6-torsion and integral points on quartic surfaces

Working Paper

Author(s):
Chan, Stephanie; Koymans, Peter; Pagano, Carlo; Sofos, Efthymios

Publication date:
2024-03-20

Permanent link:
https://doi.org/10.3929/ethz-b-000684649

Originally published in:
arXiv

6-TORSION AND INTEGRAL POINTS ON QUARTIC SURFACES

S. CHAN, P. KOYMANS, C. PAGANO, AND E. SOFOS

Abstract. We prove matching upper and lower bounds for the average of the 6-torsion of class groups of quadratic fields. Furthermore, we count the number of integer solutions on an affine quartic surface.

Contents

1. Introduction

One of the main invariants of class groups of quadratic fields $\mathbb{Q}(\sqrt{D})$ is the size $h_n(D)$ of their n-torsion. It has been investigated by several mathematicians: By the work of Gauss [32] in 1801 the average of $h_2(D)$ for $D < 0$ is a constant multiple of $\log |D|$ when ordering the number fields by $-D$. Davenport and Heilbronn [19] proved in 1971 that $h_3(D)$ has a constant average, while, Fouvry and Klüners [27, 28] in 2007 showed that $h_4(D)$ is on average a constant multiple of $\log |D|$. The influential work of Smith [50] in 2017 established the complete distribution of $h_{2^k}(D)$. There are no other values of n for which the right order of magnitude is known. For general n, there is work on bounds for $h_n(D)$ on average by Soundararajan [52], Heath-Brown–Pierce [34], Frei–Widmer [30] and Koymans–Thorner [40].

The Cohen–Lenstra conjectures [17] predict that $h_n(D)$ is of constant average for n odd and is $\log |D|$ on average for n even. Let $\mathcal{D}^+(X)$ and $\mathcal{D}^-(X)$ be the set of respectively positive and negative fundamental discriminants with absolute value up to X. In this paper we establish the right order of magnitude for the 6-torsion:

Theorem 1.1. *For all $X \geqslant 5$ we have*

$$X \log X \ll \sum_{D \in \mathcal{D}^+(X)} h_6(D) \ll X \log X \qquad \text{and} \qquad X \log X \ll \sum_{D \in \mathcal{D}^-(X)} h_6(D) \ll X \log X.$$

This marks the first time that Nair–Tenenbaum techniques are applied in arithmetic statistics, clarified in the subsequent remark:

Remark 1.2 (Idea of the proof of Theorem 1.1)**.** Using the Davenport–Heilbronn parametrisation we turn the sum $\sum_D h_6(D)$ into an average of the function $2^{\omega(m)}$ over the values m assumed by a polynomial in 4 variables, where the integer vectors lie in a subset of $\mathbb{R}^4$ with spikes. This average is a special instance of sums of the following form:

$$\sum_{a \in \mathcal{A}} f(c_a) \chi(c_a), \tag{1.1}$$

where

2020 *Mathematics Subject Classification.* 11N37 (primary), 11R29, 11D45 (secondary).

- $\mathcal{A}$ is a countable set,
- $\chi : \mathcal{A} \to [0, \infty)$ is any function of finite support,
- c_a is an "equidistributed" sequence of positive integers,
- f is a non-negative arithmetic function being multiplicative or more general.

In our companion paper [10] we prove upper bounds for such sums; here we provide its applications.

1.1. **Applications to arithmetic statistics.** The following is a more general version of Theorem 1.1 on mixed moments:

Theorem 1.3. *Fix any $s > 0$. Then for all $X \geqslant 5$ we have*

$$X(\log X)^{2^s-1} \ll \sum_{D \in \mathcal{D}^+(X)} h_2(D)^s h_3(D) \ll X(\log X)^{2^s-1}$$

and

$$X(\log X)^{2^s-1} \ll \sum_{D \in \mathcal{D}^-(X)} h_2(D)^s h_3(D) \ll X(\log X)^{2^s-1},$$

where the implied constant depends at most on s.

Remark 1.4 (Independence)**.** Theorems 1.1 and 1.3 are the first results establishing the right order of magnitude for h_n when n has more than one prime factor. Since $h_6 = h_2 h_3$, the underlying problems are related to independent behavior of h_2 and h_3. One cannot exclude a priori that $h_2(D)$ and $h_3(D)$ correlate in a way that $h_3(D)$ attains very large values when $h_2(D)$ is large.

Davenport and Heilbronn [19] proved that $h_3(D)$ has constant average when D ranges in $\mathcal{D}^+(X)$. We show that the D responsible for this fact are those for which $h_2(D)$ is essentially $(\log |D|)^{\log 2}$. For a real number $\varepsilon > -1$ define

$$\mathfrak{c}(\varepsilon) := -\frac{\varepsilon}{1+\varepsilon} + \log(1+\varepsilon) \tag{1.2}$$

and note that $\mathfrak{c}(\varepsilon) > 0$.

Theorem 1.5. *For every fixed constants $\varepsilon_1 \in (0,1)$ and $\varepsilon_2 > 0$ and all $X, z_3, z_4 \geqslant 1$ with $(\log X)^{(1+\varepsilon_2)\log 2} \leqslant z_4$ we have*

$$\sum_{\substack{D \in \mathcal{D}^+(X) \cup \mathcal{D}^-(X) \\ h_2(D) \notin (z_3, z_4)}} (h_3(D) - 1) \ll X \left(\left(\frac{z_3^{1/\log 2}}{(\log X)^{(1-\varepsilon_1)}} \right)^{\log(1+\varepsilon_1)} + \frac{1}{z_4^{\mathfrak{c}(\varepsilon_2)/\log 2}} \right),$$

where the implied constant depends only on ε_i.

We will use the companion paper to give certain bounds for the frequency of atypical values of additive functions in Theorem 3.2. This has certain algebraic applications that we describe now.

Malle's conjecture [44] regards the number of extensions $K/\mathbb{Q}$ with prefixed Galois group when ordered by their discriminant Δ_K. The case of the full symmetric group S_n has attracted special attention; here, the largest n for which asymptotics are known is $n = 5$ due to Bhargava [3]; this was later extended and generalized by Shankar–Tsimerman [49] and Bhargava–Shankar–Wang [5].

We will prove that for the vast majority of S_5-extensions, the cardinality of ramified primes can only lie in a specific interval. This was first studied by Lemke Oliver–Thorne [42], who proved that the cardinality of ramified primes is distributed according to the Gaussian distribution of approximate centre $\log\log|\Delta_K|$ and length $(\log\log|\Delta_K|)^{1/2}$. Our work complements this by proving that the cardinality can only lie outside the interval with probability that decays exponentially fast.

Theorem 1.6. *For every fixed constants $\varepsilon_1 \in (0,1)$ and $\varepsilon_2 > 0$ and all $X, z_1, z_2 \geqslant 1$ with $(1+\varepsilon_2)\log\log X \leqslant z_2$ we have*

$$\sharp\{K \text{ quintic } S_5 : |\Delta_K| \leqslant X, \omega(\Delta_K) \geqslant z_2\} \ll X \exp\left(-\mathfrak{c}(\varepsilon_2) z_2\right)$$

and

$$\sharp\{K \text{ quintic } S_5 : |\Delta_K| \leqslant X, \omega(\Delta_K) \leqslant z_1\} \\ \ll X \exp\left(-\log(1+\varepsilon_1)\left((1-\varepsilon_1)\log\log X - z_1\right)\right),$$

where the implied constants depend only on ε_i.

Since $\sharp\{K \text{ quintic } S_5 : |\Delta_K| \leqslant X\}$ has order X due to Bhargava [3], one sees from Theorem 1.6 that $\omega(\Delta_K)$ must typically lie in the interval (z_1, z_2).

Malle's conjecture for cubic S_3 fields was first established by Davenport–Heilbronn [19]. The error term was later greatly improved by Bhargava–Shankar–Tsimerman [4] and Taniguchi–Thorne [53]. Our next result shows that for 100% of S_3 fields, the number of ramified primes $\omega(\Delta_K)$ lies in a prescribed interval, giving an analog of Theorem 1.6.

Theorem 1.7. *For every fixed constants $\varepsilon_1 \in (0,1)$ and $\varepsilon_2 > 0$ and all $X, z_1, z_2 \geqslant 1$ with $(1+\varepsilon_2)\log\log X \leqslant z_2$ we have*

$$\sharp\{K \text{ cubic } S_3 : |\Delta_K| \leqslant X, \omega(\Delta_K) \geqslant z_2\}| \ll X \exp\left(-\mathfrak{c}(\varepsilon_2) z_2\right)$$

and

$$\sharp\{K \text{ cubic } S_3 : |\Delta_K| \leqslant X, \omega(\Delta_K) \leqslant z_1\} \\ \ll X \exp\left(-\log(1+\varepsilon_1)\left((1-\varepsilon_1)\log\log X - z_1\right)\right),$$

where the implied constants depend only on ε_i.

1.2. **Applications to Diophantine equations.** We count the number of integer solutions of certain Diophantine equations, examples of which are the quartic affine surface

$$x_1^2 x_2^2 + x_3^2 + x_4^2 = N$$

and the affine quartic threefold $x_1^2 x_2^2 + x_3^2 x_4^2 + x_5^2 = N$. More generally, our work will cover the equation

$$(x_1 \cdots x_k)^2 + x_{k+1}^2 + x_{k+2}^2 = N, \tag{1.3}$$

whose number of variables is roughly half the degree of the equation.

For $N \in \mathbb{N}$ let

$$L(1, \chi_{-N}) = \sum_{m=1}^{\infty} \left(\frac{-N}{m}\right)\frac{1}{m} \qquad \text{and} \qquad \mathfrak{b}(N) = \prod_{p \mid N}\left(1 + \left(\frac{-1}{p}\right)\frac{1}{p}\right).$$

Theorem 1.8. *Fix $k \in \mathbb{N}$ and let N range through positive square-free integers $3 \pmod 8$.*

- *The number of $\mathbf{x} \in \mathbb{Z}^{k+2}$ satisfying (1.3) is*

$$\asymp \mathfrak{b}(N)^{k-1} L(1, \chi_{-N}) N^{\frac{1}{2}} (\log N)^{k-1},$$

where the implied constant depends only on k.

- *The number of $\mathbf{x} \in \mathbb{Z}^{2k+1}$ satisfying*

$$(x_1 \cdots x_k)^2 + (x_{k+1} \cdots x_{2k})^2 + x_{2k+1}^2 = N$$

is $\asymp \mathfrak{b}(N)^{2(k-1)} L(1, \chi_{-N}) N^{\frac{1}{2}} (\log N)^{2(k-1)}$, where the implied constant depends only on k.

- *The number of* $\mathbf{x} \in \mathbb{Z}^{3k}$ *satisfying*

$$(x_1 \cdots x_k)^2 + (x_{k+1} \cdots x_{2k})^2 + (x_{2k+1} \cdots x_{3k})^2 = N$$

is $\asymp \mathfrak{b}(N)^{3(k-1)} L(1, \chi_{-N}) N^{\frac{1}{2}} (\log N)^{3(k-1)}$*, where the implied constant depends only on* k*.*

The upper bound in the first bullet point in Theorem 1.8 follows from earlier work of Henriot [35, Theorem 3]. All cases of Theorem 1.8 are special cases of the more general Theorem 4.1, which allows us to put general multiplicative weights on the integer solutions x_i of

$$x_1^2 + x_2^2 + x_3^2 = N.$$

Its proof is given in §4.1 and is based on Theorem 2.3 and deep estimates of Duke [21] for the Fourier coefficients of cusp forms. It is worth mentioning that matching upper and lower bounds for the number of solutions of

$$x_1^2 + x_2^2 + p^2 = N, \quad (x_1, x_2 \in \mathbb{Z}, p \text{ prime}),$$

were given via the semi-linear sieve by Friedlander and Iwaniec [31, Theorem 14.5] on the assumption of the Generalized Riemann hypothesis and the Elliott–Halberstam conjecture.

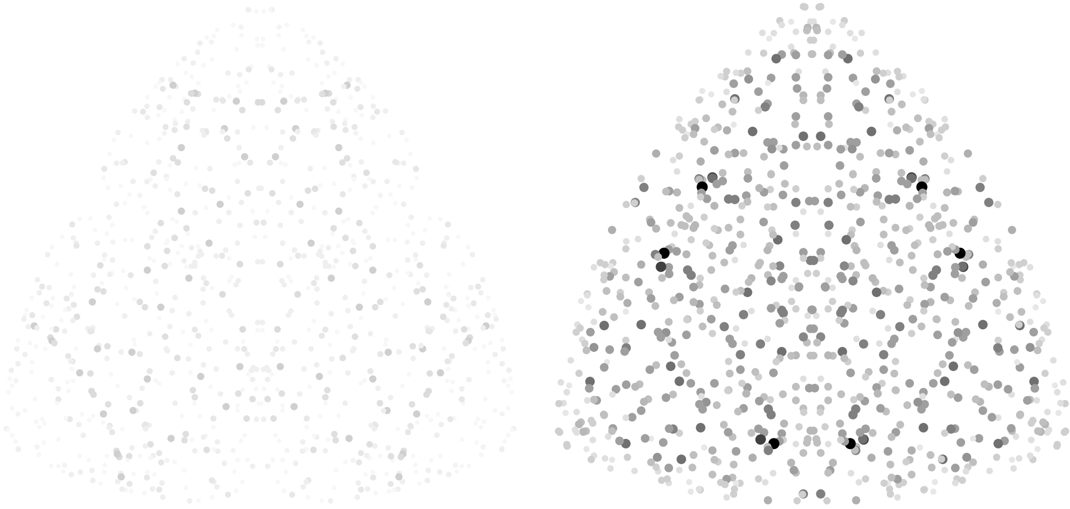

Figure 1.1. Weighted points on the sphere for $N = 1716099$ and $N = 1707035$

Remark 1.9 (Bias)**.** The term $L(1, \chi_{-N}) N^{1/2}$ corresponds to the number of terms in the sum by a classical result of Gauss, whereas, $(\log N)^{k-1}$ is the average of the k-th divisor function. The shape of $\mathfrak{b}(N)$ is biased towards integers N having many prime divisors $p \equiv 1 \pmod 4$ below $\log N$. It is possible to combine this with the work of Granville–Soundararajan [33, Theorem 5b], to find infinitely many N such that

$$\sharp\{\mathbf{x} \in \mathbb{N}^6 : (x_1 x_2)^2 + (x_3 x_4)^2 + (x_5 x_6)^2 = N\} \gg (\log \log N)^{5/2} (\log N)^3 N^{1/2}.$$

This is in constrast with the typical size, which is $(\log N)^3 N^{1/2}$ because $L(1, \chi_{-N})$ and $\mathfrak{b}(N)$ possess a limiting distribution due to the work of Chowla–Erdős [16] and Erdős–Wintner [25].

The bias is illustrated in the three-dimensional plots in Figure 1.1. They depict points $\mathbf{x} \in \mathbb{N}^3$ with $\sum_{i=1}^3 x_i^2 = N$, where each $\mathbf{x}$ is colored based on the magnitude of $\prod_{i=1}^3 \tau(x_i)$. The equations respectively have 960 and 936 solutions in $\mathbb{N}^3$. Among the six primes that divide 1716099, only one is $1 \pmod 4$. However, in the factorization of 1707035, four primes are involved, and all except one are $1 \pmod 4$.

The ideas behind the proof of Theorem 1.8 are not specific to sums of three squares. We generalise the results to equations without specific shape, with the only provision that they have enough variables compared to the degree. The end result is to study multiplicative functions over the coordinates of integer solutions of these Diophantine equations. Problems of this type have been considered by Cook–Magyar [18] and Yamagishi [56] in the case of the von Mangoldt function.

Theorem 1.10. *Fix any $s > 0$ and assume that $f : \mathbb{N} \to [0, \infty)$ is a multiplicative function satisfying $\tau(m)^{-s} \leqslant f(m) \leqslant \tau(m)^s$ for all m, where τ is the divisor function. Assume that $F \in \mathbb{Z}[x_0, \ldots, x_n]$ is a smooth homogeneous polynomial of degree d with $n \geqslant 4 + (d-1)2^d$ such that $F = 0$ has a non-zero integer solution. Then ones has for all $B \geqslant 1$*

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^{n+1} \\ \max |x_i| \leqslant B, F(\mathbf{x}) = 0}} f(|x_0 \cdots x_n|) \asymp B^{n+1-d} \left\{ \exp\left((n+1) \sum_{p \leqslant B} \frac{(f(p) - 1)}{p} \right) \right\},$$

where the implied constants depend at most on F and s.

The proof is based on Birch's circle method [7]. The term B^{n+1-d} represents the number of terms in the sum over $\mathbf{x}$. It will be clear from the proof that the assumption $f(m) \geqslant \tau(m)^{-s}$ is only needed for the lower bound.

1.3. **Polynomial values.**

Nair and Tenenbaum proved upper bounds for the average of arithmetic functions evaluated over values of polynomials in [46]. Such sums are omnipresent in number theory: their work was crucial in many different problems. Examples include

- Equidistribution of CM points ([39]),
- Manin's conjecture for counting rational points on surfaces ([14], [13]),
- Mass equidistribution ([36]),
- Unit fractions ([24]).

Such sums are of type (1.1) as can be seen by taking χ to be the indicator function of an interval and c_a to be the value of the polynomial at an integer a. However, (1.1) also covers any polynomial in any number of variables. In §5 we shall prove Theorems 1.15 and 1.16. These results respectively give matching upper and lower bounds for

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{R} \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|), \tag{1.4}$$

where $\mathcal{R}$ is a bounded subset of $\mathbb{R}^n$ and $Q(\mathbf{x})$ is an arbitrary polynomial in any number of variables. This is straightforward for polynomials without too many singularities but when Q is very singular there is high probability that a small power of a prime divides its values. The new ingredient needed is an estimate by Pierce–Schindler–Wood [47, Lemma 4.10] giving elementary proofs to statements regarding the Igusa zeta function.

We state one of the corollaries first. For $k, m \in \mathbb{N}$ denote the number of representations of m as the product of k natural numbers by $\tau_k(m)$.

Theorem 1.11. *Let $k \geqslant 2, n \geqslant 1$ be arbitrary integers, let ℓ be any positive real number and let Q be an integer irreducible polynomial in n variables. Then for all $X \geqslant 2$ we have*

$$X^n(\log X)^{k^\ell-1} \ll \sum_{\substack{\mathbf{x}\in(\mathbb{Z}\cap[-X,X])^n\\ Q(\mathbf{x})\neq 0}} \tau_k(|Q(\mathbf{x})|)^\ell \ll X^n(\log X)^{k^\ell-1}.$$

Remark 1.12 (Previously known cases)**.** Erdős [26] dealt with the case $\ell = n = 1, k = 2$ and Linnik [43] with $n = 1$, all $\ell, k \geqslant 1$ and $\deg(Q) = 1$. Later, Delmer [20] worked in the cases when $\ell \geqslant 1, k = 2, Q$ is irreducible and $n = 1$, Nair–Tenenbaum [46] for all ℓ, k, Q and when $n = 1$, and de la Bretèche–Browning [11] whenever $n = 2, \ell \geqslant 1, k \geqslant 1$ and Q is homogeneous. Asymptotics for the divisor function over the values of polynomials in more than one variable have been achieved by various authors, see, for example the work of de la Bretèche–Browning [12], Zhou–Ding [57] and the list of references therein. It is worth mentioning that in the case $k = 2, \ell = 1$ and Q a single irreducible polynomial in one variable, only the cases corresponding to linear and quadratic polynomials are known to satisfy an asymptotic, see the work of Hooley [37] and Bykovskiĭ [15].

Theorem 1.11 follows directly from our next two results. Denote

$$\varrho_Q(q) := \frac{\sharp\{\mathbf{y} \in (\mathbb{Z}/q\mathbb{Z})^n : Q(\mathbf{y}) \equiv 0 \,(\mathrm{mod}\ q)\}}{q^n}$$

for any $Q \in \mathbb{Z}[x_1, \dots, x_n]$ and $q \in \mathbb{N}$.

Definition 1.13. Let $\mathcal{D} \subset \mathbb{R}^n$ be bounded. We let

$$X(\mathcal{D}) := \sup\left\{\max_{1\leqslant i\leqslant n} |x_i| : \mathbf{x} \in \mathcal{D}\right\}.$$

Definition 1.14 (A class of functions)**.** Fix $A \geqslant 1, \epsilon > 0, C > 0$. The set $\mathcal{M}(A, \epsilon, C)$ of functions $f : \mathbb{N} \to [0, \infty)$ is defined by the property that for all coprime m, n one has

$$f(mn) \leqslant f(m) \min\{A^{\Omega(n)}, Cn^\epsilon\}.$$

Theorem 1.15. *Fix $A \geqslant 1$ and let $\mathcal{D}$ be a bounded set. Let Q be an arbitrary non-constant integer polynomial in n variables without repeated polynomial factors over $\mathbb{Q}$ and let f be a function such that for every $\epsilon > 0$ there exists $C > 0$ for which $f \in \mathcal{M}(A, \epsilon, C)$. Then*

$$\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{D}\\ Q(\mathbf{x})\neq 0}} f(|Q(\mathbf{x})|) \ll X(\mathcal{D})^n(\log 2X(\mathcal{D}))^{-r} \sum_{a\leqslant X(\mathcal{D})^n} f(a)\varrho_Q(a),$$

where r denotes the number of distinct irreducible polynomial factors of Q over $\mathbb{Q}$ and the implied constant depends at most on A, f, n and Q.

