



Browning, T. D., & Prendiville, S. M. (2015). Improvements in Birch's theorem on forms in many variables. *Journal für die reine und angewandte Mathematik*. DOI: [10.1515/crelle-2014-0122](https://doi.org/10.1515/crelle-2014-0122)

Publisher's PDF, also known as Version of record

License (if available):  
CC BY-NC-ND

Link to published version (if available):  
[10.1515/crelle-2014-0122](https://doi.org/10.1515/crelle-2014-0122)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

# Improvements in Birch's theorem on forms in many variables

By *T. D. Browning* at Bristol and *S. M. Prendiville* at Bristol

---

**Abstract.** We show that a non-singular integral form of degree  $d$  is soluble over the integers if and only if it is soluble over  $\mathbb{R}$  and over  $\mathbb{Q}_p$  for all primes  $p$ , provided that the form has at least  $(d - \frac{1}{2}\sqrt{d})2^d$  variables. This improves on a longstanding result of Birch.

## 1. Introduction

Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a homogeneous polynomial of degree  $d \geq 3$ . A fundamental ambition in number theory is to determine when the Diophantine equation

$$(1.1) \quad F(x_1, \dots, x_n) = 0$$

has a non-trivial integral solution. The Hardy–Littlewood circle method has been extraordinarily effective at answering this question for typical polynomials  $F$  when the number of variables is sufficiently large in terms of  $d$ . An obvious necessary condition for the solubility of (1.1) in integers is that it should be *everywhere locally soluble*, by which we mean that it has non-trivial solutions over  $\mathbb{R}$  and  $\mathbb{Q}_p$  for every prime  $p$ . According to a renowned result of Birch [1], these conditions are sufficient provided that  $F$  is non-singular and  $n > (d - 1)2^d$ . It is possible to relax the non-singularity condition by imposing stronger constraints on  $n$  and local solubility. For the latter, Birch asks instead for the system

$$(1.2) \quad F(x_1, \dots, x_n) = 0, \quad \nabla F(x_1, \dots, x_n) \neq \mathbf{0}$$

to be everywhere locally soluble. We say that  $F$  satisfies the *smooth Hasse principle* if this condition is sufficient to ensure that this system also has a non-trivial integral solution. Allowing  $\sigma$  to denote the (affine) dimension of the *singular locus* cut out by the system of equations

$$\nabla F(x_1, \dots, x_n) = \mathbf{0},$$

it follows from Birch's investigation [1] that  $F$  satisfies the smooth Hasse principle provided that

$$(1.3) \quad n - \sigma > (d - 1)2^d.$$

Note that  $\sigma \in \{0, \dots, n - 1\}$ , with  $\sigma = 0$  if and only if  $F$  is non-singular.

---

While working on this paper the authors were supported by the Leverhulme Trust and ERC grant 306457.

Birch's theorem has had an extensive impact on number theory, with the underlying tools being adapted to handle numerous problems. This includes, but is not limited to:

- the vanishing of  $F$  on general  $\mathbb{Z}$ -linear subspaces (Brandes [2]),
- a generalisation to the function field  $\mathbb{F}_q(t)$  (Lee [11]),
- a generalisation to bihomogeneous forms (Schindler [13]),
- a generalisation to arbitrary number fields (Skinner [15]).

Activity around reducing the lower bound (1.3) for  $n - \sigma$  in Birch's original result, however, has not been so vigorous.

The most impressive improvement to date arises in the case  $d = 3$  of cubic forms. Thus, it follows from work of Hooley [10] that the smooth Hasse principle holds for integral cubic forms provided that  $n - \sigma \geq 9$ . Moreover, Heath-Brown [8] has shown that any integral cubic form has a non-trivial integer zero provided that  $n \geq 14$ , with no restriction on the singular locus, the question of local solubility being automatic. The only other improvement to date pertains to the case  $d = 4$ . In this setting, Browning and Heath-Brown [3] have established the smooth Hasse principle for integral quartic forms provided that  $n - \sigma \geq 41$ , saving eight variables over the approach taken by Birch. Finally, this inequality has been sharpened to  $n - \sigma \geq 40$  by Hanselmann [6].

Our main result improves on (1.3) for every degree.

**Theorem 1.1.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a form of degree  $d \geq 3$  with singular locus of dimension  $\sigma$ . Suppose that*

$$n - \sigma \geq \left(d - \frac{1}{2}\sqrt{d}\right)2^d.$$

*Then the smooth Hasse principle holds for  $F$ .*

As we shall see shortly the proof of this result is based on a generalisation of the method in [3]. One verifies that the admissible range for  $n$  is weaker than that provided by [3, Theorem 1] when  $d = 4$ . In fact, for smaller values of  $d$  we are able to get a much more significant improvement, as in the following result.

