ON HYPERBOLIC CORNERS AND UNIT-AREA TRIANGLES IN PLANAR SETS OF LARGE MEASURE

ALEKSANDAR BULJ AND VJEKOSLAV KOVAČ

Abstract. For large R, we consider measurable sets $A \subseteq [0,R]^2$ that avoid triples of points of the form (x,y), $(x+t,y)$, $(x,y+1/t)$ with $x,y \in \mathbb{R}$ and $t>0$, i.e., the vertices of upward-oriented, axis-aligned right triangles of area $1/2$. We prove that the measures of such sets satisfy $|A| = O(R^2/(\log R)^c)$ for some constant $c>0$. An ingredient in the proof is a hyperbolic variant of the two-dimensional trilinear smoothing inequality by Christ, Durcik, and Roos, which follows from the work of Hsu and Lin. The aforementioned upper bound is complemented with an example of a set of measure $\Omega(R^{4/3})$ avoiding the same point configuration.

Next, we study measurable sets $A \subseteq [0,R]^2$ that avoid triples of points spanning a general triangle of area 1 and establish the upper bound $O(R)$ on their measures. This makes partial progress on a question by Erdős, who conjectured a constant upper bound, and improves over a quantitatively weak $o(R^2)$ result by Graham.

Contents

1. Introduction 1
1.1. Sets with no hyperbolic corners 2
1.2. Sets with no unit-area triangles 3
1.3. Notation 4
2. Proof of the upper bound in Theorem 1 5
2.1. The structured part 6
2.2. The error part 7
2.3. The uniform part 8
2.4. Proof completion 10
3. Proof of the lower bound in Theorem 1 12
4. Proof of Theorem 2 12
Appendix A. Proof of the hyperbolic trilinear smoothing 14
Appendix B. Graham's argument 17
Declaration of AI usage 19
Acknowledgments and funding 19
References 20

1. Introduction

We refer to Subsection 1.3 for the quite standard asymptotic notation that we use in this paper.

Date: November 15, 2026.
2020 *Mathematics Subject Classification.* Primary 28A75; Secondary 05D10, 42B20.

1.1. **Sets with no hyperbolic corners.** The first main result of the present paper deals with bounds on measures of subsets A of a large square $[0,R]^2$ that avoid a certain three-point configuration, which we call a *hyperbolic corner*.

Theorem 1. *Let $M(R)$ be the supremum of $|A|$ over measurable sets $A \subseteq [0,R]^2$ containing no triples of the form*

$$(x,y),\ (x+t,y),\ (x,y+t^{-1}) \tag{1.1}$$

with $x,y \in \mathbb{R}$ and $t>0$. Then for $R \geqslant 10$ we have

$$R^{4/3} \lesssim M(R) \lesssim R^2 \Big(\frac{\log\log R}{\log R}\Big)^{1/4}. \tag{1.2}$$

In particular,

$$M(R) \lesssim_\epsilon \frac{R^2}{(\log R)^{1/4-\epsilon}}$$

for every $0<\epsilon<1/4$. We do not know of any published nontrivial (i.e., better than $O(R^2)$) upper estimate on $M(R)$. The proof of the upper estimate in (1.2) will be given in Section 2. The exponent $1/4$ in the upper bound, like the Estimate (1.2) itself, is unlikely to be optimal, but it reflects the limitations of currently available techniques.

Theorem 1 naturally fits into an active line of research investigating *nonlinear corners*, which are configurations of the form

$$(x,y),\ (x+\gamma_1(t),y),\ (x,y+\gamma_2(t)) \tag{1.3}$$

contained either in continuous structures, like $[0,1]^2$ and $\mathbb{R}^2$, or in discrete structures like $\{0,1,\dots,n-1\}^2$ and $\mathbb{F}_q^2$. These are "curved" counterparts of the usual (linear) corners, studied extensively since [1, 44], with the recent state of the art being [27]. Most of the literature is concerned with the cases when γ_1 and γ_2 are polynomials (often of distinct degrees). Bergelson and Leibman [4] (also see [3]) showed a density theorem on the integer lattice with non-effective bounds. This has been successively quantitatively improved all the way to polylogarithmic bounds like ours; see [9, 32, 23]. Finite-field results were being developed simultaneously [14, 24, 33, 34, 31, 37].

However, the present configuration (1.1) is not a polynomial corner, which rules out purely polynomial tools, such as the PET induction and the degree lowering of Gowers norms. The paper closest in technique to ours is the groundbreaking work of Christ, Durcik, and Roos [11], who showed a continuous-parameter density theorem for

$$(x,y),\ (x+t,y),\ (x,y+t^a),$$

where $a>1$ is fixed. (For simplicity of notation, only the case $a=2$ was worked out explicitly in [11].) The proof in [11] does not fully extend to the case $a=-1$ (or, in fact, any $a<0$), even though we are primarily influenced by their work, which, in turn, reflects ideas from the paper by Bourgain [7] on the nonlinear Roth theorem in $\mathbb{R}$. One should also remark that Lim [37] studied (1.3) for certain rational functions γ_1,γ_2, but in the finite-field setting $\mathbb{F}_q^2$, where rational functions are close in spirit to polynomials.[1]

The problem of estimating $M(R)$ also belongs to the broader theme of forcing finite configurations inside large subsets of Euclidean spaces; see [29] for a brief and accessible introduction to the topic. The two-point subpattern $(x+t,y)$, $(x,y+1/t)$ of (1.1) has already attracted

[1] After the first preprint version of the present paper was announced, Hsu and Lin [26, Thm. 1.40] showed a Bourgain-type result for nonlinear corners (x,y), $(x+\gamma_1(t),y)$, $(x,y+\gamma_2(t))$ with rather general nonlinear γ_1,γ_2 such that $\gamma_1(0)=0=\gamma_2(0)$. It does not cover our Theorem 1.

some attention in the literature. Indeed, after a rotation and a translation, it becomes the configuration

$$(x, y),\ (x, y) + (t, t^{-1}) \tag{1.4}$$

studied by Bardestani and Mallahi-Karai [2] (also see [30, Thm. 10]), who proved an estimate of the same polylogarithmic type as (1.2), namely $O(R^2/\log R)$, but under the stronger hypothesis that a set $A \subseteq [0, R]^2$ avoids every pair of the form (1.4). Two-point patterns of the form

$$(x, y),\ (x, y) + (t, \gamma(t))$$

for appropriate curves γ have also been studied in [35, 17, 10, 8]. Configurations with more points typically require additional ideas, and known approaches often work in higher ambient dimensions [6, 39, 41, 40]. A result in that direction is [30, Thm. 1]. Its special case concerning axis-aligned right triangles is much weaker than Theorem 1: it asserts that a set $A \subseteq [0, R]^3$ avoiding every triple of points

$$(x, y, z),\ (x + t, y, z),\ (x, y + t^{-1}u, z + t^{-1}v)$$

for $u^2 + v^2 = 1$ satisfies the measure bound $|A| = O(R^3/(\log R)^{36})$. The important point is not the different power of the logarithm, but rather the fact that an additional degree of freedom $(u, v) \in \mathrm{S}^1$ allows one to use the Fourier decay of the surface measure, which is clearly not available in the plane.

The basic outline of the proof of the upper bound in (1.2) only superficially resembles those in recent papers in geometric Ramsey theory [12, 16, 28, 15, 30]. The beginning is quite standard: we define an appropriate pattern-counting form (2.1) associated with a parameter $\lambda > 0$. In our case λ determines an approximate scale of eccentricity of the configuration (1.1) that we aim to detect. A certain novelty (employed also in [30]) is that this parameter does not represent any usual notion of configuration "size." In particular, it makes no sense to let $\lambda \to 0^+$ or $\lambda \to +\infty$. In classical scaling arguments, letting λ approach 0 simply means that one also detects arbitrarily small instances of the desired configuration, which are in fact trivially found by the Lebesgue density theorem. Despite the absence of this interpretation, we can still meaningfully decompose the counting form into the structured, error, and uniform parts (2.4), just as in, e.g., [12, 28, 15]. Another notable difference here is that all values of λ contribute equally to the overall pattern count. Thus, to control the error part, we can simply integrate in λ over $[1/R, R]$ and avoid using Bourgain's trick of pigeonholing in the scales [6, 7]. This additional symmetry is responsible for the better upper bound in (1.2) compared with [11], with a power of $\log R$, rather than a power of $\log\log R$, in the denominator. The final ingredient in the proof of Theorem 1 is to handle the uniform part by reducing its bound to a hyperbolic variant of the trilinear smoothing estimate by Christ, Durcik, and Roos [11]; see Lemma 5.

A trivial lower bound for $M(R)$ is $M(R) \gtrsim R$, and it is realized by a single skew strip of width ~ 1 intersecting the square $[0, R]^2$. The proof of the lower estimate in (1.2) uses carefully chosen congruent equidistant strips. The details are presented in Section 3.

1.2. **Sets with no unit-area triangles.** Observe that the configuration (1.1) consists of the vertices of an upward-oriented, axis-aligned right triangle of area 1/2. We turn to the study of sets that avoid the vertices of any (possibly rotated and scalene) triangle of a fixed area. Part of the motivation comes from the following problem by Erdős on unit-area triangles, as quoted from [5, Prob. 352].

> *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?*

Erdős posed this problem in [18, p. 122–123], [19, p. 30], and [20, p. 323], and the problem was also popularized by Croft, Falconer, and Guy [13, Prob. G13], Mauldin [42, Sec. 1], [43, Sec. 5], and, most recently, the *Erdős problems* community [5, Prob. 352]. Our second main result makes partial progress by proving an upper bound on measures of sets A contained in a large square $[0,R]^2$ that do not contain vertices of a unit-area triangle.

Theorem 2. *Let $M_\Delta(R)$ be the supremum of $|A|$ over measurable sets $A \subseteq [0,R]^2$ containing no triples of points that span a triangle of area 1. Then $M_\Delta(R) \lesssim R$ for every $R \geqslant 1$.*

To the best of our knowledge, there are no published upper bounds on $M_\Delta(R)$ of the shape $o(R^2)$ as $R \to \infty$. However, the corresponding discrete density result was commented rather cryptically by Graham [22, p. 96], in relation with his coloring result for triangles.