For the corresponding lower bound to hold it is necessary that $\mathcal{D}$ is not too small; this explains the condition on $\mathcal{D}$ in our next result:

Theorem 1.16. *Keep the notation and assumptions of Theorem 1.15. Assume, in addition, that $\mathcal{D}$ contains an open sphere of radius at least $X \geqslant 1$ and that $f : \mathbb{N} \to [0, \infty)$ is a multiplicative function such that*

$$\textit{for each } T \geqslant 1 \textit{ one has } \inf\{f(m) : \Omega(m) \leqslant T\} > 0.$$

Then there exists a positive constant θ_Q that depends on Q such that

$$\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap\mathcal{D}\\ Q(\mathbf{x})\neq 0}} f(|Q(\mathbf{x})|) \gg X^n(\log 2X)^{-r} \sum_{a\leqslant X} f(a)\varrho_Q(a),$$

where the implied constants depend at most on A, f, n and Q.

Notation. For a non-zero integer m define

$$\Omega(m) := \sum_{p \mid m} v_p(m),$$

where v_p is the standard p-adic valuation. Define $P^+(m)$ and $P^-(m)$ respectively to be the largest and the smallest prime factor of a positive integer m and let $P^+(1) = 1$ and $P^-(1) = +\infty$. For a real number x we reserve the notation $[x]$ for the largest integer not exceeding x. Throughout the paper we use the standard convention that empty products are set equal to 1. Throughout the paper we shall also make use of the convention that when iterated logarithm functions $\log t, \log\log t$, etc., are used, the real variable t is assumed to be sufficiently large to make the iterated logarithm well-defined.

Acknowledgements. The work started during the research stay of SC, PK and CP during the workshop *Problèmes de densité en Arithmétique* at CIRM Luminy in 2023. We would like to thank the organisers Samuele Anni, Peter Stevenhagen and Jan Vonk. We thank Levent Alpöge, Robert J. Lemke Oliver and Frank Thorne for useful discussions on §3.1. PK gratefully acknowledges the support of Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation. Part of the work of SC was supported by the National Science Foundation under Grant No. DMS-1928930, while the author was in residence at the MSRI in Spring 2023.

Structure of the paper. In §2 we recall the necessary results from [10]. Sections 3.1-3.2 respectively contain the proofs of Theorems 1.1 and 1.3 on the 6-rank. In §3.3 we prove Theorem 3.2 that provides tail bounds for the probability of large values of additive functions in the general setting of Theorem 2.3. This is then applied in Sections 3.4-3.5 and 3.6 to prove Theorems 1.5-1.6 and 1.7 respectively on h_3, S_5 and S_3 extensions. Sections 4.1-4.6 contain the proof of Theorem 4.1 on sums of three squares; this is more general than Theorem 1.8. The proof of Theorem 1.10 on general Diophantine equations is located in §4.7. Lastly, in §5.2 and §5.3 we prove respectively Theorems 1.15 and 1.16 on averages of arithmetic functions over values of arbitrary polynomials. They are then applied in §5.4 to prove Theorem 1.11.

2. Prerequisite lemmas

In this section we recall the required bounds proved in [10].

Definition 2.1 (Density functions)**.** Fix $\kappa, \lambda_1, \lambda_2, B, K > 0$. We define the set $\mathcal{D}(\kappa, \lambda_1, \lambda_2, B, K)$ of multiplicative functions $h : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ by the properties

- for all $B < w < z$ we have

$$\prod_{\substack{p \text{ prime} \\ w \leqslant p < z}} (1 - h(p))^{-1} \leqslant \left(\frac{\log z}{\log w}\right)^{\kappa} \left(1 + \frac{K}{\log w}\right), \tag{2.1}$$

- for every prime $p > B$ and integers $e \geqslant 1$ we have

$$h(p^e) \leqslant \frac{B}{p}, \tag{2.2}$$

- for every prime p and $e \geqslant 1$ we have

$$h(p^e) \leqslant p^{-e\lambda_1 + \lambda_2}. \tag{2.3}$$

Let $\mathcal{A}$ be an infinite set and for each $T \geqslant 1$ let $\chi_T : \mathcal{A} \to [0, \infty)$ be any function for which

$$\{a \in \mathcal{A} : \chi_T(a) > 0\} \text{ is finite for every } T \geqslant 1. \tag{2.4}$$

We also assume that

$$\lim_{T\to+\infty}\sum_{a\in\mathcal{A}}\chi_T(a)=+\infty. \tag{2.5}$$

Assume that we are given a sequence of strictly positive integers $(c_a)_{a\in\mathcal{A}}$ indexed by $\mathcal{A}$ and denoted by

$$\mathfrak{C}:=\{c_a : a\in\mathcal{A}\}.$$

We will be interested in estimating sums of the form

$$\sum_{a\in\mathcal{A}}\chi_T(a)f(c_a),$$

where f is an arithmetic function.

We will need the following notion of 'equi-distribution' of the values of the integer sequence c_a in arithmetic progressions. For a non-zero integer d and any $T\geqslant 1$, let

$$C_d(T)=\sum_{\substack{a\in\mathcal{A}\\ c_a\equiv 0 \pmod d}}\chi_T(a).$$

Definition 2.2 (Equidistributed sequences)**.** We say that $\mathfrak{C}$ is equidistributed if there exist positive real numbers $\theta,\xi,\kappa,\lambda_1,\lambda_2,B,K$ with $\max\{\theta,\xi\}<1$, a function $M:\mathbb{R}_{\geqslant 1}\to\mathbb{R}_{\geqslant 1}$ and a function $h_T\in\mathcal{D}(\kappa,\lambda_1,\lambda_2,B,K)$ such that

$$C_d(T)=h_T(d)M(T)\left\{1+O\left(\prod_{\substack{B<p\leqslant M(T)\\ p\nmid d}}(1-h_T(p))^2\right)\right\}+O(M(T)^{1-\xi}) \tag{2.6}$$

for every $T\geqslant 1$ and every $d\leqslant M(T)^\theta$, where the implied constants are independent of d and T.

It is worth emphasizing that in this definition the constants $\theta,\xi,\kappa,\lambda_1,\lambda_2,B,K$ are all assumed to be independent of T. For example, the bound $h_T(p^e)=O(1/p)$ in (2.2) holds with an implied constant that is independent of e,p as well as T. From now on we shall write M for $M(T)$. We are now ready to state the main result in [10].

Theorem 2.3. *Let $\mathcal{A}$ be an infinite set and for each $T\geqslant 1$ define $\chi_T:\mathcal{A}\to[0,\infty)$ to be any function such that both* (2.4) *and* (2.5) *hold. Take a sequence of strictly positive integers $\mathfrak{C}=(c_a)_{a\in\mathcal{A}}$. Assume that $\mathfrak{C}$ is equidistributed with respect to some positive constants $\theta,\xi,\kappa,\lambda_1,\lambda_2,B,K$ and functions $M(T)$ and $h_T\in\mathcal{D}(\kappa,\lambda_1,\lambda_2,B,K)$ as in Definition 2.2. Fix any $A>1$ and assume that f is a function such that for every $\epsilon>0$ there exists $C>0$ for which $f\in\mathcal{M}(A,\epsilon,C)$, which is introduced in Definition 1.14. Assume that there exists $\alpha>0$ and $\widetilde{B}>0$ such that for all $T\geqslant 1$ one has*

$$\sup\{c_a : a\in\mathcal{A},\chi_T(a)>0\}\leqslant\widetilde{B}M^\alpha, \tag{2.7}$$

where $M=M(T)$ is as in Definition 2.2. Then for all $T\geqslant 1$ we have

$$\sum_{a\in\mathcal{A}}\chi_T(a)f(c_a)\ll M\prod_{B<p\leqslant M}(1-h_T(p))\sum_{a\leqslant M}f(a)h_T(a),$$

where the implied constant is allowed to depend on $\alpha,A,B,\widetilde{B},\theta,\xi,K,\kappa$, λ_i, the function f and the implied constants in (2.6)*, but is independent of T and M.*

Let us now recall the corresponding lower bound proved in [10].

Theorem 2.4. *Keep the notation and assumptions of Theorem 2.3. Assume, in addition, that $f:\mathbb{N}\to[0,\infty)$ is a multiplicative function for which*

$$\text{for each } L\geqslant 1 \text{ one has } \inf\{f(m):\Omega(m)\leqslant L\}>0$$

and that the error term in Definition 2.2 satisfies

$$C_d(T) = h_T(d)M(T)\left\{1 + o_{T\to\infty}\left(\prod_{\substack{B<p\leqslant M(T)\\ p\nmid d}} (1-h_T(p))^2\right)\right\} + O(M(T)^{1-\xi})$$

whevever $d \leqslant M(T)^\theta$*. Then for all* $T \geqslant 1$ *we have*

$$\sum_{a\in\mathcal{A}} \chi_T(a)f(c_a) \gg M(T)\prod_{p\leqslant M(T)}(1-h_T(p))\sum_{a\leqslant M(T)} f(a)h_T(a),$$

where the implied constants are independent of T *and* M*.*

We finish this section with lemmas that will be needed in the forthcoming applications.

Lemma 2.5. *Let* $Q \in \mathbb{Z}[x_1,\dots,x_n]$ *be non-constant and without repeated factors over* $\mathbb{Q}$*. Then as* $x\to\infty$ *one has*

$$\sum_{p\leqslant x} \frac{\sharp\{\mathbf{x}\in\mathbb{F}_p^n : Q(\mathbf{x})=0\}}{p^{n-1}} = r\frac{x}{\log x}(1+o(1)),$$

where r *is the number of distinct irreducible factors of* Q *in* $\mathbb{Q}[x_1,\dots,x_n]$*. Furthermore, there exists a constant* $c=c(Q)$ *such that for* $x\geqslant 2$ *one has*

$$\prod_{p\ \text{prime}, p\leqslant x}\left(1-\frac{\sharp\{\mathbf{x}\in\mathbb{F}_p^n : Q(\mathbf{x})=0\}}{p^n}\right) = c(\log x)^{-r} + O((\log x)^{-r-1}),$$

where the implied constant depends on Q*.*

Proof. We factor Q over $\overline{\mathbb{Q}}$ as $c_0\prod_{i=1}^t Q_i$ with c_0 in $\mathbb{Q}^*$, with $Q_i\in\overline{\mathbb{Z}}[x_1,\dots,x_n]$ irreducible and with the property that if Q_i occurs in the factorisation, then so does each of its Galois conjugates. We write $S=\{Q_1,\dots,Q_t\}$ for the set of factors obtained in this way. Let $K/\mathbb{Q}$ be the number field obtained by adding all the coefficients appearing in the factorisation. Since the factors come as Galois orbits, the field K must be Galois. The group $\mathrm{Gal}(K/\mathbb{Q})$ acts on S by permuting the factors. By the Lang–Weil bounds [41] we have that

$$\frac{\sharp\{\mathbf{x}\in\mathbb{F}_p^n : Q(\mathbf{x})=0\}}{p^n} = \frac{c_Q(p)}{p} + O\left(\frac{1}{p^{\frac{3}{2}}}\right),$$

where $c_Q(p)$ denotes the number of distinct irreducible factors of Q defined over $\mathbb{F}_p$, when one factorizes the polynomial $Q \bmod p$ in $\overline{\mathbb{F}_p}[x_1,\dots,x_n]$: in other words the irreducible factors in the $\mathbb{F}_p$-factorisation that remain irreducible factors in the $\overline{\mathbb{F}_p}$-factorisation. We now wish to express the function $c_Q(p)$ as a function of the Artin symbol of p in $K/\mathbb{Q}$, for sufficiently large primes p.

For a prime p that is also unramified in $K/\mathbb{Q}$, $\mathrm{Art}(p,K/\mathbb{Q})$ defines a conjugacy class in $\mathrm{Gal}(K/\mathbb{Q})$, which we view as permutations on S via the action. The number of fixed points of the resulting permutation is independent of the element in the conjugacy class, and we denote this function of $\mathrm{Gal}(K/\mathbb{Q})$ as $g\mapsto \mathrm{Fix}(g)$. This defines a function on sufficiently large primes via $p\mapsto \mathrm{Fix}(\mathrm{Art}(p,K/\mathbb{Q}))$. We claim that for p sufficiently large we have that $c_Q(p)=\mathrm{Fix}(\mathrm{Art}(p,K/\mathbb{Q}))$. Indeed, observe that since Q has no repeated factors over $\mathbb{Q}$, it follows that its reduction modulo p has no repeated factors in $\mathbb{F}_p[x_1,\dots,x_n]$ provided that we take p sufficiently large. Furthermore, for p sufficiently large, choosing any prime $\bar{\mathfrak{p}}$ above p in $\overline{\mathbb{Z}}$, we have that all of the elements of S remain irreducible when reduced modulo $\bar{\mathfrak{p}}$.

We claim that if:

(P1) Q has no repeated factors modulo p,

(P2) all of the elements of S remain irreducible modulo $\bar{\mathfrak{p}}$,
(P3) p is unramified in $K/\mathbb{Q}$,
(P4) c_0 is coprime to p,

then $c_Q(p) = \mathrm{Fix}(\mathrm{Art}(p, K/\mathbb{Q}))$. To see this, let us fix a prime $\mathfrak{p}$ of $\mathcal{O}_K$ lying above p. Recall that there is a unique element $\sigma \in \mathrm{Art}(p, K/\mathbb{Q})$ such that $\sigma(\mathfrak{p}) = \mathfrak{p}$ and $\sigma(\alpha) \equiv \alpha^p \bmod \mathfrak{p}$ for each α in $\mathcal{O}_K$. Now let us reduce each Q_i modulo $\mathfrak{p}$. The factors Q_i remain distinct thanks to (P1) and also remain irreducible thanks to (P2). Since p is unramified thanks to (P3), we can find the unique element σ as above. Furthermore, c_0 being non-zero makes sure that Q is not 0 modulo $\mathfrak{p}$. Thus, the reduction of the Q_i is truly the factorisation of Q modulo $\mathfrak{p}$.

If $\sigma(Q_i) = Q_i$, then all of the coefficients γ of Q_i satisfy $\gamma^p = \gamma$ when reduced modulo $\mathfrak{p}$, i.e. they are all in $\mathbb{F}_p$. So each fixed point of σ gives an irreducible factor of Q over $\overline{\mathbb{F}_p}$ that is already in $\mathbb{F}_p$. Conversely, suppose that $\sigma(Q_i) \neq Q_i$. Since the factors remain distinct modulo $\mathfrak{p}$ by (P1), then $\sigma(Q_i)$ and Q_i are also distinct factors modulo $\mathfrak{p}$. But this means that the polynomial Q_i modulo $\mathfrak{p}$ and the same polynomial with all coefficients raised to the power p are different factors of Q modulo $\mathfrak{p}$. In other words, Q_i is not defined over $\mathbb{F}_p$. Hence we have precisely proved that under the assumptions (P1)-(P4), the quantity $\mathrm{Fix}(\mathrm{Art}(p, K/\mathbb{Q}))$ equals the number of $\overline{\mathbb{F}_p}$-irreducible components of Q that are defined over $\mathbb{F}_p$, i.e. it equals $c_Q(p)$.

Recall that each of (P1)-(P4) is satisfied for all sufficiently large primes. By the Chebotarev density theorem we obtain the following for all $x \geqslant 2$,

$$\sum_{p \leqslant x} \frac{\sharp\{\mathbf{x} \in \mathbb{F}_p^n : Q(\mathbf{x}) = 0\}}{p^{n-1}} = \left(\frac{\sum_{g \in \mathrm{Gal}(K/\mathbb{Q})} \mathrm{Fix}(g)}{\sharp \mathrm{Gal}(K/\mathbb{Q})} \right) \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2} \right).$$

Using partial summation we obtain a constant B depending only on Q such that

$$\sum_{p \leqslant x} \frac{\sharp\{\mathbf{x} \in \mathbb{F}_p^n : Q(\mathbf{x}) = 0\}}{p^{n}} = \left(\frac{\sum_{g \in \mathrm{Gal}(K/\mathbb{Q})} \mathrm{Fix}(g)}{\sharp \mathrm{Gal}(K/\mathbb{Q})} \right) (\log\log x) + B + O(1/\log x), \quad (2.8)$$

from which one can deduce an asymptotic for the product over primes $p \leqslant x$ in the statement of the lemma by taking logarithms.

To complete the proof, it suffices to observe that if G is a finite group acting on a finite set X and if $\mathrm{Fix}(g)$ denotes the number of fixed points of an element g in G viewed as permutation of X, then

$$\frac{\sum_{g \in G} \mathrm{Fix}(g)}{\sharp G}$$

equals the number of orbits of G acting on X. In our case the number of $\mathrm{Gal}(K/\mathbb{Q})$-orbits acting on S is r, thus completing the argument. □

Lemma 2.6. *Let $Q \in \mathbb{Z}[x_1, \dots, x_n]$ be non-constant and without repeated factors over $\mathbb{Q}$. Then for any prime p the number of $\mathbf{x} \in (\mathbb{Z}/p^2\mathbb{Z})^n$ for which $Q(\mathbf{x}) \equiv 0 \pmod{p^2}$ is $O(p^{n-2})$, where the implied constant depends at most on Q.*

Proof. For a point $\mathbf{t}$ in $(\mathbb{Z}/p\mathbb{Z})^n$ satisfying $Q(\mathbf{t}) \equiv 0 \pmod{p}$, we denote

$$N(\mathbf{t}) = \sharp \left\{ \mathbf{x} \in (\mathbb{Z}/p^2\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \left(\bmod\ p^2 \right) \text{ and } \mathbf{x} \equiv \mathbf{t} \pmod{p} \right\}.$$

By definition, we have that

$$\sharp \left\{ \mathbf{x} \in (\mathbb{Z}/p^2\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \left(\bmod\ p^2 \right) \right\} = \sum_{\substack{\mathbf{t} \in (\mathbb{Z}/p\mathbb{Z})^n \\ Q(\mathbf{t}) \equiv 0 (\bmod\ p)}} N(\mathbf{t}).$$

Suppose that $\mathbf{t}$ in $(\mathbb{Z}/p\mathbb{Z})^n$ satisfies both $Q(\mathbf{t}) \equiv 0 \,(\bmod\, p)$ and $\nabla Q(\mathbf{t}) \not\equiv \mathbf{0} \pmod p$. Then Hensel's lemma implies that $N(\mathbf{t}) = p^{n-1}$. The Lang–Weil estimates [41] imply that the number of $\mathbf{t} \in (\mathbb{Z}/p\mathbb{Z})^n$ for which $Q(\mathbf{t}) \equiv 0 \,(\bmod\, p)$ is $O(p^{n-1})$, hence,

$$\sharp\left\{\mathbf{x} \in (\mathbb{Z}/p^2\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \left(\bmod\, p^2\right)\right\} = O(p^{2n-2}) + \sum_{\substack{\mathbf{t}\in(\mathbb{Z}/p\mathbb{Z})^n \\ Q(\mathbf{t})\equiv 0(\bmod\ p) \\ \nabla Q(\mathbf{t})\equiv \mathbf{0}(\bmod\ p)}} N(\mathbf{t}).$$

Since one has the trivial bound $N(\mathbf{t}) \leqslant \sharp\{\mathbf{x} \in (\mathbb{Z}/p^2\mathbb{Z})^n : \mathbf{x} \equiv \mathbf{t} \,(\bmod\, p)\} = p^n$, it is sufficient for the proof to show that

$$\sharp\{\mathbf{t} \in (\mathbb{Z}/p\mathbb{Z})^n : Q(\mathbf{t}) \equiv 0 \,(\bmod\, p)\,, \nabla Q(\mathbf{t}) \equiv \mathbf{0} \,(\bmod\, p)\} = O(p^{n-2}).$$

Since Q has no repeated polynomial factors over $\mathbb{Q}$, it follows that it has no repeated polynomial factors over $\mathbb{F}_p$ for all sufficiently large primes p. For these p we therefore see that the intersection $Q(\mathbf{t}) = \nabla Q(\mathbf{t}) = \mathbf{0}$ defines a subvariety of $\mathbb{A}^n_{\mathbb{F}_p}$ of codimension at least 2. The Lang–Weil estimates [41] therefore provide the required bound $O(p^{n-2})$. □

Lemma 2.7. *Fix any positive real numbers c_0, c_1, c_2, c_3 and assume that $F : \mathbb{N} \to [0,\infty)$ is a multiplicative function such that*

$$F(p^e) \leqslant \min\left\{\frac{c_0}{p}, \frac{p^{c_1}}{p^{ec_2}}\right\} \tag{2.9}$$

for all primes p and $e \geqslant 1$ and $F(p^e) \leqslant c_3/p^2$ for all $p, e \geqslant 2$. Fix any $C, C' > 0$ and assume that $G : \mathbb{N} \to [0,\infty)$ is a function such that for all coprime positive integers a, b one has $G(ab) \leqslant G(a)\min\{C^{\Omega(b)}, C'b^{c_2/2}\}$.

