Geometric and o-minimal Littlewood–Offord problems

Jacob Fox[*]   Matthew Kwan[†]   Hunter Spink[‡]

Abstract

The classical Erdős–Littlewood–Offord theorem says that for nonzero vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, any $x \in \mathbb{R}^d$, and uniformly random $(\xi_1, \ldots, \xi_n) \in \{-1, 1\}^n$, we have $\Pr(a_1\xi_1 + \cdots + a_n\xi_n = x) = O(n^{-1/2})$. In this paper we show that $\Pr(a_1\xi_1 + \cdots + a_n\xi_n \in S) \le n^{-1/2+o(1)}$ whenever S is definable with respect to an o-minimal structure (for example, this holds when S is any algebraic hypersurface), under the necessary condition that it does not contain a line segment. We also obtain an inverse theorem in this setting.

1 Introduction

Consider a random variable $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $a_1, \ldots, a_n \in \mathbb{R}$ are real numbers and $\xi_1, \ldots, \xi_n \in \{-1, 1\}$ are i.i.d Rademacher random variables (meaning that each ξ_i is independent and identically distributed with $\Pr(\xi_i = 1) = \Pr(\xi_i = -1) = 1/2$). Broadly speaking, the classical Littlewood–Offord problem asks for *anti-concentration* estimates for X: what can we say about the maximum probability that X is equal to a single value, or falls in an interval of prescribed length? We refer the reader to [27] for a thorough survey of the Littlewood–Offord problem and its applications, beyond the brief (and selective) history in what follows.

In connection with their work on random polynomials, Littlewood and Offord [22] proved that if each $|a_i| \ge 1$, then $\max_{x \in \mathbb{R}} \Pr(|X - z| \le 1) = O(\log n/\sqrt{n})$. Erdős [12] later obtained the optimal $O(1/\sqrt{n})$ bound, in what is now known as the Erdős–Littlewood–Offord theorem (to see that the $O(1/\sqrt{n})$ bound cannot be improved, consider the case where each $a_i = 1$). Note that if one is concerned about the probability of taking a particular value instead of falling in an interval, a simple rescaling argument shows that with only the assumption that each $a_i \ne 0$ we have $\max_{x \in \mathbb{R}} \Pr(X = x) = O(1/\sqrt{n})$.

The Littlewood–Offord problem naturally generalises to higher dimensions. For any $d \in \mathbb{N}$, consider $a_1, \ldots, a_n \in \mathbb{R}^d$, and again let $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Kleitman [21] proved that if each $\|a_i\| = \sqrt{a_i \cdot a_i} \ge 1$, then $\max_{x \in \mathbb{R}^d} \Pr(\|X - z\| \le R) \le C_{R,d}/\sqrt{n}$ for some $C_{R,d}$ depending only on R, d, and Frankl and Füredi [15] later obtained asymptotically optimal bounds on $C_{R,d}$ (see [27, Section 2] for more about the history of this problem). The point-concentration version of the Littlewood–Offord problem generalises much more easily to higher dimensions: if each $a_i \ne 0$, we can project to a generic 1-dimensional subspace and apply the Erdős–Littlewood–Offord theorem to see that the optimal point concentration bound $\max_{x \in \mathbb{R}^d} \Pr(X = x) = O(1/\sqrt{n})$ still holds.

In this paper we consider a different generalisation of the Littlewood–Offord problem to $\mathbb{R}^d$.

Question 1. *Assuming only that each $a_i \ne 0$, what geometric constraints can be imposed on a set $S \subset \mathbb{R}^d$ which imply non-trivial uniform upper bounds on $\Pr(X \in S)$?*

Since we do not make any assumption about the size of the a_i (other than that they are nonzero), this question is much more about the geometry of S than the spatial anti-concentration of X, and therefore has a very different flavour to previous Littlewood-Offord variants. For example, when S is a smooth

---

manifold, any answer to this question would have to be applicable both at a local scale (where S resembles an affine subspace), and at a global scale (where the geometry of S becomes more visible).

It seems that questions of this type have not been systematically studied before. However the case where $S \subset \mathbb{R}^2$ is a parabola was considered by Costello in his study of the *quadratic Littlewood–Offord problem*, where he showed that $\Pr(X \in S) \le n^{-1/2+o(1)}$ by observing a connection to incidence geometry and applying a weighted version of the Szemerédi–Trotter theorem [6, Lemma 13].

If S contains a line segment (between $x, y \in \mathbb{R}^d$, say), then taking $a_1 = \frac{x+y}{2}, a_2 = \cdots = a_n = \frac{x-y}{2n}$ yields $\Pr(X \in S) \ge \frac{1}{2}$, so no non-trivial uniform upper bound on $\Pr(X \in S)$ can be established. Highly "oscillatory" behaviour is also a problem for us: for example, if $a_1 = \cdots = a_n = (1, 0) \in \mathbb{R}^2$, then X always lies on the graph of the function $x \mapsto \sin(2\pi x)$.

A very general property that precludes this kind of oscillatory behaviour is the property of being *definable with respect to an o-minimal structure*. We postpone the formal definition to Section 1.4, but for now we note that this property holds for any set that can be expressed via a Boolean combination of equalities and inequalities involving compositions of polynomials, the exponential function, and restrictions of real analytic functions to compact boxes (sets obtainable in this way are said to be *explicitly definable*, see for example the survey of Scanlon [38]). For example, any affine variety, and more generally any semi-algebraic set, is definable with respect to an o-minimal structure.

If S is definable with respect to an o-minimal structure, and does not contain any line segment, we prove that $\Pr(X \in S) \le n^{-1/2+o(1)}$, nearly matching Erdős' $O(1/\sqrt{n})$ point concentration bound for the usual Littlewood–Offord problem.

Theorem 1.1. *Let $S \subseteq \mathbb{R}^d$ be a set which is definable with respect to an o-minimal structure and does not contain any line segment, and let $\alpha > 0$. If n is sufficiently large in terms of α, S, then the following holds. Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Then*

$$\Pr(X \in S) \le n^{-1/2+\alpha}.$$

In fact, n only has to be sufficiently large in terms of α, d, and the "complexity" of S (a notion for definable sets we will introduce in Definition 1.12). Every semi-algebraic set (i.e., every Boolean combination of sets in $\mathbb{R}^d$ defined by polynomial equalities and inequalities) is definable with respect to every o-minimal structure, and in this case being of bounded complexity amounts to a bound on the degrees of the polynomials and the number of regions used in the Boolean combination. So, for example, if n is sufficiently large in terms of α, d we have

$$\Pr(\|X - z\| = R) \le n^{-1/2+\alpha},$$

because for fixed d, the spheres of the form $\{x \in \mathbb{R}^d : \|x - z\|^2 = R^2\}$, for $(R, z) \in \mathbb{R}_{>0} \times \mathbb{R}^d$, have bounded complexity.

We remark that the assumption that each ξ_i has a Rademacher distribution (or even that the ξ_i have the same distribution) is actually not essential. Indeed, using standard conditioning arguments (see [24, Section 4]) one can deduce from Theorem 1.1 a stronger result that holds whenever X is a sum of independent random variables which each have reasonably non-degenerate distributions.

We also remark that another restriction that rules out the possibility of "highly oscillatory behaviour" is the restriction that S is a set of points in convex position (for example, the boundary of a strictly convex set). We are able to prove a weaker version of Theorem 1.1 in this setting; see Section 1.3.

It seems plausible that the optimal point concentration bound $O(1/\sqrt{n})$ given by the Erdős–Littlewood–Offord theorem also holds in the setting of Theorem 1.1.

Conjecture 1.2. *In Theorem 1.1 we have $\Pr(X \in S) \le C_S/\sqrt{n}$, for some C_S depending only on S.*

We are able to prove Conjecture 1.2 when S has a certain "generic intersection property" (see Definition 1.6), which in particular holds when $S \subset \mathbb{R}^2$ is a convex or irreducible algebraic plane curve, when $S \subset \mathbb{C}^2 \cong \mathbb{R}^4$ is an irreducible complex algebraic curve (which is of real dimension 2), and when $S \subset \mathbb{R}^3$ is the boundary of a 3-dimensional ball (see Theorem 1.7).

In addition to high-dimensional extensions, many other variants and generalisations of the Littlewood–Offord problem have been proposed and studied. One especially influential direction is the *inverse theory* of the Littlewood–Offord problem, which studies the relationship between the concentration behaviour of X and the arithmetic structure of $a_1, \ldots, a_n$. In particular, inverse Littlewood–Offord theorems were essential tools for some of the landmark results in random matrix theory (see for example [40, 42, 44]).

Recall that a *generalised arithmetic progression* of *rank q* is a subset of $\mathbb{R}^d$ of the form

$$Q_{b,v_1,\ldots,v_q,m_1,\ldots,m_q} := \{b + r_1 v_1 + \cdots + r_q v_q : r_i \in \{0, \ldots, m_i - 1\} \text{ for all } i\},$$

where $b, v_1, \ldots, v_q \in \mathbb{R}^d$ and $m_1, \ldots, m_q \in \mathbb{N}$. Perhaps the most famous inverse theorem, proved by Tao and Vu [42] (see also the sharpenings in [25, 41]), says that if any point probability $\Pr(X = x)$ is polynomially large (that is, at least n^{-C} for any C), then there must be a very strong arithmetic reason: all but n^ε of the vectors $a_1, \ldots, a_n$ must lie in a rank q generalised arithmetic progression of size at most n^B (with q, B bounded in terms of $C, \varepsilon > 0$).

We prove an analogous inverse theorem (Theorem 9.1) in the setting of Theorem 1.1: if $\Pr(X \in S)$ is polynomially large, then almost all of the vectors $a_1, \ldots, a_n$ lie in a small, low-rank generalised arithmetic progression. This follows from the following results for arbitrary (not necessarily definable) $S \subseteq \mathbb{R}^d$.

Theorem 1.3. *Consider any subset $S \subseteq \mathbb{R}^d$, and let $\xi_1, \xi_2, \ldots$ be i.i.d. Rademacher random variables. Suppose that there exists $m \geq 1$ such that either*

(a) $\Pr(a_1 \xi_1 + \cdots + a_{m+1} \xi_{m+1} \in S) < 1/2$ *for all nonzero $a_1, \ldots, a_{m+1} \in \mathbb{R}^d$, or, more generally*

(b) $\Pr(a_1 \xi_1 + \cdots + a_m \xi_m \in S') < 1$ *for all nonzero $a_1, \ldots, a_m \in \mathbb{R}^d$ and all translates S' of S.*

Then, for all n and nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, we have $\Pr(X \in S) \leq C_m \rho^{1/(m2^{m-1})}$, where $X = a_1 \xi_1 + \cdots + a_n \xi_n$ and $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$, and C_m depends only on m.

Proof that assumption (a) implies assumption (b). Consider any translate $S' = S - b$ of S. If $b = 0$, we note that $\Pr(a_1 \xi_1 + \ldots + a_{m-1} \xi_{m-1} + (a_m/2)\xi_m + (a_m/2)\xi_{m+1} \in S) < 1/2$ implies that $\Pr(a_1 \xi_1 + \ldots + a_m \xi_m \in S') < 1$, and if $b \neq 0$ we note that $\Pr(a_1 \xi_1 + \ldots + a_m \xi_m + b\xi_{m+1} \in S) < 1/2$ implies that $\Pr(a_1 \xi_1 + \ldots + a_m \xi_m \in S') < 1$. $\qquad\square$

In other words, Theorem 1.3 says there is a polynomial relationship between probabilities of the form $\Pr(X \in S)$ and the maximum point concentration probability $\max_{x \in \mathbb{R}^d} \Pr(X = x)$, whenever we have a nontrivial uniform bound on $\Pr(X \in S)$ for any individual n.

Theorem 1.1 and Theorem 1.3 together imply that whenever $S \subseteq \mathbb{R}^d$ is a definable set not containing a line segment, then $\Pr(X \in S) = O(\rho^{c_S})$ for some c_S depending on S. One can then combine this statement with the Tao–Vu inverse theorem (or any other inverse theorem for the Littlewood–Offord problem). We remark that the logical structure of the paper does not actually proceed in quite this way: we first directly prove an inequality (Theorem 1.14) of the form $\Pr(X \in S) = O(\rho^{c_S})$, where c_S only depends on S via its dimension. We then use this inequality, in addition to the Tao–Vu inverse theorem, to deduce an inverse theorem (Theorem 9.1) in our setting, and then use this inverse theorem in the proof of Theorem 1.1.

We also remark that the condition in Theorem 1.3 that $\Pr(a_1 \xi_1 + \cdots + a_m \xi_m \in S') < 1$ (for all translates S' of S, and all nonzero $a_1, \ldots, a_m \in \mathbb{R}^d$) is equivalent to the condition that S does not include a generalised arithmetic progression $Q_{b,v_1,\ldots,v_m,2,\ldots,2}$ for nonzero $v_1, \ldots, v_m \in \mathbb{R}^d$, which is in turn equivalent to the condition that S does not include a Minkowski sum $A_1 + \cdots + A_m$ with each $|A_i| = 2$. As it turns out, the interaction between Minkowski sums and the geometry of S will play a crucial role in what follows.

In the next few subsections we describe the ideas in the proofs of Theorem 1.1 and Theorem 1.3, and present a few additional related results.

