Poissonian correlations of αn^d mod 1

Chris Lutsko, Nick Rome, and Niclas Technau

Abstract

Let $x(n) := \alpha n^d \mod 1$ for integer $d > 1$ and non-zero real α. Then, we show that $\{x(n)\}_{n>0}$ has Poissonian ℓ-point correlations for almost all choices of α when d is large (depending on ℓ). This falls in line with the expected behavior from the Berry-Tabor conjecture. Further, in the spirit of a conjecture of Rudnick–Sarnak [RS98], we show Poissonian ℓ-point correlations for a set of badly approximable α of full Hausdorff dimension by a Fourier analytic transference principle.

The proof makes use of an application of the determinant method to count points on a diagonal hypersurface of degree d in such a way as to capture the contribution of points belonging to lower dimensional varieties. As d grows, these 'special solutions' dominate the count and non-special solutions become increasingly rare. This stratified counting statement allows us to control the number of points on average very effectively.

1 Introduction

1.1 Poissonian Correlations of monomials mod 1

Fix a degree, $d \in \mathbb{N}_{>1}$, and a dilation, $\alpha \in \mathbb{R}_{\neq 0}$, and consider the sequence $\mathcal{X}_\alpha^d := \{x(n)\}_{n>0}$, where

$$x(n) = x_{\alpha,d}(n) := \alpha n^d \mod 1.$$

Then, given a test function, $f \in C_c^\infty(\mathbb{R}^{\ell-1})$, we denote the ℓ-point correlation function of $\mathcal{X}_\alpha^d$ by

$$R_\ell(\mathcal{X}_\alpha^d) := \lim_{N\to\infty} R_\ell^N(\mathcal{X}_\alpha^d) := \lim_{N\to\infty} \frac{1}{N} \sum_{\mathbf{n}\in[N]^\ell}^{*} \sum_{\mathbf{k}\in\mathbb{Z}^{\ell-1}} f(N(x(n_1)-x(n_2),\ldots,x(n_{\ell-1})-x(n_\ell))), \tag{1.1}$$

if it exists, where $*$ indicates that the coordinates of $\mathbf{n}$ are distinct. This measures the probability of finding ℓ points of $\{x(n)\}_{n<N}$ in a uniformly thrown $\frac{1}{N}$-neighborhood on the torus, $[0,1]$. A sequence has Poissonian ℓ-point correlation if $R_\ell(\mathcal{X}_\alpha^d)$ converges to $\mathbf{E}(f)$. A well-known theorem of Rudnick and Sarnak showed that

Theorem 1 ([RS98, Theorem 1])**.** *For $d \in \mathbb{Z}_{>1}$ and almost all choices of α, the sequence $\mathcal{X}_\alpha^d$ has Poissonian pair ($\ell = 2$) correlation.*

Moreover, the theorem is true if d is a non-zero real [RT22, AEBM21]. Our main result is the following extension of Theorem 1 to higher order correlation functions.

Theorem 2. *For almost all choices of α, the sequence $\mathcal{X}_\alpha^d$ has Poissonian ℓ-point correlation for all d satisfying*

$$d > (2\ell)^{4\ell} =: d_\ell. \tag{1.2}$$

Theorem 1 is one of the seminal results in the field. When $d = 2$, this is particularly interesting since the values $x(n)$ represent the energy levels for the semi-classical boxed harmonic oscillator. Thus, Theorem 1 provided (almost everywhere) evidence towards the Berry-Tabor conjecture which connects the local statistics of such energy levels to the dynamical properties of the underlying classical system (i.e integrable vs chaotic). For $d > 2$, while one can produce a Hamiltonian which gives these eigenvalues, we are unable to connect these Hamiltonians to classical systems.

Aside from this application to quantum chaos, building on work of van der Corput among many others (see [KN74] for history), it is very common to study the local statistics of sequences. A key challenge is to show that the correlations of such sequences converge to the Poissonian limit. In fact, it has long been expected (see for example [Mar00, Rud08, RS98, RSZ01, RS24]) that for almost any

MSC2020: 11K06, 11L07, 15B48

Key words and phrases: Local Statistics; Sequences Modulo 1; Determinant method

choice of α the sequence $x(n)$ should have Poissonian local statistics for any $d > 2$. However, for $\ell > 2$, while there are results for real valued sequences [ABR24, BL23, Mar03, EMM05], for sequences modulo 1 very little progress has been made. In part, that is the case because higher ℓ-point correlation functions do not concentrate in an L^2-sense in many cases, precluding canonical concentrations arguments right from the start. This was first observed by Rudnick and Sarnak [RS98]; see also [TW22, Appendix C] for more discussion.

When the exponent d is a small real (e.g $\leq 1/3$) the correlations are known to be Poissonian [LST25, LT21, LT25] and when the sequence grows with a lacunary rate then the sequence is known to have Poissonian statistics almost everywhere [RZ02]. Hence, Theorem 2 is the only direct progress towards this central conjecture in the field since [RS98] from the late 90's.

The following extension of Theorem 2 shows that we can replace the Lebesgue measure on $[0, 1)$ with any measure satisfying a growth condition on its Fourier coefficients.

Theorem 3. *Let $d > d_{4\ell}$. Suppose μ is a finite Borel measure on $[0, 1]$ whose Fourier transform $\widehat{\mu}(\xi) = \int_{\mathbb{R}} e(-\xi\alpha)d\mu(\alpha)$ decays on average like $|\xi|^{-1/2}$ in the following sense. If $0 < s < 1$, then*

$$|\widehat{\mu}(0)|^2 + \sum_{u\neq 0} |\widehat{\mu}(u)|^2 \cdot |u|^{s-1} < \infty. \tag{1.3}$$

Then, $\mathcal{X}_\alpha^d$ has Poissonian ℓ-point correlations for μ-almost any $\alpha \in [0, 1]$ as soon as $d > d_{4\ell}$.

We say θ is diophantine of type τ if there exists a c_θ such that $|\theta - \frac{p}{q}| \geq c_\theta q^{-\tau}$ for all relatively prime p and q. It is immediate that the diophantine properties of α will dictate the behavior of the correlations (see for example [RSZ01, Theorem 2]). Rudnick, Sarnak and Zaharescu [RSZ01, p. 38] conjecture that any diophantine number with sufficiently square-free denominators in its continued fraction convergents should lead to Poissonian behavior of the gaps in $\mathcal{X}_\alpha^2$. An application of Theorem 3 shows that for almost every diophantine α, the sequence $\mathcal{X}_\alpha^d$ will have Poissonian ℓ-correlations for $d > d_\ell$.

In fact, while Theorem 2 applies to a much larger set of α, Rudnick and Sarnak [RS98, Abstract] conjecture that, if α is badly approximable and $d \geq 2$ is a fixed integer, then the normalized spacings between elements of $\mathcal{X}_\alpha^d$ have 'Poissonian statistics'. While there have been some steps towards this conjecture (see e.g [HB10]), no one has confirmed it for any degree or statistic. The following corollary states the existence of a 'large' subset of badly approximable numbers which satisfy this conjecture with respect to the ℓ-point correlation function.

Corollary 4. *Fix $\ell > 1$ and let $d > d_{4\ell}$. Then there exists a Hausdorff dimension 1 subset, $\mathcal{A}$ of the badly approximable numbers such that $\mathcal{X}_\alpha^d$ has Poissonian ℓ-point correlations for all $\alpha \in \mathcal{A}$.*

Proof. The s-energy

$$I_s(\mu) = \iint_{[0,1]^2} |x - y|^{-s} d\mu(x)d\mu(y)$$

of μ is well-known to be comparable to its discretization (1.3), see [HMR07, Corollary 2.7]. Hence, Theorem 3 and Frostman's Lemma implies directly that any set $\mathcal{S} \subseteq [0, 1]$ of Hausdorff dimension 1 contains a set $\mathcal{G} \subseteq \mathcal{S}$ so that $\mathcal{X}_\alpha^d$ has Poissonian ℓ-point correlations for any $\alpha \in \mathcal{G}$ once $d > d_{4\ell}$. Choosing $\mathcal{S}$ to be the set of badly approximable numbers proves the corollary.

