

HEILBRONN'S TRIANGLE PROBLEM IN THREE DIMENSIONS

DOMINIQUE MALDAGUE, HONG WANG, DMITRII ZAKHAROV

ABSTRACT. We show that among any n points in the unit cube one can find a triangle of area at most $n^{-2/3-c}$ for some absolute constant $c > 0$. This gives the first non-trivial upper bound for the three-dimensional version of Heilbronn's triangle problem. This estimate is a consequence of the following result about configurations of point-line pairs in $\mathbb{R}^3$: for $n \geq 2$ let $p_1, \dots, p_n \in [0, 1]^3$ be a collection of points and let ℓ_i be a line through p_i for every i such that $d(p_i, \ell_j) \geq \delta$ for all $i \neq j$. Then we have $n \lesssim \delta^{-3+\gamma}$ for some absolute constant $\gamma > 0$. The analogous result about point-line configurations in the plane was previously established by Cohen, Pohoata and the last author.

1. INTRODUCTION

In late 1940-s, Heilbronn asked the following question in discrete geometry. For a set of points P let $\Delta = \Delta(P)$ be the minimum area of a triangle spanned by a triple of points in P (collinear triples count as a 0 area triangles). Now define $\Delta(n) = \max_P \Delta(P)$ where the maximum is taken over all subsets $P \subset [0, 1]^2$ of size n . How does the function $\Delta(n)$ grow with n ? Two constructions (one algebraic and one probabilistic) due to Erdős show that $\Delta(n) \geq cn^{-2}$ for some constant $c > 0$ and Heilbronn conjectured that this lower bound is sharp. In 1982, Komlós–Pintz–Szemerédi [6] disproved his conjecture by showing that $\Delta(n) \gtrsim \frac{\log n}{n^2}$. In the other direction, it is not hard to see that $\Delta(n) \leq Cn^{-1}$ holds for some constant C . In 1951, Roth [13] managed to improve this bound by a small factor tending to zero using a clever density increment argument. Later on, in 1972-73, after a slight improvement by Schmidt [14], Roth [11], [12] developed a new powerful analytic method and used it to get a polynomial saving $\Delta(n) \lesssim n^{-1-c}$ for some (explicit) constant $c > 0$. In 1981, Komlós–Pintz–Szemerédi [5] optimized Roth's method and proved that $\Delta(n) \lesssim n^{-8/7+o(1)}$. This was the best known upper bound for about 40 years, until in 2023-24, Cohen, Pohoata and the third author managed to improve the exponent from $8/7$ to $8/7 + c$ in [1] and then to $7/6$ in [2]. These improvements were made possible by discovering new connections between Roth's approach and recent work in harmonic analysis and fractal geometry. In particular, Roth's analytic method turns out to be closely related to the high-low method in harmonic analysis pioneered by Guth–Solomon–Wang [3] and the key new input in the improved exponent $8/7 + c$ in [1] comes from the work on radial projections estimates [9].

The following is a natural generalization of Heilbronn's triangle problem to higher dimensions. For integers $d \geq 2$ and $k \in [3, d+1]$ and a set of points $P \subset \mathbb{R}^d$ we can define $\Delta_k(P)$ to be the smallest $k-1$ -dimensional volume of a k -vertex simplex defined by P . We then define a function $\Delta_{k,d}(n) = \max_P \Delta_k(P)$ where the maximum is over all subsets $P \subset [0, 1]^d$ of size n . In particular, we have $\Delta_{3,2}(n) = \Delta(n)$. One can also consider values $k > d+1$ and let $\Delta_k(P)$ to be the smallest possible d -volume of a convex hull of a k -tuple of points in P . For any given pair of parameters k, d we can investigate the asymptotic behavior of the function $\Delta_{k,d}(n)$. Perhaps the two most interesting special cases of this question are the cases of the full-dimensional simplices, $k = d+1$, and triangles, $k = 3$ (we refer to [18] for the treatment of other values of k). In the case of simplices, the 'easy' bounds are $n^{-d} \lesssim \Delta_{d+1,d}(n) \lesssim n^{-1}$. The lower bound was improved by Lefmann [7] to $\Delta_{d+1,d}(n) \gtrsim \frac{\log n}{n^d}$ and the upper bound was improved by the third author [18] to $\Delta_{d+1,d}(n) \lesssim n^{-\log d + C}$ for some constant C (along with a number of bounds in low dimensions, e.g. that $\Delta_{4,3}(n) \leq Cn^{-4/3}$). In the case of triangles in $\mathbb{R}^d$, the 'easy' bounds are $n^{-\frac{2}{d-1}} \lesssim \Delta_{3,d}(n) \lesssim n^{-\frac{2}{d}}$. The lower bound follows by taking a uniformly random set of $2n$ points and removing from it all triangles with area less than $cn^{-\frac{2}{d-1}}$ for

a small constant $c > 0$. This lower bound was improved to $\Delta_{3,d}(n) \gtrsim (\log n)^{\frac{1}{d-1}} n^{-\frac{2}{d-1}}$ by Lefmann [7] following the semi-random approach of Komlós–Pintz–Szemerédi [6]. For the upper bound one can use the following packing argument. Since $[0, 1]^d$ can be covered by at most $n/3$ cubes with sides of length at most $10n^{-1/d}$, by pigeonhole principle, there is a cube containing at least 3 points of P . These points then form a triangle of area less than $Cn^{-2/d}$. To the best of our knowledge, this is still asymptotically the best known upper bound on the function $\Delta_{3,d}(n)$ for any $d \geq 3$.

We make the first progress in this direction and prove a new estimate for Heilbronn’s triangle problem in $\mathbb{R}^3$:

Theorem 1.1. *Let $P \subset [0, 1]^3$ be a set of n points. Then P contains three points forming a triangle of area at most $Cn^{-2/3-c}$ for some absolute constants $c > 0, C \geq 1$.*

In other words, we have the new upper bound $\Delta_{3,3}(n) \leq Cn^{-2/3-c}$. Our proof produces an explicit, albeit very tiny, value of c . By tracking the estimates, one should be able to take $c = 10^{-100}$ or so in Theorem 1.1, however we did not make any attempt to optimize this value.

In [2], the estimate $\Delta(n) \leq n^{-7/6+o(1)}$ was proven using the following result about configurations of incident point-line pairs.

Theorem 1.2 ([2]). *For all $\varepsilon > 0$ there is C_ε such that the following holds for all $n \geq 2$. Let $p_1, \dots, p_n \in [0, 1]^2$ be a collection of points and let ℓ_i be a line through p_i . Suppose that $d(p_i, \ell_j) \geq \delta$ for all $i \neq j$ and some $\delta > 0$. Then we have $n \leq C_\varepsilon \delta^{-\varepsilon} \delta^{-3/2}$.*

In the same spirit, we reduce Theorem 1.1 to the following three dimensional version of Theorem 1.2:

Theorem 1.3. *There exists an absolute constant $\gamma > 0$ such that for all $\varepsilon > 0$ there is C_ε such that the following holds. Let $p_1, \dots, p_n \in [0, 1]^3$ be a collection of points and let ℓ_i be a line through p_i . Suppose that $d(p_i, \ell_j) \geq \delta$ for all $i \neq j$ and some $\delta > 0$. Then we have $n \leq C_\varepsilon \delta^{-\varepsilon} \delta^{-3+\gamma}$.*

Proof of Theorem 1.1. Let $P \subset [0, 1]^3$ be a set of size n . By covering the unit cube with $10/n^{1/3}$ -size cubes and applying pigeonhole principle, we can find a pair of points $\{p_1, q_1\} \subset P$ such that $d(p_1, q_1) \leq Cn^{-1/3}$. Repeating this argument, we can then find a pair $\{p_2, q_2\} \subset P$ disjoint from $\{p_1, q_1\}$ so that $d(p_2, q_2) \leq C(n-2)^{-1/3}$. Continuing in this manner, we can find a collection of $m = \lfloor n/4 \rfloor$ pairwise disjoint pairs of points $\{p_i, q_i\} \subset P$ such that $d(p_i, q_i) \leq 2Cn^{-1/3}$ for all i . Let $\ell_i = \overline{p_i q_i}$ and consider the collection of point-line pairs $X = \{(p_i, \ell_i), i = 1, \dots, m\}$. Let $\delta = \min_{i \neq j} d(p_i, \ell_j)$. By Theorem 1.3, for all $\varepsilon > 0$ we have $m \lesssim_\varepsilon \delta^{-3+\gamma-\varepsilon}$. Thus, there exists $i \neq j$ such that $d(p_i, \ell_j) \leq \delta \lesssim n^{-\frac{1}{3-\gamma}+o(1)}$. The triangle $p_i p_j q_j$ has area at most $n^{-\frac{1}{3}-\frac{1}{3-\gamma}+o(1)} \leq Cn^{-2/3-c}$, finishing the proof. $\square$

To put results stated in Theorem 1.3 and Theorem 1.2 into perspective, it is helpful to introduce some notation. A *point-line pair* in $\mathbb{R}^d$ is a pair (p, ℓ) where $p \in [0, 1]^d$ and ℓ is a line passing through p . We denote by Ω_d the set of all point-line pairs in $\mathbb{R}^d$ (note that $\dim \Omega_d = 2d - 1$). We are interested in properties of finite collections of point-line pairs, i.e. subsets $X \subset \Omega_d$. For any such X , we define the *minimal distance* of X as

$$d(X) = \min_{(p, \ell) \neq (p', \ell') \in X} d(p, \ell'),$$

i.e. it is the minimal distance between a point in the configuration X and a line passing through a different point in the configuration.

Let $\text{PL}_d(\gamma)$ denote the following statement: for any $\varepsilon > 0$ there is C_ε such that for any point-line configuration $X \subset \Omega_d$ with $d(X) \geq \delta$, it follows that $|X| \leq C_\varepsilon \delta^{-\varepsilon} \delta^{-d+\gamma}$ for every $\varepsilon > 0$. With this definition, Theorem 1.2 is the statement $\text{PL}_2(1/2)$ and Theorem 1.3 is equivalent to the statement that $\text{PL}_3(\gamma)$ holds for some $\gamma > 0$.

Note that for any $d \geq 2$, the statement $\text{PL}_d(0)$ is trivially true. Indeed, for a configuration $X \subset \Omega_d$, consider the set of points $P[X] = \{p : (p, \ell) \in X\} \subset [0, 1]^d$ (taking points with multiplicity if there

are overlaps). By covering the unit cube with subcubes of size $C_d|X|^{-1/d}$ and applying the pigeonhole principle, we can find some $p \neq p' \in P[X]$ so that $d(p, p') \leq C_d|X|^{-1/d}$. Now if $(p', \ell') \in X$ is the corresponding point-line pair, then we get $d(X) \leq d(p, \ell') \leq d(p, p') \leq C_d|X|^{-1/d}$. So if $d(X) \geq \delta$ then it follows that $|X| \lesssim \delta^{-d}$, showing $\text{PL}_d(0)$.

On the other hand, we can consider a collection of $\sim \delta^{1-d}$ points in $[0, 1]^{d-1} \times \{0\}$ which are pairwise δ -separated and for each point draw the line in the vertical direction through it. The resulting point-line configuration X satisfies $|X| \sim \delta^{1-d}$ and $d(X) \geq \delta$. This implies that $\text{PL}_d(1 + \varepsilon)$ is false for every $\varepsilon > 0$. In an upcoming note [8], it will be shown that exists some constant $c > 0$ such that for all $d \geq 2$ the statement $\text{PL}_d(1 - c)$ does not hold.

In the plane, the statement $\text{PL}_2(1/2)$ appears to be the limit of the high-low method as demonstrated by the Szemerédi–Trotter example and the hermitian unital example in the finite field model, we refer to [2] for more discussion on this. There are some tentative reasons to think that $\text{PL}_d(1/d)$ is going to be at the limit for analytic approaches and reaching this limit would be very interesting. However the proof presented in this paper is quite far from this limit as the explicit value of γ in Theorem 1.3 is extremely poor. The proof presented in this paper relies on a certain numerical coincidence specific to $\mathbb{R}^3$ (see the next section for more details) and it appears that the problem becomes significantly harder in dimension four. Thus, we leave proving that $\text{PL}_4(\gamma)$ holds for some $\gamma > 0$ as a direction for future research.

The rest of the paper is organized as follows. In Section 2 we provide an exposition of the proof of Theorem 1.3. We start by giving a proof of $\text{PL}_2(\gamma)$ for some $\gamma > 0$ using the two-dimensional high-low inequality. We then discuss the adjustments to the strategy in three dimensions. In Section 3 we prove two results about incidences between points and tubes. These results are based on Wolff's two-ends and hairbrush arguments and do not require any 'advanced' geometric measure theory (though those might become handy when trying to optimize the value of γ). In Section 4, we prove several high-low inequalities for point-tube incidences in $\mathbb{R}^3$, starting from the basic estimate, followed by two refined versions. The incidence results from the previous section are a key ingredient here. In Section 5 we give some preliminary results needed in Section 6, in which we employ the high-low inequalities from Section 4 in an iterative scheme and prove Theorem 1.3.

1.1. Notation. We use the following asymptotic notation. We write $A \lesssim B$ if $|A| \leq CB$ for some constant C , we also write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$ (and similarly for the following notation). We write $A \lesssim_K B$ if $A \leq CK^C B$ for some constant C and similarly if there more parameters in the subscript. We sometimes write $A \lesssim_{\delta} B$ to denote the fact that $A \lesssim_{\delta^{-\varepsilon}} B$ where $\delta > 0$ is a parameter which will be clear from context. Sometimes we write $O(1)$ for a quantity bounded in absolute value by a constant and $o(1)$ for a quantity tending to 0 with δ . Finally, we somewhat informally use the notation $A \ll B$ to say that 'A is much smaller than B', e.g. it will typically mean that $A \lesssim \delta^c B$ for some $c > 0$ unless specified otherwise.

1.2. Acknowledgements. The third author thanks Alex Cohen and Larry Guth for helpful conversations about the problem.

2. PROOF OVERVIEW.

In this section we outline the main ideas in the proof of Theorem 1.3. We begin by recalling the two-dimensional high-low inequality and use it to give a short proof of the two-dimensional version of Theorem 1.3. Then we discuss the new difficulties that arise when we go to three dimensions and how to overcome them.

2.1. High-low inequality in two dimensions. The key tool in polynomial upper bounds for Heilbronn's triangle problem is an analytic inequality relating the incidences between points and lines on different scales. We follow the setup used in [1] and [2].

Let P and L be finite collections of points and lines in $[0, 1]^2$. For a certain symmetric bump function $\varphi : \mathbb{R} \rightarrow [0, 1]$ supported on $[-2, 2]$ we define

$$I(\delta; P, L) = \sum_{p \in P, \ell \in L} \varphi(\delta^{-1}d(p, \ell)).$$

This is the number of approximate incidences between P and L on scale δ , smoothed out by a bump function φ . It is essentially counting the number of pairs $(p, \ell) \in P \times L$ such that $d(p, \ell) \leq \delta$.

Note that if P and L are taken randomly inside $[0, 1]^2$, then we expect $I(\delta; P, L) \approx \delta|P||L|$. So it is natural to introduce the normalization

$$B(\delta) = B(\delta; P, L) = \frac{I(\delta; P, L)}{\delta|P||L|}.$$

Roth's key idea was to relate the incidence counts $I(\delta; P, L)$ between many different scales. Roughly speaking, he observed that the normalized incidence function $B(\delta)$ is essentially constant on scales where points and lines are sufficiently well distributed. Namely, let us define concentration numbers

$$\begin{aligned} M_P(\delta) &= \max_{Q: \delta \times \delta \text{ square}} |P \cap Q|, \\ M_L(\delta \times 1) &= \max_{T: \delta \times 1 \text{ tube}} |L \cap T| \end{aligned}$$

where $L \cap T$ is the set of lines $\ell \in L$ so that $|T \cap \ell| \geq 1/2$. With this notation, we have the following estimate.

Proposition 2.1. *There is a bump function φ such that any P and L and $\delta > 0$ we have:*

$$(1) \quad |B(\delta) - B(\delta/2)|^2 \lesssim \delta^{-3} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times 1)}{|L|}$$

See Appendix A in [2] for a proof. In Section 4 we will use the Fourier-analytic approach of Guth–Solomon–Wang [3] to prove versions of (1) in three dimensions. In particular, by following the proof of Lemma 4.2 one can recover (1) with an $\delta^{-\varepsilon}$ -loss. For now, let us just briefly describe the proof strategy of (1). To each P and L we can associate (blurred) indicator functions g and f . Then the incidence number $B(\delta)$ can be computed as a scalar product $\langle f, g \rangle$ of functions f and g . By applying Plancherel, we can rewrite this as a scalar product of their Fourier transforms $\hat{f}$ and $\hat{g}$. Now we can split the Fourier transform $\hat{f}$ into the high frequency and low frequency parts, giving us a decomposition $f = f^{high} + f^{low}$. The low frequency is obtained by restricting $\hat{f}$ on a ball of radius δ^{-1}/K around the origin. On the physical side, this corresponds to blurring f by a bump function of radius $K\delta$. This means that the low part of the scalar product $\langle g, f^{low} \rangle$ is essentially the normalized incidence count at scale $K\delta$, i.e. it is essentially $B(K\delta)$. Now we can use this to estimate using Cauchy–Schwarz $|B(\delta) - B(K\delta)| \leq \langle g, f^{high} \rangle \leq \|g\|_2 \|f^{high}\|_2$. We then convert each L_2 -norm using orthogonality and $L_1 - L_\infty$ to information about distribution of P and L on scale δ . After some computation, we get that the high part is controlled by the right hand side of (1) (with an extra $\delta^{-\varepsilon}$ loss).

Let us close this section by giving three sharp constructions for (1):

- (i) Let $P = \{0\}$ and let L be a collection of $\sim \delta^{-1}$ directionally separated lines through the origin. Then we have $|B(\delta) - B(2\delta)| \sim B(\delta) \sim \delta^{-1}$ and $\delta^{-3} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta)}{|L|} \sim \delta^{-2}$, matching (1).
- (ii) Dually, Let $L = \{\ell_0\}$ for some line ℓ_0 and let P be a collection of $\sim \delta^{-1}$ points on ℓ_0 . Then we have $|B(\delta) - B(2\delta)| \sim B(\delta) \sim \delta^{-1}$ and $\delta^{-3} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta)}{|L|} \sim \delta^{-2}$, matching (1).
- (iii) A much less trivial example comes from sharp examples for the Szemerédi–Trotter theorem: let P and L be sets of N points and lines defining $\sim N^{4/3}$ incidences. Then for $\delta = cN^{-2/3}$ one can verify that both sides of (1) are ~ 1 . See [1] for more details.

So to make an improvement over (1) we would need to put a spacing constraint on (P, L) which would separate it from all 3 types of examples. The Szemerédi–Trotter example appears to make this

a very challenging task and this is the key bottleneck for further improvements on Heilbronn's triangle problem in the plane.

Rather surprisingly, the situation changes in three dimensions: for a direct analogue of (1) in $\mathbb{R}^3$ (which we will discuss below) we only have two types of sharp examples roughly matching (i) and (ii) above, but the grid-like examples giving (iii) are no longer sharp. This allows us to give simple criteria under which the simple high-low estimate can be improved.

2.2. Point-line problem in two dimensions. Let us now give a proof of the two dimensional version of Theorem 1.3.

Let $X = \{(p_i \in \ell_i), i = 1, \dots, n\}$ be a collection of point-line incident pairs inside $[0, 1]^2$. Let us define the *minimal distance* of X as

$$d(X) = \min_{(p, \ell) \neq (p', \ell') \in X} d(p, \ell').$$

Proposition 2.2. *Let $\delta > 0$ and X be a collection of point-line pairs in $[0, 1]^2$ with $d(X) \geq \delta$. Then we have $|X| \lesssim \delta^{-2+\gamma}$ for some absolute constant $\gamma > 0$. In other words, the property $\text{PL}_2(\gamma)$ holds for some $\gamma > 0$.*

Recall that [2] showed that $\text{PL}_2(1/2)$ holds. Here we present a much simpler argument than the one given in [2]. By tracking down the estimates one can check that the proof below gives something like $\gamma = 0.1$.

Proof. For the sake of contradiction let us assume that $|X| \approx_\gamma \delta^{-2}$ holds. Let $P = P[X]$ and $L = L[X]$ be the sets of points and lines of the point-line configuration X . Note that we have $|P| = |L| = |X|$.

The condition $d(X) \geq \delta$ implies that points in P are δ -separated: indeed, for $(p, \ell), (p', \ell') \in X$ we have $d(p, p') \geq d(p, \ell') \geq \delta$. So for any $w \in (\delta, 1)$ we get

$$(2) \quad M_P(w) = \max_{Q \text{ } w\text{-square}} |P \cap Q| \lesssim (w/\delta)^2 \lesssim_\gamma w^2 |P|.$$

Similarly, lines in L must be δ -separated and for $w \in (\delta, 1)$ we get

$$(3) \quad M_L(w \times 1) = \max_{T \text{ } w \times 1 \text{ tube}} |L \cap T| \lesssim (w/\delta)^2 \lesssim_\gamma w^2 |L|.$$

So by Proposition 2.1, we obtain

$$(4) \quad |B(w) - B(w/2)|^2 \lesssim w^{-3} \frac{M_P(w)}{|P|} \frac{M_L(w \times 1)}{|L|} \lesssim_\gamma w.$$

By the assumption, we have $d(p, \ell') \geq \delta$ for any $p \in P$ and $\ell' \in L$ for which $(p, \ell') \notin X$. This means that P and L define an abnormally small amount of incidences on scale δ : for a small constant $c > 0$, we have $I(c\delta; P, L) \lesssim |X| \ll \delta |P| |L|$ and so $B(c\delta) \lesssim \frac{|X|}{\delta |X|^2} \lesssim_\gamma \delta$.

On the other hand, we are now going to show that $B(C\delta^{1/2}) \approx_\gamma 1$ holds. Using (4) for $w \in (c\delta, C\delta^{1/2})$ this will then lead to a contradiction for small enough γ .