1.1 Polynomial anti-concentration and semi-algebraic sets

Recall that $S \subseteq \mathbb{R}^d$ is a semi-algebraic set if it can be expressed as the set of points satisfying a Boolean combination of equations $f = 0$ and inequalities $g > 0$ for various polynomials f, g. We start by

introducing a standard notion of complexity for semi-algebraic sets (this is not quite the same as the o-minimal notion of complexity mentioned previously, but as we will see later it is closely related).

Definition 1.4. *Say that a semi-algebraic subset $S \subset \mathbb{R}^d$ has* description complexity *at most r if it can be written as a Boolean combination of at most r sets of the form $\{x \in \mathbb{R}^d : f(x) = 0\}$ and $\{x \in \mathbb{R}^d : f(x) > 0\}$, for polynomials f of degree at most r.*

We next discuss how to prove Theorem 1.1 for semi-algebraic sets, using some of the ideas that have been developed in the study of polynomial anti-concentration. These ideas will also play a key role in the proof of the full (o-minimal) version of Theorem 1.1.

The *polynomial Littlewood–Offord problem* is concerned with upper bounds on probabilities of the form $\Pr(p(\xi_1, \ldots, \xi_n) = x)$, where p is an n-variable polynomial satisfying certain conditions (and, as always, $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables). This topic has its roots in random matrix theory (see [7, 26]), but there are also a number of connections between the polynomial Littlewood–Offord problem and the theory of Boolean functions (see [24, 37]). The strongest and most general result in this area is due to Meka, Nguyen and Vu [24], establishing near-optimal anti-concentration estimates in terms of a certain combinatorial parameter of the polynomial p (see also [13]).

As observed by Kane (see the discussion in [24]), the result of Meka, Nguyen and Vu can actually be deduced from estimates related to the *Gotsman–Linial conjecture*. Recall that a *degree-r polynomial threshold function* $f : \{-1, 1\}^n \to \{-1, 1\}$ is a Boolean function of the form $\mathrm{sgn}(p(x))$ for some polynomial $p : \mathbb{R}^d \to \mathbb{R}$ of degree at most r, where

$$\mathrm{sgn}(y) = \begin{cases} 1 & y > 0 \\ -1 & y \leq 0. \end{cases}$$

The i-th *influence* $\mathrm{Inf}_i(f)$ of a Boolean function f is the probability that, at a random evaluation point $(\xi_1, \ldots, \xi_n) \in \{-1, 1\}^n$, changing the value of ξ_i would change the value of $f(\xi_1, \ldots, \xi_n)$. The *total influence* or *average sensitivity* of f is

$$\mathrm{AS}(f) = \mathrm{Inf}_1(f) + \cdots + \mathrm{Inf}_n(f).$$

Then, the Gotsman–Linial conjecture is that[1] for any degree-r polynomial threshold function f, we have $\mathrm{AS}(f) = O(r\sqrt{n})$. Kane [19] proved the weaker result $\mathrm{AS}(f) \leq C_r(\log n)^{C_r}\sqrt{n}$ for constant C_r depending only on r (see Theorem 8.1), from which he was able to deduce a strong bound on the polynomial Littlewood–Offord problem (see the discussion in [24]). We are similarly able to deduce the following quantitative version of Theorem 1.1 in the case of semi-algebraic sets.

Theorem 1.5 (cf. Theorem 1.1). *Suppose $S \subset \mathbb{R}^d$ is a k-dimensional semi-algebraic set of description complexity r not containing a line segment. Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Then there are $C_{d,r}$ depending only on d, r, and C_r depending only on r, such that $\Pr(X \in S) \leq C_{d,r}(\log n)^{C_r}/\sqrt{n}$.*

We remark that a full resolution of the Gotsman–Linial conjecture would remove the $(\log n)^{C_r}$ factor, implying Conjecture 1.2 for semi-algebraic sets not including a line segment. In fact, there is a natural special case of the Gotsman–Linial conjecture that would already suffice, and does not appear to have previously been studied (see Section 10).

1.2 The geometry and additive combinatorics of incidence graphs

To prove Theorem 1.14, and to complete the proof of Theorem 1.1 for general sets definable with respect to an o-minimal structure, we study the geometry of $S \subset \mathbb{R}^d$ via the additive combinatorics of finite subsets of S. In this subsection we explain how to encode certain combinatorial information in a weighted graph, and prove special cases of Conjecture 1.2 by considering forbidden subgraphs of this weighted graph. This will serve as a warm-up to the next few subsections, where we extend these ideas to hypergraphs.

For a point set $\mathcal{F} \subseteq \mathbb{R}^d$ and a family of geometric objects $\mathcal{G}$ in $\mathbb{R}^d$, the *bipartite incidence graph* $G(\mathcal{F}, \mathcal{G})$ is the graph with parts $\mathcal{F}$ and $\mathcal{G}$, with an edge between $f \in \mathcal{F}$ and $g \in \mathcal{G}$ whenever $f \in g$. This notion is

---

[1]Gotsman and Linial actually conjectured a stronger bound, which has since been disproved; see [3, 20].

fundamental in incidence geometry: one important way of proving bounds on the number of incidences between $\mathcal{F}$ and $\mathcal{G}$ is to study the subgraphs that $G(\mathcal{F}, \mathcal{G})$ cannot contain, due to the geometry of $\mathcal{F}, \mathcal{G}$, and to then apply a theorem from extremal graph theory. For example, one can obtain a non-trivial upper bound for the Erdős unit distance problem by taking $\mathcal{F} \subset \mathbb{R}^2$ a set of points, and $\mathcal{G}$ the set of unit circles centred at these points. Then $G(\mathcal{F}, \mathcal{G})$ contains no $K_{2,3}$, and the number of edges in $G(\mathcal{F}, \mathcal{G})$ can be upper-bounded with the Kővári–Sós–Turán theorem (see [23, Chapter 4.5]).

To study probabilities of the form $\Pr(X \in S)$ using incidence graphs, we consider a partition $\{1, \ldots, n\} = I_1 \sqcup I_2$, and defining the random variables $X_i = \sum_{j \in I_i} a_j \xi_j$ (where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables), we let V_i be the support of X_i. Then we consider the bipartite graph G with parts V_1, V_2 and an edge between $x \in V_1$ and $y \in V_2$ whenever $x + y \in S$. Note that G can be interpreted as the bipartite incidence graph $G(\mathcal{F}, \mathcal{G})$ for $\mathcal{F} = V_1$ and $\mathcal{G} = \{S - x : x \in V_2\}$. By assigning each vertex $x \in V_i$ the weight $w(x) := \Pr(X_i = x)$, we have

$$\Pr(X \in S) = \Pr(X_1 + X_2 \in S) = \sum_{xy \in E(G)} w(x)w(y).$$

Now, under certain geometric conditions on S, we can show that G is $K_{a,b}$-free for some a, b. We can then bound $\Pr(X \in S)$ by proving and applying a weighted version of the Kővári–Sós–Turán theorem (Theorem 3.1) to G. Specifically, we note that G is $K_{a,b}$-free if either of the following two equivalent conditions hold.

1. (Additive combinatorics interpretation). For any $A, B \subset \mathbb{R}^d$ with $|A| = a$, $|B| = b$, we have $A + B \not\subseteq S$.

2. (Incidence geometry interpretation). For any $A \subset \mathbb{R}^d$ with $|A| = a$, we have $\left| \bigcap_{a \in A} (S - a) \right| < b$.

As an application of these ideas, suppose that $S \subset \mathbb{R}^d$ has dimension $\dim(S) \leq d - 1$. Then, heuristically, we would expect that d generic translates $S - x_1, S - x_2, \ldots, S - x_d$ intersect in a 0-dimensional set, i.e., finitely many points. If there is a constant D such that for *every* choice of translates, we have $|(S - x_1) \cap \cdots \cap (S - x_d)| \leq D$, then G is $K_{d,D+1}$-free. To prove an optimal anti-concentration bound for $\Pr(X \in S)$, we will need a version of this condition which is inherited by intersections with affine subspaces.

Definition 1.6. *Say that $S \subseteq \mathbb{R}^d$ satisfies the* generic intersection property *if there is $D \in \mathbb{N}$ such that the following holds. For any affine subspace $\Lambda \subseteq \mathbb{R}^d$ (including $\Lambda = \mathbb{R}^d$), we have*

$$\left| \bigcap_{i=1}^{\dim \Lambda} ((S \cap \Lambda) - x_i) \right| \leq D$$

for any distinct $x_1, \ldots, x_{\dim \Lambda} \in \mathbb{R}^d$.

Using the strategy described above together with a strong Littlewood–Offord-type theorem due to Hálasz [17] (see Theorem 2.6), we are able to prove Conjecture 1.2 for sets S satisfying the generic intersection property.

Theorem 1.7. *Suppose that $S \subset \mathbb{R}^d$ has the generic intersection property. Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1 \xi_1 + \cdots + a_n \xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Then there is C_S depending only on S such that:*

1. $\Pr(X \in S) \leq C_S n^{-1/2}$, and
2. $\Pr(X \in S) \leq C_S \rho^{1/(d+1)}$, where $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$.

The first part of Theorem 1.7 is best possible, but the second part may not be. Some interesting examples of sets S satisfying the generic intersection property are strictly convex plane curves, irreducible complex algebraic curves other than lines, and 2-dimensional spheres. Also, note that if S can be represented as a finite union of sets satisfying Conjecture 1.2, then S satisfies Conjecture 1.2 itself. So, Theorem 1.7 actually implies that Conjecture 1.2 holds for plane curves with any finite number of inflection points and vertical tangent lines.

As a final remark, we note that there are very close connections between our weighted Kővári–Sós–Turán theorem (Theorem 3.1) and some other inequalities previously used in connection with the Littlewood–Offord problem. For example, in [6], a weighted version of the closely-related Szemerédi–Trotter theorem [18] was applied to study the quadratic Littlewood–Offord problem. There is also a close connection to *decoupling inequalities* (see for example [6, 7, 37]) commonly applied to the polynomial Littlewood–Offord problem, which are themselves closely related to *Sidorenko's conjecture* (see for example [39]) in extremal graph theory. See Remark 3.2 for further discussion.

1.3   Sum hypergraphs, and an application to points in convex position

In Section 1.2, we explained how to prove bounds on probabilities of the form $\Pr(X \in S)$ under the assumption that S does not include a sumset $A + B$, for small sets A, B. In this subsection we explain how to generalise this to prove Theorem 1.3, giving bounds on probabilities of the form $\Pr(X \in S)$ under the assumption that S does not include an iterated sumset $A_1 + \cdots + A_k$. The ideas in this subsection will be extended further in the next subsection, when we begin to discuss the geometry of sets definable with respect to an o-minimal structure.

We remark that some other natural questions on the additive combinatorics of finite subsets of $S \subset \mathbb{R}^d$ can be phrased in terms of the intersection of S with sumsets $A_1 + \cdots + A_k$. For example, if $k = d$ and A_i consists of multiples of the ith standard basis vector, then $A_1 + \cdots + A_k$ is a *combinatorial box* (which may also be expressed in the form $X_1 \times \cdots \times X_d \subseteq \mathbb{R}^d$). The *Elekes–Rónyai–Szabó problem* [9, 10] asks what structure is imposed on S if $S \subseteq \mathbb{R}^d$ is the zero locus of a single polynomial and has large intersection with a combinatorial box $X_1 \times \cdots \times X_d \subset \mathbb{R}^d$. See [33, 34, 35, 36] for some recent work on this problem. If we specialise further, requiring $X_i = [a_i, b_i] \cap \frac{1}{m}\mathbb{Z}$ for some fixed m, then the size of the intersection with S determines the number of rational points of S with denominator m in a box $\prod_i [a_i, b_i]$. This direction of study has been closely connected to recent developments in *o-minimal geometry* (see Section 1.4).

As an example of a natural geometric condition which precludes the inclusion of an iterated sumset, it is known that if a set of points in $\mathbb{R}^d$ lies in convex position (meaning no point is a convex combination of the others), then it cannot contain a sumset of the form $A_1 + \cdots + A_{d+1}$, where each $|A_i| \geq 2$ (see Lemma 6.1).

Definition 1.8. *For $S \subset \mathbb{R}^d$ and subsets $V_1, \ldots, V_k \subset \mathbb{R}^d$, we define the k-uniform k-partite sum hypergraph $H_S(V_1, \ldots, V_k)$ to have vertex sets $V_1, \ldots, V_k$, and an edge $v_1 \ldots v_k$ whenever $v_1 + \cdots + v_k \in S$.*

For $a_1, \ldots, a_n \in \mathbb{R}^d$, and a partition $\{1, \ldots, n\} = I_1 \sqcup \cdots \sqcup I_k$, we define the random variables $X_i = \sum_{j \in I_i} a_j \xi_j$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables, and let V_i be the support of X_i. Then we define $H_{X,S}(I_1, \ldots, I_k)$ to be the hypergraph $H_S(V_1, \ldots, V_k)$, with associated vertex-weight function w given by $w(x) := \Pr(X_i = x)$ for $x \in V_i$.

Note that with this definition, we have

$$\Pr(X \in S) = \Pr(X_1 + \cdots + X_k \in S) = \sum_{x_1 \ldots x_k \in E(H_{X,S}(I_1, \ldots, I_k))} w(x_1) \ldots w(x_k).$$

Erdős [11] proved a hypergraph extension of the Kővári–Sós–Turán theorem, which in particular gives an upper bound on the number of edges in a k-uniform hypergraph which is $K_{2,\ldots,2}^{(k)}$-free. We prove a weighted version of this (actually a slight generalisation; see Corollary 5.2) and we use this to prove Theorem 1.3. As an application we prove the following theorem on sets of points in convex position.

