Mathematika 65 (2019) 1–56
doi:10.1112/S002557931800030X

SARNAK'S SATURATION PROBLEM FOR COMPLETE INTERSECTIONS

D. SCHINDLER AND E. SOFOS

Abstract. We study almost prime solutions of systems of Diophantine equations in the Birch setting. Previous work shows that there exist integer solutions of size B with each component having no prime divisors below $B^{1/u}$, where u equals $c_0 n^{3/2}$, n is the number of variables and c_0 is a constant depending on the degree and the number of equations. We improve the polynomial growth $n^{3/2}$ to the logarithmic $(\log n)(\log\log n)^{-1}$. Our main new ingredients are the generalization of the Brüdern–Fouvry vector sieve in any dimension and the incorporation of smooth weights into the Davenport–Birch version of the circle method.

Contents

§1. *Introduction.* Let $f_1, \dots, f_R \in \mathbb{Z}[x_1, \dots, x_n]$ be forms of degree d and write $\mathbf{f} = (f_1, \dots, f_R)$. We consider the affine variety defined by

$$V_{\mathbf{f}} : f_i(x_1, \dots, x_n) = 0, \quad 1 \leqslant i \leqslant R. \tag{1.1}$$

We are interested in Sarnak's saturation problem, that is to find a Zariski-dense set of integer zeros $(x_1, \dots, x_n) \in V_{\mathbf{f}}(\mathbb{Z})$ where each x_i is either a prime or has a small number of prime divisors. Recent work of Cook and Magyar [11] is concerned with finding prime solutions to the Diophantine system $\mathbf{f}(\mathbf{x}) = \mathbf{s}$ for $\mathbf{s} \in \mathbb{Z}^R$, i.e. solutions for which every variable x_i is a prime number. They succeed in establishing a local to global principle, including an asymptotic formula, via the circle method if the Birch rank $\mathfrak{B}(\mathbf{f})$, that will be defined at the beginning of §2, satisfies $\mathfrak{B}(\mathbf{f}) \geqslant \chi(R, d)$ for some function $\chi(R, d)$ which only depends on the degree d and the number of polynomials R. However, the value of $\chi(R, d)$, as it would result from the current proof in [11], is expected to

Received 8 March 2018, published online 24 August 2018.
MSC (2010): 11D72 (primary), 11P32 (secondary).

be tower exponential in d and R. For systems of quadratic forms one has

$$\chi(R,2) \leqslant 2^{2^{CR^2}}.$$

For more general systems we do not have any explicit upper bounds on this function.

It is therefore natural to ask whether one can find an explicit condition which ensures the existence of a Zariski dense set of integer solutions with all coordinates being almost prime; this is usually referred to as Sarnak's saturation problem. Let $\Omega(m)$ denote the number of prime factors of m counted with multiplicity. Almost primes have zero density in the integers owing to the generalized prime number theorem: for each fixed integer $k \geqslant 1$ we have

$$\frac{1}{x}\sharp\{m \in \mathbb{N} \cap [1,x] : \Omega(m) \leqslant k\} \sim \frac{(\log\log x)^{k-1}}{(k-1)!\log x}, \quad \text{as } x \to \infty.$$

The fact that one seeks solutions in thin subsets of integers places problems of this type in a higher level of difficulty than studying the number of all integer solutions in expanding regions. Yamagishi [31] showed the existence of infinitely many integer solutions in the case $R = 1$ and for large n, with every solution having exactly 2 prime factors. This corresponds to taking $k = 2$ in the asymptotic above.

In this paper we are interested in a harder question than that of almost primes, namely in finding solutions within sets that have asymptotically zero density compared to the set of almost primes. Let $P^-(m)$ denote the least prime divisor of a positive integer $m \neq 1$ and define $P^-(1) := 1$. Integers m with $P^-(m) \geqslant m^{1/u}$ for some $u > 1$ are almost primes, however their density is *arbitrarily smaller* in comparison. Indeed, by Buchstab's theorem [9] one has the following for all fixed $k \in \mathbb{N}_{\geqslant 2}$ and $u \in \mathbb{R}_{>1}$:

$$\frac{\sharp\{m \in \mathbb{N} \cap [1,x] : P^-(m) \geqslant x^{1/u}\}}{\sharp\{m \in \mathbb{N} \cap [1,x] : \Omega(m) \leqslant k\}} \sim \frac{(k-1)!uw(u)}{(\log\log x)^{k-1}} \ll_{k,u} \frac{1}{(\log\log x)^{k-1}}, \quad \text{as } x \to \infty,$$

where $w(u)$ is the Buchstab function. Progress on the saturation problem within this thinner set of solutions was recently made by Magyar and Titichetrakun [27]. They managed to treat systems of equations where the number of variables is the same as in Birch's work [2], i.e. assuming that the Birch rank exceeds $R(R+1)(d-1)2^{d-1}$. They proved lower bounds of the correct order of magnitude regarding the number of integer solutions with each coordinate x_i satisfying $P^-(|x_i|) \geqslant |x_i|^{1/u}$, where u is any constant in the range

$$u \geqslant 2^8 n^{3/2} d(d+1)R^2(R+1)(R+2). \tag{1.2}$$

The ultimate goal of showing that all variables x_i can simultaneously be prime corresponds to the value $u > 2 - \epsilon$ for some $\epsilon > 0$, hence any result decreasing the admissible value for u in (1.2) is an equivalent reformulation of progress towards this goal. Our aim in this paper is to decrease the admissible value for u when the degree and the number of equations is fixed so as to have at most logarithmic growth in terms of n rather than polynomial.

1.1. *Summary of our results.* In order to prove quantitative or qualitative results for the system of equations (1.1) one typically needs n to be sufficiently large in terms of d and R and the singular locus of $V_{\mathbf{f}}$. Thus, for example, the Hasse principle is known for non-singular cubic hypersurfaces when $n \geqslant 9$ (Hooley [20]), for non-singular quartics when $n \geqslant 40$ (Hanselmann [17]) and for non-singular quintics in at least $n \geqslant 101$ variables (Browning and Prendiville [6]). One may expect that the dependence of (1.2) on n should decrease when n increases; we are not able to provide a bound that is independent of n but we shall provide a bound that depends logarithmically on n rather than polynomially. For this we shall use the vector sieve of Brüdern and Fouvry to show that for fixed d, R one can improve (1.2) to

$$u \gg \frac{\log n}{\log \log n}.$$

This constitutes a major improvement over (1.2) and it applies to almost all situations in the Birch setting, see Theorem 1.1. This is the main result in this paper.

As an additional result we shall provide an improvement in all situations in the Birch setting, however, this will not be of logarithmic nature. Namely, using the Rosser–Iwaniec sieve we shall prove that one can take $u \gg n$ in all of the remaining cases, see Theorem 1.5, while, in some situations covered by Theorem 1.5 but not by Theorem 1.1 we shall show via the weighted sieve that there are many integer zeros $(x_1, \ldots, x_n)$ where the total number of prime factors of $|x_1 \cdots x_n|$ is $\ll n \log n$, while at the same time every prime factor of each $|x_i|$ is at least $|x_i|^\alpha$ for some $0 < \alpha < 1$ independent of $\mathbf{x}$, see Theorem 1.6.

1.2. *The vector sieve in arbitrary dimension.* The *vector sieve* was brought into light by Brüdern and Fouvry [7] to show that for all sufficiently large positive integers N satisfying $N \equiv 4 \pmod{24}$ the Lagrange equation

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

has many solutions $\mathbf{x} \in \mathbb{N}^4$ with each x_i being indivisible by any prime of size at most $N^{1/u}$ with $u \geqslant 68.86$. Problems of Waring–Goldbach type become less hard the more variables are available and the expectation is that one can take each x_i to be a prime for N as above—this is still open while the case of representations by 5 squares of primes was settled by Hua [21]. The vector sieve was later used to make improvements on the admissible value for u in Lagrange's equation by Heath-Brown and Tolev [19], Tolev [30] and Cai [10], as well as in other sieving problems [3, 8, 18].

The main idea of the vector sieve is to use a combinatorial inequality that replaces the usual lower bound sieve by a linear combination of products of sieving functions each of dimension 1, one of the advantages being an improvement over the admissible value for u. There are other applications of the vector sieve in the literature but to our knowledge it has not been applied for sieves of arbitrarily large sieve dimension (the reader is referred to the book of Friedlander and Iwaniec [16] for the terminology).

Let us now proceed to the statement of our main theorem. Denoting the p-adic units by $\mathbb{Z}_p^\times$ we will always make the assumption that

$$\mathbf{f} = \mathbf{0} \text{ has non-singular solutions in } (0,1)^n \text{ and in } (\mathbb{Z}_p^\times)^n \text{ for every prime } p. \tag{1.3}$$

We shall define the quantity $K = K(\mathbf{f})$ in (2.1) using the notion of the Birch rank $\mathfrak{B}(\mathbf{f})$. Let

$$\Upsilon := \frac{d\mathfrak{B}(\mathbf{f})}{(d-1)2^{d-1}}\left(d - \frac{1}{R}\right) + R, \tag{1.4}$$

as well as

$$\theta' := \min\left\{\frac{1}{\rho}, \frac{\epsilon_{1,1} - dR}{\epsilon_{1,2} + \epsilon_{1,3}}, \frac{\epsilon_{2,1} - dR}{\epsilon_{2,2} + \epsilon_{2,3}}, \frac{\epsilon_{3,1} - dR}{\epsilon_{3,2} + \epsilon_{3,3}}\right\}, \tag{1.5}$$

where

$$\rho := 4R(R+1)d\left(1 + \frac{d}{2R(d-1)+1} + \frac{3Rd}{3R(d-1)+1}\right) \tag{1.6}$$

and the vectors $(\epsilon_{i,1}, \epsilon_{i,2}, \epsilon_{i,3})$ are defined as the columns of the following matrix:

$$\boldsymbol{\epsilon} := \begin{bmatrix} Rd + 1/2 & K & \dfrac{(K - R(R+1)(d-1))}{4R(R+1)d} + Rd \\ Rd + 1/2 & K & \dfrac{d(K - R^2(d-1))}{2R(d-1)} + R + K + \dfrac{2dK}{d-1} - Rd \\ 0 & 0 & \max\left\{0, \dfrac{K - R(R+1)(d-1)}{4R(R+1)d} - R - K + Rd\right\} \end{bmatrix}. \tag{1.7}$$

Here follows the main result of our paper.

THEOREM 1.1. *There exists a positive absolute constant c_0 such that whenever the forms $f_1, \ldots, f_R \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree $d \geqslant 2$ satisfy (1.3) and*

$$\mathfrak{B}(\mathbf{f}) > \max\{2^{d-1}(d-1)R(R+1), 2^{d-1}(d-1)R^2 + (R+1)(\Upsilon+1), 2^{d-1}(d^2-1)R^2\},$$

then we have for all large enough $B \geqslant 1$,

$$\sharp\{\mathbf{x} \in ((0, B] \cap \mathbb{N})^n : \mathbf{f}(\mathbf{x}) = \mathbf{0}, P^-(x_1 \cdots x_n) > B^{\theta'((\log\log n)/(c_0 \log n))}\} \gg \frac{B^{n-Rd}}{(\log B)^n},$$

where the constant θ' satisfies $\theta' \gg_{d,R} 1$.

This provides a lower bound $\log P^-(x_1 \cdots x_n)/\log B$ in terms of n that vanishes logarithmically slow as $n \to +\infty$, which constitutes a large improvement over the previously best known result that gave a polynomial decay [27]. The proof of Theorem 1.1 will be given in §6. A crucial input for the sieving arguments will be a general version of Birch's theorem that we shall prove in §2, see Theorem 2.1. Note that a similar result for one quadratic form is proved in the work of Browning and Loughran [5, Theorem 4.1], whereas our result aims at general complete intersections. More importantly, Theorem 2.1 allows situations where congruence conditions are imposed to every integer coordinate with a different moduli for every coordinate, while in their result one is only allowed to consider the same moduli for every coordinate. This extra feature will be of central importance for the vector sieve.

An inspection of the argument at the end of §6 shows that we can take $c_0 = 3$ in Theorem 1.1 when the number of variables n is sufficiently large. For $s \in \mathbb{R}_{>2}$ let $0 < f(s) \leqslant 1 \leqslant F(s)$ be the sieve functions associated with the linear Rosser–Iwaniec sieve, defined for example in [24], which satisfy $F(s), f(s) = 1 + O(s^{-s})$. One can improve the lower bound for $\log P^-(x_1 \cdots x_n)/\log B$ given by Theorem 1.1 by replacing the term $c_0 \log n/\log\log n$ by any value $s > 2$ that satisfies

$$F(s)^n < \left(1 + \frac{1}{n-1}\right) f(s).$$

A special case of Theorem 1.1 is the case of non-singular hypersurfaces.

COROLLARY 1.2. *There exists a positive absolute constant c_1 such that whenever f is an integer non-singular form of degree $d \geqslant 5$ in more than $2^{d-1}(d^2-1)$ variables that fulfils (1.3) then the following estimate holds for all large enough $B \geqslant 1$:*

$$\sharp\{\mathbf{x} \in ((0, B] \cap \mathbb{N})^n : f(\mathbf{x}) = 0, P^-(x_1 \cdots x_n) > B^{(c_1 \log\log n)/(d \log n)}\} \gg \frac{B^{n-d}}{(\log B)^n}.$$

Our results require a few more variables than in the Birch setting, which for non-singular hypersurfaces requires $n > 2^d(d-1)$. The reason for this is rooted to the way that the vector sieve works: in introducing n linear sieving functions in place of a single n-dimensional lower bound sieve the technique requires that we have a good control on the independency of the events that a large prime p divides several coordinates of an integer zero, this is related to the function δ that will be studied in §3. The Birch assumption

$$\mathfrak{B}(\mathbf{f}) > 2^{d-1}(d-1)R(R+1)$$

does not always allow a good bound for δ, however a slightly stronger geometric assumption will be shown to be sufficient via a version of Weyl's inequality that is uniform in the coefficients of the underlying polynomials. It must be noted that the work of Yamagishi [31, Theorem 1.3] only applies to smooth hypersurfaces in $n > 8^d(4d-2)$ variables, which ought to be compared with the assumption $n > 2^{d-1}(d^2-1)$ of Corollary 1.2.

1.3. *Applications to the saturation problem.* One further advantage of Theorem 2.1 is that it allows the use of any smooth weight with compact support. We can therefore establish a version of Theorem 1.1 where one counts solutions near an arbitrary non-singular point in $V_{\mathbf{f}}(\mathbb{R})$. This allows us to settle Sarnak's problem for the complete intersections under consideration. To phrase our result we first need the following definition. Each $x \in \mathbb{P}^{n-1}(\mathbb{Q})$ can be written uniquely up to sign in the form $x = [\pm\mathbf{x}]$, where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ and $\gcd(x_1, \ldots, x_n) = 1$. We can then define the function $\mathscr{L} : \mathbb{P}^{n-1}(\mathbb{Q}) \to \mathbb{R}_{\geqslant 0}$ through

$$\mathscr{L}(x) := \max_{\substack{1\leqslant i\leqslant n\\ x_i\neq 0}} \max\left\{\frac{\log|x_i|}{\log p} : p \text{ is a prime dividing } x_i\right\}.$$

Thus $\mathscr{L}(x) \leqslant u$ holds for some $x = [\pm(x_1, \ldots, x_n)] \in \mathbb{P}^{n-1}(\mathbb{Q})$ and $u \in \mathbb{R}_{\geqslant 0}$ if and only if

$$x_i \neq 0 \Rightarrow P^-(|x_i|) \geqslant |x_i|^{1/u}.$$

Definition 1.3 (Level of saturation). Assume that $X \subset \mathbb{P}^{n-1}$ is a variety defined over $\mathbb{Q}$. The level of saturation of X is the infimum of all real non-negative numbers u such that

$$\{x \in X(\mathbb{Q}) : \mathscr{L}(x) \leqslant u\}$$

is Zariski dense in X.

Note that in this definition the level of saturation is allowed to be infinite, for example if $X(\mathbb{Q})$ is not Zariski dense. Recalling the definition of the number of prime divisors $\Omega_{\mathbb{P}^{n-1}(\mathbb{Q})}(x)$ of a rational point $x \in \mathbb{P}^{n-1}(\mathbb{Q})$ in the paragraph before [29, Definition 1.1], we observe that if $\prod_i x_i \neq 0$ then

$$\Omega_{\mathbb{P}^{n-1}(\mathbb{Q})}(x) \leqslant n\mathscr{L}(x).$$

Therefore, according to [29, Definition 1.1], if X has a finite level of saturation then it has a finite saturation number. Therefore one could perceive Definition 1.3 as a refinement of the standard notion of saturation.

THEOREM 1.4. *There exists a positive absolute constant c_0 such that whenever the forms $f_1, \ldots, f_R \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree $d \geqslant 2$ satisfy (1.3) and*

$$\mathfrak{B}(\mathbf{f}) > \max\{2^{d-1}(d-1)R(R+1), 2^{d-1}(d-1)R^2 + (R+1)(\Upsilon+1), 2^{d-1}(d^2-1)R^2\}$$

and the complete intersection in $\mathbb{P}^{n-1}$ that is defined through

$$V_{\mathbf{f}} : f_1 = f_2 = \cdots = f_R = 0$$

is geometrically irreducible then $V_{\mathbf{f}}$ has finite level of saturation. In addition, the level of saturation is at most

$$\frac{c_0 \log n}{\theta' \log\log n},$$

where the constant θ' satisfies $\theta' \gg_{d,R} 1$.

1.4. *Results via the Rosser–Iwaniec sieve.* We next provide an almost prime result that covers all situations in the Birch setting, thus completing the treatment of the cases not covered by Theorem 1.1. This will provide a lower bound for $\log P^-(x_1\cdots x_n)/\log B$ that is worse than the one in Theorem 1.1 but still better than (1.2); this is due to the strength of the level of distribution result implied by Theorem 2.1.

THEOREM 1.5. *For any forms $f_1,\ldots,f_R\in\mathbb{Z}[x_1,\ldots,x_n]$ of degree $d\geqslant 2$ satisfying (1.3) and $K>R(R+1)(d-1)$ we have for all large enough $B\geqslant 1$,*

$$\sharp\{\mathbf{x}\in((0,B]\cap\mathbb{N})^n:\mathbf{f}(\mathbf{x})=\mathbf{0},\ P^-(x_1\cdots x_n)>B^{\theta'/3.75n}\}\gg\frac{B^{n-Rd}}{(\log B)^n},$$

where θ' is given in (1.5) and satisfies $\theta'\gg_{d,R}1$.

1.5. *Results via the weighted sieve.* Theorem 1.5 supplies a polynomially fast convergence to zero for $\log P^-(x_1\cdots x_n)/\log B$ with respect to n. This is slightly undesired, thus we shall provide a complementary result that furnishes many integer zeros satisfying a bound of similar quality for $\log P^-(x_1\cdots x_n)/\log B$ with the additional desired property that $x_1\cdots x_n$ has few prime factors. This will be implemented via the weighted sieve. We choose to include this result here because along the proof we shall provide a potentially useful reformulation of the weighted sieve given in the book of Diamond and Halberstam [**15**]. This reformulation allows the incorporation of further weights and will be given in Theorem 4.4.

Define

$$u'':=(n-Rd)\max\left\{\frac{(2\epsilon_{i,2}-1)}{\epsilon_{i,1}-Rd}:1\leqslant i\leqslant 3\right\},\tag{1.8}$$

$$\widehat{u}:=\max\{u'',1/\theta',2(n-Rd)\rho\},\qquad \widehat{v}:=\frac{nc_n-1}{\theta'-1/\widehat{u}},\tag{1.9}$$

where c_n is a sequence that satisfies $\lim_{n\to+\infty}c_n=2.44\ldots$. We furthermore let

$$r_0:=\frac{n\widehat{u}}{n-Rd}-1+n\left(1+\frac{\widehat{u}}{\widehat{v}}c_n\right)\log\frac{\widehat{v}}{\widehat{u}}-n\left(1-\frac{\widehat{u}}{\widehat{v}}\right).\tag{1.10}$$

THEOREM 1.6. *For any forms $f_1,\ldots,f_R\in\mathbb{Z}[x_1,\ldots,x_n]$ of degree $d\geqslant 2$ satisfying (1.3) and $\mathfrak{B}(\mathbf{f})>\max\{(d-1)R(R+1)2^{d-1},(d^2-1)R2^{d-1},(d-1)R^22^{d-1}+2(R+1)\}$ we have for all $r_1>r_0$ and all large enough $B\geqslant 1$,*

$$\sharp\{\mathbf{x}\in((0,B]\cap\mathbb{N})^n:\mathbf{f}(\mathbf{x})=\mathbf{0},\ P^-(x_1\cdots x_n)>B^{1/\widehat{v}},\ \Omega(x_1\cdots x_n)\leqslant r_1\}\gg\frac{B^{n-Rd}}{(\log B)^n},$$

where $\widehat{v}=O_{d,R}(n)$ and $r_0=O_{d,R}(n\log n)$.

A simple consequence of Theorem 1.1 is that it provides many integer zeros $\mathbf{x}$ with

$$\Omega(x_1 \cdots x_n) \ll \frac{n \log n}{\log \log n},$$

which constitutes an asymptotic saving compared to the estimate

$$\Omega(x_1 \cdots x_n) \ll n \log n$$

supplied by Theorem 1.6. This is surely surprising to those familiar with the weighted sieve and its applications to higher dimensional sieve problems. The reason that the vector sieve gives a better saturation number here is the strong level of distribution supplied by Theorem 2.1, which is a result of using smooth weights. Indeed, Theorem 2.1 allows us to estimate asymptotically the number of integer solutions of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ subject to divisibility conditions of the form $k_i | x_i$ for $\mathbf{x}$ in a region having the shape $\mathbf{x} \in B[-1, 1]^n$ and vectors $\mathbf{k} \in \mathbb{N}^n$ of size $|\mathbf{k}| \leqslant B^{1/s}$, where $s > 1$ depends on d and R but not on n. Such a level of distribution is usually not available in other problems related to the weighted sieve.

Notation. We shall reserve the symbol $\nu(m)$ for the counting function of distinct prime factors of a positive integer m. For vectors $\mathbf{x} \in \mathbb{R}^n$, $n \in \mathbb{N}$, we shall reserve the symbols $|\mathbf{x}|$ and $|\mathbf{x}|_1$ for the supremum and the ℓ^1 norm respectively. For vectors $\mathbf{k}, \mathbf{x} \in \mathbb{N}^n$ we shall abbreviate the simultaneous conditions $k_i | x_i$ by $\mathbf{k}|\mathbf{x}$. Similarly we write $\mathbf{k} \leqslant \mathbf{x}$ or $\mathbf{k} < \mathbf{x}$ or $|\mathbf{k}| \leqslant \mathbf{x}$ for the simultaneous conditions $k_i \leqslant x_i$ (respectively $k_i < x_i$ and $|k_i| \leqslant x_i$) for $1 \leqslant i \leqslant n$. We shall furthermore find it convenient to introduce the notation

$$\widetilde{\mathbf{k}} := k_1 \cdots k_n,$$

as well as

$$\langle \mathbf{kx} \rangle = (k_1 x_1, \ldots, k_n x_n).$$

For $q \in \mathbb{N}$, $z \in \mathbb{C}$ we shall write

$$e_q(z) := \mathrm{e}^{2\pi \mathrm{i} z/q} \quad \text{and} \quad e(z) := \mathrm{e}^{2\pi \mathrm{i} z}.$$

The letter ϵ will refer to an arbitrarily small positive fixed constant and to ease the notation we shall not record the dependence of the implied constant in the $\ll$ and $O(\cdot)$ notation. The letter w will be reserved to denote certain weight functions that will be considered constant throughout our work, thus we shall not record the dependence of the implied constant in the $\ll$ and $O(\cdot)$ notation. Throughout our work the forms $\mathbf{f}$ are considered to be constant, thus each implied constant in the $\ll$ and $O(\cdot)$ notation will depend on the coefficients of $\mathbf{f}$, d, n, z_0 and W, where the constants z_0, W are functions of $\mathbf{f}$ whose meaning will become clear in due course. Any extra dependencies will be specified by the use of a subscript.