**Theorem 1.2.** *Let  $F \in \mathbb{Z}[x_1, \dots, x_n]$  be a form of degree  $3 \leq d \leq 9$  with singular locus of dimension  $\sigma$ . Suppose that*

$$n - \sigma > \frac{3}{4}d2^d - 2d.$$

*Then the smooth Hasse principle holds for  $F$ .*

The question of determining when the system (1.2) is everywhere locally soluble is far from being decided. Denoting by  $v_d(p)$  the least integer  $n$  such that every degree  $d$  form  $F \in \mathbb{Z}[x_1, \dots, x_n]$  has a non-trivial zero in  $\mathbb{Q}_p$ , Artin conjectured that  $v_d(p) = d^2 + 1$  for every prime  $p$  (this is known to be false for even  $d$  but is still open for forms of odd degree). Specialising to the case  $d = 5$  of quintic forms, where solubility over  $\mathbb{R}$  is automatic, it was shown by Leep and Yeoman [12] that  $v_5(p) = 26$  for  $p \geq 47$ . This was strengthened by Heath-Brown [9], so that this equality holds for  $p \geq 17$ . In particular, when Theorem 1.2 is applied to non-singular quintic forms, it suffices to check the solubility over the field  $\mathbb{Q}_p$  for primes  $p \leq 13$ . In this range, the best result we have is due to Zahid [16], who establishes that  $v_5(p) \leq 4562912$  for  $p \leq 13$ .

We shall give an overview of the proof of our main results in Section 2. Taking  $d = 3$  in Theorem 1.2, we obtain the Hasse principle for non-singular cubic forms in at least 13 variables. It is no coincidence that this matches the constraint arising in Skinner's work [14] on non-singular cubic forms over number fields. Indeed, when  $d = 3$ , our proof reduces to the argument in [14] (which over  $\mathbb{Q}$  is Heath-Brown's seminal work [7]—without a Kloosterman refinement). When  $d = 4$ , the inequality in Theorem 1.2 recovers the conclusions of [3] precisely. When  $d = 5$ , for example, we witness a saving of 18 variables over Birch's result. With more work it is likely that the extra saving obtained by Hanselman [6] could be extended to give further modest improvements in Theorem 1.2.

Birch [1] has an analogous result for general systems of integral forms  $F_1, \dots, F_R$  of equal degree. It would be interesting to determine whether the methods of this paper can be developed to produce comparable improvements for  $R > 1$ . Similarly, once suitably modified, it is natural to hope that our argument yields corresponding improvements in the generalisations [2, 11, 13, 15] discussed above.

**Acknowledgement.** The authors are very grateful to the anonymous referee for numerous helpful comments.

## 2. Preliminaries

Our proof of Theorems 1.1 and 1.2 proceeds via the Hardy–Littlewood circle method. In this section, we outline the strategy of the proof, together with some conventions regarding notation and some preliminary technical results.

The overall goal is to establish an asymptotic formula for the quantity

$$(2.1) \quad N_\omega(F; P) := \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ F(\mathbf{x})=0}} \omega(\mathbf{x}/P),$$

as  $P \rightarrow \infty$ , for a suitable weight function  $\omega : \mathbb{R}^n \rightarrow [0, \infty)$ . We show that under the assumptions of Theorem 1.1 or 1.2 on  $n - \sigma$ , there is a constant  $c_F > 0$  such that

$$N_\omega(F; P) \sim c_F P^{n-d},$$

provided that the system (1.2) is everywhere locally soluble.

Our starting point is the identity

$$N_\omega(F; P) = \int_{\mathbb{T}} S(\alpha, P) d\alpha,$$

where  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  and

$$(2.2) \quad S(\alpha, P) := \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega(\mathbf{x}/P) e(\alpha F(\mathbf{x})).$$

The idea is then to divide the torus  $\mathbb{T}$  into a set of major arcs  $\mathfrak{M}$  and minor arcs  $\mathfrak{m}$ . Given  $\Delta > 0$  we define the major arcs  $\mathfrak{M} = \mathfrak{M}(\Delta)$  to be the set

$$(2.3) \quad \mathfrak{M} = \bigcup_{q \leq P^\Delta} \bigcup_{\substack{a \leq q \\ (a,q)=1}} \left\{ \alpha \in \mathbb{T} : \left\| \alpha - \frac{a}{q} \right\| \leq P^{\Delta-d} \right\}.$$

Here we have written  $\|\beta\|$  for the distance from  $\beta \in \mathbb{R}$  to the nearest integer, a function which induces a metric on  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  via  $d(\alpha, \beta) = \|\alpha - \beta\|$ . The intervals in (2.3) are non-overlapping provided that  $\Delta < \frac{d}{3}$ . We define the minor arcs to be their complement  $\mathfrak{m} = \mathbb{T} \setminus \mathfrak{M}$ . In the usual way we seek to prove an asymptotic formula

$$(2.4) \quad \int_{\mathfrak{M}} S(\alpha, P) d\alpha \sim c_F P^{n-d},$$

as  $P \rightarrow \infty$ , together with a satisfactory bound on the minor arcs

$$(2.5) \quad \int_{\mathfrak{m}} S(\alpha, P) d\alpha = o(P^{n-d}).$$

Here the constant  $c_F$  turns out to be a product of local densities which will be positive if the system (1.2) is everywhere locally soluble. The treatment of (2.4) is standard and is the focus of Section 5.

