the inclusion

$$S \subseteq C \cap ((b + t\Gamma) \cap A) \tag{5.1}$$

and its reverse hold. This will imply that $VC(\mathcal{T}_t(A,B)) \geqslant 3$. We check this case by case, allowing ourselves the flexibility to redefine b as the context requires it. If S is a 3-element set, i.e., $S = C$, note that for $j \in \{1,2,3\}$, we have $(x + v_j) - x = v_j \in t\Gamma$ and so, by taking $b = x$, it follows that (5.1) holds, while in this case the reverse inclusion holds trivially. Next we consider the 2-element case, i.e., we assume that $S = \{x + v_i, x + v_j\}$ for some mutually distinct $i, j \in \{1,2,3\}$. Since

$$(x + v_i) - (x + v_i + v_j) = -v_j \in t\Gamma, \qquad (x + v_j) - (x + v_i + v_j) = -v_i \in t\Gamma,$$

because Γ is centrally symmetric, it follows that (5.1) holds with $b = x + v_i + v_j$. To check the reverse inclusion, note that if it were to fail, then the curve segment $(b + t\Gamma) \cap A$ would also have to contain the point $x + v_k$, for $k \in \{1,2,3\} \setminus \{i,j\}$ and so we have that all 3 points of C lie on the curve $x + t\Gamma$. However, since by Theorem 1.7, we know that x and b are mutually distinct, we have arrived at a contradiction with Lemma 2.6 (a). Lastly, we consider the 1-element case, i.e., we assume that $S = \{x + v_j\}$ for some $j \in \{1,2,3\}$. Because $(x + v_j) - (x + v_j + s_j) = -s_j \in t\Gamma$ and because $t\Gamma$ is centrally symmetric, it follows that (5.1) holds with $b = x + v_j + s_j$. To see that the reverse inclusion also holds, we again suppose otherwise, i.e., we suppose that for some $i \in \{1,2,3\} \setminus \{j\}$, we have that $x_i + v_i \in (b + t\Gamma) \cap A$. However, by everything we have shown so far, we can conclude that

$$x + v_i, x + v_j \in x + t\Gamma, (x + v_i + v_j) + t\Gamma, (x + v_j + s_j) + t\Gamma.$$

Since by Theorem 1.7, the points $x, x + v_i + v_j, x + v_j + s_j$ are mutually distinct, we obtain a contradiction with Lemma 2.6 (b). It remains to show that $VC(\mathcal{T}_t(A,B)) = 3$, which we do by showing that no 4-element set can shatter $\mathcal{T}_t(A,B)$. Assume otherwise and let

$$C = \{c_1, c_2, c_3, c_4\} \subseteq A$$

be a 4-element set which shatters $\mathcal{T}_t(A,B)$. However, by definition, this means that there exist mutually distinct $b_1, b_2, b_3 \in B$ such that

$$((b_1+t\Gamma)\cap A) = \{c_1,c_2,c_3,c_4\},\ ((b_2+t\Gamma)\cap A) = \{c_1,c_2,c_3\},\ ((b_3+t\Gamma)\cap A) = \{c_1,c_2,c_4\},$$

but this contradicts Lemma 2.6 (b), because c_1 and c_2 both lie on 3 distinct translates of the curve $t\Gamma$. $\square$

Theorem 5.1. *Let $\delta \in (0,1)$ be arbitrary. Then, there exists $J = J(\delta)$ such that for all $\lambda_1, \lambda_2, \dots, \lambda_J > 1$ which satisfy $2\lambda_k \leqslant \lambda_{k+1}$ for all $k = 1,2,\dots,J-1$, for all $R > 0$ such that $R \geqslant 2\lambda_J$ and for all $M \geqslant 1$ so that $\lambda_1, \lambda_2, \dots, \lambda_J \in [M, 2^{-6}R]$, the following holds: For an arbitrary $z \in \mathbb{R}^2$ and for all $A, B \subseteq z + [0,R]^2$ which satisfy*

$$\inf_{M\leqslant r\leqslant 2^{-6}R} \frac{1}{(2r)^6}\int_{z+[0,R]^2}\int_{([-r,r]^2)^6} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\,\mathrm{d}v_1\,\mathrm{d}v_2\,\mathrm{d}v_3\,\mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3\,\mathrm{d}x \gtrsim \delta R^2 \tag{5.2}$$

there exists $j \in \{1,2,\dots,J\}$ such that for $\lambda = \lambda_j$ it follows that

$$\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\,\mathrm{d}\mu_\lambda(v_1)\mathrm{d}\mu_\lambda(v_2)\mathrm{d}\mu_\lambda(v_3)\mathrm{d}\mu_\lambda(s_1)\mathrm{d}\mu_\lambda(s_2)\mathrm{d}\mu_\lambda(s_3)\mathrm{d}x \gtrsim_\delta R^2, \tag{5.3}$$

where μ is as before, the probabilistically normalized (uniform) arclength measure supported on a smooth, closed, centrally symmetric planar curve Γ which is the boundary of a convex and compact region with non-empty interior.