□

1.2 Stratified point counting

Fix a degree $d \geq 2$ and dimension n, and consider the homogeneous diagonal form

$$Q_{\mathbf{a}}(\mathbf{x}) := a_0 x_0^d + \cdots + a_n x_n^d \tag{1.4}$$

with coefficient vector $\mathbf{a} \in (\mathbb{Z}_{\neq 0})^{n+1}$. A natural question in diophantine analysis is to count solutions to the equation $Q_{\mathbf{a}}(\mathbf{x}) = 0$ with entries taken from a box $[0, N]^{n+1}$. To that end, let $X_n \subset \mathbb{P}^n$ denote the diagonal hypersurface defined by $Q_{\mathbf{a}}(\mathbf{x}) = 0$. For any point, $x \in X_n(\mathbb{Q})$, we take a representative $\mathbf{x} = (x_0, \dots, x_n) \in \mathbb{Z}^{n+1}_{\text{prim.}}$ and we write

$$H(x) = ||\mathbf{x}||_\infty.$$

Define the counting function

$$\mathcal{N}_{Q_{\mathbf{a}}}(N) := \#\{x \in X_n(\mathbb{Q}) : H(x) \leq N\}.$$

For general hypersurfaces, Browning–Heath-Brown [BHB06] showed that

$$\mathcal{N}_{Q_{\mathbf{a}}}(N) = O_{d,n}(N^{n-1+\varepsilon}) \quad \text{for all } \varepsilon > 0, \tag{1.5}$$

uniformly in $\mathbf{a}$. However, one expects that the diagonal structure of the form should allow for improvements. Indeed, Salberger-Wooley [SW10] showed that once d is large, diagonal solutions dominate. In particular, once $d > (2n+2)^{4n+4}$, this provides the bound

$$\mathcal{N}_{Q_{\mathbf{a}}}(N) = O_{d,n}(N^{\lfloor \frac{n+1}{2} \rfloor + \varepsilon}) \quad \text{for all } \varepsilon > 0. \tag{1.6}$$

The purpose of this paper is to improve this upper bound on $\mathcal{N}_{Q_{\mathbf{a}}}(N)$ by showing that the count depends on the coefficients $\mathbf{a}$ and can be massively reduced for generic[1] $\mathbf{a}$. Before stating the main theorem, let $m = m(\mathbf{a})$ denote the largest $0 \le k \le \frac{n+1}{2}$ such that there exist k disjoint pairs of distinct indices $(i_1, j_1), \ldots, (i_k, j_k)$ with

$$-a_{i_\ell} a_{j_\ell} = \square \quad \forall \ell = 1, \ldots, k$$

where $\square$ denotes the set of squares. Furthermore, let

$$L(d,n) := 1 + \frac{2(d-n+1)}{n^2 - n} \tag{1.7}$$

and let

$$\phi(m,d,n) := \max_{0 \le s \le m} \left(s + \sum_{r=1}^{n-2s-1} \frac{r+1}{\sqrt[r]{L(d,n)}} \right).$$

Then our main theorem states that the count is dominated not just by diagonal terms, but all the "quasi-diagonal" terms enumerated by m.

Theorem 5. *Let $n > d > 3$. Let W denote the subset of points $x \in X_n(\mathbb{Q})$ for which there exists a pair of distinct indices i, j such that*

$$a_i x_i^d + a_j x_j^d = 0.$$

Then, for all $\varepsilon > 0$

$$\#\{x \in X_n(\mathbb{Q}) \setminus W : H(x) \le N\} \ll N^{\sum_{r=1}^{n-1} \frac{r+1}{L(d,n)^{1/r}} + \epsilon}.$$

Moreover, for the total number of points we have the bound

$$\mathcal{N}_{Q_{\mathbf{a}}}(N) = O\left(N^{\phi(m,d,n)+\varepsilon}\right). \tag{1.8}$$

In both bounds, the implied constant is uniform in $\mathbf{a}$.

Remark. In particular, observe that for a fixed n, once $d > d_n$ we have that

$$\#\{x \in X_n(\mathbb{Q}) \setminus W : H(x) \le N\} = o(N),$$

demonstrating that the quasi-diagonal solutions dominate the count in a very strong manner. Similarly, another immediate consequence is that for any $\epsilon > 0$ there exists a $d(\epsilon)$ such that $d > d(\epsilon)$ yields

$$N_{Q_{\mathbf{a}}}(N) = O_\epsilon(N^{m+\epsilon}),$$

where m is defined as above.

The condition that $-a_i a_j \in \square$ characterizes when quasi-diagonal solutions coming from linear spaces within the hypersurface arise, and in particular, m is the dimension of the largest linear space within the hypersurface. Note that the exponent ϕ varies significantly with m. In the generic case one expects $m = 0$, hence the exponent is dominated by the r sum, which decays as d gets large and n is fixed. Conversely, in the worst case scenario where the hypersurface contains $\frac{n-1}{2}$ dimensional planes, the first term in the exponent dominates and we recover the asymptotic bound from Salberger-Wooley (1.6).

In the circle method literature, these kinds of paucity results have been established for many diagonal systems of equations (see for instance [Woo23, BW22, BW03, SW97]). One distinction in our approach is that it works in the orthogonal regime, where the degree is much larger than the number of variables. Another is that our bounds allow for any choice of coefficients whereas most previous works handle specific equations with coefficients in $\{\pm 1\}$. These distinctions we have in common with the work of Salberger–Wooley and stem from the principal tool being the determinant method. The major input that allows us to control the rational points on diagonal hypersurfaces outside of linear subspaces is the fact that such points cannot lie on subvarieties of very small degree (c.f. Lemma 10).

[1] In the sense of being drawn uniformly at random from a homogeneous expanding box.

1.3 Additional applications

1.3.1 Matrix counting

Rational points on diagonal hypersurfaces appear in a variety of contexts. For instance, in the representation of a number as a sum of d^{th} powers [HB09]. Another application is to counting matrices. For instance, suppose one wanted to count matrices with small rank. To that end, consider the class of $k \times n$ matrices with degree d, bounded entries:

$$\mathcal{M}_{d,k,n}(N) := \left\{ \begin{pmatrix} x_{1,1}^d & \dots & x_{1,k}^d \\ & \vdots & \\ x_{n,1}^d & \dots & x_{n,k}^d \end{pmatrix} \; : \; x_{i,j} < N \right\}. \tag{1.9}$$

Further, let

$$\mathcal{M}(r, N) = \mathcal{M}_{d,k,n}(r, N) := \{M \in \mathcal{M}_{d,k,n}(N) \; : \; \operatorname{rank}(M) = r\}.$$

Theorem 5 then allows us to provide strong bounds for the number of matrices in $\mathcal{M}(r, N)$.

Theorem 6. *Let, $k > n > 2$ be fixed and let $r < n$. Further, assume $d > d_k$, then for any $\varepsilon > 0$ we have*

$$\#\mathcal{M}(r, M) = O(N^{nr+(r/2+1)(k-r)+\varepsilon}). \tag{1.10}$$

The statistics of matrices whose entries come from arithmetically interesting sets has become a hot topic in number theory and combinatorics in recent years. One early seminal result in this area is [DRS93]. However, there have been numerous improvements and extensions over the years to questions about matrices of fixed rank, determinant, and characteristic polynomial [BL24, GNY17, DHP25, CM25, BL23, MOS24, AKOS25]. Theorem 6 is an improvement on recent work of Mohammadi-Ostafe-Shparlinski [MOS24] in the case of monomial entries of the same degree (although they treat polynomial matrices of any form). For comparison, the exponent therein is very hard to write down succinctly though we can compare (1.10) to $nr + (r - 5/4)(k - r)$.

