Mathematische Zeitschrift (2021) 297:855–865
https://doi.org/10.1007/s00209-020-02537-1

Mathematische Zeitschrift

Existence of similar point configurations in thin subsets of $\mathbb{R}^d$

Allan Greenleaf[1] · Alex Iosevich[1] · Sevak Mkrtchyan[1]

Received: 8 January 2019 / Accepted: 16 April 2020 / Published online: 5 June 2020

Abstract

We prove the existence of similar and multi-similar point configurations (or simplexes) in sets of fractional Hausdorff dimension in Euclidean space. Let $d \geq 2$ and $E \subset \mathbb{R}^d$ be a compact set. For $k \geq 1$, define

$$\Delta_k(E) = \left\{ \left(|x^1 - x^2|, \ldots, |x^i - x^j|, \ldots, |x^k - x^{k+1}| \right) : \left\{ x^i \right\}_{i=1}^{k+1} \subset E \right\} \subset \mathbb{R}^{k(k+1)/2},$$

the $(k+1)$-*point configuration set* of E. For $k \leq d$, this is (up to permutations) the set of congruences of $(k+1)$-point configurations in E; for $k > d$, it is the edge-length set of $(k+1)$-graphs whose vertices are in E. Previous works by a number of authors have found values $s_{k,d} < d$ so that if the Hausdorff dimension of E satisfies $\dim_{\mathcal{H}}(E) > s_{k,d}$, then $\Delta_k(E)$ has positive Lebesgue measure. In this paper we study more refined properties of $\Delta_k(E)$, namely the existence of similar or multi–similar configurations. For $r \in \mathbb{R}$, $r > 0$, let

$$\Delta_k^r(E) := \{ \mathbf{t} \in \Delta_k(E) : r\mathbf{t} \in \Delta_k(E) \} \subset \Delta_k(E).$$

We show that if $\dim_{\mathcal{H}}(E) > s_{k,d}$, for a natural measure ν_k on $\Delta_k(E)$, one has $\nu_k\left(\Delta_k^r(E)\right) > 0$ all $r \in \mathbb{R}_+$. Thus, in E there exist many pairs of $(k+1)$-point configurations which are similar by the scaling factor r. We extend this to show the existence of multi–similar configurations of any multiplicity. These results can be viewed as variants and extensions, for compact thin sets, of theorems of Furstenberg, Katznelson and Weiss [7], Bourgain [2] and Ziegler [11] for sets of positive density in $\mathbb{R}^d$.

Keywords Point configurations · Similarity classes · Hausdorff dimension

Revision of 15 April 2020. The first listed author was partially supported by NSF DMS-1362271 and -1906186. The second listed author was partially supported by NSA H98230-15-1-0319. The third listed author was partially supported by the Simons Foundation Collaboration Grant No. 422190.

✉ Allan Greenleaf
allan@math.rochester.edu

Alex Iosevich
iosevich@math.rochester.edu

Sevak Mkrtchyan
sevak.mkrtchyan@rochester.edu

[1] Department of Mathematics, University of Rochester, Rochester, NY 14627, USA

Mathematics Subject Classification 52C10 · 05D05

1 Introduction

Furstenberg, Katznelson and Weiss [7] proved that if $A \subset \mathbb{R}^2$ has positive upper Lebesgue density and A_δ denotes its δ-neighborhood, then, given vectors u, v in $\mathbb{R}^2$, there exists r_0 such that, for all $r > r_0$ and any $\delta > 0$, there exists $\{x, y, z\} \subset A_\delta$ forming a triangle similar to $\{\mathbf{0}, u, v\}$, i.e., congruent to $\{\mathbf{0}, ru, rv\}$ for some scaling factor $r > 0$. Under the same assumptions, Bourgain [2] proved in $\mathbb{R}^d$ that if $u^1, \dots, u^k \in \mathbb{R}^d$, $k \le d$, there exists r_0 such that, for all $r > r_0$ and any $\delta > 0$, there exists $\{x^1, x^2, \dots, x^{k+1}\} \subset A_\delta$ forming a simplex similar to $\{\mathbf{0}, u^1, \dots, u^k\}$ via scaling r. Perhaps the most general result of this approximate similarity type is due to Ziegler [11].

Bourgain also showed that if $k < d$ and the simplex is non-degenerate, i.e., of positive k-dimensional volume, then the conclusion in [2] holds for exact similarities ($\delta = 0$): for r sufficiently large, there exists $\{x^1, x^2, \dots, x^{k+1}\} \subset A$ similar to $\{\mathbf{0}, u^1, \dots, u^k\}$ via scaling r. The purpose of this paper is to prove variants of such exact similarity results for compact sets E of Hausdorff dimension $\dim_{\mathcal{H}}(E) < d$, sometimes referred to as *thin sets*. The statements necessarily need to be different in this context than for sets of positive upper Lebesgue density since, e.g., in a compact E the distances between points are bounded above by $\mathrm{diam}(E)$. More fundamental restrictions are known: e.g., there exist compact subsets of $\mathbb{R}^2$ of full Hausdorff dimension that do not contain vertices of any equilateral triangle (Falconer [6]). Nevertheless, we are able to prove that if $\dim_{\mathcal{H}}(E)$ is above a threshold, then given *any* $r > 0$ there exist many pairs (in fact, a set of positive measure in an appropriate sense) of $(k+1)$-point configurations which are similar via the scaling factor r and whose vertices are, as in [2], in E itself and not just in a δ-neighborhood of E. We also treat what we call *multi-similarities*: multiple configurations in E which are jointly similar to each other via multiple scaling factors.

We now turn to the results of this paper; to state them, some definitions are needed.

Definition 1 Let $d \ge 2$, $E \subset \mathbb{R}^d$ be a compact set, and μ a Frostman measure on E. For $1 \le k \le d$, denote points of $\mathbb{R}^{\binom{k+1}{2}} = \mathbb{R}^{k(k+1)/2}$ by $\mathbf{t} := (t^{ij})$, and for $x^1, \dots, x^{k+1} \in \mathbb{R}^d$, let $\mathbf{v}_{k,d}(x^1, \dots, x^{k+1}) \in \mathbb{R}^{k(k+1)/2}$ be the vector with entries $|x^i - x^j|$, $1 \le i < j \le k+1$, listed in the lexicographic order.