Then for all $x \geqslant 1$ we have

$$\sum_{\substack{n\leqslant x \\ P^-(n)>c_0}} F(n)G(n) \ll \exp\left(\sum_{c_0<p\leqslant x} F(p)G(p)\right),$$

where the implied constant depends at most on c_i and C, C'.

Proof. We define a multiplicative function H' such that when p is prime and $e \geqslant 2$ one has $H'(p^e) = \min\{C^e, C'p^{c_2e/2}\}$ while $H'(p) = G(p)$. It is not difficult to show that for all coprime positive integers a, b we have $G(ab) \leqslant G(a)H'(b)$. Hence, $G(b) \leqslant H'(b)$ for all b and therefore the sum in the lemma is at most

$$\sum_{\substack{n\leqslant x \\ P^-(n)>c_0}} F(n)H'(n) \leqslant \prod_{\substack{n\leqslant x \\ P^-(n)>c_0}} \left(1+\sum_{e\geqslant 1} F(p^e)H'(p^e)\right) \leqslant \exp\left(\sum_{c_0<p\leqslant x, e\geqslant 1} F(p^e)H'(p^e)\right)$$

due to the inequality $1 + z \leqslant \mathrm{e}^z$ valid for all $z \in \mathbb{R}$. Let $\mathfrak{E}$ be a positive integer that will be specified later. The contribution of $e > \mathfrak{E}$ is at most

$$p^{c_1}\sum_{e>\mathfrak{E}} p^{-ec_2}H'(p^e) \leqslant C'p^{c_1}\sum_{e>\mathfrak{E}} p^{-ec_2/2} \leqslant C'p^{c_1-\mathfrak{E}c_2/2}(1-2^{-c_2/2})^{-1} \ll p^{c_1-\mathfrak{E}c_2/2}.$$

Taking $\mathfrak{E}$ to be the least positive integer satisfying $2(c_1+2)/c_2 \leqslant \mathfrak{E}$ yields the bound $\ll p^{-2}$. The contribution of the terms in the interval $[2,\mathfrak{E}]$ is

$$\leqslant \sum_{2\leqslant e\leqslant\mathfrak{E}} F(p^e)H'(p^e) \leqslant \sum_{2\leqslant e\leqslant\mathfrak{E}} F(p^e)C^e \leqslant \frac{c_3}{p^2}\sum_{2\leqslant e\leqslant\mathfrak{E}} C^e \ll \frac{1}{p^2} \ll \frac{1}{p^2}.$$

Thus, the overall bound becomes

$$\exp\left(\sum_{c_0<p\leqslant x, e\geqslant 1} F(p^e)H'(p^e)\right) \leqslant \exp\left(\sum_{c_0<p\leqslant x} F(p)H'(p)\right)\exp\left(\sum_{c_0<p\leqslant x} O(1/p^2)\right),$$

which is sufficient because $H'(p) = G(p)$. □

Lemma 2.8. *Assume that $Q \in \mathbb{Z}[x_1, \dots, x_n]$ is non-constant. Then for all $e \geqslant 1$ and primes p we have*

$$p^{-en} \sharp\{\mathbf{x} \in (\mathbb{Z}/p^e\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \pmod{p^e}\} \ll p^{-e/\deg(Q)},$$

where the implied constant depends at most on Q.

Proof. If Q is homogeneous then the bound follows from [47, Lemma 4.10]. If not, then we can work with the homogenized version Q_1 of Q, which is a homogeneous polynomial in $n+1$ variables having the same degree satisfying $Q_1(\mathbf{x}, 1) = Q(\mathbf{x})$. Thus, using the homogeneity of Q_1, one has

$$\sharp\{\mathbf{x} \in (\mathbb{Z}/p^e\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0\} = \frac{\sharp\{z \in (\mathbb{Z}/p^e\mathbb{Z})^*, \mathbf{x} \in (\mathbb{Z}/p^e\mathbb{Z})^n : Q_1(\mathbf{x}, z) \equiv 0\}}{(p-1)p^{e-1}}.$$

Applying [47, Lemma 4.10] to Q_1 shows that the numerator in the right hand-side is $\ll p^{e(n+1)-e/\deg(Q_1)}$, which is sufficient. □

Lemma 2.9. *Let $Q \in \mathbb{Z}[x_1, \dots, x_n]$ be a non-constant polynomial having no repeated factors over $\mathbb{Q}$. Fix any $\lambda > 0, C \in [1, 2)$ and assume that $G : \mathbb{N} \to \mathbb{R} \cap [0, \infty)$ is multiplicative, that $G(p) = \lambda$ for every prime p, that $G(p^e) \leqslant C^e$ for all $e \in \mathbb{N}$ and primes p and that for all $\epsilon > 0$ there exists $C'(\epsilon) > 0$ such that $G(b) \leqslant C'(\epsilon) b^\epsilon$ for all $b \in \mathbb{N}$.*

Then there exists a positive constant c that depends on Q and G, such that when $x \to \infty$ we have

$$\sum_{1 \leqslant m \leqslant x} G(m) \frac{\sharp\{\mathbf{x} \in (\mathbb{Z}/m\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \pmod{m}\}}{m^n} \sim c(\log x)^{\lambda r},$$

where r is the number of distinct irreducible factors of Q in $\mathbb{Q}[x_1, \dots, x_n]$.

Proof. We employ Wirsing's result [55, Satz 1] with

$$f_0(m) = G(m) \frac{\sharp\{\mathbf{x} \in (\mathbb{Z}/m\mathbb{Z})^n : Q(\mathbf{x}) \equiv 0 \pmod{m}\}}{m^{n-1}}.$$

By [51, Lemma 2.7] we have $f_0(p^e) \leqslant G(p^e) \deg(Q) \leqslant C^e \deg(Q)$ if Q is primitive, which implies a similar bound for non-primitive Q. This means that the assumption [55, Equation (3)] is met. To verify [55, Equation (4)] we note that

$$\sum_{p \leqslant x} G(p) \frac{\sharp\{\mathbf{x} \in \mathbb{F}_p^n : Q(\mathbf{x}) = 0\}}{p^{n-1}} = \lambda \sum_{p \leqslant x} \frac{\sharp\{\mathbf{x} \in \mathbb{F}_p^n : Q(\mathbf{x}) = 0\}}{p^{n-1}}$$

is asymptotic to $\lambda r x / \log x$ by Lemma 2.5 and the assumption that G is constantly λ on the primes. Hence, as $x \to \infty$, [55, Equation (5)] gives

$$\sum_{m \leqslant x} f_0(m) \sim c' \frac{x}{\log x} \prod_{p \leqslant x} \left(1 + \sum_{e \geqslant 1} G(p^e) \frac{\sharp\{\mathbf{x} \in (\mathbb{Z}/p^e\mathbb{Z})^n : Q(\mathbf{x}) = 0\}}{p^{en}}\right)$$

for some positive constant c'. Finally, using an argument that is similar to the ones in the proof of Lemma 2.7 and making use of Lemmas 2.6 and 2.8 to control the contribution of terms with $e \geqslant 2$ shows that when $x \to \infty$ the last product over $p \leqslant x$ is asymptotic to

$$c'' \exp\left(\sum_{p \leqslant x} G(p) \frac{\sharp\{\mathbf{x} \in (\mathbb{Z}/p\mathbb{Z})^n : Q(\mathbf{x}) = 0\}}{p^n}\right)$$

for some positive c''. Injecting (2.8) shows that

$$\sum_{m\leqslant x} f_0(m) \sim c''' x(\log x)^{\lambda r-1}$$

for some positive constant c'''. Noting that

$$G(m)\frac{\sharp\{\mathbf{x}\in(\mathbb{Z}/m\mathbb{Z})^n : Q(\mathbf{x})\equiv 0 \,(\mathrm{mod}\, m)\}}{m^n}=\frac{f_0(m)}{m}$$

and using partial summation concludes the proof. □

We shall later need the following substitute of Wirsing's theorem for multiplicative functions for which the average over the primes is not known.

Lemma 2.10. *Fix any $k\in\mathbb{N}$ and assume that f is a multiplicative function satisfying $0\leqslant f(p^e)\leqslant \tau(p^e)^k p^{-e}$ for all $e\geqslant 1$ and primes p. Then for all $x\geqslant 2$ we have*

$$\sum_{n\leqslant x} f(n) \asymp \exp\left(\sum_{p\leqslant x} f(p)\right),$$

where the implied constants depend at most on k.

Proof. The upper bound is evident. For the lower bound our plan is to prove that there exists $\delta=\delta(k)\in(0,1)$ such that

$$\exp\left(\sum_{p\leqslant x^\delta} f(p)\right)\ll \sum_{n\leqslant x} f(n). \tag{2.10}$$

This is clearly sufficient since

$$\sum_{x^\delta<p\leqslant x} f(p)\ll_k \sum_{x^\delta<p\leqslant x}\frac{1}{p}\ll_k 1.$$

To prove (2.10) we start by noting that for each $y\in[2,x]$ one has

$$\sum_{n\leqslant x} f(n)\geqslant \sum_{\substack{n\leqslant x\\ P^+(n)\leqslant y}} f(n)\mu(n)^2=\sum_{P^+(n)\leqslant y} f(n)\mu(n)^2-\sum_{\substack{n>x\\ P^+(n)\leqslant y}} f(n)\mu(n)^2.$$

Since there exists $C(k)>0$ such that

$$\sum_{P^+(n)\leqslant y} f(n)\mu(n)^2\geqslant C(k)\exp\left(\sum_{p\leqslant y} f(p)\right),$$

it suffices to show that

$$\sum_{\substack{n>x\\ P^+(n)\leqslant y}} f(n)\mu(n)^2\leqslant \frac{C(k)}{2}\exp\left(\sum_{p\leqslant y} f(p)\right).$$

We will see that this holds when $y=x^\delta$, where δ is a small positive constant that depends on k. Define $\sigma=1/\log y$ so that by Rankin's trick we have

$$\begin{aligned}\sum_{\substack{n>x\\ P^+(n)\leqslant y}} f(n)\mu(n)^2 &\leqslant x^{-\sigma}\sum_{P^+(n)\leqslant y} f(n)\mu(n)^2 n^\sigma\\ &=x^{-\sigma}\prod_{p\leqslant y}(1+f(p)p^\sigma)\leqslant x^{-\sigma}\exp\left(\sum_{p\leqslant y} f(p)p^\sigma\right).\end{aligned}$$

Since $\mathrm{e}^t \leqslant 1 + \mathrm{e}\cdot t$ for $0 \leqslant t \leqslant 1$, we see that $p^\sigma \leqslant 1 + \mathrm{e}\cdot \sigma \log p$ for all primes $p \leqslant y$, hence, the sum inside the exponential is at most

$$\sum_{p\leqslant y} f(p) + O\left(\sigma \sum_{p\leqslant y} f(p)\log p\right) \leqslant \sum_{p\leqslant y} f(p) + O\left(\sigma \sum_{p\leqslant y} \frac{\log p}{p}\right) = \sum_{p\leqslant y} f(p) + O(1).$$

Hence, there exists a positive constant $C_1(k)$ such that

$$\sum_{\substack{n>x \\ P^+(n)\leqslant y}} f(n)\mu(n)^2 \leqslant C_1(k) x^{-\sigma} \exp\left(\sum_{p\leqslant y} f(p)\right).$$

Denote $C_2(k) = C(k)/(2C_1(k))$. We want to make sure that $x^{-\sigma} \leqslant C_2(k)$; this can be achieved by taking $y = x^\delta$ with $\delta = \max\{1/2, (-\log C_2(k))^{-1}\}$. This is because we have $x^\sigma = \mathrm{e}^{1/\delta}$ due to $y = x^\delta$. □

3. Arithmetic statistics

3.1. **6-torsion.** Here we prove Theorems 1.1 using Theorem 2.3. This has an assumption related to a level of distribution result. Similar results have been obtained by [6, Theorem 1.2], [23, Section 6] and [42, Theorem 2.1]. Here we use the one by Belabas [2, Théorème 1.2]. Let g_1 be the multiplicative function defined as

$$g_1(p^e) = \begin{cases} p/(p+1), & \text{if } p \geqslant 2 \text{ and } e = 1 \\ 0, & \text{if } p > 2 \text{ and } e \geqslant 2 \\ 4/3, & \text{if } p = 2 \text{ and } e = 2 \\ 4/3, & \text{if } p = 2 \text{ and } e = 3 \\ 0, & \text{if } p = 2 \text{ and } e \geqslant 4. \end{cases}$$

It is not difficult to see that

$$\prod_{p\leqslant X}\Big(1 - \frac{g_1(p)}{p}\Big) \sum_{a\leqslant X} 2^{\omega(a)} \frac{g_1(a)}{a} \leqslant \prod_{p\leqslant X}\Big(1 - \frac{g_1(p)}{p}\Big) \sum_{a\leqslant X} 2^{\omega(a)} \frac{4}{3a} \ll \log X. \tag{3.1}$$

Lemma 3.1. *Fix any $\epsilon > 0$. Then for all $q \in \mathbb{N}$ and $X \geqslant 2$ with $q < X^{\frac{1}{15}-\epsilon}$ we have*

$$\sum_{\substack{D\in\mathcal{D}^+(X) \\ q|D}} (h_3(D) - 1) = \frac{1}{\pi^2}\frac{g_1(q)}{q} X + O\left(\frac{X}{q(\log X)^2(\log\log X)^{2-\epsilon}} + X^{\frac{15}{16}+\epsilon} q^{-\frac{1}{16}}\right)$$

and

$$\sum_{\substack{D\in\mathcal{D}^-(X) \\ q|D}} (h_3(D) - 1) = \frac{3}{\pi^2}\frac{g_1(q)}{q} X + O\left(\frac{X}{q(\log X)^2(\log\log X)^{2-\epsilon}} + X^{\frac{15}{16}+\epsilon} q^{-\frac{1}{16}}\right),$$

where the implied constants are independent of q and X.

Proof. This follows from [2, Théorème 1.2] and the remark immediately thereafter. □

We are now ready to begin the proof of Theorem 1.1. The lower bounds follow from $h_6(D) \geqslant h_2(D)$ and genus theory. This idea for the lower bound was further exploited and investigated in [29, Section 5]. For the upper bounds, we use Theorem 2.3 with

$$\mathcal{A} = \{D \text{ fundamental discriminant}\},\ \chi_T(D) = (h_3(D) - 1)\mathbb{1}_{|D|\leqslant T}(D),\ c_D = |D|.$$

We let $h(q) = g_1(q)/q$ and f be the multiplicative function $f(n) = 2^{\omega(n)}$. Lemma 3.1 shows that the level of distribution assumption in Definition 2.2 is satisfied with

$$M(T) = \frac{4}{\pi^2}T, \ \ \theta = \frac{1}{30} \quad \text{and} \quad \xi = \frac{1}{32}.$$

Let $h_n^+(D)$ be the size of the n-torsion subgroup of the narrow class group. We have $h_n(D) \leqslant h_n^+(D)$. Since $h_2^+(D) = 2^{\omega(D)-1}$ and $h_6^+(D) = h_2^+(D)h_3^+(D)$, we obtain

$$\sum_{D \in \mathcal{D}^+(X)} h_6^+(D) + \sum_{D \in \mathcal{D}^-(X)} h_6^+(D)$$

$$= \sum_{D \in \mathcal{D}^+(X)} h_2^+(D) + \sum_{D \in \mathcal{D}^-(X)} h_2^+(D) + \sum_{D \in \mathcal{D}^+(X) \cup \mathcal{D}^-(X)} (h_3(D) - 1)\, h_2^+(D).$$

The first two sums are readily estimated as $O(X \log X)$. For the final sum, the application of Theorem 2.3 and (3.1) yields

$$\sum_{D \in \mathcal{D}^+(X) \cup \mathcal{D}^-(X)} (h_3(D) - 1)\, h_2^+(D) = \sum_{D \in \mathcal{A}} \chi_X(D) f(c_D) \ll X \log X,$$

as required.

3.2. **Proof of Theorem 1.3.** The proof is as that of Theorem 1.1 with the only difference being that the bound

$$\sum_{a \leqslant X} \frac{2^{s\omega(a)}}{a} \ll \prod_{p \leqslant X} \Big(1 + \frac{2^s}{p}\Big) \ll (\log X)^{2^s}$$

must be used in place of (3.1).

3.3. **Tail bounds for additive functions.** We next study S_5 and S_3-extensions for which it will be necessary to turn to the distribution of additive functions. Perhaps the most famous additive function is $\omega(n)$, which is roughly speaking normally distributed with mean $\log\log n$ and standard deviation $\sqrt{\log\log n}$ by the Erdős–Kac theorem. However, these type of Erdős–Kac results do not give any tail bounds for the frequency of very large values of additive functions. We will prove strong tail bounds in the general setting of Theorem 2.3. Recall the definition of $\mathfrak{c}(\epsilon)$ from (1.2).

Theorem 3.2. *Let $\mathcal{A}$ be a set and $\chi_T : \mathcal{A} \to [0, \infty)$ be a function for which both* (2.4) *and* (2.5) *hold. Fix any positive constants $\kappa, \lambda_1, \lambda_2, B, K, C, r$ and let $h \in \mathcal{D}(\kappa, \lambda_1, \lambda_2, B, K)$ be such that*

$$\sum_{p \leqslant x} h(p) = r \log\log x + O(1) \tag{3.2}$$

and such that for all primes p and integers e $\geqslant 2$ we have $h(p^e) \leqslant C/p^2$. Fix any θ, ξ and let $\mathfrak{C} = (c_a)_{a \in \mathcal{A}}$ be an equidistributed sequence as in Definition 2.2. Assume that there exists $\alpha > 0$ and $\widetilde{B} > 0$ such that for all $T \geqslant 1$ one has (2.7) *and $M = M(T)$ is as in Definition 2.2. Suppose $\psi : \mathbb{N} \to \mathbb{R}$ is an additive function such that there exists some constant $\tilde{A} \geqslant 1$ such that*

$$\psi(n) \leqslant \tilde{A}\Omega(n),$$

and that for every $\epsilon > 0$ there exists $\tilde{C} \geqslant 0$ such that $\psi(n) \leqslant \epsilon \log m + \tilde{C}$.

Then for every fixed constants $\varepsilon_1 \in (0, 1)$ and $\varepsilon_2 > 0$ and all $T, z_2 \geqslant 1$ with

$$r\tilde{A}(1 + \varepsilon_2) \log\log M \leqslant z_2,$$

we have

$$\sum_{\substack{a\in\mathcal{A}\\ \psi(c_a)\geqslant z_2}} \chi_T(a) \ll M\exp\left(-\frac{\mathfrak{c}(\varepsilon_2)}{\tilde{A}}z_2\right). \tag{3.3}$$

Further assume that there exists some constant $\tilde{A}_1>0$ such that

$$\sum_{p\leqslant x}(1-\varepsilon_1)^{\psi(p)/\tilde{A}_1}h(p)\leqslant r(1-\varepsilon_1)\log\log x+O(1), \tag{3.4}$$

then for all $T, z_1\geqslant 1$

$$\sum_{\substack{a\in\mathcal{A}\\ \psi(c_a)\leqslant z_1}} \chi_T(a) \ll M\exp\left(-\frac{\log(1+\varepsilon_1)}{\tilde{A}_1}\left(r\tilde{A}_1\,(1-\varepsilon_1)\log\log M - z_1\right)\right), \tag{3.5}$$

where the implied constants depend on $\varepsilon_i, \tilde{A}, C, \tilde{C}, \alpha, B, \tilde{B}, \theta, \xi, K, \kappa, \lambda_i, \psi$ and the implied constants in (2.6)*, but are independent of z_i, T and M.*

Proof. Fix a number $\beta>1$, which will be specified later. Define the multiplicative function $f_\beta(c)=\beta^{\psi(c)}$. If $\psi(c)\geqslant z_2$ then $1\leqslant \beta^{-z_2}f_\beta(c)$, hence,

$$\sum_{\substack{a\in\mathcal{A}\\ f_\beta(c_a)\geqslant z_2}} \chi_T(a)\leqslant \beta^{-z_2}\sum_{a\in\mathcal{A}}\chi_T(a)f_\beta(c_a).$$

The assumptions on ψ imply that $f_\beta\in\mathcal{M}\left(\beta^{\tilde{A}},\epsilon\log\beta,\beta^{\tilde{C}}\right)$ for every $\epsilon>0$. We can thus bound the sum in the right-hand side by Theorem 2.3, hence,

$$\sum_{\substack{a\in\mathcal{A}\\ \psi(c_a)\geqslant z_2}} \chi_T(a)\ll \frac{M}{\beta^{z_2}}\left(\prod_{B<p\leqslant M}(1-h(p))\sum_{k\leqslant M}\beta^{\psi(k)}h(k)\right), \tag{3.6}$$

where the implied constant is independent of z_2, T and M. We estimate the sum over k in (3.6) by applying Lemma 2.7 with $F=h$ and $G=\beta^\psi$. This yields, by combining with (3.2), the upper bound

$$\sum_{k\leqslant M}\beta^{\psi(k)}h(k)\ll\exp\left(\beta^{\tilde{A}}\sum_{p\leqslant M}h(p)\right)\ll(\log M)^{r\beta^{\tilde{A}}}.$$

By (3.2), we also have

$$\prod_{B<p\leqslant M}(1-h(p))\ll\exp\left(-\sum_{p<M}h(p)\right)\ll(\log M)^{-r}. \tag{3.7}$$

This allows us to bound the right-hand side of (3.6) by

$$\ll M\frac{(\log M)^{r(\beta^{\tilde{A}}-1)}}{\beta^{z_2}}\leqslant M\exp\left(\left(\frac{\beta^{\tilde{A}}-1}{\tilde{A}(1+\varepsilon_2)}-\log\beta\right)z_2\right)$$

due to our assumption $(\log M)^r\leqslant\exp(z_2/(\tilde{A}(1+\varepsilon_2)))$. Define $\beta=(1+\varepsilon_2)^{1/\tilde{A}}$. Then

$$\frac{\beta^{\tilde{A}}-1}{\tilde{A}(1+\varepsilon_2)}-\log\beta=\frac{1}{\tilde{A}}\left(\frac{\varepsilon_2}{1+\varepsilon_2}-\log(1+\varepsilon_2)\right)=-\frac{\mathfrak{c}(\varepsilon_2)}{\tilde{A}}.$$

This concludes the proof of (3.3).