§2. *A version of Birch's theorem with lopsided boxes and smooth weights.* In our applications of sieve methods it will be important to be able to count integer zeros of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ such that each integer coordinate x_i is divisible by a fixed integer $k_i \leqslant |x_i|$. A change of variables makes clear that a version of Birch's theorem with lopsided boxes and with uniformity of the error term in the coefficients of the polynomials is sufficient. One can do this without smooth weights, however the resulting error terms will give a weak level of distribution for our sieve applications. We shall instead use smooth weights and as a result we shall later be able to take k_i much closer to the size of x_i.

We now proceed to describe the version of Birch's theorem that we shall need. Assume that we are given any finite collection of polynomials

$$g_i \in \mathbb{Z}[x_1, \dots, x_n], \quad 1 \leqslant i \leqslant R,$$

denote the homogeneous part of g_i by $g_i^\natural$ and assume that there exists $d \in \mathbb{N}_{\geqslant 2}$ such that

$$1 \leqslant i \leqslant R \Rightarrow \deg(g_i^\natural) = d.$$

The Birch rank, denoted by $\mathfrak{B}(\mathbf{g}^\natural)$, is defined as the codimension of the affine variety in $\mathbb{C}^n$ which is given by

$$\mathrm{rank}\left(\left(\frac{\partial g_i^\natural(\mathbf{x})}{\partial x_j}\right)_{1 \leqslant i \leqslant R, 1 \leqslant j \leqslant n}\right) < R.$$

We set

$$K := 2^{-(d-1)}\mathfrak{B}(\mathbf{g}^\natural). \tag{2.1}$$

Let us fix any smooth compactly supported weight function $w : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ with the property $\mathrm{supp}(w) \subset [-2, 2]$. For $\mathbf{P} = (P_1, \dots, P_n) \in (\mathbb{R}_{\geqslant 1})^n$ we denote

$$\widetilde{\mathbf{P}} := \prod_{i=1}^n P_i, \qquad P_{\max} := \max_{1 \leqslant i \leqslant n} P_i \quad \text{and} \quad P_{\min} := \min_{1 \leqslant i \leqslant n} P_i$$

and fix an element $\mathbf{z} \in [-1, 1]^n$. Our aim is to find an asymptotic formula for the counting function

$$N_w(\mathbf{P}) := \sum_{\substack{\mathbf{y} \in \mathbb{Z}^n \\ \mathbf{g}(\mathbf{y}) = \mathbf{0}}} \prod_{i=1}^n w\left(\frac{y_i}{P_i} - z_i\right).$$

Birch's influential work [2] treated the case where w is replaced by the characteristic function of a finite interval and

$$K > R(R+1)(d-1), \; P_{\min} = P_{\max}.$$

For our applications of sieving methods a result that is uniform in the size of each P_i as well as the coefficients of each g_i is required. For $h \in \mathbb{C}[x_1, \dots, x_n]$ we denote by $\|\mathbf{h}\|$ the maximum of the absolute values of its coefficients and for $h_1, \dots, h_R \in \mathbb{C}[x_1, \dots, x_n]$ we let

$$\|\mathbf{h}\| := \max\{\|h_i\| : 1 \leqslant i \leqslant R\}.$$

THEOREM 2.1. *Let g_i, w, $\mathbf{z}$, P_i be as above, assume that $K > R(R+1)(d-1)$ and*

$$\frac{P_{\max}}{P_{\min}} < \|\mathbf{g}\|^{-1/(2R(d-1)+1)} \|\mathbf{g}^\natural\|^{-3R/(3R(d-1)+1)} P_{\max}^{1/4R(R+1)d}. \tag{2.2}$$

Then one has for each $\epsilon > 0$,

$$\begin{aligned} N_w(\mathbf{P}) - \mathfrak{S} J_w \ll \widetilde{\mathbf{P}} (P_{\max}/P_{\min})^R P_{\max}^{-Rd-1/2} + \widetilde{\mathbf{P}}^{1+\epsilon} (P_{\max}/P_{\min})^K P_{\max}^{-K} \\ + \|\mathbf{g}^\natural\|^{2K/(d-1)-R} \|\mathbf{g}\|^{(K-R^2(d-1))/2R(d-1)} \\ \times \widetilde{\mathbf{P}}^{1+\epsilon} (P_{\max}/P_{\min})^{R+K} P_{\max}^{-Rd-(K-R(R+1)(d-1))/4R(R+1)d}, \end{aligned}$$

where the implied constant depends at most on $\epsilon > 0$. Here $\mathfrak{S}$ and J_ω are the usual circle method singular series and singular integral and are defined in (2.6) and (2.7) respectively.

Our sole aim in this section is to establish Theorem 2.2. All implied constants may depend on n, R, d but not on the coefficients of the polynomials $g_i(\mathbf{y})$, $1 \leqslant i \leqslant n$. We start by introducing the exponential sum

$$S_w(\boldsymbol{\alpha}) := \sum_{\mathbf{y} \in \mathbb{Z}^n} \prod_{i=1}^n w\left(\frac{y_i}{P_i} - z_i\right) e(\boldsymbol{\alpha} \cdot \mathbf{g}(\mathbf{y})),$$

where we use the vector notation $\boldsymbol{\alpha} \cdot \mathbf{g}(\mathbf{y}) = \sum_{i=1}^R \alpha_i g_i(\mathbf{y})$. By orthogonality we now have

$$N_w(\mathbf{P}) = \int_{[0,1]^R} S_w(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}.$$

We shall follow Birch's approach [2] to approximate $N_w(\mathbf{P})$. Our first step is to produce a Weyl type inequality for $S_w(\boldsymbol{\alpha})$. Recall that $g_i^\natural(\mathbf{y})$ are homogeneous polynomials of degree d, which can be written as

$$g_i^\natural(\mathbf{y}) = d! \sum_{1 \leqslant j_1, \ldots, j_d \leqslant n} g_{j_1,\ldots,j_d}^{(i)} y_{j_1} \ldots y_{j_d},$$

with symmetric coefficients $g_{j_1,\ldots,j_d}$ (i.e. such that $g_{j_1,\ldots,j_d} = g_{\sigma(j_1),\ldots,\sigma(j_d)}$ for a permutation σ of the indices). We associate its multilinear forms

$$\Phi_i(\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(d)}) = d! \sum_{1 \leqslant j_1, \ldots, j_d \leqslant n} g_{j_1,\ldots,j_d}^{(i)} y_{j_1}^{(1)} \ldots y_{j_d}^{(d)},$$

and set

$$\Phi(\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(d)}) := \sum_{i=1}^R \alpha_i \Phi_i(\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(d)}).$$

LEMMA 2.2. *With the notation above we have*

$$\frac{|S_w(\boldsymbol{\alpha})|^{2^{d-1}}}{\widetilde{\mathbf{P}}^{2^{d-1}}} \ll \widetilde{\mathbf{P}}^{-d} \sum_{-2\mathbf{P}<\mathbf{h}^{(1)}<2\mathbf{P}} \cdots \sum_{-2\mathbf{P}<\mathbf{h}^{(d-1)}<2\mathbf{P}} \times \prod_{i=1}^{n} \min\{P_i, \|\Phi(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \mathbf{e}^{(i)})\|^{-1}\}.$$

Proof. For $w(x)$ a weight function and $h \in \mathbb{R}$ we introduce the notation

$$w_h(x) = w(x+h)w(x).$$

Moreover, for $h_1, \ldots, h_m \in \mathbb{R}$, we iteratively define

$$w_{h_1,\ldots,h_m} = w_{h_1,\ldots,h_{m-1}}(x+h_m)w_{h_1,\ldots,h_{m-1}}(x).$$

The same Weyl differencing process as in the proof of Lemma 3.3 (in particular equation (3.5)) in [6] or in Lemma 2.1 in [2] leads to

$$\frac{|S_w(\boldsymbol{\alpha})|^{2^{d-1}}}{\widetilde{\mathbf{P}}^{2^{d-1}}} \ll \widetilde{\mathbf{P}}^{-d} \sum_{-2\mathbf{P}<\mathbf{h}^{(1)}<2\mathbf{P}} \cdots \sum_{-2\mathbf{P}<\mathbf{h}^{(d-1)}<2\mathbf{P}} |S_w(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \boldsymbol{\alpha})|,$$

where

$$\begin{aligned} S_w(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \boldsymbol{\alpha}) &= \sum_{\mathbf{y}\in\mathbb{Z}^n} \left\{ \prod_{i=1}^{n} w_{h_i^{(1)}/P_i,\ldots,h_i^{(d-1)}/P_i}\left(\frac{y_i}{P_i} - z_i\right)\right\} \\ &\quad \times e\left(\sum_{i=1}^{R} \alpha_i \Phi_i(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \mathbf{y}) + c(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)})\right), \end{aligned}$$

with integers $c(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)})$ independent of $\mathbf{y}$. Hence

$$\begin{aligned} |S_w(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \boldsymbol{\alpha})| = \Bigg| &\sum_{\mathbf{y}\in\mathbb{Z}^n} \left\{ \prod_{i=1}^{n} w_{h_i^{(1)}/P_i,\ldots,h_i^{(d-1)}/P_i}\left(\frac{y_i}{P_i} - z_i\right)\right\} \\ &\times e(\Phi(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \mathbf{y}))\Bigg|. \end{aligned}$$

The estimate

$$S_w(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \boldsymbol{\alpha}) \ll \prod_{i=1}^{n} \min\{P_i, \|\Phi(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \mathbf{e}^{(i)})\|^{-1}\}$$

can then be obtained via partial summation. □

We define the counting function

$$M(\boldsymbol{\alpha}, \mathbf{P}) := \sharp\{-2\mathbf{P} \leqslant \mathbf{h}^{(i)} \leqslant 2\mathbf{P}, 1 \leqslant i \leqslant d-1 : \|\Phi(\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(d-1)}, \mathbf{e}^{(j)})\| < P_j^{-1} \forall 1 \leqslant j \leqslant n\}.$$

As Lemma 3.2 is deduced from [12, Lemma 3.1] we obtain the following lemma.

LEMMA 2.3. *One has*

$$|S_w(\boldsymbol{\alpha})|^{2^{d-1}} \ll \widetilde{\mathbf{P}}^{2^{d-1}-d+1+\epsilon} M(\boldsymbol{\alpha}, \mathbf{P}).$$

Next we need a version of [13, Lemma 12.6] which is modified for lopsided boxes.

LEMMA 2.4. *Let $L_1, \ldots, L_n$ be symmetric linear forms given by $L_i = \gamma_{i1}u_1 + \cdots + \gamma_{in}u_n$ for $1 \leqslant i \leqslant n$, i.e. such that $\gamma_{ij} = \gamma_{ji}$ for $1 \leqslant i, j \leqslant n$. Let $a_1, \ldots, a_n > 1$ be real numbers. We denote by $N(Z)$ the number of integers solutions $u_1, \ldots, u_{2n}$ of the system of inequalities*

$$|u_i| < a_i Z, \quad 1 \leqslant i \leqslant n, \qquad |L_i - u_{n+i}| < a_i^{-1} Z, \quad 1 \leqslant i \leqslant n.$$

Then for $0 < Z_1 \leqslant Z_2 \leqslant 1$ we have

$$\frac{N(Z_2)}{N(Z_1)} \ll \left(\frac{Z_2}{Z_1}\right)^n.$$

Proof. Let Λ be the $2n$-dimensional lattice defined by

$$\begin{aligned} x_i &= a_i^{-1} u_i, \quad 1 \leqslant i \leqslant n \\ x_{n+i} &= a_i(\gamma_{i1}u_1 + \cdots + \gamma_{in}u_n + u_{n+i}), \quad 1 \leqslant i \leqslant n. \end{aligned}$$

As in [13, proof of Lemma 12.6] we note that the inequalities describing $N(Z)$ are equivalent to

$$|x_i| < Z, \quad 1 \leqslant i \leqslant 2n,$$

for a point $(x_1, \ldots, x_{2n})$ in the lattice Λ. We identify the lattice Λ with its matrix

$$\Lambda = \begin{pmatrix} a_1^{-1} & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & a_n^{-1} & 0 & \ldots & 0 \\ a_1\gamma_{11} & \ldots & a_1\gamma_{1n} & a_1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_n\gamma_{n1} & \ldots & a_n\gamma_{nn} & 0 & \ldots & a_n \end{pmatrix}$$

and we find that the adjoint lattice is given by

$$M = (\Lambda^t)^{-1} = \begin{pmatrix} a_1 & \dots & 0 & -a_1\gamma_{11} & \dots & -a_1\gamma_{n1} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & a_n & -a_n\gamma_{1n} & \dots & -a_n\gamma_{nn} \\ 0 & \dots & 0 & a_1^{-1} & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & a_n^{-1} \end{pmatrix}.$$

Since $\gamma_{ij} = \gamma_{ji}$ for all $1 \leqslant i, j \leqslant n$ the two lattices Λ and M can be transformed into one another by interchanging the order of $x_1, \dots, x_{2n}$ and $u_1, \dots, u_{2n}$ and changing signs at some variables. Hence they have the same successive minima. Now the proof of Lemma 12.6 in [13] applies to our situation and an identical argument concludes our proof. □

We now apply Lemma 2.4 to the counting function $M(\boldsymbol{\alpha}, \mathbf{P})$. Let $0 < \theta < 1$ and set $Z = P_{\max}^{\theta-1}$. We then obtain the following bound:

$$|S_w(\boldsymbol{\alpha})|^{2^{d-1}} \ll \frac{\widetilde{\mathbf{P}}^{2^{d-1}-d+1+\epsilon}}{Z^{(d-1)n}} \sharp \mathscr{I},$$

where $\mathscr{I}$ is defined by

$$\{(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \in \mathbb{Z}^{(d-1)n} : |\mathbf{x}^{(i)}| \leqslant Z\mathbf{P}, \|\Phi(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}, \mathbf{e}_j)\| \\ < Z^{d-1} P_j^{-1}, \forall 1 \leqslant j \leqslant n\}.$$

We are now in a position to obtain a form of Weyl's inequality for $S_w(\boldsymbol{\alpha})$ (for a Weyl's inequality in a similar setting see for example [2, Lemma 4.3]). Let V^* be the affine variety defined by

$$\operatorname{rank}\left(\frac{\partial g_i^\natural(\mathbf{x})}{\partial x_j}\right)_{1\leqslant i\leqslant R, 1\leqslant j\leqslant n} < R,$$

and recall that

$$K = \frac{n - \dim V^*}{2^{d-1}}.$$

LEMMA 2.5. *Assume that* $0 < \theta < 1$. *Then one has either*

(i)

$$S_w(\boldsymbol{\alpha}) \ll \widetilde{\mathbf{P}}^{1+\epsilon} \left(\frac{P_{\max}}{P_{\min}}\right)^K P_{\max}^{-\theta K},$$

or

(ii) *there are integers* $1 \leqslant q \leqslant \|\mathbf{g}^\natural\|^R P_{\max}^{R(d-1)\theta}$, *and* $0 \leqslant a_1, \dots, a_R < q$ *with* $\gcd(\mathbf{a}, q) = 1$ *and*

$$|q\alpha_i - a_i| \leqslant \|\mathbf{g}^\natural\|^{R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+R(d-1)\theta}, \quad 1 \leqslant i \leqslant R.$$

Proof. First assume that $P_{\max}^{\theta-1}P_{\min} \geqslant 1$. We start with the bound

$$|S_w(\boldsymbol{\alpha})|^{2^{d-1}} \ll \widetilde{\mathbf{P}}^{2^{d-1}-d+1+\epsilon} P_{\max}^{(1-\theta)(d-1)n} \sharp \mathscr{I}.$$

Consider the affine variety $\mathscr{Y} \subset \mathbb{A}^{n(d-1)}$ given by

$$\mathscr{Y} : \operatorname{rank}(\Phi_i(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}, \mathbf{e}_j))_{1\leqslant i\leqslant R, 1\leqslant j\leqslant n} < R.$$

We set

$$\mathscr{E} := \{(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \in \mathbb{Z}^{n(d-1)} \cap \mathscr{Y} : |\mathbf{x}^{(i)}| \leqslant P_{\max}^{\theta-1}\mathbf{P},\ \forall 1 \leqslant i \leqslant d-1\}.$$

Now we distinguish two cases.

(i) Assume that $\mathscr{I} \subset \mathscr{E}$. Then we bound the cardinality of $\mathscr{E}$ by dimension bounds. We dissect the region given by the conditions that $|\mathbf{x}^{(i)}| \leqslant P_{\max}^{\theta-1}\mathbf{P}$ into boxes where all the side lengths are equal (at the boundaries we allow for overlapping boxes which will result in slight overcounting) and of size $P_{\max}^{\theta-1}P_{\min}$. The number of such boxes is bounded by

$$\ll \left(\prod_{i=1}^n \frac{P_i}{P_{\min}}\right)^{d-1}.$$

On each of the boxes we apply a linear transformation to move the box to the origin. Then we apply [4, Theorem 3.1]. Note that this bound is independent of the coefficients of the variety (only depending on the dimension and degree) and hence uniform in the shift. We obtain

$$\sharp\mathscr{E} \ll \left(\prod_{i=1}^n \frac{P_i}{P_{\min}}\right)^{d-1} (P_{\max}^{\theta-1}P_{\min})^{\dim \mathscr{Y}}.$$

By [2, Lemma 3.3] we have $\dim \mathscr{Y} \leqslant \dim V^* + (d-2)n$, hence we obtain the bound

$$\sharp\mathscr{E} \ll \left(\prod_{i=1}^n \frac{P_i}{P_{\min}}\right)^{d-1} (P_{\max}^{\theta-1}P_{\min})^{\dim V^*+(d-2)n}.$$

Together with our assumption $\mathscr{I} \subset \mathscr{E}$ we obtain

$$\begin{aligned}|S_w(\boldsymbol{\alpha})|^{2^{d-1}} &\ll \widetilde{\mathbf{P}}^{2^{d-1}+\epsilon} P_{\max}^{(\theta-1)(\dim V^*+(d-2)n)} P_{\max}^{-(d-1)n(\theta-1)} P_{\min}^{-n+\dim V^*} \\ &\ll \widetilde{\mathbf{P}}^{2^{d-1}+\epsilon} P_{\max}^{(1-\theta)(n-\dim V^*)} P_{\min}^{-n+\dim V^*} \\ &\ll \widetilde{\mathbf{P}}^{2^{d-1}+\epsilon} P_{\max}^{-\theta(n-\dim V^*)} \left(\frac{P_{\max}}{P_{\min}}\right)^{n-\dim V^*}.\end{aligned}$$

This estimate gives option (i) in the statement of our lemma.

Next we assume that $\mathscr{I} \setminus \mathscr{E} \neq \emptyset$. Let $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)})$ be such a point in the difference set, i.e.

$$\operatorname{rank}(\Phi_i(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}, \mathbf{e}_j))_{1\leqslant i\leqslant R, 1\leqslant j\leqslant n} = R.$$

With no loss of generality we assume that the leading $R \times R$ minor is of full rank, and set

$$q := |\det(\Phi_i(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d-1)}, \mathbf{e}_j))_{1 \leqslant i, j \leqslant R}|.$$

Note that

$$q \ll \|\mathbf{g}^\natural\|^R P_{\max}^{R(d-1)\theta}.$$

Moreover, we have the system of equations

$$\sum_{i=1}^{R} \alpha_i \Phi_i(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d-1)}, \mathbf{e}_j) = \widetilde{a}_j + \delta_j, \quad 1 \leqslant j \leqslant R,$$

with $\widetilde{a}_1, \ldots, \widetilde{a}_R$ integers and

$$|\delta_j| \ll P_{\max}^{(\theta-1)(d-1)} P_j^{-1}, \quad 1 \leqslant j \leqslant n.$$

We now obtain (after changing θ by ϵ for ϵ arbitrarily small) as in the proof of [2, Lemma 2.5] an approximation $1 \leqslant a_1, \ldots, a_R \leqslant q$ to the real numbers α_i of the quality

$$|q\alpha_i - a_i| \leqslant \|\mathbf{g}^\natural\|^{R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+R(d-1)\theta}, \quad 1 \leqslant i \leqslant R.$$

Note that alternative (i) in Lemma 2.5 trivially holds if $P_{\max}^{\theta-1} P_{\min} \leqslant 1$. □

Next we come to the definition of the major arcs. Let $0 < \theta < 1$ and assume that

$$P_{\max}^{\theta-1} P_{\min} \geqslant 1. \tag{2.3}$$

For $q \in \mathbb{N}$ and $1 \leqslant a_1, \ldots, a_R \leqslant q$ we define the major arc

$$\mathfrak{M}_{\mathbf{a},q}(\theta) := \{\boldsymbol{\alpha} \in [0,1]^R : |q\alpha_i - a_i| \leqslant \|\mathbf{g}^\natural\|^{R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+R(d-1)\theta}, \\ 1 \leqslant i \leqslant R\}.$$

Moreover we define the major arcs $\mathfrak{M}(\theta)$ as the union

$$\mathfrak{M}(\theta) = \bigcup_{1 \leqslant q \leqslant \|\mathbf{g}^\natural\|^R P_{\max}^{R(d-1)\theta}} \bigcup_{\substack{1 \leqslant a_1, \ldots, a_R \leqslant q \\ \gcd(\mathbf{a}, q) = 1}} \mathfrak{M}_{\mathbf{a},q}(\theta)$$

and set $\mathfrak{m}(\theta) := [0,1]^R \setminus \mathfrak{M}(\theta)$.

A short calculation gives the following bound for the measure of the major arcs $\mathfrak{M}(\theta)$.

LEMMA 2.6. *Assume that $0 < \theta < 1$ such that (2.3) holds. Then one has*

$$\operatorname{meas}(\mathfrak{M}(\theta)) \ll \|\mathbf{g}^\natural\|^{R^2} P_{\min}^{-R} P_{\max}^{-R(d-1)+R(R+1)(d-1)\theta}.$$

We are now ready to provide an L^1-bound for the exponential sum $S_w(\boldsymbol{\alpha})$ over the minor arcs, which is a modification of [2, Lemma 4.4] and proved in the very same way.