(i) The *k-simplex set* or *$(k+1)$-point configuration set* of E is

$$\Delta_k(E) := \left\{ \mathbf{v}_{k,d}(x^1, \dots, x^{k+1}) : x^j \in E \right\} \subset \mathbb{R}^{k(k+1)/2}. \tag{1.1}$$

Note that the k-simplex will necessarily be degenerate if $k > d$.

(ii) Define a measure ν_k on $\mathbb{R}^{k(k+1)/2}$, induced by a Frostman measure [1] μ on E, by the relation, for $f(\mathbf{t}) \in C_0\left(\mathbb{R}^{k(k+1)/2}\right)$,

$$\int_{\mathbb{R}^{k(k+1)/2}} f(\mathbf{t})\, d\nu_k(\mathbf{t}) := \int \dots \int f\left(\mathbf{v}_{k,d}\left(x^1, \dots, x^{k+1}\right)\right) d\mu(x^1) \cdots d\mu(x^{k+1}). \tag{1.2}$$

Remark 1 For $k \le d$, $\Delta_k(E)$ can be considered, modulo the symmetric group S_{k+1} acting on the x^i, as the set of congruence classes of $(k+1)$-point configurations in E, or equivalently

[1] Recall that a compact $E \subset \mathbb{R}^d$ always supports a Frostman measure; see [10] for the definition of Frostman measures and their basic properties.

the set of (possibly degenerate) k-simplexes in $\mathbb{R}^d$ spanned by points of E. The measure ν_k is supported on $\Delta_k(E)$ and has total mass at most $\mu(E)^{k+1}$. The action of the finite group S_{k+1} will be irrelevant for our results, which are expressed in terms of certain sets of configurations having positive $k(k+1)/2$-dimensional Lebesgue measure. (The situation when $k > d$ is discussed below.)

The study of the Lebesgue measure of the distance set $\Delta_1(E)$ for thin sets was begun in 1986 by Falconer [5] and has led to many important results, comprising too large a literature to summarize here. Most directly relevant, the best results currently known for the positivity of the measure of $\Delta_k(E)$, $1 < k \le d$, are due to Erdoğan, Hart and Iosevich [4] and Greenleaf, Iosevich, Liu and Palsson [8]. The former proved that the Lebesgue measure $\mathcal{L}^{k(k+1)/2}(\Delta_k(E)) > 0$ if $\dim_{\mathcal{H}}(E) > \frac{d+k+1}{2}$, while the latter obtained the threshold $\frac{dk+1}{k+1}$, improved to $\frac{8}{5}$ for $k = d = 2$. We note that all of these results, except for [9], are proven by establishing that the measure ν_k defined by (1.2) has a density in $L^2(\mathbb{R}^{k(k+1)/2})$.

Definition 2 Let $d \ge 2,\ k \ge 1$. The L^2-*threshold* for the k-simplex problem (or $(k+1)$-point configuration problem) in $\mathbb{R}^d$ is

$$s_{k,d} := \inf \left\{ s : \dim_{\mathcal{H}}(E) > s \implies \nu_k \text{ is absolutely continuous and } \int_{\Delta_k(E)} \nu_k^2(\mathbf{t})\, d\mathbf{t} < \infty \right\},$$

where E runs over all compact sets $E \subset \mathbb{R}^d$.

Thus, $s_{k,d}$ is less than or equal to the values mentioned in the paragraph above.

The case of $k > d$ needs to be treated somewhat differently, since in that range the set $\Delta_k(E) \subset \mathbb{R}^{k(k+1)/2}$ has lower dimension than $\mathbb{R}^{k(k+1)/2}$ and so cannot have positive Lebesgue measure, regardless of the Hausdorff dimension of E. This stems from the fact that, when $k > d$, specifying the $k(k+1)/2$ pairwise distances between $k+1$ points in $\mathbb{R}^d$ gives an over-determined system: knowing only some of the distances determines the rest. Thus, although $\Delta_k(E)$ still makes sense, the setup has to be modified. In Chatzikonstantinou, Iosevich, Mkrtchyan and Pakianathan [3] it was shown that for $k > d$ the set of congruence classes of $(k+1)$-tuples of elements of E can be naturally viewed as a subset of $\mathbb{R}^{d(k+1)-\binom{d+1}{2}}$; if $m := d(k+1) - \binom{d+1}{2}$ appropriately chosen distances are specified, then the other distances are determined, up to finitely many possibilities. Let P be such a collection of m edges. In the terminology of [3], P is a maximally independent (in $\mathbb{R}^d$) subset of the edges of the complete graph on $k+1$ vertices. Extend the definition of $\mathbf{v}_{k,d}$ to the case $k > d$ by setting $\mathbf{v}_{k,d}(x^1, \dots, x^{k+1}) = \left(|x^i - x^j|\right)_{(i,j)\in P} \in \mathbb{R}^m$, where the entries in the range are ordered lexicographically. Using this, we can define a measure ν_k on $\mathbb{R}^m$ and a set $\Delta_k(E) \subset \mathbb{R}^m$ similarly to (1.2) in Definition 1. Note that ν_k and $\Delta_k(E)$ will depend on the choice of P, but for our purposes this is irrelevant, so we will fix a particular P once and for all.

While $\mathbf{v}_{k,d}(x^1, \dots, x^{k+1})$ doesn't determine the congruence class of $(x^1, \dots, x^{k+1})$ uniquely, it identifies it up to a finite number of possibilities. The number of these possibilities is bounded above by a constant $u_{d,k}$, depending only on d and k. In this sense, congruence classes of $(k+1)$-tuples of elements of a compact set E in $\mathbb{R}^d$ for $k > d \ge 2$ can be naturally viewed as a subset of $\mathbb{R}^m$. It was shown in [3] that if $k > d$ and the Hausdorff dimension of E is greater than $d - \frac{1}{k+1}$, then the m-dimensional Lebesgue measure of the set of congruence classes of $(k+1)$-point configurations with endpoints in E is positive, and, as with most of the results for $k \le d$, this was shown by first establishing that the measure ν_k defined by (1.2) has a density in $L^2(\mathbb{R}^m)$.