Our main innovation lies in our treatment of (2.5). The plan is to develop extensively the approach adopted in [3] to estimate  $S(\alpha, P)$  when  $F$  is a quartic form. This relied on a single application of van der Corput differencing to get a family of exponential sums involving cubic polynomials. These were then estimated directly using Poisson summation, rather than through further differencing operations. In our work, which deals with forms of degree  $d$ , we produce two key estimates for  $S(\alpha, P)$  in Section 3. The first (Proposition 3.6) is obtained via  $d - k$  applications of van der Corput differencing together with an application of Birch's bound from [1] (suitably modified), as it applies to exponential sums with underlying polynomials of degree at most  $k$ . The second result (Proposition 3.8) is proved using  $d - 3$  applications of van der Corput differencing together with the bound for cubic exponential sums from [3] obtained using Poisson summation.

The final treatment of (2.5) is carried out in Section 4. It is somewhat disappointing that we are unable to cover all of the minor arcs when  $n - \sigma > \frac{3}{4}d2^d - 2d$  for any  $d \geq 3$ . As we shall see in Remark 4.8, however, the criterion that emerges from our deliberations requires  $n - \sigma$  to be asymptotically  $d2^d$ .

The remainder of this section is taken up with introducing notation and proving some preliminary technical results. We use absolute values  $|\mathbf{x}|$  to denote the norm  $\max_i |x_i|$ . Given  $\epsilon \in (\mathbb{N} \cup \{0\})^n$  and a sufficiently differentiable function  $g : \mathbb{R}^n \rightarrow \mathbb{C}$ , put

$$\partial^\epsilon g = \frac{\partial^{\epsilon_1 + \dots + \epsilon_n} g}{\partial x_1^{\epsilon_1} \dots \partial x_n^{\epsilon_n}}.$$

We begin with the following standard results.

**Lemma 2.1** (Partial summation formula). *Let  $\varphi : \{1, \dots, N\}^n \rightarrow \mathbb{C}$  be a function and let*

$$T_\varphi(\mathbf{t}) := \sum_{1 \leq x_1 \leq t_1} \dots \sum_{1 \leq x_n \leq t_n} \varphi(\mathbf{x}).$$

*Then for any  $g \in C^n(\mathbb{R}^n)$  we have*

$$\sum_{1 \leq \mathbf{x} \leq N} g(\mathbf{x})\varphi(\mathbf{x}) = \sum_{\epsilon \in \{0,1\}^n} \frac{(-N)^{\epsilon_1 + \dots + \epsilon_n}}{N^n} \int_{[0,N]^n} \partial^\epsilon g(N\bar{\epsilon} + \mathbf{t}_\epsilon) T_\varphi(N\bar{\epsilon} + \mathbf{t}_\epsilon) d\mathbf{t},$$

*where  $\bar{\epsilon} = (1, 1, \dots, 1) - \epsilon$  and  $\mathbf{t}_\epsilon$  denotes the vector whose  $i$ th coordinate equals zero if  $\epsilon_i = 0$  and equals  $t_i$  if  $\epsilon_i = 1$ .*

*Proof.* This is a consequence of partial summation and induction on the dimension.  $\square$

**Lemma 2.2** (Shrinking lemma). *Given a symmetric  $n \times n$  real matrix  $A$ , define  $N_A(H, \lambda)$  to be the number of  $\mathbf{h} \in \mathbb{Z}^n$  satisfying  $|\mathbf{h}| \leq H$  and  $\|(A\mathbf{h})_j\| \leq \lambda$  for all  $j$ . Then for any  $H \geq 1$ ,  $\lambda \in (0, 1/H]$  and  $\theta \in (0, 1]$ , we have the estimate*

$$N_A(H, \lambda) \ll (H^{-1} + \theta)^{-n} N_A(\theta H, \theta \lambda).$$

*Proof.* This is a consequence of a result by Davenport [4, Lemma 12.6], which is proved using the geometry of numbers. To be precise the statement of [4, Lemma 12.6] gives the bound

$$N_A(H, \lambda) \ll \theta^{-n} N_A(\theta H, \theta \lambda).$$

To see how this implies the lemma, first note that if  $\theta \geq H^{-1}$ , then

$$\theta^{-1} \leq 2(H^{-1} + \theta)^{-1}$$

and we are done. Next, if  $\theta < H^{-1}$ , then  $N_A(\theta H, \theta \lambda) = 1$  and

$$(H^{-1} + \theta)^{-1} > \frac{1}{2}H.$$

Therefore the trivial estimate gives

$$(H^{-1} + \theta)^{-n} N_A(\theta H, \theta \lambda) \geq 2^{-n} H^n \gg N_A(H, \lambda),$$

