MULTIVARIATE NORMAL DISTRIBUTION FOR INTEGRAL POINTS ON VARIETIES

DANIEL EL-BAZ, DANIEL LOUGHRAN, AND EFTHYMIOS SOFOS

ABSTRACT. Given a variety with coefficients in $\mathbb{Z}$, we study the distribution of the number of primes dividing the coordinates as we vary an integral point. Under suitable assumptions, we show that this has a multivariate normal distribution. We generalise this to more general Weil divisors, where we obtain a geometric interpretation of the covariance matrix. For our results we develop a version of the Erdős–Kac theorem that applies to fairly general integer sequences and does not require a positive exponent of level of distribution.

Contents

1. Introduction 1
2. A multivariate Erdős–Kac theorem 7
3. Application to integral points 25
4. Examples 28
References 37

1. Introduction

1.1. **Erdős–Kac.** To study the prime factorisation of a non-zero integer m, Erdős and Kac [12] considered the distribution of the function

$$\omega(m) = \text{number of distinct primes } p \text{ such that } p \text{ divides } m.$$

They showed that $\omega(m)$ behaves like a normal distribution with mean $\log\log m$ and variance $\log\log m$. More precisely, let $\Omega_B = \{m \in \mathbb{N} : m \leqslant B\}$ be equipped with the uniform probability measure for $B \geqslant 1$. Then as $B \to \infty$ the sequence of random variables

$$\Omega_B \to \mathbb{R}, \quad m \mapsto \frac{\omega(m) - \log\log B}{\sqrt{\log\log B}}$$

converges in distribution to the normal distribution with mean 0 and variance 1. Their work is a foundational result in probabilistic number theory and opened up many new research directions; we refer to the paper [15] and the references therein for various generalisations.

In our paper we study prime divisors of integers in sparse sequences, with an emphasis on solutions to Diophantine equations. A very special case of our results is as follows.

Theorem 1.1. *Let $f \in \mathbb{Z}[x_1, \dots, x_n]$ be a non-singular homogeneous polynomial with $n > (\deg(f) - 1)2^{\deg(f)}$. Let $\Omega_B = \{\mathbf{x} \in \mathbb{Z}^n : f(\mathbf{x}) = 0, \max_i |x_i| \leqslant B, \gcd(x_1, \dots, x_n) = 1\}$ be equipped with the uniform probability measure. If $f(\mathbf{x}) = 0$ has a non-trivial integer solution, then as $B \to \infty$ the random vectors*

$$\Omega_B \to \mathbb{R}^n, \quad \mathbf{x} = (x_1, \dots, x_n) \mapsto \left(\frac{\omega(x_1) - \log\log B}{\sqrt{\log\log B}}, \dots, \frac{\omega(x_n) - \log\log B}{\sqrt{\log\log B}} \right)$$

converge in distribution to the standard multivariate normal distribution on $\mathbb{R}^n$.

Date: August 20, 2021.
2010 *Mathematics Subject Classification.* 14G05; 60F05, 11N36.

By the *standard multivariate normal distribution*, we mean the multivariate normal distribution with zero mean vector and identity covariance matrix. We refer the reader to §1.5 for a reminder on multivariate normal distributions.

One knows how to count the number of solutions to the equation $f(\mathbf{x}) = 0$ using the circle method. Our motivation comes from trying to understand the more subtle arithmetic properties of the solutions, and is partly motivated by Sarnak's saturation problem [6], which asks whether there are solutions with coordinates being prime or almost prime.

Theorem 1.1 shows that the coordinate x_i typically has $\log\log|x_i|$ prime factors. Moreover, it compares the numbers of prime factors of different coordinates; the fact that we obtain the identity covariance matrix means that the number of prime factors of different coordinates is 'uncorrelated', something which is not a priori obvious. We have a purely geometric interpretation of this phenomenon, which we explain in more detail later (Theorem 1.8).

We are only aware of a few papers in the literature in probabilistic number theory which deal with a multivariate distribution: LeVeque [20, §4] on $(\omega(m), \omega(m+1))$ (stated by Erdős without proof [13]), Halberstam [17] again, on $(\omega(m), \omega(m+1))$, and Tanaka [26] on the distribution of $(\omega(f_1(m)), \dots, \omega(f_n(m)))$, where f_i are restricted to be pairwise coprime integer univariate polynomials. These can all be obtained as special cases of our most general result on a multivariate version of the Erdős–Kac theorem for integer sequences satisfying certain hypotheses (see §2, in particular Theorem 2.5). This more general result allows one to prove a general version of Tanaka's result, with no restrictions on f_i and, furthermore, to replace ω by any strongly additive function in Theorem 1.1. It may be viewed as a multidimensional version of Billingsley's work [1, §3].

1.2. **Distribution of the prime divisors of the coordinates.**

Let $X \subset \mathbb{P}^{n-1}_{\mathbb{Q}}$ be a projective variety over $\mathbb{Q}$. For $x \in \mathbb{P}^{n-1}(\mathbb{Q})$ we choose a representative $\mathbf{x} \in \mathbb{Z}^n$ with $\gcd(x_1, \dots, x_n) = 1$ such that $x = (x_1 : \cdots : x_n)$. Recall that the naive height of x is defined through $H(x) = \max\{|x_1|, \dots, |x_n|\}$. We are interested in the distribution of $\omega(x_i)$, which only depends on $x \in \mathbb{P}^{n-1}(\mathbb{Q})$ and is well-defined providing $x_i \neq 0$.

1.2.1. *Complete intersections.* For $R \geqslant 1$ and $1 \leqslant i \leqslant R$, let $f_i \in \mathbb{Z}[X_1, \dots, X_n]$ be homogeneous of the same degree D. The Birch rank, denoted by $\mathfrak{B}(\mathbf{f})$, is defined to be the codimension of the affine variety in $\mathbb{C}^n$ given by

$$\operatorname{rk}\left(\left(\frac{\partial f_i(\mathbf{x})}{\partial x_j}\right)_{1 \leqslant i \leqslant R, 1 \leqslant j \leqslant n}\right) < R. \tag{1.1}$$

Theorem 1.2. *Let $X \subset \mathbb{P}^{n-1}_{\mathbb{Q}}$ be the complete intersection given by $f_1 = \ldots = f_R = 0$ as above and let $\Omega_B = \{x \in X(\mathbb{Q}) : H(x) \leqslant B, x_1 \cdots x_n \neq 0\}$ be equipped with the uniform probability measure. Assume that X is smooth and $\mathfrak{B}(\mathbf{f}) > 2^{D-1}(D-1)R(R+1)$. If $X(\mathbb{Q}) \neq \varnothing$ then as $B \to \infty$ the random vectors*

$$\Omega_B \to \mathbb{R}^n, \quad x = (x_1 : \ldots : x_n) \mapsto \left(\frac{\omega(x_1) - \log\log B}{\sqrt{\log\log B}}, \dots, \frac{\omega(x_n) - \log\log B}{\sqrt{\log\log B}}\right)$$

converge in distribution to the standard multivariate normal distribution on $\mathbb{R}^n$.

1.2.2. *Homogeneous spaces.* Another class of examples to which our main result applies is given by certain symmetric varieties in affine space. We defer the precise definition of this class to §4.2 and instead present our results for two explicit families of such varieties.

Let Q be a non-degenerate, indefinite integral quadratic form in $n \geqslant 3$ variables. For each $k \in \mathbb{Z}\backslash\{0\}$, we consider the variety

$$L_k : \qquad Q(\mathbf{x}) = k \quad \subset \mathbb{A}^n_{\mathbb{Z}}$$

equipped with the usual height function $H(\mathbf{x}) = \max_i |x_i|$.

Theorem 1.3. *Let $k \in \mathbb{Z}\backslash\{0\}$ and $n \geqslant 3$. If $n = 3$, assume that $-k\operatorname{disc}(Q)$ is not a perfect square. Let $\Omega_B = \{\mathbf{x} \in L_k(\mathbb{Z}) : H(\mathbf{x}) \leqslant B, x_1 \cdots x_n \neq 0\}$ be equipped with the uniform probability measure. If $L_k(\mathbb{Z}) \neq \varnothing$ then as $B \to \infty$ the random vectors*

$$\Omega_B \to \mathbb{R}^n, \quad \mathbf{x} \mapsto \left(\frac{\omega(x_1) - \log\log B}{\sqrt{\log\log B}}, \dots, \frac{\omega(x_n) - \log\log B}{\sqrt{\log\log B}} \right)$$

converge in distribution to the standard multivariate normal distribution on $\mathbb{R}^n$.

For $n \geqslant 2$ and $k \in \mathbb{Z}\backslash\{0\}$, consider the variety

$$V_{n,k} : \qquad \det(M) = k \quad \subset \mathbb{A}^{n^2}_{\mathbb{Z}},$$

where det denotes the determinant, viewed as a homogeneous polynomial of degree n (in particular, $V_{n,1} = \mathrm{SL}_n$).

Theorem 1.4. *Let $k \in \mathbb{Z}\backslash\{0\}$, $n \geqslant 2$ and $\Omega_B = \{M = (m_{i,j}) \in V_{n,k}(\mathbb{Z}) : H(M) \leqslant B, m_{i,j} \neq 0\}$ be equipped with the uniform probability measure. As $B \to \infty$ the random vectors*

$$\Omega_B \to \mathbb{R}^{n^2} \quad M = (m_{i,j}) \mapsto \left(\frac{\omega(m_{i,j}) - \log\log B}{\sqrt{\log\log B}} \right)_{i,j \in \{1,\dots,n\}}$$

converge in distribution to the standard multivariate normal distribution on $\mathbb{R}^{n^2}$.

1.2.3. *Conics.* In all the above cases, we obtained the identity covariance matrix, meaning that the random variables given by each coordinate are independent. In the case of plane conics however, we obtain a very different result. Firstly, we need to choose a different normalisation, as it turns out that $\omega(x_i)$ need not have average order $\log\log B$ in general. Secondly, there may be non-trivial correlations.

Theorem 1.5. *Let $C \subset \mathbb{P}^2_{\mathbb{Q}}$ be a smooth plane conic with $C(\mathbb{Q}) \neq \varnothing$. Let $\Omega_B = \{x \in C(\mathbb{Q}) : H(x) \leqslant B, x_1x_2x_3 \neq 0\}$ be equipped with the uniform probability measure. Let $c_{i,j}$ denote the number of common irreducible components (counted without multiplicity) of the divisors $x_i = 0$ and $x_j = 0$ on C. Then the random vectors*

$$\Omega_B \to \mathbb{R}^3, \quad x \mapsto \left(\frac{\omega(x_1) - c_{1,1}\log\log B}{\sqrt{c_{1,1}\log\log B}}, \frac{\omega(x_2) - c_{2,2}\log\log B}{\sqrt{c_{2,2}\log\log B}}, \frac{\omega(x_3) - c_{3,3}\log\log B}{\sqrt{c_{3,3}\log\log B}} \right)$$

converge in distribution to a central multivariate normal distribution with covariance matrix whose (i,j)-entry is $(c_{i,j}/\sqrt{c_{i,i}c_{j,j}})$.

Example 1.6.

(1) Take $C : x_1^2 + x_2^2 = x_3^2$. The divisors $x_1 = 0$ and $x_2 = 0$ have two irreducible components while $x_3 = 0$ is irreducible, and these have no components in common. Hence we just obtain the identity matrix for the covariance matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so there are no correlations between the number of prime divisors of the coordinates.

(2) Take $C : x_1x_2 + x_2x_3 + x_3x_1 = 0$. Every divisor $x_i = 0$ is a union of two rational points, and they each contain one point in common. We obtain the covariance matrix

$$\begin{pmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{pmatrix}.$$

This is not the identity matrix, which is reflected by the fact that there is a non-trivial relation between the prime divisors of x_i and x_j, as is clear from the equation.

(3) Take $C : x_1x_2 = x_3^2$. Here $x_1 = 0$ and $x_2 = 0$ are both irreducible and are the irreducible components of $x_3 = 0$ (we do not count irreducible components with multiplicity). The covariance matrix is therefore

$$\begin{pmatrix} 1 & 0 & 1/\sqrt{2} \\ 0 & 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 1 \end{pmatrix}.$$

This matrix is singular; this means that the associated probability measure is supported on a proper linear subspace of $\mathbb{R}^3$. From the equation it is also clear that the prime divisors of x_3 are completely determined by those of x_1 and x_2.

The example with singular covariance matrix is essentially the only example for conics.

Theorem 1.7. *Let $C \subset \mathbb{P}^2_{\mathbb{Q}}$ be a smooth plane conic for which the associated covariance matrix in Theorem 1.5 is singular. Then, up to permuting coordinates, the conic has the equation $x_1x_2 = cx_3^2$ for some $c \in \mathbb{Q}$.*

1.3. **A geometric reformulation.** We now come to our most general results. To state them we require some notation.

Let $X \subset \mathbb{P}^d_{\mathbb{Q}}$ be a quasi-projective variety over $\mathbb{Q}$. Then the usual height on projective space induces a height function $H : X(\mathbb{Q}) \to \mathbb{R}_{>0}$. Let $\mathcal{X}$ be a choice of model for X over $\mathbb{Z}$. Then the model allows us to define the set of integral point $\mathcal{X}(\mathbb{Z})$, which is naturally a subset of $X(\mathbb{Q}) = \mathcal{X}(\mathbb{Q})$.

We assume that $\mathcal{X}$ and the height H satisfy the following properties. There exists a bound $A > 0$ and constants $M, \eta > 0$ such that for $Q \in \mathbb{N}$ square-free with $\gcd(Q, \prod_{p \leqslant A} p) = 1$ and for $\Upsilon \subset \mathcal{X}(\mathbb{Z}/Q\mathbb{Z})$, we have

$$\frac{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B, x \bmod Q \in \Upsilon\}}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} = \frac{\#\Upsilon}{\#\mathcal{X}(\mathbb{Z}/Q\mathbb{Z})} + O(Q^M B^{-\eta}) \tag{1.2}$$

as $B \to \infty$. We call this condition *effective equidistribution*, as it says that the solutions are equidistributed in congruence classes with an explicit error term. For our applications it does not matter how large M is, since we will take Q with $\log Q = o(\log B)$. This property holds for example for affine space, projective space [21, Prop. 2.1], Birch range complete intersections (see §4.1) and a general class of symmetric varieties (see §4.2).

Let $\mathcal{Z} \subset \mathcal{X}$ be a closed subscheme. For $x \in \mathcal{X}(\mathbb{Z})\backslash\mathcal{Z}(\mathbb{Z})$, we define

$$\omega_{\mathcal{Z}}(x) = \#\{p : x \bmod p \in \mathcal{Z}(\mathbb{F}_p)\}. \tag{1.3}$$

The condition $x \notin \mathcal{Z}(\mathbb{Z})$ is easily seen to imply that the number of such primes is finite, hence this is well-defined. Note that $\omega_{\mathcal{Z}}(x) = \omega_{\mathcal{Z}_{\text{red}}}(x)$ where $\mathcal{Z}_{\text{red}}$ denotes the reduced subscheme underlying $\mathcal{Z}$. In particular, we may always assume that $\mathcal{Z}$ is reduced.

Taking $\mathcal{X} = \mathbb{A}^1_{\mathbb{Z}}$ and $\mathcal{Z}$ the origin, this recovers the classical number of prime divisors function ω used in §1.1. Taking $\mathcal{Z}$ to be the coordinate hyperplane $x_i = 0$, we obtain the function $\omega(x_i)$ studied in §1.2. This is an important change of viewpoint, which makes clear that $\omega(x_i)$ actually has an intrinsic geometric definition. A natural question is how the geometry affects the distribution of $\omega_{\mathcal{Z}}$; as we shall soon see, the geometry determines everything and there is a natural geometric interpretation for all the results in §1.2.

In Proposition 3.1 we study the average order of this function for a flat closed subscheme $\mathcal{Z} \subset \mathcal{X}$. If $\mathcal{Z}$ is *not* a divisor, then $\omega_{\mathcal{Z}}$ has constant average order. The more interesting case is where $\mathcal{Z} = \mathcal{D}$ is a divisor: here $\omega_{\mathcal{D}}$ has average order $c_{\mathcal{D}} \log\log B$, where $c_{\mathcal{D}}$ denotes the number of irreducible components of $\mathcal{D}$. In particular this behaves strikingly like the usual number of primes divisors of an integer.

Our main theorem on integral points is an analogue of Erdős–Kac's result for our function $\omega_{\mathcal{D}}$. However, given that there are many possible choices for $\mathcal{D}$ it is also natural to simultaneously consider finitely many $\mathcal{D}$, and study the correlations between these divisors. The result we obtain shows that there is in fact a multivariate normal distribution, whose covariance matrix is given explicitly in terms of the geometry of the divisors.

Theorem 1.8. *Let $X \subset \mathbb{P}^d_{\mathbb{Q}}$ be a quasi-projective variety with induced height function H and $\mathcal{X}$ a choice of model for X over $\mathbb{Z}$ which satisfy* (1.2). *Let $\Omega_B = \{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}$ be equipped with the uniform probability measure.*

Let $D_1, \dots, D_n \subset X$ be a collection of reduced divisors, $\mathcal{D}_i$ their closures in $\mathcal{X}$ and $\mathcal{D}$ the union of the $\mathcal{D}_i$. Let $c_{i,j}$ denote the number of common irreducible components of D_i and D_j. Then as $B \to \infty$, the random vectors

$$\Omega_B \backslash \mathcal{D}(\mathbb{Z}) \to \mathbb{R}^n, \quad x \mapsto \left(\frac{\omega_{\mathcal{D}_1}(x) - c_{1,1}\log\log B}{\sqrt{c_{1,1}\log\log B}}, \dots, \frac{\omega_{\mathcal{D}_n}(x) - c_{n,n}\log\log B}{\sqrt{c_{n,n}\log\log B}} \right)$$

converge in distribution to a central multivariate normal distribution with covariance matrix whose (i,j)-entry is $(c_{i,j}/\sqrt{c_{i,i}c_{j,j}})$.

Moreover, let $\mathcal{R} = \langle D_1, \dots, D_n \rangle \subset \operatorname{Div} X$ be the group of divisors of X generated by the D_i and let r be the rank of $\mathcal{R}$. Then the covariance matrix has rank r.

Theorem 1.8 gives a much more general setting than the results mentioned earlier in the introduction; it allows one to also obtain results where the x_i are replaced by arbitrary polynomials. For example, we obtain the following immediate corollary of Theorem 1.8.