Denote $\Delta = \delta^{1/2}$ and let $\mathbb{T}$ be the set of essentially distinct $\Delta \times 1$ tubes which cover L and let $\mathcal{Q}$ be the set of disjoint Δ -boxes Q covering P . Using $|P|, |L| \approx_\gamma \delta^{-2}$ we see that $|\mathbb{T}|, |\mathcal{Q}| \approx_\gamma \delta^{-1}$. After a passing to a large subset in X we may assume that every $T \in \mathbb{T}$ contains $\approx_\gamma \delta^{-1}$ lines $\ell \in L$ and every $Q \in \mathcal{Q}$ contains $\approx_\gamma \delta^{-1}$ points from P . Using this, we can lower bound

$$(5) \quad I(C\Delta; P, L) \gtrsim_\gamma \delta^{-2} I(\mathcal{Q}, \mathbb{T})$$

where $I(\mathcal{Q}, \mathbb{T}) = \#\{(Q, T) : Q \cap T \neq \emptyset\}$. Fix $Q \in \mathcal{Q}$ and let us estimate the number of $T \in \mathbb{T}$ which intersect with Q . Summing over all Q will give us a lower bound on the number of incidences between $\mathcal{Q}$ and $\mathbb{T}$. For every point $p \in P \cap Q$ we have a unique line $\ell = \ell(p)$ so that $(p, \ell) \in X$. Every such line $\ell(p)$ is covered by some tube $T = T(p) \in \mathbb{T}$. In particular, we have $T(p) \cap Q \neq \emptyset$ for every $p \in P \cap Q$. For a fixed $T \in \mathbb{T}$ let $p_1, \dots, p_m$ be the set of points in $P \cap Q$ such that $T(p_j) = T$ for every j . Note that the directions of lines $\ell(p_j)$, $j = 1, \dots, m$, lie in an arc of length $\sim \Delta$ determined by the main

axis of T . So by pigeonhole principle we can find a direction $\theta_0 \in S^1$ so that $|\theta(\ell(p_j)) - \theta_0| \leq \Delta/100$ for $\gtrsim m$ indices $j \in [m]$. Let j, j' be any two such indices. Since $d(p_{j'}, \ell(p_j)) \geq \delta$ and $p_j, p_{j'} \in Q$, simple geometry implies that $|\pi_{\theta_0}(p_{j'}) - \pi_{\theta_0}(p_j)| \geq \delta/2$, where π_{θ_0} is the orthogonal projection in the direction of θ_0 . So since $\pi_{\theta_0}(p_j) \in \pi_{\theta_0}(Q)$ for any such j , we conclude that $m \lesssim \Delta/\delta = \delta^{-1/2}$. Thus,

$$\#\{T \in \mathbb{T} : T \cap Q \neq \emptyset\} \gtrsim |P \cap Q|/m \gtrsim_{\gamma} \delta^{-1/2}.$$

It follows that $I(\mathcal{Q}, \mathbb{T}) \gtrsim_{\gamma} \delta^{-3/2}$ and so by (5), $I(C\Delta; P, L) \approx_{\varepsilon} \Delta|P||L|$, i.e. $B(C\Delta) \gtrsim_{\varepsilon} 1$. By combining this with $B(c\delta) \lesssim_{\varepsilon} \delta$ and (4), we arrive at a contradiction for sufficiently small γ . $\square$

2.3. Higher dimensions. Let us see how the high-low approach would work in $\mathbb{R}^d$. Given collections of points P and lines L in $[0, 1]^d$, we can define incidence functions $I(\delta; P, L) = \sum_{p \in P, \ell \in L} \varphi(d(p, \ell))$ and $B(\delta) = \frac{I(\delta; P, L)}{\delta^{d-1}|P||L|}$ similarly to $\mathbb{R}^2$ (note that in $\mathbb{R}^d$ we expect $I(\delta) \approx \delta^{d-1}|P||L|$). We can also define the concentration numbers

$$\begin{aligned} M_P(\delta) &= \max_{Q: \delta \text{ cube}} |P \cap Q|, \\ M_L(\delta \times \dots \times \delta \times 1) &= \max_{T: \delta \times 1 \text{ tube}} |L \cap T| \end{aligned}$$

With these definitions, a direct analogue of Proposition 2.1, is the following estimate.

Proposition 2.3. *For an appropriate bump function φ and all P and L in $\mathbb{R}^d$ we have*

$$(6) \quad |B(\delta) - B(2\delta)|^2 \lesssim_{\varepsilon} \delta^{-\varepsilon} \delta^{-3(d-1)} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times \dots \times \delta \times 1)}{|L|}.$$

See Lemma 4.2 for a proof of this bound in $\mathbb{R}^3$ (the proof in other dimensions is identical). An estimate like this was first proven by Guth–Solomon–Wang [3] to study well-spaced collections of tubes in $\mathbb{R}^3$. The proof strategy goes along the same lines as the outline we gave in Section 2.1 for the two-dimensional estimate. The main difference is in the estimate of the L_2 -norm of the high term $\|f^{high}\|_2$: in the plane one can use separation in Fourier space to split f^{high} into a sum of (essentially) pairwise orthogonal pieces f_{θ}^{high} (essentially by grouping lines according to their direction). Then each piece f_{θ}^{high} can be bounded by orthogonality using physical separation. In higher dimensions however, the corresponding pieces f_{θ}^{high} are no longer separated in the Fourier space. The Fourier transform of each f_{θ}^{high} looks like a great circle on the $(d-1)$ -dimensional sphere S^{d-1} and so the overlap pattern of these great circles controls the size of the L_2 -norm. In general, we do not have that much information about the great circles, but by using a crude estimate on the overlap, one can use this geometric description to deduce (6).

There are two main cases where this estimate is sharp:

- (i) Let $P = \{0\}$ and L be a collection of $\sim \delta^{-(d-1)}$ directionally separated lines passing through the origin. Then we have $|T_{\delta}(\ell) \cap P| = 1 = \delta^{-(d-1)}(\delta^{d-1}|P|)$ and so $|B(\delta) - B(2\delta)| \sim B(\delta) \sim \delta^{-(d-1)}$. On the other hand,

$$\delta^{-3(d-1)} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times \dots \times \delta \times 1)}{|L|} \sim \delta^{-2(d-1)}$$

matching (6).

Furthermore, by taking a union of $c\delta^{-(d-1)}$ random translates of this construction, we get a collection of $\sim \delta^{-(d-1)}$ points and $\sim \delta^{-2(d-1)}$ lines such that $|B(\delta) - B(2\delta)| \sim B(\delta) \sim 1$ and the high-low error (6) is also $\approx_{\varepsilon} 1$.

- (ii) Let P and L be sets of $\sim \delta^{-(d-1)}$ points and $\sim \delta^{-2(d-2)}$ lines contained in a fixed hyperplane. Then we typically have $|T_{\delta}(\ell) \cap P| \sim \delta^{d-2}|P| = \delta^{-1}(\delta^{d-1}|P|)$ and so $|B(\delta) - B(2\delta)| \sim B(\delta) \sim$

δ^{-1} . On the other hand,

$$\delta^{-3(d-1)} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times \dots \times \delta \times 1)}{|L|} \sim \delta^{-2},$$

matching (6).

Furthermore, by taking a union of $c\delta^{-1}$ random translates of this construction, we get a collection of $\sim \delta^{-d}$ points and $\sim \delta^{-2d+3}$ lines such that $|B(\delta) - B(2\delta)| \sim B(\delta) \sim 1$ and the high-low error is also ~ 1 .

Now let $X = \{(p \in \ell)\}$ be a point-line configuration in $[0, 1]^d$ and suppose that

$$d(X) = \min_{(p, \ell) \neq (p', \ell') \in X} d(p, \ell') \geq \delta.$$

From this condition it follows that $P = P[X]$ is a δ -separated set of points and $L = L[X]$ is a δ -separated set of lines in $[0, 1]^d$. So we get $|X| \leq |P| \lesssim \delta^{-d}$. Our goal would be to improve on this estimate by a polynomial factor δ^κ for some constant $\kappa > 0$. For the sake of contradiction let us suppose that $|X| \approx_\kappa \delta^{-d}$. As in the two-dimensional case, we have very few incidences on scale δ :

$$I(c\delta; P, L) \leq |X| \ll \delta^{d-1} |P| |L|$$

and so we get $B(c\delta) = \frac{I(c\delta; P, L)}{\delta^{d-1} |P| |L|} \lesssim_\kappa \delta$.

On the other hand, let us assume that we managed to prove that $B(\delta^\alpha) \approx_\kappa 1$ for some $\alpha \in (0, 1)$. We would like to use the high-low method to reach a contradiction by showing that the difference $|B(w) - B(w/2)|$ for all $w \in (\delta, \delta^\alpha)$ is much less than 1. For this approach to work, we at the very least have to be able to show that $|B(\delta) - B(2\delta)| \ll 1$. The high-low inequality (6) gives us:

$$(7) \quad |B(\delta) - B(2\delta)|^2 \lesssim_\varepsilon \delta^{-\varepsilon} \delta^{-3(d-1)} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times \dots \times \delta \times 1)}{|L|} \approx_\kappa \delta^{3-d},$$

where we used the estimates $M_P, M_L \lesssim 1$ coming from the fact that points in P and lines in L are pairwise δ -separated. If $d = 2$ then this bound gives $|B(\delta) - B(2\delta)| \lesssim_\kappa \delta^{1/2} \ll 1$ and this allowed us to run the high-low argument in Section 2.2. But for $d \geq 3$ the right hand side is no longer small enough. Examples (i), (ii) above show that this bound is sharp in general so this is a serious obstruction to this approach. The situation in case of $d = 3$ is quite special however: the right hand side of (7) exceeds 1 only by a factor of the form $\delta^{-C\kappa+o(1)}$ where κ is our initial error term in the assumption $|X| \approx_\kappa \delta^{-d}$. So if we could squeeze out a tiny saving from the bound (6) then it might give us a fighting chance.

This idea turns out to work: in Section 4 we prove two different estimates which improve (6) by a polynomial factor for $d = 3$ under certain geometric constraints on L . Note that in example (ii) there are lots of lines contained in a fixed plane. One can also construct similar examples where lines concentrate a lot in neighborhoods of several planes. So a necessary condition to improve on (6) is that L does not have too many lines in a neighborhood of a plane. Let $M_L(u \times w \times 1)$ be the maximum number of lines $\ell \in L$ contained in a $u \times w \times 1$ box Π (in the sense that $|\Pi \cap \ell| \geq 1/2$). The example (ii) then satisfies $M_L(\delta \times 1 \times 1) \sim \delta^{-2}$ which is the maximum possible value for a set of δ -separated lines L .

Luckily for us, a situation like this is impossible for the set of lines $L = L[X]$ coming from our point-line configuration X satisfying $d(X) \geq \delta$ and $|X| \approx_\kappa \delta^{-3}$. Namely, using the 2-dimensional case of the point-line problem (i.e. that $\text{PL}_2(3/2)$ holds by Theorem 1.2), we can get a good estimate on the number of lines near any plane. In particular, we can get a non-concentration estimate $M_L(\delta \times 1 \times 1) \leq \delta^{-3/2+o(1)} \ll \delta^{-2}$. Indeed, for a $\delta \times 1 \times 1$ slab Π we may essentially treat the set of pairs $(p, \ell) \in X$ such that $|\ell \cap \Pi| \geq 1/2$ as a point-line configuration in the plane.

A slight generalization of this idea implies that we have the following spacing condition of the set of lines L :

$$(8) \quad M_L(u \times w \times 1) \ll (u/\delta) \times (w/\delta)^2, \text{ for all } \delta \leq u \leq w \leq 1$$

(where $\ll$ refers to some polynomial gain, possibly involving powers of u, w , see Proposition 5.1 for the precise claim).

Our first improvement over (6) for $d = 3$ has the following shape:

$$(9) \quad \text{If } |\theta(L)|_\delta \ll \delta^{-2} \text{ and (8), holds, then } |B(\delta) - B(2\delta)| \ll 1,$$

where $\theta(L) \subset S^2$ is the set of directions of lines $\ell \in L$, $|\cdot|_\delta$ denotes the δ -covering number. Note that the restriction on the number of directions in (9) implies that (P, L) cannot look like the sharp bush example (i) (since the number of lines passing through a fixed point is upper bounded by the number of directions) and condition (8) implies that L cannot look like the sharp plane example (ii). One drawback of this estimate is that it only works if L does not span all possible directions. There is no reason to expect that this is going to hold for the set of lines L coming from a point-line configuration X with $d(X) \geq \delta$. The good news is that restrictions on the direction set appear naturally when we start rescaling into smaller boxes and this is related to how we use the information about directions in small boxes in the ‘initial estimate’ argument from Section 2.2.

The second improvement has the following shape:

$$(10) \quad \text{If } L \text{ is well-spaced on scale } \delta^{1/2} \text{ and (8) holds, then } |B(\delta) - B(2\delta)| \ll 1.$$

Here the condition ‘well-spaced on scale $\Delta = \delta^{1/2}$ ’ means that for any $\delta^{1/2} \times \delta^{1/2} \times 1$ tube T we have an estimate $|T \cap L| \lesssim \Delta^4 |L|$. In other words, we require an almost maximum possible number of Δ -tubes to cover L . Note that the bush example (i) does not satisfy this property: we only need $\sim \Delta^{-2}$ Δ -tubes to cover each bush and there are only $\sim \delta^{-1}$ bushes overall. Again the condition (8) makes sure that L does not look like example (ii). The ‘well-spaced’ condition originates from Guth–Solomon–Wang [3] and is used there in a similar manner to improve the bound on the high term.

Coming back to our point-line configuration X , there is no reason for lines to concentrate in any prisms or tubes, so it is natural to expect that L satisfies both premises of (10). So at least in a ‘typical’ situation we can now get the desired improvement in (7). However there is a significant complication: in order to get a contradiction and show $|X| \ll \delta^{-3}$ we need to improve over the simple high-low estimate (6) for every $w \in (\delta, \delta^\alpha)$, not just the end-point case $w = \delta$, and we also have not yet explained how do we obtain the initial estimate $B(\delta^\alpha) \approx 1$.

If (10) does not apply, then this means that L is not well-spaced on scale Δ , i.e. we can cover L by $\lesssim_\kappa \Delta^{-4+\beta}$ Δ -tubes for some $\beta > 0$. Let $\mathbb{T}$ denote the family of tubes covering L (by pigeonholing, we know that each tube contain about the same number of lines). It follows that for a typical Δ -box Q intersecting P we then have $\approx_\kappa \Delta^2 |\mathbb{T}| \lesssim \Delta^{-2+\beta}$ tubes from $\mathbb{T}$ which intersect with Q (cf. with the initial estimate part in Section 2.2). Consider an affine rescaling map $\psi : Q \rightarrow [0, 1]^3$ and let $P' = \psi(P \cap Q)$ and $X' = \{\psi(p \in \ell), p \in P \cap Q\}$ the corresponding point-line configuration. We then have $|X'| \approx_\kappa \Delta^{-3}$ and $d(X') \geq \Delta$, i.e. X' is an essentially sharp example for the point-line problem on scale Δ . On the other hand, observe that for $L' = L[X']$ we have

$$|\theta(L')|_\Delta \lesssim \#\{T \in \mathbb{T} : T \cap Q \neq \emptyset\} \lesssim \Delta^{-2+\beta}.$$

So we constructed a new point-line configuration X' essentially matching the trivial upper bound and which now determines a reduced amount of directions on scale Δ . So we can use (9) on X' to improve the simple high-low estimate at scale $w = \Delta$. This is not the end of the story – we still need to improve the high-low error for every scale $w > \Delta$. The restricted direction set property will ensure that (9) is still applicable for w within a small multiple of Δ (say $w \leq \Delta^{1-\tau}$ for some small constant $\tau > 0$). For larger w we no longer have useful direction information, so we can only rely on the estimate (10). If this estimate fails then this means that L' is not well-spaced on scale $w^{1/2}$. So by rescaling into a $w^{1/2}$ box we get a new configuration X'' with restricted direction set on scale $w^{1/2}$. Continuing in this manner we slowly increase the range of scales where the bound (9) applies and eventually we rescale into a final set X^* for which we can apply (9) for all $w \in [\delta^*, \delta^{*\alpha}]$ (where $\delta^{*\alpha}$ is a scale on which we have available a sufficiently strong initial estimate). By carefully tracking the parameters, we show that we only need perform constantly many rescaling steps and so this procedure eventually terminates. In the

end, we obtain a configuration X^* for which we know that $B(\delta^{*\alpha}) \approx_\kappa 1$, $B(\delta^*) \lesssim_\kappa \delta^*$ and the high-low error is $\ll 1$ for every $w \in [\delta^*, \delta^{*\alpha}]$, which gives us a contradiction for sufficiently small κ . (In the actual proof in Section 6, this argument is arranged slightly differently in order to make tracking of parameters a little cleaner.)

3. INCIDENCE ESTIMATES

3.1. Two-ends decomposition. In the next section we prove high-low inequalities by analyzing neighborhoods of lines. On the Fourier side, $\delta \times \delta \times 1$ tubes become $\delta^{-1} \times \delta^{-1} \times 1$ slabs through the origin. If we study the overlap patterns of these slabs on S^2 then we see $\sim \delta \times 1$ tubes on the surface of the sphere. The following lemma allows us to cut these tubes into a small number of pieces so that the resulting pieces do not overlap too much. This estimate on the overlap will be a key input into an L^2 estimate of the high frequency part in the proof of Theorem 4.1.

This lemma is a version of Wolff's 'two-ends' argument which gives an upper bound on the set of very rich points for a collection of tubes provided that those points are well-spread out on every tube. In our situation, we do not necessarily have such a condition in place and instead we reduce to such a condition but cutting out short segments where rich points concentrate too much.

Lemma 3.1. *Let $1 > \Delta > \delta > 0$. For any collection of distinct $1 \times \delta$ tubes $\mathbb{T}$, there exists a collection of essentially distinct $8\delta \times \Delta$ tubes $\mathbb{U}$ with the following property. For $T \in \mathbb{T}$ there is a set $\mathbb{U}(T) \subset \mathbb{U}$ of size at most $C \log_{2/\Delta}(1/\delta)$ such that the collection of sets*

$$2T \setminus \bigcup_{U \in \mathbb{U}(T)} U$$

is at most $C\Delta^{-2}|\mathbb{T}|^{1/2}$ overlapping. Moreover, each $U \in \mathbb{U}(T)$ has the property that T intersects the boundary of U in the sides of length 8δ .

Proof. Fix a set of tubes $\mathbb{T}$, without loss of generality we may assume that $\mathbb{T}$ is contained in $[-2, 2]^2$. For $i = 1, \dots, m$ and $T \in \mathbb{T}$ we will construct a sequence $U_i(T)$ of $\frac{1}{2}\Delta \times 4\delta$ tubes coaxial with T as follows. For $i \geq 0$ suppose we already constructed $U_j(T)$ for $j \leq i$.

Fix a $\delta/10$ -net $\mathcal{P}$ in $[-2, 2]^2$. For a collection of sets in the plane $\mathbb{S}$ let $\mathcal{P}_r(\mathbb{S})$ denote the set of points $p \in \mathcal{P}$ which are contained in at least r sets from $\mathbb{S}$. Fix $r = C_1 \Delta^{-2} |\mathbb{T}|^{1/2}$ for some constant C_1 which we will specify.

For $T \in \mathbb{T}$ denote

$$T_i := 4T \setminus \bigcup_{j=1}^i U_j(T),$$

and $T_0 = 2T$. Given $T \in \mathbb{T}$, we define $U_{i+1}(T)$ to be a $\frac{1}{2}\Delta \times 4\delta$ tube U coaxial with T which maximizes the size of the intersection

$$|\mathcal{P}_r(\mathbb{T}_i) \cap T_i \cap \frac{1}{2}U|.$$

Here $\frac{1}{2}U$ refers to the $\frac{1}{4}\Delta \times 4\delta$ tube with the same center as U . Now we inductively estimate the number of r -rich points for $\mathbb{T}_i$. For $i \geq 0$ we have by Cauchy-Schwarz

$$(11) \quad I(\mathcal{P}_r(\mathbb{T}_i), \mathbb{T}_{i+1}) = \sum_{T \in \mathbb{T}} |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}| \leq |\mathbb{T}|^{1/2} \left(\sum_{T \in \mathbb{T}} |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}|^2 \right)^{1/2}$$

For $T \in \mathbb{T}$ observe that by the choice of $U_{i+1}(T)$ we have

$$|\mathcal{P}_r(\mathbb{T}_i) \cap \frac{1}{2}U_{i+1}(T)| \geq \frac{1}{8}\Delta |\mathcal{P}_r(\mathbb{T}_i) \cap T_i| \geq \frac{1}{8}\Delta |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}|.$$

Thus, we have

$$|\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}|^2 \leq 8\Delta^{-1} |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}| |\mathcal{P}_r(\mathbb{T}_i) \cap \frac{1}{2}U_{i+1}(T)|$$

and since $\frac{1}{2}U_{i+1}(T)$ and T_{i+1} are $\Delta/4$ -separated, we conclude:

$$\sum_{T \in \mathbb{T}} |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}|^2 \leq 8\Delta^{-1} \sum_{p, p' \in \mathcal{P}_r(\mathbb{T}_i): d(p, p') \geq \Delta/4} \# \left\{ T \in \mathbb{T} : p \in T_{i+1}, p' \in \frac{1}{2}U_{i+1}(T) \right\},$$

there are at most $C\Delta^{-1}$ distinct $4 \times 4\delta$ tubes passing through any pair of $\Delta/4$ -separated points, so we get

$$\sum_{T \in \mathbb{T}} |\mathcal{P}_r(\mathbb{T}_i) \cap T_{i+1}|^2 \leq C_0 \Delta^{-2} |\mathcal{P}_r(\mathbb{T}_i)|^2$$

for some absolute constant C_0 , thus,

$$I(\mathcal{P}_r(\mathbb{T}_i), \mathbb{T}_{i+1}) \leq C_0^{1/2} \Delta^{-1} |\mathbb{T}|^{1/2} |\mathcal{P}_r(\mathbb{T}_i)|.$$

So we get by the choice of $r = C_1 \Delta^{-2} |\mathbb{T}|^{1/2}$:

$$|\mathcal{P}_r(\mathbb{T}_{i+1})| \leq \frac{I(\mathcal{P}_r(\mathbb{T}_i), \mathbb{T}_{i+1})}{r} \leq C_1^{-1} C_0^{1/2} \Delta |\mathcal{P}_r(\mathbb{T}_i)|.$$

Let $C_1 = C_0^{1/2}$, then we obtain

$$|\mathcal{P}_r(\mathbb{T}_i)| \leq (\Delta/2)^i |\mathcal{P}_r(\mathbb{T}_0)| \leq (\Delta/2)^i C \delta^{-2}$$

so $\mathcal{P}_r(\mathbb{T}_m) = \emptyset$ for $m = C \log_{2/\Delta}(1/\delta)$.