Furthermore, estimates for $\#\mathcal{M}(r, M)$ play a key role in the analysis of [BL23] where Blomer-Li use these counting estimates to show that the ℓ-point correlations between points on the real line given by diagonal equations in k variables with power d are almost surely Poissonian.

In the context of Blomer-Li, Theorem 6 gives a slightly improved range of validity. However, the regime where our bound produces the best gain over the estimates used therein is when the degree is large. In that regime, we improve the term in Blomer-Li's result coming from the small rank matrices. Counting matrices with full rank is more difficult since there is no diophantine equation coming from the linear combination of rows in the small rank case. Thus, while our results give an improvement to the range of validity, we cannot naïvely fill in this missing range (when d is large). That said, perhaps a more precise application of Theorem 5 and a closer analysis of the integrals arising in Blomer-Li's proof could still fill in this range.

1.3.2 Nearest neighbor spacing distribution

Correlation statistics of a sequence allow us to obtain information about the gap distribution. To define this, we relabel the points $\{x_n\}_{n=1}^N$ so that the labels correspond to the position on the torus $u_1^{(N)} \le u_2^{(N)} \le \cdots \le u_N^{(N)}$. We define the cumulative gap distribution to be the limit (if it exists)

$$P(\mathcal{X}_\alpha^d)(s) = \lim_{N\to\infty} P_N(\mathcal{X}_\alpha^d)(s) := \frac{\#\{i \le N \; : \; u_{i+1}^{(N)} - u_i^{(N)} < s/N\}}{N}.$$

The scaling s/N ensures that this is measuring the gaps on the level of the mean spacing. From Theorem 2 one can derive the following corollary which states that the gap distribution of $\mathcal{X}_\alpha^d$ can be bounded by Taylor approximations of $1 - e^{-x}$. Thus, the gap distribution is approximately Poissonian.

Corollary 7. *Let $K \ge 1$, and $d > d_{2K}$ be integers. For almost all α, we have*

$$\sum_{1\le k\le 2K} (-1)^{k+1}\frac{s^k}{k!} \le \liminf_{N\to\infty} P_N(\mathcal{X}_\alpha^d)(s) \le \limsup_{N\to\infty} P_N(\mathcal{X}_\alpha^d)(s) \le \sum_{1\le k\le 2K-1} (-1)^{k+1}\frac{s^k}{k!} \tag{1.11}$$

for all $s > 0$. Consequently, $\lim_{d\to\infty} P(\mathcal{X}_\alpha^d)(s) = 1 - e^{-s}$ for almost all α.

Proof. (1.11) is a direct consequence of Theorem 2 and known relations between the gap distribution and correlation functions, see [KR99, Lemma 11 and (A.2)] and compare [TY20, Corollary 1.6]. □

1.4 Further directions

The main novelty in this paper is the application of the determinant method to the counting problem arising in the correlations mod 1 context; together with the stratified counting which allows us to control averages far more effectively. We have made some effort to tighten the screws in this argument but there are still some sources of loss which we anticipate could be improved. This would likely improve the range of possible degrees in Theorem 2.

In addition, these kinds of stratified counting theorems could be used in a variety of arithmetic contexts. Indeed, there are the obvious analogues of the various representation problems considered in Marmon [Mar11] and potential applications to lower order terms in the mean values of Weyl sums. Moreover, all our counting results should be extendable to global fields via the work of Paredes–Sasyk [PS22][2]. As far as future efforts go, it would be particularly interesting to see results of this kind for non-diagonal hypersurfaces.

Notation

Throughout this paper we use typical Bachmann-Landau notation: for functions $f, g : X \to \mathbb{R}$ defined on some set X, we write $f \ll g$ (or $f = O(g)$) to denote that there exists a constant $C > 0$ such that $|f(x)| < C|g(x)|$ for all $x \in X$. We write $f \asymp g$ if $f \ll g$ and $g \ll f$ and let $f = o(g)$ imply that $\frac{f(x)}{g(x)} \to 0$. We let $[N] := \{1, \ldots, N\}$.

Given a Schwartz function $f : \mathbb{R}^m \to \mathbb{R}$, let $\widehat{f}$ denote the m-dimensional Fourier transform:

$$\widehat{f}(\mathbf{k}) := \int_{\mathbb{R}^m} f(\mathbf{x}) e(-\mathbf{x} \cdot \mathbf{k}) \mathrm{d}\mathbf{x}, \qquad \text{for } \mathbf{k} \in \mathbb{Z}^m.$$

Here, and throughout we let $e(x) := e^{2\pi i x}$. Finally, we use $\varepsilon > 0$ as a small constant which can change from line to line.

Plan of the paper: we prove the counting theorem, Theorem 5, in section 2. Then we prove the Lebesgue almost everywhere statement, Theorem 2, in section 3. Then we prove the extension to fractal measures, Theorem 3, in section 4. Then, we finish with the proof of Theorem 6 in section 5. We have taken care to use notation that is coherent and remains close enough to the literature to avoid committing any sins. However, since the different sections are rather separate, we will not import the notation between the different sections.

2 Proof of Theorem 5

Fix $a_0, \ldots, a_n \in \mathbb{Z}_{\neq 0}$ and let $X_n \subset \mathbb{P}^n$ denote the projective hypersurface defined by the diagonal equation $a_0 x_0^d + \ldots + a_n x_n^d = 0$. A *quasi-diagonal* subvariety on X_n are those subsets of the rational points on X_n satisfying an equation of the form

$$\sum_{j \in P_1} a_j x_j^d = 0 = \sum_{j \in P_2} a_j x_j^d,$$

for some partition $P_1 \cup P_2 = \{0, 1, \ldots, n\}$ into parts of size at least 2. The lines contained within such subvarieties are called the *standard lines* and are the only lines on X_n (c.f. [Deb01, Ex. 2.5.3]).

The following result of Salberger controls the degree of curves on X_n and is the major geometric engine behind our results.

Theorem 8 ([Sal23, Thm. 9.1])**.** *Let C be a closed integral curve on X_n of degree δ which does not lie on any other diagonal hypersurface of degree d in $\mathbb{P}^n$. Then*

$$\delta \geq 1 - \frac{2}{n-1} + \frac{2d}{n(n-1)}.$$

Salberger [Sal23] and Marmon [Mar11] showed that the only curves of small degree belonging to X_n are the standard lines when $n = 3$ and 4, respectively. We start by establishing a slightly weaker variant that holds for all n.

Lemma 9. *Let $C \subset X_n$ be a curve of degree less than $L(d, n)$ (see (1.7)). Then C is contained within a quasi-diagonal subvariety.*

[2] We thank Matteo Verzobio for this observation.

Proof. We proceed by induction on n. The base case when $n = 3$ corresponds to a result of Salberger [Sal23, Thm. 9.4]. Let us then suppose that the conclusion holds for all natural numbers between 3 and $n-1$ and let $C \subset X_n$ be an integral curve of degree less than $L(d,n)$.

Applying Theorem 8 we see that C must lie on another diagonal hypersurface. Let us call this hypersurface Y and suppose that it is cut out by the vanishing of the form $G(\mathbf{x}) = \sum_{i=0}^{n} b_i x_i^d$. Moreover, denote by $F(\mathbf{x})$ the form whose vanishing defines X_n. Then there exists a linear combination of F and G such that at least one of the coefficients is 0. Hence without loss of generality, we may assume that $b_n = 0$. Now we project $\mathbb{P}^n \to \mathbb{P}^{n-1}$ onto the first n coordinates. Let Z denote the image of Y under this projection, and D the image of C. Note that Z is again a diagonal hypersurface, now inside $\mathbb{P}^{n-1}$. The image D must be either a point or a curve, however if it were a point then C would be the line passing through the point $[0 : \ldots : 0 : 1]$ which cannot be since $a_n \neq 0$. Therefore D is a curve and since it has degree smaller than $L(d,n) \leq L(d,n-1)$, we conclude by the inductive hypothesis that there exists a subsum $\sum_{i\in I} b_i x_i^d$ vanishing on C.