Theorem 5.1 implies Theorem 1.7. Assume that Theorem 5.1 is true and that Theorem 1.7 is not. Then, there exist measurable sets $A, B \subseteq \mathbb{R}^2$ such that $\overline{\delta}_{VC}(A,B) > 0$ and there exists a sequence of scales $(\lambda_m)_{m\in\mathbb{N}}$ such that for all $m \in \mathbb{N}$ and for all $x \in B$ and all $v_1, v_2, v_3, s_1, s_2, s_3 \in \lambda_m\Gamma$ either (1.1) or (1.2) do not hold. By passing to a subsequence, we can assume that $\lambda_1 > 1$ and $2\lambda_m \leqslant \lambda_{m+1}$ for all $m \in \mathbb{N}$. We now apply Theorem 5.1 with $\delta = \frac{\overline{\delta}_{VC}(A,B)}{4}$ and let J be as in the conclusion of Theorem 5.1. Unpacking Definition 1.5, we first find $M \geqslant 1$ large enough as well as a sequence of positive real numbers $(R_m)_{m\in\mathbb{N}} \in$ converging to infinity and a sequence $(z_m)_{m\in\mathbb{N}}$ of points in $\mathbb{R}^2$ such that for all $m \in \mathbb{N}$

$$\inf_{M\leqslant r\leqslant 2^{-6}R_m} \frac{1}{(2r)^6}\int_{z_m+[0,R_m]^2}\int_{([-r,r]^2)^6} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\,\mathrm{d}v_1\,\mathrm{d}v_2\,\mathrm{d}v_3\,\mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3\,\mathrm{d}x > \delta R_m^2$$

Since $(\lambda_k)_{k\in\mathbb{N}}$ is a lacunary sequence, it follows that there exists some $i \in \mathbb{N}$ such that $\lambda_i > M$ and since R_m converges to infinity, there exists some $m \in \mathbb{N}$ such that $\lambda_{i+J-1} < 2^{-6}R_m$ However, Theorem 5.1. now implies that there exists some $j \in \{0,1,\dots,J-1\}$ such that (5.3) holds for $\lambda = \lambda_{i+j}$. A standard measure-theoretic argument now provides us with the existence of $x \in B$ and $v_1, v_2, v_3, s_1, s_2, s_3 \in \lambda_j\Gamma$ which satisfy (1.1) and (1.2). Thus, we have arrived at a contradiction. $\square$

Finally, we begin with the proof of Theorem 5.1. We assume that for $J \in \mathbb{N}$, which we specify bellow, we are given scales $\lambda_1, \lambda_2, \dots, \lambda_J > 1$ which satisfy $2\lambda_k \leqslant \lambda_{k+1}$ for all

$k=1,2,\dots,J-1$, and also $R>0$ such that $R\geqslant 2\lambda_J$, as well as $M\geqslant 1$ so that $\lambda_1,\lambda_2,\dots,\lambda_J\in[M,2^{-6}R]$. As before, we introduce the following notation for a given scale λ

$$\mathcal{N}_\lambda^0 := \int_{(\mathbb{R}^2)^7} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\,\mathrm{d}\mu_\lambda(v_1)\,\mathrm{d}\mu_\lambda(v_2)\,\mathrm{d}\mu_\lambda(v_3)\,\mathrm{d}\mu_\lambda(s_1)\,\mathrm{d}\mu_\lambda(s_2)\,\mathrm{d}\mu_\lambda(s_3)\,\mathrm{d}x$$

and, for $\varepsilon>0$, we also consider

$$\mathcal{N}_\lambda^\varepsilon := \int_{(\mathbb{R}^2)^7} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\prod_{i=1}^{3}(\mu_\lambda * \mathrm{g}_{\varepsilon\lambda})(v_i)\prod_{i=1}^{3}(\mu_\lambda * \mathrm{g}_{\varepsilon\lambda})(s_i)$$
$$\mathrm{d}\mu(v_1)\,\mathrm{d}\mu(v_2)\,\mathrm{d}\mu(v_3)\,\mathrm{d}\mu(s_1)\,\mathrm{d}\mu(s_2)\,\mathrm{d}\mu(s_3)\,\mathrm{d}x.$$

We split the counting form $\mathcal{N}_\lambda^0$ into three parts:

$$\mathcal{N}_\lambda^0 = I_s + I_e + I_u,$$

where

$$I_s = \mathcal{N}^1, \quad I_e = \mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^1, \quad I_u = \mathcal{N}_\lambda^0 - \mathcal{N}_\lambda^\varepsilon$$

and call them once more, the structured part, the error part and the uniform part, respectively. Throughout the proof, we will again use the letter λ to mean any one of the scales $\lambda_1,\lambda_2,...,\lambda_J$ whenever the claim in question does not depend on which particular scale we are working with.

5.1 The Structured Part bound

By once more employing (4.4) we conclude that

$$\mathcal{N}_\lambda^1 \gtrsim R^2(R\lambda)^{-2}\int_{(\mathbb{R}^2)^7} \mathcal{F}_{A,B}(x;v_1,v_2,v_3;s_1,s_2,s_3)\prod_{i=1}^{3}\mathbb{1}_{[-\lambda,\lambda]^2}(v_i)\prod_{i=1}^{3}\mathbb{1}_{[-\lambda,\lambda]^2}(s_i)$$
$$\mathrm{d}\mu(v_1)\,\mathrm{d}\mu(v_2)\,\mathrm{d}\mu(v_3)\,\mathrm{d}\mu(s_1)\,\mathrm{d}\mu(s_2)\,\mathrm{d}\mu(s_3)\,\mathrm{d}x.$$

and using assumption (5.2), we deduce

$$\mathcal{N}_\lambda^1 \geqslant C_s\delta R^2, \tag{5.4}$$

for some constant $C_s>0$.