Now define $\mathbb{U}$ to be a collection of distinct $\Delta \times 8\delta$ -tubes such that for every $T \in \mathbb{T}$ and $i = 1, \dots, m$ we have $U_i(T) \subset 0.9U$ for some $U \in \mathbb{U}$. Define $\mathbb{U}(T) \subset \mathbb{U}$ to be a set of at most m tubes covering $\{U_i(T)\}$. Observe that sets $2T \setminus \bigcup \mathbb{U}(T)$ have overlap at most r . Indeed, if there is a r -rich point for $\{T \setminus \bigcup \mathbb{U}(T)\}$ then the $\delta/10$ -neighborhood of p is r -rich for $\{4T \setminus \bigcup U_i(T)\}$. But this means that $\mathcal{P}_r(\mathbb{T}_m) \neq \emptyset$, contradicting the construction. This completes the proof. $\square$

For reals $b \in [0, 2\pi]$ and $a \in [0, b]$, we define a spherical $a \times b$ -rectangle to be a region $R \subset S^2$ which is the a -neighbourhood of a great circle arc of length b . We will need an analogue of Lemma 3.1 for spherical rectangles instead of tubes. For a spherical $a \times b$ -rectangle R we denote cR the spherical $ca \times cb$ rectangle with the same axis and center.

Corollary 3.2. *There is a constant $c > 0$ such that the following holds for all $b \in (0, 2\pi]$, $\Delta < c$ and $\delta \leq c\Delta b$. For any collection of essentially distinct spherical $\delta \times b$ -rectangles $\mathbb{T}$, there exists a collection of essentially distinct spherical $10\delta \times \Delta b$ -rectangles $\mathbb{U}$ with the following property. For $T \in \mathbb{T}$ there is a set $\mathbb{U}(T) \subset \mathbb{U}$ of size at most $C \log_{2/\Delta}(1/\delta)$ such that the collection of sets*

$$1.5T \setminus \bigcup_{U \in \mathbb{U}(T)} U$$

is at most $C\Delta^{-2} |\mathbb{T}|^{1/2}$ overlapping. Moreover, each $U \in \mathbb{U}(T)$ has the property that $U \cap T$ contains a spherical $\delta \times (\Delta b - 10\delta)$ -rectangle.

Here by essentially distinct rectangles we mean that $|U \cap U'| \leq 0.9|U|$ for every $U \neq U'$.

Proof. Let $r > 0$ be a small constant. Note that a spherical $\delta \times b$ rectangle can be covered by at most $16/r$ many spherical $\delta \times \min(r/8, b)$ -rectangles. So by replacing rectangles in $\mathbb{T}$ by the covering we may assume that $b \leq r/8$ holds (at the cost of increasing C by a factor of $O(1/r)$).

We can cover S^2 by $\sim r^{-2}$ many r -balls B_i in such a way that any $r/2$ -ball is covered by at least one B_i . So is covered $\mathbb{T}$ by families $\mathbb{T}_{B_i} = \{T \in \mathbb{T} : T \subset B_i\}$ and we have the property that if $T, T' \in \mathbb{T}$ satisfy $T \cap T' \neq \emptyset$ then there is some i so that $T, T' \in \mathbb{T}_{B_i}$. Thus, it is enough to prove the statement for one $\mathbb{T}_{B_i}$ at a time. Now let Π be the plane tangent to S^2 at the center point of B_i and consider the radial projection map $\pi : B_i \rightarrow \Pi$ (with projection center at the origin). Observe that under this map, spherical $a \times b$ -rectangles contained in B_i essentially map to $a \times b$ rectangles in the plane Π . More precisely, for any $\varepsilon > 0$ and $r < r(\varepsilon)$ the image $\pi(R)$ of an $a \times b$ spherical rectangle is contained

in a $(1 + \varepsilon)a \times (1 + \varepsilon)b$ plane rectangle and contains a $(1 - \varepsilon)a \times (1 - \varepsilon)b$ plane rectangle. So using the projection π we may apply Lemma 3.1 to a family of $(1 - \varepsilon)\delta \times (1 - \varepsilon)b$ rectangles obtained from $\pi(\mathbb{T}_{B_i})$. Taking ε a small constant, say $\varepsilon = 0.1$, we get the desired conclusion. $\square$

3.2. Hairbrush estimate. For a set of $\delta \times 1$ tubes in the plane $\mathbb{T}$ we say that $\mathbb{T}$ is (t, C) -Katz-Tao of tubes, if for every $w \leq 1$ and every $w \times 1$ box B we have

$$|\mathbb{T} \cap B| \leq C(w/\delta)^t.$$

This is a convenient spacing condition for studying incidence geometry of tubes and it appears naturally in the application to our point-line incidence problem.

We say that a set of $\delta \times 1$ -tubes $\mathbb{T}$ in $\mathbb{R}^3$ satisfies the Katz-Tao Convex Wolff axiom with exponents (t_1, t_2) and error C if for any $u \leq w \leq 1$ and $u \times w \times 1$ box B we have

$$|\mathbb{T} \cap B| \leq C(u/\delta)^{t_1}(w/\delta)^{t_2}.$$

For shorthand, we say $\mathbb{T}$ is a (t_1, t_2, C) -Katz-Tao set of tubes. Katz-Tao and similar notions of concentration were introduced in the works of Katz and Tao [4] and are a ubiquitous tool in fractal geometry and harmonic analysis. The notion of (t_1, t_2, C) -Katz-Tao sets of tubes is closely related to Wolff axioms [17] which is often used to study Kakeya and Furstenberg sets [15]. In particular, Convex Wolff Katz-Tao axioms used in the recent resolution of the Kakeya conjecture in $\mathbb{R}^3$ [16] correspond to $(1, 1, C)$ -Katz-Tao sets.

In Proposition 5.1 below we show that the set of δ -tubes defined by a point-line configuration X with $d(X) \geq \delta$ forms a $(1 + \gamma, 2 - \gamma, \delta^{-\varepsilon})$ -Katz-Tao set of tubes where $\gamma \in [0, 1]$ is a number such that $\text{PL}_2(\gamma)$ holds. In this section we use Wolff's hairbrush argument originating from [17] to prove an incidence estimate on tubes satisfying the (t_1, t_2, C) -Katz-Tao condition. Our argument only relies on the 'classical' techniques and does not use any of the recent developments in the area. This makes the paper essentially self-contained but leaves open a direction for potential optimization of our methods.

Given a set of tubes $\mathbb{T}$ a shading Y on $\mathbb{T}$ is a collection of subsets $Y(T) \subset T$ for every $T \in \mathbb{T}$. We say that $Y(T)$ is λ -dense if $|Y(T)| \geq \lambda|T|$. First we prove an estimate on the volume of the union of a shading $\bigcup_{T \in \mathbb{T}} Y(T)$ in terms of shading density provided that $\mathbb{T}$ is t -Katz-Tao. The estimate we prove follows from Wolff's two-ends argument and is far from optimal. The Furstenberg set estimate [10] can be used to prove a sharp estimate. Since it is not necessary for our application we opted for a more elementary argument instead.

Lemma 3.3. *Let $\mathbb{T}$ be a set of δ -tubes in the plane which is a (t, K) -Katz-Tao for some $t \geq 1$ and $K > 0$. Let $Y(T) \subset T$ be a λ -dense set for some $\lambda \in [\delta, 1]$. Then for every $\varepsilon > 0$:*

$$\left| \bigcup_{T \in \mathbb{T}} Y(T) \right| \gtrsim_{\varepsilon} \delta^{\varepsilon} K^{-1} \lambda^2 \delta^t |\mathbb{T}|$$

Proof. By dyadic pigeonholing, we can find a dyadic number μ and a subset $Y'(T) \subset Y(T)$ for every $T \in \mathbb{T}$ so that each point in $U' = \bigcup_{T \in \mathbb{T}} Y'(T)$ is contained in $\sim \mu$ shadings $Y'(T)$ and so that we have $\sum |Y'(T)| \gtrsim \frac{1}{\log(1/\delta)} \sum |Y(T)|$. By double counting, we have

$$(12) \quad \mu |U'| \sim \sum_{T \in \mathbb{T}} |Y'(T)| \gtrsim \frac{1}{\log(1/\delta)} |\mathbb{T}| \lambda \delta.$$

Next, we perform a two-ends reduction. Let $\varepsilon > 0$ be arbitrarily small. We may assume that $\mu \geq C\delta^{-\varepsilon}$ since otherwise (12) already gives a sufficiently good bound. For every $p \in U'$ let $\mathbb{T}_{Y'}(p)$ be the set of tubes $T \in \mathbb{T}$ so that $p \in Y'(T)$. For $p \in U'$, let $T_{p,\theta}$ be a $1 \times \theta$ tube for some dyadic $\theta \in [\delta, 1]$ which maximizes the quantity

$$(13) \quad |\mathbb{T}_{Y'}(p)[T_{p,\theta}]| \theta^{-\varepsilon},$$

where $\mathbb{T}_{Y'}(p)[T_p]$ consists of tubes $T \in \mathbb{T}_{Y'}(p)$ so that $|T \cap T_p| \geq |T|/2$. Note that we have by definition

$$(14) \quad |\mathbb{T}_{Y'}(p)[T_p]| \gtrsim \theta^\varepsilon |\mathbb{T}_{Y'}(p)| \geq \delta^\varepsilon |\mathbb{T}_{Y'}(p)|.$$

Let $Y''(T) \subset Y'(T)$ be a new shading where we include p in $Y''(T)$ iff $T \in \mathbb{T}_{Y'}(p)[T_p]$. Then we get

$$\sum |Y''(T)| = \int_{U'} |\mathbb{T}_{Y'}(p)[T_p]| \gtrsim \delta^\varepsilon \int_{U'} |\mathbb{T}_{Y'}(p)| = \delta^\varepsilon \sum |Y'(T)|.$$

Finally, we can pigeonhole the value of θ and define a shading $\tilde{Y}(T) = Y''(T) \cap U'_\theta$ where $p \in U'_\theta$ if T_p has dimensions $1 \times \theta$. By choosing θ appropriately, we can guarantee that $\sum |\tilde{Y}(T)| \gtrsim \frac{1}{\log(1/\delta)} \sum |Y''(T)|$.

After these reductions, it suffices to lower bound the volume of $\tilde{U} = \bigcup \tilde{Y}(T)$.

Note that for any $p \in \tilde{U}$, the set of tubes $\mathbb{T}_{\tilde{Y}}(p)$ is contained in some $C\theta \times C$ tube (for some constant C). Using this, we can construct a collection of $C\theta \times C$ tubes $\mathbb{T}_\theta$ with the following two properties: for every $T, T' \in \mathbb{T}$ so that $\tilde{Y}(T) \cap \tilde{Y}(T') \neq \emptyset$ there exists $T_\theta \in \mathbb{T}_\theta$ so that $T, T' \subset T_\theta$, and for every $T \in \mathbb{T}$ there are only constantly many $T_\theta \in \mathbb{T}_\theta$ so that $T \subset T_\theta$. For $T_\theta \in \mathbb{T}_\theta$ let $\mathbb{T}[T_\theta]$ be the set of tubes contained in T_θ and $\tilde{U}_{T_\theta} = \bigcup_{\mathbb{T}[T_\theta]} \tilde{Y}(T)$. We get $|\tilde{U}| \gtrsim \sum_{\mathbb{T}_\theta} |\tilde{U}_{T_\theta}|$.

Consider the following set

$$S = \{(p, T, T') \in \tilde{U} \times \mathbb{T} \times \mathbb{T} : \exists T_\theta \in \mathbb{T}_\theta : T, T' \in \mathbb{T}[T_\theta], p \in \tilde{Y}(T) \cap \tilde{Y}(T'), \angle T, T' \geq \alpha\theta\}$$

where $\alpha = 2^{-C\varepsilon^{-1}}$ is a constant depending on ε . Using the maximality of T_p in (13), every $\alpha\theta \times 1$ tube T'_p contains at most $\alpha^\varepsilon |\mathbb{T}_{Y'}(p)[T_p]|$ tubes $T \in \mathbb{T}_{Y'}(p)$. By definition, for $p \in \tilde{U}$ we have $\mathbb{T}_{Y'}(p)[T_p] = \mathbb{T}_{Y''}(p) = \mathbb{T}_{\tilde{Y}}(p)$, so for every $p \in \tilde{U}$, we get that there are at least $|\mathbb{T}_{\tilde{Y}}(p)|(|\mathbb{T}_{\tilde{Y}}(p)| - C\alpha^\varepsilon |\mathbb{T}_{\tilde{Y}}(p)|) \geq |\mathbb{T}_{\tilde{Y}}(p)|^2/2$ pairs $(T, T') \in \mathbb{T}_{\tilde{Y}}(p) \times \mathbb{T}_{\tilde{Y}}(p)$ which form an angle at least $\alpha\theta$. We conclude that

$$\begin{aligned} |S| &:= \int_{\tilde{U}} \#\{(T, T') : (p, T, T') \in S\} \geq \frac{1}{2} \int_{\tilde{U}} |\mathbb{T}_{\tilde{Y}}(p)|^2 \\ &\geq \frac{1}{2} \frac{1}{|\tilde{U}|} \left(\int_{\tilde{U}} |\mathbb{T}_{\tilde{Y}}(p)| \right)^2 = \frac{1}{2} \frac{1}{|\tilde{U}|} \left(\sum_{T \in \mathbb{T}} |\tilde{Y}(T)| \right)^2 \\ &\gtrsim \delta^{3\varepsilon} \frac{1}{|\tilde{U}|} \lambda^2 \delta^2 |\mathbb{T}|^2. \end{aligned}$$

On the other hand, for any T, T' forming an angle at least $\alpha\theta$, we have $|\tilde{Y}(T) \cap \tilde{Y}(T')| \leq |T \cap T'| \lesssim \delta^2/\alpha\theta$. So we obtain

$$|S| \lesssim \sum_{T_\theta \in \mathbb{T}_\theta} |\mathbb{T}[T_\theta]|^2 (\delta^2/\alpha\theta) \lesssim_\varepsilon (\delta^2/\theta) |\mathbb{T}| \max_{T_\theta \in \mathbb{T}_\theta} |\mathbb{T}[T_\theta]|.$$

By the Katz–Tao axiom, we have $|\mathbb{T}[T_\theta]| \lesssim K(\theta/\delta)^t$. So by combining upper and lower bounds on S , we obtain

$$\begin{aligned} \delta^{3\varepsilon} \frac{1}{|\tilde{U}|} \lambda^2 \delta^2 |\mathbb{T}|^2 &\lesssim (\delta^2/\theta) |\mathbb{T}| K(\theta/\delta)^t \\ |\tilde{U}| &\gtrsim \delta^{3\varepsilon} K^{-1} \lambda^2 \theta^{1-t} \delta^t |\mathbb{T}| \gtrsim \delta^{3\varepsilon} K^{-1} \lambda^2 \delta^t |\mathbb{T}| \end{aligned}$$

since $t \geq 1$, as desired. $\square$

It will be convenient to use the following corollary.

Corollary 3.4. *Let $u \in (\delta, 1)$ and let $\mathbb{T}$ be a set of δ -tubes contained in a $Cu \times C$ rectangle R and which is a (t, K) -Katz–Tao for some $t \geq 1$ and $K > 0$. Let $Y(T) \subset T$ be a λ -dense set for some $\lambda \in [\delta, 1]$. Then for every $\varepsilon > 0$:*

$$\left| \bigcup Y(T) \right| \gtrsim_\varepsilon \delta^\varepsilon K^{-1} \lambda^2 \delta^t u^{1-t} |\mathbb{T}|$$

Proof. Let $\psi : R \rightarrow [0, C]^2$ be the natural rescaling map and consider the set of $\sim \delta/u$ -tubes $\mathbb{T}' = \psi(\mathbb{T})$. It is easy to check that $\mathbb{T}'$ is $(C'K, t)$ -Katz-Tao, so Lemma 3.3 gives

$$|\bigcup \psi(Y(T))| \gtrsim_{\varepsilon} \delta^{\varepsilon} K^{-1} \lambda^2 (\delta/u)^t |\mathbb{T}|$$

Since ψ increases volume by a factor of u^{-1} , the result follows. $\square$

Now we use Lemma 3.3 (and actually Corollary 3.4) to prove a volume estimate for shadings on a (t_1, t_2, C) -Katz-Tao set of tubes in $\mathbb{R}^3$. This estimate is a generalization of Wolff's hairbrush [17] estimate which was developed by Wolff to prove that Kakeya sets in $\mathbb{R}^3$ have dimension at least $5/2$. Needless to say that this estimate is far from being optimal and sharp volume estimates for (t_1, t_2, C) -Katz-Tao sets are currently unavailable: even in case of Kakeya $t_1 = t_2 = 1$ it is essentially equivalent to the maximal Kakeya conjecture.

Lemma 3.5. *Let $\mathbb{T}$ be a set of δ -tubes in $\mathbb{R}^3$ which is (t_1, t_2, K) -Katz-Tao for some $t_1, t_2 \geq 1$ and $K > 0$. Let Y be a λ -dense shading on $\mathbb{T}$, then we have*

$$(15) \quad \left| \bigcup Y(T) \right| \gtrsim_{\varepsilon} \delta^{\varepsilon} K^{-\frac{2+t_1}{2t_1+2t_2}} \lambda^{5/2} \delta^2 |\mathbb{T}|^{\frac{2+t_1}{2t_1+2t_2}}.$$

Proof. By repeating the dyadic pigeonholing and two-ends reduction from Lemma 3.3, we can find a dyadic scale $\theta \in [\delta, 1]$ and a subset $\tilde{Y}(T) \subset Y(T)$ for every $T \in \mathbb{T}$ so that for every $p \in \tilde{U} = \bigcup_{T \in \mathbb{T}} \tilde{Y}(T)$ there is a θ -tube T_p so that $\mathbb{T}_{\tilde{Y}}(p) \subset \mathbb{T}_Y(p)[T_p]$ and T_p maximizes (13). As previously, define a collection of $C\theta \times C$ -tubes $\mathbb{T}_{\theta}$ covering $\mathbb{T}$ and define $\tilde{U}_{T_{\theta}} = \bigcup_{T \in \mathbb{T}_{\theta}} \tilde{Y}(T)$. After these refinements, we will have $|\tilde{U}| \gtrsim \sum_{\mathbb{T}_{\theta}} |\tilde{U}_{T_{\theta}}|$ and $\sum |\tilde{Y}(T)| \gtrsim \delta^{\varepsilon} (\log 1/\delta)^{-O(1)} \sum |Y(T)|$.

For a $C\delta \times C\theta \times C$ slab H we can view the set of tubes $\mathbb{T}[H]$ contained in H as essentially a planar configuration of tubes: we can slice H by a random plane P parallel to the two major axes of H and study the intersections $\tilde{Y}(T) \cap P$, $T \in \mathbb{T}[H]$. Note that for any $C\delta \times Cw \times C$ subslab $B \subset H$ the (t_1, t_2, K) -Katz-Tao axiom implies $|\mathbb{T}[B]| \lesssim K(w/\delta)^{t_2}$. So we get that the collection of tubes we obtain inside the plane P is (t_2, CK) -Katz-Tao. So for any H and subsets $Y'(T) \subset T$ we get the following estimate using Corollary 3.4:

$$(16) \quad \left| \bigcup_{T \in \mathbb{T}[H]} Y'(T) \right| \gtrsim \int_0^{C\delta} \left| \bigcup_{T \in \mathbb{T}[H]} Y'(T) \cap P_x \right| dx \gtrsim_{\varepsilon} \delta^{\varepsilon} K^{-1} \delta^{1+t_2} \theta^{1-t_2} |\mathbb{T}[H]| \lambda_{H,Y'}^2,$$

where P_x denotes a plane distance x away from a fixed face of H and $\lambda_{H,Y'}$ is the average density of shadings $Y'(T)$ over $T \in \mathbb{T}[H]$.

Now we fix $T_{\theta} \in \mathbb{T}_{\theta}$. Let $r = \delta^{C\varepsilon} \lambda$ be a parameter. Similarly to the proof of Lemma 3.3 consider the following set:

$$S_{T_{\theta}} = \{(p, T, T') \in \tilde{U}_{T_{\theta}} \times \mathbb{T}[T_{\theta}] \times \mathbb{T}[T_{\theta}] : p \in \tilde{Y}(T) \cap \tilde{Y}(T'), \angle T, T' \geq \alpha\theta\}.$$

Then we analogously to before, get for $\alpha = 2^{-C\varepsilon-1}$:

$$(17) \quad |S_{T_{\theta}}| \geq \frac{1}{2} \int_{\tilde{U}_{T_{\theta}}} |\mathbb{T}_{\tilde{Y}}(p)|^2 \gtrsim \frac{1}{|\tilde{U}_{T_{\theta}}|} \left(\sum_{T \in \mathbb{T}[T_{\theta}]} |\tilde{Y}(T)| \right)^2 \sim \frac{1}{|\tilde{U}_{T_{\theta}}|} \lambda_{T_{\theta}}^2 \delta^4 |\mathbb{T}[T_{\theta}]|^2,$$

where $\lambda_{T_{\theta}}$ is the average density of shadings $\tilde{Y}(T)$ for $T \in \mathbb{T}[T_{\theta}]$.

For every $T \in \mathbb{T}[T_{\theta}]$ let $\mathbb{T}(T)$ be the set of $T' \in \mathbb{T}[T_{\theta}]$ such that $\tilde{Y}(T) \cap \tilde{Y}(T') \neq \emptyset$, $\angle T, T' \geq \alpha\theta$. By definition, we have

$$(18) \quad |S_{T_{\theta}}| = \sum_{T \in \mathbb{T}[T_{\theta}]} \sum_{T' \in \mathbb{T}(T)} |\tilde{Y}(T) \cap \tilde{Y}(T')| \lesssim \sum_{T \in \mathbb{T}[T_{\theta}]} |\mathbb{T}(T)| (\delta^3 / \alpha\theta).$$

Note that for $T' \in \mathbb{T}(T)$ we have $|\tilde{Y}(T') \setminus \mathcal{N}_{r\theta}(T)| \geq |\tilde{Y}(T')| - C\delta^2 r / \alpha$, where $\mathcal{N}_{r\theta}(T_0)$ is the r -neighbourhood of T .