Theorem 1.9. *Suppose that $S \subset \mathbb{R}^d$ is a set of points in convex position. Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Then there is C_d depending only on d such that:*

1. $\Pr(X \in S) \leq C_d n^{-d/2^d}$, and
2. $\Pr(X \in S) \leq C_d \rho^{1/(d2^{d-1})}$, where $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$.

If S is any affine variety not containing a line segment, then one can show using Bézout's theorem that for sufficiently large k there is no sumset of the form $A_1 + \cdots + A_k$ included in S, where each $|A_i| \geq 2$.

However, we instead consider a more sophisticated variation of this fact, in the context of sets definable with respect to o-minimal structures, in the following subsection.

1.4 o-minimal geometry and irreducible components

Recall that affine varieties in $\mathbb{R}^d$ have notions of "irreducible component" and "degree", which control the complexity of intersections. More precisely, the following properties are satisfied.

1. Every k-dimensional affine variety $S \subset \mathbb{R}^d$ with degree r is a union of at most r irreducible varieties with dimension at most k.

2. (Bézout's Theorem) Every pair of distinct irreducible k-dimensional varieties $S_1, S_2 \subset \mathbb{R}^d$ of degree at most r intersect in an affine variety with dimension at most $k-1$ and degree at most r^2.

Recalling the k-uniform sum hypergraph $H_S(V_1, \ldots, V_k)$ from Definition 1.8, the combinatorial manifestation of the above facts is that if S is an affine variety not containing a line segment then this hypergraph satisfies a certain recursive property (we note that the hypothesis of not containing a line segment ensures that no irreducible variety is equal to a nontrivial translate of itself). In a hypergraph H with edge set E, recall that the *link* of a vertex v is the $(k-1)$-uniform hypergraph with edge set $\{e : e \cup \{v\} \in E\}$. The *common link* of a pair of vertices is the intersection of the links of those vertices.

Definition 1.10. *Say that a 1-uniform hypergraph H has* complexity M *if it has at most M edges. For $k > 1$, say that a k-uniform k-partite hypergraph H with vertex sets $V_1, \ldots, V_k$ has complexity M if we can write $H = H_1 \cup \ldots H_M$ such that for $1 \leq i \leq M$, the common link in H_i of any two elements $x, y \in V_1$ is a $(k-1)$-uniform $(k-1)$-partite hypergraph on the vertex sets $V_2, \ldots, V_k$ with complexity M.*

Note that if a k-uniform k-partite hypergraph H is $K_{2,\ldots,2}^{(k)}$-free, then it has complexity 1. As it happens, a bounded-complexity assumption is enough to prove a (weighted) Kővári–Sós–Turán-type theorem (Corollary 5.2), which implies Theorems 1.1 and 1.14 for affine varieties. Actually, one can prove that the sum hypergraph H has bounded complexity under a much weaker condition on S, namely the condition that S does not contain a line segment and is *definable with respect to an o-minimal structure*. We will introduce this notion after discussing some recent developments in the area (see [45] for a historical introduction and [38] for a more recent survey).

Following earlier results for curves in $\mathbb{R}^2$ due to Bombieri and Pila [2], Pila [28] proved that any affine variety $S \subset \mathbb{R}^d$ of degree r and dimension k satisfies

$$\left| S \cap \frac{1}{n} \mathbb{Z}^d \cap \prod_{i=1}^d [a_i, b_i] \right| = n^{k-1+1/r+o(1)}$$

for any constants $a_1 < b_1, \ldots, a_d < b_d$. Extending beyond the algebraic setting, Pila and Wilkie [29] considered the setting where $S \subset \mathbb{R}^d$ is definable with respect to an o-minimal structure, in which case one can obtain similar results by decomposing S into an "algebraic part" and a complementary "transcendental part". Exploiting these rational point counting theorems, powerful results in arithmetic dynamics such as the Manin–Mumford conjecture and parts of the André–Oort conjecture have recently been attacked with great success (see for example [31, 32]).

On the more combinatorial side, a number of results bounding the number of edges in incidence graphs and hypergraphs (such as the work of Fox, Pach, Sheffer, Suk and Zahl [14], and Do [8] on the semialgebraic Zarankiewicz problem) have been generalised in the o-minimal direction. In particular, Basu and Raz [1] have generalised the Szemerédi–Trotter theorem to certain o-minimal incidence graphs and Chernikov, Galvin, and Starchenko [4] have generalised this to more general incidence hypergraphs.

To proceed further, we finally define the notion of an o-minimal structure. Recall that a first-order formula over $\mathbb{R}$ is a well-formed sentence involving parentheses $(,)$, quantifiers $\exists, \forall$, predicates (such as $>, <, =$), connectives (such as $\vee, \wedge, \rightarrow, \neg, \top, \bot$), variables (which we give single-letter names such as x, y, z), operations (such as $+$, $\times$, exp), and constants (one for every element of $\mathbb{R}$), such that we can specialise the unbound variables to elements of $\mathbb{R}$ and interpret the sentence as either true $\top$ or false $\bot$. For example, if $\phi(x)$ is the formula $(\exists y)(y^2 = x) \wedge (\exists z)(z^2 = 1 - x)$, then $\phi(x)$ is true exactly when the unbound variable x has $x \in [0, 1]$.

Definition 1.11. *A collection $\mathcal{F}$ of functions $f : \mathbb{R}^{\ell_f} \to \mathbb{R}$ of various arities ℓ_f generates an o-minimal structure $\mathbb{R}_{\mathcal{F}}$ (over the real field) if any first-order formula ϕ involving one free parameter x, operations in $\mathcal{F} \cup \{+, \times\}$ and predicates $>, <, =$, has the property that the set*

$$S_\phi := \{x : \phi(x) \text{ is true}\} \subseteq \mathbb{R}$$

is a finite union of points and (possibly unbounded) intervals.

We say a subset $S \subseteq \mathbb{R}^d$ is definable *(with respect to $\mathcal{F}$) if there is a first-order formula ϕ with d free parameters $x_1, \ldots, x_d$, operations in $\mathcal{F} \cup \{+, \times\}$, and predicates $>, <, =$, such that*

$$S = S_\phi := \{(x_1, \ldots, x_d) : \phi(x_1, \ldots, x_d) \text{ is true}\} \subseteq \mathbb{R}^d.$$

Example 0: $\mathcal{F} = \{f\}$ with $f(x, y) = \sin(x + y)$ does not generate an o-minimal structure. Indeed, the formula $\phi(x)$ given by $f(x, 0) = 0$ has $S_\phi = 2\pi\mathbb{Z}$, which is not a finite union of points and intervals.

Example 1: Take $\mathcal{F} = \emptyset$. Obviously, semi-algebraic sets (Boolean combinations of subsets of $\mathbb{R}^d$ given by polynomial equalities and inequalities) are definable. *A priori*, it may seem that other more exotic sets are definable as well (for example, images of semi-algebraic sets under polynomial maps are definable, and one can define sets using functions that are themselves only defined implicitly in terms of polynomials). However, the Tarski–Seidenberg theorem [43] states that in this setting actually the only definable sets are semi-algebraic subsets of $\mathbb{R}^d$ (so, the "exotic" sets described above are actually semi-algebraic, though this isn't obvious), and taking $d = 1$ shows that $\mathcal{F}$ generates an o-minimal structure (commonly denoted $\mathbb{R}_{\mathrm{alg}}$).

Example 2: Wilkie [47] showed that $\mathcal{F} = \{\exp\}$ generates an o-minimal structure. In this setting, definable sets are rather more complicated to describe, but in [47] it is shown that they are all of the form $\pi(\{(x_1, \ldots, x_n) : f(x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}) = 0\})$, where f is a polynomial in $2n$ variables and π is a linear projection $\mathbb{R}^n \to \mathbb{R}^m$ for some $m \in \mathbb{N}$.

Example 3: van den Dries and Miller [46] showed that exp, together with all "restricted analytic functions" ($f|_{[0,1]^d}$ for f an analytic function on an open neighborhood of $[0, 1]^d$) generate an o-minimal structure.

In the setting of o-minimal geometry, the following notion plays a similar role to the notion of degree for affine varieties, and description complexity for semi-algebraic sets (recall Definition 1.4).

Definition 1.12. *We say an o-minimal structure $\mathbb{R}_{\mathcal{F}}$ is said to be* finitely generated *if $\mathcal{F}$ is finite.*

For a finitely generated o-minimal structure $\mathbb{R}_{\mathcal{F}}$, the complexity *of a formula ϕ is the number of symbols used in its description (every constant from $\mathbb{R}$ counts as one symbol). The complexity of a definable set S is then defined to be the minimum complexity of a formula defining S.*

Examples 1 and 2 above describe finitely generated o-minimal structures, but Example 3 does not. For an arbitrary o-minimal structure $\mathbb{R}_{\mathcal{F}}$, note that every definable set only uses a finite number of the functions in $\mathcal{F}$ in its description. Hence restricting our attention to finitely generated o-minimal structures is not actually a restriction at all, and in fact our notion of complexity is equivalent to the notion of "definable families" more commonly used in the o-minimality literature (see Remark 7.7).

By the Tarski–Seidenberg theorem (see [43]), the degree of an affine variety and the description complexity of a semi-algebraic set are bounded from above and below by a function of its complexity with respect to the o-minimal structure $\mathbb{R}_{\emptyset} = \mathbb{R}_{\mathrm{alg}}$.

We show that the above notion of complexity controls the hypergraph notion of complexity in Definition 1.10. This is done by constructing an analogue of irreducible components we call *self-irreducible components*, for definable sets containing no line segment, which satisfies a restricted analogue of Bézout's theorem in which we only consider intersections $S_1 \cap S_2$ for which S_1, S_2 are translates of each other.

Proposition 1.13. *For a finitely generated o-minimal structure $\mathbb{R}_{\mathcal{F}}$ and any $r, d \in \mathbb{N}$, there is $C_{\mathcal{F},d,r}$ such that for any k-dimensional definable subset $S \subset \mathbb{R}^d$ with complexity r not containing a line segment, any $(k+1)$-uniform $(k+1)$-partite intersection hypergraph $H_S(V_1, \ldots, V_{k+1})$ has complexity at most $C_{\mathcal{F},d,r}$.*

Using Proposition 1.13 and a Kővári–Sós–Turán-type theorem, we deduce the following strong bound comparing $\Pr(X \in S)$ with the maximum point concentration probability.

Theorem 1.14. *Let $S \subseteq \mathbb{R}^d$ be a set which is definable with respect to an o-minimal structure and does not contain any line segment. Consider nonzero d-dimensional vectors $a_1, \dots, a_n \in \mathbb{R}^d$, and write $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \dots, \xi_n$ are i.i.d. Rademacher random variables. Then there is a constant C_S such that for $k = \dim(S)$ and $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$, we have*

$$\Pr(X \in S) \leq C_S \rho^{1/((k+1)2^k)}.$$

Here the constant C_S depends only on the dimension d and the complexity of S.

The proof of Theorem 1.1 is a little more involved, and combines most of the ideas we have discussed so far. First, using the Tao–Vu inverse theorem we observe that the desired result follows from Theorem 1.14 unless most of the coefficients $a_1, \dots, a_n$ lie in a small generalised arithmetic progression with low rank. After conditioning on a small number of ξ_i, the random variable X is conditionally supported in a low-rank generalised arithmetic progression Q. We then adapt parts of a general theory due to Pila [30] to cover $S \cap Q$ with a small number of projections of open subsets of semi-algebraic sets. Although the resulting covering sets are not themselves semi-algebraic, we can use similar ideas as in the proof of Theorem 1.5 to deduce Theorem 1.1.

We note that in the works mentioned earlier in this subsection, counting the number of rational points [29] and algebraic points [30] of bounded height on definable sets S, one has very poor control on the number of such points lying on the "algebraic" part of S. One therefore obtains the strongest bounds by excising this part from S. It is worth highlighting that our situation is quite different: we have bounds of the form $\Pr(X \in S_{\mathrm{alg}}) \leq n^{-1/2+o(1)}$ for the algebraic part S_{alg}, nearly matching the point concentration bound $\Pr(X = x) = O(n^{-1/2})$ given by the Erdős–Littlewood–Offord theorem.

2 Preliminaries on point concentration

In this section we collect a few results about point concentration, which will be useful throughout the paper. For $a_1, \dots, a_n \in \mathbb{R}^d$, let $\rho(a_1, \dots, a_n) = \max_{x \in \mathbb{R}^d} \Pr(a_1\xi_1 + \cdots + a_n\xi_n = x)$, where $\xi_1, \dots, \xi_n$ are i.i.d. Rademacher random variables. We start with the following simple facts.

Fact 2.1. *For any subset $I \subseteq \{1, \dots, n\}$ we have $\rho(a_1, \dots, a_n) \leq \rho((a_i)_{i \in I})$.*

Fact 2.2. *For any subset $I \subseteq \{1, \dots, n\}$ we have $\rho(a_1, \dots, a_n) \geq \rho((a_i)_{i \in I})\rho((a_i)_{i \notin I})$.*

Fact 2.3. *We have $\rho(a_1, \dots, a_n) \geq \rho(a_1, \dots, a_{n-1})/2$.*

We also need a "partite" notion of point concentration.