5.2 The Error Part bound

We will show

$$\sum_{j=1}^{J}\left|\mathcal{N}_{\lambda_j}^1 - \mathcal{N}_{\lambda_j}^\varepsilon\right| \leqslant C_e\varepsilon^{-18}\log(\varepsilon^{-1})R^2, \tag{5.5}$$

for some constant $C_e>0$. By using the fundamental theorem of calculus and the reparametrized heat equation (2.6), we can write

$$\mathcal{N}_\lambda^1 - \mathcal{N}_\lambda^\varepsilon = \sum_{i=1}^{6}\mathcal{L}_\lambda^{i,\varepsilon,1},$$

where for a fixed $k \in \{1, 2, \ldots, 6\}$,

$$\mathcal{L}_{\lambda}^{k,\varepsilon,1} := \int_{\varepsilon}^{1} \int_{(\mathbb{R}^2)^7} \mathcal{F}_6(x; v_1, v_2, v_3, v_4, v_5, v_6)\, (\mu * \mathbb{k}_{t\lambda})(v_k) \prod_{\substack{i=1 \\ i \neq k}}^{n} (\mu_\lambda * \mathrm{g}_{t\lambda})(v_i)\, \mathrm{d}v_1 \ldots \mathrm{d}v_6\, \mathrm{d}x \frac{\mathrm{d}t}{t}$$

and where $\mathcal{F}_6(x; v_1, v_2, v_3, v_4, v_5, v_6)$ is as in Definition 3.1 for the particular choice of functions described in (3.2). Our first goal is to bound away all appearances of measure μ in a given $\mathcal{L}_{\lambda}^{k,\varepsilon,1}$. We introduce two additional parameters, which should be understood as auxiliary convolutions scales. For $e' = 10^{-2}e^{-1}$, which is just a suitably chosen small constant, and any $\lambda, t > 0$, we have

$$\int_{e't\lambda}^{ee't\lambda} \frac{\mathrm{d}s}{s} = 1,$$

so we can introduce an additional integral into the form $\mathcal{L}_{\lambda}^{k,\theta,\varepsilon}$, giving us slightly more flexibility. We also denote

$$r = \sqrt{(t\lambda)^2 - 2s^2}.$$

and observe that $t\lambda \sim r \sim s$. Using (2.2) and (2.3), we get that

$$\mu_t * \mathbb{k}_{t\lambda} = \frac{(t\lambda)^2}{s^2} \sum_{l=1}^{2} \mathbb{h}_s^{(l)} * \mathbb{h}_s^{(l)} * \mathrm{g}_r * \mu_t,$$

and using this decomposition, followed by application of (2.4) and (2.5), we bound $|\mathcal{L}_{\lambda}^{k,\varepsilon,1}|$ by

$$\varepsilon^{-3n} \sum_{l=1}^{2} \int_{[1,\infty)^6} \int_{\varepsilon}^{1} \int_{e't\lambda}^{ee't\lambda} \int_{(\mathbb{R}^2)^7} \int_{(\mathbb{R}^2)^2} \mathcal{F}_6(x; v_1, \ldots, v_6) \mathbb{h}_s^{(l)}(v_k - y_1) \mathbb{h}_s^{(l)}(y_1 - y_2)$$
$$\mathrm{g}_{s\gamma_k}(y_2) \prod_{\substack{i=1 \\ i \neq k}}^{n} (\mathrm{g}_{s\gamma_i})(v_i)\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}v_1 \ldots \mathrm{d}v_6\, \mathrm{d}x$$
$$\frac{\mathrm{d}s}{s} \frac{\mathrm{d}t}{t} \frac{\mathrm{d}\gamma_1}{\gamma_1^2} \cdots \frac{\mathrm{d}\gamma_6}{\gamma_6^2}.$$

We can now collapse the convolution back, using (2.2) and (2.3) once more, in order to get that the above equals

$$\int_{\varepsilon}^{1} \sum_{j=1}^{J} \int_{e't\lambda_j}^{ee't\lambda_j} \int_{[1,\infty)^6} \int_{(\mathbb{R}^2)^7} \mathcal{F}_6(x; v_1, \ldots, v_6) \mathbb{k}_{s\sqrt{2+\gamma_k^2}}(v_k)$$
$$\prod_{\substack{i=1 \\ i \neq k}}^{n} (\mathrm{g}_{s\gamma_i})(v_i)\, \mathrm{d}v_1 \ldots \mathrm{d}v_6\, \mathrm{d}x \frac{\mathrm{d}\gamma_1}{\gamma_1^2} \cdots \frac{\mathrm{d}\gamma_k}{\gamma_k^2(2+\gamma_k^2)} \cdots \frac{\mathrm{d}\gamma_6}{\gamma_6^2} \frac{\mathrm{d}s}{s} \frac{\mathrm{d}t}{t}.$$

