

COLORING AND DENSITY THEOREMS FOR CONFIGURATIONS OF A GIVEN VOLUME

VJEKOSLAV KOVAČ

ABSTRACT. This is a treatise on finite point configurations spanning a fixed volume to be found in a single color-class of an arbitrary finite (measurable) coloring of the Euclidean space \mathbb{R}^n , or in a single large measurable subset $A \subseteq \mathbb{R}^n$. More specifically, we study vertex-sets of simplices, rectangular boxes, and parallelotopes, attempting to make progress on several open problems posed in the 1970s and the 1980s. As one of the highlights, we give the negative answer to a question of Erdős and Graham, by coloring the Euclidean plane \mathbb{R}^2 in 25 colors without creating monochromatic rectangles of unit area. More generally, we construct a finite coloring of the Euclidean space \mathbb{R}^n such that no color-class contains the 2^m vertices of any (possibly rotated) m -dimensional rectangular box of volume 1. A positive result is still possible if rectangular boxes of merely sufficiently large volumes are sought in a single color-class of a finite measurable coloring of \mathbb{R}^n , and we establish it under an additional assumption $n \geq m + 1$. Also, motivated by a question of Graham on reasonable bounds in his result on monochromatic axes-aligned right-angled m -dimensional simplices, we establish its measurable coloring and density variants with polylogarithmic bounds, again in dimensions $n \geq m + 1$. Next, we generalize a result of Erdős and Mauldin, by constructing an infinite measure set $A \subseteq \mathbb{R}^n$ such that every n -parallelotope with vertices in A has volume strictly smaller than 1. Finally, some results complementing the literature on isometric embeddings of hypercube graphs and on the hyperbolic analogue of the Hadwiger–Nelson problem also follow as byproducts of our approaches.

CONTENTS

1. Introduction	2
1.1. Triangles and simplices	3
1.2. Rectangles and rectangular boxes	6
1.3. Parallelograms and parallelotopes	7
1.4. Distance graphs in the plane	8
1.4.1. Embeddings of hypercube graphs	9
1.4.2. Hyperbolic embeddings	10
1.5. Paper outline and the main ideas	12
2. Rectangles: proof of Theorem 3	12
3. Rectangular boxes: proof of Theorem 4	14
4. Parallelograms: proof of Theorem 6	19
5. Hypercube graphs: proof of Theorem 8	19
6. Parallelotopes: proof of Theorem 7	20
7. Notation and preliminaries from harmonic analysis	21

2020 *Mathematics Subject Classification.* Primary 05D10; Secondary 52C10, 28A75, 42B20.

Key words and phrases. Euclidean Ramsey theory, geometric measure theory, regularity decomposition, heat flow, multilinear singular integral.

8. Right simplices: proofs of Theorem 1 and Corollary 2	24
9. Large rectangular boxes: proof of Theorem 5	32
10. Hyperbolic embeddings: proof of Theorem 10	39
11. Remarks for harmonic analysts	43
Acknowledgments	45
References	45

1. INTRODUCTION

Euclidean Ramsey theory mainly investigates existence of a given point configuration in an arbitrary finite coloring of the Euclidean space \mathbb{R}^n or its subsets. Its systematic study was initiated by Erdős, Graham, Montgomery, Rothschild, Spencer, and Straus [27, 28, 29] in the 1970s. On the other hand, *geometric measure theory* took a similar route in the 1980s and it started investigating point configurations that are found in arbitrary large measurable subsets of \mathbb{R}^n . Here we call a set $A \subseteq \mathbb{R}^n$ *large* if it has a positive density (as defined below) or, at least, a positive measure. Besides obvious connections to the Ramsey theory, the latter topic also shares the aesthetics with Szemerédi-type theorems from additive combinatorics. In fact, it can even be thought of as the arithmetic combinatorics in a geometric setting, typically formulated in the continuous parameter rather than a discrete one.

The present paper studies coloring and density theorems for rather flexible point configurations, from which we only require that they span a body of a given shape and a given volume. This particular subtopic has been initiated by Graham [34] and Erdős [25]. This treatise does not exhaust the subject, but it rather concentrates on the simplest configurations, such as simplices, rectangular boxes, and parallelotopes, as they exhibit different behaviors already. Our main goal is to show that many natural questions (or their variants) posed in the classical collections of open problems [26, 25] (also see [26, 12, 33]) are manageable by the techniques of harmonic analysis that have been developed in the meantime. However, quite a few basic questions are still left open.

Let us agree on the terminology. A *finite coloring* of \mathbb{R}^n (or its subset S) is any partition of \mathbb{R}^n (or S) into finitely many sets $\mathcal{C}_1, \dots, \mathcal{C}_r$, which are then called *color-classes*. A set T is *monochromatic* if $T \subseteq \mathcal{C}_j$ for some index $1 \leq j \leq r$. We say that the coloring is *measurable* if each of the color-classes \mathcal{C}_j is a Lebesgue-measurable set. Similarly, we say that the coloring is *Jordan-measurable* if the boundary of \mathcal{C}_j has Lebesgue measure 0 for every index j . Results that hold for an arbitrary finite (measurable) coloring of \mathbb{R}^n (or an $S \subseteq \mathbb{R}^n$), perhaps with a prescribed number of colors, are sometimes referred to as *coloring theorems*.

Next, the Lebesgue measure of a measurable set $A \subseteq \mathbb{R}^n$ (i.e., the length, the area, or, generally, the volume) will simply be written as $|A|$. If A is a measurable subset of the cube $[0, R]^n$ for some $R > 0$, then we will refer to its *density* as the number $|A|/R^n$. On the other hand, for a general measurable $A \subseteq \mathbb{R}^n$ the relevant notion of largeness is the *upper density*, defined as

$$\bar{d}_n(A) := \limsup_{R \rightarrow \infty} \frac{|A \cap [-R/2, R/2]^n|}{R^n}.$$

Results that hold for an arbitrary measurable $A \subseteq \mathbb{R}^n$ (or $A \subseteq [0, R]^n$) of strictly positive (upper) density are sometimes called *density theorems*. A slightly more convenient notion is

the *upper Banach density*,

$$\bar{\delta}_n(A) := \limsup_{R \rightarrow \infty} \sup_{x \in \mathbb{R}^n} \frac{|A \cap (x + [0, R]^n)|}{R^n}.$$

It is clearly always true that $\bar{\delta}_n(A) \geq \bar{d}_n(A)$, so we make a density result stronger by assuming that A merely has positive upper Banach density.

The tools from harmonic analysis employed in this work can be traced back to a paper by Bourgain [7] and they have been strengthened or modified by a number of authors over the years. Some of the work that has been influential to the author of the present paper had been done by Lyall and Magyar [43, 46, 44] and by Cook, Magyar, and Pramanik [11]. Instead of listing the extensive literature we refer the reader to a partially expository paper [40] or a brief survey of recent results [41]. Harmonic analysis tools are well-suited for establishing density theorems in a fixed ambient dimension n , which, in turn, imply their measurable coloring variants (cf. Theorems 1, 5, and 10 below). Indeed, if $\mathcal{C}_1, \dots, \mathcal{C}_r$ is any measurable coloring of \mathbb{R}^n , then

$$\bar{\delta}_n(\mathcal{C}_1) + \dots + \bar{\delta}_n(\mathcal{C}_r) \geq \bar{\delta}_n(\mathbb{R}^n) = 1,$$

so we must have

$$\bar{\delta}_n(\mathcal{C}_j) \geq \frac{1}{r} > 0$$

for at least one index $1 \leq j \leq r$. (An analogous claim holds for measurable colorings of the cube $[0, R]^n$.) Contrapositively, if we construct a measurable counterexample to a potential coloring result, then the corresponding density result is not possible either (cf. Theorems 3, 4, 6, and 8 below).

Admittedly, the measurability assumption is a seriously restricting one. There are many examples in the literature where the coloring result becomes significantly easier when only measurable colorings are considered; see Soifer's book [50, Ch. 9]. For instance, the well-known *Hadwiger–Nelson problem* asks for a minimal number of colors needed to color \mathbb{R}^2 so that all pairs of points at distance 1 apart receive different colors. It is easy to see that 7 colors suffice. The exact minimal number has naturally been called the *chromatic number of the plane*. Back in 1981, Falconer [30] showed that the measurable chromatic number of \mathbb{R}^2 is at least 5, but it took another 37 years until de Grey [16] proved the same for the ordinary chromatic number. Also, from an old result by Furstenberg, Katznelson, and Weiss [32] it follows that a measurable set $A \subseteq \mathbb{R}^2$ satisfying $\bar{d}_2(A) > 0$ contains pairs of points at every sufficiently large distance from each other. This has been reproved independently by Falconer and Marstrand [31] and also by Bourgain [7]. As opposed to that, only recently Davies [14] solved a 1994 conjecture of Rosenfeld [50, §60.1] (popularized by Erdős), by showing that there exists a pair of points at an odd distance from each other for every (not necessarily measurable) finite coloring of \mathbb{R}^2 . This was generalized to prime and polynomial distances by Davies, McCarty, and Pilipczuk [15]. A question by Kahle about distances of the form $n!$ [50, Problem 60.14] and the one by Soifer on distances of the form 2^n [50, Problem 60.16] remain open.

1.1. Triangles and simplices. Graham [34] answered positively a question of Gurevich, who asked whether, for positive integers $m \leq n$ and every finite coloring of \mathbb{R}^n , there exists an m -dimensional simplex of volume 1 with all of its $m + 1$ vertices colored the same [34, p. 89]. In fact, Graham could choose the simplices to be right-angled with axes-aligned perpendicular edges, i.e., the edges coming from a single vertex are parallel to the coordinate axes. Graham

combined his observations with a general product theorem for colorings [34, Theorem 3], which he developed particularly for this purpose; also see the generalization that he proved jointly with Spencer [35]. His most general coloring theorem for simplices, namely [34, Theorem 4], reads as follows.

For all finite colorings of \mathbb{R}^n , some color-class contains, for all $V > 0$ and lines ℓ_1, \dots, ℓ_n which span \mathbb{R}^n , a simplex having volume V and edges through one vertex parallel to the ℓ_i .

Indeed, the “slanted” variant of this result is clearly equivalent to the axes-aligned one, simply by applying a volume-preserving linear transformation. Several variants of Graham’s result were studied in [17, 3, 1].

By inspecting Graham’s proof (already for triangles) one notices that he first works on a huge part of the integer lattice \mathbb{Z}^n [34, Theorem 1], the size of which is an iterated van der Waerden number in the number of colors used. Then he scales the lattice to obtain the result in \mathbb{R}^n , but also wonders [34, p. 97] if a proof using only a reasonable number of integer points (or, strictly speaking, only reasonably large integer-point triangles) is possible. A natural continuous-parameter version of Graham’s question is the one that considers only colorings of a sufficiently large cube $[0, R]^n \subseteq \mathbb{R}^n$.

Problem 1. *Let m, n be fixed positive integers such that $n \geq m$. Is there a reasonable lower bound (i.e., not of the Ackermann type) on the number $R = R(r) \in (0, \infty)$ such that, for every coloring of the n -cube $[0, R]^n$ in r colors, there exists a right-angled m -dimensional simplex of unit volume with monochromatic vertices?*

We are able to give the positive answer for measurable colorings only and under the additional assumption $n \geq m + 1$ in part (b) of the following theorem. At the same time, we establish its corresponding density variant in part (a).

Theorem 1. *For every integer $m \geq 2$ there exists a constant $C_m \in (0, \infty)$ with the following properties for every $n \geq m + 1$.*

(a) *If $R \in (1, \infty)$ and $A \subseteq [0, R]^n$ is a measurable set with density*

$$\frac{|A|}{R^n} \geq \left(\frac{C_m}{\log R} \right)^{1/(9m^2)}, \quad (1.1)$$

then A contains $m + 1$ vertices of a right-angled m -dimensional simplex of unit volume.

(b) *If $R \in (0, \infty)$ and if the n -cube $[0, R]^n$ is measurably colored in r colors, then*

$$R \geq \exp(C_m r^{9m^2})$$

is sufficient to guarantee the existence of a right-angled m -dimensional simplex of unit volume with monochromatic vertices.

In both parts of the theorem, the m -simplex can be chosen such that $m - 1$ of its edges from the right-angled vertex are parallel to the coordinate vectors $\mathbf{e}_1, \dots, \mathbf{e}_{m-1}$, while the remaining edge is parallel to the linear span of $\mathbf{e}_m, \dots, \mathbf{e}_n$.

Some possible positions of such right triangles in \mathbb{R}^3 are depicted in Figure 1. Clearly, part (b) is a consequence of part (a), since, by our introductory comments, at least one color-class has density δ at least $1/r$ and since condition (1.1) can be rewritten as

$$R \geq \exp(C_m \delta^{-9m^2})$$

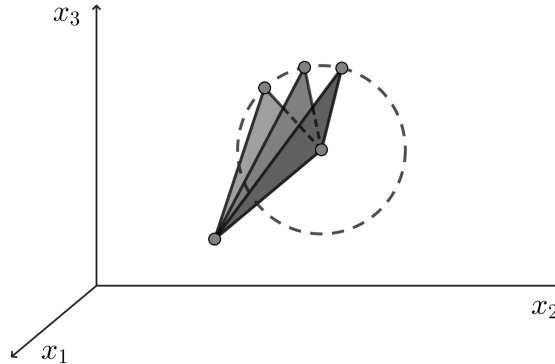


FIGURE 1. Illustration of Theorem 1.

in terms of the density δ . The emphasis in Theorem 1 is on its effective bounds: a polylogarithmic one in part (a) and an exponential one (in the power of r) in part (b). The order of quantifiers is important: we want a density theorem in a fixed dimension n . It is not difficult to fix a number $\delta > 0$ and find a sufficiently large n such that the result holds for regular simplices in subsets of \mathbb{R}^n with density at least δ . In this setting, quantitative interdependence of n and δ has been studied by Castro-Silva, de Oliveira Filho, Slot, and Vallentin [10].

Graham suggested (without sharing many details) the possibility of a density variant of his discrete result [34, Theorem 1] in Concluding Remarks [34, p. 96], but he mentioned general triangles only, i.e., not axes-aligned or even right-angled ones, which makes the problem easier. Indeed, the corresponding analogue in the plane \mathbb{R}^2 is indeed quite easy: it is an old observation of Erdős, first published as an exercise in the popular magazine *Matematikai Lapok*; see the comments in [24, p. 122].

A measurable set $A \subseteq \mathbb{R}^2$ of infinite area (and even an unbounded set of positive area) necessarily contains the vertices of a triangle of area 1.

The proof is a simple consequence of the one-dimensional Steinhaus theorem [53] and it was elaborated on in [12, p. 182]. The corresponding problem for right-angled triangles Erdős mentioned as being open [25, p. 324].

Problem 2 (Erdős [25]). *Let $A \subseteq \mathbb{R}^2$ be a measurable set with infinite area. Must A contain the vertices of a right triangle of area 1?*

In the last problem we cannot also ask from the triangle to be axes-aligned. Indeed, the example of a rotated thin infinite strip/tube shows that, if $A \subseteq \mathbb{R}^n$ merely has infinite volume, then the perpendicular sides of right-angled unit triangles/simplices with vertices in A can avoid any prescribed tuple of orthogonal directions. The same example also shows that the simplices with vertices in a set A of merely positive area can all have disproportional ratios of their edge-lengths. For this reason we find non-trivial the following consequence of Theorem 1(a), which deals with sets of positive upper Banach density.

Corollary 2. *Let $m \geq 2$ and $n \geq m + 1$ be positive integers. There exists a constant $C'_m \in (0, \infty)$, depending only on m , with the following property: if a measurable set $A \subseteq \mathbb{R}^n$ satisfies $\bar{\delta}_n(A) > 0$, then for every number $V > 0$ there exists a right-angled m -dimensional simplex*

of m -volume V with all $m + 1$ vertices in A and with the ratio of lengths of any two of its perpendicular edges being at most $\exp(C'_m \bar{\delta}_n(A)^{-9m^2})$.

Even a positive solution to Problem 2 and its higher-dimensional generalizations still would not cover the quantitative aspect of Corollary 2.

An interesting related problem is an old question of Erdős on triangles in planar sets A of merely sufficiently large finite area [24, p. 122–123].

Problem 3 (Erdős [24]). *Is it true that there is an absolute constant C so that, if $A \subseteq \mathbb{R}^2$ has area greater than C , then A contains the vertices of a triangle of area 1?*

Problem 3 has since been featured in numerous collections of open problems, such as the book *Unsolved problems in geometry* by Croft, Falconer, and Guy [12, Problem G13], Bloom's website *Erdős problems* [6, #352], or the expository paper by Mauldin [47]. The techniques in this paper are insufficient to detect configurations in sets that are only known to have large measure.

1.2. Rectangles and rectangular boxes. Graham [34] naturally wondered if his result that we discussed in the previous section

(...) can be generalized to configurations which are not simplexes. [34, p. 96].

Erdős and Graham posed a natural follow-up question in their well-known problem book [26].

The question is: Is this also true for rectangles? Or perhaps parallelograms? [26, p. 15].

To the best of author's knowledge, these problems, either for rectangles or for parallelograms, have not been addressed in the literature so far. They were explicitly mentioned as being open in the papers by Adhikari and Rath [2, 3] and in *Mathematical Reviews* MR0558877 by K. Steffens and MR2106573 by S. Oates-Williams. They were also recently stated among the open questions in the 2015 edition of the book *Rudiments of Ramsey theory* by Graham and Butler [33, p. 56].

Let us first discuss the variant for rectangles. The following standalone wording of the problem is taken from Bloom's website *Erdős problems* [6, #189].

Problem 4 (Erdős and Graham [34, 26]). *If \mathbb{R}^2 is finitely colored then must there exist some color-class which contains the vertices of a rectangle of every given area? ✓*

Note that the rectangles can be arbitrary rotated. We are able to show that Problem 4 has the negative answer.

Theorem 3. *There exists a Jordan-measurable coloring of the plane in 25 colors such that no color-class contains the vertices of a rectangle of area 1.*

With a bit more work we will also establish a natural generalization to higher dimensions; see Theorem 4 below. However, we have singled out the two-dimensional case in Theorem 3 above, because its proof is elegant and constructive.

Theorem 4. *For every positive integer n there exists a finite Jordan-measurable coloring of the Euclidean space \mathbb{R}^n with the following property: for every positive integer $m \leq n$ there is no m -dimensional rectangular box of m -volume equal to 1 with all of its 2^m vertices colored the same.*

The order of quantifiers is important again. The analogue of Theorem 4 would not hold and Theorem 5 below would become trivial if the ambient dimension n could depend on the number of colors used. One could find sufficiently large n simply by the known fact that vertices of a unit cube (or any other rectangular box) form a so-called *Ramsey configuration*; see [27].

After the negative result in Theorem 4 it is perhaps still possible that, for every finite coloring of \mathbb{R}^n , there exist monochromatic m -dimensional rectangular boxes of all sufficiently large m -volumes. We are able to show that this is indeed the case as long as $n \geq m + 1$ and only measurable colorings are considered. In fact, we can even prove a density version of that result thanks to the fact that the techniques for studying large copies of finite configurations in sets of positive density have been developed substantially over the last 40 years (cf. the literature mentioned at the beginning of the introductory section).

Theorem 5. *Take positive integers m and n such that $n \geq m + 1$.*

- (a) *If $A \subseteq \mathbb{R}^n$ is a measurable set such that $\bar{\delta}_n(A) > 0$, then there exists a number $V_0 > 0$ (depending on the set A) such that for every number $V \geq V_0$ there exists an m -dimensional rectangular box of m -volume V with all 2^m vertices in the set A .*
- (b) *For every finite measurable coloring of \mathbb{R}^n there exist a color-class \mathcal{C} and a number $V_0 > 0$ such that for every number $V \geq V_0$ there exists an m -dimensional rectangular box of m -volume V with all vertices in \mathcal{C} .*

In both parts of the theorem, rectangular boxes can be chosen to have $m - 1$ edges parallel to the first $m - 1$ coordinate vectors and the last edge parallel to the linear span of the remaining $n - m + 1$ coordinate vectors.