Corollary 1.9. *Let $f_1, \dots, f_n \in \mathbb{Z}[x_1, \dots, x_d]$ and $c_{i,j}$ denote the number of irreducible primitive non-constant polynomials f with $f \mid f_i$ and $f \mid f_j$. Let $\Omega_B = \{\mathbf{x} \in \mathbb{Z}^d : H(\mathbf{x}) \leqslant B, f_1(\mathbf{x}) \cdots f_n(\mathbf{x}) \neq 0\}$ be equipped with the uniform probability measure. As $B \to \infty$, the random vectors*

$$\Omega_B \to \mathbb{R}^n, \quad \mathbf{x} \mapsto \left(\frac{\omega(f_i(\mathbf{x})) - c_{i,i}\log\log B}{\sqrt{c_{i,i}\log\log B}} \right)_{i=1,\dots,n}$$

converge in distribution to a central multivariate normal distribution with covariance matrix whose (i,j)-entry is $(c_{i,j}/\sqrt{c_{i,i}c_{j,j}})$.

Corollary 1.9 generalises numerous special cases already known in the literature. The case $n = d = 1$ and f_1 is irreducible is due to Halberstam [17, Thm. 3]. The case $n = 2, d = 1$ and $f_1(x) = x, f_2(x) = x + 1$ is also due to Halberstam [17, Thm. 1] and LeVeque [20, §4]. The case $n = 1$ and f_1 is a product of geometrically irreducible polynomials is due to Xiong [29, Thm. 1]. The case $d = 1$ and the f_i pairwise coprime is due to Tanaka [26]. All these cases either concern a univariate normal distribution, or a multivariate distribution with identity covariance matrix. Our results give a unified proof of all these special cases, and apply in much greater generality.

Remark 1.10. The covariance matrix in Theorem 1.8 equals the identity matrix if and only if each pair of distinct divisors D_i and D_j have no irreducible component in common.

Remark 1.11. Our assumption (1.2) implies that the map $\mathcal{X}(\mathbb{Z}) \to \mathcal{X}(\mathbb{F}_p)$ is surjective for all but finitely many primes p; this may be viewed as a weak form of strong approximation. However (1.2) does not imply strong approximation, since our condition may fail at finitely many primes and we do need require any information modulo higher powers of p.

Remark 1.12. Our method shows that is is possible to replace $\mathcal{X}(\mathbb{Z})$ in (1.2) by the assumption that there exists some subset $\Omega \subset \mathcal{X}(\mathbb{Z})$ which satisfies (1.2). In particular, one can also consider cases in which there are accumulating subvarietes or thin subsets.

Remark 1.13. Let us emphasise that Theorem 1.8 makes clear that it is really the geometric properties of the chosen divisors, rather than the geometry of the underlying variety, which determines the covariance matrix. For example, let $X \subset \mathbb{P}^n$ be as in Theorem 1.2, with coordinates x_i. We apply the d-uple embedding $X \subset \mathbb{P}^n \subset \mathbb{P}^N$ for some $d > 1$, where $N = \binom{n+d}{d} - 1$ and we take the coordinates y_i on $\mathbb{P}^N$. Then applying Theorem 1.8 to X with respect to coordinate hyperplanes $y_i = 0$, we obtain a covariance matrix which is no longer diagonal; indeed, this is exactly the same as applying Theorem 1.8 to the divisors $x_0^{d_0} \cdots x_n^{d_n} = 0$, running over all monomials of degree d, whence it is easily seen that the covariance matrix is no longer diagonal.

1.4. **Outline of the paper.** In §2 we state our most general theorem (Theorem 2.5), which is a multivariate version of the Erdős–Kac theorem for integer sequences satisfying certain hypotheses, and may be viewed as a multidimensional version of Billingsley's work [1, §3]. The statement is very involved, in order to allow for the greatest flexibility for applications. To help the reader, we therefore state a simplified version first in Theorem 2.1. This section is dedicated to the proofs of Theorems 2.1 and 2.5, and is the technical heart of the paper.

In §3 we prove Theorem 1.8 using Theorem 2.1. The final §4 concerns various example applications of Theorem 1.8 to proving the remaining results stated in the introduction. We finish with an example of a cubic surface to which our method does not apply, but for which we expect an analogue of our results to hold.

1.5. Notation and conventions.

Number theory. We say that a function $g : \mathbb{N}^n \to \mathbb{C}$ is *multiplicative* if for all $\mathbf{a}, \mathbf{b} \in \mathbb{N}^n$ we have

$$g(a_1b_1, \ldots, a_nb_n) = g(\mathbf{a})g(\mathbf{b}), \quad \text{if } \gcd(a_1a_2\cdots a_n, b_1b_2\cdots b_n) = 1. \tag{1.4}$$

For a prime p, we denote by ν_p the p-adic valuation.

Algebraic geometry. Let X be a variety over $\mathbb{Q}$. A *model* of X over $\mathbb{Z}$ is a finite type scheme $\mathcal{X} \to \operatorname{Spec} \mathbb{Z}$ together with a choice of isomorphism $X \cong \mathcal{X}_{\mathbb{Q}}$.

Probability theory.

Definition 1.14. A random vector $(X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ has a *multivariate normal distribution* if for every $\mathbf{t} \in \mathbb{R}^n$ the random variable $\sum_{i=1}^n t_i X_i$ has a univariate normal distribution.

Note that for some $\mathbf{t}$ the random variable $\sum_{i=1}^n t_i X_i$ may follow a Dirac delta distribution; by convention one views this as a univariate normal distribution with variance 0. In this case, the associated probability measure will be supported on some affine subspace of $\mathbb{R}^n$.

A multivariate normal distribution is uniquely determined by its mean vector $\boldsymbol{\mu}$ and its covariance matrix $\boldsymbol{\Sigma}$, whose (i, j)-entry is $\mathrm{Cov}[X_i, X_j]$. We denote by $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ the associated probability measure on $\mathbb{R}^n$. A *central* multivariate normal distribution is one with zero mean vector. A *standard* multivariate normal distribution is one with zero mean vector and covariance matrix given by the identity matrix.

We use the notation $\Rightarrow$ to denote convergence in distribution of a sequence of random variables, i.e. if the corresponding sequence of probability measures convergences weakly.

1.6. **Acknowledgements.** We thank Carlo Pagano and Zeev Rudnick for helpful comments and suggestions. The first-named author is supported by the Austrian Science Fund (FWF), projects F-5512 and Y-901. The second-named author is supported by EPSRC grant EP/R021422/2. The first and third-named authors acknowledge the support of the Max Planck Institute for Mathematics, where a large part of this work was carried out. We wish to thank the anonymous referee for helpful remarks that helped to improve various parts of the paper.

2. A multivariate Erdős–Kac theorem

In this section we provide a multidimensional generalisation of the Erdős–Kac theorem for general integer sequences. Our main result (Theorem 2.5), proves that multiple additive functions evaluated at integer sequences defined on an arbitrary set and well-distributed in arithmetic progressions of very small moduli obey a multivariate normal distribution. We first give a simplified version (Theorem 2.1) which is sufficient for many applications, to help ease the reader into the more general technical statement.

2.1. **Simplified version of the main theorem.** Let Ω be an infinite set and assume that we are given a function $h : \Omega \to \mathbb{R}_{\geqslant 0}$ with

$$N(B) \text{ finite for all } B \geqslant 0, \quad \text{where } N(B) := \#\{a \in \Omega : h(a) \leqslant B\}. \tag{2.1}$$

Note that as Ω is infinite we have $N(B) \to \infty$. Moreover (2.1) implies that Ω is countable. For each $B \geqslant 0$ we equip the set Ω with the structure of a probability space using the discrete σ-algebra and probability measure

$$\mathbf{P}_B[S] := \frac{\#\{a \in S : h(a) \leqslant B\}}{\#\{a \in \Omega : h(a) \leqslant B\}}, \quad S \subseteq \Omega.$$

Note that this measure is supported on the finite set $\#\{a \in \Omega : h(a) \leqslant B\}$, where it induces the uniform measure. Next, we assume that we are given $n \in \mathbb{N}$ and a function

$$m : \Omega \to \mathbb{N}^n, a \in \Omega \mapsto (m_1(a), \dots, m_n(a)). \tag{2.2}$$

We are interested in studying the distribution of the vector

$$(\omega(m_1(a)), \dots, \omega(m_n(a))).$$

As with the classical Erdős–Kac theorem, we need to normalise by suitable factors first. We have to assume some kind of regularity among the values of $m_i(a)$, namely, that there exists $A \in \mathbb{R}$ such that the following limit exists for all $\mathbf{d} \in \mathbb{N}^n$ satisfying $p \mid d_1 \cdots d_n \Rightarrow p > A$,

$$\lim_{B\to+\infty} \frac{\#\left\{a \in \Omega : h(a) \leqslant B, d_i \mid m_i(a) \,\forall\, 1 \leqslant i \leqslant n\right\}}{N(B)} =: g(\mathbf{d}). \tag{2.3}$$

The reason for assuming (2.3) only for moduli without small prime factors is that in certain situations it is convenient to ignore small 'bad' primes. We furthermore assume that

$$g \text{ is multiplicative in the sense of (1.4)} \tag{2.4}$$

and extend g to $\mathbb{N}^n$ by setting it equal to 0 for $\mathbf{d}$ such that $d_1 \cdots d_n$ has a prime factor $p \leqslant A$. For any $1 \leqslant i, j \leqslant n$ we let

$$g_i(d) := g(1, \ldots, 1, \underset{\substack{\uparrow \\ i}}{d}, 1, \ldots, 1) \text{ and } g_{i,j}(d) := g(1, \ldots, 1, \underset{\substack{\uparrow \\ i}}{d}, 1, \ldots, 1, \underset{\substack{\uparrow \\ j}}{d}, 1, \ldots, 1).$$

We now assume that for every $1 \leqslant i \leqslant n$ we have

$$\sum_{p>T} g_i(p)^2 = O\left(\frac{1}{\log T}\right) \text{ and } \sum_{p\leqslant T} g_i(p) = c_i \log\log T + c_i' + O\left(\frac{1}{\log T}\right), \tag{2.5}$$

for some $c_i > 0, c_i' \in \mathbb{R}$. This assumption is highly typical and usually met in sieve theory problems as it corresponds to a sieve of 'dimension' c_i. In light of (2.3) the sum $\sum_{p\leqslant m_i(a)} g_i(p)$ should be thought of as approximating the expected value of $\omega(m_i(a))$ as one samples over suitably many $a \in \Omega$.

The main arithmetic input in our theorem is a statement regarding the speed of convergence in (2.3). Namely, we define $\mathcal{R}(\mathbf{d}, B)$ for each $\mathbf{d} \in \mathbb{N}^n$ and $B \geqslant 1$ via

$$\mathcal{R}(\mathbf{d}, B) := \#\left\{a \in \Omega : h(a) \leqslant B, d_i \mid m_i(a) \,\forall 1 \leqslant i \leqslant n\right\} - g(\mathbf{d})N(B). \tag{2.6}$$

We demand that $\mathcal{R}(\mathbf{d}, B)$ is asymptotically smaller than $N(B)$ for most $\mathbf{d}$ that are smaller than the 'typical size' of the $m_i(a)$. To make this notion precise, we first call $\mathcal{F}(B)$ the typical size of $\max_{1\leqslant i\leqslant n} m_i(a)$, namely we assume there exists a function $\mathcal{F} : \mathbb{R}_{\geqslant 1} \to \mathbb{R}$ with

$$\lim_{B\to\infty} \frac{1}{N(B)} \#\left\{a \in \Omega : h(a) \leqslant B, \max_{1\leqslant i\leqslant n} m_i(a) \leqslant \mathcal{F}(B)\right\} = 1. \tag{2.7}$$

It will turn out that the other assumptions in our set-up ensure that $\lim_{B\to\infty} \mathcal{F}(B) = +\infty$. Secondly, we assume that the sequences $m_i(a)$ are well-distributed in arithmetic progressions whose modulus is small compared to $\mathcal{F}(B)$. Namely, let

$$\varepsilon(B) := \frac{\log\log\log \mathcal{F}(B)}{\sqrt{\log\log \mathcal{F}(B)}} \tag{2.8}$$

and assume that for the same $A \in \mathbb{R}$ as above, the following estimate is valid for all $\gamma > 0$

$$\sum_{\substack{\mathbf{d}\in\mathbb{N}^n \\ |\mathbf{d}|\leqslant \mathcal{F}(B)^{\varepsilon(B)} \\ p\mid d_1\cdots d_n \Rightarrow p>A}} \mu(d_1)^2 \cdots \mu(d_n)^2 |\mathcal{R}(\mathbf{d}, B)| \ll_\gamma \frac{N(B)}{(\log\log \mathcal{F}(B))^\gamma} \tag{2.9}$$

with an implied constant that is independent of B. Assumption (2.9) is the main arithmetic input needed in our main results (see Remark 2.2). Define the function $\mathbf{K} : \Omega \to \mathbb{R}^n$ via

$$\mathbf{K}(a) := \left(\frac{\omega(m_1(a)) - c_1 \log\log \mathcal{F}(B)}{\sqrt{c_1 \log\log \mathcal{F}(B)}}, \ldots, \frac{\omega(m_n(a)) - c_n \log\log \mathcal{F}(B)}{\sqrt{c_n \log\log \mathcal{F}(B)}} \right). \tag{2.10}$$

This is the promised normalisation. Our result is as follows.

Theorem 2.1. *Let $n \in \mathbb{N}$ and assume that we are given a set Ω, a real number A and functions $h, m, g, \mathcal{F}$ such that* (2.1)*,* (2.2)*,* (2.3)*,* (2.4)*,* (2.5)*,* (2.7) *and* (2.9) *hold. Furthermore, assume that for every $1 \leqslant i, j \leqslant n$ the following limit exists,*

$$\lim_{T \to +\infty} \frac{\sum_{p \leqslant T} g_{i,j}(p)}{(\sum_{p \leqslant T} g_i(p))^{1/2} (\sum_{p \leqslant T} g_j(p))^{1/2}}. \tag{2.11}$$

Then the random vectors

$$(\Omega, \mathbf{P}_B) \to \mathbb{R}^n, \quad a \mapsto \mathbf{K}(a), \tag{2.12}$$

converge in distribution as $B \to \infty$ to a central multivariate normal distribution with covariance matrix $\mathbf{\Sigma}$ whose (i, j)-entry is the limit (2.11)*.*

There are three noteworthy aspects in Theorem 2.1. Firstly, the simplest case with $n = 1$ applies to functions defined on a *general* set Ω, hence it recovers normal distribution results related to irreducible polynomials [17, Thm. 3], values of irreducible polynomials at primes [18] and entries of matrices [11]. It also applies to new situations, such as the coordinates of integer zeros of affine algebraic varieties that do not necessarily have a group structure.

Secondly, Theorem 2.1 studies *multidimensional* normal laws for arithmetic functions. The only related example that we could find in the literature is due to Halberstam [17, Thm. 1] and LeVeque [20, §4] regarding $(\omega(m), \omega(m + 1))$ and its generalisation given by Tanaka [26] regarding $(\omega(f_1(m)), \ldots, \omega(f_n(m)))$ for non-constant integer irreducible polynomials f_i that are relatively coprime. These results are recovered by our theorem by taking $\Omega = \mathbb{N}, m_i(a) = f_i(a)$ for $1 \leqslant i \leqslant n$, and the covariance matrix is the $n \times n$ identity matrix.

Thirdly, the covariance matrix is the identity if and only if the sequences $\omega(m_i(a))$ and $\omega(m_j(a))$ are 'uncorrelated' for all $i \neq j$. Such a phenomenon is however not present in many situations (such as the prime factors of coordinates of affine algebraic varieties) and one must therefore obtain a general Erdős–Kac law that would apply to situations with non-vanishing correlations. This is the most important new aspect of Theorem 2.1, namely, that it covers multivariate normal distributions with *arbitrary* covariance matrix.

Remark 2.2. Assumption (2.9) resembles a *level of distribution* condition in sieve theory. In typical situations one takes $\mathcal{F}(B) = N(B)^c$ for some fixed $c > 0$, where the size condition on $\mathbf{d}$ becomes $|\mathbf{d}| \leqslant \mathcal{F}(B)^{\varepsilon(B)} = N(B)^{o(1)}$. This is much lighter than the usually stricter assumption in classical sieve theory problems, where a positive exponent of level of distribution is required, i.e. one requires the same error term but with the summation over $\mathbf{d}$ with $|\mathbf{d}| \leqslant N(B)^\alpha$ for some fixed $\alpha > 0$. Note that if there exist $\eta > 0$ and $M > 0$ such that

$$\frac{\#\{a \in \Omega : h(a) \leqslant B, d_i \mid m_i(a) \, \forall 1 \leqslant i \leqslant n\}}{\#\{a \in \Omega : h(a) \leqslant B\}} = g(\mathbf{d}) + O\left(N(B)^{-\eta} (\max_{1 \leqslant i \leqslant n} d_i)^M \right)$$

and if $\mathcal{F}(B) = N(B)^c$, then (2.9) holds due to the estimate

$$\sum_{|\mathbf{d}| \leqslant \mathcal{F}(B)^{\varepsilon(B)}} |\mathcal{R}(\mathbf{d}, B)| \ll N(B)^{1-\eta} \sum_{|\mathbf{d}| \leqslant \mathcal{F}(B)^{\varepsilon(B)}} (\max_{1 \leqslant i \leqslant n} d_i)^M \ll N(B)^{1-\eta} \mathcal{F}(B)^{\varepsilon(B)(M+n)}.$$

Since $\varepsilon(B) = o(1)$, this is $\ll N(B)^{1-\eta/2}$, which, for every $\gamma > 0$ is

$$o\left(\frac{N(B)}{(\log\log N(B))^{\gamma}}\right) = O\left(\frac{N(B)}{(\log\log \mathcal{F}(B))^{\gamma}}\right)$$

owing to the equality $\mathcal{F}(B) = N(B)^c$.

2.2. **The main theorem.** We now state the main technical result in the present paper; it is a general version of Theorem 2.1. Let Ω be an infinite set and assume that for every $B \in \mathbb{R}_{\geqslant 1}$ we are given a function $\chi_B : \Omega \to \mathbb{R}_{\geqslant 0}$ such that

$$B \geqslant 1 \Rightarrow \{a \in \Omega : \chi_B(a) > 0\} \text{ finite.} \tag{2.13}$$

In applications the function $\chi_B(x)$ will either denote the characteristic function of elements x having 'height' bounded by B or it will be a smooth 'weight' function of the form $w(x/B)$. We also demand that

$$\lim_{B\to+\infty} \sum_{a\in\Omega} \chi_B(a) = +\infty. \tag{2.14}$$

For each $B \geqslant 0$ we equip the set Ω with the structure of a probability space using the discrete σ-algebra and probability measure

$$\mathbf{P}_B\left[S\right] := \frac{\sum_{a\in S} \chi_B(a)}{\sum_{a\in\Omega} \chi_B(a)}, \quad S \subseteq \Omega.$$

Assume that $M : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is any function satisfying

$$\lim_{B\to+\infty} \frac{\sum_{a\in\Omega} \chi_B(a)}{M(B)} = 1. \tag{2.15}$$

Next, we assume that we are given $n \in \mathbb{N}$ and a function

$$m : \Omega \to \mathbb{N}^n, a \in \Omega \mapsto (m_1(a), \ldots, m_n(a)). \tag{2.16}$$

We will find general assumptions which ensure that certain functions display Gaussian behaviour simultaneously for all i when evaluated at $m_i(a)$. We first need the following function g, that contains information on the divisors of typical values of $m_i(a)$.

