POISSON SPACING STATISTICS FOR LATTICE POINTS ON CIRCLES

PÄR KURLBERG AND STEPHEN LESTER

Abstract. We show that along a density one subsequence of admissible radii, the nearest neighbor spacing between lattice points on circles is Poissonian.

1. Introduction

For a sequence of real numbers (a_n) and integer $N \geq 1$ the probability measure given by

$$\nu_N = \frac{1}{N} \sum_{1 \leq n \leq N} \delta(\{a_n\}),$$

where $\delta(\cdot)$ denotes the Dirac delta function and $\{a_n\}$ is the fractional part of a_n, describes the distribution of the points $(a_n)_{n=1}^N$ within $\mathbb{R}/\mathbb{Z}$ and we say that $(\{a_n\})$ is uniformly distributed modulo 1 if ν_N weakly converges to the Lebesgue measure λ on $\mathbb{R}/\mathbb{Z}$ as $N \to \infty$. A classical result of Weyl provides the following criterion: $(\{a_n\})$ is uniformly distributed if and only if the Fourier coefficients of ν_N converge pointwise to those of λ as $N \to \infty$.

A more refined study of the behavior of a sequence modulo 1 examines the local spacing statistics of the sequence, that is how the sequence is distributed at the scale $1/N$, which provides insight into how the elements of the sequence are spaced together. Computing the local spacing statistics of a given sequence is in general a challenging problem. Two important examples include the nearest neighbor spacing statistics of the sequence consisting of the ordinates of zeros of the Riemann zeta-function as well as the sequence of energy levels of quantized Hamiltonians; both of these are subjects of well-known open conjectures. On the other hand there are some notable instances where the local spacing statistics have, partially or fully, been computed, e.g. see [20, 21, 22, 38, 9, 27, 4, 3, 1], including works on angles of Euclidean lattice points visible from the origin [2, 30].

Bourgain, Rudnick, and Sarnak [7] initiated the study of the spacing of $\mathbb{Z}^3$-lattice points lying on the sphere and conjectured that their nearest neighbor spacing is Poissonian, whereas in dimensions ≥ 4 they observed that lattice points lying on spheres display a much more rigid behavior. In the two dimensional case, they [7, §1.5] state that the local statistical questions certainly make sense along a density one subsequence of admissible radii and leave open predicting the limiting behavior.

In this article we investigate the spacing of angles arising from $\mathbb{Z}^2$-lattice points lying on circles. Due to the symmetry of lattice points under rotation by $\pi/2$ it suffices to consider their angles modulo $\pi/2$ and given a circle of radius $\sqrt{n}$ we write $r(n) = \frac{1}{4}\#\{\vec{x} \in \mathbb{Z}^2 : |\vec{x}|^2 = n\}$. For $\vec{x} = (x, y) \in \mathbb{Z}^2$ we write $\theta_{\vec{x}} := \arg(x + iy) \pmod{\pi/2}$ and $\mathcal{S}$ for the set of natural numbers, which can be written as a sum of two integer squares. Given $n \in \mathcal{S}$ with $r(n) = N$,

Date: Friday 31st January, 2025.

1991 *Mathematics Subject Classification.* 11K36 (primary); 11E25 (secondary).

we define the probability measure ϱ_n on $[0, \pi/2)$ by

$$\varrho_n = \frac{1}{4N} \sum_{\substack{\vec{x} \in \mathbb{Z}^2 \\ |\vec{x}|^2 = n}} \delta(\theta_{\vec{x}}).$$

A striking result independently due to Kátai and Környei [25] and Erdős and Hall [17] shows that there exists a density one subsequence $\mathcal{S}_0 \subset \mathcal{S}$ such that the angles $\theta_{\vec{x}}$ with $|\vec{x}|^2 \in \mathcal{S}_0$ very strongly equidistribute as $|\vec{x}| \to \infty$, in the sense that for $n_j \in \mathcal{S}_0$

$$\sup_{I \subset \mathbb{R}/(\frac{\pi}{2}\mathbb{Z})} \left| \varrho_{n_j}(I) - \frac{|I|}{\pi/2} \right| \ll \frac{1}{N_j^{\log(\pi/2)/\log 2 - \varepsilon}}$$

where $\log(\pi/2)/\log 2 = 0.651496129\ldots$, for every $\varepsilon > 0$. The restriction to a density one subsequence of $\mathcal{S}$ is best possible since in the case where $|\vec{x}|^2$ is a prime congruent to $1 \pmod 4$ there are only two angles $\pmod{\pi/2}$; in fact, the set of possible limiting measures is very rich even after restricting to circles having a growing number of lattice points on them (cf. [29].) Interestingly, the result above shows at scales $N^{-\gamma}$ with $\frac{1}{2} < \gamma < \log(\pi/2)/\log 2$ that lattice points on circles do not behave like independent identically distributed (iid) random

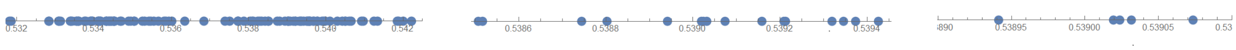


FIGURE 1. Three plots of angles of lattice points lying on the circle of radius $\sqrt{n}$ with $n = 3.7366813 \cdot 10^{35}$ and of $N = r(n) = 2^{14}$ nearby 0.5390 with windows of length $1/100, 1/1000, 1/5000$, respectively. Note that $N^{-1/2} = 0.0078125\ldots$, $N^{-\log(\pi/2)/(\log 2)} = 0.0017960\ldots$, and $N^{-1} = 0.0000610\ldots$.

variables, uniformly distributed on $[0, 1]$, U_n, $n = 1, 2, \ldots$, for which the Chung-Smirnov law of the iterated logarithm [10, 40] gives almost surely (a.s.) that

$$\limsup_{N \to \infty} \frac{\sqrt{N}}{\sqrt{2 \log \log N}} \sup_{I \subset \mathbb{T}} \left| \nu_N(I) - |I| \right| = 1.$$

While behaving differently at intermediate scales, at the local scale, $1/N$, numerical evidence

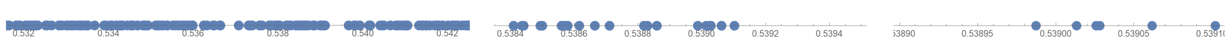


FIGURE 2. 2^{14} random points in $[0, 1]$. Three plots of the points that lie nearby 0.5390 with windows of length $1/100, 1/1000, 1/5000$, respectively.

indicates that lattice points on circles appear to typically display Poissonian statistics (cf. Figure 3 below), which would coincide with the known behavior of U_n at this scale (cf. [18, §1.7])).

We prove that the nearest neighbor spacing of angles of lattice points on circles is Poissonian along a density one subsequence of admissible radii. Listing the angles of the lattice

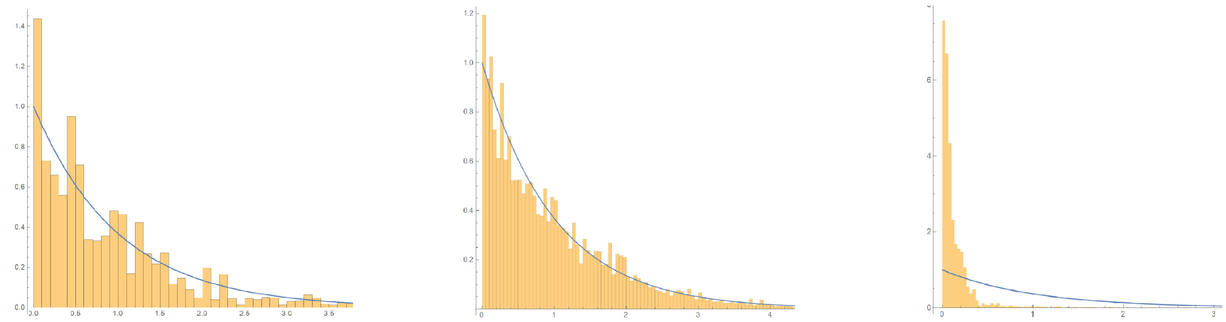


FIGURE 3. Histogram plots of the nearest neighbor spacing of 2^{14} angles of lattice points on the circle of radius $\sqrt{n}$ with $n = 3.7366813 \cdot 10^{35}$ (left), 2^{20} angles with $n = 2.1957869 \cdot 10^{54}$ (center) and 2^{16} angles with $n = 9.1943528 \cdot 10^{63}$ (right), with plot of e^{-s} (solid lines). The plot on the right is atypical; here n is a product of primes of the form $m^2 + 1$.

points on the circle of radius $\sqrt{n}$ as $0 \le \theta_1 < \theta_2 < \cdots < \theta_N < \pi/2$ for n with $r(n) = N$, we define the nearest neighbor spacing measure on $\mathbb{R}_{\ge 0}$ by

$$\mu_n = \frac{1}{N-1} \sum_{1 \le j \le N-1} \delta(\tfrac{N}{\pi/2}(\theta_{j+1} - \theta_j)). \tag{1.1}$$

Theorem 1.1. *There exists a density one subsequence $\{n_j\} \subset \mathcal{S}$ such that for any interval $I \subset \mathbb{R}_{\ge 0}$ and $N_j = r(n_j)$ as $j \to \infty$ we have*

$$\mu_{n_j}(I) = \int_I e^{-s}\, ds + o_I(1).$$

The restriction to a density one subsequence in our result is essentially best possible, since there exist subsequences $\{n_j\} \subset \mathcal{S}$, for which all the angles are highly localized near 0 or $\pi/2$. This remains true even if one requires $r(n_j) \to \infty$ as $n_j \to \infty$ (cf. the plot on the right in Figure 3.) In fact, this construction can easily be made into a rigorous argument showing that there exist subsequences of elements of $\mathcal{S}$ for which the nearest neighbor spacing measure is given by a delta mass supported at zero.

Moreover, we are also able to compute the joint limiting distribution. For n with $r(n) = N > \ell$, consider the measure on $\mathbb{R}^{\ell}_{\ge 0}$ given by

$$\mu_{\ell,n} = \frac{1}{N-\ell} \sum_{1 \le n \le N-\ell} \delta(\tfrac{N}{\pi/2}(\theta_{j+1} - \theta_j, \theta_{j+2} - \theta_{j+1}, \ldots, \theta_{j+\ell} - \theta_{j+\ell-1})).$$

Theorem 1.2. *There exists a density one subsequence $\{n_j\} \subset \mathcal{S}$ such that for any intervals $I_1, \ldots, I_\ell \subset \mathbb{R}_{\ge 0}$, where $\ell \in \mathbb{N}$ is fixed, and $N_j = r(n_j)$ as $j \to \infty$ we have*

$$\mu_{\ell,n_j}(I_1, \ldots, I_\ell) = \int_{I_1 \times \cdots \times I_\ell} e^{-(s_1 + \cdots + s_\ell)}\, ds_1 \cdots ds_\ell + o_{I_1,\ldots,I_\ell}(1).$$

We remark that the methods can be adapted to work for "thinner" subsequences inside $\mathcal{S}$, e.g. giving Poisson spacings in a density one subsequence of square free integers $n \in \mathcal{S}$ having exactly $\lfloor \frac{1}{2} \log\log n \rfloor$ prime factors.

1.1. **Discussion of the proof.** As shown in [27, 26], to show the nearest neighbor spacings have Poissonian statistics it suffices to compute the r-level correlations of the angles. As in previous works, we will work with a smoothed version of the r-correlation function and given a Schwartz function $f : \mathbb{R}^{r-1} \to \mathbb{R}$ we define

$$
\begin{aligned}
F_N(x_1, \dots, x_{r-1}) &= \sum_{j_1, \dots, j_{r-1} \in \mathbb{Z}} f\left(\frac{N}{\pi/2} \left(x_1 + j_1 \cdot \frac{\pi}{2}, \dots, x_{r-1} + j_{r-1} \cdot \frac{\pi}{2} \right) \right) \\
&= \frac{1}{N^{r-1}} \sum_{k_1, \dots, k_{r-1} \in \mathbb{Z}} \widehat{f}\left(\frac{k_1}{N}, \dots, \frac{k_{r-1}}{N} \right) e^{4i(k_1 x_1 + \cdots + k_{r-1} x_{r-1})}
\end{aligned}
\tag{1.2}
$$

where the second equality follows from applying the Poisson summation formula. The r-correlation of the angles $0 \le \theta_1 < \theta_2 < \ldots < \theta_N < \pi/2$ arising from lattice points on the circle of radius $\sqrt{n}$ is defined by

$$
R_r(n; F_N) = \frac{1}{N} \sum_{1 \le \ell_1, \dots, \ell_r \le N} F_N(\theta_{\ell_1} - \theta_{\ell_2}, \theta_{\ell_2} - \theta_{\ell_3}, \dots, \theta_{\ell_{r-1}} - \theta_{\ell_r}).
$$

We make no further assumption on f and using a standard combinatorial sieving argument we can relate the r-correlation over *distinct* angles to r'-correlations over all angles, as above, with $r' \le r$.

To evaluate the r-correlation we will apply (1.2) to see that

$$
R_r(n; F_N) = \frac{1}{N^r} \sum_{\vec{k} \in \mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N} \right) \lambda_{-4k_1}(n) \lambda_{4(k_1 - k_2)}(n) \cdots \lambda_{4(k_{r-2} - k_{r-1})}(n) \lambda_{4k_{r-1}}(n)
\tag{1.3}
$$

where $\lambda_{4k}(n) = \sum_{1 \le \ell \le N} e^{-i4k\theta_\ell}$ is the $4k$th Fourier coefficient of the measure $\sum_{1 \le \ell \le N} \delta(\theta_\ell)$. (Note that the set of lattice points on a circle is invariant under rotation by $\pi/2$, hence the kth Fourier coefficient of the measure is zero, unless $4|k$. Further, this order four symmetry implies that any r-tuple of distances occurs with multiplicity divisible by four.) Our first step is to understand the average of $R_r(n; F_N)$ over n. To explain our approach, let us first consider the pair correlation $R_2(n; F_N)$. Our strategy allows us to freeze the value of N and let n vary over all $n \le x$ with $r(n) = N$ for each $N \le (\log x)^{O(1)}$ (which is not very restrictive since $r(n) = (\log n)^{\frac{\log 2}{2} + o(1)}$ for a density one subset of $\mathcal{S}$.) At this point we use the Landau-Selberg-Delange method, which is described in detail in Sections 4.1 and 4.3, to evaluate the mean value of $\lambda_{-4k}(n)\lambda_{4k}(n) = \lambda_{4k}(n)^2$ over $n \le x$ with $r(n) = N$. Here we rely on the fact that $L(s, 4k) = \sum_{n \ge 1} \lambda_{4k}(n) n^{-s}$ is a Hecke L-function and analogously to the Riemann zeta-function the function $L(s, 4k)$, has an Euler product, analytic continuation, and zero free region. Noting that $\lambda_0(n) = r(n) = N$, arguing in this way we show in Sections 4.2-4.3 that

$$
\frac{1}{\#\{n \le x : r(n) = N\}} \sum_{\substack{n \le x \\ r(n) = N}} \lambda_{4k}(n)^2 = \begin{cases} N^2 & \text{if } k = 0 \\ N(g(k; Y) + \mathcal{E}_k) & \text{otherwise,} \end{cases}
\tag{1.4}
$$

where there error term $\mathcal{E}_k$ is typically $o(1)$, $Y = (\log N/\log 2 - 1)/\log\log x$ and $g(k; Y)$ is an Euler product such that $g(k; Y)/L(1, 8k)^Y$ is bounded. The distinction between the terms with $k = 0, k \ne 0$ arises since the Dirichlet series $\sum_{n \ge 1} \lambda_{4k}(n)^2 n^{-s}$ has a pole of order 2 at $s = 1$ if $k = 0$, whereas if $k \ne 0$ it has a pole of order 1 and the residue will be given in terms of $L(1,8k)$. It turns out that the contributions are of similar order of magnitude since the contributions from the order 2 pole occurs on a subspace of smaller codimension. At this point we use (1.4) in (1.3) to get

$$\frac{1}{\#\{n\le x: r(n)=N\}}\sum_{\substack{n\le x\\ r(n)=N}} R_2(n;F_N)=\widehat{f}(0)+\frac{1}{N}\sum_{k\in\mathbb{Z}\setminus\{0\}}\widehat{f}\left(\frac{k}{N}\right)g(k;Y)+o(1).$$

For f that approximates the indicator function of an interval I we have $\widehat{f}(0)\approx|I|$, which is consistent with a sequence that displays Poisson spacing statistics. To compute the second term above, in Section 4.2 we use analytic properties of $L(s,8k)$ to express $g(k;Y)$ in terms of a *short* Dirichlet polynomial of length $\le N^{o(1)}$ with coefficients related to $\lambda_{4k}(n)$, that is, $g(k;Y)$ is well-behaved as k varies and we can apply Poisson summation again to show the second term above equals $f(0)+o(1)$. To pass to the pair correlation function over distinct angles we subtract the contribution from the terms with $\theta_{\ell_1}=\theta_{\ell_2}$, which equals $\frac{1}{N}\cdot NF_N(0)=f(0)+o(1)$ and cancels the secondary main term above. To show that it is Poisson along a density one subsequence we will bound the variance by computing the average of $R_2(n;F_N)^2$ (following a similar strategy), similar to the argument used by Sarnak to show Poisson pair correlation for Laplace eigenvalues of generic $2d$ tori [39]. An interesting obstruction to extending this for $r>2$ in Sarnak's setting is "variance blowup" — expectations remain consistent with Poisson statistics, but variance growth makes it impossible to deduce Poisson behavior for almost all tori.

Fortunately there is no such variance blowup in our setting for larger r, but the approach described above becomes significantly more involved and further ideas are needed. As before, our first step is to apply the Landau-Selberg-Delange method to transform the problem, which gives a new expression (cf. Proposition 4.1) that depends on $\vec{k}\in\mathbb{Z}^{r-1}$ in a complex way including an arithmetic factor $g(\vec{k};Y)$ as well as an analytic factor $\alpha(\vec{k})$ which comes from the order of the pole of the Dirichlet series associated to $\lambda_{-4k_1}(n)\lambda_{4(k_1-k_2)}(n)\cdots\lambda_{4(k_{r-2}-k_{r-1})}(n)\lambda_{4k_{r-1}}(n)$, at $s=1$. The arithmetic factor $g(\vec{k};Y)$ is roughly of the shape of a product of Hecke L-functions at $s=1$, which we can approximate by a short Dirichlet polynomial (cf. Proposition 4.4), so if we were to ignore the analytic factor we would be able to evaluate the average over $\vec{k}$. While the behavior of the analytic factor is complex, its value is fixed over certain subregions of $\mathbb{Z}^{r-1}$ (e.g. for $r=2$, this corresponds to $k=0$, $k\neq 0$) and by partitioning $\mathbb{Z}^{r-1}$ into these subregions we are able to perform the average over $\vec{k}$ by means of another application of the Poisson summation formula over each nontrivial subregion, each of which is essentially a lattice (cf. Section 5.2, in particular (5.16)). This procedure essentially solves the arithmetic part of the problem.

However, the procedure described above yields a term for each subregion in the decomposition of $\mathbb{Z}^{r-1}$ that gives a contribution that is potentially the same size as our main term or potentially even larger (later on we are able to rule out any terms "blowing up" as $N\to\infty$.) We need to control the terms arising from these subregions of $\mathbb{Z}^{r-1}$ and would like to understand the behavior of the analytic factor, which counts the order of the pole at $s=1$ of the Dirichlet series mentioned above. It turns out, this is ≥ 1 for every $\vec{k}\in\mathbb{Z}^{r-1}$, so that there are poles everywhere! Moreover, the number of terms arising in the decomposition of $\mathbb{Z}^{r-1}$ into subregions where the value of the analytic factor is fixed grows exponentially with r making it difficult to analyze for large r. This complexity arises in part due to residual lower order correlations and we are left with an explicit yet intractable combinatorial expression for the r-correlation.

A similar predicament also occurs in the context of computing the r-correlation of the zeros of the Riemann zeta-function, where after solving the arithmetic part of the problem one is left with an unwieldy expression for the r-correlation. In the range corresponding to the "diagonal terms" Rudnick and Sarnak [36, 37] succeeded in solving the combinatorial problem. The "off-diagonal" terms can be analyzed at the heuristic level using Hardy-Littlewood's conjecture on correlations of primes (e.g. [32]), however for large r, this leaves a horrible combinatorial formula. By expanding the random matrix r-correlation function as a sum of cycles, Bogomolny and Keating [5, 6] were able to eventually arrive at the same expression (including the diagonal terms), thereby heuristically matching the correlations of zeros of $\zeta(s)$ with that of eigenvalues of GUE random matrices in the full range. Innovative recent works [16, 14, 8, 31] on statistics of zeros of L-functions have solved closely related combinatorial problems through relating these expressions to identical ones coming from L-functions over function fields [16] or from random matrices [14, 8, 31], for which the spacing statistics of the zeros/eigenvalues, respectively, are known (see also [13]). We pursue an approach similar in spirit by introducing a random model for the angles of lattice points — interesting on its own right — and compute the spacing statistics for this model in two ways. First, we argue directly and show that the r-level correlations of the random model are Poissonian. Second, we then compute the r-level correlation in the random case in a different way and eventually arrive at the same complicated expression as in the deterministic setting, including the exact same analytic factor. Hence, we are able to match the deterministic computation with the random one and then are able to evaluate the r-correlation by combining these two approaches.