as required.  $\square$

Our next result involves Diophantine approximation. Given  $\alpha \in \mathbb{T}$  and  $q \in \mathbb{N}$  we say that  $\alpha$  and  $q$  are *primitive* if there exists an  $a \in \mathbb{Z}$  with  $(a, q) = 1$  such that  $\|q\alpha\| = |q\alpha - a|$ . Notice that if  $q$  and  $\alpha$  are not primitive, then one can find a divisor  $q_0$  of  $q$  which is primitive to  $\alpha$  and which satisfies  $\|q_0\alpha\| \leq \frac{q_0}{q} \|q\alpha\| < \|q\alpha\|$ . The following simple result is due to Heath-Brown [8, Lemma 2.3].

**Lemma 2.3.** *Let  $\alpha \in \mathbb{T}$  and  $q \in \mathbb{N}$  be primitive. Suppose that  $m \in \mathbb{Z}$  satisfies:*

- (i)  $|m| < \frac{1}{2} \|q\alpha\|^{-1}$ ,
- (ii)  $\|m\alpha\| < \frac{1}{2} q^{-1}$ ,
- (iii)  $|m| < q$  or  $\|m\alpha\| < \|q\alpha\|$ .

*Then  $m = 0$ .*

*Proof.* Since, by assumption,  $\alpha$  and  $q$  are primitive, there exists an integer  $a$  with the property that  $(a, q) = 1$  and  $\|q\alpha\| = |q\alpha - a|$ . Our formulation of the lemma now follows from [8, Lemma 2.3] with  $P_0 = 2q$ .  $\square$

The remaining results in this section involve viewing various varieties that are defined over  $\mathbb{Q}$  over several different finite fields. To simplify the exposition, write  $\mathbb{F}_\infty$  for  $\mathbb{Q}$ . Given a form  $G \in \mathbb{Z}[x_1, \dots, x_n]$  and  $\nu$  a prime or the prime at infinity, define the *singular locus* of  $G$  over  $\mathbb{F}_\nu$  to be the algebraic set

$$(2.6) \quad \text{Sing}_\nu(G) := \{\mathbf{x} \in \mathbb{A}_{\mathbb{F}_\nu}^n : \partial^{e_i} G(\mathbf{x}) = 0 \text{ for } 1 \leq i \leq n\}.$$

Here  $\mathbf{e}_i$  denotes the  $i$ th standard basis vector. Throughout we use the notation

$$(2.7) \quad \sigma_v(G) := \dim \text{Sing}_v(G).$$

Denote the positive part of a real number  $x$  by

$$x^+ := \max\{x, 0\}.$$

The following is Lemma 1 of Browning and Heath-Brown [3].

**Lemma 2.4.** *Let  $G \in \mathbb{Z}[x_1, \dots, x_n]$  be a form of degree  $d$  whose singular locus over  $\mathbb{F}_v$  has dimension  $\sigma_v(G)$ . Define*

$$B_v(G, s) := \{\mathbf{h} \in \mathbb{A}_{\mathbb{F}_v}^n : \sigma_v(\mathbf{h} \cdot \nabla G) \geq s\}.$$

*Then, provided  $v$  is coprime to  $d$ , the set  $B_v(G, s)$  is an affine variety, defined by  $O_d(1)$  equations, each of degree  $O_d(1)$ , with*

$$\dim B_v(G, s) \leq n - (s - \sigma_v(G))^+.$$

Next, for a vector  $\mathbf{a} \in \mathbb{Z}^n$ , let us write  $[\mathbf{a}]_p$  for the image of  $\mathbf{a}$  under the natural projection  $\mathbb{Z}^n \rightarrow \mathbb{F}_p^n$ . For consistency, we write  $[\mathbf{a}]_\infty$  for  $\mathbf{a}$ . The following is a simple consequence of Browning and Heath-Brown [3, Lemma 4].

**Lemma 2.5** (Dimension growth bound). *Let  $\mathcal{P} \subset \{p : p \text{ prime}\} \cup \{\infty\}$  be a finite subset. To each  $v \in \mathcal{P}$  we associate an affine variety  $X_v \subset \mathbb{A}_{\mathbb{F}_v}^n$  defined by at most  $D$  equations with coefficients in  $\mathbb{F}_v$ , each of degree at most  $D$ . Suppose that the dimension of  $X_v$  is at most  $k_v$ . Then there exists a constant  $A(D, n) > 0$  such that for any  $T \geq 1$  we have*

$$\#\{\mathbf{a} \in \mathbb{Z}^n \cap [-T, T]^n : [\mathbf{a}]_v \in X_v \text{ for all } v \in \mathcal{P}\} \leq A(D, n)^{|\mathcal{P}|} \sum_{v \in \mathcal{P}} T^{k_v} \prod_{\substack{\mu \in \mathcal{P} \\ k_\mu < k_v}} \mu^{-(k_v - k_\mu)},$$

where we interpret  $\mu^{-1}$  to be 0 when  $\mu = \infty$ .