For a tube $T \in \mathbb{T}[T_\theta]$ let us consider a collection of $C\delta \times C\theta \times C$ slabs $H_1, \dots, H_m$, $m \sim \theta/\delta$, which pass through T and cover T_θ . Consider the shading $Y'(T') = \tilde{Y}(T') \setminus \mathcal{N}_{r\theta}(T)$. Applying (16) to each $\mathbb{T}(T)[H_i]$ with shading Y' gives:

$$(19) \quad |\tilde{U}_{T_\theta}| \gtrsim r \sum_{i=1}^m \left| \bigcup_{T \in \mathbb{T}(T)[H_i]} Y'(T') \right| \gtrsim \delta^\varepsilon r K^{-1} \delta^{1+t_2} \theta^{1-t_2} \sum_{i=1}^m |\mathbb{T}(T)[H_i]| \lambda_{H_i, Y'}^2 \\ \gtrsim \delta^\varepsilon r K^{-1} \delta^{1+t_2} \theta^{1-t_2} |\mathbb{T}(T)| \lambda_{T, Y'}^2$$

where we used the fact that sets $H_i \setminus \mathcal{N}_{r\theta}(T_0)$ are $\lesssim r^{-1}$ overlapping. Here $\lambda_{T, Y'}$ denotes the average density of the shading $Y'(T')$ for $T' \in \mathbb{T}(T)$ and we used Cauchy–Schwarz to convert the sum of squares of densities $\lambda_{H_i, Y'}^2$ to $\lambda_{T, Y'}^2$. To be more precise, we estimate it as follows: note that by definition

$$|\mathbb{T}(T)| \lambda_{T, Y'} = \sum_{i=1}^m |\mathbb{T}(T)[H_i]| \lambda_{H_i, Y'}$$

and so

$$|\mathbb{T}(T)|^2 \lambda_{T, Y'}^2 = \left(\sum_{i=1}^m |\mathbb{T}(T)[H_i]| \lambda_{H_i, Y'} \right)^2 = \left(\sum_{i=1}^m |\mathbb{T}(T)[H_i]|^{1/2} (|\mathbb{T}(T)[H_i]|^{1/2} \lambda_{H_i, Y'}) \right)^2 \\ \leq |\mathbb{T}(T)| \sum_{i=1}^m |\mathbb{T}(T)[H_i]| \lambda_{H_i, Y'}.$$

We deal with the density terms in the following computations in a similar manner. So by (18), we get

$$|\mathbb{T}[T_\theta]| |\tilde{U}_{T_\theta}| \gtrsim \delta^\varepsilon r K^{-1} \delta^{1+t_2} \theta^{1-t_2} \sum_{T \in \mathbb{T}[T_\theta]} |\mathbb{T}(T)| \lambda_{T, Y'}^2 \gtrsim \delta^\varepsilon r K^{-1} \delta^{1+t_2} \theta^{1-t_2} |S_{T_\theta}| (\alpha\theta/\delta^3) \lambda_{T_\theta, Y'}^2$$

and so (17) gives

$$|\mathbb{T}[T_\theta]| |\tilde{U}_{T_\theta}| \gtrsim \delta^\varepsilon r K^{-1} \delta^{1+t_2} \theta^{1-t_2} (\alpha\theta/\delta^3) \lambda_{T_\theta}^2 \lambda_{T_\theta, Y'}^2 \frac{1}{|\tilde{U}_{T_\theta}|} \delta^4 |\mathbb{T}[T_\theta]|^2 \\ (20) \quad |\tilde{U}_{T_\theta}|^2 \gtrsim \delta^\varepsilon r K^{-1} \delta^{2+t_2} \theta^{2-t_2} |\mathbb{T}[T_\theta]| \lambda_{T_\theta}^2 \lambda_{T_\theta, Y'}^2.$$

By the (t_1, t_2, K) -Katz–Tao axiom we have $|\mathbb{T}[T_\theta]| \lesssim M := \min\{K(\theta/\delta)^{t_1+t_2}, |\mathbb{T}|\}$, which we use to lower bound

$$|\mathbb{T}[T_\theta]|^{1/2} = |\mathbb{T}[T_\theta]| / |\mathbb{T}[T_\theta]|^{1/2} \gtrsim |\mathbb{T}[T_\theta]| M^{-1/2}.$$

Now we note that since $\lambda_{T_\theta} \geq \lambda_{T_\theta, Y'} \geq \lambda_{T_\theta} - Cr$ we can bound

$$\sum_{T_\theta \in \mathbb{T}_\theta} |\mathbb{T}[T_\theta]| \lambda_{T_\theta} \lambda_{T_\theta, Y'} \geq \sum_{T_\theta \in \mathbb{T}_\theta} |\mathbb{T}[T_\theta]| \lambda_{T_\theta, Y'}^2 \gtrsim |\mathbb{T}| \lambda_Y^2,$$

where λ_Y is the average density of $Y'(T)$ over all $T \in \mathbb{T}$. Finally, we recall that $\sum_{\mathbb{T}} |\tilde{Y}(T)| \gtrsim \delta^{2\varepsilon} \lambda \delta^2 |\mathbb{T}|$ and so if we take $r \ll \delta^{2\varepsilon} \lambda$ then we also get $\sum_{\mathbb{T}} |Y'(T)| \gtrsim \delta^{2\varepsilon} \lambda \delta^2 |\mathbb{T}|$. We conclude that

$$|\tilde{U}| \gtrsim \sum_{T_\theta \in \mathbb{T}_\theta} |\tilde{U}_{T_\theta}| \gtrsim \delta^\varepsilon r^{1/2} K^{-1/2} \delta^{(2+t_2)/2} \theta^{(2-t_2)/2} |\mathbb{T}| M^{-1/2} \lambda^2.$$

Now we observe

$$\theta^{(2-t_2)/2} |\mathbb{T}| \max\{|\mathbb{T}|^{-1/2}, K^{-1/2} (\delta/\theta)^{(t_1+t_2)/2}\} \\ \geq \theta^{(2-t_2)/2} |\mathbb{T}| \left(|\mathbb{T}|^{-1/2} \right)^{1 - \frac{2-t_2}{t_1+t_2}} \left(K^{-1/2} (\delta/\theta)^{(t_1+t_2)/2} \right)^{\frac{2-t_2}{t_1+t_2}} \\ = K^{-\frac{2-t_2}{2t_1+2t_2}} \delta^{\frac{2-t_2}{2}} |\mathbb{T}|^{\frac{2+t_1}{2t_1+2t_2}}$$

and so putting everything together gives

$$|\tilde{U}| \gtrsim \delta^{C\varepsilon} K^{-\frac{2+t_1}{2t_1+2t_2}} \lambda^{5/2} \delta^2 |\mathbb{T}|^{\frac{2+t_1}{2t_1+2t_2}}.$$

□

4. HIGH-LOW ESTIMATES IN $\mathbb{R}^3$

4.1. Basic high-low inequality in $\mathbb{R}^3$. Fix $\chi : \mathbb{R}^3 \rightarrow [0, 1]$ a compactly supported radial bump function such that $\chi(x) \geq 1/2$ for $|x| \leq 1/2$ and $\chi(x) = 0$ for $|x| \geq 2$. For $w > 0$ denote $\chi_w(x) = w^{-3} \chi(x/w)$ and define $\eta_w = \chi_w * \chi_{w/2}$.

If P and L are sets of points and lines in $\mathbb{R}^3$ then we define the smooth incidence count as

$$I(w; P, L) = w^2 \langle 1_P, \sum_{\ell \in L} 1_\ell * \eta_w \rangle$$

where $1_P = \sum_{p \in P} 1_p$ is the sum of delta functions at points of P and 1_ℓ is the delta function on ℓ . Explicitly, we have

$$\langle 1_p, 1_\ell * \eta_w \rangle = \int_{\ell} \eta_w(x - p) dx.$$

Define the normalized incidence function as

$$B(w; P, L) = \frac{I(w, P, L)}{w^2 |P| |L|} = \langle g, f * \eta_w \rangle$$

where $g = |P|^{-1} \sum_{p \in P} 1_p$ and $f = |L|^{-1} \sum_{\ell \in L} 1_\ell$ are the normalized indicator functions of the set of points and the set of lines. In what follows, we will for convenience denote $B(w) = B(w; P, L)$ in cases when the sets P, L are clear from the context.

We let $\mathcal{A}_{3,1}$ denote the set of lines in $\mathbb{R}^3$ which intersect the ball $B(0, 1)$ and we fix a metric on $\mathcal{A}_{3,1}$. For a set of lines $L \subset \mathcal{A}_{3,1}$ and $u \leq w \leq 1$ we define $M_L(u \times w \times 1)$ as the maximum over all $u \times w \times 1$ boxes Π of $|L \cap \Pi|$ where $L \cap \Pi = \{\ell \in L : |\ell \cap \Pi| \geq 1/2\}$.

Proposition 4.1. *For any L and $u \leq u' \leq 1$ and $w \leq w' \leq 1$ we have*

$$M_L(u' \times w' \times 1) \lesssim (u'/u)^2 (w'/w)^2 M_L(u \times w \times 1).$$

Proof. Given a collection of lines L , let Π' be a $u' \times w' \times 1$ prism such that $|\Pi' \cap L| = M_L(u' \times w' \times 1)$. Let $\Pi' \cap L = \{\ell_1, \dots, \ell_m\}$, we have $|\Pi' \cap \ell_i| \geq 1/2$ for all i . Consider the two planes P_1, P_2 which contain the $u' \times w'$ faces of the prism Π' and let $R'_1 \subset P_1, R'_2 \subset P_2$ be $4u' \times 4w'$ rectangles which are dilates of the corresponding faces of Π' . Then we have $\ell_i \cap P_j \in R'_j$ for $j = 1, 2$ and all i . Now we cover each R'_j by $\lesssim (u'/u)(w'/w)$ many disjoint $0.1u \times 0.1w$ rectangles $R_{j,t}$. By pigeonhole principle, there is a pair of rectangles R_{1,t_1}, R_{2,t_2} such that the number of indices i so that $\ell_i \cap P_j \in R_{j,t_j}$ is lower bounded by $c(u/u')^2 (w/w')^2 m$. One can check that the set of such lines is covered by some $u \times w \times 1$ prism Π . We conclude that $M_L(u \times w \times 1) \gtrsim (u/u')^2 (w/w')^2 m$, as desired. □

The next lemma is the simplest version of the high-low inequality in $\mathbb{R}^3$. This is essentially due to Guth–Solomon–Wang [3] who used it to prove incidence upper bounds for well-spaced tubes in $\mathbb{R}^3$. Later on, we will prove two inequalities which improve on this bound under certain restrictions on the set of lines L . Both of these improvements will be crucial ingredients in the proof of Theorem 1.3.

Lemma 4.2. *For any $\varepsilon > 0$, $\delta > 0$ and any $P \subset [-1, 1]^3$, $L \subset \mathcal{A}_{3,1}$ we have*

$$(21) \quad |B(\delta) - B(2\delta)|^2 \lesssim_{\varepsilon} \delta^{-6-\varepsilon} \frac{M_P(\delta)}{|P|} \frac{M_L(\delta \times \delta \times 1)}{|L|}.$$

Proof. Let g and f be the normalized indicator functions of P and L as defined above. By definition, we have

$$B(\delta) - B(2\delta) = \langle g * \eta_\delta, f \rangle - \langle g * \eta_{2\delta}, f \rangle = \langle g * \chi_\delta, f * (\chi_{\delta/2} - \chi_{2\delta}) \rangle$$

Since $P \subset [-1, 1]^3$ we may restrict the domain of integration to $[-2, 2]^3$. By the Cauchy-Schwarz inequality we then get

$$|B(\delta) - B(2\delta)| \leq \|g * \chi_\delta\|_{L^2([-2, 2]^3)} \|f * (\chi_{\delta/2} - \chi_{2\delta})\|_{L^2([-2, 2]^3)}.$$

We can estimate

$$\|g * \chi_\delta\|_{L^2([-2, 2]^3)}^2 \leq \|g * \chi_\delta\|_1 \|g * \chi_\delta\|_\infty = \|g * \chi_\delta\|_\infty \lesssim \frac{\delta^{-3} \mathbf{M}_P(\delta)}{|P|}$$

where we recall $\int_{\mathbb{R}^3} \chi_\delta = 1$ and $g * \chi_\delta(x) = |P|^{-1} \sum_{p \in P} \chi_\delta(x - p)$.

Letting $\psi = \chi_{\delta/2} - \chi_{2\delta}$, we now focus on $\|f * \psi\|_{L^2([-2, 2]^3)}^2$. For a large enough constant $C > 0$ we have

$$\|f * \psi\|_{L^2([-2, 2]^3)}^2 \lesssim \|[\chi_C \cdot f] * \psi\|_2^2.$$

Since χ is a fixed smooth function, we have the Fourier decay bound

$$|\hat{\psi}(\xi)| \lesssim_d (1 + \delta|\xi|)^{-d}$$

for any $d \geq 1$. On the other hand, note that $\hat{\chi}_w(0) = 1$ for all w and $\hat{\chi}$ is a smooth radially symmetric function. So the Taylor series of $\hat{\psi} = \hat{\chi}_{\delta/2} - \hat{\chi}_{2\delta}$ starts at degree 2, and so we have

$$|\hat{\psi}(\xi)| \lesssim (\delta|\xi|)^2.$$

We conclude that

$$(22) \quad |\hat{\psi}(\xi)| \lesssim_d (\delta|\xi|)^2 (1 + \delta|\xi|)^{-d}$$

for any fixed $d \geq 1$.

Thus, by Plancherel's theorem, we need to estimate

$$W(w, L) = \int_{|\xi| \sim w^{-1}} |\widehat{F}(\xi)|^2 d\xi, \quad \text{where } F = \chi_C \cdot f,$$

for all w . Note that F really is a distribution, not a function, but the quantity $W(w, L)$ still has a well-defined meaning. Consider a w -net $\Theta \subset S^2$ and decompose $L = \bigsqcup_{\theta \in \Theta} L_\theta$ where L_θ consists of lines ℓ with direction $\theta(\ell)$ (arbitrarily choosing between $\pm\theta(\ell)$) satisfying $d(\theta(\ell), \theta) \leq w$. Let $f_\theta = |L|^{-1} \sum_{\ell \in L_\theta} 1_\ell$ and $F_\theta = \chi_C \cdot f_\theta$.

For each $\theta \in \Theta$ we can upper bound

$$(23) \quad \begin{aligned} \int_{|\xi| \sim w^{-1}} |\widehat{F_\theta}(\xi)|^2 d\xi &\lesssim \int_{|\xi| \sim w^{-1}} |\widehat{F_\theta}(\xi)|^2 |\widehat{\chi}_{w/C}|^2 d\xi \leq \int_{\mathbb{R}^3} |\widehat{F_\theta}(\xi)|^2 |\widehat{\chi}_{w/C}|^2 d\xi = \\ &= \|F_\theta * \chi_{w/C}\|_2^2 \leq \|F_\theta * \chi_{w/C}\|_1 \|F_\theta * \chi_{w/C}\|_\infty \lesssim w^{-2} |L_\theta| \mathbf{M}_L(w \times w \times 1) |L|^{-2}. \end{aligned}$$

For $|\xi| \sim w^{-1}$, let $\theta(\xi) = \frac{\xi}{|\xi|}$. By the rapid decay of $\widehat{\chi}_C$, the function $\widehat{\chi_C \cdot 1_\ell}$ is essentially supported on the 1-neighborhood of the plane $\theta^\perp \subset \mathbb{R}^3$. This implies that we have

$$\widehat{F}(\xi) = \sum_{\theta \in \Theta: d(\theta, \theta(\xi)^\perp) \leq Kw} \widehat{F_\theta}(\xi) + O_d(w^{-2} K^{-d})$$

for any fixed $d \geq 1$. Since there are at most CKw^{-1} many $\theta \in \Theta$ in a strip of width Kw on S^2 , for any fixed ξ we have by Cauchy-Schwarz:

$$\left| \sum_{\theta \in \Theta: d(\theta, \theta(\xi)^\perp) \leq Kw} \widehat{F_\theta}(\xi) \right|^2 \leq CKw^{-1} \sum_{\theta \in \Theta: d(\theta, \theta(\xi)^\perp) \leq Kw} |\widehat{F_\theta}(\xi)|^2$$

and so

$$|\widehat{F}(\xi)|^2 \lesssim Kw^{-1} \sum_{\theta \in \Theta: d(\theta, \theta(\xi)^\perp) \leq Kw} |\widehat{F}_\theta(\xi)|^2 + O_d(w^{-4}K^{-d}) \leq Kw^{-1} \sum_{\theta \in \Theta} |\widehat{F}_\theta(\xi)|^2 + O_d(w^{-4}K^{-d}).$$

So by integrating this over all ξ and applying (23) this gives

$$\begin{aligned} \int_{|\xi| \sim w^{-1}} |\widehat{F}(\xi)|^2 &\lesssim Kw^{-1} \sum_{\theta} w^{-2} |L_\theta| M_L(w \times w \times 1) |L|^{-2} + O_d(w^{-7}K^{-d}) \\ &\lesssim Kw^{-3} \frac{M_L(w \times w \times 1)}{|L|} + O_d(w^{-7}K^{-d}). \end{aligned}$$

Thus, we obtain the following estimate for any $d \geq 1$ and any $K \geq 1$:

$$W(w, L) \lesssim Kw^{-3} \frac{M_L(w \times w \times 1)}{|L|} + O_d(w^{-7}K^{-d}).$$

By Proposition 4.1 for dyadic values of $w \in (K^{-1}\delta, C)$ we have $M_L(w \times w \times 1) \lesssim (1 + (w/\delta)^4) M_L(\delta \times \delta \times 1)$. So by (22) we can estimate $\|F * \psi\|_2^2$ by:

$$\begin{aligned} \|F * \psi\|_2^2 &\lesssim \int_{|\xi| \lesssim 1} |\widehat{F}\widehat{\psi}|^2 + \sum_{K^{-1}\delta < w \lesssim 1} W(w, L)(\delta w^{-1})^4 + \int_{|\xi| > K\delta^{-1}} |\widehat{F}\widehat{\psi}|^2 \\ &\lesssim \int_{|\xi| \lesssim 1} |\widehat{F}\widehat{\psi}|^2 + \sum_{K^{-1}\delta < w \lesssim 1} (K^2 w^{-3} \frac{M_L(w \times w \times 1)}{|L|})(\delta w^{-1})^4 + O_d(\delta^{-7}K^{-d}) \\ &\lesssim \int_{|\xi| \lesssim 1} |\widehat{F}\widehat{\psi}|^2 + \sum_{K^{-1}\delta < w \lesssim 1} (K^6 w \delta^{-4} \frac{M_L(\delta \times \delta \times 1)}{|L|})(\delta w^{-1})^4 + O_d(\delta^{-7}K^{-d}) \\ &\lesssim \int_{|\xi| \lesssim 1} |\widehat{F}\widehat{\psi}|^2 + K^7 \delta^{-3} \frac{M_L(\delta \times \delta \times 1)}{|L|} + O_d(\delta^{-7}K^{-d}). \end{aligned}$$

The first term in the upper bound is controlled by

$$\int_{|\xi| \lesssim 1} |\widehat{F}\widehat{\psi}|^2 \lesssim \delta^4 \int_{|\xi| \lesssim 1} |\widehat{F}|^2 \lesssim \delta^4 \int F^2 * \chi_c \lesssim \delta^4$$

which is negligible (in particular it is smaller than the second term), so altogether we have

$$\|f * \psi\|_{L^2([-2, 2]^3)}^2 \lesssim K^7 \delta^{-3} \frac{M_L(\delta \times \delta \times 1)}{|L|} + O_d(\delta^{-7}K^{-d}).$$

By choosing $K = \delta^{-\varepsilon/10}$ and $d = 10\varepsilon^{-1}$ we conclude the proof. $\square$

4.2. Improved high-low using two-ends decomposition on Fourier side. For a set of lines L we let $\theta(L) \subset S^2$ be the set of directions spanned by lines in L and for $w > 0$ we write $|\theta(L)|_w$ to be the w -covering number of the set of directions.

In the next theorem, we prove a refined version of the high-low inequality which takes into account information about the set of directions of L and the concentration of L in prisms.

Theorem 4.1. *Let $P \subset [-1, 1]^3$ and $L \subset \mathcal{A}_{3,1}$ be finite collections of points and lines and let $\delta > 0$. Then we have*

$$(24) \quad |B(\delta) - B(2\delta)|^2 \lesssim_\varepsilon \delta^{-6-\varepsilon} \frac{M_P(\delta)}{|P|} \max_{u \in (\delta, 1)} \left(\min\{|\theta(L)|_\delta^{1/2} \delta, u\} u \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \right).$$

To prove this estimate, we run the proof of Lemma 4.2 up until the point where we upper bound the number of angles θ is a strip of width Kw by the trivial upper bound CKw^{-1} . At this stage, if we know that $|\Theta| \ll w^{-2}$, we can apply the two-ends decomposition (Lemma 3.1) to obtain that

multiplicity is upper bounded by $|\Theta|^{1/2} \ll w^{-1}$, apart from a few shorter segments where multiplicity can still be large. To deal with those segments, we can rescale the picture and repeat the argument.

Before the proof, let us formulate a useful corollary which gives sufficient conditions to improve over the basic high-low estimate from Lemma 4.2.