Definition 2.4. *Say that a sequence $(a_1, \dots, a_n) \in (\mathbb{R}^d)^n$ is $(\lambda_1, \dots, \lambda_m)$-anti-concentrated if there is a partition $I_1 \sqcup \cdots \sqcup I_m = \{1, \dots, n\}$ such that $\max_{x \in \mathbb{R}^d} \Pr(\sum_{i \in I_j} a_i\xi_i = x) \leq \lambda_j$ for all j. If $\lambda_1 = \dots \lambda_m = \lambda$, we abbreviate this by saying $(a_1, \dots, a_n)$ is m-part λ-anti-concentrated.*

For example, note that if each $a_i \neq 0$ then $(a_1, \dots, a_n)$ is m-part $O(m^{1/2}n^{-1/2})$-anti-concentrated. Indeed, this follows from the Erdős–Littlewood–Offord theorem, and any equipartition $I_1 \sqcup \cdots \sqcup I_m$ will do.

We next show that upper bounds on point concentration imply similar upper bounds on partite point concentration.

Lemma 2.5. *For $0 \leq \lambda_1, \dots, \lambda_m < 1/2$, if $\rho(a_1, \dots, a_n) \leq \prod_{i=1}^m \lambda_i$, then $(a_1, \dots, a_n)$ is $(2\lambda_1, \dots, 2\lambda_m)$-anti-concentrated.*

Proof. We prove this by induction on m, noting that the case $m = 1$ is trivial. Let $p_i = \rho(a_1, \dots, a_i)$ and let $q_i = \rho(a_{i+1}, \dots, a_n)$. Note that $p_0 = 1$ and $p_n = \rho(a_1, \dots, a_n)$.

Now, for $i = 0, 1, \dots, n - 1$, we have $p_i \geq p_{i+1} \geq p_i/2$ by Fact 2.1 and Fact 2.3. There exists by discrete continuity an i for which $\lambda_1 \leq p_i \leq 2\lambda_1$. Note that $p_i q_i \leq \rho(a_1, \dots, a_n) \leq \prod_{i=1}^m \lambda_i$ by Fact 2.2, so $q_i \leq \prod_{i=2}^m \lambda_i$. Therefore, by the induction hypothesis, $(a_{i+1}, \dots, a_n)$ is $(2\lambda_2, \dots, 2\lambda_m)$-anti-concentrated (with partition $I_2 \sqcup \cdots \sqcup I_m = \{i+1, \dots, n\}$, say). Taking $I_1 = \{1, \dots, i\}$, the partition $I_1 \sqcup \cdots \sqcup I_m = \{1, \dots, n\}$ shows that $(a_1, \dots, a_n)$ is $(2\lambda_1, \dots, 2\lambda_m)$-anti-concentrated, as desired. $\square$

Finally, we will need the following theorem of Halász, which gives strong bounds for the Littlewood–Offord problem in the case where the coefficients "robustly" span the entire space $\mathbb{R}^d$.

Theorem 2.6 ([17, Theorem 1]). *Let $a_1, \ldots, a_n \in \mathbb{R}^d$ be vectors such that no proper linear subspace of $\mathbb{R}^d$ contains half of the vectors a_i. Let $\xi_1, \ldots, \xi_n$ be i.i.d. Rademacher random variables, and let $X = a_1\xi_1 + \cdots + a_n\xi_n$. Then $\max_x \Pr(X = x) \leq C_d n^{-d/2}$, where C_d depends only on d.*

3 A weighted version of the Kővári–Sós–Turán theorem

In this section we prove a weighted version of the Kővári–Sós–Turán theorem, which we will use frequently to bound the total weight of various graphs and hypergraphs.

Theorem 3.1. *Consider a bipartite graph G with bipartition $A \cup B$, where the vertices in A, B are weighted (with total weights $w(A) = w(B) = 1$). Let $w(G) = \sum_{a \sim b} w(a)w(b)$. Assume every set of t vertices in A has common neighbourhood with total weight at most q, and assume every vertex in A has weight at most ρ. Then*
$$w(G) \leq q^{1/t} + t\rho.$$

We remark that the unweighted Kővári–Sós–Turán theorem follows from Theorem 3.1 if we set $w(a) = 1/|A|$ for each $a \in A$ and $w(b) = 1/|B|$ for each $b \in B$.

Remark 3.2. A weighted bipartite graph G as in Theorem 3.1 can be viewed as encoding an event $\mathcal{E}(X_1, X_2)$ depending on two random variables X_1, X_2. Given independent copies $X_1^{(1)}, \ldots, X_1^{(t)}$ of X_1, it follows from Hölder's inequality (or alternatively Jensen's inequality) that
$$\Pr(\mathcal{E}(X_1, X_2)) \leq \Pr(\mathcal{E}(X_1^{(1)}, X_2) \cap \ldots \cap \mathcal{E}(X_1^{(t)}, X_2))^{1/t}.$$

This can be interpreted as a *decoupling inequality* (see for example [7, Lemma 4.7] for a related inequality). Let $\mathcal{F}$ be the event that some $X_1^{(i)} = X_1^{(j)}$, so in the setting of Theorem 3.1,
$$\Pr(\mathcal{E}(X_1^{(1)}, X_2) \cap \ldots \cap \mathcal{E}(X_1^{(t)}, X_2)) \leq \Pr(\mathcal{E}(X_1^{(1)}, X_2) \cap \ldots \cap \mathcal{E}(X_1^{(t)}, X_2) \,|\, \overline{\mathcal{F}}) + \Pr(\mathcal{F}) \leq q + \binom{t}{2}\rho,$$

and therefore $w(G) \leq \left(q + \binom{t}{2}\rho\right)^{1/t}$. This is comparable to the bound $q^{1/t} + t\rho$ in Theorem 3.1, unless ρ is substantially larger than q. We note that this weaker inequality is sufficient for all applications in this paper, except the second part of Theorem 1.7.

To prove Theorem 3.1, first we prove a version where A is not weighted.

Lemma 3.3. *Consider a bipartite graph $G = (A, B)$ where the vertices in B are weighted (with total weight 1). Let $w(G) = \sum_{a \sim b} w(b)$. Assume every set of t vertices in A has common neighbourhood with total weight at most q. Then*
$$w(G) \leq q^{1/t}n + t - 1.$$

Proof. We assume that $w(G) \geq t - 1$, or else we are immediately done. For a vertex $b \in B$, let d_b be the degree of b. We have $\sum_{b \in B} w(b)\binom{d_b}{t} = \sum_{a_1, \ldots, a_t} w(N(a_1, \ldots, a_t)) \leq q\binom{n}{t}$, where the latter sum is over all t-sets of distinct vertices $a_1, \ldots, a_t \in A$. Now, the function $x \mapsto \binom{x}{t}$ is convex for $x \geq t - 1$ so by Jensen's inequality we have
$$\binom{w(G)}{t} = \binom{\sum_{b \in B} w(b)d_b}{t} \leq q\binom{n}{t}.$$

Observing that $\binom{w(G)}{t} > (w(G) - t + 1)^t/t!$ and $\binom{n}{t} < n^t/t!$, the desired result follows. $\square$

Now we deduce Theorem 3.1.

Proof of Theorem 3.1. Note that $w(G)$ can be viewed as a linear function of $w \in [0, \rho]^A$. So, the maximum value of $w(G)$, under the constraint $w(A) = 1$ and holding $w|_B$ constant, is attained when $w(u) \in \{0, \rho\}$ for all but one $u \in A$.

We can therefore assume that after deleting all the weight-zero vertices in A, and possibly one additional vertex, and multiplying all weights in A by $1/\rho$, we arrive at a graph G' with parts A' and B, where each of the $|A'| = \lfloor 1/\rho \rfloor$ vertices in A' are unweighted, and the vertices in B have the same weights as in G.

By Lemma 3.3, we have $w(G') \leq q^{1/t}(1/\rho) + t - 1$. Multiplying by ρ and observing that the deleted vertices can contribute a weight of at most ρ, the desired result follows. $\qquad\square$

4 Sets with the generic intersection property

In this section we prove Theorem 1.7. The following proposition will actually imply both parts of Theorem 1.7.

Proposition 4.1. *Let $S \subseteq \mathbb{R}^d$ have the property that every t translates of S intersect in at most M points. Suppose $(a_1, \ldots, a_n)$ is (λ_1, λ_2)-anti-concentrated. Then for $X = a_1 \xi_1 + \cdots + a_n \xi_n$, where $\xi_1, \ldots, \xi_n$ are independent Rademacher random variables, we have $\Pr(X \in S) \leq M^{1/t} \lambda_2^{1/t} + t\lambda_1$.*

Proof. Let $\{1, \ldots, n\} = I_1 \sqcup I_2$ be a partition witnessing the fact that $(a_1, \ldots, a_n)$ is (λ_1, λ_2)-anti-concentrated, and let $H = H_{X,S}(I_1, I_2)$ be the bipartite sum graph from Definition 1.8.

Now, the common neighbourhood of a t-tuple of vertices $x_1, \ldots, x_t \in A$ is the set of all $y \in B$ lying in the intersection $(S - x_1) \cap \cdots \cap (S - x_t)$, so by assumption this common neighbourhood has at most M vertices, and therefore has total weight at most $M\lambda_2$. The desired result then follows from Theorem 3.1. $\qquad\square$

We now prove Theorem 1.7.

Proof of Theorem 1.7. Recall that S having the generic intersection property means that there is a constant D_S such that for any affine subspace $\Lambda \subset \mathbb{R}^d$ (including $\Lambda = \mathbb{R}^d$) and distinct $x_1, \ldots, x_{\dim \Lambda} \in \mathbb{R}^d$ we have

$$\left| \bigcap_{i=1}^{\dim \Lambda} ((S \cap \Lambda) - x_i) \right| \leq D_S.$$

For the first part of the theorem, we show that $\Pr(X \in S) = O(n^{-1/2})$, where the implicit constant only depends on d, D_S. We induct on d. The result for $d = 1$ follows from the Erdős–Littlewood–Offord theorem, so assume $d \geq 2$ and that the result is true for all dimensions less than d.

Let C_d be as in Theorem 2.6. First, if $(a_1, \ldots, a_n)$ is 2-part λ-anti-concentrated, for $\lambda = C_d/\lfloor n/2 \rfloor^{-d/2} = O(n^{-d/2})$, then we can apply Proposition 4.1 to obtain $\Pr(X \in S) = O((n^{-d/2})^{1/d}) = O(n^{-1/2})$ as desired.

Otherwise, if $\rho(a_1, \ldots, a_{\lfloor n/2 \rfloor}) > C_d/\lfloor n/2 \rfloor^{-d/2}$ or $\rho(a_{\lfloor n/2 \rfloor+1}, \ldots, a_n) > C_d/\lfloor n/2 \rfloor^{-d/2}$ then by Theorem 2.6 at least $\lfloor n/2 \rfloor/2$ of the coefficients a_i lie in a linear subspace $W \subset \mathbb{R}^d$ of dimension $d-1$ (without loss of generality, suppose this is the case for the first $\lfloor n/2 \rfloor/2$ coefficients). Let $X_1 = \sum_{i \leq \lfloor n/2 \rfloor/2} a_i \xi_i$ and $X_2 = \sum_{i > \lfloor n/2 \rfloor/2} a_i \xi_i$. Then

$$\Pr(X \in S) = \Pr(X_1 \in S - X_2) = \Pr(X_1 \in (S - X_2) \cap W).$$

But, if we condition on any outcome of X_2, the set $(S - X_2) \cap W \subseteq W$ itself satisfies the generic intersection property (inside $W \cong \mathbb{R}^{d-1}$), so by the inductive hypothesis we know that

$$\Pr(X_1 \in (S - X_2) \cap W) = O((\lfloor n/2 \rfloor/2)^{-1/2}) = O(n^{-1/2}),$$

implying the desired result.

For the second part of the theorem, we have to show that $\Pr(X \in S) = O(\rho^{1/(d+1)})$, where $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$. To do this we observe that $(a_1, \ldots, a_n)$ is $(2\rho^{1/(d+1)}, 2\rho^{d/(d+1)})$-anti-concentrated,

by Lemma 2.5, and then we apply Proposition 4.1. (We remark that here we actually only used the generic intersection property for $\Lambda = \mathbb{R}^d$.) $\qquad\square$

5 An extremal theorem for bounded-complexity hypergraphs

Recall the definition of hypergraph complexity from Definition 1.10. We prove the following Kővári–Sós–Turán-type theorem for hypergraphs of bounded complexity, and deduce Theorem 1.3 from it.

Theorem 5.1. *If H is an m-partite m-uniform hypergraph with complexity M on the vertex sets $V_1, \ldots, V_m$, and there is a vertex weighting function w such that $w(V_j) = 1$ and $w(x) \le \lambda$ for all vertices x, then*
$$w(H) \le C_{M,m} \lambda^{1/2^{m-1}},$$
for some $C_{M,m}$ depending only on M and m.

We remark that there are certain similarities between Theorem 5.1 and Costello, Tao and Vu's decoupling lemma [7, Lemma 6.3].

Proof. We proceed by induction. The result is clearly true for $m = 1$, so, we consider some $m \ge 2$ and assume the result is true for uniformities less than m.

