DISTANCE GRAPHS AND SETS OF POSITIVE UPPER DENSITY IN $\mathbb{R}^d$

NEIL LYALL ÁKOS MAGYAR

Abstract. We present a refinement and sharp extension of a result of Bourgain on finding configurations of $k+1$ points in general position in measurable subset of $\mathbb{R}^d$ of positive upper density whenever $d \geq k+1$ to all proper k-degenerate distance graphs.

1. Introduction

1.1. **Background.** A result of Katznelson and Weiss [5] states that if $A \subseteq \mathbb{R}^2$ has positive upper Banach density, then its distance set $\{|x - x'| : x, x' \in A\}$ contains all large sufficiently numbers. Recall that the *upper Banach density* of a measurable set $A \subseteq \mathbb{R}^d$ is defined by

$$\delta^*(A) = \lim_{N\to\infty} \sup_{t\in\mathbb{R}^d} \frac{|A \cap (t + Q_N)|}{|Q_N|}, \tag{1}$$

where $|\cdot|$ denotes Lebesgue measure on $\mathbb{R}^d$ and Q_N denotes the cube $[-N/2, N/2]^d$.

Note that the distance set of any set of positive Lebesgue measure in $\mathbb{R}^d$ automatically contains all sufficiently small numbers (by for example the Lebesgue density theorem) and that it is easy to construct a set of positive upper density which does not contain a fixed distance by placing small balls centered on an appropriate square lattice.

This result was later reproved using Fourier analytic techniques by Bourgain in [2] where he established the following more general result for all finite point configurations $V = \{v_0, v_1, \ldots, v_k\}$ with the property that $\{v_1 - v_0, \ldots, v_k - v_0\}$ forms a linearly independent collection of vectors in $\mathbb{R}^d$, namely to all non-degenerate simplices, in the sequel we shall simply refer to such finite point configurations as being in *general position*.

Theorem 1 (Bourgain [2]). *Let $\Delta_k \subseteq \mathbb{R}^d$ be a fixed collection of $(k+1)$ points in general position.*

If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq k+1$, then there exists a threshold $\lambda_0 = \lambda_0(A, \Delta_k)$ such that A contains an isometric copy of $\lambda \cdot \Delta_k$ for all $\lambda \geq \lambda_0$.

Recall that a point configuration Δ'_k is said to be an isometric copy of $\lambda \cdot \Delta_k$ if there exists a bijection $\phi : \Delta_k \to \Delta'_k$ such that $|\phi(v) - \phi(w)| = \lambda\, |v - w|$ for all $v, w \in \Delta_k$.

Bourgain further demonstrated in [2] that no result along the lines of Theorem 1 can hold for configurations that contain any three points in arithmetic progression on a line, specifically showing that for any $d \geq 1$ there are sets of positive upper Banach density in $\mathbb{R}^d$ which do not contain an isometric copy of configurations of the form $\{0, y, 2y\}$ with $|y| = \lambda$ for all sufficiently large λ. However, in [13], Ziegler showed that if $A \subseteq \mathbb{R}^d$ has positive upper density and $V = \{0, v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$, then there does exists a threshold $\lambda_0 = \lambda_0(A, V)$ such that A_ε contains an isometric copy of $\lambda \cdot V$ for all $\lambda \geq \lambda_0$ and any $\varepsilon > 0$, where A_ε denotes the ε-neighborhood of A.

Together these results may be viewed as initial results in *geometric Ramsey theory* where, roughly speaking, one shows that "large" but otherwise arbitrary sets necessarily contain certain geometric configurations. Recently there has been a number of results in this direction in various contexts, see [3], [1], and [7]. This objective of this article is to present a common extension in the setting of measurable subsets of Euclidean spaces of positive upper Banach density, while simultaneously presenting a new approach to (and refinement of) Theorem 1 based on a simple notion of uniform distribution attached to an appropriate scale. For another instance of this new approach see [8] where configurations of points that form the vertices of a rigid geometric

2010 *Mathematics Subject Classification.* 11B30.
The first author was partially supported by grant NSF-DMS 1702411 and the Simons Foundation Collaboration Grant for Mathematicians 245792. The second author was partially supported by grant NSF-DMS 1600840 and ERC-AdG 321104.

square, and more generally the direct product of any two finite point configurations in general position, are addressed.

1.2. **Distance Graphs and Main Result.** A *distance graph* $\Gamma = \Gamma(V, E)$ is a connected finite graph with vertex set V contained in $\mathbb{R}^d$ for some $d \geq 1$. We say that Γ is *k-degenerate* if each of its subgraphs contain a vertex with degree at most k; that is, some vertex in the subgraph touches k or fewer of the subgraphs edges. It is thus straightforward to verify, by induction, that if a given graph is k-degenerate, then there exists an ordering of its vertex set $V = \{v_0, v_1, \ldots, v_n\}$ in such a way that $|V_j| \leq k$ for all $1 \leq j \leq n$, where

$$V_j := \{v_i \,:\, (v_i, v_j) \in E \text{ with } 0 \leq i < j\} \tag{2}$$

denotes the set of predecessors of the vertex v_j. In this article we shall always assume that the vertices of any given k-degenerate graph have been ordered as such. The *degeneracy* of a graph is defined to be the smallest k for which it is k-degenerate. Finally, we shall refer to a distance graph as *proper* if for every $1 \leq j \leq n$, the set of vertices $v_j \cup V_j$, namely v_j together with its direct predecessors, are in general position.

Given a distance graph $\Gamma = \Gamma(V, E)$ and $\lambda > 0$ we will say that $\Gamma' = \Gamma'(V', E')$ is *isometric* to $\lambda \cdot \Gamma$ if there exists a bijection $\phi : V \to V'$ such that $(v, w) \in E$ if and only if $(\phi(v), \phi(w)) \in E'$ and $|\phi(v) - \phi(w)| = \lambda |v - w|$, and say that Γ' is an *δ-close isometric* copy of $\lambda \cdot \Gamma$ if one has the additional "angular closeness" property that

$$\frac{(\phi(v) - \phi(w)) \cdot (v - w)}{|\phi(v) - \phi(w)|\,|v - w|} > 1 - \delta \tag{3}$$

for all $(v, w) \in E$. Finally, we say that $A \subseteq \mathbb{R}^d$ contains a distance graph $\Gamma = \Gamma(V, E)$ if its vertex set $V \subseteq A$.

