Geometric Littlewood–Offord problems via lattice point counting

Alexandr Grebennikov* Matthew Kwan†

May 30, 2025

Abstract

Consider nonzero vectors $a_1, \cdots, a_n \in \mathbb{C}^k$, independent Rademacher random variables $\xi_1, \dots, \xi_n$, and a set $S \subseteq \mathbb{C}^k$. What upper bounds can we prove on the probability that the random sum $\xi_1 a_1 + \cdots + \xi_n a_n$ lies in S? We develop a general framework that allows us to reduce problems of this type to counting lattice points in S. We apply this framework with known results from diophantine geometry to prove various bounds when S is a set of points in convex position, an algebraic variety, or a semialgebraic set. In particular, this resolves conjectures of Fox–Kwan–Spink and Kwan–Sauermann.

We also obtain some corollaries for the *polynomial Littlewood–Offord problem*, for polynomials that have bounded *Chow rank* (i.e., can be written as a polynomial of a bounded number of linear forms). For example, one of our results confirms a conjecture of Nguyen and Vu in the special case of polynomials with bounded Chow rank: if a bounded-degree polynomial $F \in \mathbb{C}[x_1, \dots, x_n]$ has bounded Chow rank and "robustly depends on at least b of its variables", then $\mathbb{P}[F(\xi_1, \dots, \xi_n) = 0] \leqslant O(1/\sqrt{b})$. We also prove significantly stronger bounds when F is "robustly irreducible", towards a conjecture of Costello.

1 Introduction

Throughout this paper $\xi_1, \dots, \xi_n$ will always denote a sequence of independent Rademacher random variables (that is, taking values 1 or -1 with probability $1/2$).

In 1943, motivated by their study of random algebraic equations, Littlewood and Offord [26] considered the following question: given a sequence of n nonzero real numbers $c_1, \dots, c_n$, what is the maximum probability that the random variable $\xi_1 c_1 + \ldots + \xi_n c_n$ equals a given value[1]? They proved an upper bound of the form $O(\log n/\sqrt{n})$, which was sharpened by an elegant argument of Erdős [14] to the following precise result (now called the Erdős–Littlewood–Offord theorem):

$$\sup_{z \in \mathbb{R}} \mathbb{P}[\xi_1 c_1 + \ldots + \xi_n c_n = z] \leqslant 2^{-n} \binom{n}{\lfloor n/2 \rfloor} = \big(\sqrt{2/\pi} + o(1)\big) \frac{1}{\sqrt{n}} = O\Big(\frac{1}{\sqrt{n}}\Big). \tag{1}$$

Since then, the Erdős–Littlewood–Offord theorem has been generalised in many different directions, and these results have found applications in a wide variety of different fields (e.g., random matrix theory, the theory of Boolean functions, extremal combinatorics; see [29] for a survey). In this paper we introduce a general approach to attack several different questions in Littlewood–Offord theory of a "geometric" flavour.

*Institute of Science and Technology Austria (ISTA), aleksandr.grebennikov@ist.ac.at.

†Institute of Science and Technology Austria (ISTA), matthew.kwan@ist.ac.at.

Both authors are supported by ERC Starting Grant "RANDSTRUCT" No. 101076777.

[1]Here and for the rest of this introduction, we specialise to the "discrete" form of the Littlewood–Offord problem. (There is also a "small-ball" form, where one assumes that the c_i have absolute value at least 1, and studies the likelihood that $\xi_1 c_1 + \ldots + \xi_n c_n$ falls in a given interval of radius 1.)

1.1 Polynomial Littlewood–Offord problem

First, a natural direction of generalisation is to replace the linear form $\xi_1 c_1 + \ldots + \xi_n c_n$ by a polynomial $F(\xi_1, \ldots, \xi_n)$ of higher degree. This direction was first considered by Rosiński and Samorodnitsky [33], Costello, Tao and Vu [11], and Razborov and Viola [32] in the contexts of Lévy chaos, discrete random matrices and Boolean functions, respectively.

It is widely believed that if F is an n-variable degree-d polynomial that is "robustly nonzero" then a bound analogous to (1) should hold. For example, it was conjectured by Nguyen and Vu (see [27, 32]) that if F has at least bn^{d-1} nonzero coefficients, then[2]

$$\mathbb{P}[F(\xi_1, \ldots, \xi_n) = 0] \leqslant O_d\left(\frac{1}{\sqrt{b}}\right). \tag{2}$$

This is known to hold for $d = 1$ (thanks to the Erdős–Littlewood–Offord theorem) and for $d = 2$ (thanks to recent work of Kwan and Sauermann [25]). For general d, the best available bound (due to Meka, Nguyen and Vu [27], via a result of Kane [24]) falls short of (2) by a factor of $(\log b)^{O_d(1)}$.

If true, the bound in (2) is best-possible: for example, one can see this by considering the polynomial $(x_1 + \cdots + x_n)^d$. However, it is natural to wonder whether one can prove much stronger bounds if one makes assumptions to rule out this kind of example. Indeed, it was conjectured by Costello [10] that (2) can be significantly improved when the polynomial F is "robustly irreducible". Though his original conjecture was recently disproved by Kwan, Sah and Sawhney (see [23, Appendix B]), it is plausible that the following "repaired" version of his conjecture still holds: consider any polynomial $F \in \mathbb{C}[t_1, \ldots, t_n]$ of degree $d \geqslant 2$, and suppose that for any reducible[3] polynomial G of degree at least d, the difference $F - G$ has at least bn^{d-1} nonzero coefficients. Then,

$$\mathbb{P}[F(\xi_1, \ldots, \xi_n) = 0] \leqslant O_{d,\varepsilon}(b^{-1+\varepsilon}).$$

We remark that a slight variation on this "repaired" conjecture was also suggested by Jin, Kwan, Sauermann and Wang [23]; they observed that this bound, if true, would be best possible.

As our first results in this paper, we essentially resolve the above conjectures, under an assumption that F has "bounded complexity". Formally, for a polynomial $F \in \mathbb{F}[t_1, \ldots, t_n]$ of degree d (over some field $\mathbb{F} \subseteq \mathbb{C}$), define its *Chow rank (over $\mathbb{F}$)* to be the minimal number c such that F can be represented as $\sum_{i=1}^{c} P_i$, where each P_i is a product of d (not necessarily homogeneous) linear forms with coefficients in $\mathbb{F}$. One can check that the Chow rank of any homogeneous polynomial of a fixed degree d is "equivalent" to the tensor rank of its coefficient tensor (in the sense that each of them is bounded by some function of the other).

Theorem 1.1. *Let $1 \leqslant b \leqslant n$ and $d, c \geqslant 1$ be integers. Let $F \in \mathbb{C}[t_1, \ldots, t_n]$ be a polynomial of degree d and Chow rank at most c.*

(1) Suppose that F "robustly depends on at least b of its variables", in the sense that it is not possible to make $F(\xi_1, \ldots, \xi_n)$ identically zero by fixing fewer than b of the ξ_i to ± 1 values. (In particular, this holds whenever F has at least bn^{d-1} nonzero degree-d coefficients[4].) Then

$$\mathbb{P}[F(\xi_1, \ldots, \xi_n) = 0] \leqslant O_{d,c}(b^{-1/2}).$$

(2) Suppose that $d \geqslant 2$ and $d \neq 3$. Also, suppose that for any reducible polynomial $G \in \mathbb{C}[x_1, \ldots, x_n]$ of degree at most d, the difference $F - G$ has at least bn^{d-1} nonzero coefficients. Then for any $\varepsilon > 0$,

$$\mathbb{P}[F(\xi_1, \ldots, \xi_n) = 0] = O_{d,c,\varepsilon}(b^{-1+\varepsilon}).$$

[2] Subscripts on asymptotic notation indicate quantities that should be treated as constants.
[3] In this paper, we use the convention that the zero polynomial is reducible.
[4] Note that the total number of terms in F with degree less than d is at most dn^{d-1}, so these terms can be essentially ignored.

We will discuss our proof approach properly later in this introduction, but very briefly: we prove Theorem 1.1 via a connection to diophantine geometry. Specifically, we leverage various known estimates for the number of lattice points on algebraic varieties; the reason for the "$d \neq 3$" exception in Theorem 1.1(2) is that uniform estimates for Heath-Brown's so-called *dimension growth conjecture* are available for affine algebraic varieties of all degrees except 3.

It is worth remarking that related geometric considerations have already played an important role in the resolution of the quadratic Littlewood–Offord problem (i.e., the proof of (2) in the case $d = 2$). Indeed, the first bound of the form $b^{-1/2+o(1)}$, due to Costello [10], was proved via a low-rank/high-rank dichotomy, using geometric techniques (related to the Szemerédi–Trotter theorem) in the low-rank case, and using completely different "decoupling" techniques in the high-rank case. For Kwan and Sauermann's recent $O(1/\sqrt{b})$ bound [25], they also took a geometric point of view for low-rank quadratic polynomials (though their proof did not as cleanly split into a low rank and high rank case). In general, it seems to us (in some very vague sense) that the worst-case behaviour for polynomial anticoncentration is driven by geometric considerations for "low-complexity polynomials", and driven by statistical/mixing considerations for "high-complexity polynomials".

Regarding the assumptions on F in Theorem 1.1: in the earliest work on the polynomial Littlewood–Offord problem, the usual assumption was that F has many nonzero coefficients (as in the Nguyen–Vu conjecture at the start of the section). It was first observed by Razborov and Viola [32][5] that one can state polynomial Littlewood–Offord bounds with a much weaker assumption that F "robustly depends on many variables", as in Theorem 1.1(1). One can attempt to restate (the repaired version of) Costello's conjecture with an analogous type of assumption (namely, that it is not possible to make F reducible by fixing fewer than b variables), but such an assumption leads to a different worst-case bound. Indeed, by considering the polynomial $(x_1 + \cdots + x_{n/2})^d - (x_{n/2+1} + \cdots + x_n)$, it is not hard to see that in this setting we cannot hope for a bound stronger than about $b^{-1+1/(2d)}$. We are able to match this lower bound up to logarithmic factors, as follows.

Theorem 1.2. *Let $2 \leqslant b \leqslant n$ and $d, c \geqslant 1$ be integers, and let $\mathbb{F}$ be a subfield of $\mathbb{C}$. Let $F \in \mathbb{F}[t_1, \dots, t_n]$ be a degree-d polynomial which is irreducible (over $\mathbb{F}$) and has Chow rank at most c (over $\mathbb{F}$). Suppose that F remains an irreducible degree-d polynomial (over $\mathbb{F}$) after any substitution of ± 1 values into fewer than b of its variables. Then*

$$\mathbb{P}[F(\xi_1, \dots, \xi_n) = 0] \leqslant O_{d,c}(b^{-1+1/(2d)} (\log b)^{C_{d,c}}),$$

for some constant $C_{d,c}$ depending only on d and c.

We remark that Theorem 1.2 also illustrates that our methods are applicable for polynomials that are reducible over $\mathbb{C}$ but irreducible over a smaller field $\mathbb{F} \subseteq \mathbb{C}$ (provided that certain necessary lattice point enumeration estimates are available).

1.2 Littlewood–Offord problem for algebraic varieties

It turns out that the polynomial Littlewood–Offord problem, under a bounded Chow rank assumption, can be naturally interpreted as a geometric problem in low-dimensional space. Indeed, recall that a polynomial $F \in \mathbb{F}[t_1, \dots, t_n]$ of degree d and Chow rank c can be represented as $P_1 + \cdots + P_c$, where each P_i is a product of d (not necessarily homogeneous) linear forms. Alternatively, one can write this as

$$F(t_1, \dots, t_n) = f(L_1(t_1, \dots, t_n), \dots, L_k(t_1, \dots, t_n)).$$

for $k = dc$, some homogeneous linear forms $L_1, \dots, L_k$ with coefficients in $\mathbb{F}$, and a polynomial $f \in \mathbb{F}[x_1, \dots, x_k]$. (In fact, the polynomial f obtained this way has a certain specific form, but this turns out not to be useful for us.) Now, if we write $a_{ij} \in \mathbb{F}$ for the coefficient of t_j in $L_i(t_1, \dots, t_n)$, and let $a_j = (a_{1j}, \dots, a_{kj}) \in \mathbb{F}^k$, then the event that $F(\xi_1, \dots, \xi_n) = 0$ can be interpreted as the event that $\xi_1 a_1 + \ldots + \xi_n a_n$ lies in the algebraic variety $S = \{x \in \mathbb{C}^k : f(x) = 0\} \subseteq \mathbb{C}^k$. This observation was implicitly leveraged in [10, 25].

[5] The Razborov–Viola assumption only took multilinear degree-d terms into account; the specific assumption in Theorem 1.1(1) was first considered by Kwan and Sauermann [25].

In connection with their work on the polynomial Littlewood–Offord problem, Kwan and Sauermann made a general conjecture along these lines [25, Conjecture 12.1]. Specifically, they conjectured that if one can form at least b disjoint bases of $\mathbb{C}^k$ from the vectors $a_1, \dots, a_n \in \mathbb{C}^k$ (this is a measure of how "robustly" $a_1, \dots, a_n$ span the space $\mathbb{C}^k$), then for any affine algebraic variety $S \in \mathbb{C}^k$ of dimension ℓ and degree d, we have

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant O_{d,k}(b^{-(k-\ell)/2}).$$

We remark that the $d = 1$ case of this conjecture is essentially equivalent to a classical theorem of Halász [18], and in the course of their resolution of the quadratic Littlewood–Offord problem (for quadratic polynomials of not necessarily bounded Chow rank), Kwan and Sauermann proved this conjecture for quadrics inside affine-linear subspaces [25, Theorem 4.2]. We prove this conjecture in full generality.

Theorem 1.3. *Let $0 \leqslant \ell \leqslant k$ and $d \geqslant 1$ be integers. Let $S \subseteq \mathbb{C}^k$ be a (possibly reducible) affine algebraic variety of dimension at most ℓ and degree at most d. Consider vectors $a_1, \dots, a_n \in \mathbb{C}^k$, and assume that one can form b disjoint bases from them. Then*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant O_{d,k}\Big(b^{-(k-\ell)/2}\Big).$$

In fact, Theorem 1.3 is deduced as a corollary of the following more refined estimate.

Theorem 1.4. *Let $0 \leqslant \ell \leqslant k$ and $d \geqslant 2$ be integers. Let $S \subseteq \mathbb{C}^k$ be an irreducible affine algebraic variety of dimension ℓ and degree d. Consider vectors $a_1, \dots, a_n \in \mathbb{C}^k$, and assume that one can form $b \geqslant 2$ disjoint bases from them. Then*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant O_{d,k}\Big(b^{-(k-\ell+1-\frac{1}{d})/2}(\log b)^{C_{d,k}}\Big),$$

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

We remark that although these results are stated over the field of complex numbers, one can deduce similar results for its smaller subfields (this is necessary for the proof of Theorem 1.2). Indeed, for a *real* algebraic set $S_{\mathbb{R}} \subseteq \mathbb{R}^k$, one can consider its "complexification" $S_{\mathbb{C}} \subseteq \mathbb{C}^k$ (that is, the minimal complex variety containing $S_{\mathbb{R}}$). It satisfies $S_{\mathbb{C}} \cap \mathbb{R}^k = S_{\mathbb{R}}$, and the (complex) dimension of $S_{\mathbb{C}}$ coincides with the (real) dimension of $S_{\mathbb{R}}$.

The main ingredient in the proof of Theorem 1.4 is a general theorem (Theorem 8.1) estimating probabilities of the form $\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S]$ in terms of a certain "lattice point density" of S. Theorem 1.4 in particular is proved using an estimate of Pila [30, 31], but in general one can plug other number-theoretic results about counting lattice points on varieties into Theorem 8.1 to obtain analogous results (all the theorems mentioned so far are proved in this way). The following example illustrates why the connection to number theory is not surprising in this context.