By the definition of hypergraph complexity, H is a union of at most C hypergraphs H_i, in such a way that the common link $N_i(x, y)$ of any $x, y \in V_1$ in H_i is an $(m-1)$-partite $(m-1)$-uniform hypergraph of complexity at most C. So, by the inductive hypothesis $w(N_i(x, y)) = O(\lambda^{1/2^{m-2}})$. Consider the weighted bipartite graph G_i with vertex sets V_1 and $V_2 \times \cdots \times V_m$, with an edge from v_1 to $(v_2, \ldots, v_m)$ if $\{v_1, \ldots, v_m\} \in E(G_i)$. Define the weighting function w' by $w'(v_1) = w(v_1)$ for all $v_1 \in V_1$ and $w'((v_2, \ldots, v_m)) = w(v_2) \ldots w(v_m)$ for all $(v_2, \ldots, v_m) \in V_2 \times \cdots \times V_m$. Then, in G_i, the common neighbourhood of $x, y \in V_1$ has weight exactly $w(N_i(x, y))$. Hence we may apply Theorem 3.1 with $\rho = \lambda$, $q = O(\lambda^{1/2^{m-2}})$, and $t = 2$, to obtain
$$w(H_i) = w'(G_i) = O(\lambda^{1/2^{m-1}}).$$

Summing this over all i gives us the desired bound on $w(H)$. $\qquad\square$

We can use the same ideas as in the proof of Theorem 1.7 to prove two corollaries on probabilities of the form $\Pr(X \in S)$ (the first of which implies Theorem 1.3). For both of these corollaries, we consider nonzero $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1 \xi_1 + \cdots + a_n \xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Recall the definitions of the sum hypergraph $H_{X,S}(I_1, \ldots, I_k)$ and $H_S(V_1, \ldots, V_m)$ from Definition 1.8.

Corollary 5.2. *Suppose $S \subset \mathbb{R}^d$ is such that any m-partite m-uniform sum hypergraph of the form $H_S(V_1, \ldots, V_m)$ has complexity at most M. Let $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$. Then we have*
$$\Pr(X \in S) \le C_{M,m} \rho^{1/(m 2^{m-1})},$$
for some $C_{M,m}$ depending only on M and m.

Proof. By Lemma 2.5, $(a_1, \ldots, a_n)$ is m-part $2\rho^{1/m}$-anti-concentrated. Let $I_1 \sqcup \cdots \sqcup I_m$ be the corresponding partition, and apply Theorem 5.1 to $H_{X,S}(I_1, \ldots, I_m)$, with $\lambda = 2\rho^{1/m}$. $\qquad\square$

Proof of Theorem 1.3. Recall that we showed assumption (a) implies assumption (b). So, we suppose assumption (b) is satisfied, meaning that S includes no Minkowski sum $A_1 + \cdots + A_m$ with each $|A_i| = 2$. Then, sum hypergraphs of the form $H_S(V_1, \ldots, V_m)$ contain no $K_{2,\ldots,2}^{(m)}$, meaning that they have complexity 1. Hence, by Corollary 5.2 we conclude that $\Pr(X \in S) \le C_{1,m} \rho^{1/(m 2^{m-1})}$. $\qquad\square$

Corollary 5.3. *Suppose S has the property that for any affine subspace $\Lambda \subseteq \mathbb{R}^d$ with dimension $\ell = \dim(\Lambda) \leq d$, any sum hypergraph of the form $H_{X,S\cap\Lambda}(I_1,\ldots,I_\ell)$ has complexity at most M. Then*

$$\Pr(X \in S) \leq C_{M,k,d} n^{-d/2^d},$$

for some $C_{M,k,d}$ depending only on M, d.

Proof. We induct on d. In the case $d = 1$, the fact that the 1-uniform hypergraph $H_{X,S}(\{1,\ldots,n\})$ has complexity at most M implies that $|S \cap X| \leq M$, so the desired result follows from the Erdős–Littlewood–Offord Theorem. So, consider some $d \geq 2$, and assume the theorem statement holds in dimensions lower than d.

Let C_d be as in Theorem 2.6, and first suppose $(a_1,\ldots,a_n)$ is d-part λ-anti-concentrated, for $\lambda = C_d|I_j|^{-d/2}$. Then, the desired result follows from Theorem 5.1 with $m = d$.

Otherwise, fixing an arbitrary equipartition $\{1,\ldots,n\} = I_1 \sqcup \cdots \sqcup I_d$, there is some j such that $\rho((a_i)_{i\in I_j}) \geq C_d|I_j|^{-d/2}$. By Theorem 2.6, there is a linear subspace $W \subset \mathbb{R}^d$ of dimension $d-1$ containing half of the a_i in I_1 (say $a_1,\ldots,a_q$ for some $q \geq \lfloor n/d \rfloor/2$). Let $X_1 = a_1\xi_1 + \cdots + a_q\xi_q$ and $X_2 = a_{q+1}\xi_{q+1} + \cdots + a_n\xi_n$, so that

$$\Pr(X \in S) = \Pr(X_1 \in S - X_2) = \Pr(X_1 \in (S - X_2) \cap W).$$

But, if we condition on any outcome of X_2, then applying the induction hypothesis to the set $(S - X_2) \cap W \subseteq W$ (lying inside $W \cong \mathbb{R}^{d-1}$) shows that

$$\Pr(X_1 \in (S - X_2) \cap W) = O\left((\lfloor n/d\rfloor/2)^{-(d-1)/2^{d-1}}\right) = O\left(n^{-d/2^d}\right),$$

implying the desired result. $\square$

6 Sets of points in convex position

In this section we prove Theorem 1.9. We start with the following simple fact about sets of points in convex position.

Lemma 6.1. *There is no finite set of points $Z = A_1 + \cdots + A_{d+1} \subseteq \mathbb{R}^d$, with $|A_i| = 2$, that is in convex position.*

Proof. Consider such a set Z, and for the purpose of contradiction assume its elements are in convex position. We may assume that the affine span of Z is full-dimensional (otherwise we can induct on the dimension, observing that the $d = 0$ case is trivial). First we show that $|Z| = 2^{d+1}$. Indeed, if not, let i be the first index such that $|A_1 + \cdots + A_i| = 2^i$ and $|A_1 + \cdots + A_{i+1}| < 2^{i+1}$. Then writing $A_{i+1} = \{x, y\}$, there exist $z, w \in A_1 + \cdots + A_i$ such that $z + x = y + w$. But then $v_1 = w + x$, $v_2 = z + x = w + y$ and $v_3 = z + y$ are an arithmetic progression (with common difference $z - w = y - x$), and are hence not in convex position. But for any $v \in A_{i+2} + \cdots + A_{d+1}$, we have $v_1 + v, v_2 + v, v_3 + v \in Z$, contradicting that the elements of Z are in convex position.

Now, the convex hull $\text{conv}(Z)$ is a Minkowski sum of intervals $\text{conv}(A_1) + \cdots + \text{conv}(A_{d+1})$, also called a *zonotope*. But it is known (see for example [16]) that any zonotope in d dimensions with m generators has at most $2\sum_{i=0}^{d-1}\binom{m-1}{i}$ vertices, and when $m = d+1$ this tells us that $|Z| \leq 2^{d+1} - 1$, a contradiction. $\square$

To obtain the bounds in Theorem 1.9 we will actually need the following slight variant of Lemma 6.1.

Proposition 6.2. *If $S \subset \mathbb{R}^d$ is a set of points in convex position, then there exists a cover $S = S_1 \cup \cdots \cup S_{2d}$ such that for any subsets $A_1,\ldots,A_d \subset \mathbb{R}^d$ with $|A_i| = 2$, and any S_j, we have*

$$A_1 + \cdots + A_d \not\subseteq S_j.$$

13

Proof. Let $S \subseteq \mathbb{R}^d$ be a set of points in convex position. We need to cover S with $2d$ sets, none of which contains a Minkowski sum $A_1 + \cdots + A_d$ with each $|A_i| = 2$ (call such a Minkowski sum a *bad configuration*).

For each point $x \in S$ let v_x be the outward normal vector to a supporting hyperplane H_x for S. Then $\langle v_x, y - x \rangle < 0$ for every $y \in S \backslash \{x\}$. Let $e_1, \ldots, e_n$ be the standard basis vectors of $\mathbb{R}^d$, and let S_i^+ (respectively S_i^-) be the set of $x \in S$ with $\langle v_x, e_i \rangle > 0$ (respectively, $\langle v_x, e_i \rangle < 0$). Clearly these sets cover S; we will prove that none of them contains a bad configuration.

Without loss of generality, we consider S_1^+. Imagining that e_1 points "downwards", one should think of S_1^+ as being "concave up". Suppose $Z = \{a_1, b_1\} + \cdots + \{a_d, b_d\}$ is a bad configuration. We may assume that Z does not lie in a $(d-1)$-dimensional affine subspace (in which case by Lemma 6.1 it would not be in convex position and would therefore not be a subset of S). It suffices to prove that there is some point $x \in Z$ which is situated vertically above some other point $y' = x + \lambda e_1 \in \text{conv}(Z)$ (here $\lambda > 0$). Indeed, this would imply that $\langle v_x, y' - x \rangle = \lambda \langle v_x, e_1 \rangle > 0$, and hence there is some $y \in Z \subseteq S_1^+$ such that $\langle v_x, y - x \rangle > 0$, contradicting the defining property of v_x.

Let $\pi(Z)$ be the projection of Z onto the hyperplane perpendicular to e_1, so by Lemma 6.1, either some $b_i - a_i$ is parallel to e_1 or else there is some $\pi(z) \in \pi(Z)$ which is a convex combination of other points in $\pi(Z)$. In the former case, the desired fact is obvious, and in the latter case we may take x to be z, if z is above another point of $\text{conv}(Z)$, or the vertex diametrically opposite z in Z, otherwise. $\qquad\square$

We finally prove Theorem 1.9.

Proof of Theorem 1.9. If $S \subseteq \mathbb{R}^d$ is a set of points in convex position, then Proposition 6.2 tells us that we can write $S = S_1 \cup \ldots S_{2d}$ such that any sum hypergraph of the form $H_{S_i}(V_1, \ldots, V_d)$ is $K_{2,\ldots,2}^{(d)}$-free and therefore has complexity 1. The first part of Theorem 1.9 then follows from Corollary 5.3 and the second part follows from Corollary 5.2. $\qquad\square$

7 Hypergraph complexity for definable sets

In this section we prove Proposition 1.13, bounding the complexity of sum hypergraphs associated with a definable set. This implies Theorem 1.14 by Corollary 5.2. As mentioned in the introduction, sets definable with respect to an o-minimal structure have a well-behaved notion of "dimension". There are many equivalent ways of defining dimension; we choose one for concreteness.

Definition 7.1 ([5, Definition 3.14]). *Say a formula ϕ with $d+e$ unbound variables $x_1, \ldots, x_d, y_1, \ldots, y_e$ defines a function $f_\phi : \mathbb{R}^e \to \mathbb{R}^d$ if for every choice of $(x_1, \ldots, x_e) \in \mathbb{R}^e$, there is exactly one choice of $(y_1, \ldots, y_d) \in \mathbb{R}^d$ such that ϕ is true, in which case we set $f_\phi(x_1, \ldots, x_e) = (y_1, \ldots, y_d)$. Then for a definable set $S \subset \mathbb{R}^d$, $\dim S$ is the largest $e \in \mathbb{N}$ such that there is an injective function of the form $f_\phi : \mathbb{R}^e \to \mathbb{R}^d$ with range contained in S (in which case we write $f_\phi : \mathbb{R}^e \hookrightarrow S$).*

Remark 7.2. Applying [5, Proposition 3.17(4)], [5, Corollary 3.16], and [5, Proposition 2.5] shows that we may take f_ϕ to be continuous in the above definition, although we will not need this.

We refer the reader to [5, Section 3.3] for additional basic properties of dimension (such as dimension being well-defined, $\dim \mathbb{R}^d = d$, dimension 0 sets are exactly the finite subsets of $\mathbb{R}^d$, $\dim(X \cup Y) = \max(\dim X, \dim Y)$, etc.)

The following new notion will be essential to the proof.

Definition 7.3. *We say that a definable set T not containing a line segment is* self-irreducible *if $\dim((T - x) \cap (T - y)) < \dim T$ for all distinct $x, y \in \mathbb{R}^d$. We denote by $\mathcal{I}$ the collection of all self-irreducible definable sets.*

Now, Proposition 1.13 is an immediate consequence of the following proposition.

Proposition 7.4. *For any finitely generated o-minimal structure $\mathbb{R}_\mathcal{F}$, there is a function $f : \mathbb{N} \to \mathbb{N}$ such that the following hold.*

1. *Every definable set S with complexity r, not containing a line segment, can be written as a union of at most $f(r)$ self-irreducible sets $T \in \mathcal{I}$, each with complexity at most $f(r)$.*
2. *For each self-irreducible $T \in \mathcal{I}$ with complexity r, and any distinct elements $x, y \in \mathbb{R}^d$, the intersection $(T - x) \cap (T - y)$ is definable, has lower dimension than T, and has complexity at most $f(r)$.*

Actually, the second fact in Proposition 7.4 is trivial by the definition of $\mathcal{I}$. Indeed, we always have $\dim((T - x) \cap (T - y)) < \dim T$, and for the complexity bound we note that if $\phi_T(z)$ is a formula defining T, then $\phi_T(x + z) \wedge \phi_T(y + z)$ is a formula defining $(T - x) \cap (T - y)$.