We also have to sum over the scales λ_j. At this point, we use the lacunary nature of the scales in question. For a fixed $t \in (\varepsilon, 1)$, each interval $[e't\lambda_j, ee't\lambda_j]$ can overlap at most two intervals of the form $[e't\lambda_k, ee't\lambda_k]$ for $k = 1, 2, \ldots, J$. Therefore, a fixed $s \in [0, \infty)$ can lie only in two such intervals. Thus, after a routine application of Fubini's theorem, we can integrate in the t-variable and conclude that $\sum_{j=1}^{J}|\mathcal{L}_{\lambda_j}^{i,\varepsilon,1}|$ is bounded by

$$2\varepsilon^{-18}\log(\varepsilon^{-1})\int_{[1,\infty)^6}\int_0^\infty\int_{(\mathbb{R}^2)^7}\mathcal{F}_6(x;v_1,\dots,v_6)\Bbbk_{s\sqrt{2+\gamma_k^2}}(v_k)$$
$$\prod_{\substack{i=1\\ i\neq k}}^{n}(\mathrm{g}_{s\gamma_i})(v_i)\,\mathrm{d}v_1\dots\mathrm{d}v_6\,\mathrm{d}x\frac{\mathrm{d}s}{s}\frac{\mathrm{d}\gamma_1}{\gamma_1^2}\cdots\frac{\mathrm{d}\gamma_k}{\gamma_k^2(2+\gamma_k^2)}\cdots\frac{\mathrm{d}\gamma_6}{\gamma_6^2}.$$

However, we now recognize that the above is actually a superposition of, so-called, singular Brascamp–Lieb forms, for which, in the following subsection, we obtain bounds which are uniform in the kernel scales. That is to say, Theorem 5.2 below, allows us to conclude

$$\begin{aligned}\sum_{j=1}^{J}|\mathcal{L}_{\lambda_j}^{i,\varepsilon,1}| &\lesssim \varepsilon^{-18}\log(\varepsilon^{-1})\int_{[1,\infty)^6}\int_{[0,\infty\rangle}\prod_{(r_1,\dots,r_6)\in\{0,1\}^6}\|f_{(r_1,\dots,r_6)}\|_{\mathrm{L}^{2^n}}\frac{\mathrm{d}\gamma_1}{\gamma_1^2}\cdots\frac{\mathrm{d}\gamma_k}{\gamma_k^2(2+\gamma_k^2)}\cdots\frac{\mathrm{d}\gamma_6}{\gamma_6^2}\\ &\lesssim \varepsilon^{-18}\log(\varepsilon^{-1})R^2,\end{aligned}$$

which implies the desired bound.

5.3 A particular case of singular Brascamp–Lieb inequalities

We start by fixing notation which shall be used throughout this subsection. For a fixed positive integer $n\in\mathbb{N}$, and $j\in\{1,2,\dots,n\}$ and fixed scales $\alpha_1,\alpha_2,\dots,\alpha_n>0$ we consider the integral kernel $K_j^{\alpha_1,\alpha_2,\dots,\alpha_n}:(\mathbb{R}^2)^n\to[0,+\infty)$

$$K_j^{\alpha_1,\alpha_2,\dots,\alpha_n}(v_1,v_2,\dots,v_n)=\Bbbk_{\alpha_j}(v_j)\prod_{\substack{i=1\\ i\neq j}}^{n}\mathrm{g}_{\alpha_i}(v_i).$$

We also consider a finite family of functions $(f_r)_{r\in\{0,1\}^n}$ in $\mathrm{L}^{2^n}(\mathbb{R}^2)$. Our main object of interest is the following multilinear integral form

$$\mathcal{J}_j^{\alpha_1,\alpha_2,\dots,\alpha_n}((f_r)_{r\in\{0,1\}^n}):=-\int_0^\infty\int_{(\mathbb{R}^2)^{n+1}}\prod_{r\in\{0,1\}^n}f_r(x+r\cdot v)$$
$$K_j^{s\alpha_1,s\alpha_2,\dots,s\alpha_n}(v_1,v_2,\dots,v_n)\,\mathrm{d}v_1\dots\mathrm{d}v_n\,\mathrm{d}x\frac{\mathrm{d}s}{s}.$$

The main result of this subsection is

Theorem 5.2. *For any functions $(f_r)_{r\in\{0,1\}^n}$ in $\mathrm{L}^{2^n}(\mathbb{R}^2)$, any $\alpha_1,\alpha_2,\dots,\alpha_n>0$ and any $j\in\{1,2,\dots,n\}$, it follows that*

$$|\mathcal{J}_j^{\alpha_1,\alpha_2,\dots,\alpha_n}((f_r)_{r\in\{0,1\}^n})|\lesssim_n\prod_{r\in\{0,1\}^n}\|f_r\|_{\mathrm{L}^{2^n}(\mathbb{R}^2)}$$

The key idea of the proof is to gradually apply the Cauchy–Schwarz inequality, in order to "symmetrize" the form under consideration. Note that a priori we do not even know that our forms are well defined. However, we will prove the desired bound in the Proposition 5.3 by assuming that all functions under consideration belong to the Schwartz class. Theorem 5.2 then easily follows by applying classic functional-analytic tools and standard approximation arguments, which both guarantee that our form is well defined and bounded.