The main result of this article is the following

Theorem 2. *Let $\Gamma = \Gamma(V, E)$ be a proper k-degenerate distance graph and $\delta > 0$.*

(i) *If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq k + 1$, then there exists $\lambda_0 = \lambda_0(A, \Gamma, \delta)$ such that A contains an δ-close isometric copy of $\lambda \cdot \Gamma$ for all $\lambda \geq \lambda_0$.*

(ii) *If $A \subseteq [0, 1]^d$ with $|A| > 0$ and $d \geq k + 1$, then A will contain an δ-close isometric copy of $\lambda \cdot \Gamma$ for all λ in some interval of length at least* $\exp(-C_{\Gamma,\delta} |A|^{-C|V|})$.

Intuitively one should visualize a distance graph with edges made of rigid rods which can freely turn around the vertices, an isometric copy of a distance graph in a set $A \subseteq \mathbb{R}^d$ as a folding of the graph so that all of its vertices are supported on A, and an δ-close isometric copy of a distance graph in a set $A \subseteq \mathbb{R}^d$ as a suitably small perturbation of the graph so that all of its vertices are supported on A.

Part (i) of Theorem 2 already constitutes a refinement of Theorem 1 when Γ is simply taken to be a complete distance graph on $(k+1)$ vertices in general position. In this special case it establishes that positive upper density subsets of $\mathbb{R}^d$ not only contain an isometric copy of all sufficiently large dilates of a given non-degenerate simplex, as already guaranteed by Theorem 1, but that these copies can in fact be found as sufficiently large dilates of a "small rotation" of the original simplex. We further note that in both parts of Theorem 2 the dimension d is restricted only by the "level of degeneracy" of the given distance graph and not on the number of its vertices which could in fact be arbitrarily large. It is important to further observe that the length of the interval of dilations guaranteed by Part (ii) of Theorem 2 depends only on the measure of A and not on the set A itself.

Allowing the edges to rotate around the vertices is essential in our arguments. For example, the authors are unaware of any proof that there are k-equally spaced points along a line in a subset of positive density of $\mathbb{R}^2$, with *arbitrary* large gaps that does not invoking Szemerédi's theorem [11], and that such a result is in fact not possible for *all* sufficiently large gaps[1]. The reason being that the linear relations between the points of the pattern are no longer there when we allow for rotations of the edges around the vertices. A crucial observation of this note is that in this case the frequency of isometric copies in a given set is controlled by a simple norm, which may be we viewed as a Euclidean analogue of the so-called U^1-seminorm [12, Chapter 11], utilized in additive combinatorics. In the context of finite field geometries, a geometric analogue of the Gowers U^2-uniformity norm was developed in [9] and used to prove that sets of positive density contain isometric

[1]For $k = 3$ a result of this type was obtained in [4], with the Euclidean distance replaced by the ℓ^p-distance, for all $p \neq 2$.

copies of all circular quadrilaterals. We hope to address such problems for subsets of positive upper density of Euclidean spaces in the future.

As mentioned above, various special cases of our main result have been established, albeit in different contexts. Indeed, in [3] the embedding of large copies of trees (1-degenerate distance graphs) was shown for dense subsets of the integer lattice. In [1] it was shown that measurable subsets $A \subseteq [0,1]^d$ of Hausdorff dimension larger than $\frac{d+1}{2}$ contain an isometric copy of $\lambda \cdot \Gamma$ for all λ in some interval, in the special case when Γ is a finite path. Very recently, parallel to our work, embedding of bounded degree distance graphs was addressed for subsets of vector spaces over finite fields [7].

Examples of Distance Graphs

1. A non-empty connected graph is 1-degenerate if and only if it is a tree (contains no cycles). Any tree with vertices in $\mathbb{R}^d$ with $d \geq 1$ is isometric to a proper 1-degenerate distance graph in $\mathbb{R}^2$.
2. Cycles with vertices in $\mathbb{R}^d$ with $d \geq 1$ form 2-degenerate distance graphs, but these are not necessarily isometric to a proper 2-degenerate distance graph in $\mathbb{R}^d$ for any $d \geq 1$. Indeed, if $V = \{0, 1, 2\} \subseteq \mathbb{R}$ and $E = \{(0,1), (1,2), (0,2)\}$, then this defines just such a distance graph.
3. If $V = \{(i,j) : 0 \leq i,j \leq n\} \subseteq \mathbb{R}^2$ and $E = \{((i,j),(i'j')) : |i-i'| + |j-j'| = 1\}$, then this "2-dimensional grid" forms a proper 2-degenerate distance graph in $\mathbb{R}^2$.

   In general, one can construct a proper 2-degenerate distance graph in $\mathbb{R}^3$ as follows: Start with any proper cycle with vertices in $\mathbb{R}^3$, such as a proper triangle (three vertices in general position) or four vertices forming a "non-rigid" square (no diagonal edges), and at every step attach an edge (or vertex) of another proper cycle (or tree) to any of the edges (or vertices) of the graph constructed at the previous step.

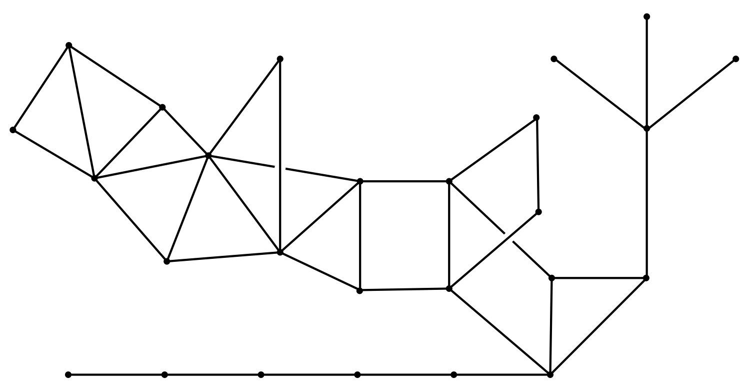

FIGURE 1. An example of a proper 2-degenerate distance graph in $\mathbb{R}^3$

4. A complete graph with vertices $\{v_0, \ldots, v_k\} \subseteq \mathbb{R}^k$ forms a proper k-degenerate distance graph if and only if $\{v_0, \ldots, v_k\}$ are in general position. Another example of a proper k-degenerate distance graph in $\mathbb{R}^k$ is the "k-dimensional grid" with vertices $V = \{(i_1, \ldots, i_k) : 0 \leq i_1, \ldots, i_k \leq n\} \subseteq \mathbb{R}^k$ and edges $E = \{((i_1, \ldots, i_k), (i'_1, \ldots, i'_k)) : |i_1 - i'_1| + \cdots + |i_k - i'_k| = 1\}$.