> *There is a density version of Theorem 1 which can be proved, based on using Szemerédi's theorem in place of van der Waerden's theorem. This density version asserts that for any $\epsilon > 0$, there is an integer $T(\epsilon)$ so that if $n \geqslant n(\epsilon)$ and $R \subseteq \{(i,j) : 1 \leqslant i,j \leqslant n\}$ with $|R| > \epsilon n^2$, then R contains the vertices of a triangle of area $T(\epsilon)$. We will not discuss those directions in this note, however.*

In Appendix B we decipher what Graham actually meant and use this to derive

$$M_\Delta(R) = o(R^2) \quad \text{as } R \to \infty. \tag{1.5}$$

However, the latter proof does not come anywhere near concrete effective estimates on $M_\Delta(R)$.

Theorem 1 and $M_\Delta(R) \lesssim M(R)$ together already imply a concrete estimate. However, Theorems 1 and 2 reveal that the numbers $M(R)$ and $M_\Delta(R)$ necessarily have very different asymptotic behavior. The proof of Theorem 2 in Section 4 uses the lengths of horizontal sections of rotated copies of A. Triangle avoidance gives a pointwise inequality for these lengths. We average it over all rotations and apply Plancherel's theorem in polar coordinates.

We do not know of any non-trivial lower bounds on $M_\Delta(R)$ and the original Erdős problem is still open. In fact, Erdős even suspected that $M_\Delta(R) = 4\pi/\sqrt{27}$ for every sufficiently large R, witnessed by the open disk circumscribed around the equilateral triangle of area 1.

1.3. **Notation.** Let $\mathcal{A} \lesssim \mathcal{B}$, $\mathcal{B} \gtrsim \mathcal{A}$, $\mathcal{A} = O(\mathcal{B})$, and $\mathcal{B} = \Omega(\mathcal{A})$ all mean that there exists a positive constant C such that $|\mathcal{A}| \leqslant C|\mathcal{B}|$. If the constant is allowed to depend on a set of parameters P, we emphasize this as $\mathcal{A} \lesssim_P \mathcal{B}$, $\mathcal{B} \gtrsim_P \mathcal{A}$, $\mathcal{A} = O_P(\mathcal{B})$, or $\mathcal{B} = \Omega_P(\mathcal{A})$. The meaning of $\mathcal{A} \sim \mathcal{B}$ (resp. $\mathcal{A} \sim_P \mathcal{B}$) is that both $\mathcal{A} \lesssim \mathcal{B}$ and $\mathcal{A} \gtrsim \mathcal{B}$ (resp. $\mathcal{A} \lesssim_P \mathcal{B}$ and $\mathcal{A} \gtrsim_P \mathcal{B}$) hold.

We write $|A|$ for the Lebesgue measure of a measurable subset A of the Euclidean plane.

The *Fourier transform* of an integrable function $f\colon \mathbb{R}^2 \to \mathbb{C}$ will be normalized as

$$\widehat{f}(\xi,\eta) := \int_{\mathbb{R}^2} f(x,y) e^{-2\pi i(x\xi+y\eta)} \,\mathrm{d}x\,\mathrm{d}y, \quad (\xi,\eta) \in \mathbb{R}^2.$$

Throughout the paper, $\mathcal{F}_1$ will denote the Fourier transform in the first variable and $\mathcal{F}_2$ the Fourier transform in the second variable only:

$$(\mathcal{F}_1 f)(\xi,y) := \int_{\mathbb{R}} f(x,y) e^{-2\pi i x\xi} \,\mathrm{d}x, \quad (\mathcal{F}_2 f)(x,\eta) := \int_{\mathbb{R}} f(x,y) e^{-2\pi i y\eta} \,\mathrm{d}y.$$

Convolution of integrable functions $f, g\colon \mathbb{R}^2 \to \mathbb{C}$ is defined as

$$(f * g)(x,y) := \int_{\mathbb{R}^2} f(x-u, y-v) g(u,v) \,\mathrm{d}u\,\mathrm{d}v$$

for almost every $(x,y)\in\mathbb{R}^2$. Also, $*_1$ (resp. $*_2$) will denote the convolution of a two-variable function f in the first (resp. second) variable with a given single-variable function φ:

$$(f *_1 \varphi)(x,y) := \int_{\mathbb{R}} f(x-s,y)\varphi(s)\,\mathrm{d}s, \quad (f *_2 \varphi)(x,y) := \int_{\mathbb{R}} f(x,y-s)\varphi(s)\,\mathrm{d}s.$$

2. Proof of the upper bound in Theorem 1

Fix a nonnegative C^∞ function ζ that is not identically zero and has compact support contained in the interval $[1/2,2]$. Also let ϕ be an even nonnegative C^∞ function with $\int_{\mathbb{R}}\phi = 1$, supported in $[-20,20]$ and such that ϕ is bounded from below by a positive constant on $[-10,10]$. For $t>0$ write

$$\phi_t(x) := \frac{1}{t}\phi\Big(\frac{x}{t}\Big).$$

If $f_0, f_1, f_2\colon \mathbb{R}^2\to\mathbb{R}$ are bounded, compactly supported, and measurable, define the *exact counting form* associated with a certain scaling parameter $\lambda\in(0,\infty)$ by

$$\mathcal{N}^0_\lambda(f_0,f_1,f_2) := \int_0^\infty \int_{\mathbb{R}^2} f_0(x,y) f_1(x+\lambda u, y) f_2\Big(x, y+\frac{1}{\lambda u}\Big)\zeta(u)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u. \tag{2.1}$$

In the case $\lambda=1$ we simply write $\mathcal{N}$ in place of $\mathcal{N}^0_1$:

$$\mathcal{N}(f_0,f_1,f_2) := \int_0^\infty \int_{\mathbb{R}^2} f_0(x,y) f_1(x+u, y) f_2\Big(x, y+\frac{1}{u}\Big)\zeta(u)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u. \tag{2.2}$$

For $0<\varepsilon\leqslant 1$, we also introduce the *smoothed counting form*

$$\mathcal{N}^\varepsilon_\lambda(f_0,f_1,f_2) := \mathcal{N}^0_\lambda\Big(f_0, f_1 *_1 \phi_{\lambda\varepsilon}, f_2 *_2 \phi_{\lambda^{-1}\varepsilon}\Big).$$

The dominated convergence theorem easily gives

$$\lim_{\varepsilon\to 0^+} \mathcal{N}^\varepsilon_\lambda(f_0,f_1,f_2) = \mathcal{N}^0_\lambda(f_0,f_1,f_2).$$

These forms correspond to a localized piece of the hyperbola $uv=1$ determined by the parameter λ, namely the piece in a scaled neighborhood of the point $(\lambda,1/\lambda)$. The other parameter, ε, is viewed as the smoothing parameter, which "blurs out" the picture and thus typically overcounts the occurrences of the sought point configuration.

When $f_0=f_1=f_2=\mathbb{1}_A$, we simply write $\mathcal{N}^0_\lambda(A)$ and $\mathcal{N}^\varepsilon_\lambda(A)$, respectively. The form $\mathcal{N}^0_\lambda(A)$ counts triples

$$(x,y),\ (x+\lambda u,y),\ \Big(x,y+\frac{1}{\lambda u}\Big)$$

with $u\in\operatorname{supp}\zeta\subseteq[1/2,2]$. Hence, if A contains no triples of the form in Theorem 1, then

$$\mathcal{N}^0_\lambda(A) = 0 \quad \text{for every } \lambda>0. \tag{2.3}$$

We decompose

$$\mathcal{N}^0_\lambda(A) = \mathcal{N}^1_\lambda(A) + \Big(\mathcal{N}^\varepsilon_\lambda(A) - \mathcal{N}^1_\lambda(A)\Big) + \Big(\mathcal{N}^0_\lambda(A) - \mathcal{N}^\varepsilon_\lambda(A)\Big) \tag{2.4}$$

for a number $\varepsilon\in(0,1)$ that will be chosen later. In analogy with the literature in graph theory and arithmetic combinatorics, the three terms are respectively called:

- $\mathcal{N}^1_\lambda(A)$ = the structured part,
- $\mathcal{N}^\varepsilon_\lambda(A) - \mathcal{N}^1_\lambda(A)$ = the error part, and
- $\mathcal{N}^0_\lambda(A) - \mathcal{N}^\varepsilon_\lambda(A)$ = the uniform part;

the same names were also used for somewhat analogous quantities in [15], [28], and in several later papers. The basic strategy is proving a lower bound for the structured part (i.e., on the very "coarse" scale $\varepsilon = 1$) that is independent of λ taken from the most reasonable scale range $\lambda \in [1/R, R]$. The uniform part should satisfy an upper bound that is again uniform in λ and makes it negligible in comparison with the structured part as soon as ε is chosen to be sufficiently small. Then we are left with the error part, which will be controlled, but rather in the L^2 norm sense with respect to the measure $\mathrm{d}\lambda/\lambda$. This will contradict (2.3), not for a single concrete value of λ, but "on the average," i.e., once we integrate (2.4) in λ. Once again, we remark that this strategy proves to be efficient, despite the fact that one cannot "interpolate" the proof with some trivial limit $\lambda \to 0^+$, as is usually the case with configurations that can shrink to a point so that the Lebesgue density theorem applies there.

2.1. **The structured part.** We begin with the study of the first term appearing in decomposition (2.4).