Theorem 5 is only new when the ambient dimension n satisfies $m + 1 \leq n \leq 2m - 1$. It follows from the existing literature in the cases $n \geq 2m$, as Lyall and Magyar [46, Theorem 1.1(i)] showed that a set $A \subseteq \mathbb{R}^{2m}$ of positive upper Banach density contains the vertices of an m -cube with every sufficiently large edge-length. The critical case $n = m \geq 2$ is not covered by our theorem and it remains open. The analogous result for $n = m = 1$ clearly fails, as is seen by taking $A = (-1/5, 1/5) + \mathbb{Z}$. Thus, the following problem remains.

Problem 5. *If $n \geq 2$ and \mathbb{R}^n is measurably colored in finitely many colors, must there exist a color-class which contains the 2^n vertices of a rectangular box of every sufficiently large volume?*

It is also unclear if the measurability condition can be removed from the statement of Theorem 5(b).

1.3. Parallelograms and parallelotopes. The situation changes radically if the boxes are “slanted”, i.e., replaced with parallelograms, parallelepipeds, or, more generally, parallelotopes. Currently, we are not able to solve the parallelogram variant of Problem 4.

Problem 6 (Erdős and Graham [34, 26]). *If \mathbb{R}^2 is finitely colored then must there exist some color-class which contains the vertices of a parallelogram of every given area?*

Unable to answer the question in Problem 6, we only state a partial result that will be obtained by modifying easily the ideas used for Theorem 3.

Theorem 6. *Suppose that we are given finitely many lines, $\ell_1, \dots, \ell_m \subset \mathbb{R}^2$, and a positive number ε . There exists a Jordan-measurable coloring of \mathbb{R}^2 such that there is no parallelogram*

of area 1 with monochromatic vertices that, additionally, has one side parallel to some line ℓ_i or it has all angles greater than ε .

In other words, one should concentrate on finding almost degenerate parallelograms and infinitely many directions should be considered. For these reasons Problem 6 could be difficult, especially if it turns out that the claim has a positive answer.

As a related observation, Erdős remarked [25, p. 324] (without details or a reference) that he and Mauldin constructed a set $A \subseteq \mathbb{R}^2$ of infinite area that does not contain vertices of a parallelogram of area 1 (also see the comments under Problem [6, #353]). Even if this has no implications to finite colorings of \mathbb{R}^2 or to positive density subsets of \mathbb{R}^2 , it still hints that sets that avoid the vertices of parallelograms of unit area can be quite large. Without knowing which set Erdős and Mauldin had in mind, we observe that the example can be quite simple: an (x, y) -region between the positive parts of the coordinate axes $x = 0$ and $y = 0$ and the hyperbola $2xy = 1$. In fact, we will easily prove the following stronger higher-dimensional result.

Theorem 7. *Let $n \geq 2$ be a positive integer and take $\varepsilon > 0$. There exists a Jordan-measurable set (i.e., a set with boundary of measure 0) $A \subseteq \mathbb{R}^n$ of infinite volume such that every n -dimensional parallelotope with all 2^n vertices in A has volume less than ε .*

One has to note that Theorem 7 is a rather simple result, using a more straightforward construction than Theorem 4.

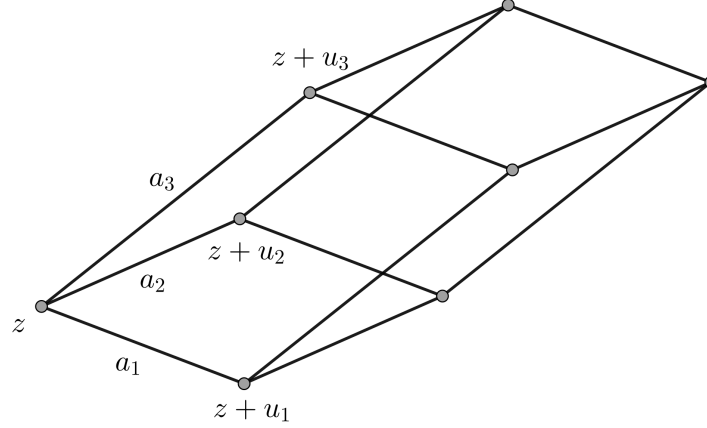
1.4. Distance graphs in the plane. Here we discuss two lines of research that, at first, do not seem to be related to the paper topic, as they will be concerned with certain non-rigid structures, as opposed to polytopes. However, it will turn out they are, in fact, related to Theorem 3 from Subsection 1.2 and to Graham's theorem stated in Subsection 1.1, so they fit naturally in the present paper too. For this reason, unsurprisingly, we will be able to reuse the same techniques applied to our previous results, both positive and negative ones.

Lyall and Magyar [44] initiated the study of embeddings of distance graphs in measurable sets $A \subseteq \mathbb{R}^n$ of positive upper Banach density. The most definite embedding result is available for distance trees, which were also studied (in similar contexts) in [9, 37, 45], so let us spend a few words to formulate it here.

Let $\mathcal{T} = (V, E)$ be a finite tree with the set of vertices V and the set of edges E . To every edge $k \in E$ we assign a number $a_k \in (0, \infty)$ interpreted as *edge-length*, so we call the resulting structure a *distance tree*. It follows from a more general result by Lyall and Magyar [44, Theorem 2(i)] that

for every distance tree \mathcal{T} and every set $A \subseteq \mathbb{R}^2$ of positive upper Banach density, there exists a number $\lambda_0 > 0$ (depending on \mathcal{T} and A) such that for every number $\lambda \geq \lambda_0$ the set A contains the vertices of an isometric copy of \mathcal{T} scaled by the factor λ .

Note that a non-scaled tree (or any other graph) need not be isometrically embeddable in A , simply because there is no reason why A should contain two points at distance 1 apart (cf. the Hadwiger–Nelson problem mentioned before). The above result is also dimensionally sharp: sufficiently large distances need not be found in a positive density subset of \mathbb{R} ; take $A = [-1/5, 1/5] + \mathbb{Z}$.


 FIGURE 2. Embedding of a 1-skeleton of an n -box.

1.4.1. *Embeddings of hypercube graphs.* Embeddings of more general distance graphs are much more complicated and only partially resolved. Whenever the graph has a cycle, one can choose its edge-lengths to violate the triangle inequality, so it will not be embeddable in the Euclidean space. Lyall and Magyar [44] resolved this problem by starting from a distance graph that already is embedded in some other Euclidean space. Then they gave sufficient conditions for it to have all large dilates found in every set $A \subseteq \mathbb{R}^d$ of positive upper Banach density. However, even when their [44, Theorem 2] applies, it is rarely known what the minimal ambient dimension d can be. Surprisingly, this is open already in the case of the equilateral (or any other non-degenerate) triangle.

Problem 7 (implicit in [32], stated as [12, Problem G14]). *Do all sets $A \subseteq \mathbb{R}^2$ of positive upper density contain a congruent copy of the vertices of every sufficiently large equilateral triangle?*

Note that the corresponding result for triangles in \mathbb{R}^3 is known, as it is a special case of Bourgain's [7, Theorem 2].

Here we study a particular case of a graph, which is called a *hypercube graph* in the combinatorial sense, but as a distance graph it is better thought of as a 1-skeleton of an n -dimensional box

$$[0, a_1] \times [0, a_2] \times \cdots \times [0, a_n]. \quad (1.2)$$

Its embedding in \mathbb{R}^2 is simply a set of 2^n mutually distinct points

$$z + r_1 u_1 + r_2 u_2 + \cdots + r_n u_n \quad \text{for } (r_1, r_2, \dots, r_n) \in \{0, 1\}^n \quad (1.3)$$

such that $z \in \mathbb{R}^2$ and $u_1, \dots, u_n \in \mathbb{R}^2$ are vectors with lengths $|u_k| = a_k$ for $1 \leq k \leq n$; see Figure 2.

This case of an embedding in \mathbb{R}^2 was not covered by [44, Theorem 2]; their result would require the ambient dimension to be at least $n + 1$. Still, Predojević and the author [42, Theorem 1 and Remark 3] showed that

a measurable subset $A \subseteq \mathbb{R}^2$ of positive upper Banach density contains embeddings of all sufficiently large dilates of a fixed 1-skeleton of an n -box; for

instance, in A we can find points (1.3) with $|u_k| = a_k = \lambda$ for all sufficiently large numbers $\lambda \in (0, \infty)$.

This result generalizes the aforementioned result on distance trees, since every tree is a subgraph of a hypercube graph. A natural question is what happens when we assume a more flexible constraint relating the edge-lengths than $a_1 = a_2 = \dots = a_n = \lambda$, such as the product of the numbers a_k being 1.

Theorem 8. *There exists a Jordan-measurable finite coloring of \mathbb{R}^2 such that no color-class contains an embedding (1.3) of the 1-skeleton of an n -box (1.2) satisfying $a_1 \dots a_n = 1$.*

In particular, a measurable set $A \subseteq \mathbb{R}^2$ with $\bar{\delta}_2(A) > 0$ need not contain such an embedding. Theorem 8 will be shown by a simple modification of the construction used to prove Theorem 3 above. This might sound a bit surprising, since the quantity $a_1 \dots a_n$ is the volume of the original box (1.2), but it has no such interpretation for an embedding of its 1-skeleton (1.3).

1.4.2. Hyperbolic embeddings. Bardestani and Mallahi-Karai [5] studied a generalization of the Hadwiger–Nelson problem in which the points v and w of the same color are not allowed to satisfy the equation $Q(v-w) = 1$, where Q is a quadratic form over some local field. Under the additional measurability assumption, they were able to characterize when the corresponding chromatic number is infinite. Bardestani also informed the author that an easy modification of the argument from [5], which in turn uses a bound on the so-called independence ratio by Bachoc, DeCorte, de Oliveira Filho, and Vallentin [4], also implies a density variant of the same result in the plane \mathbb{R}^2 in the case of real hyperbolic forms; see Theorem 9 below. Davies [13] was motivated by a few unresolved questions in the particular hyperbolic case

$$Q: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad Q((x, y), (x', y')) = (x - x')^2 - (y - y')^2$$

and its higher-dimensional space-time generalizations. He observed that Graham’s result on right triangles is relevant in that matter and it already implies that the corresponding (non-measurable) chromatic number is infinite as well.

It is natural to consider a coloring of a large square $[0, R]^2$ and impose bounds on R sufficient to find equally colored points (x, y) and (x', y') such that

$$(x - x')^2 - (y - y')^2 = 1. \tag{1.4}$$

The arguments of Bardestani and Mallahi-Karai [5, p. 325–327] are also quantitative and they showed that this is true for measurable colorings in r colors under the condition of the form $R \geq \exp(Cr)$. Let us formulate this result in the spirit of Theorem 1.

Theorem 9 (Bardestani and Mallahi-Karai [5]). (a) *If $R \in (1, \infty)$ and $A \subseteq [0, R]^2$ is a measurable set with density*

$$\frac{|A|}{R^2} \geq \frac{1280}{\log R},$$

then there exist $(x, y), (x', y') \in A$ such that (1.4) holds.

(b) *If $R \in (0, \infty)$ and if the square $[0, R]^2$ is measurably colored in r colors, then*

$$R \geq e^{1280r}$$

guarantees that there exist equally colored points (x, y) and (x', y') satisfying (1.4).

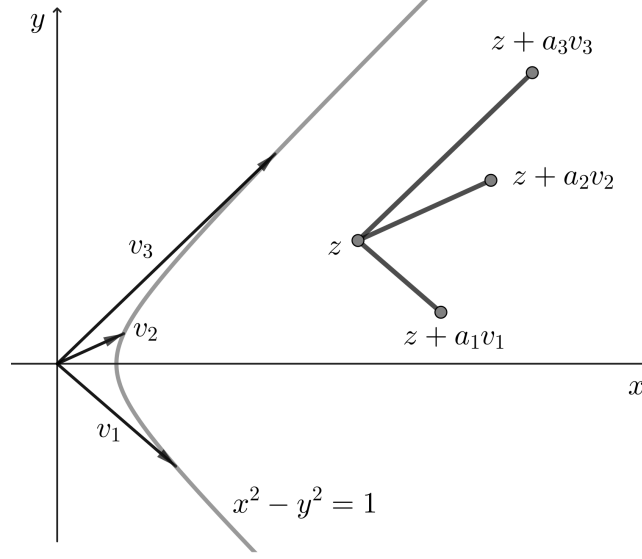


FIGURE 3. Illustration of Theorem 10.

In comparison, the argument of Davies [13] based on Graham's theorem applies to non-measurable colorings too, but it requires an Ackermann-type bound on R in terms of r .

We remark, once again, that Theorem 9 is not an original result that we show here. As we have mentioned, it was not explicitly formulated in [5], but it can be deduced rather easily from that paper, as observed by its authors. Our techniques are better suited to larger patterns, so we show the following generalization to a configuration of $n + 1$ points, with a slightly weaker bound for $n = 1$ than Theorem 9. Let \mathcal{H} denote the standard hyperbola in \mathbb{R}^2 with the Cartesian equation

$$\mathcal{H} : x^2 - y^2 = 1.$$

It will replace the standard unit circle, $x^2 + y^2 = 1$, which plays a role in the usual Euclidean graph embedding theorems.

Theorem 10. *For every positive integer n and distinct numbers $a_1, \dots, a_n \in (0, \infty)$ there exists a constant $C \in (0, \infty)$ with the following properties.*

- (a) *If $R \in (1, \infty)$ and $A \subseteq [0, R]^2$ is a measurable set with density*

$$\frac{|A|}{R^2} \geq \left(\frac{C}{\log R} \right)^{1/(2n+1)},$$

then there exist $z \in \mathbb{R}^2$ and vectors $v_1, \dots, v_n \in \mathcal{H}$ such that the $n + 1$ points

$$z, z + a_1v_1, \dots, z + a_nv_n \tag{1.5}$$

all belong to the set A .

- (b) *If $R \in (0, \infty)$ and if the square $[0, R]^2$ is measurably colored in r colors, then*

$$R \geq \exp(Cr^{2n+1})$$

guarantees that there exist $n + 1$ equally colored points of the form (1.5) for some vectors $v_1, \dots, v_n \in \mathcal{H}$.

Part (b) is, once again, a clear consequence of part (a). We will use a similar outline and the same ideas employed in the proofs of Theorems 1 and 5 to establish Theorem 10 too. See Figure 3 for an illustration of the case $n = 3$. Note that, in this hyperbolic setting, we are not allowing the freedom to scale the hyperbola \mathcal{H} by some large factor λ . It is a non-compact curve, so it leaves more freedom than the circle to find non-scaled copies of the desired configuration. In the language of distance trees introduced at the beginning of Subsection 1.4, we are looking for certain hyperbolic embeddings of *star trees* (or simply *stars*).

As a consequence of Theorem 10(a), one immediately obtains the same property of sets $A \subseteq \mathbb{R}^2$ of positive upper Banach density. Clearly, it is not enough for A to merely have infinite measure, as is seen from the trivial example: $n = 1$, $a_1 = 1$, and $A = [-1/3, 1/3] \times \mathbb{R}$.

1.5. Paper outline and the main ideas. Each of Sections 2–6 and 8–10 is dedicated to the proof of a single theorem. We first establish the “negative” results, since they are more elementary. Theorem 3 relies on an explicit partition of the complex plane \mathbb{C} , by coloring the number z based on the relative location of z^2 to the (scaled) Gaussian integers $\mathbb{Z} + i\mathbb{Z}$. The construction could be thought of as a complex modification of the approach of Erdős at al. [27, §3] who used $|z|^2$ in place of z^2 . The proof of Theorem 6 uses the same idea, only with roughly $1/\sin \varepsilon$ many colors, while the proof of Theorem 8 extends it to higher complex powers, namely z^n for $n \geq 3$. The proof of Theorem 4 proceeds differently, by constructing the coloring that prevents monochromatic boxes that are only slightly rotated from their “standard” position and then using compactness of the group of rotations $\text{SO}(n)$. The trick of using compactness of the underlying transformation group is essentially due to Straus [54], but its applicability is far from automatic: for instance, it can be used for rectangles but not for general parallelograms. Finally, Theorem 7 has a rather straightforward inductive proof, by estimating the determinant representing the volume of a parallelotope.

Proofs of the “positive” density results rely on harmonic analysis, so a short Section 7 serves as an interlude, containing a few basic facts and some notation. The proof of Theorem 5 follows a similar outline as [20, §7] by Durcik and the author, which, in turn, stems from a general outline of performing a certain regularity decomposition of the counting form [7, 8, 11]. Similar proofs also appeared in a couple of other recent papers [40, 42], but here we also capitalize on the fact that the “fixed volume” condition is more flexible than fixing edge-lengths, so a lower ambient dimension suffices for that reason. The proof of Theorem 1 brings an additional novelty, by reinventing the notion of configuration size and defining it according to the lengths of some edges only, leaving the room for pigeonholing and achieving the unit volume by fixing the length of the remaining edge. In the proof of Theorem 10 we need to take care of several edges simultaneously, making it also similar to [40, §3].

Section 11, the closing one, contains a short philosophical discussion on why area 1 triangles allow a density theorem (in sufficiently large dimensions), while area 1 rectangles do not. We draw analogies with singular integral forms from multilinear harmonic analysis.

The author hopes that the present paper will revive the topic of studying existence theorems for fixed-volume point configurations, which seems to have been explored in only a few papers [47, 17, 3, 1, 48] since the early work of Graham [34].

2. RECTANGLES: PROOF OF THEOREM 3

We are about to give a coloring of \mathbb{R}^2 that uses 25 colors and has a slightly stronger property: no color-class will contain the vertices of a parallelogram such that the product of lengths of

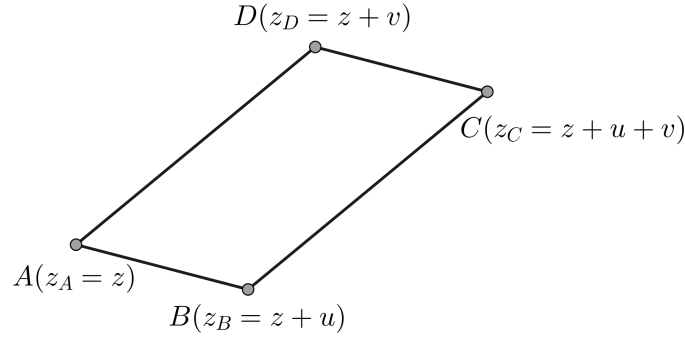


FIGURE 4. Coordinatization of a parallelogram.

its two consecutive sides equals 1. For rectangles this clearly specializes to the property of their area being equal to 1.

Proof of Theorem 3. Let us place a (possibly degenerate) parallelogram $\mathcal{P} = ABCD$ in the complex plane, so that its vertices A, B, C, D are respectively coordinatized by the complex numbers z_A, z_B, z_C, z_D as in Figure 4. Consider a complex quantity $\mathcal{J}(\mathcal{P})$ defined as

$$\mathcal{J}(\mathcal{P}) := z_A^2 - z_B^2 + z_C^2 - z_D^2. \quad (2.1)$$

In this definition we specify the vertex A to be the one with the smallest coordinate z_A in the lexicographic ordering of $\mathbb{C} \equiv \mathbb{R}^2$. Otherwise, $\mathcal{J}(\mathcal{P})$ would have only been determined up to multiplication by ± 1 .