   More generally, one can construct a proper k-degenerate distance graph in $\mathbb{R}^{k+1}$ as follows: Start with any known proper k-degenerate distance graph with vertices in $\mathbb{R}^{k+1}$ and at every step attach another proper ℓ-degenerate distance graph with $\ell \leq k$ to any of the faces, edges, or vertices of the graph constructed at the previous step.

Remark on the Sharpness of the dimension condition in Theorem 2

Let $e_1, \ldots, e_k$ be the standard basis vectors of $\mathbb{R}^k$ and Δ_+ and Δ_- denote the complete graphs with vertices $\{0, e_1, e_2, \ldots, e_k\}$ and $\{0, -e_1, e_2, \ldots, e_k\}$ respectively. It is clear that $\Gamma = \Delta_+ \cup \Delta_-$ then defines a proper k-degenerate distance graph with the property that any isometric copy of $\lambda \cdot \Gamma$ in $\mathbb{R}^k$ must contain three collinear points, i.e. a copy of $\{-\lambda e_1, 0, \lambda e_1\}$ obtained by a translation and a rotation. As mentioned above, it was shown in [2] that there are sets of positive upper Banach density in $\mathbb{R}^k$, for any k, which do not contain such configurations for all large λ. This example shows the sharpness of the dimension condition $d \geq k+1$ in Theorem 2.

1.3. **Outline of the paper.** In Section 2 we introduce a norm which measures the uniformity of distribution with respect to a scale L. We prove that this norm controls the frequency with which isometric copies of a given distance graph occur in a subset of the unit cube. This is analogous to the so-called von-Neumann type inequalities in additive combinatorics, see for example [12, Chapter 11]. In Section 3 we observe that sets of positive density are uniformly distributed with respect to sufficiently large scales which immediately implies Part (i) of Theorem 2. The proof of Part (ii) is also provided in Section 3 and based on a decomposition of a set into uniformly distributed parts. Section 4 contains an alternative approach inspired by Bourgain's argument in [2]. We include this in order to highlight the simplicity and directness of our approach to Part (i) of Theorem 2, but also with the hope that this will serve to clarify Bourgain's approach and emphasize that our approach to Part (ii) of Theorem 2 is in essence a physical space reinterpretation of Bourgain's original.

2. A Counting Function and Generalized von-Neumann inequality

Let $\Gamma = \Gamma(V, E)$ be a fixed proper k-degenerate distance graph with vertex set $V = \{v_0, v_1, \dots, v_n\}$ with $v_0 = 0$ in $\mathbb{R}^d$ with $d \geq k+1$. As our arguments are analytic, we need to define a measure on the configuration space of all isometric copies of Γ. For each $(v_i, v_j) \in E$ let $t_{ij} = |v_i - v_j|^2$. The configuration space of all isometric copies of Γ, with the vertex v_0 remaining fixed at 0, namely

$$S_\Gamma := \{(x_0, x_1, \dots, x_n) \in \mathbb{R}^{d(n+1)} : x_0 = 0 \text{ and } |x_i - x_j|^2 = t_{ij} \text{ for all } i,j \text{ for which } (v_i, v_j) \in E\} \tag{4}$$

is clearly a real subvariety. Since Γ is proper, there exists a points $(x_0, x_1, \dots, x_n) \in S_\Gamma$, for example one could simply take $x_j = v_j - v_0$ for $0 \leq j \leq n$, with the property that for all $1 \leq j \leq n$, the sets $\overline{X}_j := x_j \cup X_j$ where $X_j := \{x_i : v_i \in V_j\}$ are in general position and each of the spheres

$$S_j = S_j(X_j) := \{x \in \mathbb{R}^d : |x - x_i|^2 = t_{ij} \text{ for all } x_i \in X_j\} \tag{5}$$

depending on the set of points X_j have dimension $d - |X_j|$, which is at least 1 if our distance graph Γ is k-degenerate and $d \geq k+1$. Note further that

$$(0, x_1, \dots, x_n) \in S_\Gamma \iff x_j \in S_j(X_j) \ \text{ for all } \ 1 \leq j \leq n. \tag{6}$$

Since the affine subspaces spanned by S_j and X_j respectively are orthogonal (and the points of $\overline{X}_j$ are in general position) it follows that that the radius r_j of S_j is positive and equal to the distance from x_j to the affine subspace spanned by X_j. Specifically, if $X_j = \{x_{i_1}, \dots, x_{i_\ell}\}$, then the fact that $\overline{X}_j = \{x_{i_1}, \dots, x_{i_\ell}, x_j\}$ is in general position ensures that the volume of the ℓ-dimensional fundamental parallelotope determined by the vectors $\{x_j - x_{i_1}, \dots, x_j - x_{i_\ell}\}$ is non-zero. Since the volume of this parallelotope is equal to the square root of the so-called Gram determinant, namely the determinant of the (Gram) inner product matrix

$$\det\{(x_j - x_{i_{m_1}}) \cdot (x_j - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell}$$

it follows that

$$r_j = \sqrt{\frac{\det\{(x_j - x_{i_{m_1}}) \cdot (x_j - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell}}{\det\{(x_{i_\ell} - x_{i_{m_1}}) \cdot (x_{i_\ell} - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell - 1}}} \tag{7}$$

as r_j is the height of our parallelotope if we take its base to be the $(\ell-1)$-dimensional parallelotope determined by the vectors $\{x_{i_\ell} - x_{i_1}, \dots, x_{i_\ell} - x_{i_{\ell-1}}\}$. It follows that $r_j = r_j(X_j)$, in addition to being positive, in fact depends continuously on the points in X_j.

An important consequence of the discussion above that there exists a constant $r_\Gamma > 0$ and compactly supported functions $\eta_j \in C^\infty(\mathbb{R}^d)$ with $0 \leq \eta_j \leq 1$ such that for all $(x_0, x_1, \dots, x_n) \in S_\Gamma$ with $x_j \in \operatorname{supp} \eta_j$, the corresponding spheres S_j will all have radius $r_j \geq r_\Gamma$.

Definition 2.1 (Localized Counting Function)**.** For any $0 < \lambda \ll 1$ and functions

$$f_0, f_1, \dots, f_n : [0,1]^d \to \mathbb{R}$$

with $d \geq k+1$ we define

$$T_\Gamma(f_0, f_1, \dots, f_n)(\lambda) = \iint \cdots \int f_0(x) f_1(x - \lambda x_1) \cdots f_n(x - \lambda x_n)\, d\mu_n(x_n) \cdots d\mu_1(x_1)\, dx \tag{8}$$

where $d\mu_j(x_j) = \eta_j(x_j)\, d\sigma_j(x_j)$ and σ_j denotes the normalized surface measure on S_j.