With the theorems of [2,7,11] in mind, obtaining more refined structural information about $\Delta_k(E)$ is of natural interest. Since E is compact in our setting, the questions need to reflect the fact, that all the pairwise distances are bounded. We will show that, if $\dim_{\mathcal{H}}(E) > s_{k,d}$, then among the k-simplexes of E, all possible similarity scaling factors, $0 < r < \infty$, occur, and each one does so with positive ν_k-measure. Furthermore, we show that what we call *multi-similarities* of arbitrarily high multiplicity occur as well. Thus, this holds for the values of $\dim_{\mathcal{H}}(E)$ in the settings of all of the previous results referred to above where $\mathcal{L}^{k(k+1)/2}(\Delta_k(E)) > 0$ has been obtained, with the possible exception of [9].

To make this more precise, for $r \in \mathbb{R}_+ := (0, \infty)$ let

$$\Delta_k^r(E) := \{\mathbf{t} \in \Delta_k(E) : r\mathbf{t} \in \Delta_k(E)\} \subset \Delta_k(E), \tag{1.3}$$

the set of all k-simplexes $\mathbf{t}$ in E for which there is also a simplex in E similar to $\mathbf{t}$ via the scaling factor r. Interchanging the roles of the two simplexes in such a pair, $\{\mathbf{t}, r\mathbf{t}\} \subset \Delta_k(E)$, note for later use that

$$\Delta_k^{1/r}(E) = \Delta_k^r(E). \tag{1.4}$$

One can not only look for similar pairs $\{\mathbf{t}, r\mathbf{t}\} \subset \Delta_k(E)$, but more generally for similarities of higher multiplicity.

Definition 3 A collection $\{\mathbf{t}, r_1\mathbf{t}, \ldots, r_{n-1}\mathbf{t}\} \subset \Delta_k(E)$, with $\{1, r_1, \ldots, r_{n-1}\}$ pairwise distinct, is an *n-similarity of k-simplexes in E*, also referred to as a *multi-similarity of multiplicity n*.

Our main results are the following. All are obtained under the assumptions that $d \geq 2$, $k \geq 1$, $E \subset \mathbb{R}^d$ is compact, μ is a Frostman measure on E and ν_k is the measure induced by μ as in Definition 1. The following three results establish (quantitatively) the existence in E of many multi-similarities of multiplicities 2, 3 and $n \geq 3$, resp.

Theorem 1 *Let $d \geq 2$, $E \subset \mathbb{R}^d$ compact and $k \geq 1$. Suppose that* $\dim_{\mathcal{H}}(E) > s_{k,d}$, *the L^2-threshold for the k-simplex problem from Def. 2. Then there is a lower bound, uniform in r, $0 < r < \infty$,*

$$\nu_k(\Delta_k^r(E)) \geq C(k, E) > 0.$$

With the same notation and assumptions as in Theorem 1, we also have:

Theorem 2 *Suppose that* $\dim_{\mathcal{H}}(E) > s_{k,d}$. *Then there exist distinct $r_1, r_2 > 0$, with*

$$\nu_k\left(\Delta_k^{r_1}(E) \cap \Delta_k^{r_2}(E)\right) > 0.$$

In fact, for any partition $\mathbb{R}_+ = \coprod_{\alpha \in A} R_\alpha$ with each $R_\alpha \neq \emptyset$ and countable, there exist distinct $\alpha_1, \alpha_2 \in A$ and $r_1 \in R_{\alpha_1}$, $r_2 \in R_{\alpha_2}$, such that $\nu_k\left(\Delta_k^{r_1}(E) \cap \Delta_k^{r_2}(E)\right) > 0$.

Theorem 3 *Suppose that* $\dim_{\mathcal{H}}(E) > s_{k,d}$. *Then for all $n \in \mathbb{N}$, there exists an $M = M(n, k, E) \in \mathbb{N}$ such that for any distinct $r_1, \ldots, r_M \in \mathbb{R}_+$, there exist distinct $r_{i_1}, \ldots, r_{i_n}$ such that*

$$\nu_k\left(\Delta_k^{r_{i_1}}(E) \cap \Delta_k^{r_{i_2}}(E) \cap \cdots \cap \Delta_k^{r_{i_n}}(E)\right) > 0.$$

Remark 2 More explicitly, Theorem 1 says that, given any $r > 0$, there exist $(k+1)$-point configurations $\{x^1, x^2, \ldots, x^{k+1}\}$ and $\{y^1, y^2, \ldots, y^{k+1}\}$ in E which are similar via the scaling factor r, i.e., there exists a translation $\tau \in \mathbb{R}^d$ and a rotation $\theta \in O_d(\mathbb{R})$ such that $y^j = r\theta(x^j + \tau)$, $1 \leq j \leq k+1$. Thus, among the $(k+1)$-point configurations or k-simplexes in E, in the terminology of Definition 3, there exist (many) 2-similarities.

Furthermore, Theorem 2 states that there exist 3-similarities in E, i.e., triples of $(k+1)$-point configurations, $\{x^j\}$, $\{y^j\}$, $\{z^j\}$ in E and scalings r_1, r_2 so that $y^j = r_1\theta_1(x^j + \tau_1)$, $z^j = r_2\theta_2(x^j + \tau_2)$ for appropriate $\theta_1, \theta_2 \in O_d(\mathbb{R})$ and $\tau_1, \tau_2 \in \mathbb{R}^d$, and furthermore that r_1, r_2 can be arranged to lie in different subsets of a partition of $\mathbb{R}_+$ as stated. For example, decomposing $\mathbb{R}_+$ into the multiplicative cosets of $\mathbb{Q}_+$, there exist similarities of multiplicity 3 with r_2/r_1 irrational; similarly, replacing the positive rationals with the positive algebraic numbers, there exist such with r_2/r_1 transcendental.