Proposition 5.3. *For any $n \in \mathbb{N}$, any functions $(f_r)_{r\in\{0,1\}^n}$ belonging to the Schwartz class $\mathcal{S}(\mathbb{R}^2)$ and for any $0 \leqslant k \leqslant n-1$ it follows that for mutually distinct $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$ and for all $a_1, a_2, \ldots, a_k \in \{0, 1\}$ and all $j \in \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\}$, the following bound holds*

$$|\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)})_{r\in\{0,1\}^n})| \lesssim_n \prod_{r\in\{0,1\}^n} \|f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)}\|_{\mathrm{L}^{2^n}(\mathbb{R}^n)}$$

The proof of Proposition 5.3 will be done by the backward induction on $0 \leqslant k \leqslant n$. We note that the base case of the induction $k = n$, follows by combining Lemmas 5.4 and 5.5 below, in a way similar to how the estimate (5.7) below is established. Notice that Theorem 5.2 is precisely recovered as the case $k = 0$, where we agree that the cardinality of the empty set equals 0, modulo the aforementioned functional analysis density argument. The assumption that in a given inductive step, the coordinate which carries the Laplacian has not appeared among the coordinates which have been fixed by the freeze function is precisely what will meaningfully allow us to reduce the number of functions we work with at a given step. However, we shall also need to know what happens in the complementary case, i.e., if the "Laplacian coordinate" is actually fixed by the freeze function. This is done in Lemma 5.4 below whose proof is following the approaches from [24] and [6] and a variant of which has been used in a similar context in [8], [9], [25] and [26]. In order to be able to connect the two cases, we shall need Lemma 5.5 below, which we state without proof because it is formally identical to the proof of the analogous statement in [26]. We also advise the interested reader to look at Lemma 3 in [6] and Lemma 4 in [10].

Lemma 5.4. *For fixed scales $\alpha_1, \alpha_2, \ldots, \alpha_n > 0$, for functions $(f_r)_{r\in\{0,1\}^n}$ in $\mathrm{L}^{2^n}(\mathbb{R}^2)$, for fixed $k \leqslant n$ and any coordinates $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$, we claim that for each $j' \in \{i_1, i_2, \ldots, i_k\}$, we have that*

$$\mathcal{J}_{j'}^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)})_{r\in\{0,1\}^n}) \geqslant 0.$$

Proof. For a fixed $m \in \mathbb{N}$ and $w_1, w_2, \ldots, w_m \in \mathbb{R}^2$, we denote

$$\mathcal{F}^m_{(w_1,w_2,\ldots,w_l)}(y) := \prod_{r\in\{0,1\}^m} f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)}(y + w_1r_1 + w_2r_2 + \cdots + w_mr_m).$$

Using (2.2) combined with the fact that a convolution of two odd functions is even, after a simple calculation, we obtain that for some $C > 0$, we have

$$\mathcal{J} = C\sum_{l=1}^{2}\int_0^\infty \int_{(\mathbb{R}^2)^n} \Big\|\mathcal{F}^{n-1}_{v_1,\ldots,v_{j-1},v_{j+1},\ldots,v_n} * \mathrm{h}^{(l)}_{\frac{s\alpha_j}{\sqrt{2}}}\Big\|^2_{\mathrm{L}^2(\mathbb{R}^2)} \prod_{\substack{i=1\\ i\neq j}}^{n} \mathrm{g}_{s\alpha_i}(v_i) \prod_{\substack{i=1\\ i\neq j}} \mathrm{d}v_i\,\mathrm{d}x\frac{\mathrm{d}s}{s}$$

which implies the desired claim. $\square$

Lemma 5.5. *For a finite family of functions $(f_r)_{r\in\{0,1\}^n}$ in $\mathcal{S}(\mathbb{R}^2)$, and $\alpha_1,\alpha_2,\ldots,\alpha_n>0$, it holds that*

$$\sum_{i=1}^{n}\mathcal{J}_i = 2\pi\int_{\mathbb{R}^2}\prod_{r\in\{0,1\}^n} f_r(x)\,\mathrm{d}x,$$

where we have briefly denoted $\mathcal{J}_i=\mathcal{J}_i^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_r)_{r\in\{0,1\}^n})$.

Proof of Proposition 5.3. As has already been mentioned the proof proceeds by induction on k. We show how to make an inductive step, i.e., how we reduce a number of different functions appearing in our form by a factor of 2. Let $k<n$ be arbitrary and also let $j\in\{1,2,\ldots,n\}$, $i_1,i_2,\ldots,i_k\in\{1,2,\ldots,n\}\setminus\{j\}$ and $a_1,a_2,\ldots,a_k\in\{0,1\}$ be arbitrary. By using (2.2), introducing the change of variables $y=x+v_j$, $\mathrm{d}y=\mathrm{d}v_j$ and expanding the convolution, we conclude that $\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)})_{r\in\{0,1\}^n})$ is, up to a constant, equal to

$$\begin{aligned}\sum_{l=1}^{d}\int_0^\infty\int_{(\mathbb{R}^2)^n}&\Big(\int_{\mathbb{R}^2}\prod_{\substack{r\in\{0,1\}\\ r_j=0}} f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)}(x+r\cdot v)\,\mathbb{h}^{(l)}_{\frac{s\alpha_j}{\sqrt{2}}}(u-x)\,\mathrm{d}x\Big)\\ &\Big(\int_{\mathbb{R}^2}\prod_{\substack{r\in\{0,1\}\\ r_j=1}} f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)}(y+C_{(j,0)}(r)\cdot v)\,\mathbb{h}^{(l)}_{\frac{s\alpha_j}{\sqrt{2}}}(y-u)\,\mathrm{d}y\Big)\\ &\prod_{\substack{i=1\\ i\neq j}}^{n}\mathrm{g}_{s\alpha_i}(v_i)\,\mathrm{d}u\prod_{\substack{i=1\\ i\neq j}}^{n}\mathrm{d}v_i\frac{\mathrm{d}s}{s}\end{aligned}$$