Definition 2.3 (The density function g)**.** We assume that there exists $A \in \mathbb{R}$ such that the following limit exists for all $\mathbf{d} \in \mathbb{N}^n$ with $p \mid d_1 \cdots d_n \Rightarrow p > A$,

$$\lim_{B\to+\infty} \frac{1}{M(B)} \sum_{\substack{a\in\Omega \\ d_i \mid m_i(a)\, \forall 1\leqslant i\leqslant n}} \chi_B(a). \tag{2.17}$$

We define $g : \{\mathbf{d} \in \mathbb{N}^n : p \mid d_1 \cdots d_n \Rightarrow p > A\} \to \mathbb{R}$ as the value of this limit. We extend g to $\mathbb{N}^n$ by setting it equal to 0 for $\mathbf{d}$ such that $d_1 \cdots d_n$ has a prime factor $p \leqslant A$. Furthermore, we assume that

$$g \text{ is multiplicative in the sense of (1.4).} \tag{2.18}$$

Let us introduce the arithmetic functions whose values at $m_i(a)$ we shall study. These functions will be of the form $\sum_{p\mid m_i(a)} \theta_i(p)$, where the sum is taken over prime divisors p and $\theta_i(p)$ are bounded functions. These function clearly generalise ω as can be seen by taking $A = 0$ and $\theta_i(p) = 1$ for all p. To be precise, we assume that we are given functions $\theta_1, \ldots, \theta_n$ defined on the primes, taking values on $\mathbb{R}$ and that there exists $\Theta \in \mathbb{R}$ with

$$|\theta_i(p)| \leqslant \Theta \text{ for all } 1 \leqslant i \leqslant n \text{ and primes } p. \tag{2.19}$$

For any $S \subset \{1, \ldots, n\}$ and $b \in \mathbb{N}$ we define $g_S : \mathbb{N} \to \mathbb{R}$ via

$$g_S(b) := g\left(1 + (b-1)\mathbb{1}_S(1), \ldots, 1 + (b-1)\mathbb{1}_S(i), \ldots, 1 + (b-1)\mathbb{1}_S(n)\right), \tag{2.20}$$

i.e. we put b in position i if $i \in S$ and we put 1 otherwise. We furthermore define

$$\mathcal{M}_i(T) := \sum_{p \leqslant T} \theta_i(p) g_i(p), \quad (T \geqslant 0, 1 \leqslant i \leqslant n). \tag{2.21}$$

The function $\mathcal{M}_i(T)$ approximates the 'mean' of $\sum_{p | m_i(a)} \theta_i(p)$ as one samples over suitably many $a \in \Omega$. In addition to these means we shall also need to consider the analogous of 'variances' $\mathcal{V}_i(T)^2$, thus we let

$$\mathcal{V}_i(T) := \Big(\sum_{p \leqslant T} \theta_i(p)^2 g_i(p) \left(1 - g_i(p)\right) \Big)^{1/2}, \quad (T \geqslant 0, 1 \leqslant i \leqslant n). \tag{2.22}$$

We assume that for all $i = 1, \ldots, n$ we have

$$\lim_{T \to +\infty} \mathcal{V}_i(T) = +\infty. \tag{2.23}$$

Let us define the function $\mathbf{K} : \Omega \to \mathbb{R}^n$ via

$$\mathbf{K}(a) := \left(\frac{\left(\sum_{p | m_1(a)} \theta_1(p)\right) - \mathcal{M}_1(m_1(a))}{\mathcal{V}_1(m_1(a))}, \ldots, \frac{\left(\sum_{p | m_n(a)} \theta_n(p)\right) - \mathcal{M}_n(m_n(a))}{\mathcal{V}_n(m_n(a))} \right). \tag{2.24}$$

If $\mathcal{V}_i(m_i(a)) = 0$, then by convention we take the ith entry to be 1 (note that our later assumptions will imply that for any i the event $\mathcal{V}_i(m_i(a)) = 0$ has probability 0)

We will study the behaviour of the functions $\sum_{p | m_i(a)} \theta_i(p)$ simultaneously for all i and as a ranges over Ω. To make the notation easier in what follows we normalised these functions by first centering around their 'expected mean' $\mathcal{M}_i$ and then dividing by the 'standard deviation' $\mathcal{V}_i$. We define $\mathcal{R}(\mathbf{d}, B)$ for each $\mathbf{d} \in \mathbb{N}^n$ and $B \geqslant 0$ via

$$\mathcal{R}(\mathbf{d}, B) := \left(\sum_{\substack{a \in \Omega \\ d_i | m_i(a)\, \forall 1 \leqslant i \leqslant n}} \chi_B(a) \right) - g(\mathbf{d}) M(B). \tag{2.25}$$

Our result will hold if the size of $\mathcal{R}(\mathbf{d}, B)$ is relatively small compared to $\mathcal{M}_i(B)$ and $\mathcal{V}_i(B)$ as one averages over small $\mathbf{d}$. To make this precise we need the following piece of notation.

Definition 2.4 (Truncation pairs)**.** We say that a pair of functions $(\mathcal{F}, \psi)$ with $\mathcal{F} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ and $\psi : \mathbb{R} \to (0, 1]$ is a *truncation pair* if the following is satisfied. First

$$\lim_{B \to +\infty} \frac{1}{\sum_{a \in \Omega} \chi_B(a)} \sum_{\substack{a \in \Omega \\ m_i(a) \leqslant \mathcal{F}(B)\, \forall i}} \chi_B(a) = 1. \tag{2.26}$$

Next

$$\lim_{B \to +\infty} \mathcal{F}(B)^{\psi(B)} = +\infty, \tag{2.27}$$

$$\lim_{B \to +\infty} \frac{1}{\psi(B) \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} = 0, \quad \frac{\mathcal{M}_i(m_i(a)) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \Rightarrow 0 \tag{2.28}$$

and

$$\frac{\mathcal{V}_i(m_i(a))}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \Rightarrow 1. \tag{2.29}$$

Lastly, we assume that for every $k_1, \dots, k_n \in \mathbb{N}$ we have

$$\lim_{B \to +\infty} \frac{\prod_{i=1}^n \left(\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}) + \Theta\right)^{k_i}}{M(B) \prod_{i=1}^n \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^{k_i}} \sum_{\substack{\mathbf{d} \in \mathbb{N}^n \\ (2.31)}} (g(\mathbf{d}) |\mathcal{R}((1, \dots, 1), B)| + |\mathcal{R}(\mathbf{d}, B)|) = 0, \quad (2.30)$$

where the summation is over $\mathbf{d} \in \mathbb{N}^n$ with

$$\begin{cases} p \mid d_1 \cdots d_n \Rightarrow A < p \leqslant \mathcal{F}(B)^{\psi(B)}, \\ \mu(d_i)^2 = 1, \quad \forall 1 \leqslant i \leqslant n, \\ \omega(d_i) \leqslant k_i, \quad \forall 1 \leqslant i \leqslant n. \end{cases} \quad (2.31)$$

Theorem 2.5. *Assume that we are given $n \in \mathbb{N}$, an infinite set Ω, a function $M : \mathbb{R}_{\geqslant 1} \to \mathbb{R}_{\geqslant 0}$ and for all $B \in \mathbb{R}_{\geqslant 1}$ a function $\chi_B : \Omega \to \mathbb{R}_{\geqslant 0}$ such that for any $B \in \mathbb{R}_{\geqslant 1}$ the assumptions* (2.13)*,* (2.14) *and* (2.15) *are satisfied. Assume further that we are given a function $m : \Omega \to \mathbb{N}^n$ satisfying* (2.17)*, a real number A and a map $g : \mathbb{N}^n \to \mathbb{R}$ satisfying* (2.18) *and functions $\theta_1, \dots, \theta_n$ defined on the primes that take values in $\mathbb{R}$ that fulfil* (2.19) *and* (2.23)*. Assume that there exists a truncation pair $(\mathcal{F}, \psi)$ satisfying* (2.26)*-*(2.30) *and that for every $1 \leqslant i, j \leqslant n$ the following limit exists,*

$$\lim_{T \to +\infty} \frac{\sum_{p \leqslant T} \theta_i(p) \theta_j(p) \left(g_{\{i,j\}}(p) - g_i(p) g_j(p)\right)}{\mathcal{V}_i(T) \mathcal{V}_j(T)}. \quad (2.32)$$

Then the random vectors

$$(\Omega, \mathbf{P}_B) \to \mathbb{R}^n, \quad a \mapsto \mathbf{K}(a), \quad (2.33)$$

converge in distribution as $B \to \infty$ to a central multivariate normal distribution with covariance matrix $\mathbf{\Sigma}$ whose (i, j)-entry is the limit (2.32)*.*

2.3. **The proof of Theorem 2.5.** To prove the result, we shall use the *method of moments*. Specifically, the normal distribution has the special property that it is completely determined by its moments. Therefore it suffices to calculate the moments in our case. Our precise application is slightly more delicate, and we instead approximate with a sum of random variables, and use a version of the method of moments due to Billingsey (Lemma 2.14).

Our strategy consists of showing that for all $\mathbf{t} \in \mathbb{R}^n$ the random variable on Ω given by

$$\sum_{i=1}^n t_i \left(\frac{\left(\sum_{p \mid m_i(a)} \theta_i(p)\right) - \mathcal{M}_i(m_i(a))}{\mathcal{V}_i(m_i(a))} \right) \quad (2.34)$$

converges in distribution as $B \to \infty$ to a suitable linear combination of univariate normal distributions. To be able to use the level-of-distribution property (2.30), we show that we can restrict the size of the primes $p \mid m_i(a)$ to the range $p \leqslant \mathcal{F}(B)^{\psi(B)}$.

Lemma 2.6. *For all $\mathbf{t} \in \mathbb{R}^n$ we have*

$$\sum_{i=1}^n t_i \frac{\left(\sum_{p \mid m_i(a)} \theta_i(p)\right) - \mathcal{M}_i(m_i(a))}{\mathcal{V}_i(m_i(a))} - \sum_{i=1}^n t_i \frac{\left(\sum_{A < p \mid m_i(a), p \leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\right) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \Rightarrow 0.$$

Proof. By Slutsky's theorem [10, 3.2.13], if X_m, Y_m are sequences of random variables with $X_m \Rightarrow 0$ and $Y_m \Rightarrow 0$ then $X_m + Y_m \Rightarrow 0$. Therefore, it suffices to prove that

$$\frac{\left(\sum_{p \mid m_i(a)} \theta_i(p)\right) - \mathcal{M}_i(m_i(a))}{\mathcal{V}_i(m_i(a))} - \frac{\left(\sum_{p \mid m_i(a), A < p \leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\right) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \Rightarrow 0$$

for every i. Using (2.29) we see that it is sufficient to prove that

$$\frac{\left(\sum_{p\mid m_i(a)}\theta_i(p)\right)-\left(\sum_{p\mid m_i(a),A<p\leqslant \mathcal{F}(B)^{\psi(B)}}\theta_i(p)\right)+\left(\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})-\mathcal{M}_i(m_i(a))\right)}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})}\Rightarrow 0. \quad (2.35)$$

To prove this we use (2.19) to see that $\sum_{p\leqslant A}\theta_i(p)\ll A\Theta$ and

$$\left(\sum_{p\mid m_i(a)}\theta_i(p)\right)-\left(\sum_{p\mid m_i(a),p\leqslant \mathcal{F}(B)^{\psi(B)}}\theta_i(p)\right)\ll\Theta\#\left\{p>\mathcal{F}(B)^{\psi(B)}:p\mid m_i(a)\right\}.$$

Let $\Omega_0=\{a\in\Omega:m_i(a)\leqslant\mathcal{F}(B)\}$; note that $\lim_{B\to+\infty}\mathbf{P}_B(\Omega_0)=1$ by (2.26). Thus we can use the bound $\#\{p>z:p\mid m\}\leqslant(\log m)/(\log z)$ to see that the numerator in (2.35) is

$$\begin{aligned}&\ll_{A,\Theta} 1+\frac{\log m_i(a)}{\log(\mathcal{F}(B)^{\psi(B)})}+\left(\mathcal{M}_i(m_i(a))-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})\right)\\&\ll\frac{1}{\psi(B)}+\left(\mathcal{M}_i(m_i(a))-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})\right).\end{aligned}$$

The proof is concluded by using (2.28). □

For a function $h:\Omega\to\mathbb{C}$ we define $\mathbb{E}_B$ as follows,

$$\mathbb{E}_B[h]:=\frac{1}{\sum_{a\in\Omega}\chi_B(a)}\sum_{a\in\Omega}\chi_B(a)h(a),$$

i.e. the expected value of h with respect to $\mathbf{P}_B$. We begin by reducing the evaluation of moments to averages over $\mathbf{d}$ of the error term functions $\mathcal{R}(\mathbf{d},B)$ introduced in (2.25).

Lemma 2.7. *For all $B\geqslant 1$ and $k_1,\dots,k_n\in\mathbb{N}$ the following estimate holds with an absolute implied constant,*

$$\begin{aligned}&\mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{\substack{A<p\leqslant\mathcal{F}(B)^{\psi(B)}\\p\mid m_i(a)}}\theta_i(p)\right)^{k_i}\right]-\sum_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant k_i\\A<p_{i,j}\leqslant\mathcal{F}(B)^{\psi(B)}}}g\left(P_1,\dots,P_n\right)\prod_{\substack{1\leqslant u\leqslant n\\1\leqslant v\leqslant k_u}}\theta_u(p_{u,v})\\&\ll\frac{\Theta^{k_1+\cdots+k_n}}{M(B)}\sum_{\substack{\mathbf{d}\in\mathbb{N}^n\\(2.31)}}\left(g(\mathbf{d})|\mathcal{R}((1,\dots,1),B)|+|\mathcal{R}(\mathbf{d},B)|\right),\end{aligned}$$

where P_u is the radical of $\prod_{1\leqslant v\leqslant k_u}p_{u,v}$ and the sum over $p_{i,j}$ is over primes.

Proof. Expanding the k_i-th powers gives

$$\left(\sum_{a\in\Omega}\chi_B(a)\right)\mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{\substack{A<p\leqslant\mathcal{F}(B)^{\psi(B)}\\p\mid m_i(a)}}\theta_i(p)\right)^{k_i}\right]=\sum_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant k_i\\A<p_{i,j}\leqslant\mathcal{F}(B)^{\psi(B)}}}\left\{\prod_{\substack{1\leqslant u\leqslant n\\1\leqslant v\leqslant k_u}}\theta_u(p_{u,v})\right\}\sum_{\substack{a\in\Omega\\\forall i:P_i\mid m_i(a)}}\chi_B(a).$$

By (2.25) this equals

$$E+M(B)\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} g(P_1,\dots,P_n)\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}}\theta_u(p_{u,v}),$$

where E is given by

$$E:=\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} \mathcal{R}((P_1,\dots,P_n),B)\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}}\theta_u(p_{u,v}).$$

By (2.19) we infer that

$$|E|\leqslant \Theta^{k_1+\cdots+k_n}\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}}|\mathcal{R}((P_1,\dots,P_n),B)|\leqslant \Theta^{k_1+\cdots+k_n}\sum_{\substack{\mathbf{d}\in\mathbb{N}^n\\ (2.31)}}|\mathcal{R}(\mathbf{d},B)|.$$

To conclude the proof, it follows from (2.15) that

$$\frac{M(B)}{\sum_{a\in\Omega}\chi_B(a)}=\frac{\sum_{a\in\Omega}\chi_B(a)-\mathcal{R}((1,\dots,1),B)}{\sum_{a\in\Omega}\chi_B(a)}=1+O\left(\frac{|\mathcal{R}((1,\dots,1),B)|}{M(B)}\right).$$

We deduce that

$$\left|\mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{\substack{p\leqslant\mathcal{F}(B)^{\psi(B)}\\ p\mid m_i(a)}}\theta_i(p)\right)^{k_i}\right]-\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} g(P_1,\dots,P_n)\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}}\theta_u(p_{u,v})\right|$$

$$\ll\frac{|\mathcal{R}((1,\dots,1),B)|}{M(B)}\left(\sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} g(P_1,\dots,P_n)\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}}\theta_u(p_{u,v})\right)+\frac{\Theta^{k_1+\cdots+k_n}}{M(B)}\sum_{\substack{\mathbf{d}\in\mathbb{N}^n\\ (2.31)}}|\mathcal{R}(\mathbf{d},B)|.$$

The last sum over $p_{i,j}$ is at most $\Theta^{k_1+\cdots+k_n}\sum_{\mathbf{d}\in\mathbb{N}^n,(2.31)}g(\mathbf{d})$, thus concluding the proof. □

We need to understand the expression $\sum_{p_{i,j}}\prod_{i,j}\theta_i(p_{i,j})g(P_1,\dots,P_n)$ in Lemma 2.7 before proceeding. This will be based on interpreting the function $g_S(p)$ in (2.17) as the 'probability' that p divides each component of the vector $(m_i(a))_{i\in S}$ as a ranges through Ω. We do this by introducing some auxiliary random vectors.