LEMMA 2.7. *Let $0 < \theta < 1$ such that (2.3) holds. Assume that*

$$K > R(R+1)(d-1).$$

Then one has

$$\int_{\mathfrak{m}(\theta)} |S_w(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} \ll \widetilde{\mathbf{P}}^{1+\epsilon} \left(\frac{P_{\max}}{P_{\min}}\right)^K P_{\max}^{-K} + \widetilde{P}^{1+\epsilon} \|\mathbf{g}^\natural\|^{R^2} \left(\frac{P_{\max}}{P_{\min}}\right)^{R+K} \times P_{\max}^{-Rd-(K-R(R+1)(d-1))\theta+\epsilon},$$

for $\epsilon > 0$ arbitrarily small.

For technical convenience we introduce the slightly larger major arcs

$$\mathfrak{M}'_{\mathbf{a},q}(\theta) := \{\boldsymbol{\alpha} \in [0,1]^R : |q\alpha_i - a_i| \leqslant q\|\mathbf{g}^\natural\|^{R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+R(d-1)\theta}, \\ 1 \leqslant i \leqslant R\},$$

and

$$\mathfrak{M}'(\theta) = \bigcup_{1 \leqslant q \leqslant \|\mathbf{g}^\natural\|^R P_{\max}^{R(d-1)\theta}} \; \bigcup_{\substack{1 \leqslant a_1, \ldots, a_R \leqslant q \\ \gcd(\mathbf{a},q)=1}} \mathfrak{M}'_{\mathbf{a},q}(\theta).$$

We record that the major arcs $\mathfrak{M}'_{\mathbf{a},q}(\theta)$ are disjoint for θ small enough and that

$$\mathrm{meas}(\mathfrak{M}'(\theta)) \ll \|\mathbf{g}^\natural\|^{2R^2} P_{\min}^{-R} P_{\max}^{-R(d-1)+(2R^2+R)(d-1)\theta}.$$

A minor modification of the proof of Lemma 4.1 in [2] gives the following result.

LEMMA 2.8. *Assume that*

$$\|\mathbf{g}^\natural\|^{3R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+3R(d-1)\theta} < 1. \tag{2.4}$$

Then for $1 \leqslant q \leqslant \|\mathbf{g}^\natural\|^R P_{\max}^{R(d-1)\theta}$ and $1 \leqslant a_1, \ldots, a_R \leqslant q$, $\gcd(\mathbf{a}, q) = 1$ the major arcs $\mathfrak{M}'_{\mathbf{a},q}(\theta)$ are disjoint.

We now come to the major arc approximation of $S_w(\boldsymbol{\alpha})$. Let $q \in \mathbb{N}$ and $1 \leqslant a_1, \ldots, a_R \leqslant q$. We define the exponential sum

$$S_{\mathbf{a},q} := \sum_{\mathbf{y} \pmod q} e\left(\frac{\mathbf{a}}{q} \cdot \mathbf{g}(\mathbf{y})\right)$$

and the integral

$$I_w(\boldsymbol{\gamma}) := \int_{\mathbb{R}^n} e(\boldsymbol{\gamma} \cdot \mathbf{g}(\mathbf{u})) \prod_{i=1}^n w\left(\frac{u_i}{P_i} - z_i\right) d\mathbf{u}.$$

LEMMA 2.9. *Let $q \in \mathbb{N}$ and $0 \leqslant a_1, \ldots, a_R < q$. Write $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$. Assume that $q < P_{\min} P_{\max}^{-\varepsilon}$ and*

$$|\boldsymbol{\beta}| q P_{\max}^{d-1} \|\mathbf{g}\| < P_{\max}^{-\varepsilon}.$$

Then one has the following approximation for any real $N \geqslant 1$:

$$S_w(\boldsymbol{\alpha}) = q^{-n} S_{\mathbf{a},q} I_w(\boldsymbol{\beta}) + O_N(\widetilde{\mathbf{P}} P_{\max}^{-N}).$$

Proof. We recall the definition of the exponential sum $S_w(\boldsymbol{\alpha})$ as

$$S_w(\boldsymbol{\alpha}) = \sum_{\mathbf{x} \in \mathbb{Z}^n} \prod_{i=1}^{n} w\left(\frac{x_i}{P_i} - z_i\right) e(\boldsymbol{\alpha} \cdot \mathbf{g}(\mathbf{x})).$$

We split the summation variables $\mathbf{x}$ into residue classes modulo q and obtain

$$S_w(\boldsymbol{\alpha}) = \sum_{\mathbf{y} \pmod q} e\left(\frac{\mathbf{a}}{q} \cdot \mathbf{g}(\mathbf{y})\right) \sum_{\mathbf{w} \in \mathbb{Z}^n} \prod_{i=1}^{n} w\left(\frac{y_i + w_i q}{P_i} - z_i\right) e(\boldsymbol{\beta} \cdot \mathbf{g}(\mathbf{y} + q\mathbf{w})).$$

We now consider the inner sum for a fixed vector $\mathbf{y}$ modulo q. Let

$$\psi(\mathbf{w}) := \prod_{i=1}^{n} w\left(\frac{y_i + w_i q}{P_i} - z_i\right) e(\boldsymbol{\beta} \cdot \mathbf{g}(\mathbf{y} + q\mathbf{w})).$$

We apply Euler–Maclaurin's summation formula (see [28, Theorem B.5]) of order $\tilde{\kappa}$ into each coordinate direction. If we choose $\tilde{\kappa}$ large enough depending only on ε, n and N we obtain

$$\sum_{\mathbf{w} \in \mathbb{Z}^n} \psi(\mathbf{w}) = \int_{\mathbf{w} \in \mathbb{R}^n} \psi(\mathbf{w})\, d\mathbf{w} + O_N(\widetilde{\mathbf{P}} P_{\max}^{-N}).$$

Note that all the boundary terms in Euler–Maclaurin's summation formula vanish due to the smooth weight function w. Since N was arbitrary we find after even enlarging $\tilde{\kappa}$ that

$$S_w(\boldsymbol{\alpha}) = S_{\mathbf{a},q} \int_{\mathbf{w} \in \mathbb{R}^n} \psi(\mathbf{w})\, d\mathbf{w} + O_N(\widetilde{\mathbf{P}} P_{\max}^{-N}).$$

A variable substitution now gives the statement of the lemma. □

Next we consider the singular integral. Note that in contrast to most approaches we defined the integral $I_w(\boldsymbol{\gamma})$ with the inhomogeneous polynomials $\mathbf{g}(\mathbf{y})$ instead of taking their homogenizations. We now replace $\mathbf{g}(\mathbf{y})$ by $\mathbf{g}^\natural(\mathbf{y})$ in $I_w(\boldsymbol{\gamma})$ which will simplify the discussion of absolute convergence. Define

$$I_w^\natural(\boldsymbol{\gamma}) = \int_{\mathbb{R}^n} e(\boldsymbol{\gamma} \cdot \mathbf{g}^\natural(\mathbf{u})) \prod_{i=1}^{n} w\left(\frac{u_i}{P_i} - z_i\right) d\mathbf{u}.$$

LEMMA 2.10. *Assume that* $|\mathbf{z}| \leqslant 1$. *Then one has*

$$I_w(\boldsymbol{\gamma}) - I_w^{\natural}(\boldsymbol{\gamma}) \ll \widetilde{\mathbf{P}}|\boldsymbol{\gamma}|\|\mathbf{g}\| P_{\max}^{d-1}.$$

The proof of the lemma follows from directly comparing the integrands of the two integrals. Under the assumptions of Lemma 2.9 we observe that

$$S_w(\boldsymbol{\alpha}) = q^{-n} S_{\mathbf{a},q} I_w^{\natural}(\boldsymbol{\beta}) + O_N(\widetilde{\mathbf{P}} P_{\max}^{-N}) + O(\widetilde{\mathbf{P}}|\boldsymbol{\beta}|\|\mathbf{g}\| P_{\max}^{d-1}).$$

We define the truncated singular series

$$\mathfrak{S}(Q) := \sum_{q \leqslant Q} q^{-n} S_{\mathbf{a},q},$$

and the truncated singular integral

$$J_w(Q) := \int_{|\boldsymbol{\gamma}| \leqslant Q} I_w^{\natural}(\boldsymbol{\gamma})\, d\boldsymbol{\gamma}.$$

With these definitions we can write the major arc contribution in the following way.

LEMMA 2.11. *Assume* $|\mathbf{z}| \leqslant P_{\max}$ *and that (2.4) holds, as well as*

$$\max\{\|\mathbf{g}^{\natural}\|^R P_{\max}^{R(d-1)\theta} P_{\min}^{-1}, \|\mathbf{g}\|\|\mathbf{g}^{\natural}\|^{2R-1} P_{\min}^{-1} P_{\max}^{2R(d-1)\theta}\} < P_{\max}^{-\varepsilon}. \tag{2.5}$$

Then one has

$$\begin{aligned}\int_{\mathfrak{M}'(\theta)} S_w(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = {} & \mathfrak{S}(\|\mathbf{g}^{\natural}\|^R P_{\max}^{R(d-1)\theta}) \\ & \times J_w(\|\mathbf{g}^{\natural}\|^{R-1} P_{\min}^{-1} P_{\max}^{-(d-1)+R(d-1)\theta}) + O_N(\widetilde{\mathbf{P}} P_{\max}^{-N}) \\ & + O(\|\mathbf{g}^{\natural}\|^{2R^2+R}\|\mathbf{g}\|\widetilde{\mathbf{P}} P_{\min}^{-R-1} P_{\max}^{-R(d-1)+(2R^2+2R)(d-1)\theta}),\end{aligned}$$

for any real $N \geqslant 1$.

Proof. By Lemma 2.8 the major arcs are disjoint thus the proof follows from Lemma 2.9. $\square$

Next we aim to complete the singular series. We recall [**22**, Lemma 2.2] (see also [**23**, §2, Lemma 2.14]).

LEMMA 2.12. *For any* $\varepsilon > 0$ *one has*

$$|S_{\mathbf{a},q}| \ll \|\mathbf{g}^{\natural}\|^{K/(d-1)} q^{n-K/R(d-1)+\varepsilon}.$$

We shall soon see that the truncated singular series $\mathfrak{S}(Q)$ is converging for $Q \to \infty$, thus we shall set

$$\mathfrak{S} = \lim_{Q \to \infty} \mathfrak{S}(Q). \tag{2.6}$$

Lemma 2.12 gives the following speed of convergence.

LEMMA 2.13. *Assume that $K > R(d-1)$. Then $\mathfrak{S}$ is absolutely convergent. Moreover one has*

$$\mathfrak{S} - \mathfrak{S}(Q) \ll \|\mathbf{g}^\natural\|^{K/(d-1)} Q^{1-K/R(d-1)+\varepsilon},$$

for any $\varepsilon > 0$ and $|\mathfrak{S}| \ll \|\mathbf{g}^\natural\|^{K/(d-1)}$.

In preparation for the proof of the absolute convergence of the singular integral, we note the following lemma, which is a consequence of Lemma 2.5.

LEMMA 2.14. *Assume that $|\boldsymbol{\alpha}|^3 \|\mathbf{g}^\natural\|^2 P_{\min}^2 P_{\max}^{2(d-1)} < 1$. Then one has*

$$S_w(\boldsymbol{\alpha}) \ll \widetilde{\mathbf{P}}^{1+\epsilon} \left(\frac{P_{\max}}{P_{\min}}\right)^K (|\boldsymbol{\alpha}| \|\mathbf{g}^\natural\|^{-R+1} P_{\min} P_{\max}^{d-1})^{-K/R(d-1)},$$

for any positive ϵ.

LEMMA 2.15. *Assume that $P_i \geqslant 1$ for $1 \leqslant i \leqslant n$ and that $|\mathbf{z}| \leqslant 1$. Then*

$$I_w^\natural(\boldsymbol{\gamma}) \ll \widetilde{\mathbf{P}} \min \left\{1, \widetilde{\mathbf{P}}^\epsilon \left(\frac{P_{\max}}{P_{\min}}\right)^{K(1+1/R(d-1))} (P_{\max}^d |\boldsymbol{\gamma}| \|\mathbf{g}^\natural\|^{-R+1})^{-K/R(d-1)}\right\}.$$

The proof of Lemma 2.15 is relatively standard (see [**2**, Lemma 5.2]), with the exception that we compare the oscillatory integral $I_w^\natural(\boldsymbol{\gamma})$ with parameters $P_1, \ldots, P_n$ to an exponential sum with box length $B_1, \ldots, B_n$ such that $B_i/B_{\max} = P_i/P_{\max}$ for all $1 \leqslant i \leqslant n$.

We shall show that the truncated singular integral $J_w(Q)$ converges for $Q \to \infty$, we will therefore let

$$J_w := \lim_{Q \to \infty} J_w(Q), \tag{2.7}$$

and call it the singular integral.

In the following we will always assume that $1 \leqslant P_i$ for $1 \leqslant i \leqslant n$ and that $|\mathbf{z}| \leqslant 1$. As a consequence of Lemma 2.15 we obtain the following result.

LEMMA 2.16. *Assume that $K > R^2(d-1)$. Then J_w is absolutely convergent and*

$$J_w - J_w(Q) \ll \widetilde{\mathbf{P}}^{1+\epsilon} \left(\frac{P_{\max}}{P_{\min}}\right)^{K(1+1/R(d-1))} \times (P_{\max}^d \|\mathbf{g}^\natural\|^{-R+1})^{-K/R(d-1)} Q^{-K/R(d-1)+R}.$$

Moreover, we have

$$J_w \ll \widetilde{\mathbf{P}}^{1+\epsilon} \left(\frac{P_{\max}}{P_{\min}}\right)^{R^2(d-1)+R} P_{\max}^{-Rd} \|\mathbf{g}^\natural\|^{R(R-1)}.$$

We can now complete both the singular series and singular integral in our major arc analysis. According to Lemmas 2.13 and 2.16 we obtain the following result.

LEMMA 2.17. *Assume that equations (2.3), (2.4) and (2.5) hold and* $|\mathbf{z}| \leqslant P_{\max}$*, as well as* $K > R^2(d-1)$*. Then the following holds for any real* $N \geqslant 1$*:*

$$\int_{\mathfrak{M}'(\theta)} S_w(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \mathfrak{S} J_w + O(\|\mathbf{g}^\natural\|^{2R^2+R}\|\mathbf{g}\|\widetilde{\mathbf{P}} P_{\min}^{-R-1} P_{\max}^{-R(d-1)+(2R^2+2R)(d-1)\theta})$$
$$+ O_N\left(\widetilde{\mathbf{P}} P_{\max}^{-N} + \|\mathbf{g}^\natural\|^{K/(d-1)+R^2-R}\widetilde{\mathbf{P}}^{1+\epsilon}\left(\frac{P_{\max}}{P_{\min}}\right)^{R+K} P_{\max}^{-Rd-K\theta+R^2(d-1)\theta}\right).$$

Theorem 2.1 is now a consequence of the major arc analysis in Lemma 2.17 in combination with the minor arc analysis from Lemma 2.7. For this, we choose θ by

$$P_{\max}^{\theta} = \|\mathbf{g}\|^{-1/(2R(d-1)+1)}\|\mathbf{g}^\sharp\|^{-3R/(3R(d-1)+1)} P_{\max}^{1/4R(R+1)d}.$$

Then we clearly have $0 < \theta < 1$ and equation (2.3) reduces to the assumption (2.2). Moreover, one quickly sees that with this choice of θ both of the conditions (2.4) and (2.5) are satisfied. It remains to understand that the error terms in Lemma 2.17 and Lemma 2.7 are both majorized by the error term in Theorem 2.1. We bound the first error term in Lemma 2.17 by

$$\|\mathbf{g}^\sharp\|^{2R^2+R}\|\mathbf{g}\|\widetilde{\mathbf{P}}\left(\frac{P_{\max}}{P_{\min}}\right)^R P_{\max}^{-Rd-1} P_{\max}^{(2R^2+2R)d\theta} \ll \widetilde{\mathbf{P}}\left(\frac{P_{\max}}{P_{\min}}\right)^R P_{\max}^{-Rd-1/2}.$$

Note that the last error term in Lemma 2.17 as well as the second error term in Lemma 2.7 are bounded by

$$\|\mathbf{g}^\natural\|^{2K/(d-1)-R}\|\mathbf{g}\|^{(K-R^2(d-1))/2R(d-1)}\widetilde{\mathbf{P}}^{1+\epsilon}(P_{\max}/P_{\min})^{R+K}$$
$$\times P_{\max}^{-Rd-(K-R(R+1)(d-1))/4R(R+1)d}.$$

The first error term in Lemma 2.7 is also present in the statement of Theorem 2.1.

Lastly let us remark that it is a well-known fact that the singular series factorizes as $\mathfrak{S} = \prod_p \sigma_p(\mathbf{g})$, where for any prime p we have

$$\sigma_p(\mathbf{g}) := \lim_{l\to\infty} p^{-l(n-R)}\sharp\{1 \leqslant \mathbf{x} \leqslant p^l : p^l|\mathbf{g}(\mathbf{x})\}.$$

§3. *Local densities.* Throughout this section we will have R forms of degree $d > 1$,

$$f_1, \ldots, f_R \in \mathbb{Z}[x_1, \ldots, x_n]$$

and we will always assume that the Birch rank satisfies

$$\mathfrak{B}(\mathbf{f}) > 2^{d-1}(d-1)R(R+1).$$

For a prime p and a vector $\mathbf{j} = (j_1, \ldots, j_n) \in (\mathbb{Z}_{\geqslant 0})^n$ we shall be concerned with bounding the quantities

$$\delta(\mathbf{j}) := \lim_{l \to \infty} p^{-l(n-R)} \sharp\{1 \leqslant x_1, \ldots, x_n \leqslant p^l : p^l | \mathbf{f}(p^{j_1} x_1, \ldots, p^{j_n} x_n)\},$$

these estimates will be applied later towards the proof of Theorems 1.1, 1.5 and 1.6. We suppress the letter p from the notation for δ to make the presentation easier to follow. The forms $\mathbf{f}$ will be considered constant, however the prime p and the vector $\mathbf{j}$ will not, thus we shall require uniformity of our bounds with respect to p and $\mathbf{j}$. For later applications we only have to consider all big enough primes $p > z_0$, where z_0 is a constant depending at most on the coefficients of $\mathbf{f}$ and n, d, R. This constant will be enlarged, if needed, with no further comment. Let us emphasize that the entities $\delta(\mathbf{j})$ encode the probability of the events

$$p^{j_1} | x_1, \ldots, p^{j_n} | x_n$$

as $\mathbf{x} \in \mathbb{Z}^n$ sweeps through the zeros of $\mathbf{f} = \mathbf{0}$, therefore, they are intimately connected with certain closed subvarieties of $\mathbf{f} = \mathbf{0}$. This is manifested even in the most simple of situations: for a primitive integer zero of $x_1 x_2 = x_3^2$ and a prime $p | x_3$ we always have $p^2 | x_1$ or $p^2 | x_2$ as a result of the subvariety $x_1 x_2 = x_3^2$, $x_3 = 0$ being reducible. We shall give geometric conditions that prevent $\delta(\mathbf{j})$ to attain large values for general systems $\mathbf{f} = \mathbf{0}$.

For every $\mathbf{j} \in \{0, 1\}^n$ we define the system $\mathbf{f}^{\mathbf{j}} = \mathbf{0}$ of R forms in $n - |\mathbf{j}|_1$ variables via

$$f_\xi^{\mathbf{j}}(\mathbf{x}) = f_\xi(x_1, \ldots, x_n)|_{x_i = 0 \text{ if } j_i = 1}, \quad \xi \in \mathbb{N} \cap [1, R].$$

We later need a lower bound for the Birch rank of the new systems, as for example obtained in [**11**, Lemma 3]. As there is a slight oversight in the proof of [**11**, Lemma 3], we give here the statement and proof of the corrected lemma where the quantity R in [**11**, Lemma 3] is replaced by $R + 1$.

LEMMA 3.1. *One has*

$$\mathfrak{B}(\mathbf{f}^{\mathbf{j}}) \geqslant \mathfrak{B}(\mathbf{f}) - (R+1)|\mathbf{j}|_1. \tag{3.1}$$

It is important to note here that we view $\mathbf{f}^{\mathbf{j}}$ *as a system of* R *equations in* $n - |\mathbf{j}|_1$ *variables.*

Proof. Let $\widetilde{V}^* \subset \mathbb{P}_{\mathbb{C}}^{n-1}$ be the projective variety given by

$$\operatorname{rank}\left(\frac{\partial f_\xi}{\partial x_i}\right)_{\xi, i} < R,$$

and note that this is well defined as all of the polynomials f_ξ are homogeneous. Then the Birch rank of $\mathbf{f}$ is given by

$$\mathfrak{B}(\mathbf{f}) = n - \dim(\widetilde{V}^*) - 1.$$

Similarly, let $\widetilde{V^{*,\mathbf{j}}} \subset \mathbb{P}_{\mathbb{C}}^{n-|\mathbf{j}|_1-1}$ be the projective variety given by

$$\operatorname{rank}\left(\frac{\partial f_\xi^{\mathbf{j}}}{\partial x_i}\right)_{\xi,i} < R,$$

such that we have

$$\mathfrak{B}(\mathbf{f}^{\mathbf{j}}) = n - |\mathbf{j}|_1 - \dim(\widetilde{V^{*,\mathbf{j}}}) - 1.$$

The variety $\widetilde{V^{*,\mathbf{j}}}$ naturally embeds into the linear subspace of $\mathbb{P}_{\mathbb{C}}^{n-1}$ given by $x_i = 0$ for $j_i = 1$. We write $\iota(\widetilde{V^{*,\mathbf{j}}})$ for this embedding. Then we observe that

$$\iota(\widetilde{V^{*,\mathbf{j}}}) \cap \bigcap_{1\leqslant \xi \leqslant R} \bigcap_{\substack{1\leqslant i\leqslant n\\ j_i=1}} \left\{\frac{\partial f_\xi}{\partial x_i} = 0\right\} \subset \widetilde{V^*}.$$

Hence we obtain

$$\dim(\widetilde{V^{*,\mathbf{j}}}) - R|\mathbf{j}|_1 \leqslant \dim(\widetilde{V^*}).$$

Finally, this implies

$$\mathfrak{B}(\mathbf{f}^{\mathbf{j}}) \geqslant n - |\mathbf{j}|_1 - (\dim(\widetilde{V^*}) + R|\mathbf{j}|_1) - 1 = \mathfrak{B}(\mathbf{f}) - (R+1)|\mathbf{j}|_1. \qquad \square$$

This is a convenient place to introduce the helpful notation

$$\Theta(\mathbf{j}) := \frac{\mathfrak{B}(\mathbf{f}^{\mathbf{j}})}{R(d-1)2^{d-1}}$$

and $\Theta(\mathbf{0})$ will be denoted by Θ. For non-negative integers $j_1, \ldots, j_n$, any prime p and a vector $\mathbf{x}$ we use the notation

$$p^{\mathbf{j}}|\mathbf{x} \iff p^{j_i}|x_i \quad \text{for all } 1 \leqslant i \leqslant n.$$

This enables us to introduce the densities

$$\sigma_p(p^{\mathbf{j}}|\mathbf{x}) = \lim_{l\to\infty} p^{-l(n-R)} \sharp\{1 \leqslant x_1, \ldots, x_n \leqslant p^l : p^l|\mathbf{f}(\mathbf{x}),\, p^{\mathbf{j}}|\mathbf{x}\}$$

and from the definition of δ we infer that

$$\frac{\delta(\mathbf{j})}{p^{|\mathbf{j}|_1}} = \sigma_p(p^{\mathbf{j}}|\mathbf{x}).$$

LEMMA 3.2. *Let t, d be integers with $2 \leqslant d < t$. Then for each $\mathbf{a} \in (\mathbb{Z}/p^{t-d}\mathbb{Z})^R$ with $p \nmid \mathbf{a}$ and any vector polynomial $\mathbf{g} \in \mathbb{Z}[\mathbf{x}]^R$ with $\max_{1\leqslant i\leqslant R} \deg(g_i) \leqslant d-1$ we have*

$$\sum_{\mathbf{x} \pmod{p^{t-1}}} e_{p^t}(p^d \mathbf{a}\cdot\mathbf{f}(\mathbf{x}) + p\mathbf{a}\cdot\mathbf{g}(\mathbf{x})) \ll_\epsilon p^{(t-1)(n-((t-d)/(t-1))\Theta+\epsilon)},$$

where the implied constant is independent of p, t, $\mathbf{g}$ and $\mathbf{a}$.