Example 1.5. Consider the standard basis vectors $e_1, \dots, e_k$ of $\mathbb{C}^k$, and consider the sequence of vectors $a_1, \dots, a_{2mk}$ consisting of $2m$ copies of $e_i/2$ for each $i \in \{1, \dots, k\}$ (where m is sufficiently large with respect to k). Then each coordinate of the random variable $X := \xi_1 a_1 + \dots + \xi_{2mk} a_{2mk}$ is equal to $t \in \mathbb{Z}$ with probability $\binom{2m}{m+t}/2^{2m}$, independently. Thus, by standard estimates on binomial coefficients, X is essentially equidistributed (up to a multiplicative constant factor depending on k) over the integer points of the box $[-\sqrt{m}, \sqrt{m}]^k$. Therefore, the probability that X lies in a variety $S \subseteq \mathbb{C}^k$ is closely related to the proportion of integer points in this box that lie on S.

We emphasise that the assumption in Theorem 1.4 only guarantees that the vectors $a_1, \dots, a_n$ "robustly span $\mathbb{C}^k$" as a vector space. In particular, they may be very far from resembling the standard generators of the integer grid. The main goal of this paper is to establish a connection between Littlewood–Offord-type questions and counting lattice points on varieties in this general setting.

1.3 Littlewood–Offord problem for general sets

In the above subsection, we have been discussing probabilities of the form $\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S]$, where S is an algebraic variety. It is natural to wonder whether one can obtain similar upper bounds with more general (or completely different) assumptions on S: what are the geometric properties of a set S which ensure that random sums are unlikely to fall in them?

This general direction was recently initiated by Fox, Kwan and Spink [16], who investigated several very general conditions on S: namely, the condition that S is a set of points in convex position (i.e., no point in S can be represented as a convex combination of the others), and the condition that S is "definable with respect to an o-minimal structure" (this is a very general model-theoretic notion which ensures that S does not have "infinitely oscillating" structure).

First, we discuss the case when S is a set of points in convex position (which includes, in particular, boundaries of strictly convex bodies). In this setting, Fox, Kwan and Spink proved that for any nonzero vectors $a_1, \dots, a_n$ in $\mathbb{R}^k$ the probability that $\xi_1 a_1 + \dots + \xi_n a_n$ lies in S is at most $O_k(n^{-k/2^k})$ [16, Theorem 1.9(1)], and conjectured that the stronger bound $n^{-1/2+o_k(1)}$ should hold [16, Conjecture 10.1]. We show that this is indeed the case, and provide an asymptotically sharp bound.

Theorem 1.6. *Let $S \subseteq \mathbb{R}^k$ be a set of points in convex position. Consider arbitrary nonzero vectors $a_1, \dots, a_n \in \mathbb{R}^k$. Then, as k is fixed and n tends to infinity,*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant \Big(2\sqrt{2/\pi} + o_k(1)\Big) n^{-1/2}.$$

The following example shows that this bound is asymptotically sharp: let $a \in \mathbb{R}^k$ be an arbitrary nonzero vector, and take $S = \{-a, a\}$, $a_1 = a_2 = \dots = a_{2n+1} = a$. However, this example is essentially one-dimensional. We obtain a stronger bound under the assumption that the vectors $a_1, \dots, a_n$ "robustly span $\mathbb{R}^k$" for $k \geqslant 2$ (and use it to deduce Theorem 1.6).

Theorem 1.7. *Let $S \subseteq \mathbb{R}^k$ be a set of points in convex position. Consider vectors $a_1, \dots, a_n \in \mathbb{R}^k$, and assume that one can form $b \geqslant 2$ disjoint bases from them. Then*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant O_k\big(b^{-1+1/(k+1)} (\log b)^{C_k}\big)$$

for some constant C_k depending only on k.

In the current work, we do not pursue the most general situation when $S \subseteq \mathbb{R}^k$ is a set "definable with respect to an o-minimal structure"[6]. Instead, we highlight the special case of semialgebraic sets: sets defined by a collection of polynomial equations and inequalities. When S is a semialgebraic set which does not contain a line segment, Fox, Kwan and Spink proved that $\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant n^{-1/2} (\log n)^{O_k(1)}$ [16, Theorem 1.5]. We observe that our approach allows us to remove the logarithmic factor.

Theorem 1.8. *Let $S \subseteq \mathbb{R}^k$ be a semialgebraic set, which does not contain a line segment. Consider arbitrary nonzero vectors $a_1, \dots, a_n \in \mathbb{R}^k$. Then*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_n a_n \in S] \leqslant O_S(n^{-1/2}).$$

This bound is sharp up to a multiplicative constant factor, but can be further improved under the assumption that the vectors $a_1, \dots, a_n$ "robustly span a high-dimensional subspace", see Remark 8.5.

[6] It does seem to be possible to adapt our methods to this setting by combining them with the tools from [16], but this would require us to introduce and explain various concepts from o-minimal geometry, which are outside the scope of the present paper.

1.4 Organization of the paper

This paper is organized as follows. In Section 2 we provide a high-level outline of the proofs of Theorems 1.4 and 1.7. In Section 3 we introduce notation that will be used throughout the paper, and prove several preliminary lemmas. In Section 4, we prove a convenient intermediate result (Theorem 4.1) that reduces the case of "polynomial point probabilities" to lattice point counting. Next, in Section 5, we present the proofs of the results for sets of points in convex position (Theorems 1.6 and 1.7).

For the rest of the paper (Sections 6, 7 and 8) we turn to the algebraic setting. In Section 6 we review several useful ingredients from algebraic geometry and number theory. Section 7 contains our most technical theorem (Theorem 7.1), which provides a decomposition of the ambient vector space into subspaces via an iterative decoupling procedure. Finally, in Section 8 we prove our general result for the algebraic setting (Theorem 8.1), and use it to deduce Theorems 1.1, 1.2, 1.3, 1.4 and 1.8.

Basic notation. For a positive integer n, we write $[n]$ to denote the set $\{1, \dots, n\}$. Our use of asymptotic notation is standard: for functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leqslant C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f(n) \geqslant c|g(n)|$ for sufficiently large n, $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and $f = o(g)$ to mean that $f/g \to 0$ as $n \to \infty$. Subscripts on asymptotic notation indicate quantities that should be treated as constants. All logarithms are assumed to be in base 2.

Acknowledgements. We would like to thank Tim Browning and Matteo Verzobio for helpful comments and references from the number theory literature.

2 Proof outline

In this section we provide a high-level sketch of the proofs of Theorem 1.7 for sets of points in convex position and Theorem 1.4 for algebraic varieties (all our other main results are either deduced from one of these two theorems, or from the intermediate lemmas in their proofs).

The proofs of Theorem 1.7 and Theorem 1.4 are based on similar ideas, though the proof of Theorem 1.7 is simpler.

2.1 Sets of points in convex position

To prove Theorem 1.7 we need to obtain an upper bound on the probability that $X := \xi_1 a_1 + \ldots + \xi_n a_n$ lies in our set S of points in convex position. We separately treat the cases when

$$\rho := \sup_{x \in \mathbb{R}^k} \mathbb{P}(X = x) < n^{-C}$$

and when $\rho \geqslant n^{-C}$ (for some appropriately chosen C depending on the dimension k).

The "spread-out" case ($\rho < n^{-C}$). In this case, we just apply a result of Fox, Kwan and Spink [16, Theorem 1.9(2)] which implies that

$$\mathbb{P}[X \in S] \leqslant O_k(\rho^{1/(k2^{k-1})}).$$

That is to say, the probability of lying in S can be bounded in terms of the maximum *point* concentration probability. This directly implies the conclusion of Theorem 1.7, if C is sufficiently large. The proof of the above bound (in [16]) is based on a reduction to a Kővári–Sós–Turán-type theorem in an auxiliary hypergraph, and some simple combinatorial consequences of the fact that the points in S lie in convex position.

The "concentrated" case ($\rho \geqslant n^{-C}$). In this second case, we take advantage of the *inverse theory* for the linear Littlewood–Offord problem. Roughly speaking, the philosophy of this theory is that the only way for ρ to be large is for the coefficients $a_1, \dots, a_n$ to have strong additive structure.

Specifically, our main tool will be the "optimal inverse theorem" for the linear Littlewood–Offord problem, proved by Nguyen and Vu ([28, Theorem 2.5], stated below as Theorem 4.3). It says that if $\rho \geqslant n^{-C}$, then almost all of the vectors $a_1, \dots, a_n$ are contained in a common *generalized arithmetic progression* ("GAP" for short; see Definition 4.2), whose rank is bounded in terms of C, and whose volume depends in an "optimal" way on ρ.

In Theorem 4.4, we show how to *iterate* this optimal inverse theorem, to prove that in fact the random variable $X = \xi_1 a_1 + \ldots + \xi_n a_n$ is *approximately equidistributed* in a certain GAP of bounded rank. In other words, if we eliminate[7] a few "exceptional" vectors a_i, then up to an affine-linear transformation, we can think of X as being approximately the uniform distribution on the integer points in a box of the form $[-B_1, B_1] \times \cdots \times [-B_q, B_q]$.

As a result, the problem of upper-bounding $\mathbb{P}[X \in S]$ reduces to the problem of upper-bounding the number of integer points in a box $[-B_1, B_1] \times \cdots \times [-B_q, B_q]$ which lie in a certain affine-linear transformation of S. For this, we can take advantage of classical estimates in discrete geometry (in particular, we use an estimate due to Andrews [1], stated in this paper as Theorem 5.1).

The full details of the proof of Theorem 1.7 appear in Section 5.

2.2 Algebraic varieties

To prove Theorem 1.4 we need to obtain an upper bound on the probability that $X := \xi_1 a_1 + \ldots + \xi_n a_n$ lies in our algebraic variety S.

If we try to proceed via the same dichotomy as for the proof of Theorem 1.7, the "concentrated" case ($\rho \geqslant n^{-C}$) works in exactly the same way: the only change is that Andrews' theorem should be replaced by a result of Pila [30, 31] (stated in this paper as Theorem 6.7), counting integer points on an affine algebraic variety.

Unfortunately, we encounter some issues in the "spread-out" case ($\rho < n^{-C}$). Recall that for Theorem 1.7 we used a result of Fox, Kwan and Spink bounding $\mathbb{P}[X \in S]$ in terms of ρ. Fox, Kwan and Spink also proved a similar result that can be applied to algebraic varieties ([16, Theorem 1.14]), but it requires the additional assumption that the variety S does not contain any affine lines. In general, without such an assumption on S there is no nontrivial bound[8] on $\mathbb{P}[X \in S]$ in terms of ρ.

Therefore, we take a different point of view. Instead of separately considering two extreme cases, our argument can be seen as an *interpolation* between these two cases. Namely, in Theorem 7.1 we obtain a decomposition of the ambient vector space $\mathbb{C}^k$ into a direct sum $U \oplus W$ of a "disordered" subspace U and a "structured" subspace W (where the decomposition is chosen with respect to the sequence of vectors $a_1, \dots, a_n$ and the variety S).

Roughly speaking, the property we will guarantee for our "structured" subspace W is that, after eliminating a few "exceptional" vectors a_i, the *projection of X onto W* concentrates on some point with polynomially large probability (at least n^{-C} for some constant C). Let $\pi_W : \mathbb{C}^k = U \oplus W \to W$ be the projection map, and let S' be the maximal subset of W such that we have $\pi_W^{-1}(S') \subseteq S$. Then X lies in $\pi_W^{-1}(S')$ if and only if its projection $\pi_W(X)$ lies in S', so $\mathbb{P}\big[X \in \pi_W^{-1}(S')\big]$ can be estimated using the same approach as for the "concentrated" case described in the previous subsection (replacing Andrews' theorem with Pila's theorem, as described at the beginning of this subsection).

By construction of S', knowing the value of the projection $\pi_W(X)$ *cannot* allow us to conclude that X lies in $S \setminus \pi_W^{-1}(S')$: this always depends on the "disordered" coordinate of X (corresponding to the subspace U) as well. So, the property we will guarantee for our "disordered" subspace U is simply that X is very unlikely to lie in $S \setminus \pi_W^{-1}(S')$. Together with the above considerations, this gives the desired upper bound on the probability that X lies in S.

[7] To "eliminate" a vector a_i just means to fix an outcome of the corresponding random variable ξ_i, and work in the resulting conditional probability space.

[8] For example, suppose that S is the line $\{(x, y) : x = 0\} \subseteq \mathbb{R}^2$, and suppose $a_1, \dots, a_n \in \mathbb{R}^d$ are defined by $a_i = (1, 2^i)$. Then $\xi_1 a_1 + \ldots + \xi_n a_n$ lies in S with probability $\Theta(n^{-1/2})$, while ρ is exponentially small.

The proof of Theorem 7.1 is by an iterative procedure: we begin with $U = \mathbb{C}^k$, $W = \{0\}$, and then repeatedly enlarge W while keeping it "structured" (shrinking U correspondingly). At each step of this procedure, we use a *decoupling* argument (Lemma 7.2), which relates the probability that $X = \xi_1 a_1 + \ldots + \xi_n a_n$ lies in a variety S with the probability that it lies in certain linear subspaces. We refer the reader to the discussion in Section 7 for more details.

3 Notation and preliminaries

Let $A = (a_1, \ldots, a_n)$ be a sequence of vectors in a finite-dimensional vector space V (over some subfield $\mathbb{F}$ of $\mathbb{C}$). We note that the order of the vectors $a_1, \ldots, a_n$ is irrelevant for us in this work, and the word "sequence" is used as a synonym for the word "multiset".

For a subset $I \subseteq [n]$, $I = \{i_1, \ldots, i_m\}$ we define the subsequence $A[I] = (a_{i_1}, \ldots, a_{i_m})$. We also say that A' is a subsequence of A of size m if there exists $I \subseteq [n]$, $|I| = m$ such that $A' = A[I]$.

We define the *basis packing number* of a sequence A to be the maximum number of disjoint bases of V one can form from the vectors of A. Formally, the basis packing number of A is the largest integer b for which there exist b pairwise disjoint subsets $I_1, \ldots, I_b \subseteq [n]$ such that for each $1 \leqslant j \leqslant b$ the subsequence $A[I_j]$ is a basis of V.

Definition 3.1. Define the maximum point probability $\rho(A)$ by

$$\rho(A) = \sup_{x \in V} \mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n = x].$$

More generally, for any set $S \subseteq V$ we define the maximum S-translate probability $\rho(A, S)$ by

$$\rho(A, S) = \sup_{x \in V} \mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S + x],$$

where $S + x = \{s + x : s \in S\}$.

In these terms, $\rho(A) = \rho(A, \{0\})$. This general definition turns out to be convenient for us, due to the following simple properties.