We would like to draw a similar conclusion but for a subsum involving the a_i not the b_i. We project down now from $\mathbb{P}^n$ to $\mathbb{P}^{\#I-1}$ onto the coordinates lying in I. Let W denote the subvariety given by the subsum found above $\sum_{i\in I} b_i x_i^d = 0$. Either the image of C under this projection is a point or a curve. If it is a curve then as before we may find another diagonal form in $\mathbb{P}^{\#I-1}$ which also vanishes on the image of C. By the exact argument just used, there is a subset $I' \subset I$ and a form $\sum_{i\in I'} c_i x_i^d$ vanishing on the image of C. We can therefore project again onto $\mathbb{P}^{\#I'-1}$.

We repeat this process, reducing the dimension of the ambient projective space by projection, until one of two things happens. Either our curve gets projected onto a point or we eventually project onto a projective space of dimension ≤ 4. In the latter case, we know by the work of Salberger and Marmon, that the image of C must be a line. In particular, this image must be a line on the projection of X_n, which remains a Fermat hypersurface. Since all lines are standard lines, we conclude that there must be a vanishing subsum.

Otherwise, this procedure has produced an integer r and a projection $\mathbb{P}^n \xrightarrow{\pi} \mathbb{P}^r$ such that the image of C under this projection is a point. Let us suppose, without loss of generality, that point is $[y_0 : \ldots : y_{r-1} : 1]$. The fibre under π above this point is a linear space L inside $\mathbb{P}^n$. Points on the intersection $L \cap X_n$ must satisfy an equation of the shape

$$cx_r^d + a_{r+1}x_{r+1}^d + \ldots + a_n x_n^d = 0, \tag{2.1}$$

where

$$c = a_0 y_0^d + \ldots a_{r-1} y_{r-1}^d + a_r$$

If $c = 0$ then we are done. Otherwise, we may apply the inductive hypothesis to (2.1). This will produce a subset I of indices whose subsum vanishes. If $r \notin I$ then we have found a vanishing subsum of the original equation. Otherwise the sum of the indices not in I must also sum to 0 and thus we are have the claim. □

We now extend this conclusion to subvarieties of greater dimension.

Lemma 10. *Let $Y \subset X_n$ be a subvariety of positive codimension and of degree less than $L(d,n)$. Then there exists a subset $I \subset \{0, \ldots, n\}$ such that $\#I \geq 2$ and such that the form $\sum_{i\in I} a_i x_i^d$ vanishes on Y.*

Proof. This proof is more or less exactly the same as [SW10, Lemma 4.2], we just have stronger input for the base case of our induction hence arrive at a stronger conclusion for each dimension. □

With this result in hand we may now apply the determinant method.

Lemma 11. *There is a collection of*

$$O_{d,n,\epsilon}\left(B^{\sum_{r=3}^{n-1} \frac{r+1}{\sqrt[r]{L(d,n)}}}\right)$$

surfaces of degree $O_{d,n,\epsilon}(1)$ such that every point of X_n of height at most B lies in one of these subvarieties.

Our result is based on [SW10, Lemma 3.4], however in that result the authors only cover the points of bounded height with subvarieties of codimension $(n-2)/2$. We are able to go deeper because Lemma 10 affords us control over the degree of *all* subvarieties which avoid vanishing subsums.

Proof. This proof is exactly like that of [SW10, Lemma 3.4], inductively applying [SW10, Lemma 3.3]. □

Now that the points have been restricted to surfaces, we may apply the following result of Salberger. Note that the below is not how the result is stated in [Sal23] but the same proof provides this result.

Lemma 12 ([Sal23, Thm. 6.1])**.** *Let $\epsilon > 0$ and let $S \subset \mathbb{P}^n_{\mathbb{Q}}$ be a projective irreducible surface of degree d. Then the number of rational points on S of height at most B which do not lie on any irreducible curves of degree at most $e-1$ is bounded by*

$$\ll_{n,d,e,\epsilon} B^{\frac{3}{\sqrt{d}}+\epsilon} + B^{\frac{3}{2\sqrt{d}}+\frac{2}{e}+\epsilon}.$$

As a consequence, we are able to count points which do not satisfy $\sum_{i\in I} a_i x_i^d = 0$ for any subset $I \subset \{0,\dots,n\}$ for which $2 \le \#I \le n-1$. We say that for such points, there is *no vanishing subsum.*

Corollary 13. *The number of points of height at most B on X_n for which there does not exist a vanishing subsum is*

$$O_\epsilon\left(B^{\sum_{r=3}^{n-1}\frac{r+1}{\sqrt[r]{L(d,n)}}}\left(B^{\frac{3}{\sqrt{L(d,n)}}+\epsilon} + B^{\frac{3}{2\sqrt{L(d,n)}}+\frac{2}{L(d,n)}+\epsilon}\right)\right).$$

This affirms the first claim of Theorem 5.

Proof. By the previous result, the rational points lie on at most

$$O_{d,n,\epsilon}\left(B^{\sum_{r=3}^{n-1}\frac{r+1}{\sqrt[r]{L(d,n)}}}\right)$$

surfaces. By Lemma 10, we know that the points with no vanishing subsums must lie on surfaces of degree at least $L(d,n)$ and by the lemma before that we have that any curve on such a surface also has degree at least $L(d,n)$. Therefore, we conclude by applying Lemma 12. □

Finally, we recall Heath-Brown's result counting points on curves.

Lemma 14 ([HB02, Thm. 5])**.** *Let C be an irreducible non-singular curve in $\mathbb{P}^3$ of degree d. Then the number of points of height at most B on C is $O_\epsilon(B^{\frac{2}{d}+\epsilon})$.*

With all this in hand, we can conclude our upper bound for the point count.

Theorem 15. *Let $n > d > 3$ and consider $X_n \subset \mathbb{P}^n$ the diagonal hypersurface defined by the equation $\sum_{i=0}^n a_i x_i^d = 0$. Let m be the largest number of disjoint pairs of distinct indices $\{i,j\}$ such that $-a_i a_j$ is a square. Then the number of rational points on X_n of height bounded by B is at most*

$$\ll_\epsilon \sum_{s=0}^{m} B^{s+\sum_{r=1}^{n-2s-1}\frac{r+1}{\sqrt[r]{L(d,n)}}+\epsilon}.$$

Proof. We proceed via strong induction on n. To set the base case we show that the desired bound holds for $n = 3$ and 4.

Suppose that $n = 3$. Then the number of points avoiding a vanishing subsum is counted by Lemma 12 and is at most

$$B^{\frac{3}{\sqrt{d}}+\epsilon} + B^{\frac{3}{2\sqrt{d}}+\frac{2}{L(d,n)}+\epsilon}.$$

Should there exist a subsum it must be a pair of binary subsums. In this case we have $m = 2$ and we apply the bound

$$\#\{|x|,|y| \le B : ax^2 + by^2 = 0\} \ll B. \tag{2.2}$$

Hence the number of points is bounded by

$$\begin{cases} B^{\frac{3}{\sqrt{d}}+\epsilon} + B^{\frac{3}{2\sqrt{d}}+\frac{2}{L(d,n)}+\epsilon} \ll B^{\frac{3}{\sqrt{L(d,n)}}+\frac{2}{L(d,n)}+\epsilon} \text{ if } m=0 \\ B^2 \text{ if } m = 2, \end{cases}$$

completing the proof in this case.