So, we focus on proving the first of the two facts in Proposition 7.4. We will need one technical lemma about o-minimal sets not containing a line segment, whose proof we defer until later in this section.

Definition 7.5. *Given a definable set $S \subset \mathbb{R}^d$, let $\mathrm{Bad}(S) = \{t \in \mathbb{R}^d \setminus \{\vec{0}\} : \dim(S \cap (S - t)) = \dim S\}$.*

Lemma 7.6. *For any finitely generated o-minimal structure $\mathbb{R}_{\mathcal{F}}$, there is a function $g : \mathbb{N} \to \mathbb{N}$ such that the following holds. Let $S \subset \mathbb{R}^d$ be a definable set with complexity r not containing a line segment. Then $\mathrm{Bad}(S)$ is a finite set, and $|\mathrm{Bad}(S)| \leq g(r)$.*

The proof of Lemma 7.6 requires a few facts about definable sets. For each of these we fix a specific finitely generated o-minimal structure $\mathbb{R}_{\mathcal{F}}$. We will use [5] as our reference for o-minimal geometry.

Remark 7.7. We make a note about our use of "complexity" (which doesn't appear in [5]). Given a formula ϕ in a unbound variables with b real constants, we can create a formula ω in $a + b$ unbound variables where all real constants have been replaced with unbound variables. For a given complexity, there are only finitely many such formulas ω, and the definable set $S_\omega \subset \mathbb{R}^{a+b}$ has the property that for the projection $\pi : \mathbb{R}^{a+b} \to \mathbb{R}^b$ onto the last b coordinates, the fibers are the definable sets S_ϕ for all ϕ corresponding to ω. This then gives a "family of definable sets", which is the context where the theorems in [5] apply. There is no loss of generality in our formulation because every "family of definable sets" is defined by a formula of some complexity, which then bounds the complexity of each fiber of such a family.

Fact 7.8 (The uniform finiteness theorem [5, Theorem 2.9]). *For $A \subset \mathbb{R}^d$ a definable set and a linear projection $L : \mathbb{R}^d \to \mathbb{R}^k$, if all fibers of $A \to L(A)$ are finite, then their sizes are bounded by a function of the complexity of A.*

Corollary 7.9. *If $S \subset \mathbb{R}^d$ is a definable set, then either $|S|$ is bounded above by a function of the complexity of S, or S is uncountably infinite.*

Proof. If S were countably infinite, then a generic projection to a 1-dimensional subspace would be a definable countably infinite set, which contradicts that definable subsets of $\mathbb{R}$ are finite unions of points and intervals.

If $|S|$ is finite, then the complexity bound follows by applying Fact 7.8 to the projection to $\mathbb{R}^0$. $\square$

Corollary 7.10. *If $S \subset \mathbb{R}^d$ is a definable set not containing a line segment, then for every line $\ell \subset \mathbb{R}^d$, $|S \cap \ell|$ is bounded by a function of the complexity of S.*

Proof. First, note that $S \cap \ell$ not containing a line segment implies that $S \cap \ell$ is finite, by the defining property of o-minimality after a generic projection to $\mathbb{R}$. Now, consider the definable set $A = \{(v, w, \lambda) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} : w \neq 0$ and $v + \lambda w \in S\}$ and the linear projection $L : (v, w, \lambda) \mapsto (v, w)$. Then A has complexity bounded as a function of the complexity of S, and the result follows from Fact 7.8. $\square$

Fact 7.11 (Constant-dimension loci are definable [5, Theorem 3.18]). *For a definable set $S \subset \mathbb{R}^d$, a linear projection $L : \mathbb{R}^d \to \mathbb{R}^q$, and any $k \in \mathbb{N}$, the subset $\{b \in L(S) : \dim(L|_S^{-1}(b)) = k\} \subseteq L(S)$ is definable and has complexity bounded by the complexity of S.*

Fact 7.12. *Suppose $X, Y \subset \mathbb{R}^d$ are definable sets and $\dim X \cap Y = \dim X = \dim Y$. Then there is a Euclidean ball $B(x, r) \subset \mathbb{R}^d$ with $x \in X \cap Y$ such that $X \cap B(x, r) = Y \cap B(x, r)$.*

Proof. Set $k = \dim(X) = \dim(X \cap Y) = \dim(Y)$, and let $W = X \cup Y \setminus (X \cap Y)$. Then $\dim(\overline{W} \setminus W) \leq \dim(W) - 1$ by [5, Theorem 3.22], so $\dim(\overline{W} \setminus W) \leq \dim(X \cup Y) - 1 = \max(\dim(X), \dim(Y)) - 1 = k - 1$. Hence $(X \cap Y) \setminus \overline{W} = (X \cap Y) \setminus (\overline{W} \setminus W)$ is nonempty. Since $\overline{W}$ is closed, any $x \in (X \cap Y) \setminus \overline{W}$ has an open ball $B(x, r)$ not intersecting W, meaning $B(x, r) \cap (X \cup Y) \subset X \cap Y$. Therefore $X \cap B(x, r) = (X \cap Y) \cap B(x, r) = Y \cap B(x, r)$ as desired. $\qquad\square$

Now, we prove Proposition 7.4, assuming Lemma 7.6.

Proof of Proposition 7.4. As we have already established the second fact, we only prove the first. We induct on $\dim S$, noting that the theorem is trivial for $\dim S = 0$ because then S is a union of a bounded (in terms of r) number of points by Corollary 7.9. Recalling Lemma 7.6, let $v_1, \ldots, v_s$ be an enumeration of $\mathrm{Bad}(S)$, where $s = |\mathrm{Bad}(S)|$ is bounded in terms of the complexity r of S. Let $L_i : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be the linear projection onto the orthogonal complement of v_i. Then for any $x \in L_i(S)$, the fiber $L_i|_S^{-1}(x) \subset S$ is a definable subset of the line $L_i^{-1}(x)$, and is therefore a finite union of points and intervals. But S does not contain any intervals, so it is a finite union of points.

By the uniform finiteness theorem (Fact 7.8), there is some C_r depending on r such that each $|L_i^{-1}(x) \cap S| \leq C_r$. Then, for $0 \leq i_1, \ldots, i_s \leq C_r$, consider the set

$$S_{i_1, \ldots, i_s} := \{x \in S : |(x + \mathbb{R}_{\geq 0} v_j) \cap S| = i_j \text{ for } 1 \leq j \leq s\}.$$

That is to say, $S_{i_1, \ldots, i_s}$ is the set of points in S for which exactly i_j points of S lie "ahead" of S in the v_j-direction. Note that this set is definable, because the condition $|(x + \mathbb{R}_{\geq 0} v) \cap S| = i$ may be interpreted as the statement that there exist distinct nonnegative real numbers $\lambda_1 = 0, \lambda_2, \ldots, \lambda_i$ such that $x + \lambda_i v \in S$ and there does not exist a non-negative real number λ distinct from $\lambda_1, \ldots, \lambda_i$ such that $x + \lambda v \in S$.

By the choice of C_r, the sets $S_{i_1, \ldots, i_s}$ partition S. If $\dim S_{i_1, \ldots, i_s} < \dim S$ we can use the inductive hypothesis to decompose $S_{i_1, \ldots, i_s}$ into a bounded (in terms of r) number of sets in $\mathcal{I}$. On the other hand, when $\dim S_{i_1, \ldots, i_s} = \dim S$ then we claim that $\mathrm{Bad}(S_{i_1, \ldots, i_s}) = \emptyset$, meaning that we already have $S_{i_1, \ldots, i_s} \in \mathcal{I}$. To see this, note that $\mathrm{Bad}(S_{i_1, \ldots, i_s}) \subseteq \mathrm{Bad}(S) = \{v_1, \ldots, v_s\}$, but by construction $S_{i_1, \ldots, i_s} \cap (S_{i_1, \ldots, i_s} + v_j) = \emptyset$ for all j. $\qquad\square$

It remains to prove Lemma 7.6.

Proof of Lemma 7.6. We first claim that $\mathrm{Bad}(S)$ is definable with complexity bounded in terms of the complexity of S. Consider the definable set $A = \{(x, t) : t \in \mathbb{R}^d, x \in S \cap (S - t)\} \subset \mathbb{R}^d \times \mathbb{R}^d$, and the projection $\pi : A \to \mathbb{R}^d$ defined by $(x, t) \mapsto t$. We can then write $\mathrm{Bad}(S) = \{t \in \pi(A) : \dim \pi^{-1}(t) = \dim(S)\} \subseteq \pi(A)$, so by Fact 7.11, $\mathrm{Bad}(S)$ is definable and has complexity bounded in terms of the complexity of A (which in turn is bounded in terms of the complexity of S). By Corollary 7.9, it follows that either $|\mathrm{Bad}(S)|$ is bounded in terms of the complexity of S, or is uncountably infinite.

So, it suffices to show that $\mathrm{Bad}(S)$ is countable. Say that $x, y \in S$ are *r-isometric* if $(B(x, r) \cap S) + (y - x) = B(y, r) \cap S$, or equivalently $B(x, r) \cap S = B(x, r) \cap (S - (y - x))$. We next claim that for each $t \in \mathrm{Bad}(S)$ there is $x_t \in S$ and $r_t > 0$ such that $x_t + t \in S$ and such that x_t and $x_t + t$ are r_t-isometric.

For any $t \in \mathrm{Bad}(S)$ we have $\dim(S \cap (S - t)) = \dim S = \dim(S - t)$, so by Fact 7.12 we obtain $x_t \in S \cap (S - t)$ and $r_t > 0$ such that $S \cap B(x_t, r_t) = (S - t) \cap B(x_t, r_t)$, meaning that x_t and $x_t + t$ are r_t-isometric, as desired.

Now, to show $\mathrm{Bad}(S)$ is countable, we show that the set of $t \in \mathrm{Bad}(S)$ with $r_t \geq \varepsilon$ is finite for all ε (we may then consider a countable sequence $\varepsilon_1, \varepsilon_2, \ldots$ converging to zero). Let $F_\varepsilon \subset \mathbb{R}^d \times \mathbb{R}^d$ be the definable set

$$F_\varepsilon := \{(x, t) \in \mathbb{R}^d \times (\mathbb{R}^d \setminus 0) : x, x + t \in S \text{ are } \varepsilon\text{-isometric}\} \subset \mathbb{R}^d \times \mathbb{R}^d.$$

Then F_ε contains (x_t, t) for every $t \in \mathrm{Bad}(S)$ with $r_t \geq \varepsilon$. Again, let π be the projection $(x, t) \mapsto t$, so it suffices to show that the definable set $\pi(F_\varepsilon)$ is finite.

For every $t \in \pi(F_\varepsilon)$, let $x_t \in \mathbb{R}^d$ be some element such that $(x_t, t) \in F_\varepsilon$ (we have already specified such x_t when $t \in \mathrm{Bad}(S)$ and $r_t \geq \epsilon$). Suppose for the purpose of contradiction that $\pi(F_\varepsilon)$ is uncountably infinite.

Then for every $\eta < \varepsilon/2$, by the pigeonhole principle there are distinct t_1, t_2 such that $\|x_{t_1} - x_{t_2}\| \le \eta$ and $\|t_1 - t_2\| \le \eta$. We claim that x_{t_1} and $x_{t_1} + (t_1 - t_2)$ are $(\varepsilon - 2\eta)$-isometric. Indeed, note that for any subsets $D_1 \subset B(x_{t_1}, \varepsilon)$ and $D_2 \subset B(x_{t_2}, \varepsilon)$ we have $D_1 \cap S = ((D_1 + t_1) \cap S) - t_1$ and $D_2 \cap S = ((D_2 + t_2) \cap S) - t_2$. Taking $D_1 = B(x_{t_1}, \varepsilon - 2\eta)$ we have

$$B(x_{t_1}, \varepsilon - 2\eta) \cap S = (B(x_{t_1} + t_1, \varepsilon - 2\eta) \cap S) - t_1,$$

and taking $D_2 = B(x_{t_1} + t_1 - t_2, \varepsilon - 2\eta) \subset B(x_{t_2}, \varepsilon)$, we have

$$B(x_{t_1} + t_1 - t_2, \varepsilon - 2\eta) \cap S = (B(x_{t_1} + t_1, \varepsilon - 2\eta) \cap S) - t_2.$$

Adding $t_2 - t_1$ to this last equality and comparing with the previous one, we obtain that x_{t_1} and $x_{t_1} + (t_1 - t_2)$ are $(\varepsilon - 2\eta)$-isometric as desired.

Iteratively applying this fact, we see that $x_{t_1} + \ell(t_1 - t_2) \in S$ for $\ell \in \mathbb{N}$ with $0 \le \ell \le (\varepsilon - 2\eta)/\eta$. Taking $\eta \to 0$, this gives arbitrarily long arithmetic progressions contained in S, contradicting Corollary 7.10. $\square$

8 Semi-algebraic sets

In this section we prove Theorem 1.5, giving a bound of the form $\Pr(X \in S) \le (\log n)^{O(1)}/\sqrt{n}$ when S is semi-algebraic. As discussed in the introduction, the main tool we use in the proof of Theorem 1.5 will be the following theorem of Kane [19] on the total influence (average sensitivity) of polynomial threshold functions.