By applying the Cauchy–Schwarz inequality, expanding the square in each of the two factors, making changes of variables analogous to the one already made, exploiting the fact that $\mathbb{h}^{(l)}$ is odd and finally once more collapsing the two convolutions in each of the two factors by again invoking 2.2, we conclude that we can bound

$$|\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)})_{r\in\{0,1\}^n})|^2$$

by the product

$$\begin{aligned}&|\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k),(j,0)}(r)})_{r\in\{0,1\}^n})|\\ \cdot&|\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k),(j,1)}(r)})_{r\in\{0,1\}^n})|\end{aligned}\tag{5.6}$$

Note that we cannot immediately apply the inductive hypothesis on

$$\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k),(j,a_j)}(r)})_{r\in\{0,1\}^n}),$$

for $a_j\in\{0,1\}$, because the coordinate j appears as one of the $k+1$ fixed coordinates in our freeze function and we explicitly excluded such cases in our inductive statement. Still, we will circumvent this problem by using Lemma 5.5. Let (i_{k+1},a_{k+1}) stand for either one of the two pairs $(j,0)$ or $(j,1)$ above. Denote the following

$$\mathcal{S}_1 := \sum_{\substack{l=1 \\ l\in\{i_1,i_2,\ldots,i_{k+1}\}}}^{n} \mathcal{J}_l^{\alpha_1,\alpha_2,\ldots,\alpha_n}\big((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_{k+1},a_{k+1})}(r)})_{r\in\{0,1\}^n}\big),$$

$$\mathcal{S}_2 := \sum_{\substack{l=1 \\ l\notin\{i_1,i_2,\ldots,i_{k+1}\}}}^{n} \mathcal{J}_l^{\alpha_1,\alpha_2,\ldots,\alpha_n}\big((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_{k+1},a_{k+1})}(r)})_{r\in\{0,1\}^n}\big).$$

We can bound $|\mathcal{S}_2|$ by using the inductive hypothesis and we can also bound $|\mathcal{S}_1+\mathcal{S}_2|$ by using Lemma 5.5 and Hölder's inequality. Therefore, we obtain a bound for $|\mathcal{S}_1|$

$$\begin{aligned}|\mathcal{S}_1| &\leqslant |\mathcal{S}_1+\mathcal{S}_2| + |\mathcal{S}_2| \\ &\leqslant (2\pi+(n-k)) \prod_{r\in\{0,1\}^n} \|f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_{k+1},a_{k+1})}(r)}\|_{\mathrm{L}^{2^n}(\mathbb{R}^n)}.\end{aligned}$$

In particular, using positivity guaranteed to us by Lemma 5.4, we conclude that

$$\begin{aligned}&|\mathcal{J}_j^{\alpha_1,\alpha_2,\ldots,\alpha_n}\big((f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_{k+1},a_{k+1})}(r)})_{r\in\{0,1\}^n}\big)| \\ &\leqslant |\mathcal{S}_1| \lesssim_n \prod_{r\in\{0,1\}^n} \|f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_{k+1},a_{k+1})}(r)}\|_{\mathrm{L}^{2^n}(\mathbb{R}^n)},\end{aligned} \tag{5.7}$$

which, after a brief calculation, gives us that (5.6) is bounded by

$$\prod_{r\in\{0,1\}^n} \|f_{C_{(i_1,a_1),(i_2,a_2),\ldots,(i_k,a_k)}(r)}\|^2_{\mathrm{L}^{2^n}(\mathbb{R}^n)}.$$

By taking square roots, we obtain the desired claim.

□

5.4 The Uniform Part bound

For $0<\theta<\varepsilon<1$, we focus on bounding

$$|\mathcal{N}_\lambda^\theta - \mathcal{N}_\lambda^\varepsilon|.$$

Similar to before, using the fundamental theorem of calculus and the re-parametrized heat equation, we can write

$$\mathcal{N}_\lambda^\varepsilon - \mathcal{N}_\lambda^\theta = \sum_{i=1}^{6} \mathcal{L}_\lambda^{i,\theta,\varepsilon},$$

where for a fixed $k\in\{1,2,\ldots,6\}$, we have denoted, similarly as before

$$\mathcal{L}_\lambda^{k,\theta,\varepsilon} := \int_\varepsilon^1 \int_{(\mathbb{R}^2)^7} \mathcal{F}_6(x;v_1,v_2,v_3,v_4,v_5,v_6)\,(\mu * \mathrm{k}_{t\lambda})(v_k) \prod_{\substack{i=1\\ i\neq k}}^{n} (\mu_\lambda * \mathrm{g}_{t\lambda})(v_i)\,\mathrm{d}v_1\ldots\mathrm{d}v_6\,\mathrm{d}x\frac{\mathrm{d}t}{t}.$$

For fixed variables $v_1,\dots,v_{k-1},v_{k+1},\dots,v_6$, we denote

$$F_k(u)=\prod_{r\in\{0,1\}^5}\mathbb{1}_{z+[0,R]^2}(x+r_1v_1+\cdots+r_{k-1}v_{k-1}+r_{k+1}v_{k+1}+\cdots+r_6v_6).$$