Corollary 4.3. *Let $P \subset [-1, 1]^3$ and $L \subset \mathcal{A}_{3,1}$ be finite collections of points and lines and let $\delta > 0$. Suppose that for some $\kappa \in [0, 1]$, $\nu > 0$ and $M \geq 1$ we have $|\theta(L)|_\delta \leq \nu\delta^{-2}$ and $M_L(\delta \times \delta/u \times 1) \leq u^{-2+\kappa}M$ for any $u \in (\delta, 1]$. Then we have*

$$|B(\delta) - B(2\delta)|^2 \lesssim_\varepsilon \nu^{\kappa/4} \delta^{-6-\varepsilon} \frac{M_P(\delta)}{|P|} \frac{M}{|L|}$$

Proof. Note that taking $u = 1$ implies that $M_L(\delta \times \delta \times 1) \leq M$. We use (24) and split into two ranges: if $u \in [\nu^{1/4}, 1]$ then

$$\min\{|\theta(L)|_\delta^{1/2} \delta, u\} u \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \leq \nu^{1/2} u \frac{M_L(\delta \times \delta/u \times 1)}{|L|}$$

we have $M_L(\delta \times \delta/u \times 1) \lesssim u^{-2} M_L(\delta \times \delta \times 1)$ (by Proposition 4.1) and so the right hand side is at most $\nu^{1/4} \frac{M_L(\delta \times \delta \times 1)}{|L|} \lesssim \nu^{\kappa/4} \frac{M}{|L|}$. For $u \in [\delta, \nu^{1/4}]$ we have

$$\min\{|\theta(L)|_\delta^{1/2} \delta, u\} u \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \leq u^2 \frac{u^{-2+\kappa} M}{|L|} \leq \nu^{\kappa/4} \frac{M}{|L|}.$$

□

Proof of Theorem 4.1. For $\varepsilon > 0$ take $K = \delta^{-\varepsilon/100}$ and $\Delta = \delta^{\varepsilon/100}$ and $d = 10^4 \varepsilon^{-1}$.

Let $g = \frac{1}{|P|} 1_P$ and $f = \frac{1}{|L|} \sum_{\ell \in L} 1_\ell$ be the normalized indicator functions of P and L , respectively. By repeating the proof of Lemma 4.2, we let $F = \chi_C \cdot f$ and $\psi = \chi_{\delta/2} - \chi_{2\delta}$ and estimate

$$(25) \quad |B(\delta) - B(2\delta)| \leq \|g * \chi_\delta\|_{L^2([-2, 2]^3)} \|f * \psi\|_{L^2([-2, 2]^3)} \lesssim (\delta^{-3} M_P(\delta) |P|^{-1})^{1/2} \|F * \psi\|_2$$

and reduce the problem to understanding the dyadic weights

$$W(w, L) = \int_{|\xi| \sim w^{-1}} |\widehat{F}|^2$$

for an arbitrary $w \in (K^{-1}\delta, 1)$. We split $L = \bigsqcup_{\theta \in \Theta} L_\theta$ where $\Theta = \Theta_w$ is a w -separated set of directions whose w -neighborhood covers $\theta(L)$ and $|\Theta| \sim |\theta(L)|_w$. Let $f_\theta = |L|^{-1} \sum_{\ell \in L_\theta} 1_\ell$ and $F_\theta = \chi_C \cdot f_\theta$ be the corresponding functions. As in the proof of Lemma 4.2 we have the estimate

$$(26) \quad \int_{|\xi| \sim w^{-1}} |\widehat{F}_\theta|^2 \lesssim w^{-2} |L_\theta| M_L(w \times w \times 1) |L|^{-2}$$

and we have an approximation

$$\widehat{F}(\xi) = \sum_{\theta \in \Theta: d(\theta, \theta(\xi)^\perp) \leq Kw} \widehat{F}_\theta(\xi) + O_d(w^{-2} K^{-d})$$

At this point in the proof of Lemma 4.2, we used the fact that the Kw -neighborhood of a great circle contains at most $\sim Kw^{-1}$ directions $\theta \in \Theta$. Here we observe that if $|\Theta|_w \ll w^{-2}$ then for most ξ we have much fewer such directions. To make this precise, we iteratively apply Corollary 3.2.

For $a \leq b \leq 2\pi$ we define a $a \times b$ rectangle on the sphere S^2 to be the a -neighborhood of a great circle segment of length b . Let Given an rectangle $U \subset S^2$, we can define a smooth bump function ρ_U supported on the rectangular set $\{|\xi| \in (w^{-1}/4, 2w^{-1}), \theta(\xi) \in 1.1U\}$ and equals 1 on the set $\{|\xi| \in (w^{-1}/2, w^{-1}), \theta(\xi) \in U\}$.

Denote $\delta_0 = Kw$. For $\theta \in \Theta$ let $U = U_\theta$ be the δ_0 -neighborhood of $\theta^\perp$ on S^2 and let $U_0 = \{U_\theta, \theta \in \Theta\}$. Fix some $\Delta > 0$ and let $m = C \frac{\log(2/\delta_0)}{\log(2/\Delta)}$. For $j = 1, \dots$ we are going to construct a collection

of essentially distinct $20^j \delta_0 \times \Delta^j$ -rectangles $\mathbb{U}_j$ on S^2 inductively as follows. Given a collection of rectangles $\mathbb{U}_j$ with $j \geq 0$, we do the following:

- Consider a collection of balls $\{B_{j,i}\}$ of radius $C\Delta^j$ on S^2 which are $O(1)$ -overlapping and every $3\Delta^j$ -ball is contained in at least one $B_{j,i}$, and set $\mathbb{U}_j[B_i] = \{U \in \mathbb{U}_j : U \subset B_i\}$,
- Apply Corollary 3.2 (with $\Delta/1.1$ in place of Δ) to each $\mathbb{U}_j[B]$ to obtain a set of essentially distinct $10 \cdot 20^j \delta_0 \times \Delta^{j+1}/1.1$ rectangles $\mathbb{U}_{j+1,B}$ inside B . For each $U \in \mathbb{U}_j[B]$ we thus get some $\mathbb{U}(U, B) \subset \mathbb{U}_{j+1,B}$ of size at most m so that sets

$$1.5U \setminus \bigcup_{U' \in \mathbb{U}(U, B)} U'$$

are at most $C\Delta^{-2}|\mathbb{U}_j[B]|^{1/2}$ -overlapping.

- We let $\mathbb{U}_{j+1}$ be a set of essentially distinct $20^{j+1} \delta_0 \times \Delta^{j+1}$ rectangles covering the union of sets $\mathbb{U}_{j+1,B}$ over all balls B (here ‘essentially distinct’ means that, say, $|U \cap U'| \leq (1-c)|U|$ for any $U \neq U' \in \mathbb{U}_{j+1}$ and some constant $c > 0$). Further we let $\mathbb{U}(U) \subset \mathbb{U}_{j+1}$ be the subset of rectangles $U \in \mathbb{U}_{j+1}$ which contain at least one rectangle in $\bigcup_B \mathbb{U}(U, B)$ where the union is over all balls B containing U .

We run this process for $j = 0, \dots, j_0$ where j_0 is the minimum index such that $(\Delta/100)^j < \delta_0$ holds. We obtain collections of rectangles $\mathbb{U}_j$, $j = 0, \dots, j_0$ such that for each $j < j_0$ and $U \in \mathbb{U}_j$ we have a selected subset $\mathbb{U}(U) \subset \mathbb{U}_{j+1}$ of size at most m . It follows from construction that for any $j < j_0$, the collection of sets $1.1U \setminus (\bigcup_{U' \in \mathbb{U}(U)} U')$ over $U \in \mathbb{U}_j$ is at most

$$(27) \quad C\Delta^{-2} \max_i \{|\mathbb{U}_j[B_{j,i}]|^{1/2}\} \leq C\Delta^{-2}|\mathbb{U}_j|^{1/2} \leq C\Delta^{-2}(m^j|\Theta|)^{1/2}.$$

(This bound is quite wasteful but it will suffice for us.) For $j = j_0$ let us put for convenience $\mathbb{U}(U) = \emptyset$.

Now we use this data to decompose functions $\widehat{F}_\theta$ into pieces with good overlapping properties. For $U \in \mathbb{U}_j$ we define a function F_U essentially supported on the region $\bar{U} = \{|\xi| \sim w^{-1}, \theta(\xi) \in U\}$ inductively as follows. For $j = 0$ we define $G_U = \rho_U \cdot \widehat{F}_\theta$ for each $U = U_\theta \in \mathbb{U}_0$. Given $U \in \mathbb{U}_j$ for some $j \geq 0$ let us write $\mathbb{U}(U) = \{U_1, \dots, U_k\} \subset \mathbb{U}_{j+1}$ and define functions

$$(28) \quad F_U = G_U \cdot \prod_{t=1}^k (1 - \rho_{U_t}),$$

$$(29) \quad G_U^{(t)} = G_U \cdot \rho_{U_t} \prod_{t'=1}^{t-1} (1 - \rho_{U_{t'}}), \quad t = 1, \dots, k$$

so that $G_U = F_U + G_U^{(1)} + \dots + G_U^{(k)}$. We then define for each $U' \in \mathbb{U}_{j+1}$:

$$G_{U'} = \sum_{U \in \mathbb{U}_j, t: U_t = U'} G_U^{(t)}$$

and proceed with the construction to the next index j . By the definition of ρ_U , the function F_U is then supported on the set

$$\{\xi : |\xi| \sim w^{-1}, \theta(\xi) \in 1.1U \setminus \bigcup_{U' \in \mathbb{U}(U)} U'\}.$$

So by (27), the supports of functions F_U , $U \in \mathbb{U}_j$ for $j < j_0$ are at most $C\Delta^{-2}m^j|\Theta|^{1/2}$ overlapping. For any $j \in \{0, \dots, j_0\}$ the supports of F_U , $U \in \mathbb{U}_{j_0}$, are trivially at most $C(\Delta^j/20^j\delta_0)$ -overlapping (since this is a collection of essentially distinct $20^j\delta_0 \times \Delta^j$ -rectangles). By the choice of j_0 , we have $C(\Delta^{j_0}/20^{j_0}\delta_0) \leq C100^{j_0}$. Thus, for any $j \in \{0, \dots, j_0\}$ the supports of F_U , $U \in \mathbb{U}_j$ are $C \min(\Delta^{-2}(100m)^j|\Theta|^{1/2}, \Delta^j/20^j\delta_0)$ -overlapping (note that for $j = j_0$ the second term dominates).

By this construction, we obtain a decomposition

$$\begin{aligned}\widehat{F} &= \sum \rho_{U_\theta} \widehat{F}_\theta + \sum (1 - \rho_{U_\theta}) \widehat{F}_\theta, \\ \sum \rho_{U_\theta} \widehat{F}_\theta &= \sum_{j=0}^{j_0} \sum_{U \in \mathbb{U}_j} F_U.\end{aligned}$$

We can estimate

$$\int_{|\xi| \sim w^{-1}} \left| \sum (1 - \rho_{U_\theta}) \widehat{F}_\theta \right|^2 \lesssim_d w^{-3} K^{-d}$$

by the rapid decay of $1_{|\xi| \sim w^{-1}} \widehat{F}_\theta$ outside U_θ . The remainder sum $\sum \rho_{U_\theta} \widehat{F}_\theta$ can be estimated as

$$\begin{aligned}(30) \quad & \int_{|\xi| \sim w^{-1}} \left| \sum \rho_{U_\theta} \widehat{F}_\theta \right|^2 \leq (j_0 + 1) \sum_{j=0}^{j_0} \int_{|\xi| \sim w^{-1}} \left| \sum_{\mathbb{U}_j} F_U \right|^2 \\ & \lesssim j_0 \sum_j C \min(\Delta^{-2} (100m)^j |\Theta|^{1/2}, \Delta^j / 20^j \delta_0) \sum_j \int_{|\xi| \sim w^{-1}} |F_U|^2,\end{aligned}$$

where we used the overlap property of supports of functions F_U . It remains to estimate the L_2 -norm of the function F_U . Let L_U be the set of lines ‘contributing’ to F_U , namely L_U is the union of the line sets L_θ over all θ such that there exists a chain

$$U_\theta = U_0, U_1, \dots, U_j = U$$

where $U_{i+1} \in \mathbb{U}(U_i)$ for every $i = 0, \dots, j-1$. Observe that since $|\mathbb{U}(U)| \leq m$ we have

$$(31) \quad \sum_{\mathbb{U}_j} |L_U| \leq m^j |L|.$$

We can write

$$F_U = |L|^{-1} \sum_{\ell \in L_U} (\widehat{1_\ell \cdot \chi_C}) \cdot \rho_{U,\ell},$$

where $\rho_{U,\ell}$ is a smooth 1-bounded function supported on $\overline{U} = \{\xi : |\xi| \sim w^{-1}, \theta(\xi) \in 1.1U\}$ and composed of products and sums of various functions $\rho_{U'}$ over some $U' \in \mathbb{U}_{j'}$ with $j' \in \{1, \dots, j+1\}$, see (28), (29). This implies that $\check{\rho}_{U,\ell}$ is a linear combination of convolutions of various $\check{\rho}_{U'}$. For $U' \in \mathbb{U}_j$ recall that $\rho_{U'}$ was chosen to be a smooth bump function on $\overline{U'}$. Since U' is a spherical $20^j \delta_0 \times \Delta^j$ rectangle, it follows that $\overline{U'}$ is approximately a $w^{-1} 20^j \delta_0 \times w^{-1} \Delta^j \times w^{-1}$ box. So if we let $\overline{U'}^\vee$ be the $w/(20^j \delta_0) \times w/\Delta^j \times w$ box which is dual to $\overline{U'}$ and is centrally symmetric around the origin, then we have the decay estimate

$$|\check{\rho}_{U'}(x)| \lesssim_d K^{-d} |\overline{U'}|, \quad x \in \mathbb{R}^3 \setminus K \overline{U'}^\vee.$$

Now note that if $\rho_{U'}$ appears in the expression for $\rho_{U,\ell}$ then rectangles U and U' have approximately aligned major axes (by the last part of Corollary 3.2) and so the union of boxes $\overline{U'}^\vee$ over all such U' is contained in the symmetric around the origin $w/\delta_0 \times w/\Delta^{j+1} \times w$ box $\tilde{U}^\vee$ (where the axes of $\tilde{U}^\vee$ coincide with the axes of $\overline{U}$). So we conclude that

$$|\check{\rho}_{U,\ell}(x)| \lesssim_{d,m} K^{-d} |\overline{U}|, \quad x \in \mathbb{R}^3 \setminus K \tilde{U}^\vee.$$

By similar analysis we can show that $\|\check{\rho}_{U,\ell}\|_1 \lesssim_m 1$ and $\|\check{\rho}_{U,\ell}\|_\infty \lesssim |\overline{U}|$. Using this we estimate

$$\int_{|\xi| \sim w^{-1}} |F_U|^2 \lesssim \int_{\mathbb{R}^3} |F_U|^2 |\widehat{\chi}_{w/C}|^2 = \|\check{F}_U * \chi_{w/C}\|_2^2 \leq \|\check{F}_U * \chi_{w/C}\|_1 \|\check{F}_U * \chi_{w/C}\|_\infty$$

we have

$$\check{F}_U = |L|^{-1} \sum_{\ell \in L_U} (1_\ell \cdot \chi_C) * \check{\rho}_{U,\ell}$$

and so the L_1 -norm is upper bounded as

$$\|\check{F}_U * \chi_{w/C}\|_1 \leq C|L|^{-1} \sum_{\ell \in L_U} \|\check{\rho}_{U,\ell}\|_1 \lesssim \frac{|L_U|}{|L|}.$$

For the L_∞ -norm we estimate

$$|\check{F}_U * \chi_{w/C}(x)| \lesssim \frac{1}{|L|} \sum_{\ell \in L_U} 1_{x \in (\ell \cap B^3(0,C) + K\tilde{U}^\vee)} \|\check{\rho}_{U,\ell}\|_\infty + O(|\bar{U}|K^{-d})$$

since the direction of $\ell \in L_U$ is aligned with the longest axis of the box $K\tilde{U}^\vee$ (which by definition has length $w/\delta_0 = K$), we conclude that the set $(\ell \cap B^3(0,C) + K\tilde{U}^\vee)$ is contained in $C'K^2\tilde{U}^\vee$ for some constant C' . So the number of lines $\ell \in L_U$ such that x is contained in $C'K^2\tilde{U}^\vee$ is upper bounded by $M_L(C'K^2w \times C'K^2w/\Delta^{j+1} \times 1) \lesssim K^8\Delta^{-2}M_L(w \times w/\Delta^j \times 1)$. Using this we obtain

$$\|\check{F}_U * \chi_{w/C}\|_\infty \lesssim K^8 \frac{1}{|L|} M_L(w \times w/\Delta^{j+1} \times 1) |\bar{U}| + O(|\bar{U}|K^{-d}).$$

Take $K = \delta^{-\varepsilon/100}$ and $\Delta = \delta^{\varepsilon/100}$ and $d = 10^4\varepsilon^{-1}$. Note that then we have $m, j_0 \lesssim \varepsilon^{-1}$ and so in particular, terms of the shape $100^{j_0}, m^j$ are all $\lesssim_\varepsilon 1$. Recall that $|\bar{U}| \sim w^{-3}(20\Delta)^j\delta_0 \lesssim_\varepsilon \Delta^j w^{-2}$. Thus, we obtain

$$\int_{|\xi| \sim w^{-1}} |F_U|^2 \lesssim_\varepsilon \delta^{-\varepsilon/3} \Delta^j w^{-2} \frac{|L_U| M_L(w \times w/\Delta^j \times 1)}{|L|^2}$$

(cf. (26)). So plugging this into (30) and using (31) gives

$$(32) \quad W(w, L) = \int_{|\xi| \sim w^{-1}} |\hat{F}|^2 \lesssim_\varepsilon \delta^{-\varepsilon/2} \max_{u \in (w,1)} \min\{|\Theta|^{1/2}, uw^{-1}\} uw^{-2} \frac{M_L(w \times w/u \times 1)}{|L|}$$

Now we use the fact that $w^{-4}M_L(w \times w/u \times 1) \lesssim K^4\delta^{-4}M_L(\delta \times \delta/u \times 1)$ (by Proposition 4.1) we get that

$$W(w, L) \lesssim_\varepsilon \delta^{-2\varepsilon/3} \delta^{-4} \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \max_{u \in (w,1)} \min\{|\Theta|^{1/2}, uw^{-1}\} uw^2.$$

Recall that $|\Theta| = |\Theta_w| \sim |\theta(L)|_w \lesssim K^2|\theta(L)|_\delta$. So by the decay estimates (22) on $\hat{\psi}$ we conclude that

$$\begin{aligned} \|f * \psi\|_{L^2([-2,2]^3)}^2 &\lesssim \sum_w W(w, L) (\delta w^{-1})^2 (1 + \delta w^{-1})^{-d} \\ &\lesssim_\varepsilon \sum_w \delta^{-2\varepsilon/3} \delta^{-2} \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \max_{u \in (w,1)} \min\{|\theta(L)|_\delta^{1/2}, uw^{-1}\} u (1 + \delta w^{-1})^{-d} \\ &\lesssim_\varepsilon \delta^{-\varepsilon} \max_{u \in (\delta,1)} \left(\min\{|\theta(L)|_\delta^{1/2}, u\delta^{-1}\} u \delta^{-2} \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \right) + O(\delta^{-7}K^{-d}). \end{aligned}$$

so plugging this back into (25) gives

$$|B(\delta) - B(2\delta)|^2 \lesssim_\varepsilon \delta^{-3-\varepsilon} \frac{M_P(\delta)}{|P|} \max_{u \in (\delta,1)} \left(\min\{|\theta(L)|_\delta^{1/2}, u\delta^{-1}\} u \delta^{-2} \frac{M_L(\delta \times \delta/u \times 1)}{|L|} \right)$$

which concludes the proof. $\square$

4.3. Improved high-low using rescaling and hairbrush estimate. The next lemma provides an alternative approach to improving Lemma 4.2. The idea is to split δ -tubes into $\delta \times \delta \times \delta^{1/2}$ segments and to apply the simple high-low (Lemma 4.2) to the short tubes inside $\delta^{1/2}$ boxes. The crux of this approach is estimating how much can the $\delta \times \delta \times \delta^{1/2}$ segments cluster inside the $\delta^{1/2}$ boxes. By decomposing the set of δ -tubes into a union over thick $\delta^{1/2}$ tubes this can be restated as a lower bound on the volume of some shading of tubes inside a fixed thick tube. Assuming the original set of tubes is (t_1, t_2, C) -Katz–Tao we can use Lemma 3.5 to estimate this volume.

We say that a set of lines L is C_0 -uniform at scale δ if for every $\ell \in L$ the number of lines $\ell' \in L$ such that $\ell' \cap B(0, 1) \subset T_\ell(\delta)$ is at least $M_L(\delta \times \delta \times 1)/C_0$. Later on, we will introduce a more systematic approach to uniformity but this will suffice for now.

Theorem 4.2. *Let $P \subset [0, 1]^3$ and $L \subset \mathcal{A}_{3,1}$ be finite collections of points and lines. Suppose that $M_P(\delta) \leq A\delta^3|P|$ for some $A \geq 1$.*

Suppose that L is C_0 -uniform at scale δ and that L defines a (t_1, t_2, K) -Katz-Tao set of δ -tubes. Denote $\alpha = \frac{t_1+2}{2t_1+2t_2}$. Then we have

$$(33) \quad |B(\delta/2) - B(\delta)|^{9/2} \lesssim_\varepsilon C_0^{O(1)} \delta^{-\varepsilon} \delta^{-5/2} K^\alpha A^{7/2} \frac{M_L(\delta^{1/2} \times \delta^{1/2} \times 1)^{1-\alpha} M_L(\delta \times \delta \times 1)^\alpha}{|L|}$$

We will apply this estimate in the situations of the following type. Suppose that P, L are δ -separated sets of points and lines (say, coming from a point-line configuration X with $d(X) \geq \delta$) and that we have $|P|, |L| \approx_\varepsilon \delta^{-3}$. Say that we know from Proposition 5.1 that L defines a $(1 + \gamma, 2 - \gamma, \delta^{-\varepsilon})$ -Katz-Tao set of δ -tubes. Moreover, let us suppose that L is *well-spaced*, namely that lines in L are well distributed on scale $\delta^{1/2}$, i.e. $M_L(\delta^{1/2} \times \delta^{1/2} \times 1) \lesssim \delta^2 |L| \approx \delta^{-1}$ holds. Under these assumptions, we have $\alpha = \frac{1}{2} + \frac{\gamma}{6}$ and (33) gives

$$|B(\delta/2) - B(\delta)|^{9/2} \lesssim_\varepsilon \delta^{-5/2} \frac{M_L(\delta^{1/2} \times \delta^{1/2} \times 1)^{1-\alpha}}{|L|} \lesssim \delta^{-\frac{1}{2}+\alpha} = \delta^{\frac{7}{6}}$$

and so we get a polynomial improvement over Lemma 4.2 provided that $\gamma > 0$.