Theorem 8.1 ([19, Theorem 1.2]). *Every degree-r threshold function f has*

$$\mathrm{AS}(f) \le \sqrt{n}(\log n)^{O(r \log r)} 2^{O(r^2 \log r)}.$$

The application of Theorem 8.1 will be very similar to an argument due to Kane used to prove bounds for the polynomial Littlewood–Offord problem (see [24, Section 3]). We will require two lemmas concerning semi-algebraic sets. The first is a bound on the average sensitivity of the indicator function of a semi-algebraic set.

Lemma 8.2. *There is a constant C_r so that every semi-algebraic set $S \subset \mathbb{R}^d$ of description complexity r has*

$$\mathrm{AS}(1_S) \le C_r \sqrt{n}(\log n)^{C_r}.$$

Proof. Given threshold functions $g_1, \ldots, g_m$ of degree r and a sign pattern $(\varepsilon_1, \ldots, \varepsilon_m) \in \{-1, 1\}^m$, it is easy to see that the function

$$g(x) = \begin{cases} 1 & (g_1(x), \ldots, g_m(x)) = (\varepsilon_1, \ldots, \varepsilon_m) \\ 0 & \text{otherwise} \end{cases}$$

satisfies $\mathrm{AS}(g) \le \sum_{i=1}^{m} \mathrm{AS}(g_i) \le m\sqrt{n}(\log n)^{O(r \log r)} 2^{O(r^2 \log r)}$ by Theorem 8.1. The result follows from applying this to the threshold functions associated to the polynomial conditions exhibiting the degree-complexity r (noting that for a polynomial $f : \mathbb{R}^d \to \mathbb{R}$, the function $1_{f=0}$ is the threshold function for f^2). $\square$

The second is the following very weak Littlewood–Offord-type bound, which we will "boost" in the course of the proof of Theorem 1.5.

Lemma 8.3. *There is some $N_{r,d}$ depending only on r and d such that the following holds. Let $S \in \mathbb{R}^d$ be a semi-algebraic set with description complexity r not containing a line segment. Then if $n \ge N_{r,d}$ we have $\Pr(X \in S) < 1$, where $X = a_1 \xi_1 + \cdots + a_n \xi_n$ for independent Rademacher $\xi_1, \ldots, \xi_n$.*

Proof. This follows from Proposition 1.13 and Corollary 5.3 (we remark that the proof could be greatly simplified in this particular case using the existing machinery of irreducible components for algebraic sets and Bézout's theorem). $\square$

Proof of Theorem 1.5. Let $N = N_{r,d}$, in the notation of Lemma 8.3. We claim that

$$\Pr(X \in S) \le 2^N \sum_{i=1}^{N} \mathrm{Inf}_i(1_S). \tag{1}$$

If we can show this, then by symmetry, the same fact will then hold for the corresponding sum over any set of N indices. Applying this fact to n/N disjoint sets of N indices, it will follow that

$$(n/N)\Pr(X \in S) \le 2^N \sum_{i=1}^{n} \mathrm{Inf}_i(1_S) = 2^N \, \mathrm{AS}(1_S),$$

at which point we can conclude with Lemma 8.2.

To prove Equation (1), let $\mathcal{G}$ be the set of all $\xi \in \{0,1\}^n$ such that $X \in S$, and let $\mathcal{G}'$ be the set of all $\xi \in \{0,1\}^n$ such that the status of the event $X \in S$ changes when we change some ξ_i with $i \le N$. It suffices to prove that $|\mathcal{G}| \le 2^N |\mathcal{G}'|$. In turn, to prove this it suffices to show that for any $x \in \mathcal{G}$ there is some $\xi' \in \mathcal{G}'$ agreeing with ξ on all but the first N coordinates.

To see this, start with some $\xi \in \mathcal{G}$, meaning that $X(\xi) \in S$. By Lemma 8.3, there is some $\xi'' \in S$, agreeing with ξ on all but the first N coordinates, such that $X(\xi'') \notin S$. Now, switch from ξ to ξ'' by flipping bits one-by-one; along the way we must visit some $\xi' \in \mathcal{G}'$. $\qquad \square$

9   A strong bound for definable sets

In this section we prove Theorem 1.1, giving a bound of the form $\Pr(X \in S) \le n^{-1/2+o(1)}$ when S is definable. We start with the following consequence of Theorem 1.14 and the Tao–Vu inverse theorem [42], which gives a strong bound on $\Pr(X \in S)$ unless most of the coefficients a_i lie in a small generalised arithmetic progression with low rank.

Theorem 9.1. *Let $\alpha, C > 0$, and $d \in \mathbb{N}$. Then there are constants $q = q_{d,C,\alpha} \in \mathbb{N}$ and $B = B_{d,C,\alpha} > 0$ such that the following is true.*

Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, write $X = a_1 \xi_1 + \cdots + a_n \xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables, and let $S \subseteq \mathbb{R}^d$ be a set which is definable with respect to a finitely generated o-minimal structure $\mathcal{F}$ and does not contain any line segment. Then for n sufficiently large (in terms of $\mathcal{F}$ and the complexity of S), if $\Pr(X \in S) \ge n^{-C}$, then there is a generalised arithmetic progression Q with rank at most q and size at most n^B such that $a_i \in Q$ for all but at most n^α indices i.

We will also need the following generalisation of Lemma 8.3, in which line segments are allowed, but only in certain directions. In the proof of Theorem 1.1, we will use this to deduce a slight strengthening of Theorem 1.5, which will be applied to a semi-algebraic set $\tilde{S}$ and coefficients $\tilde{a}_i$ obtained by "lifting" $S \subseteq \mathbb{R}^d$ and $a_1, \ldots, a_n \in \mathbb{R}^d$ into a higher-dimensional space $\mathbb{R}^q$.

Proposition 9.2. *Let $W \subseteq \mathbb{R}^q$ be a linear subspace and $\tilde{S} \subseteq \mathbb{R}^q$ a semi-algebraic set of description complexity r such that every line segment in S has direction vector contained in W. Then there is $N_{r,q}$ depending only on r, q such that if $n \ge N_{r,q}$, then for any collection of vectors $\tilde{a}_1, \ldots, \tilde{a}_n$ with $\tilde{a}_i \notin W$, and Rademacher random variables $\xi_1, \ldots, \xi_n$, we have $\Pr(\tilde{a}_1 \xi_1 + \cdots + \tilde{a}_n \xi_n \in \tilde{S}) < 1$.*

Proof. Instead of description complexity, it is convenient to work with complexity in the o-minimal structure $\mathbb{R}_\emptyset = \mathbb{R}_{\mathrm{alg}}$ (by the Tarski–Seidenberg theorem, this is equivalent to working with semi-algebraic description complexity). We show that there is $N_{k,r,q}$ which works for all $\tilde{S} \subset \mathbb{R}^q$ of complexity at most r and dimension at most k, by induction on k.

For the base case $k = 0$, every 0-dimensional semi-algebraic set $\tilde{S}$ is a finite union of points, and the number of points in $\tilde{S}$ is bounded in terms of r and q. So, the existence of suitable $N_{0,r,q}$ follows from the Erdős–Littlewood–Offord theorem. Now, consider $\tilde{S}$ with dimension $k \ge 1$.

Let $\pi : \mathbb{R}^q \to \mathbb{R}^d$ be a linear projection with kernel W. We first claim that $\pi(\mathrm{Bad}(\tilde{S}))$ is finite. Indeed, suppose it is not. Then, as in the proof of Lemma 7.6, there exists $\varepsilon > 0$ and an uncountable collection

of $\tilde{t} \in \tilde{S}$ with $\pi(\tilde{t})$ distinct, for which there are points $\tilde{x}_{\tilde{t}} \in \tilde{S}$ such that $\tilde{x}_{\tilde{t}}$ and $\tilde{x}_{\tilde{t}} + \tilde{t}$ are ε-isometric. Then, again as in the proof of Lemma 7.6, for any $\eta > 0$ we can find an arithmetic progression of length $(\varepsilon - 2\eta)/\eta$ contained in $\tilde{S}$, in some direction $\tilde{t}_1 - \tilde{t}_2$. Since $\pi(\tilde{t}_1) \neq \pi(\tilde{t}_2)$, this direction is not contained in W. Taking $\eta > 0$ arbitrarily small, we have found arbitrarily many points of $\tilde{S}$ contained in a line ℓ not parallel to W. But this is a contradiction, as the number of such points must be bounded as a function of r and q, by the same proof as for Corollary 7.10 (with w constrained not to lie in W).

Let $M = M_{r,q}$ be large enough (in terms of r and q) such that $|\pi(\text{Bad}(\tilde{S}))| \leq M$ (the existence of such an M follows from Corollary 7.9 as $\pi(\text{Bad}(\tilde{S}))$ is finite and definable with complexity bounded in terms of r). Write $X_M = \tilde{a}_1 \xi_1 + \cdots + \tilde{a}_M \xi_M$ and $X' = \tilde{a}_{M+1} \xi_{M+1} + \ldots \tilde{a}_n \xi_n$, so $X = X_M + X'$. Now, given vectors $\tilde{w}_0, \ldots, \tilde{w}_{M+1}$ with $\pi(\tilde{w}_i)$ distinct, we then have

$$\dim \left(\bigcap_{i=0}^{M+1} (\tilde{S} - \tilde{w}_i) \right) < \dim \tilde{S} \leq k.$$

Indeed, by the pigeonhole principle, for some i we have $\tilde{w}_i - \tilde{w}_0 \notin \text{Bad}(\tilde{S})$, meaning that $\bigcap_{i=0}^{M} (\tilde{S} - \tilde{w}_i) \subseteq -\tilde{w}_0 + \left(\tilde{S} \cap (\tilde{S} - (\tilde{w}_i - \tilde{w}_0)) \right)$ has dimension less than $\dim \tilde{S}$. But the random variable $\pi(X_M) = \pi(\tilde{a}_1)\xi_1 + \cdots + \pi(\tilde{a}_M)\xi_M$ attains at least $M + 1$ different values. So, for $X = \tilde{a}_1 \xi_1 + \ldots \tilde{a}_n \xi_n$ to be supported in S, it must be the case that X' is supported in

$$\tilde{S}' := \bigcap_{\tilde{w} \in \text{supp}(X_M)} (\tilde{S} - \tilde{w}),$$

which has dimension at most $k - 1$. The complexity of $\tilde{S}'$ is bounded above by a function r' of r and q, so for this r' we can choose $N_{k,r,q} = M_{r,q} + N_{k-1,r',q}$. $\qquad \square$

The final ingredient we will need for the proof of Theorem 1.1 is the following powerful theorem of Pila. In this statement, "open" is with respect to the Euclidean topology on $\mathbb{R}^q$.

Theorem 9.3 ([30, Theorem 3.5(2)]). *Fix $\beta > 0$, $q \in \mathbb{N}$ and a finitely generated o-minimal structure $\mathbb{R}_{\mathcal{F}}$, and suppose N is sufficiently large in terms of $\beta, q, \mathcal{F}$. For any definable set $\tilde{S} \subseteq \mathbb{R}^q$ of complexity r, there are $s \leq N^\beta$ open subsets $\tilde{U}_1, \ldots, \tilde{U}_s$ of semi-algebraic sets $\tilde{T}_1, \ldots, \tilde{T}_s$, with each $\tilde{U}_i \subseteq \tilde{S}$ and each $\tilde{T}_i$ with description complexity bounded in terms of $\mathcal{F}, r, q$, such that $\{0, \ldots, N\}^q \cap \tilde{S} \subseteq \tilde{U}_1 \cup \cdots \cup \tilde{U}_s$.*

Remark 9.4. The statement of Theorem 9.3 above is not exactly the same as [30, Theorem 3.5(2)]. Indeed, Pila states his theorem in terms of a "basic block family", and technically this statement only implies that the semi-algebraic sets $\tilde{T}_1, \ldots, \tilde{T}_s$ have complexity bounded in terms of r, q, when interpreted as sets definable with respect to $\mathbb{R}_{\mathcal{F}}$. This is not quite enough to guarantee a bound on description complexity (such a bound would follow from the Tarski–Seidenberg theorem only if $\mathcal{F} = \emptyset$). However, in the construction in the proof of [30, Theorem 3.5(2)], one can see directly that $\tilde{T}_1, \ldots, \tilde{T}_s$ have bounded description complexity.

Now we prove Theorem 1.1.

Proof of Theorem 1.1. Let $\mathcal{F}$ be the generators for the o-minimal structure that S is defined with respect to. Applying Theorem 9.1 with $C = 1/2$ and say $\alpha = 1/2$, we immediately have $\Pr(X \in S) \leq 1/\sqrt{n}$ (for large n) unless there is a generalised arithmetic progression Q with rank at most $q = O(1)$ and size at most $n^B = n^{O(1)}$, such that $a_i \in Q$ for at least $n/2$ indices i. So, we may assume there is such a Q, and without loss of generality we may assume $a_1, \ldots, a_{\lfloor n/2 \rfloor} \in Q$. Let $X_1 = \sum_{i \leq \lfloor n/2 \rfloor} a_i \xi_i$ and $X_2 = \sum_{i > \lfloor n/2 \rfloor} a_i \xi_i$, and condition on any outcome of X_2 and let $S' = S - X_2$. For the remainder of the proof our goal is to show that $\Pr(X_1 \in S') \leq n^{-1/2 + o(1)}$. Let $m = \lfloor n/2 \rfloor$ in what follows.