To prove (3.5) we fix $\beta=(1-\varepsilon_1)^{1/\tilde{A}_1}\in(0,1)$ so that

$$\frac{\beta^{\tilde{A}_1}-1}{\tilde{A}_1(1-\varepsilon_1)}-\log\beta=-\frac{\mathfrak{c}(-\varepsilon_1)}{\tilde{A}_1}<0. \tag{3.8}$$

If $\psi(c) \leqslant z_1$ then $\beta^{-z_1} f_\beta(c) \geqslant 1$, hence,

$$\sum_{\substack{a\in\mathcal{A}\\ f_\beta(c_a)\leqslant z_1}} \chi_T(a) \leqslant \beta^{-z_1} \sum_{a\in\mathcal{A}_1} \chi_T(a) f_\beta(c_a).$$

The assumptions on ψ imply that $f_\beta \in \mathcal{M}(1,\epsilon,1)$ for every $\epsilon > 0$. We can thus bound the sum in the right-hand side by Theorem 2.3, hence,

$$\sum_{\substack{a\in\mathcal{A}\\ \psi(c_a)\leqslant z_1}} \chi_T(a) \ll \frac{M}{\beta^{z_1}} \left(\prod_{B<p\leqslant M} (1-h(p)) \sum_{k\leqslant M} \beta^{\psi(k)} h(k) \right), \tag{3.9}$$

where the implied constant is independent of z_1, T and M. We estimate the sum over k in (3.9) by applying Lemma 2.7 with $F = h$ and $G = \beta^\psi$. By (3.4), this yields the upper bound

$$\sum_{k\leqslant M} \beta^{\psi(k)} h(k) \ll \exp\left(\sum_{p\leqslant M} \beta^{\psi(p)} h(p) \right) \ll (\log M)^{r\beta^{\tilde{A}_1}}.$$

Combining with (3.7), this allows us to bound the right-hand side of (3.9) by

$$\begin{aligned} &\ll \frac{M}{\beta^{z_1}} (\log M)^{(\beta^{\tilde{A}_1}-1)r} \leqslant M \exp\left((\beta^{\tilde{A}_1}-1) r \log\log M - z_1 \log\beta \right) \\ &\leqslant M \exp\left((\log\beta) \left(r\tilde{A}_1 (1-\varepsilon_1) \log\log M - z_1 \right) \right), \end{aligned}$$

due to $-\log\beta < \frac{1-\beta^{\tilde{A}_1}}{\tilde{A}_1(1-\varepsilon_1)}$ from (3.8). Finally observe that

$$\log\beta = \frac{1}{\tilde{A}_1} \log(1-\varepsilon_1) < -\frac{1}{\tilde{A}_1} \log(1+\varepsilon_1)$$

since $1/(1-\varepsilon_1) > 1+\varepsilon_1$. This concludes the proof of (3.3). □

3.4. **Proof of Theorem 1.5.** Recall that $1/2 \leqslant h_2(D) 2^{-\omega(D)} \leqslant 1$. Hence, $h_2(D) \geqslant z_4$ implies that $\omega(D) \geqslant \frac{\log z_4}{\log 2}$. We use Theorem 3.2 with $\psi = \omega$, $z_2 = (\log z_4)/(\log 2)$ and

$$\mathcal{A} = \{D \text{ fundamental discriminant}\},\ \chi_T(D) = (h_3(D)-1) \mathbb{1}_{|D|\leqslant T}(D),\ c_D = D.$$

As in the proof of Theorem 1.1, we pick h to be the multiplicative function defined by $h(q) = g_1(q)/q$. We can check that (3.2) is satisfied with $r = 1$. Lemma 3.1 shows that the level of distribution assumption in Definition 2.2 is satisfied with $M(T) = \frac{4}{\pi^2} T$, $\theta = \frac{1}{30}$ and $\xi = \frac{1}{32}$. The rest of the assumptions of Theorem 3.2 can be readily verified. To bound the contribution of the cases with $h_2(D) \leqslant z_3$ we use (3.5) with $z_1 = (\log z_3)/(\log 2)$.

3.5. **The proof of Theorem 1.6.** We use Theorem 3.2 with

$$\mathcal{A} = \{K \text{ quintic } S_5\},\ \chi_T(K) = \mathbb{1}_{|\Delta_K|\leqslant T}(K),\ c_K = \Delta_K,\ \psi = \omega.$$

To show that (c_K) is equidistributed, we use Bhargava's parametrization of S_5-extensions. The estimates from [45, Theorem 6] implies that

$$C_d(T) = h(d) \frac{13}{120} T + O(T^{\frac{399}{400}})$$

for all $d \leqslant T^{\frac{3}{400}}$, where h is multiplicative, and satisfies the properties that

$$h(p) = \frac{(p+1)(p^2+p+1)}{p^4+p^3+2p^2+2p+1} = \frac{1}{p} + O\Big(\frac{1}{p^2}\Big)$$

for $p > 5$, and $h(p^e) \leqslant h(p^2) \ll 1/p^2$ for all $e \geqslant 2$. Moreover $h(p^e) = 0$ for all $e \geqslant 5$, $p > 5$, and also $h(p^e) = 0$ for $p \in \{2,3,5\}$ and e sufficiently large ($e \geqslant 100$ suffices). Hence (2.6) is satisfied with $M(T) = \frac{13}{120} T$, $\theta = \frac{3}{400}$ and $\xi = \frac{1}{400}$. We let $\psi = \omega$ and $r = \tilde{A} = \tilde{A}_1 = 1$.

The assumption $\psi(n) \leqslant \epsilon \log n + \tilde{C}$ is met due to the bound $\omega(n) \ll (\log n)/(\log\log n)$ that is a consequence of the Prime Number Theorem. An application of Theorem 3.2 then concludes the proof.

3.6. **The proof of Theorem 1.7.** The arguments are similar to the ones in the proof of Theorem 1.6. The only difference is that the uniformity estimates are imported from the work of Bhargava–Taniguchi–Thorne [6, Theorem 1.3]. Specifically, with the notation

$$\mathcal{A} = \{K \text{ cubic } S_3\},\ \chi_T(K) = \mathbb{1}_{|\Delta_K| \leqslant T}(K),\ c_K = \Delta_K$$

one has

$$C_d(T) = h(d)\frac{4}{12\zeta(3)}T + O(T^{\frac{5}{6}})$$

for all $d \leqslant T^{\frac{1}{20}}$, where the implied constant is absolute and h is multiplicative satisfying

$$h(p^e) = \begin{cases} (p+1)/(p^2+p+1) & \text{if } e = 1, \\ O(1/p^2) & \text{if } e = 2, \\ 0 & \text{if } e \geqslant 3 \end{cases}$$

for every prime $p > 3$. We also have $h(2^e) = 0$ and $h(3^e) = 0$ for sufficiently large e. Since $h(p) = 1/p + O(1/p^2)$ we can employ Theorem 3.2 with $\psi = \omega$ and $r = \tilde{A} = \tilde{A}_1 = 1$.

4. Diophantine equations

4.1. **Three squares.** Denote $L(1, \chi_{-N}) = \sum_{m=1}^{\infty} \left(\frac{-N}{m}\right) m^{-1}$, where $(\frac{\cdot}{\cdot})$ is the Legendre symbol. A theorem of Gauss states that for positive square-free $N \equiv 3 \,(\mathrm{mod}\ 8)$ one has

$$\sharp\{\mathbf{x} \in \mathbb{Z}^3 : x_1^2 + x_2^2 + x_3^2 = N\} = \frac{8}{\pi} L(1, \chi_{-N}) N^{1/2}.$$

The main result of this section allows to put multiplicative weights on each variable. For $N \in \mathbb{N}$ and an arithmetic function f we define

$$c_f(N) := \prod_{p \mid N} \left(1 + \left(\frac{-1}{p}\right) \frac{(f(p) - 1)}{p}\right). \tag{4.1}$$

Theorem 4.1. *Fix any $A > 0, s > 0, \alpha_1, \alpha_2, \alpha_3 \in \{0, 1\}$ and let $\alpha = \sum_{i=1}^{3} \alpha_i$. Assume that $f : \mathbb{N} \to [0, \infty)$ is a multiplicative function such that $f(ab) \leqslant \tau(a)^s f(b)$ holds for all $a, b \in \mathbb{N}$. Then for all positive square-free integers $N \equiv 3 \,(\mathrm{mod}\ 8)$ we have*

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z}\setminus\{0\})^3 \\ x_1^2 + x_2^2 + x_3^2 = N}} \prod_{i=1}^{3} f(|x_i|)^{\alpha_i} \ll L(1, \chi_{-N}) N^{1/2} \left(c_f(N)^\alpha \exp\left(\alpha \sum_{p \leqslant N} \frac{f(p) - 1}{p}\right) + \frac{1}{(\log N)^A}\right),$$

where the implied constant is independent of N.

If, in addition, for each $L \geqslant 1$ one has $\inf\{f(m) : \Omega(m) \leqslant L\} > 0$, then for all positive square-free integers $N \equiv 3 \,(\mathrm{mod}\ 8)$ we have

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z}\setminus\{0\})^3 \\ x_1^2 + x_2^2 + x_3^2 = N}} \prod_{i=1}^{3} f(|x_i|)^{\alpha_i} \gg L(1, \chi_{-N}) N^{1/2} \left(c_f(N)^\alpha \exp\left(\alpha \sum_{p \leqslant N} \frac{f(p) - 1}{p}\right) - \frac{1}{(\log N)^A}\right),$$

where the implied constant is independent of N.

The case $\alpha = 1$ corresponds to imposing weights on one coefficient only; its proof is given in §4.3. It is a straightforward combination of Theorem 2.3 and work of Duke [21], the details of which are recalled in §4.2. The proof in the cases with $\alpha = 2, 3$ require additional sieving arguments and for reasons of space we give the full details only in the harder case $\alpha = 3$. Specifically, in §4.4 we transform the sums into ones where $\prod_{i=1}^3 f(|x_i|)$ is replaced by $f(\prod_{i=1}^3 |x_i|)$. Subsequently, in §4.5 we prove the requisite level of distribution for the transformed sums. Finally, in §4.6 we prove Theorem 4.1.

4.2. **Input from cusp forms.** The main result in this subsection is Lemma 4.4; it regards the number of solutions of $x_1^2 + x_2^2 + x_3^2 = N$, with each x_i divisible by an arbitrary integer d_i. This is closely related to work of Brüdern–Blomer [8, Lemma 2.2] in the case where each d_i is square-free. The proof of Lemma 4.4 combines the work of Duke [21] with that of Jones [38]. We recall [21, Theorem 2, Equation (3)]:

Lemma 4.2 (Duke). *There exists a positive constant κ such that for every positive definite quadratic integer ternary form q and every square-free integer N one has*

$$\sharp\{\mathbf{x} \in \mathbb{Z}^3 : q(\mathbf{x}) = N\} = \kappa L(1, \chi_{q,N}) \mathfrak{S}(q, N) \frac{\sqrt{N}}{\sqrt{D}} + O(D^6 N^{1/2-1/30}),$$

where the implied constant is absolute, D is the determinant of the matrix $(\partial^2 q/\partial x_i \partial x_j)$, $\chi_{q,N}$ is the Dirichlet character $\chi_{q,N}(m) = (\frac{-2D\mathrm{disc}(\mathbb{Q}(\sqrt{N}))}{m})$ and

$$\mathfrak{S}(q, N) := \prod_{p|2D} \lim_{\lambda\to\infty} \frac{\sharp\left\{\mathbf{x} \in (\mathbb{Z}/p^\lambda\mathbb{Z})^3 : q(\mathbf{x}) \equiv N \left(\mathrm{mod}\ p^\lambda\right)\right\}}{p^{2\lambda}}.$$

Note that the definition of $\mathfrak{S}$ in [21, Equation (4)] involves a finite value of λ, however, this is equivalent since these densities stabilise owing to the fact that N is square-free. We now specify the constant κ. When $q = \sum_{i=1}^3 x_i^2$ we have $D = 8$, hence,

$$\sharp\{\mathbf{x} \in \mathbb{Z}^3 : x_1^2 + x_2^2 + x_3^2 = N\} = \frac{\kappa \mathcal{N}(16^3)}{2\sqrt{2}} \sqrt{N} \sum_{m=1}^{\infty} \left(\frac{-4N}{m}\right) \frac{1}{m} + O(N^{1/2-1/30}),$$

where $\mathcal{N}(m) = \sharp\{\mathbf{x} \in (\mathbb{Z}/m\mathbb{Z})^3 : x_1^2 + x_2^2 + x_3^2 \equiv N\ (\mathrm{mod}\ m)\} m^{-2}$.

Lemma 4.3. *For any integer $N \equiv 3\ (\mathrm{mod}\ 8)$ and $t \geqslant 3$, the number of solutions of $x_1^2 + x_2^2 + x_3^2 \equiv N\ (\mathrm{mod}\ 2^t)$ is 4^t.*

Proof. Since $N \equiv 3\ (\mathrm{mod}\ 8)$ every x_i must be odd. Let x_1, x_2 run through all odd elements $(\mathrm{mod}\ 2^t)$ and then count the number of x_3 for which $x_3^2 \equiv a\ (\mathrm{mod}\ 2^t)$, where $a \equiv N - x_1^2 - x_2^2\ (\mathrm{mod}\ 2^t)$. Here $N \equiv 3\ (\mathrm{mod}\ 8)$, hence, $a \equiv 1\ (\mathrm{mod}\ 8)$. Now we use the following fact: for $t \geqslant 3$ and each $a \in \mathbb{Z}/2^t\mathbb{Z}$ with $a \equiv 1\ (\mathrm{mod}\ 8)$, the number of solutions of $x^2 \equiv a\ (\mathrm{mod}\ 2^t)$ is 4. This gives a total number of solutions $2^{t-1} \cdot 2^{t-1} \cdot 4 = 4^t$. □

In particular, $\mathcal{N}(16^3) = 1$. We obtain

$$\sharp\{\mathbf{x} \in \mathbb{Z}^3 : x_1^2 + x_2^2 + x_3^2 = N\} = \frac{\kappa}{2\sqrt{2}} \sqrt{N} \sum_{m=1}^{\infty} \left(\frac{-4N}{m}\right) \frac{1}{m} + O(N^{1/2-1/30}).$$

By [1, Theorem B, page 99] this equals $\frac{16}{\pi} \mathcal{L} \sqrt{N}$, where $\mathcal{L} := \sum_{m=1}^\infty \left(\frac{-4N}{m}\right) \frac{1}{m}$. By Siegel's theorem we have $\mathcal{L} \gg N^{-1/60}$, hence,

$$\frac{\kappa}{2\sqrt{2}} - \frac{16}{\pi} = O\left(\frac{1}{N^{1/30}\mathcal{L}}\right) = O\left(\frac{1}{N^{1/60}}\right),$$

which shows that $\kappa = 32\sqrt{2}/\pi$.

Lemma 4.4. *For each* $\mathbf{c} \in \mathbb{N}^3$ *and positive square-free* $N \equiv 3 \,(\mathrm{mod}\ 8)$ *we have*

$$\sharp\left\{\mathbf{x} \in \mathbb{Z}^3 : \sum_{i=1}^{3}(c_i x_i)^2 = N\right\} = \frac{8}{\pi} L(1, \chi_{-N}) h_N(\mathbf{c}) N^{1/2} + O((c_1c_2c_3)^{12} N^{1/2-1/30}),$$

where the implied constant is absolute, $h_N(\mathbf{c})$ *is given by*

$$\frac{2^{\sharp\{p \mid c_1c_2c_3\}}}{c_1c_2c_3} \prod_{p \mid c_1c_2c_3} \left(1 - \frac{(\frac{-N}{p})}{p}\right) \mathbb{1}((4.2) - (4.4)) \prod_{\substack{p \mid c_1c_2c_3 \\ p \text{ divides exactly one } c_i}} 2^{-\mathbb{1}(p \nmid N)} \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right)$$

and

$$2 \nmid c_1c_2c_3, \tag{4.2}$$

$$\gcd(c_1, c_2, c_3) = 1, \tag{4.3}$$

$$p \text{ divides exactly two } c_i \Rightarrow \left(\frac{N}{p}\right) = 1, \tag{4.4}$$

$$p \mid N, p \text{ divides exactly one } c_i \Rightarrow p \equiv 1 \,(\mathrm{mod}\ 4)\,. \tag{4.5}$$

Proof. We use Lemma 4.2 with $q = \sum_{i=1}^3 c_i^2 x_i^2$ so that $D = 8(c_1c_2c_3)^2$. Let us note that $\mathrm{disc}(\mathbb{Q}(\sqrt{N})) = 4N$, hence, the character $\chi_{q,N}(m)$ is given by

$$\left(\frac{-2 \cdot 8(c_1c_2c_3)^2 \cdot 4N}{m}\right) = \left(\frac{-N}{m}\right) \mathbb{1}(\gcd(2c_1c_2c_3, m) = 1).$$

Therefore, the value of the corresponding L-function at 1 is

$$\prod_{p \nmid 2c_1c_2c_3} \frac{1}{\left(1 - \frac{(\frac{-N}{p})}{p}\right)} = \frac{L(1, \chi_{-N})}{2} \prod_{\substack{p \mid c_1c_2c_3 \\ p \neq 2}} \left(1 - \frac{(\frac{-N}{p})}{p}\right).$$

To work out the term $\mathfrak{S}$ we use the work of Jones [38]. In the terminology of [38, Theorem 1.3] we take $Q = \sum_{i=1}^3 (c_i x_i)^2, m = N$. When $p \neq 2$ divides exactly one of the c_i, say, c_3, then we take $a = c_1^2, b_1 = 0$ and [38, Equation (1.5)] shows that the p-adic factor in $\mathfrak{S}$ equals

$$\mathbb{1}(p \mid N) 2 \left(1 - \frac{1}{p}\right) \mathbb{1}(p \equiv 1 \,(\mathrm{mod}\ 4)) + \mathbb{1}(p \nmid N) \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right).$$

If p divides exactly two of the c_i's, say c_2 and c_3 then by taking $a = c_1^2$ in [38, Equation (1.4)] shows that the p-adic factor in $\mathfrak{S}$ becomes 2 or 0, according to whether $(\frac{N}{p}) = 1$ or not. Finally, since N is square-free, there is no prime p that divides every c_i since that would imply that p^2 divides N. Furthermore, by Lemma 4.3 the 2-adic density equals 1. □

4.3. **The one-variable case.** The case $\alpha_1 = 1, \alpha_2 = \alpha_3 = 0$ can be treated in a straightforward manner and we deal with it in this subsection. We use Theorem 2.3 with

$$\mathcal{A} = (\mathbb{Z} \setminus \{0\})^3, c_a = |y_1|, T = N, \chi_N(a) = \mathbb{1}_{\{N\}}(y_1^2 + y_2^2 + y_3^2), M(N) = \frac{8}{\pi} L(1, \chi_{-N}) N^{1/2}.$$

To show that assumption (2.6) holds we use Lemma 4.4 to infer that

$$\sum_{\substack{a \in \mathcal{A} \\ d \mid y_1}} \chi_N(a) = \frac{8}{\pi} L(1, \chi_{-N}) G_N(d) N^{1/2} + O(d^{12} N^{1/2-1/30}),$$

where G_N is multiplicative and defined as

$$p^e G_N(p^e) = \mathbb{1}(p \neq 2) 2^{\mathbb{1}(p|N)} \mathbb{1}(p \mid N \Rightarrow p \equiv 1 \,(\mathrm{mod}\ 4)) \left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right)\left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right).$$

We have $N^{1/2-\epsilon} \ll M(N) \ll N^{1/2+\epsilon}$ for every fixed $\epsilon > 0$ by Siegel's theorem. Now let θ, ξ be positive constants that will be fixed later. For all $d \leqslant M(N)^\theta$ and any fixed positive constant ϵ, we have

$$d^{12} N^{1/2-1/30} \ll M(N)^{12\theta+1-1/15+\epsilon}.$$

Hence, (2.6) holds for some $\xi > 0$ as long as $12\theta < 1/15$. In particular, it holds when $\theta = 10^{-3}$. Assumptions (2.2)-(2.3) hold due to the bound $G_N(p^e) \ll p^{-e}$ that is valid with an absolute implied constant. One can take $\kappa = 10$ in (2.1) due to the estimate $G_N(p) \leqslant 10/p$ that holds for all primes p.