Proof. We shall use [2, Lemma 4.3] with $P = p^{t-1}$ and $\boldsymbol{\alpha} = p^{-t+d}\mathbf{a}$; in doing so we observe that lower degree polynomials leave the strength of the bounds in [2] unaffected. Recall that the constant K in [2, equation (8)] is given via $\mathfrak{B}(\mathbf{f})/2^{d-1}$. Our aim is to acquire a constant $\eta > 0$, as large as possible, such that $\boldsymbol{\alpha} \notin \mathscr{M}(\eta)$, where $\mathscr{M}(\eta)$ is given in [2, §4, equation (5)]. This would then imply that the sum in our lemma is

$$\ll p^{(t-1)(n-\mathfrak{B}(\mathbf{f})\eta/2^{d-1}+\epsilon)}.$$

The assumption $\boldsymbol{\alpha} \in \mathscr{M}(\eta)$ provides non-negative integers $q', a'_1, \ldots, a'_R$ fulfilling

$$\gcd(a'_1, \ldots, a'_R, q') = 1, \quad 1 \leqslant q' \leqslant p^{(t-1)R(d-1)\eta}$$

and such that for all $i = 1, \ldots, R$ the succeeding inequality is valid,

$$2|q'a_i - a'_i p^{t-d}| \leqslant p^{t-d+(t-1)(-d+R(d-1)\eta)}. \tag{3.2}$$

As explained in [2, Lemma 4.1], we need to assume $2R(d-1)\eta < d$ in order to ensure that the major arcs are disjoint. It is straightforward to infer that this condition is met upon choosing

$$\eta := \eta(\epsilon) = \frac{t-d}{(t-1)R(d-1)} - \epsilon$$

for any small enough $\epsilon > 0$. Furthermore, this choice of η makes the exponent of p in (3.2) non-positive, thus giving birth to the equalities $q'a_i = a'_i p^{t-d}$ for all i. In particular, we obtain $p^{t-d} = q' \leqslant p^{(t-1)R(d-1)\eta}$, thus $t-d \leqslant (t-1)R(d-1)\eta$, which constitutes a violation to the the definition of η. $\square$

For $\mathbf{j} \in \{0, 1\}^n$, $c \in \mathbb{N}$ and any prime p define

$$E(p^c; \mathbf{j}) := \sharp\{\mathbf{x} \pmod{p^c} : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \pmod{p^c},\ p^{j_i}|x_i \forall i\}.$$

This quantity is intimately related to the geometry of $\mathbf{f}^{\mathbf{j}} = \mathbf{0}$ and we begin by using it to approximate $\delta(\mathbf{j})$.

LEMMA 3.3. *Let $\mathbf{j} \in \{0, 1\}^n$ and assume that $\Theta > R$. Then there is some $z_0 > 0$, such that for $p > z_0$ and each sufficiently small $\epsilon > 0$, we have*

$$\delta(\mathbf{j}) = p^{d(R-n)+|\mathbf{j}|_1} E(p^d; \mathbf{j}) + O(p^{-\Theta+R(d+1)+\epsilon}),$$

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

Proof. For $t \geqslant 1$, $\mathbf{j} \in \{0, 1\}^n$ and any $\mathbf{a} \in \mathbb{Z}^R$ we bring into play the entities

$$W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x}) := \sum_{\substack{\mathbf{x} \pmod{p^t} \\ p^{\mathbf{j}}|\mathbf{x}}} e_{p^t}(\mathbf{a} \cdot \mathbf{f}(\mathbf{x}))$$

and

$$G(\mathbf{j};\, p^t) := p^{-tn} \sum_{\mathbf{a} \pmod{p^t}}^{*} W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x}), \tag{3.3}$$

where the summation $\sum^{*}_{\mathbf{a} \pmod q}$ is over vectors $\mathbf{a} \in (\mathbb{Z}/q\mathbb{Z})^R$ with $\gcd(\mathbf{a}, q) = 1$. We have

$$\begin{aligned}
\delta(\mathbf{j}) p^{-|\mathbf{j}|_1} &= \lim_{l\to\infty} p^{-l(n-R)} \sum_{\mathbf{a} \pmod{p^l}} \frac{1}{p^{Rl}} \sum_{\substack{\mathbf{x} \pmod{p^l}\\ p^{\mathbf{j}}|\mathbf{x}}} e_{p^l}(\mathbf{a}\cdot\mathbf{f}(\mathbf{x}))\\
&= \lim_{l\to\infty} \sum_{\mathbf{a} \pmod{p^l}} p^{-ln} \sum_{\substack{\mathbf{x} \pmod{p^l}\\ p^{\mathbf{j}}|\mathbf{x}}} e_{p^l}(\mathbf{a}\cdot\mathbf{f}(\mathbf{x}))\\
&= \lim_{l\to\infty} \left(\sum_{t=1}^{l} G(\mathbf{j};\, p^t) + p^{-ln} \sharp\{\mathbf{x} \pmod{p^l} : p^{\mathbf{j}}|\mathbf{x}\} \right)\\
&= p^{-|\mathbf{j}|_1} + \lim_{l\to\infty} \sum_{t=1}^{l} G(\mathbf{j};\, p^t),
\end{aligned}$$

whence

$$\delta(\mathbf{j}) = 1 + p^{|\mathbf{j}|_1} \sum_{t=1}^{\infty} G(\mathbf{j};\, p^t). \tag{3.4}$$

Observe that for each form $F \in \mathbb{Z}[\mathbf{x}]$, any prime p and any fixed integer vector $\mathbf{y}$ there exists an integer polynomial $F_{\mathbf{y}} \in \mathbb{Z}[\mathbf{x}]$ of degree strictly smaller than $\deg(F)$, such that

$$F(\mathbf{y} + p\mathbf{x}) = p^{\deg(F)} F(\mathbf{x}) + F(\mathbf{y}) + pF_{\mathbf{y}}(\mathbf{x}).$$

Hence, if $t \geqslant d+1$, this allows us to rewrite the exponential sum $W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x})$ as

$$\begin{aligned}
&\sum_{\substack{\mathbf{y}\in(\mathbb{N}\cap[1,p])^n\\ p^{\mathbf{j}}|\mathbf{y}}} \sum_{\mathbf{h}\in(\mathbb{N}\cap[1,p^{t-1}])^n} e_{p^t}(\mathbf{a}\cdot\mathbf{f}(\mathbf{y}+p\mathbf{h}))\\
&\quad= \sum_{\substack{\mathbf{y}\in(\mathbb{N}\cap[1,p])^n\\ p^{\mathbf{j}}|\mathbf{y}}} e(p^{-t}\mathbf{a}\cdot\mathbf{f}(\mathbf{y})) \sum_{\mathbf{h}\in(\mathbb{N}\cap[1,p^{t-1}])^n} e(p^{d-t}\mathbf{a}\cdot\mathbf{f}(\mathbf{h}) + p^{-t+1}\mathbf{a}\cdot\mathbf{g}_{\mathbf{y}}(\mathbf{h})),
\end{aligned}$$

where the polynomials $\mathbf{g}_{\mathbf{y}}(\mathbf{h})$ have degree strictly smaller than d in $\mathbf{h}$. Invoking Lemma 3.2 endows us with the following bound for the inner sum over $\mathbf{h}$:

$$\ll p^{(t-1)(n+\epsilon)-(t-d)\Theta},$$

where the implicit constant is independent of $p, t, \mathbf{y}$ and $\mathbf{a}$. Hence, for $t > d$ we deduce that

$$W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x}) \ll p^{t(n+\epsilon)-|\mathbf{j}|_1-(t-d)\Theta},$$

thereby procuring the validity of

$$\sum_{t=d+1}^{\infty} |G(\mathbf{j};\, p^t)| \ll p^{-|\mathbf{j}|_1+d\Theta} \sum_{t=d+1}^{\infty} p^{-t(\Theta-R-\epsilon)}.$$

Our assumption $R < \Theta$ shows that for each $0 < \epsilon < (\Theta - R)/2$ the sum over t has the value

$$\frac{p^{-(d+1)(\Theta-R-\epsilon)}}{1-p^{-(\Theta-R-\epsilon)}}$$

and increasing the value of z_0 to ensure that $z_0^{(\Theta-R)/2} \geqslant 2$ shows that

$$\sum_{t=d+1}^{\infty} |G(\mathbf{j};\, p^t)| \ll p^{-|\mathbf{j}|_1-\Theta+R(d+1)+\epsilon(d+1)}.$$

To control the contribution of the terms with $t \leqslant d$ we note that

$$p^{-|\mathbf{j}|_1} + \sum_{t=1}^{d} G(\mathbf{j};\, p^t) = p^{d(R-n)} \sharp\{\mathbf{x} \pmod{p^d} : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \pmod{p^d},\, p^{\mathbf{j}} | \mathbf{x}\},$$

thus concluding our proof. □

Observe that, at least when $|\mathbf{j}|_1$ is relatively small, the quantity $E(p^d; \mathbf{j})$ regards the number of zeros (mod p) of a variety in sufficiently many variables; thus the estimates of Birch yield the required estimation of $E(p^d; \mathbf{j})$.

LEMMA 3.4. *Let $\mathbf{j} \in \{0, 1\}^n$ and assume that $\Theta(\mathbf{j}) > R$ is fulfilled. Then for all $\epsilon > 0$ and primes $p > z_0$ we have*

$$E(p^d; \mathbf{j}) = p^{d(n-R)-|\mathbf{j}|_1} + O_\epsilon(p^{d(n-R)-|\mathbf{j}|_1-(\Theta(\mathbf{j})-R)+\epsilon}),$$

with an implicit constant that is independent of p.

Proof. We initiate our argument by slicing the counting function $E(p^d; \mathbf{j})$ along the variables which are divisible by p. Let $I = \{1 \leqslant i \leqslant n : j_i = 1\}$ and for $\mathbf{x}' = (x_i)_{i \in I} \in (\mathbb{Z}/p^d\mathbb{Z})^{|I|}$ we define

$$E(p^d; \mathbf{j}; \mathbf{x}') := \sharp\{x_i \pmod{p^d},\, i \notin I : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \pmod{p^d}\}.$$

We rewrite this counting function with exponential sums as follows,

$$E(p^d; \mathbf{j}; \mathbf{x}') = p^{d(n-|\mathbf{j}|_1)-dR} + p^{-dR} \sum_{t=1}^{d} p^{(n-|\mathbf{j}|_1)(d-t)} \times \sideset{}{^*}\sum_{\mathbf{a} \pmod{p^t}} \sum_{\substack{x_i \pmod{p^t} \\ i \notin I}} e_{p^t}(\mathbf{a} \cdot \mathbf{f}(\mathbf{x})).$$

Note that the degree d part of the polynomial $\mathbf{f}(\mathbf{x})$ when viewed as a polynomial in the variables $x_i, i \notin I$, is $\mathbf{f}^{\mathbf{j}}(\mathbf{x})$. We now apply [2, Lemma 5.4], the strength of which is unaffected by lower degree polynomials, to obtain for any $\epsilon > 0$ and uniformly for all $p > z_0$,

$$\sum_{\substack{x_i \pmod{p^t} \\ i \notin I}} e_{p^t}(\mathbf{a} \cdot \mathbf{f}(\mathbf{x})) \ll_\epsilon p^{t(n-|\mathbf{j}|_1-\Theta(\mathbf{j}))+\epsilon}.$$

We use this to estimate $E(p^d; \mathbf{j}; \mathbf{x}')$ as follows,

$$\begin{aligned} E(p^d; \mathbf{j}; \mathbf{x}') - p^{d(n-|\mathbf{j}|_1-R)} &\ll_\epsilon p^{d(n-|\mathbf{j}|_1-R)+\epsilon} \sum_{t=1}^{d} p^{t(R-\Theta(\mathbf{j}))} \\ &\ll_\epsilon p^{d(n-|\mathbf{j}|_1-R)-(\Theta(\mathbf{j})-R-\epsilon)}. \end{aligned}$$

We can now evaluate $E(p^d; \mathbf{j})$ as

$$\sum_{\substack{x_i \pmod{p^d}, i \in I \\ p|x_i}} E(p^d; \mathbf{j}; \mathbf{x}') = p^{dn-|\mathbf{j}|_1-Rd} + O_\epsilon(p^{d(n-R)-|\mathbf{j}|_1-(\Theta(\mathbf{j})-R)+\epsilon}),$$

which concludes our proof. $\square$

Tying Lemmas 3.3 and 3.4 together provides the succeeding estimate.

COROLLARY 3.5. *Assume that* $\mathbf{j} \in \{0, 1\}^n$, $\min\{\Theta, \Theta(\mathbf{j})\} > R$ *and that* p *is a prime in the range* $p > z_0$. *Then the following holds for each* $\epsilon > 0$ *with an implied constant depending only on* $\mathbf{f}$ *and* ϵ*:*

$$\delta(\mathbf{j}) = 1 + O(p^{R-\min\{\Theta-dR, \Theta(\mathbf{j})\}+\epsilon}).$$

Utilizing (3.1) to find lower bounds for $\Theta(\mathbf{j})$ gives the following consequence of Corollary 3.5.

COROLLARY 3.6. *Assume that for some* $\mathbf{j} \in \{0, 1\}^n$ *we have*

$$\mathfrak{B}(\mathbf{f}) > \max\{(d-1)R^2 2^{d-1} + (R+1)|\mathbf{j}|_1, (d^2-1)R^2 2^{d-1}\}.$$

Then there exists $\lambda > 0$ *such that for all large enough primes* $p > z_0 = z_0(\mathbf{f})$*, we have*

$$\delta(\mathbf{j}) = 1 + O(p^{-\lambda}),$$

with an implied constant depending only on $\mathbf{f}$.

We can see that the bound $\delta(\mathbf{j}) \ll 1$ fails when $|\mathbf{j}|_1$ approaches n hence the assumption $\Theta(\mathbf{j}) > R$ of Corollary 3.5 is no longer applicable. Indeed, a moment's thought reveals that $\delta(1, \ldots, 1) = p^{dR}\sigma_p$ and that whenever $h_i \geqslant j_i$ for all $1 \leqslant i \leqslant n$ then $\delta(\mathbf{j}) \geqslant \delta(\mathbf{h}) p^{|\mathbf{j}|_1-|\mathbf{h}|_1}$. The bound $\sigma_p \gg 1$, valid with an implied constant independent of p when p is sufficiently large, reveals that for such p we have

$$n-\frac{dR}{2}<|\mathbf{j}|_1\leqslant n\Rightarrow\delta(\mathbf{j})\gg p^{dR/2}$$

with an implied constant independent of p. Therefore we need to provide (necessarily weaker) bounds for the densities $\delta(\mathbf{j})$ which are however valid through the whole range $1\leqslant|\mathbf{j}|_1\leqslant n$. The crucial import will be bounds for the exponential sums in Birch's work with the additional property that the dependence on the coefficients of the underlying forms is explicitly recorded.

LEMMA 3.7. *Assume that $\Theta>R$. Then there exists a large $z_0=z_0(\mathbf{f})$ such that for each $\mathbf{j}\in\{0,1\}^n$, $\epsilon>0$ and prime $p>z_0$ the following holds with an implicit constant depending at most on ϵ and $\mathbf{f}$:*

$$\delta(\mathbf{j})\ll p^{dR\Theta+R-\Theta+\epsilon}.$$

Proof. We start by rewriting

$$W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x})=p^{-|\mathbf{j}|_1}\sum_{\mathbf{x}\ (\mathrm{mod}\ p^t)}e_{p^t}(\mathbf{a}\cdot\mathbf{f}(p^{j_1}x_1,\ldots,p^{j_n}x_n))$$

and considering $\mathbf{f}(p^{j_1}x_1,\ldots,p^{j_n}x_n)$ as a system of homogeneous polynomials in the variables $x_1,\ldots,x_n$. Note that the maximum of the coefficients is bounded by C_1p^d for a positive constant $C_1=C_1(\mathbf{f})$ that is independent of p. Moreover, the Birch rank of the system $\mathbf{f}(\mathbf{x})=\mathbf{0}$ equals the Birch rank of the system $\mathbf{f}(p^{j_1}x_1,\ldots,p^{j_n}x_n)=\mathbf{0}$. Alluding to the estimate [22, Lemma 2.2] supplies us with the bound

$$W_{\mathbf{a},p^t}(p^{\mathbf{j}}|\mathbf{x})p^{|\mathbf{j}|_1}\ll_\epsilon p^{dR\Theta+t(n-\Theta+\epsilon)},$$

which, once injected into (3.4), offers the validity of

$$\delta(\mathbf{j})-1\ll p^{dR\Theta}\sum_{t=1}^{\infty}p^{t(R-\Theta+\epsilon)}.$$

Enlarging z_0 and $1/\epsilon$ if needed, ensures the convergence of the sum over t to a value that is $\ll_{z_0}p^{R-\Theta+\epsilon}$, independently of p. □

For a prime p and a vector $\mathbf{j}\in(\mathbb{Z}_{\geqslant 0})^n$ we define

$$\varpi(p^{j_1},\ldots,p^{j_n}):=\frac{\delta(\mathbf{j})}{\sigma_p(\mathbf{f})}. \tag{3.5}$$

The standard estimate $\sigma_p=1+O(p^{-1-\epsilon(\mathbf{f})})$ holds for some $\epsilon(\mathbf{f})>0$. Alluding to Lemma 3.7 supplies us with the following corollary.

COROLLARY 3.8. *Assume that $\mathfrak{B}(\mathbf{f})>R^2(d-1)2^{d-1}$ and recall the definition of Υ in (1.4). Then the following bound holds uniformly for each $\mathbf{j}\in\{0,1\}^n$ and $p>z_0$:*

$$\varpi(p^{\mathbf{j}})\ll p^{\Upsilon}.$$

§4. *Proof of Theorems 1.5 and 1.6.*

4.1. *Preparations.* Owing to (1.3), there exists positive integers $z_0 = z_0(\mathbf{f})$, $m = m(\mathbf{f})$ such that if we let

$$W := \prod_{p \leqslant z_0} p^m,$$

then there exists $\mathbf{y} \in (\mathbb{N} \cap [1, W])^n$ fulfilling the following:

$$\gcd(y_1 \cdots y_n, W) = 1 \tag{4.1}$$

and

$$p \leqslant z_0 \Rightarrow \sigma_p(\mathbf{f}(\mathbf{y} + W\mathbf{x})) > 0. \tag{4.2}$$

Define

$$\mathscr{A} := \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{f}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \equiv \mathbf{y} \pmod{W}\}. \tag{4.3}$$

Let us now choose a non-singular point $\boldsymbol{\zeta} \in V_{\mathbf{f}}(\mathbb{R})$ (whose existence is guaranteed by (1.3)) and we let $\eta \in (0, \min_i \{\min\{\zeta_i/2, (1-\zeta_i)/2\}\})$ be arbitrary. Defining

$$\mathscr{B}_\eta := \left\{\mathbf{x} \in \mathbb{R}^n : \left|\mathbf{x} - \frac{\boldsymbol{\zeta}}{2|\boldsymbol{\zeta}|}\right| < \eta\right\}, \tag{4.4}$$

we see that for any such η, one has $\mathscr{B}_\eta \subset (0, 1)^n$. Now we choose any smooth function $w : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ of compact support in $[-\eta/2, \eta/2]$ and such that if $|t| \leqslant \eta/4$ then $w(t) > 0$. Letting $w_0 := \sup\{w(t) : t \in \mathbb{R}\}$ we have $\mathbf{1}_{\{0<t\leqslant B\}}(t) \geqslant w_0^{-1} w(t/B - \zeta_i/(2|\boldsymbol{\zeta}|))$ and therefore for every $\mathbf{x} \in \mathbb{Z}^n$,

$$\prod_{i=1}^n \mathbf{1}_{\{0<x_i\leqslant B\}}(\mathbf{x}) \geqslant w_0^{-n} \prod_{i=1}^n w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right). \tag{4.5}$$

4.2. *A level of distribution result.* Let us now take the opportunity to record a *level of distribution* result that will be the main input in the forthcoming sieving arguments. For $\mathbf{k} \in \mathbb{N}^n$ with $\gcd(\widetilde{\mathbf{k}}, W) = 1$ and each k_i being square-free let $w : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ be a smooth weight as above. We let

$$N_w(B; \mathbf{k}) := \sum_{\substack{\mathbf{x} \in \mathscr{A} \\ k_i | x_i}} \prod_{i=1}^n w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right). \tag{4.6}$$

Recall the definition of the matrix $\boldsymbol{\epsilon}$ in (1.7). Our result will involve an error term related to the following function, defined for $\mathbf{m} \in \mathbb{N}^n$ and $B \geqslant 1$:

$$E(B; \mathbf{m}) := \sum_{i=1}^3 B^{-\epsilon_{i,1}} |\mathbf{m}|^{\epsilon_{i,2}} \min\{m_j\}^{\epsilon_{i,3}}.$$

Furthermore, extend the function ϖ defined in (3.5) to $\mathbb{N}^n$ by letting for $\mathbf{k} \in \mathbb{N}^n$,

$$\varpi(\mathbf{k}) := \prod_{p | k_1 \cdots k_n} \varpi(p^{\nu_p(k_1)}, \dots, p^{\nu_p(k_n)})$$

and if $\gcd(k_1\cdots k_n, W) = 1$ we define $\boldsymbol{\tau} \in (\mathbb{Z} \cap [0, W))^n$ via $\langle \boldsymbol{\tau}\mathbf{k}\rangle \equiv \mathbf{y} \pmod W$. Finally, we let

$$\mathfrak{S}(\mathbf{f}, W) := \prod_{p\mid W} \sigma_p(\mathbf{f}(\boldsymbol{\tau} + W\mathbf{s})) \prod_{p\nmid W} \sigma_p(\mathbf{f})$$

and

$$\mathscr{J}_w(\mathbf{f}, W) := \frac{1}{W^n} \int_{\mathbb{R}^R} \int_{\mathbb{R}^n} e(\boldsymbol{\gamma} \cdot \mathbf{f}(\mathbf{u})) \prod_{i=1}^n w\left(u_i - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right) d\mathbf{u}\, d\boldsymbol{\gamma}.$$