1.2. **Acknowledgments.** P.K. was partially supported by the Swedish Research Council (grant numbers 2016-03701,2020-04036). S.L. was partially supported by an EPSRC Standard Grant EP/T028343/1. We would like to thank Jon Keating and Peter Sarnak for comments on an earlier version of this manuscript. We would also like to thank the two referees for their careful reading of the manuscript, and many helpful comments greatly improving the exposition.

2. The set-up

Let $\mathcal{N}_M$ be the set of square free integers with prime divisors all congruent to 1 (mod 4) and total number of prime divisors equal to M i.e.

$$\mathcal{N}_M = \{n \in \mathbb{N} : \mu^2(2n)b(n) = 1 \text{ and } \Omega_1(n) = M\},$$

where $b(\cdot)$ is the indicator function of the set of integers which can be represented as a sum of two squares, μ denotes the Möbius function (which should not be confused with the measure μ_n), and

$$\Omega_1(n) = \sum_{\substack{p^a || n \\ p \equiv 1 \ (\mathrm{mod}\ 4)}} a.$$

Throughout we use the notation 1_P, which we define as $1_P = 1$ if the statement P is true and $1_P = 0$ otherwise. We also write $\mathbb{1}_{\mathcal{A}}$ for the indicator function of the set $\mathcal{A}$ e.g. $\mathbb{1}_{\mathcal{S}} = b$.

Note that for each $n \in \mathcal{N}_M$ we have $r(n) = 2^M$. Also let

$$\mathcal{N}_M(x) = \{n \in \mathcal{N}_M : n \le x\}.$$

Let $\mathcal{O} = \mathbb{Z}[i]$ denote the Gaussian integers. For n with $r(n) = 2^M = N$ we express the r-correlation of the lattice points on the circle with radius $\sqrt{n}$ by

$$R_r(n; F_N) = \frac{1}{N} \sum_{\substack{(\beta_1),\dots,(\beta_r)\subset\mathcal{O} \\ |\mathcal{O}/(\beta_1)|=\dots=|\mathcal{O}/(\beta_r)|=n}} F_N(\theta_{\beta_1} - \theta_{\beta_2}, \dots, \theta_{\beta_{r-1}} - \theta_{\beta_r}),$$

with F_N as defined in (1.2), and $\theta_\beta = \arg(\beta)$ — note that the angle $\arg(\beta)$ is well defined modulo $\pi/2$ since it is independent of choice of generator of (β). We recall that the reason for working with period $\pi/2$, rather than 2π, is that the set of lattice points on a given circle is invariant under rotation by $\pi/2$.

Because we will restrict to a density one set of integers which are sums of two squares, it suffices to consider sums of two squares $n = ef^2$ with e square free and $f \le F$ for F which grows arbitrarily slowly as $x \to \infty$ (say $F = F(x) \le x^{1/3}$) since the number of such sums of two squares $\le x$ with $f \ge F$ is

$$\sum_{f\ge F} \sum_{e \le x/f^2} b(ef^2)\mu^2(e) \ll \frac{x}{F\sqrt{\log x}}. \tag{2.1}$$

(this follows easily from the observation that $b(f^2e)\mu^2(e) = b(e)\mu^2(e)$ and the bound $\sum_{n\le x} b(n) \ll x/\sqrt{\log x}$, $x \ge 2$.) Hence, it suffices to consider $n = n_0 n_1$ where $n_0 = 2^a f^2 d$, $1 \le f \le F$, $a \in \{0, 1\}$, d is a divisor of f with $\mu^2(2d)b(d) = 1$ and $n_1 \in \mathcal{N}_M$ with $(n_0, n_1) = 1$ and $1 \le M \le A \log\log x$ for a fixed, yet sufficiently large $A > 0$. We remark that $r(n) = (\log n)^{(\log 2)/2+o(1)}$ holds for a full density subset of the set integers for which $r(n) > 0$ (cf. [28, Prop. 3.4]), hence the condition $M \in [1, A\log\log x]$ is not very restrictive; in fact any $A > 1/2$ suffices to capture a full density subset. It is then sufficient to average over *odd* square free numbers n_1; this partitioning simplifies some of the analysis below but is not essential to our argument. Write $\mathcal{N}_{M,n_0} = \{n \in \mathcal{N}_M : (n, n_0) = 1\}$ and $\mathcal{N}_{M,n_0}(x) = \{n \in \mathcal{N}_{M,n_0} : n \le x\}$. We also define

$$R_{r,n_0}(n; F_N) := R_r(n_0 n; F_N).$$

3. A random model

Motivated by Hecke's result on the equidistribution of angles of Gaussian primes, we introduce a simple, purely probabilistic, model for the angular gaps between lattice points $(x, y) \in \mathbb{Z}^2$ on the circle $x^2 + y^2 = n$, for n square free. Namely, for each prime $p|n$, pick a random uniformly distributed angle $\vartheta_p \in \mathbb{T} := \mathbb{R}/((\pi/2)\mathbb{Z})$ (which should not be confused with the deterministic angle θ_p associated with p.) Letting M denote the number of prime divisors of n, we obtain $r(n) = 2^M$ angles

$$\vartheta = \sum_{p|n} \pm\vartheta_p$$

by varying the M signs, and with probability one these angles are distinct.

It is convenient to parameterize these angles as follows: order the set of prime divisors $p|n$ according to size, i.e., $p_1 < p_2 < \ldots < p_M$ and (abusing notation) put $\vartheta_i = 2\vartheta_{p_i}$ for

$i = 1, \dots, M$ (note that $\vartheta_1, \dots, \vartheta_M$ are then independent and identically distributed uniform random variables on $\mathbb{T}$), and for each subset $J \subset \{1, \dots, M\}$ define a random variable

$$x_J := \sum_{j \in J} \vartheta_j \pmod{\pi/2}; \tag{3.1}$$

note that these points and the original points differ by a translation by $-\sum_{p|n} \vartheta_p$ and hence their spacings as well as correlations are identical.

We define the r-correlation of the points $(\theta_\beta + x_J)_{\substack{(\beta)\subset\mathcal{O}:|\mathcal{O}/\beta|=n_0 \\ J\subset\{1,\dots,M\}}}$ by

(3.2)

$$\mathcal{R}_{r,n_0}(F_N) = \frac{1}{N} \sum_{\substack{J_1,\dots,J_r \subset \{1,\dots,M\} \\ |\mathcal{O}/(\beta_1)|=\dots=|\mathcal{O}/(\beta_r)|=n_0}} F_N(\theta_{\beta_1} - \theta_{\beta_2} + x_{J_1} - x_{J_2}, \dots, \theta_{\beta_{r-1}} - \theta_{\beta_r} + x_{J_{r-1}} - x_{J_r}),$$

where $N = 2^M r(n_0)$. We then show that analogues of the mean and the variance in the deterministic model are well matched by the mean and variance of the random model.

Theorem 3.1. *Let $A > 0$ be fixed and $n_0 \in \mathcal{S}$. Suppose that $1 \le M \le A \log\log x$. We have that*

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n \in \mathcal{N}_{M,n_0}(x)} R_{r,n_0}(n; F_N) = \mathbb{E}\big(\mathcal{R}_{r,n_0}(F_N)\big) + O\left(\frac{1}{(\log\log x)} + \frac{1}{M^{10}}\right). \tag{3.3}$$

Additionally,

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n \in \mathcal{N}_{M,n_0}(x)} R_{r,n_0}(n; F_N)^2 = \mathbb{E}\big(\mathcal{R}_{r,n_0}(F_N)^2\big) + O\left(\frac{1}{(\log\log x)} + \frac{1}{M^{10}}\right). \tag{3.4}$$

The implied constants depend at most on f, r, n_0 and A.

4. Applying the LSD method

Let $k_1, \dots, k_{r-1} \in \mathbb{Z}$. Here and in what follows we use the convention that $k_0 = k_r = 0$. To compute the r-correlation we apply (1.2). Switching the order of summation we will need to understand the sum

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n \in \mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n) \quad \text{where} \quad \lambda_{4k}(n) = \sum_{|\mathcal{O}/(\beta)|=n} e^{i4k\theta_\beta} \tag{4.1}$$

for each given $k_1, \dots, k_{r-1}$ (note that while β is only defined up to multiplication by i, the angle $4k\theta_\beta$ is well defined modulo 2π.) This will be done by applying the Landau-Selberg-Delange method, which requires some analytic input.

Let us introduce the following Hecke L-functions. Recall that $\mathcal{O} = \mathbb{Z}[i]$ denotes the ring of integers of $\mathbb{Q}(i)$; recall also that $\mathcal{O}$ is a principal ideal domain. Further, for $k \in \mathbb{Z}$ and an ideal $\mathfrak{b} = (\beta) \subset \mathcal{O}$, let $\Xi_{4k}(\mathfrak{b}) = e^{4ki\theta_\beta}$ and define (note that $|\mathcal{O}/\mathfrak{b}|$ is the norm of the ideal $\mathfrak{b}$)

$$L(s, 4k) = \sum_{0 \neq \mathfrak{b} \subset \mathcal{O}} \frac{\Xi_{4k}(\mathfrak{b})}{|\mathcal{O}/\mathfrak{b}|^s} = \prod_{\mathfrak{p} \subset \mathcal{O}} \left(1 - \frac{\Xi_{4k}(\mathfrak{p})}{|\mathcal{O}/\mathfrak{p}|^s}\right)^{-1} = \prod_p L_p(s, 4k), \tag{4.2}$$

where the sum is over all nonzero ideals of $\mathcal{O}$, the product is over the set of prime ideals in $\mathcal{O}$, and

$$L_p(s,4k)=\prod_{\substack{\mathfrak{p}\subset\mathcal{O}\\ p|(|\mathcal{O}/\mathfrak{p}|)}}\left(1-\frac{\Xi_{4k}(\mathfrak{p})}{p^s}\right)^{-1}.$$

Separately considering the split, inert, and ramified primes we see that

$$\begin{aligned}L_p(s,4k)^{-1}=&\begin{cases}\left(1-\frac{\Xi_{4k}(\mathfrak{p})}{p^s}\right)\left(1-\frac{\Xi_{4k}(\overline{\mathfrak{p}})}{p^s}\right) & \text{if } p\equiv 1 \pmod 4,\\ 1-\frac{1}{p^{2s}} & \text{if } p\equiv 3 \pmod 4,\\ 1-\frac{e^{\pi ik}}{2^s} & \text{if } p=2\end{cases}\\ =&1-\frac{\lambda_{4k}(p)}{p^s}+\frac{\psi_4(p)}{p^{2s}},\end{aligned}\tag{4.3}$$

where ψ_4 denotes the non-principal character modulo 4; we have also used that $(2)=((1+i)^2)$ as ideals in $\mathcal{O}$. The L-function $L(s,4k)$ admits an analytic continuation to all of $\mathbb{C}$ provided that $k\neq 0$; for $k=0$, $L(s,0)=\zeta_{\mathbb{Q}(i)}(s)=\zeta(s)L(s,\psi_4)$, which has a simple pole at $s=1$. Additionally, for primes $p\equiv 1 \pmod 4$ we write $\theta_p=\theta_{(\pi)}=\arg(\pi)$ where π is the Gaussian prime with $|\pi|^2=p$ and $0<\arg(\pi)<\pi/4$, so that for any $a\in\mathbb{N}$ we have

$$\lambda_{4k}(p^a)=\frac{1}{p^{2ka}}\sum_{l=0}^{a}\pi^{4k}\overline{\pi}^{4k(a-l)}=e^{-4kai\theta_p}\sum_{l=0}^{p}e^{8kli\theta_p}.$$

Hence, we conclude that for any $a,k\in\mathbb{N}$ that

$$\lambda_{4k}(p^a)=\begin{cases}e^{-4kai\theta_p}\sum_{l=0}^{a}e^{8kli\theta_p} & \text{if } p\equiv 1 \pmod 4,\\ 1_{2|a} & \text{if } p\equiv 3 \pmod 4,\\ (-1)^{ka} & \text{if } p=2.\end{cases}\tag{4.4}$$

Given $\mathbf{S}\subset\{0,1,\ldots,r-1\}$ and $k_1,\ldots,k_{r-1}\in\mathbb{Z}$ define

$$k_{\mathbf{S}}=k_{\mathbf{S}}(k_1,k_2,\ldots,k_{r-1})=\sum_{j\in\mathbf{S}}(k_{j+1}-k_j)\in\mathbb{Z}$$

and observe $k_{\mathbf{S}}+k_{\mathbf{S}^c}=\sum_{j=0}^{r-1}(k_{j+1}-k_j)=0$ (recall $k_0=k_r=0$ by convention.) Consequently for $p\equiv 1 \pmod 4$, we have that $\lambda_{4k_{\mathbf{S}}}(p)=e^{i4k_{\mathbf{S}}\theta_p}+e^{-i4k_{\mathbf{S}}\theta_p}=e^{i4k_{\mathbf{S}}\theta_p}+e^{i4k_{\mathbf{S}^c}\theta_p}$. Motivated by the previous observation, let $\mathcal{M}_r=\mathcal{P}(\{0,\ldots,r-1\})/\sim$ where for $\mathbf{S},\mathbf{T}\in\mathcal{P}(\{0,\ldots,r-1\})$ we write $\mathbf{S}\sim\mathbf{T}$ if $\mathbf{S}=\mathbf{T}^c$ where $\mathbf{T}^c=\{0,\ldots,r-1\}\setminus\mathbf{T}$ and for a set $\mathcal{A}$ we use the notation $\mathcal{P}(\mathcal{A})$ to denote the power set of $\mathcal{A}$. Using the statements above we find that

$$\begin{aligned}\prod_{j=0}^{r-1}\lambda_{4(k_{j+1}-k_j)}(p)&=\prod_{j=0}^{r-1}e^{-4(k_{j+1}-k_j)i\theta_p}(1+e^{8(k_{j+1}-k_j)i\theta_p})\\ &=\prod_{j=0}^{r-1}(1+e^{8(k_{j+1}-k_j)i\theta_p})=\sum_{\mathbf{S}\subset\{0,1,\ldots,r-1\}}e^{8k_{\mathbf{S}}i\theta_p}=\sum_{[\mathbf{S}]\in\mathcal{M}_r}\lambda_{8k_{\mathbf{S}}}(p).\end{aligned}\tag{4.5}$$

With $\vec{k} = (k_1, \dots, k_{r-1}) \in \mathbb{Z}^{r-1}$ define

$$\alpha = \alpha(\vec{k}) = \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} = 0}} 1, \tag{4.6}$$

and note that $\#\mathcal{M}_r \le 2^{r-1}$. Further, as $k_\emptyset(\vec{k}) = 0$ there is at least one term in the sum in (4.6), and thus we have $1 \le \alpha(\vec{k}) \le 2^{r-1}$ for all $\vec{k} \in \mathbb{Z}^{r-1}$. Using a classical result of Hecke (see [24, p. 130 & Eq'n (5.52)] or [33, Exercise 14, p. 385]), which gives equidistribution of angles of Gaussian primes, there exists an absolute constant $c > 0$ such that

$$\sum_{\substack{\mathfrak{p} \subset \mathcal{O} \\ |\mathcal{O}/\mathfrak{p}| \le x}} \Xi_{4k}(\mathfrak{p}) \log |\mathcal{O}/\mathfrak{p}| = 1_{k=0}\, x + O\left((|k|+1) x e^{-c\sqrt{\log x}} \right).$$

Hence, we have that

$$\sum_{\substack{p \le x \\ p \nmid n_0}} \left(\prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(p) \right) \log p = \alpha(\vec{k}) x + O(x/(\log x)^B) \tag{4.7}$$

uniformly for $\max_j |k_j| \le e^{(\log x)^{1/3}}$ and every $B > 0$, where the implied constant depends at most on B, r and n_0.[1]

We now state the main result of this section, which provides an estimate for (4.1).

Proposition 4.1. *Let $Y = (M-1)/\log\log x$. Also let both $A > 0$ and $n_0 \in \mathcal{S}$ be fixed. Let $\alpha(\vec{k})$ be as in (4.6). Suppose $1 \le M \le A \log\log x$. Then there exist absolute constants $A', c_0 > 0$ such that for $\max_{[\mathbf{S}] \in \mathcal{M}_r} |k_{\mathbf{S}}| \le e^{(\log x)^{c_0}}$ we have that*

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n \in \mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n) = (2\alpha(\vec{k}))^M \left(g(\vec{k}; Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x} \right) \right),$$

where

(4.8)

$$g(\vec{k}; Y) = \prod_{\substack{p \equiv 1 \pmod 4 \\ (p, n_0) = 1}} \left(1 + \frac{Y \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \ne 0}} \lambda_{8k_{\mathbf{S}}}(p)}{\alpha(\vec{k}) p (1 + \frac{2Y}{p})} \right) \quad \text{and} \quad \mathcal{L}(\vec{k}) = \sum_{\pm} \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \ne 0}} |L(1, 8k_{\mathbf{S}})|^{\pm A'}.$$

4.1. **Sums of multiplicative functions.** To prove Proposition 4.1 we will use the Landau-Selberg-Delange method to estimate the mean value of certain multiplicative functions following the arguments of Tenenbaum [41, Ch. II.5-II.6]. For more background on the LSD method, including refinements of the classical approach see the works of Granville-Koukoulopoulos [19] and de La Bretéche- Tenenbaum [15].

[1] Hence, it can be shown that the order of the pole of the Dirichlet series associated to $\prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n)$ at $s = 1$ equals $\alpha(\vec{k})$.

For $w \in \mathbb{C}$ with $|w| \le A$ and $r \ge 2$ define the multiplicative function $f(\cdot; w)$ supported on integers co-prime to n_0 given by

$$f(n;w) = w^{\Omega_1(n)} b(n)\mu^2(2n) \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n).$$

For $\mathrm{Re}(s) > 1$ let

$$F(s;w) = \sum_{n\ge 1} \frac{f(n;w)}{n^s} = \prod_{(p,2n_0)=1} \left(1 + \frac{wb(p)\sum_{[\mathbf{S}]\in\mathcal{M}_r} \lambda_{8k_{\mathbf{S}}}(p)}{p^s}\right). \tag{4.9}$$

Given $w \in \mathbb{C}$ with $|w| \le A$ we also define

$$\begin{aligned} \widetilde{c}_0(\alpha(\vec{k})w) &= \prod_p \left(1 + \frac{f(p;w)}{p}\right)\left(1 - \frac{1}{p}\right)^{\alpha(\vec{k})w}, \\ \widetilde{C}_0(w) &= \prod_p \left(1 + \frac{1_{(p,2n_0)=1} wb(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{w/2}. \end{aligned} \tag{4.10}$$

Recall (4.7) and that $b(p) = \frac{1}{2}(1+\psi_4(p))$ where ψ_4 is the nonprincipal character modulo 4. Noting that $(1-\frac{1}{p})^z = 1 - \frac{z}{p} + \cdots$ and recalling (4.7) we see the Euler products defined in (4.10) converge. Using the LSD method we will prove the following result.

Lemma 4.2. *Let $A > 0$ and $n_0 \in \mathcal{S}$. Suppose $w \in \mathbb{C}$ satisfies $|w| \le A$. For $x \ge 2$ we have*

$$\sum_{\substack{n \le x \\ (n,2n_0)=1}} w^{\Omega_1(n)} b(n)\mu^2(n) = x(\log x)^{w/2-1}\left(\frac{\widetilde{C}_0(w)}{\Gamma(w/2)} + O\left(\frac{1}{\log x}\right)\right). \tag{4.11}$$

and

$$\sum_{n\le x} f(n;w) = x(\log x)^{\alpha(\vec{k})w-1}\left(\frac{\widetilde{c}_0(\alpha(\vec{k})w)}{\Gamma(\alpha(\vec{k})w)} + O\left(\frac{(\log K)^{O(1)}}{\log x}\right)\right), \tag{4.12}$$

where the implied constants depend on n_0, r, A.

We also we require the following result, which is a special case of [41, Th'm II.6.1.3].