Thus, Theorem 2.3 shows that

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z}\setminus\{0\})^3 \\ x_1^2+x_2^2+x_3^2=N}} f(|x_1|) \ll M(N) \prod_{1 \ll p \leqslant M(N)} (1 - G_N(p)) \sum_{a \leqslant M(N)} f(a) G_N(a).$$

By Lemma 2.10 we infer that

$$\prod_{1 \ll p \leqslant M(N)} (1 - G_N(p)) \sum_{a \leqslant M(N)} f(a) G_N(a) \asymp \exp\left(\sum_{p \leqslant M(N)} (f(p) - 1) G_N(p)\right),$$

where the implied constants are independent of N. The sum over p equals

$$\sum_{p|N} \left(\frac{-1}{p}\right) \frac{f(p)-1}{p} + \sum_{p \leqslant N} \frac{f(p)-1}{p} + O\Big(1 + \sum_{\substack{p > M(N) \\ p|N}} \frac{1}{p} + \sum_{M(N) < p \leqslant N} \frac{1}{p}\Big).$$

We have $M(N) \gg N^{1/4}$ by Siegel's theorem, thus, the error term is $\ll 1 + N^{-1/4}\omega(N)$ is bounded, something that suffices for the proof of the upper bound. To prove the lower bound we apply Theorem 2.4 in the same manner.

4.4. **Transformation.** To transform the sums in Theorem 4.1 a preliminary step is to show that for most integer solutions of $x_1^2 + x_2^2 + x_3^2 = N$ the common divisors of each pair (x_i, x_j) are typically small. In Lemma 4.5 we show that these divisors are frequently smaller than any fixed power of N, while in Lemmas 4.6-4.7 we show that these divisors are smaller than a power of $\log N$. The latter task combines equidistribution in the form of Lemma 4.4 with a "level-lowering" mechanism that is grounded on work of Brady [9].

Lemma 4.5. *Fix any $s > 0$ and $\delta \in (0, 1/6)$. Then for any positive square-free integer $N \equiv 3 \,(\mathrm{mod}\ 4)$ we have*

$$\sum_{c > N^\delta} \sum_{\substack{\mathbf{x} \in (\mathbb{Z}\setminus\{0\})^3, c|(x_1,x_2) \\ x_1^2+x_2^2+x_3^2=N}} (\tau(x_1)\tau(x_2)\tau(x_3))^s \ll N^{1/2-\delta/2} L(1, \chi_{-N}),$$

where the implied constant depends at most on δ and s.

Proof. Since $c^2 \mid x_1^2 + x_2^2 = N - x_3^2$, we obtain the upper bound

$$\ll_{\epsilon,s} N^\epsilon \sum_{c > y} \sum_{\substack{|x_3| \leqslant N^{1/2} \\ c^2 | N - x_3^2}} r_2(N - x_3^2) \ll_\epsilon N^{2\epsilon} \sum_{c > y} \sharp\{|x_3| \leqslant N^{1/2} : c^2 \mid N - x_3^2\}.$$

We shall now split in two ranges:

$$N^\delta < c \leqslant N^{1/2-\delta} \quad \text{and} \quad c > N^{1/2-\delta}.$$

For the second range we write $D = (N - x_3^2)/c^2$ and note that $D \leqslant N^{2\delta}$. Swapping summation thus leads to

$$\sum_{c > N^{1/2-\delta}} \sharp\{|x_3| \leqslant N^{1/2} : c^2 \mid N - x_3^2\} \leqslant \sum_{1 \leqslant D \leqslant N^{2\delta}} \sharp\{c, x_3 \in \mathbb{Z} : N = x_3^2 + Dc^2\}.$$

Since $D > 0$, the unit group of $\mathbb{Q}(\sqrt{-D})$ is bounded independently of D. From the theory of binary quadratic forms we can then infer that

$$\sharp\{c, x_3 \in \mathbb{Z} : N = x_3^2 + Dc^2\} \ll \sum_{m \mid N} \left(\frac{D}{m}\right) \ll_\epsilon N^\epsilon,$$

where the implied constant depends only on ϵ. This gives the overall bound

$$\ll N^{3\epsilon+2\delta} \ll N^{4\epsilon+2\delta} L(1, \chi_{-N})$$

by Siegel's estimate. Using $\delta < 1/6$ we see that $2\delta < 1/2 - \delta$, thus, taking $\epsilon = \delta/8$ gives the bound $N^{1/2-\delta/2} L(1, \chi_{-N})$, which is satisfactory.

We next deal with the first range. Splitting in progressions we get

$$\sharp\{|x_3| \leqslant N^{1/2} : c^2 \mid N - x_3^2\} \leqslant \sum_{\substack{t \in \mathbb{Z}/c^2\mathbb{Z} \\ c^2 \mid N - t^2}} \left(\frac{N^{1/2}}{c^2} + 1\right) \ll N^\epsilon \left(\frac{N^{1/2}}{c^2} + 1\right).$$

Summing over the range $N^\delta < c \leqslant N^{1/2-\delta}$ this gives $\ll N^{1/2-\delta+\epsilon}$, which is acceptable upon choosing a suitably small value for ϵ. □

The proof of the next two results uses crucially that the range $c > N^\delta$ has already been dealt with.

Lemma 4.6. *Fix arbitrary $s > 0$ and let $\beta = 60(100s - 1)/7$. For any $\mathbf{c} \in \mathbb{N}^3$ and positive square-free integer $N \equiv 3 \pmod 8$ we have*

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^3, c_i \mid x_i \forall i \\ x_1^2 + x_2^2 + x_3^2 = N}} (\tau(x_1)\tau(x_2)\tau(x_3))^s \ll L(1, \chi_{-N}) N^{1/2} (\log\log N)^2 (\log N)^{3 \cdot 2^{\beta+1}} \prod_{i=1}^3 \frac{\tau(c_i)^{\beta+3}}{c_i}$$
$$+ N^{1/2-1/100} (c_1 c_2 c_3)^{12},$$

where the implied constant depends at most on β and s.

Proof. The function $H(\delta) = \delta \log_2(\delta^{-1}) + (1-\delta) \log_2(1-\delta)^{-1}$ satisfies $H(7/6000) > 1/100$. Taking $\delta = 7/6000$ we see that the assumption $7\beta + 60 = 6000s$ allows us to use [9, Theorem 4]. This yields the following bound for the sum over $\mathbf{x}$ in the lemma:

$$\ll_{\beta,s} \sum_{\substack{\mathbf{d} \in \mathbb{N}^3 \\ d_i \leqslant N^{\delta/2} \forall i}} (\tau(d_1)\tau(d_2)\tau(d_3))^\beta \sharp\{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^3 : x_1^2 + x_2^2 + x_3^2 = N, [c_i, d_i] \mid x_i \forall i\},$$

where $[\cdot, \cdot]$ denotes the least common multiple. By Lemma 4.4 this can be bounded by

$$\ll \sum_{\substack{\mathbf{d} \in \mathbb{N}^3 \\ d_i \leqslant N^{\delta/2} \forall i}} \Big(\prod_{i=1}^3 \tau(d_i)^\beta\Big) (\log\log N)^2 \Big(L(1, \chi_{-N}) N^{1/2} \prod_{i=1}^3 \frac{\tau([c_i, d_i])}{[c_i, d_i]} + N^{1/2-1/30} \prod_{i=1}^3 [c_i, d_i]^{12}\Big),$$

where we used the following standard bound for $t = c_1 c_2 c_3$,

$$\prod_{p \mid t} \left(1 + \frac{1}{p}\right) \ll \prod_{p \mid t} \left(1 - \frac{1}{p}\right)^{-1} = \frac{t}{\phi(t)} \ll \log\log t. \tag{4.6}$$

Using $[c_i, d_i] \leqslant c_i d_i$ we can see that the second part of this sum is

$$\ll N^{1/2-1/30} \prod_{i=1}^{3} c_i^{12} \left(\sum_{d \leqslant N^{\delta/2}} \tau(d)^{\beta} d^{12} \right)^3 \ll N^{1/2-1/30+19.5\delta} (\log N)^{3(2^{\beta}-1)} \prod_{i=1}^{3} c_i^{12},$$

which is $\ll N^{1/2-1/30+20\delta} \prod c_i^{12}$. Our choice $\delta = 7/6000$ makes sure that this is $\ll N^{1/2-1/100} \prod c_i^{12}$. The first part of the sum is $\ll L(1, \chi_{-N})(\log \log N)^2 N^{1/2} \prod \mathcal{S}(c_i)$, where

$$\mathcal{S}(c) := \sum_{d \leqslant N} \tau(d)^{\beta} \frac{\tau([c,d])}{[c,d]} \leqslant \frac{\tau(c)}{c} \sum_{d \leqslant N} \tau(d)^{\beta+1} \frac{\gcd(c,d)}{d}.$$

Writing $m = \gcd(c,d)$ and $d = mt$, the sum over d can be seen to be at most

$$\sum_{m|c} m \sum_{\substack{d \leqslant N \\ m|d}} \frac{\tau(d)^{\beta+1}}{d} \leqslant \sum_{m|c} \tau(m)^{\beta+1} \sum_{t \leqslant N} \frac{\tau(t)^{\beta+1}}{t} \ll \tau(c)^{\beta+2} (\log N)^{2^{\beta+1}},$$

which is sufficient. □

Lemma 4.7. *Fix any positive A and s. Then for any positive square-free $N \equiv 3 \pmod 8$ we have*

$$\sum_{c > (\log N)^A} \sum_{\substack{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^3, c | (x_1, x_2) \\ x_1^2 + x_2^2 + x_3^2 = N}} (\tau(x_1)\tau(x_2)\tau(x_3))^s \ll_{A,s} L(1, \chi_{-N}) N^{1/2} (\log N)^{\rho(s) - A/2},$$

where $\rho(s) = 6 \cdot 2^{(6000s-60)/7}$ and the implied constant depends at most on A and s.

Proof. Fix any $\delta > 0$. By Lemma 4.5 we can discard the contribution of $c > N^{\delta}$. For the remaining c we employ Lemma 4.6 with β defined by $7\beta + 60 = 6000s$. We obtain

$$\begin{aligned} &\ll \sum_{(\log N)^A < c \leqslant N^{\delta}} \left(L(1, \chi_{-N}) N^{1/2} (\log \log N)^2 (\log N)^{3 \cdot 2^{\beta+1}} \frac{\tau(c)^{2\beta+6}}{c^2} + N^{1/2-1/100} c^{24} \right) \\ &\ll L(1, \chi_{-N}) N^{1/2} (\log N)^{3 \cdot 2^{\beta+1} - A/2} + N^{1/2-1/100+25\delta}. \end{aligned}$$

Choosing sufficiently small δ and using Siegel's bound we obtain

$$N^{-1/100+25\delta} \ll N^{-1/1000} \ll L(1, \chi_{-N}) (\log N)^{3 \cdot 2^{\beta+1} - A/2},$$

which is sufficient. □

Define for $N, m_1, m_2, m_3 \in \mathbb{N}$ the function

$$\mathcal{R}_{\mathbf{m}}(N) := \sum_{\substack{\mathbf{y} \in (\mathbb{Z} \setminus \{0\})^3 : \sum_i (m_j m_k y_i)^2 = N \\ \gcd(y_i, y_j) = 1 \forall i \neq j}} f(y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3}).$$

Lemma 4.8. *Fix any $A > 0$. In the setting of Theorem 4.1 we have*

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^3 \\ x_1^2 + x_2^2 + x_3^2 = N}} \prod_{i=1}^{3} f(x_i)^{\alpha_i} \ll \sum_{\substack{\mathbf{m} \in \mathbb{N}^3, (4.7) \\ \max m_i \leqslant (\log N)^A}} \mathcal{R}_{\mathbf{m}}(N) \prod_{i=1}^{3} \tau(m_i)^{2s} + \frac{L(1, \chi_{-N}) N^{1/2}}{(\log N)^{A/2 - \rho(s)}},$$

where $\rho(s)$ is as in Lemma 4.7, the implied constant depends at most on s, A and

$$\gcd(m_i, 2m_j) = 1 \forall i \neq j \quad p \mid m_1 m_2 m_3 \Rightarrow \left(\frac{N}{p} \right) = 1. \tag{4.7}$$

Proof. By our assumption $f \leqslant \tau^s$ and Lemma 4.7 we may write the sum over $\mathbf{x}$ as

$$\sum_{\substack{\mathbf{x}\in(\mathbb{Z}\setminus\{0\})^3, x_1^2+x_2^2+x_3^2=N \\ \gcd(x_i,x_j)\leqslant(\log N)^A \forall i\neq j}} \prod_{i=1}^3 f(x_i)^{\alpha_i} + O(L(1,\chi_{-N})N^{1/2}(\log N)^{-A/2+\rho(s)}).$$

For $\{i,j,k\} = \{1,2,3\}$ we let $m_i = \gcd(x_j,x_k)$ so that the m_i are coprime in pairs due to $\gcd(x_1,x_2,x_3)=1$ that can be inferred from the fact that N is square-free and $N=\sum_i x_i^2$. Hence, letting $y_i := x_i/(m_j m_k)$ we see that $m_i = \gcd(x_j,x_k)$ is equivalent to $1=\gcd(m_k y_j, m_j y_k)$. We obtain

$$\sum_{\substack{\mathbf{m}\in\mathbb{N}^3 \\ \gcd(m_i,m_j)=1\forall i\neq j \\ m_i\leqslant(\log N)^A\forall i}} \sum_{\substack{\mathbf{y}:\sum_i (m_j m_k y_i)^2=N \\ \gcd(y_i,y_j)=1\forall i\neq j}} \prod_{i=1}^3 f(m_j m_k y_i)^{\alpha_i} + O(L(1,\chi_{-N})N^{1/2}(\log N)^{-A/2+\rho(s)}).$$

We omitted the condition $\gcd(y_i,m_i)=1$ as it is implied by the fact that N is square-free and a sum of integer multiples of m_i^2 and y_i^2. Our assumption $f(ab)\leqslant\tau(a)^s f(b)$ allows us to write

$$\prod_{i=1}^3 f(m_j m_k y_i)^{\alpha_i} \leqslant \prod_{i=1}^3 \tau(m_j)^s\tau(m_k)^s f(y_i)^{\alpha_i} = f(y_1^{\alpha_1}y_2^{\alpha_2}y_3^{\alpha_3})\prod_{i=1}^3\tau(m_i)^{2s},$$

since the y_i are pairwise coprime and f is multiplicative. The condition that each prime divisor p of m_i must satisfy $(\frac{N}{p})=1$ comes from the fact that each m_i divides two of the coefficients of $\sum_i (m_j m_k y_i)^2$ and is coprime to the third. Finally, if one of the m_i is even, then 4 divides $N-(m_j m_k y_i)^2$, which is impossible owing to $N\equiv 3 \pmod 4$. □

4.5. **Level of distribution.** Throughout this subsection $\mathbf{m}$ is a fixed vector in $\mathbb{N}^3$ satisfying (4.7). For positive integers d,N define

$$C_d(N) := \sharp\left\{\mathbf{y}\in\mathbb{Z}^3 : \begin{array}{l}(m_2m_3y_1)^2+(m_1m_3y_2)^2+(m_1m_2y_3)^2=N, \\ \gcd(y_i,y_j)=1\forall i\neq j, \quad d\mid y_1y_2y_3\end{array}\right\}.$$

The main result is Lemma 4.11; it gives a level of distribution result for $C_d(N)$ that will subsequently be fed into Theorem 2.3 to bound $\mathcal{R}_{\mathbf{m}}(N)$.

We start with a sieving argument that deals with the coprimality of the y_i.

Lemma 4.9. *Keep the setting of Theorem 4.1 and fix any $\delta\in(0,1/9)$. For all $\mathbf{m}$ as in (4.7) and all $d\in\mathbb{N}$ we have*

$$C_d(N) = \sum_{\substack{\mathbf{d}\in\mathbb{N}^3, d=d_1d_2d_3 \\ \gcd(d_i,d_j)=1\forall i\neq j \\ \gcd(d_i,m_i)=1\forall i}} \sum_{\substack{\mathbf{b}\in\mathbb{N}^3, \max b_i\leqslant N^{\delta_i}\forall i \\ \gcd(b_i,b_j)=1\forall i\neq j \\ \gcd(b_i,d_i m_j)=1\forall i\neq j}} \mu(b_1)\mu(b_2)\mu(b_3)C_{\mathbf{b},d}(N) + O(N^{1/2-\delta/400}L(1,\chi_{-N})),$$

where $\delta_1=\delta/100, \delta_2=\delta/10, \delta_3=\delta$, the quantity $C_{\mathbf{b},d}(N)$ is given by

$$\sharp\{\mathbf{t}\in\mathbb{Z}^3 : N=(m_2m_3[d_1,b_2b_3]t_1)^2+(m_1m_3[d_2,b_1b_3]t_2)^2+(m_1m_2[d_3,b_1b_2]t_3)^2\}$$

and the implied constant depends at most on δ.

Proof. Since y_i are coprime in pairs in $C_d(N)$, we can write $d=d_1d_2d_3$ where $d_i\mid y_i$ and the d_i are coprime in pairs. Then, $C_d(N)$ becomes

$$\sum_{\substack{\mathbf{d}\in\mathbb{N}^3, d=d_1d_2d_3 \\ \gcd(d_i,m_id_j)=1\forall i\neq j}} \sharp\left\{\mathbf{y}\in(\mathbb{Z}\setminus\{0\})^3 : \begin{array}{l}(m_2m_3y_1)^2+(m_1m_3y_2)^2+(m_1m_2y_3)^2=N, \\ \gcd(y_i,y_j)=1\forall i\neq j, \quad d_i\mid y_i\forall i\end{array}\right\}.$$

The condition $\gcd(d_i, m_i) = 1$ comes from the fact that N is square-free and the sum of squares equation. We now use the expression $\sum_{b_1 \mid (y_2,y_3)} \mu(b_1)$ to detect the coprimality of y_2 and y_3. The contribution of $b_1 > N^{\delta/100}$ will then be at most

$$\tau_3(d) \sum_{b_1 > N^{\delta/100}} \sharp\{\mathbf{t} \in \mathbb{Z}^3 : t_1^2 + t_2^2 + t_3^2 = N, b_1 \mid (t_2, t_3)\} \ll \tau_3(d) N^{1/2-\delta/200} L(1, \chi_{-N})$$

by Lemma 4.7. We have $d \leqslant N^3$ due to $d \mid y_1y_2y_3$, hence, the bound $\tau_3(d) \ll N^{\delta/400}$ shows that the contribution is $\ll N^{1/2-\delta/400} L(1, \chi_{-N})$. Next, we use $\sum_{b_2 \mid (y_1,y_3)} \mu(b_2)$ to detect the coprimality of y_1 and y_3. The contribution of $b_2 > N^{\delta/10}$ is

$$\ll \tau_3(d) \sum_{\substack{b_1 \leqslant N^{\delta/100} \\ b_2 > N^{\delta/10}}} \sharp\{\mathbf{t} \in \mathbb{Z}^3 : t_1^2 + t_2^2 + t_3^2 = N, b_2 \mid (y_1, y_3)\} \ll \tau_3(d) N^{1/2+\delta/100-\delta/20} L(1, \chi_{-N})$$

by Lemma 4.7. This can be seen to be $\ll N^{1/2-\delta/150} L(1, \chi_{-N})$ as before. Finally, using $\sum_{b_3 \mid (y_1,y_2)} \mu(b_3)$, we can see that the range $b_3 > N^{\delta}$ contributes

$$\ll \tau_3(d) \sum_{\substack{b_1 \leqslant N^{\delta/100}, b_2 \leqslant N^{\delta/10} \\ b_3 > N^{\delta}}} \sharp\{\mathbf{t} \in \mathbb{Z}^3 : t_1^2 + t_2^2 + t_3^2 = N, b_3 \mid (t_1, t_2)\} \ll N^{1/2-\delta/10} L(1, \chi_{-N}).$$

We thus obtain the expression claimed in the lemma. The conditions of the form $\gcd(b_1, b_2b_3d_1m_2m_3) = 1$ in the lemma come from the fact that N is square-free. Finally, the vectors $\mathbf{t}$ having $t_i = 0$ for some i contribute at most

$$\ll \tau_3(d) N^{\delta_1+\delta_2+\delta_3} r_2(N) \ll N^{2\delta},$$

which is acceptable by the assumption $\delta < 1/9$. □

We next apply Lemma 4.4. Denote

$$b = b_1b_2b_3, \quad m = m_1m_2m_3 \quad \text{and} \quad \mathfrak{C}_d := \prod_{\substack{p \equiv 3 \pmod 4 \\ p \mid (d,N)}} p.$$

Lemma 4.10. *Keep the setting of Lemma 4.9 and fix any $\varpi > 0$. For all $d \in \mathbb{N}$ and $\mathbf{m} \in \mathbb{N}^3$ as in* (4.7) *with the additional restriction $\max m_i \leqslant (\log N)^{\varpi}$ we have*

$$C_d(N) = \frac{8}{\pi} L(1, \chi_{-N}) N^{1/2} M_1 M_2 + O(d^{12} N^{1/2+\max\{50\delta-1/30, -\delta/800\}} (\log N)^{100\varpi} L(1, \chi_{-N})),$$

where the implied constant depends at most on δ and ϖ. Here

$$M_1 = \frac{\mathbb{1}(2 \nmid d)}{d} \frac{2^{\omega(m)}}{m^2} \prod_{p \mid m} \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right)$$

and

$$M_2 = \sum_{\mathbf{b} \in \mathbb{N}^3} 2^{\sharp\{p \mid b : p \nmid m\}} \frac{\mu(b)}{b^2} 2^{\sharp\{p \mid d : p \nmid bm\}} \prod_{p \mid b, p \nmid m} \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right) \frac{2^{\sharp\{p \mid (d,b)\}}}{2^{\sharp\{p \mid d : p \nmid bmN\}}}$$
$$\times \gcd(d, b) 3^{\sharp\{p \mid d : p \nmid bm\}} 2^{\sharp\{p \mid d, p \mid m, p \nmid b\}} \prod_{\substack{p \mid d \\ p \nmid bm}} \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right)\left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right),$$

where the sum is over $\mathbf{b}$ satisfying the further conditions

$$\mathfrak{C}_d \mid b_1b_2b_3m, \gcd(b_i, 2b_jm_j) = 1 \forall i \neq j,$$

and $(\frac{N}{p}) = 1$ for all primes $p \mid b_1b_2b_3$ with $p \nmid m$.