Lemma 2.8. *For every prime $p>A$ there exists a random vector $\mathbf{X}_p=(X_{1,p},\dots,X_{n,p})$, such that*

$$\text{The random vectors } \mathbf{X}_p \text{ are independent for all primes } p. \tag{2.36}$$

$$X_{i,p} \text{ is Bernoulli and takes values in } \{0,1\}, \tag{2.37}$$

$$\mathrm{Prob}\left[\bigcap_{i\in S}\{X_{i,p}=1\}\right]=g_S(p) \text{ for all } S\subset\{1,\dots,n\}. \tag{2.38}$$

Proof. We first show that for a fixed prime p, there exists a random vector satisfying (2.37) and (2.38). To do so, let $S \subset \{1, \dots, n\}$ with complement S^c. Then we define

$$\mathrm{Prob}\left[\bigcap_{i\in S}\{X_{i,p}=1\}\bigcap_{i\in S^c}\{X_{i,p}=0\}\right]=\sum_{S'\subset S^c}(-1)^{|S'|}g_{S\cup S'}(p). \tag{2.39}$$

To see that this gives a well-defined random vector, it suffices to show that each probability (2.39) is non-negative (that the sum of all probabilities equals 1 follows from a simple inclusion-exclusion argument and the fact that $g_{\varnothing}(p) = g(1, \dots, 1) = 1$). However, by inclusion-exclusion we have

$$\begin{aligned}0 \leqslant \lim_{B\to+\infty}\frac{1}{M(B)}\sum_{\substack{a\in\Omega\\ p\mid m_i(a),\, i\in S\\ p\nmid m_i(a),\, i\in S^c}}\chi_B(a) &= \lim_{B\to+\infty}\frac{1}{M(B)}\sum_{S'\subset S^c}(-1)^{|S'|}\sum_{\substack{a\in\Omega\\ p\mid m_i(a),\, i\in S'}}\chi_B(a)\\ &= \sum_{S'\subset S^c}(-1)^{|S'|}g_{S\cup S'}(p),\end{aligned}$$

by (2.17) and (2.20), as required. The properties (2.37) and (2.38) then follow easily. Then (2.36) follows from Kolmogorov's extension theorem [10, Thm. 2.1.21]. $\square$

For $1 \leqslant i \leqslant n$ and $B > 1$ define the random variable

$$S_{i,B} := \sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}}\theta_i(p)X_{i,p}.$$

A very special case of the definition of the $X_{i,p}$ is that

$$\mathrm{Prob}[X_{i,p}=1]=g_i(p) \text{ and } \mathrm{Prob}[X_{i,p}=0]=1-g_i(p),$$

hence, recalling (2.21) and (2.22), for all $1 \leqslant i \leqslant n$ and $T \geqslant 0$ we get

$$\mathbb{E}\left[S_{i,B}\right]=\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}) \text{ and } \mathrm{Var}\left[S_{i,B}\right]^{1/2}=\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)}).$$

In verifying the last two equalities we have implicitly used that $g_i(p) = 0$ for $p \leqslant A$, as can be seen by Definition 2.3. In our next lemma we use Lemma 2.7, that regards moments of

$$\sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}}\theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a)),$$

to study the moments of

$$\sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}}\theta_i(p)\left(\mathbb{1}_{p\mathbb{Z}}(m_i(a))-g_i(p)\right),$$

which are closer to $\mathbf{K}(a)$ in (2.24).

Lemma 2.9. *For each $\mathbf{k} \in \mathbb{N}^n$ the following holds with an absolute implied constant,*

$$\begin{aligned}&\mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}}\theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a))-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})\right)^{k_i}\right]-\mathbb{E}\left[\prod_{i=1}^n\left(S_{i,B}-\mathbb{E}\left[S_{i,B}\right]\right)^{k_i}\right]\\ &\ll\frac{\prod_{i=1}^n\left(|\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|+\Theta\right)^{k_i}}{M(B)}\left\{\sum_{\mathbf{d}\in\mathbb{N}^n,(2.31)}\left(g(\mathbf{d})|\mathcal{R}((1,\dots,1),B)|+|\mathcal{R}(\mathbf{d},B)|\right)\right\}.\end{aligned}$$

Proof. Note that by (2.18) and (2.36)-(2.37) we get

$$\begin{aligned}\mathbb{E}\left[\prod_{i=1}^n S_{i,B}^{k_i}\right] &= \sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} \left\{\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}} \theta_u(p_{u,v})\right\} \mathbb{E}\left[\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}} X_{u,p_{u,v}}\right]\\ &= \sum_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant k_i\\ A<p_{i,j}\leqslant \mathcal{F}(B)^{\psi(B)}}} \left\{\prod_{\substack{1\leqslant u\leqslant n\\ 1\leqslant v\leqslant k_u}} \theta_u(p_{u,v})\right\} g\left(\mathrm{rad}\left(\prod_{v=1}^{k_1} p_{1,v}\right),\ldots,\mathrm{rad}\left(\prod_{v=1}^{k_n} p_{n,v}\right)\right),\end{aligned}$$

therefore the difference in Lemma 2.7 equals

$$\mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a))\right)^{k_i}\right] - \mathbb{E}\left[\prod_{i=1}^n S_{i,B}^{k_i}\right].$$

Using the binomial theorem we see that

$$\begin{aligned}&\mathbb{E}_B\left[\prod_{i=1}^n\left(-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}) + \sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a))\right)^{k_i}\right]\\ &= \sum_{\substack{0\leqslant j_1\leqslant k_1\\ \cdots\\ 0\leqslant j_n\leqslant k_n}} \left\{\prod_{i=1}^n (-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}))^{k_i-j_i}\binom{k_i}{j_i}\right\} \mathbb{E}_B\left[\prod_{i=1}^n\left(\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a))\right)^{j_i}\right]\end{aligned}$$

and

$$\mathbb{E}\left[\prod_{i=1}^n(-\mathbb{E}[S_{i,B}]+S_{i,B})^{k_i}\right] = \sum_{\substack{0\leqslant j_1\leqslant k_1\\ \cdots\\ 0\leqslant j_n\leqslant k_n}} \left\{\prod_{i=1}^n (-\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}))^{k_i-j_i}\binom{k_i}{j_i}\right\} \mathbb{E}\left[\prod_{i=1}^n S_{i,B}^{j_i}\right].$$

Alluding to Lemma 2.7 shows that the difference in our lemma is

$$\ll \frac{1}{M(B)} \sum_{\substack{0\leqslant j_1\leqslant k_1\\ \cdots\\ 0\leqslant j_n\leqslant k_n}} \left\{\prod_{i=1}^n |\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|^{k_i-j_i}\binom{k_i}{j_i}\Theta^{j_i}\right\} \sum_{\mathbf{d}\in\mathbb{N}^n} (g(\mathbf{d})|\mathcal{R}((1,\ldots,1),B)| + |\mathcal{R}(\mathbf{d},B)|),$$

where the sum over $\mathbf{d}$ is subject to the same conditions as in (2.31), except that $\omega(d_i)\leqslant k_i$ must be replaced by $\omega(d_i)\leqslant j_i$. Noting that $j_i\leqslant k_i$ and that each term in the sum over $\mathbf{d}$ is non-negative, we may bound the sum over $\mathbf{d}$ by the same one where the summation is over those $\mathbf{d}$ that satisfy (2.31). Therefore, the last quantity is at most

$$\frac{1}{M(B)} \sum_{\substack{0\leqslant j_1\leqslant k_1\\ \cdots\\ 0\leqslant j_n\leqslant k_n}} \left\{\prod_{i=1}^n |\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|^{k_i-j_i}\binom{k_i}{j_i}\Theta^{j_i}\right\} \sum_{\substack{\mathbf{d}\in\mathbb{N}^n\\ (2.31)}} (g(\mathbf{d})|\mathcal{R}((1,\ldots,1),B)| + |\mathcal{R}(\mathbf{d},B)|).$$

The proof is concluded by noting that the sum over j_i is $\left(|\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|+\Theta\right)^{k_i}$. ☐

For the rest of this section we fix an arbitrary vector $\mathbf{t} \in \mathbb{R}^n$. We are now in place to study the linear combination in (2.34) by modelling it via a linear combination of the random variables $S_{i,B}$. More specifically, for every prime $p > A$ we define the random variable

$$Y_p := \sum_{i=1}^{n} \frac{t_i \theta_i(p) \left(X_{i,p} - g_i(p)\right)}{\mathrm{Var}\left[S_{i,B}\right]^{1/2}}. \tag{2.40}$$

We next reformulate the previous lemmas using the variables Y_p.

Lemma 2.10. *For every $k \in \mathbb{N}$ the function of B given by*

$$\mathbb{E}_B\left[\left(\sum_{i=1}^{n} t_i \frac{\left(\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p) \mathbb{1}_{p\mathbb{Z}}(m_i(a))\right) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})}\right)^k\right]$$
$$- \mathbb{E}\left[\left(\frac{\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p}{\mathrm{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]^{1/2}}\right)^k\right]$$

tends to 0 *as* $B \to \infty$.

Proof. Using the multinomial theorem we see that the quantity $\mathbb{E}_B\left[\cdot\right]$ in the lemma equals

$$\sum_{\substack{\mathbf{k}\in(\mathbb{Z}_{\geqslant 0})^n \\ k_1+\cdots+k_n=k}} \frac{k! t_1^{k_1}\cdots t_n^{k_n}}{k_1!\cdots k_n!} \mathbb{E}_B\left[\prod_{i=1}^{n}\left(\frac{\left(\sum_{p\mid m_i(a), A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \theta_i(p)\right) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})}\right)^{k_i}\right].$$

By (2.40) we see that the term $\mathbb{E}\left[\cdot\right]$ in the lemma can similarly be written as

$$\sum_{\substack{\mathbf{k}\in(\mathbb{Z}_{\geqslant 0})^n \\ k_1+\cdots+k_n=k}} \frac{k! t_1^{k_1}\cdots t_n^{k_n}}{k_1!\cdots k_n!} \mathbb{E}\left[\prod_{i=1}^{n}\left(\frac{S_{i,B} - \mathbb{E}\left[S_{i,B}\right]}{\mathrm{Var}\left[S_{i,B}\right]^{1/2}}\right)^{k_i}\right].$$

Subtracting the last two equations and invoking Lemma 2.9 and (2.30) concludes the proof. □

Our plan is to use the Central Limit Theorem to study the distribution of $\sum_p Y_p$. Before that we need to study some basic properties of Y_p.

Lemma 2.11.

(1) The random variables Y_p are independent;
(2) For every prime p we have $\mathbb{E}\left[Y_p\right] = 0$;
(3) For every prime p the quantity $\mathrm{Var}\left[Y_p\right]$ equals

$$\sum_{i=1}^{n} \frac{t_i^2 \theta_i(p)^2 g_i(p)\left(1 - g_i(p)\right)}{\mathrm{Var}\left[S_{i,B}\right]} + 2 \sum_{1\leqslant i<j\leqslant n} \frac{t_i t_j \theta_i(p)\theta_j(p)\left(g_{\{i,j\}}(p) - g_i(p)g_j(p)\right)}{\mathrm{Var}\left[S_{i,B}\right]^{1/2}\mathrm{Var}\left[S_{j,B}\right]^{1/2}};$$

(4) We have

$$\lim_{B\to+\infty} \mathrm{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right] = Q(\mathbf{t}), \quad \textit{where } Q(\mathbf{t}) := \sum_{i=1}^{n} t_i^2 + 2\sum_{1\leqslant i<j\leqslant n} \sigma_{ij} t_i t_j,$$

and σ_{ij} are given by (2.32)*. In particular, we have $Q(\mathbf{t}) \geqslant 0$.*

Proof.

(1) This follows directly from (2.36).
(2) This follows from linearity of expectation and the fact that $\mathbb{E}[X_{i,p}] = g_i(p)$.
(3) Recall that the covariance of two random variables W_1, W_2 is defined by

$$\mathrm{Cov}[W_1, W_2] := \mathbb{E}[W_1 W_2] - \mathbb{E}[W_1]\,\mathbb{E}[W_2].$$

Using the standard formula $\mathrm{Var}\left[\sum_{i=1}^n X_i\right] = \sum_i \mathrm{Var}[X_i] + 2\sum_{i<j}\mathrm{Cov}[X_i, X_j]$ shows that $\mathrm{Var}[Y_p]$ equals

$$\sum_{i=1}^n \frac{t_i^2 \theta_i(p)^2 \,\mathrm{Var}[X_{i,p} - g_i(p)]}{\mathrm{Var}[S_{i,B}]}$$
$$+2 \sum_{1 \leqslant i<j \leqslant n} \frac{t_i t_j \theta_i(p)\theta_j(p)\,\mathrm{Cov}[(X_{i,p} - g_i(p)), (X_{j,p} - g_j(p))]}{\mathrm{Var}[S_{i,B}]^{1/2}\,\mathrm{Var}[S_{j,B}]^{1/2}}.$$

Using the rules $\mathrm{Var}[X + c] = \mathrm{Var}[X]$ and $\mathrm{Cov}[X - c, Y - c'] = \mathrm{Cov}[X, Y]$ this becomes

$$\sum_{i=1}^n \frac{t_i^2 \theta_i(p)^2 \,\mathrm{Var}[X_{i,p}]}{\mathrm{Var}[S_{i,B}]} + 2 \sum_{1 \leqslant i<j \leqslant n} \frac{t_i t_j \theta_i(p)\theta_j(p)\,\mathrm{Cov}[X_{i,p}, X_{j,p}]}{\mathrm{Var}[S_{i,B}]^{1/2}\,\mathrm{Var}[S_{j,B}]^{1/2}}.$$

By (2.37)-(2.38) we have $\mathrm{Var}[X_{i,p}] = g_i(p)(1 - g_i(p))$ and

$$\mathrm{Cov}[X_{i,p}, X_{j,p}] = \mathbb{E}[X_{i,p}X_{j,p}] - \mathbb{E}[X_{i,p}]\,\mathbb{E}[X_{j,p}] = g_{\{i,j\}}(p) - g_i(p)g_j(p),$$

which concludes the proof.
(4) Using the third part of the present lemma shows that

$$\mathrm{Var}\left[\sum_{A<p \leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right] = \sum_{i=1}^n t_i^2 + 2 \sum_{1 \leqslant i<j \leqslant n} t_i t_j \sum_{p \leqslant \mathcal{F}(B)^{\psi(B)}} \frac{\theta_i(p)\theta_j(p)\left(g_{\{i,j\}}(p) - g_i(p)g_j(p)\right)}{\mathrm{Var}[S_{i,B}]^{1/2}\,\mathrm{Var}[S_{j,B}]^{1/2}},$$

where we used that if $p \leqslant A$ then $g_{\{i,j\}}(p) = 0 = g_i(p)$. This equals

$$\sum_{i=1}^n t_i^2 + 2 \sum_{1 \leqslant i<j \leqslant n} t_i t_j \,\mathrm{Cov}\left[\frac{S_{i,B} - \mathbb{E}[S_{i,B}]}{\mathrm{Var}[S_{i,B}]^{1/2}}, \frac{S_{j,B} - \mathbb{E}[S_{j,B}]}{\mathrm{Var}[S_{j,B}]^{1/2}}\right].$$

One of the assumptions of Theorem 2.5 is that the limits in (2.32) exist. A direct comparison with the last expression here shows that

$$\lim_{B \to +\infty} \mathrm{Var}\left[\sum_{A<p \leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right] = Q(\mathbf{t})$$

due to (2.27). As a consequence, we obtain that $Q(\mathbf{t})$ is non-negative. □

We are now in position to apply the Central Limit Theorem to $\sum_p Y_p$.

Lemma 2.12. *For all $\mathbf{t} \in \mathbb{R}^n$ with $Q(\mathbf{t}) > 0$ the sequence of random variables*

$$\frac{1}{Q(\mathbf{t})^{1/2}} \sum_{A<p \leqslant \mathcal{F}(B)^{\psi(B)}} Y_p$$

converges in distribution to the standard normal distribution as $B \to \infty$. If $Q(\mathbf{t}) = 0$, then

$$\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p$$

converges in distribution to 0 *as $B \to \infty$.*

Proof. The first two parts of Lemma 2.11 imply that

$$\mathbb{E}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right] = 0.$$

Therefore, if $Q(\mathbf{t}) = 0$ then the last part of Lemma 2.11, implies that

$$\mathbb{E}\left[\left(\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right)^2\right] = \mathrm{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right] \to 0, \text{ as } B \to +\infty.$$

Chebyshev's inequality then yields $\sum_p Y_p \Rightarrow 0$.

We now assume that $Q(\mathbf{t}) > 0$. It is clear from the last part of Lemma 2.11 that we only have to show that one can apply the Central Limit Theorem to the sum $\sum_p Y_p$ in the present lemma. To do this we shall verify that Lindeberg's condition for the Central Limit Theorem for triangular arrays [2, Theorem 27.2] is satisfied. Recalling that $\mathbb{E}[Y_p] = 0$, this condition can be written as

$$\lim_{B\to+\infty} \frac{1}{\mathrm{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]} \sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \mathbb{E}\left[Y_p^2 \mathbb{1}\left(\left\{|Y_p| > \delta\, \mathrm{Var}\left[\sum_p Y_p\right]\right\}\right)\right] = 0.$$

By the last part of Lemma 2.11 it is clear that this is equivalent to showing that for all $\delta > 0$

$$\lim_{B\to+\infty} \sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \mathbb{E}\left[Y_p^2 \mathbb{1}\left(\{|Y_p| > \delta\}\right)\right] = 0. \tag{2.41}$$

To prove (2.41), note by the definition (2.40) and the bounds (2.19), $g_i(p) \leqslant 1$, we obtain

$$|Y_p| \ll_{\delta,n,\mathbf{t}} \frac{1}{\min_i \mathrm{Var}\left[S_{i,B}\right]^{1/2}}, \tag{2.42}$$

where the implied constant is independent of p and B. Hence, for any fixed $\delta > 0$ we see that by assumption (2.23) we have $\mathbb{1}(\{|Y_p| > \delta\}) = 0$ for all sufficiently large B. This is sufficient for (2.41). □

Lemma 2.10 shows that the moments of the number-theoretic objects $\sum_p \theta_i(p)\mathbb{1}_{p\mathbb{Z}}(m_i(a))$ essentially behave like the moments of certain random variables related to $\sum_p Y_p$, and in Lemma 2.12 we saw that $\sum_p Y_p$ has a limiting distribution. To pass from this to limiting distributions for the number-theoretic objects we first need to prove certain growth estimates for the moments of the related random variables. This is the goal of the next lemma.

Lemma 2.13. *Assume that $\mathbf{t} \in \mathbb{R}^n$ is fixed and that $Q(\mathbf{t}) > 0$. Then there exists $L > 0$ (that is independent of B and k) such that for all $k \in \mathbb{N}$ and $B \geqslant 1$ one has*

$$\left|\mathbb{E}\left[\left(\frac{\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p}{\operatorname{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]^{1/2}}\right)^k\right]\right| \ll k! L^k,$$

where the implied constant is independent of B, k and L.