LEMMA 4.1. *Assume* $\mathfrak{B}(\mathbf{f}) > 2^{d-1}R(R+1)(d-1)$ *and that* $\mathbf{k} \in \mathbb{N}^n$ *satisfies*

$$\gcd(k_1\cdots k_n, W) = 1 \quad \text{and} \quad |\mathbf{k}| \leqslant B^{1/\rho}(\log B)^{-1},$$

where $B \in \mathbb{R}_{\geqslant 1}$ *and the constant* ρ *was defined in (1.6). Then for each* $\epsilon > 0$ *we have*

$$N_w(B; \mathbf{k}) = \mathscr{J}_w(\mathbf{f}, W)\mathfrak{S}(\mathbf{f}, W)\frac{\varpi(\mathbf{k})}{\widetilde{\mathbf{k}}}B^{n-Rd} + O\left(\frac{B^{n+\epsilon}}{\widetilde{\mathbf{k}}}E(B; \mathbf{k})\right).$$

Proof. Defining $\mathbf{g}(\mathbf{s}) := \mathbf{f}(\langle \mathbf{k}(\boldsymbol{\tau} + W\mathbf{s})\rangle)$ gives

$$N_w(B; \mathbf{k}) = \sum_{\substack{\mathbf{s}\in\mathbb{Z}^n \\ \mathbf{g}(\mathbf{s})=\mathbf{0}}} \prod_{i=1}^n w\left(\frac{s_i}{B/k_i W} - \left(\frac{\zeta_i}{2|\boldsymbol{\zeta}|} - \frac{\tau_i}{B/k_i}\right)\right).$$

We shall apply Theorem 2.1 at this point; before doing so we need to verify that

$$\left|\frac{\zeta_i}{2|\boldsymbol{\zeta}|} - \frac{\tau_i}{B/k_i}\right| \leqslant 1$$

and that condition (2.2) is met. The former is easy to verify due to $|\boldsymbol{\tau}| \leqslant W \ll 1$ and $\rho > 1$, which implies that $B/k_i \geqslant B^{1-1/\rho}(\log B) \to +\infty$. Regarding (2.2), the obvious equality $\mathbf{g}^\natural(\mathbf{s}) = W^d\mathbf{f}(\langle \mathbf{k}\mathbf{s}\rangle)$ presents us with $\max\{\|\mathbf{g}^\natural\|, \|\mathbf{g}\|\} \ll |\mathbf{k}|^d$, thus the growth condition on $|\mathbf{k}|$ in our lemma is sufficient. The last issue to be commented regards the real densities. The real density provided by the application of Theorem 2.1 is

$$\int_{\mathbb{R}^R} \int_{\mathbb{R}^n} e(W^d\boldsymbol{\beta} \cdot \mathbf{f}(\langle \mathbf{k}\mathbf{s}\rangle)) \prod_{i=1}^n w\left(\frac{s_i}{B/Wk_i} - \left(\frac{\zeta_i}{2|\boldsymbol{\zeta}|} - \frac{\tau_i}{B/k_i}\right)\right) d\mathbf{s}\, d\boldsymbol{\beta}.$$

Note that the proof of Theorem 2.1 in fact shows that the real density can also be replaced by its inhomogeneous version,

$$\int_{\mathbb{R}^R} \int_{\mathbb{R}^n} e(\boldsymbol{\beta} \cdot \mathbf{f}(\langle \mathbf{k}\boldsymbol{\tau}\rangle + W\langle \mathbf{k}\mathbf{s}\rangle)) \prod_{i=1}^n w\left(\frac{s_i}{B/Wk_i} - \left(\frac{\zeta_i}{2|\boldsymbol{\zeta}|} - \frac{\tau_i}{B/k_i}\right)\right) d\mathbf{s}\, d\boldsymbol{\beta}.$$

For this we note that the major arc analysis initially came in its inhomogeneous form, namely having $\mathbf{f}(\langle \mathbf{k}\boldsymbol{\tau}\rangle + W\langle \mathbf{k}\mathbf{s}\rangle)$ in the exponential. Moreover, by shifting the centre of the weight functions, one sees that Lemma 2.15 still applies to the inhomogeneous form and then everything stays exactly the same with regard to the error terms.

To continue the proof of our lemma we perform the linear change of variables $s_i \mapsto u_i$ and $\beta_i \mapsto \gamma_i$ given by $k_i(\tau_i + W s_i) = B u_i$, $B^d \beta_i = \gamma_i$. This leads to the following expression for the real density in our lemma:

$$\frac{B^{n-Rd}}{W^n \widetilde{\mathbf{k}}} \int_{\mathbb{R}^R} \int_{\mathbb{R}^n} e(\boldsymbol{\gamma}\cdot \mathbf{f}(\mathbf{u})) \prod_{i=1}^{n} w\left(u_i - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right) d\mathbf{u}\, d\boldsymbol{\gamma},$$

which equals $\mathscr{J}_w(\mathbf{f}, W)\widetilde{\mathbf{k}}^{-1} B^{n-Rd}$. □

The most noteworthy property of Lemma 4.1 is related to the presence of $\widetilde{\mathbf{k}}^{-1}$ in the error term; this allows us to drastically improve the level of distribution in the forthcoming applications.

4.3. *Using the Rosser–Iwaniec sieve.* By (4.5) we have the following whenever z satisfies $z_0 < z < B$:

$$\sum_{\substack{\mathbf{x}\in(\mathbb{N}\cap[-B,B])^n \\ \mathbf{f}(\mathbf{x})=\mathbf{0},\, P^-(\widetilde{\mathbf{x}})>z}} 1 \geqslant w_0^{-n} \sum_{\substack{\mathbf{x}\in\mathscr{A} \\ P^-(\widetilde{\mathbf{x}})>z}} \prod_{i=1}^{n} w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right).$$

Let us now bring into play a lower bound sieve sequence λ_k^- of dimension n. Recall the definition of θ' in (1.5). We shall make use of the terminology in [16, §11.8]; in doing so we shall call the support of λ^- by $D := B^\delta$, for some constant $\delta \in (0, \theta')$. Using $(1 * \mu)(l) \geqslant (1 * \lambda^-)(l)$ for $l = \gcd(P(z_0, z), \widetilde{\mathbf{x}})$ yields

$$\sum_{\substack{\mathbf{x}\in(\mathbb{N}\cap[-B,B])^n \\ \mathbf{f}(\mathbf{x})=\mathbf{0},\, P^-(\widetilde{\mathbf{x}})>z}} 1 \geqslant w_0^{-n} \sum_{\substack{k \mid P(z_0,z) \\ k\leqslant B^\delta}} \lambda_k^- \sum_{\substack{\mathbf{x}\in\mathscr{A} \\ k\mid\widetilde{\mathbf{x}}}} \prod_{i=1}^{n} w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right).$$

The proof of [7, Lemma 8] can be directly adapted in the setting of arbitrary dimension, thus providing the equality of the inner sum over $\mathbf{x}$ to

$$\mu(k) \sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ p\mid\widetilde{\mathbf{k}} \Leftrightarrow p\mid k}} \mu(\mathbf{k}) N_w(B; \mathbf{k}),$$

where here and throughout the rest of the paper we will use the notation

$$\mu(\mathbf{k}) := \mu(k_1)\cdots\mu(k_n).$$

A moment's thought reveals that the succeeding function is multiplicative,

$$g(k) := \mathbf{1}_{(k,W)=1}(k)\mu(k) \sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ p\mid\widetilde{\mathbf{k}} \Leftrightarrow p\mid k}} \mu(\mathbf{k})\varpi(\mathbf{k})\widetilde{\mathbf{k}}^{-1}, \tag{4.7}$$

a notation which allows us to assort our conclusions so far in the following form:

$$\sum_{\substack{\mathbf{x}\in(\mathbb{N}\cap[-B,B])^n\\ \mathbf{f}(\mathbf{x})=\mathbf{0},\, P^-(\widetilde{\mathbf{x}})>z}} 1 \gg B^{n-Rd} \sum_{\substack{k\mid P(z_0,z)\\ k\leqslant B^\delta}} \lambda_k^- g(k) + O\Bigg(B^{n+\epsilon} \sum_{k\leqslant B^\delta} |\mu(k)| \sum_{\substack{\mathbf{k}\in\mathbb{N}^n\\ p\mid\widetilde{\mathbf{k}}\Leftrightarrow p\mid k}} \frac{|\mu(\mathbf{k})|}{\widetilde{\mathbf{k}}} E(B;\mathbf{k})\Bigg).$$

In bounding the error term we will be confronted with sums of the form

$$b_k := |\mu(k)| \sum_{\substack{\mathbf{k}\in\mathbb{N}^n\\ p\mid\widetilde{\mathbf{k}}\Leftrightarrow p\mid k}} \frac{|\mu(\mathbf{k})|}{\widetilde{\mathbf{k}}} |\mathbf{k}|^{\alpha_1} \min\{k_i\}^{\alpha_2},$$

where $\alpha_i \geqslant 0$. Each $\mathbf{k}$ making a contribution to b_k satisfies $|\mathbf{k}| \leqslant k \leqslant \widetilde{\mathbf{k}}$, therefore

$$b_k \ll |\mu(k)| k^{\alpha_1+\alpha_2-1+\epsilon}.$$

We deduce that for each $1 \leqslant j \leqslant 3$, the quantity

$$B^{-\epsilon_{j,1}} \sum_{k\leqslant B^\delta} k^{\epsilon_{j,2}+\epsilon_{j,3}-1} \ll B^{-\epsilon_{j,1}+\delta(\epsilon_{j,2}+\epsilon_{j,3})}$$

becomes $\ll B^{-Rd-\epsilon'}$ for some $\epsilon' > 0$ due to $\delta < \theta'$. Therefore, we can see that for each $\delta \in (0, \theta')$ and $\epsilon > 0$ there exists $\eta = \eta(\epsilon, \delta) > 0$ such that

$$B^{n+\epsilon} \sum_{k\leqslant B^\delta} |\mu(k)| \sum_{\substack{\mathbf{k}\in\mathbb{N}^n\\ p\mid\widetilde{\mathbf{k}}\Leftrightarrow p\mid k}} \frac{|\mu(\mathbf{k})|}{\widetilde{\mathbf{k}}} E(B;\mathbf{k}) \ll B^{n-Rd-\eta}.$$

This leads to the conclusion that subject to the assertion

$$\sum_{\substack{k\mid P(z_0,z)\\ k\leqslant B^\delta}} \lambda_k^- g(k) \gg (\log B)^{-n} \tag{4.8}$$

we can establish Theorem 1.5 due to

$$\sum_{\substack{\mathbf{x}\in(\mathbb{N}\cap[-B,B])^n\\ \mathbf{f}(\mathbf{x})=\mathbf{0},\, P^-(\widetilde{\mathbf{x}})>z}} 1 \gg \frac{B^{n-Rd}}{(\log B)^n}.$$

To prove (4.8) we shall use [16, Theorem 11.12]. To this end, for any polynomials $h_i \in \mathbb{Z}[x_1, \ldots, x_n]$ we abbreviate

$$\sigma_p(p|\mathbf{h}(\mathbf{x})) := \lim_{l\to+\infty} p^{-(n-R)l} \sharp\{1 \leqslant \mathbf{x} \leqslant p^l : p^l|\mathbf{f}(\mathbf{x}),\, p|\mathbf{h}(\mathbf{x})\}.$$

LEMMA 4.2. *For each prime $p > z_0$ we have $g(p)\sigma_p = \sigma_p(p|x_1 \cdots x_n)$.*

Proof. The definition (4.7) furnishes

$$g(p)\sigma_p = \sum_{m=1}^{n} \frac{(-1)^{m-1}}{p^m} \sum_{\substack{\mathbf{j}\in\{0,1\}^n \\ |\mathbf{j}|_1=m}} \delta(\mathbf{j}),$$

thus, letting $N_{\mathbf{j}}(p^l) := \sharp\{1 \leqslant x_1, \ldots, x_l \leqslant p^l : \mathbf{f}((p^{j_i}x_i)) \equiv \mathbf{0} \pmod{p^l}\}$, we conclude that

$$g(p) \lim_{l\to+\infty} \frac{N_{\mathbf{0}}(p^l)}{p^{(n-R)l}} = \lim_{l\to+\infty} \sum_{m=1}^{n} \frac{(-1)^{m-1}}{p^m} \sum_{\substack{\mathbf{j}\in\{0,1\}^n \\ |\mathbf{j}|_1=m}} \frac{N_{\mathbf{j}}(p^l)}{p^{(n-R)l}}. \tag{4.9}$$

Obviously, if $j_i = 1$ and $y_i \equiv x_i \pmod{p^{l-1}}$ then $p^{j_i}y_i \equiv p^{j_i}x_i \pmod{p^l}$. Therefore we may split the interval $[1, p^l]$ into p subintervals of length p^{l-1} to obtain

$$\begin{aligned} N_{\mathbf{j}}(p^l) = p^{|\mathbf{j}|_1}\sharp\{&j_i = 1 \Rightarrow 1 \leqslant x_i \leqslant p^{l-1}, \\ &j_i = 0 \Rightarrow 1 \leqslant x_i \leqslant p^l : \mathbf{f}((p^{j_i}x_i)) \equiv \mathbf{0} \pmod{p^l}\}. \end{aligned}$$

One can see that this entity equals $\sharp\{\mathbf{x} \leqslant p^l : \mathbf{f}(\mathbf{x}) \equiv \mathbf{0} \pmod{p^l}, j_i = 1 \Rightarrow p|x_i\}$, hence, combining this with (4.9) yields the desired result. □

LEMMA 4.3. *There exists $\epsilon_0 \in (0, 1)$ such that one has*

$$g(p) = \frac{n}{p} + O(p^{-1-\epsilon_0}).$$

Proof. For a prime p and $t \in \mathbb{N}$ let $M(p^t) := \sharp\{1 \leqslant \mathbf{x} \leqslant p^t : p^t|\mathbf{f}(\mathbf{x}), p \nmid x_1 \cdot \ldots \cdot x_n\}$. Then [11, Lemmas 11–12] imply that there exists a positive $\epsilon_0 > 0$ such that

$$\left(1 - \frac{1}{p}\right)^{-n} \lim_{t\to\infty} p^{-t(n-R)} M(p^t) = 1 + O(p^{-1-\epsilon_0}).$$

We observe that $\lim_{t\to\infty} p^{-t(n-R)} M(p^t) = \sigma_p - \sigma_p(p|x_1 \ldots x_n)$, thus Lemma 4.2 reveals that

$$\begin{aligned} g(p) &= \frac{\sigma_p(p|x_1 \ldots x_n)}{\sigma_p} \\ &= 1 - \sigma_p^{-1} \lim_{t\to\infty} p^{-t(n-R)} M(p^t) \\ &= 1 - \left(1 - \frac{1}{p}\right)^n \frac{1}{\sigma_p} + O(\sigma_p^{-1} p^{-1-\epsilon_0}). \end{aligned}$$

The work of Birch [2] establishes the existence of a positive ϵ_1 such that $\sigma_p = 1 + O(p^{-1-\epsilon_1})$. This is sufficient for our lemma. □

Enlarging z_0 if necessary, ensures that for all primes p we have

$$0 \leqslant g(p) < 1 \quad \text{and} \quad g(p) \leqslant \frac{n}{p} + O(p^{-1-\epsilon_0}).$$

This means that one can take $\kappa = n$ in [**16**, equation (11.129)], hence our sieve problem has dimension n. By [**15**, Theorem 17.2, Proposition 17.3], the sieving limit β fulfils $\beta \leqslant 3.75n$, thus [**16**, Theorem 11.12], in combination with Lemma 4.3, guarantees the veracity of (4.8) under the condition

$$\frac{\log D}{\log z} > 3.75n.$$

This concludes the proof of Theorem 1.5.

4.4. *Using the weighted sieve.* In the last section we saw that sieving out small prime divisors of $x_1 \cdots x_n$ for integer zeros of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ gives rise to a sieve of dimension n. When the dimension of the sieve increases then the weighted sieve gives better results for the number of prime divisors in our sequence. We would like to use the weighted sieve in the form given in the Cambridge Tract of Diamond and Halberstam [**15**, Theorem 11.1], however we shall need a more flexible version of their work; one that allows the use of smooth weights. This will follow from a generalization of the weighted sieve that will be given in §4.4.1. This generalization permits the use of any suitable non-negative function rather than just a smooth weight as well as sieving in multisets.

4.4.1. *The weighted sieve with smooth weights.* We assume that $\mathscr{M}$ is any set equipped with two functions $\pi : \mathscr{M} \to \mathbb{N}$, $h : \mathscr{M} \to \mathbb{R}$ such that

$$h(\mathscr{M}) \subset [0, 1], \quad h \neq 0, \sharp\mathscr{M} < \infty. \tag{4.10}$$

For convenience of presentation we shall prefer the notation $\overline{m} = \pi(m)$. We also assume that there exists a set of primes $\mathscr{P}$, a constant $X \in \mathbb{R}_{>0}$ and a multiplicative function $\omega : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ such that, when letting

$$r_{\mathscr{M},h}(k) := \sum_{\substack{m\in\mathscr{M}\\ b|\overline{m}}} h(m) - \frac{\omega(b)}{b} X \quad (b \in \mathbb{N}),$$

there exist constants $\tau \in (0, 1]$, $\kappa \in \mathbb{N}$, $A_1 \geqslant 1$ and $A_2 \geqslant 1$ such that

$$\sum_{1\leqslant b\leqslant X^{\tau}(\log X)^{-A_1}} \mu(b)^2 4^{\nu(b)} |r_{\mathscr{M},h}(b)| \leqslant A_2 \frac{X}{(\log X)^{\kappa+1}}, \tag{4.11}$$

where the function ω enjoys the following properties for some constants $\kappa \geqslant 1$, $A > 1$:

$$0 \leqslant \omega(p) < p \ (p \in \mathscr{P}), \quad \omega(p) = 0 \ (p \notin \mathscr{P}) \tag{4.12}$$

$$\prod_{w_1\leqslant p<w} \left(1 - \frac{\omega(p)}{p}\right)^{-1} \leqslant \left(\frac{\log w}{\log w_1}\right)^{\kappa} \left(1 + \frac{A}{\log w_1}\right), \quad 2 \leqslant w_1 < w. \tag{4.13}$$

We furthermore demand that

$$m \in \mathscr{M}, \quad p|\overline{m} \Rightarrow p \in \mathscr{P}, \tag{4.14}$$

and that there exists a constant $\mu_0 > 0$ such that

$$\max\{|\overline{m}| : m \in \mathscr{M}\} \leqslant X^{\tau\mu_0}. \tag{4.15}$$

Lastly, we shall say that the property $\mathbf{Q}(u, v)$ holds for two real positive numbers $u < v$ if

$$\mathbf{Q}(u, v) : \sum_{\substack{X^{1/v} \leqslant p \leqslant X^{1/u} \\ p \in \mathscr{P}}} \sum_{\substack{m \in \mathscr{M} \\ p^2 | \overline{m}}} h(m) \ll \frac{X}{\log X} \prod_{\substack{p \in \mathscr{P} \\ p < X^{1/v}}} \left(1 - \frac{\omega(p)}{p}\right). \tag{4.16}$$

Before stating the main theorem in this section recall the definition of $f = f_\kappa$, $F = F_\kappa$ and β_κ in [**15**, Theorem 6.1] through certain differential equations. The inequality $\beta_\kappa < \nu_\kappa$ is proved for $\kappa \geqslant 200$ in [**14**, Theorem 2]; here ν_κ is the Ankeni–Onishi sieving limit [**1**] that satisfies $\nu_\kappa \sim c\kappa$ as $\kappa \to +\infty$, where

$$c = \frac{2}{e \log 2} \exp\left(\int_0^2 \frac{\mathrm{e}^u - 1}{u}\, du\right) = 2.445 \ldots.$$

In particular there exists an absolute positive constant c_0 such that $\beta_\kappa \leqslant c_0 \kappa$ for all $\kappa \geqslant 1$.

THEOREM 4.4 (Diamond–Halberstam–Richert). *Assume that $\kappa \geqslant 1$, $\mathscr{M}$, X, ω, μ_0 are as above, that each one of the conditions (4.10)–(4.15) holds, that r is a natural number satisfying $r > N(u, v; \kappa, \mu_0, \tau)$, where*

$$N(u, v; \kappa, \mu_0, \tau) := \tau\mu_0 u - 1 + \frac{\kappa}{f_\kappa(\tau v)} \int_u^v F_\kappa\left(v\left(\tau - \frac{1}{s}\right)\right)\left(1 - \frac{u}{s}\right)\frac{ds}{s}$$

and u, v satisfy $\mathbf{Q}(\mathbf{u}, \mathbf{v})$, $\tau v > \beta_\kappa$ as well as $1/\tau < u < v$. Then we have

$$\sharp\{m \in \mathscr{M}, P^-(\overline{m}) \geqslant X^{1/v}, \Omega(\overline{m}) \leqslant r\} \gg X \prod_{\substack{p \in \mathscr{P} \\ p < X^{1/v}}} \left(1 - \frac{\omega(p)}{p}\right).$$

Proof. The proof is merely a careful recast of the proof of Theorem 11.1 in [**15**, §11]. In place of the function defined in [**15**, equation (11.6)] we shall use the following function that combines the classical weights related to the weighted sieve in addition to the new weight h:

$$W_h(\mathscr{M}, \mathscr{P}, z, y, \lambda) := \sum_{\substack{m \in \mathscr{M} \\ \gcd(\overline{m}, P(z)) = 1}} h(m) \left\{\lambda - \sum_{\substack{p \in \mathscr{P}, p | \overline{m} \\ z \leqslant p < y}} \left(1 - \frac{\log p}{\log y}\right)\right\},$$

where $P(z) := \prod\{p : p \in \mathscr{P}, p < z\}$. A statement analogous to [15, equation (11.9)] can be verified once the entities $S(\mathscr{A}, \mathscr{P}, X^{1/v})$ and $S(\mathscr{A}_p, \mathscr{P}, X^{1/v})$ are replaced by

$$\sum_{\substack{m\in\mathscr{M}\\ \gcd(\overline{m}, P(X^{1/v}))=1}} h(m) \quad \text{and} \quad \sum_{\substack{m\in\mathscr{M}, p|\overline{m}\\ \gcd(\overline{m}, P(X^{1/v}))=1}} h(m)$$

respectively. The rest of the arguments in [15, §11.2] are carried automatically to our setting since, once the level of distribution result (4.11) is applied, all information regarding $\mathscr{M}$ and h is absorbed into X. The only point of departure is the use of various sieve estimates from previous chapters of the book. These sieve estimates boil down to the use of the fundamental lemma of sieve theory and the Selberg sieve, both of which can be adapted to our setting. This is due to the non-negativity of the function h, which allows various combinatorial inequalities to be adapted once multiplied by h. One example of this is in the case of an upper bound sieve, say λ^+: opening up the convolution in the right side of $(1 * \mu) \leqslant (1 * \lambda^+)$ gives

$$\sum_{\substack{m\in\mathscr{M}\\ \gcd(\overline{m}, P(z))=1}} h(m) \leqslant \sum_{k|P(z)} \lambda_k^+ \sum_{\substack{m\in\mathscr{M}\\ k|\overline{m}}} h(m),$$

and one can now use (4.11) to absorb $\mathscr{M}$ and h in X for the rest of the argument.