Lemma 3. *For every measurable $A \subseteq [0,R]^2$ and every λ with*

$$R \geqslant \max\{\lambda, \lambda^{-1}\},$$

one has

$$\mathcal{N}^1_\lambda(A) \gtrsim \frac{|A|^3}{R^4} = R^2\Big(\frac{|A|}{R^2}\Big)^3.$$

Proof. First, we rewrite the smoothed form with $\varepsilon = 1$ as

$$\mathcal{N}^1_\lambda(f_0,f_1,f_2) = \int_{\mathbb{R}^4} f_0(x,y)f_1(x',y)f_2(x,y')K^+_\lambda(x-x',y-y')\,\mathrm{d}x\,\mathrm{d}x'\,\mathrm{d}y\,\mathrm{d}y',$$

where

$$K^+_\lambda(x,y) := \int_0^\infty \phi\Big(\frac{x}{\lambda}+u\Big)\phi\Big(\lambda y+\frac{1}{u}\Big)\zeta(u)\,\mathrm{d}u.$$

Since ϕ is bounded from below on $[-10,10]$ and ζ is supported in $[1/2,2]$, we have the pointwise kernel bound

$$K^+_\lambda \gtrsim \mathbb{1}_{[-\lambda,\lambda]\times[-\lambda^{-1},\lambda^{-1}]},$$

which we can plug into the last representation of $\mathcal{N}^1_\lambda(A) = \mathcal{N}^1_\lambda(\mathbb{1}_A,\mathbb{1}_A,\mathbb{1}_A)$. That way we obtain

$$\mathcal{N}^1_\lambda(A) \gtrsim \sum_{Q=I\times J}\int_{I\times I\times J\times J} \mathbb{1}_A(x,y)\mathbb{1}_A(x',y)\mathbb{1}_A(x,y')\,\mathrm{d}x\,\mathrm{d}x'\,\mathrm{d}y\,\mathrm{d}y', \tag{2.5}$$

where the sum runs over the partition of $[0,R]^2$ into $N \sim R^2$ many rectangles $Q = I\times J$ with $|I| \leqslant \lambda$ and $|J| \leqslant 1/\lambda$.

Fix one such rectangle $Q = I\times J$. Abbreviate

$$F(x) := \int_J \mathbb{1}_A(x,y')\,\mathrm{d}y', \quad G(y) := \int_I \mathbb{1}_A(x',y)\,\mathrm{d}x'.$$

By Hölder's inequality,

$$\begin{aligned}|A\cap Q|^3 &= \Big(\int_{I\times J}\mathbb{1}_A(x,y)\,\mathrm{d}x\,\mathrm{d}y\Big)^3\\ &\leqslant \Big(\int_{I\times J}\mathbb{1}_A(x,y)F(x)G(y)\,\mathrm{d}x\,\mathrm{d}y\Big)\underbrace{\Big(\int_{I\times J}\frac{\mathbb{1}_A(x,y)}{F(x)}\,\mathrm{d}x\,\mathrm{d}y\Big)}_{\leqslant|I|}\underbrace{\Big(\int_{I\times J}\frac{\mathbb{1}_A(x,y)}{G(y)}\,\mathrm{d}x\,\mathrm{d}y\Big)}_{\leqslant|J|}.\end{aligned}$$

Here we interpret $\mathbb{1}_A(x,y)/F(x)$ and $\mathbb{1}_A(x,y)/G(y)$ as 0 wherever the denominator vanishes. Hence, the corresponding summand in (2.5) satisfies

$$\int_{I\times J} \mathbb{1}_A(x,y)F(x)G(y)\,\mathrm{d}x\,\mathrm{d}y \geqslant \frac{|A\cap Q|^3}{|I||J|} \geqslant |A\cap Q|^3,$$

since $|I||J| \leqslant \lambda\lambda^{-1} = 1$. By convexity of $t\mapsto t^3$ and Jensen's inequality,

$$\sum_Q |A\cap Q|^3 \geqslant N^{-2}\Big(\sum_Q |A\cap Q|\Big)^3 \gtrsim \frac{|A|^3}{R^4}.$$

Combining the last two displayed estimates proves the claim. □

2.2. **The error part.** The second needed ingredient is an L^2 control of the error part from (2.4).

Lemma 4. *For every bounded measurable $A\subseteq\mathbb{R}^2$ and every $0<\varepsilon\leqslant 1$,*

$$\Big(\int_0^\infty |\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A)|^2\,\frac{\mathrm{d}\lambda}{\lambda}\Big)^{1/2} \lesssim |A|\Big(\log\frac{1}{\varepsilon}\Big)^{1/2}. \tag{2.6}$$

Proof. Using

$$\phi_{\lambda\varepsilon}(x)\phi_{\lambda^{-1}\varepsilon}(y) - \phi_\lambda(x)\phi_{\lambda^{-1}}(y) = (\phi_{\lambda\varepsilon}-\phi_\lambda)(x)\,\phi_{\lambda^{-1}\varepsilon}(y) + \phi_\lambda(x)\,(\phi_{\lambda^{-1}\varepsilon}-\phi_{\lambda^{-1}})(y)$$

the form definitions yield

$$\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A) = E_1(\lambda,\varepsilon) + E_2(\lambda,\varepsilon),$$

where

$$E_1(\lambda,\varepsilon) := \mathcal{N}_\lambda^0\Big(\mathbb{1}_A, \mathbb{1}_A *_1 (\phi_{\lambda\varepsilon}-\phi_\lambda), \mathbb{1}_A *_2 \phi_{\lambda^{-1}\varepsilon}\Big),$$
$$E_2(\lambda,\varepsilon) := \mathcal{N}_\lambda^0\Big(\mathbb{1}_A, \mathbb{1}_A *_1 \phi_\lambda, \mathbb{1}_A *_2 (\phi_{\lambda^{-1}\varepsilon}-\phi_{\lambda^{-1}})\Big).$$

For the first term, we use $0\leqslant \mathbb{1}_A *_2 \phi_{\lambda^{-1}\varepsilon}\leqslant 1$. For each fixed u and λ, the Cauchy–Schwarz inequality in (x,y) gives

$$\int_{\mathbb{R}^2} \mathbb{1}_A(x,y)\,\Big|\Big(\mathbb{1}_A *_1 (\phi_{\lambda\varepsilon}-\phi_\lambda)\Big)(x+\lambda u, y)\Big|\,\mathrm{d}x\,\mathrm{d}y \leqslant \|\mathbb{1}_A\|_{\mathrm{L}^2}\,\big\|\mathbb{1}_A *_1 (\phi_{\lambda\varepsilon}-\phi_\lambda)\big\|_{\mathrm{L}^2},$$

so that, integrating this estimate against $\zeta(u)\,\mathrm{d}u$, we obtain

$$|E_1(\lambda,\varepsilon)| \lesssim |A|^{1/2}\big\|\mathbb{1}_A *_1 (\phi_{\lambda\varepsilon}-\phi_\lambda)\big\|_{\mathrm{L}^2}.$$

Consequently, by Plancherel's theorem,

$$\begin{aligned}\int_0^\infty |E_1(\lambda,\varepsilon)|^2\frac{\mathrm{d}\lambda}{\lambda} &\lesssim |A|\int_0^\infty \big\|\mathbb{1}_A *_1 (\phi_{\lambda\varepsilon}-\phi_\lambda)\big\|_{\mathrm{L}^2}^2\frac{\mathrm{d}\lambda}{\lambda}\\ &\lesssim |A|\int_{\mathbb{R}^2} |\widehat{\mathbb{1}_A}(\xi,\eta)|^2\Big(\int_0^\infty \big|\widehat{\phi}(\varepsilon\lambda\xi)-\widehat{\phi}(\lambda\xi)\big|^2\frac{\mathrm{d}\lambda}{\lambda}\Big)\,\mathrm{d}\xi\,\mathrm{d}\eta \qquad (2.7)\\ &\lesssim |A|\log\frac{1}{\varepsilon}\int_{\mathbb{R}^2}|\widehat{\mathbb{1}_A}(\xi,\eta)|^2\,\mathrm{d}\xi\,\mathrm{d}\eta = |A|^2\log\frac{1}{\varepsilon}.\end{aligned}$$

Namely, the inner integral in (2.7) is $\lesssim \log(1/\varepsilon)$ uniformly in $\xi\neq 0$, which is seen by changing variables $s=\lambda|\xi|$ and observing

$$\int_0^\infty \big|\widehat{\phi}(\varepsilon s)-\widehat{\phi}(s)\big|^2\frac{\mathrm{d}s}{s} \lesssim \log\frac{1}{\varepsilon}. \tag{2.8}$$

Estimate (2.8) is, in turn, a consequence of

$$\big|\widehat{\phi}(\varepsilon s)-\widehat{\phi}(s)\big| \lesssim \begin{cases} s & \text{for } 0<s\leqslant 1,\\ 1 & \text{for } 1<s\leqslant \varepsilon^{-1},\\ (\varepsilon s)^{-1} & \text{for } s>\varepsilon^{-1}, \end{cases}$$

which follows from $\widehat{\phi}(0)=1$, smoothness, and the rapid decay of $\widehat{\phi}$. The same argument in the second variable gives

$$\int_0^\infty |E_2(\lambda,\varepsilon)|^2\,\frac{\mathrm{d}\lambda}{\lambda} \lesssim |A|^2\log\frac{1}{\varepsilon},$$

which finally proves (2.6). □

2.3. **The uniform part.** For $s_1,s_2\in\mathbb{R}$ we use the anisotropic Sobolev norm

$$\|f\|_{\mathrm{H}^{s_1,s_2}} := \Big(\int_{\mathbb{R}^2} |\widehat{f}(\xi,\eta)|^2\Big(1+\xi^2\Big)^{s_1}\Big(1+\eta^2\Big)^{s_2}\,\mathrm{d}\xi\,\mathrm{d}\eta\Big)^{1/2}.$$

A genuinely non-elementary ingredient in the paper is the following local estimate, which is a hyperbolic analogue of the trilinear smoothing inequality of Christ, Durcik, and Roos [11].

Lemma 5 (Hyperbolic trilinear smoothing)**.** *There exists $\sigma>0$ such that the trilinear form* (2.2) *satisfies*

$$|\mathcal{N}(f_0,f_1,f_2)| \lesssim \|f_0\|_{\mathrm{L}^\infty}\|f_1\|_{\mathrm{H}^{-\sigma,0}}\|f_2\|_{\mathrm{H}^{0,-\sigma}}$$

for bounded measurable functions $f_0,f_1,f_2\colon\mathbb{R}^2\to\mathbb{R}$ supported in $[-10,10]^2$.

Lemma 5 will be proved in Section A of the appendix. Here we return to decomposition (2.4) and manage to control its last term.