Note that if $A \subseteq [0,1]^d$ with $d \geq k+1$ and $T_\Gamma(1_A, 1_A, \dots, 1_A)(\lambda) > 0$ with the support of each function η_j taken suitably small with respect to some initially prescribed $\delta > 0$ and containing the vertices v_j of Γ, then A must contain a point configuration $\Gamma' = \{x, x + \lambda x_1, \dots, x + \lambda x_n\}$ with each $x_j \in S_j(X_j)$, and hence by (6) an δ-close isometric copy of $\lambda \cdot \Gamma$.

The key to showing that $T_\Gamma(1_A, 1_A, \dots, 1_A)(\lambda)$ is positive for certain sets A, and functions η_j, is to estimate (8) in terms of a suitable uniformity norm localized to a scale L (related to λ).

Definition 2.2 ($U^1(L)$-norm)**.** For $0 < L \ll 1$ and functions $f : [0,1]^d \to \mathbb{R}$ we define

$$\|f\|_{U^1(L)} = \|f * \varphi_L\|_2$$

where $\varphi_L(x) = L^{-d}\, \varphi(L^{-1}x)$ with $\varphi = 1_{[-1/2,1/2]^d}$.

Note that if $A \subseteq [0,1]^d$ with $\alpha = |A| > 0$ and we define $f_A := 1_A - \alpha 1_{[0,1]^d}$, then

$$\|f_A\|^2_{U^1(L)} = \int_{\mathbb{R}^d} \left| \frac{|A \cap (t + Q_L)|}{|Q_L|} - |A| \right|^2 dt, \tag{9}$$

where $Q_L = [-L/2, L/2]^d$.

Evidently the $U^1(L)$-norm is measuring the mean-square uniform distribution of A on scale L. The engine that drives our approach to Theorem 2 is the following

Proposition 1 (Generalized von-Neumann)**.** *Let $0 < \varepsilon, \lambda \ll 1$. For any $L \leq \varepsilon^6 \lambda$, $0 \leq m \leq n$ and functions*

$$f_0, f_1, \dots, f_m : [0,1]^d \to [-1,1]$$

we have that

$$|T_\Gamma(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)| \leq \|f_m\|_{U^1(L)} + O_\Gamma(\varepsilon).$$

Here 1 stands for the indicator function of the unit cube $[0,1]^d$ and $O_\Gamma(\varepsilon)$ means a quantity bounded by $C_\Gamma\, \varepsilon$ with a constant C_Γ depending only on Γ. We will also use the notation $f \ll_\Gamma g$ to indicate that $|f| \leq c_\Gamma\, g$ with a constant $c_\Gamma > 0$ *sufficiently small* for our purposes.

The above proposition immediately implies the following result for uniformly distributed sets from which we will deduce both parts of Theorem 2 in Section 3 below.

Corollary 1. *Let Γ be a proper k-degenerate distance graph with $n+1$ vertices in $\mathbb{R}^d$ with $d \geq k+1$.*

Let $\alpha \in (0,1)$ and $0 < \lambda \leq \varepsilon \ll_\Gamma \alpha^{n+1}$. If $A \subseteq [0,1]^d$ with $|A| = \alpha$ satisfies $\|f_A\|_{U^1(\varepsilon^6\lambda)} \ll \varepsilon$, then

$$T_\Gamma(1_A, 1_A, \dots, 1_A)(\lambda) \geq \frac{c_0}{2}\, \alpha^{n+1}$$

where

$$c_0 = \iint \cdots \int d\mu_n(x_n) \cdots d\mu_1(x_1)\, dx$$

Proof. The result follows immediately from Proposition 1 since

$$T_\Gamma(1_A, \dots, 1_A)(\lambda) = c_0\, \alpha^{n+1} + \sum_{m=0}^{n} \alpha^{n-m}\, T_\Gamma(\underbrace{1_A, \dots, 1_A}_{m \text{ copies}}, f_A, 1, \dots, 1)(\lambda)$$

where $f_A = 1_A - \alpha 1_{[0,1]^d}$. □

We conclude this section with the proof of Proposition 1.

Proof of Proposition 1. Fix $0 \le m \le n$. We have

$$|T_\Gamma(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)| \le \int \cdots \int \left(\int \left| \int f_m(x - \lambda x_m)\, c_{m+1}(x_1, \dots, x_m)\, d\mu_m(x_m) \right| dx \right) d\mu_{m-1}(x_{m-1}) \cdots d\mu_1(x_1)$$

where

$$c_{m+1}(x_1, \dots, x_m) = \int \cdots \int d\mu_n(x_n) \cdots d\mu_{m+1}(x_{m+1}) \tag{10}$$

if $0 \le m \le n-1$ and $c_{n+1} = 1$. It follows from an application of Cauchy-Schwarz and Plancherel that

$$|T_\Gamma(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)|^2 \le \int |\widehat{f_m}(\xi)|^2 I_m(\lambda\, \xi)\, d\xi \tag{11}$$

where

$$I_m(\xi) = \int \cdots \int \left| \widehat{c_{m+1}\mu_m}(\xi) \right|^2 d\mu_{m-1}(x_{m-1}) \cdots d\mu_1(x_1) \tag{12}$$

with

$$\widehat{c_{m+1}\mu_m}(\xi) = \int c_{m+1}(x_1, \dots, x_m) \eta_m(x_m)\, e^{-2\pi i x_m \cdot \xi}\, d\sigma_m(x_m)$$

if $2 \le m \le n$ and $I_1 = |\widehat{c_2\mu_1}|^2$. In light of the trivial uniform bound $0 \le I_m(\xi) \le 1$ and the fact that

$$\|f_m\|^2_{U^1(L)} = \int |\widehat{f_m}(\xi)|^2 |\widehat{\varphi}(L\xi)|^2\, d\xi$$

it suffices to establish that

$$I_m(\lambda\xi)(1 - \widehat{\varphi}(L\xi)^2) = O_\Gamma(\varepsilon^2). \tag{13}$$

Since $0 \le \widehat{\varphi}(\xi)^2 \le 1$ for all $\xi \in \mathbb{R}^d$ and $\widehat{\varphi}(0) = 1$ it follows that $0 \le 1 - \widehat{\varphi}(L\xi)^2 \le \min\{1, 4\pi L|\xi|\}$. The uniform bound (13) thus reduces to establishing the decay estimate

$$I_m(\xi) \le \min\{1, C_\Gamma\, |\xi|^{-1/2}\} \tag{14}$$

since this would in turn imply that

$$I_m(\lambda\xi)(1 - \widehat{\varphi}(L\xi)^2) \le C_\Gamma \min\{(\lambda|\xi|)^{-1/2}, \varepsilon^6 \lambda|\xi|\} \le C_\Gamma\, \varepsilon^2$$

whenever $L \le \varepsilon^6 \lambda$.