We may assume that Q is homogeneous (having base point $b = 0$), by adding an additional generator if necessary. Consider the projection $\pi : \mathbb{R}^q \to \mathbb{R}^d$ sending the standard basis vectors in $\mathbb{R}^q$ to the generators of Q, and let $\tilde{a}_1, \ldots, \tilde{a}_m \in \{0, \ldots, n^B\}^q$ be such that $\pi(\tilde{a}_i) = a_i$ for each i. Then with $\tilde{X}_1$ the random variable $\tilde{a}_1 \xi_1 + \cdots + \tilde{a}_m \xi_m$, and $\tilde{S}'$ the definable set $\pi^{-1}(S')$, we have

$$\Pr(X_1 \in S') = \Pr(\tilde{X}_1 \in \tilde{S}') = \Pr(\tilde{X}_1 \in \tilde{S}' \cap \{0, \ldots, n^{B+1}\}^q).$$

Now, fix $\varepsilon > 0$. From now on, the implicit constants in asymptotic notation are allowed to depend on $q, r, \mathcal{F}, \varepsilon$. By Theorem 9.3 (with $N = n^{B+1}$ and $\beta = \varepsilon/(2B + 2)$), for some $s \leq n^{\varepsilon/2}$ there are open subsets $\tilde{U}_1, \ldots, \tilde{U}_s$ of semi-algebraic sets $\tilde{T}_1, \ldots, \tilde{T}_s$, with each $\tilde{U}_i \subseteq \tilde{S}'$ and each $\tilde{T}_i$ with description complexity $O(1)$, such that $\tilde{S}' \cap \{0, \ldots, n^{B+1}\}^q \subseteq \tilde{U}_1 \cup \cdots \cup \tilde{U}_s$.

Since each $\pi(\tilde{U}_i) \subseteq S$ and S contains no line segment, any line segment that might be contained in $\tilde{U}_i$ must have direction vector in the kernel W of π. Now, for each i let $\tilde{T}_i'$ be the subset of $\tilde{T}_i$ obtained by removing all line segments with direction vectors not in W. Note that $\tilde{T}_i'$ is itself a semi-algebraic set with complexity $O(1)$: indeed, $\tilde{T}_i'$ can be described by a first-order formula with length $O(1)$, so this follows from the Tarski–Seidenberg theorem. If an open subset of $\tilde{T}_i$ contains a single element of a line segment in $\tilde{T}_i$, then it contains an entire sub-segment. So, $\tilde{U}_i \subseteq \tilde{T}_i'$. Now, using Proposition 9.2 instead of Lemma 8.3, the proof of Theorem 1.5 shows that

$$\Pr(\tilde{X}_1 \in \tilde{U}_i) \leq \Pr(\tilde{X}_1 \in \tilde{T}_i') \leq m^{-1/2+o(1)} = n^{-1/2+o(1)}$$

for each i. Taking the union bound over all i, and assuming n is sufficiently large, we have

$$\Pr(X_1 \in S') = \Pr(\tilde{X}_1 \in \tilde{S}' \cap \{0, \ldots, n^{B+1}\}^q) \leq \sum_{i=1}^{s} \Pr(\tilde{X}_1 \in \tilde{U}_i) \leq sn^{-1/2+o(1)} \leq n^{-1/2+\varepsilon}.$$

We conclude by taking $\varepsilon \to 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

10   Concluding remarks

In this paper we have proposed some natural geometric variants of the Littlewood–Offord problem, and proved some bounds in several cases. Of course, Conjecture 1.2 remains open for many natural choices of S. The simplest open case is where $S \in \mathbb{R}^4$ is the three-dimensional unit sphere in four-dimensional space. It would also be of interest to improve the bounds in Theorem 1.9 for sets of points in convex position: can we at least get bounds of the form $\Pr(X \in S) \leq n^{-1/2+o(1)}$ in this case?

Conjecture 10.1. *Let $S \subseteq \mathbb{R}^d$ be a set of points in convex position, and let $\varepsilon > 0$. If n is sufficiently large in terms of ε, d, then the following holds. Consider nonzero d-dimensional vectors $a_1, \ldots, a_n \in \mathbb{R}^d$, and write $X = a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables. Then*

$$\Pr(X \in S) \leq n^{-1/2+\varepsilon}.$$

The analogue of Conjecture 1.2 also plausibly holds in this case.

Regarding Theorem 1.14, it would be good to understand the best possible bound on $\Pr(X \in S)$ in terms of the point probability $\rho = \max_{x \in \mathbb{R}^d} \Pr(X = x)$. If one were able to prove that $\Pr(X \in S) = O(\rho^{1/k})$, this would actually imply Conjecture 1.2, using Halász' theorem (Theorem 2.6) in a similar way to the proof of Theorem 1.7.

Also, recall that the proof of Theorem 1.5 proceeded via a bound on the average sensitivity of polynomial threshold functions, related to the Gotsman–Linial conjecture. As we discussed in Section 1.1, a full resolution of the Gotsman–Linial conjecture would imply Conjecture 1.2 for affine varieties, but actually the following "bounded" version of the Gotsman–Linial conjecture would suffice, and may be of independent interest.

Conjecture 10.2. *Let $S \subseteq \mathbb{R}^d$ be a semi-algebraic set with description complexity r, and for nonzero coefficients $a_1, \ldots, a_n \in \mathbb{R}^n$ let $f(\xi_1, \ldots, \xi_n)$ be the Boolean function measuring whether $a_1\xi_1 + \cdots + a_n\xi_n \in S$. Then f has average sensitivity at most $C_{r,d}\sqrt{n}$, for some $C_{r,d}$ depending only on r and d.*

Conjecture 10.2 is equivalent to the special case of the Gotsman–Linial conjecture where we only consider polynomials of the form $f \circ \pi$, where f is a d-variable polynomial and $\pi : \mathbb{R}^n \to \mathbb{R}^d$ is a linear projection.

We remark that in the special case of dimension $d = 2$, even though we were able to prove Conjecture 1.2, it is not clear how to prove Conjecture 10.2. It might also be interesting to consider the case where S is a convex or definable set.

It would also be interesting to better understand which sets have the generic intersection property in Definition 1.6. In particular, it would be nice to know whether a random algebraic hypersurface in $\mathbb{R}^d$ (conditioned on not containing a line) has the generic intersection property almost surely. This would imply that Conjecture 1.2 holds for "almost all algebraic hypersurfaces".

References

[1] S. Basu and O. E. Raz, *An o-minimal Szemerédi-Trotter theorem*, Q. J. Math. **69** (2018), no. 1, 223–239.

[2] E. Bombieri and J. Pila, *The number of integral points on arcs and ovals*, Duke Math. J. **59** (1989), no. 2, 337–357.

[3] B. Chapman, *The Gotsman-Linial conjecture is false*, Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2018, pp. 692–699.

[4] A. Chernikov, D. Galvin, and S. Starchenko, *Cutting lemma and Zarankiewicz's problem in distal structures*, Selecta Math. (N.S.) **26** (2020), no. 2, Paper No. 25, 27.

[5] M. Coste, *An introduction to o-minimal geometry*, Istituti editoriali e poligrafici internazionali Pisa, 2000.

[6] K. P. Costello, *Bilinear and quadratic variants on the Littlewood-Offord problem*, Israel J. Math. **194** (2013), no. 1, 359–394.

[7] K. P. Costello, T. Tao, and V. Vu, *Random symmetric matrices are almost surely nonsingular*, Duke Math. J. **135** (2006), no. 2, 395–413.

[8] T. Do, *Zarankiewicz's problem for semi-algebraic hypergraphs*, J. Combin. Theory Ser. A **158** (2018), 621–642.

[9] G. Elekes and L. Rónyai, *A combinatorial problem on polynomials and rational functions*, J. Combin. Theory Ser. A **89** (2000), no. 1, 1–20.

[10] G. Elekes and E. Szabó, *How to find groups? (and how to use them in Erdős geometry?)*, Combinatorica **32** (2012), no. 5, 537–571.

[11] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel J. Math. **2** (1964), 183–190.

[12] P. Erdős, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51** (1945), 898–902.

[13] J. Fox, M. Kwan, and L. Sauermann, *Combinatorial anti-concentration inequalities, with applications*, Math. Proc. Cambridge Philos. Soc. (to appear).

[14] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl, *A semi-algebraic version of Zarankiewicz's problem*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 6, 1785–1810.

[15] P. Frankl and Z. Füredi, *Solution of the Littlewood-Offord problem in high dimensions*, Ann. of Math. (2) **128** (1988), no. 2, 259–270.

[16] K. Fukuda, *Zonotopes*, lecture notes for a course on polyhedral computation, `https://www.cs.mcgill.ca/~fukuda/760B/handouts/expoly3.pdf`, 2002.

[17] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Period. Math. Hungar. **8** (1977), no. 3-4, 197–211.

[18] A. Iosevich, S. Konyagin, M. Rudnev, and V. Ten, *Combinatorial complexity of convex sequences*, Discrete Comput. Geom. **35** (2006), no. 1, 143–158.

[19] D. M. Kane, *The correct exponent for the Gotsman-Linial conjecture*, Comput. Complexity **23** (2014), no. 2, 151–175.

[20] H. W. Kim, C. Maldonado, and J. Wellens, *On graphs and the Gotsman-Linial conjecture for $d = 2$*, arXiv preprint arXiv:1709.06650 (2017).

[21] D. J. Kleitman, *On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors*, Advances in Math. **5** (1970), 155–157 (1970).

[22] J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation. III*, Rec. Math. [Mat. Sbornik] N.S. **12(54)** (1943), 277–286.

[23] J. Matoušek, *Lectures on discrete geometry*, Graduate Texts in Mathematics, vol. 212, Springer-Verlag, New York, 2002.

[24] R. Meka, O. Nguyen, and V. Vu, *Anti-concentration for polynomials of independent random variables*, Theory Comput. **12** (2016), Paper No. 11, 16 pages.

[25] H. Nguyen and V. Vu, *Optimal inverse Littlewood–offord theorems*, Advances in Mathematics **226** (2011), no. 6, 5298–5319.

[26] H. H. Nguyen, *Inverse Littlewood-Offord problems and the singularity of random symmetric matrices*, Duke Math. J. **161** (2012), no. 4, 545–586.

[27] H. H. Nguyen and V. H. Vu, *Small ball probability, inverse theorems, and applications*, Erdős centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 409–463.

[28] J. Pila, *Density of integral and rational points on varieties*, Astérisque (1995), no. 228, 4, 183–187, Columbia University Number Theory Seminar (New York, 1992).

[29] J. Pila and A. J. Wilkie, *The rational points of a definable set*, Duke Math. J. **133** (2006), no. 3, 591–616.

[30] J. Pila, *On the algebraic points of a definable set*, Selecta Math. (N.S.) **15** (2009), no. 1, 151–170.

[31] J. Pila, *O-minimality and the André-Oort conjecture for $\mathbb{C}^n$*, Ann. of Math. (2) **173** (2011), no. 3, 1779–1840.

[32] J. Pila and U. Zannier, *Rational points in periodic analytic sets and the Manin-Mumford conjecture*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. **19** (2008), no. 2, 149–162.

[33] O. E. Raz, M. Sharir, and F. De Zeeuw, *Polynomials vanishing on Cartesian products: the Elekes-Szabó theorem revisited*, Duke Math. J. **165** (2016), no. 18, 3517–3566.

[34] O. E. Raz, M. Sharir, and F. de Zeeuw, *The Elekes-Szabó theorem in four dimensions*, Israel J. Math. **227** (2018), no. 2, 663–690.

[35] O. E. Raz, M. Sharir, and J. Solymosi, *Polynomials vanishing on grids: the Elekes-Rónyai problem revisited*, Amer. J. Math. **138** (2016), no. 4, 1029–1065.

[36] O. E. Raz and Z. Shem-Tov, *Expanding polynomials: a generalization of the Elekes-Rónyai theorem to d variables*, Combinatorica **40** (2020), no. 5, 721–748.

[37] A. Razborov and E. Viola, *Real advantage*, ACM Trans. Comput. Theory **5** (2013), no. 4, Art. 17.

[38] T. Scanlon, *O-minimality as an approach to the André-Oort conjecture*, Around the Zilber-Pink conjecture/Autour de la conjecture de Zilber-Pink, Panor. Synthèses, vol. 52, Soc. Math. France, Paris, 2017, pp. 111–165.

[39] A. Sidorenko, *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), no. 2, 201–204.

[40] T. Tao and V. Vu, *From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices*, Bull. Amer. Math. Soc. (N.S.) **46** (2009), no. 3, 377–396.

[41] T. Tao and V. Vu, *A sharp inverse Littlewood-Offord theorem*, Random Structures Algorithms **37** (2010), no. 4, 525–539.

[42] T. Tao and V. H. Vu, *Inverse Littlewood-Offord theorems and the condition number of random discrete matrices*, Ann. of Math. (2) **169** (2009), no. 2, 595–632.

[43] A. Tarski, *A decision method for elementary algebra and geometry*, University of California Press, Berkeley and Los Angeles, Calif., 1951, 2nd ed.

[44] K. Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), no. 2, 593–634.

[45] L. van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998.

[46] L. van den Dries and C. Miller, *On the real exponential field with restricted analytic functions*, Israel J. Math. **85** (1994), no. 1-3, 19–56.

[47] A. J. Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function*, J. Amer. Math. Soc. **9** (1996), no. 4, 1051–1094.