Lemma 6. *There exists $\sigma>0$ such that, for every bounded measurable set $A\subseteq\mathbb{R}^2$, every $\lambda>0$, and every $0<\varepsilon\leqslant 1$,*

$$|\mathcal{N}^0_\lambda(A)-\mathcal{N}^\varepsilon_\lambda(A)| \lesssim |A|\varepsilon^\sigma.$$

Proof. It is sufficient to work with $0<\varepsilon\leqslant 1/20$. Since $\lambda>0$ is fixed, we can define the anisotropically rescaled indicator function

$$f(x,y) := \mathbb{1}_A\Big(\lambda x,\frac{y}{\lambda}\Big),$$

which satisfies

$$\|f\|_{\mathrm{L}^\infty}\leqslant 1,\quad \|f\|_{\mathrm{L}^2}=|A|^{1/2}.$$

We will make sure that the estimates below depend on no other properties of f, and in particular do not depend on λ. This is a place in our proof where the scaling (t,t^{-1}) comes as very convenient. Recall the definitions of the counting forms. Substitutions $x=\lambda x'$, $y=y'/\lambda$ give

$$\mathcal{N}^0_\lambda(A)=\mathcal{N}(f,f,f)$$

and

$$\mathcal{N}^\varepsilon_\lambda(A)=\mathcal{N}\big(f,\,f*_1\phi_\varepsilon,\,f*_2\phi_\varepsilon\big).$$

Our goal is to prove a slightly more general estimate than needed:

$$\big|\mathcal{N}(f,g,h)-\mathcal{N}\big(f,\,g*_1\phi_\varepsilon,\,h*_2\phi_\varepsilon\big)\big| \lesssim \varepsilon^\sigma\|f\|_{\mathrm{L}^\infty}\|g\|_{\mathrm{L}^2}\|h\|_{\mathrm{L}^2}, \tag{2.9}$$

which will prove the lemma, once one takes g and h to be the same as f defined above.

First, we decompose the difference on the left-hand side of (2.9) as $T_1(\varepsilon) + T_2(\varepsilon)$, where

$$T_1(\varepsilon) := \mathcal{N}\big(f,\, g - g *_1 \phi_\varepsilon,\, h\big),$$
$$T_2(\varepsilon) := \mathcal{N}\big(f,\, g *_1 \phi_\varepsilon,\, h - h *_2 \phi_\varepsilon\big).$$

We now localize the two terms T_1 and T_2. The point of the localization is only to put the local smoothing estimate on bounded "windows," but no estimate on the number of windows will be used. All constants in this localization argument are independent of R, λ, and ε. Choose a partition of unity on $\mathbb{R}^2$,

$$1 = \sum_{m\in\mathbb{Z}^2} \rho(z-m), \quad z \in \mathbb{R}^2,$$

for some C^∞ function ρ supported in the square $[-2,2]^2$. For $m \in \mathbb{Z}^2$ put

$$f_m(z) := f(z)\rho(z-m), \quad g_m(z) := g(z)\rho(z-m), \quad h_m(z) := h(z)\rho(z-m),$$

so that only finitely many pieces f_m, g_m, h_m are nonzero and bounded overlap gives

$$\sum_{m\in\mathbb{Z}^2} \|g_m\|_{\mathrm{L}^2}^2 \lesssim \|g\|_{\mathrm{L}^2}^2, \quad \sum_{m\in\mathbb{Z}^2} \|h_m\|_{\mathrm{L}^2}^2 \lesssim \|h\|_{\mathrm{L}^2}^2. \tag{2.10}$$

We expand the first term as

$$T_1(\varepsilon) = \sum_{k\in\mathbb{Z}^2} \sum_{l\in\mathbb{Z}^2} \sum_{m\in\mathbb{Z}^2} \mathcal{N}\big(f_m, g_{m+k} - g_{m+k} *_1 \phi_\varepsilon, h_{m+l}\big), \tag{2.11}$$

noting that the sums in k and l are both effectively finite, since only finitely many summands are nonzero. In fact, the integrand

$$f_m(x,y)\,(g_{m+k} - g_{m+k} *_1 \phi_\varepsilon)(x+u, y)\, h_{m+l}\Big(x, y + \frac{1}{u}\Big)\,\zeta(u)$$

is identically zero unless absolute values of both coordinates of k and l are at most 7, which amounts to having to sum only over $O(1)$ many pairs of indices (k,l). For that reason, it is sufficient to bound the innermost sum in (2.11), the one in m, uniformly in the fixed indices k and l. We next apply Lemma 5 to each such summand, with the smoothing exponent $\sigma \in (0,1]$ as in the lemma. Its estimate applies to the functions f_m, $g_{m+k} - g_{m+k} *_1 \phi_\varepsilon$, h_{m+l}, all translated by $-m$ (so that the translates become supported in $[-10,10]^2$), giving

$$\big|\mathcal{N}\big(f_m, g_{m+k} - g_{m+k} *_1 \phi_\varepsilon, h_{m+l}\big)\big| \lesssim \|f_m\|_{\mathrm{L}^\infty} \|g_{m+k} - g_{m+k} *_1 \phi_\varepsilon\|_{\mathrm{H}^{-\sigma,0}} \|h_{m+l}\|_{\mathrm{H}^{0,-\sigma}}.$$

Here we used the fact that the Sobolev norms remain unchanged by the translation. Using the Cauchy–Schwarz inequality in m, we obtain

$$\begin{aligned}&\Big|\sum_{m\in\mathbb{Z}^2} \mathcal{N}\big(f_m, g_{m+k} - g_{m+k} *_1 \phi_\varepsilon, h_{m+l}\big)\Big| \\ &\lesssim \|f\|_{\mathrm{L}^\infty} \Big(\sum_{m\in\mathbb{Z}^2} \|g_m - g_m *_1 \phi_\varepsilon\|_{\mathrm{H}^{-\sigma,0}}^2\Big)^{1/2} \Big(\sum_{m\in\mathbb{Z}^2} \|h_m\|_{\mathrm{H}^{0,-\sigma}}^2\Big)^{1/2}.\end{aligned} \tag{2.12}$$

Since $\widehat{\phi}(0) = 1$ and $\widehat{\phi}$ is smooth,

$$\big|1 - \widehat{\phi}(\varepsilon\xi)\big| \lesssim \min\{1, \varepsilon|\xi|\} \leqslant \varepsilon^\sigma |\xi|^\sigma \lesssim \varepsilon^\sigma (1+\xi^2)^{\sigma/2},$$

so

$$\|g_m - g_m *_1 \phi_\varepsilon\|_{\mathrm{H}^{-\sigma,0}}^2 = \int_{\mathbb{R}^2} \big|\widehat{g_m}(\xi,\eta)\big|^2 \big|1-\widehat{\phi}(\varepsilon\xi)\big|^2 (1+\xi^2)^{-\sigma}\,\mathrm{d}\xi\,\mathrm{d}\eta$$
$$\lesssim \varepsilon^{2\sigma} \int_{\mathbb{R}^2} |\widehat{g_m}(\xi,\eta)|^2\,\mathrm{d}\xi\,\mathrm{d}\eta = \varepsilon^{2\sigma}\|g_m\|_{\mathrm{L}^2}^2.$$

Hence, by (2.10) and $\|h\|_{\mathrm{H}^{0,-\sigma}} \leqslant \|h\|_{\mathrm{L}^2}$, the right hand side in (2.12) is

$$\lesssim \|f\|_{\mathrm{L}^\infty} \Big(\varepsilon^{2\sigma} \sum_{m\in\mathbb{Z}^2} \|g_m\|_{\mathrm{L}^2}^2\Big)^{1/2} \Big(\sum_{m\in\mathbb{Z}^2} \|h_m\|_{\mathrm{L}^2}^2\Big)^{1/2} \lesssim \varepsilon^\sigma \|f\|_{\mathrm{L}^\infty}\|g\|_{\mathrm{L}^2}\|h\|_{\mathrm{L}^2}.$$

The second term, T_2, is decomposed in the same way, leading to the study of

$$\Big|\sum_{m\in\mathbb{Z}^2} \mathcal{N}\big(f_m,\, g_{m+k} *_1 \phi_\varepsilon,\, h_{m+l} - h_{m+l} *_2 \phi_\varepsilon\big)\Big|$$
$$\lesssim \|f\|_{\mathrm{L}^\infty} \Big(\sum_{m\in\mathbb{Z}^2} \|g_m *_1 \phi_\varepsilon\|_{\mathrm{H}^{-\sigma,0}}^2\Big)^{1/2} \Big(\sum_{m\in\mathbb{Z}^2} \|h_m - h_m *_2 \phi_\varepsilon\|_{\mathrm{H}^{0,-\sigma}}^2\Big)^{1/2}$$
$$\lesssim \|f\|_{\mathrm{L}^\infty} \Big(\sum_{m\in\mathbb{Z}^2} \|g_m\|_{\mathrm{L}^2}^2\Big)^{1/2} \Big(\varepsilon^{2\sigma} \sum_{m\in\mathbb{Z}^2} \|h_m\|_{\mathrm{L}^2}^2\Big)^{1/2},$$

where we used

$$\|h_m - h_m *_2 \phi_\varepsilon\|_{\mathrm{H}^{0,-\sigma}} \lesssim \varepsilon^\sigma \|h_m\|_{\mathrm{L}^2}$$

and (2.10). □

The key point in the above proof is that the global Estimate (2.9) was reduced to the local estimate from Lemma 5 via almost orthogonality (2.10). We have not seen such a sharp localization argument in previous literature on point configurations, possibly because the target estimate is rarely independent of λ. In our case this owes to the fact that the anisotropic dilation $(x,y) \mapsto (\lambda x, y/\lambda)$ is area-preserving.