There exist $u, v, z \in \mathbb{C}$ such that the vertices of \mathcal{P} have complex coordinates

$$z_A = z, \quad z_B = z + u, \quad z_C = z + u + v, \quad z_D = z + v;$$

see Figure 4 again. The quantity $\mathcal{J}(\mathcal{P})$ now simplifies as

$$\mathcal{J}(\mathcal{P}) = z^2 - (z + u)^2 + (z + u + v)^2 - (z + v)^2 = 2uv.$$

Consecutive side lengths of \mathcal{P} are $|u|$ and $|v|$, so we have

$$|\mathcal{J}(\mathcal{P})| = 2$$

whenever their product equals 1. As we have already mentioned, this holds in particular if \mathcal{P} is a rectangle of area 1. Therefore, it remains to find a coloring of \mathbb{C} such that, if all vertices of \mathcal{P} are assigned the same color, then the complex number $\mathcal{J}(\mathcal{P})$ does not lie on the circle

$$\{w \in \mathbb{C} : |w| = 2\}. \quad (2.2)$$

For each pair $(j, k) \in \{0, 1, 2, 3, 4\}^2$ define a color-class $\mathcal{C}_{j,k}$ as

$$\mathcal{C}_{j,k} := \left\{ z \in \mathbb{C} : z^2 \in \frac{10}{3} \left(\mathbb{Z} + \mathbf{i}\mathbb{Z} + \frac{j + \mathbf{i}k}{5} + \left[0, \frac{1}{5}\right] + \mathbf{i} \left[0, \frac{1}{5}\right] \right) \right\}.$$

If the four vertices of $\mathcal{P} = ABCD$ belonged to the same color-class, then, by the definition (2.1), we would clearly have

$$\mathcal{J}(\mathcal{P}) \in \frac{10}{3} \left(\mathbb{Z} + \mathbf{i}\mathbb{Z} + \left(-\frac{2}{5}, \frac{2}{5} \right) + \mathbf{i} \left(-\frac{2}{5}, \frac{2}{5} \right) \right).$$

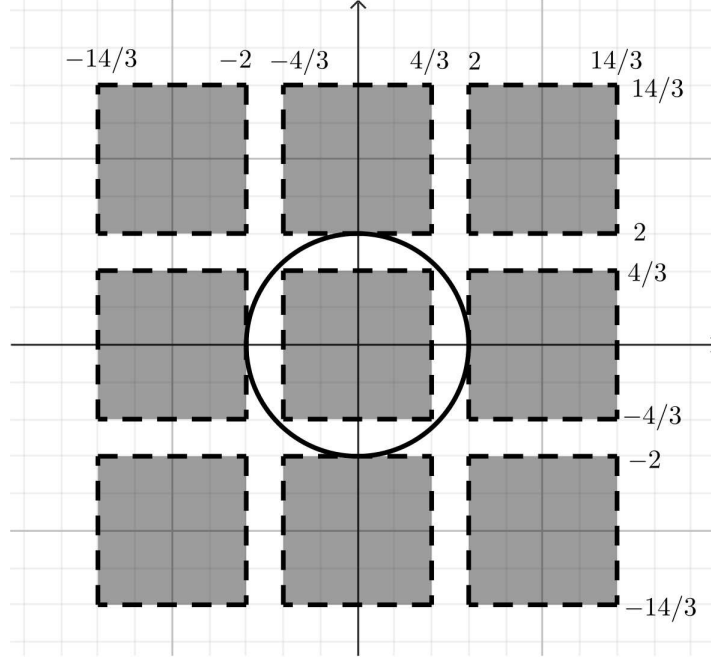


FIGURE 5. The circle misses the squares.

The above set does not intersect the circle (2.2); see Figure 5. Indeed, the central square lies fully inside (2.2) because of $4\sqrt{2}/3 < 2$, while all remaining open squares clearly belong to its exterior. \square

We can say that the above solution relies on the invariant quantity $|\mathcal{I}(\mathcal{P})|$ assigned to rectangles of area 1. It does not generalize to all higher dimensions, since it uses multiplication of complex numbers, so we will resort to an “almost invariant” quantity in the next section.

Let us illustrate the coloring constructed in the previous proof. Boundaries of the color-classes are given in the (x, y) -coordinate system by the equations

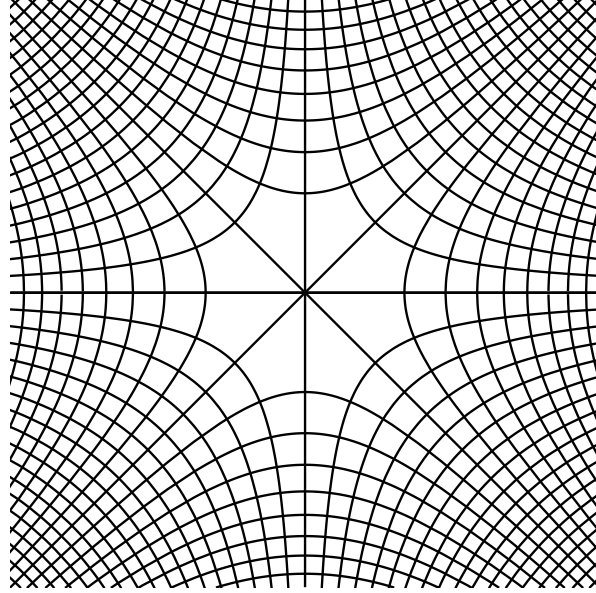
$$x^2 - y^2 = \frac{2a}{3} \quad \text{and} \quad xy = \frac{b}{3}$$

for arbitrary $a, b \in \mathbb{Z}$. These are two mutually orthogonal families of hyperbolas (including degenerate ones for $a = 0$ or $b = 0$), depicted in Figure 6.

3. RECTANGULAR BOXES: PROOF OF THEOREM 4

The idea is based on the following simple observation. Let us first take an m -dimensional rectangular box in \mathbb{R}^n with edge-lengths $a_1, a_2, \dots, a_m \in (0, \infty)$ and with sides parallel to the first m coordinate axes of \mathbb{R}^n . Its vertices can then be enumerated by subsets T of $\{1, 2, \dots, m\}$ as

$$\left(q + \sum_{j \in T} a_j \oplus_j : T \subseteq \{1, 2, \dots, m\} \right), \quad (3.1)$$


 FIGURE 6. Boundaries of color-classes $\mathcal{C}_{j,k}$.

where $q = (q_1, \dots, q_n) \in \mathbb{R}^n$ is some point, namely, one of its vertices. Let us compute the alternating sum of the product of the first m coordinates of the box vertices,

$$\sum_{T \subseteq \{1, 2, \dots, m\}} (-1)^{m-|T|} \left(\prod_{j \in T^c} q_j \right) \left(\prod_{j \in T} (q_j + a_j) \right) = \prod_{j=1}^m (-q_j + q_j + a_j) = a_1 a_2 \cdots a_m,$$

and notice that we have obtained precisely the box volume.

The crucial part of the proof will be to show that the same quantity is almost invariant for slightly rotated rectangular boxes, where the “slightness” can be prescribed uniformly over all box eccentricities. Note that this property is not quite obvious and it is sensitive to the pattern shape, as the uniformity will fail already for parallelograms in \mathbb{R}^2 . After we construct a coloring that prohibits those slightly tilted boxes, in the last step we will rotate it by finitely many matrices $U_1, \dots, U_L \in \text{SO}(n)$, thanks to compactness of the rotation group.

Before the proof we need a simple identity.

Lemma 3.1. *The identity*

$$\sum_{T \subseteq \{1, 2, \dots, m\}} (-1)^{m-|T|} \prod_{k=1}^m \left(p_k + \sum_{j \in T} v_{j,k} \right) = \sum_{\sigma \in \mathbb{S}_m} v_{1,\sigma(1)} v_{2,\sigma(2)} \cdots v_{m,\sigma(m)} \quad (3.2)$$

holds for complex numbers $(p_k)_{1 \leq k \leq m}$ and $(v_{j,k})_{1 \leq j, k \leq m}$.

Here \mathbb{S}_m denotes the group of permutations on the set $\{1, 2, \dots, m\}$. It is peculiar to notice that the right hand side RHS of (3.2) is the permanent of the matrix $(v_{j,k})_{j,k}$.

Proof of Lemma 3.1. We will prove the identity by induction on m . The basis case $m = 1$ is trivial as then the identity reads

$$-p_1 + (p_1 + v_{1,1}) = v_{1,1}.$$

Take a positive integer $m \geq 2$. Observe that the left hand side LHS of (3.2) is a homogeneous polynomial of degree m in $m^2 + m$ variables p_k and $v_{j,k}$, but the degree of each of the variables in it is at most 1. Differentiating it with respect to the variable $v_{m,l}$ for some $l \in \{1, \dots, m\}$ we obtain

$$\begin{aligned} \frac{\partial}{\partial v_{m,l}} \text{LHS} &= \sum_{T \subseteq \{1, \dots, m-1\}} (-1)^{m-1-|T|} \prod_{\substack{1 \leq k \leq m \\ k \neq l}} \left(p_k + v_{m,k} + \sum_{j \in T} v_{j,k} \right) \\ &= \sum_{\substack{\sigma \in \mathbb{S}_m \\ \sigma(m)=l}} v_{1,\sigma(1)} v_{2,\sigma(2)} \cdots v_{m-1,\sigma(m-1)}, \end{aligned}$$

where in the last equality we applied the induction hypothesis to a smaller collection of numbers

$$(p_k + v_{m,k})_{1 \leq k \leq m, k \neq l}, \quad (v_{j,k})_{1 \leq j \leq m-1, 1 \leq k \leq m, k \neq l}$$

and renamed the indices. Therefore

$$\text{LHS} - \sum_{l=1}^m v_{m,l} \frac{\partial}{\partial v_{m,l}} \text{LHS} = \text{LHS} - \text{RHS}$$

and this quantity does not depend on $v_{m,1}, \dots, v_{m,m}$, so it can be evaluated simply by plugging $v_{m,1} = \dots = v_{m,m} = 0$ into the left hand side:

$$\sum_{T \subseteq \{1, \dots, m\}} (-1)^{m-|T|} \prod_{k=1}^m \left(p_k + \sum_{\substack{1 \leq j \leq m-1 \\ j \in T}} v_{j,k} \right) = 0.$$

The last sum equals 0 because its terms can be paired to cancel each other: adding/subtracting the element m to/from the set T gives exactly the same term, only with the opposite signs. We have obtained $\text{LHS} - \text{RHS} = 0$. \square

Now we turn to the proof of the announced result.

Proof of Theorem 4. It is sufficient to fix integers $m \leq n$, since later we can consider a joint refinement (i.e., mutual class-wise intersections) of the constructed colorings of \mathbb{R}^n obtained for $m = 1, 2, \dots, n$. To each m -dimensional rectangular box \mathcal{R} in \mathbb{R}^n we assign a real-valued quantity $\mathcal{J}(\mathcal{R})$ defined as

$$\mathcal{J}(\mathcal{R}) := \sum_{x=(x_1, \dots, x_n) \text{ is a vertex of } \mathcal{R}} (-1)^{m-\text{par}(x)} x_1 \cdots x_m. \quad (3.3)$$

Note that only the first m out of n coordinates of a point x appear in the product $x_1 \cdots x_m$. Also, here $\text{par}(x)$ denotes the *parity* of a vertex x , which is computed as follows. Choose the base vertex of \mathcal{R} to be the one with the smallest coordinate representation in the lexicographic ordering of \mathbb{R}^n . The parity of any vertex of \mathcal{R} is defined to be the parity of its distance from the base vertex in the 1-skeleton graph of \mathcal{R} . This number is either 0 or 1 and it changes as we move from a vertex to its neighbor along a 1-edge of \mathcal{R} . Clearly, the expression (3.3) has 2^m terms, half of them get the $+$ sign and half of them come with the $-$ sign; see the illustration of these signs in Figure 7.

Every m -dimensional rectangular box $\mathcal{R} \subseteq \mathbb{R}^n$ can be obtained by rotating an axes-aligned box \mathcal{R}_0 about the origin,

$$\mathcal{R} = U\mathcal{R}_0, \quad (3.4)$$

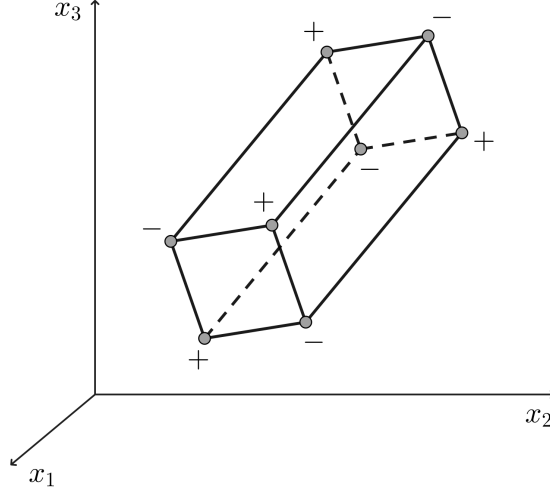


FIGURE 7. Signs attached to box vertices.

where the vertices of \mathcal{R}_0 are given by (3.1), while the rotation U is given by a special orthogonal transformation, $U \in \text{SO}(n)$. The vertices of \mathcal{R} can then be written as

$$\left(p + \sum_{j \in T} v_j : T \subseteq \{1, 2, \dots, m\}\right),$$

where $p = Uq$ and $v_j = a_j U e_j$ for each index $1 \leq j \leq m$. Also, write coordinate-wise:

$$p = (p_k)_{1 \leq k \leq n}, \quad v_j = (v_{j,k})_{1 \leq k \leq n} \quad (3.5)$$

and note that Lemma 3.1 gives

$$\mathcal{J}(\mathcal{R}) = \pm \sum_{\sigma \in \mathbb{S}_m} v_{1,\sigma(1)} v_{2,\sigma(2)} \cdots v_{m,\sigma(m)}. \quad (3.6)$$

First, suppose that the rotation U satisfies $\|U - I\|_{\text{op}} < \varepsilon$, where $\|T\|_{\text{op}}$ denotes the operator norm of a matrix T and

$$\varepsilon := \frac{1}{2^{m+2} m!}.$$

In particular,

$$\left| \frac{1}{a_j} v_j - e_j \right| = |(U - I)e_j| < \varepsilon,$$

so that

$$|v_{j,j} - a_j| < \varepsilon a_j$$

for $1 \leq j \leq m$ and

$$|v_{j,k}| < \varepsilon a_j$$

for $1 \leq j \leq m$, $1 \leq k \leq n$, $j \neq k$. Consequently,

$$\begin{aligned}
& \left| \sum_{\sigma \in \mathbb{S}_m} v_{1,\sigma(1)} v_{2,\sigma(2)} \cdots v_{m,\sigma(m)} - a_1 a_2 \cdots a_m \right| \\
& \leq |v_{1,1} v_{2,2} \cdots v_{m,m} - a_1 a_2 \cdots a_m| + \sum_{\substack{\sigma \in \mathbb{S}_m \\ \sigma \neq \text{id}}} |v_{1,\sigma(1)} v_{2,\sigma(2)} \cdots v_{m,\sigma(m)}| \\
& < m\varepsilon(1+\varepsilon)^{m-1} a_1 a_2 \cdots a_m + (m! - 1)\varepsilon(1+\varepsilon)^{m-1} a_1 a_2 \cdots a_m \\
& \leq 2^m m! \varepsilon a_1 a_2 \cdots a_m \leq \frac{1}{4} a_1 a_2 \cdots a_m.
\end{aligned} \tag{3.7}$$

Let us now additionally assume that \mathcal{R} has volume 1. Then (3.6) combined with (3.7) and $a_1 a_2 \cdots a_m = 1$ gives

$$\mathcal{J}(\mathcal{R}) \in \left(-\frac{5}{4}, -\frac{3}{4}\right) \cup \left(\frac{3}{4}, \frac{5}{4}\right). \tag{3.8}$$

Partition \mathbb{R}^n into the sets

$$\mathcal{S}_l := \left\{ (x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_1 x_2 \cdots x_m \in \frac{3}{2} \left(\mathbb{Z} + \left[\frac{l}{3 \cdot 2^m}, \frac{l+1}{3 \cdot 2^m} \right) \right) \right\}$$

for $0 \leq l \leq 3 \cdot 2^m - 1$. We claim that the vertices of \mathcal{R} cannot all belong to the same set \mathcal{S}_l . Namely, if they did, then the definition of $\mathcal{J}(\mathcal{R})$ would give

$$\mathcal{J}(\mathcal{R}) \in \frac{3}{2}\mathbb{Z} + \left(-\frac{1}{4}, \frac{1}{4}\right) = \cdots \cup \left(-\frac{7}{4}, -\frac{5}{4}\right) \cup \left(-\frac{1}{4}, \frac{1}{4}\right) \cup \left(\frac{5}{4}, \frac{7}{4}\right) \cup \cdots. \tag{3.9}$$

However, (3.8) and (3.9) together lead to a contradiction as the sets on their right hand sides are disjoint.

Finally, we handle completely arbitrary rectangular boxes \mathcal{R} of unit volume. Consider an open neighborhood \mathcal{O} of the identity I in the rotation group $\text{SO}(n)$ defined as

$$\mathcal{O} := \{V \in \text{SO}(n) : \|V - I\|_{\text{op}} < \varepsilon\}.$$

Then the family $\{U\mathcal{O} : U \in \text{SO}(n)\}$ constitutes an open cover of the compact space $\text{SO}(n)$, so it can be reduced to a finite subcover $\{U_1\mathcal{O}, U_2\mathcal{O}, \dots, U_L\mathcal{O}\}$. The color-classes of the desired coloring of \mathbb{R}^n can now be defined as

$$\mathcal{C}_{l_1, l_2, \dots, l_L} := (U_1 \mathcal{S}_{l_1}) \cap (U_2 \mathcal{S}_{l_2}) \cap \cdots \cap (U_L \mathcal{S}_{l_L}),$$

where (l_1, l_2, \dots, l_L) run over all L -tuples of elements from $\{0, 1, 2, \dots, 3 \cdot 2^m - 1\}$. In words, this is a joint refinement of L mutually rotated colorings. Suppose that the vertices of \mathcal{R} belong to the same color-class $\mathcal{C}_{l_1, l_2, \dots, l_L}$. Let $U \in \text{SO}(n)$ be as in (3.4), but without any assumption on the norm of $U - I$. Take an index $i \in \{1, \dots, L\}$ such that $U \in U_i \mathcal{O}$. Then the box

$$\mathcal{R}' := U_i^{-1} \mathcal{R}$$

satisfies

$$\mathcal{R}' = U_i^{-1} U \mathcal{R}_0, \quad \|U_i^{-1} U - I\|_{\text{op}} < \varepsilon$$

and all of its vertices are in the set

$$U_i^{-1} \mathcal{C}_{l_1, l_2, \dots, l_L} \subseteq U_i^{-1} U_i \mathcal{S}_{l_i} = \mathcal{S}_{l_i}.$$

This contradicts the construction from the previous part of the proof. \square

The number of colors needed in the above proof grows superexponentially in n .

4. PARALLELOGRAMS: PROOF OF THEOREM 6

This will be an easy modification of the proof of Theorem 3. Thus, we work in the complex plane and keep the notation from Section 2.

Proof of Theorem 6. Fix $\varepsilon \in (0, \pi/2)$, take a positive integer N greater than $6/\sin \varepsilon$, and this time define the color-classes as

$$\mathcal{C}'_{j,k} := \left\{ z \in \mathbb{C} : z^2 \in \frac{4}{\sin \varepsilon} \left(\mathbb{Z} + i\mathbb{Z} + \frac{j + ik}{N} + \left[0, \frac{1}{N}\right) + i\left[0, \frac{1}{N}\right) \right) \right\}$$

for integers $0 \leq j, k \leq N-1$. For every monochromatic parallelogram \mathcal{P} coordinatized as in Section 2 (see Figure 4) we conclude

$$uv = \frac{1}{2} \mathcal{J}(\mathcal{P}) \in \frac{2}{\sin \varepsilon} \left(\mathbb{Z} + i\mathbb{Z} + \left(-\frac{2}{N}, \frac{2}{N}\right) + i\left(-\frac{2}{N}, \frac{2}{N}\right) \right)$$

and, in particular,

$$|uv| \notin \left[1, \frac{1}{\sin \varepsilon}\right].$$

If all angles of \mathcal{P} are at least ε , then its area is comparable to the product of its side-lengths. More precisely,

$$|uv| \sin \varepsilon \leq |\mathcal{P}| \leq |uv|$$

and we see that $\text{area}(\mathcal{P})$ cannot equal 1.

Next, we can partition \mathbb{C} into

$$\mathcal{C}''_j := \left\{ z \in \mathbb{C} : \text{Im}(z^2) \in 4 \left(\mathbb{Z} + \frac{j}{5} + \left[0, \frac{1}{5}\right) \right) \right\}$$

for $0 \leq j \leq 4$. For every parallelogram \mathcal{P} that is monochromatic with respect to this coloring we certainly have

$$\begin{aligned} \text{Im}(uv) &= \frac{1}{2} (\text{Im}(z^2) - \text{Im}((z+u)^2) + \text{Im}((z+u+v)^2) - \text{Im}((z+v)^2)) \\ &\in 2\mathbb{Z} + \left(-\frac{4}{5}, \frac{4}{5}\right). \end{aligned}$$

If one side of \mathcal{P} is parallel to the real axis, then $u \in \mathbb{R}$ or $v \in \mathbb{R}$, so the area of \mathcal{P} is

$$|\mathcal{P}| = |\text{Im}(uv)|$$

and we again conclude that this area cannot be equal to 1. This is not surprising because we have in fact been working with the quantity \mathcal{J} from (3.3) in Section 3 and we have just shown that it is invariant, but only for these special parallelograms. Rotate the coloring $\mathcal{C}''_0, \dots, \mathcal{C}''_4$ in a way that it prevents monochromatic area 1 parallelograms from having one side parallel to one of the given lines ℓ_i ; repeat this for $i = 1, 2, \dots, m$.