Proof. By the last part of Lemma 2.11 we have $\lim_B \operatorname{Var}[\sum_p Y_p] = Q(\mathbf{t})$, therefore there exists $B_0 \geqslant 0$ such that $\operatorname{Var}\left[\sum_p Y_p\right]$ is strictly positive for all $B \geqslant B_0$. For such B we let

$$Z_p := Y_p \operatorname{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]^{-1/2},$$

so that (recall that the Y_p are independent by the first part of Lemma 2.11)

$$\mathbb{E}\left[\left(\frac{\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p}{\operatorname{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]^{1/2}}\right)^k\right] = \sum_{\substack{1\leqslant u\leqslant k \\ (k_1,\dots,k_u)\in\mathbb{N}^u \\ k_1+\dots+k_u=k}} \frac{k!}{\prod_{i=1}^u k_i!} {\sum}^* \prod_{i=1}^u \mathbb{E}\left[Z_{p_i}^{k_i}\right], \tag{2.43}$$

where the sum $\sum^*$ is over prime tuples satisfying $A < p_1 < \ldots < p_u \leqslant \mathcal{F}(B)^{\psi(B)}$. Note that we have $\mathbb{E}\left[Z_p\right] = 0$, therefore we can add the restriction that every k_i is strictly larger than 1. By the bound (2.42) we deduce that there exists $\mathcal{L}$ such that for all $B \geqslant 1$ one has $|Z_p| \leqslant \mathcal{L} \leqslant 1 + \mathcal{L}$. Therefore, for all $k_i \geqslant 2$ we have

$$|\mathbb{E}\left[Z_p^{k_i}\right]| \leqslant (1+\mathcal{L})^{k_i-2}\mathbb{E}\left[Z_p^2\right] \leqslant (1+\mathcal{L})^{k_i}\mathbb{E}\left[Z_p^2\right].$$

Thus, using $k_1 + \cdots + k_u = k$, we obtain

$$\sum_{A<p_1<\ldots<p_u\leqslant \mathcal{F}(B)^{\psi(B)}} \prod_{i=1}^u \mathbb{E}\left[Z_{p_i}^{k_i}\right] \leqslant (1+\mathcal{L})^k \sum_{A<p_1<\ldots<p_u\leqslant \mathcal{F}(B)^{\psi(B)}} \prod_{i=1}^u \mathbb{E}\left[Z_{p_i}^2\right],$$

which is at most

$$\frac{(1+\mathcal{L})^k}{u!}\left(\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \mathbb{E}\left[Z_p^2\right]\right)^u.$$

Now note that

$$\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \mathbb{E}\left[Z_p^2\right] = \frac{1}{\operatorname{Var}\left[\sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} Y_p\right]} \sum_{A<p\leqslant \mathcal{F}(B)^{\psi(B)}} \operatorname{Var}\left[Y_p\right] = 1.$$

Thus, by (2.43) we get

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(\frac{\sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}} Y_p}{\operatorname{Var}\left[\sum_{A<p\leqslant\mathcal{F}(B)^{\psi(B)}} Y_p\right]^{1/2}}\right)^k\right]\right| \\
&\leqslant (1+\mathcal{L})^k \sum_{1\leqslant u\leqslant k/2} \frac{1}{u!} \sum_{\substack{(k_1,\dots,k_u)\in(\mathbb{N}_{\geqslant 2})^u \\ k_1+\cdots+k_u=k}} \frac{k!}{k_1!\cdots k_u!} \\
&\leqslant (1+\mathcal{L})^k k! \sum_{1\leqslant u\leqslant k/2} \frac{k^u}{u!} \leqslant (1+\mathcal{L})^k k! \mathrm{e}^k.
\end{aligned}$$

This concludes the proof. □

2.3.1. *Conclusion of the proof of Theorem 2.5.* The following is the version of the method of moments we shall be using. It is the reason we proved Lemmas 2.10, 2.12, and 2.13.

Lemma 2.14 (Billinglsey, [1, Thm. 11.2])**.** *Let $\zeta, \xi_n, \zeta_n, (n \in \mathbb{N})$, be random variables. Suppose that $\zeta_n \Rightarrow \zeta$ and that*

$$\lim_{n\to+\infty} \left|\mathbb{E}\left[\xi_n^k\right] - \mathbb{E}\left[\zeta_n^k\right]\right| = 0 \text{ for all } k \in \mathbb{N} \quad \text{and} \quad \sup_{n,k\in\mathbb{N}} \frac{\left|\mathbb{E}\left[\zeta_n^k\right]\right|}{k!L^k} \leqslant 1 \text{ for some } L \in \mathbb{R}.$$

Then $\xi_n \Rightarrow \zeta$.

Note that by the fourth part of Lemma 2.11 the matrix $\boldsymbol{\Sigma}$ defined in Theorem 2.5 is positive semi-definite, so the multivariate normal distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ is well-defined. Let $\mathbf{X}$ be a random vector in $\mathbb{R}^n$ with distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$. Using the Cramér–Wold theorem [3, Thm. 29.4] we see that the convergence of the sequence (2.33) to a multivariate normal distribution follows if we show that for every $\mathbf{t} \in \mathbb{R}^n$ one has

$$\sum_{i=1}^n t_i \frac{\left(\sum_{p\mid m_i(a)} \theta_i(p)\right) - \mathcal{M}_i(m_i(a))}{\mathcal{V}_i(m_i(a))} \Rightarrow \sum_{i=1}^n t_i X_i.$$

In light of Lemma 2.6 this is equivalent to proving

$$\sum_{i=1}^n t_i \frac{\left(\sum_{A<p\mid m_i(a), p\leqslant\mathcal{F}(B)^{\psi(B)}} \theta_i(p)\right) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \Rightarrow \sum_{i=1}^n t_i X_i.$$

This can be deduced by injecting Lemmas 2.10, 2.12 and 2.13 into Lemma 2.14. □

2.4. **The proof of Theorem 2.1.** The proof is a combination of the Fundamental Lemma of the Combinatorial Sieve and Theorem 2.5. As a first step we show that the function $\mathcal{F}$ tends to infinity.

Lemma 2.15. *In the setting of Theorem 2.1 we have $\lim_{B\to\infty} \mathcal{F}(B) = +\infty$.*

Proof. Let $1 \leqslant i \leqslant n$. By (2.5) there are infinitely many primes p with $g_i(p) > 0$. Let p be such a prime. By (2.3) with $\mathbf{d} = (p\mathbb{1}_{\{i\}}(j) + \mathbb{1}_{\{1,\dots,n\}\setminus\{i\}}(j))_{j=1}^n$ we get

$$\liminf_{B\to+\infty} \frac{\#\left\{a \in \Omega : h(a) \leqslant B, p \leqslant m_i(a)\right\}}{N(B)} > 0.$$

Thus, by (2.7) we obtain $p \leqslant \liminf_{B\to+\infty} \mathcal{F}(B)$. Taking $p \to \infty$ concludes the proof. □

Before proceeding we must show that the typical size of $m_i(a)$ is not too small. This will require the Fundamental Lemma of the Combinatorial Sieve as given in [27, Thm. 3, p. 60].

Lemma 2.16. *Let $\mathcal{A}$ be a finite set of integers and $\mathfrak{P}$ a set of primes. If there exist real numbers $A, \kappa > 0$ and a multiplicative function $g : \mathbb{N} \to \mathbb{R} \cap [0, \infty)$ such that*

$$\prod_{\eta \leqslant p \leqslant \xi} \left(1 - \frac{g(p)}{p}\right)^{-1} \leqslant \left(\frac{\log \xi}{\log \eta}\right)^{\kappa} \left(1 + \frac{A}{\log \eta}\right), \ (\forall\, 2 \leqslant \eta \leqslant \xi), \tag{2.44}$$

then for all $X, y, u > 1$ the number of $a \in \mathcal{A}$ that are coprime to every $p \in \mathfrak{P} \cap (0, y]$ is

$$\ll X \prod_{\substack{p \in \mathfrak{P} \\ p \leqslant y}} \left(1 - \frac{g(p)}{p}\right) + \sum_{\substack{d \leqslant y^u \\ p \mid d \Rightarrow p \in \mathfrak{P}}} \mu(d)^2 \left| \# \{a \in \mathcal{A} : d \mid a\} - \frac{g(d)}{d} X \right|,$$

where the implied constant depends at most on κ and A.

Lemma 2.17. *Let ε be as in (2.8) and let $\varepsilon_1 : [1, \infty) \to [0, \infty)$ satisfy*

$$\lim_{B \to \infty} \varepsilon_1(B) = 0. \tag{2.45}$$

Define $z_0(B) := \mathcal{F}(B)^{\varepsilon_1(B)\varepsilon(B)}$ and assume that $\lim_{B \to \infty} z_0(B) = \infty$. Then

$$\lim_{B \to \infty} \mathbf{P}_B \left[a \in \Omega : z_0(B) \geqslant \min_{1 \leqslant i \leqslant n} m_i(a) \right] = 0.$$

Proof. By Boole's inequality it is sufficient to show that for all $1 \leqslant i \leqslant n$ one has

$$\lim_{B \to \infty} \mathbf{P}_B \left[a \in \Omega : z_0(B) \geqslant m_i(a) \right] = 0. \tag{2.46}$$

Let $z(B) := \mathcal{F}(B)^{\varepsilon(B)/2}$. To prove (2.46) we note that the inequality $z_0(B) \geqslant m_i(a)$ implies that for every prime $p \in (z_0(B), z(B))$ we have $p \nmid m_i(a)$. Therefore, letting $W := \prod_{z_0(B) < p < z(B)} p$, we get

$$\#\{a \in \Omega : h(a) \leqslant B, z_0(B) \geqslant m_i(a)\} \leqslant \#\{a \in \Omega, h(a) \leqslant B, \gcd(m_i(a), W) = 1\}.$$

We use Lemma 2.16 with $\mathcal{A} := \{m_i(a) : a \in \Omega, h(a) \leqslant B\}$ and

$$\mathfrak{P} := \{p > z_0(B) : p \text{ prime}\}, X := N(B), \kappa := c_i, g(p) := p g_i(p), u := 2, y := z(B) - 1.$$

Assumption (2.5) and the estimate $\log(1 - z)^{-1} = z + O(z^2), |z| < 1$ show that

$$\log \prod_{\eta \leqslant p \leqslant \xi} (1 - g_i(p))^{-1} = \sum_{\eta \leqslant p \leqslant \xi} g_i(p) + O\left(\sum_{\eta \leqslant p \leqslant \xi} g_i(p)^2\right) = c_i \log\left(\frac{\log \xi}{\log \eta}\right) + O\left(\frac{1}{\log \eta}\right), \tag{2.47}$$

from which we infer (2.44) by using $\exp(\varepsilon) = 1 + O(\varepsilon)$ for $\varepsilon = O(1/\log \eta)$. Lemma 2.16 gives the following bound for $\#\{a \in \Omega, h(a) \leqslant B, \gcd(m_i(a), W) = 1\}$,

$$\ll N(B) \prod_{z_0(B) < p \leqslant z(B)} (1 - g_i(p)) + \sum_{\substack{d \leqslant z(B)^2 \\ p \mid d \Rightarrow p > z_0(B)}} \mu(d)^2 \left| \# \{a \in \Omega : h(a) \leqslant B, d \mid m_i(a)\} - g_i(d) N(B) \right|.$$

Exponentiating (2.47) shows that the first term is

$$\ll N(B) \left(\frac{\log z_0(B)}{\log z(B)}\right)^{c_i} \ll N(B) \varepsilon_1(B)^{c_i} = o(N(B)).$$

To bound the second term we note that $z_0(B) > A$ for B sufficiently large due to the assumption $\lim_{B\to\infty} z_0(B) = \infty$. Therefore, we can replace the condition $p > z_0(B)$ by $p > A$, thus, the second term is

$$\ll \sum_{\substack{d\leqslant z(B)^2\\ p\mid d\Rightarrow p>A}} \mu(d)^2 |\mathcal{R}((1,\dots,1,d,1\dots,1),B)|,$$

where every component in the vector in $\mathcal{R}$ equals 1, except the i-th entry, which equals d. By (2.9) and Lemma 2.15 we immediately find that this is $o(N(B))$. This verifies (2.46) and hence concludes the proof. □

We will use Theorem 2.5 with

$$\chi_B(a) := \mathbb{1}_{[0,B]}(h(a)),\ \theta_i(p) = 1 \text{ and } M(B) = N(B).$$

With these choices we see that (2.13), (2.14) and (2.15) are satisfied due to (2.1). The assumption (2.19) obviously holds with $\Theta = 1$. We next show that $(\mathcal{F},\psi)$ fulfils the truncation-pair Definition 2.4, where $\mathcal{F}$ is as in (2.7) and

$$\psi(B) := \frac{\varepsilon(B)}{\sqrt{\log\log\log \mathcal{F}(B)}} = \frac{\sqrt{\log\log\log \mathcal{F}(B)}}{\sqrt{\log\log \mathcal{F}(B)}}, \tag{2.48}$$

where the equality is by (2.8). Firstly, (2.26) follows directly from (2.7). Secondly, to verify (2.27) it is clearly sufficient to show that

$$\lim_{B\to+\infty} \psi(B)\log \mathcal{F}(B) = +\infty.$$

This, however, follows by (2.48) and Lemma 2.15.

Before proceeding, note that by (2.5) we have

$$\begin{aligned}\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}) &= c_i \log\log \mathcal{F}(B) - c_i \log \frac{1}{\psi(B)} + O(1)\\ &= c_i \log\log \mathcal{F}(B) + O(\log\log\log \mathcal{F}(B)),\end{aligned}$$

where the last estimate is due to Lemma 2.15 and (2.48). We similarly have

$$\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^2 = c_i \log\log \mathcal{F}(B) + O(\log\log\log \mathcal{F}(B)).$$

The first part of (2.28) follows from

$$\frac{1}{\psi(B)\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \ll \frac{1}{\psi(B)\sqrt{\log\log \mathcal{F}(B)}} = \frac{1}{\sqrt{\log\log\log \mathcal{F}(B)}} = o(1),$$

which goes to 0 as $B \to \infty$ by Lemma 2.15. To verify the second part of (2.28) we use Lemma 2.17 with

$$\varepsilon_1(B) = (\log\log\log \mathcal{F}(B))^{-1/2}.$$

This choice shows that the function $z_0(B)$ of Lemma 2.17 coincides with $\mathcal{F}(B)^{\psi(B)}$. By Lemma 2.17 we deduce that there exists a set $S \subset \Omega$ with $\lim_{B\to\infty} \mathbf{P}_B[S] = 1$ and such that whenever $a \in S$ then we have for all $1 \leqslant i \leqslant n$ that

$$\mathcal{F}(B)^{\psi(B)} \leqslant m_i(a) \leqslant \mathcal{F}(B). \tag{2.49}$$

Hence, for $a \in S$ we get

$$\begin{aligned} 0 \leqslant \frac{\mathcal{M}_i(m_i(a)) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} &\leqslant \frac{\mathcal{M}_i(\mathcal{F}(B)) - \mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \\ &= \frac{c_i(\log\log\mathcal{F}(B)) + O(1) - c_i(\log\log\mathcal{F}(B)) + O(\log\log\log\mathcal{F}(B))}{\sqrt{c_i(\log\log\mathcal{F}(B)) + O(1)}}, \end{aligned}$$

which, by Lemma 2.15, tends to 0 as $B \to \infty$. We are thus left with verifying (2.29) and (2.30). For the former we observe that for $a \in S$ we have the following by (2.49),

$$1 \leqslant \frac{\mathcal{V}_i(m_i(a))}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} \leqslant \frac{\mathcal{V}_i(\mathcal{F}(B))}{\mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})} = \frac{\sqrt{c_i \log\log\mathcal{F}(B) + O(1)}}{\sqrt{c_i \log\log\mathcal{F}(B) + O(\log\log\log\mathcal{F}(B))}},$$

which, by Lemma 2.15, tends to 1 as $B \to \infty$. We are left with verifying (2.30). Note that in our setting one has $\mathcal{R}((1,\dots,1),B) = 0$ due to $M(B) = N(B)$ and (2.6), hence we only have to show

$$\lim_{B\to+\infty} \frac{\prod_{i=1}^n \left(1 + |\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|^{k_i}\right)}{N(B)\prod_{i=1}^n \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^{k_i}} \sum_{\substack{\mathbf{d}\in\mathbb{N}^n \\ (2.31)}} |\mathcal{R}(\mathbf{d},B)| = 0 \tag{2.50}$$

in order to verify (2.30). Note that the bounds

$$\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)}) \ll \log\log\mathcal{F}(B) \text{ and } \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^2 \gg \log\log\mathcal{F}(B)$$

imply that

$$\frac{\prod_{i=1}^n \left(|\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})|\right)^{k_i}}{\prod_{i=1}^n \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^{k_i}} \ll (\log\log\mathcal{F}(B))^{\frac{k_1+\cdots+k_n}{2}}.$$

Now note that the d_i in (2.50) satisfy for all sufficiently large B,

$$d_i \leqslant \mathcal{F}(B)^{k_i\psi(B)} \leqslant \mathcal{F}(B)^{\varepsilon(B)},$$

therefore,

$$\begin{aligned} &\frac{\prod_{i=1}^n \left(|\mathcal{M}_i(\mathcal{F}(B)^{\psi(B)})| + 1\right)^{k_i}}{N(B)\prod_{i=1}^n \mathcal{V}_i(\mathcal{F}(B)^{\psi(B)})^{k_i}} \sum_{\substack{\mathbf{d}\in\mathbb{N}^n \\ (2.31)}} |\mathcal{R}(\mathbf{d},B)| \\ &\ll \frac{(\log\log\mathcal{F}(B))^{\frac{k_1+\cdots+k_n}{2}}}{N(B)} \sum_{\substack{\mathbf{d}\in\mathbb{N}^n \\ |\mathbf{d}|\leqslant\mathcal{F}(B)^{\varepsilon(B)} \\ p|d_1\cdots d_n \Rightarrow p>A}} \mu(d_1)^2\cdots\mu(d_n)^2|\mathcal{R}(\mathbf{d},B)|, \end{aligned}$$

which is $o(1)$ as can be seen by taking $\gamma = 1 + \frac{1}{2}(k_1 + \cdots + k_n)$ in (2.9). This confirms (2.50). Having verified all assumptions of Theorem 2.5, the result is that the random vector

$$\left(\frac{\omega(m_1(a)) - \sum_{p\leqslant m_1(a)} g_1(p)}{\left(\sum_{p\leqslant m_1(a)} g_1(p)(1-g_1(p))\right)^{1/2}}, \dots, \frac{\omega(m_n(a)) - \sum_{p\leqslant m_n(a)} g_n(p)}{\left(\sum_{p\leqslant m_n(a)} g_n(p)(1-g_n(p))\right)^{1/2}} \right)$$

has the limiting distribution as in Theorem 2.1. We next deduce the analogous distribution result for the function given in (2.10). We define the random vectors on Ω by

$$\mathbf{V} = \left(\frac{\mathcal{V}(m_1(a))}{\sqrt{c_1 \log\log \mathcal{F}(B)}}, \dots, \frac{\mathcal{V}(m_n(a))}{\sqrt{c_n \log\log \mathcal{F}(B)}} \right),$$

$$\mathbf{T} = \left(\frac{\omega(m_1(a)) - \mathcal{M}(m_1(a))}{\mathcal{V}(m_1(a))}, \dots, \frac{\omega(m_n(a)) - \mathcal{M}(m_n(a))}{\mathcal{V}(m_n(a))} \right),$$

and

$$\mathbf{M} = \left(\frac{\mathcal{M}(m_1(a)) - c_1 \log\log \mathcal{F}(B)}{\mathcal{V}(m_1(a))}, \dots, \frac{\mathcal{M}(m_n(a)) - c_n \log\log \mathcal{F}(B)}{\mathcal{V}(m_n(a))} \right).$$

Recalling (2.10), we have

$$\mathbf{K} = \mathbf{V}(\mathbf{T} + \mathbf{M})$$

where the product is taken coordinate-wise. Theorem 2.5 implies that $\mathbf{T} \Rightarrow \mathcal{N}(\mathbf{0}, \Sigma)$. Moreover by (2.5), the fact that $\mathcal{F}(B) \to \infty$ and (2.49), we have $\mathbf{M} \Rightarrow \mathbf{0}$ and $\mathbf{V} \Rightarrow \mathbf{1}$ (where $\mathbf{1}$ is the n-dimensional vector all of whose coordinates are 1). Slutsky's theorem therefore implies that $\mathbf{K} \Rightarrow \mathcal{N}(\mathbf{0}, \Sigma)$. Furthermore, the limit in (2.32) becomes (2.11). This is due to (2.5), which ensures that

$$\sum_{p \leqslant T} \theta_i(p)\theta_j(p) g_i(p) g_j(p) = \sum_{p \leqslant T} g_i(p) g_j(p) \leqslant \left(\sum_{p_1, p_2 \leqslant T} g_i(p_1)^2 g_j(p_2)^2 \right)^{1/2} = O(1),$$

by Cauchy–Schwarz, and

$$\mathcal{V}_i(T)^2 = \sum_{p \leqslant T} g_i(p)\left(1 - g_i(p)\right) = \sum_{p \leqslant T} g_i(p) + O(1). \qquad \square$$

3. Application to integral points

In this section we prove Theorem 1.8 using Theorem 2.1.