To establish (14) we will use the fact that in addition to being trivially bounded by 1, the Fourier transform of $c_{m+1}\mu_m$ also decays for large ξ in certain directions, specifically

$$|\widehat{c_{m+1}\mu_m}(\xi)| \le \min\{1, (r_\Gamma \cdot (\operatorname{dist}(\xi, \operatorname{span} X_m))^{-1/2}\} \tag{15}$$

uniformly over all $x_1, \dots, x_{m-1}$ with $x_j \in \operatorname{supp} \eta_j$. This estimate is an easy consequence of the well-known asymptotic behavior of the Fourier transform of the measure on the unit sphere $S^{d-|X_m|} \subseteq \mathbb{R}^{d-|X_m|+1}$ induced by Lebesgue measure, see for example [10].

Using the fact that the measure $d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1)$ is clearly invariant under the rotations

$$(x_1, \dots, x_m) \to (Ux_1, \dots, Ux_m),$$

for any $U \in SO(d)$, together with (15) and the fact that $0 \le \eta_j \le 1$ for $1 \le j \le m$, then gives

$$\begin{aligned}
I_m(\xi) &\le C \int \cdots \int (1 + r_\Gamma \cdot \operatorname{dist}(\xi, \operatorname{span} X_m))^{-1}\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1) \\
&= C \int \cdots \int \int_{SO(d)} (1 + r_\Gamma \cdot \operatorname{dist}(\xi, \operatorname{span} UX_m))^{-1}\, d\mu(U)\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1) \\
&= C \int \cdots \int \int_{S^{d-1}} (1 + r_\Gamma |\xi| \cdot \operatorname{dist}(y, \operatorname{span} X_m))^{-1}\, d\sigma(y)\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1)
\end{aligned}$$

where σ denote normalized measure on the unit sphere S^{d-1} in $\mathbb{R}^d$ induced by Lebesgue measure. Estimate (14) then follows from the easy observation that the inner integral above satisfies the uniform estimate

$$\int_{S^{d-1}} (1 + r_\Gamma|\xi| \cdot \text{dist}(y, \text{span } X_m))^{-1}\, d\sigma(y) = O((1 + r_\Gamma|\xi|)^{-1/2}). \tag{16}$$ □

3. Proof of Theorem 2

We will deduce Theorem 2 from Corollary 1 by localizing to cubes on which our set in suitably uniformly distributed. In the case of Part (i) this is achieved as a direct consequence of the definition of upper Banach density, while for Part (ii) this is achieved via an energy increment argument.

3.1. **Direct Proof of Part (i) of Theorem 2.** Let $\varepsilon > 0$ and $A \subseteq \mathbb{R}^d$ with $\delta^*(A) > 0$.

The following two facts follow immediately from the definition of upper Banach density, see (1):

(i) There exist $M_0 = M_0(A, \varepsilon)$ such that for all $M \geq M_0$ and all $t \in \mathbb{R}^d$

$$\frac{|A \cap (t + Q_M)|}{|Q_M|} \leq (1 + \varepsilon^4/3)\, \delta^*(A).$$

(ii) There exist arbitrarily large $N \in \mathbb{R}$ such that

$$\frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq (1 - \varepsilon^4/3)\, \delta^*(A)$$

for some $t_0 \in \mathbb{R}^d$.

Combining (i) and (ii) above we see that for any $\lambda \geq \lambda_0 := \varepsilon^{-6} M_0$, there exist $N \geq \varepsilon^{-6}\lambda$ and $t_0 \in \mathbb{R}^d$ such that

$$\frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 + \varepsilon^4) \frac{|A \cap (t_0 + Q_N)|}{|Q_N|}$$

for all $t \in \mathbb{R}^d$. Consequently, Theorem 2 reduces, via a rescaling of $A \cap (t_0 + Q_N)$ to a subset of $[0, 1]^d$, to establishing that if Γ is a proper k-degenerate distance graph, $0 < \lambda \leq \varepsilon \ll 1$ and $A \subseteq [0, 1]^d$ is measurable with $|A| > 0$ and the property that

$$\frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 + \varepsilon^4)\, |A|$$

for all $t \in \mathbb{R}^d$, then A contains an isometric copy of $\lambda \cdot \Gamma$.

Now since $A \cap (t + Q_{\varepsilon^6\lambda})$ is only supported in $[-\varepsilon^6\lambda, 1 + \varepsilon^6\lambda]^d$ and

$$|A| = \int_{\mathbb{R}^d} \frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|}\, dt$$

it easily follows that

$$\left|\left\{t \in \mathbb{R}^d \,:\, 0 < \frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 - \varepsilon^2)\, |A|\right\}\right| = O(\varepsilon^2)$$

and hence that

$$\|f_A\|^2_{U^1(\varepsilon^6\lambda)} = \int_{\mathbb{R}^d} \left|\frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} - |A|\right|^2 dt = O(\varepsilon^2).$$

The result thus follows from Corollary 1 above provided $\varepsilon \ll \delta^*(A)^{n+1}$. □

3.2. **Proof of Part (ii) of Theorem 2.**

Lemma 1 (Localization Principle)**.** *Let $A\subseteq[0,1]^d$ with $d\geq k+1$ and $|A|=\alpha>0$.*

Let $\varepsilon>0$ and $\varepsilon^7\gg L_1\gg L_2\gg\cdots$ be any decreasing sequence with $L_1^{-1}\in\mathbb{N}$ and $L_{j+1}\leq c\,\varepsilon^7 L_j$ with $L_{j+1}|L_j$ for all $j\geq 1$. If we let $\mathcal{G}_j$ denote the partition of $[0,1]^d$ into cubes of sidelength L_j, then there exists $1\leq j\leq C\varepsilon^{-2}$ such that for all but at most εL_j^{-d} of the cubes Q in $\mathcal{G}_j$ the set A will be uniform distributed on the smaller scale L_{j+1} inside Q in the sense that

$$\frac{1}{|Q|}\int_Q\left|\frac{|A\cap Q\cap(t+Q_{L_{j+1}})|}{|Q_{L_{j+1}}|}-\frac{|A\cap Q|}{|Q|}\right|^2dt\leq\varepsilon. \tag{17}$$