It remains to take a joint refinement of all colorings of \mathbb{C} constructed in this proof. \square

5. HYPERCUBE GRAPHS: PROOF OF THEOREM 8

Let us conveniently work in the complex plane again. The 1-skeleton of (1.2) is then embedded in \mathbb{C} as in Figure 2 given in the introductory section.

Proof of Theorem 8. By taking $p_k = z$ and $v_{j,k} = u_j$ for each index k in the identity (3.2) in Lemma 3.1 from Section 3, we obtain a simpler (and well-known) polynomial identity

$$\sum_{T \subseteq \{1,2,\dots,n\}} (-1)^{n-|T|} \left(z + \sum_{j \in T} u_j \right)^n = n! u_1 u_2 \cdots u_n \quad (5.1)$$

for $z, u_1, u_2, \dots, u_n \in \mathbb{C}$. Now we partition \mathbb{C} into color-classes defined as

$$\mathcal{C}_{j,k} := \left\{ z \in \mathbb{C} : z^n \in 2n! \left(\mathbb{Z} + i\mathbb{Z} + \frac{j + ik}{2^{n+1}} + \left[0, \frac{1}{2^{n+1}} \right) + i \left[0, \frac{1}{2^{n+1}} \right) \right) \right\}$$

for $0 \leq j, k \leq 2^{n+1} - 1$. If all 2^n points (1.3) belong to the same class $\mathcal{C}_{j,k}$, then (5.1) implies

$$u_1 u_2 \cdots u_n \in 2(\mathbb{Z} + i\mathbb{Z}) + \left(-\frac{1}{2}, \frac{1}{2} \right) + i \left(-\frac{1}{2}, \frac{1}{2} \right).$$

From this it is clear that $a_1 \cdots a_n = |u_1 \cdots u_n|$ cannot be equal to 1. \square

6. PARALLELOTOPES: PROOF OF THEOREM 7

Proof of Theorem 7. By the induction on a positive integer n we will prove that every n -parallelotope with vertices in the set

$$A_{n,\theta} := \{(x_1, x_2, \dots, x_n) \in (0, \infty)^n : x_1 x_2 \cdots x_n \leq \theta\},$$

for some $\theta > 0$, has n -volume strictly smaller than $n!\theta$. See Figure 8 for an illustration of the case $n = 2$. Afterwards, it will remain to choose $\theta = \varepsilon/n!$ and observe that $A_{n,\theta}$ has volume

$$|A_{n,\theta}| = \int_{(0,\infty)^{n-1}} \frac{\theta dx_1 \cdots dx_{n-1}}{x_1 \cdots x_{n-1}} = \infty$$

for every $n \geq 2$.

The induction basis $n = 1$ is trivial since the observed 1-parallelotopes are just segments in the interval $(0, \theta]$, so their length is clearly less than θ . Now take $n \geq 2$ and suppose that the claim holds for the sets $A_{n-1,\vartheta}$ for every $\vartheta > 0$. Let $\theta > 0$ and let \mathcal{P} be an n -parallelotope with vertices in $A_{n,\theta}$. Suppose that p is the vertex of \mathcal{P} with the smallest first coordinate (i.e., coordinate x_1). Also suppose that \mathcal{P} is spanned from p by vectors v_1, \dots, v_n and introduce their coordinates just as in formula (3.5). We have

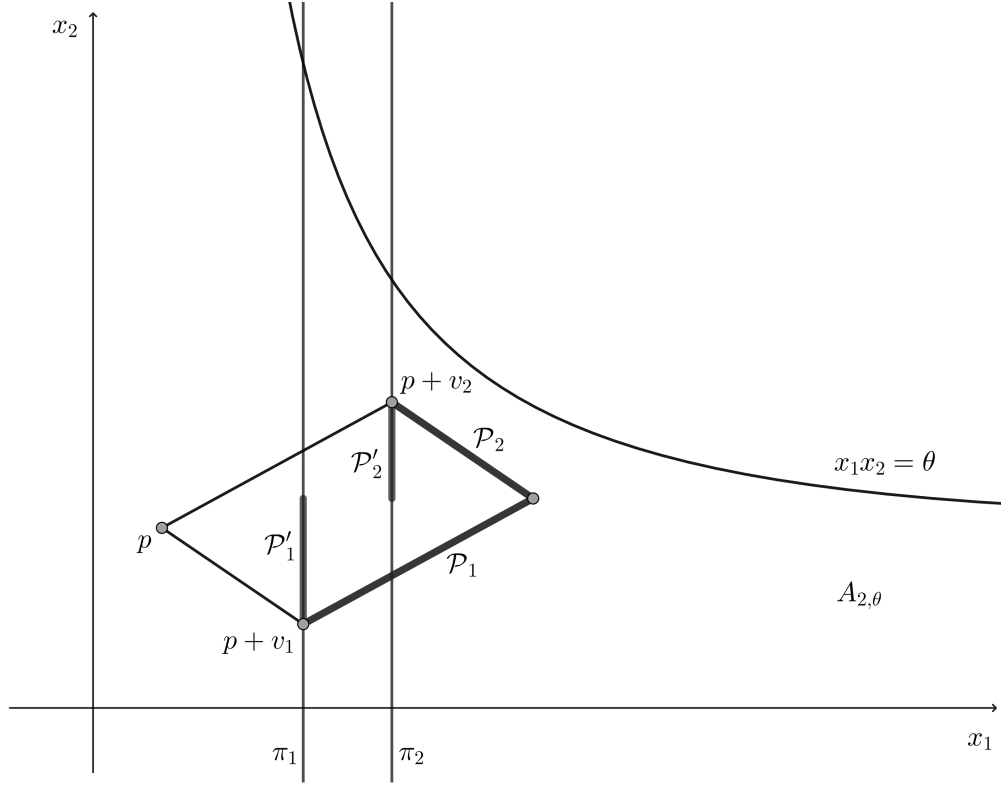
$$v_{j,1} \geq 0 \quad \text{for } 1 \leq j \leq n \quad (6.1)$$

by the choice of the vertex p . Finally, for each $1 \leq i \leq n$, let π_i denote the hyperplane orthogonal to the first coordinate axis and passing through the point $p + v_i$ (i.e., with Cartesian equation $x_1 = p_1 + v_{i,1}$), let \mathcal{P}_i denote the $(n-1)$ -parallelotope spanned from the point $p + v_i$ by the vectors $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$, and let \mathcal{P}'_i be the orthogonal projection of \mathcal{P}_i onto π_i ; see Figure 8. From (6.1) we see that all vertices of \mathcal{P}_i have x_1 -coordinates at least $p_1 + v_{i,1}$, so the vertices of \mathcal{P}'_i lie in the set

$$A_{n,\theta} \cap \pi_i = \{p_1 + v_{i,1}\} \times A_{n-1,\theta/(p_1+v_{i,1})}.$$

By the induction hypothesis applied inside $\pi_i \cong \mathbb{R}^{n-1}$, its volume satisfies

$$|\mathcal{P}'_i| < \frac{(n-1)!\theta}{p_1 + v_{i,1}}. \quad (6.2)$$


 FIGURE 8. A parallelogram with vertices in $A_{2,\theta}$.

Using the Laplace expansion of a determinant (with respect to the first column), (6.1), and (6.2) we can finally estimate the volume of \mathcal{P} as

$$\begin{aligned} |\mathcal{P}| &= |\det(v_{j,k})_{1 \leq j,k \leq n}| \leq \sum_{i=1}^n v_{i,1} \left| \det(v_{j,k})_{1 \leq j \leq n, j \neq i} \right| \\ &= \sum_{i=1}^n v_{i,1} |\mathcal{P}'_i| \leq \sum_{i=1}^n v_{i,1} \frac{(n-1)! \theta}{p_1 + v_{i,1}} < n! \theta, \end{aligned}$$

which concludes the induction step and completes the proof. \square

7. NOTATION AND PRELIMINARIES FROM HARMONIC ANALYSIS

From this section on we introduce the notation

$$\mathcal{A} \lesssim_P \mathcal{B} \quad \text{and} \quad \mathcal{B} \gtrsim_P \mathcal{A},$$

which means that inequality $\mathcal{A} \leq C_P \mathcal{B}$ holds with some constant $C_P \in (0, \infty)$ depending only on a set of parameters P . It is always understood that inequality constants depend on the (configuration or ambient) dimensions, so these are not explicitly listed in the set P . Next,

$$\mathcal{A} \sim_P \mathcal{B}$$

means that both $\mathcal{A} \lesssim_P \mathcal{B}$ and $\mathcal{B} \lesssim_P \mathcal{A}$ hold simultaneously.

When we have a function f of two (or more) variables mapping $(x, y) \mapsto f(x, y)$, then we write $f(\cdot, y)$ for the single-variable function that maps $x \mapsto f(x, y)$ for a fixed y . Similarly, $f(x, \cdot)$ denotes a single-variable function that assigns $y \mapsto f(x, y)$.

The standard Gaussian function on \mathbb{R}^d is

$$g(x) := e^{-\pi|x|^2},$$

its partial derivatives are

$$h^{(l)}(x) := \frac{\partial}{\partial x_l} g(x) = -2\pi x_l e^{-\pi|x|^2}$$

for $1 \leq l \leq d$, while its Laplacian will be denoted as

$$k(x) := \Delta g(x) = (4\pi^2|x|^2 - 2d\pi) e^{-\pi|x|^2}.$$

Dimension d is not visible from the notation $g, h^{(l)}, k$, but it will always be understood from a context.

If the *Fourier transformation* operator $f \mapsto \widehat{f}$ is normalized such that

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

for every Schwartz function f on \mathbb{R}^d , then the Fourier transforms of the above Gaussian functions are given as:

$$\widehat{g}(\xi) = e^{-\pi|\xi|^2}, \quad \widehat{h^{(l)}}(\xi) = 2\pi i \xi_l e^{-\pi|\xi|^2}, \quad \widehat{k}(\xi) = -4\pi^2|\xi|^2 e^{-\pi|\xi|^2}. \quad (7.1)$$

Also the Fourier transform of a finite positive (or even complex) Borel measure on \mathbb{R}^d is defined as

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

The *convolution* of μ and f is a new function $\mu * f$ given by the formula

$$(\mu * f)(x) := \int_{\mathbb{R}^d} f(x - y) d\mu(y)$$

and the property

$$\widehat{\mu * f} = \widehat{\mu} \widehat{f}$$

is well-known to hold. For every $\alpha \in (0, \infty)$ the scaled function f_α and the scaled measure μ_α are defined by

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

and their Fourier transforms are computed as

$$\widehat{f_\alpha}(\xi) = \widehat{f}(\alpha\xi), \quad \widehat{\mu_\alpha}(\xi) = \widehat{\mu}(\alpha\xi).$$

All of these basic properties of Fourier transforms are found in every textbook on real harmonic analysis, such as [52].

Anisotropic scalings are also sometimes useful. We define them for functions f and measures μ on \mathbb{R}^2 for every pair $(\alpha, \beta) \in (0, \infty)^2$ as

$$f_{\alpha,\beta}(x, y) := \alpha^{-1}\beta^{-1} f(\alpha^{-1}x, \beta^{-1}y), \quad \mu_{\alpha,\beta}(E) := \mu(\{(\alpha^{-1}x, \beta^{-1}y) : (x, y) \in E\})$$

and then we have

$$\widehat{f_{\alpha,\beta}}(\zeta, \eta) = \widehat{f}(\alpha\zeta, \beta\eta), \quad \widehat{\mu_{\alpha,\beta}}(\zeta, \eta) = \widehat{\mu}(\alpha\zeta, \beta\eta).$$

Note that all of the functions $f_\alpha, f_{\alpha,\beta}$ have the same L^1 -norm as f , while the measures $\mu_\alpha, \mu_{\alpha,\beta}$ have the same total mass as μ .

Immediate consequences of (7.1) are the Gaussian convolution identities:

$$\sum_{l=1}^d \mathfrak{h}_\alpha^{(l)} * \mathfrak{h}_\beta^{(l)} = \frac{\alpha\beta}{\alpha^2 + \beta^2} \mathbb{k}_{\sqrt{\alpha^2 + \beta^2}}, \quad (7.2)$$

$$\mathbb{k}_\alpha * \mathfrak{g}_\beta = \frac{\alpha^2}{\alpha^2 + \beta^2} \mathbb{k}_{\sqrt{\alpha^2 + \beta^2}} \quad (7.3)$$

for $\alpha, \beta \in (0, \infty)$. Similarly, in two dimensions,

$$\sum_{l=1}^2 \mathfrak{h}_{\alpha,\beta}^{(l)} * \mathfrak{h}_{\alpha,\beta}^{(l)} = \frac{1}{2} \mathbb{k}_{\alpha\sqrt{2}, \beta\sqrt{2}} \quad (7.4)$$

$$\mathbb{k}_{\alpha,\beta} * \mathfrak{g}_{\alpha,\beta} = \frac{1}{2} \mathbb{k}_{\alpha\sqrt{2}, \beta\sqrt{2}} \quad (7.5)$$

for $\alpha, \beta \in (0, \infty)$. If μ is a compactly supported finite positive Borel measure on \mathbb{R}^d , then

$$\mu * \mathfrak{g} \gtrsim_\mu \mathbb{1}_{[-1,1]^d}. \quad (7.6)$$

It is allowed that the implicit constant in this pointwise inequality depends on the choice of the measure μ . In fact, it can be taken to be the minimum of \mathfrak{g} on the set

$$\{x - y : x \in [-1, 1]^d, y \in \text{supp } \mu\}$$

multiplied by the total mass of μ . In particular, this holds when $\mu = \delta_0$ is the Dirac measure at the origin, when it reads $\mathfrak{g} \gtrsim \mathbb{1}_{[-1,1]^d}$, which is quite obvious anyway. Slightly more generally, we also have

$$\mathfrak{g}_\alpha \gtrsim_P \mathbb{1}_{[-1,1]^d}. \quad (7.7)$$

when $\alpha \sim_P 1$. When μ is a smooth truncation of the surface measure of a hyper-surface with non-vanishing Gaussian curvature, then its Fourier transform decays as

$$|\widehat{\mu}(\xi)| \lesssim_\mu (1 + |\xi|)^{-(d-1)/2}; \quad (7.8)$$

see [51, §VIII.3]. In particular, this holds when $\mu = \sigma$ is the spherical measure supported on the standard unit sphere $\mathbb{S}^{d-1} \subseteq \mathbb{R}^d$, also called the circle measure when $d = 2$. Note that an easy consequence of (7.8) and (7.1) is

$$\sup_{\xi \in \mathbb{R}^d} |\widehat{\mu}(\xi) \widehat{\mathbb{k}}(t\xi)| \lesssim_\mu \sup_{\xi \in \mathbb{R}^d} |\xi|^{-1/2} |\widehat{\mathbb{k}}(t\xi)| \sim t^{1/2} \quad (7.9)$$

when $d \geq 2$. The *heat equation* in the present reparametrization of the time t reads

$$\frac{\partial}{\partial t} \mathfrak{g}_t(x) = \frac{1}{2\pi t} \mathbb{k}_t(x) \quad (7.10)$$

for $(t, x) \in (0, \infty) \times \mathbb{R}^d$. It can be verified by direct differentiation.

For a measurable function $f: \mathbb{R}^d \rightarrow \mathbb{C}$ and $p \in [1, \infty)$ we denote (the so-called *Lebesgue seminorm*)

$$\|f\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^p dx \right)^{1/p}.$$

The endpoint case $p = \infty$ is interpreted as the essential supremum of the image of $|f|$,

$$\|f\|_{L^\infty(\mathbb{R}^d)} := \operatorname{ess\,sup}_{x \in \mathbb{R}^d} |f(x)|.$$

The standard L^2 inner product is

$$\langle f, g \rangle_{L^2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} f(x) \overline{g(x)} \, dx.$$

Young's convolution inequality reads

$$\|f * g\|_{L^r(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)} \quad (7.11)$$

when the exponents $p, q, r \in [1, \infty]$ satisfy $1 + 1/r = 1/p + 1/q$; see [52, §V.1]

8. RIGHT SIMPLICES: PROOFS OF THEOREM 1 AND COROLLARY 2

We only need to prove part (a) of Theorem 1. It is sufficient to work in dimension $n = m + 1$, since, in higher dimensions n , Fubini's theorem guaranties that at least one the sections parallel to a fixed $(m + 1)$ -dimensional coordinate plane has density at least that of A . The parameter m is fixed from now on. Take an $R \in (0, \infty)$ and a measurable set $A \subseteq [0, R]^{m+1}$ with density

$$\delta := |A|/R^{m+1} \in (0, 1].$$

Furthermore, denote

$$\theta = m^{-1} 2^{-m^2 - m - 1} \delta^{m+1} \quad (8.1)$$

throughout the proof. For a parameter $\lambda \in (0, \infty)$, interpreted as a certain “scale” that we consider one at a time, we define a *configuration-counting form* as

$$\begin{aligned} \mathcal{N}_\lambda^0(A; R) &:= \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \mathbb{1}_A(x, y) \left(\prod_{k=1}^{m-1} \mathbb{1}_A(x + u_k \mathbf{e}_k, y) \right) \mathbb{1}_A(x, y + v) \\ &\quad d\sigma_{m!|u_1 \cdots u_{m-1}|^{-1}}(v) \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) du R^{-m-1} dy dx, \end{aligned}$$

where σ is the circle measure in \mathbb{R}^2 normalized to have total mass 1 on \mathbb{S}^1 and we write

$$u = (u_1, \dots, u_{m-1}) \in \mathbb{R}^{m-1}$$

for shortness. Observe that $\mathcal{N}_\lambda^0(A; R)$ is a certain density of a subcollection of axes-aligned right-simplices inside A : it detects when all of the points

$$(x, y), (x + u_1 \mathbf{e}_1, y), \dots, (x + u_{m-1} \mathbf{e}_{m-1}, y), (x, y + v) \quad (8.2)$$

belong to A by performing normalized integration over all simplex vertices $(x, y) \in [0, R]^{m+1}$, over $m - 1$ simplex edges parallel to the coordinate axes with lengths $\theta\lambda \leq |u_k| \leq \lambda$, and over the last perpendicular edge parallel to the remaining coordinate 2-plane with its length precisely equal to $|v| = m!|u_1 \cdots u_{m-1}|^{-1}$, so that the m -dimensional simplex volume equals 1. Merely knowing that $\mathcal{N}_\lambda^0(A; R) > 0$ guarantees existence of the desired point configuration inside A . The freedom to choose the parameter λ will allow certain pigeonholing. The idea is to show that \mathcal{N}_λ^0 is positive for at least one out of sufficiently many choices for λ , i.e., from a sufficiently large interval on the logarithmic scale, which is available as soon as R is large enough (in terms of δ).

The main idea, traced back to Bourgain [7] and emphasized by Cook, Magyar, and Pramanik [11], is to consider a *certain smoothed counting form*, which we here define for every $\varepsilon \in (0, 1]$ by convolving the circle measure σ with a scaled Gaussian \mathfrak{g}_ε :

$$\begin{aligned} \mathcal{N}_\lambda^\varepsilon(A; R) &:= \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \mathbb{1}_A(x, y) \left(\prod_{k=1}^{m-1} \mathbb{1}_A(x + u_k \mathfrak{e}_k, y) \right) \mathbb{1}_A(x, y + v) \\ &\quad \times (\sigma * \mathfrak{g}_\varepsilon)_{m!|u_1 \dots u_{m-1}|^{-1}}(v) \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dv du R^{-m-1} dy dx. \end{aligned}$$

It approximates the original counting form in the limit when $\varepsilon \rightarrow 0$, as is shown by the following simple lemma.