2.4. **Proof completion.**

Proof of the upper estimate in Theorem 1. Let

$$\delta := \frac{|A|}{R^2} \in (0,1]$$

and suppose that $A \subseteq [0,R]^2$ does not contain the pattern (1.1) for any $x, y \in \mathbb{R}$, $t > 0$. Fix the exponent $\sigma > 0$ from Lemma 6. Choose

$$\varepsilon := c_0 \delta^{2/\sigma},$$

where $c_0 \in (0,1]$ will be chosen later. By (2.3) and Lemma 6,

$$\mathcal{N}_\lambda^\varepsilon(A) = \mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^0(A) \lesssim R^2\delta\varepsilon^\sigma \leqslant \frac{1}{2} c_1 R^2 \delta^3$$

for every $\lambda > 0$, provided c_0 was chosen sufficiently small, where $c_1 > 0$ is the constant from Lemma 3, i.e.,

$$\mathcal{N}_\lambda^1(A) \geqslant c_1 R^2 \delta^3$$

for every $\lambda \in [1/R, R]$. In particular,

$$\mathcal{N}_\lambda^1(A) - \mathcal{N}_\lambda^\varepsilon(A) \gtrsim R^2\delta^3.$$

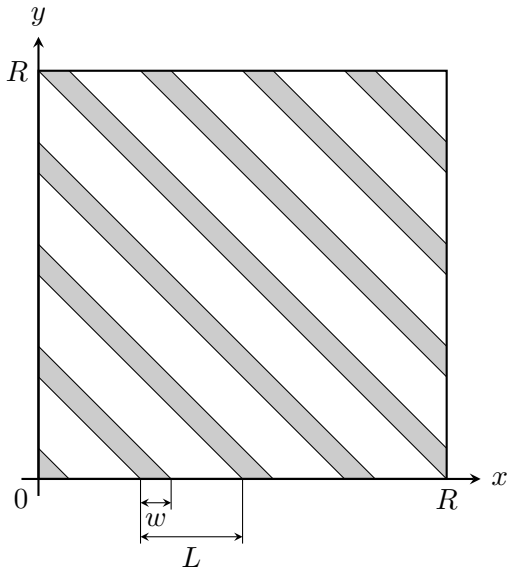

FIGURE 1. Example for the lower bound.

We square this and integrate over $[1/R, R]$ with respect to $\mathrm{d}\lambda/\lambda$, so that Lemma 4 then gives

$$R^4\delta^6 \log R \lesssim \int_{1/R}^{R} \left|\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A)\right|^2 \frac{d\lambda}{\lambda} \lesssim R^4\delta^2 \log\frac{1}{\varepsilon}.$$

By the choice of ε we have obtained

$$\delta^4 \log R \lesssim 1 + \log\frac{1}{\delta},$$

i.e.,

$$\frac{1}{\delta^4}\Big(1 + \log\frac{1}{\delta}\Big) \gtrsim \log R.$$

This implies[2]

$$\frac{1}{\delta} \gtrsim \Big(\frac{\log R}{\log\log R}\Big)^{1/4},$$

which proves the more difficult half of Theorem 1. □

Remark 7. Note that Lemmas 4 and 6 did not use any information about the localization of the set A. For instance, the proof of latter lemma owes this to the fact that transformations $(x, y) \mapsto (\lambda x, y/\lambda)$ are area-preserving, so one can in fact turn attention to a single value of λ at a time, and all underlying estimates are exactly the same. An ultimate consequence is that we are able to obtain a better upper bound (1.2) than the doubly-logarithmic bounds in [7] and [11]. The only place where the concentration of A on $[0, R]^2$ was, in fact, needed is the lower bound in Lemma 3.

3. Proof of the lower bound in Theorem 1

Proof of the lower estimate in Theorem 1. The set A is defined to be the intersection of the square $[0,R]^2$ with equidistant bands,

$$A := \Big\{(x,y)\in[0,R]^2 \,:\, x+y \in \bigcup_{j\in\mathbb{Z}} (jL, jL+w)\Big\}.$$

It is illustrated schematically in Figure 1. Here, we put $N = \lceil R^{2/3}\rceil$ and choose parameters L, w that determine the band spacing and their common width in the $x+y$ coordinate by

$$L - w = \frac{1}{\sqrt{N+1}}, \quad L + w = \frac{1}{\sqrt{N}}.$$

More explicitly,

$$w = \frac{1}{2\sqrt{N}\sqrt{N+1}(\sqrt{N}+\sqrt{N+1})} \sim R^{-1}, \quad L = \frac{\sqrt{N}+\sqrt{N+1}}{2\sqrt{N}\sqrt{N+1}} \sim R^{-1/3}.$$

Also note that $w < N^{-3/2} \leqslant R^{-1}$.

Suppose that A contains a hyperbolic corner

$$(x,y),\ (x+t,y),\ (x,y+t^{-1}), \qquad t>0.$$

Its two leg lengths t, t^{-1} belong to $[R^{-1}, R]$, so they both exceed w. Consequently, the first two points cannot lie in the same band, and the same is true for the first and the third point. Differences of the corresponding band indices therefore give positive integers m, n such that

$$|t - mL| < w, \quad |t^{-1} - nL| < w.$$

Since $m, n \geqslant 1$ and $L \pm w$ have the values above, this further implies

$$m(L-w) < t < m(L+w), \quad n(L-w) < t^{-1} < n(L+w),$$

i.e.,

$$\frac{m}{\sqrt{N+1}} < t < \frac{m}{\sqrt{N}}, \quad \frac{n}{\sqrt{N+1}} < t^{-1} < \frac{n}{\sqrt{N}}.$$

Multiplication gives $N < mn < N+1$, which is impossible.

By L-periodicity, each horizontal section of A has length at least $\lfloor R/L\rfloor w$. Integrating these lengths gives

$$|A| \geqslant R\Big\lfloor \frac{R}{L}\Big\rfloor w \gtrsim \frac{R^2 w}{L} \sim R^{4/3}. \qquad \square$$

4. Proof of Theorem 2

Proof of Theorem 2. We prove $|A| \lesssim R$ for every $R>0$ and every measurable set $A \subseteq [0,R]^2$ containing no triple of points spanning a triangle of area $1/2$. Namely, dilation by $2^{-1/2}$ then deduces the same estimate (up to a different constant) for sets avoiding triangles of area 1. By inner regularity of Lebesgue measure, it suffices to consider compact sets A. We write $f := \mathbb{1}_A$.

Let $\mathcal{R}_\theta$ denote the planar rotation about the origin through angle θ,

$$\mathcal{R}_\theta(x,y) := (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta),$$

[2] If $x^4(1+\log x) = y$ is solved for $x>0$ when $y>0$ is sufficiently large, then $x = ((4+o(1))y/\log y)^{1/4}$, $y \to +\infty$.

and write $G_\theta(y)$ for the length of the horizontal section of $\mathcal{R}_\theta A$ at height y,

$$G_\theta(y) := \int_{\mathbb{R}} \mathbb{1}_{\mathcal{R}_\theta A}(x,y)\,\mathrm{d}x.$$

For every θ, the rotated set $\mathcal{R}_\theta A$ again has measure $|A|$ and still avoids triangles of area $1/2$. Moreover, for each θ, the function G_θ is supported in an interval of length at most $2^{1/2}R$, so Hölder's inequality gives

$$|A|^3 = |\mathcal{R}_\theta A|^3 = \Big(\int_{\mathbb{R}} G_\theta(y)\,\mathrm{d}y\Big)^3 \leqslant 2R^2 \int_{\mathbb{R}} G_\theta(y)^3\,\mathrm{d}y. \tag{4.1}$$

Fix $\theta \in [0,2\pi)$ and a height $y \in \mathbb{R}$ at which the section of $\mathcal{R}_\theta A$ is nonempty. By compactness it has a leftmost point (x_0,y). The set of $t>0$ for which $(x_0+t,y) \in \mathcal{R}_\theta A$ has one-dimensional Lebesgue measure $G_\theta(y)$. For each such t, the horizontal line at height $y+t^{-1}$ must miss $\mathcal{R}_\theta A$, since any point on that line would form a triangle of area $1/2$ with (x_0,y) and (x_0+t,y). Thus $G_\theta(y+t^{-1})=0$ for all such t, and consequently

$$\begin{aligned} G_\theta(y)^3 &= \int_0^\infty G_\theta(y)^2 \mathbb{1}_{\mathcal{R}_\theta A}(x_0+t,y)\,\mathrm{d}t \\ &= \int_0^\infty \big|G_\theta(y+t^{-1}) - G_\theta(y)\big|^2 \mathbb{1}_{\mathcal{R}_\theta A}(x_0+t,y)\,\mathrm{d}t \\ &\leqslant \int_0^\infty \big|G_\theta(y+t^{-1}) - G_\theta(y)\big|^2\,\mathrm{d}t. \end{aligned} \tag{4.2}$$

The same inequality is trivial when the section is empty. Recalling (4.1), integrating (4.2) in y, and applying Fubini's and Plancherel's theorems gives

$$\begin{aligned} |A|^3 &\lesssim R^2 \int_0^\infty \int_{\mathbb{R}} |\widehat{G_\theta(\cdot + t^{-1})}(\eta) - \widehat{G_\theta}(\eta)|^2\,\mathrm{d}\eta\,\mathrm{d}t \\ &= R^2 \int_{\mathbb{R}} |\widehat{G_\theta}(\eta)|^2 \Big(\int_0^\infty \big|e^{2\pi i \eta/t} - 1\big|^2\,\mathrm{d}t\Big)\,\mathrm{d}\eta. \end{aligned} \tag{4.3}$$

By Fubini's theorem and the change of variables $(x,y) = \mathcal{R}_\theta(u,v)$,

$$\begin{aligned} \widehat{G_\theta}(\eta) &= \int_{\mathbb{R}^2} \mathbb{1}_{\mathcal{R}_\theta A}(x,y) e^{-2\pi i y\eta}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^2} f(u,v) e^{-2\pi i (u\sin\theta + v\cos\theta)\eta}\,\mathrm{d}u\,\mathrm{d}v = \widehat{f}(\eta\sin\theta, \eta\cos\theta). \end{aligned}$$

We also have the elementary estimate

$$\int_0^\infty \big|e^{2\pi i\eta/t} - 1\big|^2\,\mathrm{d}t \lesssim \int_0^\infty \min\Big\{1, \frac{\eta^2}{t^2}\Big\}\,\mathrm{d}t \lesssim |\eta|$$

for every $\eta \in \mathbb{R}$, so that (4.3) becomes

$$|A|^3 \lesssim R^2 \int_{\mathbb{R}} |\widehat{f}(\eta\sin\theta, \eta\cos\theta)|^2 |\eta|\,\mathrm{d}\eta.$$

Averaging this over θ and integrating in polar coordinates we then obtain

$$|A|^3 \lesssim \frac{R^2}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} |\widehat{f}(\eta\sin\theta, \eta\cos\theta)|^2 |\eta|\,\mathrm{d}\eta\,\mathrm{d}\theta = \frac{R^2}{\pi} \big\|\widehat{f}\big\|_{\mathrm{L}^2}^2 = \frac{R^2}{\pi} \|f\|_{\mathrm{L}^2}^2 \lesssim R^2 |A|.$$

This proves $|A| \lesssim R$, establishing the theorem. □

Appendix A. Proof of the hyperbolic trilinear smoothing

The main part of the proof of Lemma 5 can be derived from the existing literature. A rather general smoothing theorem for fiberwise bilinear multipliers was sketched by Hsu and Lin [25].[3] However, we will prefer to use a less general formulation of this result from Lin's dissertation [38], which contains a very detailed proof in the form of a blueprint for a future Lean formalization. Alternative proofs of Lemma 5 could be worked out by closely following the proofs of the trilinear smoothing inequalities by Christ, Durcik, and Roos [11] and Gaitan and Lie [21], on which the proof in [38] is based.