Fact 3.2. *Let S be a subset of V. If A' is a subsequence of A then $\rho(A, S) \leqslant \rho(A', S)$.*

Proof. Let $A' = A[I]$, and fix $x \in V$. Denote $X = \sum_{i \in I} \xi_i a_i$, $Y = \sum_{i \notin I} \xi_i a_i$. Conditioning on the outcome of Y, we have

$$\mathbb{P}[X + Y \in S - x] = \mathbb{E}_Y \Big[\mathbb{P}[X \in S - x - Y \mid Y]\Big] \leqslant \sup_{z \in V} \mathbb{P}[X \in S - z] = \rho(A', S). \qquad \square$$

Fact 3.3. *Let S_1, S_2 be two subsets of V. Then*

$$\max(\rho(A, S_1), \rho(A, S_2)) \leqslant \rho(A, S_1 \cup S_2) \leqslant \rho(A, S_1) + \rho(A, S_2).$$

Proof. The first inequality holds trivially. For the second one, denote $X = \xi_1 a_1 + \ldots + \xi_n a_n$. Then for any $x \in V$

$$\mathbb{P}[X \in (S_1 \cup S_2) - x] = \mathbb{P}[X \in (S_1 - x) \cap (S_2 - x)] \leqslant \sup_{x \in V} \mathbb{P}[X \in S_1 - x] + \sup_{x \in V} \mathbb{P}[X \in S_2 - x]. \qquad \square$$

Fact 3.4. *Let $\pi : V \to U$ is a surjective linear map. Then for any sequence $A = (a_1, \ldots, a_n)$ of vectors in V, and any set $S \subseteq U$, we have*

$$\rho(\pi(A), S) = \rho(A, \pi^{-1}(S)).$$

Proof. One can verify that for any outcomes of $\xi_1, \ldots, \xi_n$ and any $x \in V$ we have

$$\xi_1\pi(a_1) + \ldots + \xi_n\pi(a_n) + \pi(x) \in S \quad \text{if and only if} \quad \xi_1 a_1 + \ldots + \xi_n a_n + x \in \pi^{-1}(S).$$

The statement then follows by taking the supremum over $x \in V$ of the probabilities of both these events. □

We also record the following "dropping to a subspace" lemma, which also appeared in [25] (in a slightly different form). Roughly speaking, it says that any sequence of nonzero vectors in a vector space V contains a large subsequence, which has linear basis packing number inside a possibly smaller subspace $V' \subseteq V$.

Lemma 3.5. *Let n, k and b be positive integers, such that $n - (b-1)k(k+1)/2 > 0$. Consider a sequence of nonzero vectors $A = (a_1, \ldots, a_n)$ in a vector space V of dimension k. Then there exists a linear subspace $V' \subseteq V$, and a subsequence A' of A of size at least $n-(b-1)k(k+1)/2$, such that all elements of A' lie in V', and the basis packing number of A' (as a sequence in V') is at least b.*

Proof. We argue by induction on k. In the case $k = 1$ each of $n \geqslant b$ vectors forms a one-element basis.

Let b' be the basis packing number of A. If $b' \geqslant b$, then we are done. Otherwise, let $I_1, \ldots, I_{b'}$ be the disjoint sets of indices corresponding to $b' \leqslant b-1$ bases. Then the subsequence $A_0 = A[[n] \setminus (I_1 \cup \ldots \cup I_{b'})]$ of size $n - b'k$ does not contain a basis of V. Therefore, its vectors lie in a linear subspace $V_0 \subsetneq V$ of dimension $k-1$. As

$$n - b'k - (b-1)\frac{(k-1)k}{2} \geqslant n - (b-1)\frac{k(k+1)}{2} > 0,$$

we can apply the induction hypothesis to the sequence A_0 in the vector space V_0 to obtain the desired subsequence. □

Note that $\mathbb{F}$ (which is a subfield of $\mathbb{C}$) contains the integers. Much of our analysis will focus on counting lattice points within subsets of $\mathbb{F}^k$, so we introduce the following notation.

Definition 3.6. For a set $S \subseteq \mathbb{F}^k$ and a real number $B \geqslant 0$, we define the integer point counting function as

$$N_S(B) = |\{(x_1, \ldots, x_k) \in \mathbb{Z}^k \cap S : |x_i| \leqslant B \text{ for } 1 \leqslant i \leqslant k\}|,$$

and the integer points density function as

$$d_S(B) = \sup_{\varphi}\left(\frac{N_{\varphi(S)}(B)}{|\mathbb{Z}^k \cap [-B, B]^k|}\right) = \sup_{\varphi}\left(\frac{N_{\varphi(S)}(B)}{(2\lfloor B\rfloor + 1)^k}\right),$$

where the supremum is taken over all bijective affine-linear maps $\varphi : \mathbb{F}^k \to \mathbb{F}^k$.

Although slightly non-standard, this definition of $d_S(B)$ is convenient for our purposes. Note that $d_S(B)$ is invariant under bijective affine-linear transformations of S. Therefore, it does not depend on the choice of basis in $\mathbb{F}^k$, and makes sense for a set S in an abstract finite-dimensional vector space V over $\mathbb{F}$.

Furthermore, we observe that it is also invariant under taking preimages of projections.

Proposition 3.7. *Let $\psi : \mathbb{F}^r \to \mathbb{F}^k$ be a surjective affine-linear map. Then for any set $S \subseteq \mathbb{F}^k$ and any $B \geqslant 0$ we have*

$$d_{\psi^{-1}(S)}(B) = d_S(B).$$

Proof. By replacing B with its integer part $\lfloor B\rfloor$, we can assume that B is a non-negative integer.

First, we prove the inequality $d_{\psi^{-1}(S)}(B) \geqslant d_S(B)$. Given a bijective affine-linear map $\varphi_1 : \mathbb{F}^k \to \mathbb{F}^k$, there exists a bijective affine-linear map $\varphi_2 : \mathbb{F}^r \to \mathbb{F}^r$ such that $\varphi_1 \circ \psi \circ \varphi_2 = p$, where $p : \mathbb{F}^r \to \mathbb{F}^k$ is the projection onto the first k coordinates. Then

$$\varphi_2^{-1}(\psi^{-1}(S)) = p^{-1}(\varphi_1(S)),$$

and therefore

$$(2B+1)^r d_{\psi^{-1}(S)}(B) \geqslant N_{\varphi_2^{-1}(\psi^{-1}(S))}(B) = N_{p^{-1}(\varphi_1(S))}(B) = (2B+1)^{r-k} N_{\varphi_1(S)}(B).$$

Taking the supremum over φ_1 gives the desired inequality.

Next, we prove the converse inequality $d_{\psi^{-1}(S)}(B) \leqslant d_S(B)$. Let $e_1, \dots, e_r$ be the standard basis vectors of $\mathbb{F}^r$. Consider a bijective affine-linear map $\varphi_2 : \mathbb{F}^r \to \mathbb{F}^r$. The kernel of $\psi \circ \varphi_2$ has dimension $r-k$, thus we can choose k standard basis vectors $e_{j_1}, \dots, e_{j_k}$ such that the subspace W spanned by them has trivial intersection with this kernel.

Then the restriction of $\varphi_2 \circ \psi$ to each translate of W is a bijective affine-linear map. We consider "slices" of the box $[-B, B]^r$ by the translates of W, and bound the number of integer points on each of them in terms of $d_S(B)$.

Let $J = \{j_1, \dots, j_k\} \subseteq [r]$. Then

$$\begin{aligned}
N_{\varphi_2^{-1}(\psi^{-1}(S))}(B) &= \left|\left\{(c_j)_{j\in[r]} \in \mathbb{Z}^r : |c_j| \leqslant B,\ \sum_{j\in[r]} c_j e_j \in \varphi_2^{-1}(\psi^{-1}(S))\right\}\right| \\
&= \sum_{\substack{(c_j)_{j\in[r]\setminus J}\in\mathbb{Z}^{r-k},\\ |c_j|\leqslant B}} \left|\left\{(c_j)_{j\in J} \in \mathbb{Z}^k : |c_j| \leqslant B,\ \sum_{j\in J} c_j e_j \in W \cap \left(\varphi_2^{-1}(\psi^{-1}(S)) - \sum_{j\in[r]\setminus J} c_j e_j\right)\right\}\right| \\
&= \sum_{\substack{(c_j)_{j\in[r]\setminus J}\in\mathbb{Z}^{r-k},\\ |c_j|\leqslant B}} \left|\left\{(c_j)_{j\in J} \in \mathbb{Z}^k : |c_j| \leqslant B,\ \sum_{j\in J} c_j e_j \in \varphi_{(c_j)}^{-1}(S)\right\}\right|,
\end{aligned}$$

where $\varphi_{(c_j)} : W \to \mathbb{F}^k$ is the bijective affine-linear map defined by $\varphi_{(c_j)}(w) = \psi(\varphi_2(w + \sum_{j\in[r]\setminus J} c_j e_j))$. Recalling the definition of the density function d_S, we conclude that

$$N_{\varphi_2^{-1}(\psi^{-1}(S))}(B) \leqslant \sum_{\substack{(c_j)_{j\in[r]\setminus J}\in\mathbb{Z}^{r-k},\\ |c_j|\leqslant B}} N_{\varphi_{(c_j)}^{-1}(S)}(B) \leqslant (2B+1)^{r-k} \cdot (2B+1)^k d_S(B) = (2B+1)^r d_S(B).$$

Taking the supremum over φ_2 implies that $d_{\psi^{-1}(S)}(B) \leqslant d_S(B)$, completing the proof. □

We also note that the value of $d_S(B)$ does not "jump too much" when B changes: namely, if B_1 and B_2 differ by at most a multiplicative constant factor, then the same holds true for $d_S(B_1)$ and $d_S(B_2)$.

Proposition 3.8. *Suppose that $0 \leqslant B_1 \leqslant B_2 \leqslant cB_1$ for some $c \geqslant 1$. Then for any set $S \subseteq \mathbb{F}^k$ we have*

$$\frac{1}{(3c)^k} d_S(B_1) \leqslant d_S(B_2) \leqslant 2^k d_S(B_1).$$

Proof. The first inequality follows by observing that $N_{\varphi(S)}(B_1) \leqslant N_{\varphi(S)}(B_2)$ and

$$2\lfloor B_2 \rfloor + 1 \leqslant 2B_2 + 1 \leqslant c(2B_1+1) \leqslant 3c(2\lfloor B_1 \rfloor + 1).$$

For the second inequality, we cover the integer points of the box $[-B_2, B_2]^k$ by $M = \left\lceil (2\lfloor B_2 \rfloor + 1)/(2\lfloor B_1 \rfloor + 1) \right\rceil^k$ boxes with side lengths $2\lfloor B_1 \rfloor$ centered at points with integer coordinates. Then for any bijective affine-linear map $\varphi : \mathbb{F}^k \to \mathbb{F}^k$ we have

$$N_{\varphi(S)}(B_2) \leqslant M \cdot \sup_{x\in\mathbb{Z}^k} N_{\varphi(S)-x}(B_1) \leqslant M \cdot (2\lfloor B_1 \rfloor + 1)^k d_S(B_1) \leqslant (2(2\lfloor B_2 \rfloor + 1))^k d_S(B_1).$$

Thus, $N_{\varphi(S)}(B_2)/(2\lfloor B_2 \rfloor + 1)^k \leqslant 2^k d_S(B_1)$, and taking the supremum over φ completes the proof. □

4 Reduction to lattice point counting

In this section we prove Theorem 4.1, stated below. For a general set S, this theorem provides an upper bound on the maximum S-translate probability $\rho(A, S)$ in terms of the integer point density function from Definition 3.6, under the assumption that the maximum point probability $\rho(A)$ is polynomially large (this is the "concentrated" case described in Section 2).

Theorem 4.1. *Fix $\delta \in (0,1)$ and $C, C_1 > 0$. Let $A = (a_1, \dots, a_n)$ be a sequence of vectors in a finite-dimensional vector space V (over a field $\mathbb{F} \subseteq \mathbb{C}$), such that the basis packing number of A is at least δn and $\rho(A) \geqslant n^{-C}$. Then there exists $r = O_C(1)$ such that for any subset $S \subseteq V$ we have*

$$\rho(A, S) \leqslant O_{\delta, C, C_1}\Big(d_S(\sqrt{n \log n}) \cdot (\log n)^r + n^{-C_1}\Big). \tag{3}$$

Example 1.5 shows that the bound in this theorem is sharp up to logarithmic factors.

Definition 4.2. A subset Q of an abelian group G is called a *proper symmetric generalized arithmetic progression* (*proper symmetric GAP*, for short) of rank r if there exist $v_1, \dots, v_r \in G$ and $q_1, \dots, q_r \in \mathbb{N}$ such that

$$Q = \{c_1 v_1 + \dots + c_r v_r : c_i \in \mathbb{Z},\ |c_i| \leqslant q_i \text{ for } 1 \leqslant i \leqslant r\},$$

and each element of Q can be represented as $c_1 v_1 + \dots + c_r v_r$ in a unique way.

A key ingredient in the proof of Theorem 4.1 is an *inverse theorem* for the linear Littlewood–Offord problem. The first theorem of this kind was proved in seminal work of Tao and Vu [36]; it states that if $\rho(A) \geqslant n^{-C}$, then almost all of the elements of A are contained in a common GAP whose volume is at most n^B and whose rank is at most r, for some r and B depending only on C. The quantitative aspects of this theorem were subsequently improved in theorems of Tao and Vu [35] and Nguyen and Vu [28]; we state the latter theorem below.

Theorem 4.3 (Optimal inverse Littlewood–Offord theorem; Nguyen and Vu [28, Theorem 2.5])**.** *Fix $\varepsilon \in (0,1)$ and $C > 0$. Let A be a sequence of n elements of an abelian torsion-free group, satisfying $\rho(A) \geqslant n^{-C}$. Then for any $n^\varepsilon \leqslant s \leqslant n$, there exists a proper symmetric GAP Q of rank $r = O_{C,\varepsilon}(1)$, such that it contains all but at most s elements of A, and*

$$|Q| = O_{C,\varepsilon}\Big(\rho(A)^{-1} s^{-r/2}\Big).$$

A significant shortcoming of Theorem 4.3 is that the bound on $|Q|$ is in terms of $\rho(A)$, which can be much smaller than $\rho(A')$ (we wish to "discard" the elements in $A \setminus A'$, and it is important to avoid a dependence on the discarded elements). We can address this issue by *iterating* Theorem 4.3, yielding the following result.

Theorem 4.4. *Fix $\varepsilon \in (0,1)$ and $C > 0$. Let A be a sequence of n elements of an abelian torsion-free group, satisfying $\rho(A) \geqslant n^{-C}$. Then for any $n^\varepsilon \leqslant s_1 \leqslant n$, there exists a proper symmetric GAP Q of rank $r = O_{C,\varepsilon}(1)$ and a subsequence A' of A of size at least $n - s_1$, such that all elements of A' lie in Q, and*

$$|Q| = O_{C,\varepsilon}\Big(\rho(A')^{-1} (s_1/\log n)^{-r/2}\Big).$$

Remark. We suspect that in the setting of Theorem 4.4, a stronger bound of the form $O_{C,\varepsilon}(\rho(A)^{-1} s_1^{-r/2})$ might hold (this would yield a common generalisation of Theorems 4.3 and 4.4).

Proof of Theorem 4.4. We assume that n is sufficiently large compared to C and ε. By decreasing s_1, we may assume that $s_1 \leqslant n/2$. Then by increasing C, we may also assume that $\rho(A) \geqslant (n - s_1)^{-C}$. Let $L = \lceil C \log_2 n \rceil$, and let $s = s_1/L$.

We construct a descending chain $A_0, A_1, \dots, A_L$ of subsequences of A with sizes $n_0 \geqslant n_1 \geqslant \dots \geqslant n_L$ satisfying $n_i \geqslant n - is$, as follows. Set $A_0 = A$. To obtain A_{i+1} from A_i, first note that $n^{\varepsilon/2} \leqslant s \leqslant n$, and that by Fact 3.2

$$\rho(A_i) \geqslant \rho(A) \geqslant (n - s_1)^{-C} \geqslant (n - is)^{-C} \geqslant n_i^{-C}.$$

Therefore, we can apply the optimal inverse theorem (Theorem 4.3) to the sequence A_i. As a result, we obtain a proper symmetric GAP Q_i of rank $r_i = O_{C,\varepsilon}(1)$, such that all but at most s elements of A_i lie in Q_i, and

$$|Q_i| = O_{C,\varepsilon}\Big(\rho(A_i)^{-1} s^{-r_i/2}\Big). \tag{4}$$

Let A_{i+1} be the subsequence of A_i consisting of all elements that lie in Q_i. Then we indeed have

$$n_{i+1} \geqslant n_i - s \geqslant n - (i+1)s.$$

Suppose that for some $0 \leqslant i < L$ we have $\rho(A_{i+1}) \leqslant 2\rho(A_i)$. In this case we can replace $\rho(A_i)$ by $\rho(A_{i+1})$ in the estimate (4). Then we are done by taking $A' = A_{i+1}$ and $Q = Q_i$.