Lemma 8.1. *For every $R, \lambda \in (0, \infty)$ and a measurable set $A \subseteq [0, R]^{m+1}$ we have*

$$\lim_{\varepsilon \rightarrow 0+} \mathcal{N}_\lambda^\varepsilon(A; R) = \mathcal{N}_\lambda^0(A; R).$$

Proof of Lemma 8.1. Denoting

$$F(u, v) := \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \mathbb{1}_A(x, y) \left(\prod_{k=1}^{m-1} \mathbb{1}_A(x + u_k \mathfrak{e}_k, y) \right) \mathbb{1}_A(x, y + v) dy dx$$

and expanding out the convolution in the integral

$$\int_{\mathbb{R}^2} F(u, v) (\sigma_{m!|u_1 \dots u_{m-1}|^{-1}} * \mathfrak{g}_{\varepsilon m!|u_1 \dots u_{m-1}|^{-1}})(v) dv,$$

we see that $\mathcal{N}_\lambda^\varepsilon(A; R)$ can be rewritten as

$$\begin{aligned} \mathcal{N}_\lambda^\varepsilon(A; R) &= \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} (F(u, \cdot) * \mathfrak{g}_{\varepsilon m!|u_1 \dots u_{m-1}|^{-1}}) d\sigma_{m!|u_1 \dots u_{m-1}|^{-1}} \\ &\quad \times R^{-m-1} \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) du. \end{aligned}$$

Observe that $F: \mathbb{R}^{m-1} \times \mathbb{R}^2 \rightarrow [0, \infty)$ is bounded and continuous. The dominated convergence theorem, applied first to the inner integral for each fixed u and then to the outer integral in u , proves that $\mathcal{N}_\lambda^\varepsilon(A; R)$ converges as $\varepsilon \rightarrow 0$ to

$$\int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} F(u, \cdot) d\sigma_{m!|u_1 \dots u_{m-1}|^{-1}} R^{-m-1} \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) du,$$

which is precisely $\mathcal{N}_\lambda^0(A; R)$. □

The strategy is to decompose:

$$\mathcal{N}_\lambda^0(A; R) = \underbrace{\mathcal{N}_\lambda^1(A; R)}_{\text{structured part}} + \underbrace{(\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R))}_{\text{error part}} + \underbrace{(\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R))}_{\text{uniform part}}, \quad (8.3)$$

where the names of the *structured*, *error*, and *uniform* parts are borrowed from regularity lemmas in extremal combinatorics, which they mimic. The structured part is a dominant and positive one, the uniform part is always small, while the error part can be made small only after successful pigeoholing in sufficiently many choices of λ ; see a general discussion of the approach in [40, §1.3].

In line with that, as the first ingredient we show that the structured part is bounded from below only in terms of δ , uniformly in the parameter λ from a certain range.

Lemma 8.2. *For every $R, \lambda \in (0, \infty)$ such that $R^{-1/(m-1)} \leq \lambda \leq R$ and a measurable set $A \subseteq [0, R]^{m+1}$ of density $\delta = |A|/R^{m+1}$ we have*

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \delta^{(m+1)(2m-1)}.$$

Note that the implicit constant depends on m alone.

Proof of Lemma 8.2. Recall the bound (7.6) for the convolution $\sigma * g$, which can be scaled to give

$$\begin{aligned} (\sigma * g)_{m!|u_1 \dots u_{m-1}|^{-1}} &\gtrsim |u_1 \dots u_{m-1}|^2 \mathbb{1}_{[-m!|u_1 \dots u_{m-1}|^{-1}, m!|u_1 \dots u_{m-1}|^{-1}]^2} \\ &\geq (\theta\lambda)^{2(m-1)} \mathbb{1}_{[-\lambda^{-m+1}, \lambda^{-m+1}]^2} \end{aligned}$$

for $u_1, \dots, u_{m-1} \in [\theta\lambda, \lambda]$. Consequently,

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \theta^{2(m-1)} R^{-m-1} \lambda^{m-1} \int_{D \setminus E} \Phi(u, v, x, y) \, d(u, v, x, y), \quad (8.4)$$

where

$$\begin{aligned} D := \{ &(u, v, x, y) : |u_k| \leq \lambda \text{ for every } 1 \leq k \leq m-1, \\ &|v_k| \leq \lambda^{-m+1} \text{ for every } 1 \leq k \leq 2, \\ &0 \leq x_k \leq R \text{ for every } 1 \leq k \leq m-1, \\ &0 \leq y_k \leq R \text{ for every } 1 \leq k \leq 2\}, \end{aligned}$$

$$E := \{(u, v, x, y) \in D : |u_k| < \theta\lambda \text{ for some } 1 \leq k \leq m-1\},$$

and we also denoted

$$\Phi(u, v, x, y) := \mathbb{1}_A(x, y) \left(\prod_{k=1}^{m-1} \mathbb{1}_A(x + u_k e_k, y) \right) \mathbb{1}_A(x, y + v).$$

Let \mathcal{I} denote a partition of $[0, \lceil R/\lambda \rceil \lambda] \supseteq [0, R]$ into intervals of length λ and let \mathcal{Q} be a partition of $[0, \lceil R\lambda^{m-1} \rceil \lambda^{-m+1}]^2 \supseteq [0, R]^2$ into squares of side-length λ^{-m+1} . Clearly, by the construction and the assumption $R \geq \max\{\lambda, \lambda^{-m+1}\}$,

$$\text{card } \mathcal{I} = \left\lceil \frac{R}{\lambda} \right\rceil \leq \frac{2R}{\lambda}, \quad \text{card } \mathcal{Q} = \lceil R\lambda^{m-1} \rceil^2 \leq (2R\lambda^{m-1})^2. \quad (8.5)$$

The integral

$$\int_D \Phi(u, v, x, y) \, d(u, v, x, y) \quad (8.6)$$

is greater than or equal to the sum of integrals

$$\begin{aligned} \int_{I_1 \times \dots \times I_{m-1} \times Q} &\left(\prod_{k=1}^{m-1} \int_{I_k} \mathbb{1}_A(x_1, \dots, x_{k-1}, x', x_{k+1}, \dots, x_{m-1}, y) \, dx' \right) \\ &\times \left(\int_Q \mathbb{1}_A(x, y') \, dy \right) \mathbb{1}_A(x, y) \, d(x, y) \end{aligned} \quad (8.7)$$

over all choices of $I_1, \dots, I_{m-1} \in \mathcal{I}$ and $Q \in \mathcal{Q}$. The total number of such choices is, by (8.5),

$$N := (\text{card } \mathcal{I})^{m-1} (\text{card } \mathcal{Q}) \leq (2R)^{m+1} \lambda^{m-1}. \quad (8.8)$$

By Hölder's inequality, the integral (8.7) multiplied by

$$\prod_{k=1}^{m-1} \int_{I_1 \times \dots \times I_{m-1} \times Q} \frac{\mathbb{1}_A(x, y)}{\int_{I_k} \mathbb{1}_A(\dots, x_{k-1}, x', x_{k+1}, \dots, y) dx'} d(x, y) = (\lambda^{m-2}(\lambda^{-m+1})^2)^{m-1}$$

(where the fraction is interpreted as 0 if the denominator equals 0) and

$$\int_{I_1 \times \dots \times I_{m-1} \times Q} \frac{\mathbb{1}_A(x, y)}{\int_Q \mathbb{1}_A(x, y') dy'} d(x, y) = \lambda^{m-1}$$

is at least

$$\left(\int_{I_1 \times \dots \times I_{m-1} \times Q} \mathbb{1}_A(x, y) d(x, y) \right)^{m+1} = |A \cap (I_1 \times \dots \times I_{m-1} \times Q)|^{m+1}.$$

Thus, (8.6) is bounded from below by

$$\begin{aligned} & \lambda^{(m-1)^2} \sum_{\substack{I_1, \dots, I_{m-1} \in \mathcal{I} \\ Q \in \mathcal{Q}}} |A \cap (I_1 \times \dots \times I_{m-1} \times Q)|^{m+1} \\ & \geq \lambda^{(m-1)^2} N \left(\frac{1}{N} \sum_{\substack{I_1, \dots, I_{m-1} \in \mathcal{I} \\ Q \in \mathcal{Q}}} |A \cap (I_1 \times \dots \times I_{m-1} \times Q)| \right)^{m+1} \\ & = \lambda^{(m-1)^2} N^{-m} |A|^{m+1} \\ & \geq (2R)^{-m(m+1)} \lambda^{-m+1} |A|^{m+1}, \end{aligned}$$

by Jensen's inequality and (8.8). Finally, using the assumption $|A| = \delta R^{m+1}$, we get

$$\int_D \Phi \geq 2^{-m(m+1)} \delta^{m+1} R^{m+1} \lambda^{-m+1}. \quad (8.9)$$

Next, the integral

$$\int_E \Phi(u, v, x, y) d(u, v, x, y)$$

needs to be bounded from above. On the part of the set E where $|u_k| < \theta \lambda$ for a fixed given index $1 \leq k \leq m-1$ we simply estimate Φ by 1 and integrate out in all variables, bounding that part of the integral by

$$(\theta \lambda) \lambda^{m-2} (\lambda^{-m+1})^2 R^{m-1} R^2 = \theta \lambda^{-m+1} R^{m+1}.$$

Therefore,

$$\int_E \Phi \leq m \theta R^{m+1} \lambda^{-m+1}. \quad (8.10)$$

From (8.4), (8.9), and (8.10) we conclude

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \theta^{2(m-1)} (2^{-m(m+1)} \delta^{m+1} - m \theta),$$

so choosing θ as in (8.1) we end up with

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \delta^{2(m+1)(m-1)+m+1} = \delta^{(m+1)(2m-1)},$$

as desired. \square

Next, we turn to the uniform part, for which we show that it decays as $\varepsilon \rightarrow 0$ uniformly in λ .

Lemma 8.3. *For every $\lambda, R \in (0, \infty)$, $\varepsilon \in (0, 1]$, and a measurable set $A \subseteq [0, R]^{m+1}$ we have*

$$|\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \lesssim \varepsilon^{1/2}.$$

Proof of Lemma 8.3. In this proof we conveniently denote

$$f(x, y; u) := \mathbb{1}_A(x, y) \prod_{k=1}^{m-1} \mathbb{1}_A(x + u_k \mathbf{e}_k, y)$$

and only remember that $f(\cdot; u)$ is bounded by 1 and supported on $[0, R]^{m+1}$. Also define

$$a = a(u_1, \dots, u_{m-1})$$

to be an abbreviation for

$$a := m! |u_1 \cdots u_{m-1}|^{-1}.$$

For any $0 < \tau < \varepsilon$, by using the heat equation (7.10) in the integral form

$$(\sigma * \mathbf{g}_\tau)(v) - (\sigma * \mathbf{g}_\varepsilon)(v) = - \int_\tau^\varepsilon (\sigma * \mathbf{k}_t)(v) \frac{dt}{2\pi t}, \quad (8.11)$$

we can write

$$\begin{aligned} \mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R) &= - \int_\tau^\varepsilon \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x, y; u) \mathbb{1}_A(x, y + v) R^{-m-1} \lambda^{-m+1} \\ &\quad \times (\sigma * \mathbf{k}_t)_a(v) \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dv dy du dx \frac{dt}{2\pi t}. \end{aligned}$$

If $\widehat{f}(x, \xi; u)$ denotes the Fourier transform of $y \mapsto f(x, y; u)$ evaluated at ξ and if we similarly define $\widehat{\mathbb{1}}_A(x, \xi)$, then we can rewrite

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x, y; u) \mathbb{1}_A(x, y + v) (\sigma * \mathbf{k}_t)_a(v) dv dy$$

using Plancherel's theorem on the frequency side as

$$\int_{\mathbb{R}^2} \widehat{\mathbb{1}}_A(x, \xi) \overline{\widehat{f}(x, \xi; u)} \widehat{\sigma}(a\xi) \widehat{\mathbf{k}}(ta\xi) d\xi.$$

This is, by (7.9), the Cauchy-Schwarz inequality, and another application of Plancherel's theorem, bounded in the absolute value by a constant multiple of

$$\begin{aligned} &t^{1/2} \left(\int_{\mathbb{R}^2} |\widehat{\mathbb{1}}_A(x, \xi)|^2 d\xi \right)^{1/2} \left(\int_{\mathbb{R}^2} |\widehat{f}(x, \xi; u)|^2 d\xi \right)^{1/2} \\ &= t^{1/2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(x, y)^2 dy \right)^{1/2} \left(\int_{\mathbb{R}^2} f(x, y; u)^2 dy \right)^{1/2} \leq t^{1/2} R^2. \end{aligned}$$

Thus,

$$\begin{aligned} |\mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| &\lesssim \int_\tau^\varepsilon \int_{[0, R]^{m-1}} \int_{[-\lambda, \lambda]^{m-1}} t^{1/2} R^{-m+1} \lambda^{-m+1} du dx \frac{dt}{t} \\ &\lesssim \int_\tau^\varepsilon t^{-1/2} dt \lesssim \varepsilon^{1/2}. \end{aligned}$$

The desired estimate is obtained in the limit as $\tau \rightarrow 0$, thanks to Lemma 8.1. \square

The error part is not controlled for a single scale λ , but rather “on the average” over all scales $\lambda = e^\alpha$, $\alpha \in \mathbb{R}$.

Lemma 8.4. *For every $R \in (0, \infty)$, $\varepsilon \in (0, 1]$, and a measurable set $A \subseteq [0, R]^{m+1}$ we have*

$$\int_{\mathbb{R}} (\mathcal{N}_{e^\alpha}^\varepsilon(A; R) - \mathcal{N}_{e^\alpha}^1(A; R))^2 d\alpha \lesssim \theta^{-4(m-1)} \left(\log \frac{1}{\varepsilon} \right)^2.$$

Proof of Lemma 8.4. Using the same notation as in the previous proof and applying (8.11) again, we write

$$\begin{aligned} \mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R) &= - \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x, y; u) \mathbb{1}_A(x, y + v) \\ &\quad R^{-m-1} \lambda^{-m+1} (\sigma * \mathbb{k}_t)_a(v) \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dv dy du dx \frac{ds}{s} \frac{dt}{2\pi t}. \end{aligned}$$

Here we also introduced the new integration variable,

$$s \in [e^{-1}t\lambda^{-m+1}, t\lambda^{-m+1}]. \quad (8.12)$$

The corresponding integral of $1/s$ equals 1, so the above equality is still valid. Whenever $\lambda \in (0, \infty)$, $t \in [\varepsilon, 1]$, s is as in (8.12), and $\theta\lambda \leq |u_k| \leq \lambda$ for $1 \leq k \leq m-1$, then we define

$$r = r(\lambda, t, s, u_1, \dots, u_{m-1}) \in (0, \infty) \quad (8.13)$$

as the unique solution to

$$2s^2 + r^2 = (ta)^2. \quad (8.14)$$

It is well-defined since $ta \geq m!t\lambda^{-m+1} \geq 2s$. Then convolution identities (7.3) and (7.2) yield

$$\mathbb{k}_{ta} = \left(\frac{ta}{s} \right)^2 \sum_{l=1}^2 \mathbb{h}_s^{(l)} * \mathbb{h}_s^{(l)} * \mathbb{g}_r,$$

so

$$\begin{aligned} & - \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x, y; u) \mathbb{1}_A(x, y + v) (\sigma * \mathbb{k}_t)_a(v) dv dy \\ &= \left(\frac{ta}{s} \right)^2 \sum_{l=1}^2 \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} f(x, y; u) \mathbb{h}_s^{(l)}(y - z) dy \right) \\ &\quad \times \left(\int_{\mathbb{R}^2} \mathbb{1}_A(x, y') \mathbb{h}_s^{(l)}(y' - z') dy' \right) (\sigma_a * \mathbb{g}_r)(z - z') dz dz'. \end{aligned}$$

If the numbers t, s, u_k are as before, then $ta/s \lesssim \theta^{-m+1}$, which in turn implies

$$\begin{aligned} |\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R)| &\lesssim (\theta^2 \lambda)^{-m+1} R^{-m-1} \sum_{l=1}^2 \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \\ &\quad \left| \int_{\mathbb{R}^2} f(x, y; u) \mathbb{h}_s^{(l)}(y - z) dy \right| \left| \int_{\mathbb{R}^2} \mathbb{1}_A(x, y') \mathbb{h}_s^{(l)}(y' - z') dy' \right| \\ &\quad \times (\sigma_a * \mathbb{g}_r)(z - z') \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dz dz' du dx \frac{ds}{s} \frac{dt}{t}. \end{aligned}$$

By the Cauchy–Schwarz inequality (in all variables but y and y') we estimate

$$(\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R))^2 \lesssim (\theta^4 \lambda^2)^{-m+1} R^{-2m-2} \mathcal{A} \mathcal{B}, \quad (8.15)$$

where

$$\begin{aligned} \mathcal{A} := & \sum_{l=1}^2 \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{[0,R]^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} f(x, y; u) \mathfrak{h}_s^{(l)}(y - z) \, dy \right)^2 \\ & \times (\sigma_a * \mathfrak{g}_r)(z - z') \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) \, dz \, dz' \, du \, dx \, \frac{ds}{s} \, \frac{dt}{t} \end{aligned}$$

and

$$\begin{aligned} \mathcal{B} := & \sum_{l=1}^2 \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{[0,R]^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(x, y') \mathfrak{h}_s^{(l)}(y' - z') \, dy' \right)^2 \\ & \times (\sigma_a * \mathfrak{g}_r)(z - z') \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) \, dz \, dz' \, du \, dx \, \frac{ds}{s} \, \frac{dt}{t}. \end{aligned}$$

First, we turn our attention to \mathcal{A} . Integrating out in z' , estimating crudely by Young's inequality (7.11),

$$\begin{aligned} & \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} f(x, y; u) \mathfrak{h}_s^{(l)}(y - z) \, dy \right)^2 \, dz \\ & \leq \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x, y; u) f(x, y'; u) (|\mathfrak{h}_s^{(l)}| * |\mathfrak{h}_s^{(l)}|)(y - y') \, dy \, dy' \\ & = \langle f(x, \cdot; u) * |\mathfrak{h}_s^{(l)}| * |\mathfrak{h}_s^{(l)}|, f(x, \cdot; u) \rangle_{L^2(\mathbb{R}^2)} \\ & \leq \|f(x, \cdot; u)\|_{L^2(\mathbb{R}^2)}^2 \|\mathfrak{h}^{(l)}\|_{L^1(\mathbb{R}^2)}^2 \lesssim R^2, \end{aligned}$$

and finally integrating out in u, x, s, t , we bound

$$\mathcal{A} \lesssim R^{m+1} \lambda^{m-1} \log \frac{1}{\varepsilon}.$$

On the other hand, \mathcal{B} is easily controlled by integrating immediately in z and u :

$$\begin{aligned} \mathcal{B} & \lesssim \lambda^{m-1} \sum_{l=1}^2 \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(x, y') \mathfrak{h}_s^{(l)}(y' - z') \, dy' \right)^2 \, dz' \, dx \, \frac{ds}{s} \, \frac{dt}{t} \\ & = -\frac{1}{2} \lambda^{m-1} \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{\mathbb{R}^{m-1}} \langle \mathbb{1}_A(x, \cdot) * \mathbb{K}_{s\sqrt{2}}, \mathbb{1}_A(x, \cdot) \rangle_{L^2(\mathbb{R}^2)} \, dx \, \frac{ds}{s} \, \frac{dt}{t} \\ & = 4\pi^2 \lambda^{m-1} \int_\varepsilon^1 \int_{e^{-1}t\lambda^{-m+1}}^{t\lambda^{-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(x, \xi)|^2 e^{-2\pi s^2 |\xi|^2} |\xi|^2 \, d\xi \, dx \, s \, ds \, \frac{dt}{t}, \end{aligned}$$

where we also used (7.2), Plancherel's theorem, and the third formula from (7.1). Combine the obtained estimates for \mathcal{A} and \mathcal{B} with (8.15), substitute $\lambda = e^\alpha$, and integrate in α to obtain

$$\begin{aligned} \int_{\mathbb{R}} (\mathcal{N}_{e^\alpha}^\varepsilon(A; R) - \mathcal{N}_{e^\alpha}^1(A; R))^2 d\alpha &\lesssim \theta^{-4m+4} R^{-m-1} \left(\log \frac{1}{\varepsilon} \right) \\ &\times \int_{\varepsilon}^1 \int_{\mathbb{R}} \int_{te^{-(m-1)\alpha-1}}^{te^{-(m-1)\alpha}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(x, \xi)|^2 e^{-2\pi s^2 |\xi|^2} |\xi|^2 d\xi dx s ds d\alpha \frac{dt}{t}. \end{aligned}$$