Before proving Lemma 1 we first show that it, together with Corollary 1 (after rescaling), is sufficient to establish Part (ii) of Theorem 2. Let $\varepsilon\ll_\Gamma\alpha^{n+1}$ and $\{Q_i\}$ denote the cubes of sidelength L_j in the partition $\mathcal{G}_j$ of $[0,1]^d$ that we obtain from Lemma 1. If we then let $A_i=A\cap Q_i$ and set $\alpha_i=|A\cap Q_i|/|Q_i|$ it follows from Corollary 1 (after rescaling) and Hölder's inequality that for any $\lambda\in(\varepsilon^{-6}L_{j+1},\varepsilon L_j)$ we have

$$T_\Gamma(1_A,\ldots,1_A)(\lambda)\geq\sum_{i=1}^{L_j^{-d}}T_\Gamma(1_{A_i},\ldots,1_{A_i})(\lambda)\geq\frac{c_0}{4}\,L_j^d\sum_{i=1}^{L_j^{-d}}\alpha_i^{n+1}\geq\frac{c_0}{4}\left(L_j^d\sum_{i=1}^{L_j^{-d}}\alpha_i\right)^{n+1}=\frac{c_0}{4}\,|A|^{n+1}. \tag{18}$$

Proof of Lemma 1. Let $\{Q_i\}$ denote the cubes of sidelength L_j in the partition $\mathcal{G}_j$ of $[0,1]^d$ and

$$g_j=1_A-\mathbb{E}(1_A\,|\,\mathcal{G}_j)$$

where

$$\mathbb{E}(1_A\,|\,\mathcal{G}_j)(x)=\frac{|A\cap Q_i|}{|Q_i|}$$

for each $x\in Q_i$. If $\|g_j\|_{U^1(L_{j+1})}\geq\varepsilon$, then by definition

$$\int\left|\frac{1}{|Q_{L_{j+1}}|}\int_{x+Q_{L_{j+1}}}g_j(y)\,dy\right|^2dx\geq c\,\varepsilon^2.$$

It follows that there must exist a $x_0\in[0,1]^d$ for which the shifted grid $x_0+\mathcal{G}_{j+1}$ satisfies

$$\int\left|\mathbb{E}(g_j\,|\,x_0+\mathcal{G}_{j+1})\right|^2dx\geq c\,\varepsilon^2$$

from which one can easily conclude that the (unshifted) refined grid $\mathcal{G}_{j+2}$ satisfies

$$\int\left|\mathbb{E}(g_j\,|\,\mathcal{G}_{j+2})\right|^2dx\geq c\,\varepsilon^2 \tag{19}$$

provided $L_{j+2}\ll\varepsilon^2L_{j+1}$. By orthogonality, it follows immediately from (19) and the definition of g_j that

$$\|\mathbb{E}(1_A\,|\,\mathcal{G}_{j+2})\|_2^2\geq\|\mathbb{E}(1_A\,|\,\mathcal{G}_j)\|_2^2+c\varepsilon^2 \tag{20}$$

and hence that there must exist $1\leq j\leq C\varepsilon^{-2}$ such that $\|g_j\|_{U^1(L_{j+1})}\leq\varepsilon$ from which it follows that

$$\sum_{i=1}^{L_j^{-d}}\int\left|\frac{1}{|Q_{L_{j+1}}|}\int_{x+Q_{L_{j+1}}}(1_{A_i}-\alpha_i1_{Q_i})(y)\,dy\right|^2dx\leq C\varepsilon^2$$

provided $L_{j+1}\ll\varepsilon^2L_j$. □

4. A Second Proof of Theorem 2

We conclude by presenting a second proof of Theorem 2 which is closer in spirit to Bourgain's original proof of Theorem 1. We include this in order to highlight the simplicity and directness of our approach to Part (i) of Theorem 2 above, but also to emphasize that our approach to Part (ii) of Theorem 2 is in essence a physical space reinterpretation of Bourgain original approach.

4.1. **Reducing Theorem 2 to a Dichotomy between Randomness and Structure.** Let Γ be a proper k-degenerate distance graph in $[0,1]^d$ with $d \geq k+1$. As we shall see, Theorem 2 is an immediate consequence of the following proposition which reveals that if $A \subseteq [0,1]^d$ has positive measure but does not contain an isometric copy of $\lambda \cdot \Gamma$ for all λ in a given interval, then this "non-random" behavior is detected by the Fourier transform of the characteristic function of A and results in "structural information", specifically a concentration of its L^2-mass on appropriate annuli.

Proposition 2 (Dichotomy)**.** *Let $\Gamma = \Gamma(V,E)$ be a proper k-degenerate distance graph in $[0,1]^d$ with $d \geq k+1$.*

If $A \subseteq [0,1]^d$ with $|A| > 0$, $0 < a \leq b \ll \varepsilon^4$ with $0 < \varepsilon \ll_\Gamma |A|^{n+1}$, and A does not contain an isometric copy of $\lambda \cdot \Gamma$ for some λ in $[a,b]$, then

$$\int_{\varepsilon^2/b \leq |\xi| \leq 1/\varepsilon^2 a} |\widehat{1_A}(\xi)|^2 \, d\xi \gg |A|^{2n+2} \tag{21}$$

with the implied constant above independent of a, b, and ε.

Proof that Proposition 2 implies Theorem 2. We shall first establish Part (ii) of Theorem 2, so we start by letting $A \subseteq [0,1]^d$ with $|A| > 0$. For any fixed $0 < \varepsilon \ll_\Gamma |A|^{n+1}$, let $\{\mathcal{I}_j\}_{j=1}^{J(\varepsilon)}$ denote a sequence of intervals with $\mathcal{I}_j := [a_j, b_j]$ satisfying

$$b_{j+1} \ll \varepsilon^4 a_j \tag{22}$$

and $b_1 \ll \varepsilon^4$ with the property that for each $1 \leq j \leq J(\varepsilon)$ there exists a $\lambda \in \mathcal{I}_j$ such that

$$x + \lambda \cdot U(\Delta) \not\subseteq A \tag{23}$$

for all $x \in A$ and $U \in SO(d)$. Proposition 2, together with (22), would then imply that

$$J(\varepsilon)\, \varepsilon^2 \leq \sum_{j=1}^{J(\varepsilon)} \int_{\varepsilon^2/b_j \leq |\xi| \leq 1/\varepsilon^2 a_j} |\widehat{1_A}(\xi)|^2 \, d\xi \leq \int |\widehat{1_A}(\xi)|^2 \, d\xi \tag{24}$$

a contradiction if $J(\varepsilon) \gg \varepsilon^{-2}$ since by Plancherel we know that $\int |\widehat{1_A}(\xi)|^2 \, d\xi = |A| \leq 1$.