Otherwise, we have $\rho(A_{i+1}) > 2\rho(A_i)$ for all $0 \leqslant i < L$. Then

$$\rho(A_L) > 2^L \rho(A) \geqslant 2^{C \log_2 n} n^{-C} = 1.$$

But $\rho(A_L)$ is a supremum of probabilities, a contradiction. □

We will also need a simple concentration inequality (a consequence of Hoeffding's inequality [21]). For positive reals $q_1, \dots, q_r$ we let $Q_r(q_1, \dots, q_r) = \{(x_1, \dots, x_r) \in \mathbb{Z}^r : |x_i| \leqslant q_i\}$.

Proposition 4.5. *For any vectors $a_1, \dots, a_m$ in $Q_r(q_1, \dots, q_r)$ and $t > 0$ we have*

$$\mathbb{P}[\xi_1 a_1 + \dots + \xi_m a_m \notin Q_r(tq_1, \dots, tq_r)] \leqslant 2r \exp\left(-\frac{t^2}{2m}\right).$$

Proof. For each $1 \leqslant j \leqslant r$, the j-th coordinate of $\xi_1 a_1 + \dots + \xi_m a_m$ is a sum of m independent random variables; each of them has expected value equal to zero and is supported in $[-q_j, q_j]$. By Hoeffding's inequality,

$$\mathbb{P}[|(\xi_1 a_1 + \dots + \xi_m a_m)_j| > tq_j] \leqslant 2\exp\left(-\frac{2 \cdot (tq_j)^2}{m(2q_j)^2}\right) \leqslant 2\exp\left(-\frac{t^2}{2m}\right).$$

The union bound over $1 \leqslant j \leqslant r$ completes the proof. □

Proof of Theorem 4.1. We apply Theorem 4.4 to A with $s_1 = \delta n/2$. As a result, we obtain a subsequence A' of size at least $(1 - \delta/2)n$ and a proper symmetric GAP Q of rank $r = O_C(1)$ such that all elements of A' lie in Q, and

$$|Q| \leqslant O_{\delta,C}\Big(\rho(A')^{-1} (n/\log n)^{-r/2}\Big).$$

Let $v_1, \dots, v_r$ be the generators of Q, and let $e_1, \dots, e_r$ be the standard basis vectors of $\mathbb{F}^r$. Let $\psi : \mathbb{F}^r \to V$ be the linear map defined by $\psi(e_i) = v_i$ for all i. Then Q is the image of the set of integer points of some box $Q_r(q_1, \dots, q_r)$ under this map ψ. Since Q is proper, the restriction $\psi|_{Q_r(q_1,\dots,q_r)}$ provides a bijection between $Q_r(q_1, \dots, q_r)$ and Q. In particular,

$$|Q_r(q_1, \dots, q_r)| = |Q| \leqslant O_{\delta,C}\left(\frac{(\log n)^{r/2}}{\rho(A') \cdot n^{r/2}}\right). \tag{5}$$

As the vectors of A' lie in Q, we can define A^* to be the sequence of vectors in $\mathbb{C}^r$ obtained by taking preimages of the elements of A' under this bijection.

By assumption, the basis packing number of A is at least δn. Hence, the basis packing number of A' is at least $\delta n/2 > 0$. In particular, the sequence $A' = \psi(A^*)$ contains a basis of V, thus the map ψ is surjective.

Then by Facts 3.2 and 3.4 we have

$$\rho(A, S) \leqslant \rho(A', S) = \rho(A^*, \psi^{-1}(S)), \tag{6}$$

and (by Facts 3.3 and 3.4)

$$\rho(A^*) \leqslant \rho(A^*, \ker \psi) = \rho(A'). \tag{7}$$

Let $a_1^*, \dots, a_m^*$ be the vectors of the sequence A^*, where $m \geqslant (1-\delta/2)n$, and consider $x \in \mathbb{F}^r$. By (6), it suffices to estimate the probability that $\xi_1 a_1^* + \ldots + \xi_m a_m^*$ lies in the set $S_x^* = (\psi^{-1}(S) - x)$.

By increasing C_1, we may assume that $\sqrt{2C_1}$ is a positive integer. Define the dilated box

$$Q^* = Q_r(\sqrt{2C_1 n \log n} \cdot q_1, \dots, \sqrt{2C_1 n \log n} \cdot q_n).$$

By Proposition 4.5, combined with the fact that $n/2 \leqslant m \leqslant n$, we have

$$\mathbb{P}[\xi_1 a_1^* + \ldots + \xi_m a_m^* \notin Q^*] \leqslant 2r \cdot m^{-C_1} \leqslant O_{C,C_1}(n^{-C_1}).$$

This corresponds to the n^{-C_1} term in the desired bound (3).

For the rest of the argument we focus on the probability that $\xi_1 a_1^* + \ldots + \xi_m a_m^*$ lies in the set $S_x^* \cap Q^*$. We estimate this probability by the union bound: that is, we view it as the sum of $\mathbb{P}[\xi_1 a_1^* + \ldots + \xi_m a_m^* = y]$ over all $y \in S_x^* \cap Q^*$. By definition, $\rho(A^*)$ is the maximum point probability. Then (7) implies that for any $y \in \mathbb{F}^r$

$$\mathbb{P}[\xi_1 a_1^* + \ldots + \xi_m a_m^* = y] \leqslant \rho(A^*) \leqslant \rho(A'). \tag{8}$$

To estimate the number of points in $|S_x^* \cap Q^*|$, we partition Q^* into boxes with side lengths $2\sqrt{n \log n}$. The number of points of S_x^* inside each constituent box may be expressed as a product of the total number of integer points in this box and the proportion of them that lie in S_x^*. This proportion, in turn, is bounded above by the relevant value of the density function $d_{S_x^*}(\sqrt{n \log n})$. Recalling that $S_x^* = \psi^{-1}(S) - x$ and ψ is a surjective linear map, by Proposition 3.7 we have $d_{S_x^*} = d_S$.

Taking the sum over all boxes in our partition, we obtain

$$|S_x^* \cap Q^*| \leqslant |Q^*| \cdot d_{S_x^*}(\sqrt{n \log n}) = |Q^*| \cdot d_S(\sqrt{n \log n}). \tag{9}$$

From (5) we have

$$|Q^*| \leqslant (n \log n)^{r/2} \cdot |Q_r(q_1, \dots, q_r)| \leqslant O_{\delta,C,C_1}\big(\rho(A')^{-1} (\log n)^r\big). \tag{10}$$

Finally, we combine the inequalities (8), (9) and (10) to conclude that

$$\begin{aligned}\mathbb{P}[\xi_1 a_1^* + \ldots + \xi_m a_m^* \in S_x^* \cap Q^*] &\leqslant \rho(A') \cdot |S_x^* \cap Q^*| \leqslant \rho(A') \cdot |Q^*| \cdot d_S(\sqrt{n \log n}) \\ &\leqslant O_{\delta,C,C_1}\Big(d_S(\sqrt{n \log n}) \cdot (\log n)^r\Big).\end{aligned}$$

This completes the proof. □

5 Sets of points in convex position: proof of Theorem 1.7

In this section we prove Theorem 1.7 and deduce Theorem 1.6. As described in Section 2, our strategy to prove Theorem 1.7 is to reduce the problem to counting integer points in a certain preimage of S. Therefore, we need an estimate on the maximum possible number of integer points in convex position inside $[-B, B]^k$. The following theorem is due to Andrews [1] (generalising an earlier result of Jarník [22] for the case $k = 2$).

Theorem 5.1 (Andrews [1]; see also [2, Theorem 2]). *Let $S \subseteq \mathbb{R}^k$ be a set of points in convex position. Then for any $B \geqslant 1$*

$$N_S(B) \leqslant O_k(B^{k-\frac{2k}{k+1}}).$$

Remark. Theorem 5.1 can be viewed as an upper bound on the number of vertices of a lattice polytope contained in $[-B, B]^k$. Bárány and Larman [2, Theorem 1] proved that this bound is tight up to a multiplicative constant factor: a lower bound of the same order of magnitude is achieved by the convex hull of integer points inside the ball of radius B.

As described in the outline, we use the following result of Fox, Kwan and Spink [16] to handle the "spread-out" case.

Theorem 5.2 (Fox, Kwan and Spink [16, Theorem 1.9(2)]). *Let $S \subseteq \mathbb{R}^k$ be a set of points in convex position, and let A be a sequence of nonzero vectors in $\mathbb{R}^k$. Then*

$$\rho(A, S) \leqslant O_k\Big(\rho(A)^{1/(k2^{k-1})}\Big).$$

Proof of Theorem 1.7. We have a sequence A of vectors in $\mathbb{R}^k$ with the basis packing number at least b. Consider the subsequence $A_0 = A[I_0]$ containing only the vectors of b disjoint bases. It has size $m := bk$ and basis packing number equal to b.

We need to estimate the probability that $\xi_1 a_1 + \ldots + \xi_n a_n$ lies in S. It is bounded by $\rho(A, S)$, which is at most $\rho(A_0, S)$ by Fact 3.2.

Let $C = k2^{k-1}$. First, suppose that $\rho(A_0) < m^{-C}$. Then, by Theorem 5.2, we have

$$\rho(A_0, S) \leqslant O_k\Big(\rho(A_0)^{1/(k2^{k-1})}\Big),$$

which is at most $O_k(m^{-1})$. Since $m = bk$, this gives the desired bound.

Therefore, we may assume that $\rho(A_0) \geqslant m^{-C}$. Applying Theorem 4.1 to A_0 with $\mathbb{F} = \mathbb{R}$, $\delta = 1/k$ and $C_1 = 1$, we obtain

$$\rho(A_0, S) \leqslant O_k\Big(d_S(\sqrt{m \log m}) \cdot (\log m)^r + m^{-1}\Big) \tag{11}$$

for some $r = O_k(1)$. Observe that for any bijective affine-linear map $\varphi : \mathbb{R}^k \to \mathbb{R}^k$ the set $\varphi(S)$ is also in convex position. Thus, by the definition of the density function d_S combined with Andrews' theorem (Theorem 5.1), for any $B \geqslant 1$ we have

$$d_S(B) = \sup_{\varphi}\left(\frac{N_{\varphi(S)}(B)}{(2\lfloor B \rfloor + 1)^k}\right) \leqslant O_k(B^{-\frac{2k}{k+1}}).$$

We substitute this into (11), recalling that $m = bk$, to conclude that

$$\rho(A_0, S) \leqslant O_k(b^{-\frac{k}{k+1}} (\log b)^{r-\frac{k}{k+1}}).$$

This completes the proof. □

Next, we combine Theorem 1.7 with the "dropping to a subspace" argument (Lemma 3.5) to deduce Theorem 1.6.

Proof of Theorem 1.6. Fix an arbitrary $\varepsilon \in (0, 1)$. We will prove that if n is sufficiently large in terms of k and ε then $\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S] \leqslant (2\sqrt{2/\pi} + \varepsilon) n^{-1/2}$.

We apply Lemma 3.5 to the sequence $A = (a_1, \ldots, a_n)$ with $b = \lfloor \varepsilon n/(k(k+1)) \rfloor + 1$. As a result, we obtain a subsequence $A' = A[I]$ of size at least $(1 - \varepsilon/2)n$, such that all elements of A' lie in a linear subspace $V' \subseteq \mathbb{R}^k$,

and the basis packing number of A' inside V' is at least $b = \Omega_{k,\varepsilon}(n)$. Conditioning on the outcomes of the random variables $(\xi_i)_{i\in[n]\setminus I}$, we have

$$\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S] \leqslant \sup_{x\in\mathbb{R}^k} \mathbb{P}\left[\sum_{i\in I} \xi_i a_i \in (S-x)\cap V'\right].$$

Fix an arbitrary $x \in \mathbb{R}^k$, and let $\ell = \dim V'$. First we consider the case $\ell \geqslant 2$. Applying Theorem 1.7 to A' and $(S-x)\cap V'$, we conclude that

$$\mathbb{P}\left[\sum_{i\in I} \xi_i a_i \in (S-x)\cap V'\right] \leqslant O_\ell\Big(b^{-\frac{\ell}{\ell+1}}(\log b)^{C_\ell}\Big).$$

Since $2 \leqslant \ell \leqslant k$ and $b = \Omega_{k,\varepsilon}(n)$, this bound is at most $2\sqrt{2/\pi}\cdot n^{-1/2}$ when n is sufficiently large (in terms of k and ε).

Next, we deal with the case $\ell = 1$. In this case, $(S-x)\cap V'$ is a set of points in convex position on a line. Then it contains at most 2 points, and the desired probability can be bounded by the classical Erdős–Littlewood–Offord theorem:

$$\mathbb{P}\left[\sum_{i\in I} \xi_i a_i \in (S-x)\cap V'\right] \leqslant 2\cdot 2^{-|I|}\binom{|I|}{\lfloor |I|/2\rfloor} = \Big(2\sqrt{2/\pi}+o(1)\Big)|I|^{-1/2}.$$

Since $|I| \geqslant (1-\varepsilon/2)n$, one can check that this expression is at most $(2\sqrt{2/\pi}+\varepsilon)n^{-1/2}$ when n is sufficiently large (in terms of ε). □

6 Algebraic preliminaries

Now, for the rest of the paper we turn our attention to Theorem 1.4 and its corollaries. We start with some preliminaries from algebraic geometry and number theory.

6.1 Algebraic geometry

In this subsection, we review some basic concepts and facts from algebraic geometry. We loosely follow the exposition in [8, Section 7.1], and refer to [19, Chapter 1] for more details.

An *affine algebraic variety* $S \subseteq \mathbb{C}^k$ (a *variety*, for short) is the set of common zeros of a finite collection of polynomials $f_1, \ldots, f_m \in \mathbb{C}[x_1, \ldots, x_k]$:

$$S = \{(x_1, \ldots, x_k) \in \mathbb{C}^k : f_1(x_1, \ldots, x_k) = \ldots = f_m(x_1, \ldots, x_k) = 0\}.$$

A variety is called *irreducible* if it cannot be written as a union of two proper subvarieties. Each variety S can be uniquely written as the union of irreducible subvarieties $S_1 \cup \ldots \cup S_m$, such that $S_i \not\subseteq S_j$ for any $i \neq j$. Varieties $S_1, \ldots, S_m$ are called the *irreducible components* of S.

The *dimension* of an irreducible variety S is the maximal integer ℓ such that there exists a chain of non-empty irreducible subvarieties $S_0 \subsetneq S_1 \subsetneq \ldots \subsetneq S_\ell = S$. The dimension $\dim S$ of an arbitrary variety S is defined as the maximum dimension of its irreducible components. By convention, the empty variety is reducible and has dimension $-\infty$.

The *codimension* $\operatorname{codim} S$ of a variety $S \subseteq \mathbb{C}^k$ is defined as $k - \dim S$.

The *degree* of an irreducible variety S is the cardinality of the intersection of S with a "generic" affine subspace of dimension $\operatorname{codim} S$ (a well-defined positive integer). The degree $\deg S$ of an arbitrary variety S is defined as the sum of the degrees of its irreducible components. By convention, the degree of the empty variety is 0.

Fact 6.1. *For any two varieties $S, T \subseteq \mathbb{C}^k$*

$$\deg(S \cup T) \leqslant \deg(S) + \deg(T).$$

Fact 6.2. *For any variety $S \subseteq \mathbb{C}^k$ and a surjective affine-linear map $\pi : \mathbb{C}^r \to \mathbb{C}^k$ we have*

$$\operatorname{codim} \pi^{-1}(S) = \operatorname{codim} S, \quad \deg \pi^{-1}(S) = \deg S.$$

Furthermore, if S is irreducible, then $\pi^{-1}(S)$ is also irreducible.