Theorem 4.3 (Tenenbaum [41])**.** *Let $z \in \mathbb{C}$. Let $a_z(\cdot)$ be an arithmetic function for which there exists $A > 0$ such that for each $n \in \mathbb{N}$ there is a power series expansion in the disc $|z| \le A$ of the form*

$$a_z(n) = \sum_{M=0}^{\infty} c_M(n) z^M.$$

Suppose there exists a function $h(z)$ that is holomorphic in $|z| \le A$ such that for $x \ge 3$ and $|z| \le A$

$$\sum_{n\le x} a_z(n) = x(\log x)^{z-1}(zh(z) + O(R(x)))$$

for some quantity $R(x)$, which does not depend on z and the implied constant depends at most on A. Then for $x \geq 3$ and $1 \leq M \leq A \log\log x$ we have

$$\sum_{n \leq x} c_M(n) = \frac{x}{\log x} \frac{(\log\log x)^{M-1}}{(M-1)!} \left(h\left(\frac{M-1}{\log\log x}\right) + O\left(\frac{(M-1)}{(\log\log x)^2} \max_{|z| \leq A} |h''(z)| + \frac{\log\log x}{M} R(x)\right)\right)$$

where the implied constant depends at most on A.

4.2. **Analytic estimates.** Before proceeding to the proofs of Lemma 4.2 and Proposition 4.1 we require further analytic results. Using the generalized binomial theorem, we see that for $z \in \mathbb{C}$ and $\mathrm{Re}(s) > 1$ that

$$L(s, 4k)^z = \prod_{\mathfrak{p} \subset \mathcal{O}} \left(\sum_{j=0}^{\infty} \binom{z+j-1}{j} \frac{\Xi_{4k}(\mathfrak{p}^j)}{|\mathcal{O}/\mathfrak{p}^j|^s} \right) = \sum_{n \geq 1} \frac{\lambda_{4k}(n; z)}{n^s}$$

where $\binom{\cdot}{\cdot}$ denotes the generalized binomial coefficients (in particular, we have $(1-t)^{-z} = \sum_{k=0}^{\infty} \binom{z+k-1}{k} t^k$ for $|t| < 1$) and $\lambda_{4k}(\cdot; z)$ is the multiplicative function given by

$$\lambda_{4k}(n; z) = \sum_{\mathfrak{a} \subset \mathcal{O} : |\mathcal{O}/\mathfrak{a}| = n} b_z(\mathfrak{a}) \Xi_{4k}(\mathfrak{a}), \tag{4.13}$$

where the sum is over ideals in $\mathcal{O}$ having norm n, and for $\mathfrak{a} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_w^{a_w}$, we let $b_z(\mathfrak{a}) = \binom{z+a_1-1}{a_1} \cdots \binom{z+a_w-1}{a_w}$. We have for $|z| \leq A$ that

$$b_z(\mathfrak{a}) \ll |\mathcal{O}/\mathfrak{a}|^{o(1)}, \qquad \lambda_{4k}(n; z) \ll n^{o(1)}, \tag{4.14}$$

uniformly for $k \in \mathbb{Z}$. For $Y \geq 0$, $\vec{k} \in \mathbb{Z}^{r-1}$ define the multiplicative functions $w(\cdot; Y), s(\cdot; \vec{k})$ supported on square free numbers given by

$$w(p; Y) = \frac{Y b(p)}{1 + \frac{2Y}{p}}, \qquad s(p; \vec{k}) = \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \neq 0}} \lambda_{8k_{\mathbf{S}}}(p). \tag{4.15}$$

The following key proposition will allows us to express $g(\vec{k}; Y)$ (cf. (4.8) for its definition) as a short Dirichlet polynomial (later we will choose $y = \exp((\log \|\vec{k}\|_\infty)^{3/4})$.)

Proposition 4.4. *Let $K = \max_{[\mathbf{S}] \in \mathcal{M}_r} |k_{\mathbf{S}}| + 2$, $A > 0$, and $n_0 \in \mathcal{S}$. Then for $y \geq \exp((\log K)^{2/3+o(1)})$ and $0 \leq Y \leq A$, we have that*

$$g(\vec{k}; Y) = \sum_{\substack{m \leq y \\ (m, 2n_0) = 1}} \frac{\alpha(\vec{k})^{-\Omega_1(m)} w(m; Y) s(m; \vec{k})}{m} + O\left(\exp\left(\frac{-\log y}{(\log(K+y))^{2/3+o(1)}}\right)\right), \tag{4.16}$$

where the implied constant depends at most on n_0, A. Additionally, for $k \neq 0$, $y \geq \exp((\log k)^{2/3+o(1)})$, and $|\mathrm{Re}(z)| \leq \frac{\log y}{(\log(k+y))^{2/3+o(1)}}$ we have

$$L(1, 4k)^z = \sum_{m \leq y} \frac{\lambda_{4k}(m; z)}{m} + O\left(\exp\left(\frac{-\log y}{(\log(k+y))^{2/3+o(1)}}\right)\right). \tag{4.17}$$

Before proceeding to the proof let us note that for each $1 \leq j \leq r-1$ choosing $\mathbf{S} = \{0, 1, \ldots, j-1\}$ yields $k_{\mathbf{S}} = k_j$, so that $\|\vec{k}\|_\infty \leq K$.

Proof. Since (4.16) and (4.17) follow from similar arguments we will only prove (4.16). For $\mathrm{Re}(s) > 1$ define $Z(s) = Z(s, \vec{k})$ by
(4.18)

$$Z(s) = \prod_{(p,2n_0)=1} \left(1 + \frac{Y b(p) \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \neq 0}} \lambda_{8k_{\mathbf{S}}}(p)}{\alpha(\vec{k})(1 + \frac{2Y}{p}) p^s} \right) = \sum_{\substack{m \geq 1 \\ (m,2n_0)=1}} \frac{\alpha(\vec{k})^{-\Omega_1(m)} w(m;Y) s(m;\vec{k})}{m^s}.$$

Further, for $\mathrm{Re}(s) > 1$, define $\mathcal{G}(s) = \mathcal{G}(s, \vec{k})$ by

$$\mathcal{G}(s) = Z(s) \prod_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \neq 0}} L(s, 8k_{\mathbf{S}})^{-Y/\alpha(\vec{k})}.$$

Recalling (4.3), we see that in the region $\mathrm{Re}(s) \geq 1/3$, say, we have for each prime p that

$$\prod_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \neq 0}} L_p(s, 8k_{\mathbf{S}})^{-Y/\alpha(\vec{k})} = 1 - \frac{Y}{\alpha(\vec{k})} \sum_{\substack{[\mathbf{S}] \in \mathcal{M}_r \\ k_{\mathbf{S}} \neq 0}} \frac{\lambda_{8k_{\mathbf{S}}}(p)}{p^s} + O\left(\frac{1}{p^{2\,\mathrm{Re}(s)}} \right), \tag{4.19}$$

where the implied constant depends at most on A, r. Hence, using (4.18) and (4.19) we see for $\mathrm{Re}(s) > 1$ that

$$\mathcal{G}(s) = \prod_p (1 + \mathcal{G}_p(s)) \tag{4.20}$$

where $\mathcal{G}_p(s)$ is holomorphic in $\mathrm{Re}(s) \geq 1/3$, say, and satisfies $\mathcal{G}_p(s) \ll p^{-2\,\mathrm{Re}(s)}$ in this region. By (4.20), it follows that $\mathcal{G}(s)$ admits an analytic continuation to $\mathrm{Re}(s) \geq 1/2 + \varepsilon$ and in this region $|\mathcal{G}(s)| \ll 1$ where the implied constant depends at most on A, n_0, r, ε (recall that $1 \leq \alpha(\vec{k}) \leq 2^{r-1}$.)

Let us state the Vinogradov-Korobov bounds

$$|L(s, 8k)|^{\pm 1} \ll (\log(|k| + |s|))^{2/3+o(1)} \tag{4.21}$$

due to Coleman [12, Thm. 1, 2], which are valid in the region $s = \sigma + it$ with

$$\sigma \geq 1 - \frac{1}{(\log(|k| + |t|))^{2/3+o(1)}}$$

(see also [12, Lem. 8] and [42, Thm. 3.10-3.11], cf. [23, Lem. 10,11].)

Hence, the result follows using a standard contour integration argument by applying Perron's formula [33, Corollary 5.3] as follows. Let $T = y^3 + K$. We get that

$$\sum_{\substack{m \leq y \\ (m,2n_0)=1}} \frac{\alpha(\vec{k})^{-\Omega_1(m)} w(m;Y) s(m;\vec{k})}{m} = \frac{1}{2\pi i} \int_{1-iT}^{1+iT} Z(s+1) \frac{y^s}{s} \, ds + O(y^{-1/2}).$$

Shifting contours to $\sigma = -1/(\log T)^{2/3+o(1)}$ we pick up a simple pole at $s = 0$ with residue $Z(1) = g(\vec{k}; Y)$. The horizontal contour integrals are easily seen to be $\ll y^{-2}(\log T)^{O(1+Y)}$, using (4.21). The vertical contour on the line $\sigma = -1/(\log T)^{2/3+o(1)}$ is

$$\ll (\log T)^{O(1+Y)} \exp\left(\frac{-\log y}{3(\log(K+y))^{2/3+o(1)}} \right)$$

also using (4.21). Collecting estimates completes the proof. $\square$

4.3. **Applications of the LSD method.** The proof of Lemma 4.2 broadly follows the proof of [41, Th'm II.5.3.3] due to Tenenbaum and we require some additional notation. Given $Z > \rho > 0$ and $c \in \mathbb{C}$, we define the partial Hankel contour $\mathcal{H}_c(\rho; Z) := I^- \cup I^+ \cup \mathcal{C}$ as follows. The contour $\mathcal{H}_c(\rho; Z)$ begins with the line segment $I^- = [c - Z, c - \rho]$ oriented left to right, then connects to $\mathcal{C}$ which is the counter-clockwise oriented circle $|s - c| = \rho$ omitting the point $s = c - \rho$, then connects to $I^+ = [c - Z, c - \rho]$ oriented right to left. We cite the following result [41, Corollary II.5.2.2.1].

Lemma 4.5 (Tenenbaum [41])**.** *For $Z > 1$ we have uniformly for $z \in \mathbb{C}$ and $\rho > 0$ that*

$$\frac{1}{2\pi i}\int_{\mathcal{H}_0(\rho;Z)} s^{-z}e^s ds = \frac{1}{\Gamma(z)} + O(47^{|z|}\Gamma(1+|z|)e^{-\frac{1}{2}Z}),$$

where for $s \in I^{\pm}$ we have $\arg(s) = \pm\pi$.

Proof of Lemma 4.2. We will only prove (4.12), since the proof of (4.11) is similar, yet simpler. Arguing as in (4.18), (4.19), and (4.20) we see that for $|w| \le A$ the function

$$L(s, 4k)^{-w} \prod_{(p,2n_0)=1} \left(1 + \frac{wb(p)\lambda_{4k}(p)}{p^s}\right)$$

may be expressed as an Euler product that is absolutely convergent in $\mathrm{Re}(s) \ge 1/2+\varepsilon$ and is $\ll 1$ in that region, where the implied constant depends at most on n_0, A and, in particular, not on k since $|\lambda_{4k}(n)| \le r(n) \ll n^{o(1)}$. Also, recall $L(s, 0) = \zeta(s)L(s, \psi_4)$ where ψ_4 is the nonprincipal character modulo 4. Consequently, we see that we can write (cf. (4.9))

$$F(s; w) = H_{n_0}(s; w) \prod_{[\mathbf{S}]\in\mathcal{M}_r} L(s, 8k_{\mathbf{S}})^w = H_{n_0}(s; w)\zeta(s)^{\alpha(\vec{k})w}L(s, \psi_4)^{\alpha(\vec{k})w} \prod_{\substack{[\mathbf{S}]\in\mathcal{M}_r\\ k_{\mathbf{S}}\neq 0}} L(s, 8k_{\mathbf{S}})^w, \tag{4.22}$$

where $H_{n_0}(s; w)$ is an analytic function with $|H_{n_0}(s; w)| \ll 1$ for $\mathrm{Re}(s) \ge \frac{1}{2} + \varepsilon$ and $|w| \le A$, where the implied constant depends at most on A, ε, n_0, r (recall $1 \le \alpha(\vec{k}) \le 2^{r-1}$.)

Write $K = \max_{[\mathbf{S}]\in\mathcal{M}_r} |k_{\mathbf{S}}| + 2$. Using (4.22) along with (4.21) we have that $F(s; w)$ is holomorphic in the region $\mathrm{Re}(s) \ge 1 - (\log(|\mathrm{Im}(s)| + K))^{-2/3-o(1)}$, $s \neq 1$ and satisfies the bound

$$F(s; w) \ll \big(\log(K + |\mathrm{Im}(s)|)(|s-1|^{-1} + 1)\big)^{O(1)}. \tag{4.23}$$

Additionally, by a similar argument and noting $\zeta(s) \sim (s-1)^{-1}$ as $s \to 1$ we have for $|s - 1| \le (\log K)^{-1}$ that

$$(s-1)^{\alpha(\vec{k})w}F(s; w) \ll (\log K)^{O(1)}. \tag{4.24}$$

By (4.22), $(s-1)^{\alpha(\vec{k})w}F(s; w)$ is analytic in a neighborhood of $s = 1$ and for each nonnegative integer j we define

$$\widetilde{c}_j(\alpha(\vec{k})w) = \frac{1}{j!}\frac{d^j}{ds^j}\bigg|_{s=1} \frac{(s-1)^{\alpha(\vec{k})w}F(s; w)}{s}.$$

For $j=0$ this is the same as (4.10) since $\zeta(s) \sim (s-1)^{-1}$ as $s \to 1$. Using Cauchy's theorem and (4.24) we see that

$$\widetilde{c}_j(\alpha(\vec{k})w) = \frac{1}{2\pi i} \int_{|s-1|=\frac{1}{\log K}} (s-1)^{\alpha(\vec{k})w} F(s;w) \frac{ds}{s(s-1)^{j+1}} \ll (\log K)^{O(j)}, \tag{4.25}$$

where the implied constant depends on n_0, r, A.

We note that unless $K \leq \exp((\log x)^{c_0})$ for some $c_0 > 0$, (4.12) holds trivially upon noting that

$$|f(n;w)| \leq A^{\Omega_1(n)} 2^{r\Omega(n)}.$$

We assume $K \leq \exp((\log x)^{c_0})$ for some $c_0 > 0$ where $c_0 > 0$ is sufficiently small (in terms of A, r.) By Perron's formula (cf. [33, Corollary 5.3]) for $x \geq 2$ and $T = e^{\sqrt{\log x}}$ we have that

$$\sum_{n \leq x} f(n;w) = \frac{1}{2\pi i} \int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT} F(s;w) \frac{x^s}{s}\, ds + O(xe^{-\frac{1}{2}\sqrt{\log x}}). \tag{4.26}$$

Let $\rho = 1/(2\log x)$ and $Z = (\log T)^{-1}$. Consider the positively oriented contour $\mathcal{D}$, which consists of the line segments that connect $1 + \frac{1}{\log x} - iT$ to $1 + \frac{1}{\log x} + iT$, $1 + \frac{1}{\log x} + iT$ to $1 - Z + iT$, $1 - Z + iT$ to $1 - Z$, together with the contour $\mathcal{H}_1(\rho, Z)$ with reversed orientation, along with the line segments that connect $1 - Z$ to $1 - Z - iT$ and $1 - Z - iT$ to $1 + \frac{1}{\log x} - iT$. By Cauchy's theorem

$$\frac{1}{2\pi i} \int_{\mathcal{D}} F(s;w) \frac{x^s}{s} = 0.$$

Consequently using (4.23) to bound the horizontal contours that connect $1 + \frac{1}{\log x} \pm iT$ to $1 - Z \pm iT$ and the left-hand side, vertical contours, we see that

$$\begin{aligned} \frac{1}{2\pi i} \int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT} F(s;w) \frac{x^s}{s}\, ds =& \frac{1}{2\pi i} \int_{\mathcal{H}_1(\rho,Z)} F(s;w) \frac{x^s}{s}\, ds \\ &+ O\left(\frac{x}{T} (\log x)^{O(1)} + x^{1-Z} (\log x)^{O(1)} \right) \end{aligned} \tag{4.27}$$

where the contour $\mathcal{H}_1(\rho, Z)$ is oriented counter-clockwise. The error term is $\ll xe^{-\sqrt{\log x}/2}$.

Recalling (4.25), we have for $s \in \mathcal{H}_1(\rho, Z)$ (so in particular $|s-1| = o((\log K)^{-1})$) and a fixed integer $J \geq 0$ that

$$\begin{aligned} \frac{(s-1)^{\alpha(\vec{k})w} F(s;w)}{s} =& \sum_{j=0}^{\infty} \widetilde{c}_j(\alpha(\vec{k})w)(s-1)^j \\ =& \sum_{j=0}^{J} \widetilde{c}_j(\alpha(\vec{k})w)(s-1)^j + O(|s-1|^{J+1}(\log K)^{O(1)}), \end{aligned}$$

where the implied constants depend on n_0, A, r, J. Hence, we have

$$\begin{aligned} \frac{1}{2\pi i} \int_{\mathcal{H}_1(\rho,Z)} F(s;w) \frac{x^s}{s}\, ds =& \sum_{j=0}^{J} c_j(\alpha(\vec{k})w) \frac{1}{2\pi i} \int_{\mathcal{H}_1(\rho,Z)} (s-1)^{j-\alpha(\vec{k})w} x^s\, ds \\ &+ O\left((\log K)^{O(1)} \int_{\mathcal{H}_1(\rho,Z)} x^{\mathrm{Re}(s)} |s-1|^{J+1-\alpha(\vec{k})\,\mathrm{Re}\, w} |ds| \right). \end{aligned} \tag{4.28}$$

To bound the error term we separately consider the contributions from the portions of $\mathcal{H}_1(\rho, Z)$ that correspond to $I^{\pm}$ and $\mathcal{C}$ to see that

(4.29)
$$\int_{\mathcal{H}_1(\rho,Z)} x^{\mathrm{Re}(s)}|s-1|^{J+1-\alpha(\vec{k})\,\mathrm{Re}\,w}|ds| \ll \int_{1-Z}^{1-\rho} x^{\sigma}(1-\sigma)^{J+1-\alpha(\vec{k})\,\mathrm{Re}\,w}d\sigma + x(\log x)^{\alpha(\vec{k})\,\mathrm{Re}\,w-J-2}$$
$$\ll x(\log x)^{\alpha(\vec{k})\,\mathrm{Re}\,w-J-2},$$

where in the last step we made the change of variables $u = (1-\sigma)\log x$ to bound the integral.

Let $U = Z\log x = (\log x)^{1/2}$. We now make the change of variables $z = (s-1)\log x$ then apply Lemma 4.5 to get for each $0 \le j \le J$ that

$$\frac{1}{2\pi i}\int_{\mathcal{H}_1(\rho,Z)} (s-1)^{j-\alpha(\vec{k})w}x^s\,ds = \frac{x}{(\log x)^{j-\alpha(\vec{k})w+1}}\frac{1}{2\pi i}\int_{\mathcal{H}_0(1/2,U)} z^{j-\alpha(\vec{k})w}e^z\,dz$$
$$= \frac{x}{(\log x)^{j-\alpha(\vec{k})w+1}}\left(\frac{1}{\Gamma(\alpha(\vec{k})w-j)} + O(e^{-\frac{1}{2}U})\right),$$

where the implied constant depends on J, r, A. Use the preceding estimate along with (4.29) in (4.28) then apply the resulting formula in (4.27). With this formula in hand, choose $J = A2^r$ (recall $0 \le \alpha(\vec{k}) \le 2^{r-1}$) to obtain

$$\frac{1}{2\pi i}\int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT} F(s;w)\frac{x^s}{s}\,ds = x\sum_{j=0}^{J}\widetilde{c}_j(\alpha(\vec{k})w)\frac{(\log x)^{\alpha(\vec{k})w-j-1}}{\Gamma(\alpha(\vec{k})w-j)} + O(x(\log x)^{-J/2}).$$

Applying this in (4.26) and using (4.25) completes the proof. □

Proof of Proposition 4.1. We will first obtain estimates for $\#\mathcal{N}_{M,n_0}(x)$ and $\sum_{n\in\mathcal{N}_{M,n_0}(x)}\prod_{j=0}^{r-1}\lambda_{4(k_{j+1}-k_j)}(n)$. We will only provide a detailed proof of the latter as the argument for estimating $\#\mathcal{N}_{M,n_0}(x)$ is similar, yet simpler.