Proof of Theorem 4.2. Using C_0 -uniformity of L we can find a collection essentially distinct δ -tubes $\mathbb{T}$ and a finitely overlapping covering $L = \bigcup_{T \in \mathbb{T}} L[T]$ so that $|\mathbb{T}| \lesssim C_0 \frac{|L|}{M_L(\delta \times \delta \times 1)}$ and $1/C_0 \lesssim \frac{|L[T]|}{M_L(\delta \times \delta \times 1)} \lesssim 1$ for every $T \in \mathbb{T}$. Then $\mathbb{T}$ is a $(t_1, t_2, C_0 K)$ -Katz-Tao set of δ -tubes.

Denote $\Delta = \delta^{1/2}$. Decompose $P = \bigsqcup_{Q \in \mathcal{Q}} P[Q]$ where $\mathcal{Q} = \mathcal{Q}_\Delta$ is a collection of disjoint Δ -boxes covering P . Fix some parameter $\tau > 0$ and for each $\ell \in L$ consider the set $\mathcal{Q}(\ell)$ of Δ -boxes $Q \in \mathcal{Q}$ such that

$$(34) \quad \langle 1_\ell * (\eta_\delta - \eta_{\delta/2}), 1_{P[Q]} \rangle \geq \tau \Delta |P|.$$

Using the point-wise bound $|1_\ell * \eta_\delta| \lesssim \delta^{-2}$, for any $Q \in \mathcal{Q}$ intersecting $T_{2\delta}(\ell)$ we have crude bounds

$$|\langle 1_\ell * \eta_\delta, 1_{P[Q]} \rangle| + |\langle 1_\ell * \eta_{\delta/2}, 1_{P[Q]} \rangle| \lesssim \delta^{-2} |T_{2\delta}(\ell) \cap Q \cap P| \lesssim \delta^{-5/2} M_P(\delta) = A\Delta |P|$$

where we used that $T_{2\delta}(\ell) \cap Q$ can be covered by $\sim \delta^{-1/2}$ δ -balls. So we can estimate

$$\begin{aligned} \langle 1_\ell * (\eta_\delta - \eta_{\delta/2}), 1_P \rangle &= \sum_{Q \in \mathcal{Q}: Q \cap T_{2\delta}(\ell) \neq \emptyset} \langle 1_\ell * (\eta_\delta - \eta_{\delta/2}), 1_{P[Q]} \rangle \\ &\leq \sum_{Q \notin \mathcal{Q}(\ell)} \tau \Delta |P| + \sum_{Q \in \mathcal{Q}(\ell)} C A \Delta |P| \leq C |P| (\tau + |\mathcal{Q}(\ell)| A \Delta) \end{aligned}$$

where we estimated the number of $Q \cap T_{2\delta}(\ell) \neq \emptyset$ with $Q \notin \mathcal{Q}(\ell)$ by $C\Delta^{-1}$. Summing this over all lines $\ell \in L$ gives for some constant C

$$(35) \quad B(\delta) - B(\delta/2) \leq C\tau + C A |L|^{-1} \sum_{\ell \in L} \Delta |\mathcal{Q}(\ell)|.$$

For the sake of contradiction, let us assume that $B(\delta) - B(\delta/2) \geq 2C\tau$ holds. From (35) we then get

$$(36) \quad |L|^{-1} \sum_{\ell \in L} \Delta |\mathcal{Q}(\ell)| \geq \tau A^{-1} =: \lambda_0.$$

For $T \in \mathbb{T}$ let us define $\mathcal{Q}(T) = \bigcup_{\ell \in L[T]} \mathcal{Q}(\ell)$. By pigeonholing, we can find some dyadic $\lambda \in [\lambda_0, C]$ and a subset $\mathbb{T}' \subset \mathbb{T}$ such that $|\mathcal{Q}(T)| \sim \lambda \Delta^{-1}$ for $T \in \mathbb{T}'$ and $|\mathbb{T}'| \gtrsim \frac{\lambda_0}{\lambda \log(C/\lambda_0)} |\mathbb{T}|$.

Decompose $\mathbb{T}' = \bigsqcup_{T_\Delta \in \mathbb{T}_\Delta} \mathbb{T}'[T_\Delta]$ where $\mathbb{T}_\Delta$ is a collection of essentially distinct Δ -tubes covering $\mathbb{T}'$. Let $L[T_\Delta]$ be the union of $L[T]$ over all $T \in \mathbb{T}'[T_\Delta]$.

Let $\mathbb{U}$ be a maximal by inclusion set of essentially distinct $2\delta \times 2\delta \times 4\Delta$ -tubes such that for every $U \in \mathbb{U}$ there exists $T \in \mathbb{T}'$ and $Q \in \mathcal{Q}(T)$ so that $T \cap 2Q \subset U$. We estimate the size of $\mathbb{U}$ in two different ways. First, for $T_\Delta \in \mathbb{T}_\Delta$ define $\mathbb{U}[T_\Delta]$ to be the set of segments U contained in $2T$. Observe that since tubes in $\mathbb{T}_\Delta$ are essentially distinct, the sets $\mathbb{U}[T_\Delta]$ are finitely-overlapping and so

$$(37) \quad |\mathbb{U}| \gtrsim \sum_{T_\Delta} |\mathbb{U}[T_\Delta]|.$$

Second, for $Q \in \mathcal{Q}$ set $\mathbb{U}[Q]$ be the set of U contained in $2Q$, then we have an upper bound

$$(38) \quad |\mathbb{U}| \lesssim \sum_Q |\mathbb{U}[Q]|.$$

Fix $T_\Delta \in \mathbb{T}_\Delta$. Define a shading $Y = Y_{T_\Delta}$ on $\mathbb{T}'[T_\Delta]$ as follows: for a tube T we set

$$Y(T) = \bigcup_{Q \in \mathcal{Q}(T)} T \cap 2Q$$

By the construction of $\mathbb{T}'$, the shading Y has density at least λ .

Let $\psi = \psi_{T_\Delta}$ be an affine map which rescales T_Δ into the unit cube and define $\tilde{\mathbb{T}}$ to be the set of $\Delta = \delta/\Delta$ -tubes $\psi(T)$, $T \in \mathbb{T}'[T_\Delta]$. Since $\mathbb{T}' \subset \mathbb{T}$ and $\mathbb{T}$ is (t_1, t_2, CK) -Katz–Tao by assumption, it follows that $\tilde{\mathbb{T}}$ is a (t_1, t_2, CK) -Katz–Tao set of Δ -tubes.

By Lemma 3.5 applied to $\tilde{\mathbb{T}}$ and the rescaled shading $\tilde{Y}$ (which still has density at least λ) we obtain

$$\left| \bigcup_{\tilde{\mathbb{T}}} \tilde{Y}(\tilde{T}) \right| \gtrsim_\epsilon \delta^\epsilon K^{-\frac{2+t_1}{2t_1+2t_2}} \lambda^{5/2} \Delta^2 |\tilde{\mathbb{T}}|^{\frac{2+t_1}{2t_1+2t_2}}.$$

Since ψ increases the volume by a factor of Δ^{-2} we have

$$|\mathbb{U}[T_\Delta]|(\delta \times \delta \times \Delta) \gtrsim \left| \bigcup_{\mathbb{T}'[T_\Delta]} Y(T) \right| \sim \Delta^2 \left| \bigcup_{\tilde{\mathbb{T}}} \tilde{Y}(\tilde{T}) \right|$$

and so we get a lower bound

$$(39) \quad |\mathbb{U}[T_\Delta]| \gtrsim_\epsilon \delta^\epsilon K^{-\frac{2+t_1}{2t_1+2t_2}} \lambda^{5/2} \Delta^{-1} |\mathbb{T}'[T_\Delta]|^{\frac{2+t_1}{2t_1+2t_2}}$$

We have for every $T_\Delta \in \mathbb{T}_\Delta$ that $|\mathbb{T}'[T_\Delta]| \lesssim \frac{M_L(\Delta \times \Delta \times 1)}{M_L(\delta \times \delta \times 1)} =: M$. So since $\alpha := \frac{2+t_1}{2t_1+2t_2} \in [0, 1]$, we obtain $\sum |\mathbb{T}'[T_\Delta]|^\alpha \gtrsim M^\alpha (|\mathbb{T}'|/M) = M^{\alpha-1} |\mathbb{T}'|$ and so

$$(40) \quad |\mathbb{U}| \gtrsim \sum_{T_\Delta} |\mathbb{U}[T_\Delta]| \gtrsim K^{-\alpha} \lambda^{5/2} \delta^{-1/2+2\epsilon} M^{\alpha-1} |\mathbb{T}'|.$$

On the other hand, fix $Q \in \mathcal{Q}$ and let $\psi = \psi_Q : Q \rightarrow [0, 1]^3$ be the affine rescaling map. Let $P^Q = \psi(P[Q])$. Let L_Q to be the set of lines $\ell \in L'$ is such that $Q \in \mathcal{Q}(\ell)$ and let $L[Q] \subset L_Q$ be a maximal $\sim \delta$ -separated collection. Note that we have $|\mathbb{U}[Q]| \lesssim |L[Q]|$. Define L^Q to be the set of lines $\psi(\ell)$ with $\ell \in L[Q]$. Observe that by a change of variables for any $w > 0$:

$$B(w; P^Q, L^Q) = \left\langle |P^Q|^{-1} 1_{P^Q} * \eta_w, |L^Q|^{-1} \sum_{L^Q} 1_\ell \right\rangle = \Delta^{-2} \left\langle |P[Q]|^{-1} 1_{P[Q]} * \eta_{w\Delta}, |L[Q]|^{-1} \sum_{L[Q]} 1_\ell \right\rangle$$

so using (34) we get

$$(41) \quad B(\Delta; P^Q, L^Q) - B(\Delta/2; P^Q, L^Q) \geq \tau \frac{\Delta^3 |P|}{|P[Q]|}.$$

By Lemma 4.2 we can upper bound this difference as

$$\begin{aligned} |B(\Delta; P^Q, L^Q) - B(\Delta/2; P^Q, L^Q)|^2 &\lesssim_{\epsilon'} \Delta^{-6-\epsilon'} \frac{M_{P^Q}(\Delta)}{|P^Q|} \frac{M_{L^Q}(\Delta \times \Delta \times 1)}{|L^Q|} \lesssim \\ &\lesssim \Delta^{-6-\epsilon'} \frac{M_P(\delta)}{|P[Q]|} |L[Q]|^{-1} \lesssim \Delta^{-\epsilon'} A \frac{|P|}{|P[Q]|} |\mathbb{U}[Q]|^{-1}. \end{aligned}$$

since $L[Q]$ is a maximal δ -separated collection of lines. We obtain from (41) that

$$\begin{aligned} \tau^2 \frac{\Delta^6 |P|^2}{|P[Q]|^2} &\lesssim \Delta^{-\epsilon'} A \frac{|P|}{|P[Q]|} |\mathbb{U}[Q]|^{-1}, \\ |\mathbb{U}[Q]| &\lesssim \tau^{-2} \Delta^{-6-\epsilon'} |P[Q]| A |P|^{-1}, \end{aligned}$$

so summing over Q and using (38) gives

$$|\mathbb{U}| \lesssim \tau^{-2} \delta^{-3-\epsilon'} A.$$

So by (40) we have

$$\begin{aligned} \tau^{-2} \delta^{-3-\epsilon'} A &\gtrsim |\mathbb{U}| \gtrsim K^{-\alpha} \lambda^{5/2} \delta^{-1/2+2\epsilon} M^{\alpha-1} |\mathbb{T}'|, \\ (42) \quad \lambda^{5/2} &\lesssim K^{\alpha} \tau^{-2} \delta^{-5/2-4\epsilon} M^{1-\alpha} A |\mathbb{T}'|^{-1}. \end{aligned}$$

Using $\lambda \geq \lambda_0 = \tau A^{-1}$ and $|\mathbb{T}'| \gtrsim \delta^{\epsilon} \frac{\lambda_0}{\lambda} |\mathbb{T}|$ we have

$$\lambda^{5/2} |\mathbb{T}'| \gtrsim \delta^{\epsilon} \lambda_0^{5/2} |\mathbb{T}| = \delta^{\epsilon} \tau^{5/2} A^{-5/2}$$

and so (42) gives

$$\begin{aligned} \delta^{\epsilon} \tau^{5/2} A^{-5/2} |\mathbb{T}| &\lesssim K^{\alpha} \tau^{-2} \delta^{-5/2-4\epsilon} M^{1-\alpha} A \\ \tau^{9/2} &\lesssim \delta^{-5\epsilon} K^{\alpha} M^{1-\alpha} A^{7/2} \delta^{-5/2} |\mathbb{T}|^{-1} \end{aligned}$$

Now recall that $M = \frac{M_L(\Delta \times \Delta \times 1)}{M_L(\delta \times \delta \times 1)}$ and $|\mathbb{T}| \gtrsim \frac{|L|}{C_0 M_L(\delta \times \delta \times 1)}$. Using this we get

$$\tau^{9/2} \lesssim C_0 \delta^{-5\epsilon} \delta^{-5/2} K^{\alpha} A^{7/2} \frac{M_L(\Delta \times \Delta \times 1)^{1-\alpha} M_L(\delta \times \delta \times 1)^{\alpha}}{|L|}$$

This matches the right hand side of (33), up to changing ϵ . So the difference $B(\delta) - B(\delta/2)$ is upper bounded by the desired quantity from above. Exactly the same argument shows that $B(\delta) - B(\delta/2)$ is lower bounded by the same quantity from below, giving us the desired upper bound on $|B(\delta/2) - B(\delta)|^2$. $\square$

5. FINAL PREPARATIONS

Recall from Section 1 that Ω_d is the set of pairs (p, ℓ) where $p \in [0, 1]^d$ and ℓ is a line passing through p . For a line ℓ in $\mathbb{R}^d$ we let $\theta(\ell) \in S^{d-1}$ denote a unit vector in the direction of ℓ (choosing arbitrarily between the two possibilities).

Given a configuration $X \subset \Omega_d$, we let $P[X] = \{p, (p, \ell) \in X\}$ and $L[X] = \{\ell, (p, \ell) \in X\}$ be the (multi-)sets of points and lines in $[0, 1]^d$ determined by X . Furthermore, we let $\theta[X] = \{\theta(\ell), (p, \ell) \in X\} \subset S^{d-1}$ be the multiset of directions of lines defined by X .

Recall the definition of the property $\text{PL}_d(\gamma)$ from Section 1. In the next proposition we show that if we know $\text{PL}_2(\gamma)$ for some $\gamma > 0$ then for any point-line configuration $X \subset \Omega_3$ with $d(X) \geq \delta$ the set of δ tubes defined by X satisfies a Katz–Tao axiom.

Proposition 5.1. *Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma \in [0, 1]$.*

If $X \subset \Omega_3$ is a point-line configuration such that $d(X) \geq \delta$, then for all $u \leq w \in (0, 1)$ with $uw \geq \delta$ we have:

$$M_{L[X]}(u \times w \times 1) \lesssim_{\epsilon} \delta^{-3-\epsilon} u^{1+\gamma} w^{2-\gamma}.$$

In other words, the set of δ -tubes determined by $L[X]$ is $(1 + \gamma, 2 - \gamma, C_{\epsilon} \delta^{-\epsilon})$ -Katz–Tao.

Proof. Let $\Pi \subset \mathbb{R}^3$ be an $u \times w \times 1$ prism and let X_Π be the set of pairs $(p, \ell) \in X$ with $p \in \Pi$ and $|\ell \cap \Pi| \geq 1/2$. Let $P_\Pi = P[X_\Pi] \subset \Pi$ be the corresponding set of points. Denote $\alpha = c\delta/u$ for a small constant $c > 0$. By pigeonhole principle, there exists a $\alpha u \times \alpha w \times \alpha$ box $\Pi' \subset \Pi$ homothetic to Π such that

$$|P_{\Pi'}| = |P_\Pi \cap \Pi'| \gtrsim \alpha^3 |P_\Pi| \gtrsim (\delta/u)^3 |P_\Pi|.$$

Let $X_{\Pi'}$ be the set of pairs $(p, \ell) \in X$ so that $p \in \Pi'$ and $|\ell \cap \Pi| \geq 1/2$. Let $\psi : B' \rightarrow [0, 1]^2$ be an affine map which collapses the side of length αu and rescales the other two by factors $(\alpha w)^{-1} \sim u/\delta w$ and $\alpha^{-1} \sim u/\delta$. Observe that for $(p, \ell) \neq (p', \ell') \in X_{\Pi'}$ we have

$$(43) \quad d(\psi(p), \psi(\ell')) \sim (u/\delta w) d(p, \ell') \gtrsim u/w.$$

So the point-line configuration $\tilde{X} = \psi(X_{\Pi'})$ in $[0, 1]^2$ satisfies $d(\tilde{X}) \gtrsim u/w$.

We conclude using the property $\text{PL}_2(\gamma)$ that

$$|X_{\Pi'}| \sim |\tilde{X}| \lesssim_\varepsilon (u/w)^{-2+\gamma-\varepsilon}.$$

So we get

$$M_{L[X]}(u \times w \times 1) \leq \max_{\Pi} |X_\Pi| \lesssim \alpha^{-3} |X_{\Pi'}| \lesssim (\delta/u)^{-3} (u/w)^{-2+\gamma-\varepsilon} \lesssim_\varepsilon \delta^{-3-\varepsilon} u^{1+\gamma} w^{2-\gamma}.$$

□

Proposition 5.1 will be used to estimate the error terms in the high-low estimates from Section 4. In order to prove $\text{PL}_3(\gamma)$ for some $\gamma > 0$ it is enough to assume that $\text{PL}_2(\gamma_0)$ holds for some $\gamma_0 > 0$. So Proposition 2.2 will suffice to prove Theorem 1.3 but if one were to optimize the value of γ one could of course use the stronger statement $\text{PL}_2(1/2)$ given by Theorem 1.2.

5.1. Uniformity. It will be convenient to assume some nice regularity properties of X . The following setup and lemmas are closely related to the framework in [2]. Some technical details are simplified here though (mainly due to the fact that we do not need to rescale into tubes). Let $X \subset \Omega_d$ be a point-line configuration in $\mathbb{R}^d$. For $u, v, w > 0$ let us define a concentration number

$$M_X(u, v, w) = \max_{(p_0, \ell_0) \in \Omega_d} \#\{(p, \ell) \in X : d(p, p_0) \leq u, d(\theta, \theta_0) \leq v, d(\ell, \ell_0) \leq w\},$$

where $\theta = \theta(\ell)$ and $\theta_0 = \theta(\ell_0)$. For the purposes of this definition it is convenient to define a metric on the set of lines as follows:

$$d(\ell, \ell') = d(\theta(\ell), \pm\theta(\ell')) + \min_{p \in \ell, p' \in \ell'} d(p, p'),$$

where we choose the sign so that this distance is minimized. In what follows we will omit this sign with the convention that $\theta(\ell)$ is always chosen appropriately. From this definition it is clear that we always have $d(\theta(\ell), \theta(\ell')) \leq d(\ell, \ell')$ which implies the identity

$$M_X(u, v, w) = M_X(u, \min(v, w), w).$$

The quantity M_X generalizes the concentration constants of the point-set $P[X]$ and the line set $L[X]$, namely, we have the approximate formulas:

$$\begin{aligned} M_{P[X]}(u) &\sim M_X(u, 1, 1), \\ M_{L[X]}(w \times w \times 1) &\sim M_X(1, 1, w) = M_X(1, w, w). \end{aligned}$$

Observe that the numbers M_X satisfy the following ‘Lipschitz’ property for any $A, B, C \geq 1$:

$$(44) \quad M_X(Au, Bv, Cw) \lesssim A^d B^{d-1} C^{2d-2} M_X(u, v, w).$$

The proof is analogous to the proof of Proposition 4.1. For parameters $K > 0$ and $\Delta_1 > \dots > \Delta_m$, let us say that a configuration $X \subset \Omega_d$ is K -uniform on a sequence of scales $\Delta_1 > \dots > \Delta_m$ if for all $i, j, k \in \{1, \dots, m\}$ and for any $(p_0, \ell_0) \in X$ the number of $(p, \ell) \in X$ with $d(p, p_0) \leq \Delta_i$, $d(\theta, \theta_0) \leq \Delta_j$ and $d(\ell, \ell_0) \leq \Delta_k$ is at least $M_X(\Delta_i, \Delta_j, \Delta_k)/K$.

Lemma 5.2. *For any δ and $m \geq 1$ there exists some $K \leq (C \log 1/\delta)^{10m^3}$ such that the following holds. For any $X \subset \Omega$ and any $\Delta_1 > \dots > \Delta_m \geq \delta$ there exists $X' \subset X$ of size at least $K^{-1}|X|$ so that X' is K -uniform on scales $\Delta_1, \dots, \Delta_m$.*

Proof. This lemma is the same as Lemma 3.6 in [2], so let us only give a brief sketch. For each triple $(\Delta_i, \Delta_j, \Delta_k)$ be consider a metric on the space Ω_d given by

$$d_{i,j,k}((p, \theta), (p', \theta')) = \max \left\{ \frac{d(p, p')}{\Delta_i}, \frac{d(\theta, \theta')}{\Delta_j}, \frac{d(\ell, \ell')}{\Delta_k} \right\}$$

Note that $M_X(\Delta_i, \Delta_j, \Delta_k)$ is essentially the maximum number of points of X contained in a unit ball in metric $d_{i,j,k}$. Now we fix a maximal ~ 1 -net $\mathcal{P}_{i,j,k} \subset \Omega_d$ in the metric $d_{i,j,k}$ and define a hypergraph $H \subset X \times \prod \mathcal{P}_{i,j,k}$ as follows. For each $(p, \ell) \in X$ we put a tuple of the form $((p, \ell), (x_{i,j,k}))$ in H where $x_{i,j,k}$ is a closest (in metric $d_{i,j,k}$) element of the net $\mathcal{P}_{i,j,k}$ to (p, ℓ) . Now using a combinatorial lemma (see Lemma 3.5 in [2]) we can pass to a large subgraph $H' \subset H$ which is approximately regular (the degrees of vertices in each component of H' are the same up to a logarithmic factor). We then let $X' \subset X$ be the corresponding set of point-line pairs. $\square$

For $\delta > 0, K \geq 2$, we say that a point-line pair configuration X is (δ, K) -uniform if the following two conditions hold:

- (i) X is K -uniform on the sequence of scales $K^{-1} > K^{-2} > \dots > K^{-m}$ with $m = \lfloor \log_K(1/\delta) \rfloor$,
- (ii) For each $j = 1, \dots, m$, the set of points $P[X]$ can be covered by K^{-j} -cubes which are pairwise $C_d K^{-j}$ -separated for some fixed constant C_d depending on the dimension.