Fact 6.3. *Let $f \in \mathbb{C}[x_1, \ldots, x_k]$ be an irreducible (over $\mathbb{C}$) polynomial of degree $d \geqslant 1$. Then $S = \{x \in \mathbb{C}^k : f(x) = 0\}$ is an irreducible variety of dimension $k-1$ and degree d.*

Fact 6.4 (Generalized Bézout's theorem [17, Example 12.3.1]; see also [8, 7])**.** *For any two varieties $S, T \subseteq \mathbb{C}^k$,*

$$\deg(S \cap T) \leqslant \deg(S) \cdot \deg(T).$$

Proposition 6.5. *Let $\{S_i\}_{i \in I}$ be a (not necessarily finite) collection of varieties in $\mathbb{C}^k$. Suppose that $\deg(S_i) \leqslant d$ for all $i \in I$. Then the set $S = \bigcap_{i \in I} S_i$ is a variety of degree at most d^k.*

Proof. We will prove a stronger statement: for any irreducible variety $T \subseteq \mathbb{C}^k$ of dimension at most ℓ and degree at most d_0, the intersection $T \cap S$ is a variety of degree at most $d_0 d^\ell$. The proposition then follows from this statement applied with $T = \mathbb{C}^k$.

We argue by induction on $\ell = \dim T$. If $T \cap S = T$, there is nothing to prove. Otherwise, there exists $i \in I$ such that $T \cap S_i \subsetneq T$. Let $T_1, \ldots, T_m$ be the irreducible components of $T \cap S_i$. Each of them has dimension at most $\ell - 1$, and by Bézout's theorem (Fact 6.4) we have

$$\sum_{j=1}^{m} \deg(T_j) = \deg(T \cap S_i) \leqslant d_0 d.$$

By the induction hypothesis, we know that for each $1 \leqslant j \leqslant m$ the set $T_j \cap S$ is a variety, and that

$$\deg(T_j \cap S) \leqslant \deg(T_j) \cdot d^{\ell - 1}.$$

As $T \cap S = \bigcup_{j=1}^{m} (T_j \cap S)$, by Fact 6.1 we conclude that

$$\deg(T \cap S) \leqslant \sum_{j=1}^{m} \deg(T_j \cap S) \leqslant d^{\ell-1} \sum_{j=1}^{m} \deg(T_j) \leqslant d_0 d^\ell. \qquad \square$$

6.2 Number theory

A large area of research in number theory is concerned with counting integer solutions to polynomial equations or, more generally, integer points on algebraic varieties. The most basic result in this direction is the Schwartz–Zippel lemma.

Proposition 6.6 (Schwartz–Zippel lemma for varieties, see for example [7, Lemma 14])**.** *Let $S \subseteq \mathbb{C}^k$ be a variety of dimension ℓ and degree d. Then for any $B \geqslant 1$,*

$$N_S(B) \leqslant d \cdot (2B+1)^\ell.$$

Since bijective affine-linear maps preserve dimension and degree (by Fact 6.2), in terms of the density function this means $d_S(B) \leqslant d \cdot (2B+1)^{-(k-\ell)}$.

Proposition 6.6 is sharp when S is a union of d axis-parallel affine subspaces of dimension ℓ (and it can be approximately sharp whenever S contains a dimension-ℓ affine subspace as one of its irreducible components). However, one can obtain much stronger bounds by making certain assumptions on S. We will need a few different results in this direction.

First, we can make the assumption that S is irreducible. The following bound in this setting was proved by Pila [31] (improving on his slightly weaker bound in [30]). It was proved using the so-called *Bombieri–Pila determinant method* (famously introduced by Bombieri and Pila [4] to prove a similar theorem for curves in $\mathbb{R}^2$). We refer the reader to [3] for a recent survey on this topic.

Theorem 6.7 (Pila [30, 31])**.** *Let $S \subseteq \mathbb{C}^k$ be an irreducible variety of dimension ℓ and degree d. Then for any $B \geqslant 2$,*

$$N_S(B) = O_{d,k}\big(B^{\ell-1+1/d}(\log B)^{C_d}\big)$$

for some constant C_d depending only on d.

Again, Fact 6.2 allows to rewrite this in terms of the density function as

$$d_S(B) \leqslant O_{d,k}\big(B^{-(k-\ell+1-1/d)}(\log B)^{C_d}\big).$$

The example $S = \{x \in \mathbb{C}^k : x_1 = x_2^d\}$ shows that Theorem 6.7 is sharp up to logarithmic factors, i.e., one cannot hope to remove the "$1/d$" term in the exponent, in general. However, there is a general belief that if one makes a mild assumption ruling out examples of this type, one should expect a bound of the form $N_S(B) \leqslant B^{\ell-1+o(1)}$. A conjecture along these lines was first proposed by Heath-Brown [20], and this conjecture (and its variants) are usually collectively referred to as the "dimension growth conjecture". Various partial results are available, see [37] and the references therein. In this paper we use the following result, due to Vermeulen [37] for $d \geqslant 4$ and due to Browning and Gorodnik [5] for $d = 2$ (using ideas of Browning, Heath-Brown and Salberger [6]; see also [34]), which settles the (uniform) "affine dimension growth conjecture" for affine hypersurfaces of degree $d \neq 3$.

Theorem 6.8 (Affine dimension growth conjecture for hypersurfaces; Vermeulen [37, Theorem 1.2] and Browning–Gorodnik [5, Theorem 1.11])**.** *Consider an irreducible polynomial $f \in \mathbb{C}[x_1, \dots, x_k]$ of degree $d \neq 3$. Suppose that f cannot be represented as a polynomial of two linear forms. Then, with $S \subseteq \mathbb{C}^k$ as the zero set of f, and for any $B \geqslant 1$, $\varepsilon > 0$, we have*

$$N_S(B) \leqslant O_{d,k,\varepsilon}(B^{k-2+\varepsilon}).$$

Note that if the assumption of Theorem 6.8 holds for a polynomial f, then it also holds for $f \circ \varphi$ for any bijective affine-linear map $\varphi : \mathbb{C}^k \to \mathbb{C}^k$. Therefore, we can rewrite the resulting bound in terms of the density function as $d_S(B) \leqslant O_{d,k,\varepsilon}(B^{-2+\varepsilon})$.

Vermeulen [37] and Browning–Gorodnik [5] state their results only for polynomials with rational coefficients. The reason is that this is the hardest case, which is also most natural to consider from number-theoretic point of view. Since we would like to apply Theorem 6.8 to an arbitrary polynomial f, we provide a short argument which handles the case when f is not proportional to a polynomial with coefficients in $\mathbb{Q}$. Namely, Proposition 6.9 below (applied with $\mathbb{F} = \mathbb{Q}$) implies that in this case the set of integer zeros of f lies in a variety of codimension at least 2. Then the desired bound on its size follows directly from the Schwartz–Zippel lemma (Proposition 6.6).

Proposition 6.9. *Consider a field $\mathbb{F} \subseteq \mathbb{C}$, and consider an irreducible polynomial $f \in \mathbb{C}[x_1, \dots, x_k]$ of degree d, which is not proportional to a polynomial with coefficients in $\mathbb{F}$. Let $S \subseteq \mathbb{C}^k$ be the zero set of f. Then there exists another variety $T \subseteq S$ of dimension at most $k-2$ and degree at most d^2 such that $T \cap \mathbb{F}^k = S \cap \mathbb{F}^k$.*

Proof. Rescale f so that one of its coefficients is equal to 1. Then it has a coefficient $z \in \mathbb{C} \setminus \mathbb{F}$.

Recall the following simple fact: for any $z \in \mathbb{C} \setminus \mathbb{F}$ there exists an automorphism σ of $\mathbb{C}$ which acts as the identity on $\mathbb{F}$ but does not fix z. To prove this fact, one can first define this automorphism on $\mathbb{F}(z)$ by sending z to a different root of the minimal polynomial of z over $\mathbb{F}$ if it is algebraic over $\mathbb{F}$, and to (say) $z+1$ if z is transcendental over $\mathbb{F}$. Then one can extend this automorphism to the whole of $\mathbb{C}$ (see for example [38]).

Applying this automorphism σ to each coefficient of f, we obtain a polynomial f^σ, which is not proportional to f but still satisfies $f(q) = f^\sigma(q)$ for any $q \in \mathbb{F}^k$. Therefore, the variety T defined as

$$T = \{x \in \mathbb{C}^k : f(x) = f^\sigma(x) = 0\}$$

indeed satisfies $T \cap \mathbb{F}^k = S \cap \mathbb{F}^k$. By Fact 6.3, T is an intersection of two distinct irreducible varieties of dimension $k-1$ and degree d. Then it has dimension at most $k-2$ and, by Bézout's theorem (Fact 6.4), degree at most d^2. □

7 Decomposition into subspaces

In this section we prove Theorem 7.1, stated below. As outlined in Section 2, this provides a decomposition of $\mathbb{C}^k$ into a "structured" and subspace W and a "disordered" subspace U.

Theorem 7.1. *Fix $\delta, C_1 > 0$. Let $S \subseteq \mathbb{C}^k$ be a variety of degree at most d. Consider a sequence A of vectors in $\mathbb{C}^k$ with basis packing number at least δn.*

Then there exists a decomposition of $\mathbb{C}^k$ as $U \oplus W$ for some linear subspaces U and W, a variety $S' \subseteq W$ of degree at most d^k, and a subsequence A' of A satisfying all the following conditions:

(1) The basis packing number of A' is at least $(\delta/(2(2k)^k)) \cdot n$;

(2) Let $\pi_W : \mathbb{C}^k \to W$ be the projection map. Then $\rho(\pi_W(A')) \geqslant n^{-C}$ for some constant C not depending on n (but possibly depending on δ, C_1, d, k);

(3) $\pi_W^{-1}(S') \subseteq S$, and $\rho(A', S \setminus \pi_W^{-1}(S')) \leqslant n^{-C_1}$.

Intuitively, condition (2) says that the projection of (a subsequence of) A onto W has polynomially large point probabilities (which allows us to apply an inverse Littlewood–Offord theorem such as Theorem 4.3). The set $\pi_W^{-1}(S') \subseteq S$ can be viewed as "the part of S which we can control via its projection onto W". Condition (3) then gives us control over the complementary part of S (which cannot be studied via its projection onto W).

Our strategy to prove Theorem 7.1 is to consider the following procedure. We begin with $A' = A$ and $U = \mathbb{C}^k$, where conditions (1) and (2) hold automatically. Then we show that the only way for condition (3) to fail is if A' contains a "linear-size" subsequence A_1 for which $\rho(A_1, U_1)$ remains polynomial in n for some proper subspace $U_1 \subsetneq U$. In that case, we set $A' = A_1$ and $U = U_1$ (while maintaining conditions (1) and (2)), and repeat the process. As the dimension of U decreases on each step, the procedure terminates after at most k steps.

Later in this section, we will state and prove Proposition 7.4, which describes a single step of the above procedure. Its proof relies on Lemma 7.2, stated below, which allows one to bound the "variety probability" $\rho(A, S)$ in terms of certain "subspace probabilities" $\rho(A', V)$ (for subsequences A' of A and subspaces V contained in a translate of S).

Lemma 7.2. *Let $S \subseteq \mathbb{C}^k$ be a variety of dimension at most ℓ and degree at most d. Consider a sequence A of vectors in $\mathbb{C}^k$, partitioned into $\ell+1$ subsequences $A_0, \dots, A_\ell$. Then*

$$\rho(A, S) \leqslant (\ell+1) \cdot d \cdot \left(\sup_{i,V} \rho(A_i, V)\right)^{1/2^\ell},$$

where the supremum is taken over $0 \leqslant i \leqslant \ell$ *and over linear subspaces* $V \subseteq \mathbb{C}^k$*, such that* $V \subseteq S - y$ *for some* $y \in \mathbb{C}^k$*.*

The proof of Lemma 7.2 is based on an "iterative decoupling argument", inspired by the approach of [25]. For context, *decoupling* is a general term for a large body of techniques in probability and statistics for "reducing from dependent situations to independent ones" (see for example the monograph [13]). In Littlewood–Offord theory, "decoupling" usually refers to a class of techniques to reduce polynomial anticoncentration to linear anticoncentration (popularised by Costello, Tao and Vu [11]), via inequalities such as Lemma 7.3 below. The particular statement of Lemma 7.3 appears (for example) as [12, Lemma 8.4], but for the convenience of the reader we provide the short proof.

Lemma 7.3. *If an event* $\mathcal{E}(X, Y)$ *depends on independent random objects* X, Y*, and* X' *is an independent copy of* X*, then*

$$\mathbb{P}[\mathcal{E}(X,Y)] \leqslant \Big(\mathbb{P}[\mathcal{E}(X,Y) \text{ and } \mathcal{E}(X',Y)]\Big)^{1/2}.$$

Proof. By the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\mathbb{P}[\mathcal{E}(X,Y) \text{ and } \mathcal{E}(X',Y)] &= \mathbb{E}_Y\Big[\mathbb{P}[\mathcal{E}(X,Y) \text{ and } \mathcal{E}(X',Y) \,|\, Y]\Big] = \mathbb{E}_Y\Big[\mathbb{P}[\mathcal{E}(X,Y) \,|\, Y]^2\Big] \\ &\geqslant \mathbb{E}_Y\Big[\mathbb{P}[\mathcal{E}(X,Y) \,|\, Y]\Big]^2 = \mathbb{P}[\mathcal{E}(X,Y)]^2.\end{aligned}$$

Taking square roots on both sides completes the proof. □

Proof of Lemma 7.2. Let $S_1, \dots, S_m$ be the irreducible components of S. Recall that (by definition) $\deg S = \sum_{j=1}^m \deg S_j$, and that (by Fact 3.3) $\rho(A, S) \leqslant \sum_{j=1}^m \rho(A, S_j)$. Therefore, by treating each irreducible component separately, we may assume that S is irreducible.

We argue by induction on ℓ. In the base case $\ell = 0$ we consider only one subsequence $A_0 = A$, the variety S consists of a single point, and the only linear subspace appearing in the supremum has dimension zero. So, in this case both sides of the inequality are equal to $\rho(A)$.

Let n be the size of A, and let $I_0 \subseteq [n]$ be the set of indices corresponding to the subsequence A_0. Consider independent random variables

$$X = \sum_{i \in I_0} \xi_i a_i, \quad Y = \sum_{i \in [n] \setminus I_0} \xi_i a_i,$$

and let X' be an independent copy of X.

Fix any $x \in \mathbb{C}^k$. Then by the decoupling lemma (Lemma 7.3),

$$\begin{aligned}\mathbb{P}[X + Y \in S - x] &\leqslant \mathbb{P}[X + Y \in S - x \text{ and } X' + Y \in S - x]^{1/2} \\ &= \mathbb{P}[Y \in (S - x - X) \cap (S - x - X')]^{1/2}\end{aligned}$$

Define a (random) variety $T = (S - x - X) \cap (S - x - X')$.