Finally, Theorem 3 shows that there exist multi-similar $(k+1)$-point configurations in E of arbitrarily high multiplicity, and that the scaling factors can be chosen to come from an arbitrary set of distinct elements of $\mathbb{R}_+$, as long as that set has large enough cardinality relative to the desired similarity multiplicity.

Remark 3 We note that the conclusions of Remark 2 hold when $k > d$ as well. Denoting $\mathbf{x} := (x^1, \ldots, x^{k+1})$, the fact that $\mathbf{v}_{k,d}(\mathbf{x}) = \mathbf{v}_{k,d}(\mathbf{y})$ does not imply that $\mathbf{x}$ and $\mathbf{y}$ are congruent and hence $\mathbf{v}_{k,d}(\mathbf{x}) = r\mathbf{v}_{k,d}(\mathbf{y})$ does not imply $\mathbf{x}$ and $\mathbf{y}$ are similar. However, the conclusions of Remark 2 still hold as follows. Recall from the introduction that by results of Chatzikonstantinou, Iosevich, Mkrtchyan and Pakianathan [3], $\mathbf{v}_{k,d}(\mathbf{x})$ determines the congruence type of $\mathbf{x}$ up to at most a bounded number $u_{d,k}$ of choices. Using Theorem 3 with $n \cdot u_{d,k}$ instead of n we see that there exist $\mathbf{x}, \mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_{nu_{d,k}}}$ and $r_{i_1}, \ldots, r_{i_{nu_{d,k}}}$ such that $\mathbf{v}_{k,d}(\mathbf{x}), r_{i_1}\mathbf{v}_{k,d}(\mathbf{x}_1), \ldots, r_{i_{nu_{d,k}}}\mathbf{v}_{k,d}(\mathbf{x}_{nu_{d,k}})$ are all congruent. It follows that $\mathbf{x}, \mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_{nu_{d,k}}}$ all fall within at most $u_{d,k}$ congruence classes; thus, by the pigeon hole principle, at least $n+1$ of them must be in some congruence class. This argument applies to the conclusions of Remark 2 for the other Theorems as well.

2 Proofs of Theorems 2 and 3

We start by showing that Theorems 2 and 3 follow from Theorem 1 by measure-theoretic arguments. To prove Theorem 2, let $\mathbb{R}_+ = \coprod_{\alpha \in A} R_\alpha$ be a partition of $\mathbb{R}_+$ into a (necessarily uncountable) collection of nonempty countable subsets. From the definition (1.3), it follows that each $\Delta_k^r(E)$ is ν_k-measurable. Hence, if for each $\alpha \in A$, with slight abuse of notation we define the set

$$\Delta_k^\alpha(E) := \bigcup_{r \in R_\alpha} \Delta_k^r(E),$$

then, being a countable union of measurable sets, each $\Delta_k^\alpha(E)$ is ν_k-measurable. Furthermore, combining $R_\alpha \neq \emptyset$, the monotonicity of ν_k and Theorem 1, one sees that each $\nu_k(\Delta_k^\alpha(E)) > 0$. However, $\nu_k(\Delta_k(E)) \leq \mu(E)^{k+1} < \infty$, and no finite (or even σ-finite) measure space can be the pairwise disjoint union of an uncountable collection of measurable subsets of positive measure. Thus, there must exist $\alpha_1 \neq \alpha_2$ such that $\Delta_k^{\alpha_1}(E) \cap \Delta_k^{\alpha_2}(E) \neq \emptyset$; it follows that there are $r_j \in R_{\alpha_j}$, $j = 1, 2$, such that $\Delta_k^{r_1}(E) \cap \Delta_k^{r_2}(E) \neq \emptyset$. For the full claim of Theorem 2, that there exist distinct $\alpha_1, \alpha_2 \in A$ and $r_1 \in R_{\alpha_1}$, $r_2 \in R_{\alpha_2}$, such that $\nu_k\left(\Delta_k^{r_1}(E) \cap \Delta_k^{r_2}(E)\right) > 0$, first make a choice of one representative from each of the R_α, then choose an arbitrary countably infinite subset of these, and finally apply Theorem 3.

For the proof of Theorem 3, we use the uniform lower bound from Theorem 1, $\nu_k(\Delta_k^r(E)) \geq C(E,k) > 0$, $\forall r \in \mathbb{R}_+$, combined with $\nu_k(\Delta_k(E)) < \infty$. Theorem 3 then follows from the following measure-theoretic pigeon-hole principle, which might be of independent interest and whose proof is deferred to the Appendix, Sect. 1.

Lemma 1 *Let $\mathcal{X} = (X, \mathcal{M}, \sigma)$ be a finite measure space. For $0 < c < \sigma(X)$, let $\mathcal{M}_c = \{A \in \mathcal{M} : \sigma(A) \geq c\}$. Then, for every $n \in \mathbb{N}$, there exists an $N = N(\mathcal{X}, c, n) \in \mathbb{N}$ such that for any collection $\{A_1, \ldots, A_N\} \subset \mathcal{M}_c$ of cardinality N, there is a subcollection $\{A_{i_1}, \ldots, A_{i_n}\}$ of cardinality n such that $\sigma(A_{i_1} \cap \cdots \cap A_{i_n}) > 0$ and hence $A_{i_1} \cap \cdots \cap A_{i_n} \neq \emptyset$.*

3 Proof of Theorem 1

To keep the exposition simple, we first prove Theorem 1 in the case $k \leq d$. In Sects. 3 and 4 we will assume $k \leq d$. In the case $k > d$ the arguments are very similar. Since $\mathbf{v}_{k,d}(x^1, \ldots, x^{k+1})$ determines the congruence type of $(x^1, \ldots, x^{k+1})$ up to at most $u_{d,k}$ choices, the constant $u_{k,d}$ will appear throughout the proof. However, since the results here are up to multiplicative constants, this doesn't play any essential role.