We will apply Theorem 4.3 together with Lemma 4.2 and to match notation we write $z = \alpha(\vec{k})w$ (recall that $1 \le \alpha(\vec{k}) \le 2^{r-1}$.) Let $h(z) = \frac{\widetilde{c}_0(z)}{z\Gamma(z)}$. For $|z| \le A$, (4.12) implies

$$\sum_{n\le x} f(n; z/\alpha(\vec{k})) = x(\log x)^{z-1}\left(zh(z) + O\left(\frac{(\log K)^{O(1)}}{\log x}\right)\right). \tag{4.30}$$

Since $K \le e^{(\log x)^{c_0}}$ and c_0 is sufficiently small the error term above is negligible. By (4.10) (with $\alpha(\vec{k})w = z$) and (4.22) we find that

(4.31)
$$h(z) = \frac{1}{z\Gamma(z)}\prod_p\left(1 + \frac{f(p;\frac{z}{\alpha(\vec{k})})}{p}\right)\left(1-\frac{1}{p}\right)^z = \frac{H_{n_0}(1; z/\alpha(\vec{k}))L(1,\psi_4)^z}{z\Gamma(z)}\prod_{\substack{[\mathbf{S}]\in\mathcal{M}_r\\ k_{\mathbf{S}}\neq 0}} L(1, 8k_{\mathbf{S}})^{z/\alpha(\vec{k})}$$

where $H_{n_0}(1;\frac{z}{\alpha(\vec{k})}) \ll 1$ for $|z| \le A$ and is analytic in that region. Hence for $|z| \le A$ we have that

$$|h''(z)| \ll \prod_{\substack{[\mathbf{S}]\in\mathcal{M}_r\\ k_{\mathbf{S}}\neq 0}} \max_{\pm}|L(1,8k_{\mathbf{S}})|^{\pm(A+2)} \ll \sum_{\pm}\sum_{\substack{[\mathbf{S}]\in\mathcal{M}_r\\ k_{\mathbf{S}}\neq 0}} |L(1,8k_{\mathbf{S}})|^{\pm A'} =: \mathcal{L}(\vec{k}) \tag{4.32}$$

where the last step follows from the inequality of arithmetic and geometric means and A' and the implied constant depend at most on n_0, r, A (here we also used that $|\log t| \le \max(t, 1/t)$ for $t > 0$.) Taking $a_z(n) = f(n; \frac{z}{\alpha(\vec{k})})$ we have for $|z| \le A$ that

$$a_z(n) = \sum_{M=0}^{\infty} c_M(n) z^M, \qquad \text{where} \qquad c_M(n) = \begin{cases} f(n; \alpha(\vec{k})^{-1}) & \text{if } M = \Omega_1(n), \\ 0 & \text{otherwise.} \end{cases}$$

Write $X = \log\log x$. Hence, using (4.30), (4.31), and (4.32) we may apply Theorem 4.3 with $R(x) = (\log K)^{O(1)}/\log x$ to get that for $Y = \frac{M-1}{X}$

(4.33)

$$\sum_{n \in \mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n) = \alpha(\vec{k})^M \sum_{n \le x} c_M(n)$$

$$= \alpha(\vec{k})^M \frac{x}{\log x} \frac{X^{M-1}}{(M-1)!} \left(\frac{1}{Y\Gamma(Y)} \prod_p \left(1 + \frac{f(p; \frac{Y}{\alpha(\vec{k})})}{p} \right) \left(1 - \frac{1}{p} \right)^Y + O\left(\frac{M\mathcal{L}(\vec{k})}{X^2} \right) \right),$$

where we have used that for $K \le e^{(\log x)^{c_0}}$ with c_0 sufficiently small $R(x)X/M \ll M\mathcal{L}(\vec{k})/X^2$ (cf. (4.21)) to simplify the error term. By a similar, yet simpler argument that we will omit we also get for $1 \le M \le A \log\log x$ that

(4.34)

$$\#\mathcal{N}_{M,n_0}(x) = 2^{-M} \frac{x}{\log x} \frac{X^{M-1}}{(M-1)!} \left(\frac{1}{Y\Gamma(Y)} \prod_p \left(1 + \frac{2Y 1_{(p,2n_0)=1} b(p)}{p} \right) \left(1 - \frac{1}{p} \right)^Y + O\left(\frac{M}{X^2} \right) \right).$$

The Euler product on the right-hand side of (4.34) is $\asymp 1$. Also, the ratio of the Euler product on the right-hand side of (4.33) to that on the right-hand side of (4.34) is

$$\prod_{(p,2n_0)=1} \left(1 + \frac{Y b(p) \sum_{[\mathbf{S}] \in \mathcal{M}_r} \lambda_{8k_{\mathbf{S}}}(p)}{\alpha(\vec{k}) p} \right) \Big/ \left(1 + \frac{2Y b(p)}{p} \right) = g(\vec{k}; Y),$$

which follows from grouping the terms in the summation over $[\mathbf{S}] \in \mathcal{M}_r$ with $k_{\mathbf{S}} = 0$, for which $\lambda_{8k_{\mathbf{S}}}(p) = 2$, and those with $k_{\mathbf{S}} \ne 0$. Hence, combining (4.33) and (4.34) we have that

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n \in \mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n) = (2\alpha(\vec{k}))^M \left(g(\vec{k}; Y) + O\left(\frac{M}{(\log\log x)^2} \mathcal{L}(\vec{k}) \right) \right).$$

□

5. Averaging over $(k_1, \ldots, k_{r-1})$

The main result of this section is the following proposition, which reduces the computation of the r-correlation to a combinatorial expression. For $\vec{k} = (k_1, \ldots, k_{r-1}) \in \mathbb{Z}^{r-1}$, and using the convention $k_0 = k_r = 0$, let us define

$$\ell_{\vec{k}}(n_0) = \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n_0). \tag{5.1}$$

Proposition 5.1. *Let $A > 0$ and $n_0 \in \mathcal{S}$. For $1 \le M \le A \log\log x$ we have that*

(5.2)
$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n\in\mathcal{N}_{M,n_0}(x)} R_{r,n_0}(n;F_N) = \frac{1}{N^r} \sum_{\vec{k}\in\mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0)(2\alpha(\vec{k}))^M + O\left(\frac{1}{\log\log x} + \frac{1}{M^{10}}\right)$$

where $\alpha(\vec{k})$ is as given in (4.6) *and the implied constant depends on at most f, n_0, r, and A.*

We recall that F_N is defined in (1.2), and that f is a Schwartz function. To prove the proposition we first apply the main result of the previous section. In the left-hand side of (5.2) we apply (1.2) (see also (1.3)) and observe that by the rapid decay of $\widehat{f}$ and the trivial bound $|\lambda_{4k}(n)| \le 2^M$ for $n \in \mathcal{N}_{M,n_0}$ the terms with $\|\vec{k}\|_\infty \ge N^{1+o(1)}$ contribute $\ll N^{-B}$ for any $B > 0$, so that we can add or remove these terms at the cost of a negligible error term. Using the previous observation and applying Proposition 4.1 the left-hand side of (5.2) equals

$$\begin{aligned} &= \frac{1}{N^r} \sum_{\vec{k}\in\mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0) \frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n\in\mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n) \\ &= \frac{1}{N^r} \sum_{\vec{k}\in\mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0)(2\alpha(\vec{k}))^M \left(g(\vec{k};Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right) + O\left(\frac{1}{N^{10}}\right). \end{aligned} \tag{5.3}$$

Since $\alpha(\vec{k}) = \sum_{\substack{[\mathbf{S}]\in\mathcal{M}_r \\ k_{\mathbf{S}}=0}} 1$ depends on $\vec{k}$ in a complex way it is not immediately clear how to perform the average over $\vec{k}$. We will decompose $\mathbb{R}^{r-1}$ into certain "subspaces" on which the value of $\alpha(\vec{k})$ is constant (outside other such subspaces of lower dimension.)

5.1. **The decomposition.** Given $S \subset \mathcal{M}_r$ fix an ordering $S = \{[\mathbf{S}_1], \dots, [\mathbf{S}_{|S|}]\}$ as well as choice of equivalence class representatives $\mathbf{S}_i$, $i = 1, \dots, |S|$, and let us define the matrix $M_S = (a_{ij}) \in \mathrm{Mat}_{|S|\times r-1}(\mathbb{R})$ where

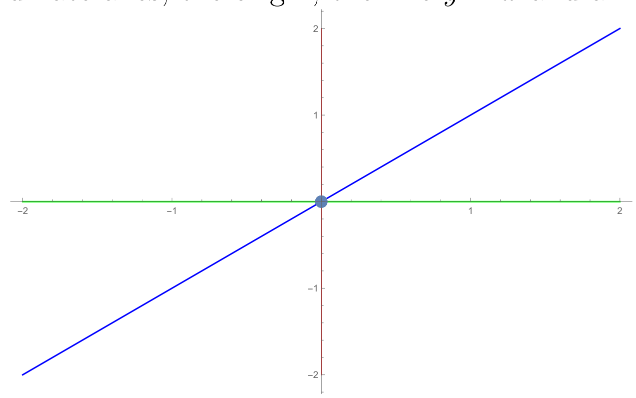
$$a_{ij} = \begin{cases} -1 & \text{if} \quad j \in \mathbf{S}_i, j-1 \notin \mathbf{S}_i, \\ 0 & \text{if} \quad j, j-1 \in \mathbf{S}_i \text{ or } j, j-1 \notin \mathbf{S}_i, \\ 1 & \text{if} \quad j-1 \in \mathbf{S}_i, j \notin \mathbf{S}_i. \end{cases} \tag{5.4}$$

While M_S depends on the ordering and choice of equivalence class representatives of the elements of S the kernel of M_S does not and since we will only be concerned with the latter this resolves the issue of indeterminacy for us. For example, for $r = 3$ listing the elements of $S = \mathcal{M}_3$ as $\{[\emptyset], [\{0\}], [\{1\}], [\{2\}]\}$ gives

$$M_{\mathcal{M}_3} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}. \tag{5.5}$$

The kernels of the matrices M_S as S ranges over subsets of $\mathcal{M}_r$ encode the structure of the r-correlation of the lattice points. For example, for the triple correlation (i.e. the left-hand side of (3.3) with $r = 3$) it turns out that without specifying further conditions on the Schwartz function f, such as $f(0, y) = 0$, there are five leading order terms. The kernels of the matrices M_S, $S \subset \mathcal{M}_3$, correspond to the x, y-axes, the line $y = x$, the origin, and all of $\mathbb{R}^2$ and these five subspaces give rise to the five main terms in the triple correlation. After specifying f to detect only *distinct* angles four of the five main terms vanish and the one remaining main term is $\widehat{f}(0)$, which is consistent with Poisson spacing statistics. For larger r, directly computing the r-correlation in this way becomes difficult since the number of subspaces that need to be analyzed grows exponentially with r. We will pursue a different approach that uses a random model to indirectly solve this combinatorial problem.

FIGURE 4. Plot of the kernels of M_S as S varies over $\mathcal{M}_3$, which correspond to the x, y-coordinate axes, the origin, the line $y = x$ and all of $\mathbb{R}^2$.



We now partition $\mathbb{R}^{r-1}$ into subspaces given by kernels of M_S as S varies over $S \subset \mathcal{M}_r$. Since this is a partition, we need to remove the intersection with other kernels. We define

$$\ker(M_S)^\star = \ker(M_S) \setminus \bigcup_{S_1 \subset \mathcal{M}_r : S_1 \supsetneq S} \ker(M_{S_1}).$$

Let us note that we may have $\ker(M_S) = \ker(M_{S_1})$ for $S_1 \supsetneq S$, in which case $\ker(M_S)^\star$ is empty and we say that S is *non-maximal*, we say S is *maximal* if $\ker(M_S)^\star$ is non-empty. We now let

$$\mathcal{V}_r = \{\ker(M_S)^\star : S \subset \mathcal{M}_r\}, \tag{5.6}$$

which gives us our decomposition, into disjoint subsets,

$$\coprod_{V^\star \in \mathcal{V}_r} V^\star = \mathbb{R}^{r-1}. \tag{5.7}$$

Example. *In the case $r = 3$, the partition is determined by considering the kernels of the submatrices which consist of collections of rows of $M_{\mathcal{M}_3}$ in* (5.5) *and our partition of $\mathbb{R}^2$ is*

$$\begin{aligned} V_1^\star =& \{(x, y) : x \neq 0, y \neq 0, x \neq y\}, \\ V_2^\star =& \{(x, y) : x = 0, y \neq 0\}, \\ V_3^\star =& \{(x, y) : y = x, x \neq 0\}, \\ V_4^\star =& \{(x, y) : y = 0, x \neq 0\}, \end{aligned}$$

and $V_5^\star = \{(0, 0)\}$ which correspond to the kernels of $M_{\{[\emptyset]\}}$, $M_{\{[\emptyset],[\{0\}]\}}$, $M_{\{[\emptyset],[\{1\}]\}}$, $M_{\{[\emptyset],[\{2\}]\}}$, $M_{\mathcal{M}_3}$ respectively, after removing intersections.

5.2. **Finding cancellation.** Given $V^\star \in \mathcal{V}_r$ let V denote the linear span of the elements in $V^\star$ and define

$$d = \dim(V). \tag{5.8}$$

In order to bound the analytic factor $\alpha(\vec{k})$ (as defined in (4.6)) in terms of d we begin with a simple geometric lemma.

Lemma 5.2. *Let $W \subset \mathbb{R}^n$ be a linear subspace of dimension d. Then*

$$|W \cap \{0,1\}^n| \leq 2^d,$$

i.e., W can intersect the corners of the n-dimensional hypercube, with one corner at the origin, in at most 2^d points.

Proof. Choose a basis $\vec{w}_1, \dots, \vec{w}_d$ for W and define a $n \times d$ matrix $A = [\vec{w}_1 \dots \vec{w}_d]$. Since A has column rank d, there exists d independent rows in A, and by deleting all other rows we may form an invertible $d \times d$ matrix B. Now, any point $\vec{w} \in W$ can be written as $\vec{w} = A\vec{x}$ for a unique $\vec{x} \in \mathbb{R}^d$, and we further note that $A\vec{x} \in \{0,1\}^n$ implies that $B\vec{x} \in \{0,1\}^d$. Thus, since B is invertible there are exactly 2^d possible choices of $\vec{x}$ so that $B\vec{x} \in \{0,1\}^d$, and the claimed upper bound follows. □

Lemma 5.3. *Given $V^\star \in \mathcal{V}_r$ there exists $\alpha_{V^\star} \geq 1$ such that*

$$\alpha(\vec{k}) = \alpha_{V^\star}, \qquad \forall \vec{k} = (k_1, \dots, k_{r-1}) \in V^\star. \tag{5.9}$$

Further, with d as in (5.8) *we have*

$$\alpha_{V^\star} \leq 2^{r-d-1}. \tag{5.10}$$

Proof. We begin by introducing some convenient notation to parameterize the sets of $\vec{k}$'s on which $\alpha(\vec{k})$ is constant. Define a bilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^r \times \mathbb{R}^{r-1}$ by

$$\langle \vec{l}, \vec{k} \rangle = \sum_{j=0}^{r-1} l_{j+1}(k_{j+1} - k_j). \tag{5.11}$$

Let $W_r = \mathcal{P}(\{0, 1, \dots, r-1\})$. For $\mathbf{S} \in W_r$ define $\vec{l}_{\mathbf{S}} = (l_j) \in \mathbb{R}^r$ by $l_j = 1$ if $j \in \mathbf{S}$ and $l_j = 0$ if $j \notin \mathbf{S}$. Observe that

$$2 \cdot \alpha(\vec{k}) = \#\{\mathbf{S} \in W_r : \langle \vec{l}_{\mathbf{S}}, \vec{k} \rangle = 0\}, \tag{5.12}$$

where the factor of 2 accounts for the fact that on the right-hand side for each $\mathbf{S} \in W_r$ we have also counted $\mathbf{S}^c$ whereas these sets have been identified in $\mathcal{M}_r$.

Each element $\mathbf{S} \in W_r$ corresponds to an element of $\{0,1\}^r \subset \mathbb{R}^r$ under the bijection $\mathbf{S} \to (1_{j \in \mathbf{S}})_{0 \leq j \leq r-1}$. Using this and (5.12) we see that

$$2\alpha(\vec{k}) = \#\{\vec{l} \in \{0,1\}^r : \langle \vec{l}, \vec{k} \rangle = 0\}. \tag{5.13}$$

Also for $S \subset W_r$ we define M_S as in (5.4) (the only difference is that we do not need to fix equivalence class representatives) and also define $\ker(M_S)^\star$ analogously (which is the same as before since choosing equivalence class representatives does not affect the kernels.) Additionally, under the bijection described above each $S \subset W_r$ is associated to some $\widetilde{S} \subset \{0,1\}^r$ so that we can re-express $\ker(M_S) \subset \mathbb{R}^{r-1}$ as

$$\ker(M_S) = \{\vec{k} \in \mathbb{R}^{r-1} : \langle \vec{l}, \vec{k} \rangle = 0, \ \forall \vec{l} \in \widetilde{S}\}. \tag{5.14}$$

Let
$$\ker(M_S)^\perp := \{\vec{l} \in \mathbb{R}^r : \langle \vec{l}, \vec{k} \rangle = 0,\ \forall \vec{k} \in \ker(M_S)\}.$$

Note that if $d = \dim(\ker(M_S))$, then $\dim(\ker(M_S)^\perp) = r - d$. To see this, write $\langle \vec{l}, \vec{k} \rangle = \vec{k} \cdot \phi(\vec{l})$, where $\cdot$ denotes the standard inner product on $\mathbb{R}^{r-1}$, and $\phi : \mathbb{R}^r \to \mathbb{R}^{r-1}$ is given by the *surjection* $\phi(\vec{l}) = (l_2 - l_1, l_3 - l_2, \ldots, l_r - l_{r-1})$. Thus, if S is maximal so that $\ker(M_S)^\star$ is non-empty (and $V^\star = \ker(M_S)^\star \in \mathcal{V}_r$) we find that for any $\vec{k} \in \ker(M_S)^\star$, using (5.13) and (5.14), that

$$2\alpha(\vec{k}) = \#\widetilde{S}; \tag{5.15}$$

in particular $\alpha(\vec{k})$ is constant for $\vec{k} \in V^\star$. Further, by Lemma 5.2, $\#\widetilde{S} \le 2^{r-d}$ since $\dim(\ker(M_S)^\perp) = r - d$, and the proof is concluded. $\square$

By Lemma 5.3 we have $2\alpha_{V^\star} \le 2^{r-d}$ and hence, by (5.3), (5.7), and (5.9), to establish Proposition 5.1 it suffices to show for each $V^\star \in \mathcal{V}_r$ that

$$\begin{aligned}\sum_{\vec{k} \in V^\star \cap \mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0) \left(g(\vec{k}; Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x} \right) \right) \\ = \sum_{\vec{k} \in V^\star \cap \mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0) + O\left(\frac{N^d}{\log\log x} + \frac{N^d}{M^{10}} \right).\end{aligned} \tag{5.16}$$

Using Proposition 4.4 we can express $g(\vec{k}; Y)$ and $\mathcal{L}(\vec{k})$ in terms of short Dirichlet polynomials of lengths $y = N^{o(1)}$ and since $V \cap \mathbb{Z}^{r-1}$ is a lattice this will allow us to use Poisson summation to establish (5.16).

We first require that the angles $(\theta_p)_{p \equiv 1 \pmod 4}$ are linearly independent $\pmod{2\pi}$ over $\mathbb{Q}$. We also need a quantitative bound for how close these combinations are to multiples of 2π, which is a consequence of repulsion of angles of Gaussian integers.

Lemma 5.4. *Let $J \ge 1$. For each $j = 1, \ldots, J$ suppose $p_j \equiv 1 \pmod 4$, and $p_i \ne p_j$ if $i \ne j$. Then for $c_1, \ldots, c_j \in \mathbb{Z}$*

$$\exp\left(i \sum_{j=1}^{J} c_j \theta_{p_j} \right) \ne 1 \tag{5.17}$$

unless $c_j = 0$ for each $j = 1, \ldots, J$. Moreover,

$$\left| \exp\left(i \sum_{j=1}^{J} c_j \theta_{p_j} \right) - 1 \right| \ge \frac{1}{\sqrt{p_1^{|c_1|} \cdots p_J^{|c_J|}}}. \tag{5.18}$$

Proof. Recall that $p_1, \ldots, p_J$ are distinct. We split the proof into two cases. First if $J = 1$ we cannot have that $e^{ic\theta_p} = 1$ since θ_p is not a rational multiple of π by Niven's theorem [34, Th'm 3.11].

Let $m = p_2^{|c_2|/2} \cdots p_J^{|c_J|/2}$. For $J \ge 2$ we observe that if equality in (5.17) holds then there exist $x + iy, u + iv \in \mathbb{Z}[i]$ such that

$$\frac{x + iy}{p_1^{|c_1|/2}} = \exp(ic_1\theta_{p_1}) = \exp\left(-i \sum_{j=2}^{J} c_j \theta_{p_j} \right) = \frac{u + iv}{m}$$

which implies that

$$x + iy = \frac{p_1^{|c_1|/2}}{m}(u + iv)$$

consequently c_j is even for each j so $m \in \mathbb{Z}$ and since $(m, p_1) = 1$, m divides both u and v. Also $m^2 = u^2 + v^2$ so that u or v equals zero. This implies that x or y equals zero but this is not possible since it would imply that $e^{ic\theta_{p_1}} = 1$ for some integer c. This proves the first claim.