First we deal with the case when $\dim T \leqslant \ell - 1$. By Bézout's theorem (Fact 6.4), we have $\deg T \leqslant d^2$. Applying the induction hypothesis to the variety T and the subsequences $A_1, \dots, A_\ell$, we obtain that

$$\mathbb{P}_Y[Y \in T \mid \dim T \leqslant \ell - 1] \leqslant \ell d^2 \Big(\sup_{i', V'} \rho(A_{i'}, V')\Big)^{1/2^{\ell-1}}. \tag{12}$$

where the supremum is taken over $1 \leqslant i' \leqslant \ell$ and over linear subspaces V' contained in a translate of T. Since T itself is contained in a translate of S, this supremum is bounded above by the supremum appearing in the statement of the lemma:

$$\sup_{1 \leqslant i' \leqslant \ell, V' \subseteq T - y'} \rho(A_{i'}, V') \leqslant \sup_{0 \leqslant i \leqslant \ell, V \subseteq S - y} \rho(A_i, V).$$

Next, we consider the case when $\dim T = \dim S = \ell$. As S is irreducible, this can happen only if $S - x - X = S - x - X'$. Define

$$V_S = \{v \in \mathbb{C}^k : S = S - v\}.$$

Equivalently, V_S consists of all vectors $v \in \mathbb{C}^k$ such that for any $x \in S$ the point $x + v$ also lies in S. We claim that V_S is a linear subspace.

It is clear that if v_1, v_2 lie in V_S then $v_1 + v_2$ also lies in V_S. Thus, for any $y \in S$ and $v \in V_S$ the point $y + tv$ lies in S for any $t \in \mathbb{N}$. So, the variety S has infinitely many intersection points with the line $y + tv$. Then it contains the whole line, and $y + tv \in S$ for any $t \in \mathbb{C}$. Hence, if $v \in V_S$ then $tv \in V_S$ for any $t \in \mathbb{C}$.

Therefore, V_S is a linear subspace contained in $S - y$ for any $y \in S$. From its definition, we have

$$\mathbb{P}_{X,X'}[\dim T = \ell] = \mathbb{P}_{X,X'}[(X + x) - (X' + x) \in V_S] = \mathbb{E}_{X'}\Big[\mathbb{P}_X[X \in V_S + X' \mid X']\Big] \leqslant \rho(A_0, V_S). \tag{13}$$

Combining (12) and (13), we conclude that

$$\begin{aligned}
\mathbb{P}[Y \in T] &\leqslant \mathbb{P}_{X,X'}[\dim T = \ell] + \mathbb{E}_{X,X'}\Big[\mathbb{P}_Y[Y \in T \mid \dim T \leqslant \ell - 1]\Big] \\
&\leqslant \rho(A_0, V_S) + \ell d^2 \left(\sup_{1 \leqslant i' \leqslant \ell, V' \subseteq T - y'} \rho(A_{i'}, V') \right)^{1/2^{\ell-1}} \\
&\leqslant (\ell + 1) d^2 \left(\sup_{0 \leqslant i \leqslant \ell, V \subseteq S - y} \rho(A_i, V) \right)^{1/2^{\ell-1}}.
\end{aligned}$$

So, for any $x \in \mathbb{C}^k$ we have

$$\mathbb{P}[X + Y \in S - x] \leqslant \mathbb{P}[Y \in T]^{1/2} \leqslant (\ell + 1) d \left(\sup_{i,V} \rho(A_i, V) \right)^{1/2^{\ell}}.$$

As $\rho(A, S) = \sup_{x \in \mathbb{C}^k} \mathbb{P}[X + Y \in S - x]$, this completes the proof. □

Now we state and prove Proposition 7.4, which constitutes one step of the iterative procedure in the proof of Theorem 7.1.

Proposition 7.4. *Fix $\delta, C, C_1 > 0$. Let $S \subseteq \mathbb{C}^k$ be a variety of degree at most d. Consider a sequence A of **at most** n vectors in $\mathbb{C}^k$ with basis packing number at least δn. Let U be a subspace of $\mathbb{C}^k$ satisfying $\rho(A, U) \geqslant n^{-C}$. Fix a decomposition of $\mathbb{C}^k$ as $U \oplus W$ for some linear subspace W, and let $\pi_W : \mathbb{C}^k \to W$ be the projection map. Then at least one of the following holds:*

(a) There exists a subsequence A' of A with basis packing number at least $(\delta/2)n$, and a variety $S' \subseteq W$ of degree at most d^k, such that $\pi_W^{-1}(S') \subseteq S$ and

$$\rho(A', S \setminus \pi_W^{-1}(S')) \leqslant n^{-C_1}.$$

(b) There exists a subsequence A'' of A with basis packing number at least $(\delta/(2k))n$, and a linear subspace $U' \subsetneq U$, such that

$$\rho(A'', U') \geqslant n^{-C''}$$

for some constant C'' not depending on n (but possibly depending on δ, C, C_1, d, k).

Proof. First, we define a variety $S' \subseteq W$ in the following way:

$$S' = \{w \in W : u + w \in S \text{ for every } u \in U\}$$

In other words, $S' = \bigcap_{u \in U} (S - u) \cap W$. By Bézout's theorem (Fact 6.4), the degree of each $(S - u) \cap W$ is at most d. Therefore, by Proposition 6.5, the set S' is indeed a variety of degree at most d^k.

Next we define a subsequence A'. Let $\pi_W : \mathbb{C}^k \to W$ be the projection map. Then by Fact 3.4,

$$\rho(\pi_W(A)) = \rho(A, U) \geqslant n^{-C}.$$

The basis packing number of A is at least δn, thus, in particular, it contains at least δn vectors. Since $\rho(\pi_W(A)) \geqslant n^{-C} \geqslant (\delta n)^{-C'}$ for some constant $C' = C'(\delta, C)$, we can apply the optimal inverse theorem (Theorem 4.3) to the sequence $\pi_W(A)$, with $s = \delta n/2$. As a result, we obtain a proper symmetric GAP $Q \subseteq W$ of rank $r = O_{\delta,C}(1)$, containing all but at most $\delta n/2$ elements of $\pi_W(A)$, with

$$|Q| \leqslant K_{\delta,C} \cdot \rho(\pi_W(A))^{-1} n^{-r/2}$$

for some constant $K_{\delta,C}$ depending only on δ and C.

We define A' to be the subsequence of vectors a in A whose projections $\pi_W(a)$ lie in Q. Since A' is obtained by removing at most $\delta n/2$ elements from A, its basis packing number is at least $\delta n/2$.

Let $a'_1, \ldots, a'_m$ be the vectors of A', and let $X = \xi_1 a'_1 + \ldots + \xi_m a'_m$. Suppose that condition (a) does not hold: that is, for some $x \in \mathbb{C}^k$

$$\mathbb{P}\big[X + x \in S \setminus \pi_W^{-1}(S')\big] > n^{-C_1}. \tag{14}$$

The projection of each vector in the sequence A' onto W lies in Q, thus for any outcomes of independent Rademacher random variables $\xi_1, \ldots, \xi_m$, the sum $\xi_1 \pi_W(a'_1) + \ldots + \xi_m \pi_W(a'_m)$ lies in the dilated GAP nQ. (In fact, with high probability it lies in the smaller dilated GAP $\sqrt{n \log n} Q$ by Proposition 4.5, but here this is not important for us.) In particular, the random variable $\pi_W(X + x)$ is supported on some finite set $H \subseteq W$ satisfying

$$|H| \leqslant n^r |Q| \leqslant n^r \cdot K_{\delta,C} \cdot \rho(\pi_W(A))^{-1} n^{-r/2} \leqslant K_{\delta,C} \cdot n^{C + r/2}. \tag{15}$$

Choose C'' such that $K_{\delta,C} \cdot k \cdot d \cdot n^{-C''/2^{k-1} + C + r/2 + C_1} < 1$ for any $n \geqslant 2$. Suppose that condition (b) also does not hold: that is, any subsequence A'' of A with basis packing number at least $(\delta/(2k))n$ and any proper subspace $U' \subsetneq U$ satisfy

$$\rho(A'', U') < n^{-C''}.$$

Our goal is to show that this leads to a contradiction.

For any $w \in W$ let S_w be the intersection of S with the affine subspace $U + w$. Then

$$S \setminus \pi_W^{-1}(S') = \bigsqcup_{w \in W, w \notin S'} S_w.$$

Moreover, the probability that $X + x$ lies in S_w is positive only if $w = \pi_W(S_w)$ lies in H. Therefore,

$$\mathbb{P}\big[X + x \in S \setminus \pi_W^{-1}(S')\big] = \sum_{w \in H, w \notin S'} \mathbb{P}[X + x \in S_w] \leqslant \sum_{w \in H, w \notin S'} \rho(A', S_w). \tag{16}$$

In order to estimate each summand we use Lemma 7.2. As the basis packing number of the sequence A' is at least $\delta n/2$, we can partition it into k subsequences $A_0, \ldots, A_{k-1}$, such that each of them has basis packing number at least $\delta n/(2k)$.

Fix $w \in H \setminus S'$. Then, by definition of S', we have $S_w \subsetneq U + w$. Note that $\dim S_w \leqslant k-1$ and (by Bézout's theorem (Fact 6.4)) $\deg S_w \leqslant \deg S \leqslant d$. Therefore, by Lemma 7.2,

$$\rho(A', S_w) \leqslant k \cdot d \cdot \left(\sup_{0 \leqslant i \leqslant k-1,\ V \subseteq S_w - y} \rho(A_i, V) \right)^{1/2^{k-1}}.$$

Crucially, since $S_w \subsetneq U + w$, the supremum in the right hand side is taken only over proper subspaces of U. As we assumed that condition (b) does not hold, we conclude that

$$\rho(A', S_w) \leqslant k \cdot d \cdot n^{-C''/2^{k-1}}.$$

Combining this with (15) and (16), we obtain

$$\mathbb{P}\big[X + x \in S \setminus \pi_W^{-1}(S')\big] \leqslant |H| \cdot k \cdot d \cdot n^{-C''/2^{k-1}} \leqslant K_{\delta,C} \cdot k \cdot d \cdot n^{-C''/2^{k-1}+C+r/2}.$$

This is less than n^{-C_1} by our choice of C'', which contradicts our assumption (14). □

Now we show how to iterate Proposition 7.4 to prove Theorem 7.1.

Proof of Theorem 7.1. We will describe an iterative process to construct a descending chain $A_0, A_1, \ldots$ of subsequences of A, and a descending chain $U_0 \supsetneq U_1 \supsetneq \ldots$ of linear subspaces of $\mathbb{C}^k$, maintaining the following two properties:

- The basis packing number of A_i is at least $(\delta/(2k)^i)n$;
- $\rho(A_i, U_i) \geqslant n^{-C_i'}$ for some constant C_i' not depending on n.

We start with $A_0 = A$, $U_0 = \mathbb{C}^k$ and $C_0' = 1$. Now, suppose we have already constructed A_i and U_i. We will attempt to find a decomposition $U \oplus W$ with $U = U_i$ (and some A', S'), satisfying the desired properties (1), (2) and (3) (actually, only (3) is nontrivial). If this is not possible, we will show how to construct A_{i+1} and U_{i+1}, to continue the process. Since this process can continue for at most k steps (as it is not possible to have a descending chain of more than $k+1$ linear subspaces of $\mathbb{C}^k$), this is sufficient to prove Theorem 7.1.

Let W_i be an arbitrary linear complement of U_i, and apply Proposition 7.4 to the sequence A_i and the decomposition $\mathbb{C}^k = U_i \oplus W_i$. Suppose that condition (a) of Proposition 7.4 holds. Then there exists a variety $S' \subseteq W_i$ of degree at most d^k, and a subsequence A' of A_i with basis packing number at least $(\delta/(2(2k)^i))n$ such that

$$\rho(A', S \setminus \pi_{W_i}^{-1}(S')) \leqslant n^{-C_1}.$$

In this case we are done by setting $U = U_i$, $W = W_i$, and $C = C_i'$. Indeed, properties (1) and (3) are satisfied by the above. Recalling Facts 3.2 and 3.4, we note that $\rho(A', U_i) \geqslant \rho(A_i, U_i) = \rho(\pi_{W_i}(A_i)) \geqslant n^{-C_i'}$, thus property (2) holds as well.

Otherwise, condition (b) of Proposition 7.4 holds, which gives us a subsequence A_{i+1} with basis packing number at least $(\delta/(2k)^{i+1})n$, and a subspace $U_{i+1} \subsetneq U_i$ such that

$$\rho(A_{i+1}, U_{i+1}) \geqslant n^{-C_{i+1}'}$$

for some C_{i+1}' depending only on δ, C_i', C_1, d, k. This allows us to proceed to the next step of the process. □

8 Proofs of the main results

In this section we prove a general result (Theorem 8.1 below), which allows us to reduce Littlewood–Offord-type statements about estimating $\rho(A, S)$ for an algebraic variety S to questions about counting lattice points on S (in the sense of the integer point density function from Definition 3.6). We will then show how to apply Theorem 8.1 along with the number-theoretic results presented in Section 6.2 to derive Theorems 1.1, 1.3 and 1.4. After that, we deduce Theorems 1.2 and 1.8 from Theorem 1.4.

Theorem 8.1. *Let $S \subseteq \mathbb{C}^k$ be a variety of dimension at most ℓ and degree at most d. Consider a sequence A of vectors in $\mathbb{C}^k$ with basis packing number at least $b \geqslant 2$. Then there exists $r = O_{d,k}(1)$ such that*

$$\rho(A, S) \leqslant O_{d,k}\Big(\Big(d_S(\sqrt{b \log b}) + (b \log b)^{-(k-\ell+1)/2}\Big) \cdot (\log b)^r\Big).$$

In comparison to Theorem 4.1, note that Theorem 8.1 does not require that $\rho(A) \geqslant n^{-C}$, but instead requires that S is an algebraic variety of bounded degree (and it gives a slightly worse bound).

Proof of Theorem 8.1. Let $S_1, \ldots, S_t$ (where $t \leqslant d$) be the irreducible components of S. By Fact 3.3, we have $\rho(A, S) \leqslant \sum_{j=1}^{t} \rho(A, S_j)$. Therefore, by treating each irreducible component separately, we may assume that S is irreducible.

We have a sequence A of vectors in $\mathbb{C}^k$ with basis packing number at least b. Consider a subsequence A_0 of A, which contains only the vectors of the b bases. It has size $m = bk$ and basis packing number equal to b. Applying Theorem 7.1 to the sequence A_0 and the variety S with $\delta = 1/k$ and $C_1 = k + 1$, we obtain a decomposition $\mathbb{C}^k = U \oplus W$, a variety $S' \subseteq W$ of degree at most d^k, and a subsequence A' satisfying the following conditions:

(1) The basis packing number of A' is at least αm for some $\alpha = \alpha(k) > 0$;

(2) Let $\pi_W : \mathbb{C}^k \to W$ be the projection map. Then for some $C = C(d, k)$ we have $\rho(\pi_W(A')) \geqslant m^{-C}$;

(3) $\pi_W^{-1}(S') \subseteq S$, and $\rho(A', S \setminus \pi_W^{-1}(S')) \leqslant m^{-(k+1)}$.

Assuming that b is sufficiently large with respect to d and k, condition (3) guarantees that

$$\rho(A', S \setminus \pi_W^{-1}(S')) \leqslant m^{-(k+1)} \leqslant b^{-(k+1)} \leqslant (b \log b)^{-(k-\ell+1)/2} (\log b)^r. \tag{17}$$

By Facts 3.2 and 3.3,

$$\rho(A, S) \leqslant \rho(A', S) \leqslant \rho(A', \pi_W^{-1}(S')) + \rho(A', S \setminus \pi_W^{-1}(S')).$$

Then, by (17), the second summand $\rho(A', S \setminus \pi_W^{-1}(S'))$ is small compared to the desired upper bound. Therefore, it is sufficient to focus on the first summand $\rho(A', \pi_W^{-1}(S'))$. Fact 3.4 implies that

$$\rho(A', \pi_W^{-1}(S')) = \rho(\pi_W(A'), S').$$

By condition (1), the basis packing number of the sequence A' is at least αm, thus the same holds for $\pi_W(A')$ (as a sequence of vectors in W). Let m' be the size of A', so we have $\alpha m \leqslant m' \leqslant m$. Condition (2) then implies that $\rho(\pi_W(A')) \geqslant m^{-C} \geqslant (m')^{-C'}$ for some constant $C' = C'(C, k)$. Therefore, we can apply Theorem 4.1 to the sequence $\pi_W(A')$ and the variety S' with $\mathbb{F} = \mathbb{C}$, $\delta = \alpha$ and $C_1 = k + 1$. As a result, we obtain a positive integer $r = O_{d,k}(1)$ such that

$$\rho(\pi_W(A'), S') \leqslant O_{d,k}\Big(d_{S'}(\sqrt{m' \log m'}) \cdot (\log m')^r + (m')^{-(k+1)}\Big).$$

Since $\alpha b \leqslant \alpha m \leqslant m' \leqslant m = bk$ and the density function does not "jump too much" by Proposition 3.8, we conclude that

$$\rho(\pi_W(A'), S') \leqslant O_{d,k}\Big(d_{S'}(\sqrt{b \log b}) \cdot (\log b)^r + b^{-(k+1)}\Big). \tag{18}$$

Again, $b^{-(k+1)}$ is small compared to the desired upper bound. The expression in the first summand is quite similar to the one in the statement of the theorem, except that it involves the density function of S' instead of S. We consider two cases depending on whether $\pi_W^{-1}(S') = S$ or not.