To establish Part (i) of Theorem 2 with this approach we will argue indirectly and thus suppose that $A \subseteq \mathbb{R}^d$ is a set with $\delta^*(A) > 0$ for which the conclusion of Part (i) of Theorem 2 fails to hold, namely that there exist arbitrarily large $\lambda \in \mathbb{R}$ for which A does not contain an isometric copy of $\lambda \cdot \Gamma$.

We now let $0 < \alpha < \delta^*(A)$, $0 < \varepsilon \ll_\Gamma \alpha^{n+1}$, and fix $J \gg \varepsilon^{-2}$ as above. By our indirect assumption we can choose a sequence $\{\lambda_j\}_{j=1}^J$ with the property that $\lambda_{j+1} \ll \varepsilon^4 \lambda_j$ for all $1 \leq j \leq J-1$ and A does not contain an isometric copy of $\lambda_j \cdot \Gamma$ for each $1 \leq j \leq J$. It follows from the definition of upper Banach density that exist $N \in \mathbb{R}$ with $N \gg \lambda_1$ and $t_0 \in \mathbb{R}^d$ for which

$$\frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq \alpha.$$

Rescaling $A \cap (t_0 + Q_N)$ to a subset of $[0,1]^d$ and arguing as in the proof of Part (ii) above but this time with $b_j = \lambda_j / N$ again leads to a contradiction. □

4.2. **Proof of Proposition 2.** Let $f = 1_A$ and $\Gamma = \{0, v_1, \ldots, v_n\}$ be a fixed proper k-degenerate distance graph. We will utilize the existence of a suitably *smoothed* version of f with the certain properties, specifically

Lemma 2. *For any $\varepsilon > 0$ there exists a function $g : \mathbb{R}^d \to (0,1]$, an appropriate smoothing of f, such that*

$$|g(x - \lambda z) - g(x)| \ll \varepsilon \tag{25}$$

uniformly in $x \in [0,1]^d$ and $|z| \leq 1$. Moreover, if $\varepsilon \ll |A|^{n+1}$, then

$$\int f(x) g(x)^n \, dx \gg |A|^{n+1}. \tag{26}$$

The proof of Lemma 2 is straightforward and presented in Section 4.3 below. Assuming for now the existence of a function g with property (25) it follows that

$$T_\Gamma(f,f,\dots,f)(\lambda)=\int f(x)g(x)^n\,dx+\sum_{m=1}^{n}T_\Gamma(fg^{n-m},f,\dots,f,f-g,\underbrace{1,\dots,1}_{n-m\text{ copies}})(\lambda)+O(n\varepsilon)\tag{27}$$

where, as in (8) in Section 2, we define

$$T_\Gamma(f_0,f_1,\dots,f_n)(\lambda)=\iint\cdots\int f_0(x)f_1(x-\lambda x_1)\cdots f_n(x-\lambda x_n)\,d\mu_n(x_n)\cdots d\mu_1(x_1)\,dx$$

with $d\mu_j(x_j)=\eta_j(x_j)\,d\sigma_j(x_j)$ and σ_j denotes the normalized surface measure on S_j.

If A does not contain an isometric copy of $\lambda\cdot\Gamma$ for some λ in $[a,b]$, then it clearly follows that

$$T_\Gamma(f,f,\dots,f)(\lambda)=0.$$

In light of (26) and (27) it follows that if $\varepsilon\ll|A|^{n+1}/n$ then there must exist $1\le m\le n$ such that

$$\int\cdots\int\left(\int\left|\int[f-g](x-\lambda x_m)\,c_{m+1}(x_1,\dots,x_m)\,d\mu_m(x_m)\right|dx\right)d\mu_{m-1}(x_{m-1})\cdots d\mu_1(x_1)\gg|A|^{n+1}\tag{28}$$

with c_{m+1} defined as before in equation (10) above. It then follows from an application of Cauchy-Schwarz and Plancherel that

$$\int|\widehat{f}(\xi)-\widehat{g}(\xi)|^2\,I_m(\lambda\xi)\,d\xi\gg|A|^{2n+2}\tag{29}$$

with I_m again defined as before in equation (12) above. The fact that g will be taken to be a sufficient smoothing of f ensures that its Fourier transform satisfies

$$|\widehat{f}(\xi)-\widehat{g}(\xi)|\le\varepsilon\,|\widehat{f}(\xi)|\tag{30}$$

provided $|\xi|\le\varepsilon^2b^{-1}$, see Section 4.3 below. This, together with the fact that $I_m(\xi)$ is bounded by 1 uniformly in ξ, and Plancherel, ensures that (29) implies

$$\int_{\varepsilon^2/b\le|\xi|}|\widehat{f}(\xi)|^2\,I_m(\lambda\xi)\,d\xi\gg|A|^{2n+2}.\tag{31}$$

Estimate (21), and hence Proposition 2, then follows easily from estimate (31) and our previously established estimates for I_m, namely (14).

4.3. **A smooth cutoff function and Proof of Lemma 2.**

4.3.1. *A smooth cutoff function.* Let $\psi:\mathbb{R}^d\to(0,\infty)$ be a Schwartz function that satisfies

$$1=\widehat{\psi}(0)\ge\widehat{\psi}(\xi)\ge0\qquad\text{and}\qquad\widehat{\psi}(\xi)=0\ \text{ for }\ |\xi|>1.$$

As usual, for any given $t>0$, we define

$$\psi_t(x)=t^{-d}\psi(t^{-1}x).\tag{32}$$

First we record the trivial observation that

$$\int\psi_t(x)\,dx=\int\psi(x)\,dx=\widehat{\psi}(0)=1\tag{33}$$

as well as the simple, but important, observation that ψ may be chosen so that

$$\left|1-\widehat{\psi}_t(\xi)\right|=\left|1-\widehat{\psi}(t\xi)\right|\ll\min\{1,t|\xi|\}.\tag{34}$$

Finally we record a formulation, appropriate to our needs, of the fact that for any given small parameter ε, our cutoff function $\psi_t(x)$ will be essentially supported where $|x|\le\varepsilon^{-1}t$ and is approximately constant on smaller scales. More precisely,

Lemma 3. *Let $\varepsilon > 0$ and $t > 0$, then*

$$\int_{|y|\geq \varepsilon^{-1}t} \psi_t(y)\,dy \ll \varepsilon. \tag{35}$$

and

$$\int \big|\psi_t(y-\lambda z) - \psi_t(y)\big|\,dy \ll \varepsilon \tag{36}$$

uniformly for $|z|\leq 1$, provided $t \gg \varepsilon^{-1}\lambda$.