*Proof.* We describe how to deduce the above from [3, Lemma 4]. Let  $N(T)$  denote the cardinality that is to be estimated. Define the set

$$(2.8) \quad \{p_1, \dots, p_r\} := \{v \in \mathcal{P} : k_v < k_\infty\}.$$

Notice that  $p_1, \dots, p_r$  are all necessarily finite primes. Writing  $\kappa_i$  for  $k_{p_i}$ , let us order the  $p_i$  so that

$$(2.9) \quad \kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_r.$$

We can then apply [3, Lemma 4], with  $l = k_\infty$ , to conclude that there exists a constant  $A(D, n) > 0$  such that

$$N(T) \leq A(D, n)^{r+1} \left( T^l \prod_{i=1}^r p_i^{\kappa_i - l} + \sum_{i=1}^r T^{\kappa_i} \prod_{j=i}^r p_j^{\kappa_j - \kappa_i} \right).$$

By (2.8) we have  $r + 1 \leq |\mathcal{P}|$ . It therefore remains to show that

$$(2.10) \quad T^l \prod_{i=1}^r p_i^{\kappa_i - l} + \sum_{i=1}^r T^{\kappa_i} \prod_{j=i}^r p_j^{\kappa_j - \kappa_i} \leq \sum_{v \in \mathcal{P}} T^{k_v} \prod_{\substack{\mu \in \mathcal{P} \\ k_\mu < k_v}} \mu^{-(k_v - k_\mu)}.$$

We have defined  $l$  to be  $k_\infty$ . Furthermore, by (2.8) we have  $k_\mu < k_\infty$  if and only if  $\mu = p_i$  for some  $i$ . Therefore

$$T^l \prod_{i=1}^r p_i^{\kappa_i - l} = T^{k_\infty} \prod_{\substack{\mu \in \mathcal{P} \\ k_\mu < k_\infty}} \mu^{-(k_\infty - k_\mu)}.$$

Next, fix  $i \in \{1, \dots, r\}$ . Then by (2.9) we have  $k_\mu < k_{p_i}$  if and only if  $\mu = p_j$  for some  $j > i$ . Thus

$$T^{\kappa_i} \prod_{j=i}^r p_j^{\kappa_j - \kappa_i} = T^{\kappa_i} \prod_{j>i} p_j^{\kappa_j - \kappa_i} = T^{k_{p_i}} \prod_{\substack{\mu \in \mathcal{P} \\ k_\mu < k_{p_i}}} \mu^{-(k_{p_i} - k_\mu)}.$$

We have shown that each term of summation on the left-hand side of (2.10) has an identical term of summation on the right-hand side, which therefore completes the derivation of our lemma from [3, Lemma 4]. □

### 3. Exponential sum estimates

This section is the heart of our paper and is concerned with estimating a very general family of multi-dimensional exponential sums with polynomial arguments. We begin by introducing the following class of weight functions.

**Definition 3.1** (Smooth weights  $\mathcal{S}^+(\mathbf{c})$ ). Let  $\mathbf{c} = (c, c_0, c_1, \dots)$  be an increasing infinite tuple of positive absolute constants which are *super-exponential* in the sense that for any non-negative integers  $i$  and  $j$  we have  $c_{i+j} \geq c_i c_j$ . We define  $\mathcal{S}^+(\mathbf{c})$  to be the set of smooth weight functions  $\omega : \mathbb{R}^n \rightarrow [0, \infty)$  satisfying

- (i)  $\text{supp}(\omega) \subset [-c, c]^n$ ,
- (ii) for any  $\epsilon \in (\mathbb{N} \cup \{0\})^n$  we have  $\|\partial^\epsilon \omega\|_{L^\infty(\mathbb{R}^n)} \leq c_{\epsilon_1 + \dots + \epsilon_n}$ .

Of central concern to us is the *exponential sum*

$$S(\alpha, P) := \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega(\mathbf{x}/P) e(\alpha f(\mathbf{x})),$$

where  $\omega \in \mathcal{S}^+(\mathbf{c})$  and  $f \in \mathbb{Z}[x_1, \dots, x_n]$  is the *underlying polynomial*. Throughout, we write  $f^{[k]}$  for the homogeneous part of  $f$  of degree  $k$ . The height of  $f$ , written  $\text{Height}(f)$ , is the maximum absolute value of the coefficients of  $f$ . We henceforth assume that the underlying polynomial  $f$  has degree at most  $d$ , with leading form  $f^{[d]}$  having singular locus (2.6) over  $\mathbb{Q}$  of dimension  $\sigma := \sigma_\infty(f^{[d]})$ . In the statement of all results in this section we assume that  $\alpha \in \mathbb{T}$  and  $q \in \mathbb{N}$  are primitive.