For the proof of the present theorem it remains to adapt the arguments in [15, §11.3]. First, the contribution towards $\sum_m h(m)$ of those $m \in \mathscr{M}$ such that $\overline{m}$ is divisible by the square of a prime $p \in \mathscr{P}$ in the range $X^{1/v} \leqslant p \leqslant X^{1/u}$ can be safely ignored due to condition (4.16). An inspection of [15, §11] reveals that condition $\mathbf{Q_0}$ in [15, equation (11.2)] is used in the proof of [15, Theorem 11.1] only to deal with this particular sum over primes in $\mathscr{P} \cap [X^{1/v}, X^{1/u}]$. We are thus free to focus our attention exclusively on the contribution of the elements of the set

$$\mathscr{M}' := \{m' \in \mathscr{M} : \text{ there is no prime } p \in \mathscr{P} \cap [X^{1/v}, X^{1/u}] \text{ such that } p^2 \mid \overline{m}'\}.$$

The last inequality in [15, p. 140] becomes

$$\sum_{\substack{X^{1/v}\leqslant p<X^{1/u}\\ p\in\mathscr{P}, p|\overline{m}'}} \left(1 - \frac{u \log p}{\log X}\right) \geqslant \Omega(\overline{m}') - \frac{u \log |\overline{m}'|}{\log X},$$

which, when multiplied by $h(m')$, gives, as in [15, p. 141],

$$W_h(\mathscr{M}', \mathscr{P}, z, y, \lambda) \leqslant (r+1) \sum_{\substack{m'\in\mathscr{M}', \Omega(\overline{m}')\leqslant r\\ \gcd(\overline{m}', P(X^{1/v}))=1}} h(m')$$

for the choice of λ and r made in [15, p. 141]. The property $h(\mathscr{M}) \subset [0,1]$ shows that

$$\begin{aligned} &\sharp\{m \in \mathscr{M} : P^-(\overline{m}) \geqslant X^{1/v}, \Omega(\overline{m}) \leqslant r\} \\ &\quad \geqslant \sharp\{m' \in \mathscr{M}' : P^-(\overline{m}') \geqslant X^{1/v}, \Omega(\overline{m}') \leqslant r\} \\ &\quad \geqslant \sum_{\substack{m' \in \mathscr{M}, \Omega(\overline{m}') \leqslant r \\ \gcd(\overline{m}', P(X^{1/v}))=1}} h(m') \geqslant \frac{1}{r+1} W_h(\mathscr{M}', \mathscr{P}, z, y, \lambda), \end{aligned}$$

which allows the rest of the proof of [15, Theorem 11.1] to be adapted to our case. Finally, the choice of the constants v and r given in our theorem is borrowed from the inequalities succeeding [15, equation (11.22)]. □

Remark 4.5. The setting of Theorem 4.4 includes that of [15, Theorem 1.1]; indeed, one can choose $(\mathscr{M}, \pi, h) = (\mathscr{A}, \mathrm{id}, 1)$.

Remark 4.6. In most cases it is easy to verify $\mathbf{Q}(u, v)$ for all $u, v > 0$, however this is not the case for the problem of prime factors of $x_1 \cdots x_n$ for integer solutions $\mathbf{x} = (x_1, \ldots, x_n)$ of general Diophantine equations, since, as explained in §3, quite often a prime could divide two coordinates of $\mathbf{x}$.

Remark 4.7. A table of estimates for β_κ for $1 \leqslant \kappa \leqslant 10$ is given in [15, p. 227]. Furthermore, [15, equation (11.21)] contains estimates for r that are slightly weaker but simpler than that of [15, Theorem 11.1]. For example, the choice $\xi = \beta_\kappa$ in [15, equation (11.21)] shows that, as long as $\mathbf{Q}((2\beta_\kappa - 1)/\tau\beta_\kappa, (2\beta_\kappa - 1)/\tau)$ holds, then the conclusion of Theorem 4.4 remains valid with $v = (2\beta_\kappa - 1)/\tau$ and for all natural numbers r satisfying

$$r \geqslant \mu_0 - 1 + (\mu_0 - \kappa)(1 - 1/\beta_\kappa) + (\kappa + 1)\log \beta_\kappa. \tag{4.17}$$

In fact [15, equation (11.21)] with $\xi = \beta_\kappa$ shows that if $\mathbf{Q}(u, v)$ holds for some $u > 1/\tau$ and any $v > u$, then letting

$$v' := \frac{\beta_\kappa - 1}{\tau - 1/u}$$

we deduce that the conclusion of Theorem 4.4 still holds with any r satisfying

$$r \geqslant \tau\mu_0 u - 1 + \left(\kappa + \frac{u}{v'}\beta_\kappa\right)\log\frac{v'}{u} - \kappa\left(1 - \frac{u}{v'}\right). \tag{4.18}$$

To prove Theorem 1.6 we take

$$\mathscr{M} := \{\mathbf{x} \in \mathbb{N}^n : \mathbf{f}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \equiv \mathbf{y} \ (\mathrm{mod}\ W), |\mathbf{x}| \leqslant B\}, \pi(\mathbf{x}) := \widetilde{\mathbf{x}},$$

and we let

$$h(\mathbf{x}) := \prod_{i=1}^n w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right).$$

Then for $\mathscr{P}$ being the set of all primes $p > z_0$, g as in (4.7), θ' as in (1.5) and any $0 < \epsilon < \theta'$ we can verify all conditions (4.11)–(4.15) with

$$X := \mathscr{J}_w(\mathbf{f}, W)\mathfrak{S}(\mathbf{f}, W)B^{n-Rd}, \qquad \omega(b) := bg(b),$$
$$\kappa := n, \qquad \tau := \theta' - \epsilon, \qquad \mu_0 = \frac{n}{n-Rd}\frac{1+\epsilon}{\theta'-\epsilon}$$

with an argument that is identical to that in §4.3. It remains to check condition $\mathbf{Q}(u, v)$ and for this we note that in our setting, the sum in (4.16) is at most

$$\sum_{X^{1/v}<p\leqslant X^{1/u}} \sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ \widetilde{\mathbf{k}}=p^2}} \sum_{\substack{\mathbf{x}\in\mathscr{A} \\ k_i\mid x_i}} \prod_{i=1}^n w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right).$$

Invoking Lemma 4.1 we see that, if $u > 2(n-Rd)\rho$, where ρ is defined in (1.6), this is

$$\ll B^{n-Rd}\Bigg(\sum_{X^{1/v}\leqslant p\leqslant X^{1/u}} p^{-2}\sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ \widetilde{\mathbf{k}}=p^2}} \varpi(\mathbf{k})\Bigg)$$
$$+ B^{n+\epsilon}\Bigg(\sum_{X^{1/v}<p\leqslant X^{1/u}} p^{-2}\sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ \widetilde{\mathbf{k}}=p^2}} E(B;\mathbf{k})\Bigg).$$

Assuming $\max\{(d^2-1)R^2 2^{d-1}, (d-1)R^2 2^{d-1}+2(R+1)\} < \mathfrak{B}(\mathbf{f})$, we obtain via Corollary 3.6 that the first sum over $\mathbf{k}$ above is $\ll 1$, thus, when $v > 0$, the first term contributes

$$\ll B^{n-Rd-(n-Rd)/v} \ll \frac{B^{n-Rd}}{(\log B)^n}.$$

It remains to verify that there exists $\epsilon' > 0$ such that

$$\sum_{X^{1/v}<p\leqslant X^{1/u}} p^{-2}\sum_{\substack{\mathbf{k}\in\mathbb{N}^n \\ \widetilde{\mathbf{k}}=p^2}} E(B;\mathbf{k}) \ll B^{-\epsilon'-Rd}. \tag{4.19}$$

For this we note that each $\epsilon_{i,2}$ is at least $1/2$ owing to $K \geqslant \max\{Rd, R^2(d-1)\}$ and $K \geqslant 1$. Thus the error term above becomes

$$\ll \sum_{i=1}^3 B^{-\epsilon_{i,1}} \sum_{X^{1/v}<p\leqslant X^{1/u}} p^{-2+2\epsilon_{i,2}} \ll \sum_{i=1}^3 B^{-\epsilon_{i,1}+((n-Rd)/u)(2\epsilon_{i,2}-1)}.$$

Therefore, if

$$u > \max\left\{\frac{(n-Rd)(2\epsilon_{i,2}-1)}{\epsilon_{i,1}-Rd} : 1 \leqslant i \leqslant 3\right\}$$

then (4.19) holds. Now define $u_0 := (1+\epsilon)\max\{u_1, 1/(\theta'-\epsilon), 2(n-Rd)\rho\}$, where

$$u_1 := \max\left\{\frac{(n-Rd)(2\epsilon_{i,2}-1)}{\epsilon_{i,1}-Rd} : 1 \leqslant i \leqslant 3\right\}.$$

Then applying (4.18) with $u = u_0$ and $v' := (nc_n - 1)/(\theta' - \epsilon - 1/u_0)$, allows us to take

$$r \geqslant \frac{n}{n-Rd}(1+\epsilon)u_0 - 1 + n\left(1 + \frac{u_0}{v'}c_n\right)\log\frac{v'}{u_0} - n\left(1 - \frac{u_0}{v'}\right),$$

where $c_n := \beta_n/n$ satisfies $\lim_{n\to+\infty} c_n = 2.44\ldots$. Letting $\epsilon > 0$ be arbitrarily close to zero concludes the proof of the lower bound claimed in Theorem 1.6. This is because the quantities u'', $\widehat{u}$, $\widehat{v}$ introduced in (1.8) and (1.9) are such that for fixed $\mathbf{f}$, n, d, R we have

$$\lim_{\epsilon\to 0}(u_1, u_0, v) = (u'', \widehat{u}, \widehat{v}).$$

To complete the proof of Theorem 1.6 it remains to verify the estimates regarding $\widehat{v}$ and r_0, where r_0 is defined in (1.10). It is easy to see that $u''/(n-Rd)$ is a function of K that is bounded away from 0 and $+\infty$, while a similar remark applies to ρ and θ'. This implies that $\widehat{u} \ll_{d,R} n$ and noting that $\widehat{u} < \widehat{v}$, one has

$$r_0 \ll_{d,R} \widehat{u} + n\log\frac{\widehat{v}}{\widehat{u}} \ll_{d,R} n\left(1 + \log\frac{\widehat{v}}{\widehat{u}}\right),$$

where the implied constant is independent of K and n. The identity $nc_n - 1 = \tau\widehat{v} - \widehat{v}/\widehat{u}$ shows that

$$\widehat{v}/\widehat{u} \ll n + \widehat{v} \ll n + \frac{n}{\theta' - 1/\widehat{u}} \ll n,$$

therefore $r_0 = O_{d,R}(n\log n)$, with an implied constant depending at most on d and R.

§5. *Multidimensional vector sieve.* The next lemma constitutes a generalization of the vector sieve of Brüdern and Fouvry [7] to arbitrarily many variables.

LEMMA 5.1 (Multidimensional vector sieve). *Let $n \in \mathbb{N}$ and assume that we are given 2 sequences λ_i^-, λ_i^+, $(i = 1, \ldots, n)$ such that for each $m \in \mathbb{N}$ and $1 \leqslant i \leqslant n$ we have*

$$(1 * \lambda_i^-)(m) \leqslant (1 * \mu)(m) \leqslant (1 * \lambda_i^+)(m). \tag{5.1}$$

Then the following inequality holds for each $\mathbf{m} \in \mathbb{N}^n$:

$$\prod_{i=1}^{n}(1*\mu)(m_i) \geqslant \sum_{i=1}^{n}(1*\lambda_i^-)(m_i)\prod_{\substack{1\leqslant j\leqslant n\\ j\neq i}}(1*\lambda_j^+)(m_j) - (n-1)\prod_{i=1}^{n}(1*\lambda_i^+)(m_i).$$

Proof. In light of (5.1) it is sufficient to verify

$$\prod_{i=1}^{n}(1*\mu)(m_i) \geqslant -(n-1)\prod_{i=1}^{n}(1*\lambda_i^+)(m_i)+\sum_{i=1}^{n}(1*\mu)(m_i)\prod_{\substack{1\leqslant j\leqslant n\\ j\neq i}}(1*\lambda_j^+)(m_j). \tag{5.2}$$

If $m_i = 1$ for all $i = 1, \ldots, n$ then $(1 * \lambda_i^+)(m_i) \geqslant 1$, thus the entities $x_i := 1/(1 * \lambda_i^+)(m_i)$ fulfil $0 < x_i \leqslant 1$. The inequality (5.2) becomes $x_1 \cdots x_n \geqslant -n + 1 + (x_1 + \cdots + x_n)$. Letting $A_i = 1 - x_i$ the last inequality becomes $(1 - A_1) \cdots (1 - A_n) \geqslant 1 - (A_1 + \cdots + A_n)$, which is the Weierstrass product inequality, see [**26**, equation (1)]. In the remaining case where there exists i with $m_i \neq 1$ we can assume that $(1 * \lambda_i^+)(m_i) \neq 0$ for each such i, for otherwise both sides of (5.2) vanish. We may now introduce for each $1 \leqslant i \leqslant n$ the variables $x_i := 1/(1 * \lambda_i^+)(m_i)$; then (5.2) becomes

$$n - 1 \geqslant \sum_{\substack{1\leqslant i\leqslant n\\ m_i=1}} x_i.$$

The proof is concluded upon observing that the condition $m_i = 1$ implies $x_i \leqslant 1$. □

Our aim now becomes to prove a version of the fundamental lemma of sieve theory in the context of prime divisors of coordinates of integer zeros in varieties. The exact form is given in Proposition 5.5 and the rest of this section is devoted to its proof. The quantity under consideration is the weighted density of vectors $\mathbf{x} \in \mathscr{A}$ with $|\mathbf{x}| \leqslant B$ such that $\widetilde{\mathbf{x}}$ does not have prime divisors in the range $p \leqslant z_1$ for any z_1 with $z_0 < z_1 \leqslant B$. We prefer to keep the choice of z_1 unspecified in this section and we shall only need the value $z_1 = (\log B)^A$ for $A > 0$ independent of B in §6.

For $\mathbf{k} \in \mathbb{N}^n$ and $y_1, y_2 \in \mathbb{R}$ with $y_1 < y_2$ we define

$$\mu(\mathbf{k}) := \prod_{i=1}^{n}\mu(k_i) \quad \text{and} \quad P(y_1, y_2) := \prod_{y_1<p\leqslant y_2} p.$$

For a smooth function $w : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ that is as in §4.1, any $z_1 > z_0$ and any $\mathbf{l} \in \mathbb{N}^n$ we let

$$G(B, z_1; \mathbf{l}) := \sum_{\substack{\mathbf{x}\in\mathscr{A}, l_i|x_i,\\ p|x_1\cdots x_n\Rightarrow p>z_1}} \prod_{i=1}^{n} w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right). \tag{5.3}$$

We are interested in estimating $G(B, z_1; \mathbf{l})$ whenever $\mathbf{l} \in \mathbb{N}^n$ fulfils $l_i | P(z_1, z)$, where z is any constant satisfying $z > z_1$. This is analogous to [**7**, Proposition, p. 83] and we shall also begin by proving the upper bound. We shall use the upper and lower bound sieves, λ^+ and λ^-, as defined at the bottom of [**7**, p. 84].

Assume that λ^+ is an upper bound sieve supported in $[1, D_1]$ and note that the condition $\mathbf{x} \equiv \mathbf{y} \pmod{W}$ ensures that $p \nmid \widetilde{\mathbf{x}}$ for all $p \leqslant z_0$. Recalling definition (4.6) we see that whenever $l_i | P(z_1, z)$ then

$$G(B, z_1; \mathbf{l}) \leqslant \sum_{\substack{\mathbf{k} \in \mathbb{N}^n \\ k_i | P(z_0, z_1)}} N_w(B; (k_1 l_1, \ldots, k_n l_n)) \prod_{i=1}^{n} \lambda^+_{k_i}.$$

Note that all $\mathbf{k}$ and $\mathbf{l}$ above must satisfy

$$\gcd(\widetilde{\mathbf{k}}, \widetilde{\mathbf{l}}) = 1 = \gcd\left(\widetilde{\mathbf{k}} \quad \widetilde{\mathbf{l}}, \prod_{p \leqslant z_0} p\right), \qquad \mu(k_i)^2 = 1 = \mu(l_i)^2.$$

Recall definition (1.6) and assume that

$$|\mathbf{l}| \leqslant \frac{B^{1/\rho}}{D_1 \log B}. \tag{5.4}$$

Then Lemma 4.1 shows that if $K > R(R+1)(d-1)$ and (5.4) holds then

$$N_w(B; (k_1 l_1, \ldots, k_n l_n)) = \frac{\varpi(\mathbf{k})}{\widetilde{\mathbf{k}}} X_{\mathbf{l}} + O\left(\frac{B^{n+\epsilon}}{\widetilde{\mathbf{l}}} \frac{E(B; (k_1 l_1, \ldots, k_n l_n))}{\widetilde{\mathbf{k}}}\right),$$

where

$$X_{\mathbf{l}} := \mathfrak{S}(\mathbf{f}) \mathscr{J}_w(\mathbf{f}, W) \frac{\varpi(\mathbf{l})}{\widetilde{\mathbf{l}}} B^{n-Rd}.$$

We may now set

$$\Sigma(D_1, z_1) = \sum_{\substack{\mathbf{k} \in \mathbb{N}^n \\ k_i | P(z_0, z_1)}} \frac{\varpi(\mathbf{k})}{\widetilde{\mathbf{k}}} \prod_{i=1}^{n} \lambda^+_{k_i}$$

to obtain

$$\begin{aligned} G(B, z_1; \mathbf{l}) &\leqslant \Sigma(D_1, z_1) X_{\mathbf{l}} \\ &\quad + O\left(\frac{B^{n+\epsilon}}{\widetilde{\mathbf{l}}} \sum_{\substack{|\mathbf{k}| \leqslant D_1 \\ p | \widetilde{\mathbf{k}} \Rightarrow z_0 < p \leqslant z_1}} \frac{\mu(\mathbf{k})^2}{\widetilde{\mathbf{k}}} E(B; (k_1 l_1, \ldots, k_n l_n))\right). \end{aligned} \tag{5.5}$$

5.1. *Bounds for ϖ.* One has to be careful when adapting the approach [7] to homogeneous equations. The reason is that in the case of Lagrange's equation there exists a multiplicative function $\widetilde{\varpi}$ satisfying

$$\varpi(\mathbf{m}) \leqslant \prod_{i=1}^{n} \widetilde{\varpi}(m_i)$$

and such that for all large primes p one has $\widetilde{\varpi}(p) \leqslant 2$, see [7, Lemma 12, part (iii)]. It is easy to see that bounds of this quality fail to hold rather spectacularly for systems of forms $\mathbf{f} = \mathbf{0}$ as in Theorem 1.1. Indeed,

$$\begin{aligned}
&\varpi(p, \ldots, p)\\
&\quad = \sigma_p^{-1} \lim_{l\to\infty} p^{-l(n-R)} \sharp\{\mathbf{x} \ (\mathrm{mod}\ p^l) : \mathbf{f}(p\mathbf{x}) \equiv 0 \ (\mathrm{mod}\ p^l)\}\\
&\quad = p^{Rd}\sigma_p^{-1} \lim_{l\to\infty} p^{-(l-d)(n-R)} \sharp\{\mathbf{x} \ (\mathrm{mod}\ p^{l-d}) : \mathbf{f}(\mathbf{x}) \equiv 0 \ (\mathrm{mod}\ p^{l-d})\}\\
&\quad = p^{Rd}.
\end{aligned}$$

To confront this issue our first task is to control the contribution towards $\Sigma(D_1, z_1)$ of integer vectors $\mathbf{k}$ such that there exists $i < j$ with $k_{ij} := \gcd(k_i, k_j)$ attaining a large value. Define

$$\Sigma^*(D_1, z_1) = \sum_{\substack{\mathbf{k}\in\mathbb{N}^n\\ k_i \mid P(z_0,z_1)\\ \max k_{ij}\leqslant\Delta}} \frac{\varpi(\mathbf{k})}{\widetilde{\mathbf{k}}} \prod_{i=1}^{n} \lambda_{k_i}^+$$

and recall the definition of Υ in (1.4).

LEMMA 5.2. *Assuming* $\max\{(d-1)R^2 2^{d-1} + (R+1)(\Upsilon+1), (d^2-1)R^2 2^{d-1}\} < \mathfrak{B}(\mathbf{f})$, *one has*

$$\Sigma(D_1, z_1) - \Sigma^*(D_1, z_1) \ll \Delta^{-1+\epsilon} (\log z_1)^n.$$

Proof. The quantity under investigation is $\ll \sum_{1\leqslant l_1<l_2\leqslant n} \mathscr{E}(l_1, l_2)$, where

$$\mathscr{E}(l_1, l_2) := \sum_{\delta>\Delta} \mu(\delta)^2 \sum_{\substack{k_i \mid P(z_0,z_1)\\ \delta\mid k_{l_1}, \delta\mid k_{l_2}}} \frac{\varpi(\mathbf{k})}{\widetilde{\mathbf{k}}}.$$

We may now use the multiplicative properties of ϖ to deduce that

$$\mathscr{E}(l_1, l_2) \ll \sum_{\delta>\Delta} \Bigg(\prod_{\substack{z_0<p\leqslant z_1\\ p\mid\delta}} \sum_{\substack{\mathbf{j}\in\{0,1\}^n\\ j_{l_1}=j_{l_2}=1}} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} \Bigg) \Bigg(\prod_{\substack{z_0<p\leqslant z_1\\ p\nmid\delta}} \sum_{\mathbf{j}\in\{0,1\}^n} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} \Bigg).$$

Fix $\eta \in (0, 1/4)$ and let us denote $s_0 := \Upsilon + 1 + \eta$. By Corollary 3.8 we obtain

$$\sum_{|\mathbf{j}|_1\geqslant s_0} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} \ll p^{\Upsilon} \sum_{s_0\leqslant s\leqslant n} \binom{n}{s} p^{-s} \ll p^{-1-\eta}.$$

The assumptions of our lemma allow us to apply Corollary 3.6 whenever $|\mathbf{j}|_1 \leqslant s_0$. Thus it supplies us with some $\lambda > 0$ such that $\varpi(p^{\mathbf{j}}) = 1 + O(p^{-\lambda})$, which yields

$$\sum_{\mathbf{j}\in\{0,1\}^n} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} = 1 + \frac{n}{p} + O(p^{-1-\epsilon})$$

and

$$\sum_{\substack{\mathbf{j}\in\{0,1\}^n\\ j_{l_1}=j_{l_2}=1}} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} = p^{-2} + O(p^{-2-\epsilon}),$$

for some $\epsilon > 0$. Assorting all related estimates we obtain for square-free δ that

$$\prod_{\substack{z_0 < p \leqslant z_1 \\ p \mid \delta}} \sum_{\substack{\mathbf{j} \in \{0,1\}^n \\ j_{l_1} = j_{l_2} = 1}} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} \ll \delta^{-2+\epsilon},$$

and

$$\prod_{\substack{z_0 < p \leqslant z_1 \\ p \nmid \delta}} \sum_{\mathbf{j} \in \{0,1\}^n} \frac{\varpi(p^{\mathbf{j}})}{p^{|\mathbf{j}|_1}} \ll \prod_{z_0 < p \leqslant z_1} \left(1 + \frac{n}{p} + O(p^{-1-\epsilon})\right) \ll (\log z_1)^n.$$

These estimates prove that $\mathscr{E}(l_1, l_2) \ll (\log z_1)^n \sum_{\delta > \Delta} \delta^{-2+\epsilon}$, which is sufficient. □

For any square-free integer m and index $1 \leqslant i \leqslant n$ define

$$\varpi_i(m) := \varpi(1, \ldots, 1, m, 1, \ldots, 1),$$

where m appears in the ith position. For $\epsilon > 0$ define the multiplicative function

$$\phi_\epsilon(m) := \prod_{\substack{p \mid m \\ p > z_0}} (1 + p^{-\epsilon}).$$

Note that if assumptions of Corollary 3.6 hold for $|\mathbf{j}|_1 = 1$ then there exists $\epsilon = \epsilon(\mathbf{f}) > 0$ such that $\sigma_p(p^{\mathbf{e}_i} | \mathbf{x}) = 1/p + O(p^{-1-\epsilon})$. Enlarging z_0 and replacing ϵ by a smaller positive constant if needed yields the following result.