Suppose that $n = 4$. The total number of points avoiding vanishing subsums is provided by Corollary 13 which yields

$$B^{\frac{4}{\sqrt[3]{L(d,n)}}+\epsilon}\left(B^{\frac{3}{\sqrt{L(d,n)}}} + B^{\frac{3}{2\sqrt{L(d,n)}}+\frac{2}{L(d,n)}}\right).$$

If a quaternary subsum vanishes then we have one choice for the remaining variable and the bound

$$O\left(B^2 + B^{\frac{3}{\sqrt{L(d,n)}}+\frac{2}{L(d,n)}}\right)$$

for the four relevant variables, by the previous paragraph. If a point is not enumerated by either of these counts then there must be a set of 3 indices $\{i_1, i_2, i_3\} \subset \{0, \dots, 4\}$ such that

$$a_{i_1} x_{i_1}^d + a_{i_2} x_{i_2}^d + a_{i_3} x_{i_3}^d = 0.$$

The number of potential solutions by Lemma 14 is $O(B^{\frac{2}{L(d,n)}})$. Then there are $O(B)$ choices for the remaining two variables, by (2.2). In total then we produce a bound of

$$O\left(B^2 + B^{1+\frac{2}{L(d,n)}+\epsilon} + B^{\frac{4}{\sqrt[3]{L(d,n)}}+\frac{3}{\sqrt{L(d,n)}}+\frac{2}{L(d,n)}+\epsilon}\right).$$

Now suppose the claim holds in all dimensions less than n. Points for which no proper subsum vanishes are counted by Corollary 13 which provides the bound

$$O\left(B^{\sum_{r=1}^{n-1} \frac{r+1}{(L(d,n))^{\frac{1}{r}}}+\epsilon}\right).$$

For any remaining point $\mathbf{x}$, there exists a minimal partition $I_1, \dots, I_j$ of $\{1, \dots, n\}$ such that each subsum $\sum_{i\in I_k} a_i x_i^d$ vanishes and such that there is no partition satisfying the same property with a greater number of parts in the partition. If I_k is a singleton then there is only 1 solution. For any I_k of size 2, we apply the estimate (2.2). For any I_k of size 3, we apply Lemma 14. Any other part in the partition has size between 4 and n, thus we may apply the inductive hypothesis.

In particular, the number of points which have exactly s binary vanishing subsums is bounded by B^s times the number of ways that the remaining $n+1-2s$ variables can vanish without there being smaller vanishing subsum. This provides precisely the s^{th} term in the sum. □

Remark. The conclusion of Theorem 15 is probably suboptimal as currently written. This is because at each step when one removes variables corresponding to subsums, one should be able to conclude that the degree bound on subvarieties of the subsums is $L(d, \dim(\text{subsum}))$ rather than the full $L(d,n)$.

3 Proof of Theorem 2

Theorem 2 follows a standard procedure, first applied in this context by Rudnick and Sarnak [RS98] but later improved for pair correlations by Aistleitner, Larcher, Lewko [ALL17] and generalized to higher correlations by Technau and Yesha [TY20]. Namely, [TY20, Proposition 7.1] implies that Theorem 2 follows from the following moment bounds via an approximation argument. Let $\mathcal{F}_\ell \subset C_c^\infty(\mathbb{R}^{\ell-1})$ denote the set of rectangular smooth functions:

$$\mathcal{F}_\ell := \left\{ f(\mathbf{x}) = \prod_{i=1}^{\ell-1} f_i(x_i), \quad \text{where } f_1, \dots, f_{\ell-1} \in C_c^\infty(\mathbb{R}^{\ell-1}) \right\}.$$

Lemma 16. *For $\mathcal{X}_\alpha^d$ as defined in (1.1), ℓ satisfying $d > d_\ell$ and $f \in \mathcal{F}_\ell$ there exists an $\varepsilon > 0$ such that*

$$\mathbf{E}\left(R_\ell^N(\mathcal{X}_\alpha^d)\right) = \mathbf{E}(f) + O(N^{-\varepsilon}) \tag{3.1}$$

and

$$\mathbf{Var}\left(R_\ell^N(\mathcal{X}_\alpha^d)\right) = O(N^{-\varepsilon}). \tag{3.2}$$

Proof. First, apply Poisson summation and a change of variables to the sum in $\mathbf{k}$ in the definition of the ℓ-correlation, that is

$$\begin{aligned} R_\ell^N(\mathcal{X}_\alpha^d) &= \frac{1}{N} \sideset{}{^*}\sum_{\mathbf{n}\in[N]^\ell} \sum_{\mathbf{k}\in\mathbb{Z}^{\ell-1}} f(N(x(n_1) - x(n_2) + k_1, \dots, x(n_{\ell-1}) - x(n_\ell) + k_{\ell-1})), \\ &= \frac{1}{N^\ell} \sideset{}{^*}\sum_{\mathbf{n}\in[N]^\ell} \sum_{\mathbf{a}\in\mathcal{A}_\ell(N)} \widehat{f}\left(\frac{\mathbf{a}}{N}\right) e(\alpha(\mathbf{a}\cdot\mathbf{n}^d)) + O(N^{-\varepsilon}), \end{aligned}$$

where

$$\mathcal{A}_\ell(N) := \{(\mathbf{a}, a_\ell) \ : \mathbf{a} \in (\mathbb{Z}^{\ell-1} \setminus \{\mathbf{0}\}) \cap [-N^{1+\varepsilon}, N^{1+\varepsilon}]^{\ell-1},\ a_\ell = -a_1 - \dots - a_{\ell-1}\}. \tag{3.3}$$

Note that the fast decay of Fourier coefficients allows us to keep $|a_i| < N^{1+\varepsilon}$. Hence,

$$R_\ell^N(\mathcal{X}_\alpha^d) = \mathbf{E}(f) + \frac{1}{N^\ell} \sum_{\mathbf{n}\in[N]^\ell}^{*} \sum_{\mathbf{a}\in\mathcal{A}_\ell(N)} \widehat{f}\left(\frac{\mathbf{a}}{N}\right) e(\alpha Q_{\mathbf{a}}(\mathbf{n})) + O(N^{-\varepsilon}), \tag{3.4}$$

where $\mathbf{n}^d := (n_1^d, \ldots, n_\ell^d)$ and $Q_{\mathbf{a}}(\mathbf{n}) := \mathbf{a} \cdot \mathbf{n}^d$.

Henceforth, denote

$$\mathcal{E} := \frac{1}{N^\ell} \sum_{\mathbf{n}\in[N]^\ell}^{*} \sum_{\mathbf{a}\in\mathcal{A}_\ell(N)} \widehat{f}\left(\frac{\mathbf{a}}{N}\right) e(\alpha Q_{\mathbf{a}}(\mathbf{n})).$$

Note that the integral is 0 unless $Q_{\mathbf{a}}(\mathbf{n})$ is 0. Hence, on fixing a small $\varepsilon > 0$, we have

$$\mathcal{E} \ll \frac{1}{N^\ell} \sum_{\mathbf{a}\in\mathcal{A}_\ell(N)} \#\{\mathbf{n} \in [N^{1+\varepsilon}]^\ell \text{ distinct}, \ \ Q_{\mathbf{a}}(\mathbf{n}) = 0\},$$

By an inductive argument on the different correlation orders we can take $a_i > 0$ for all i. In fact, this is where the condition that the x_i be distinct appears, if not, the lower order correlations with some of the a_i being 0 cannot be bounded by $O(N^{-\varepsilon})$ (see [LT21, Section 3] for a detailed explanation of this argument). Given $\mathbf{a} \in \mathcal{A}_\ell(N)$, let $0 \le m(\mathbf{a}) \le \ell/2$ denote the largest number of disjoint distinct pairs (i, j) such that $-a_i a_j = \square$. Further, let

$$\mathcal{A}_\ell(N, m) := \{\mathbf{a} \in \mathcal{A}_\ell(N) : \ m(\mathbf{a}) = m, \ |a_i| > 0 \ \forall_{1\le i<\ell}\}.$$

Then, by Theorem 5,

$$\mathcal{E}_m := \sum_{\mathbf{a}\in\mathcal{A}_\ell(N,m)} \#\{\mathbf{n} \in [N^{1+\varepsilon}]^\ell \text{ distinct} : Q_{\mathbf{a}}(\mathbf{n}) = 0\} \ll N^{\phi(m,d,\ell)+\varepsilon} \#\mathcal{A}_\ell(N, m).$$