3.1. **Inner product of divisors.**

Let X be an integral Noetherian scheme and $\operatorname{Div} X$ the free abelian group generated by the integral (Weil) divisors on X. To simplify some of the statements and proofs in what follows, we introduce an inner product on $\langle \cdot, \cdot \rangle$ on $\operatorname{Div} X$ as follows. For integral divisors D, E we define

$$\langle D, E \rangle = \begin{cases} 1, & D = E, \\ 0, & D \neq E. \end{cases}$$

As the integral divisors form a basis of $\operatorname{Div} X$, this extends to an inner product on $\operatorname{Div} X$. Explicitly $\langle D, E \rangle$ is the number of common irreducible components of $D, E \in \operatorname{Div} X$ counted with multiplicity. This extends to $\operatorname{Div}_{\mathbb{R}} X := (\operatorname{Div} X) \otimes_{\mathbb{Z}} \mathbb{R}$ and we let $\| \cdot \| : \operatorname{Div}_{\mathbb{R}} X \to \mathbb{R}_{\geqslant 0}$ be the induced norm. Our inner product is a convenient piece of notation which should not be confused with more subtle geometric information like intersection numbers of divisors.

3.2. **Points over finite fields.** In the statement, we implicitly only sum over those p with $\mathcal{X}(\mathbb{F}_p) \neq \varnothing$. (A similar convention applies to Corollary 3.2.)

Proposition 3.1. *Let X be a geometrically integral variety over $\mathbb{Q}$ of dimension n and $\mathcal{X}$ a model of X over $\mathbb{Z}$. Let $Z \subsetneq X$ be a reduced closed subscheme of pure dimension d and $\mathcal{Z}$ its closure in $\mathcal{X}$.*

(1) We have

$$\sum_{p \geqslant T} \left(\frac{\#\mathcal{Z}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} \right)^2 \ll \frac{1}{T}.$$

(2) We have

$$\sum_{p \leqslant T} \frac{\#\mathcal{Z}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} = \begin{cases} C_Z + O(1/T), & \text{if } d < n-1, \\ \langle Z, Z \rangle \log\log T + C'_Z + O(1/\log T), & \text{if } d = n-1. \end{cases}$$

for some $C_Z > 0$ and $C'_Z \in \mathbb{R}$.

Proof. By the Lang–Weil estimates [19] we have

$$\#\mathcal{Z}(\mathbb{F}_p) \ll p^d \ll p^{n-1}, \quad \#\mathcal{X}(\mathbb{F}_p) = p^n + O(p^{n-1/2}).$$

Then (1) follows from the estimate

$$\sum_{p \geqslant T} \left(\frac{\#\mathcal{Z}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} \right)^2 \ll \sum_{p \geqslant T} p^{-2} \ll \frac{1}{T},$$

while, the case $d < n-1$ of (2) follows from

$$\sum_{p \geqslant T} \frac{\#\mathcal{Z}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} \ll \sum_{p \geqslant T} \frac{p^d}{p^n} \leqslant \sum_{p \geqslant T} p^{-2} \ll \frac{1}{T}.$$

It thus suffices to prove (2) when $d = n-1$. For a number field k we denote by $z_p(k)$ the number of prime ideals of k of degree 1 over p. Let I be the set of irreducible components of Z, and for each $i \in I$ let k_i be the algebraic closure of $\mathbb{Q}$ in the function field of the corresponding irreducible component; this is a number field. For all sufficiently large primes p, the irreducible components of $\mathcal{Z}_{\mathbb{F}_p}$ which are geometrically integral correspond exactly to those prime ideals of k_i of degree 1 over p. Moreover the components which are not geometrically integral contain $O(p^{n-2})$ points over $\mathbb{F}_p$, by Lang–Weil. Thus applying the Lang–Weil estimates to each irreducible component of $\mathcal{Z}_{\mathbb{F}_p}$ gives

$$\#\mathcal{Z}(\mathbb{F}_p) = \sum_{i \in I} z_p(k_i) p^{n-1} + O(p^{n-3/2})$$

and hence

$$\frac{\#\mathcal{Z}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} = \sum_{i \in I} \frac{z_p(k_i)}{p} + O\left(\frac{1}{p^{3/2}} \right).$$

However turning this into a sum over the non-zero prime ideals of the number field, we have

$$\sum_{p \leqslant T} z_p(k) = \sum_{\substack{N(\mathfrak{p}) \leqslant T \\ N(\mathfrak{p}) \text{ prime}}} 1 = \sum_{N(\mathfrak{p}) \leqslant T} 1 + O(T^{1/2}) = \mathrm{Li}(T) + O(T \exp(-c\sqrt{\log T}))$$

for some constant $c > 0$, where the second equality is by [25, Lem. 9.3] and the last by the prime ideal theorem [25, Thm. 3.1]. The result now follows from partial summation. ☐

Corollary 3.2. *Let D_1, D_2 be reduced divisors on X and $\mathcal{D}_i$ their closures in $\mathcal{X}$. Then*

$$\lim_{T\to\infty} \frac{\sum_{p\leqslant T} \#(\mathcal{D}_1 \cap \mathcal{D}_2)(\mathbb{F}_p)/\#\mathcal{X}(\mathbb{F}_p)}{(\sum_{p\leqslant T} \#\mathcal{D}_1(\mathbb{F}_p)/\#\mathcal{X}(\mathbb{F}_p))^{1/2}(\sum_{p\leqslant T} \#\mathcal{D}_2(\mathbb{F}_p)/\#\mathcal{X}(\mathbb{F}_p))^{1/2}} = \frac{\langle D_1, D_2\rangle}{\|D_1\|\|D_2\|}.$$

3.3. **Proof of Theorem 1.8.** We now take the notation and set up of Theorem 1.8.

3.3.1. *Application of Theorem 2.1.* We take $\Omega = \mathcal{X}(\mathbb{Z})\backslash\mathcal{D}(\mathbb{Z}), h = H$ and

$$m : \Omega \to \mathbb{N}^n, \quad x \mapsto \left(\prod_{\substack{p \\ x \bmod p \in \mathcal{D}_i(\mathbb{F}_p)}} p \right)_{i=1,\dots,n}.$$

That (2.1) and (2.2) hold is clear. We next show (2.3) and (2.4) using (1.2). For this we require the following.

Lemma 3.3. *We have*

$$\#\{x \in \mathcal{D}(\mathbb{Z}) : H(x) \leqslant B\} = o(\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}).$$

Proof. Let $0 < \varepsilon < \eta$ and let p be a prime with $B^{(\eta-\varepsilon)/M} < p < 2B^{(\eta-\varepsilon)/M}$ (this exists by Bertrand's postulate). Then applying (1.2) we obtain

$$\begin{aligned} \frac{\#\{x \in \mathcal{D}(\mathbb{Z}) : H(x) \leqslant B\}}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} &\leqslant \frac{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B, x \bmod p \in \mathcal{D}(\mathbb{F}_p)\}}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} \\ &= \frac{\#\mathcal{D}(\mathbb{F}_p)}{\#\mathcal{X}(\mathbb{F}_p)} + O\left(\frac{p^M}{B^\eta}\right) \ll \frac{1}{p} + B^{-\varepsilon} = o(1) \end{aligned}$$

where the penultimate line is by the Lang–Weil estimates. □

Let $d_1, \dots, d_n$ be square-free and let $d = [d_1, \dots, d_n]$ be their least common multiple. Let

$$\Upsilon_{\mathbf{d}} = \{x \in \mathcal{X}(\mathbb{Z}/d\mathbb{Z}) : x \bmod d_i \in \mathcal{D}(\mathbb{Z}/d_i\mathbb{Z}), i = 1, \dots, n\}.$$

Providing each d_i is coprime to every $p \leqslant A$, Lemma 3.3 and (1.2) imply that

$$\frac{\#\{x \in \Omega : H(x) \leqslant B, x \bmod d_i \mid m_i(x), i = 1, \dots, n\}}{\#\{x \in \Omega : H(x) \leqslant B\}} = \frac{\#\Upsilon_{\mathbf{d}}}{\#\mathcal{X}(\mathbb{Z}/d\mathbb{Z})} + O(d^M B^{-\eta}).$$

Thus (2.3) holds with

$$g(\mathbf{d}) = \frac{\#\Upsilon_{\mathbf{d}}}{\#\mathcal{X}(\mathbb{Z}/d\mathbb{Z})},$$

where g is supported on vectors $\mathbf{d}$ with square-free entries such that $p \mid d_i \implies p > A$. To see that g is multiplicative, let $\gcd(d_1 \dots d_n, d'_1 \dots d'_n) = 1$. Then

$$g(\mathbf{d}\mathbf{d}') = \frac{\#\Upsilon_{\mathbf{d}\mathbf{d}'}}{\#\mathcal{X}(\mathbb{Z}/[d_1 d'_1, \dots, d_n d'_n]\mathbb{Z})} = \frac{\#\Upsilon_{\mathbf{d}}}{\#\mathcal{X}(\mathbb{Z}/d\mathbb{Z})} \cdot \frac{\#\Upsilon_{\mathbf{d}'}}{\#\mathcal{X}(\mathbb{Z}/d'\mathbb{Z})} = g(\mathbf{d})g(\mathbf{d}')$$

by the Chinese remainder theorem and our coprimality assumption. This shows (2.4). Next (2.5) follows from Proposition 3.1 with $c_i = \langle D_i, D_i\rangle$. To show (2.7) we use the following.

Lemma 3.4. *There exists $c > 0$ such that for all $x \in \mathcal{X}(\mathbb{Z})\backslash\mathcal{D}(\mathbb{Z})$ we have*

$$\prod_{\substack{p \\ x \bmod p \in \mathcal{D}(\mathbb{F}_p)}} p \ll H(x)^c.$$

Proof. Let $\overline{\mathcal{D}}$ be the closure of D in $\mathbb{P}^d_{\mathbb{Z}}$. Choose homogeneous polynomials $f_1, \dots, f_r$ over $\mathbb{Z}$ which generate the ideal of $\overline{\mathcal{D}}$. As $x \notin \mathcal{D}(\mathbb{Z})$, we have $f_i(x) \neq 0$ for some i. Moreover $x \bmod p \in \mathcal{D}(\mathbb{F}_p)$ implies that $p \mid f_i(x)$. Thus the quantity in question is at most $|f_i(x)| \ll H(x)^{\deg f_i} \ll H(x)^{\deg \overline{\mathcal{D}}}$, as required. ☐

We find that (2.7) holds with $\mathcal{F}(B) \ll B^c$. Then (2.9) follows from (1.2) (see Remark 2.2). Finally the limit (2.11) exists and equals $\langle D_i, D_j \rangle / \|D_i\| \|D_j\|$ by Corollary 3.2. Thus all assumptions of Theorem 2.1 hold and we deduce the first part of Theorem 1.8.

3.3.2. *Rank of the matrix.* It remains to prove the final part of Theorem 1.8, regarding the formula for the rank of the covariance matrix. As the D_i are reduced, the matrix $(c_{i,j}/\sqrt{c_{i,i}c_{j,j}})$ is exactly the Gram matrix of the divisors $D_1/\|D_1\|, \dots, D_n/\|D_n\|$ with respect to the inner product on $\operatorname{Div}_{\mathbb{R}} X$ defined in §3.1. However the rank of the Gram matrix is the dimension of the vector subspace of $\operatorname{Div}_{\mathbb{R}} X$ generated by the $D_i/\|D_i\|$. But this is also equal to the rank of the subgroup of $\operatorname{Div} X$ generated by the D_i. This completes the proof. ☐

4. Examples

We now give various examples illustrating our results and use Theorem 1.8 to prove the special cases stated in the introduction.

4.1. **Complete intersections.** Here we explain the proof of Theorem 1.2. We apply Theorem 1.8 with $X : f_1 = \dots = f_R = 0$, $d = n - 1$ and $D_i = X \cap (x_i = 0)$. We take $\mathcal{X}$ to be the model given by taking the closure of X in $\mathbb{P}^d_{\mathbb{Z}}$. It suffices to verify (1.2).

Choose $A = A(\mathbf{f}) > 0$ such that X has good reduction at all primes $p > A$. Let Q be square-free and supported on primes greater than A. Let $\Upsilon \subset \mathcal{X}(\mathbb{Z}/Q\mathbb{Z})$ and

$$N(\Upsilon, B) := \#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B, x \bmod Q \in \Upsilon\}.$$

We first note that the leading term of (1.2) is known to hold, and follows from equidistribution results of Peyre and standard properties of Tamagawa measures [23, 24].

Lemma 4.1. *We have*

$$\lim_{B \to \infty} \frac{N(\Upsilon, B)}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} = \frac{\#\Upsilon}{\#\mathcal{X}(\mathbb{Z}/Q\mathbb{Z})}.$$

Proof. By [23, Prop. 5.5.3], Manin's conjecture holds here with respect to arbitrary choices of height function. This implies that the rational points are equidistributed with respect to Peyre's Tamagawa measure [23, Prop. 3.3]. The measure of the resulting adelic volumes is calculated in [24, Thm. 2.14(b)] (cf. [24, Cor. 2.15]), and gives the stated result. ☐

It therefore suffices to show that we can obtain an asymptotic formula for $N(\Upsilon, B)$ with an effective error term. Denote the affine cone of Υ by

$$\widehat{\Upsilon} = \{\mathbf{y} \in (\mathbb{Z}/Q\mathbb{Z})^n : \mathbf{y} \not\equiv \mathbf{0} \bmod p \, \forall p \mid Q, \, \mathbf{y} \bmod Q \in \Upsilon\}.$$

We begin with a Möbius inversion. The key observation in the following lemma is that we may take the Möbius variable k to be small.

Lemma 4.2. *Fix an arbitrary $\eta_1 > 0$. Then for all $B \geqslant 1$ we have that $N(\Upsilon, B)$ equals*

$$\frac{1}{2} \sum_{\mathbf{y} \in \widehat{\Upsilon}} \sum_{\substack{k \in \mathbb{N} \cap [1, B^{\eta_1}] \\ \gcd(k,Q)=1}} \mu(k) \# \left\{ \mathbf{x} \in \left(\mathbb{Z} \cap \left[-\frac{B}{k}, \frac{B}{k} \right] \right)^n : \mathbf{f}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \equiv \frac{\mathbf{y}}{k} \bmod Q \right\} + O_{\mathbf{f}} \left(Q^n B^{n-RD-\eta_1} \right),$$

where the implied constant depends at most on $\mathbf{f}$*.*

Proof. Using Möbius inversion we see that $N(\Upsilon, B)$ equals

$$\begin{aligned}&\frac{1}{2}\#\left\{\mathbf{x}\in\mathbb{Z}^n : \gcd(x_1,\dots,x_n)=1, \max_i |x_i| \leqslant B, \mathbf{f}(\mathbf{x})=\mathbf{0}, \mathbf{x} \bmod Q \in \widehat{\Upsilon}\right\}\\ =&\frac{1}{2}\sum_{\substack{1\leqslant k\leqslant B\\ \gcd(k,Q)=1}} \mu(k)\#\{\mathbf{x}\in\mathbb{Z}^n : \max_i |x_i| \leqslant B/k, \mathbf{f}(\mathbf{x})=\mathbf{0}, k\mathbf{x} \bmod Q \in \widehat{\Upsilon}\}\\ =&\frac{1}{2}\sum_{\mathbf{y}\in\widehat{\Upsilon}}\sum_{\substack{k\leqslant B\\ \gcd(k,Q)=1}} \mu(k)\#\{\mathbf{x}\in\mathbb{Z}^n : \max_i |x_i| \leqslant B/k, \mathbf{f}(\mathbf{x})=\mathbf{0}, \mathbf{x}\equiv k^{-1}\mathbf{y} \bmod Q\},\end{aligned}$$

where the inverse is taken modulo Q. We note that Birch's estimate [4, Thm. 1] ensures that for all $P \geqslant 1$ one has

$$\#\{\mathbf{x}\in\mathbb{Z}^n : \max_i |x_i| \leqslant P, \mathbf{f}(\mathbf{x})=\mathbf{0}\} = O_{\mathbf{f}}\left(P^{n-RD}\right).$$

Therefore, ignoring the condition $\mathbf{x}\equiv k^{-1}\mathbf{y} \bmod Q$ we obtain

$$\#\{\mathbf{x}\in\mathbb{Z}^n : \max_i |x_i| \leqslant B/k, \mathbf{f}(\mathbf{x})=\mathbf{0}, \mathbf{x}\equiv k^{-1}\mathbf{y} \bmod Q\} \ll_{\mathbf{f}} (B/k)^{n-RD}.$$

Noting that our assumptions ensure that $n-RD\geqslant 2$, hence this is $\ll_{\mathbf{f}} B^{n-RD}k^{-2}$. Using the trivial bound $\#\widehat{\Upsilon}\leqslant Q^n$ we therefore see that for all $L\geqslant 1$ we have

$$\begin{aligned}&\sum_{\mathbf{y}\in\widehat{\Upsilon}}\sum_{\substack{L<k\leqslant B\\ \gcd(k,Q)=1}} \mu(k)\#\{\mathbf{x}\in\mathbb{Z}^n : \max_i |x_i| \leqslant B/k, \mathbf{f}(\mathbf{x})=\mathbf{0}, \mathbf{x}\equiv k^{-1}\mathbf{y} \bmod Q\}\\ \ll_{\mathbf{f}}& Q^nB^{n-RD}\sum_{k>L}k^{-2} \ll B^{n-RD}Q^nL^{-1}.\end{aligned}$$

Taking $L=B^{\eta_1}$ concludes the proof. □

We next record the case $\boldsymbol{\nu}=\mathbf{0}$ of the work by van Ittersum [28, Thm. 2.15]. It gives an effective error term for the number of integer zeros of bounded height on a complete intersection of polynomials which need not be homogeneous. For a polynomial g let $\widetilde{g}$ denote the homogeneous part of g.