**Remark 3.2** (Implicit constants). Throughout this section, all implicit constants may depend on  $\epsilon, d, n$  and  $c_i$ , where  $c_i$  is a term of the super-exponential sequence appearing in  $\mathcal{S}^+(\mathbf{c})$ . We determine  $\mathbf{c}$  for our particular choice of  $\omega$  in Section 5, from which it follows that  $c_i = O_{i,F}(1)$ , where  $F$  is the form occurring in Theorem 1.1 or 1.2. However, since the  $c_i$  increase with  $i$ , we emphasise that the dependence of any implicit constant on  $c_i$  is subject to the caveat that  $i = O_{d,n}(1)$ , which can be guaranteed in all our results.



Our first estimate is the classical Weyl bound for  $S(\alpha, P)$ . However, unlike the standard treatment found in, say, [1] or [4], we work with smoothly weighted exponential sums and we do not assume that  $\text{Height}(f^{[d]}) = O(1)$ .

**Lemma 3.3** (Weyl bound). *Suppose that  $\text{Height}(f^{[d]}) \leq H$ . Then*

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll (\log P)^n (P^{1-d} + \|q\alpha\|H + qP^{-d} + \min\{Hq^{-1}, \frac{1}{\|q\alpha\|P^d}\})^{\frac{n-\sigma}{d-1}}.$$

Note that  $\sigma = n$  when  $\deg(f) < d$ , so that this estimate is trivial in that case. Rather than giving a suitably modified sketch of the standard proof, we opt for a detailed account based on van der Corput differencing. This affords us the opportunity to introduce, in a less technical context, some of the key ideas behind our later arguments.

*Proof of Lemma 3.3.* Let  $1 \leq H_1, \dots, H_{d-1} \leq P$  be parameters to be determined later. Set  $\chi(\mathbf{x}) := \omega(\mathbf{x}/P)e(\alpha f(\mathbf{x}))$ . Changing variables and averaging gives the identity

$$S(\alpha, P) = \frac{1}{[H_1]^n} \sum_{1 \leq \mathbf{h} \leq H_1} \sum_{\mathbf{x} \in \mathbb{Z}^n} \chi(\mathbf{x} + \mathbf{h}).$$

The number of  $\mathbf{x} \in \mathbb{Z}^n$  for which there exists some  $1 \leq \mathbf{h} \leq H_1$  with  $\chi(\mathbf{x} + \mathbf{h}) \neq 0$  is  $O(P^n)$ . Interchanging the order of summation and applying Cauchy–Schwarz, it follows that

$$|S(\alpha, P)|^2 \ll \frac{P^n}{H_1^{2n}} \sum_{1 \leq \mathbf{h}, \mathbf{h}' \leq H_1} \sum_{\mathbf{x} \in \mathbb{Z}^n} \chi(\mathbf{x} + \mathbf{h}) \overline{\chi(\mathbf{x} + \mathbf{h}')}.$$

Let

$$(3.1) \quad \omega_{\mathbf{h}}(\mathbf{x}) := \omega(\mathbf{x} + \mathbf{h})\omega(\mathbf{x}) \quad \text{and} \quad f_{\mathbf{h}}(\mathbf{x}) := f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}).$$

Applying the triangle inequality, it follows that

$$|S(\alpha, P)|^2 \ll \frac{P^n}{H_1^{2n}} \sum_{\mathbf{h}_1} r_{H_1}(\mathbf{h}_1) \left| \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega_{\mathbf{h}_1/P}(\mathbf{x}/P) e(\alpha(f_{\mathbf{h}_1}(\mathbf{x}))) \right|,$$

where

$$r_{H_1}(\mathbf{h}_1) := \#\{(\mathbf{h}, \mathbf{h}') : 1 \leq \mathbf{h}, \mathbf{h}' \leq H_1 \text{ and } \mathbf{h} - \mathbf{h}' = \mathbf{h}_1\}.$$

Notice that  $r_{H_1}(\mathbf{h}_1) \leq H_1^n$  and  $r_{H_1}(\mathbf{h}_1) = 0$  if  $|\mathbf{h}_1| \geq H_1$ . Thus

$$(3.2) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^2 \ll \frac{1}{H_1^n} \sum_{-H_1 < \mathbf{h}_1 < H_1} \left| \frac{S_{\mathbf{h}_1}(\alpha, P)}{P^n} \right|,$$

where

$$(3.3) \quad S_{\mathbf{h}_1}(\alpha, P) := \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega_{\mathbf{h}_1/P}(\mathbf{x}/P) e(\alpha(f_{\mathbf{h}_1}(\mathbf{x}))).$$

We call the derivation of (3.2) the method of *van der Corput differencing*. We claim that  $\omega_{\mathbf{h}_1/P} \in \mathcal{S}^+(\mathbf{c}')$  with  $c' = c$  and  $c'_m = c_m 2^m$ . Since  $\omega_{\mathbf{h}_1/P}$  is a product of two non-negative smooth functions, it is itself non-negative and smooth. Since one of the factors which

comprise  $\omega_{\mathbf{h}_1/P}$  is  $\omega$ , we have

$$\text{supp}(\omega_{\mathbf{h}_1/P}) \subset \text{supp}(\omega) \subset [-c, c]^n.$$