Set

$$T(f_1,f_2)(x,y) := \int_{\mathbb{R}} f_1(x+t,y) f_2\Big(x, y+\frac{1}{t}\Big)\zeta(t)\,\mathrm{d}t,$$

so that the desired estimate reads

$$\|T(f_1,f_2)\|_{\mathrm{L}^1} \lesssim \|f_1\|_{\mathrm{H}^{-\sigma,0}} \|f_2\|_{\mathrm{H}^{0,-\sigma}} \tag{A.1}$$

for an appropriately small $\sigma > 0$ and f_1, f_2 supported in $[-10,10]^2$.

The fiberwise multiplier formulation of T is

$$T(f_1,f_2)(x,y) = \int_{\mathbb{R}^2} (\mathcal{F}_1 f_1)(\xi,y)(\mathcal{F}_2 f_2)(x,\eta) m(\xi,\eta) e^{2\pi i(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta,$$

where

$$m(\xi,\eta) := \int_{\mathbb{R}} e^{2\pi i(t\xi+\eta/t)}\zeta(t)\,\mathrm{d}t.$$

We localize the multiplier in frequency. Let $\varphi \in C^\infty(\mathbb{R})$ be a function such that $\mathbb{1}_{[-1,1]} \leqslant \varphi \leqslant \mathbb{1}_{[-2,2]}$. Let $\widetilde{\varphi}_0 := \varphi$ and for $j \geqslant 1$ let $\widetilde{\varphi}_j(u) = \varphi(2^{-j}u) - \varphi(2^{-j+1}u)$. Let $S(u) := \sum_{j\geqslant 0} \widetilde{\varphi}_j^2(u)$. At each point $u \in \mathbb{R}$, the sum defining $S(u)$ has at most two nonzero elements, so $S(u) \geqslant \frac{1}{2}$. Therefore, if we define

$$\psi_j(u) := \frac{\widetilde{\varphi}_j(u)}{S(u)^{1/2}},$$

then $\psi_j \in C_c^\infty(\mathbb{R})$ for every $j \geqslant 0$,

$$\operatorname{supp}\psi_0 \subseteq [-2,2], \quad \operatorname{supp}\psi_j \subseteq 2^j([-2,-1/2]\cup[1/2,2]), \quad j \geqslant 1,$$

and

$$\sum_{j\geqslant 0} \psi_j^2 \equiv 1.$$

For $j,k \in \mathbb{N}_0$, let

$$m_{j,k}(\xi,\eta) := \psi_j^2(\xi)\psi_k^2(\eta) \int_{\mathbb{R}} e^{2\pi i(\xi t+\eta/t)}\zeta(t)\,\mathrm{d}t$$

and define

$$T_{j,k}(f_1,f_2)(x,y) := \int_{\mathbb{R}^2} (\mathcal{F}_1 f_1)(\xi,y)(\mathcal{F}_2 f_2)(x,\eta) m_{j,k}(\xi,\eta) e^{2\pi i(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta.$$

The required Estimate (A.1) will follow from the following two lemmas.

[3] General corner-type trilinear smoothing estimates were shown by Hsu and Lin [26] after the first preprint version of the present paper was announced. Lemma 5 can now be deduced more directly from their [26, Thm. 1.1].

Lemma 8 (Non-stationary part)**.** *Let $j, k \in \mathbb{N}_0$ be integers such that $|j-k| > 10$. Then the estimate*

$$\|T_{j,k}(f_1, f_2)\|_{\mathrm{L}^1} \lesssim_N 2^{-N\max\{j,k\}} \|f_1\|_{\mathrm{L}^2} \|f_2\|_{\mathrm{L}^2}$$

holds for every $N > 0$.

Proof. The proof follows the proof of a single scale version of the high–low Coifman–Meyer paraproduct estimate.

Let $\Phi_{\xi,\eta}(t) = \xi t + \eta/t$. When $|j-k| > 10$, it follows that for $t \in \operatorname{supp}\zeta \subset [\frac{1}{2}, 2]$, $\xi \in \operatorname{supp}\psi_j$ and $\eta \in \operatorname{supp}\psi_k$ the following estimate holds

$$|\Phi'_{\xi,\eta}(t)| = \left|\xi - \frac{\eta}{t^2}\right| \gtrsim 2^{\max\{j,k\}}. \tag{A.2}$$

Repeated integration by parts with

$$L_{\xi,\eta} := \frac{1}{2\pi i\, \Phi'_{\xi,\eta}(t)} \frac{\mathrm{d}}{\mathrm{d}t}$$

gives, for every fixed finite range of α, β and every $N > 0$,

$$\left|\partial_\xi^\alpha \partial_\eta^\beta m(\xi,\eta)\right| \lesssim_{N,\alpha,\beta} 2^{-N\max\{j,k\}}. \tag{A.3}$$

Indeed, differentiating the integral in ξ and η only inserts powers of t and t^{-1} with $\zeta(t)$ and the integration by parts gives the required decay. Since $(\xi,\eta) \mapsto m(\xi,\eta)\psi_j(\xi)\psi_k(\eta)$ is a smooth function supported in a $[-2^{\max\{j,k\}+1}, 2^{\max\{j,k\}+1}]^2$ square (which has side length $2^{\max\{j,k\}+2}$), using classical results for Fourier series, it follows that

$$m(\xi,\eta)\psi_j(\xi)\psi_k(\eta) = \sum_{n_1,n_2\in\mathbb{Z}} c_{n_1,n_2} e^{2\pi i(n_1\xi + n_2\eta)/2^{\max\{j,k\}+2}},$$

where the coefficients, because of Estimate (A.3), satisfy

$$|c_{n_1,n_2}| \lesssim_\varphi 2^{-N\max\{j,k\}} (1 + |n_1| + |n_2|)^{-10}.$$

Therefore, using Cauchy–Schwarz and Young convolution inequalities, we have

$$\begin{aligned}
&\|T_{j,k}(f_1,f_2)\|_{\mathrm{L}^1} \\
&\quad\leqslant \sum_{n_1,n_2\in\mathbb{Z}} |c_{n_1,n_2}| \left\| f_1 *_1 \psi_j^\vee(x + n_1/2^{\max\{j,k\}+2}, y) f_2 *_2 \psi_k^\vee(x, y + n_2/2^{\max\{j,k\}+2}) \right\|_{\mathrm{L}^1} \\
&\quad\lesssim_\varphi 2^{-N\max\{j,k\}} \sum_{n_1,n_2\in\mathbb{Z}} (1 + |n_1| + |n_2|)^{-10} \|f_1\|_{\mathrm{L}^2} \|f_2\|_{\mathrm{L}^2} \\
&\quad\lesssim 2^{-N\max\{j,k\}} \|f_1\|_{\mathrm{L}^2} \|f_2\|_{\mathrm{L}^2}.
\end{aligned}$$

□

Lemma 9 (Stationary part)**.** *There exists a constant $c > 0$, depending only on ζ and ψ in the definition of $m_{j,k}$, such that the following holds for every $j, k \in \mathbb{N}_0$ such that $|j-k| \leqslant 10$:*

$$\|T_{j,k}(f_1,f_2)\|_{\mathrm{L}^1([0,1]^2)} \lesssim_{\zeta,\psi} 2^{-c\max\{j,k\}} \|f_1\|_{\mathrm{L}^2} \|f_2\|_{\mathrm{L}^2}.$$

Proof. We claim that this is essentially what Theorem [38, Ch. 3, Thm. 3.1.12] claims for $\gamma(t) = -1/t$, but a few comments are due. First, we observe that

$$\gamma'(t) = \frac{1}{t^2}, \quad \gamma''(t) = -\frac{2}{t^3}, \quad \theta(t) := \frac{\gamma'(t)}{\gamma''(t)} = -\frac{t}{2},$$

and hence

$$\theta'(t) = -\frac{1}{2}, \quad \theta(t)\theta''(t) - \theta'(t)(1 + \theta'(t)) = \frac{1}{4},$$

so the quantitative nondegeneracy hypotheses in [38, Thm. 3.1.12] hold.

Theorem [38, Thm. 3.1.12] is stated only for the case $j = k$ with the functions ψ_j having support in smaller intervals (not annuli) $2^j[1-\delta, 1+\delta]$ for $\delta > 0$ sufficiently small depending on γ. The assumption that $\delta > 0$ is small enough is imposed because in the statement of theorem [38, Thm. 3.1.12], the function γ is defined only on $[1/2, 3/2]$ and they have to make sure that the stationary point of the phase Φ lies in that interval so that the decomposition of the multiplier in Proposition [38, Prop. 3.1.17], which uses the method of stationary phase, holds true.

In our case, the function $\gamma = -1/t$ is smooth on $(0,\infty)$ and satisfies the assumption

$$\inf_{t\in I} |\gamma''(t)| \neq 0$$

for any compact interval $I \subset (0,\infty)$. The choice $I = [2^{-6}, 2^6]$ will satisfy the requirements.