Lemma 4.3 (van Ittersum)**.** *Let* $g_1,\dots,g_R\in\mathbb{Z}[x_1,\dots,x_n]$ *be arbitrary polynomials of common degree* D *and assume that* $\mathfrak{B}(\mathbf{g}) > 2^{D-1}(D-1)R(R+1)$*. Then there exist positive* M_1,η_2 *that depend at most on* $\mathfrak{B}(\mathbf{g}), R$ *and* D *such that*

$$\#\left\{\mathbf{z}\in\mathbb{Z}^n : \max_{1\leqslant i\leqslant n}|z_i|\leqslant B : \mathbf{g}(\mathbf{z})=\mathbf{0}\right\} = \mathfrak{S}(\mathbf{g})J(\widetilde{\mathbf{g}})B^{n-RD} + O\left(B^{n-RD-\eta_2}C\widetilde{C}^{M_1}\right),$$

where the implied constant depends at most on $n, D, R, \mathfrak{B}(\mathbf{g})$ *and where* C *and* $\widetilde{C}$ *respectively denote the maximum absolute value of the coefficient of all* g_i *and* $\widetilde{g}_i$*. Here* $\mathfrak{S}(\mathbf{g})$ *is the Hardy–Littlewood singular series associated to the system* $\mathbf{g}=\mathbf{0}$ *and* $J(\widetilde{\mathbf{g}})$ *is the Hardy–Littlewood singular integral associated to the system* $\widetilde{\mathbf{g}}=\mathbf{0}$*.*

Using this, we obtain the following.

Lemma 4.4. *There exist $M_2, \eta_3 > 0$ that only depend on $\mathbf{f}$ such that for all $\mathbf{t} \in (\mathbb{Z}/Q\mathbb{Z})^n$ the quantity*

$$\#\{\mathbf{x} \in \mathbb{Z}^n : \max_i |x_i| \leqslant B, \mathbf{f}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \equiv \mathbf{t} \bmod Q\}$$

equals

$$\sigma_\infty \left(\prod_{p \mid Q} \sigma_p \left(\mathbf{t}, p^{\nu_p(Q)} \right) \right) \left(\prod_{p \nmid Q} \sigma_p \right) B^{n-RD} + O_{\mathbf{f}} \left(B^{n-RD-\eta_3} Q^{M_2} \right),$$

where the implied constant depends at most on $\mathbf{f}$. Here, σ_∞ is the standard Hardy–Littlewood singular integral associated to the system $\mathbf{f} = \mathbf{0}$,

$$\sigma_p := \lim_{m \to +\infty} \frac{\# \{ \mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \bmod p^m \}}{p^{m(n-R)}},$$

and for all $e \geqslant 1$ and $\mathbf{s} \in (\mathbb{Z}/p^e\mathbb{Z})^n$ we denote

$$\sigma_p(\mathbf{s}, p^e) := \lim_{m \to +\infty} \frac{\# \{ \mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \bmod p^m, \mathbf{x} \equiv \mathbf{s} \bmod p^e \}}{p^{m(n-R)}}.$$

Proof. We first deal with the case $Q > B$. In this instance we plainly have

$$\#\{\mathbf{x} \in \mathbb{Z}^n : \max_i |x_i| \leqslant B, \mathbf{x} \equiv \mathbf{t} \bmod Q\} \ll 1,$$

which is clearly $\ll B^{n-RD}$ since the Birch rank assumption implies $n > RD$. The estimate $\sigma_p(\mathbf{s}, p^e) \leqslant \sigma_p$ shows that we always have

$$\sigma_\infty \left(\prod_{p \mid Q} \sigma_p \left(\mathbf{t}, p^{\nu_p(Q)} \right) \right) \left(\prod_{p \nmid Q} \sigma_p \right) B^{n-RD} \leqslant \sigma_\infty \left(\prod_{p} \sigma_p \right) B^{n-RD} \ll_{\mathbf{f}} B^{n-RD}.$$

Therefore, for all $\eta \in (0, 1]$ and $M \geqslant 1$ one has

$$B^{n-RD} < B^{n-RD-1} Q \leqslant B^{n-RD-\eta} Q^M.$$

We are then free to assume that $Q \leqslant B$ for the rest of the proof.

Without loss of generality we can assume that $\mathbf{t} \in (\mathbb{Z} \cap [0, Q))^n$. We then use the change of variables $\mathbf{x} = \mathbf{t} + Q\mathbf{z}$ to write the counting function in our lemma as

$$= \# \left\{ \mathbf{z} \in \mathbb{Z}^n : \max_i \left| z_i + \frac{t_i}{Q} \right| \leqslant \frac{B}{Q}, \mathbf{f}(\mathbf{t} + Q\mathbf{z}) = \mathbf{0} \right\}.$$

We now apply Lemma 4.3 with $g(\mathbf{z}) := \mathbf{f}(\mathbf{t} + Q\mathbf{z})$. We have

$$|z_i| \leqslant \frac{B}{Q} - 1 \Rightarrow \left| z_i + \frac{y_i}{Q} \right| \leqslant \frac{B}{Q} \Rightarrow |z_i| \leqslant \frac{B}{Q} + 1,$$

therefore, if we let $B_1 = \frac{B}{Q} - 1$ and $B_2 = \frac{B}{Q} + 1$ we see that

$$\#\{\mathbf{z} \in \mathbb{Z}^n : \max_i |z_i| \leqslant B_j, \mathbf{g}(\mathbf{z}) = \mathbf{0}\}, j = 1, 2$$

give lower and upper bounds for the counting function in our lemma, respectively. We note that $\mathbf{f}$ and $\mathbf{g}$ are related via a non-singular linear change of variables, hence $\mathfrak{B}(\mathbf{g}) = \mathfrak{B}(\mathbf{f})$. By Lemma 4.3 and $B_j = B/Q + O(1)$ we therefore obtain

$$J(\widetilde{\mathbf{g}}) \left(\prod_p \tau_p \right) (B/Q + O(1))^{n-RD} + O \left(B^{n-rD-\eta_2} C \widetilde{C}^{M_1} \right),$$

where

$$J(\widetilde{\mathbf{g}})=\int_{\boldsymbol{\gamma}\in\mathbb{R}^R}\left(\int_{\boldsymbol{\zeta}\in\mathbb{R}^n}\exp\left(2\pi i\sum_{i=1}^R\gamma_i\widetilde{g}_i(\boldsymbol{\zeta})\right)\mathrm{d}\boldsymbol{\zeta}\right)\mathrm{d}\boldsymbol{\gamma} \tag{4.1}$$

and

$$\tau_p:=\lim_{m\to+\infty}\frac{\#\left\{\mathbf{z}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{g}(\mathbf{z})\equiv\mathbf{0}\bmod p^m\right\}}{p^{m(n-R)}}. \tag{4.2}$$

We note that $\widetilde{\mathbf{g}}(\mathbf{z})=\mathbf{f}(Q\mathbf{z})=Q^D\mathbf{f}(\mathbf{z})$, therefore, $\widetilde{C}\ll_{\mathbf{f}}Q^D$. The bound $0\leqslant t_i\leqslant Q$ and the identity $(Qz_i+t_i)^k=\sum_{i=0}^k\left(\binom{k}{j}Q^jt_i^{k-j}\right)z_i^j$ imply that $C\ll_{\mathbf{f}}Q^D$. We conclude that $C\widetilde{C}^{M_1}\ll_{\mathbf{f}}Q^{D(1+M_1)}$. Using the bound $Q\leqslant B$ we have

$$(B/Q+O(1))^{n-RD}=(B/Q)^{n-RD}+O((B/Q)^{n-RD-1}),$$

which leads to the quantity in our lemma being equal to

$$J(\widetilde{\mathbf{g}})\left(\prod_p\tau_p\right)\left(\frac{B}{Q}\right)^{n-RD}+O\left(B^{n-rD-\eta_2}Q^{D(1+M_1)}+\left|\left(\prod_p\tau_p\right)J(\widetilde{\mathbf{g}})\right|(B/Q)^{n-RD-1}\right). \tag{4.3}$$

Using $\widetilde{\mathbf{g}}(\boldsymbol{\zeta})=\mathbf{f}(Q\boldsymbol{\zeta})=Q^D\mathbf{f}(\boldsymbol{\zeta})$ and the change of variables $Q^D\boldsymbol{\gamma}=\boldsymbol{\beta}$ shows that $J(\widetilde{\mathbf{g}})$ is

$$\begin{aligned}&\int_{\boldsymbol{\gamma}\in\mathbb{R}^R}\left(\int_{\boldsymbol{\zeta}\in\mathbb{R}^n}\exp\left(2\pi i\sum_{i=1}^R(Q^D\gamma_i)f_i(\boldsymbol{\zeta})\right)\mathrm{d}\boldsymbol{\zeta}\right)\mathrm{d}\boldsymbol{\gamma}\\&=Q^{-RD}\int_{\boldsymbol{\gamma}\in\mathbb{R}^R}\left(\int_{\boldsymbol{\zeta}\in\mathbb{R}^n}\exp\left(2\pi i\sum_{i=1}^R\beta_if_i(\boldsymbol{\zeta})\right)\mathrm{d}\boldsymbol{\zeta}\right)\mathrm{d}\boldsymbol{\beta}.\end{aligned}$$

This is clearly $Q^{-RD}J(\mathbf{f})$, in other words, we have seen that $J(\widetilde{\mathbf{g}})=Q^{-RD}\sigma_\infty$. This converts (4.3) into

$$Q^{-n}\sigma_\infty\left(\prod_p\tau_p\right)B^{n-RD}+O_{\mathbf{f}}\left(B^{n-RD-\eta_2}Q^{D(1+M_1)}+\left(\prod_p\tau_p\right)B^{n-RD-1}Q^{-n+1}\right). \tag{4.4}$$

For a prime $p\nmid Q$ the change of variables $(\mathbb{Z}/p^m\mathbb{Z})^n\to(\mathbb{Z}/p^m\mathbb{Z})^n$, $\mathbf{z}\mapsto\mathbf{x}$ that is given by $\mathbf{x}\equiv\mathbf{t}+Q\mathbf{z}\bmod p^m$ is invertible modulo p^m, therefore, the numerator within the limit in (4.2) equals

$$\#\left\{\mathbf{z}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{f}(\mathbf{t}+Q\mathbf{z})\equiv\mathbf{0}\bmod p^m\right\}=\#\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{f}(\mathbf{x})\equiv\mathbf{0}\bmod p^m\right\}.$$

In particular, (4.2) agrees with σ_p. If $p\mid Q$ a similar argument, with the map under consideration being $\mathbf{x}\equiv Qp^{-\nu_p(Q)}\mathbf{z}\bmod p^k$, shows that

$$\#\left\{\mathbf{z}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{f}(\mathbf{t}+Q\mathbf{z})\equiv\mathbf{0}\bmod p^m\right\}=\#\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{f}(\mathbf{t}+p^{\nu_p(Q)}\mathbf{x})\equiv\mathbf{0}\bmod p^m\right\}.$$

We can clearly rewrite this as

$$\sum_{\substack{\mathbf{w}\in(\mathbb{Z}/p^m\mathbb{Z})^n\\\mathbf{f}(\mathbf{w})\equiv\mathbf{0}\bmod p^m}}\#\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^n:\mathbf{w}\equiv\mathbf{t}+p^{\nu_p(Q)}\mathbf{x}\bmod p^m\}.$$

Now assume that $m > \nu_p(Q)$. Then the inner sum contribution is non-zero only when $\mathbf{w} \equiv \mathbf{t} \bmod p^{\nu_p(Q)}$. We thus obtain

$$\sum_{\substack{\mathbf{w}\in(\mathbb{Z}/p^m\mathbb{Z})^n \\ \mathbf{f}(\mathbf{w})\equiv\mathbf{0} \bmod p^m \\ \mathbf{w}\equiv\mathbf{t} \bmod p^{\nu_p(Q)}}} \#\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^n : \mathbf{x}\equiv\frac{\mathbf{w}-\mathbf{t}}{p^{\nu_p(Q)}} \bmod p^{m-\nu_p(Q)}\right\}.$$

The new inner cardinality clearly equals $p^{n\nu_p(Q)}$ since every $x_i \bmod p^m$ is uniquely determined modulo $p^{m-\nu_p(Q)}$. This gives the following for all $p \mid Q$,

$$\tau_p = p^{n\nu_p(Q)} \lim_{m\to+\infty} \frac{\#\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^n : \mathbf{f}(\mathbf{x})\equiv\mathbf{0} \bmod p^m, \mathbf{x}\equiv\mathbf{t} \bmod p^{\nu_p(Q)}\right\}}{p^{m(n-R)}}.$$

This is clearly at most $p^{n\nu_p(Q)}\sigma_p$, therefore,

$$\prod_p \tau_p = \prod_{p\nmid Q}\sigma_p \prod_{p\mid Q} p^{n\nu_p(Q)}\sigma_p(\mathbf{t}, p^{\nu_p(Q)}) \leqslant Q^n \prod_p \sigma_p,$$

which, when injected in (4.4), shows that the asymptotic in our lemma holds with an error term

$$\ll_{\mathbf{f}} B^{n-RD-\eta_2}Q^{D(1+M_1)} + B^{n-RD-1}Q \ll B^{n-RD-\min\{\eta_2,1\}}Q^{D(1+M_1)}.$$

Letting $M_2 = \max\{1, D(1+M_1)\}$ and $\eta_3 := \min\{\eta_2, 1\}$ concludes the proof of the lemma. □

Lemma 4.5. *There exist $M_3, \eta_4 > 0$ that only depend on $\mathbf{f}$ such that for all $Q \in \mathbb{N}$ and $\Upsilon \subset \mathcal{X}(\mathbb{Z}/Q\mathbb{Z})$ we have*

$$N(\Upsilon, B) = B^{n-RD}\frac{\sigma_\infty}{2}\left(\prod_{p\nmid Q}\sigma_p\left(1-\frac{1}{p^{n-RD}}\right)\right)\sum_{\mathbf{y}\in\widehat{\Upsilon}}\prod_{p\mid Q}\sigma_p(\mathbf{y}, p^{\nu_p(Q)}) + O_{\mathbf{f}}\left(B^{n-RD-\eta_4}Q^{M_3}\right),$$

where the implied constant depends at most on $\mathbf{f}$.

Proof. Injecting Lemma 4.4 into Lemma 4.2 shows that $N(\Upsilon, B)$ equals

$$\frac{1}{2}\sum_{\mathbf{y}\in\widehat{\Upsilon}}\sum_{\substack{k\in\mathbb{N}\cap[1,B^{\eta_1}] \\ \gcd(k,Q)=1}} \mu(k)\left(\sigma_\infty\prod_{p\nmid Q}\sigma_p\prod_{p\mid Q}\sigma_p(\mathbf{y}/k, p^{\nu_p(Q)})\left(\frac{B}{k}\right)^{n-RD} + O\left(Q^{M_2}\left(\frac{B}{k}\right)^{n-RD-\eta_3}\right)\right)$$
$$+ O_{\mathbf{f}}\left(Q^n B^{n-RD-\eta_1}\right).$$

The fact that $\gcd(k, Q) = 1$ shows that for $p \mid Q$ we have $\sigma_p(\mathbf{y}/k, p^{\nu_p(Q)}) = \sigma_p(\mathbf{y}, p^{\nu_p(Q)})$. Furthermore, the trivial estimate $\#\widehat{\Upsilon} \leqslant Q^n$ shows that $N(\Upsilon, B)$ equals

$$B^{n-RD}\frac{\sigma_\infty}{2}\left(\prod_{p\nmid Q}\sigma_p\right)\left(\sum_{\substack{k\in\mathbb{N}\cap[1,B^{\eta_1}] \\ \gcd(k,Q)=1}}\frac{\mu(k)}{k^{n-RD}}\right)\sum_{\mathbf{y}\in\widehat{\Upsilon}}\prod_{p\mid Q}\sigma_p(\mathbf{y}, p^{\nu_p(Q)})$$

up to a term whose modulus is

$$\ll Q^{n+M_2}\sum_{k\in\mathbb{N}\cap[1,B^{\eta_1}]}\left(\frac{B}{k}\right)^{n-RD-\eta_3} + Q^n B^{n-RD-\eta_1} \ll Q^{n+M_2}B^{n-RD-\eta_3+\eta_1}$$

due to the bound $n - RD - \eta_3 \geqslant 0$ and $\sum_{k \leqslant B^{\eta_1}} 1 \ll B^{\eta_1}$. This is admissible as can be seen by taking $\eta_1 = \eta_3/2$ in Lemma 4.2. The main term contains a sum over $k \in [1, B^{\eta_1}]$ that can be written as

$$\prod_{p \nmid Q} \left(1 - \frac{1}{p^{n-RD}}\right) + O\left(\sum_{k > B^{\eta_1}} \frac{1}{k^2}\right) = \prod_{p \nmid Q} \left(1 - \frac{1}{p^{n-RD}}\right) + O(B^{-\eta_1}),$$

because our assumptions on the Birch rank ensure that $n - RD \geqslant 2$. We plainly have $\sigma_p(\mathbf{y}, p^{\nu_p(Q)}) \leqslant \sigma_p$ hence the contribution of the last term $O(B^{-\eta_1})$ is

$$\ll_{\mathbf{f}} \left(\prod_{p \nmid Q} \sigma_p\right) B^{n-RD-\eta_1} \sum_{\mathbf{y} \in \hat{\Upsilon}} \prod_{p \mid Q} \sigma_p(\mathbf{y}, p^{\nu_p(Q)}) \leqslant \left(\prod_{p \nmid Q} \sigma_p\right) B^{n-RD-\eta_1} \sum_{\mathbf{y} \in \hat{\Upsilon}} \prod_{p \mid Q} \sigma_p,$$

which is $(\prod_p \sigma_p) B^{n-RD-\eta_1} \# \Upsilon \ll_{\mathbf{f}} B^{n-RD-\eta_1} \# \Upsilon \leqslant B^{n-RD-\eta_1} Q^n$. This is admissible. The main term is

$$B^{n-RD} \frac{\sigma_\infty}{2} \left(\prod_{p \nmid Q} \sigma_p \left(1 - \frac{1}{p^{n-RD}}\right)\right) \sum_{\mathbf{y} \in \hat{\Upsilon}} \prod_{p \mid Q} \sigma_p(\mathbf{y}, p^{\nu_p(Q)}),$$

which is as stated in our lemma. □

We now record the end result of our investigation, which may be of independent interest. For completeness, we recall our assumptions.