By noticing that $F_6(x;v_1,\dots,v_6)\leqslant F_k(x+v_6)F_k(x)$ and by making a change of variables $u=x+v_k$, $\mathrm{d}u=\mathrm{d}v_k$, we obtain the bound

$$\mathcal{L}_\lambda^{k,\theta,\varepsilon}\leqslant\int_\varepsilon^1\int_{(\mathbb{R}^2)^5}\prod_{\substack{i=1\\i\neq k}}^n(\mu_\lambda*\mathrm{g}_{t\lambda})(v_i)\int_{\mathbb{R}^2}(F_k*\mu_\lambda*\Bbbk_{t\lambda})(u)F_k(u)\,\mathrm{d}u\,\mathrm{d}v_1\dots\mathrm{d}v_{k-1}\mathrm{d}v_k\dots\mathrm{d}v_6\,\frac{\mathrm{d}t}{t}.$$

By applying Plancherel's formula in the inner integral, followed by Lemma 2.2 (b) and then integrating in all the remaining variables, we conclude that

$$|\mathcal{L}_\lambda^{k,\theta,\varepsilon}|\lesssim R^2\int_\theta^\varepsilon t^{\frac{-1}{2}}\,\mathrm{d}t,$$

By integrating in t, using the dominated convergence theorem to let $\theta\to0$ we obtain the bound

$$|\mathcal{N}_\lambda^\varepsilon-\mathcal{N}_\lambda^0|\leqslant C_u\varepsilon^{\frac{1}{2}}R^2, \tag{5.8}$$

for some constant $C_u>0$.

5.5 Combining the estimates

Finally, we choose

$$\varepsilon=\frac{1}{2}\Big(\frac{\delta C_sR^2}{3C_u}\Big)^2,\qquad J=\left\lceil\frac{3C_e\log(\varepsilon^{-1})}{\delta\varepsilon^{18}C_sR^2}\right\rceil \tag{5.9}$$

From (5.5), by pigeonholing in j, we can find a "good scale" λ_j such that

$$|I_e|\leqslant\frac{C_s}{3}\delta R^2.$$

Combining this with (5.4) and (5.8), we obtain the desired claim

$$\mathcal{N}_{\lambda_j}^0\geqslant I_s-|I_e|-|I_u|\geqslant\frac{C_s}{3}\delta R^2.$$

Acknowledgements

The work was supported by the *Croatian Science Foundation* under the project number HRZZ-IP-2022-10-5116 (FANAP). The author is grateful to Professor Vjekoslav Kovač for his patient guidance, helpful comments and stimulating discussions and also to Shobu Shiraki for plenty of insightful discussions, comments and suggestions. The author further wishes to express gratitude to Professor Alex Iosevich for the suggestion that studying the connection between VC dimension and flexible configurations in Euclidean spaces might lead to interesting results.

References

[1] Ruben Ascoli, Livia Betti, Justin Cheigh, Alex Iosevich, Ryan Jeong, Xuyan Liu, Brian McDonald, Wyatt Milgrim, Steven J. Miller, Francisco Romero Acosta, and Santiago Velazquez Iannuzzelli. VC-dimension and distance chains in $\mathbb{F}_q^d$. *Korean J. Math.*, 32(1):43–57, 2024.

[2] Ruben Ascoli, Livia Betti, Justin Cheigh, Alex Iosevich, Ryan Jeong, Xuyan Liu, Brian McDonald, Wyatt Milgrim, Steven J. Miller, Francisco Romero Acosta, and Santiago Velazquez Innuzzelli. VC-dimension of hyperplanes over finite fields. *Graphs Combin.*, 41(2):Paper No. 47, 13, 2025. `doi:10.1007/s00373-025-02909-6`.

[3] Jean Bourgain. A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$. *Israel J. Math.*, 54(3):307–316, 1986. `doi:10.1007/BF02764959`.

[4] Brian Cook, Ákos Magyar, and Malabika Pramanik. A Roth-type theorem for dense subsets of $\mathbb{R}^d$. *Bull. Lond. Math. Soc.*, 49(4):676–689, 2017. `doi:10.1112/blms.12043`.

[5] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces: Revised and Updated*. Dover Publications, Mineola, NY, revised and updated edition, 2016.

[6] Polona Durcik. An L^4 estimate for a singular entangled quadrilinear form. *Math. Res. Lett.*, 22(5):1317–1332, 2015. `doi:10.4310/MRL.2015.v22.n5.a3`.

[7] Polona Durcik. L^p estimates for a singular entangled quadrilinear form. *Trans. Amer. Math. Soc.*, 369(10):6935–6951, 2017. `doi:10.1090/tran/6850`.

[8] Polona Durcik and Vjekoslav Kovač. Boxes, extended boxes and sets of positive upper density in the Euclidean space. *Math. Proc. Cambridge Philos. Soc.*, 171(3):481–501, 2021. `doi:10.1017/S0305004120000316`.

[9] Polona Durcik and Vjekoslav Kovač. A Szemerédi-type theorem for subsets of the unit cube. *Anal. PDE*, 15(2):507–549, 2022. `doi:10.2140/apde.2022.15.507`.

[10] Polona Durcik, Vjekoslav Kovač, and Luka Rimanić. On side lengths of corners in positive density subsets of the Euclidean space. *Int. Math. Res. Not. IMRN*, 14(22):6844–6869, 2018. `doi:10.1093/imrn/rnx093`.

[11] Polona Durcik, Lenka Slavíková, and Christoph Thiele. Local bounds for singular Brascamp-Lieb forms with cubical structure. *Math. Z.*, 302(4):2375–2405, 2022. `doi:10.1007/s00209-022-03148-8`.