By a classical bound (see, for instance, [AKOS25, Thm. 2.6]) the number of ways to choose a_1, a_2 so that $-a_1 a_2 \in \square$ is $O(N^{1+\epsilon} \log N)$ hence we save a factor of N. Moreover, recall that a_ℓ is determined once the first $\ell - 1$ components of $\mathbf{a}$ are fixed. Suppose $a_{\ell-1} a_\ell \in \square$, then, if a_ℓ is of size $O(N)$, there are $\sqrt{N}$ many choices for $a_{\ell-1}$. Hence,

$$|\mathcal{A}_\ell(N, m)| \ll \begin{cases} N^{\ell-1+\varepsilon} & \text{if } m = 0 \\ N^{\ell-m-1+\varepsilon} & \text{if } 1 \le m < \ell/2 \\ N^{\ell/2-1/2+\varepsilon} & \text{if } m = \ell/2. \end{cases}$$

Thus

$$\begin{aligned} \mathcal{E} &\ll \frac{1}{N^\ell} \sum_{m=0}^{\ell/2} \mathcal{E}_m \\ &\ll \frac{N^\varepsilon}{N^\ell} \sum_{m=0}^{\ell/2} N^{\phi(m,d,\ell)} |\mathcal{A}_\ell(N, m)| \\ &\ll \frac{N^\varepsilon}{N^\ell} \left(N^{\phi(0,d,\ell)+\ell-1} + N^{\ell-1/2} + \sum_{m=1}^{\ell/2-1} N^{\phi(m,d,\ell)} N^{\ell-m-1} \right). \end{aligned}$$

Note that, for m fixed

$$\begin{aligned} \phi(m, d, \ell) + \ell - m - 1 &\le \max_{0\le s\le m} (s - m) + \ell - 1 + \sum_{r=1}^{\ell-1} \frac{r+1}{\sqrt[r]{L(d, n)}} \\ &\le \ell - 1 + \sum_{r=1}^{\ell-1} \frac{r+1}{\sqrt[r]{L(d, n)}} = \phi(0, d, \ell) + \ell - 1. \end{aligned}$$

Thus, the boundary terms are the dominant ones. Hence,

$$\mathcal{E} \ll N^{\phi(0,d,\ell)-1} + N^{-1/2+\varepsilon}.$$

We conclude that the expected value converges whenever

$$\phi(0, d, \ell) < 1. \tag{3.5}$$

Turning now to the variance, the proof is very similar. Using (3.1), the second moment equals

$$\mathbf{E}\left(R_\ell^N(\mathcal{X}_\alpha^d)^2\right) = \frac{1}{N^{2\ell}}\int_0^1 \left|\sum_{\mathbf{n}\in[N]^\ell}^{*}\sum_{\mathbf{a}\in\mathbb{Z}^{\ell-1}} \widehat{f}\left(\frac{\mathbf{a}}{N}\right)e(\alpha(\mathbf{a}\cdot\mathbf{n}^d))\right|^2 \mathrm{d}\alpha$$
$$= \mathbf{E}(f)^2 + \frac{1}{N^{2\ell}}\int_0^1 \left|\sum_{\mathbf{n}\in[N]^\ell}^{*}\sum_{\mathbf{a}\in\mathcal{A}_\ell} \widehat{f}\left(\frac{\mathbf{a}}{N}\right)e(\alpha(\mathbf{a}\cdot\mathbf{n}^d))\right|^2 \mathrm{d}\alpha + O(N^{-\varepsilon}).$$

Expanding the square and computing the integral in α leaves us to bound

$$\mathcal{E} := \frac{1}{N^{2\ell}}\sum_{\mathbf{a}\in\mathcal{A}_\ell}\sum_{\mathbf{b}\in\mathcal{A}_\ell}\#\{\mathbf{m},\mathbf{n}\in[N]^{\ell,*} \ : \ Q(\mathbf{a},\mathbf{b},\mathbf{m},\mathbf{n})=0\}, \tag{3.6}$$

where $[N]^{\ell,*}$ is the set of $[N]^\ell$ with distinct entries and $Q(\mathbf{a},\mathbf{b},\mathbf{m},\mathbf{n}) = Q(\mathbf{a},\mathbf{m}) - Q(\mathbf{b},\mathbf{n})$.

If we apply the same estimates leading up to (3.5), we arrive at

$$\mathcal{E} \ll N^{-1+\varepsilon} + N^{\phi(0,d,2\ell)-2+\varepsilon}.$$

Note that, since $d > d_\ell$ we have $\phi(0,d,2\ell) < 2$.

□

4 Proof of Theorem 3

The following lemmas allow us to control the growth of trigonometric polynomials coming from expanding the measure.

Lemma 17. *Let $T(\xi) = \sum_{|u|\le U} c_u e(-\xi u)$ be a trigonometric polynomial of degree U where all $c_u \in \mathbb{C}$. Let μ be as in Theorem 3. Then for all $\rho \in (0,1)$,*

$$\int_{-\infty}^{\infty} T(\xi)\, d\mu(\xi) \ll_{\rho,\mu} U^\rho \|T\|_{L^2}. \tag{4.1}$$

Proof. Denote by X the left-hand side of (4.1). Notice, that for $0 < s < 1$ we have

$$X = \sum_{|u|\le U} c_u \widehat{\mu}(u) = \sum_{|u|\le U} c_u u^{\frac{1-s}{2}}\widehat{\mu}(u) u^{\frac{s-1}{2}}.$$

By the Cauchy-Schwarz inequality,

$$|X|^2 \le \left(\sum_{|u|\le U} |c_u|^2 u^{1-s}\right)\cdot\left(\sum_{|u|\le U} |\widehat{\mu}(u)|^2 |u|^{s-1}\right).$$

The summation in the first bracket is at most $U^{1-s}\|T\|_{L^2}^2$. By (1.3) the summation in the second bracket is at most a constant $C_{s,\mu} > 0$. Choosing $s = 1-2\rho$ and taking roots, completes the proof. □

We prove now a slightly more general statement than needed, as we believe the transference mechanism is of independent interest.

Lemma 18. *Let $C > 1$ be fixed. For each integer $N \ge 1$, let $T_N : [0,1] \to \mathbb{C}$ be a trigonometric polynomial of degree $U_N = O(N^C)$. Assume $\mathbf{E}(T_N) = o(1)$. Suppose that there exists an $\varepsilon > 0$ so that*

$$\int_0^1 |T_N(\alpha) - \mathbf{E}(T_N)|^4 d\alpha \ll_\varepsilon N^{-2\varepsilon}. \tag{4.2}$$

Let μ be as in Theorem 3. Then, there exists an increasing sequence of integers $(N_m)_m$ with $N_m \sim m^{3/\varepsilon}$ such that

$$\lim_{m\to\infty} T_{N_m}(\alpha) = 0 \tag{4.3}$$

for μ-almost all $\alpha \in [0,1]$.