LEMMA 5.3. *Assume that* $\mathfrak{B}(\mathbf{f}) > \max\{R^2 2^{d-1}(d^2 - 1), R^2 2^{d-1}(d - 1) + (R + 1)\}$. *Then there exists* $\epsilon = \epsilon(\mathbf{f}) > 0$ *such that for all square-free integers m,*

$$\max_{1 \leqslant i \leqslant n} \varpi_i(m) \leqslant \phi_\epsilon(m).$$

Observe that for all $\mathbf{d} \in \mathbb{N}^n$ with $\mu(\mathbf{d})^2 = 1$ the expression

$$\frac{\varpi(\mathbf{d})}{\prod_{i=1}^n \varpi_i(d_i)}$$

is a function of the vector $(\gcd(d_i, d_j))_{1 \leqslant i < j \leqslant n}$. To see this, it is enough to consider the case when $\widetilde{\mathbf{d}}$ is divisible by a single prime, say p. We need to show that if $\mathbf{h}, \mathbf{k} \in \{0, 1\}^n$ and

$$i \neq j \Rightarrow \min(k_i, k_j) = \min(h_i, h_j) \tag{5.6}$$

then

$$\frac{\varpi(p^{\mathbf{k}})}{\prod_{i=1}^n \varpi_i(p^{k_i})} = \frac{\varpi(p^{\mathbf{h}})}{\prod_{i=1}^n \varpi_i(p^{h_i})}. \tag{5.7}$$

Obviously this holds in the case that $\mathbf{k} = \mathbf{h}$ and we can therefore assume that $\mathbf{k} \neq \mathbf{h}$. A little thought reveals that in this case (5.6) guarantees that there exist $l, m, i \neq j$ such that $(\mathbf{k}, \mathbf{h})$ equals one of the following:

$$(\boldsymbol{e}_l, \mathbf{0}), (\mathbf{0}, \boldsymbol{e}_m), (\boldsymbol{e}_i, \boldsymbol{e}_j).$$

For any such instance we can verify that both sides of (5.7) equal 1, hence our claim holds. We have proved that there exists a function $\widehat{g} : \mathbb{N}^{\binom{n}{2}} \to \mathbb{R}_{\geqslant 0}$ such that

$$\mu(\mathbf{d})^2 = 1 \Rightarrow \varpi(\mathbf{d}) = \widehat{g}((d_{i,j})) \prod_{i=1}^{n} \varpi_i(d_i).$$

The function $\varpi_i(d_i)$ keeps track of the probability that $d_i | x_i$ and the function $\widehat{g}((d_{i,j}))$ takes values close to 1 when the events $d_i | x_i$ are independent (in a suitable sense) but can obtain larger values in general.

Defining

$$S((u_{i,j})) := \sum_{\substack{\mathbf{k} \in \mathbb{N}^n \\ k_i | P(z_0, z_1) \\ (k_i, k_j) = u_{i,j}}} \prod_{i=1}^{n} \frac{\lambda_{k_i}^{+} \varpi_i(k_i)}{k_i}$$

enables us to write

$$\Sigma^*(D_1, z_1) = \sum_{\substack{u_{i,j} \leqslant \Delta \\ 1 \leqslant i < j \leqslant n}} \widehat{g}((u_{i,j})) S((u_{i,j})). \tag{5.8}$$

We may now use the expression $(\mu * 1)((k_i/u_{i,j}, k_j/u_{i,j}))$ to detect the condition $(k_i, k_j) = u_{i,j}$, thus inferring

$$S((u_{i,j})) = \sum_{\substack{(l_{i,j}) \in \mathbb{N}^{\binom{n}{2}} \\ 1 \leqslant i \neq j \leqslant n \\ u_{i,j} l_{i,j} | P(z_0, z_1)}} \mu(\mathbf{l}) \prod_{i=1}^{n} \Bigg(\sum_{\substack{k \in \mathbb{N} \\ k | P(z_0, z_1) \\ \xi_i | k}} \frac{\lambda_k^{+} \varpi_i(k)}{k} \Bigg), \tag{5.9}$$

where

$$\xi_i := \operatorname{rad}\Bigg(\prod_{\substack{1 \leqslant j \leqslant n \\ j \neq i}} u_{i,j} l_{i,j} \Bigg)$$

and rad stands for the radical of a positive integer. Under the assumptions of Lemma 5.3 we thus obtain the following estimate for all square-free integers δ:

$$\Bigg| \sum_{\substack{k \in \mathbb{N} \\ k | P(z_0, z_1) \\ \delta | k}} \frac{\lambda_k^{+} \varpi_i(k)}{k} \Bigg| \leqslant \frac{\phi_\epsilon(\delta)}{\delta} \prod_{z_0 < p \leqslant z_1} (1 + p^{-1} + p^{-1-\epsilon}) \ll \frac{\phi_\epsilon(\delta)}{\delta} \log z_1.$$

Note that the succeeding inequality holds for all divisors m' of m,

$$\frac{\phi_\epsilon(m)}{m} \leqslant \frac{\phi_\epsilon(m')}{m'}.$$

Letting ξ_i^* be the radical of $\prod_{j\neq i} l_{i,j}$ and using the last inequalities with $\delta = m = \xi_i$ and $m' = \xi_i^*$ allows us to truncate the sum in (5.9) to the range $l_{i,j} \leqslant \Delta^{B_1}$, where $B_1 > 0$ is a constant that will be chosen in due course. The contribution of $l_{1,2} > \Delta^{B_1}$ is

$$\ll (\log z_1)^n \sum_{\substack{l_{i,j}\leqslant D_1, l_{i,j}\mid P(z_0,z_1)\\ l_{1,2}>\Delta^{B_1}}} \frac{\mu(\mathbf{l})^2}{\widetilde{\boldsymbol{\xi}}^*} \prod_{i=1}^n \phi_\epsilon(\xi_i^*),$$

where D_1 is the support of λ^+.

We may now use the inequality

$$\phi_\epsilon(\xi_i^*) \leqslant \prod_{j\neq i} \phi_\epsilon(l_{i,j})$$

to obtain

$$\prod_{i=1}^n \phi_\epsilon(\xi_i^*) \leqslant \prod_{1\leqslant i\neq j\leqslant n} \phi_\epsilon(l_{i,j})^2.$$

Hence the last sum is

$$\leqslant \sum_{l_{1,2}>\Delta^{B_1}} \frac{\mu(\mathbf{l})^2}{\xi_1^*\cdots\xi_n^*} \prod_{1\leqslant i\neq j\leqslant n} \phi_\epsilon(l_{i,j})^2.$$

This is really a summation over the variables $l_{1,2},\ldots,l_{n-1,n}$ because each expression ξ_i^* is a function of some of these variables. We first perform a summation over $l_{n-1,n}$. Recalling that

$$\xi_i^* = \mathrm{rad}\bigg(\prod_{j\neq i} l_{i,j}\bigg)$$

we see that only ξ_{n-1}^* and ξ_n^* depend on $l_{n-1,n}$, since they satisfy

$$\xi_n^* = [l_{n-1,n}, \xi_n^{**}], \qquad \xi_{n-1}^* = [l_{n-1,n}, \xi_{n-1}^{**}],$$

where both ξ_{n-1}^{**} and ξ_n^{**} are defined as ξ_{n-1}^* and ξ_n^* but with the variable $l_{n-1,n}$ missing, i.e.

$$\xi_{n-1}^{**} := \mathrm{rad}\bigg(\prod_{j\neq n-1,n} l_{n-1,j}\bigg), \qquad \xi_n^{**} := \mathrm{rad}\bigg(\prod_{j\neq n-1,n} l_{n,j}\bigg).$$

Hence the sum over $l_{n-1,n}$ is

$$\sum_{l_{n-1,n}} \frac{\mu(l_{n-1,n})^2 \phi_\epsilon(l_{n-1,n})^2}{[l_{n-1,n}, \xi^{**}_{n-1}][l_{n-1,n}, \xi^{**}_n]},$$

which equals

$$\frac{1}{\xi^{**}_{n-1}\xi^{**}_n} \sum_{l_{n-1,n}} \frac{\mu(l_{n-1,n})^2 \phi_\epsilon(l_{n-1,n})^2}{l^2_{n-1,n}} \gcd(\xi^{**}_{n-1}, l_{n-1,n}) \gcd(\xi^{**}_n, l_{n-1,n}).$$

The last sum is

$$\leqslant \prod_{p \mid \xi^{**}_{n-1}\xi^{**}_n} (2 + 2p^{-\epsilon} + p^{-2\epsilon}) \prod_p (1 + p^{-2}(1 + p^{-\epsilon})^2) \ll \tau(\xi^{**}_{n-1})^A \tau(\xi^{**}_n)^A,$$

where $A = 3$. Of course we can bound any ξ^{**}_k by the product of all available variables except $l_{n-1,n}$, i.e. $\prod_{\{i,j\} \neq \{n-1,n\}} l_{i,j}$, thus we obtain

$$\ll \frac{1}{\xi^{**}_{n-1}\xi^{**}_n} \prod_{\{i,j\} \neq \{n-1,n\}} \tau(l_{i,j})^{2A}.$$

The process above is the first step of a finite induction that eliminates all variables $l_{i,j}$, beginning from $l_{n-1,n}$ and terminating with $l_{1,2}$. At each step expressions of the form

$$\sum_{l_{1,2} > \Delta^{B_1}} \frac{\mu(\mathbf{l})^2}{\xi'_1 \cdots \xi'_n} \prod^{\flat}_{1 \leqslant i \neq j \leqslant n} \tau(l_{i,j})^A$$

are bounded by

$$\ll \sum_{l_{1,2} > \Delta^{B_1}} \frac{\mu(\mathbf{l})^2}{\xi''_1 \cdots \xi''_n} \prod^{\flat\flat}_{1 \leqslant i \neq j \leqslant n} \tau(l_{i,j})^{100A},$$

where the notation $\xi', \prod^\flat$ means that some of the variables $l_{i,j}$ have been eliminated, the notation $\xi'', \prod^{\flat\flat}$ that one further variable has been eliminated and the constant A' depends at most on A and $\mathbf{f}$. At the last step of the induction we will arrive at the expression

$$\sum_{l_{1,2} > \Delta^{B_1}} \frac{\mu(l_{1,2})^2}{l^2_{1,2}} \tau(l_{1,2})^C,$$

where $C = C(\mathbf{f})$. Obviously this is $\ll \Delta^{-B_1/2}$. The arguments above show that

$$S((u_{i,j})) = \sum_{\substack{(l_{i,j}) \in \mathbb{N}^{\binom{n}{2}} \\ l_{i,j} \leqslant \Delta^{B_1} \\ u_{i,j} l_{i,j} \mid P(z_0, z_1)}} \mu(\mathbf{l}) \prod_{i=1}^n \left(\sum_{\substack{k \in \mathbb{N} \\ k \mid P(z_0,z_1) \\ \xi_i \mid k}} \frac{\lambda^+_k \varpi_i(k)}{k} \right) + O((\log z_1)^n \Delta^{-B_1/2}), \tag{5.10}$$

where the implied constant is independent of the $u_{i,j}$.

We now aim to use a consequence of the linear case of the Rosser–Iwaniec sieve (in fact the linear case was settled first by Jurkat and Richert [25]) that is given in [7, Lemma 11]. We shall find it convenient to use the error term appearing in [24, Theorem 1], this will lead us to replace the term $\mathrm{e}^{\sqrt{L-s}}(\log D)^{-1/3}$ in [7, Lemma 10] and [7, Lemma 11] by

$$\mathrm{e}^{\sqrt{L}}Q(s)(\log D)^{-1/3}$$

where, as stated in [24, equation (1.6)], the function $Q(s)$ satisfies

$$Q(s) < \exp\{-s\log s + s\log\log 3s + O(s)\}, \quad s \geqslant 3.$$

The constant L in our case will depend at most on the coefficients of $\mathbf{f}$, which is considered constant throughout our paper, thus we can assume that the terms above are $\ll_{\mathbf{f}} s^{-s}(\log D)^{-1/3}$, with an implied constant depending at most on $\mathbf{f}$. Let us choose the set of primes

$$\mathscr{P} := \{p \text{ prime: } p > z_0\}.$$

Moreover, we observe that $\varpi_i(k)$ is a multiplicative function for all $1 \leqslant i \leqslant n$. We define the modified multiplicative function $\widetilde{\varpi}_i(k)$ by

$$\widetilde{\varpi}_i(p) := \begin{cases} \varpi_i(p) & \text{if } p > z_0, \\ 0 & \text{if } p \leqslant z_0. \end{cases}$$

So far we can only assume that $\varpi_i(p) \leqslant 1 + p^{-\epsilon}$, whereas in [7] they work with the stronger statement that $\varpi(p) \leqslant 1 + 1/(p-1)$. However, we still get the bound present in [7, equation (3.10)] for a uniform L. For this we observe that

$$\begin{aligned} \log \prod_{w_1 < p \leqslant w_2} \left(1 - \frac{\widetilde{\varpi}_i(p)}{p}\right)^{-1} &\leqslant \sum_{w_1 < p \leqslant w_2} \log\left(1 - \frac{1}{p} - \frac{C}{p^{1+\epsilon}}\right)^{-1} \\ &\leqslant \sum_{w_1 < p \leqslant w_2} \left(\frac{1}{p} + \frac{C}{p^{1+\epsilon}}\right) + O(w_1^{-1}) \\ &\leqslant \log\log w_2 - \log\log w_1 + O\left(\frac{1}{\log w_1}\right), \end{aligned}$$

by Mertens' theorem. This leads to the bound

$$\prod_{w_1 < p \leqslant w_2} \left(1 - \frac{\widetilde{\varpi}_i(p)}{p}\right)^{-1} \leqslant \left(\frac{\log w_2}{\log w_1}\right)\left(1 + \frac{L}{\log w_1}\right),$$

for a uniform constant $L = L(\mathbf{f})$. We can now directly apply [7, Lemma 11] to the inner sums appearing in (5.10). Introduce the constant s_0 through $s_0 := (\log D_1)/(\log z_1)$, which we demand fulfils $s_0 \geqslant 3$, and set

$$U_i(z_1, \xi_i) := \mu(\xi_i) \prod_{\substack{p \mid \xi_i \\ p > z_0}} \frac{\varpi_i(p)}{p - \varpi_i(p)} \prod_{z_0 < p \leqslant z_1} \left(1 - \frac{\varpi(p)}{p}\right).$$

This provides us with

$$\sum_{\substack{k\in\mathbb{N}\\ k|P(z_0,z_1)\\ \xi_i|k}} \frac{\lambda_k^+\varpi_i(k)}{k} = U_i(z_1,\xi_i) + O(\tau(\xi_i)s_0^{-s_0}).$$

Owing to the apparent bounds $0 \leqslant \varpi_i(p) < p/2$, valid for $p > z_0$ (as long as z_0 is enlarged) we deduce that $|U_i(z_1,\xi_i)| \leqslant 1$ for all $1 \leqslant i \leqslant n$ and divisors $\xi_i|P(z_0,z_1)$. We use this approximation in (5.10), to obtain

$$S((u_{i,j})) - \sum_{\substack{l_{i,j}\leqslant\Delta^{B_1}\\ u_{i,j}l_{i,j}|P(z_0,z_1)}} \mu(\mathbf{l})\prod_{i=1}^{n} U_i(z_1,\xi_i)$$
$$\ll (\log z_1)^n\Delta^{-B_1/2} + \Delta^{B_1\binom{n}{2}+1/100}(s_0^{-s_0} + s_0^{-s_0}(\log D_1)^{-1/3}).$$

Assume that the assumptions in Lemma 5.2 are satisfied. Together with equation (5.8) we now obtain

$$\Sigma(D_1,z_1) = \Sigma^{MT}(D_1,z_1) + \Sigma^{ET}(D_1,z_1),$$

with a main term given by

$$\Sigma^{MT}(D_1,z_1) = \sum_{\substack{u_{i,j}\leqslant\Delta\\ 1\leqslant i<j\leqslant n}} \widehat{g}((u_{i,j})) \sum_{\substack{(l_{i,j})\in\mathbb{N}^{\binom{n}{2}}\\ l_{i,j}\leqslant\Delta^{B_1}\\ u_{i,j}l_{i,j}|P(z_0,z_1)}} \mu(\mathbf{l})\prod_{i=1}^{n} U_i(z_1,\xi_i),$$

and an error term satisfying

$$\Sigma^{ET}(D_1,z_1) \ll \frac{(\log z_1)^n}{\Delta^{1-\epsilon}} + \frac{\Delta^{C+\binom{n}{2}-B_1/2}}{(\log z_1)^{-n}} + \Delta^{C+(B_1+1)\binom{n}{2}+1/100}s_0^{-s_0},$$

where $C = C(\mathbf{f}) > 0$ is such that

$$|\widehat{g}((u_{i,j}))| \ll \max\{u_{i,j}\}^C.$$

We will assume that such a C exists for the moment, this will be proved later in Lemma 5.4. Therefore we may choose $B_1 > 0$ large enough so that $C + \binom{n}{2} - B_1/2 < -1$. We can then obtain

$$\Sigma(D_1,z_1) - \Sigma^{MT}(D_1,z_1) \ll \frac{(\log z_1)^n}{\Delta^{1-\epsilon}} + \Delta^c s_0^{-s_0}, \tag{5.11}$$

where $c = c(\mathbf{f}) > 0$. Note that here we implicitly assume that $s_0 \geqslant 3$, thus $\log D_1/\log z_1 \geqslant 3$.

LEMMA 5.4. *Assume that* $\mathfrak{B}(\mathbf{f}) > \max\{R^2 2^{d-1}(d^2-1), R^2 2^{d-1}(d-1) + (R+1)\}$. *Let* $\mathbf{u} \in \mathbb{N}^{\binom{n}{2}}$ *be such that* $\mu^2(\mathbf{u}) = 1$ *and such that* $p | \widetilde{\mathbf{u}}$ *implies that* $p > z_0$. *Then, for* z_0 *sufficiently large one has*

$$\widehat{g}((u_{i,j})) \ll \left(\prod_{i \neq j} u_{i,j} \right)^{(d\mathfrak{B}(\mathbf{f})/(d-1)2^{d-1})(d-1/R)+R+\epsilon}.$$

Proof. First we recall that

$$\widehat{g}((u_{i,j})) \prod_{i=1}^{n} \varpi_i(u_i) = \varpi(\mathbf{u}), \tag{5.12}$$

where we have $u_{i,j} = \gcd(u_i, u_j)$. For bounding $\widehat{g}((u_{i,j}))$ we may make the following assumption: if p is a prime with $p | u_i$ for some $1 \leqslant i \leqslant n$, then there is a $1 \leqslant j \leqslant n$, $j \neq i$ such that $p | u_{i,j}$. Otherwise we could replace in (5.12) the vector $\mathbf{u}$ with a vector $\widetilde{\mathbf{u}}$ where $\widetilde{u_k} = u_k$ for $k \neq i$ and $\widetilde{u_k} = u_k/p$ for $k = i$. In particular, we may assume that

$$u_i \leqslant \prod_{j \neq i} u_{ij},$$

for every $1 \leqslant i \leqslant n$.

Next we observe that

$$\prod_{i=1}^{n} \varpi_i(u_i) = \prod_{i=1}^{n} \prod_{p | u_i} \varpi_i(p).$$

We recall the identity $\varpi_i(p) = p\sigma(p^{\mathbf{e}_i} | \mathbf{x}) \sigma_p^{-1}$. By Corollary 3.6 we have

$$\sigma(p^{\mathbf{e}_i} | \mathbf{x}) = \frac{1}{p} + O(p^{-1-\epsilon}).$$

Therefore we obtain

$$\prod_{i=1}^{n} \varpi_i(u_i) = \prod_{i=1}^{n} \prod_{p | u_i} (1 + O(p^{-1-\epsilon}))^{-1} (1 + O(p^{-\epsilon})),$$

and $\prod_{i=1}^{n} \varpi_i(u_i)^{-1} \ll_{\mu} (u_1 \cdots u_n)^{\mu}$, for any $\mu > 0$. By Corollary 3.8 we have

$$\varpi(p^{\mathbf{j}}) \ll p^{(d\mathfrak{B}(\mathbf{f})/(d-1)2^{d-1})(d-1/R)+R}.$$

Injecting these bounds into (5.12) yields

$$\begin{aligned} \widehat{g}((u_{i,j})) &\ll \left(\prod_{p | \widetilde{\mathbf{u}}} p \right)^{(d\mathfrak{B}(\mathbf{f})/(d-1)2^{d-1})(d-1/R)+R+\epsilon} \\ &\ll \left(\prod_{i \neq j} u_{i,j} \right)^{(d\mathfrak{B}(\mathbf{f})/(d-1)2^{d-1})(d-1/R)+R+\epsilon}, \end{aligned}$$

thus concluding the proof. □

As in [7, p. 90], we now observe that $\Sigma^{MT}(D_1, z_1)$ is independent of D_1. We set

$$D_2 := \max(D_1, 3^{z_1})$$

and with equation (5.11) applied to D_2 instead of D_1, we obtain that

$$\Sigma(D_1, z_1) - \Sigma(D_2, z_1) \ll \frac{(\log z_1)^n}{\Delta^{1-\epsilon}} + \Delta^c s_0^{-s_0}.$$

For this choice of D_2 we have $\lambda_d^+ = \mu(d)$ for $d|P(z_0, z_1)$ (note that with the change of D_1 to D_2 also the sieve weights λ change). Hence we can compute $\Sigma(D_2, z_1)$ as

$$\Sigma(D_2, z_1) = \sum_{\substack{\mathbf{d}\in\mathbb{N}^n \\ d_i|P(z_0,z_1)}} \frac{\varpi(\mathbf{d})}{\widetilde{\mathbf{d}}} \prod_{i=1}^n \mu(d_i) = \prod_{z_0<p\leqslant z_1} \left(1 - \frac{g(p)}{p}\right),$$

with $g(p)$ defined as in (4.7). Injecting our estimates for $\Sigma(D_1, z_1)$ into (5.5) yields the upper bound in the next result.