Finally, by the product rule for differentiation and the super-exponential nature of the  $c_m$ , for any  $\epsilon \in (\mathbb{N} \cup \{0\})^n$  with  $\epsilon_1 + \dots + \epsilon_n = m$  we have

$$\begin{aligned} |\partial^\epsilon \omega_{\mathbf{h}_1/P}(\mathbf{x})| &\leq \sum_{\lambda + \mu = \epsilon} \binom{\epsilon_1}{\lambda_1} \cdots \binom{\epsilon_n}{\lambda_n} |\partial^\lambda \omega(\mathbf{x} + P^{-1}\mathbf{h}_1) \partial^\mu \omega(\mathbf{x})| \\ &\leq c_m \sum_{\lambda + \mu = \epsilon} \binom{\epsilon_1}{\lambda_1} \cdots \binom{\epsilon_n}{\lambda_n} \\ &= c_m 2^m. \end{aligned}$$

The claim follows.

Let us define  $\omega_{\mathbf{h}_1, \dots, \mathbf{h}_r}$  via (3.1) with  $\omega = \omega_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}$  and  $\mathbf{h} = \mathbf{h}_r$ . Then by induction and our previous claim, we have

$$\omega_{(\mathbf{h}_1, \dots, \mathbf{h}_r)/P} \in \mathcal{S}^+(\mathbf{c}')$$

with  $\mathbf{c}'$  the super-exponential sequence given by  $c' = c$  and  $c'_m = c_m 2^{rm}$ .

Define  $f_{\mathbf{h}_1, \dots, \mathbf{h}_r}$  analogously, so that if  $g = f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}$  is defined, then we set

$$f_{\mathbf{h}_1, \dots, \mathbf{h}_r}(\mathbf{x}) := g_{\mathbf{h}_r}(\mathbf{x}) = f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}(\mathbf{x} + \mathbf{h}_r) - f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}(\mathbf{x}).$$

Notice that

$$f_{\mathbf{h}_1, \mathbf{h}_2}(\mathbf{x}) = f(\mathbf{x} + \mathbf{h}_1 + \mathbf{h}_2) - f(\mathbf{x} + \mathbf{h}_1) - f(\mathbf{x} + \mathbf{h}_2) + f(\mathbf{x}) = f_{\mathbf{h}_2, \mathbf{h}_1}(\mathbf{x}).$$

It follows from this, and the inductive definition, that  $f_{\mathbf{h}_1, \dots, \mathbf{h}_r}$  is invariant under permutations of the  $\mathbf{h}_i$ . Furthermore, by Taylor's formula we have

$$(3.4) \quad f_{\mathbf{h}_1, \dots, \mathbf{h}_r}^{[d-r]} = \mathbf{h}_r \cdot \nabla f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}^{[d-(r-1)]}.$$

Consequently,  $f_{\mathbf{h}_1, \dots, \mathbf{h}_r}^{[d-r]}$  is linear in each  $\mathbf{h}_i$ .

Iterating the argument that led to (3.2) and (3.3), we deduce that for each  $1 \leq r \leq d-1$  we have

$$(3.5) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^{2^r} \ll \frac{1}{(H_1 \cdots H_r)^n} \sum_{-H_1 < \mathbf{h}_1 < H_1} \cdots \sum_{-H_r < \mathbf{h}_r < H_r} \left| \frac{S_{\mathbf{h}_1, \dots, \mathbf{h}_r}(\alpha, P)}{P^n} \right|,$$

where

$$S_{\mathbf{h}_1, \dots, \mathbf{h}_r}(\alpha, P) := \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega_{(\mathbf{h}_1, \dots, \mathbf{h}_r)/P}(\mathbf{x}/P) e(\alpha(f_{\mathbf{h}_1, \dots, \mathbf{h}_r}(\mathbf{x}))).$$

Since  $f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]}(\mathbf{x})$  is linear in  $\mathbf{x}$ , we have

$$f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}(\mathbf{x}) = \sum_i (\partial^{e_i} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]}) x_i + c_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}},$$

where  $\partial^{e_i} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]}$  and  $c_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}$  are integers independent of  $\mathbf{x}$ . Define the function

$$\varphi : [-cP, cP]^n \cap \mathbb{Z}^n \rightarrow \mathbb{C}$$

via  $\varphi(\mathbf{x}) = e(\alpha f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}(\mathbf{x}))$ . Then, in the notation of Lemma 2.1, we have

$$\begin{aligned} |T_\varphi(\mathbf{t})| &= \left| \prod_{i=1}^n \sum_{-cP \leq x_i \leq t_i} e(\alpha \partial^{\mathbf{e}_i} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]} x_i) \right| \\ &\ll \prod_{i=1}^n \min\{P, \|\alpha \partial^{\mathbf{e}_i} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]}\|^{-1}\}. \end{aligned}$$