If $j, k \leqslant 20$, the statement trivially follows from Cauchy–Schwarz and Young convolution inequalities as in the proof of Lemma 8.[4]

For $\max\{j,k\} \geqslant 20$, the operators associated with

$$m_{j,k}(\xi,\eta)\mathbb{1}_{(0,\infty)}(\xi)\mathbb{1}_{(-\infty,0)}(\eta) \quad \text{and} \quad m_{j,k}(\xi,\eta)\mathbb{1}_{(-\infty,0)}(\xi)\mathbb{1}_{(0,\infty)}(\eta)$$

can be handled using the same method as in Lemma 8. Indeed, the phase satisfies (A.2), allowing us to perform integration by parts.

Finally, to handle operators associated with

$$m_{j,k}(\xi,\eta)\mathbb{1}_{(0,\infty)}(\xi)\mathbb{1}_{(0,\infty)}(\eta) \quad \text{and} \quad m_{j,k}(\xi,\eta)\mathbb{1}_{(-\infty,0)}(\xi)\mathbb{1}_{(-\infty,0)}(\eta),$$

we observe that since $|j-k| \leqslant 10$, the stationary point of the phase Φ satisfies $t_0 = \sqrt{\eta/\xi} \in [2^{-6}, 2^6]$. Therefore, the proof of [38, Theorem 3.1.12] with the given choice of I gives the desired bound. □

Proof of Lemma 5. As we have already observed, it remains to prove estimate (A.1).

Let $\widetilde{\psi}_0 \in C_c^\infty(\mathbb{R})$ be a function that is equal to 1 on the support of ψ_0, let $\widetilde{\psi}_1 \in C^\infty(\mathbb{R})$ be a function supported in $[-4,-1/4]\cup[1/4,4]$ that is equal to 1 on $\operatorname{supp}\psi_1 \subset [-2,-1/2]\cup[1/2,2]$ and let $\widetilde{\psi}_j(t) = \widetilde{\psi}_1(2^{-j}t)$. For $j \in \mathbb{N}_0$, let $P_j^{(1)}$ be the operator defined with

$$\big(P_j^{(1)} f\big)(x,y) = \mathcal{F}_1^{-1}\big((\widetilde{\psi}_j \otimes \mathbb{1}_{\mathbb{R}})(\mathcal{F}_1 f)\big)(x,y) = \int_{\mathbb{R}} \widetilde{\psi}_j(\xi)(\mathcal{F}_1 f)(\xi,y)e^{2\pi i\xi x}\,\mathrm{d}\xi$$

and similarly let $P_j^{(2)}$ be the operator defined with

$$\big(P_j^{(2)} f\big)(x,y) = \mathcal{F}_2^{-1}\big((\mathbb{1}_{\mathbb{R}} \otimes \widetilde{\psi}_j)(\mathcal{F}_2 f)\big)(x,y) = \int_{\mathbb{R}} \widetilde{\psi}_j(\eta)(\mathcal{F}_2 f)(x,\eta)e^{2\pi i\eta y}\,\mathrm{d}\eta.$$

Then, because of the support assumption, we conclude that

$$T_{j,k}(f_1,f_2) = T_{j,k}(P_j^{(1)} f_1, P_k^{(2)} f_2).$$

Observe that for functions f_1, f_2 supported on $[-10,10]^2$, we have $\operatorname{supp} T(f_1,f_2) \subset [-20,20]^2$. Therefore, by partitioning $[-20,20]^2$ into $O(1)$ squares of sidelength 1, it is sufficient to bound $\|T(f_1,f_2)\|_{L^1(Q)}$ for a square Q of sidelength 1. Without loss of generality, we can assume that $Q = [0,1]^2$.

[4] We prove this case separately to avoid the technicality that for $j = 0$ or $k = 0$, the truncated function $\psi_0\mathbb{1}_{(0,\infty)}$, which we use in the rest of the proof is no longer smooth.

Using the multiplier decomposition, previous observations, Lemmas 8 and 9, and the fact that $2\max\{j,k\} \geqslant j+k$, there exists $c>0$ such that the following chain of inequalities is true

$$\begin{aligned}\|T(f_1,f_2)\|_{\mathrm{L}^1([0,1]^2)} &\leqslant \sum_{j,k\geqslant 0} \|T_{j,k}(f_1,f_2)\|_{\mathrm{L}^1([0,1]^2)} \\ &= \sum_{j,k\geqslant 0} \|T_{j,k}(P_j^{(1)}f_1, P_k^{(2)}f_2)\|_{\mathrm{L}^1([0,1]^2)} \\ &\lesssim_{\zeta,\varphi} \sum_{j,k\geqslant 0} 2^{-c(j+k)} \|P_j^{(1)}f_1\|_{\mathrm{L}^2}\|P_k^{(2)}f_2\|_{\mathrm{L}^2} \\ &= \Big(\sum_{j\geqslant 0} 2^{-cj}\|P_j^{(1)}f_1\|_{\mathrm{L}^2}\Big)\Big(\sum_{k\geqslant 0} 2^{-ck}\|P_k^{(2)}f_2\|_{\mathrm{L}^2}\Big).\end{aligned}$$

Finally, using the Cauchy–Schwarz inequality and the Littlewood–Paley characterization of the Sobolev norm, we have

$$\sum_{j\geqslant 0} 2^{-cj}\|P_j^{(1)}f_1\|_{\mathrm{L}^2} \leqslant \underbrace{\Big(\sum_{j\geqslant 0} 2^{-cj}\Big)^{1/2}}_{\lesssim 1}\Big(\sum_{j\geqslant 0} 2^{-cj}\|P_j^{(1)}f_1\|_{\mathrm{L}^2}^2\Big)^{1/2} \lesssim_c \|f_1\|_{\mathrm{H}^{-c/2,0}}.$$

Using the analogous inequality for f_2, we conclude the proof. □

Appendix B. Graham's argument

The following lemma is the most likely reconstruction of Graham's comment in [22] based on Szemerédi's theorem.

Lemma 10. *For every $\beta \in (0,1]$, there exist a positive integer n and a number $T>0$ such that every set $B \subseteq \{0,1,\dots,n-1\}^2$ with at least βn^2 elements contains three points spanning a triangle of area exactly T.*

Proof. Take a positive integer $r \geqslant 4/\beta$. By Szemerédi's theorem there exists a positive integer N such that every subset of $\{0,1,\dots,N-1\}$ of size at least $(\beta/2)N$ contains a non-degenerate arithmetic progression of length $r!+1$. We choose the integer T mentioned in the lemma formulation to be $T := r!N!/2$. Finally, let n be a multiple of N so large that $n \geqslant (r+1)N!$.

Let $B \subseteq \{0,1,\dots,n-1\}^2$ satisfy $|B| \geqslant \beta n^2$. A row

$$\big\{(x,y) : x \in \{0,1,\dots,n-1\}\big\}$$

indexed by $y \in \{0,1,\dots,n-1\}$ is said to be *populated* if at least $(\beta/2)n$ of its n elements belong to B. There are at least $(\beta/2)n$ populated rows, since otherwise

$$|B| \leqslant \frac{\beta}{2}n^2 + \Big(n - \frac{\beta}{2}n\Big)\frac{\beta}{2}n < \beta n^2,$$

which would be in a contradiction with our assumption on the size of B. Partition all rows from the $n\times n$ grid into n/N strips of height N. At least one strip contains at least $(\beta/2)N$ populated rows. By the choice of N according to Szemerédi's theorem, there are $r!+1$ equally spaced populated rows inside that strip; let them be determined with their vertical coordinates

$$y_0,\ y_0+l,\ y_0+2l,\ \dots,\ y_0+r!l. \tag{B.1}$$

We define $k := N!/l \in \mathbb{N}$.

Now consider the lowest populated row, the one indexed by y_0: the set B occupies at least $(\beta/2)n$ of its points. Consider all arithmetic progressions of length $r+1$,

$$x_0,\ x_0+k,\ x_0+2k,\dots,\ x_0+rk,$$

that fully belong to $\{0,1,\dots,n-1\}$. We claim that at least one of these progressions contains at least two numbers from $\{x : (x,y_0)\in B\}$. If that was not the case, then we would fully cover (with possible overlaps) $\{0,1,\dots,n-1\}$ by $\lceil n/(r+1)\rceil$ such arithmetic progressions to conclude that the row indexed by y_0 contains at most

$$\left\lceil \frac{n}{r+1}\right\rceil < \frac{2n}{r} \leqslant \frac{\beta}{2}n$$

points from B, which contradicts the choice of y_0. Choose such a progression and such two numbers x_0+ik, x_0+jk for some integers x_0 and $0\leqslant i<j\leqslant r$. Thus,

$$(x_0+ik,y_0),\ (x_0+jk,y_0)\in B. \tag{B.2}$$

The number

$$y_1 = y_0 + \frac{r!}{j-i}l$$

is one of the numbers (B.1), so we can find a point

$$(x_1,y_1)\in B. \tag{B.3}$$

The three points (B.2), (B.3) determine a triangle of area

$$\frac{1}{2}(j-i)k\frac{r!}{j-i}l = \frac{1}{2}r!N! = T. \qquad \square$$

Note that we could not have been very picky about the area value T in the previous lemma. Namely, it had to satisfy certain divisibility properties: if $2T$ had too few divisors, then it would not allow us to consider sufficiently many potential triangles with vertices on the integer lattice $\mathbb{Z}^2$. However, the actual values of n and T will not matter when we pass from the discrete setting to $\mathbb{R}^2$.

We briefly argue how Lemma 10 gives an alternative proof of (1.5), as this transference is easy and standard. Namely, fix $\delta\in(0,1]$. Apply Lemma 10 with $\beta=\delta/2$, which then gives $n\in\mathbb{N}$ and $T>0$ with the stated properties. We will show that, for every $R\geqslant 8nT^{-1/2}\delta^{-1}$, every subset $A\subseteq[0,R]^2$ of measure greater than δR^2 contains a triple of points that span a triangle of area 1. Namely, for every

$$(u,v)\in Q := \Big[0, R-\frac{n-1}{T^{1/2}}\Big]^2$$

define

$$B_{u,v} := \big\{(k,l)\in\{0,1,\dots,n-1\}^2 : (u,v)+T^{-1/2}(k,l)\in A\big\}.$$

Then

$$\fint_Q |B_{u,v}|\,\mathrm{d}u\,\mathrm{d}v \geqslant \frac{1}{R^2}n^2\Big|A\cap\Big[\frac{n}{T^{1/2}},R-\frac{n}{T^{1/2}}\Big]^2\Big| > \frac{\delta}{2}n^2,$$

so there exist u,v such that $|B_{u,v}|>(\delta/2)n^2$. From Lemma 10 we know that $B_{u,v}$ contains vertices of a triangle with area T, whence A contains vertices of a triangle with area 1. Consequently, $M_\Delta(R)\leqslant\delta R^2$.