Proposition 4.6 (Effective equidistribution for Birch systems). *Assume that $f_1, \ldots, f_R$ are integer homogeneous polynomials in n variables, all of the same degree D and that the Birch rank satisfies $\mathfrak{B}(\mathbf{f}) > 2^{D-1}(D-1)R(R+1)$. Assume that $f_1 = \cdots = f_R = 0$ is smooth, that it has a $\mathbb{Q}$-rational point and denote by $\mathcal{X}$ the model given by taking its closure in $\mathbb{P}_{\mathbb{Z}}^{n-1}$. Then there exist positive constants A, M, η that only depend on $\mathbf{f}$ such that for all $Q \in \mathbb{N}$ only divisible by primes $p > A$ and all $\Upsilon \subset \mathcal{X}(\mathbb{Z}/Q\mathbb{Z})$, we have*

$$\frac{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B, x \bmod Q \in \Upsilon\}}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} = \frac{\#\Upsilon}{\#\mathcal{X}(\mathbb{Z}/Q\mathbb{Z})} + O(B^{-\eta} Q^M),$$

where the implied constant is independent of B and Q.

Proof. Using Birch's theorem [4, Thm. 1] and Möbius inversion, there exists $\eta_5 > 0$ that only depends on $\mathbf{f}$ such that

$$\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\} = \frac{1}{2\zeta(n - RD)} \sigma_\infty \left(\prod_p \sigma_p\right) B^{n-RD} + O_{\mathbf{f}}\left(B^{n-RD-\eta_5}\right).$$

Moreover, our assumptions that $X(\mathbb{Q}) \neq \varnothing$ and that X is smooth implies that $\sigma_\infty > 0$ and $\sigma_p > 0$ for all primes p. Then Lemma 4.5 gives

$$\frac{N(\Upsilon, B)}{\#\{x \in \mathcal{X}(\mathbb{Z}) : H(x) \leqslant B\}} = \sum_{\mathbf{y} \in \hat{\Upsilon}} \prod_{p \mid Q} \frac{\sigma_p(\mathbf{y}, p^{\nu_p(Q)})}{\sigma_p \left(1 - \frac{1}{p^{n-RD}}\right)} + O_{\mathbf{f}}\left(B^{-\min\{\eta_4, \eta_5\}} Q^{n+M_3}\right)$$

with an implied constant depending at most on $\mathbf{f}$. The result now follows from Lemma 4.1. □

Proposition 4.6 proves the equidistribution property (1.2), hence, we may apply Theorem 1.8. To finish, it suffices to explain why we obtain the identity covariance matrix.

Lemma 4.7. *Let k be a field of characteristic zero and $Y \subset \mathbb{P}^d$ a smooth complete intersection with $\dim Y \geqslant 3$, which is not contained in a hyperplane. Then $Y \cap H$ is irreducible for any hyperplane $H \subset \mathbb{P}^d$.*

Proof. The Lefschetz hyperplane section theorem implies that $\operatorname{Pic} Y \cong \mathbb{Z}$ generated by the hyperplane class [16, Exposé XII, Cor. 3.7]. Thus if $H \cap Y = D_1 + D_2$ for effective divisors D_1 and D_2, we must have $[D_i] = 0$ for some i; the result follows. □

Thus the intersections with the coordinate hyperplanes are irreducible, so they contain no common irreducible components. The result therefore follows from Theorem 1.8. This completes the proof of Theorem 1.2, and Theorem 1.1 follows immediately.

4.2. **Homogeneous spaces.** Counting integral points on homogeneous spaces has a long history and we only mention a few relevant milestones: Duke, Rudnick and Sarnak [9] used spectral analysis to deal with a class of (affine) symmetric varieties, Gorodnik and Nevo [14] used the mean ergodic theorem in order to obtain error terms with a power saving; Nevo and Sarnak [22] have recently applied such counting results to the problem of finding (and estimating the number of) prime or almost-prime points on such varieties.

In this paper we consider the class of symmetric varieties studied by Browning and Gorodnik in [7] and begin by recalling their set-up, which is more general than [22]. Let G be a connected semisimple algebraic group defined over $\mathbb{Q}$ and let $\iota\colon G \to \mathrm{GL}_n$ be a linear representation defined over $\mathbb{Q}$ with finite kernel. Let $Y \subset \mathbb{A}^n_{\mathbb{Q}}$ be a subvariety which is left invariant under the action of G via ι. We assume that G acts transitively on Y, so that Y has the form G/L where L is an algebraic subgroup defined over $\mathbb{Q}$. We denote by $Y(\mathbb{Z}) = \mathcal{Y}(\mathbb{Z})$, where $\mathcal{Y} \subset \mathbb{A}^n_{\mathbb{Z}}$ is the model given by the closure of Y in $\mathbb{A}^n_{\mathbb{Z}}$. We assume that $Y(\mathbb{Z}) \neq \varnothing$. Moreover, the following assumptions are made:

1. L is a symmetric subgroup of G, meaning the Lie algebra of L is the fixed locus of a non-trivial involution defined over $\mathbb{Q}$;
2. the connected component of L has no non-trivial $\mathbb{Q}$-rational characters;
3. the group G is $\mathbb{Q}$-simple;
4. the group $G(\mathbb{R})$ is connected and has no compact factors.

This is the class of symmetric varieties Y which we shall be interested in. For $\mathbf{y} \in Y(\mathbb{Z})$, we define its height by $H(\mathbf{y}) = \max_{i \in \{1,\dots,n\}} |y_i|$. We use the following result, stated in [7, Prop. 3.1].

Proposition 4.8. *There exists $\delta > 0$ such that for every $\ell \in \mathbb{N}$ and every $\xi \in Y(\mathbb{Z}/\ell\mathbb{Z})$ we have*

$$\begin{aligned}\#\{\mathbf{y} \in Y(\mathbb{Z}) : H(\mathbf{y}) \leqslant B, \mathbf{y} \equiv \xi \bmod \ell\}& \\ = \mu_\infty(Y;B) \prod_{p \text{ prime}} \hat{\mu}_p(Y;\xi,\ell) &+ O(\ell^{\dim(L)+2\dim(G)} \mu_\infty(Y;B)^{1-\delta})\end{aligned}$$

as $B \to \infty$, where

$$\hat{\mu}_p(Y;\xi,\ell) = \lim_{t\to\infty} p^{-t\dim(Y)} \#\{\mathbf{y} \in Y(\mathbb{Z}/p^t\mathbb{Z}) : \mathbf{y} \equiv \xi \bmod p^{v_p(\ell)}\}$$

is the p-adic density and $\mu_\infty(Y;B)$ is the real density, as defined in [7, (1.6)].

It follows from this and Hensel's lemma that

$$\begin{aligned}&\#\{\mathbf{y} \in Y(\mathbb{Z}) : H(\mathbf{y}) \leqslant B, \mathbf{y} \equiv \xi \bmod \ell\} \\ &= \frac{\#\{\mathbf{y} \in Y(\mathbb{Z}/\ell\mathbb{Z}) : \mathbf{y} = \xi\}}{\#Y(\mathbb{Z}/\ell\mathbb{Z})} N(Y,B) + O\left(\ell^{\dim(L)+2\dim(G)} N(Y,B)^{1-\delta}\right),\end{aligned}$$

where $N(Y,B)(\mathbb{Z}) = \#\{\mathbf{x} \in Y(\mathbb{Z}) : H(\mathbf{x}) \leqslant B\}$. The effective equidistribution property (1.2) is therefore easily seen to hold in this case. Theorem 1.8 thus shows the following.

Theorem 4.9. *Let $Y \subset \mathbb{A}^n$ be a symmetric variety in the class described above and let $\Omega_B = \{\mathbf{y} \in Y(\mathbb{Z}) : H(\mathbf{y}) \leqslant B, y_1 \cdots y_n \neq 0\}$ be equipped with the uniform probability measure. Then as $B \to \infty$ the random vectors*

$$\Omega_B \to \mathbb{R}^n, \quad \mathbf{y} \mapsto \left(\frac{\omega(y_1) - c_{1,1} \log\log B}{\sqrt{c_{1,1} \log\log B}}, \dots, \frac{\omega(y_n) - c_{n,n} \log\log B}{\sqrt{c_{n,n} \log\log B}} \right)$$

converge in distribution to the central multivariate distribution with covariance matrix whose (i,j)-entry is $c_{i,j}$, the number of common irreducible components of $y_i = 0$ and $y_j = 0$ in Y.

Of course Theorem 1.8 also gives a version for general divisors D_i. We now explain how Theorem 1.3 and Theorem 1.4 are corollaries of the above theorem and why the covariance matrix is the identity in these examples.

Proof of Theorem 1.3. That the varieties in Theorem 1.3 fall under the setting of this section is explained in [7, Rem. 1.3]. The conclusion now follows immediately from Theorem 4.9, as it is easily checked that the intersection with each coordinate hyperplane is irreducible (for $n = 3$ this follows from our assumption that $-k \operatorname{disc}(Q)$ is not a perfect square). □

Proof of Theorem 1.4. That the varieties in Theorem 1.4 fall under the setting of this section can be seen as follows. Let $G = \mathrm{SL}_n \times \mathrm{SL}_n$ act on the space $\mathcal{M}_n$ of $n \times n$ matrices by mapping $M \in \mathcal{M}_n$ to $g^{-1}Mh$, for $(g,h) \in G$. Then $V_{n,k} = G/L$, with $L = \mathrm{SL}_n$ being diagonally embedded in G. Here again the conclusion easily follows from the fact that the intersection with each coordinate hyperplane is irreducible. This can be proved, for example, by applying a suitable version of the Lefschetz hyperplane section to the intersection of a hyperplane with the projectivised hypersurface $\det(M) = kz^n$. □

Remark 4.10. Let us note that for general choices of symmetric varieties $Y \subset \mathbb{A}^n$ in Theorem 4.9, one can obtain non-identity covariance matrices. For example, let $\sigma_d : \mathrm{SL}_n \to \mathrm{GL}_N$ be the dth symmetric power representation and take $G = \mathrm{SL}_n \times \mathrm{SL}_n$ with the representation $G \to \mathrm{GL}_N, (g,h) \mapsto \sigma_d(g)^{-1}\sigma_d(h)$. Then Y, given by the orbit of the identity matrix, has the stated property for $d > 1$ (this is a variant of the construction in Remark 1.13).

4.3. **Conics.** We now prove our results on conics from §1.2.3.

4.3.1. *Proof of Theorem 1.5.* Any smooth conic with a rational point is isomorphic to the projective line. The effective equidistribution property (1.2) is known to hold for the projective line [21, Prop. 2.1]. The result *loc. cit.* is proved for the standard height on $\mathbb{P}^1_{\mathbb{Q}}$, but a minor modification shows that property (1.2) in fact holds for more general choices of height function, for some choice of M and η, which in particular shows that the hypotheses of Theorem 1.8 hold in this case. This therefore immediately gives the result. □

4.3.2. *Proof of Theorem 1.7.* Let C be as in Theorem 1.7. Let $D_i' : x_i = 0$ and $D_i = D'_{i,\mathrm{red}}$. As the covariance matrix is singular, Theorem 1.8 shows that there is a linear relation between the divisors D_i in $\operatorname{Div} C$. But we have $D_i' = b_i D_i$ for some $b_i \in \{1, 2\}$. Thus this also gives a relation

$$a_0 D_0' + a_1 D_1' + a_2 D_2' = 0$$

between the D_i'. We take the minimal such relation, so that $\gcd(a_0, a_1, a_2) = 1$. Moreover, as $\deg D_i' = 2$, we find that $a_0 + a_1 + a_2 = 0$. Changing signs as required and permuting coordinates, we obtain the relation

$$x_1^{a_1} x_2^{a_2} = c x_0^{a_1 + a_2}$$

in the homogeneous coordinate ring of C, for some $c \in \mathbb{Q}$. But the only relation in this ring is the equation of the conic $C : Q(x_0, x_1, x_2) = 0$, hence $Q \mid x_1^{a_1} x_2^{a_2} - c x_0^{a_1+a_2}$. But $\gcd(a_1, a_2) = 1$, implies that this polynomial is irreducible, hence we must have $Q = c'(x_1^{a_1} x_2^{a_2} - c x_0^{a_1+a_2})$ for some $c' \in \mathbb{Q}$. As $\deg Q = 2$ we have $a_1 = a_2 = 1$, as required. □

4.4. **A cubic surface.** Consider the cubic surface

$$X : \quad x_1 x_2 x_3 = x_0^3 \qquad \subset \mathbb{P}^3_{\mathbb{Q}}.$$

With respect to the coordinate hyperplanes $x_i = 0$, we conjecture that an analogue of Theorem 1.8 holds with covariance matrix

$$\begin{pmatrix} 1 & \sqrt{5}/3 & \sqrt{5}/3 & \sqrt{5}/3 \\ \sqrt{5}/3 & 1 & 2/5 & 2/5 \\ \sqrt{5}/3 & 2/5 & 1 & 2/5 \\ \sqrt{5}/3 & 2/5 & 2/5 & 1 \end{pmatrix}. \tag{4.5}$$

Let us explain how we obtained this. First X is singular, and the counting problem should really take place on the minimal desingularisation $\widetilde{X}$ of X. We then naively apply the formula from Theorem 1.8 with respect to the divisors D_i on $\widetilde{X}$ given by the pull backs of the (reduced) hyperplanes $H_i : x_i = 0$. For $i \neq 0$ the H_i are the lines $x_i = x_0 = 0$, whereas H_0 is the union of these three lines. Any two lines meet in a singular point of type A_2, and these are all the singular points. The singularities are resolved by blowing-up twice, which introduces 2 new exceptional curves. As any line contains two singular points, a calculation using the above considerations shows that

$$c_{i,i} = \begin{cases} 5, & i \in \{1,2,3\}, \\ 9, & i = 0. \end{cases} \qquad c_{i,j} = \begin{cases} 2, & i \neq j \in \{1,2,3\}, \\ 5, & i = 0, j \in \{1,2,3\}. \end{cases}$$

The formula (4.5) now easily follows. Note that this matrix is singular, due to the obvious relation $H_0 = H_1 + H_2 + H_3$.

How would one go about proving this? Firstly, it follows from [8, §3.10] that the rational points on $\widetilde{X}$ are equidistributed, which gives the main term in (1.2). One can prove that

$$\#\{x \in \widetilde{\mathcal{X}}(\mathbb{Z}) : H(x) \leqslant B, x \bmod Q \in \Upsilon\} = B P_\Upsilon(\log B) + O_\Upsilon(B^{1-\eta}), \tag{4.6}$$

where $\eta > 0$ and P_Υ is a polynomial of degree 6 whose coefficients depend on Υ. By equidistribution one understands the leading coefficient of P_Υ; the challenge lies with controlling the dependence on the lower order terms of P_Υ and whether an asymptotic formula of the shape (4.6), with powers of $\log B$ appearing, can be used to obtain an Erdős–Kac law.

References

[1] P. Billingsley, The probability theory of additive arithmetic functions. *Ann. Probability*, **2**, (1974), 749–791.

[2] ———, *Probability and measure.* Third Edition, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1995.

[3] ———, *Convergence of probability measures.* Second edition, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.

[4] B. J. Birch, Forms in many variables. *Proc. Roy. Soc. London Ser. A* **265** (1961/62), 245–263.

[5] M. Borovoi, Z. Rudnick, Hardy–Littlewood varieties and semisimple groups. *Invent. Math.* **119**, (1995), 37–66.

[6] J. Bourgain, A. Gamburd, P. Sarnak, Affine linear sieve, expanders, and sum-product. *Invent. Math.* **179**, (2010), 559–644.

[7] T.D. Browning, A. Gorodnik, Power-free values of polynomials on symmetric varieties. *Proc. London Math. Soc.* **114**, (2017), 1044–1080.

[8] A. Chambert-Loir, Y. Tschinkel, Integral points of bounded height on toric varieties. `arxiv:1006.3345`.

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

[10] R. Durret, *Probability: Theory and Examples*, Fifth edition. Cambridge Series in Statistical and Probabilistic Mathematics, **49**. Cambridge University Press, Cambridge, 2019.

[11] D. El-Baz, An analogue of the Erdős-Kac theorem for the special linear group over the integers. *Acta Arith*, to appear.

[12] P. Erdős, M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, **62**, (1940), 738–742.

[13] P. Erdős, On the distribution function of additive functions. *Ann. of Math.*, **47**, (1946), 1–20.

[14] A. Gorodnik, A. Nevo, Quantitative ergodic theorems and their number-theoretic applications. *Bull. Amer. Math. Soc. (N.S.)*, **52**, (2015), 65–113.

[15] A. Granville, K. Soundararajan, Sieving and the Erdős-Kac theorem. *Equidistribution in number theory, an introduction*, NATO Sci. Ser. II Math. Phys. Chem., Springer, Dordrecht, **237**, (2007), 15–27.

[16] A. Grothendieck, Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux (SGA 2). Documents Mathématiques (Paris), 4. Société Mathématique de France, Paris, 2005.

[17] H. Halberstam, On the distribution of additive number-theoretic functions. II. *J. London Math. Soc.*, **31**, (1956), 1–14.

[18] ———, On the distribution of additive number-theoretic functions. III. *J. London Math. Soc.*, **31**, (1956), 14–27.

[19] S. Lang, A. Weil, Number of points of varieties in finite fields. *Amer. J. Math.* **76** (1954), 819–827.

[20] W. LeVeque, On the size of certain number-theoretic functions. *Trans. Amer. Math. Soc.* **66** (1949), 440–463.

[21] D. Loughran, E. Sofos, An Erdős-Kac law for local solubility in families of varieties. *Selecta Math.*, to appear.

[22] A. Nevo, P. Sarnak, Prime and almost prime integral points on principal homogeneous spaces. *Acta Math.* **205**, (2010), 361–402.

[23] E. Peyre, *Hauteurs et mesures de Tamagawa sur les variétés de Fano*, Duke Math. J. **79** (1995), no. 1, 101–218.

[24] P. Salberger, Tamagawa measures on universal torsors and points of bounded height on Fano varieties. *Astérisque* **251** (1998), 91–258.

[25] J.-P. Serre, *Lectures on* $N_X(p)$. Chapman & Hall/CRC Research Notes in Mathematics, 11. CRC Press, Boca Raton, FL, 2012.

[26] M. Tanaka, On the number of prime factors of integers. *Jap. J. Math.*, **25** (1956), no. 8, 1–20.

[27] G. Tenenbaum, *Introduction to analytic and probabilistic number theory.* Cambridge Studies in Advanced Mathematics, **46**, Cambridge University Press, Cambridge, 1995.

[28] J.-W. M. van Ittersum, Quantitative results on Diophantine equations in many variables. *Acta Arithmetica*, to appear, *arXiv:1709.05126*, (2017).

[29] M. Xiong, The Erdős-Kac theorem for polynomials of several variables. *Proc. Amer. Math. Soc.*, **137** (2009), no. 8, 2601–2608.

Graz University of Technology, Institute of Analysis and Number Theory, Steyrergasse 30/II, 8010 Graz, Austria
Email address: el-baz@math.tugraz.at

Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK
URL: https://sites.google.com/site/danielloughran

Department of Mathematics, University of Glasgow, University Place, Glasgow, G12 8QQ, UK
Email address: efthymios.sofos@glasgow.ac.uk