[12] Polona Durcik and Christoph Thiele. Singular Brascamp-Lieb inequalities with cubical structure. *Bull. Lond. Math. Soc.*, 52(2):283–298, 2020. `doi:10.1112/blms.12310`.

[13] Tanja Eisner and Terence Tao. Large values of the Gowers-Host-Kra seminorms. *J. Anal. Math.*, 117:133–186, 2012. `doi:10.1007/s11854-012-0018-2`.

[14] Paul Erdős. Some combinatorial, geometric and set theoretic problems in measure theory. In D. Kölzow and D. Maharam-Stone, editors, *Measure Theory, Oberwolfach 1983: Proceedings of the Conference held at Oberwolfach, June 26–July 2, 1983*, volume 1089 of *Lecture Notes in Mathematics*, pages 321–327. Springer, Berlin, Heidelberg, 1984. `doi:10.1007/BFb0072626`.

[15] Kenneth J. Falconer and John M. Marstrand. Plane sets with positive density at infinity contain all large distances. *Bull. London Math. Soc.*, 18(5):471–474, 1986. `doi:10.1112/blms/18.5.471`.

[16] David Fitzpatrick, Alex Iosevich, Brian McDonald, and Emmett Wyman. The VC-dimension and point configurations in $\mathbb{F}_q^2$. *Discrete Comput. Geom.*, 71(4):1167–1177, 2024. doi:10.1007/s00454-023-00570-5.

[17] Hillel Furstenberg, Yitzchak Katznelson, and Benjamin Weiss. Ergodic theory and configurations in sets of positive density. In *Mathematics of Ramsey theory*, volume 5 of *Algorithms Combin.*, pages 184–198. Springer, Berlin, 1990. doi:10.1007/978-3-642-72905-8_13.

[18] William Timothy Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. doi:10.1007/s00039-001-0332-9.

[19] Ronald Lewis Graham. Recent trends in Euclidean Ramsey theory. *Discrete Math.*, 136(1-3):119–127, 1994. Trends in discrete mathematics. doi:10.1016/0012-365X(94)00110-5.

[20] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)*, 161(1):397–488, 2005. doi:10.4007/annals.2005.161.397.

[21] Alex. Iosevich, Gail Jardine, and Brian McDonald. Cycles of arbitrary length in distance graphs on $\mathbb{F}_q^d$. *Tr. Mat. Inst. Steklova*, 314:31–48, 2021. English version published in Proc. Steklov Inst. Math. **314** (2021), no. 1, 27–43. doi:10.4213/tm4189.

[22] Alex Iosevich, Akos Magyar, Alex McDonald, and Brian McDonald. The VC-dimension and point configurations in $\mathbb{R}^d$. *arXiv preprint*, 2025. URL: https://arxiv.org/abs/2510.13984, arXiv:arXiv:2510.13984.

[23] Alex Iosevich, Brian McDonald, and Maxwell Sun. Dot products in $\mathbb{F}_q^3$ and the Vapnik-Chervonenkis dimension. *Discrete Math.*, 346(1):Paper No. 113096, 9, 2023. doi:10.1016/j.disc.2022.113096.

[24] Vjekoslav Kovač. Boundedness of the twisted paraproduct. *Rev. Mat. Iberoam.*, 28(4):1143–1164, 2012. doi:10.4171/RMI/707.

[25] Vjekoslav Kovač. Density theorems for anisotropic point configurations. *Canad. J. Math.*, 74(5):1244–1276, 2022. doi:10.4153/S0008414X21000225.

[26] Vjekoslav Kovač and Bruno Predojević. Large dilates of hypercube graphs in the plane. *Anal. Math.*, 50(3):893–915, 2024. doi:10.1007/s10476-024-00045-6.

[27] Neil Lyall and Ákos Magyar. Product of simplices and sets of positive upper density in $\mathbb{R}^d$. *Math. Proc. Cambridge Philos. Soc.*, 165(1):25–51, 2018. doi:10.1017/S0305004117000184.

[28] Neil Lyall and Ákos Magyar. Distance graphs and sets of positive upper density in $\mathbb{R}^d$. *Anal. PDE*, 13(3):685–700, 2020. doi:10.2140/apde.2020.13.685.

[29] Neil Lyall and Ákos Magyar. Weak hypergraph regularity and applications to geometric Ramsey theory. *Trans. Amer. Math. Soc. Ser. B*, 9:160–207, 2022. doi:10.1090/btran/61.

[30] Brian McDonald. *Configuration Problems Over Finite Fields and the Vapnik-Chervonenkis Dimension*. ProQuest LLC, Ann Arbor, MI, 2023. Thesis (Ph.D.)–University of Rochester. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqm&rft_dat=xri:pqdiss:30529399.

[31] Brian McDonald, Anurag Sahay, and Emmett L. Wyman. The VC dimension of quadratic residues in finite fields. *Discrete Math.*, 348(1):Paper No. 114192, 13, 2025. doi:10.1016/j.disc.2024.114192.

[32] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[33] Terence Tao. Exploring the toolkit of Jean Bourgain. *Bulletin of the American Mathematical Society*, 58(2):155–171, 2021. Published electronically January 27, 2021. `doi:10.1090/bull/1716`.

[34] Vladimir Naumovich Vapnik and Alexey Yakovlevich Chervonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Teor. Verojatnost. i Primenen.*, 16:264–279, 1971.