Proof. First, we show $\mathbf{E}(T_N,\mu) = o(1)$. Let $\widetilde{T}_N = T_N - \mathbf{E}(T_N)$, as $\mathbf{E}(T_N,\mu) - \mathbf{E}(T_N) = \int_{-\infty}^{\infty}\widetilde{T}_N(\xi)\, d\mu(\xi)$. Here

$$\widetilde{T}_N(\xi) = \sum_{|u|\le U_N} c_{u,N} e(-\xi u).$$

The Cauchy-Schwarz inequality and (4.2) imply $\|\widetilde{T}_N\|_{L^2} \le \|\widetilde{T}_N\|_{L^4}^{1/2} \ll_\varepsilon N^{-\varepsilon}$. Thus, (4.1) yields for $\rho \in (0,1)$ that

$$\int_{-\infty}^{\infty} \widetilde{T}_N(\xi)\, d\mu(\xi) \ll_{\rho,\mu} N^{C\rho} \|\widetilde{T}_N\|_{L^2} \ll N^{C\rho} N^{-\varepsilon}.$$

Upon choosing $\rho \in (0, \varepsilon/C)$, we deduce $\mathbf{E}(T_N, \mu) = o(1)$. Next, we upper-bound the μ-measure of $P_N = \{\alpha \in [0,1] : |\widetilde{T}_N(\alpha)| > N^{-\frac{\varepsilon}{10}}\}$. Then,

$$|\widetilde{T}_N(\xi)|^2 = \sum_{|u| \le 2U_N} k_{u,N} e(-\xi u) \quad \text{where} \quad k_{u,N} = \sum_{v<u} c_{v,N} \cdot c_{u-v,N}.$$

Applying (4.1) with the trigonometric polynomial $|\widetilde{T}_N(\xi)|^2$ and using (4.2) produces the estimate

$$\int_{-\infty}^{\infty} |\widetilde{T}_N(\xi)|^2\, d\mu(\xi) \ll_{\rho,\mu} (2U)^\rho \|\widetilde{T}_N\|_{L^4}^2 \ll_\mu N^{C\rho} N^{-\varepsilon}.$$

Upon choosing ρ small, we see that the right hand side is $O(N^{-\varepsilon/2}|\widetilde{T}_N(\xi)|^2)$. Chebychev's inequality yields $\mu(P_N) \ll N^{-\varepsilon/2}$. The convergence Borel-Cantelli lemma shows that the limsup set P of the P_{N_m} is a μ-null set because

$$\sum_{m\ge1} \mu(P_{N_m}) \ll \sum_{m \ge 1} N_m^{-\varepsilon/2} \ll \sum_{m\ge 1} m^{-3/2}$$

converges. Taking $\mathcal{G} = S \setminus P$ and noting that (4.3) holds for any $\alpha \in \mathcal{G}$ completes the proof. □

Now we are in a position to prove Theorem 3.

Proof of Theorem 3. By the Fourier series expansion (3.4), we have

$$R_\ell^N(\mathcal{X}_\alpha^d) = \mathbf{E}(f) + c_{0,N} + O(N^{-\varepsilon}) + \sum_{u \ne 0} c_{u,N} e(-u\alpha) \text{ where } c_{u,N} := \frac{1}{N^\ell} \sum^{*}_{\mathbf{n} \in [N]^\ell} \sum_{\mathbf{a} \in \mathbb{Z}^{\ell-1} \setminus \{\mathbf{0}\}} \widehat{f}\left(\frac{\mathbf{a}}{N}\right) 1(Q_{\mathbf{a}}(\mathbf{n}) = u).$$

By (3.1), we know $c_{0,N} = o(1)$. Let $U_N = N^{1+d+\varepsilon}$ and

$$T_N = \sum_{1 \le |u| \le U_N} c_{u,N} e(-u\alpha).$$

Then the rapid decay of $\widehat{f}$ implies

$$R_\ell^N(\mathcal{X}_\alpha^d) = \mathbf{E}(f) + c_{0,N} + O(N^{-\varepsilon}) + T_N + O(N^{-\varepsilon}).$$

To use Lemma 18, we bound $\int_0^1 |T_N(\alpha) - \mathbf{E}(T_N)|^4 d\alpha$.

Indeed,

$$\int_0^1 |T_N(\alpha) - \mathbf{E}(T_N)|^4 d\alpha \ll \frac{1}{N^{4\ell}} \sum_{j \le 4} \sum^{*}_{\mathbf{n}_j \in [N]^\ell} \sum_{\substack{\mathbf{a}_j \in \mathbb{Z}^\ell \setminus \{\mathbf{0}\} \\ \|\mathbf{a}_j\|_\infty \le N^{1+\varepsilon}}} 1\big(Q_{\mathbf{a}_1}(\mathbf{n}_1) - Q_{\mathbf{a}_2}(\mathbf{n}_2) = Q_{\mathbf{a}_3}(\mathbf{n}_3) - Q_{\mathbf{a}_4}(\mathbf{n}_4)\big).$$

The counting condition is now relaxed to count non-zero $\mathbf{b} \in \mathbb{Z}^{4\ell}$ and $\mathbf{m} \in [N]^{4\ell}$ so that $Q_{\mathbf{b}}(\mathbf{m}) = 0$ while $\|\mathbf{b}\|_\infty, \|\mathbf{m}\|_\infty \le N^{1+\varepsilon}$. Because $d > d_{4\ell}$, we obtain this information from (3.1). The upshot by a usual Borel-Cantelli and Chebychev argument is that for μ-almost any $\alpha \in [0,1]$ we have $R_\ell^{N_m}(\mathcal{X}_\alpha^d) \to \mathbf{E}(f)$ as $m \to \infty$, for any fixed $(N_m)_m$ with $N_m \sim m^{4/\varepsilon}$. Upgrading the convergence along a sub-lacunary subsequence to convergence as $N \to \infty$ is a routine sandwiching argument, see [TY20, Lemma 7.2]. □

5 Proof of Theorem 6

Now we turn to the matrix counting problem described in section 1.3.1. Let

$$\mathcal{M}(r,N) = \mathcal{M}_{d,k,n}(r,N) := \{M \in \mathcal{M}_{d,k,n}(N) \ : \ \mathrm{rank}(M) = r\}.$$

To prove Theorem 6 we follow the example of [BL23, Section 4]. Let

$$T := \left\{ \begin{pmatrix} x_{1,1}^d & \dots & x_{1,k}^d \\ & \vdots & \\ x_{n,1}^d & \dots & x_{n,k}^d \end{pmatrix} \ : \ x_{i,j} < N \right\}.$$

First, we assume $\operatorname{rank}(T) < n$. Because of the small rank, we can assume (w.l.o.g.) that the bottom $n-r$ rows are linear combinations of the first r rows. That is,

$$x_{j,i}^d = \sum_{\nu=1}^{r} \rho_{j,\nu} x_{\nu,i}^d, \qquad r+1 \le j \le n, \qquad 1 \le i \le k, \tag{5.1}$$

for some $\rho_{j,\nu} \in \mathbb{Q}$ (at least two of which are non-zero to ensure that $\mathbf{x}_j \neq \mathbf{x}_\nu$ for $\nu \le r$).

To populate T, first choose the first r columns indiscriminately which contributes $O(N^{rn})$. This choice will uniquely determine $\boldsymbol{\rho}$. Indeed, if we write

$$T = \begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix} \in \mathbb{R}^{(r+(n-r))\times(r+(k-r))},$$

then since T_1 is invertible, we have for each $r+1 \le j \le n$ that

$${}^tT_1 \begin{pmatrix} \rho_{j,1} \\ \vdots \\ \rho_{j,r} \end{pmatrix} = \begin{pmatrix} x_{j,1}^d \\ \vdots \\ x_{j,r}^d \end{pmatrix}.$$

Now, consider (5.1), which is a non-degenerate form of degree d in at least 3 variables. A priori, the coefficients of this form are rational numbers but we may simply clear denominators. Applying the worst case scenario of Theorem 5 then yields

$$O(N^{\frac{r}{2}+1+\epsilon})$$

choices for the $(r+1)$-tuple $x_{j,i}, x_{1,i}, \ldots, x_{r,i}$. Note that this bound is uniform in the coefficients of the form and so in particular is insensitive to our having cleared denominators and to the fact that we have no control over the $\rho_{j,\nu}$. Hence, we conclude

$$\mathcal{M}(r,N) = O(N^{nr+(r/2+1)(k-r)+\varepsilon}),$$

the claimed bound in Theorem 6.

Remark. The first term is the same as in [BL23, display below (4.7)], however the $(k-r)$ dependence is now $r/2$ (as opposed to $r-1$). One could hope to use the finer information available in Theorem 5 to further improve this bound however with the present approach we can say very little about the coefficients $\rho_{j,\nu}$.