For $\epsilon > 0$, define a smooth approximation of ν_k on $\mathbb{R}^{k(k+1)/2}$ by the density

$$\nu_k^\epsilon(\mathbf{t}) = \int \cdots \int \prod_{1 \leq i < j \leq k+1} \sigma_{t_{ij}}^\epsilon(x^i - x^j) \prod_{l=1}^{k+1} d\mu(x^l), \tag{3.1}$$

where σ_t is the normalized surface measure on the sphere of radius t in $\mathbb{R}^d$ and $\sigma_t^\epsilon(x) := \sigma_t * \rho_\epsilon(x)$, with $\rho \in C_0^\infty(\mathbb{R}^d)$, $\rho \geq 0$, $\text{supp}(\rho) \subset \{|t| < 1\}$, $\int \rho = 1$ and $\rho_\epsilon(x) = \epsilon^{-d}\rho(\epsilon^{-1}x)$. Then each $\nu_k^\epsilon \in C_0^\infty$ and $\nu_k^\epsilon \to \nu_k$ weak* as $\epsilon \to 0$. Thus,

$$\nu_k(\Delta_k^r(E)) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^{k(k+1)/2}} \nu_k^\epsilon(r\mathbf{t})\, d\nu_k(\mathbf{t}).$$

By (1.2), for ϵ fixed,

$$\int_{\mathbb{R}^{k(k+1)/2}} \nu_k^\epsilon(r\mathbf{t})\, d\nu_k(\mathbf{t}) = \int \nu_k^\epsilon\left(r(x^1 - x^2), \ldots, r(x^k - x^{k+1})\right) d\mu(x^1) \ldots d\mu(x^{k+1}).$$

Using the definition in (3.1), we see that this is

$$\approx \epsilon^{-\binom{k+1}{2}} \int \cdots \int_{\{||x^i - x^j| - r|y^i - y^j|| < \epsilon;\, 1 \leq i < j \leq k+1\}} d\mu(x^1) \ldots d\mu(x^{k+1}) d\mu(y^1) \ldots d\mu(y^{k+1}), \tag{3.2}$$

which we denote by I_ϵ. Here, and throughout, we write $X \lesssim Y$ (resp. $X \approx Y$) if there exist constants $0 < c < C$, depending only on k, E and the choice of ρ (and thus implicitly on d), such that $X \lesssim CY$ (resp., $cY \leq X \leq CY$). Also, we denote the $2(k+1)$-fold product measure in (3.2) and similar occurences by $\mu^{2(k+1)}$.

For each rotation $\theta \in O_d(\mathbb{R})$, define a measure $\lambda_{r,\theta}$ on $\mathbb{R}^d$ by

$$\int f(z)\, d\lambda_{r,\theta}(z) = \int\int f(u - r\theta v)\, d\mu(u) d\mu(v), \;\; f \in C_0\left(\mathbb{R}^d\right).$$

This is the push-forward of $\mu \times \mu$ under the map $(u, v) \to u - r\theta v$, has total mass $||\lambda_{r,\theta}|| = \mu(E)^2$, and is supported in $E - r\theta E$. We show below that, if $\dim_{\mathcal{H}}(E) > s_{k,d}$, for a.e. θ, the measure $\lambda_{r,\theta}$ is absolutely continuous with a density in $L^{k+1}(\mathbb{R}^d)$, which we denote by $\lambda_{r,\theta}(\cdot)$. Let $d\theta$ denote the Haar probability measure on $O_d(\mathbb{R})$.

Proposition 1 *With the notation above,*

$$\liminf_{\epsilon\to 0} I_\epsilon \approx \int\int \left(\lambda_{r,\theta}(z)\right)^{k+1} dz\, d\theta. \tag{3.3}$$

By definition, the quantity on the right hand side of (3.3) is finite if $\dim_{\mathcal{H}}(E) > s_{k,d}$, the L^2-threshold for the k-simplex problem. Prop. 1 was proved in [8] in the case $r = 1$; the proof in the general case is similar, but we supply it in the next section for the sake of completeness.

Continuing with the proof of Theorem 1, by Hölder we have

$$\begin{aligned} \mu(E)^2 &= \int\int \lambda_{r,\theta}(z)\cdot 1\, dz\, d\theta \\ &\leq \left(\iint \left(\lambda_{r,\theta}(z)\right)^{k+1} dz\, d\theta\right)^{\frac{1}{k+1}} \times \left(\iint_{\mathrm{supp}(\lambda_{r,\theta})\times O_d(\mathbb{R})} 1^{\frac{k+1}{k}}\, dz\, d\theta\right)^{\frac{k}{k+1}}. \end{aligned} \tag{3.4}$$

Since $\mathrm{supp}(\lambda_{r,\theta})$, being contained in $E - r\theta E$, has Lebesgue measure $\lesssim (1 + r^d)\mu(E)$, we divide both sides of (3.4) by the second factor on the right hand side and raise to the $k+1$ power to obtain

$$\mu(E)^{k+1}(1+r^d)^{-(k+1)} \lesssim \iint \left(\lambda_{r,\theta}(z)\right)^{k+1} dz\, d\theta.$$

Combining this with Proposition 1, we conclude that, for $\dim_{\mathcal{H}}(E) > s_{k,d}$ and $0 < r \leq 1$,

$$\liminf_{\epsilon\to 0} \int \nu_k^\epsilon(r\mathbf{t})\, d\nu_k(\mathbf{t}) \approx \int\int \left(\lambda_{r,\theta}(z)\right)^{k+1} dz\, d\theta \gtrsim 1. \tag{3.5}$$

It follows that $\liminf_{\epsilon\to 0} \nu_k(\{\mathbf{t} : r\mathbf{t} \in \Delta_{k,\epsilon}(E)\}) \gtrsim 1$, where $\Delta_{k,\epsilon}(E)$ is the ϵ-neighborhood of $\Delta_k(E)$. Since the sets $\{\mathbf{t} : r\mathbf{t} \in \Delta_{k,\epsilon}(E)\}$ are nested as $\epsilon \searrow 0$, we conclude that, for $0 < r \leq 1$,

$$\nu_k(\{\mathbf{t} : r\mathbf{t} \in \Delta_k(E)\}) \gtrsim 1. \tag{3.6}$$

However, by (1.4), $\nu_k(\{\mathbf{t} : r\mathbf{t} \in \Delta_k(E)\}) = \nu_k\left(\left\{\mathbf{t} : r^{-1}\mathbf{t} \in \Delta_k(E)\right\}\right)$; therefore, (3.6) holds for $1 \leq r < \infty$ as well. This completes the proof of Theorem 1, up to the verification of Proposition 1.