Proof. We employ Lemma 4.4 with $c_i = m_j m_k [d_i, b_j b_k]$ to estimate $C_d(N)$ in Lemma 4.9. The error term is

$$\ll \tau_3(d) d^{12} N^{1/2-1/30} (\log N)^{100\varpi} \prod_{i=1}^{3} \sum_{b \leqslant N^{\delta_i}} b^{24} \ll d^{12} N^{1/2-1/30+50\delta} (\log N)^{100\varpi} L(1, \chi_{-N})$$

by Siegel's bound and $\tau_3(d) \ll N^{\delta-\delta_1-\delta_2} L(1,\chi_{-N})$ that is implied by $d \leqslant N^3$.

To deal with the main term let us recall that the m_i are pairwise coprime and use the coprimality conditions on the b_i, d_i to see that (4.3) is always met. Denote $b := b_1 b_2 b_3$. Note that a prime p divides exactly two of the c_i if and only if $p \mid bm$. In addition, p divides exactly one of c_i if and only if p divides d but not bm. We get the main term

$$\frac{8}{\pi} L(1, \chi_{-N}) N^{1/2} \frac{\mathbb{1}(2 \nmid m)}{m^2} \mathfrak{K} \prod_{p \mid m} \left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right),$$

where $\mathfrak{K}$ is the sum

$$\sum\nolimits^{*} 2^{\omega(mb)} \frac{\mu(b_1)\mu(b_2)\mu(b_3)}{(b_1 b_2 b_3)^2} \prod_{\substack{p \mid b_1 b_2 b_3 \\ p \nmid m}} \left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right) \frac{\mathbb{1}(2 \nmid d) 2^{\sharp\{p \mid d : p \nmid bm\}}}{d} \mathfrak{F}(d)$$

$$\times \mathbb{1}(p \mid (d,N), p \nmid bm \Rightarrow p \equiv 1 \,(\mathrm{mod}\ 4)) \prod_{\substack{p \mid d \\ p \nmid bm}} \left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right)\left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right)\left(\frac{1}{2}\right)^{\mathbb{1}(p \nmid N)}$$

with $\sum^{*}$ taken over $\mathbf{b} \in \mathbb{N}^3$ satisfying $b_i \leqslant N^{\delta_i}$ for all i, $\gcd(b_i, 2b_j m_j) = 1$ for all $i \neq j$, and, with the further property that each prime divisor p of b that does not divide m must satisfy $(\frac{N}{p}) = 1$. The multiplicative function $\mathfrak{F}(d)$ is defined as

$$\sum_{\mathbf{d} \in \mathbb{N}^3, d = d_1 d_2 d_3} \gcd(d_1, b_2 b_3) \gcd(d_2, b_1 b_3) \gcd(d_3, b_1 b_2),$$

where the sum is subject to $\gcd(d_i, d_j m_i b_i) = 1$ for all $i \neq j$. To analyse it at prime powers p^α we use that d_i are coprime to infer that $\mathfrak{F}(p^\alpha)$ equals

$$\gcd(p^\alpha, b_2 b_3)\mathbb{1}(p \nmid b_1 m_1) + \gcd(p^\alpha, b_1 b_3)\mathbb{1}(p \nmid b_2 m_2) + \gcd(p^\alpha, b_1 b_2)\mathbb{1}(p \nmid b_3 m_3).$$

Since b_i are coprime in pairs and square-free we see that if $p \mid b_1 b_2 b_3$ then the above becomes $2p$ because $\gcd(b_i, m_j) = 1$ for all $i \neq j$. If $p \nmid b_1 b_2 b_3 m_1 m_2 m_3$ then the sum becomes 3. If $p \nmid b_1 b_2 b_3$ and $p \mid m_1 m_2 m_3$ then it becomes 2. Thus, $\mathfrak{F}(d)$ equals

$$2^{\sharp\{p \mid (b,d)\}} \gcd(b,d) 3^{\sharp\{p \mid d : p \nmid bm\}} 2^{\sharp\{p \mid (d,m) : p \nmid b\}}.$$

Using (4.6) we see that the contribution of $\mathbf{b}$ with $b_i > N_i^\delta$ for some i is

$$\ll L(1,\chi_{-N}) N^{1/2} (\log\log N)^3 \tau(d) 6^{\omega(d)} d \sum_{\substack{\mathbf{b} \in \mathbb{N}^3 \\ \exists i : b_i > N^{\delta_i}}} \frac{2^{\omega(b_1 b_2 b_3)}}{(b_1 b_2 b_3)^2} \ll L(1,\chi_{-N}) N^{1/2} d^2 N^{-\frac{1}{2 \min \delta_i}},$$

which is acceptable since $\min \delta_i > \delta/800$. To conclude the proof we note that the condition $p \mid d, p \mid N, p \nmid bm \Rightarrow p \equiv 1 \,(\mathrm{mod}\ 4)$ is equivalent to $\mathfrak{C}_d \mid bm$. □

Finally, we simplify the main term in Lemma 4.10. The error term will be obtained by taking $\delta = 80/120003$. Denote for a prime p,

$$c_p = 1 - \frac{1}{p}\left(\frac{-1}{p}\right).$$

Lemma 4.11. *Fix any $\varpi > 0$. For all $\mathbf{m} \in \mathbb{N}^3$ as in (4.7) with $\max m_i \leqslant (\log N)^{\varpi}$, all square-free positive integers $N \equiv 3 \pmod 8$ and all $d \in \mathbb{N}$ we have*

$$C_d(N) = M(N) g_N(d) + O(d^{12} N^{1/2-1/1200030} (\log N)^{100\varpi} L(1, \chi_{-N})),$$

where the implied constant depends at most on ϖ. Further,

$$M(N) = \frac{8}{\pi} L(1, \chi_{-N}) N^{1/2} \frac{2^{\omega(m_1 m_2 m_3)}}{(m_1 m_2 m_3)^2} \prod_{p \mid m_1 m_2 m_3} c_p \left(1 - \frac{1}{p^2}\right) \prod_{\substack{p \nmid m_1 m_2 m_3 \\ (\frac{N}{p})=1}} \left(1 - \frac{6 c_p}{p^2}\right)$$

and

$$\begin{aligned} g_N(d) =& \mathbb{1}(p \mid (d, N) \Rightarrow p \equiv 1 \pmod 4) \frac{\mathbb{1}(2 \nmid d)}{d} 2^{\sharp\{p \mid d : p \mid m_1 m_2 m_3 N\}} 3^{\sharp\{p \mid d : p \nmid m_1 m_2 m_3\}} \\ & \times \prod_{\substack{p \mid m_1 m_2 m_3 \\ p \mid d}} \left(1 + \frac{1}{p}\right)^{-1} \prod_{\substack{p \nmid m_1 m_2 m_3 \\ p \mid d, (\frac{N}{p})=1}} \frac{1}{\left(1 - \frac{6 c_p}{p^2}\right)\left(1 - \frac{4}{p c_p}\right)} \prod_{\substack{p \mid d \\ p \nmid m_1 m_2 m_3}} c_p \left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right). \end{aligned}$$

Proof. Let $\mathfrak{G}(b)$ be the number of $\mathbf{b} \in \mathbb{N}^3$ with $b = b_1 b_2 b_3$ and $\gcd(b_i, b_j m_j) = 1$ for all $i \neq j$. Using $2^{\sharp\{p \mid d : p \nmid bm\}} 2^{\sharp\{p \mid d, p \mid m, p \nmid b\}} 2^{\sharp\{p \mid (d,b)\}} = 2^{\omega(d)}$ we can write

$$M_2 = M_3 3^{\sharp\{p \mid d : p \nmid m\}} \frac{2^{\omega(d)}}{2^{\sharp\{p \mid d : p \nmid mN\}}} \prod_{p \mid d, p \nmid m} c_p \left(1 - \frac{1}{p}\left(\frac{-N}{p}\right)\right),$$

where M_3 is given by

$$\sum_{\substack{b \in \mathbb{N}, 2 \nmid b, \mathfrak{C}_d \mid bm \\ p \mid b, p \nmid m \Rightarrow (\frac{N}{p})=1}} 2^{\sharp\{p \mid b : p \nmid m\}} \frac{\mu(b) \gcd(d, b)}{b^2} \frac{2^{\sharp\{p \mid b : p \nmid mN, p \mid d\}}}{3^{\sharp\{p \mid b : p \nmid m, p \mid d\}}} \mathfrak{G}(b) \prod_{\substack{p \mid b, p \nmid m \\ p \mid d}} c_p^{-2} \prod_{p \mid b, p \nmid m} c_p.$$

For a prime p we have $\mathfrak{G}(p) = \mathbb{1}(p \nmid m_1 m_2) + \mathbb{1}(p \nmid m_1 m_3) + \mathbb{1}(p \nmid m_2 m_3)$. Since the m_i are coprime in pairs, $\mathfrak{G}(p)$ becomes 1 or 3 according to whether p divides m or not. Hence, $\mathfrak{G}(b) = 3^{\sharp\{p \mid b : p \nmid m\}}$ for all square-free b, thus, M_3 can be written as

$$\sum_{\substack{b \in \mathbb{N}, 2 \nmid b, \mathfrak{C}_d \mid bm \\ p \mid b, p \nmid m \Rightarrow (\frac{N}{p})=1}} 6^{\sharp\{p \mid b : p \nmid m\}} \frac{\mu(b) \gcd(d, b)}{b^2} \frac{2^{\sharp\{p \mid b : p \nmid mN, p \mid d\}}}{3^{\sharp\{p \mid b : p \nmid m, p \mid d\}}} \prod_{\substack{p \mid b, p \nmid m \\ p \mid d}} c_p^{-2} \prod_{p \mid b, p \nmid m} c_p.$$

Let us show that if the sum over b is non-empty then $\mathfrak{C}_d = 1$. To see that, assume there is a prime $p \mid \mathfrak{C}_d$. Then the condition $p \mid \mathfrak{C}_d \mid bm$ implies that $p \mid m$ or $p \nmid m$ and $p \mid b$. In the first case, the condition present in M_1 shows that $(\frac{N}{p}) = 1$, which violates the condition $p \mid \mathfrak{C}_d \mid N$. In the second case, we have $p \nmid m$ and $p \mid b$, hence, the condition in the sum over b shows that $(\frac{N}{p}) = 1$, which is a contradiction.

Now, factor the square-free b as $b_0 b_1$, where $b_0 \mid m$ and b_1 is coprime to m. We can thus write $M_3 = M_4 M_5$, where

$$M_4 = \sum_{b_0 \mid m} \frac{\mu(b_0) \gcd(d, b_0)}{b_0^2} = \prod_{\substack{p \mid m \\ p \mid d}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \mid m \\ p \nmid d}} \left(1 - \frac{1}{p^2}\right)$$

and M_5 is given by

$$\sum_{\substack{b_1 \in \mathbb{N}, \gcd(b_1, 2m)=1 \\ p \mid b_1 \Rightarrow (\frac{N}{p})=1}} 6^{\sharp\{p \mid b_1\}} \frac{\mu(b_1) \gcd(d, b_1)}{b_1^2} \left(\frac{2}{3}\right)^{\sharp\{p \mid b_1 : p \mid d\}} \prod_{p \mid b_1} c_p \prod_{\substack{p \mid b_1 \\ p \mid d}} c_p^{-2},$$

where we used the conditions $b_0 \mid m$ and $\gcd(b_1, m) = 1$ to infer that b_0, b_1 are coprime and thus split $\mu(b_0 b_1)$. The Euler product for M_5 equals

$$\prod_{\substack{p\nmid 2m\\ (\frac{N}{p})=1}} \left(1 - \frac{6c_p}{p^2}\right) \prod_{\substack{p\mid d, p\nmid 2m\\ (\frac{N}{p})=1}} \frac{1}{\left(1 - \frac{6c_p}{p^2}\right)\left(1 - \frac{4}{pc_p}\right)},$$

which concludes the proof. □

4.6. **The proof of Theorem 4.1.** We apply Theorem 2.3 with $\mathcal{A}$ being the set of vectors $\mathbf{y} \in (\mathbb{Z} \setminus \{0\})^3$ satisfying $\gcd(y_i, y_j) = 1$ for all $i \neq j$ and $c_a = |y_1 y_2 y_3|$. Further, we let $T = N$ and $\chi_N(a) = \mathbb{1}_{\{N\}}((m_2 m_3 y_1)^2 + (m_1 m_3 y_2)^2 + (m_1 m_2 y_3)^2)$. To verify assumption (2.6) we use Lemma 4.11. Note that $2^{\omega(m)} \prod_{p\mid m}(1 - 1/p) \geqslant 1$, hence

$$\frac{N^{1/2} L(1, \chi_{-N})}{(\log N)^{6\varpi}} \leqslant \frac{N^{1/2} L(1, \chi_{-N})}{(m_1 m_2 m_3)^2} \ll M \ll N^{1/2} L(1, \chi_{-N}) \qquad (4.8)$$

with absolute implied constants. Fix any strictly positive constants ξ and θ satisfying $12\theta + \xi < 1/600015$. For any positive integer $d \leqslant M^\theta$, the error term in Lemma 4.11 is

$$\ll M^{12\theta} N^{1/2} L(1, \chi_{-N}) N^{-1/1200030} (\log N)^{100\varpi} \ll M^{1+12\theta} N^{-1/1200030} (\log N)^{106\varpi}$$

by the lower bound (4.8). Using the upper bound of the same inequality we obtain

$$\ll M^{1-1/600015+12\theta} L(1, \chi_{-N})^{1/2400060} (\log N)^{106\varpi} \ll M^{1-1/600015+12\theta} (\log N)^{1+106\varpi}$$

by the bound $L(1, \chi_{-N}) \ll \log N$. The error term is $O(M^{1-\xi})$ as we have chosen ξ so that $12\theta + \xi < 1/600015$. This verifies assumption (2.6) of Theorem 2.3. The remaining assumptions are easily seen to hold since the function g_N in Lemma 4.11 satisfies $p^e g_N(p^e) = O(1)$ for all $e \geqslant 1$ and primes p with an absolute implied constant. Hence, one has for all $\mathbf{m}$ with $\max m_i \leqslant (\log N)^A$

$$\sum_{\substack{\mathbf{y} \in (\mathbb{Z}\setminus\{0\})^3 : \sum_i (m_j m_k y_i)^2 = N\\ \gcd(y_i, y_j) = 1 \forall i \neq j}} f(|y_1 y_2 y_3|) \ll M(N) T(M(N)),$$

where

$$T(y) = \prod_{1 \ll p \leqslant y} (1 - g_N(p)) \sum_{a \leqslant y} f(a) g_N(a)$$

and the implied constant depends at most on s and ϖ. We have $T(y) \asymp \exp(S(y))$ by Lemma 2.10, where $S(y)$ is

$$6 \sum_{\substack{p\mid N\\ p\equiv 1 \pmod 4}} \frac{f(p) - 1}{p} + 3 \sum_{\substack{p \leqslant y\\ p\nmid N}} \frac{f(p) - 1}{p} + O\Big(1 + \sum_{p \mid m_1 m_2 m_3} \frac{1}{p} + \sum_{\substack{p\mid N\\ p > y}} \frac{1}{p}\Big).$$

The main term is

$$3 \sum_{p\mid N} \left(\frac{-1}{p}\right) \frac{f(p) - 1}{p} + 3 \sum_{p \leqslant y} \frac{f(p) - 1}{p} + O\Big(\sum_{\substack{p\mid N\\ p > y}} \frac{1}{p}\Big).$$

The sum over $p \mid N, p > y$ is $\ll \omega(N)/y \ll (\log N)/y$. Thus, with $c_f(N)$ as in (4.1), there exists a positive constant $\nu = \nu(s)$ such that for all $y \geqslant \log N$ one has

$$\prod_{p \mid m_1 m_2 m_3} \left(1 + \frac{1}{p}\right)^{-\nu} \ll T(y) c_f(y)^{-3} \exp\left(-3 \sum_{p \leqslant y} \frac{f(p) - 1}{p}\right) \ll \prod_{p \mid m_1 m_2 m_3} \left(1 + \frac{1}{p}\right)^{\nu}.$$

Injecting the upper bound into Lemma 4.8 and using that $M(N) \leqslant N$ we get

$$\sum_{\substack{\mathbf{x}\in(\mathbb{Z}\setminus\{0\})^3\\ x_1^2+x_2^2+x_3^2=N}} \prod_{i=1}^3 f(|x_i|) \ll \frac{L(1,\chi_{-N})N^{1/2}}{(\log N)^{A/2-\rho(s)}} + L(1,\chi_{-N})N^{1/2}c_f(N)^3 \exp\left(3\sum_{p\leqslant N}\frac{f(p)-1}{p}\right)\mathcal{S}$$

where

$$\mathcal{S} = \sum_{\substack{\mathbf{m}\in\mathbb{N}^3,(4.7)\\ \max m_i\leqslant(\log N)^A}} \prod_{i=1}^3 \frac{\tau(m_i)^{2s}}{m_i^2}\prod_{p\mid m_i}\left(1+\frac{1}{p}\right)^\nu .$$

Since $\prod_{p\mid m}(1+1/p) \leqslant \tau(m)$ we can see that $\mathcal{S}$ is bounded. Enlarging the value of A allows the logarithmic exponent $A/2-\rho(s)$ to exceed any given number and it thus completes the proof of the upper bound in Theorem 4.1 when each α_i is 1.

To prove the lower bound in Theorem 4.1 we note that

$$\sum_{\substack{\mathbf{x}\in(\mathbb{Z}\setminus\{0\})^3\\ x_1^2+x_2^2+x_3^2=N}} \prod_{i=1}^3 f(|x_i|) \geqslant \sum_{\substack{\mathbf{x}\in(\mathbb{Z}\setminus\{0\})^3,\ \sum_i x_i^2=N\\ \gcd(x_i,x_j)=1\forall i\neq j}} \prod_{i=1}^3 f(|x_i|)$$

and apply Theorem 2.4 to estimate the right-hand side sum. This has a level-of-distribution assumption that can be verified using the case $m_1=m_2=m_3=1$ of Lemma 4.11. The residual stages in the proof are indistinguishable to those for the upper bound.

4.7. **General Diophantine equations.** The proof of Theorem 1.10 is based on an application of Theorem 2.3. This has specific assumptions; we start by verifying the ones related to the level of distribution in Lemma 4.13 and proceed by verifying the ones related to the growth of the sieve density function in Lemmas 4.14-4.15.

Lemma 4.12. *Let $F \in \mathbb{Z}[x_0,\dots,x_n]$ be as in Theorem 1.10. Then for every $0 \leqslant i \leqslant n$ we have $\sharp\{\mathbf{x}\in(\mathbb{Z}\cap[-B,B])^{n+1} : F(\mathbf{x})=0, x_i=0\} \ll B^{n-d}$, where the implied constant depends only on i and F.*

Proof. It is necessary to recall the definition of the Birch rank $\mathfrak{B}(g)$ of a polynomial $g\in\mathbb{Z}[x_0,x_1,\dots,x_m]$, where $m\in\mathbb{N}$. Denoting the homogeneous part of g by $g^\flat$, the number $\mathfrak{B}(g)$ is defined as the codimension of the affine variety in $\mathbb{C}^{m+1}$ given by $\nabla g^\flat(\mathbf{x})=0$. Note that $\mathfrak{B}(g)=m+1$ when $g^\flat$ is smooth. Returning to the proof of our theorem we note that setting $x_i=0$ in the polynomial $F(\mathbf{x})$ will produce a homogeneous polynomial F_i in at most $n-1$ variables. We claim that F_i will have degree d. If not, then F_i must vanish identically, which can only happen when each monomial of F_i contains x_i; hence, x_i would divide $F(\mathbf{x})$ and this would contradict the assumed smoothness of F.