First, suppose that $\pi_W^{-1}(S') = S$. Then, by Proposition 3.7, we have $d_{S'} = d_S$. Substituting this into (18) gives the desired bound, completing the proof in this case.

Otherwise, we have $\pi_W^{-1}(S') \subsetneq S$. As S is irreducible, we combine this with Fact 6.2 to conclude that

$$\operatorname{codim} S' = \operatorname{codim}(\pi_W^{-1}(S')) \geqslant \operatorname{codim} S + 1 \geqslant k - \ell + 1.$$

Recall that degree of S' is at most d^k. Then from the Schwartz–Zippel lemma (Proposition 6.6) we have

$$d_{S'}(\sqrt{b \log b}) \leqslant O_{d,k}\Big((b \log b)^{-(k-\ell+1)/2}\Big).$$

Again, substituting this into (18) completes the proof. □

Proof of Theorem 1.4. The probability $\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S]$ is bounded by $\rho(A, S)$, and we would like to prove that

$$\rho(A, S) \leqslant O_{d,k}\big(b^{-(k-\ell+1-1/d)} (\log b)^{C_{d,k}}\big).$$

By Theorem 8.1, for some $r = O_{d,k}(1)$ we have

$$\rho(A, S) \leqslant O_{d,k}\Big(\Big(d_S(\sqrt{b \log b}) + (b \log b)^{-(k-\ell+1)/2}\Big) \cdot (\log b)^r\Big), \tag{19}$$

Recall that $S \subseteq \mathbb{C}^k$ is an irreducible variety of dimension ℓ and degree d. Then, by Pila's bound (Theorem 6.7),

$$d_S(\sqrt{b \log b}) = \sup_{\varphi}\left(\frac{N_{\varphi(S)}(\sqrt{b \log b})}{(2\lfloor \sqrt{b \log b} \rfloor + 1)^k}\right) \leqslant O_{d,k}\Big(b^{-(k-\ell+1-1/d)/2} (\log b)^{C_d - (k-\ell+1-1/d)/2}\Big).$$

Substituting this into (19) gives $\rho(A, S) \leqslant O_k\big(b^{-(k-l+1-1/d)/2} (\log b)^{C_d + r - (k-l+1-1/d)/2}\big)$, completing the proof. □

Next, we deduce Theorem 1.3 by combining Theorem 1.4 with the following result of Ferber, Jain and Zhao [15] (which is a refined version of Halász' theorem [18]).

Theorem 8.2 ([15, Theorem 1.11])**.** *Let A be a sequence of vectors in $\mathbb{R}^k$, and let $I_1, \ldots, I_s$ (for some even s) be a partition of $[n]$. Denote $t := \frac{1}{s}\sum_{j=1}^{s} \dim_{\mathbb{R}} \operatorname{span}_{\mathbb{R}}\{a_i : i \in I_j\}$. Then*

$$\sup_{x \in \mathbb{R}^k} \mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n = x] \leqslant \left(2^{-s}\binom{s}{s/2}\right)^t.$$

Proof of Theorem 1.3. Let S_1 be an irreducible component of S. If $\deg S_1 \geqslant 2$, we apply Theorem 1.4 to conclude that

$$\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S_1] \leqslant O_{d,k}\Big(b^{-(k - \dim S_1 + 1 - 1/\deg S_1)/2} (\log b)^C\Big).$$

for some constant C depending only on $\deg S_1$ and k. Since

$$\dim S_1 \leqslant \dim S \leqslant \ell \quad \text{and} \quad 2 \leqslant \deg S_1 \leqslant \deg S \leqslant d,$$

this bound is at most $O_{d,k}(b^{-(k-\ell)/2})$. This completes the proof in this case.

If $\deg S_1 = 1$ then S_1 is an affine subspace of dimension at most ℓ (by enlarging it, we may assume that its dimension is exactly ℓ). A minor technical issue we need to handle is that Theorem 8.2 is stated only for point probabilities and only over the field $\mathbb{R}$.

Let V be the linear subspace which is a translate of S_1, and let $\pi : \mathbb{C}^k \to \mathbb{C}^k/V \simeq \mathbb{C}^{k-\ell} \simeq \mathbb{R}^{2(k-\ell)}$ be the quotient map. By assumption, there is a partition $I_1, \ldots, I_{b_1}$ of $[n]$ with $b_1 = 2\lfloor b/2 \rfloor \leqslant b$, such that each subset of vectors $\{a_i : i \in I_j\}$ contains a basis of $\mathbb{C}^k$. Then, for any $1 \leqslant j \leqslant b_1$,

$$\dim_{\mathbb{R}} \operatorname{span}_{\mathbb{R}}\{\pi(a_i) : i \in I_j\} \geqslant \dim_{\mathbb{C}} \operatorname{span}_{\mathbb{C}}\{\pi(a_i) : i \in I_j\} = k - \ell.$$

Applying Theorem 8.2 to this partition, we conclude that

$$\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S_1] \leqslant \sup_{x \in \mathbb{C}^k/V} \mathbb{P}[\xi_1 \pi(a_1) + \ldots + \xi_n \pi(a_n) = x] \leqslant \left(2^{-b_1}\binom{b_1}{b_1/2}\right)^{k-\ell} \leqslant O(b^{-(k-\ell)/2}).$$

Summing over all irreducible components of S (there are at most d of them) completes the proof. □

Next, we deduce Theorems 1.1, 1.2 and 1.8 from the results proved previously in this section. All these deductions share the same first step, which is an application of the "dropping to a subspace" lemma (Lemma 3.5).

Proof of Theorem 1.1. As F has Chow rank at most c, it can be written as

$$F(t_1, \ldots, t_n) = f(L_1(t_1, \ldots, t_n), \ldots, L_k(t_1, \ldots, t_n))$$

for $k = dc$, some $f \in \mathbb{C}[x_1, \ldots, x_k]$ and homogeneous linear forms $L_1, \ldots, L_k$. Suppose that the form L_i is given by $a_{i1}t_1 + \ldots + a_{in}t_n$ for some coefficients $a_{ij} \in \mathbb{C}$. Denoting $a_j = (a_{1j}, \ldots, a_{kj}) \in \mathbb{C}^k$, we have

$$F(t_1, \ldots, t_n) = f(t_1 a_1 + \ldots + t_n a_n).$$

Let $b_0 = \lfloor b/(k(k+1)) \rfloor + 1$. Then by Lemma 3.5 applied to the sequence $A = (a_1, \ldots, a_n)$, there exists a subsequence $A' = A[I_0]$ of size at least

$$n - (b_0 - 1)\frac{k(k+1)}{2} \geqslant n - b/2,$$

and a subspace $V' \subseteq \mathbb{C}^k$ (of dimension $k' \leqslant k$) such that all the elements of A' lie in V', and A' has basis packing number at least b_0 (as a sequence of vectors in V').

Let $I_1 = [n] \setminus I_0$, so $|I_1| \leqslant b/2 < b$. It suffices to show that for an arbitrary outcome of the Rademacher random variables $(\xi_i)_{i \in I_1}$, if we condition on $(\xi_i)_{i \in I_1}$ taking this particular outcome, then the desired bounds on $\mathbb{P}[F(\xi_1, \ldots, \xi_n) = 0]$ hold in the resulting conditional probability space. In other words, let F_* be a polynomial obtained by an arbitrary substitution of ± 1 instead of the variables $(t_i)_{i \in I_1}$; it suffices to prove the desired bounds with "F_*" in place of "F".

Note that we can write

$$F_*((t_i)_{i \in I_0}) = f_*\left(\sum_{i \in I_0} t_i a_i\right) \tag{20}$$

for some polynomial f_* defined on V' (one can take $f_*(x)$ to be a restriction of $f(x + x_0)$ to V', for certain $x_0 \in \mathbb{C}^k$). Therefore, we need to estimate the probability that $f_*\left(\sum_{i \in I_0} \xi_i a_i\right) = 0$.

Let $S \subseteq V' \simeq \mathbb{C}^{k'}$ be the variety defined by f_*.

(1) First, we prove Theorem 1.1(1). By our assumption, F_* is not identically zero. Since $b_0 \geqslant 1$, the vectors $(a_i)_{i \in I_0}$ span the vector space V', and thus the polynomial f_* is also not identically zero. By Fact 6.3 (applied to each irreducible factor of f_* over $\mathbb{C}$) combined with Fact 6.1, S has dimension $k' - 1$ and degree at most $\deg f_* \leqslant \deg f = d$.

As the basis packing number of the sequence A' is at least b_0, from Theorem 1.3 we obtain that

$$\mathbb{P}\left[f_*\left(\sum_{i \in I_0} \xi_i a_i\right) = 0\right] = \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in S\right] \leqslant O_{d,k}(b_0^{-1/2}).$$

Since $b_0 = \Omega_k(b)$, this completes the proof.

(2) Next, we prove Theorem 1.1(2). We may assume that $b > 2d$ (otherwise, the desired probability bound is trivial). Let $F_*^{=d}$ be the homogeneous degree-d part of F_*. The polynomial F_*, in turn, was obtained from F by substitution of ± 1's instead of at most $b/2$ of its variables. Observe that each variable of F is part of at most n^{d-1} monomials of degree d, and that there are at most dn^{d-1} monomials of degree less than d in total. Therefore, $F_*^{=d} - F$ has at most $(b/2+d)n^{d-1} < bn^{d-1}$ nonzero coefficients. So, by our assumption on F, the polynomial $F_*^{=d}$ is irreducible (in particular, it is not zero). Recalling (22), we conclude that the homogeneous degree-d part $f_*^{=d}$ of f_* is also irreducible and nonzero.

So, f_* is an irreducible polynomial of degree d. Moreover, we note that f_* cannot be written as a polynomial of two linear forms. Indeed, otherwise $f_*^{=d}$ can be represented as $g(L_1, L_2)$ for a homogeneous polynomial $g \in \mathbb{C}[x_1, x_2]$ and two homogeneous linear forms L_1, L_2. Recall that $\mathbb{C}$ is algebraically closed, and let $r_1, \dots, r_{d_*}$ be the complex roots of $g(1, z)$. Then

$$g(L_1, L_2) = L_1^d g(1, L_2/L_1) = z_0 L_1^d \prod_{j=1}^{d_*} (L_2/L_1 - r_j) = z_0 L_1^{d-d_*} \prod_{j=1}^{d_*} (L_2 - r_j L_1) \qquad (\text{for some } z_0 \in \mathbb{C}).$$

So, $f_*^{=d}$ splits into a product of linear forms. As $d \geqslant 1$, this contradicts its irreducibility.

Since the variety $S \subseteq \mathbb{C}^{k'}$ is defined by the irreducible polynomial f_* of degree d, it has dimension $k'-1$ and degree d by Fact 6.3. Recall that the basis packing number of the sequence A' is at least b_0. Applying Theorem 8.1 to A' and S, we obtain that

$$\mathbb{P}\left[f_*\left(\sum_{i\in I_0} \xi_i a_i\right) = 0\right] = \mathbb{P}\left[\sum_{i\in I_0} \xi_i a_i \in S\right] \leqslant O_{d,k}\Big(\Big(d_S(\sqrt{b_0 \log b_0}) + (b_0 \log b_0)^{-1}\Big) \cdot (\log b_0)^r\Big) \qquad (21)$$

for some $r = O_{d,k}(1)$. Since f_* is an irreducible polynomial of degree $d \neq 3$ which cannot be represented as a polynomial of two linear forms, we can estimate the density function d_S using the result of Vermeulen and Browning–Gorodnik (Theorem 6.8). Namely, we conclude that for any $\varepsilon > 0$

$$d_S\Big(\sqrt{b_0 \log b_0}\Big) \leqslant O_{d,k,\varepsilon}\big(b_0^{-1+\varepsilon}\big).$$

Since $b_0 = \Omega_k(b)$, substituting this into (21) completes the proof. □

Remark 8.3. In fact, the proof of Theorem 1.1(2) implies the following slightly stronger statement. Let $F \in \mathbb{C}[t_1, \dots, t_n]$ be a polynomial of degree $d \neq 3$ and Chow rank at most c. Suppose that after any substitution of ± 1's instead of fewer than b variables of F, the resulting polynomial is irreducible of degree d and cannot be written as a polynomial of two linear forms. Then for any $\varepsilon > 0$ we have $\mathbb{P}[F(\xi_1, \dots, \xi_n) = 0] = O_{d,k,\varepsilon}(b^{-1+\varepsilon})$.

Proof of Theorem 1.2. The first half of the proof is almost identical to the first half of the proof of Theorem 1.1. As F has Chow rank at most c (over $\mathbb{F}$), it can be written as

$$F(t_1, \dots, t_n) = f(L_1(t_1, \dots, t_n), \dots, L_k(t_1, \dots, t_n))$$

for $k = dc$, some $f \in \mathbb{F}[x_1, \dots, x_k]$ and homogeneous linear forms $L_1, \dots, L_k$ with coefficients in $\mathbb{F}$. Suppose that the form L_i is given by $a_{i1}t_1 + \dots + a_{in}t_n$ with $a_{ij} \in \mathbb{F}$. Denoting $a_j = (a_{1j}, \dots, a_{kj}) \in \mathbb{F}^k$, we have

$$F(t_1, \dots, t_n) = f(t_1 a_1 + \dots + t_n a_n).$$

Let $b_0 = \lfloor b/(k(k+1)) \rfloor + 1$. Then by Lemma 3.5 applied to the sequence $A = (a_1, \dots, a_n)$, there exists a subsequence $A' = A[I_0]$ of size at least $n - b/2$ and a subspace $V' \subseteq \mathbb{F}^k$ (of dimension $k' \leqslant k$) such that all the elements of A' lie in V', and A' has basis packing number at least b_0 (as a sequence of vectors in V').

Let $I_1 = [n] \setminus I_0$, $|I_1| \leqslant b/2 < b$. It suffices to show that for an arbitrary outcome of the Rademacher random variables $(\xi_i)_{i\in I_1}$, if we condition on $(\xi_i)_{i\in I_1}$ taking this particular outcome, then the desired bounds on

$\mathbb{P}[F(\xi_1, \dots, \xi_n) = 0]$ hold in the resulting conditional probability space. In other words, let F_* be a polynomial obtained by an arbitrary substitution of ± 1 instead of variables $(t_i)_{i \in I_1}$; it suffices to prove the desired bounds with "F_*" in place of "F".

Note that we can write

$$F_*((t_i)_{i \in I_0}) = f_*\left(\sum_{i \in I_0} t_i a_i\right) \tag{22}$$

for some polynomial f_* with coefficients in $\mathbb{F}$ defined on V' (one can take $f_*(x)$ to be a restriction of $f(x + x_0)$ to V', for certain $x_0 \in \mathbb{F}^k$). Therefore, we need to estimate the probability that $f_*\left(\sum_{i \in I_0} \xi_i a_i\right) = 0$.