Write $\theta = \sum_{j=1}^J c_j \theta_{p_j}$, and note that the above argument implies that $e^{i\theta} \neq \pm 1$. There exists $a + ib \in \mathbb{Z}[i]$ with $a^2 + b^2 = p_1^{|c_1|} \cdots p_J^{|c_J|}$ and $\frac{a+ib}{|a+ib|} = e^{i\theta}$. Hence,

$$0 < |1 - e^{i\theta}| = \left| \frac{e^{i\theta} - e^{-i\theta}}{1 + e^{i\theta}} \right| = \frac{1}{|1 + e^{i\theta}|} \cdot \frac{2|b|}{\sqrt{p_1^{|c_1|} \cdots p_J^{|c_J|}}}$$

so that $|b| \geq 1$ and consequently

$$|1 - e^{i\theta}| \geq \frac{1}{\sqrt{p_1^{|c_1|} \cdots p_J^{|c_J|}}}.$$

□

Before proceeding to the next lemma let us recall the bilinear form defined in (5.11). Given $\vec{l} = (l_1, \dots, l_r) \in \mathbb{Z}^r$ let

$$k_{\vec{l}} := \langle \vec{l}, \vec{k} \rangle = \sum_{j=0}^{r-1} l_{j+1}(k_{j+1} - k_j) \in \mathbb{Z}.$$

Recall that for $\vec{l}$ such that $l_{j+1} = 1$ if $j \in \mathbf{S}$ and $l_{j+1} = 0$ if $j \notin \mathbf{S}$ for some $[\mathbf{S}] \in \mathcal{M}_r$ we have $k_{\vec{l}} = k_{\mathbf{S}}$.

Lemma 5.5. *Let H be a Schwartz function and $w > v \geq 0$ be integers. Let $p_1, \dots, p_w$ be distinct primes $p_t \equiv 1 \pmod 4$, $t = 1, \dots, w$. Additionally, let $a_1, \dots, a_w$ be nonzero integers and $\vec{l}_1, \dots, \vec{l}_w \in \mathbb{Z}^{r-1}$ be such that $\|\vec{l}_t\|_\infty \leq |a_t|$ for each $t = 1, \dots, w$. Suppose $p_1^{|a_1|} \cdots p_w^{|a_w|} \leq N^\delta$ for some sufficiently small $\delta > 0$. Then we have that*

$$\sum_{\substack{\vec{k} \in V^\star \cap \mathbb{Z}^{r-1} \\ k_{\vec{l}_t} \neq 0, t = v+1, \dots, w}} \exp\left(4i \sum_{t=1}^w k_{\vec{l}_t} \theta_{p_t}\right) H\left(\frac{\vec{k}}{N}\right) \ll N^{d-1+o(1)}$$

where the implied constant depends at most on r and H, with $d = d(V^\star)$ as defined in (5.8).

Proof. Recall for each $V^\star \in \mathcal{V}_r$ that $V = \ker(M_S)$ for some maximal $S \subset \mathcal{M}_r$. Since a submodule M of a finitely generated $\mathbb{Z}$-module of rank $r-1$ is also free, and of rank $d \leq r-1$ (cf. [35, Proposition 9.7]), there exists a basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_d\}$ for $V \cap \mathbb{Z}^{r-1}$ such that

$$V \cap \mathbb{Z}^{r-1} = \left\{ \sum_{\ell=1}^d \mathbf{b}_\ell v_\ell : (v_1, \dots, v_d) \in \mathbb{Z}^d \right\}, \qquad \mathbf{b}_\ell = \begin{pmatrix} b_{\ell,1} \\ \vdots \\ b_{\ell,(r-1)} \end{pmatrix}, \tag{5.19}$$

and $b_{\ell,j} \in \mathbb{Z}$ for each $\ell = 1, \ldots, d$, $j = 1, \ldots, r-1$. Define $c_{\ell,\vec{l}} = \sum_{j=0}^{r-1} l_{j+1}(b_{\ell,(j+1)} - b_{\ell,j})$ where $l_1, \ldots, l_{r-1}$ denotes the components of $\vec{l}$, and where we use the convention $b_{\ell,0} = b_{\ell,r} = 0$; clearly $b_{\ell,j} = O_r(1)$, and thus $c_{\ell,\vec{l}} = O_r(\|\vec{l}\|_\infty)$ since the number of vector spaces V is $O_r(1)$ (cf. Lemma 7.4.) Using (5.19) we have, with k_j denoting the j-th coordinate of $\vec{k} \in V \cap \mathbb{Z}^{r-1}$, that $k_j = \sum_{\ell=1}^{d} b_{\ell,j} v_\ell$ for each $j = 1, \ldots, r-1$, and $k_{\vec{l}} = \sum_{\ell=1}^{d} v_\ell c_{\ell,\vec{l}}$ for each $\vec{l} \in \mathbb{Z}^r$. Let

$$G(x_1, \ldots, x_d) = H\left(\sum_{\ell=1}^{d} b_{\ell,1} x_\ell, \ldots, \sum_{\ell=1}^{d} b_{\ell,r-1} x_\ell\right).$$

Observe that $G : \mathbb{R}^d \to \mathbb{R}$ is a Schwartz function. Also, $\sum_{\vec{k} \in (V \setminus V^\star) \cap \mathbb{Z}^{r-1}} |H(\vec{k}/N)| \ll N^{d-1+o(1)}$ as $V \setminus V^\star$ consists of lower dimensional subspaces. Hence, we have that

(5.20)

$$\sum_{\substack{\vec{k} \in V^\star \cap \mathbb{Z}^{r-1} \\ k_{\vec{l}_t} \neq 0 \\ t = v+1, \ldots, w}} \exp\left(4i \sum_{t=1}^{w} k_{\vec{l}_t} \theta_{p_t}\right) H\left(\frac{\vec{k}}{N}\right) = \sum_{\substack{\vec{k} \in V \cap \mathbb{Z}^{r-1} \\ k_{\vec{l}_t} \neq 0 \\ t = v+1, \ldots, w}} \exp\left(4i \sum_{t=1}^{w} k_{\vec{l}_t} \theta_{p_t}\right) H\left(\frac{\vec{k}}{N}\right) + O(N^{d-1+o(1)})$$

$$= \sum_{\substack{\vec{v} \in \mathbb{Z}^d \\ \sum_{\ell=1}^{d} v_\ell c_{\ell,\vec{l}_t} \neq 0 \\ t = v+1, \ldots, w}} \prod_{\ell=1}^{d} \left(\exp\left(4i v_\ell \sum_{t=1}^{w} c_{\ell,\vec{l}_t} \theta_{p_t}\right)\right) G\left(\frac{\vec{v}}{N}\right) + O(N^{d-1+o(1)}).$$

Since $\sum_{\ell=1}^{d} v_\ell c_{\ell,\vec{l}_t} \neq 0$, WLOG assume $c_{1,\vec{l}_w} \neq 0$. Write E for the subset of $\mathbb{Z}^d$ such that $\sum_{\ell=1}^{d} v_\ell c_{\ell,\vec{l}_t} \neq 0$ for each $t = v+1, \ldots, w$. The left-hand side of (5.20) is

$$(5.21) \quad \ll \sum_{|v_2|, \ldots, |v_d| \ll N^{1+o(1)}} \left| \sum_{v_1 \in \mathbb{Z}} 1_{\vec{v} \in E} G\left(\frac{v_1}{N}, \frac{v_2}{N}, \ldots, \frac{v_d}{N}\right) \exp\left(4i v_1 \sum_{t=1}^{w} c_{\ell,\vec{l}_t} \theta_{p_t}\right) \right| + N^{d-1+o(1)}.$$

We wish to extend the inner sum to all of $\mathbb{Z}$. To do this, for each $t = 1, \ldots, w$ with $c_{1,\vec{l}_t} \neq 0$ we need to add back in the point v_1 such that $c_{1,\vec{l}_t} v_1 = -\sum_{\ell=2}^{d} v_\ell c_{\ell,\vec{l}_t}$. Since there are $\leq w$ such points and $w \leq N^{o(1)}$, as by assumption $p_1^{|a_1|} \cdots p_w^{|a_w|} \leq N^\delta$ which implies $2^w \leq N^\delta$, we can extend the inner sum to all of $\mathbb{Z}$ at the cost of an error term of size

$$(5.22) \qquad \ll N^{o(1)} \sum_{|v_2|, \ldots, |v_d| \ll N^{1+o(1)}} 1 \ll N^{d-1+o(1)}.$$

Let $\theta = 4 \sum_{t=1}^{w} c_{\ell,\vec{l}_t} \theta_{p_t}$ and WLOG we may assume $-\pi \leq \theta < \pi$ (since $v_1 \in \mathbb{Z}$.) Applying Poisson summation and using (5.17) (which implies $e^{i\theta} \neq 1$) repeatedly integrating by parts gives that

$$\sum_{v_1\in\mathbb{Z}} G\left(\frac{v_1}{N},\frac{v_2}{N},\ldots,\frac{v_d}{N}\right)e^{i\theta v_1} = N\sum_{a\in\mathbb{Z}}\int_{\mathbb{R}} G\left(x,\frac{v_2}{N},\ldots,\frac{v_d}{N}\right)e^{iN(\theta-2\pi a)x}\,dx$$
$$= N\sum_{a\in\mathbb{Z}}\left(\frac{-1}{iN(\theta-2\pi a)}\right)^B\int_{\mathbb{R}}(\partial_1^B G)\left(x,\frac{v_2}{N},\ldots,\frac{v_d}{N}\right)e^{iN(\theta-2\pi a)x}\,dx$$
$$\ll N\sum_{a\in\mathbb{Z}}\frac{1}{N^B|\theta-2\pi a|^B}$$

for every integer $B\geq 0$, where $(\partial_1^j G)(x_1,\ldots,x_d)=\frac{\partial^j}{\partial x_1^j}G(x_1,\ldots,x_d)$ (we have also used that $\partial_1^j G$ is a Schwartz function so that $(\partial_1^B G)(x_1,\ldots,x_d)\ll\frac{1}{1+|x_1|^A}\cdots\frac{1}{1+|x_d|^A}\ll\frac{1}{1+|x_1|^A}$, for $(x_1,\ldots,x_d)\in\mathbb{R}^d$ where the implied constant depends at most on B,G.) Since $c_{\ell,\vec{l_t}}\ll|a_t|$, we have by (5.18) and the assumption $p_1^{|a_1|}\cdots p_w^{|a_w|}\leq N^\delta$ that $|\theta|\gg N^{-1/2}$ since δ is sufficiently small. Hence, since $-\pi\leq\theta<\pi$ this implies that the right-hand side above is

$$\ll N^{1-B}\frac{1}{|\theta|^B}\ll N^{-100}.$$

Combining this estimate with (5.21) and (5.22) completes the proof. □

Lemma 5.6. *Let $n_0\in\mathcal{S}$. For each $V^\star\in\mathcal{V}_r$ with $d=\dim(V)$ we have that*

$$\sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}}\mathcal{L}(\vec{k})\left|\widehat{f}\left(\frac{\vec{k}}{N}\right)\right|\ll N^d,$$

with $\mathcal{L}(\vec{k})$ as given in (4.8)*. The implied constant depends on at most A',n_0, and f.*

Proof. Observe that there exists a Schwartz function H with $|\widehat{f}|\leq H$ (e.g. $H(x)=(|\widehat{f}(x)|^2+e^{-x^2})^{1/2}$.) Recalling (4.8) it suffices to show that

$$\sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}\\ k_{\mathbf{S}}\neq 0}} L(1,8k_{\mathbf{S}})^u H\left(\frac{\vec{k}}{N}\right)\ll N^d \tag{5.23}$$

for any $[\mathbf{S}]\in\mathcal{M}_r$ and $u\in\mathbb{Z}$ fixed, $u\neq 0$ (note that $L(1,4k)>0$ for every $k\in\mathbb{Z}$, $k\neq 0$.) Since H is a Schwartz function, we see that by (4.21) the sum is effectively restricted to $\|\vec{k}\|_\infty\leq N^{1+o(1)}$. Applying part 2 of Proposition 4.4 with $y=e^{(\log N)^{3/4}}$ we have for $|k_{\mathbf{S}}|\leq N^{1+o(1)}$ that

$$L(1,8k_{\mathbf{S}})^u=\sum_{m\leq y}\frac{\lambda_{8k_{\mathbf{S}}}(m;u)}{m}+O\left(e^{-(\log N)^{1/12-o(1)}}\right). \tag{5.24}$$

Applying (5.24) we see that the left-hand side of (5.23) equals

$$\sum_{m\leq y}\sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}\\ k_{\mathbf{S}}\neq 0}}\frac{\lambda_{8k_{\mathbf{S}}}(m;u)}{m}H\left(\frac{\vec{k}}{N}\right)+O\left(N^d e^{-(\log N)^{-1/12+o(1)}}\right), \tag{5.25}$$

where we have used that

$$\sum_{\vec{k}\in V\cap\mathbb{Z}^{r-1}} H\left(\frac{\vec{k}}{N}\right) \ll N^d \tag{5.26}$$

to estimate the error term. To see why (5.26) holds, note that $V\cap\mathbb{Z}^{r-1}$ is a lattice (of rank d), so by applying Poisson summation the bound follows. Using the definition of $\lambda_{8k}(m;u)$ as given in (4.13) we have for each $m\le y$ that

$$\sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}\\ k_{\mathbf{S}}\neq 0}} \lambda_{8k_{\mathbf{S}}}(m;u)H\left(\frac{\vec{k}}{N}\right) = \sum_{\mathfrak{b}\subset\mathcal{O}:|\mathcal{O}/\mathfrak{b}|=m} b_u(\mathfrak{b}) \sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}\\ k_{\mathbf{S}}\neq 0}} \Xi_{8k_{\mathbf{S}}}(\mathfrak{b})H\left(\frac{\vec{k}}{N}\right).$$

Let us first note that we may assume m is a sum of two squares, since otherwise $\lambda_{8k_{\mathbf{S}}}(m;u)=0$, which is clear from the definition (4.13). We first consider m for which there is cancellation in the $\vec{k}$-sum, namely $m\le y$ which are sums of two squares that are not of the form l^2 or $2l^2$; note that such m must be divisible by at least one prime $p\equiv 1 \pmod 4$. For each $\mathfrak{b}\subset\mathcal{O}$ with $|\mathcal{O}/\mathfrak{b}|=m$ by considering the factorization of $\mathfrak{b}$ into prime ideals, we see that in this case there exist distinct primes $p_1,\dots,p_w$ each $\equiv 1 \pmod 4$ and non-zero integers $a_1,\dots,a_w$ such that $\Xi_{8k}(\mathfrak{b})=e^{i8k\sum_{j=1}^{w}a_j\theta_{p_j}}$ and $p_1^{|a_1|}\cdots p_w^{|a_w|}\le y$ (note that $\Xi_{8k}((1+i))=1$.) Hence using these observations and recalling $b_u(\mathfrak{b})\ll|\mathcal{O}/\mathfrak{b}|^{o(1)}$ (cf. (4.14)), applying Lemma 5.5 with $v=0$ and $\vec{l}_j$ such that $k_{\vec{l}_j}=2a_jk_{\mathbf{S}}$ for each $j=1,\dots,w$, gives for $m\le y$, which is not of the form $m=l^2,2l^2$, that

$$\left|\sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}\\ k_{\mathbf{S}}\neq 0}} \lambda_{8k_{\mathbf{S}}}(m;u)H\left(\frac{\vec{k}}{N}\right)\right| \ll N^{d-1+o(1)}. \tag{5.27}$$

To bound the contribution of the terms with $m=l^2,2l^2$ we use that $\lambda_{8k_{\mathbf{S}}}(m;u)\ll m^{o(1)}$ (this follows from (4.13) and (4.14)) along with (5.26). Combining this with (5.27) we conclude that the main term in the left-hand side of (5.25) is

$$\ll N^d\sum_{l^2\le y}\frac{1}{l^{2-o(1)}}+N^{d-1+o(1)}\ll N^d,$$

which completes the proof. □

Lemma 5.7. *Let $n_0\in\mathcal{S}$. Also, let $V^\star\in\mathcal{V}_R$ and m be a square free integer such that $(m,2n_0)=1$ and $2\le m\le N^\delta$, where δ is sufficiently small. We have that*

$$\sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}} \ell_{\vec{k}}(n_0)s(m;\vec{k})\,\widehat{f}\left(\frac{\vec{k}}{N}\right)\ll N^{d-1+o(1)}, \tag{5.28}$$

where $\ell_{\vec{k}}(n_0)$ is as defined in (5.1) and $s(m;\vec{k})$ is as in (4.15). The implied constant depends at most on r,n_0 and f.

Proof. Let us write $n_0=2^aq_1^{2b_1}\cdots q_u^{2b_u}p_1^{a_1}\cdots p_v^{a_v}$ where $a,b_1,\dots,b_u,a_1,\dots,a_v$ are non-negative integers, q_j are primes $\equiv 3 \pmod 4$, $j=1,\dots,u$, and p_j are distinct primes $\equiv 1 \pmod 4$,

$j=1,\dots,v$. Recalling (4.4) and the fact that $\sum_{j=0}^{r-1}(k_{j+1}-k_j)=0$ we have

(5.29)
$$\begin{aligned}\ell_{\vec{k}}(n_0) &= \prod_{t=1}^{v}\prod_{j=0}^{r-1}\lambda_{4(k_{j+1}-k_j)}(p_t^{a_t}) = \prod_{t=1}^{v}\prod_{j=0}^{r-1}\sum_{0\le l\le a_t} e^{8l(k_{j+1}-k_j)i\theta_{p_t}} \\ &= \prod_{t=1}^{v}\sum_{0\le l_1,\dots,l_r\le a_t}\exp\left(8i\sum_{j=0}^{r-1}l_{j+1}(k_{j+1}-k_j)\theta_{p_t}\right) = \sum_{\substack{\vec{l}_1,\dots,\vec{l}_v\in\mathbb{Z}^r_{\ge 0}\\ \|\vec{l}_t\|_\infty\le a_t, t=1,\dots,v}}\exp\left(8i\sum_{t=1}^{v}k_{\vec{l}_t}\theta_{p_t}\right).\end{aligned}$$

Also, it is clear that $\ell_{\vec{k}}(n_0)\ll 1$, where the implied constant depends at most on n_0, r, since $|\lambda_{4k}(n)|\le\sum_{d|n}1$.