By [48, Lemma 3.1] one has $\mathfrak{B}(F_i)\geqslant\mathfrak{B}(F)-2$, since in the notation of [48] one has $|\mathbf{j}|_1=1$ and $R=1$. Recalling that F is smooth one sees that $\mathfrak{B}(F)=n+1$. Hence,

$$\mathfrak{B}(F_i)\geqslant n-1\geqslant(d-1)2^{d-1}=(\deg(F_i)-1)2^{\deg(F_i)-1}.$$

Hence, F_i satisfies the assumption on the number of variables for Birch's work [7]. In particular, [7, Equation (4), page 260] applied to F_i gives

$$\sharp\{\mathbf{x}\in(\mathbb{Z}\cap[-B,B])^n : F(x_0,x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_n)=0\}\ll B^{n-\deg(F_i)}=B^{n-d},$$

which is sufficient. □

For a prime p define

$$\sigma_p := \lim_{m\to+\infty}\frac{\sharp\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1} : F(\mathbf{x})\equiv 0 \,(\mathrm{mod}\ p^m)\}}{p^{mn}}.$$

Combining [7, Equation (20), page 256] with [7, Lemma 7.1] we see that the limit converges and, furthermore,

$$\sigma_p = 1 + O(p^{-1-\lambda}) \tag{4.9}$$

for some positive $\lambda = \lambda(F)$. The corresponding density over $\mathbb{R}$ is given by

$$\sigma_\infty := \int_{-\infty}^{+\infty} \int_{\mathbf{x}\in[-1,1]^{n+1}} \exp(2\pi i \gamma F(\mathbf{x}))\mathrm{d}\mathbf{x}\mathrm{d}\gamma.$$

It converges due to [7, Lemma 5.2]. Finally, we let

$$\sigma(F) := \sigma_\infty \prod_p \sigma_p.$$

Lemma 4.13. *Keep the assumptions of Theorem 1.10 and assume that $F = 0$ has a $\mathbb{Q}$-point. Then there exist positive constants Ξ, β, both of which depend on F, such that for all $q \in \mathbb{N}$ and $B \geqslant 1$ with $q \leqslant B^\Xi$ one has*

$$\sharp\{\mathbf{x} \in (\mathbb{Z}\backslash\{0\})^{n+1} : \max|x_i| \leqslant B, F(\mathbf{x}) = 0, q \mid x_0 \cdots x_n\} = \sigma(F)B^{n+1-d}(h_F(q)+O(B^{-\beta})),$$

where the function $h_F : \mathbb{N} \to [0,\infty)$ is defined by

$$h_F(q) = \prod_{p\mid q} \frac{1}{\sigma_p} \lim_{m\to+\infty} \frac{\sharp\left\{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^{n+1} : p^m \mid F(\mathbf{x}), p^{v_p(q)} \mid x_0 \cdots x_n\right\}}{p^{mn}}$$

and the implied constant is independent of q.

Proof. Adding back the terms for which $x_0 \cdots x_n = 0$ shows that the counting function equals

$$\sharp\{\mathbf{x} \in (\mathbb{Z} \cap [-B,B])^{n+1} : F(\mathbf{x}) = 0, \mathbf{x} \equiv \mathbf{t} \,(\mathrm{mod}\; q)\} + O(\mathcal{E}),$$

where $\mathcal{E} = \sharp\{\mathbf{x} \in (\mathbb{Z} \cap [-B,B])^{n+1} : F(\mathbf{x}) = 0, x_0 \cdots x_n = 0\}$. By Lemma 4.12 we have $\mathcal{E} = O(B^{n-d})$, which is satisfactory. To deal with the main term we partition in progressions to convert it into

$$\sum_{\substack{\mathbf{t}\in(\mathbb{Z}/q\mathbb{Z})^{n+1}\\ t_0\cdots t_n\equiv 0(\mathrm{mod}\;q)}} \sharp\{\mathbf{x} \in (\mathbb{Z} \cap [-B,B])^{n+1} : F(\mathbf{x}) = 0, \mathbf{x} \equiv \mathbf{t} \,(\mathrm{mod}\; q)\}.$$

We now employ [22, Lemma 4.4] to deduce that the cardinality equals

$$\sigma_\infty B^{n+1-d} \prod_{p\mid q} \sigma_p(\mathbf{t}, p^{v_p(q)}) \prod_{p\nmid q} \sigma_p + O(B^{n+1-d-\eta}q^M),$$

where η, M are positive constants that depend only on F,

$$\sigma_p(\mathbf{t}, p^k) := \lim_{m\to+\infty} \frac{\sharp\left\{\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^{n+1} : F(\mathbf{x}) \equiv 0 \,(\mathrm{mod}\; p^m), \mathbf{x} \equiv \mathbf{t} \left(\mathrm{mod}\; p^k\right)\right\}}{p^{mn}}$$

and the implied constant depends at most on F. The contribution of the error term is

$$\ll B^{n+1-d-\eta}q^M \sum_{\substack{\mathbf{t}\in(\mathbb{Z}/q\mathbb{Z})^{n+1}\\ t_0\cdots t_n\equiv 0(\mathrm{mod}\;q)}} 1 \ll B^{n+1-d-\eta}q^{M+n+1}.$$

Letting $\Xi := \frac{\eta}{2(M+n+1)}$, the assumption $q \leqslant B^\Xi$ implies that $q^{M+n+1} \leqslant B^{\eta/2}$, thus the error term is $O(B^{n+1-d-\eta/2})$, which is satisfactory. The main term contribution becomes

$$\sigma(F)B^{n+1-d} \sum_{\substack{\mathbf{t}\in(\mathbb{Z}/q\mathbb{Z})^{n+1}\\ t_0\cdots t_n\equiv 0(\mathrm{mod}\;q)}} \prod_{p\mid q} \frac{\sigma_p(\mathbf{t}, p^{v_p(q)})}{\sigma_p},$$

where we have used the fact that $\sigma_p > 0$ for all primes p. This is guaranteed by [7, Lemma 7.1] and the assumption that $F = 0$ is non-singular and has a $\mathbb{Q}$-point. It is straightforward to see that the sum over $\mathbf{t}$ forms a multiplicative function of q by using the Chinese Remainder Theorem. Its value at a prime power p^k is

$$\frac{1}{\sigma_p}\lim_{m\to+\infty}\frac{1}{p^{mn}}\sum_{\substack{\mathbf{t}\in(\mathbb{Z}/p^k\mathbb{Z})^{n+1}\\ t_0\cdots t_n\equiv 0\left(\bmod\ p^k\right)}}\ \sum_{\substack{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1},\ p^k|\mathbf{x}-\mathbf{t}\\ F(\mathbf{x})\equiv 0(\bmod\ p^m)}}1$$

which can be seen to coincide with $h_F(p^k)$ by interchanging the order of summation. □

The next result will be used to study h_F at prime powers. Its proof is analogous to [48, Lemma 3.4].

Lemma 4.14. *Keep the assumptions of Theorem 1.10. Fix any $i \neq j \in \mathbb{Z}\cap[0,n]$ and $\alpha,\beta\in\mathbb{Z}\cap[0,\infty)$. Then there exists $\mu_0>0$ that depends only on d such that*

$$\lim_{m\to+\infty}\frac{\sharp\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1}: p^m\mid F(\mathbf{x}), p^\alpha\mid x_i, p^\beta\mid x_j\right\}}{p^{mn}}=\frac{1}{p^{\alpha+\beta}}(1+O(p^{-1-\mu_0})),$$

where the implied constant depends at most on i,j and F.

Proof. Let $m\geqslant\max\{\alpha,\beta\}+1$. Then

$$\sharp\left\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1}: p^m\mid F(\mathbf{x}), p^\alpha\mid x_i, p^\beta\mid x_j\right\}=\sum_{\substack{(x_1,x_2)\in(\mathbb{Z}/p^m\mathbb{Z})^2\\ p^\alpha|x_1,p^\beta|x_2}}N(x_1,x_2), \qquad (4.10)$$

where $N(x_1,x_2)=\sharp\{\mathbf{y}\in(\mathbb{Z}/p^m\mathbb{Z})^{n-1}:F_{x_1,x_2}(\mathbf{y})\equiv 0\ (\bmod\ p^m)\}$ and

$$F_{x_1,x_2}(\mathbf{y}):=F(y_0,\dots,y_{i-1},x_1,y_{i+1},\dots,y_{j-1},x_2,y_{j+1},\dots,y_n).$$

The expression

$$\frac{1}{p^m}\sum_{a\in\mathbb{Z}/p^m\mathbb{Z}}\exp\left(2\pi i\frac{a}{p^m}F_{x_1,x_2}(\mathbf{y})\right)$$

is 1 or 0 according to whether $F_{x_1,x_2}(\mathbf{y})$ is divisible by p^m or not. Writing $a=p^{m-t}b$ for some $b\in(\mathbb{Z}/p^t\mathbb{Z})^*$ the expression becomes

$$\frac{1}{p^m}\sum_{t=0}^m\sum_{b\in(\mathbb{Z}/p^t\mathbb{Z})^*}\exp\left(2\pi i\frac{b}{p^t}F_{x_1,x_2}(\mathbf{y})\right)=\frac{1}{p^m}+\frac{1}{p^m}\sum_{t=1}^m\sum_{b\in(\mathbb{Z}/p^t\mathbb{Z})^*}\exp\left(2\pi i\frac{b}{p^t}F_{x_1,x_2}(\mathbf{y})\right).$$

Hence,

$$N(x_1,x_2)=p^{m(n-2)}+\frac{1}{p^m}\sum_{t=1}^m\sum_{b\in(\mathbb{Z}/p^t\mathbb{Z})^*}\sum_{\mathbf{y}\in(\mathbb{Z}/p^m\mathbb{Z})^{n-1}}\exp\left(2\pi i\frac{b}{p^t}F_{x_1,x_2}(\mathbf{y})\right).$$

Replacing $\mathbf{y}$ by its value (mod p^t) does not affect the exponential, thus,

$$N(x_1,x_2)=p^{m(n-2)}+p^{m(n-2)}\sum_{t=1}^m p^{-t(n-1)}\sum_{b\in(\mathbb{Z}/p^t\mathbb{Z})^*}\sum_{\mathbf{z}\in(\mathbb{Z}/p^t\mathbb{Z})^{n-1}}\exp\left(2\pi i\frac{b}{p^t}F_{x_1,x_2}(\mathbf{z})\right).$$

We can view $F_{x_1,x_2}(\mathbf{y})$ as a polynomial in $\mathbf{y}$ since x_1,x_2 are fixed. It will be non-homogeneous and its degree d part is homogeneous and equals $F_{0,0}(\mathbf{y})$. By [48, Lemma 3.1] one has $\mathfrak{B}(F_{x_1,x_2})\geqslant\mathfrak{B}(F)-4=n-3$, since one must take $|\mathbf{j}|_1=2$ and $R=1$ in [48]. Our assumptions ensure that $n-3\geqslant 1+(d-1)2^d$, thus, [7, Lemma 5.4] shows that, for every fixed positive ϵ, the sum over $\mathbf{z}$ is $\ll p^{t(n-1-\mu+\epsilon)}$, where $\mu=K(F_{x_1,x_2})/(d-1)$ and $K(g)$ is defined in [7, Equation (8), page 252] as $\mathfrak{B}(g^\flat)2^{-d+1}$. Note that [7, Lemma 5.4] is unaffected by the lower order terms coming from x_1, x_2, since the proof is based on a Weyl differencing process in [7, Lemma 2.1]. We obtain

$$N(x_1,x_2) = p^{m(n-2)} + O\Big(p^{m(n-2)}\sum_{t=1}^{m} p^{t(1-\mu+\epsilon)}\Big).$$

We aim to show that $1-\mu<-1$. Using $\mathfrak{B}(F_{x_1,x_2}) \geqslant n-3$ we obtain

$$\mu = \frac{\mathfrak{B}(F_{x_1,x_2})2^{-d+1}}{d-1} \geqslant \frac{(n-3)2^{-d+1}}{d-1} \geqslant 2 + \frac{2^{-d+1}}{d-1},$$

where we used our assumption $n \geqslant 4+(d-1)2^d$ in the last step. Let $\epsilon = 2^{-d}/(d-1)$ and $\mu_0 = -2+\mu-\epsilon$, so that $1-\mu+\epsilon = -1-\mu_0 < -1$ and

$$\sum_{t=1}^{m} p^{t(1-\mu+\epsilon)} \ll p^{1-\mu+\epsilon} = p^{-1-\mu_0}.$$

Therefore, $N(x_1,x_2) = p^{m(n-2)}(1+O(p^{-1-\mu_0}))$, which can be injected into (4.10) to obtain

$$\sharp\Big\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1} : p^m \mid F(\mathbf{x}), p^\alpha \mid x_i, p^\beta \mid x_j\Big\} = p^{mn-\alpha-\beta}(1+O(p^{-1-\mu_0}))$$

since the right-hand of (4.10) has $p^{2m-\alpha-\beta}$ terms. Dividing by p^{mn} concludes the proof. □

Lemma 4.15. *Keep the assumptions of Theorem 1.10, let p be a prime and assume that $F=0$ has a $\mathbb{Q}_p$-point. There exist positive constants δ_F, δ'_F that depend only on F such that for all $e \geqslant 1$ we have*

$$h_F(p) = \frac{n+1}{p}\left(1+O(p^{-1-\delta'_F})\right) \quad \text{and} \quad h_F(p^e) \leqslant p^{\delta_F - e/(n+1)},$$

where the implied constant depends at most on F.

Proof. We have

$$\begin{aligned}&\frac{\sharp\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1} : p^m \mid F(\mathbf{x}), p \mid x_0\cdots x_n\}}{p^{mn}}\\ &= \sum_{i=0}^{n}\frac{\sharp\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1} : p^m \mid F(\mathbf{x}), p \mid x_i\}}{p^{mn}} + O\left(\sum_{1\leqslant i<j\leqslant n}\mathcal{E}_{i,j}\right),\end{aligned}$$

where

$$\mathcal{E}_{i,j} = \frac{\sharp\{\mathbf{x}\in(\mathbb{Z}/p^m\mathbb{Z})^{n+1} : p^m \mid F(\mathbf{x}), p \mid x_i, p \mid x_j\}}{p^{mn}}.$$

By Lemma 4.14 with $\alpha=\beta=p$ we see that $\mathcal{E}_{i,j} \ll 1/p^2$. To estimate the sum over $0 \leqslant i \leqslant n$ in the main term we use Lemma 4.14 with $\alpha = p, \beta = 0$ to obtain

$$\sum_{i=0}^{n}\frac{(1+O(p^{-3/2}))}{p} + O(n^2p^{-2}) = \frac{n+1}{p} + O(p^{-1-\mu_0}).$$

Our assumptions ensure that $\sigma_p > 0$ and recalling (4.9) proves the claimed estimate on $h_F(p)$.

To bound $h_F(p^e)$ note that if $m > e \geqslant n$ and $\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^n$ is such that $x_0\cdots x_n \equiv 0 \,(\text{mod } p^e)$ then there exists $0 \leqslant i \leqslant n$ such that $v_p(x_i) \geqslant e/(n+1)$. By Lemma 4.14 with $\beta = 0$ and α given by the least integer satisfying $\alpha \geqslant e/(n+1)$ we infer that $h_F(p^e) \leqslant Cp^{-e/(n+1)}$ for some positive constant $C = C(F)$ by taking $m \to \infty$ in the definition of $h_F(p^e)$. If $e \in [1, n)$ then we use the trivial bound $h_F(p^e) \leqslant p^{n+1}$. Thus, in all cases we have shown that

$$h_F(p^e) \leqslant \frac{\max\{C, p^{n+1}\}}{p^{\frac{e}{n+1}}}.$$

Letting $\gamma = (\log C)/(\log 2)$, we infer that $C \leqslant 2^\gamma \leqslant p^\gamma$, hence,

$$h_F(p^e) \leqslant p^{\max\{\gamma, n+1\} - e/(n+1)},$$

thus, concluding the proof. □

To prove Theorem 1.10 we can assume with no loss of generality that $F = 0$ has a $\mathbb{Q}$-point so that the set $\mathfrak{C} := \{|x_0 \cdots x_n| : \mathbf{x} \in (\mathbb{Z} \setminus \{0\})^{n+1}, F(\mathbf{x}) = 0\}$ is non-empty. Let $\mathcal{A} = \{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^{n+1} : F(\mathbf{x}) = 0\}$ and for every $a = \mathbf{x} \in \mathcal{A}$ define $c_a = |x_0 \cdots x_n|$. Define $\chi_B : \mathcal{A} \to [0, \infty)$ by

$$\chi_B(\mathbf{x}) := \mathbb{1}_{[0,B]}(\max\{|x_i| : 0 \leqslant i \leqslant n\}).$$

By Lemma 4.13 there exist $\Xi, \beta > 0$ such that for all $q \leqslant B^\Xi$ one has

$$\sum_{\substack{a \in \mathcal{A} \\ q | c_a}} \chi_B(a) = \sigma(F) B^{n+1-d}(h_F(q) + O(B^{-\beta})).$$

This shows that the property in Definition 2.2 is fulfilled for $M = \sigma(F)B^{n+1-d}$. In particular,

$$\sup\{c_a : \chi_B(a) > 0\} \leqslant \sup\{|x_0 \cdots x_n| : \max|x_i| \leqslant B\} \leqslant B^{n+1} \ll M^{(n+1)/(n+1-d)},$$

thus, (2.7) holds with $\alpha = \frac{n+1}{n+1-d}$. Note that $h_F \in \mathcal{D}(n+1, 1/(n+1), \delta_F, B', K)$ for some positive constants B', K that depend only on F due to $h_F(p^e) \leqslant h_F(p)$ and Lemma 4.15. We can thus employ Theorem 2.3 to deduce that

$$\sum_{\substack{\mathbf{x} \in (\mathbb{Z} \setminus \{0\})^{n+1} \\ \max|x_i| \leqslant B, F(\mathbf{x}) = 0}} f(|x_0 \cdots x_n|) \ll B^{n+1-d} \prod_{p \leqslant M} (1 - h_F(p)) \sum_{a \leqslant M} f(a) h_F(a),$$

where $M = \sigma(F)B^{n+1-d}$. By Lemma 4.15 we can bound the product over $p \leqslant M$ by $\ll \exp(-(n+1)\sum_{p \leqslant M} 1/p)$. Combining this with the succeding lemma completes the proof of Theorem 1.10.

Lemma 4.16. *Keep the setting of Theorem 1.10. For every fixed constant $\gamma > 0$ and each $B \geqslant 1$ we have*

$$\sum_{a \leqslant B^\gamma} f(a) h_F(a) \ll \exp\left((n+1) \sum_{p \leqslant B} \frac{f(p)}{p}\right),$$

where the implied constant depends at most on f, A, γ and F.

Proof. We will apply Lemma 2.7 with $G = f$. It is clear that f satisfies the required assumptions. We next verify the required assumptions for h_F: the bound (2.9) holds for $h_F(p^e)$ due to Lemma 4.15 and the fact that $h_F(p^e) \leqslant h_F(p)$. It remains to prove the estimate $h_F(p^e) \ll 1/p^2$ for all $e \geqslant 2$. Since $h_F(p^e) \leqslant h_F(p^2)$ it suffices to bound $h_F(p^2)$. To do so we note that if $\mathbf{x} \in (\mathbb{Z}/p^m\mathbb{Z})^{n+1}$ is such that $F(\mathbf{x}) \equiv 0 \pmod{p^m}$ and $p^2 \mid x_0 \cdots x_n$, then either there exists i such that $p^2 \mid x_i$ or there are $i \neq j$ such that $p \mid x_i$ and $p \mid x_j$. In the first case we may employ Lemma 4.14 with $\alpha = 2$ and $\beta = 0$ and in the second case with $\alpha = 1$ and $\beta = 1$ in order to obtain the bound $h_F(p^2) \ll p^{-2}$. Now Lemma 2.7 implies that

$$\sum_{a \leqslant B^\gamma} f(a)h_F(a) \ll \exp\left(\sum_{p \leqslant B^\gamma} f(p)h_F(p)\right),$$

where the implied constant only depends on f and F. Using the estimate for $h_F(p)$ from Lemma 4.15 and noting that $f(p) \leqslant A$ and $\sum p^{-2} \leqslant \infty$, we can rewrite the sum over $p \leqslant B^\gamma$ as

$$\sum_{p \leqslant B^\gamma} f(p)h_F(p) = (n+1) \sum_{p \leqslant B^\gamma} \frac{f(p)}{p} + O\left((n+1)A\right).$$

Therefore

$$\sum_{a \leqslant B^\gamma} f(a)h_F(a) \ll \exp\left((n+1) \sum_{p \leqslant B^\gamma} \frac{f(p)}{p}\right),$$

where the implied constant depends at most on f, A, n and h_F. To conclude the proof we note that for all $0 < \gamma_1 < \gamma_2$ one has

$$\sum_{B^{\gamma_1} < p \leqslant B^{\gamma_2}} \frac{f(p)}{p} \leqslant A \sum_{B^{\gamma_1} < p \leqslant B^{\gamma_2}} \frac{1}{p} = A \log \frac{\gamma_2}{\gamma_1} + O(1/\log B) \ll 1,$$

thus, one can replace the condition $p \leqslant B^\gamma$ by $p \leqslant B$ at the cost of a different implied constant. □

5. Polynomial values

In this section we give upper and lower bounds for sums of the form (1.4).

5.1. **Proving equidistribution.**

Here we prove the necessary results that will be fed into Theorems 2.3-2.4 to yield Theorems 1.15-1.16.