By our assumption, F_* is irreducible (over $\mathbb{F}$) of degree d. Since $b_0 \geqslant 1$, the vectors $(a_i)_{i \in I_0}$ span the vector space V', and thus f_* is also irreducible (over $\mathbb{F}$) of degree d. First, we consider the case when f_* is, furthermore, irreducible over $\mathbb{C}$.

As we have $V' \simeq \mathbb{F}^{k'} \subseteq \mathbb{C}^{k'}$, let $S \subseteq \mathbb{C}^{k'}$ be the variety defined by f_*. By Fact 6.3, it is an irreducible variety of dimension $k' - 1$ and degree d. As the basis packing number of A' (as a sequence of vectors in V') is at least b_0, we apply Theorem 1.4 to conclude that

$$\mathbb{P}\left[f_*\left(\sum_{i \in I_0} \xi_i a_i\right) = 0\right] = \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in S\right] \leqslant O_{d,k}\left(b_0^{-1+\frac{1}{2d}} (\log b_0)^{C_{d,k}}\right).$$

Since $b_0 = \Omega_k(b)$, this completes the proof in this case.

Next, we consider the case when f_* is reducible over $\mathbb{C}$. Let g be any irreducible (over $\mathbb{C}$) factor of f_*. Since f_* is irreducible over $\mathbb{F}$, g is not proportional to a polynomial with coefficients in $\mathbb{F}$. Thus, by Proposition 6.9, there exists a variety $T \subseteq \mathbb{C}^{k'}$ of dimension at most $k' - 2$ and degree at most d^2 such that $T \cap \mathbb{F}^{k'} = S_g \cap \mathbb{F}^{k'}$ (where $S_g \subseteq \mathbb{C}^{k'}$ is the variety defined by g). Note that $\sum_{i \in I_0} \xi_i a_i$ always lies in $\mathbb{F}^{k'}$. As the basis packing number of A' is at least b_0, from Theorem 1.3 we conclude that

$$\mathbb{P}\left[g\left(\sum_{i \in I_0} \xi_i a_i\right) = 0\right] = \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in S_g\right] = \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in T\right] \leqslant O_{d,k}(b_0^{-1}).$$

Since $b_0 = \Omega_k(b)$, this is at most $O_{d,k}(b^{-1})$. Taking the sum over all irreducible (over $\mathbb{C}$) factors of f_* completes the proof. □

Before proceeding to the proof of Theorem 1.8, we record the following proposition about the Zariski closure of a semialgebraic set. Our main reference for properties of semialgebraic sets is the notes of Coste [9].

Proposition 8.4. *Let $S \subseteq \mathbb{R}^k \subseteq \mathbb{C}^k$ for $k \geqslant 1$ be a semialgebraic set which does not contain a line segment. Then the Zariski closure of S in $\mathbb{C}^k$ is a variety of dimension at most $k - 1$ (equivalently, this Zariski closure is not the whole $\mathbb{C}^k$).*

Proof. The statement follows from these three facts:

- [9, Proposition 3.15] The dimension of S as a semialgebraic set is defined as the maximum dimension of a cell in its cell decomposition. A cell of dimension k is homeomorphic to $(0, 1)^k$, and therefore contains a line segment. Therefore, the dimension of S as a semialgebraic set is at most $k - 1$.
- [9, Theorem 3.20] The dimension of S as a semialgebraic set is equal to the dimension of its *real* Zariski closure $S_{\mathbb{R}}$ (as a *real* algebraic set).
- The dimension of $S_{\mathbb{R}}$ as a *real* algebraic set is equal to the dimension of its *complex* Zariski closure $S_{\mathbb{C}}$ (as a *complex* algebraic variety). Since the dimension of a variety is equal to the Krull dimension of its coordinate ring [19, Proposition 1.7], this can be seen (for example) from the Noether normalization lemma. □

Proof of Theorem 1.8. Let $b_0 = \lfloor n/(k(k+1)) \rfloor + 1$. Then by Lemma 3.5 applied to the sequence $A = (a_1, \dots, a_n)$, there exists a subsequence $A' = A[I_0]$ of size at least $n/2$ and a subspace $V' \subseteq \mathbb{R}^k$ such that all the elements of A' lie in V', and A' has basis packing number at least b_0 (as a sequence of vectors in V'). Conditioning on the outcomes of the random variables $(\xi_i)_{i \in [n] \setminus I_0}$, we have

$$\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S] \leqslant \sup_{x \in \mathbb{R}^k} \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in (S - x) \cap V'\right]. \tag{23}$$

As $V' \subseteq \mathbb{R}^k \subseteq \mathbb{C}^k$, define $V'_{\mathbb{C}} \subseteq \mathbb{C}^k$ as be the minimal complex subspace containing V'. It satisfies $V'_{\mathbb{C}} \cap \mathbb{R}^k = V'$ and $\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V'$.

Fix any $x \in \mathbb{R}^k$. Note that $(S - x) \cap V'$ is a semialgebraic set which does not contain a line segment (since S does not contain a line segment). Then, by Proposition 8.4, the complex Zariski closure $S_{\mathbb{C}} \subseteq V'_{\mathbb{C}}$ of $(S-x) \cap V'$ has dimension at most $\dim V'_{\mathbb{C}} - 1$. Recall that the basis packing number of the sequence A' is at least b_0. Applying Theorem 1.3 to A' and $S_{\mathbb{C}}$, we conclude that

$$\mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in (S - x) \cap V'\right] \leqslant \mathbb{P}\left[\sum_{i \in I_0} \xi_i a_i \in S_{\mathbb{C}}\right] \leqslant O_{\deg(S_{\mathbb{C}}),k}(b_0^{-1/2}).$$

Since $b_0 = \Omega_k(n)$, substituting this into (23) completes the proof. □

Remark 8.5. The bound in Theorem 1.8 is sharp up to a multiplicative constant factor: if $S = \{0\}$, and $a_1 = \ldots = a_{2n} \in \mathbb{R}^k \setminus \{0\}$, then $\mathbb{P}[\xi_1 a_1 + \ldots + \xi_{2n} a_{2n} \in S] = \binom{2n}{n}/2^{2n} = \Theta(n^{-1/2})$. However, we sketch how this bound can be improved under additional conditions on the vectors $a_1, \dots, a_n$.

- Suppose that the vectors $a_1, \dots, a_n \in \mathbb{R}^k$ in Theorem 1.8 satisfy the following condition: for some $\delta > 0$ every line passing through the origin contains fewer than $(1-\delta)n$ of them (this happens, for example, when $k \geqslant 2$ and one can form δn disjoint bases from these vectors). Then one can modify the proof above to ensure that $\dim V'_{\mathbb{C}} \geqslant 2$. Since S does not contain a line segment, a simple argument based on Proposition 8.4 implies that no irreducible component of its Zariski closure $S_{\mathbb{C}} \subseteq V'_{\mathbb{C}}$ is a hyperplane in $V'_{\mathbb{C}}$. Therefore, applying Theorem 1.4 instead of Theorem 1.3, we obtain a stronger bound

$$\mathbb{P}[\xi_1 a_1 + \ldots + \xi_n a_n \in S] \leqslant O_{S,\delta}(n^{-3/4} (\log n)^{C_k}).$$

- Furthermore, suppose that the vectors $a_1, \dots, a_n \in \mathbb{R}^k$ satisfy the following condition: for some $\delta > 0$ every two-dimensional linear subspace contains fewer than $(1-\delta)n$ of them. In this case one can similarly ensure that $\dim V'_{\mathbb{C}} \geqslant 3$, and that each irreducible component of the Zariski closure $S_{\mathbb{C}} \subseteq V'_{\mathbb{C}}$ either has codimension at least two, or is not a preimage of a curve under a linear map. Applying Theorem 8.1 combined with the best known results on the affine dimension growth conjecture (which is settled for $d \neq 3$, but has only partial results for $d = 3$) stated for varieties "not cylindrical over a curve" [37, Theorem 1.2], we can get a bound of roughly $O_{S,\delta}(n^{-0.92})$.

9 Concluding remarks

In this paper we have introduced a general method to study certain geometric variants of the Littlewood–Offord problem via lattice point counting, and applied this method in several different contexts. There are a number of interesting directions for future research.

First, regarding the general polynomial Littlewood–Offord problem: one of our primary motivations to consider the bounded-rank setting was that this setting already seems to incorporate many of the most important difficulties of the general polynomial Littlewood–Offord problem. Indeed, the resolution of the quadratic Littlewood–Offord problem by Kwan and Sauermann [25] was accomplished by first solving the bounded-rank case, and then adapting and quantifying the approach for the general case.

Unfortunately, it seems challenging to adapt the techniques in this paper to general polynomials (without a bound on the Chow rank). We remark that we do not really need the Chow rank to be $O(1)$: indeed, it should be straightforward to modify our proof of Theorem 1.1(1) (by using Theorem 1.4 and the Erdős–Littlewood–Offord theorem instead of Theorem 1.3) to show that there is a slowly growing function $h(n)$ such that when the Chow rank of F is at most $h(n)$ then

$$\mathbb{P}[F(\xi_1, \dots, \xi_n) = 0] \leqslant O_d(b^{-1/2}),$$

with the implicit constant not depending on the Chow rank. However, $h(n)$ would definitely need to grow rather slowly (e.g., it seems significant new ideas would be required to handle Chow rank as large as $n^{0.01}$).

It may also be fruitful to consider different ("weaker") notions of rank/complexity than Chow rank. Indeed, while there is only really one sensible notion of rank for quadratic polynomials, for polynomials of higher degree there are several fundamentally different notions of rank. One natural candidate that often arises in analytic number theory is the *Schmidt rank* (also called *h-invariant* or *strength*): for a polynomial F of degree d, its Schmidt rank is the smallest integer s such that F can be written as $\sum_{i=1}^{s} P_i$, where each P_i is a product of two polynomials of degree strictly less than d (for a homogeneous polynomial of a fixed degree d, its Schmidt rank is equivalent to the so-called *partition rank* of its coefficient tensor).

Finally, another interesting direction is to consider "small-ball concentration" probabilities instead of "point concentration" probabilities. Namely, assuming that "sufficiently many" coefficients of F have absolute value at least 1, one can sometimes obtain similar upper bounds on $\mathbb{P}[|F(\xi_1, \dots \xi_n)| \leqslant 1]$ (see [18, 27, 15]). It seems difficult to adapt our methods to this setting. One reason for this is that the decoupling techniques used in Section 7 are not well-suited for small-ball probabilities. Another reason is that we are not aware of appropriate number-theoretic results (analogous to Theorem 6.7) sufficient to finish the proof in the small-ball setting.

References

[1] George E. Andrews. A lower bound for the volume of strictly convex bodies with many boundary lattice points. *Trans. Amer. Math. Soc.*, 106:270–279, 1963.

[2] Imre Bárány and David G. Larman. The convex hull of the integer points in a large ball. *Math. Ann.*, 312(1):167–181, 1998.

[3] Thomas F. Bloom and Jared Duker Lichtman. The Bombieri–Pila determinant method. December 2023. Preprint, arXiv:2312.12890.

[4] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[5] T. D. Browning and A. Gorodnik. Power-free values of polynomials on symmetric varieties. *Proc. Lond. Math. Soc. (3)*, 114(6):1044–1080, 2017.

[6] T. D. Browning, D. R. Heath-Brown, and P. Salberger. Counting rational points on algebraic varieties. *Duke Math. J.*, 132(3):545–578, 2006.

[7] Boris Bukh and Jacob Tsimerman. Sum-product estimates for rational functions. *Proc. Lond. Math. Soc. (3)*, 104(1):1–26, 2012.

[8] Alex Cohen and Guy Moshkovitz. Partition and analytic rank are equivalent over large fields. *Duke Math. J.*, 172(12):2433–2470, 2023.

[9] Michel Coste. An introduction to semialgebraic geometry, 2000.

[10] Kevin P. Costello. Bilinear and quadratic variants on the Littlewood-Offord problem. *Israel J. Math.*, 194(1):359–394, 2013.

[11] Kevin P. Costello, Terence Tao, and Van Vu. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006.

[12] Kevin P. Costello and Van H. Vu. The rank of random graphs. *Random Structures Algorithms*, 33(3):269–285, 2008.

[13] Victor De la Pena and Evarist Giné. *Decoupling: from dependence to independence*. Springer Science & Business Media, 1999.

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

[15] Asaf Ferber, Vishesh Jain, and Yufei Zhao. On the number of Hadamard matrices via anti-concentration. *Combin. Probab. Comput.*, 31(3):455–477, 2022.

[16] Jacob Fox, Matthew Kwan, and Hunter Spink. Geometric and o-minimal Littlewood-Offord problems. *Ann. Probab.*, 51(1):101–126, 2023.

[17] William Fulton. *Intersection Theory*, volume 3 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, 1984.

[18] G. Halász. Estimates for the concentration function of combinatorial number theory and probability. *Period. Math. Hungar.*, 8(3-4):197–211, 1977.

[19] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1977.

[20] D. R. Heath-Brown. Cubic forms in ten variables. *Proc. London Math. Soc. (3)*, 47(2):225–257, 1983.

[21] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[22] Vojtéch Jarník. Über die Gitterpunkte auf konvexen Kurven. *Math. Z.*, 24(1):500–518, 1926.

[23] Zhihan Jin, Matthew Kwan, Lisa Sauermann, and Yiting Wang. Algebraic aspects of the polynomial Littlewood–Offord problem. May 2025. Preprint, arXiv:2505.23335.

[24] Daniel M. Kane. The correct exponent for the Gotsman-Linial conjecture. *Comput. Complexity*, 23(2):151–175, 2014.

[25] Matthew Kwan and Lisa Sauermann. Resolution of the quadratic Littlewood–Offord problem. December 2023. Preprint, arXiv:2312.13826.

[26] 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.*, 12/54:277–286, 1943.

[27] Raghu Meka, Oanh Nguyen, and Van Vu. Anti-concentration for polynomials of independent random variables. *Theory Comput.*, 12:Paper No. 11, 16, 2016.

[28] Hoi Nguyen and Van Vu. Optimal inverse Littlewood-Offord theorems. *Adv. Math.*, 226(6):5298–5319, 2011.

[29] Hoi H. Nguyen and Van H. Vu. Small ball probability, inverse theorems, and applications. In *Erdös centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 409–463. János Bolyai Math. Soc., Budapest, 2013.

[30] J. Pila. Density of integral and rational points on varieties. Number 228, pages 4, 183–187. 1995. Columbia University Number Theory Seminar (New York, 1992).

[31] J. Pila. Density of integer points on plane algebraic curves. *Internat. Math. Res. Notices*, (18):903–912, 1996.

[32] Alexander Razborov and Emanuele Viola. Real advantage. *ACM Trans. Comput. Theory*, 5(4):Art. 17, 8, 2013.

[33] Jan Rosiński and Gennady Samorodnitsky. Symmetrization and concentration inequalities for multilinear forms with applications to zero-one laws for Lévy chaos. *Ann. Probab.*, 24(1):422–437, 1996.

[34] Per Salberger. Counting integral points of affine hypersurfaces, 2023.

[35] Terence Tao and Van Vu. A sharp inverse Littlewood-Offord theorem. *Random Structures Algorithms*, 37(4):525–539, 2010.

[36] Terence Tao and Van H. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)*, 169(2):595–632, 2009.

[37] Floris Vermeulen. Dimension growth for affine varieties. *Int. Math. Res. Not. IMRN*, (15):11464–11483, 2024.

[38] Paul B. Yale. Automorphisms of the Complex Numbers. *Math. Mag.*, 39(3):135–141, 1966.