Proof of Lemma 3. Estimate (35) is easily verified using the fact that ψ is a Schwartz function on $\mathbb{R}^d$ as

$$\int_{|y|\geq \varepsilon^{-1}t} \psi_t(y)\,dy = \int_{|y|\geq \varepsilon^{-1}} \psi(y)\,dy \ll \int_{|y|\geq\varepsilon^{-1}} (1+|y|)^{-d-1}\,dy \ll \varepsilon.$$

To verify estimate (36) we make use of the fact that both ψ and its derivative are rapidly decreasing, specifically

$$\int \big|\psi_t(y-\lambda z) - \psi_t(x)\big|\,dy \leq \int \big|\psi(y-\lambda z/t) - \psi(y)\big|\,dy \ll \frac{\lambda}{t}\int (1+|y|)^{-d-1}\,dy \ll \frac{\lambda}{t}. \qquad \square$$

4.3.2. *Proof of Lemma 2.* Let $g = f * \psi_{\varepsilon^{-1}b}$.

We first note that estimates (30) and (25) follow immediately from (34) and (36) respectively. In order to establishing the remaining "main term" estimate (26), we need only establish that if $\varepsilon \ll |A|^{n+1}$, then

$$\int f(x)g(x)\,dx \geq (1-C\varepsilon)\,|A|^2 \tag{37}$$

for some constant $C>0$, since by Hölder we would then obtain

$$(1-C\varepsilon)^n\,|A|^{2n} \leq \Big(\int f(x)g(x)\,dx\Big)^n \leq |A|^{n-1}\int f(x)g(x)^n\,dx$$

from which (26) clearly follows for sufficiently small $\varepsilon > 0$.

To establish (37) we first note that Parseval, the fact that $0\leq\widehat{\psi}\leq 1$, and a final application of Cauchy-Schwarz gives

$$\int f(x)g(x)\,dx = \int |\widehat{f}(\xi)|^2\widehat{\psi}(\varepsilon^{-1}b\,\xi)\,d\xi \geq \int |\widehat{f}(\xi)|^2|\widehat{\psi}(\varepsilon^{-1}b\,\xi)|^2\,d\xi = \int g(x)^2\,dx \geq \Big(\int_{[0,1]^d} g(x)\,dx\Big)^2. \tag{38}$$

Establishing (37) therefore reduces to showing that if $\varepsilon \ll |A|^3$, then

$$\int_{[0,1]^d} g(x)\,dx \geq (1-C\varepsilon)|A| \tag{39}$$

for some constant $C>0$. To establish (39) we use (33) and write

$$|A| = \int_{\mathbb{R}^d} g(x)\,dx = \int_{[0,1]^d} g(x)\,dx \;+ \int_{\{x\in\mathbb{R}^d\,:\,\mathrm{dist}(x,[0,1]^d)\geq\varepsilon^{-2}b\}} g(x)\,dx \;+ \int_{\{x\in\mathbb{R}^d\,:\,0<\mathrm{dist}(x,[0,1]^d)<\varepsilon^{-2}b\}} g(x)\,dx. \tag{40}$$

The fact that $b\leq\varepsilon^4$ ensures that

$$\big|\{x\in\mathbb{R}^d \,:\, 0<\mathrm{dist}(x,[0,1]^d)<\varepsilon^{-2}b\}\big| \ll \varepsilon^2 \tag{41}$$

and hence, since $\varepsilon\ll|A|$ and $0\leq g\leq 1$, that

$$\int_{\{x\in\mathbb{R}^d\,:\,0<\mathrm{dist}(x,[0,1]^d)<\varepsilon^{-2}b\}} g(x)\,dx \ll \varepsilon^2 \leq \varepsilon|A|$$

while (35) ensures that

$$\int_{\{x\in\mathbb{R}^d\,:\,\mathrm{dist}(x,[0,1]^d)\geq\varepsilon^{-2}b\}} g(x)\,dx \leq |A|\int_{|y|\gg\varepsilon^{-2}b} \psi_{\varepsilon^{-1}b}(y)\,dy \ll \varepsilon|A| \tag{42}$$

which completes the proof. $\square$

References

[1] M. Bennett, A. Iosevich, K. Taylor *Finite chains inside thin subsets of* $\mathbb{R}^d$, Anal. PDE. (2015); 9:597-614.

[2] J. Bourgain, *A Szemerédi type theorem for sets of positive density in* $\mathbb{R}^k$, Israel J. Math. 54 (1986), no. 3, 307–316.

[3] K. Bulinski, *Spherical Recurrence and locally isometric embeddings of trees into positive density subsets of* $\mathbb{Z}^d$, to appear in Math. Proc. Cambridge Philos. Soc.

[4] B. Cook, Á. Magyar, M. Pramanik, *A Roth-type theorem for dense subsets of* $\mathbb{R}^d$, Bull. Lond. Math. Soc. 49 (2017), no. 4, 676-689.

[5] H. Furstenberg, Y. Katznelson and B. Weiss, *Ergodic theory and configurations in sets of positive density*, Israel J. Math. 54 (1986), no. 3, 307–316.

[6] W. T. Gowers. *A new proof of Szemerdi's theorem*, Geom. Funct. Anal. 11 (2001), no. 3, 465-588.

[7] A. Iosevich and H. Parshall, *Embedding distance graphs in finite field vector spaces,* arXiv:1802.06460

[8] N. Lyall and Á. Magyar, *Product of simplices and sets of positive upper density in* $\mathbb{R}^d$, to appear in Math. Proc. Cambridge Philos. Soc.

[9] N. Lyall, Á. Magyar, H. Parshall, *Spherical configurations over finite fields*, to appear in Amer. J. Math.

[10] E. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality and Oscillatory Integrals*, Princeton University Press, Princeton, NJ., 1993.

[11] E., Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199-245.

[12] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, 105. Cambridge University Press, Cambridge, 2006. xviii+512 pp.

[13] T. Ziegler, *Nilfactors of* $\mathbb{R}^m$*-actions and configurations in sets of positive upper density in* $\mathbb{R}^m$, J. Anal. Math. 99 (2006), 249-266.

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: lyall@math.uga.edu

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: magyar@math.uga.edu