For any X we can pass to a subset $X' \subset X$ of size $\gtrsim C_d^{-dm}|X|$ which has the property (ii) above. By applying Lemma 5.2 to X' we can then find a subset $X'' \subset X'$ which is (δ, K) -uniform and satisfies $|X''| \gtrsim (C_d \log(1/\delta))^{O(\epsilon^{-3})}|X|$ provided that $K \geq \delta^{-\epsilon}$. For a (δ, K) -uniform configuration X we will typically restrict our attention to scales w of the form K^{-j} for some $j \geq 0$. Note however that due to (44), for any intermediate scales $w \in [K^{-j-1}, K^{-j}]$ we also have uniformity properties, albeit with an error $O(K^{O(1)})$ instead of K .

Let X be a (δ, K) -uniform configuration and fix an arbitrary pair $(p_0, \ell_0) \in X$. For $\Delta \in (\delta, 1)$ of the form $\Delta = K^{-j}$ we can define a rescaled configuration X_Δ as follows. Let Q_Δ be a Δ -cube covering p_0 which is $C_d K^{-j}$ -separated from other cubes covering $P[X]$ (coming from (ii) above). Let $\psi : Q_\Delta \rightarrow [0, 1]^d$ be the homothetic rescaling map and define

$$X_\Delta = \psi(X \cap Q_\Delta) := \{(\psi(p), \psi(\ell)) : (p, \ell) \in X, p \in Q_\Delta\}.$$

Note that the definition of X_Δ depends on the choice of the initial pair (p_0, ℓ_0) but this choice will be unimportant due to uniformity of X . For two scales Δ, Δ' we abuse notation and let $(X_\Delta)_{\Delta'} = X_{\Delta\Delta'}$, i.e. the rescalings of the point-line configuration X_Δ are compatible with those of X .

Claim 5.3. *Suppose that X is (δ, K) -uniform for $K \geq K_0(d)$ and let $\Delta = K^{-j} \in (\delta, 1)$ then X_Δ is $(\delta/\Delta, K^C)$ -uniform for some integer $C = C(d) \geq 1$. Furthermore, we have estimates*

$$\frac{1}{K^{C'}} M_X(u\Delta, \min(v, w), w\Delta) \lesssim M_{X_\Delta}(u, v, w) \leq M_X(u\Delta, \min(v, w), w\Delta)$$

valid for u, v, w of the form $K^{-j} \in [\delta/\Delta, 1]$.

Proof. Clearly, we have $d(\psi(p), \psi(p')) = \Delta^{-1}d(p, p')$ for $p, p' \in \mathbb{R}^d$ and for lines $\ell, \ell' \subset \mathbb{R}^d$ the metric we introduced scales by ψ as

$$d(\psi(\ell), \psi(\ell')) = d(\theta(\ell), \theta(\ell')) + \min_{p \in \ell, p' \in \ell'} d(\psi(p), \psi(p')) = d(\theta(\ell), \theta(\ell')) + \Delta^{-1} \min_{p \in \ell, p' \in \ell'} d(p, p').$$

Using these relations implies an upper bound

$$M_{X_\Delta}(u, v, w) \leq M_X(u\Delta, \min(v, w), w\Delta).$$

On the other hand, if $(p_1, \ell_1) \in X \cap Q_\Delta$ then by the separation condition (ii), points $(p, \ell) \in X \setminus (X \cap Q_\Delta)$ will not contribute to the number of $(p_2, \ell_2) \in X$ so that $d(p_1, p_2) \leq \Delta u$, $d(\theta_1, \theta_2) \leq v$, $d(\ell_1, \ell_2) \leq \Delta w$. Under these assumptions, we have $d(\psi(\ell_1), \psi(\ell_2)) \leq v + w$ and so for $(\tilde{p}_1, \tilde{\ell}_1) = \psi(p_1, \ell_1)$ we get

$$\begin{aligned} & \#\{(\tilde{p}_2, \tilde{\ell}_2) \in X_\Delta : d(\tilde{p}_1, \tilde{p}_2) \leq u, d(\tilde{\theta}_1, \tilde{\theta}_2) \leq v, d(\tilde{\ell}_1, \tilde{\ell}_2) \leq w\} \geq \\ & \geq \#\{(p_2, \ell_2) \in X : d(p_1, p_2) \leq \Delta u, d(\theta_1, \theta_2) \leq \min(v, \frac{w}{2}), d(\ell_1, \ell_2) \leq \frac{\Delta w}{2}\} \end{aligned}$$

So by (δ, K) uniformity of X for $K \geq 2$ this can be lower bounded by

$$\frac{1}{K} M_X(\Delta u, \min(v, K^{-1}w), K^{-1}\Delta w) \gtrsim K^{-O(1)} M_X(\Delta u, \min(v, w), \Delta w)$$

where the latter follows by (44). So we conclude that X_Δ is $(\delta/\Delta, K^C)$ -uniform for some (integer) constant C depending on dimension (provided that K is sufficiently large). $\square$

For $X \subset \Omega_d$, denote by $\theta[X] \subset S^{d-1}$ the multiset of directions determined by X . For a set A in a metric space and $\delta > 0$ we let $|A|_\delta$ denote the δ -covering number of A , i.e. the minimal number of δ -balls required to cover A . For uniform sets X , the covering numbers of related objects $P[X]$, $L[X]$, $\theta[X]$ can be determined using the concentration numbers M_X :

Proposition 5.4. *Let $X \subset \Omega_d$ be a (δ, K) -uniform configuration and let $w = K^{-j} \in (\delta, 1)$. Then we have*

$$\begin{aligned} \frac{|X|}{M_X(w, 1, 1)} &\lesssim |P[X]|_w \lesssim K \frac{|X|}{M_X(w, 1, 1)} \\ \frac{|X|}{M_X(1, 1, w)} &\lesssim |L[X]|_w \lesssim K \frac{|X|}{M_X(1, 1, w)} \\ \frac{|X|}{M_X(1, w, 1)} &\lesssim |\theta[X]|_w \lesssim K \frac{|X|}{M_X(1, w, 1)} \end{aligned}$$

Proof. Let us prove the first line, the others are analogous. First, it is clear that $|P[X]|_w \gtrsim \frac{|X|}{M_X(w, 1, 1)}$ since any w ball contains at most $C M_X(w, 1, 1)$ points from $P[X]$. Let $\mathcal{P} \subset P[X]$ be a maximal by inclusion $w/2$ -separated subset of points. Clearly, w -balls around $\mathcal{P}$ cover $P[X]$. Let $w' = K^{-j'}$ be a scale so that $w' \leq cw$ for a small constant c . Then for each $p \in \mathcal{P}$ by uniformity we have

$$|B_d(p, w/2) \cap P[X]| \geq \frac{1}{K} M_X(w', 1, 1) \gtrsim \frac{1}{K} M_X(w, 1, 1).$$

So it follows that $|\mathcal{P}| \lesssim K \frac{|X|}{M_X(w, 1, 1)}$, concluding the argument. $\square$

To lower bound the normalized incidence count $B(w) = \frac{I(w; P[X], L[X])}{w^{d-1}|X|^2}$ we can use the following ‘initial estimate’. Initial estimates like this were heavily used in arguments in [1] and [2].

Proposition 5.5. *Let $X \subset \Omega_d$ be a (δ, K) -uniform configuration and let $w = K^{-j} \in (C\delta^{1/2}, 1)$. Then we have*

$$(45) \quad B(w; P[X], L[X]) \gtrsim K^{-O(1)} \frac{|\theta[X_w]|_w}{w^{d-1}|L[X]|_w} \gtrsim K^{-O(1)} \frac{|\theta[X_w]|_w}{|\theta[X]|_w}.$$

Proof. Recall that $I(w; P, L)$ is defined as $\sum_{p \in P, \ell \in L} \varphi(w^{-1}d(p, \ell))$ for some smooth symmetric bump function φ . We also know that this function satisfies $\varphi(t) \geq c$ for all $t \in [-c, c]$ for some constant c depending on d . So we have a lower bound

$$(46) \quad I(w; P, L) \geq c \#\{p \in P, \ell \in L : d(p, \ell) \leq cw\}.$$

Let $w' = K^{-j'}$ be the largest scale such that $w' \leq c'w$ for some constant c' . Consider the following set

$$\Lambda = \{((p_0, \ell_0), (p_1, \ell_1), (p_2, \ell_2)) \in X \times X \times X : d(p_0, p_1) \leq w', d(\ell_1, \ell_2) \leq w'\}.$$

By uniformity, we have

$$|\Lambda| \geq K^{-2} |X| M_X(w', 1, 1) M_X(1, 1, w') \gtrsim K^{-2} \frac{|X|^3}{|P[X]|_w |L[X]|_w}.$$

On the other hand, for $((p_0, \ell_0), (p_1, \ell_1), (p_2, \ell_2)) \in \Lambda$ we have $d(p_0, \ell_2) \leq d(p_0, p_1) + d(p_1, \ell_2) \leq d(p_0, p_1) + O(d(\ell_1, \ell_2)) \leq O(w')$, where the implied constant depends on the metric we choose on $\mathcal{A}_{d,1}$. So if we choose the constant c' so that $O(w') \leq cw$, then (p_0, ℓ_2) gives a contribution to (46). Now if we fix (p_0, ℓ_0) and (p_2, ℓ_2) then the number of (p_1, ℓ_1) so that $((p_0, \ell_0), (p_1, \ell_1), (p_2, \ell_2)) \in \Lambda$ can be upper bounded by

$$\begin{aligned} & \#\{(p_1, \ell_1) \in X : d(p_0, p_1) \leq w', d(\ell_1, \ell_2) \leq w'\} \\ & \leq \#\{(p_1, \ell_1) \in X : d(p_0, p_1) \leq w', d(\ell_1, \tilde{\ell}_2) \leq Cw'\} \\ & \leq M_X(w', 1, Cw') \leq M_X(w, 1, w) \end{aligned}$$

where $\tilde{\ell}_2$ is a line passing through p_0 , parallel to ℓ_2 . So by choosing $Cw' \leq w$ we get

$$I(w; P[X], L[X]) \gtrsim \frac{|\Lambda|}{M_X(w, 1, w)}.$$

On the other hand, since $d(\ell, \ell') \leq w$ implies $d(\theta(\ell), \theta(\ell')) \leq w$ we have

$$M_X(w, 1, w) = M_X(w, w, w) \lesssim K^{C'} M_{X_w}(1, w, 1)$$

where the last inequality follows from Claim 5.3. Thus, using Proposition 5.4, we get

$$I(w; P[X], L[X]) \gtrsim K^{-2-C'} \frac{|\Lambda|}{M_{X_w}(1, w, 1)} \gtrsim K^{-5-C'} \frac{|X|^3}{|P[X]|_w |L[X]|_w} \frac{|\theta[X_w]|_w}{|X_w|}$$

we have $|P[X]|_w |X_w| \lesssim K |X|$, so after simplifying get:

$$B(w) \gtrsim K^{-O(1)} \frac{|\theta[X_w]|_w}{w^{d-1} |L[X]|_w}.$$

This gives the first inequality in (45). The second inequality follows from the general fact that $|L|_w \lesssim |\theta(L)|_w w^{1-d}$. $\square$

Proposition 5.6. *Suppose that X is (δ, K) -uniform and $w = K^{-j} \in (\delta^{1/2}, 1)$. Then we have*

$$|L[X]|_w \gtrsim K^{-O(1)} w |\theta[X_w]|_w |P[X]|_w.$$

Proof. Since a w -tube can be covered by $\sim w^{-1}$ w -balls, we have a trivial upper bound

$$I(w; P, L) \lesssim w^{-1} |L| M_P(w)$$

and so $B(w) \lesssim w^{-d} \frac{M_P(w)}{|P|}$. Comparing to (45) we get

$$K^{-O(1)} \frac{|\theta[X_w]|_w}{w^{d-1} |L[X]|_w} \lesssim w^{-d} \frac{M_{P[X]}(w)}{|P[X]|} \lesssim K w^{-d} |P[X]|_w^{-1}$$

and so by rearranging we get the desired bound. $\square$

6. POINT-LINE CONFIGURATIONS IN $\mathbb{R}^3$

In this section we prove Theorem 1.3. Let us restate the theorem in a form convenient for the proof.

Theorem 6.1. *Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma > 0$. Then there exists some $\kappa(\gamma) > 0$ such that $\text{PL}_3(\kappa)$ holds. Explicitly, let $\delta > 0$ and let $X \subset \Omega_3$ be a point-line configuration with $d(X) \geq \delta$. Then we have $|X| \lesssim_\varepsilon \delta^{-\varepsilon} \delta^{-3+\kappa}$ for every $\varepsilon > 0$.*

We will prove Theorem 6.1 using the following three lemmas, each giving a sufficient condition on X under which we can improve on the trivial upper bound $|X| \lesssim \delta^{-3}$. Recall that in Section 5 for a (δ, K) -uniform X we introduced the rescaled point-line configurations X_w for every $w = K^{-j} \in (\delta, 1)$. Note that by definition we have

$$d(X_w) \geq w^{-1}d(X)$$

and that $|X_w| \gtrsim K^{-1}M_X(w, 1, 1) \gtrsim w^3|X|$.

In the first lemma, we assume that for some fairly coarse scale w , the set of directions of X is far from uniform on scale w . This condition is designed to make Corollary 4.3 applicable on all scales and so it can be used to obtain a small improvement over the simple high-low.

Lemma 6.1. *Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma \in (0, 1]$. Then the following holds with $A = A(\gamma) = \frac{100}{\gamma}$.*

Let $X \subset \Omega_3$ be a (δ, K) -uniform point-line configuration such that $d(X) \geq \delta$. Suppose that for some $w = K^{-i} \in (\delta, 1)$ and $\beta \in (0, 1)$ we have $|\theta[X]|_w \leq w^{-2+\beta}$ and $w \geq \delta^{\frac{1}{A \log(2/\beta)}}$. Then $\delta^3|X| \lesssim_\varepsilon K^{O(1)}\delta^{-\varepsilon}w^{\beta/A}$. Here $\log$ denotes the natural logarithm.

Proof. For every $j = 0, \dots, m := \lfloor \log_w \delta \rfloor$ we may consider the rescaled configuration X_{w^j} . Denote by β_j the number such that

$$w^{\beta_j} = w^2|\theta[X_{w^j}]|_w.$$

Note that the sequence β_j is increasing: indeed, note that $\theta[X_{w^{j+1}}] \subset \theta[X_{w^j}]$ and so we have $\beta = \beta_0 \leq \beta_1 \leq \dots \leq \beta_m \leq 2$.

Let $\tau > 1$ be a constant to be determined later. Suppose that we have $\tau^{m-2}\beta > 2$. Then it follows that there is $j \in \{0, \dots, m-2\}$ such that

$$\beta_j \geq \tau^j \beta, \quad \beta_{j+1} \leq \tau \beta_j.$$

Let $P_j = P[X_{w^j}]$ and $L_j = L[X_{w^j}]$. By Proposition 5.5 we then get

$$B_j(w) = B(w; P_j, L_j) \gtrsim K^{-O(1)} \frac{|\theta[X_{w^{j+1}}]|_w}{|\theta[X_{w^j}]|_w} = K^{-O(1)} w^{\beta_{j+1}-\beta_j} \gtrsim K^{-O(1)} w^{(\tau-1)\beta_j}.$$

Now we apply Corollary 4.3 to the pair (P_j, L_j) and on scale $v \in [c\delta/w^j, w]$. Write $\delta_j = \delta/w^j$ for convenience (recall that $d(X_{w^j}) \geq \delta_j$ from rescaling). Since P_j is δ_j -separated, we have

$$M_{P_j}(v) \lesssim (v/\delta_j)^3.$$

By Proposition 5.1, property $\text{PL}_2(\gamma)$ implies for $u \in (v, 1)$:

$$M_{L_j}(v \times v/u \times 1) \lesssim_\varepsilon \delta^{-\varepsilon} \delta_j^{-3} v^3 u^{-2+\gamma},$$

giving $M_{L_j}(v \times v/u \times 1) \leq u^{-2+\gamma} M$ for $M \lesssim \delta^{-\varepsilon} \delta_j^{-3} v^3$. Lastly, we have

$$|\theta(L_j)|_v \lesssim (w/v)^2 |\theta(L_j)|_w \lesssim w^{\beta_j} v^{-2},$$

giving $|\theta(L_j)|_v \leq \nu v^{-2}$ with $\nu \sim w^{\beta_j}$. So by Corollary 4.3 we get

$$|B(v) - B(2v)|^2 \lesssim_\varepsilon \delta^{-\varepsilon} \nu^{\gamma/4} v^{-6} \frac{M_{P_j}(v)}{|P_j|} \frac{M}{|L_j|} \lesssim \delta^{-2\varepsilon} w^{\gamma\beta_j/4} (\delta_j^3 |X_j|)^{-2}.$$

So since $d(X_j) \geq \delta_j$ we have $B_j(c\delta_j) \lesssim \delta_j^{-2} |X_j|^{-1}$ and using the initial estimate $B_j(w) \gtrsim K^{-O(1)} w^{(\tau-1)\alpha_j}$ we then obtain by summing the high-low errors over all v :

$$K^{-O(1)} w^{(\tau-1)\beta_j} \lesssim \delta_j^{-2} |X_j|^{-1} + \delta^{-o(1)} w^{\gamma\beta_j/8} (\delta_j^3 |X_j|)^{-1}.$$

If we choose $\tau = 1 + \gamma/16$ then this bound implies

$$\delta_j^3 |X_j| \lesssim K^{O(1)} \max(\delta^{-o(1)} w^{\gamma\beta_j/16}, \delta_j w^{-\gamma\beta_j/16}) \leq K^{O(1)} \delta^{-o(1)} w^{\gamma\beta_j/16} \leq K^{O(1)} \delta^{-o(1)} w^{\gamma\beta/16}.$$

We have $\delta^3 |X| \lesssim \delta_j^3 |X_j|$ so this gives the desired bound. For the argument to work, we need the condition $(1 + \gamma/16)^{m-2} > 2/\beta$ where $m = \lfloor \log_w \delta \rfloor$ which is satisfied if, say, $m \geq 100\gamma^{-1} \log(2/\beta)$. $\square$

In the second lemma we consider an opposite extreme: we assume that the restriction of X onto every $\delta^{1/2}$ box Q spans almost all directions on scale $\delta^{1/2}$. This condition allows us to apply Theorem 4.2 on all intermediate scales.

Lemma 6.2. *Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma \in (0, 1]$. Then the following holds with $\beta_0 = \beta_0(\gamma) = \gamma/200$.*

Let $X \subset \Omega_3$ be a (δ, K) -uniform configuration with $d(X) \geq \delta$. Suppose that for some $\beta \in [0, \beta_0]$ we have $|\theta[X_\Delta]|_\Delta \geq \Delta^{-2}\delta^\beta$ for some $\Delta = K^{-i} \in [C\delta^{1/2}, CK\delta^{1/2}]$. Then we have $\delta^3|X| \lesssim K^{O(1)}\delta^{\beta_0}$.

Proof. By Proposition 5.5 we have

$$B(\Delta) = B(\Delta; P[X], L[X]) \gtrsim K^{-O(1)} \frac{|\theta[X_\Delta]|_\Delta}{|\theta[X]|_\Delta} \gtrsim K^{-O(1)} \frac{\Delta^{-2}\delta^\beta}{\Delta^{-2}} \gtrsim K^{-O(1)}\delta^\beta.$$

Now let $w \in [c\delta, \Delta]$ and let us estimate the difference $B(w/2) - B(w)$ using Theorem 4.2. Let $P = P[X]$ and $L = L[X]$. It follows from uniformity and (44) that L is $C_0 \sim K^{O(1)}$ uniform on scale w . Let $\mathbb{T}_w$ be the set of w -tubes determined by L . We have by Proposition 5.4 that $|\mathbb{T}_w| = K^{O(1)} \frac{|L|}{M_L(w \times w \times 1)}$.

By Proposition 5.1, property $\text{PL}_2(\gamma)$ implies that

$$M_L(u \times v \times 1) \lesssim_\varepsilon \delta^{-\varepsilon} u^{1+\gamma} v^{2-\gamma} \delta^{-3}.$$

Therefore, using uniformity of X , we get that for any $u \times v \times 1$ box Π we have

$$|\mathbb{T}_w \cap \Pi| \lesssim K^{O(1)} \frac{M_L(u \times v \times 1)}{M_L(w \times w \times 1)} \lesssim \left(\delta^{-\varepsilon} K^{O(1)} (w/\delta)^3 M_L(w \times w \times 1)^{-1} \right) (u/w)^{1+\gamma} (v/w)^{2-\gamma}.$$

We conclude that $\mathbb{T}_w$ is a $(1+\gamma, 2-\gamma, K_w)$ -Katz–Tao set of w -tubes where $K_w \sim \delta^{-\varepsilon} K^{O(1)} (w/\delta)^3 M_L(w \times w \times 1)^{-1}$.