If p_Q denotes the largest prime dividing all coefficients of Q then for $p \leqslant p_Q$ we have $\varrho_Q(p^e) \leqslant 1 \leqslant p_Q/p$ for all $e \geqslant 1$. For all other primes we use [51, Lemma 2.7] to obtain

$$\varrho_Q(p^e) \leqslant \frac{p_Q + \deg(Q)}{p}. \tag{5.1}$$

Furthermore, Lemma 2.8 gives

$$\varrho_Q(p^e) \leqslant \frac{C_0}{p^{e/\deg(Q)}} \tag{5.2}$$

for a positive constant C_0 that only depends on Q. Finally, Lemma 2.5 shows

$$\prod_{p \leqslant x} (1 - \varrho_Q(q)) = c(\log x)^{-r}(1 + O(1/\log x)), \tag{5.3}$$

where r is the number of irreducible components of Q and c is a positive number depending only on Q.

The following definition is based on [54, Definition 2.2]. Write $|\cdot|$ for the usual Euclidean norm on $\mathbb{R}^k$ for $k \in \mathbb{N}$.

Definition 5.1 (Regions)**.** Let $n, M \geqslant 1$ be integers and let $L > 0$ be a real number. We say that a subset S of $\mathbb{R}^n$ is in $\mathrm{Reg}(n, M, L)$ if

(1) S is bounded,

(2) there exist M maps $\phi_1, \ldots, \phi_M : [0,1]^{n-1} \to \mathbb{R}^n$ satisfying

$$|\phi_i(\mathbf{x}) - \phi_i(\mathbf{y})| \leqslant L|\mathbf{x} - \mathbf{y}|$$

for $\mathbf{x}, \mathbf{y} \in [0,1]^{n-1}$ and $i = 1, \ldots, M$ such that the images of ϕ_i cover ∂S.

Lemma 5.2 (Widmer, [54, Theorem 2.4]). *Each $S \in \mathrm{Reg}(n, M, L)$ is measurable and satisfies*

$$|\sharp(S \cap \mathbb{Z}^n) - \mathrm{vol}(S)| \leqslant n^{\frac{3n^2}{2}} M \max(L^{n-1}, 1).$$

Lemma 5.3. *Let $\mathcal{D} \in \mathrm{Reg}(n, M, L)$. Then for all $\mathbf{w} \in [-1, 1]^n$ and $t \geqslant 1$ one has*

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (t\mathcal{D} + \mathbf{w})\} = \mathrm{vol}(\mathcal{D})t^n + O(t^{n-1}),$$

where the implied constant depends on n, M, L, but is independent of t and $\mathbf{w}$.

Proof. This follows from Lemma 5.2. □

Lemma 5.4. *Let $\mathcal{B} \in \mathrm{Reg}(n, M, L)$ and let $Q \in \mathbb{Z}[x_1, \dots, x_n]$ be non-zero. Then for all $\mathbf{c} \in \mathbb{R}^n, q \in \mathbb{N}, R \geqslant 1$ with $q \leqslant R$ we have*

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : Q(\mathbf{x}) \neq 0, q \mid Q(\mathbf{x})\} = \varrho_Q(q)\mathrm{vol}(\mathcal{B})R^n(1 + O(q/R)) + O(R^{n-1}),$$

where the implied constant depends on n, $\mathcal{B}$ and Q, but is independent of R and q.

Proof. We may assume that Q is non-constant. We will first show that

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : Q(\mathbf{x}) = 0\} \ll R^{n-1},$$

where the implied constant depends at most on n, $\mathcal{B}$ and Q. We will proceed by induction on n. The case $n = 1$ is easy, so now suppose that $n > 1$. Observe that we may certainly assume that the variable x_n occurs in Q. Expand Q as

$$Q(x_1, \dots, x_n) = \sum_{i=0}^{m} Q_i(x_1, \dots, x_{n-1})x_n^i$$

with $m \geqslant 1$ and $Q_m(x_1, \dots, x_{n-1})$ non-zero. By the induction hypothesis we may bound the number of zeros of $Q_m(x_1, \dots, x_{n-1})$. If $Q_m(x_1, \dots, x_{n-1}) \neq 0$, then there are at most m possibilities, say $\alpha_1, \dots, \alpha_m$, for x_n. We may find B such that $\mathcal{B}$ is contained in $[-B, B]^n$. Then we can employ Lemma 5.3 in dimension $n - 1$ to see that the number of integer vectors in $\mathbf{x} \in \mathbb{Z}^n \cap (R[-B, B]^n + \mathbf{c})$ for which $x_n = \alpha_j$ is $\ll R^{n-1}$, where the implied constant depends at most on n and $\mathcal{B}$. We can thus add back the missing terms with $Q(\mathbf{x}) = 0$ at the cost of a negligible error term depending only on n, $\mathcal{B}$ and Q.

Note that for every $c \in \mathbb{R}$ and $y' \in \mathbb{Z}/q\mathbb{Z}$ there exists a unique $y \in \mathbb{Z} \cap [c, c + q)$ such that $y \equiv y' \pmod q$. Using this for every y_i and c_i allows us to to split in progressions in order to obtain

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : q \mid Q(\mathbf{x})\} = \sum_{\substack{\mathbf{y} \in \mathbb{Z}^n \cap \mathcal{I} \\ q \mid Q(\mathbf{y})}} \sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : \mathbf{x} \equiv \mathbf{y} \pmod q\}, \quad (5.4)$$

where $\mathcal{I}$ is the product of intervals $\prod_{i=1}^n [c_i, c_i + q)$. Letting $\mathbf{x} = \mathbf{y} + q\mathbf{z}$ we infer

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : \mathbf{x} \equiv \mathbf{y} \pmod q\} = \sharp\left\{\mathbf{z} \in \mathbb{Z}^n \cap \left(\frac{R}{q}\mathcal{B} + \frac{\mathbf{c} - \mathbf{y}}{q}\right)\right\}.$$

Since $y_i \in [c_i, c_i + q)$ and $R/q \geqslant 1$ we can use Lemma 5.3 with $\mathbf{w} = (\mathbf{c} - \mathbf{y})/q$ to obtain

$$\sharp\left\{\mathbf{z} \in \mathbb{Z}^n \cap \left(\frac{R}{q}\mathcal{B} + \frac{\mathbf{c} - \mathbf{y}}{q}\right)\right\} = \mathrm{vol}(\mathcal{B})\frac{R^n}{q^n} + O\left(\frac{R^{n-1}}{q^{n-1}}\right),$$

where the implied constant depends at most on n and $\mathcal{B}$. Injecting this into (5.4) gives

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap (R\mathcal{B} + \mathbf{c}) : q \mid Q(\mathbf{x})\} = \varrho_Q(q)(\mathrm{vol}(\mathcal{B})R^n + O(qR^{n-1})),$$

which completes the proof. □

5.2. **Upper bounds.** We have $[-1,1]^n \in \mathrm{Reg}(n, 2^n, 2)$. Thus, by Lemma 5.4 with $\mathcal{B} = [-1,1]^n$ and $\mathbf{c} = \mathbf{0}$ we have

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap T[-1,1]^n : Q(\mathbf{x}) \neq 0, q \mid Q(\mathbf{x})\} = \varrho_Q(q) 2^n T^n \Big(1 + O\Big(q/T\Big)\Big) + O(T^{n-1}),$$

whenever $q \leqslant T$, where the implied constant depends at most on n, Q and $\mathcal{B}$. In particular, if $q \leqslant T^{1/2}$ then the right-hand side becomes

$$\varrho_Q(q) 2^n T^n (1 + O(1/\sqrt{T})) + O(T^{n-1/2}).$$

Thus, we can use Theorem 2.3 with

$$\mathcal{A} = \{\mathbf{x} \in \mathbb{Z}^n : Q(\mathbf{x}) \neq 0\}, \mathfrak{C} = \{|Q(\mathbf{x})| : \mathbf{x} \in \mathcal{A}\}, \chi_T(\mathbf{x}) = \mathbb{1}_{[0,T]}(\max |x_i|),$$

since it shows that the assumption of Definition 2.2 holds with

$$h = \varrho_Q, M(T) = 2^n T^n, \theta = \frac{1}{4n}, \xi = \frac{1}{2n}.$$

Note that assumption (2.1) is satisfied with $\kappa = r$ due to (5.3), while assumptions (2.2)-(2.3) hold respectively due to (5.1)-(5.2). Hence for all $T \geqslant 1$ we have

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap T[-1,1]^n \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|) \ll T^n \prod_{p \leqslant T^n} (1 - \varrho_Q(p)) \sum_{a \leqslant T^n} f(a) \varrho_Q(a),$$

where the implied constant depends at most on A, Q and n. By (5.3) the product over $p \leqslant T^n$ is asymptotic to $(\log T)^{-r}$.

To complete the proof of Theorem 1.15 let us note that since $f \geqslant 0$ and $\mathcal{D}$ is contained in $X(\mathcal{D})[-1,1]^n$ we can write

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{D} \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|) \leqslant \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap X(\mathcal{D})[-1,1]^n \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|).$$

As we have seen, this is

$$\ll X(\mathcal{D})^n (\log 2X(\mathcal{D}))^{-r} \sum_{a \leqslant X(\mathcal{D})^n} f(a) \varrho_Q(a),$$

which is sufficient.

5.3. **Lower bounds.** By assumption $\mathcal{D}$ contains a set of the form $\mathbf{c} + X\mathcal{U}$, where $\mathcal{U}$ is the unit ball in $\mathbb{R}^n$. Since $f \geqslant 0$ we deduce that

$$\sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap \mathcal{D} \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|) \geqslant \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap (\mathbf{c} + X\mathcal{U}) \\ Q(\mathbf{x}) \neq 0}} f(|Q(\mathbf{x})|).$$

We may thus employ Theorem 2.4 with

$$\mathcal{A} = \{\mathbf{x} \in \mathbb{Z}^n : Q(\mathbf{x}) \neq 0\}, \mathfrak{C} = \{|Q(\mathbf{x})| : \mathbf{x} \in \mathcal{A}\}, \chi_t(\mathbf{x}) = \mathbb{1}_{\mathbf{c}+X\mathcal{U}}(\mathbf{x}).$$

Note that $\mathcal{U}$ can be parametrised by a single function in $n-1$ variables by using trigonometric functions, thus, we can use Lemma 5.4 to obtain

$$\sharp\{\mathbf{x} \in \mathbb{Z}^n \cap \mathbf{c} + X\mathcal{U} : Q(\mathbf{x}) \neq 0, q \mid Q(\mathbf{x})\} = \mathrm{vol}(\mathcal{U}) \varrho_Q(q) X^n + O(\deg(Q)^{\omega(q)} X^{n-1}),$$

whenever $q \leqslant X$, where the implied constant depends at most on n and Q. The proof can now be completed by following the same steps as in the final stage of the proof of Theorem 1.16.

5.4. **Proof of Theorem 1.11.** When $\mathcal{D} = [-X, X]^n$, the parameter $X(\mathcal{D})$ in Definition 1.13 satisfies $X(\mathcal{D}) \leqslant 2X$. Thus, Theorem 1.15 yields that

$$\sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap[-X,X]^n \\ Q(\mathbf{x})\neq 0}} \tau_k(|Q(\mathbf{x})|)^\ell \ll X^n(\log X)^{-1} \sum_{a\leqslant X^n} \tau_k(a)^\ell \varrho_Q(a).$$

A lower bound can be given by employing Theorem 1.16 with $\mathcal{D} = [-X, X]^n$. This is allowed since the sphere in $\mathbb{R}^n$ with radius X and centre at the origin is contained in $\mathcal{D}$. To conclude the proof we use Lemma 2.9 with $G = \tau_k^\ell$. To see that G satisfies the required assumptions we recall that for primes p and all $e \geqslant 1$ one has

$$\tau_k(p^e) = \frac{(e+k-1)(e+k-2)\cdots(e+1)}{(k-1)!}$$

so that $\tau_k(p)^\ell = k^\ell$ and $\tau_k(p^e)^\ell \leqslant e^{O_{k,\ell}(1)} \ll_{k,\ell} (3/2)^e$. Hence, Lemma 2.9 can be employed with $r = 1, \lambda = k^\ell$ and $C = 3/2$; it yields

$$\sum_{a\leqslant X} \tau_k(a)^\ell \varrho_Q(a) \asymp (\log X)^{k^\ell}.$$

References

[1] P. T. Bateman, On the Representations of a Number as the Sum of Three Squares. *Trans. Amer. Math. Soc.* **71**(2) (1951), 70–101.

[2] K. Belabas, Crible et 3-rang des corps quadratiques. *Ann. Inst. Fourier (Grenoble)*, **46** (1996), 909–949.

[3] M. Bhargava, The density of discriminants of quintic rings and fields. *Ann. of Math.* **172** (2010), 1559–1591.

[4] M. Bhargava, A. Shankar, J. Tsimerman, On the Davenport-Heilbronn theorems and second order terms. *Invent. Math.* **193** (2013), 439–499.

[5] M. Bhargava, A. Shankar, X. Wang, Geometry-of-numbers methods over global fields I: Prehomogeneous vector spaces. `arXiv:1512.03035`.

[6] M. Bhargava, T. Taniguchi, F. Thorne, Improved error estimates for the Davenport–Heilbronn theorems. `arXiv:2107.12819`.

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

[8] V. Blomer, J. Brüdern, A three squares theorem with almost primes. *Bull. Lond. Math. Soc.* **37** (2005), 507–513.

[9] Z. Brady, Divisor function inequalities, entropy, and the chance of being below average. *Math. Proc. Cambridge Philos. Soc.* **163** (2017), 547–560.

[10] S Chan, P. Koymans, C. Pagano, E. Sofos, Averages of multiplicative functions along equidistributed sequences. *Submitted* (2024).

[11] R. de la Bretèche, T. D. Browning, Sums of arithmetic functions over values of binary forms. *Acta Arith.* **125** (2006), 291–304.

[12] ———, Le problème des diviseurs pour des formes binaires de degré 4. *J. reine angew. Math.* **646** (2010), 1–44.

[13] R. de la Bretèche, G. Tenenbaum, Sur la conjecture de Manin pour certaines surfaces de Châtelet. *J. Inst. Math. Jussieu* **12** (2013), 759–819.

[14] T. D. Browning, E. Sofos, Counting rational points on quartic del Pezzo surfaces with a rational conic. *Math. Ann.* **373** (2019), 977–1016.

[15] V. A. Bykovskiĭ, Spectral expansions of certain automorphic functions and their number-theoretic applications. Automorphic functions and number theory, II, **134**, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 15–33, Springer, Berlin, 1984.

[16] S. Chowla, P. Erdős, A theorem on the distribution of the values of L-functions. *J. Indian Math. Soc.* **15** (1951), 11–18.

[17] H. Cohen, H. W. Lenstra, Heuristics on class groups of number fields. Lecture Notes in Math., **1068**, *Number theory, Noordwijkerhout*, 33–62, Springer, Berlin, 1984.

[18] B. Cook, Á. Magyar, Diophantine equations in the primes. *Invent. Math* **198** (2014), 701–737.

[19] H. Davenport, H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London Ser. A*, **322** (1971), 405–420.

[20] F. Delmer, Sur la somme de diviseurs $\sum_{k\leqslant x}\{d[f(k)]\}^s$. *C. R. Acad. Sci. Paris Sér. A-B* **272** (1971), A849–A852.

[21] W. Duke, On ternary quadratic forms. *J. of Number Th.* **110** (2005), 37–43.

[22] D. El-Baz, D. Loughran, E. Sofos, Multivariate normal distribution for integral points on varieties. *Trans. Amer. Math. Soc.* **375** (2022), 3089–3128.

[23] J. Ellenberg, L. B. Pierce, M. M. Wood, On ℓ-torsion in class groups of number fields. *Algebra Number Theory,* **11** (2017), 1739–1778.

[24] C. Elsholtz, T. Tao, Counting the number of solutions to the Erdős-Straus equation on unit fractions. *J. Aust. Math. Soc.* **94** (2013), 50–105.

[25] P. Erdős, A. Wintner, Additive arithmetical functions and statistical independence. *Amer. J. Math.*, **61** (1939), 713–721.

[26] P. Erdős, On the sum $\sum_{k=1}^{x} d(f(k))$. *J. London Math. Soc.* **27** (1952), 7–15.

[27] E. Fouvry, J. Klüners, On the 4-rank of class groups of quadratic number fields. *Invent. Math.*, **167** (2007), 455–513.

[28] E. Fouvry, J. Klüners, Cohen–Lenstra heuristics of quadratic number fields. *Algorithmic number theory,* 40–55. Lecture Notes in Comput. Sci., 4076, *Springer-Verlag, Berlin*, 2006.

[29] E. Fouvry, J. Klüners, Weighted distribution of the 4-rank of class groups and applications. *Int. Math. Res. Not. IMRN* **16** (2011), 3618–3656.

[30] C. Frei, M. Widmer, Averages and higher moments for the ℓ-torsion in class groups. *Math. Ann.* **379** (2021), 1205–1229.

[31] J. Friedlander, H. Iwaniec, Opera de cribro. *American Mathematical Society Colloquium Publications*, **57** American Mathematical Society, Providence, RI, xii+527, 2010.

[32] C. F. Gauss, *Disquisitiones Arithmeticae*, Translated into English by Arthur A. Clarke, S. J, Yale University Press, New Haven, Conn.-London, 1801.

[33] A. Granville, K. Soundararajan, The distribution of values of $L(1,\chi_d)$. *Geom. Funct. Anal.* **13** (2003), 992–1028.

[34] D. R. Heath-Brown, L. B. Pierce, Averages and moments associated to class numbers of imaginary quadratic fields. *Compos. Math.* **153** (2017), 2287–2309.

[35] K. Henriot, Nair-Tenenbaum bounds uniform with respect to the discriminant. *Math. Proc. Cambridge Philos. Soc.* **152** (2012), 405–424.

[36] R. Holowinsky, Sieving for mass equidistribution. *Ann. of Math. (2)* **172** (2010), 1499–1516.

[37] C. Hooley, On the number of divisors of a quadratic polynomial. *Acta Math.* **110** (1963), 97–114.

[38] E. Jones, Local densities of diagonal integral ternary quadratic forms at odd primes. *J. of Number Th.* **17** (2021), 547–575.

[39] I. Khayutin, Joint equidistribution of CM points. *Ann. of Math. (2)* **189** (2019), 145–276.

[40] P. Koymans and J. Thorner, Bounds for moments of ℓ-torsion in class groups. `arXiv:2309.07204`.

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

[42] R. J. Lemke Oliver, F. Thorne, The number of ramified primes in number fields of small degree. *Proc. Amer. Math. Soc.* **145** (2017), 3201–3210.

[43] J. V. Linnik, The dispersion method in binary additive problems. *American Mathematical Society, Providence, RI,* **53** Springer, Berlin-New York, x+186pp, 1963.

[44] G. Malle, On the distribution of Galois groups. II. *Experiment. Math.* **13** (2004), 129–135.

[45] K. J. McGown, F. Thorne, A. Tucker, Counting quintic fields with genus number one. `arXiv:2006.12991`.

[46] M. Nair, G. Tenenbaum, Short sums of certain arithmetic functions. *Acta Math.* **180** (1998), 119–144.

[47] L. B. Pierce, D. Schindler, M. Wood, Representations of integers by systems of three quadratic forms. *Proc. Lond. Math. Soc. (3)* **113** (2016), 289–344.

[48] D. Schindler, E. Sofos, Sarnak's saturation problem for complete intersections. *Mathematika* **65** (2019), 1–56.

[49] A. Shankar, J. Tsimerman, Counting S_5-fields with a power saving error term. *Forum Math. Sigma* **2** (2014), Paper No. e13, 8.

[50] A. Smith, 2^∞-Selmer Groups, 2^∞-class groups, and Goldfeld's conjecture. `arXiv:1702.02325`.

[51] E. Sofos, Y. Wang, Finite saturation for unirational varieties. *Int. Math. Res. Not. IMRN* **71** (2019), 4784–4821.

[52] K. Soundararajan, Divisibility of class numbers of imaginary quadratic fields. *J. London Math. Soc. (2)* **61** (2000), 681–690.

[53] T. Taniguchi, F. Thorne, Orbital L-functions for the space of binary cubic forms. *Canad. J. Math.* **65** (2013), 1320–1383.

[54] M. Widmer, Lipschitz class, narrow class, and counting lattice points. *Proc. Amer. Math. Soc.* **140** (2012), no. 2, 677–689.

[55] E. Wirsing, Das asymptotische Verhalten von Summen über multiplikative Funktionen. *Math. Ann.* **143** (1961), 75–102.

[56] S. Yamagishi, Diophantine equations in primes: density of prime points on affine hypersurfaces. *Duke Math. J.* **171** (2022), 831–884.

[57] G.-L. Zhou, Y. Ding, Sums of the higher divisor function of diagonal homogeneous forms. *Ramanujan J.* **59** (2022), 933–945.

ISTA, Am Campus 1, 3400 Klosterneuburg, Austria
Email address: stephanie.chan@ist.ac.at

Institute for Theoretical Studies, ETH Zürich, 8092, Switzerland
Email address: peter.koymans@eth-its.ethz.ch

Department of Mathematics, Concordia University, Montreal, H3G 1M8, Canada
Email address: carlo.pagano@concordia.ca

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