4 Proof of Proposition 1

We will follow closely the argument in [8, Sect. 2]. It will be convenient to denote an ordered $(k+1)$-tuple $(x^1, \ldots, x^{k+1})$ of elements of $\mathbb{R}^d$ by $\mathbf{x}$. If the corresponding set $\{x^1, \ldots, x^{k+1}\}$ is a nondegenerate simplex (i.e., affinely independent), then

$$\pi(\mathbf{x}) := span\{x^2 - x^1, \ldots, x^{k+1} - x^1\}$$

is a k-dimensional linear subspace of $\mathbb{R}^d$. $\Delta(\mathbf{x})$ will denote the (unoriented) simplex generated by $\{x^1, \ldots, x^{k+1}\}$, i.e., the closed convex hull, which is contained in the affine plane $x^1 + \pi(\mathbf{x})$. Both $\pi(\mathbf{x})$ and $\Delta(\mathbf{x})$ are independent of the order of the x^j. If $\{x^1, \ldots, x^{k+1}\}$ is similar to $\{y^1, \ldots, y^{k+1}\}$ by a scaling factor r, then, up to permutation of $y^1, \ldots, y^{k+1}$, there exists a $\theta \in \mathbb{O}(d)$ such that $x^j - x^1 = r\theta(y^j - y^1)$, $2 \leq j \leq k+1$, which is equivalent with $x^j - x^i = r\theta(y^j - y^i)$, $1 \leq i < j \leq k+1$, and $\Delta(\mathbf{x}) = (x^1 - r\theta y^1) + r\theta\Delta(\mathbf{y})$.

The group $\mathbb{O}(d)$ acts on the Grassmanians $G(k, d)$ and $G(d-k, d)$ of k (resp., $d-k$) dimensional linear subspaces of $\mathbb{R}^d$, and if $\mathbf{x}$ is similar to $\mathbf{y}$, one has $\pi(\mathbf{x}) = \theta\pi(\mathbf{y})$ and $\pi(\mathbf{x})^\perp = \theta\big(\pi(\mathbf{y})^\perp\big)$. The set of $\theta \in \mathbb{O}(d)$ fixing $\pi(\mathbf{x})$ is a conjugate of $\mathbb{O}(d-k) \subset \mathbb{O}(d)$, and we refer to this as the *stabilizer* of $\mathbf{x}$, denoted $\mathrm{Stab}(\mathbf{x})$.

For $\mathbf{x},\ \mathbf{y}$ similar, let $\tilde{\theta} \in \mathbb{O}(d)$ be such that it transforms $\mathbf{y}$ to $\mathbf{x}$. I.e. we have $\pi(\mathbf{x}) = \tilde{\theta}\pi(\mathbf{y})$ and $x^i - x^j = r\tilde{\theta}\omega(y^i - y^j)$ for all $\omega \in \mathrm{Stab}(\mathbf{y})$. For each $\mathbf{y}$, take a cover of $\mathbb{O}(d)/\mathrm{Stab}(\mathbf{y})$ by balls of radius ϵ (with respect to some Riemannian metric) with finite overlap. Since the dimension of $\mathbb{O}(d)/\mathrm{Stab}(\mathbf{y})$ is that of $\mathbb{O}(d)/\mathbb{O}(d-k)$, namely

$$\frac{d(d-1)}{2} - \frac{(d-k)(d-k-1)}{2} = kd - \frac{k(k+1)}{2},$$

one needs $N(\epsilon) \lesssim \epsilon^{-\left(kd - \frac{k(k+1)}{2}\right)}$ balls to cover it. In these balls, choose sample points, $\tilde{\theta}_m(\mathbf{y}),\ 1 \le m \le N(\epsilon)$.

One sees that

$$\begin{aligned}&\Big\{(\mathbf{x},\mathbf{y}) : \Big||x^i - x^j| - r|y^i - y^j|\Big| \le \epsilon,\ 1 \le i < j \le k+1\Big\}\\ &\subseteq \bigcup_{m=1}^{N(\epsilon)} \Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - x^j) - r\tilde{\theta}_m(\mathbf{y})\omega(y^i - y^j)\Big| \lesssim \epsilon,\\ &\qquad\qquad \forall\, 1 \le i < j \le k+1,\ \omega \in \mathrm{Stab}(\mathbf{y})\Big\}.\end{aligned}$$

Thus, the right hand side of (3.2) is bounded above by

$$\begin{aligned}\epsilon^{-\frac{k(k+1)}{2}} \sum_{m=1}^{N(\epsilon)} \mu^{2(k+1)}\Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - x^j) - r\tilde{\theta}_m(\mathbf{y})\omega(y^i - y^j)\Big| \lesssim \epsilon,\\ \forall\, 1\ \le i < j \le k+1,\ \omega \in \mathrm{Stab}(\mathbf{y})\Big\}.\end{aligned} \tag{4.1}$$

When picking the $N(\epsilon)$ balls, if each point of $O(d)/Stab(y)$ is covered by at most $p = p(d)$ of the balls, then the quantity in (4.1) also becomes, when multiplied by $1/p$, a lower bound for (3.2). Thus, the right hand side of (3.2) is comparable to the quantity in (4.1), which can be rewritten as

$$\begin{aligned}\epsilon^{-kd} \sum_{m=1}^{N(\epsilon)} \epsilon^{kd - \frac{k(k+1)}{2}} \mu^{2(k+1)}\Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\tilde{\theta}_m(\mathbf{y})\omega y^i) - (x^j - r\tilde{\theta}_m(\mathbf{y})\omega y^j)\Big| \lesssim \epsilon,\\ \forall\, 1 \le i < j \le k+1, \omega \in \mathrm{Stab}(\mathbf{y})\Big\}.\end{aligned} \tag{4.2}$$