Remark 11. The latter proof gives quantitatively very weak upper bounds on $M_\Delta(R)$ in comparison with our earlier results. Namely, quantitative Szemerédi's theorem is already insufficient to prove reasonable bounds. For example, the number N used in the proof by Leng, Sah, and Sawhney [36] grows like

$$N \geqslant \exp\Big(\exp\Big((\log(2/\beta))^{C(\lceil 4/\beta\rceil!+1)}\Big)\Big),$$

while the dependence of the constants $C(k)$ on the progression length k is not even easy to track down explicitly.

Declaration of AI usage

Several large language models were used as follows.

- First, OpenAI's ChatGPT-5.4 Pro improved the trivial lower bound $M(R) \gtrsim R$ to $M(R) \gtrsim R\log R$ by constructing a set A as a union of bands of decreasing lengths. This example was used in the first version of this paper. Afterwards, ChatGPT-6 Astra improved it further to $M(R) \gtrsim R^{4/3}$ by carefully choosing a union of equidistant but much narrower bands. This is the proof that we worked out in Section 3.
- First, the authors found an improvement over Theorem 1 for general triangles,

  $$M_\Delta(R) \lesssim R^2\Big(\frac{\log\log R}{\log R}\Big)^{1/3},$$

  by introducing the Riesz energy of the set A and applying the Hardy–Littlewood–Sobolev inequality. Then ChatGPT-5.5 Pro drafted a further improvement

  $$M_\Delta(R) \lesssim R^2\Big(\frac{\log\log R}{\log R}\Big)^{1/2},$$

  which additionally used the induction on scales R. Finally, ChatGPT-6 Astra gave a substantially different and shorter argument, leading to the current version of Theorem 2. We reworked it and simplified (using rotations) in Section 4.
- ChatGPT-5.5 Pro was also used to clarify Graham's cryptic remark and reconstruct its intended proof.
- Google's Gemini 3.1 Pro and ChatGPT-6 Astra were used to draw Figure 1.

The remaining ideas, the proofs, and the actual writing of the manuscript are entirely the work of the authors, who assume full responsibility for the content.

Acknowledgments and funding

We are grateful to Michael Christ for several useful discussions, to Adrian Beker and Polona Durcik for pointing us to the relevant literature, and to Fred Lin for providing a copy of the doctoral dissertation [38].

This work was supported in part by the Croatian Science Foundation under the project HRZZ-IP-2022-10-5116 (*FANAP*). The paper was also supported in part by the European Union – NextGenerationEU through the National Recovery and Resilience Plan 2021–2026, via an institutional grant of the University of Zagreb Faculty of Science, IK IA 1.1.3, *Impact4Math*.

References

[1] Miklós Ajtai and Endre Szemerédi. Sets of lattice points that form no squares. *Studia Sci. Math. Hungar.*, 9:9–11, 1974.

[2] 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.

[3] Vitaly Bergelson, Bernard Host, Randall McCutcheon, and François Parreau. Aspects of uniformity in recurrence. *Colloq. Math.*, 84/85:549–576, 2000. Dedicated to the memory of Anzelm Iwanik. doi:10.4064/cm-84/85-2-549-576.

[4] Vitaly Bergelson and Alexander Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996. doi:10.1090/S0894-0347-96-00194-4.

[5] Thomas Bloom. Erdős problems. https://www.erdosproblems.com/. Accessed: January 25, 2026.

[6] 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.

[7] Jean Bourgain. A nonlinear version of Roth's theorem for sets of positive density in the real line. *J. Analyse Math.*, 50:169–181, 1988. doi:10.1007/BF02796120.

[8] Benjamin B. Bruce and Malabika Pramanik. Two-point patterns determined by curves. *Math. Ann.*, 393(1):571–615, 2025. doi:10.1007/s00208-025-03254-y.

[9] Xuezhi Chen and Jingwei Guo. A polynomial Roth theorem for corners in $\mathbb{R}^2$ and a related bilinear singular integral operator. *Math. Ann.*, 390(1):255–301, 2024. doi:10.1007/s00208-023-02763-y.

[10] Xuezhi Chen and Changxing Miao. Two-point polynomial patterns in subsets of positive density in $\mathbb{R}^n$. *Int. Math. Res. Not. IMRN*, 2024(14):10865–10879, 2024. doi:10.1093/imrn/rnae108.

[11] Michael Christ, Polona Durcik, and Joris Roos. Trilinear smoothing inequalities and a variant of the triangular Hilbert transform. *Adv. Math.*, 390:Paper No. 107863, 60, 2021. doi:10.1016/j.aim.2021.107863.

[12] 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.

[13] 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.

[14] Dong Dong, Xiaochun Li, and Will Sawin. Improved estimates for polynomial Roth type theorems in finite fields. *J. Anal. Math.*, 141(2):689–705, 2020. doi:10.1007/s11854-020-0113-8.

[15] 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.

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

[17] Polona Durcik, Vjekoslav Kovač, and Mario Stipčić. A strong-type Furstenberg-Sárközy theorem for sets of positive measure. *J. Geom. Anal.*, 33(8):Paper No. 255, 16, 2023. doi:10.1007/s12220-023-01309-7.

[18] 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.

[19] Paul Erdős. My Scottish Book 'Problems'. In R. Daniel Mauldin, editor, *The Scottish Book*, pages 27–33. Birkhäuser, Boston, 1981.

[20] 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.

[21] Alejandra Gaitan and Victor Lie. Non-zero to zero curvature transition: operators along hybrid curves with no quadratic (quasi-)resonances. *Adv. Math.*, 478:Paper No. 110356, 123 pp., 2025. doi:10.1016/j.aim.2025.110356.

[22] 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.

[23] Jingwei Guo, Changxing Miao, and Guoqing Zhan. A multidimensional Szemerédi theorem in integers. Preprint, 2026. URL: https://arxiv.org/abs/2605.06360.

[24] Rui Han, Michael T. Lacey, and Fan Yang. A polynomial Roth theorem for corners in finite fields. *Mathematika*, 67(4):885–896, 2021. doi:10.1112/mtk.12108.

[25] Martin Hsu and Fred Yu-Hsiang Lin. A short proof on the boundedness of triangular Hilbert transform along curves. Preprint, arXiv:2410.15791v1, 2024. URL: https://arxiv.org/abs/2410.15791v1.

[26] Martin Hsu and Fred Yu-Hsiang Lin. Smoothing inequalities for corner-type bilinear averages: geometric characterization and applications. Preprint, arXiv:2410.15791v2, 2026. URL: https://arxiv.org/abs/2410.15791v2.

[27] Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni, and Mehtaab Sawhney. Quasipolynomial bounds for the corners theorem. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science—FOCS 2025*, pages 2027–2036. IEEE Comput. Soc. Press, Los Alamitos, CA, 2025. doi:10.1109/FOCS63196.2025.00108.

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

[29] 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.*, pages 1431–1434. EMS Press, EMS-Publishing House GmbH, 2023. doi:10.4171/OWR/2023/25.

[30] Vjekoslav Kovač. Coloring and density theorems for configurations of a given volume. *Proc. Lond. Math. Soc. (3)*, 132(3):Paper No. e70143, 56 pp, 2026. doi:10.1112/plms.70143.

[31] Ben Krause, Mariusz Mirek, Sarah Peluse, and James Wright. Polynomial progressions in topological fields. *Forum Math. Sigma*, 12:Paper No. e106, 51, 2024. doi:10.1017/fms.2024.104.

[32] Noah Kravitz, Borys Kuca, and James Leng. Corners with polynomial side length. Preprint, 2024. URL: https://arxiv.org/abs/2407.08637.

[33] Borys Kuca. Multidimensional polynomial patterns over finite fields: bounds, counting estimates and Gowers norm control. *Adv. Math.*, 448:Paper No. 109700, 61, 2024. doi:10.1016/j.aim.2024.109700.

[34] Borys Kuca. Multidimensional polynomial Szemerédi theorem in finite fields for polynomials of distinct degrees. *Israel J. Math.*, 259(2):589–620, 2024. doi:10.1007/s11856-023-2551-3.

[35] Borys Kuca, Tuomas Orponen, and Tuomas Sahlsten. On a continuous Sárközy-type problem. *Int. Math. Res. Not. IMRN*, 2023(13):11291–11315, 2023. doi:10.1093/imrn/rnac168.

[36] James Leng, Ashwin Sah, and Mehtaab Sawhney. Improved bounds for Szemerédi's theorem. Preprint, 2024. URL: https://arxiv.org/abs/2402.17995.

[37] Zi Li Lim. Uniform nonlinear Szemerédi theorem for corners in finite fields. Preprint, 2025. URL: https://arxiv.org/abs/2501.04887.

[38] Yu-Hsiang Fred Lin. *Singular Brascamp-Lieb Forms and Multilinear Fourier Multipliers with Rough or Oscillatory Multipliers*. Doctoral dissertation, University of Bonn, Bonn, Germany, 2025. URL: https://hdl.handle.net/20.500.11811/13375.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] R. Daniel Mauldin. Some problems and ideas of Erdős in analysis and geometry. In *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 365–376. János Bolyai Math. Soc., Budapest, 2013. doi:10.1007/978-3-642-39286-3_13.

[44] Ilya D. Shkredov. On a problem of Gowers. *Izv. Ross. Akad. Nauk Ser. Mat.*, 70(2):179–221, 2006. doi:10.1070/IM2006v070n02ABEH002316.

Email address: aleksandar.bulj@math.hr

Email address: vjekovac@math.hr

University of Zagreb Faculty of Science, Department of Mathematics, Bijenička cesta 30, 10000 Zagreb, Croatia