Acknowledgments

This work was supported by a grant from the Simons Foundation [SFI-MPS-TSM-00013410,CL]. NR is funded by FWF project ESP 441-NBL.

References

[ABR24] C. Aistleitner, V. Blomer, and M. Radziwiłł. Triple correlation and long gaps in the spectrum of flat tori. *J. Eur. Math. Soc. (JEMS)*, 26(1):41–74, 2024.

[AEBM21] C. Aistleitner, D. El-Baz, and M. Munsch. A pair correlation problem, and counting lattice points with the zeta function. *Geom. Funct. Anal.*, 31(3):483–512, 2021.

[AKOS25] M. Afifurrahman, V. Kuperberg, A. Ostafe, and I. Shparlinski. Statistics of ranks, determinants and characteristic polynomials of rational matrices. *Forum Math.*, 37(4):1259–1289, 2025.

[ALL17] C. Aistleitner, G. Larcher, and M. Lewko. Additive energy and the Hausdorff dimension of the exceptional set in metric pair correlation problems. *Israel J. Math.*, 222(1):463–485, 2017. With an appendix by Jean Bourgain.

[BHB06] T. D. Browning and D. R. Heath-Brown. The density of rational points on non-singular hypersurfaces. II. *Proc. London Math. Soc. (3)*, 93(2):273–303, 2006. With an appendix by J. M. Starr.

[BL23] V. Blomer and J. Li. Correlations of values of random diagonal forms. *Int. Math. Res. Not. IMRN*, (23):20296–20336, 2023.

[BL24] V. Blomer and C. Lutsko. Hyperbolic lattice point counting in unbounded rank. *J. Reine Angew. Math.*, 812:257–274, 2024.

[BW03] J. Brüdern and T. D. Wooley. The paucity problem for certain pairs of diagonal equations. *Q. J. Math.*, 54(1):41–48, 2003.

[BW22] Jörg Brüdern and Trevor D. Wooley. A paucity problem for certain triples of diagonal equations. *Bull. Lond. Math. Soc.*, 54(4):1396–1412, 2022.

[CM25] J. Chapman and A. Mudgal. Counting 2x2 integer matrices with a given determinant. *arXiv:2509.20259*, 2025.

[Deb01] O. Debarre. *Higher-dimensional algebraic geometry.* Universitext. Springer-Verlag, New York, 2001.

[DHP25] K. Dhanda, A. Haynes, and S. Prasala. Counting 2x2 matrices with fixed determinant and bounded coefficients. *Accepted (Proc. of the Cambridge Phil. Soc.)*, 2025.

[DRS93] W. Duke, Z. Rudnick, and P. Sarnak. Density of integer points on affine homogeneous varieties. *Duke Math. J.*, 71(1):143–179, 1993.

[EMM05] A. Eskin, G. Margulis, and S. Mozes. Quadratic forms of signature (2, 2) and eigenvalue spacings on rectangular 2-tori. *Ann. of Math.*, 161(2):679–725, 2005.

[GNY17] A. Gorodnik, A. Nevo, and G. Yehoshua. Counting lattice points in norm balls on higher rank simple Lie groups. *Math. Res. Lett.*, 24(5):1285–1306, 2017.

[HB02] D. R. Heath-Brown. The density of rational points on curves and surfaces. *Ann. of Math. (2)*, 155(2):553–595, 2002.

[HB09] D. R. Heath-Brown. Sums and differences of three kth powers. *J. Number Theory*, 129(6):1579–1594, 2009.

[HB10] D. R. Heath-Brown. Pair correlation for fractional parts of αn^2. *Math. Proc. Cambridge Philos. Soc.*, 148(3):385–407, 2010.

[HMR07] K. E. Hare, P. Mohanty, and M. Roginskaya. A general energy formula. *Mathematica Scandinavica*, pages 29–47, 2007.

[KN74] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. Pure and Applied Mathematics.

[KR99] P. Kurlberg and Z. Rudnick. The distribution of spacings between quadratic residues. *Duke Math. J.*, 100, 1999.

[LST25] C. Lutsko, A. Sourmelidis, and N. Technau. Pair correlation of the fractional parts of αn^θ. *J. Eur. Math. Soc. (JEMS)*, 27(10):4069–4082, 2025.

[LT21] C. Lutsko and N. Technau. Correlations of the fractional parts of αn^θ. *(Accepted: American Journal of Mathematics (AJM))*, 2021.

[LT25] C. Lutsko and N. Technau. Full Poissonian local statistics of slowly growing sequences. *Compos. Math.*, 161(1):148–180, 2025.

[Mar00] J. Marklof. The Berry-Tabor conjecture. In *Proceedings of the 3^{rd} European Congress of Mathematics*, volume 202, pages 421–427, Barcelona, 2000. Birkhäuser, Basel.

[Mar03] J. Marklof. Pair correlation densities of inhomogeneous quadratic forms. *Ann. of Math.*, 158(2):419–471, 2003.

[Mar11] O. Marmon. Sums and differences of four kth powers. *Monatsh. Math.*, 164(1):55–74, 2011.

[MOS24] A. Mohammadi, A. Ostafe, and I. Shparlinski. On some matrix counting problems. *Journal of the London Mathematical Society*, 110(6):e70044, 2024.

[PS22] M. Paredes and R. Sasyk. Uniform bounds for the number of rational points on varieties over global fields. *Algebra Number Theory*, 16(8):1941–2000, 2022.

[RS98] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. *Comm. Math. Phys.*, 194(1):61–70, 1998.

[RS24] M. Radziwiłłand A. Shubin. Poissonian pair correlation for $\alpha n^\theta \bmod 1$. *Int. Math. Res. Not. IMRN*, (9):7654–7679, 2024.

[RSZ01] Z. Rudnick, P. Sarnak, and A. Zaharescu. The distribution of spacings between the fractional parts of $n^2\alpha$. *Inventiones Mathematicae*, 145(1):37–57, 2001.

[RT22] Z. Rudnick and N. Technau. The metric theory of the pair correlation function for small non-integer powers. *J. Lond. Math. Soc*, 106(3):2752–2772, 2022.

[Rud08] Z. Rudnick. What is... quantum chaos? *Notices Amer. Math. Soc.*, 55(1):32–34, 2008.

[RZ02] Z. Rudnick and A. Zaharescu. The distribution of spacings between fractional parts of lacunary sequences. *Forum Math.*, 14(5):691–712, 2002.

[Sal23] P. Salberger. Counting rational points on projective varieties. *Proc. Lond. Math. Soc. (3)*, 126(4):1092–1133, 2023.

[SW97] C. M. Skinner and T. D. Wooley. On the paucity of non-diagonal solutions in certain diagonal Diophantine systems. *Quart. J. Math. Oxford Ser. (2)*, 48(190):255–277, 1997.

[SW10] P. Salberger and T. Wooley. Rational points on complete intersections of higher degree, and mean values of Weyl sums. *J. Lond. Math. Soc. (2)*, 82(2):317–342, 2010.

[TW22] N. Technau and A. Walker. The triple correlations of fractional parts of αn^2. *to appear in Candian J. Math, see arXiv:2005.01490 [math.NT]*, 74(5):1347–1384, 2022.

[TY20] N. Technau and N. Yesha. On the correlations of $\alpha n^\theta \bmod 1$. *J. Eur. Math. Soc. (JEMS)*, 25(10):4123–4154, 2020.

[Woo23] T. Wooley. Paucity problems and some relatives of Vinogradov's mean value theorem. *Math. Proc. Cambridge Philos. Soc.*, 175(2):327–343, 2023.

Department of Mathematics, University of Houston, 3551 Cullen Blvd, 77204, Houston, Texas, United States.
E-mail: ***clutsko@uh.edu***

Institute of Analysis and Number Theory, TU Graz, Steyrergasse 30/II, 8010 Graz, Austria.
E-mail: ***rome@tugraz.at***

Department of Mathematics, 480 Lincoln Drive, Madison, WI 53706, USA.
E-mail: ***technau@wisc.edu***