Since this holds for any choice of sample points $\tilde{\theta}_m(\mathbf{y})$, we can pick these points such that they minimize (up to a factor of 1/2, say) the quantity

$$\begin{aligned}\mu^{2(k+1)}\Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\tilde{\theta}_m(\mathbf{y})\omega y^i) - (x^j - r\tilde{\theta}_m(\mathbf{y})\omega y^j)\Big| \le \epsilon,\\ \forall\, 1 \le i < j \le k+1, \omega \in \mathrm{Stab}(\mathbf{y})\Big\}.\end{aligned}$$

The $N(\epsilon)$ preimages, under the natural projection from $\mathbb{O}(d)$, of the balls used to cover $\mathbb{O}(d)/\mathrm{Stab}(\mathbf{y})$ are ϵ-tubular neighborhoods of the preimages of the sample points $\tilde{\theta}_m(\mathbf{y})$, which we denote $T_1^\epsilon, \dots, T_{N(\epsilon)}^\epsilon$. Since $dim(\mathbb{O}(d)/\mathrm{Stab}(\mathbf{y})) = kd - \frac{k(k+1)}{2}$, each T_m^ϵ has measure $\sim \epsilon^{kd-\frac{k(k+1)}{2}}$ with respect to the Haar measure $d\theta$. Since the infimum over a set is less than or equal to the average over the set, it follows that

$$\mu^{2(k+1)}\Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\tilde{\theta}_m(\mathbf{y})\omega y^i) - (x^j - r\tilde{\theta}_m(\mathbf{y})\omega y^j)\Big| \leq \epsilon,$$
$$\forall\, 1 \leq i < j \leq k+1, \omega \in \mathrm{Stab}(\mathbf{y})\Big\}$$
$$\approx \frac{1}{\epsilon^{kd-\frac{k(k+1)}{2}}} \int_{T_m^\epsilon} \mu^{2(k+1)} \Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\theta y^i) - (x^j - r\theta y^j)\Big| \leq \epsilon,\ 1 \leq i < j \leq k+1\Big\}\, d\theta.$$

The quantity in (4.2) is thus

$$\approx \epsilon^{-kd} \sum_{m=1}^{N(\epsilon)} \int_{T_m^\epsilon} \mu^{2(k+1)} \Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\theta y^i) - (x^j - r\theta y^j)\Big| \leq \epsilon,\ 1 \leq i < j \leq k+1\Big\}\, d\theta,$$

which, since the collection $\{T_m^\epsilon\}$ have pointwise finite overlap (uniformly in ϵ), is

$$\approx \epsilon^{-kd} \int \mu^{2(k+1)} \Big\{(\mathbf{x},\mathbf{y}) : \Big|(x^i - r\theta y^i) - (x^j - r\theta y^j)\Big| \leq \epsilon,\ 1 \leq i < j \leq k+1\Big\}\, d\theta,$$

and taking the lim inf, we obtain a quantity comparable to the expression (3.3). This completes the proof of Proposition 1, and thus Theorems 1, 2, and 3.

5 Open question

The following is a natural question pertaining to the subject matter of Theorem 1:

In [1] it was shown that if E is a compact subset of $\mathbb{R}^d$, of Hausdorff dimension greater than $\frac{d+1}{2}$, then there exists a non-empty open interval I such that, for any $t \in I$, there exist $x^1, x^2, \ldots, x^{k+1} \in E$ such that $|x^{j+1} - x^j| = t$, $1 \leq j \leq k$. In view of Theorem 1, it seems reasonable to ask: given any $r > 0$, do there exist $x, y, z \in E$ such that $|x - z| = r|x - y|$? This can be regarded as a pinned version of the case $k = 1$ of Theorem 1, in the sense that the endpoint x is common to both segments whose length is being compared. Similar questions can be raised when $k > 1$.

Appendix: A measure-theoretic pigeon hole principle

Unable to find Lemma 1 in the literature, and believing that it should be useful for other problems, we prove it here. Without loss of generality the total measure $\sigma(X)$ can be normalized to be equal to 1, so for the proof we restate the result as

Lemma 2 *Let $\mathcal{X} = (X, \mathcal{M}, \sigma)$ be a probability space. For $0 < c < 1$, let $\mathcal{M}_c = \{A \in \mathcal{M} : \sigma(A) \geq c\}$. Then, for every $n \in \mathbb{N}$, there exists an $N = N(\mathcal{X}, c, n) \in \mathbb{N}$ such that, for any collection $\{A_1, \ldots, A_N\} \subset \mathcal{M}_c$ of cardinality at least N, there is a subcollection $\{A_{i_1}, \ldots, A_{i_n}\}$ of cardinality n such that $\sigma(A_{i_1} \cap \cdots \cap A_{i_n}) > 0$ and hence $A_{i_1} \cap \cdots \cap A_{i_n} \neq \emptyset$.*

To start the proof, first we establish the following claim, which is a quantitative strengthening of the statement for $n = 2$:

Claim 1 Let $\mathcal{X} = (X, \mathcal{M}, \sigma)$ be a probability space. Then for any $0 < c < 1$ there exists $P_c \in \mathbb{N}$ such that for any $N > P_c$, if $\{A_1, \ldots, A_N\} \subset \mathcal{M}_c$, then there exist distinct $i, j \le N$ such that $\sigma(A_i \cap A_j) \ge c^3/3$.