Interchanging the integrals in α and s ,

$$\int_{\mathbb{R}} \int_{te^{-(m-1)\alpha-1}}^{te^{-(m-1)\alpha}} \cdots ds d\alpha = \int_0^\infty \int_{-(\log(t/s)+1)/(m-1)}^{-(\log(t/s))/(m-1)} \cdots d\alpha ds,$$

and then performing the integration respectively in α , s , t , and ξ , we see that the expression from the right hand side of the previous display is actually equal to

$$\frac{1}{4\pi(m-1)} \theta^{-4m+4} R^{-m-1} \left(\log \frac{1}{\varepsilon} \right)^2 \int_{\mathbb{R}^{m-1}} \|\mathbb{1}_A(x, \cdot)\|_{L^2(\mathbb{R}^2)}^2 dx.$$

Note that the integral in the expression above is $|A|$, which is at most R^{m+1} . \square

Proof of Theorem 1. For a given $\delta \in (0, 1]$ we can choose

$$\varepsilon \sim \delta^{2(m+1)(2m-1)},$$

so that Lemmas 8.2 and 8.3 give

$$|\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \leq \frac{1}{3} \mathcal{N}_\lambda^1(A; R)$$

whenever $R \geq \lambda \geq 1$. Next, using Lemma 8.4 we can find a positive number

$$\begin{aligned} J &\sim \theta^{-4m+4} \delta^{-2(m+1)(2m-1)} \left(\log \frac{1}{\varepsilon} \right)^2 \\ &\sim \left(\frac{1}{\delta} \right)^{2(m+1)(4m-3)} \left(1 + \log \frac{1}{\delta} \right)^2 \end{aligned}$$

and then a number $\beta \in [0, J]$ such that

$$|\mathcal{N}_{e^\beta}^\varepsilon(A; R) - \mathcal{N}_{e^\beta}^1(A; R)| \leq \left(\frac{1}{J} \int_0^J (\mathcal{N}_{e^\alpha}^\varepsilon(A; R) - \mathcal{N}_{e^\alpha}^1(A; R))^2 d\alpha \right)^{1/2} \leq \frac{1}{3} \mathcal{N}_{e^\beta}^1(A; R),$$

as soon as R is large enough that we can actually apply Lemma 8.2 for any such $\lambda = e^\beta$, i.e.,

$$R \geq e^J. \tag{8.16}$$

From decomposition (8.3) we finally get

$$\mathcal{N}_{e^\beta}^0(A; R) \geq \frac{1}{3} \mathcal{N}_{e^\beta}^1(A; R) > 0,$$

which guarantees existence in A of the desired configuration (8.2) spanning a simplex of volume 1. Note that (8.16) is in turn guaranteed if, say,

$$R \geq \exp(C\delta^{-9m^2})$$

for a large constant C . The last condition can be rewritten precisely as the assumption (1.1) of Theorem 1(a):

$$\delta \geq \left(\frac{C}{\log R} \right)^{1/(9m^2)}. \quad \square$$

Proof of Corollary 2. Take a set $A \subseteq \mathbb{R}^n$ of positive upper Banach density. Let us first concentrate on the case $V = 1$. Denote

$$R_0 := \exp(2C_m \bar{\delta}_n(A)^{-9m^2}),$$

where C_m is the constant from Theorem 1. If we can find an axes-aligned cube $Q \subset \mathbb{R}^n$ of side-length R_0 on which A has density greater than $(C_m / \log R_0)^{1/(9m^2)}$, then part (a) of that same theorem will apply to this cube Q , only translated from the position $[0, R_0]^n$. It will give us a right m -simplex of volume 1 with vertices in $A \cap Q$. Moreover, the ratio of lengths of any two of its perpendicular edges will be at most

$$\frac{R_0^m}{m!} \leq \exp(2mC_m \bar{\delta}_n(A)^{-9m^2}),$$

as desired.

Thus, suppose that, on the contrary, for every such cube Q of side-length R_0 we have

$$\frac{|A \cap Q|}{|Q|} \leq \left(\frac{C_m}{\log R_0} \right)^{1/(9m^2)}. \quad (8.17)$$

Take some $R \geq R_0$ and an arbitrary $x \in \mathbb{R}^n$. We can easily partition the cube $x + [0, R]^n$ into a collection \mathcal{Q} of disjoint cubes of side-length R_0 and a remainder of volume less than nR_0R^{n-1} . By the assumption (8.17) applied to each cube $Q \in \mathcal{Q}$, we conclude that

$$\frac{|A \cap (x + [0, R]^n)|}{R^n} \leq \left(\frac{C_m}{\log R_0} \right)^{1/(9m^2)} + \frac{nR_0}{R}.$$

Taking the supremum in x and the upper limit in R , we obtain

$$\bar{\delta}_n(A) \leq \left(\frac{C_m}{\log R_0} \right)^{1/(9m^2)},$$

which contradicts the choice of R_0 .

Finally, for an arbitrary $V > 0$ the set $V^{-1/m}A$ clearly has the same upper Banach density $\bar{\delta}_n(A)$, so the previous case applies to it and gives a right simplex of volume 1 with vertices in $V^{-1/m}A$ and with controlled ratios of lengths of its perpendicular edges. It remains to scale the whole picture back by factor $V^{1/m}$ and thus obtain the desired simplex of volume V . \square

9. LARGE RECTANGULAR BOXES: PROOF OF THEOREM 5

Fix $n \geq m + 1$. If the statement in part (a) of the theorem did not hold, then there would exist a sequence of positive numbers $(\lambda_j)_{j=1}^\infty$ converging to ∞ such that for every index j the set A does not contain all vertices of any m -dimensional rectangular box (aligned as claimed in the theorem) in \mathbb{R}^n with volume precisely λ_j^n . Passing to a subsequence we could also ensure $\lambda_{j+1} \geq 2\lambda_j$ for each index j . The usual localization argument will allow us to work on a sufficiently large cube $[0, R]^n$ and with sufficiently (but finitely) many scales λ_j , $1 \leq j \leq J$.

Indeed, leaving the choice of J (depending on $\bar{\delta}_n(A)$) for later, we can find a number R larger than λ_J for which there exists $x \in \mathbb{R}^n$ such that

$$\frac{|A \cap (x + [0, R]^n)|}{R^n} \geq \delta := \frac{\bar{\delta}_n(A)}{2}.$$

The set $A' := (A - x) \cap [0, R]^n$ is then a subset of the cube $[0, R]^n$ with density at least δ and it still does not contain the required boxes of volumes λ_j^m , $1 \leq j \leq J$.

Thus, it is sufficient to work, from the very beginning, with a measurable subset A of a large cube $[0, R]^n$ with density $\delta > 0$. We will obtain a contradiction with the assumption that the m -volumes of the (appropriately aligned) boxes with vertices in A omit the numbers λ_j^m , $1 \leq j \leq J$. Here we work with a large number J that we can choose depending on δ , but the cube side-length $R > 0$ needs to be an arbitrary number larger than some threshold (which in turn is allowed to depend on δ and J).

It will be convenient to define

$$\theta = m^{-1} 2^{-2^m n} \delta^{2^m}. \quad (9.1)$$

Let σ be the spherical measure on \mathbb{R}^{n-m+1} , supported on the standard unit sphere in that space and having the total mass equal to 1. The appropriate pattern-counting form is now, for every $\lambda \in (0, \infty)$,

$$\begin{aligned} \mathcal{N}_\lambda^0(A; R) &:= \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \\ &\quad \left(\prod_{(r_1, \dots, r_m) \in \{0, 1\}^m} \mathbb{1}_A(x_1 + r_1 u_1, \dots, x_{m-1} + r_{m-1} u_{m-1}, y + r_m v) \right) \\ &\quad d\sigma_{\lambda^m |u_1 \dots u_{m-1}|^{-1}}(v) \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) du R^{-n} dy dx, \end{aligned}$$

where it is understood that

$$u = (u_1, \dots, u_{m-1}), \quad x = (x_1, \dots, x_{m-1}).$$

The smoothed counting form is defined, for every $\varepsilon \in (0, 1]$, as

$$\begin{aligned} \mathcal{N}_\lambda^\varepsilon(A; R) &:= \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \\ &\quad \left(\prod_{(r_1, \dots, r_m) \in \{0, 1\}^m} \mathbb{1}_A(x_1 + r_1 u_1, \dots, x_{m-1} + r_{m-1} u_{m-1}, y + r_m v) \right) \\ &\quad (\sigma * \mathbb{G}_\varepsilon)_{\lambda^m |u_1 \dots u_{m-1}|^{-1}}(v) dv \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) du R^{-n} dy dx, \end{aligned}$$

The analogue of Lemma 8.1 holds and it is shown in exactly the same way. We will apply the decomposition (8.3) here as well. The following three lemmas respectively handle the structured, uniform, and error parts.

Lemma 9.1. *For numbers $0 < \lambda \leq R$ and a measurable set $A \subseteq [0, R]^n$ with density $\delta = |A|/R^n$ we have*

$$\mathcal{N}_\lambda^1(A; R) \gtrsim_\delta 1.$$

The lower bound that we prove is in fact a concrete one, involving a certain power of δ , but Theorem 5 is not quantitative anyway.

Proof of Lemma 9.1. For arbitrary $u_1, \dots, u_{m-1} \in [\theta\lambda, \lambda]$ we can estimate the convolution of the measure σ with the Gaussian g using (7.6) as

$$\begin{aligned} (\sigma * g)_{\lambda^m |u_1 \dots u_{m-1}|^{-1}} &\gtrsim (\lambda^{-m} |u_1 \dots u_{m-1}|)^{n-m+1} \mathbb{1}_{[-\lambda^m |u_1 \dots u_{m-1}|^{-1}, \lambda^m |u_1 \dots u_{m-1}|^{-1}]^{n-m+1}} \\ &\geq \theta^{(m-1)(n-m+1)} \lambda^{-n+m-1} \mathbb{1}_{[-\lambda, \lambda]^{n-m+1}}. \end{aligned}$$

Similarly as in the previous section, we bound

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \theta^{(m-1)(n-m+1)} (R\lambda)^{-n} \int_{D \setminus E} \Phi(u, v, x, y) \, d(u, v, x, y), \quad (9.2)$$

where

$$\begin{aligned} D &:= [-\lambda, \lambda]^{m-1} \times [-\lambda, \lambda]^{n-m+1} \times [0, R]^{m-1} \times [0, R]^{n-m+1}, \\ E &:= \{(u, v, x, y) \in D : |u_k| < \theta\lambda \text{ for some } 1 \leq k \leq m-1\}, \\ \Phi(u, v, x, y) &:= \prod_{(r_1, \dots, r_m) \in \{0,1\}^m} \mathbb{1}_A(x_1 + r_1 u_1, \dots, x_{m-1} + r_{m-1} u_{m-1}, y + r_m v). \end{aligned}$$

Split $[0, \lceil R/\lambda \rceil \lambda] \supseteq [0, R]$ into intervals of length λ , denoting the resulting partition by \mathcal{I} , and split $[0, \lceil R/\lambda \rceil \lambda]^{n-m+1} \supseteq [0, R]^{n-m+1}$ into cubes of side-length λ , denoting the obtained partition by \mathcal{Q} . Then

$$\text{card } \mathcal{I} \leq \frac{2R}{\lambda}, \quad \text{card } \mathcal{Q} \leq \left(\frac{2R}{\lambda}\right)^{n-m+1}. \quad (9.3)$$

The integral of Φ over D is at least

$$\begin{aligned} \sum_{\substack{I_1, \dots, I_{m-1} \in \mathcal{I} \\ Q \in \mathcal{Q}}} \int_{I_1} \int_{I_1} \cdots \int_{I_{m-1}} \int_{I_{m-1}} \int_Q \int_Q \prod_{(r_1, \dots, r_m) \in \{0,1\}^m} \mathbb{1}_A(x_1^{r_1}, \dots, x_{m-1}^{r_{m-1}}, y^{r_m}) \\ dy^0 dy^1 dx_{m-1}^0 dx_{m-1}^1 \cdots dx_1^0 dx_1^1, \end{aligned}$$

(Note that here, and a few times later, zeros and ones in the exponents are simply upper indices.) This is, by the Cauchy–Schwarz inequality for the Gowers-box inner products [49, 36, 55] (or simply by repeated applications of the ordinary Cauchy–Schwarz inequality as in [20, §7.1]), followed by Jensen’s inequality for discrete sums and by (9.3), greater than or equal to

$$\begin{aligned} &\sum_{\substack{I_1, \dots, I_{m-1} \in \mathcal{I} \\ Q \in \mathcal{Q}}} (|I_1| \cdots |I_{m-1}| |Q|)^{2-2^m} \left(\int_{I_1} \cdots \int_{I_{m-1}} \int_Q \mathbb{1}_A(x_1, \dots, x_{m-1}, y) \, dy \, dx_{m-1} \cdots dx_1 \right)^{2^m} \\ &= \sum_{\substack{I_1, \dots, I_{m-1} \in \mathcal{I} \\ Q \in \mathcal{Q}}} \lambda^{(2-2^m)n} |A \cap (I_1 \times \cdots \times I_{m-1} \times Q)|^{2^m} \\ &\geq ((\text{card } \mathcal{I})^{m-1} (\text{card } \mathcal{Q}))^{1-2^m} \lambda^{(2-2^m)n} |A|^{2^m} \geq 2^{(1-2^m)n} (R\lambda)^n \delta^{2^m}. \end{aligned}$$

The integral of Φ over E is at most

$$|E| \leq (m-1)\theta(R\lambda)^n.$$

Thus, by (9.2),

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \theta^{(m-1)(n-m+1)} (2^{(1-2^m)n} \delta^{2^m} - m\theta).$$

The choice (9.1) for θ is a good one and it gives

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \delta^{2^m((m-1)(n-m+1)+1)} \gtrsim_\delta 1. \quad \square$$

Lemma 9.2. *For every $\lambda, R \in (0, \infty)$, every $\varepsilon \in (0, 1]$, and every measurable set $A \subseteq [0, R]^n$ we have*

$$|\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \lesssim \varepsilon^{1/2}.$$

Proof of Lemma 9.2. The proof is almost the same as that of Lemma 8.3. Here we find convenient to introduce the function

$$f(x, y; u) := \prod_{(r_1, \dots, r_{m-1}) \in \{0, 1\}^{m-1}} \mathbb{1}_A(x_1 + r_1 u_1, \dots, x_{m-1} + r_{m-1} u_{m-1}, y)$$

and the number

$$a = a(\lambda, u_1, \dots, u_{m-1})$$

defined as

$$a := \lambda^m |u_1 \cdots u_{m-1}|^{-1}.$$

For any $0 < \tau < \varepsilon$ integrate the heat equation (7.10) as in (8.11) to obtain

$$\begin{aligned} \mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R) &= - \int_\tau^\varepsilon \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} f(x, y; u) f(x, y + v; u) \\ &\quad \times R^{-n} \lambda^{-m+1} (\sigma * \mathbb{k}_t)_a(v) \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dv dy du dx \frac{dt}{2\pi t}. \end{aligned}$$

Write

$$\int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} f(x, y; u) f(x, y + v; u) (\sigma * \mathbb{k}_t)_a(v) dv dy$$

as

$$\int_{\mathbb{R}^{n-m+1}} |\widehat{f}(x, \xi; u)|^2 \widehat{\sigma}(a\xi) \widehat{\mathbb{k}}(ta\xi) d\xi,$$

the absolute value of which is bounded, thanks to (7.9), by a constant times

$$t^{1/2} \int_{\mathbb{R}^{n-m+1}} |\widehat{f}(x, \xi; u)|^2 d\xi = t^{1/2} \int_{\mathbb{R}^{n-m+1}} f(x, y; u)^2 dy \leq t^{1/2} R^{n-m+1}.$$

Thus,

$$\begin{aligned} |\mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| &\lesssim \int_\tau^\varepsilon \int_{[0, R]^{m-1}} \int_{[-\lambda, \lambda]^{m-1}} t^{1/2} R^{-m+1} \lambda^{-m+1} du dx \frac{dt}{t} \\ &\lesssim \int_\tau^\varepsilon t^{-1/2} dt \lesssim \varepsilon^{1/2}. \end{aligned}$$

Letting $\tau \rightarrow 0$ we prove the lemma. \square

Lemma 9.3. *For every $R \in (0, \infty)$, every $\varepsilon \in (0, 1]$, every choice of positive numbers $(\lambda_j)_{1 \leq j \leq J}$ satisfying $\lambda_{j+1} \geq 2\lambda_j$, and every measurable set $A \subseteq [0, R]^n$ we have*

$$\sum_{j=1}^J |\mathcal{N}_{\lambda_j}^\varepsilon(A; R) - \mathcal{N}_{\lambda_j}^1(A; R)| \lesssim_{\delta, \varepsilon} 1.$$

Proof of Lemma 9.3. Let f and a be defined as in the proof of Lemma 9.2. In the same way we arrive at

$$\begin{aligned} & \mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R) \\ &= - \int_\varepsilon^1 \int_{t\lambda/(10e)}^{t\lambda/10} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} f(x, y; u) f(x, y + v; u) \\ & \quad \times R^{-n} \lambda^{-m+1} (\sigma * \mathbb{k}_t)_a(v) \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dv dy du dx \frac{ds}{s} \frac{dt}{2\pi t}, \end{aligned}$$

noting that here we have also introduced the integral of $1/s$ in the variable

$$s \in [10^{-1}e^{-1}t\lambda, 10^{-1}t\lambda],$$

which equals 1, so it does not change the value of the whole expression. The shorthand notation r has the same meaning as in (8.13) and (8.14), except that s is now taken from a different interval. Once again, convolution identities (7.3) and (7.2) give

$$\mathbb{k}_{ta} = \left(\frac{ta}{s}\right)^2 \sum_{l=1}^{n-m+1} \mathbb{h}_s^{(l)} * \mathbb{h}_s^{(l)} * \mathbb{g}_r,$$

so

$$- \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} f(x, y; u) f(x, y + v; u) (\sigma * \mathbb{k}_t)_a(v) dv dy$$

can be rewritten as

$$\begin{aligned} & \left(\frac{ta}{s}\right)^2 \sum_{l=1}^{n-m+1} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} \left(\int_{\mathbb{R}^{n-m+1}} f(x, y; u) \mathbb{h}_s^{(l)}(y - z) dy \right) \\ & \quad \times \left(\int_{\mathbb{R}^{n-m+1}} f(x, y'; u) \mathbb{h}_s^{(l)}(y' - z') dy' \right) (\sigma_a * \mathbb{g}_r)(z - z') dz dz'. \end{aligned}$$

Under the current assumption on the numbers u_k we have $\lambda \leq a \leq \theta^{-m+1}\lambda$, which implies $s \sim t\lambda \sim_\delta ta$ and, taking absolute values, we bound the difference of the counting forms as

$$\begin{aligned} & |\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R)| \lesssim_\delta R^{-n} \sum_{l=1}^{n-m+1} \int_\varepsilon^1 \int_{t\lambda/(10e)}^{t\lambda/10} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} \\ & \quad \left| \int_{\mathbb{R}^{n-m+1}} f(x, y; u) \mathbb{h}_s^{(l)}(y - z) dy \right| \left| \int_{\mathbb{R}^{n-m+1}} f(x, y'; u) \mathbb{h}_s^{(l)}(y' - z') dy' \right| \\ & \quad \times (\sigma_a * \mathbb{g}_r)(z - z') \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dz dz' du dx \frac{ds}{s} \frac{dt}{t}. \end{aligned}$$

By the Cauchy–Schwarz inequality in t, s, x, u, z, z' and the symmetry between z and z' , we estimate this by

$$\begin{aligned} & R^{-n} \sum_{l=1}^{n-m+1} \int_\varepsilon^1 \int_{t\lambda/(10e)}^{t\lambda/10} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \\ & \quad \left(\int_{\mathbb{R}^{n-m+1}} f(x, y; u) \mathbb{h}_s^{(l)}(y - z) dy \right)^2 \lambda^{-m+1} \left(\prod_{k=1}^{m-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \right) dz du dx \frac{ds}{s} \frac{dt}{t}, \end{aligned}$$

where we have already eliminated $(\sigma_a * g_r)(z - z')$ by evaluating the integral in z' . Dominate $\mathbb{1}_{[-1,1]}$ by the Gaussian $g_{s/\lambda}$ using (7.7) and recalling $s \sim_\varepsilon \lambda$; then dilate by λ to obtain

$$\lambda^{-1} \mathbb{1}_{[-\lambda, -\theta\lambda] \cup [\theta\lambda, \lambda]}(u_k) \lesssim_\varepsilon g_s(u_k).$$

Then we expand out the square $(\int_{\mathbb{R}^{n-m+1}} \cdots dy)^2$ and collapse back the convolution using (7.2):

$$\sum_{l=1}^{n-m+1} h_s^{(l)} * h_s^{(l)} = \frac{1}{2} \mathbb{k}_{s\sqrt{2}}.$$

That way we finally end up with

$$\begin{aligned} |\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R)| &\lesssim_{\delta, \varepsilon} R^{-n} \int_{\varepsilon\lambda/(10e)}^{\lambda/10} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{m-1}} \int_{\mathbb{R}^{n-m+1}} \int_{\mathbb{R}^{n-m+1}} \\ &\quad f(x, y^0; u) f(x, y^1; u) \mathbb{k}_{s\sqrt{2}}(y^0 - y^1) \left(\prod_{k=1}^{m-1} g_s(u_k) \right) dy^0 dy^1 du dx \frac{ds}{s}. \end{aligned}$$

Due to our assumption on the exponential growth of the numbers λ_j , every positive number s lies in a bounded number of intervals $[\varepsilon\lambda_j/(10e), \lambda_j/10]$, $1 \leq j \leq J$, with a bound depending on ε alone. Therefore, applying the last estimate with $\lambda = \lambda_j$ and summing in j , we get

$$\sum_{j=1}^J |\mathcal{N}_{\lambda_j}^\varepsilon(A; R) - \mathcal{N}_{\lambda_j}^1(A; R)| \lesssim_{\delta, \varepsilon} R^{-n} \Theta_{1, \dots, 1, \sqrt{2}}^{(m)}(\mathbb{1}_A).$$

Here we define, more generally,

$$\begin{aligned} \Theta_{\gamma_1, \dots, \gamma_m}^{(k)}(F) &:= - \int_0^\infty \int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_m} \times \mathbb{R}^{d_m}} \left(\prod_{(r_1, \dots, r_m) \in \{0,1\}^m} F(x_1^{r_1}, \dots, x_m^{r_m}) \right) \mathbb{k}_{s\gamma_k}(x_k^0 - x_k^1) \\ &\quad \times \left(\prod_{\substack{1 \leq i \leq m \\ i \neq k}} g_{s\gamma_i}(x_i^0 - x_i^1) \right) d(x_1^0, x_1^1, x_2^0, x_2^1, \dots, x_m^0, x_m^1) \frac{ds}{s} \end{aligned}$$

for any $1 \leq k \leq m$, any parameters $\gamma_1, \dots, \gamma_m \in (0, \infty)$, and a real-valued bounded and compactly supported measurable function F on $\mathbb{R}^{d_1+d_2+\dots+d_m}$. In our case, $d_1 = \dots = d_{m-1} = 1$ and $d_m = n - m + 1$, but the greater generally will only make the arguments more elegant and symmetric.