We can write $m=p_{v+1}\cdots p_w$ where $p_{v+1},\dots,p_w$ are distinct primes $\equiv 1 \pmod 4$, which are co-prime to n_0. Let $W=\{1,\dots,w-v\}$ and for $S\subset W$ let $\epsilon_{j,S}=1$ if $j\in S$ and $\epsilon_{j,S}=-1$ if $j\notin S$. Recalling that $\lambda_{8k}(p)=e^{8ki\theta_p}+e^{-8ki\theta_p}$ for $p\equiv 1 \pmod 4$ we have that

(5.30)
$$\begin{aligned}s(m;\vec{k}) &= \sum_{\substack{[\mathbf{S}_{v+1}],\dots,[\mathbf{S}_w]\in\mathcal{M}_r\\ k_{\mathbf{S}_j}\ne 0,\, j=v+1,\dots,w}}\lambda_{8k_{\mathbf{S}_{v+1}}}(p_{v+1})\cdots\lambda_{8k_{\mathbf{S}_w}}(p_w) \\ &= \sum_{\substack{[\mathbf{S}_{v+1}],\dots,[\mathbf{S}_w]\in\mathcal{M}_r\\ k_{\mathbf{S}_j}\ne 0,\, j=v+1,\dots,w}}\sum_{S\subset W}\exp\left(8i\sum_{j=v+1}^{w}\epsilon_{j,S}k_{\mathbf{S}_j}\theta_{p_j}\right).\end{aligned}$$

To complete the proof we combine (5.29) and (5.30) then use the resulting expression in the left-hand side of (5.28) and apply Lemma 5.5, with $k_{\vec{l}_t}=2\epsilon_{j,S}k_{\mathbf{S}_t}$ for each $t=v+1,\dots,w$. The hypotheses of the lemma are satisfied since $n_0 m\ll N^\delta$. Finally note that $2^{|W|}=N^{o(1)}$ since $m\le N^\delta$. □

5.3. **Proof of Proposition 5.1.**

Proof of Proposition 5.1. Let $V^\star\in\mathcal{V}_r$ with $d=\dim(V)$. Recall $Y=(M-1)/\log\log x$ and by assumption $Y\le A$. Applying Lemma 5.6 and the bound $\ell_{\vec{k}}(n_0)\ll 1$ (cf. (5.29)) we have that

(5.31)
$$\begin{aligned}\sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}}&\ell_{\vec{k}}(n_0)\widehat{f}\left(\frac{\vec{k}}{N}\right)\left(g(\vec{k};Y)+O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right) \\ &= \sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}}\ell_{\vec{k}}(n_0)g(\vec{k};Y)\widehat{f}\left(\frac{\vec{k}}{N}\right)+O\left(\frac{N^d}{\log\log x}\right).\end{aligned}$$

Since $\widehat{f}$ is a Schwartz function the sum is effectively restricted to $\|\vec{k}\|_\infty\le N^{1+o(1)}$. Recall the definitions of $w(m;Y), s(m;\vec{k})$ as given in (4.15). By Proposition 4.4 with $y=e^{(\log N)^{3/4}}$

and (5.9) we have that

(5.32)

$$\sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}} \ell_{\vec{k}}(n_0)g(\vec{k};Y)\widehat{f}\left(\frac{\vec{k}}{N}\right) = \sum_{\substack{m\le y\\(m,2n_0)=1}} \frac{\alpha_{V^\star}^{-\Omega_1(m)}w(m;Y)}{m} \sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}} \ell_{\vec{k}}(n_0)s(m;\vec{k})\widehat{f}\left(\frac{\vec{k}}{N}\right) + O\left(N^d e^{-(\log N)^{1/12-o(1)}}\right),$$

where we have used (5.26) and that $\ell_{\vec{k}}(n_0)\ll 1$ to bound the error term (recall H is Schwartz function with $|\widehat{f}|\le H$.) Using Lemma 5.7 and noting that $w(m;Y)\ll m^{o(1)}$ the contribution of the terms with $2\le m\le y$ is $\ll yN^{d-1+o(1)}\ll N^{d-1+o(1)}$. Hence, applying this observation in (5.32) we see that the sum on the right-hand side of (5.31) equals

$$\sum_{\vec{k}\in V^\star\cap\mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right)\ell_{\vec{k}}(n_0) + O\left(N^d e^{-(\log N)^{1/12-o(1)}}\right).$$

(In particular the main term is given by $m=1$, for which $\Omega_1(m)=0$, $w(m;Y)=1$, and $s(m,k)=1$.) This along with (5.31) establishes (5.16), which, on taking (5.3) into account, completes the proof of Proposition 5.1. $\square$

5.4. **The variance.** Let $h:\mathbb{R}^{2r-1}\to\mathbb{R}$ be a Schwartz function with

$$h(x_1,\dots,x_{r-1},0,x_{r+1},\dots,x_{2r-1}) = \widehat{f}(x_1,\dots,x_{r-1})\widehat{f}(x_{r+1},\dots,x_{2r-1}). \tag{5.33}$$

Additionally, for $\vec{k}\in\mathbb{Z}^{2r-1}$ with $\vec{k}=(\vec{k}_1,0,\vec{k}_2)$ and $\vec{k}_1,\vec{k}_2\in\mathbb{Z}^{r-1}$ let

$$\widetilde{\ell}_{\vec{k}}(n_0) = \ell_{\vec{k}_1}(n_0)\ell_{\vec{k}_2}(n_0). \tag{5.34}$$

Proposition 5.8. *Let $A>0$ and $n_0\in\mathcal{S}$. Suppose that $1\le M\le A\log\log x$. We have that*

(5.35)

$$\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n\in\mathcal{N}_{M,n_0}(x)} R_{r,n_0}(F_N;n)^2 = \frac{1}{N^{2r}} \sum_{\substack{\vec{k}\in\mathbb{Z}^{2r-1}\\k_r=0}} (2\alpha(\vec{k}))^M \widetilde{\ell}_{\vec{k}}(n_0) h\left(\frac{\vec{k}}{N}\right) + O\left(\frac{1}{\log\log x}+\frac{1}{M^{10}}\right),$$

where the implied constant depends at most on f,r,n_0 and A.

Proof. Using an analogue of (1.3) (see also (5.1) and the first line of (5.3)) the left-hand side of (5.35) equals

$$\begin{aligned}
=&\frac{1}{N^{2r}} \sum_{\vec{k},\vec{l}\in\mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right)\widehat{f}\left(\frac{\vec{l}}{N}\right)\ell_{\vec{k}}(n_0)\ell_{\vec{l}}(n_0)\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n\in\mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}(n)\lambda_{4(l_{j+1}-l_j)}(n)\\
=&\frac{1}{N^{2r}} \sum_{\substack{\vec{k}\in\mathbb{Z}^{2r-1}\\k_r=0}} h\left(\frac{\vec{k}}{N}\right)\widetilde{\ell}_{\vec{k}}(n_0)\frac{1}{\#\mathcal{N}_{M,n_0}(x)} \sum_{n\in\mathcal{N}_{M,n_0}(x)} \prod_{j=0}^{2r-1} \lambda_{4(k_{j+1}-k_j)}(n),
\end{aligned}$$

where we have also used the convention $k_0=k_{2r}=0$ in the last line. As before, recall that the contribution of the terms with $\|\vec{k}\|_\infty\ge N^{1+o(1)}$ is $\ll N^{-B}$ for every $B>0$, which is negligible so we can add or remove these terms as we wish.

We now use an argument similar to that given in the proof of Proposition 5.1. As before we apply Proposition 4.1 to the inner sum on the right-hand side above (using the previous observation to restrict to $\|\vec{k}\|_\infty \le N^{1+o(1)}$.) The next step differs slightly; we apply our decomposition (5.7) to $\mathbb{R}^{2r-1}$ (as opposed to $\mathbb{R}^{r-1}$ previously) and we have the additional constraint $k_r = 0$. This gives that the right-hand side of the above equation equals

(5.36)

$$\frac{1}{N^{2r}} \sum_{\substack{\vec{k}\in\mathbb{Z}^{2r-1}\\ k_r=0}} h\left(\frac{\vec{k}}{N}\right)\widetilde{\ell}_{\vec{k}}(n_0)(2\alpha(\vec{k}))^M \left(g(\vec{k};Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right) + O(N^{-10})$$

$$= \frac{1}{N^{2r}} \sum_{V^\star\in\mathcal{V}_{2r}} \sum_{\substack{\vec{k}\in V^\star\cap\mathbb{Z}^{2r-1}\\ k_r=0}} h\left(\frac{\vec{k}}{N}\right)\widetilde{\ell}_{\vec{k}}(n_0)(2\alpha(\vec{k}))^M \left(g(\vec{k};Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right) + O(N^{-10}),$$

where in the first line we added back in the terms with $\|\vec{k}\|_\infty \ge N^{1+o(1)}$ after applying Proposition 4.1. Similar to before, our strategy is to evaluate the inner sum for each $V^\star \in \mathcal{V}_{2r}$, however we need to account for the condition $k_r = 0$ (and we are also working with higher dimensional lattices.) We next note that the condition $k_r = 0$ can be imposed by adding an extra linear relation to S. Namely, let $\mu = [\{0,\ldots,r-1\}] \in \mathcal{M}_{2r}$ and observe that $k_\mu = \sum_{j=0}^{r-1}(k_{j+1}-k_j) = k_r$. For $S \subset \mathcal{M}_{2r}$ let $\widetilde{S} = S \cup \{\mu\}$. Recall that for each $V^\star \in \mathcal{V}_{2r}$ there exists $S \subset \mathcal{M}_{2r}$ such that $V^\star = \ker(M_S)^\star$. Set $\widetilde{V}^\star = \ker(M_{\widetilde{S}})^\star$, $d = \dim(\widetilde{V})$, and note that $d \le 2r-1$. Hence the right-hand side of (5.36) equals

$$\frac{1}{N^{2r}} \sum_{V^\star\in\mathcal{V}_{2r}} (2\alpha_{\widetilde{V}^\star})^M \sum_{\vec{k}\in\widetilde{V}^\star\cap\mathbb{Z}^{2r-1}} \widetilde{\ell}_{\vec{k}}(n_0) h\left(\frac{\vec{k}}{N}\right) \left(g(\vec{k};Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right) + O(N^{-10}).$$

Since $\widetilde{V}^\star \in \mathcal{V}_{2r}$ arguing as in the proof of Proposition 5.1 gives that

$$\sum_{\vec{k}\in\widetilde{V}^\star\cap\mathbb{Z}^{2r-1}} \widetilde{\ell}_{\vec{k}}(n_0) h\left(\frac{\vec{k}}{N}\right)\left(g(\vec{k};Y) + O\left(\frac{\mathcal{L}(\vec{k})}{\log\log x}\right)\right)$$

$$= \sum_{\vec{k}\in\widetilde{V}^\star\cap\mathbb{Z}^{2r-1}} \widetilde{\ell}_{\vec{k}}(n_0) h\left(\frac{\vec{k}}{N}\right) + O\left(\frac{N^d}{\log\log x} + \frac{N^d}{M^{10}}\right),$$

which establishes an analogue of (5.16). To obtain the above estimate we have also used an analogue of Lemma 5.7 where $\widetilde{\ell}_{\vec{k}}(n_0)$ replaces $\ell_{\vec{k}}(n_0)$ in the left-hand side of (5.28). Since this result follows from a completely analogous argument to the one used to establish Lemma 5.7 we will omit the details. Recalling Lemma 5.3, reversing our decomposition completes the proof. □

6. Proof of Theorem 3.1: Matching the random model with the r-correlation

The goal of this section is to express our formulas for the r-correlation of lattice points in terms of the random model from Section 3 and complete the proof of Theorem 3.1. Our approach is to compute the smoothed r-correlation of the random model $\mathcal{R}_r(F_N)$ as given in (3.2) by following a similar strategy to the one used to compute the r-correlation for the angles of lattice points. Using independence of the random variables $(\vartheta_l)_{l=1}^M$ we will quickly arrive at the same expression that appears in Proposition 5.1.

Proof of Theorem 3.1. Recall the definition of x_J as given in (3.1). Let us define the random variable

$$\lambda_{4k} = \sum_{J \subset \{1,\dots,M\}} e^{4kix_J} = \prod_{l=1}^{M} (1 + e^{4ki\vartheta_l}). \tag{6.1}$$

We will first establish (3.3). Using (1.2), and with $\ell_{\vec{k}}(n_0)$ as in (5.1), we have that

$$\mathcal{R}_{r,n_0}(F_N) = \frac{1}{N^r} \sum_{\vec{k} \in \mathbb{Z}^{r-1}} \ell_{\vec{k}}(n_0) \widehat{f}\left(\frac{\vec{k}}{N}\right) \prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}, \tag{6.2}$$

where we recall that by convention $k_0 = k_r = 0$. Using (6.1) as well as that $(\vartheta_l)_{l=1}^M$ are independent and arguing as in (4.5) we have

$$\begin{aligned} \mathbb{E}\left(\prod_{j=0}^{r-1} \lambda_{4(k_{j+1}-k_j)}\right) &= \prod_{l=1}^{M} \mathbb{E}\left(\prod_{j=0}^{r-1} (1 + e^{i4(k_{j+1}-k_j)\vartheta_l}\right) \\ &= \prod_{l=1}^{M} \sum_{\mathbf{S} \subset \{0,1,\dots,r-1\}} \mathbb{E}\big(e^{i4k_{\mathbf{S}}\vartheta_l}\big) = (2\alpha(\vec{k}))^M. \end{aligned} \tag{6.3}$$

Hence, combining this with (6.2) gives

$$\mathbb{E}\big(\mathcal{R}_{r,n_0}(F_N)\big) = \frac{1}{N^r} \sum_{\vec{k} \in \mathbb{Z}^{r-1}} \widehat{f}\left(\frac{\vec{k}}{N}\right) \ell_{\vec{k}}(n_0) (2\alpha(\vec{k}))^M.$$

Using this along with (5.2) establishes (3.3).

It remains to prove (3.4). The argument proceeds similarly. Applying (1.2) and arguing as in the first step of the proof of Proposition 5.8 yields

$$\mathcal{R}_{r,n_0}(F_N)^2 = \frac{1}{N^{2r}} \sum_{\substack{\vec{k} \in \mathbb{Z}^{2r-1} \\ k_r = 0}} h\left(\frac{\vec{k}}{N}\right) \widetilde{\ell}_{\vec{k}}(n_0) \prod_{j=0}^{2r-1} \lambda_{4(k_{j+1}-k_j)}, \tag{6.4}$$

where h is as in (5.33), $\widetilde{\ell}_{\vec{k}}(n_0)$ is as in (5.34), and we use the convention $k_0 = k_{2r} = 0$. Hence, it follows from (6.3) (which we use with $2r$ in place of r) that

$$\mathbb{E}\big(\mathcal{R}_{r,n_0}(F_N)^2\big) = \frac{1}{N^{2r}} \sum_{\substack{\vec{k} \in \mathbb{Z}^{2r-1} \\ k_r = 0}} h\left(\frac{\vec{k}}{N}\right) \widetilde{\ell}_{\vec{k}}(n_0) (2\alpha(\vec{k}))^M.$$

Therefore, using this together with (5.35) establishes (3.4), which completes the proof. $\square$

7. Poisson correlations for the random model

We first treat the square free case separately as it is notationally simpler, and then use it to deduce the case for general n.

7.1. **The square free case.** We begin with the case of square free n, i.e., $n_0 \in \{1,2\}$. For simpler notation, let $N = N_n = r(n)$. In the random model we can directly handle summing over distinct angles, and will work with the following setup for the "standard" r-level correlation, which we define as the random variable

$$\mathcal{R}_r^*(F_N) := \frac{1}{N} \sideset{}{^*}\sum_{J_1,\dots,J_r \subset \{1,\dots,M\}} F_N(x_{J_1} - x_{J_2}, \dots, x_{J_{r-1}} - x_{J_r})$$

where $\sum^*$ indicates summing over *distinct* subsets $J_1, \dots, J_r$. Throughout this section we assume f has compact support.

7.1.1. *The pair correlation.* To illustrate ideas we begin by determining the pair correlation. Our approach is to compute the expected value, and then, via a variance bound, show that fluctuations around the mean are small.

The expectation of $\mathcal{R}_2^*(F_N)$*:* Using linearity of expectations, the expected value of $\mathcal{R}_2^*(F_N)$ is given by

$$\mathbb{E}\left(\mathcal{R}_2^*(F_N)\right) = \frac{1}{N} \sideset{}{^*}\sum_{J_1,J_2 \subset \{1,\dots,M\}} \mathbb{E}\left(F_N(x_{J_1} - x_{J_2})\right).$$

Since $\sum^*$ indicates summing over distinct subsets J_1, J_2, the symmetric difference $J_1 \Delta J_2$ is nontrivial and $x_{J_1} - x_{J_2}$ is a sum (with certain choices of signs) of $|J_1 \triangle J_2| \geq 1$ *independent* uniform random variables on the torus (here and in what follows $J_1 \triangle J_2$ denotes the symmetric difference between the sets J_1 and J_2.) In particular, for distinct subsets J_1, J_2 we have by a direct computation that $\mathbb{E}\left(F_N(x_{J_1} - x_{J_2})\right) = \hat{f}(0)/N$, and the total contribution equals

$$\frac{1}{N} \cdot N(N-1)\frac{\hat{f}(0)}{N} = \hat{f}(0)(1 - 1/N) = \hat{f}(0) + O_f(1/N).$$

Bounding the variance of $\mathcal{R}_2^*(F_N)$*:* We next show that the fluctuations around the mean, with large probability, are small in comparison with the mean, namely that

$$\mathbb{E}(\mathcal{R}_2^*(F_N)^2) - \mathbb{E}(\mathcal{R}_2^*(F_N))^2 = o_f(1)$$

as M tends to infinity. By linearity of expectations, we find that

$$\mathbb{E}(\mathcal{R}_2^*(F_N)^2) = \frac{1}{N^2} \sideset{}{^*}\sum_{J_1,J_2 \subset \{1,\dots,M\}} \sideset{}{^*}\sum_{J_3,J_4 \subset \{1,\dots,M\}} \mathbb{E}\left(F_N(x_{J_1} - x_{J_2}) F_N(x_{J_3} - x_{J_4})\right). \tag{7.1}$$

Now, for a full density subset of choices of two pairs of subsets (i.e., for $N^4(1+o(1))$ choices), the two components of $(x_{J_1} - x_{J_2}, x_{J_3} - x_{J_4})$ will contain at least one pair of independent random variables ("full rank"), and if this is so we have

$$\mathbb{E}\left(F_N(x_{J_1} - x_{J_2}) F_N(x_{J_3} - x_{J_4})\right) = \frac{\hat{f}(0)^2}{N^2}$$

and hence the main term of (7.1) equals

$$\frac{N^4(1+o(1))}{N^2} \cdot \frac{\hat{f}(0)^2}{N^2} = \hat{f}(0)^2 + o_f(1).$$

Remark 1. *A delicate issue is that if we would include pairs of non-distinct subsets — we then get "singular pairings" which give a contribution of the same size as the main term. Namely, if we take $J_1 = J_2$ and $J_3 = J_4$ (there are N^2 such choices), we find, on noting that $\mathbb{E}(F_N(x_{J_1} - x_{J_2}) \cdot F_N(x_{J_3} - x_{J_4})) = f(0)^2$, that the contribution from these terms equals*

$$\frac{1}{N^2} \cdot N^2 f(0)^2 = f(0)^2.$$

In fact, the same holds for the expectation! However, this should not be a surprise as allowing for pairs of points to be equal should give a secondary main term of the form $f(0)$ in addition to $\widehat{f}(0)$.

To bound the number of "low rank" tuples we argue as follows: each choice of subsets J_1, J_2, J_3, J_4 gives a group homomorphism

$$\mathbb{T}^M \to \mathbb{T} \times \mathbb{T}, \quad (\vartheta_1, \dots, \vartheta_M) \to (\sum_{j \in J_1} \vartheta_j - \sum_{j \in J_2} \vartheta_j, \sum_{j \in J_3} \vartheta_j - \sum_{j \in J_4} \vartheta_j),$$

which, on letting $\vec{\vartheta} = (\vartheta_1, \dots, \vartheta_M)$, can be written as

$$\vec{\vartheta} \to ((\vec{v}_1 - \vec{v}_2) \cdot \vec{\vartheta}, (\vec{v}_3 - \vec{v}_4) \cdot \vec{\vartheta})$$

where $\vec{v}_l = \sum_{i \in J_l} \vec{e}_i$ and $\vec{e}_1, \dots, \vec{e}_M$ denotes the standard basis for $\mathbb{R}^M$. (Note that the torus map is given by an integer entry matrix in $\mathrm{Mat}_{2,M}(\mathbb{Z})$ whose rows are given by $(\vec{v}_1 - \vec{v}_2)^t, (\vec{v}_3 - \vec{v}_4)^t$, and if this matrix has rank two the distribution of the image inside $\mathbb{T}^2$ is uniform; for a formal argument see Lemma 7.2.) We note that if the 4-tuple of subsets (J_1, J_2, J_3, J_4) is "generic", then the rank of this map is 2; when this does not hold we call the 4-tuple of subsets (J_1, J_2, J_3, J_4) "low rank". If the rank is zero, we must have $\vec{v}_1 = \vec{v}_2$ and $\vec{v}_3 = \vec{v}_4$, hence $J_1 = J_2$ and $J_3 = J_4$ and thus there is no contribution, and similarly there is no contribution if the rank is one due to either $\vec{v}_1 - \vec{v}_2 = 0$ or $\vec{v}_3 - \vec{v}_4 = 0$. The remaining rank one case is that there exists nonzero scalars α, β such that

$$\alpha(\vec{v}_1 - \vec{v}_2) = \beta(\vec{v}_3 - \vec{v}_4)$$

where, we may without loss of generality assume that α, β are coprime integers, and say $\alpha > 0$. We note that $\alpha > 1$ is impossible since the components of the vector $\alpha(\vec{v}_1 - \vec{v}_2)$ are in $\{-\alpha, 0, \alpha\}$, whereas the components of $\beta(\vec{v}_3 - \vec{v}_4)$ are in $\{-\beta, 0, \beta\}$ (note that at least one component in each of the two vector differences must be nonzero.) The remaining case is that $\alpha = 1$ and $\beta = \pm 1$; say $\beta = 1$ (the other case follows similarly.) In this case, for $\vec{v}_1, \vec{v}_2$ fixed, we find that $J_3 \setminus J_4 = J_1 \setminus J_2$, as well as $J_4 \setminus J_3 = J_2 \setminus J_1$, and the only choice left is specifying the intersection $J_3 \cap J_4$, which clearly must be contained in the *complement* of the symmetric difference $J_1 \triangle J_2$. We next show that the cardinality of the symmetric difference, for "generic" choices of J_1, J_2, is $M \cdot (1/2 + o(1))$.