PROPOSITION 5.5. *Assuming $l_i|P(z_1, z)$, $|\mathbf{l}|D_1 \log B \leqslant B^{1/\rho}$ and that $\mathfrak{B}(\mathbf{f})$ exceeds*

$$\max\{2^{d-1}(d-1)R(R+1), 2^{d-1}(d-1)R^2 + (R+1)(\Upsilon+1), 2^{d-1}(d^2-1)R^2\}$$

we have

$$\begin{aligned} G(B, z_1; \mathbf{l}) = X_{\mathbf{l}} &\prod_{z_0<p\leqslant z_1} \left(1 - \frac{g(p)}{p}\right) \\ &+ O\left(\varpi(\mathbf{l}) \frac{B^{n-Rd}}{\widetilde{\mathbf{l}}} \left(\frac{(\log z_1)^n}{\Delta^{1-\epsilon}} + \Delta^c s_0^{-s_0}\right)\right) \\ &+ O\left(\frac{B^{n+\epsilon}}{\widetilde{\mathbf{l}}} \sum_{\substack{|\mathbf{k}|\leqslant D_1 \\ p|\widetilde{\mathbf{k}} \Rightarrow z_0<p\leqslant z_1}} \mu(\mathbf{k})^2 \frac{E(B; (k_1 l_1, \ldots, k_n l_n))}{\widetilde{\mathbf{k}}}\right). \end{aligned}$$

The lower bound can be procured upon writing

$$G(B, z_1; \mathbf{l}) = \sum_{\substack{\mathbf{x}\in\mathscr{A} \\ \mathbf{l}|\mathbf{x}}} \left(\prod_{i=1}^n w\left(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\right)\right) \left(\prod_{i=1}^n (1 * \mu)(\gcd(P(z_0, z_1), x_i))\right)$$

and using Lemma 5.1 to obtain

$$G(B, z_1; \mathbf{l}) \geqslant \sum_{i=0}^n c_i M_i,$$

where for $1 \leqslant i \leqslant n$ we define $c_i := 1$ and

$$M_i := \sum_{k_i \mid P(z_0,z_1)} \Big(\lambda^-_{k_i} \prod_{j\neq i} \lambda^+_{k_j}\Big) N_w(B; (k_1 l_1, \ldots, k_n l_n)),$$

in addition to $c_0 := -(n-1)$ and

$$M_0 := \sum_{k_i \mid P(z_0,z_1)} \Big(\prod_{i=1}^n \lambda^+_{k_i}\Big) N_w(B; (k_1 l_1, \ldots, k_n l_n)).$$

The treatment of each individual M_i, $(i \neq 0)$, is identical to the treatment of M_0 earlier in this section. The only difference arises at the last step (the calculation of the Euler products in the main term). Here the coefficients c_i satisfy $\sum_{0\leqslant i \leqslant n} c_i = 1$, thus completing the proof of Proposition 5.5.

§6. *Proof of Theorems 1.1 and 1.4.* Recall the definition of the set $\mathscr{A}$ in (4.3). Our aim is to find a large function $z = z(B) \leqslant B$ such that

$$S(B,z) := \sharp\{\mathbf{x} \in \mathscr{A} : |\mathbf{x}| \leqslant B,\, p | x_1 \cdots x_n \Rightarrow p > z\} \gg \frac{B^{n-Rd}}{(\log B)^n}.$$

By (4.5) we have

$$S(B,z) \geqslant w_0^{-n} S_{\boldsymbol{\zeta}}(B,z), \tag{6.1}$$

where

$$S_{\boldsymbol{\zeta}}(B,z) := \sum_{\substack{\mathbf{x} \in \mathscr{A} \\ p | x_1 \cdots x_n \Rightarrow p > z}} \prod_{i=1}^n w\Big(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\Big).$$

One may now write the sum over $\mathbf{x}$ as

$$\sum_{\mathbf{x} \in \mathscr{A}} \Big(\prod_{i=1}^n w\Big(\frac{x_i}{B} - \frac{\zeta_i}{2|\boldsymbol{\zeta}|}\Big)\Big)\Big(\prod_{i=1}^n (1 * \mu)(\gcd(P(z), x_i))\Big).$$

For a parameter D let $\lambda^\pm$ be a sieve sequence supported in $[1, D]$. Letting

$$\beta(\mathbf{l}) := \sum_{i=1}^n \lambda^-_{l_i} \Big(\prod_{\substack{1\leqslant j \leqslant n \\ j\neq i}} \lambda^+_{l_j}\Big) - (n-1) \prod_{i=1}^n \lambda^+_{l_i}, \quad \mathbf{l} \in \mathbb{N}^n,$$

alluding to Lemma 5.1 and recalling (5.3) allows us to infer that for any $z > z_1$ we have

$$S_{\boldsymbol{\zeta}}(B,z) \geqslant \sum_{\substack{\mathbf{l} \in \mathbb{N}^n \\ l_i | P(z_1,z)}} \beta(\mathbf{l}) G(B, z_1; \mathbf{l}).$$

Define the entities

$$\Sigma(D, z_1, z) := \sum_{\mathbf{l} | P(z_1,z)} \beta(\mathbf{l}) \frac{\varpi(\mathbf{l})}{\widetilde{\mathbf{l}}},$$

$$B_1 := \sum_{\mathbf{l}|P(z_1,z)} \frac{\varpi(\mathbf{l})}{\widetilde{\mathbf{l}}}$$

and

$$B_2 := B^{dR+\epsilon} \sum_{\substack{|\mathbf{l}|\leqslant D \\ \mathbf{l}|P(z_1,z)}} \frac{1}{\widetilde{\widetilde{\mathbf{l}}}} \sum_{\substack{|\mathbf{k}|\leqslant D_1 \\ \mathbf{k}|P(z_0,z_1)}} \frac{E(B;(k_1l_1,\ldots,k_nl_n))}{\widetilde{\mathbf{k}}}.$$

Proposition 5.5 now leads to

$$\frac{S_\xi(B,z)}{\mathfrak{S}(\mathbf{f})\mathscr{J}_w(\mathbf{f},W)B^{n-Rd}} \geqslant \Sigma(D,z_1,z)\prod_{z_0<p\leqslant z_1}\left(1-\frac{g(p)}{p}\right) + O\left(\left(\frac{(\log z_1)^n}{\Delta^{1-\epsilon}}+\frac{\Delta^c}{s_0^{s_0}}\right)B_1+B_2\right). \quad (6.2)$$

Letting $m_i := k_il_i$ and taking advantage of the coprimality of k_i, l_i shows that

$$B_2 \leqslant B^{dR+\epsilon} \sum_{\substack{|\mathbf{m}|\leqslant DD_1 \\ \mathbf{m}|P(z_0,z)}} \frac{E(B;\mathbf{m})}{\widetilde{\mathbf{m}}}.$$

Recalling the definition of the matrix $\boldsymbol{\epsilon}$ given in (1.7), shows that, under the condition

$$DD_1 \leqslant \frac{B^{1/\rho}}{\log B},$$

the sum over $\mathbf{m}$ is

$$\ll \sum_{i=1}^{3} B^{-\epsilon_{i,1}} \sum_{1\leqslant m_1\leqslant DD_1} m_1^{\epsilon_{i,2}-1} \times \sum_{1\leqslant m_2\leqslant m_1} m_2^{-1}\ldots \sum_{1\leqslant m_{n-1}\leqslant m_{n-2}} m_{n-1}^{-1} \sum_{1\leqslant m_n\leqslant m_{n-1}} m_n^{\epsilon_{i,3}-1}.$$

Since each $\epsilon_{i,j}$ is non-negative we can use the estimate $\sum_{1\leqslant m\leqslant z} m^{\lambda-1} \ll_\lambda z^\lambda \log z$, valid for each fixed $\lambda \geqslant 0$ to deduce that for every $\epsilon > 0$ one has

$$B_2 \ll B^{dR+\epsilon} \sum_{i=1}^{3} B^{-\epsilon_{i,1}} (DD_1)^{\epsilon_{i,2}+\epsilon_{i,3}}.$$

Our remaining task will be to give a lower bound for Σ and an upper bound for B_1. We begin by studying the contribution to $\Sigma(D,z_1,z)$ of vectors $\mathbf{l}$ with $\delta := \gcd(l_{i_1}, l_{i_2}) \neq 1$; this task is similar to the one in Lemma 5.2 and we adapt its assumptions in what follows. Each such δ is a product of primes $p > z_1$, therefore this contribution is

$$\ll \sum_{\delta>z_1} \mu(\delta)^2 \sum_{\substack{\mathbf{l}|P(z_1,z) \\ \delta|l_{i_1},\delta|l_{i_2}}} \frac{\varpi(\mathbf{l})}{\widetilde{\mathbf{l}}}.$$

As in the proof of Lemma 5.2 we find that this is

$$\ll \left(\frac{\log z}{\log z_1}\right)^n \sum_{\delta > z_1} \delta^{-2+\epsilon} \ll z_1^{-1+\epsilon}(\log z)^n.$$

Note that if l_i, l_j are coprime for all $i \neq j$ guarantees that $\varpi(\mathbf{l}) = \prod_{i=1}^n \varpi_i(l_i)$. This gives

$$\Sigma(D, z_1, z) = \sum_{\substack{\mathbf{l} \mid P(z_1,z) \\ i \neq j \Rightarrow \gcd(l_i,l_j)=1}} \frac{\beta(\mathbf{l})}{\tilde{\mathbf{l}}} \prod_{i=1}^n \varpi_i(l_i) + O(z_1^{-1+\epsilon}(\log z)^n).$$

The same argument can also be used to establish

$$\sum_{\mathbf{l} \mid P(z_1,z)} \frac{\beta(\mathbf{l})}{\tilde{\mathbf{l}}} \prod_{i=1}^n \varpi_i(l_i) = \sum_{\substack{\mathbf{l} \mid P(z_1,z) \\ i \neq j \Rightarrow \gcd(l_i,l_j)=1}} \frac{\beta(\mathbf{l})}{\tilde{\mathbf{l}}} \prod_{i=1}^n \varpi_i(l_i) + O(z_1^{-1+\epsilon}(\log z)^n).$$

Letting

$$\Psi_i^{\pm} := \sum_{l \mid P(z_1,z)} \lambda_l^{\pm} \frac{\varpi_i(l)}{l}$$

shows that the sum on the left equals

$$\Psi := \sum_{i=1}^n \left(\Psi_i^- \prod_{\substack{1 \leqslant j \leqslant n \\ j \neq i}} \Psi_j^+\right) - (n-1) \prod_{i=1}^n \Psi_i^+,$$

thus providing

$$\Sigma(D, z_1, z) = \Psi + O(z_1^{-1+\epsilon}(\log z)^n).$$

Under the assumptions of Lemma 5.2 we can similarly show that the contribution of $\mathbf{l}$ with $\gcd(l_{i_1}, l_{i_2}) \neq 1$ to B_1 is

$$\ll \sum_{\delta > z_1} \mu(\delta)^2 \sum_{\substack{\mathbf{l} \mid P(z_1,z) \\ \delta \mid l_{i-1}, \delta \mid l_{i_2}}} \frac{\varpi(\mathbf{l})}{\tilde{\mathbf{l}}} \ll B^{n-Rd}(z_1^{-1+\epsilon}(\log z)^n).$$

Therefore

$$B_1 \ll z_1^{-1+\epsilon}(\log z)^n + \sum_{\substack{\mathbf{l} \mid P(z_1,z) \\ i \neq j \Rightarrow \gcd(l_i,l_j)=1}} \prod_{i=1}^n \frac{\varpi_i(l_i)}{l_i}$$

and the last sum is

$$\leqslant \prod_{i=1}^n \sum_{l \mid P(z_1,z)} \frac{\varpi_i(l)}{l} \leqslant \prod_{i=1}^n \prod_{z_1 < p \leqslant z} \left(1 + \frac{1}{p} + O(p^{-1-\epsilon})\right) \ll (\log z)^n,$$

hence $B_1 \ll (\log z)^n$. We therefore find via (6.2) the following lower bound:

$$\frac{S_\zeta(B,z)}{\mathfrak{S}(\mathbf{f})\mathscr{J}_w(\mathbf{f},W)B^{n-Rd}} \geqslant \Psi \prod_{z_0<p\leqslant z_1}\left(1-\frac{g(p)}{p}\right) + O\left(B^{dR+\epsilon}\sum_{i=1}^{3}B^{-\epsilon_{i,1}}(DD_1)^{\epsilon_{i,2}+\epsilon_{i,3}}\right) + O\left(\frac{(\log z)^n}{z_1^{1-\epsilon}(\log z_1)^n}+\left(\frac{(\log z_1)^n}{\Delta^{1-\epsilon}}+\frac{\Delta^c}{s_0^{s_0}}\right)(\log z)^n\right),$$

where a use of

$$\prod_{z_0<p\leqslant z_1}\left(1-\frac{g(p)}{p}\right) \ll (\log z_1)^{-n}$$

has been made; this can be inferred from the estimate $g(p)=n/p+O(p^{-1-\epsilon})$. Let us now fix any $\theta>0$ which satisfies $\theta<\theta'$, where θ' was defined in (1.5). Then there exists a small positive θ_1 such that if $D:=B^\theta$ and $D_1:=B^{\theta_1}$ then

$$B^{dR+\epsilon}\sum_{i=1}^{3}B^{-\epsilon_{i,1}}(DD_1)^{\epsilon_{i,2}+\epsilon_{i,3}} \ll B^{-\delta},$$

for some $\delta>0$ that is independent of B. Choosing $\Delta=z_1=(\log B)^{2n+1}$ shows that

$$s_0=\frac{\log D_1}{\log z_1}=\frac{\theta_1\log B}{(2n+1)\log\log B}\to\infty,$$

hence one can verify that

$$\frac{(\log z)^n}{z_1^{1-\epsilon}(\log z_1)^n}+\left(\frac{(\log z_1)^n}{\Delta^{1-\epsilon}}+\frac{\Delta^c}{s_0^{s_0}}\right)(\log z)^n \ll \frac{1}{(\log B)^n\log\log B}$$

and

$$\frac{S_\zeta(B,z)}{\mathfrak{S}(\mathbf{f})\mathscr{J}_w(\mathbf{f},W)B^{n-Rd}} \geqslant \Psi \prod_{z_0<p\leqslant z_1}\left(1-\frac{g(p)}{p}\right)+O((\log B)^{-n}(\log\log B)^{-1}).$$

The last product is $\gg (\log z_1)^{-n}$, thus it remains to show that $\Psi \gg (\log z_1/\log z)^n$. Let $s:=\log D/\log z$ and assume that $s>2$. Using the inequalities stated in [7, Lemma 10] one deduces that when $s=O_n(1)$ with an implied constant depending at most on n, then

$$\Psi \geqslant (\Psi_n(s)+O_n((\log B)^{-1/3}))\prod_{i=1}^{n}\prod_{z_1<p\leqslant z}\left(1-\frac{\varpi_i(p)}{p}\right),$$

where $\Psi_n(s):=nf(s)-(n-1)F(s)^n$. Here $f(s)$ and $F(s)$ denote the well-known functions associated with the linear sieve, their definition can be found in

[**16**, equations (12.1) and (12.2)], for example. Further information on f and F is located in [**16**, §§11, 12]. In light of the last lower bound for Ψ, it is sufficient to find the smallest possible value for s such that $\Psi_n(s) > 0$. This is equivalent to

$$\frac{F(s)^n}{f(s)} < 1 + \frac{1}{n-1}. \tag{6.3}$$

It is a standard fact that when $s > 2$ then $0 < f(s) \leqslant 1 \leqslant F(s)$. Therefore if s remains constant and independent of n then one cannot prove (6.3) for large n, this forces us to take s as a function of n that tends to infinity. At this point we have to employ asymptotic approximations for $f(s)$ and $F(s)$, these can be found in [**16**, equation (11.134)]. They are given by

$$F(s),\, f(s) = 1 \pm \exp\left\{-s\log s - s\log\log s + s + O\left(\frac{s\log\log s}{\log s}\right)\right\}$$

and one sees that if $s \geqslant 3(\log n)(\log\log n)^{-1}$ then

$$s\log s \geqslant 3\log n + \frac{3(\log 3)(\log n)}{\log\log n} - \frac{3(\log n)(\log\log\log n)}{\log\log n}.$$

Therefore, for all large enough n, say $n \geqslant n_0$ for some positive absolute constant n_0, we obtain

$$\frac{F(s)^n}{f(s)} < \left(1 + \frac{1}{n^{5/2}}\right)^{n+1}$$

and the inequality $1 + n^{-5/2} < (1 + (n-1)^{-1})^{1/(n+1)}$, valid for all $n \geqslant 2$, makes (6.3) available. In the case that $1 \leqslant n < n_0$ one can immediately infer from the approximations to $F(s)$ and $f(s)$ that if $s \to +\infty$ then (6.3) is automatically satisfied. This gives a constant σ_0 that depends at most on n_0 (and is therefore absolute) such that (6.3) is valid whenever $s \geqslant \sigma_0$. Hence there exists a positive absolute constant σ_0 such that if

$$s \geqslant \frac{3\log n}{\log\log n} + \sigma_0$$

then, alluding to (6.1), Theorem 1.1 holds with any constant $c_0 > 3 + \sigma$ and with $P^-(x_1 \cdots x_n)$ exceeding the sieving parameter $z = D^{1/s} = B^{\theta/s}$.

Proof of Theorem 1.4. The arguments in the present section have so far proved that

$$S_{\boldsymbol{\zeta}}(B, B^{\theta/s}) \gg B^{n-Rd}(\log B)^{-n}.$$

This is sufficient for Theorem 1.4 because to show that a subset of $V_{\mathbf{f}}(\mathbb{Q})$ is Zariski dense in an absolutely irreducible variety $V_{\mathbf{f}}$, it is sufficient to choose an arbitrary neighbourhood in the real analytic topology of a non-singular point $\boldsymbol{\zeta} \in V_{\mathbf{f}}(\mathbb{R})$ and show that any real point in the neighbourhood (on the variety) can be approximated by a rational point. In our case the neighbourhood is given by $B\mathscr{B}_\eta$ (where $\mathscr{B}_\eta$ was defined in (4.4)). □

Acknowledgements. We would like to thank Professor T. D. Browning and Dr. S. Yamagishi for their comments on an earlier version of this paper, as well as the anonymous referee for numerous helpful comments that have clarified the exposition considerably. The first author is supported by a NWO grant 016.Veni.173.016.

References

1. N. C. Ankeny and H. Onishi, The general sieve. *Acta Arith.* **10**(1) (1964/1965), 31–62.
2. B. J. Birch, Forms in many variables. *Proc. R. Soc. Ser.* A **265** (1961/1962), 245–263.
3. V. Blomer and J. Brüdern, A three squares theorem with almost primes. *Bull. Lond. Math. Soc.* **37**(4) (2005), 507–513.
4. T. D. Browning, *Quantitative Arithmetic of Projective Varieties* (*Progress in Mathematics* **277**), Birkhäuser (Basel, 2009).
5. T. D. Browning and D. T. Loughran, Sieving rational points on varieties. *Preprint*, 2017, arXiv:1705.01999.
6. T. D. Browning and S. M. Prendiville, Improvements in Birch's theorem on forms in many variables. *J. Reine Angew. Math.* **731** (2017), 203–234.
7. J. Brüdern and É. Fouvry, Lagrange's four squares theorem with almost prime variables. *J. Reine Angew. Math.* **454** (1994), 59–96.
8. J. Brüdern and É. Fouvry, Le crible à vecteurs. *Compos. Math.* **102**(3) (1996), 337–355.
9. A. A. Buchstab, Asymptotic estimates of a general number-theoretic function. *Mat. Sb.* **44**(2) (1937), 1239–1246.
10. Y. Cai, Lagrange's four squares theorem with variables of special type. *Int. J. Number Theory* **6**(8) (2010), 1801–1817.
11. B. Cook and Á. Magyar, Diophantine equations in the primes. *Invent. Math.* **198**(3) (2014), 701–737.
12. H. Davenport, Cubic forms in 32 variables. *Philos. Trans. R. Soc.* A **251**(A) (1959), 193–232.
13. H. Davenport, *Analytic Methods for Diophantine Equations and Diophantine Inequalities, (Cambridge Mathematical Library)*, 2nd edn., Cambridge University Press (Cambridge, 2005).
14. H. G. Diamond and H. Halberstam, On the sieve parameters α_κ and β_κ for large κ. *J. Number Theory* **67**(1) (1997), 52–84.
15. H. G. Diamond and H. Halberstam, *A Higher-Dimensional Sieve Method* (*Cambridge Tracts in Mathematics* **177**), Cambridge University Press (Cambridge, 2008).
16. J. Friedlander and H. Iwaniec, *Opera de Cribro* (*American Mathematical Society Colloquium Publications* **57**), American Mathematical Society (Providence, RI, 2010).
17. M. Hanselmann, Rational points on quartic hypersurfaces. *PhD Thesis*, Ludwig Maximilians Universität Munchen, 2012.
18. R. Heath-Brown and X. Li, Almost prime triples and Chen's theorem. *J. Number Theory* **169** (2016), 265–294.
19. D. R. Heath-Brown and D. I. Tolev, Lagrange's four squares theorem with one prime and three almost-prime variables. *J. Reine Angew. Math.* **558** (2003), 159–224.
20. C. Hooley, On nonary cubic forms. *J. Reine Angew. Math.* **386** (1988), 32–98.
21. L. K. Hua, Some results in the additive prime-number theory. *Quart. J. Math. Oxford Ser.* (2) **9**(1) (1938), 68–80.
22. J.-W. van Ittersum, Quantitative results on diophantine equations in many variables. *Preprint*, 2017, arXiv:1709.05126.
23. J.-W. van Ittersum, Two problems related to the circle method. *Master's Thesis*, Universiteit Utrecht, 2017.
24. H. Iwaniec, Rosser's sieve. *Acta Arith.* **36**(2) (1980), 171–202.
25. W. B. Jurkat and H.-E. Richert, An improvement of Selberg's sieve method. I. *Acta Arith.* **11** (1965), 217–240.
26. M. S. Klamkin and D. J. Newman, Extensions of the Weierstrass product inequalities. *Math. Mag.* **43** (1970), 137–141.
27. Á. Magyar and T. Titichetrakun, Almost prime solutions to Diophantine systems of high rank. *Int. J. Number Theory* **13**(6) (2017), 1491–1514.
28. H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory. I. Classical Theory* (*Cambridge Studies in Advanced Mathematics* **97**), Cambridge University Press (Cambridge, 2007).

29. E. Sofos and Y. Wang, Finite saturation for unirational varieties. *Int. Math. Res. Not. IMRN* (2018), doi:10.1093/imrn/rnx318.

30. D. I. Tolev, Lagrange's four squares theorem with variables of special type. In *Proceedings of the Session in Analytic Number Theory and Diophantine Equations* (*Bonner Math. Schriften* **360**), University of Bonn (Bonn, 2003), 17.

31. S. Yamagishi, Diophantine equations in semiprimes. *Preprint*, 2017, arXiv:1709.03605.

D. Schindler,
Universiteit Utrecht,
Mathematisch Instituut,
Budapestlaan 6,
Utrecht, 3584 CD,
Netherlands
E-mail: d.schindler@uu.nl

E. Sofos,
Max Planck Institute for Mathematics,
Vivatsgasse 7,
Bonn, 53111,
Germany
E-mail: sofos@mpim-bonn.mpg.de