The proof of Lemma 9.3 will be complete once we know the estimate

$$\Theta_{\gamma_1, \dots, \gamma_m}^{(k)}(F) \leq 2\pi \|F\|_{L^{2^m}(\mathbb{R}^d)}^{2^m} \quad (9.4)$$

and apply it to the function $F = \mathbb{1}_A$, for which

$$\|F\|_{L^{2^m}(\mathbb{R}^n)}^{2^m} = |A| \leq R^n.$$

Indeed, $\Theta_{\gamma_1, \dots, \gamma_m}^{(k)}$ is an example of an *entangled singular integral form* [38, 18, 19] and inequality (9.4) is, up to an unimportant constant, a particular case of the recently studied *singular Brascamp–Lieb estimates* [22, 23, 21]. However, in just a few more lines, we can give a self-contained simple proof of this very special case. It has already appeared in [20, §7.2] (based on the earlier ideas from [39, 18]), but we include it here for completeness.

First, observe that

$$\Theta_{\gamma_1, \dots, \gamma_m}^{(k)}(F) \geq 0$$

for each $1 \leq k \leq m$ and every real-valued function F . This is seen by denoting

$$\begin{aligned} f_k(x) &= f_k(x; x_1^0, x_1^1, \dots, x_{k-1}^0, x_{k-1}^1, x_{k+1}^0, x_{k+1}^1, \dots, x_m^0, x_m^1) \\ &:= \prod_{(r_1, \dots, r_{k-1}, r_{k+1}, \dots, r_m) \in \{0,1\}^{m-1}} F(x_1^{r_1}, \dots, x_{k-1}^{r_{k-1}}, x, x_{k+1}^{r_{k+1}}, \dots, x_m^{r_m}) \end{aligned}$$

and using (7.2) again, this time to rewrite

$$\begin{aligned} & - \int_{\mathbb{R}^{d_k}} \int_{\mathbb{R}^{d_k}} f_k(x_k^0) f_k(x_k^1) \mathbb{K}_{s\gamma_k}(x_k^0 - x_k^1) dx_k^0 dx_k^1 \\ &= 2 \sum_{l=1}^{d_k} \int_{\mathbb{R}^{d_k}} \left(\int_{\mathbb{R}^{d_k}} f_k(x) \mathbb{h}_{s\gamma_k/\sqrt{2}}^{(l)}(x - z) dx \right)^2 dz \geq 0. \end{aligned}$$

Essentially, we were merely undoing the last step of the computation that lead us to the particular case $\Theta_{1, \dots, 1, \sqrt{2}}^{(m)}$ in the first place. Next, using the product rule for differentiation and the heat equation (7.10), we obtain

$$\frac{\partial}{\partial s} \prod_{i=1}^m \mathbb{G}_{s\gamma_i}(u_i) = \sum_{k=1}^m \frac{1}{2\pi s} \mathbb{K}_{s\gamma_k}(u_k) \prod_{\substack{1 \leq i \leq m \\ i \neq k}} \mathbb{G}_{s\gamma_i}(u_i), \quad (9.5)$$

from which, integrating over $s \in (0, \infty)$, we easily get

$$\sum_{k=1}^m \Theta_{\gamma_1, \dots, \gamma_m}^{(k)}(F) = 2\pi \|F\|_{L^{2m}(\mathbb{R}^d)}^{2m}.$$

Since the above summands are nonnegative, each of them is dominated by the whole sum, so (9.4) follows. \square

Completion of the proof of Theorem 5. The standing assumption is

$$\mathcal{N}_{\lambda_j}^0(A; R) = 0$$

for every index $1 \leq j \leq J$. By Lemmas 9.1–9.3 we can first choose a sufficiently small $\varepsilon > 0$ (depending on δ) such that

$$|\mathcal{N}_{\lambda}^0(A; R) - \mathcal{N}_{\lambda}^{\varepsilon}(A; R)| \leq \frac{1}{3} \mathcal{N}_{\lambda}^1(A; R)$$

for every $0 < \lambda \leq R$, and then a sufficiently large J (depending on δ and ε) that the pigeonhole principle gives an index $1 \leq j \leq J$ such that

$$|\mathcal{N}_{\lambda_j}^{\varepsilon}(A; R) - \mathcal{N}_{\lambda_j}^1(A; R)| \leq \frac{1}{3} \mathcal{N}_{\lambda_j}^1(A; R).$$

Note that we were able to apply Lemma 9.1 as soon as $R \geq \lambda_j$. Then decomposition (8.3) gives

$$0 = \mathcal{N}_{\lambda_j}^0(A; R) \geq \frac{1}{3} \mathcal{N}_{\lambda_j}^1(A; R) > 0$$

for that same index j , which is a contradiction. \square

10. HYPERBOLIC EMBEDDINGS: PROOF OF THEOREM 10

Rotating the plane by 45° and stretching it by the factor $\sqrt{2}$ we can, more conveniently, replace the hyperbola \mathcal{H} with the rotated hyperbola

$$\mathcal{H}' : \quad xy = 1.$$

This will only affect the constant C from the theorem statement. Let μ be a Borel measure on \mathbb{R}^2 supported on \mathcal{H}' , defined via

$$\int_{\mathbb{R}^2} f(x, y) d\mu(x, y) = \int_0^\infty f\left(x, \frac{1}{x}\right) \varphi(x) dx,$$

where $\varphi: (0, \infty) \rightarrow [0, \infty)$ is a compactly supported C^∞ function. We also adjust φ in a way that μ has total mass 1. Constants in the proof (appearing implicitly in the notation \lesssim and \gtrsim) are allowed to depend on the exact choice of φ , which will not be notationally emphasized. They will also depend on the number n and on the edge-lengths $a_1, \dots, a_n \in (0, \infty)$, which are all regarded as fixed.

The pattern-counting form is, for $\lambda \in (0, \infty)$,

$$\mathcal{N}_\lambda^0(A; R) := R^{-2} \int_{\mathbb{R}^2} \int_{(\mathbb{R}^2)^n} \mathbb{1}_A(z) \left(\prod_{k=1}^n \mathbb{1}_A(z + u_k) \right) \left(\prod_{k=1}^n d\mu_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) dz$$

and its smoothed counterpart is, after also introducing $\varepsilon \in (0, 1]$,

$$\mathcal{N}_\lambda^\varepsilon(A; R) := R^{-2} \int_{\mathbb{R}^2} \int_{(\mathbb{R}^2)^n} \mathbb{1}_A(z) \left(\prod_{k=1}^n \mathbb{1}_A(z + u_k) \right) \left(\prod_{k=1}^n (\mu * \mathbb{g}_\varepsilon)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) du dz.$$

Here u is a shorter notation for the n -tuple (u_1, \dots, u_n) . As expected, $\mathcal{N}_\lambda^0(A; R)$ can again be obtained in the limit of $\mathcal{N}_\lambda^\varepsilon(A; R)$ as $\varepsilon \rightarrow 0$, which is shown easily, as in Lemma 8.1. For every $a, \lambda > 0$ the support of the anisotropically dilated measure $\mu_{a\lambda, a\lambda^{-1}}$ lies on the hyperbola \mathcal{H}' stretched by the factor a . Thus, $\mathcal{N}_\lambda^0(A; R) > 0$ for any value of $\lambda > 0$ clearly means that there exists the desired configuration (1.5) inside the set A .

Decomposition (8.3) will be applied one more time, so we proceed with bounding its three parts.

Lemma 10.1. *For every $\lambda, R \in (0, \infty)$ such that $R^{-1} \leq \lambda \leq R$, and a measurable set $A \subseteq [0, R]^2$ of density $\delta = |A|/R^2$ we have*

$$\mathcal{N}_\lambda^1(A; R) \gtrsim \delta^{n+1}.$$

Proof of Lemma 10.1. We use a slight modification of (7.6) and scale it anisotropically to obtain

$$(\mu_{a_k} * \mathbb{g}_{a_k})_{\lambda, \lambda^{-1}} \gtrsim \mathbb{1}_{[-\lambda, \lambda] \times [-\lambda^{-1}, \lambda^{-1}]}.$$

Let \mathcal{Q} be the partition of

$$[0, \lceil R/\lambda \rceil \lambda] \times [0, \lceil R\lambda \rceil \lambda^{-1}] \supseteq [0, R]^2$$

into $\lceil R/\lambda \rceil \lceil R\lambda \rceil \leq 4R^2$ rectangles of size $\lambda \times \lambda^{-1}$. Then

$$\begin{aligned} \mathcal{N}_\lambda^1(A; R) &\gtrsim R^{-2} \sum_{Q \in \mathcal{Q}} \int_{Q^{n+1}} \left(\prod_{k=0}^n \mathbb{1}_A(z_k) \right) d(z_k)_{0 \leq k \leq n} = R^{-2} \sum_{Q \in \mathcal{Q}} |A \cap Q|^{n+1} \\ &\geq R^{-2} (\text{card } \mathcal{Q})^{-n} \left(\sum_{Q \in \mathcal{Q}} |A \cap Q| \right)^{n+1} \geq R^{-2} (4R^2)^{-n} |A|^{n+1} \sim \delta^{n+1}. \quad \square \end{aligned}$$

Lemma 10.2. *For every $\lambda, R \in (0, \infty)$, $\varepsilon \in (0, 1]$, and a measurable set $A \subseteq [0, R]^2$ of density δ we have*

$$|\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \lesssim \delta \varepsilon^{1/2}.$$

Proof of Lemma 10.2. Using the fundamental theorem of calculus, the heat equation (7.10), and the product rule for differentiation, similarly as in (9.5), we can write, for $0 < \tau < \varepsilon$,

$$\begin{aligned} \mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R) &= -R^{-2} \sum_{j=1}^n \int_\tau^\varepsilon \int_{\mathbb{R}^2} \int_{(\mathbb{R}^2)^n} \mathbb{1}_A(z) \left(\prod_{k=1}^n \mathbb{1}_A(z + u_k) \right) \\ &\quad \times (\mu * \mathbb{k}_t)_{a_j \lambda, a_j \lambda^{-1}}(u_j) \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} (\mu * \mathbb{g}_t)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) du dz \frac{dt}{2\pi t}. \end{aligned}$$

A standard computation using Plancherel's theorem rewrites

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \mathbb{1}_A(z) \mathbb{1}_A(z + u_j) (\mu * \mathbb{k}_t)_{a_j \lambda, a_j \lambda^{-1}}(u_j) du_j dz$$

as

$$\int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\zeta, \eta)|^2 \widehat{\mu}(a_j \lambda \zeta, a_j \lambda^{-1} \eta) \widehat{\mathbb{k}}(ta_j \lambda \zeta, ta_j \lambda^{-1} \eta) d(\zeta, \eta).$$

By (7.9), the absolute value of the last integral is (up to a constant) at most

$$t^{1/2} \|\widehat{\mathbb{1}_A}\|_{L^2(\mathbb{R}^2)}^2 = t^{1/2} \|\mathbb{1}_A\|_{L^2(\mathbb{R}^2)}^2 = t^{1/2} |A|.$$

That way we end up with

$$\begin{aligned} &|\mathcal{N}_\lambda^\tau(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \\ &\lesssim R^{-2} |A| \sum_{j=1}^n \int_\tau^\varepsilon \int_{(\mathbb{R}^2)^{n-1}} \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} (\mu * \mathbb{g}_t)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) d(u_k)_{k \neq j} \frac{dt}{t^{1/2}} \\ &\lesssim R^{-2} |A| \varepsilon^{1/2} = \delta \varepsilon^{1/2}, \end{aligned}$$

where we first integrated freely in the variables u_k and then simply evaluated the integral in t . It remains to let $\tau \rightarrow 0$. \square

Before we finally turn to the error part, let us prove a very concrete estimate for the Fourier transform $\widehat{\mathbb{k}}$ of the Laplacian of the two-dimensional Gaussian. Despite being easy, in this property the hyperbolic scaling from the problem formulation is finally used.

Lemma 10.3. *The estimate*

$$\int_0^\infty \left| \widehat{\mathbb{k}}\left(s\zeta, \frac{\eta}{s}\right) \right| \frac{ds}{s} \lesssim 1$$

holds for every $(\zeta, \eta) \in \mathbb{R}^2$.

Proof of Lemma 10.3. From the last formula in (7.1) we see that the given integral equals

$$\mathcal{K}(\zeta, \eta) := 4\pi^2 \int_0^\infty (s^2 \zeta^2 + s^{-2} \eta^2) e^{-\pi(s^2 \zeta^2 + s^{-2} \eta^2)} \frac{ds}{s}.$$

In the case $\zeta \eta \neq 0$, substitution $t = s\sqrt{|\zeta|/|\eta|}$ gives

$$\begin{aligned} \mathcal{K}(\zeta, \eta) &= 4\pi^2 |\zeta \eta| \int_0^\infty (t^2 + t^{-2}) e^{-\pi|\zeta \eta|(t^2 + t^{-2})} \frac{dt}{t} \\ &\leq 8\pi^2 |\zeta \eta| \left(\int_1^\infty t^2 e^{-\pi|\zeta \eta|t^2} \frac{dt}{t} + \int_0^1 t^{-2} e^{-\pi|\zeta \eta|t^{-2}} \frac{dt}{t} \right) \\ &\leq 16\pi^2 |\zeta \eta| \int_0^\infty t e^{-\pi|\zeta \eta|t^2} dt = 8\pi. \end{aligned}$$

An even easier computation yields $\mathcal{K}(\zeta, 0) = 2\pi$ for $\zeta \neq 0$, $\mathcal{K}(0, \eta) = 2\pi$ for $\eta \neq 0$, and $\mathcal{K}(0, 0) = 0$. \square

We will appreciate Lemma 10.3 more after the discussion in Section 11 on complications that the hyperbolic setting brings to some related analytical problems. Now we can control the remaining term from the decomposition (8.3).

Lemma 10.4. *For every $R \in (0, \infty)$, $\varepsilon \in (0, 1]$, and a measurable set $A \subseteq [0, R]^2$ of density δ we have*

$$\int_{\mathbb{R}} (\mathcal{N}_{e^\alpha}^\varepsilon(A; R) - \mathcal{N}_{e^\alpha}^1(A; R))^2 d\alpha \lesssim \delta^2 \left(\log \frac{1}{\varepsilon} \right)^2.$$

Proof of Lemma 10.4. Using the computation from the previous proof, we rewrite $\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R)$ as

$$\begin{aligned} &-R^{-2} \sum_{j=1}^n \int_\varepsilon^1 \int_{\mathbb{R}^2} \int_{(\mathbb{R}^2)^n} \mathbb{1}_A(z) \left(\prod_{k=1}^n \mathbb{1}_A(z + u_k) \right) \\ &\quad \times (\mu * \mathbb{k}_t)_{a_j \lambda, a_j \lambda^{-1}}(u_j) \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} (\mu * \mathbb{g}_t)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) du dz \frac{dt}{2\pi t}. \end{aligned}$$

For each fixed $1 \leq j \leq n$ we introduce the variable $z' = z + u_j$, expand applying (7.5) and (7.4),

$$\begin{aligned} -(\mu * \mathbb{k}_t)_{a_j \lambda, a_j \lambda^{-1}}(u_j) &= -(\mu * \mathbb{k}_t)_{a_j \lambda, a_j \lambda^{-1}}(z' - z) \\ &= 4 \sum_{l=1}^2 \int_{\mathbb{R}^2} \mathbb{h}_{a_j t \lambda/2, a_j t \lambda^{-1}/2}^{(l)}(z - w) \mathbb{h}_{a_j t \lambda/2, a_j t \lambda^{-1}/2}^{(l)}(z' - w') \\ &\quad \times (\mu * \mathbb{g}_{t/\sqrt{2}})_{a_j \lambda, a_j \lambda^{-1}}(w - w') dw dw', \end{aligned}$$

and use the Cauchy–Schwarz inequality to obtain

$$(\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R))^2 \lesssim R^{-4} \mathcal{A} \mathcal{B}, \quad (10.1)$$

where

$$\begin{aligned} \mathcal{A} := & \sum_{j=1}^n \sum_{l=1}^2 \int_{\varepsilon}^1 \int_{(\mathbb{R}^2)^{n-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(z) \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} \mathbb{1}_A(z + u_k) \right) \mathbb{h}_{a_j t \lambda / 2, a_j t \lambda^{-1} / 2}^{(l)}(z - w) dz \right)^2 \\ & \times (\mu * \mathfrak{g}_{t/\sqrt{2}})_{a_j \lambda, a_j \lambda^{-1}}(w - w') \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} (\mu * \mathfrak{g}_t)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) dw dw' d(u_k)_{k \neq j} \frac{dt}{t} \end{aligned}$$

and

$$\begin{aligned} \mathcal{B} := & \sum_{j=1}^n \sum_{l=1}^2 \int_{\varepsilon}^1 \int_{(\mathbb{R}^2)^{n-1}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(z') \mathbb{h}_{a_j t \lambda / 2, a_j t \lambda^{-1} / 2}^{(l)}(z' - w') dz' \right)^2 \\ & \times (\mu * \mathfrak{g}_{t/\sqrt{2}})_{a_j \lambda, a_j \lambda^{-1}}(w - w') \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} (\mu * \mathfrak{g}_t)_{a_k \lambda, a_k \lambda^{-1}}(u_k) \right) dw dw' d(u_k)_{k \neq j} \frac{dt}{t}. \end{aligned}$$