Note that we have $|P|_\Delta \gtrsim (\delta/\Delta)^3 |P|_\delta \gtrsim \Delta^{-3}(\delta^3|X|)$. By Proposition 5.6 we have

$$|L|_\Delta \gtrsim K^{-O(1)} \Delta |\theta[X_\Delta]|_\Delta |P|_\Delta \gtrsim K^{-O(1)} \Delta^{-4} \delta^\beta (\delta^3|X|)$$

and so by Proposition 4.1 we obtain

$$(47) \quad |L|_{w^{1/2}} \gtrsim K^{O(1)} (\Delta/w^{1/2})^4 |L|_\Delta \gtrsim K^{-O(1)} w^{-2} \delta^\beta (\delta^3|X|).$$

Using (δ, K) -uniformity of X this implies

$$M_L(w^{1/2} \times w^{1/2} \times 1) \lesssim K^{O(1)} w^2 \delta^{-\beta} |L| (\delta^3|X|)^{-1}.$$

So by Theorem 4.2 applied with $t_1 = 1 + \gamma$, $t_2 = 2 - \gamma$ we have $\alpha = \frac{t_1+2}{2t_1+2t_2} = \frac{3+\gamma}{6}$ and

$$\begin{aligned} |B(w/2) - B(w)|^{9/2} &\lesssim \\ &\lesssim_\varepsilon K^{O(1)} \delta^{-\varepsilon} w^{-5/2} K_w^\alpha (\delta^3|X|)^{-7/2} \frac{M_L(w^{1/2} \times w^{1/2} \times 1)^{1-\alpha} M_L(w \times w \times 1)^\alpha}{|L|} \\ &\lesssim K^{O(1)} \delta^{-2\varepsilon} w^{-5/2} (w/\delta)^{3\alpha} (\delta^3|X|)^{-7/2} \frac{M_L(w^{1/2} \times w^{1/2} \times 1)^{1-\alpha}}{|L|} \\ &\lesssim K^{O(1)} \delta^{-2\varepsilon} w^{-5/2} (w/\delta)^{3\alpha} (\delta^3|X|)^{-9/2+\alpha} (w^2 \delta^{-\beta})^{1-\alpha} |L|^{-\alpha} \\ &\lesssim K^{O(1)} \delta^{-2\varepsilon} \delta^{-\beta(1-\alpha)} (\delta^3|X|)^{-9/2} w^{\alpha-\frac{1}{2}} \end{aligned}$$

Now recall that $w \leq \delta^{1/2}$ and $\alpha - \frac{1}{2} = \frac{\gamma}{6}$:

$$(48) \quad |B(w/2) - B(w)| \lesssim K^{O(1)} \delta^{-\varepsilon} \delta^{-\beta/9} (\delta^3|X|)^{-1} \delta^{\gamma/54}.$$

Recalling $B(c\delta) \lesssim \delta^{-2}|X|^{-1}$ (which follows from $d(X) \geq \delta$) we conclude that

$$\begin{aligned} B(\Delta) &\lesssim \max_w (B(c\delta), |B(w) - B(w/2)|) \\ \delta^\beta &\lesssim K^{O(1)} \delta^{-\varepsilon} \max((\delta^3|X|)^{-1} \delta^{\gamma/54-\beta/9}, \delta^{-2}|X|^{-1}) \end{aligned}$$

$$\delta^3|X| \lesssim K^{O(1)}\delta^{-\varepsilon} \max\{\delta^{\gamma/54-10\beta/9}, \delta^{1-\beta}\} \lesssim \delta^{\beta_0}$$

if we define $\beta_0 = \gamma/200$. This completes the proof. $\square$

Lastly, in the third lemma we consider an intermediate situation where X does not span all directions on some scale ρ but restrictions of X onto $\rho^{1/2}$ -boxes span almost all directions on scale $\rho^{1/2}$. This condition interpolates between the two conditions in previous lemmas and it allows us to combine Corollary 4.3 and Theorem 4.2 to cover all scales.

Lemma 6.3. *Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma \in (0, 1]$. Then the following holds with $A = A(\gamma) = \frac{100}{\gamma}$.*

Let $X \subset \Omega_3$ be a (δ, K) -uniform configuration such that $d(X) \geq \delta$. Let $w = K^{-2i} \in (\delta, 1)$ and suppose that $|\theta[X]|_w \leq w^{-2}\delta^\sigma$ and $|\theta[X_{w^{1/2}}]|_{w^{1/2}} \geq w^{-1}\delta^\beta$ holds for some $\beta, \sigma \in [0, 1]$.

Suppose that $w \leq \delta^{2A\beta}$ and $\sigma \geq A\beta$ and $\beta \leq 1/2$. Then $\delta^3|X| \lesssim K^{O(1)}\delta^{-\varepsilon} \max(w^{1/A}, \delta^{\sigma/A})$.

Proof. The proof essentially follows by combining the computations we did in Lemma 6.1 and Lemma 6.2. First, by Proposition 5.5 we have $B(Cw^{1/2}) \gtrsim K^{-O(1)}\delta^\beta$ and by repeating the proof of Lemma 6.2 with w and $w^{1/2}$ in place of δ and Δ , we can show that for $v \in [cw, Cw^{1/2}]$ we have the upper bound (cf. (48))

$$|B(v/2) - B(v)| \lesssim K^{O(1)}\delta^{-\varepsilon}\delta^{-\beta/9}(\delta^3|X|)^{-1}w^{\gamma/54}$$

For $v \in [c\delta, \rho]$ we have $|\theta[X]|_v \lesssim (w/v)^2|\theta[X]|_w \lesssim v^{-2}\delta^\sigma$. So using Corollary 4.3 we have

$$|B(v) - B(2v)|^2 \lesssim \delta^{-\varepsilon}\delta^{\sigma\gamma/4}v^{-6} \frac{M_{P[X]}(v)}{|P|} \frac{M}{|L|} \lesssim \delta^{-\varepsilon}\delta^{\sigma\gamma/4}(\delta^3|X|)^{-2}.$$

So by summing over $v \in (c\delta, C\rho^{1/2})$ we obtain

$$|B(c\delta) - B(Cw^{1/2})| \lesssim K^{O(1)}\delta^{-\varepsilon}(\delta^3|X|)^{-1} \max(\delta^{-\beta/9}w^{\gamma/54}, \delta^{\gamma\sigma/8})$$

and thus by using $B(c\delta) \lesssim \delta^{-2}|X|^{-1}$ we conclude

$$\delta^3|X| \lesssim K^{O(1)}\delta^{-\varepsilon} \max(\delta^{-10\beta/9}w^{\gamma/54}, \delta^{-\beta}\delta^{\gamma\sigma/8}, \delta^{1-\beta})$$

Using the restrictions on β, σ, ρ we get the desired bound on $\delta^3|X|$. $\square$

Now we are ready prove Theorem 6.1 (and thus Theorem 1.3 as well).

Proof of Theorem 6.1. Suppose that $\text{PL}_2(\gamma)$ holds for some $\gamma > 0$ and fix a point-line configuration X with $d(X) \geq \delta$. Let $A = \frac{100}{\gamma}$ and $\beta_0 = \gamma/200$.

Let $\varepsilon > 0$ be an arbitrarily small constant and denote $K = \delta^\varepsilon$. By Lemma 5.2 and the comments after it we can pass to a subset $X' \subset X$ which is (δ, K) -uniform and $|X'| \gtrsim (\log 1/\delta)^{O(\varepsilon^{-O(1)})}|X|$. So without loss of generality, we may replace X with X' and assume that X is (δ, K) -uniform. Let us say that a scale $w \in [\delta, 1]$ is admissible if $w = K^{-i}$ for an integer i . Then for every admissible $w \in (\delta, 1)$ we have a rescaled point-line configuration X_w which satisfies $d(X_w) \geq \delta/w$. From definition of X_w , we have

$$(49) \quad \delta^3|X| \lesssim K^{O(1)}(\delta/w)^3|X_w|$$

so it suffices to upper bound the size of $|X_w|$ for some w . Suppose that $\delta^3|X| \geq \delta^\kappa$ for some constant $\kappa > 0$, our goal is to estimate this constant. Note that by Claim 5.3 the configuration X_w is $(\delta/w, O(K^{O(1)}))$ -uniform and so the lemmas from section can be applied to X_w . When we apply lemmas to X_w , we define the rescaled configurations by $(X_w)_u = X_{wu}$.

Let $\rho_0 = \delta^r$ for some $r \in (0, 1/4)$ be a large admissible scale and let $\beta^* \in (0, 1)$ be a small constant to be determined later (see the end of the proof where we choose these two parameters). Let $\lambda = \log(2/\beta^*)$. Let w_0 be the smallest admissible scale such that $w_0 \geq \rho_0^{\frac{1}{A\lambda}}$ (so we have the upper bound $w_0 \leq K\rho_0^{\frac{1}{A\lambda}}$). Then by Lemma 6.1 applied to X_{δ/ρ_0} we get that at least one of the following two options holds:

$$(i) \quad \rho_0^3|X_{\delta/\rho_0}| \lesssim K^{O(1)}\rho_0^{-\varepsilon}w_0^{\beta^*/A} \text{ or}$$

(ii) $|\theta[X_{\delta/\rho_0}]|_{w_0} \geq w_0^{-2+\beta^*}$.

In the first case it follows that

$$\delta^\kappa \lesssim K^{O(1)} \rho_0^{-\varepsilon} w_0^{\beta^*/A} \lesssim K^{O(1)} \delta^{-\varepsilon} \rho_0^{\frac{\beta^*}{A^2 \log(2/\beta^*)}},$$

which then gives

$$(50) \quad \kappa \geq \frac{r\beta^*}{A^2\lambda} + O(\varepsilon)$$

As long as β^* and r are taken to be constants depending only on γ this will give a sufficient lower bound on κ .

Let us now consider the second case. Let $\rho_1 = \rho_0 w_0$ and $w_1 = w_0^2$. Note that using our convention we have $(X_{\delta/\rho_1})_{w_0} = X_{\delta/\rho_0}$. Define β_1 to be the number such that $\rho_1^{\beta_1} = w_0^{\beta^*}$. So by rearranging and using that $\log_{\rho_0} w_0 = \frac{1}{A\lambda} + O(\varepsilon)$, we get

$$\beta_1 = \frac{\beta^*}{A\lambda + 1} + O(\varepsilon).$$

In particular it is clear that $\beta_1 \leq 1/2$. So we get that

$$|\theta[(X_{\delta/\rho_1})_{w_1^{1/2}}]|_{w_1^{1/2}} \geq w_1^{-1} \rho_1^{\beta_1}.$$

Let σ_1 be the number such that

$$|\theta[X_{\rho_1}]|_{w_1} = w_1^{-2} \rho_1^{\sigma_1}.$$

So by Lemma 6.3 applied to X_{δ/ρ_1} and $w = w_1$: if $w_1 \leq \rho_1^{2A\beta_1}$ and $\sigma_1 \geq A\beta_1$ then we have

$$\rho_1^3 |X_{\delta/\rho_1}| \lesssim K^{O(1)} \rho_1^{-\varepsilon} \max(w_1^{1/A}, \rho_1^{\sigma_1/A}) \leq K^{O(1)} \delta^{-\varepsilon} \max(w_0^{2/A}, w_0^{\beta^*})$$

where we used $\sigma_1 \geq A\beta_1$ and $\rho_1^{\beta_1} = w_0^{\beta^*}$, so this leads to

$$(51) \quad \kappa \geq \frac{r}{A\lambda} \max(2/A, \beta^*) + O(\varepsilon).$$

Note that we have $\rho_1^{2A\beta_1} = w_0^{2A\beta^*}$ so if $\beta^* \leq 1/A$ then this implies $w_1 \leq \rho_1^{2A\beta_1}$. So if Lemma 6.3 does not apply, then we must have $\sigma_1 \leq A\beta_1$, i.e.

$$(52) \quad |\theta[X_{\delta/\rho_1}]|_{w_1} \geq w_1^{-2} \rho_1^{A\beta_1} = w_1^{-2} w_0^{A\beta^*}.$$

This condition is analogous to the one in alternative (ii) above. We can now iterate this argument as follows. Define a pair of scales (ρ_j, w_j) by the rule

$$\rho_j = \rho_{j-1} w_{j-1}, \quad w_j = w_{j-1}^2,$$

which gives $w_j = w_0^{2^j}$ and $\rho_j = \rho_0 w_0^{2^j - 1}$, $j \geq 0$. Suppose that we have for some $j \geq 1$ that

$$(53) \quad |\theta[X_{\delta/\rho_j}]|_{w_j} \geq w_j^{-2} w_0^{A^j \beta^*}$$

and let us try to prove (53) for $j+1$. Suppose that j is such that $\rho_{j+1} \geq \delta^{1/2}$. Let β_j be the number so that $\rho_{j+1}^{\beta_{j+1}} = w_0^{A^j \beta^*}$. This then implies by (53) that

$$|\theta[(X_{\delta/\rho_{j+1}})_{w_{j+1}^{1/2}}]|_{w_{j+1}^{1/2}} \geq w_{j+1}^{-1} \rho_{j+1}^{\beta_{j+1}}.$$

Let σ_{j+1} be the number so that

$$|\theta[X_{\delta/\rho_{j+1}}]|_{w_{j+1}} = w_{j+1}^{-2} \rho_{j+1}^{\sigma_{j+1}}$$

Then by Lemma 6.3, if we have $\beta_{j+1} \leq 1/2$, $w_{j+1} \leq \rho_{j+1}^{2A\beta_{j+1}}$ and $\sigma_{j+1} \geq A\beta_{j+1}$ then it follows that

$$\rho_{j+1}^3 |X_{\delta/\rho_{j+1}}| \lesssim K^{O(1)} \rho_{j+1}^{-\varepsilon} \max(w_{j+1}^{1/A}, \rho_{j+1}^{\sigma_{j+1}/A}) \leq K^{O(1)} \rho_{j+1}^{-\varepsilon} \max(w_0^{2^{j+1}/A}, w_0^{A^j \beta^*})$$

where we used the lower bound on σ_{j+1} and the definition of β_{j+1} . So this gives

$$(54) \quad \kappa \geq \frac{r}{A\lambda} \max(2^{j+1}/A, A^j \beta^*) + O(\varepsilon).$$

So we may assume that conditions of Lemma 6.3 do not apply. If $\sigma_{j+1} \geq A\beta_{j+1}$ then this precisely means that (53) holds with j replaced by $j+1$ and so we completed an iteration step. Otherwise, it could happen that one of $\beta_{j+1} > 1/2$, or $w_{j+1} > \rho_{j+1}^{2A\beta_{j+1}}$ or $\rho_{j+1} < \delta^{1/2}$ holds. Let us check what restrictions this puts on j . By definition,

$$\rho_{j+1} = \rho_0 w_0^{2^{j+1}-1} = \rho_0^{1+(2^{j+1}-1)/A\lambda+O(\varepsilon)} = \delta^{r(1+(2^{j+1}-1)/A\lambda)+O(\varepsilon)}$$

so we have $\rho_{j+1} \geq \delta^{1/2}$ as long as $r(1+2^{j+1}/A\lambda) \leq 1/3$ and ε is much smaller than all other parameters. Next, we have

$$\begin{aligned} \rho_{j+1}^{\beta_{j+1}} &= w_0^{A^j \beta^*} \\ \beta_{j+1} &= \frac{A^j \beta^*}{A\lambda + 2^j} + O(\varepsilon) \end{aligned}$$

so we will have $\beta_{j+1} \leq 1/2$ as long as, say, $A^j \beta^* \leq 1/2$ holds. Finally, the condition $w_{j+1} \leq \rho_{j+1}^{2A\beta_{j+1}}$ is equivalent to

$$w_0^{2^{j+1}} \leq w_0^{2A^{j+1}\beta^*}$$

i.e. we need $A^{j+1}\beta^* \leq 2^j$. We conclude these observations in the following proposition:

Proposition 6.4. *Suppose that $j \geq 1$ is such that $A^{j+1}\beta^* \leq 1$ and $2^j r \leq 1$. Then (53) implies that either (53) holds with $j+1$ instead of j or (54) holds.*

Now suppose that (53) holds for some j . Note that we have $\rho_j = \rho_0 w_0^{2^j-1} \leq w_0^{2^j} = w_j$ and $\rho_j/w_j = \rho_0/w_0 \geq \rho_0$. Observe that $(X_{\delta/\rho_j^2})_{\rho_j} = X_{\delta/\rho_j}$ and so we have

$$|\theta[(X_{\delta/\rho_j^2})_{\rho_j}]|_{\rho_j} = |\theta[X_{\delta/\rho_j}]|_{\rho_j} \geq |\theta[X_{\delta/\rho_j}]|_{w_j} \geq w_j^{-2} w_0^{A^j \beta^*} \geq \rho_j^{-2} \rho_0^2 w_0^{A^j \beta^*}$$

where we used a trivial relation between covering numbers. Recall that we defined $\beta_0 = \beta_0(\gamma) = \gamma/200$. Suppose that the following inequality holds

$$(55) \quad \rho_0^2 w_0^{A^j \beta^*} \geq \rho_j^{2\beta_0}.$$

So by Lemma 6.2 applied to X_{δ/ρ_j^2} (with $\Delta = \rho_j$) it then follows that $(\rho_j^2)^3 |X_{\delta/\rho_j^2}| \lesssim K^{O(1)} \rho_j^{2\beta_0}$ which leads to

$$(56) \quad \begin{aligned} \delta^\kappa &\lesssim K^{O(1)} \rho_j^{2\beta_0} \lesssim K^{O(1)} \rho_0^{2\beta_0} \\ \kappa &\geq 2\beta_0 r + O(\varepsilon). \end{aligned}$$

So it remains to verify (55). Expanding the definitions, it reduces to the inequality

$$2 + A^j \beta^* / A\lambda \leq 2\beta_0(1 + (2^j - 1)/A\lambda)$$

which would be implied by (recall that $\lambda = \log(2/\beta^*)$)

$$(57) \quad 2A \log(2/\beta^*) + A^j \beta^* \leq 2^j \beta_0.$$

We now select parameters as follows. Suppose that $\beta^* \leq \beta_0/A^2$ and let $j(\beta^*) = \lceil \log_A \beta_0/\beta^* \rceil - 1$. Then clearly $j(\beta^*) \geq 1$ and

$$A^{j(\beta^*)} \beta^* \leq (\beta_0/\beta^*) A^{-1} \beta^* \leq 2^{j(\beta^*)-1} \beta_0$$

and

$$2^{j(\beta^*)} \geq 2^{\log_A \beta_0/\beta^* - 2} \geq \frac{1}{4} (\beta_0/\beta^*)^{1/\log_2 A}.$$

This means that for $\beta^* \leq \beta_0/A^2$ and $j = j(\beta^*)$, the relation (57) would follow from

$$(58) \quad (16A\beta_0^{-1}) \log(2/\beta^*) \leq (\beta_0/\beta^*)^{1/\log_2 A}$$

Since a polynomial eventually dominates logarithm, this is satisfied for sufficiently small β^* . Explicitly, we have for any $x > 1$, $t > 0$ that $\log x = t^{-1} \log x^t \leq t^{-1} x^t$. So we can use the bound

$$\log(2/\beta^*) \leq 2(\log_2 A)(2/\beta^*)^{\frac{1}{2\log_2 A}}$$

and some easy rearrangements to reduce (58) to

$$\beta^* \leq (32A\beta_0^{-1} \log_2 A)^{-2\log_2 A} \beta_0^2/2.$$

So by setting β^* to be the right hand side we can then define $r = 2^{-j(\beta^*)}$ and conclude that this choice satisfies requirement of Proposition 6.4. So starting from $j = 1$ and (52) we can use Proposition 6.4 to show that (53) holds for $j = j(\beta^*)$.

We conclude that with this choice of parameters, at least one of the inequalities (50), (51), (54) or (56) must be satisfied. All of the expressions one right hand side are positive functions of γ , β^* and r , so it follows that $\kappa \geq \kappa(\gamma) > 0$, as desired. This completes the proof of Theorem 6.1. $\square$

REFERENCES

- [1] Alex Cohen, Cosmin Pohoata, and Dmitrii Zakharov, *A new upper bound for the heilbronn triangle problem*, arXiv preprint arXiv:2305.18253 (2023).
- [2] Alex Cohen, Cosmin Pohoata, and Dmitrii Zakharov, *Lower bounds for incidences*, *Inventiones mathematicae* (2025), 1–74.
- [3] Larry Guth, Noam Solomon, and Hong Wang, *Incidence estimates for well spaced tubes*, *Geometric and Functional Analysis* **29** (2019), 1844–1863.
- [4] Nets Hawk Katz and Terence Tao, *Some connections between falconer’s distance set conjecture, and sets of furstenberg type*, arXiv preprint math/0101195 (2001).
- [5] János Komlós, János Pintz, and Endre Szemerédi, *On heilbronn’s triangle problem*, *Journal of the London Mathematical Society* **2** (1981), 385–396.
- [6] János Komlós, János Pintz, and Endre Szemerédi, *A lower bound for heilbronn’s problem*, *Journal of the London Mathematical Society* **2** (1982), 13–24.
- [7] Hanno Lefmann, *Distributions of points in the unit square and large k -gons*, *European Journal of Combinatorics* **29** (2008), 946–965.
- [8] Alexander Logunov and Dmitrii Zakharov, *A configuration of point-line pairs with large minimal distance*, to appear (2025).
- [9] Tuomas Orponen, Pablo Shmerkin, and Hong Wang, *Kaufman and falconer estimates for radial projections and a continuum version of beck’s theorem*, *Geometric and Functional Analysis* **34** (2024), 164–201.
- [10] Kevin Ren and Hong Wang, *Furstenberg sets estimate in the plane*, arXiv preprint arXiv:2308.08819 (2023).
- [11] KF Roth, *On a problem of heilbronn, ii*, *Proceedings of the London Mathematical Society* **3** (1972), 193–212.
- [12] KF Roth, *On a problem of heilbronn, iii*, *Proceedings of the London Mathematical Society* **3** (1972), 543–549.
- [13] Klaus F Roth, *On a problem of heilbronn*, *Journal of the London Mathematical Society* **1** (1951), 198–204.
- [14] Wolfgang M Schmidt, *On a problem of heilbronn*, *Journal of the London Mathematical Society* **2** (1972), 545–550.
- [15] Hong Wang and Shukun Wu, *Restriction estimates using decoupling theorems and two-ends furstenberg inequalities*, arXiv preprint arXiv:2411.08871 (2024).
- [16] Hong Wang and Joshua Zahl, *Volume estimates for unions of convex sets, and the kakeya set conjecture in three dimensions*, arXiv preprint arXiv:2502.17655 (2025).
- [17] Thomas Wolff, *An improved bound for kakeya type maximal functions.*, *Revista Matemática Iberoamericana* **11** (1995), 651–674.
- [18] Dmitrii Zakharov, *Upper bounds for heilbronn’s triangle problem in higher dimensions*, *Bulletin of the London Mathematical Society* **56** (2024), 1687–1697.

UNIVERSITY OF CAMBRIDGE, UK AND UNIVERSITY OF CALIFORNIA, LOS ANGELES, USA
Email address: dm672@cam.ac.uk, dmal@math.ucla.edu

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY
Email address: hw3639@nyu.edu

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, USA
Email address: zakhdm@mit.edu