Lemma 7.1. *For $N^2(1 + o(1))$ choices of subsets $J_1, J_2 \subset \{1, \dots, M\}$ we have $|J_1 \triangle J_2| = M \cdot (1/2 + o(1))$, as $M \to \infty$.*

Proof. We use the following simple probabilistic argument: first note that the number of pairs of subsets with the desired property is N^2 times the probability of randomly selected subsets J_1, J_2 having the same property, where the two subsets are selected independently and the probability for each configuration is $1/2^M$. Or equivalently, each element $i \in \{1, \dots, M\}$ is independently selected to be in J_1 with probability $1/2$, and similarly for $i \in J_2$. In particular, for a fixed index i, each of the four possible containment patterns w.r.t. J_1, J_2 occurs with probability $1/4$. Hence an index $i \in \{1, \ldots, M\}$ is contained in $J_1 \triangle J_2$ with probability $1/2$ (as this occurs for two out of the four possible containment patterns.) Since the events for different indices i are independent, by the weak law of large numbers we find that $|J_1 \triangle J_2|/M = 1/2 + o(1)$ holds with probability $1 + o(1)$ as M grows, and the result follows. ☐

The Lemma immediately gives that for all $N^2(1 + o(1))$ "generic" choices of J_1, J_2 (also distinct), the symmetric difference has size roughly of order $M/2$, hence leaving $2^{M-|J_1 \triangle J_2|} = o(N)$ possibilities to choose the intersection $J_3 \cap J_4$. The total contribution to the variance is thus

$$\ll_f \frac{1}{N^2}(N^2 \cdot o_f(N))\frac{1}{N} = o_f(1)$$

(here we have used that $\mathbb{E}(|F_N(x_{J_1} - x_{J_2}) \cdot F_N(x_{J_3} - x_{J_4})|) \ll ||f||_\infty^2/N \ll_f 1/N$ since f having compact support implies that the support of F_N is contained in a ball of radius $\ll 1/N$, together with the fact that the rank is one.)

Finally, for the $o(N^2)$ non-generic choices of J_1, J_2, there can be at most $N = 2^M$ possibilities for the intersection $J_3 \cap J_4$; again the total contribution is

$$\ll_f o(N^2) \cdot N \cdot \frac{1}{N^2} \cdot \frac{1}{N} = o_f(1).$$

7.1.2. *Higher level correlations.* The general case is more involved combinatorially, but the key idea is still that the expectation of the r-level correlation for $x_{J_1}, \ldots, x_{J_r}$ of "full rank" (which holds for generic choices of subsets $J_1, \ldots, J_r$) dominates, and then to bound the contribution from low rank tuples.

The expectation of $\mathcal{R}_r^*$: We begin with a result regarding uniform distribution on tori.

Lemma 7.2. *Let $A \in \mathrm{GL}_n(\mathbb{Q}) \cap \mathrm{Mat}_{n\times n}(\mathbb{Z})$, and let $X = (X_1, \ldots, X_n)$ denote a uniformly distributed random variable on $\mathbb{T}^n$. Then $Y = AX$ is also uniform. More generally, if $n \geq m$ and $A \in \mathrm{Mat}_{m\times n}(\mathbb{Z})$ is an $m \times n$-matrix with rank m, and X is a uniformly distributed random variable on $\mathbb{T}^n$, then $Y = AX$ is uniformly distributed on $\mathbb{T}^m$.*

Proof. For $n = 1$, i.e., $Y = Y_1 = aX_1$, for nonzero $a \in \mathbb{Z}$ the result is clear (just consider the preimage of a small interval; it will consist of $|a|$ copies of intervals whose lengths are scaled by $1/|a|$.) For $n > 1$, A can be decomposed as $A = B_1 D B_2$ where $B_1, B_2 \in \mathrm{SL}_n(\mathbb{Z})$ and D is a diagonal matrix with integer entries (we can take D to be the Smith normal form of A, cf. [11, Theorem 2.4.12]). It is thus enough to prove the statement for $A \in \mathrm{SL}_n(\mathbb{Z})$ or $A = D$. The former is clear as multiplication by A does not change the measure (the determinant of the Jacobian equals ± 1.) The case $A = D$ and $n = 1$ is already done, and writing $dx = dx_1 \cdots dx_n$ the general case follows by change of measure one component at a time.

For $A \in \mathrm{Mat}_{m\times n}(\mathbb{Z})$ the argument is similar. After permuting columns we may assume that the m first columns of A are linearly independent. Decomposing the uniform measure on $\mathbb{T}^n$ as a product of two uniform measures $\mu_1 \times \mu_2$ on $\mathbb{T}^m \times \mathbb{T}^{n-m}$, the result follows by conditioning on the $\mathbb{T}^{n-m}$-component, using the first part applied to $\mathbb{T}^m$-component, together with the uniform measure being translation invariant. ☐

Before proceeding we introduce some further notation. As before, given subsets $J_i \subset \{1,\dots,M\}$ for $i=1,\dots,r$, let $x_i = \sum_{j\in J_i} \vartheta_j$. In order to discuss linear independence and rank, first recall that $\vec{v}_i = \sum_{j\in J_i} \vec{e}_j$ where $\vec{e}_1,\dots,\vec{e}_M$ denotes the standard basis of $\mathbb{R}^M$. We can then write

$$x_i = \vec{v}_i \cdot \vec{\vartheta}$$

where $\vec{\vartheta} = (\vartheta_1,\dots,\vartheta_M)$ denotes a vector of M uniform and independent random variables taking values in $\mathbb{T}^1$. Letting

$$\Delta_i := x_i - x_{i+1}, \quad \text{and} \quad \vec{w}_i := \vec{v}_i - \vec{v}_{i+1}, \quad i = 1,\dots,r-1 \tag{7.2}$$

we find that

$$\Delta_i = (\vec{v}_i - \vec{v}_{i+1}) \cdot \vec{\vartheta} = \vec{w}_i \cdot \vec{\vartheta}.$$

Lemma 7.3. *Let f be a compactly supported Schwartz function. Given any r-tuple of distinct subsets $J_1,\dots,J_r \subset \{1,\dots,M\}$ with associated difference vectors $\vec{w}_1,\dots,\vec{w}_{r-1}$ (cf. (7.2)), let d denote the dimension of the vector space spanned by these vectors. If d is maximal, i.e., $d=r-1$, then*

$$\mathbb{E}\left(F_N(x_1-x_2,\dots,x_{r-1}-x_r)\right) = \frac{1}{N^{r-1}}\widehat{f}(0).$$

If $0<d<r-1$, then

$$\mathbb{E}\left(|F_N(x_1-x_2,\dots,x_{r-1}-x_r)|\right) \ll_f 1/N^d.$$

Proof. The first part is an immediate consequence of Lemma 7.2, as $(\Delta_1,\dots,\Delta_{r-1})$ is a random variable uniformly distributed on $\mathbb{T}^{r-1}$.

The second part follows by a similar argument: choose d linearly independent vectors $\vec{w}_{j_1},\vec{w}_{j_2},\dots\vec{w}_{j_d}$, and use that

$$\mathbb{E}(|F_N(x_1-x_2,\dots,x_{r-1}-x_r)|) \ll_f \mathbb{E}(g_N(x_{j_1}-x_{j_2},\dots,x_{j_d}-x_{j_{d+1}})) \ll_f 1/N^d, \tag{7.3}$$

where $g_N(x_{j_1}-x_{j_2},\dots,x_{j_d}-x_{j_{d+1}})$ is obtained from F_N by taking the supremum of $|F_N(x_1-x_2,\dots,x_{r-1}-x_r)|$ over the coordinates corresponding to $x_{j_i}-x_{j_{i+1}}$ fixed for $i=1,\dots,d$ and the other ones ranging freely. (Here we use that the support of g_N is contained in some d-dimensional ball of radius $\ll_f 1/N$.) ☐

Before proceeding we next show that all low rank subspaces can be defined via $O_r(1)$ linear forms (i.e., the number of forms does not depend on M.)

Lemma 7.4. *With $\vec{w}_1,\dots,\vec{w}_{r-1}$ as defined in (7.2), let $W = \mathrm{Span}(\vec{w}_1,\dots,\vec{w}_{r-1})$ and assume that $d := \dim(W) < r-1$. Choosing a basis for W consisting of d linearly independent elements of $\{\vec{w}_1,\dots,\vec{w}_{r-1}\}$, the remaining vectors are given by $r-1-d$ linear forms in the basis vectors, and the number of distinct collections of such forms is then $O_r(1)$ (in particular, the estimate is uniform for all tuples $J_1,\dots,J_{r-1}$ as long as $d<r-1$.)*

Proof. After renumbering indices (there are at most $(r-1)! = O_r(1)$ ways to do this) we may assume that $\vec{w}_1,\dots,\vec{w}_d$ are linearly independent, and that for all integers $i \in [1, r-1-d]$ we have

$$\gamma_i \vec{w}_{d+i} = L_i(\vec{w}_1,\dots,\vec{w}_d)$$

with $\gamma_i \neq 0$ an integer and each L_i a linear form with integer coefficients, with the property that the gcd of γ_i and the coefficients of L_i equals one. Form a $M \times (d+1)$ matrix A having columns $\vec{w}_1,\dots,\vec{w}_d,\vec{w}_{d+i}$; the above linear relation can then be formulated in terms of the existence of a nonzero vector $\vec{l} \in \mathbb{Z}^{d+1}$ so that $A\vec{l} = 0$. Since $\vec{w}_1, \ldots, \vec{w}_d$ are linearly independent, we may form a $d \times (d+1)$ matrix B, having rank d, by selecting d linearly independent rows in A, with the property that $A\vec{l} = 0$ if and only if $B\vec{l} = 0$. Further the set of $\{\vec{l} : B\vec{l} = 0\}$ lies on a line, so if the coordinates of $\vec{l}$ have gcd one (in analogy with the above gcd condition), $\vec{l}$ is up to sign uniquely determined by A. On the other hand, each entry in B lies in $\{-1, 0, 1\}$, hence there are at most $O_d(1) = O_r(1)$ possible ways to choose B, and hence the number of ways to choose $\vec{l}$ is also $O_r(1)$.

Thus, for each i, there are $O_r(1)$ ways to choose γ_i, L_i, and thus there are in total $O_r(1)$ possible ways to select $r - 1 - d$ such linear relations. □

Proposition 7.5. *The number of $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r) \in (\{0,1\}^M)^r$ such that $\vec{v}_i \neq \vec{v}_j$ for $i \neq j$, and so that the rank of $(\vec{w}_1, \ldots, \vec{w}_{r-1})$ (cf. (7.2)) equals $r - 1$, is*

$$N^r(1 + o_r(1))$$

as $M \to \infty$.

Further, the number of $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r) \in (\{0,1\}^M)^r$ that are pairwise distinct, and such that the rank of $(\vec{w}_1, \ldots, \vec{w}_{r-1})$ equals $d < r - 1$ is

$$o_r(N^{d+1}),$$

as $M \to \infty$.

Before giving the proof we record the following simple linear algebra result.

Lemma 7.6. *With $\vec{w}_1, \ldots, \vec{w}_{r-1}$ as defined in (7.2), let $W = \mathrm{Span}(\vec{w}_1, \ldots, \vec{w}_{r-1})$ and let $d = \dim(W) < r - 1$. Then there exists a permutation of indices $i_1, i_2, \ldots, i_r$ such that if we let $\vec{w}'_j := \vec{v}_{i_j} - \vec{v}_{i_{j+i}}$ for $j = 1, \ldots, r - 1$, we have $W = \mathrm{Span}(\vec{w}'_1, \ldots, \vec{w}'_d)$.*

Proof. We begin by noting that the vectors $\vec{w}_1, \ldots, \vec{w}_{r-1}$ have the same span as the vectors $\vec{v}_1 - \vec{v}_2, \vec{v}_1 - \vec{v}_3, \ldots, \vec{v}_1 - \vec{v}_r$, since they are related by a lower triangular matrix all whose entries are one. Thus there exist a relabeling of indices $i_1, \ldots, i_r$ such that $i_1 = 1$, and that the vectors $\vec{v}_{i_1} - \vec{v}_{i_2}, \vec{v}_{i_1} - \vec{v}_{i_3}, \ldots, \vec{v}_{i_1} - \vec{v}_{i_{d+1}}$ are linearly independent. Letting $\vec{w}'_j = \vec{v}_{i_j} - \vec{v}_{i_{j+1}}$ for $j = 1, \ldots, r - 1$, and arguing as above, we find that that $\vec{w}'_1, \vec{w}'_2, \ldots, \vec{w}'_d$ are linearly independent. □

Proof of Proposition 7.5. We give a simple probabilistic proof showing that certain events occur with probability $1 - o(1)$ as M grows (we write $1 - o(1)$ rather than $1 + o(1)$ to emphasize that the probability is ≤ 1.) Namely, pick vectors $\vec{v}_1, \ldots, \vec{v}_r \in \{0,1\}^M$ by fair and independent coin flips — to later obtain asymptotics for counts of vectors with properties of interest we then multiply said probability with $2^{Mr} = N^r$. For M large, the vectors are all pairwise distinct with probability $1 - o(1)$. Further, if we consider a coordinate $j \in \{1, \ldots, M\}$ and fix $i \in \{1, \ldots, r\}$, the likelihood that the j-th coordinate of $\vec{v}_i$ is one, and that the j-th coordinate for all other vectors $\vec{v}_l$, $l \in \{1, \ldots, r\} \setminus \{j\}$, is zero is $1/2^r$. Thus, by the law of large numbers (as in the proof of Lemma 7.1) with probability $1 - o(1)$ as $M \to \infty$, there are r indices $j_1, j_2, \ldots, j_r$ so that $\vec{v}_l \cdot \vec{e}_{j_i} = \delta_{li}$ for $l = 1, \ldots, r$. Thus the rank of $\vec{v}_1, \ldots, \vec{v}_r$ is r with probability $1 - o_r(1)$, and consequently the rank of $\vec{w}_1, \ldots, \vec{w}_{r-1}$ is $r - 1$, which implies the first part.

As for the second part, we may now assume that $\vec{v}_1, \ldots, \vec{v}_r$ are pairwise distinct (this holds with probability $1 - o(1)$.) We next use Lemma 7.6 to reduce to the case of $\vec{w}_1, \ldots, \vec{w}_d$ being linearly independent, and that $\vec{w}_{d+1}$ can be written as a linear combination of the first d vectors. In particular there exists scalars α_i such that $\sum_{i=1}^{d+1} \alpha_i \vec{w}_i = \vec{0}$ (with $\alpha_{d+1} \neq 0$; also note that $\vec{v}_1, \ldots, \vec{v}_r$ being pairwise distinct implies that $\alpha_i \neq 0$ for some $i \leq d+1$), leading to a linear relation

$$\sum_{i=1}^{d+2} \beta_i \vec{v}_i = \vec{0}$$

where $\beta_1 = \alpha_1, \beta_{d+2} = -\alpha_{d+1}$, and otherwise $\beta_j = \alpha_j - \alpha_{j-1}$; note in particular that $\beta_{d+2} \neq 0$. Fix the coefficients of such a relation (note that there are $O_r(1)$ possible relations by Lemma 7.4.) Further, if l denotes the smallest integer such that $\alpha_l \neq 0$, we also find that $\beta_l \neq 0$, in particular there are at least two nonvanishing $\beta_i's$. If there are exactly two nonzero β_i's, we have $\beta_l \vec{v}_l = -\beta_{d+2} \vec{v}_{d+2}$. Since $\beta_l, \beta_{d+2} \neq 0$ (and at least one of $\vec{v}_l$ and $\vec{v}_{d+2}$ must be nonzero), we in fact have have $\vec{v}_l, \vec{v}_{d+2} \neq \vec{0}$, and since the coordinates of both $\vec{v}_l$ and $\vec{v}_{d+2}$ are in $\{0,1\}$ we must have $\beta_l = -\beta_{d+2}$, and thus $\vec{v}_l = \vec{v}_{d+2}$, contradicting the vectors being assumed to be pairwise distinct.

Now, if $\beta_i \neq 0$ for at least three values, we may write

$$\alpha \vec{v}_{i_1} + \gamma \vec{v}_{i_2} = \sum_{\substack{1 \leq i \leq d+2 \\ i \neq i_1, i_2}} \delta_i \vec{v}_i$$

with $\alpha, \gamma > 0$. In particular, the support[2] of $\vec{v}_{i_1} + \vec{v}_{i_2}$ is determined by the right-hand side, and thus (after taking the relabeling of indices into account), so is $J_{i_1} \cup J_{i_2}$. In fact, $J_{i_1} \cap J_{i_2}$ as well as the symmetric difference $J_{i_1} \Delta J_{i_2}$ is also determined by the right-hand side. Letting M_0 denote the cardinality of said determined symmetric difference, we find that there are 2^{M_0} ways to choose J_{i_1}, J_{i_2} for a fixed right-hand side (note that there are at most N^d possibilities for the right-hand side). Now, for a "generic" right-hand side, a simple probabilistic argument (i.e., using the law of large numbers as before) to count the number of indices not in any J_i for $i \neq i_1, i_2$, shows that for $N^d(1 + o_r(1))$ of the possible choices of the right-hand side, we have $|\cup_{i \neq i_1, i_2} J_i| \leq (1-\epsilon)M$ (where $\epsilon = \epsilon_r > 0$ is allowed to depend on r but not on M.) If this is the case, we have $M_0 \leq (1-\epsilon)M$ and there are at most $2^{(1-\epsilon)M} = o_r(N)$ ways to choose J_{i_1}, J_{i_2}, for a total of $o_r(N^{d+1})$ possibilities. On the other hand, if the right-hand side is one of the $o_r(N^d)$ "non-generic" choices (in particular allowing $M_0 = M$), there are $2^{M_0} \leq 2^M = N$ possible ways to choose J_{i_1}, J_{i_2}, and the total number of possibilities is also here $o_r(N^{d+1})$.

Finally, once $\vec{v}_1, \ldots, \vec{v}_{d+2}$ are chosen, *all* $\vec{w}_i$ are determined since $\vec{w}_{d+1}, \vec{w}_{d+2}, \ldots, \vec{w}_{r-1}$ depend linearly on $\vec{w}_1, \ldots, \vec{w}_d$, and thus the vectors $\vec{v}_{d+3}, \ldots \vec{v}_r$ are also uniquely determined. As, by Lemma 7.4, the number of linear relations is $O_r(1)$, we find that the total number of ways to choose $\vec{v}_1, \ldots, \vec{v}_r$ so that $\vec{w}_1, \ldots, \vec{w}_{r-1}$ has non-maximal rank $d < r-1$ is $o_r(N^{d+1})$. □

Theorem 7.7. *Let f be a compactly supported Schwartz function. As M grows we have*

$$\mathbb{E}\left(\mathcal{R}_r^*(F_N)\right) = \widehat{f}(0) + o_f(1)$$

and

$$\mathbb{E}\left(\mathcal{R}_r^*(F_N)^2\right) = \widehat{f}(0)^2 + o_f(1).$$

[2]By the support of a vector we mean the set of indices for which the corresponding coordinates are nonzero.