We are controlling \mathcal{A} rather crudely. First, we eliminate the convolution $\mu * \mathfrak{g}_{t/\sqrt{2}}$ by integrating out w' . Then we bound

$$\begin{aligned} & \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} \mathbb{1}_A(z) \left(\prod_{\substack{1 \leq k \leq n \\ k \neq j}} \mathbb{1}_A(z + u_k) \right) \mathbb{h}_{a_j t \lambda / 2, a_j t \lambda^{-1} / 2}^{(l)}(z - w) dz \right)^2 dw \\ & \leq \|\mathbb{1}_A * \mathbb{h}_{a_j t \lambda / 2, a_j t \lambda^{-1} / 2}^{(l)}\|_{L^2(\mathbb{R}^2)}^2 \leq \|\mathbb{1}_A\|_{L^2(\mathbb{R}^2)}^2 \|\mathbb{h}^{(l)}\|_{L^1(\mathbb{R}^2)}^2 \lesssim |A|, \end{aligned}$$

where we used Young's inequality (7.11). Finally, we integrate in the remaining variables to get

$$\mathcal{A} \lesssim |A| \log \frac{1}{\varepsilon}.$$

When dealing with \mathcal{B} we begin by integrating out the variables w and u_k for each $k \neq j$. Then we apply (7.4) and Plancherel's theorem to what is left, obtaining

$$\mathcal{B} \leq -\frac{1}{2} \sum_{j=1}^n \int_{\varepsilon}^1 \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\zeta, \eta)|^2 \widehat{\mathbb{k}}\left(\frac{a_j t \zeta}{\sqrt{2}} \lambda, \frac{a_j t \eta}{\sqrt{2}} \frac{1}{\lambda}\right) d(\zeta, \eta) \frac{dt}{t}.$$

This, (10.1), and Lemma 10.3 altogether imply

$$\begin{aligned} & \int_0^\infty (\mathcal{N}_\lambda^\varepsilon(A; R) - \mathcal{N}_\lambda^1(A; R))^2 \frac{d\lambda}{\lambda} \\ & \lesssim R^{-4} |A| \left(\log \frac{1}{\varepsilon} \right) \sum_{j=1}^n \int_{\varepsilon}^1 \int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\zeta, \eta)|^2 \left(\int_0^\infty \left| \widehat{\mathbb{k}}\left(\frac{a_j t \zeta}{\sqrt{2}} \lambda, \frac{a_j t \eta}{\sqrt{2}} \frac{1}{\lambda}\right) \right| \frac{d\lambda}{\lambda} \right) d(\zeta, \eta) \frac{dt}{t} \\ & \lesssim R^{-4} |A| \left(\log \frac{1}{\varepsilon} \right) \int_{\varepsilon}^1 |A| \frac{dt}{t} = \delta^2 \left(\log \frac{1}{\varepsilon} \right)^2. \end{aligned}$$

It remains to change the integration variable, $\lambda = e^\alpha$. □

Proof of Theorem 10. For a given $\delta \in (0, 1]$, Lemmas 10.1 and 10.2 allow us to find $\varepsilon \sim \delta^{2n}$ such that

$$|\mathcal{N}_\lambda^0(A; R) - \mathcal{N}_\lambda^\varepsilon(A; R)| \leq \frac{1}{3} \mathcal{N}_\lambda^1(A; R)$$

holds for every $1 \leq \lambda \leq R$. Afterwards, by taking

$$J \sim \frac{1}{\delta^{2n}} \left(1 + \log \frac{1}{\delta}\right)^2$$

and restricting the integration in Lemma 10.4 to $[0, J]$, we can find a particular $\beta \in [0, J]$ such that

$$|\mathcal{N}_{e^\beta}^\varepsilon(A; R) - \mathcal{N}_{e^\beta}^1(A; R)| \leq \left(\frac{1}{J} \int_0^J (\mathcal{N}_{e^\alpha}^\varepsilon(A; R) - \mathcal{N}_{e^\alpha}^1(A; R))^2 d\alpha\right)^{1/2} \leq \frac{1}{3} \mathcal{N}_{e^\beta}^1(A; R).$$

Decomposition (8.3) then shows

$$\mathcal{N}_{e^\beta}^0(A; R) \geq \frac{1}{3} \mathcal{N}_{e^\beta}^1(A; R) \gtrsim \delta^{n+1} > 0$$

and we are done. In the last display we needed $R \geq e^J \geq e^\beta$ in order to be able to apply Lemma 10.1. This is certainly satisfied whenever

$$R \geq \exp(C\delta^{-2n}(1 + \log \delta^{-1})^2)$$

for a sufficiently large constant $C > 0$. On the other hand, the condition from the theorem statement can be rewritten as

$$R \geq \exp(C\delta^{-2n-1}),$$

which is a stronger requirement than needed. \square

11. REMARKS FOR HARMONIC ANALYSTS

One can wonder about an analytical reason why a density result for area 1 rectangles is not possible in any dimension, while we were able to prove such a result for area 1 right-angled triangles.

Namely, take Schwartz functions φ and ψ , respectively defined on \mathbb{R}^{d_1} and \mathbb{R}^{d_2} , that have compact frequency supports and suppose that at least one of these supports does not contain the origin. Also, take some 4-tuple (p_1, p_2, p_3, p_4) of Lebesgue exponents in $[1, \infty]$ satisfying the Hölder scaling,

$$\sum_{k=1}^4 \frac{1}{p_k} = 1.$$

In what follows, $0 < r < R$ will be certain parameters that determine the range of the scales t that we want to study. The best constant $C_{r,R}$ in the quadrilinear estimate

$$\left| \int_r^R \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_2}} f_1(x, y) f_2(x, y') f_3(x', y) f_4(x', y') \right. \\ \left. \times \varphi_t(x - x') \psi_{1/t}(y - y') dy dy' dx dx' \frac{dt}{t} \right| \leq C_{r,R} \prod_{k=1}^4 \|f_k\|_{L^{p_k}(\mathbb{R}^{d_1+d_2})}$$

grows like a multiple of $\log(R/r)$ as $R/r \rightarrow \infty$. This is not any better than in the trivial estimate obtained from Hölder's inequality applied for each fixed t . Instead of giving a direct argument for this observation, one could argue that any nontrivial cancellation, i.e., a bound of the form $o(\log(R/r))$ as $R/r \rightarrow \infty$, would translate to a positive result for rectangles of area 1. We would simply adjust the proof of Theorem 5 to work for fixed area rectangles, making it similar to the proof of Theorem 1, which we know to be impossible due to Theorems 3 and 4.

A direct argument goes as follows. For simplicity we work in two dimensions, i.e., $d_1 = d_2 = 1$. For some $M > 0$ take $f_1: \mathbb{R}^2 \rightarrow \mathbb{C}$ to be

$$f_1(x, y) := e^{2\pi i xy} \mathfrak{g}\left(\frac{x}{M}\right) \mathfrak{g}\left(\frac{y}{M}\right)$$

and let $f_2 := \overline{f_1}$, $f_3 := \overline{f_1}$, $f_4 := f_1$. The right hand side of the quadrilinear estimate is comparable to $C_{r,R} M^2$, while

$$\begin{aligned} & \lim_{M \rightarrow \infty} \frac{1}{M^2} \int_r^R \int_{\mathbb{R}^4} f_1(x, y) f_2(x, y') f_3(x', y) f_4(x', y') \varphi_t(x - x') \psi_{1/t}(y - y') \, d(x, x', y, y') \frac{dt}{t} \\ & \quad [\text{substitute } u = x - x', v = y - y'] \\ & = \lim_{M \rightarrow \infty} \frac{1}{4} \int_r^R \int_{\mathbb{R}^2} e^{2\pi i uv} \varphi_t(u) \psi_{1/t}(v) \mathfrak{g}\left(\frac{u}{M}\right) \mathfrak{g}\left(\frac{v}{M}\right) \, d(u, v) \frac{dt}{t} \\ & = \frac{1}{4} \left(\log \frac{R}{r} \right) \int_{\mathbb{R}} \widehat{\varphi}(-v) \psi(v) \, dv, \end{aligned}$$

so we only need to make sure that φ and ψ are such that the last integral is nonzero. The fact that the above quadrilinear form does not satisfy any nontrivial estimates is in a sharp contrast with results for the so-called *twisted paraproduct*, where the scales are aligned as φ_t, ψ_t , rather than $\varphi_t, \psi_{1/t}$. It was shown to be bounded, at least when the exponents p_k are greater than 2, by Durcik [18, 19] and, in a dyadic model, previously also by the author [39].

The form associated with area 1/2 triangles is obtained by taking f_4 to be constantly 1 and the associated exponent p_4 to be ∞ . The resulting trilinear inequality

$$\begin{aligned} & \left| \int_r^R \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_2}} f_1(x, y) f_2(x, y') f_3(x', y) \varphi_t(x - x') \psi_{1/t}(y - y') \, dy \, dy' \, dx \, dx' \frac{dt}{t} \right| \\ & \leq C'_{r,R} \prod_{k=1}^3 \|f_k\|_{L^{p_k}(\mathbb{R}^{d_1+d_2})} \end{aligned}$$

now actually has non-trivial cancellation over the scales t . Its best constants $C'_{r,R}$ grow at most like $O((\log(R/r))^{1/2})$, which is easily seen similarly as in the proof of Lemma 8.4: by separating one of the functions and applying the classical square function bound to it. However, one could still ask if the above trilinear form is also bounded as $r \rightarrow 0+$ and $R \rightarrow \infty$, or, equivalently, if the constants $C'_{r,R}$ are even independent of the scale truncation parameters r and R . If, on the other hand, the scales were aligned as φ_t, ψ_t , this would again be the bound for the twisted paraproduct and it would follow from [39, Theorem 1].

Finally, if only functions f_1 and f_4 remain, then it is indeed easy to use Plancherel's theorem and Lemma 10.3 to prove the estimate

$$\begin{aligned} & \left| \int_r^R \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_2}} f_1(x, y) f_4(x', y') \varphi_t(x - x') \psi_{1/t}(y - y') \, dy \, dy' \, dx \, dx' \frac{dt}{t} \right| \\ & = \left| \int_r^R \langle f_1, f_4 * (\varphi_t \otimes \psi_{1/t}) \rangle_{L^2(\mathbb{R}^{d_1+d_2})} \frac{dt}{t} \right| \leq C \|f_1\|_{L^2(\mathbb{R}^{d_1+d_2})} \|f_4\|_{L^2(\mathbb{R}^{d_1+d_2})}, \end{aligned}$$

which is now uniform in r and R . A manifestation of this property has been used in the proof of Theorem 10: in the proof of Lemma 10.4 we were able to integrate over all scales e^α , even without squaring the difference first.

We intentionally do not discuss this matter further or in more detail here. It could be interesting on its own to define a class of “volume-preserving” or “time reversed” (alluding to the appearance of scale $1/t$) multilinear singular integral operators, and then attempt to classify those that are bounded and those that merely possess non-trivial cancellation over the scales.

ACKNOWLEDGMENTS

This work was supported in part by the Croatian Science Foundation under the project HRZZ-IP-2022-10-5116 (FANAP). The author is grateful to Mohammad Bardestani, Adrian Beker, James Davies, Polona Durcik, Keivan Mallahi-Karai, and Rudi Mrazović for several very useful discussions.

REFERENCES

- [1] Sukumar Das Adhikari and Yong-Gao Chen. On monochromatic configurations for finite colorings. *Discrete Math.*, 333:106–109, 2014. doi:10.1016/j.disc.2014.06.015.
- [2] Sukumar Das Adhikari and Purusottam Rath. Monochromatic configurations for finite colourings of the plane. *Note Mat.*, 22(1):59–63, 2003/04.
- [3] Sukumar Das Adhikari and Purusottam Rath. Remarks on monochromatic configurations for finite colorings of the plane. *Note Mat.*, 32(2):83–88, 2012. doi:10.1109/mcs.2012.2185888.
- [4] Christine Bachoc, Evan DeCorte, Fernando Mário de Oliveira Filho, and Frank Vallentin. Spectral bounds for the independence ratio and the chromatic number of an operator. *Israel J. Math.*, 202(1):227–254, 2014. doi:10.1007/s11856-014-1070-7.
- [5] Mohammad Bardestani and Keivan Mallahi-Karai. On a generalization of the Hadwiger-Nelson problem. *Israel J. Math.*, 217(1):313–335, 2017. doi:10.1007/s11856-017-1448-4.
- [6] Thomas Bloom. Erdős problems. <https://www.erdosproblems.com/>. Accessed: April 15, 2024.
- [7] 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.
- [8] Boris Bukh. Measurable sets with excluded distances. *Geom. Funct. Anal.*, 18(3):668–697, 2008. doi:10.1007/s00039-008-0673-8.
- [9] Kamil Bulinski. Spherical recurrence and locally isometric embeddings of trees into positive density subsets of \mathbb{Z}^d . *Math. Proc. Cambridge Philos. Soc.*, 165(2):267–278, 2018. doi:10.1017/S0305004117000470.
- [10] Davi Castro-Silva, Fernando Mário de Oliveira Filho, Lucas Slot, and Frank Vallentin. A recursive Lovász theta number for simplex-avoiding sets. *Proc. Amer. Math. Soc.*, 150(8):3307–3322, 2022. doi:10.1090/proc/15940.
- [11] 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.
- [12] Hallard T. Croft, Kenneth J. Falconer, and Richard K. Guy. *Unsolved problems in geometry*. Problem Books in Mathematics. Springer-Verlag, New York, 1991. Unsolved Problems in Intuitive Mathematics, II. doi:10.1007/978-1-4612-0963-8.
- [13] James Davies. Chromatic number of spacetime. Preprint (2024), arXiv:2308.16885.
- [14] James Davies. Odd distances in colourings of the plane. *Geom. Funct. Anal.*, 34(1):19–31, 2024. doi:10.1007/s00039-024-00659-w.
- [15] James Davies, Rose McCarty, and Michał Pilipczuk. Prime and polynomial distances in colourings of the plane. Preprint (2023), arXiv:2308.02483.
- [16] Aubrey D. N. J. de Grey. The chromatic number of the plane is at least 5. *Geombinatorics*, 28(1):18–31, 2018.
- [17] Adrian Dumitrescu and Minghui Jiang. Monochromatic simplices of any volume. *Discrete Math.*, 310(4):956–960, 2010. doi:10.1016/j.disc.2009.09.026.
- [18] 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.

- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] Polona Durcik and Christoph Thiele. Singular Brascamp-Lieb: a survey. In *Geometric aspects of harmonic analysis*, volume 45 of *Springer INdAM Ser.*, pages 321–349. Springer, Cham, 2021. URL: https://doi.org/10.1007/978-3-030-72058-2_9, doi:10.1007/978-3-030-72058-2\9.
- [24] Paul Erdős. Set-theoretic, measure-theoretic, combinatorial, and number-theoretic problems concerning point sets in Euclidean space. *Real Anal. Exchange*, 4(2):113–138, 1978/79.
- [25] 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.
- [26] Paul Erdős and Ronald Lewis Graham. *Old and new problems and results in combinatorial number theory*, volume 28 of *Monographies de L’Enseignement Mathématique*. Université de Genève, L’Enseignement Mathématique, Geneva, 1980.
- [27] Paul Erdős, Ronald Lewis Graham, Peter Lawrence Montgomery, Bruce L. Rothschild, Joel H. Spencer, and Ernst Gabor Straus. Euclidean Ramsey theorems. I. *J. Combinatorial Theory Ser. A*, 14:341–363, 1973. doi:10.1016/0097-3165(73)90011-3.
- [28] Paul Erdős, Ronald Lewis Graham, Peter Lawrence Montgomery, Bruce L. Rothschild, Joel H. Spencer, and Ernst Gabor Straus. Euclidean Ramsey theorems. II. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday)*, Vols. I, II, III, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 529–557. North-Holland, Amsterdam-London, 1975.
- [29] Paul Erdős, Ronald Lewis Graham, Peter Lawrence Montgomery, Bruce L. Rothschild, Joel H. Spencer, and Ernst Gabor Straus. Euclidean Ramsey theorems. III. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday)*, Vols. I, II, III, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 559–583. North-Holland, Amsterdam-London, 1975.
- [30] Kenneth J. Falconer. The realization of distances in measurable subsets covering \mathbb{R}^n . *J. Combin. Theory Ser. A*, 31(2):184–189, 1981. doi:10.1016/0097-3165(81)90014-5.
- [31] 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.
- [32] 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. URL: https://doi.org/10.1007/978-3-642-72905-8_13, doi:10.1007/978-3-642-72905-8\13.
- [33] Ron Graham and Steve Butler. *Rudiments of Ramsey theory*, volume 123 of *CBMS Regional Conference Series in Mathematics*. Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, second edition, 2015. doi:10.1090/cbms/123.
- [34] Ronald Lewis Graham. On partitions of \mathbb{E}^n . *J. Combin. Theory Ser. A*, 28(1):89–97, 1980. doi:10.1016/0097-3165(80)90061-8.
- [35] Ronald Lewis Graham and Joel H. Spencer. A general Ramsey product theorem. *Proc. Amer. Math. Soc.*, 73(1):137–139, 1979. doi:10.2307/2042899.
- [36] Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008. doi:10.4007/annals.2008.167.481.
- [37] A. Iosevich and K. Taylor. Finite trees inside thin subsets of \mathbb{R}^d . In *Modern methods in operator theory and harmonic analysis*, volume 291 of *Springer Proc. Math. Stat.*, pages 51–56. Springer, Cham, 2019. URL: https://doi.org/10.1007/978-3-030-26748-3_3, doi:10.1007/978-3-030-26748-3\3.
- [38] Vjekoslav Kovač. Bellman function technique for multilinear estimates and an application to generalized paraproducts. *Indiana Univ. Math. J.*, 60(3):813–846, 2011. doi:10.1512/iumj.2011.60.4784.

- [39] Vjekoslav Kovač. Boundedness of the twisted paraproduct. *Rev. Mat. Iberoam.*, 28(4):1143–1164, 2012. doi:10.4171/RMI/707.
- [40] Vjekoslav Kovač. Density theorems for anisotropic point configurations. *Canad. J. Math.*, 74(5):1244–1276, 2022. doi:10.4153/S0008414X21000225.
- [41] Vjekoslav Kovač. Large copies of large configurations in large sets (extended abstract). In T. Orponen, P. Shmerkin, and H. Wang, editors, *Incidence Problems in Harmonic Analysis, Geometric Measure Theory, and Ergodic Theory. Workshop report 25/2023 of the conference held at Oberwolfach, June 4–June 9, 2023*, volume 20 of *Oberwolfach Rep.* EMS Press, EMS-Publishing House GmbH, 2023. doi:10.4171/OWR/2023/25.
- [42] Vjekoslav Kovač and Bruno Predojević. Large dilates of hypercube graphs in the plane. *Anal. Math.*, 2024. Accepted for publication.
- [43] 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.
- [44] 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.
- [45] Neil Lyall and Ákos Magyar. Distances and trees in dense subsets of \mathbb{Z}^d . *Israel J. Math.*, 240(2):769–790, 2020. doi:10.1007/s11856-020-2079-8.
- [46] 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.
- [47] R. Daniel Mauldin. Some problems in set theory, analysis and geometry. In *Paul Erdős and his mathematics, I (Budapest, 1999)*, volume 11 of *Bolyai Soc. Math. Stud.*, pages 493–506. János Bolyai Math. Soc., Budapest, 2002.
- [48] Ian D. Morris. A note on configurations in sets of positive density which occur at all large scales. *Israel J. Math.*, 207(2):719–738, 2015. doi:10.1007/s11856-015-1187-3.
- [49] Ilya D. Shkredov. On a problem of Gowers. *Izv. Ross. Akad. Nauk Ser. Mat.*, 70(2):179–221, 2006. doi:10.1070/IM2006v070n02ABEH002316.
- [50] Alexander Soifer. *The New Mathematical Coloring Book. Mathematics of Coloring and the Colorful Life of Its Creators*. Springer New York, NY, 2024. doi:10.1007/978-1-0716-3597-1.
- [51] 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.
- [52] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.
- [53] Hugo Steinhaus. Sur les distances des points dans les ensembles de mesure positive. *Fund. Math.*, 1:93–104, 1920. doi:10.4064/fm-1-1-93-104.
- [54] Ernst Gabor Straus. A combinatorial theorem in group theory. *Math. Comp.*, 29:303–309, 1975. doi:10.2307/2005482.
- [55] Terence Tao. The ergodic and combinatorial approaches to Szemerédi’s theorem. In *Additive combinatorics*, volume 43 of *CRM Proc. Lecture Notes*, pages 145–193. Amer. Math. Soc., Providence, RI, 2007. doi:10.1090/crmpp/043/08.

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY OF ZAGREB, BIJENIČKA CESTA 30,
10000 ZAGREB, CROATIA

Email address: vjekovac@math.hr