Proof Suppose not. Let $S \subset (0, 1)$ be the set of all $c \in (0, 1)$ such that the statement of the claim is false, and suppose $c \in S$. Then for every $N \in \mathbb{N}$ there exists a subset $\{A_1, \ldots, A_{2N}\} \subset \mathcal{M}_c$ such that $\sigma(A_i) \ge c$ for all i but $\sigma(A_i \cap A_j) < c^3/3$ for all $i \ne j$. Consider the sets $A_{2i-1} \cup A_{2i}$, $i = 1, \ldots, N$. We have

$$\sigma(A_{2i-1} \cup A_{2i}) = \sigma(A_{2i-1}) + \sigma(A_{2i}) - \sigma(A_{2i-1} \cap A_{2i}) > c + c - \frac{c^3}{3} = 2c - \frac{c^3}{3}.$$

Since $\sigma(X) = 1 \ge \sigma(A_{2i-1} \cup A_{2i})$, this implies $1 > 2c - \frac{c^3}{3}$. In particular, since $c < 1$, we have $c \lesssim 0.52 < 3/5$; hence $[3/5, 1) \cap S = \emptyset$. Moreover,

$$\begin{aligned}
\sigma\big((A_{2i-1} \cup A_{2i}) \cap (A_{2j-1} \cup A_{2j})\big) \\
&= \sigma\big((A_{2i-1} \cap A_{2j}) \cup (A_{2i-1} \cap A_{2j-1}) \cup (A_{2i} \cap A_{2j}) \cup (A_{2i} \cap A_{2j-1})\big) \\
&\le \sigma(A_{2i-1} \cap A_{2j}) + \sigma(A_{2i-1} \cap A_{2j-1}) + \sigma(A_{2i} \cap A_{2j}) + \sigma(A_{2i} \cap A_{2j-1}) \\
&< 4\frac{c^3}{3} \le \frac{(2c - c^3/3)^3}{3} \quad \text{since } 0 < c < 1.
\end{aligned}$$

Thus, there exist N sets, namely $A_1 \cup A_2, \ldots, A_{2N-1} \cup A_{2N}$, such that each has measure at least $f(c) := 2c - \frac{c^3}{3}$ but all pairwise intersections have measure less than $\frac{f(c)^3}{3}$.

Thus, we have shown that if $c \in S$ then $f(c) \in S$ as well. However if $0 < c < 1$, then there exists $k \in \mathbb{N}$ such that $f^k(c) > 3/5$ and is thus $\notin S$ (where f^k denotes f composed with itself k times). It follows that S must be empty. □

We use Claim 1 as a building block for the proof of Lemma 2, which is by induction on n. If $n = 1$, then we can take $N = 1$, since any $A_{i_1} \in \mathcal{M}_c$ satisfies the statement. If $n = 2$ then any $N \ge \lceil 1/c \rceil$ suffices, since there cannot be more than $1/c$ pairwise disjoint sets of measure $\ge c$ each; alternatively, one may simply invoke Claim 1.

Now suppose that the conclusion of Lemma 2 holds for some n, $n \ge 2$. Set $N = 2N(\mathcal{X}, c^3/3, n) + P_c$, and suppose $\{A_1, \ldots, A_N\} \subset \mathcal{M}_c$ is a collection of cardinality N. Since $N > P_c$, by Claim 1 there exist distinct $i, j \le N$ such that $\sigma(A_i \cap A_j) > \frac{c^3}{3}$. Let $B_1 = A_i \cap A_j$. Removing A_i and A_j from the collection we still have $N - 2 > P_c$ sets, so can find another pair whose intersection has measure at least $\frac{c^3}{3}$. Repeating this procedure $N(\mathcal{X}, c^3/3, n)$ times, one finds sets $B_1, \ldots, B_{N(\mathcal{X}, c^3/3, n)} \in \mathcal{M}_{c^3/3}$. By the induction hypothesis there exist $0 < i_1 < i_2 < \cdots < i_n \le N(\mathcal{X}, c^3/3, n)$ such that $\sigma(B_{i_1} \cap \cdots \cap B_{i_n}) > 0$. Since $B_{i_1} \cap \cdots \cap B_{i_n}$ is the intersection of $2n$ distinct sets from the collection $\{A_1, \ldots, A_N\}$, the intersection of any $n+1$ of those $2n$ will have positive measure, completing the induction step. □

References

1. Bennett, M., Iosevich, A., Taylor, K.: Finite chains inside thin subsets of $\mathbb{R}^d$. Anal. PDE **9**(3), 597–614 (2016)
2. Bourgain, J.: A Szemeredi type theorem for sets of positive density. Israel J. Math. **54**(3), 307–331 (1986)
3. Chatzikonstantinou, N. Iosevich, A., Mkrtchyan, S., Pakianathan, J. *Rigidity, graphs and Hausdorff dimension*, https://arxiv.org/pdf/1708.05919.pdf (2017)

4. Erdoğan, B., Hart D., Iosevich, A. *Multiparameter projection theorems with applications to sums-products and finite point configurations in the Euclidean setting*, Recent advances in harmonic analysis and applications, 93-103, Springer Proc. Math. Stat., 25, Springer, New York, (2013)
5. Falconer, K.J.: On the Hausdorff dimensions of distance sets. Mathematika **32**, 206–212 (1986)
6. Falconer, K. J., *Some problems in measure combinatorial geometry associated with Paul Erdős* (2013). http://www.renyi.hu/conferences/erdos100/slides/falconer.pdf. Accessed 31 May 2020
7. Furstenberg, H., Katznelson, Y., Weiss, B.: Ergodic Theory and Configurations in Sets of Positive Density Mathematics of Ramsey Theory, Algorithms Combin, vol. 5, pp. 184–198. Springer, Berlin (1990)
8. Greenleaf, A., Iosevich, A., Liu, B., Palsson, E.: A group-theoretic viewpoint on Erdős–Falconer problems and the Mattila integral. Revista Mat. Iberamericana **31**(3), 799–810 (2015)
9. Guth, L., Iosevich, A., Ou, Y., Wang, H.: On Falconer distance set problem in the plane. Invent. Math. **219**(3), 779–830 (2020)
10. Mattila, P. Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability. Cambridge Stud .Adv. Math. **44**,112–114 (1995)
11. Ziegler, T.: Nilfactors of $\mathbb{R}^d$ actions and configurations in sets of positive upper density in $\mathbb{R}^m$. J. Anal. Math. **99**, 249–266 (2006)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.