Proof. The expectation: By linearity of the expectation, we have

$$\mathbb{E}\left(\mathcal{R}_r^*(F_N)\right) = \frac{1}{N} \sideset{}{^*}\sum_{J_1,\ldots,J_r \subset \{1,\ldots,M\}} \mathbb{E}\left(F_N(x_1 - x_2, \ldots, x_{r-1} - x_r)\right).$$

The main term arises from "full rank" choices of distinct subsets $J_1, \ldots, J_{r-1}$; by Proposition 7.5 there are $N^r(1 + o_r(1))$ such tuples, and by Lemma 7.3, each such term contributes

$$\frac{1}{N^{r-1}} \widehat{f}(0).$$

The contribution from the choices of distinct subsets having rank $d < r-1$ is then (after taking into account there being at most $O_r(1)$ possible linear relations), using the second parts of Proposition 7.5 and Lemma 7.3,

$$\ll_{r,f} \frac{1}{N} \sum_{d=1}^{r-2} \frac{1}{N^d} o_r(N^{d+1}) = o_{r,f}(1).$$

Bounding the variance: The argument to bound the variance is similar to the one used to bound the pair correlation variance. We first note that

$$\mathbb{E}\left(\mathcal{R}_r^*(F_N)^2\right) = \frac{1}{N^2} \sideset{}{^*}\sum_{J_1,\ldots,J_r \subset \{1,\ldots,M\}} \sideset{}{^*}\sum_{J_1',\ldots,J_r' \subset \{1,\ldots,M\}} \mathbb{E}\left(F_N(x_1 - x_2, \ldots, x_{r-1} - x_r) \cdot F_N(x_1' - x_2', \ldots, x_{r-1}' - x_r')\right). \tag{7.4}$$

The number of "full rank" tuples (i.e., of maximimal rank $2(r-1)$) is $N^{2r} \cdot (1 + o(1))$, and for these, Lemma 7.2 gives that

$$\mathbb{E}\left(F_N(x_1 - x_2, \ldots, x_{r-1} - x_r) \cdot F_N(x_1' - x_2', \ldots, x_{r-1}' - x_r')\right)$$

$$= \mathbb{E}\left(F_N(x_1 - x_2, \ldots, x_{r-1} - x_r)\right) \cdot \mathbb{E}\left(F_N(x_1' - x_2', \ldots, x_{r-1}' - x_r')\right) = \left(\frac{\widehat{f}(0)}{N^{r-1}}\right)^2.$$

To bound the contribution from tuples with rank $d < 2(r-1)$ we argue as follows. Letting W denote the linear span of $\vec{w}_1, \ldots, \vec{w}_{r-1}, \vec{w}_1', \ldots, \vec{w}_{r-1}'$ we have $d = \dim(W)$, and by an argument similar to the one used in Lemma 7.3 (i.e., use an inequality analogous to (7.3) together with Lemma 7.2) gives that

$$\mathbb{E}\left(\left|F_N(x_1 - x_2, \ldots, x_{r-1} - x_r) \cdot F_N(x_1' - x_2', \ldots, x_{r-1}' - x_r')\right|\right) \ll_f 1/N^d.$$

It thus suffices to show that the number of tuples of rank d is $o(N^{d+2})$. We may after relabeling the $\vec{v}_i$ as well as the $\vec{v}_i'$-indices select d_1 vectors $\vec{w}_1, \ldots, \vec{w}_{d_1}$ and d_2 vectors $\vec{w}_1', \ldots, \vec{w}_{d_2}'$ (with the vectors $\vec{w}_i$ and $\vec{w}_i'$ defined using the reordered indices) $d = d_1 + d_2$ and

$$\mathrm{Span}(\vec{w}_1, \ldots, \vec{w}_{d_1}, \vec{w}_1', \ldots, \vec{w}_{d_2}') = \mathrm{Span}(\vec{w}_1, \ldots, \vec{w}_{r-1}, \vec{w}_1', \ldots, \vec{w}_{r-1}').$$

More precisely, relabel both the $\vec{v}_i$ and the $\vec{v}_j'$ sets of indices as in the proof of Lemma 7.6, in particular with $\vec{v}_i$ given by J_i after relabeling the first set of indices put $\vec{w}_i = \vec{v}_{i+1} - \vec{v}_i$, and similarly with $\vec{v}_i'$ given by J_i' after relabeling the second set of indices put $\vec{w}_i' := \vec{v}_i' - \vec{v}_{i+1}'$. We then find that a basis for W is given by $\vec{w}_1, \ldots, \vec{w}_{d_1}, \vec{w}_1', \ldots, \vec{w}_{d_2}'$.

Consider now a minimal non-trivial linear relation, i.e.,

$$\sum_{i=1}^{d_1} \alpha_i \vec{w}_i + \sum_{i=1}^{d_2} \alpha_i' \vec{w}_i' + \gamma \vec{w} = 0$$

where $\gamma \neq 0$ and $\vec{w} = \vec{w}_{d_1+1}$ or $\vec{w} = \vec{w}'_{d_2+1}$. In case the relation purely involves either $\vec{w}_i'$ or $\vec{w}_i$-vectors the same argument used to treat the expectation suffices: if (say) the relation only involves w_i' vectors (so that $\alpha_1 = \ldots = \alpha_{d_1} = 0$ and $\vec{w} = \vec{w}'_{d_2+1}$), the number of such tuples is

$$\ll N^{d_1+1} o_r(N^{d_2+1}) = o_r(N^{d+2})$$

as there are at most N^{d_1+1} choices for the unconstrained $\vec{v}_1, \ldots, \vec{v}_{d_1+1}$, and the remaining $\vec{v}_i$ are determined by the first $d_1 + 1$ ones; by Proposition 7.5 there are $o(N^{d_2+1})$ possible $\vec{v}_i'$-tuples. In case the relation involves at least one $\vec{w}_i$-vector and at least one $\vec{w}_i'$-vector, we obtain a linear relation

$$\sum_{i=1}^{d_1+1} \beta_i \vec{v}_i + \sum_{j=1}^{d_2+1} \beta_j' \vec{v}_j' = 0$$

with at least three nonzero coefficients — the same argument used to prove Proposition 7.5 then gives that the number of such $2r$-tuples is also $o(N^{d_1+1+d_2+1}) = o(N^{d+2})$. $\square$

7.2. **Poisson correlations in the general case.** We next consider integers n_0, where n_0 is fixed and allowed to have prime power divisors, and the angle contribution from the n_0-part is entirely deterministic, whereas we use the random model for the square free part. We define

$$\mathcal{R}^*_{r,n_0}(F_N) := \frac{1}{N} \sum_{J_1,\ldots,J_r \subset \{1,\ldots,M\}}^{*} \sum_{(\beta_1),\ldots,(\beta_r) \subset \mathcal{O} : |\mathcal{O}/\beta_i| = n_0} F_N(\theta_{\beta_2} - \theta_{\beta_1} + x_{J_2} - x_{J_1}, \ldots). \tag{7.5}$$

(Note that with probability one the angles $x_{J_1} + \theta_{\beta_1}$ and $x_{J_2} + \theta_{\beta_2}$ are equal if and only if $J_1 = J_2$ and $(\beta_1) = (\beta_2)$ so that with probability one the sum above is over distinct points $x_{J_1} + \theta_{\beta_1}, \ldots, x_{J_r} + \theta_{\beta_r}$.)

Lemma 7.8. *Fix n_0 such that $N_0 := r(n_0) > 0$. Let f be a compactly supported Schwartz function. Further, given an integer $M > 0$ let $N = 2^M r(n_0)$. Then, as M grows,*

$$\mathbb{E}\big(\mathcal{R}^*_{r,n_0}(F_N)\big) = \widehat{f}(0) + o_f(1), \qquad \textit{and} \qquad \mathbb{E}\big(\mathcal{R}^*_{r,n_0}(F_N)^2\big) = \Big(\widehat{f}(0)\Big)^2 + o_f(1).$$

Proof. If $n_0 = 1$ then Theorem 7.7 immediately gives the result.

For general $n_0 > 1$ we now consider the expectation of $\mathcal{R}^*_{r,n_0}(F_N)$. First note that we, by definition, have $J_i \neq J_j$ for all $1 \leq i \neq j \leq r$ for all terms of $\sum^*$ in (7.5). With $N_1 = 2^M$ and $N_0 = r(n_0)$ we have $N = N_0 \cdot N_1$; and by the same argument used in Section 7.1 (the key point is that we may shift, and multiply by a bounded factor $1/N_0$ each coordinate of F_N) we find that for each fixed r-tuple $(\beta_1, \ldots, \beta_r)$, we have

$$\frac{1}{N} \sum_{J_1,\ldots,J_r \subset \{1,\ldots,M\}}^{*} \mathbb{E}\left(F_N(\theta_{\beta_2} - \theta_{\beta_1} + x_{J_2} - x_{J_1}, \ldots)\right) = (1/N_0^r) \left(\widehat{f}(0) + o_f(1)\right).$$

Summing over the N_0^r configurations of r-tuples $(\beta_1, \ldots, \beta_r)$ we find that (7.5) equals $\widehat{f}(0) + o_f(1)$.

The argument to bound the variance is similar — the key point is that for each fixed pair of collections of shifts $\beta_1, \dots, \beta_r$ and $\beta'_1, \dots, \beta'_r$ the corresponding main term, arising from pairs of "generic" subsets $J_1, \dots, J_r$ and $J'_1, \dots, J'_r$, can be evaluated exactly as before. To bound the contribution from non-generic choices of subsets, we can use the previous argument to obtain a bound that is only worse by a factor of N_0^{2r}. $\square$

8. Concluding the Proof

Proof of Theorems 1.1 and 1.2. First fix a compactly supported Schwartz function f. For $n \in \mathcal{S}$ with $r(n) = 2^M = N$ write

$$R_r^*(n; F_N) = \frac{1}{N} \sum_{\substack{(\beta_1),\dots,(\beta_r) \subset \mathcal{O} \\ |\mathcal{O}/\beta_1| = \dots = |\mathcal{O}/\beta_r| = n \\ (\beta_1),\dots,(\beta_r) \text{ distinct}}} F_N(\theta_{\beta_1} - \theta_{\beta_2}, \dots, \theta_{\beta_{r-1}} - \theta_{\beta_r}),$$

where the sum is over pairwise distinct ideals (i.e. pairwise distinct angles), and let $R_{r,n_0}^*(n; F_N) = R_r^*(n_0 n; F_N)$. By Theorem 3.1, asymptotics for the average and the second moment of the deterministic sums R_{r,n_0} are given, up to negligible errors, by the corresponding average and second moment of the random model (cf. (3.3) (3.4).) To pass to correlations over distinct angles, i.e. R_{r,n_0}^*, we use a standard combinatorial sieving argument (cf. [37, Section 4], note that (3.3) and (3.4) hold for arbitrary Schwartz functions f), which allows us to match the average and second moment of the r-correlation of *distinct* angles of the deterministic sum to those of the corresponding random model $\mathcal{R}_{r,n_0}^*$ (as given in (7.5).) The first and second moments of the random model $\mathcal{R}_{r,n_0}^*$ are evaluated in Section 7 (cf. Lemma 7.8.)

We will now show that the r-correlation of the deterministic angles is Poissonian along a full density subsequence of $\mathcal{S}$ for the given f. Let $\psi(x)$ be a non-increasing function that tends to zero arbitrarily slowly as $x \to \infty$. By the preceding discussion, using Markov's inequality we conclude that for any fixed $n_0 \in \mathcal{S}$ and $M \leq \log\log x$ that also tends to infinity with x that

$$\begin{aligned} &\#\{n \in \mathcal{N}_{M,n_0}(x) : |R_{r,n_0}^*(n; F_N) - \widehat{f}(0)| \geq \psi(n)\} \\ &\qquad \leq \frac{1}{\psi(x)^2} \sum_{n \in \mathcal{N}_{M,n_0}(x)} |R_{r,n_0}^*(n; F_N) - \widehat{f}(0)|^2 = o(\#\mathcal{N}_{M,n_0}(x)). \end{aligned} \tag{8.1}$$

As in Section 2, for $n \in \mathcal{S}$ with $n = ef^2$ where e is square-free write $n = n_0 n_1$ where $n_0 = 2^a f^2 d$ for some $a \in \{0, 1\}$, $d|f$ with $\mu(2d)b(d) = 1$ and n_1 satisfies $\mu^2(2n_1)b(n_1) = 1$, $(n_0, n_1) = 1$. Also, let $\Psi(x)$ be a monotone function that tends to infinity arbitrarily slowly with x. In the preceding notation, we claim that the subset

$$\mathcal{S}' = \{n \in \mathcal{S} : \Omega_1(n_1) \leq \log\log n \text{ and } f \leq \Psi(n)\}$$

has full density within $\mathcal{S}$, which we will prove later. Writing $\mathcal{S}'(x) = \{n \in \mathcal{S}' : n \leq x\}$ we note that if $n \in \mathcal{S}'(x)$ then $n \in \mathcal{N}_{M,n_0}(x)$ for some $M \leq \log\log x$ and $n_0 = 2^a f^2 d$ with $f \leq \Psi(x)$. Hence, using (8.1) we get that the set

$$\mathcal{S}'' = \mathcal{S}' \cap \{n \in \mathcal{S} : |R_{r,n_0}^*(n; F_N) - \widehat{f}(0)| < \psi(n)\},$$

has full density within $\mathcal{S}$. In the last step we used that $\Psi(x)$ grows arbitrarily slowly with x to account for the lack of an explicit dependence of the rate of decay of (8.1) on n_0.

Thus, once we remove a zero density subset of elements in $\mathcal{S}$, the r-level correlation with respect to the fixed Schwartz function f is Poissonian. To show that the same holds for all compactly supported Schwartz functions, we may take a countable and dense (say in the C^∞-norm) collection of compactly supported Schwartz functions, and a standard diagonalization argument then gives Poisson correlations (for all r) for some full density sub-subsequence of elements in $\mathcal{S}$ *provided* that the correlation functionals are continuous with respect to the C^∞-norm. This in turn can be seen as follows: if f is supported in some ball of radius δ, then $\limsup_{N\to\infty} \mathbb{E}(|\mathcal{R}_r^*(F_N)|) \ll_\delta \|f\|_\infty$.

Since the r-level correlations are Poissonian along a full density subset of $\mathcal{S}$, Theorems 1.1 and 1.2 are easily deduced using a combinatorial argument (e.g. see [27, Appendix A].)

Finally, we establish the claim that $\mathcal{S}'$ has full density. Arguing as in (2.1) we see that (8.2)

$$\#\{n \in \mathcal{S} : n \le x \,\&\, f \ge \Psi(n)\} \ll \sum_{a\in\{0,1\}} \sum_{f \ge \Psi(\sqrt{x})} \sum_{d|f} \sum_{n_1 \le \frac{x}{2^a f^2 d}} b(n_1) + \sqrt{x} \ll \frac{x \log(\Psi(\sqrt{x}))}{\Psi(\sqrt{x})(\log x)^{1/2}},$$

where to obtain the first bound we separately analyzed the $n \in \mathcal{S}$ with $n \le \sqrt{x}$ (which we estimated trivially) and $\sqrt{x} \le n \le x$ (for which $\Psi(n) \ge \Psi(\sqrt{x})$.) Next, we note that if h is a multiplicative function with $0 \le h(n) \le \sum_{d|n} 1$ then

$$\sum_{n\le x} h(n) \ll \frac{x}{\log x} \prod_{p \le x} \left(1 + \frac{h(p)}{p}\right),$$

see [24, Eq'n (1.85), p. 24]. Applying the preceding bound with $h(n) = b(n)2^{\Omega_1(n)}$, noting for $n = n_0 n_1 \in \mathcal{S}$ that $\Omega_1(n) \ge \Omega_1(n_1)$, and using Markov's inequality we conclude

$$\begin{aligned}\#\{n \in \mathcal{S} : n \le x, \,\&\, \Omega(n_1) \ge \log\log n\} &\ll \sqrt{x} + \frac{1}{2^{\log\log x}} \sum_{\sqrt{x} \le n \le x} b(n) 2^{\Omega_1(n)} \\ &\ll \frac{x}{(\log x)^{\log 2 + 1}} \prod_{p\le x} \left(1 + \frac{2b(p)}{p}\right) \ll \frac{x}{(\log x)^{\log 2}},\end{aligned}$$

and since $\log 2 > 1/2$ this is $o(x/\sqrt{\log x})$. Combining this with (8.2) concludes the proof of the claim. $\square$

References

[1] V. Augustin, F. P. Boca, C. Cobeli, and A. Zaharescu. The h-spacing distribution between Farey points. *Math. Proc. Cambridge Philos. Soc.*, 131(1):23–38, 2001.

[2] F. P. Boca, C. Cobeli, and A. Zaharescu. Distribution of lattice points visible from the origin. *Comm. Math. Phys.*, 213(2):433–470, 2000.

[3] F. P. Boca, C. Cobeli, and A. Zaharescu. A conjecture of R. R. Hall on Farey points. *J. Reine Angew. Math.*, 535:207–236, 2001.

[4] F. P. Boca and A. Zaharescu. The correlations of Farey fractions. *J. London Math. Soc. (2)*, 72(1):25–39, 2005.

[5] E. B. Bogomolny and J. P. Keating. Random matrix theory and the Riemann zeros. I. Three- and four-point correlations. *Nonlinearity*, 8(6):1115–1131, 1995.

[6] E. B. Bogomolny and J. P. Keating. Random matrix theory and the Riemann zeros. II. n-point correlations. *Nonlinearity*, 9(4):911–935, 1996.

[7] J. Bourgain, P. Sarnak, and Z. Rudnick. Local statistics of lattice points on the sphere. In *Modern trends in constructive function theory*, volume 661 of *Contemp. Math.*, pages 269–282. Amer. Math. Soc., Providence, RI, 2016.

[8] V. Chandee and Y. Lee. n-level density of the low-lying zeros of primitive Dirichlet L-functions. *Adv. Math.*, 369:107185, 70, 2020.

[9] S. Chaubey and N. Yesha. The distribution of spacings of real-valued lacunary sequences modulo one. available at arXiv:2108.00431, 2021.

[10] K.-L. Chung. An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.*, 67:36–50, 1949.

[11] H. Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.

[12] M. D. Coleman. A zero-free region for the Hecke L-functions. *Mathematika*, 37(2):287–304, 1990.

[13] B. Conrey and J. P. Keating. Moments of zeta and correlations of divisor-sums: V. *Proc. Lond. Math. Soc. (3)*, 118(4):729–752, 2019.

[14] J. B. Conrey and N. C. Snaith. In support of n-correlation. *Comm. Math. Phys.*, 330(2):639–653, 2014.

[15] R. de la Bretèche and G. Tenenbaum. Remarks on the Selberg-Delange method. *Acta Arith.*, 200(4):349–369, 2021.

[16] A. Entin, E. Roditty-Gershon, and Z. Rudnick. Low-lying zeros of quadratic Dirichlet L-functions, hyper-elliptic curves and random matrix theory. *Geom. Funct. Anal.*, 23(4):1230–1261, 2013.

[17] P. Erdős and R. R. Hall. On the angular distribution of Gaussian integers with fixed norm. volume 200, pages 87–94. 1999. Paul Erdős memorial collection.

[18] W. Feller. *An introduction to probability theory and its applications. Vol. II*. John Wiley & Sons, Inc., New York-London-Sydney, second edition, 1971.

[19] A. Granville and D. Koukoulopoulos. Beyond the LSD method for the partial sums of multiplicative functions. *Ramanujan J.*, 49(2):287–319, 2019.

[20] C. Hooley. On the difference of consecutive numbers prime to n. *Acta Arith.*, 8:343–347, 1962/63.

[21] C. Hooley. On the difference between consecutive numbers prime to n. II. *Publ. Math. Debrecen*, 12:39–49, 1965.

[22] C. Hooley. On the difference between consecutive numbers prime to n. III. *Math. Z.*, 90:355–364, 1965.

[23] B. Huang. Sup-norm and nodal domains of dihedral Maass forms. *Comm. Math. Phys.*, 371(3):1261–1282, 2019.

[24] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[25] I. Kátai and I. Környei. On the distribution of lattice points on circles. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 19:87–91 (1977), 1976.

[26] N. M. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bull. Amer. Math. Soc. (N.S.)*, 36(1):1–26, 1999.

[27] P. Kurlberg and Z. Rudnick. The distribution of spacings between quadratic residues. *Duke Math. J.*, 100(2):211–242, 1999.

[28] P. Kurlberg and H. Ueberschär. Quantum Ergodicity for Point Scatterers on Arithmetic Tori. *Geom. Funct. Anal.*, 24(5):1565–1590, 2014.

[29] P. Kurlberg and I. Wigman. On probability measures arising from lattice points on circles. *Math. Ann.*, 367(3-4):1057–1098, 2017.

[30] J. Marklof and A. Strömbergsson. The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems. *Ann. of Math. (2)*, 172(3):1949–2033, 2010.

[31] A. M. Mason and N. C. Snaith. Orthogonal and symplectic n-level densities. *Mem. Amer. Math. Soc.*, 251(1194):v+93, 2018.

[32] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, pages 181–193, 1973.

[33] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[34] I. Niven. *Irrational numbers*. The Carus Mathematical Monographs, No. 11. The Mathematical Association of America. Distributed by John Wiley and Sons, Inc., New York, N.Y., 1956.

[35] J. J. Rotman. *Advanced modern algebra*. Prentice Hall, Inc., Upper Saddle River, NJ, 2002.

[36] Z. Rudnick and P. Sarnak. The n-level correlations of zeros of the zeta function. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(10):1027–1032, 1994.

[37] Z. Rudnick and P. Sarnak. Zeros of principal L-functions and random matrix theory. volume 81, pages 269–322. 1996. A celebration of John F. Nash, Jr.

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

[39] P. Sarnak. Values at integers of binary quadratic forms. In *Harmonic analysis and number theory (Montreal, PQ, 1996)*, volume 21 of *CMS Conf. Proc.*, pages 181–203. Amer. Math. Soc., Providence, RI, 1997.

[40] N. V. Smirnov. Approximate laws of distribution of random variables from empirical data. *Uspehi Matem. Nauk*, 10:179–206, 1944.

[41] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Translated from the second French edition (1995) by C. B. Thomas.

[42] E. C. Titchmarsh. *The theory of the Riemann zeta-function*. The Clarendon Press, Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown.

Department of Mathematics, KTH, SE-100 44 Stockholm, Sweden
Email address: `kurlberg@math.kth.se`

Department of Mathematics, King's College London, London WC2R 2LS, UK
Email address: `steve.lester@kcl.ac.uk`