Let  $g(\mathbf{t}) = \omega_{(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})/P}(\mathbf{t}/P)$ . For  $\epsilon_1 + \dots + \epsilon_n = m$ , it follows from the chain rule that

$$|\partial^\epsilon g(\mathbf{t})| \leq P^{-m} \|\partial^\epsilon \omega_{(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})/P}\|_{L^\infty(\mathbb{R}^n)} \leq c_m 2^{(d-1)m} P^{-m}.$$

Hence partial summation (Lemma 2.1) yields the existence of  $\epsilon \in \{0, 1\}^n$  such that

$$\begin{aligned} S_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}(\alpha, P) &\ll \frac{P^{\epsilon_1 + \dots + \epsilon_n}}{P^n} \int_{[-cP, cP]^n} |\partial^\epsilon g(N\bar{\epsilon} + \mathbf{t}_\epsilon) T_\varphi(N\bar{\epsilon} + \mathbf{t}_\epsilon)| d\mathbf{t} \\ &\ll \prod_{j=1}^n \min\{P, \|\alpha \partial^{\mathbf{e}_j} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]}\|^{-1}\}. \end{aligned}$$

Write  $\underline{\mathbf{h}}$  as a shorthand for the vector  $(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})$  and write

$$M(\underline{\mathbf{h}}) := (\partial^{\mathbf{e}_j} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]})_{1 \leq j \leq n}.$$

Let us view the vector  $\alpha M(\underline{\mathbf{h}})$  as an element of the torus  $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ . Sub-dividing this torus into sub-cubes of side-length  $P^{-1}$ , each vector  $\alpha M(\underline{\mathbf{h}})$  has  $j$ th coordinate lying in an interval  $[\frac{r_j}{P}, \frac{r_j+1}{P})$ , for some  $\mathbf{r} \in \mathbb{Z}^n$  with  $0 \leq r_j < P$ . Let  $R(\mathbf{r})$  denote this region. If  $\alpha M(\underline{\mathbf{h}}) \in R(\mathbf{r})$ , then

$$\|\alpha M(\underline{\mathbf{h}})_j\| \geq \frac{r_j}{P} \quad \text{for each } j.$$

Letting

$$T(\mathbf{r}) := \{\underline{\mathbf{h}} : \alpha M(\underline{\mathbf{h}}) \in R(\mathbf{r}) \text{ and } |\mathbf{h}_i| < H_i \text{ for } 1 \leq i \leq d-1\},$$

we have

$$\begin{aligned} \sum_{|\mathbf{h}_1| < H_1} \dots \sum_{|\mathbf{h}_{d-1}| < H_{d-1}} \prod_{j=1}^n \min\{1, (P \|\alpha M(\underline{\mathbf{h}})_j\|)^{-1}\} &\leq \sum_{\mathbf{0} \leq \mathbf{r} < P} \#T(\mathbf{r}) \prod_{j=1}^n \min\{1, \frac{1}{r_j}\} \\ &\ll (\log P)^n \max_{\mathbf{r}} \#T(\mathbf{r}). \end{aligned}$$

Define

$$N_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}(\mathbf{r}) := \#\{\mathbf{h}_{d-1} \in \mathbb{Z}^n : \underline{\mathbf{h}} \in T(\mathbf{r})\}$$

and, in a similar fashion, let  $n_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}$  denote the number of integer vectors  $\mathbf{h}_{d-1}$  such that  $|\mathbf{h}_{d-1}| < H_{d-1}$  and  $\|\alpha M(\underline{\mathbf{h}})_j\| \leq P^{-1}$  for all  $j$ . Now if the integer vectors  $\mathbf{h}_{d-1}$  and  $\mathbf{h}'_{d-1}$  are counted by  $N_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}(\mathbf{r})$ , then  $\mathbf{h}'_{d-1} - \mathbf{h}_{d-1}$  is counted by  $n_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}$ , whence

$$N_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}(\mathbf{r}) \leq n_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}} \quad \text{for any } \mathbf{r} \in \mathbb{Z}^n.$$

It therefore follows that  $\#T(\mathbf{r}) \ll \mathcal{M}_{H_1, \dots, H_{d-1}}(P^{-1})$ , where

$$\mathcal{M}_{H_1, \dots, H_{d-1}}(P^{-1}) := \#\{\underline{\mathbf{h}} \in \mathbb{Z}^{(d-1)n} : |\mathbf{h}_i| < H_i \text{ and } \|\alpha M(\underline{\mathbf{h}})_j\| < P^{-1} \text{ for all } i, j\}.$$

Combining these deliberations, we deduce that

$$(3.6) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll \left( \frac{\log P}{H_1 \cdots H_{d-1}} \right)^n \mathcal{M}_{H_1, \dots, H_{d-1}}(P^{-1}).$$

Next, we claim that the linear map  $\mathbf{h}_i \mapsto M(\mathbf{h})$  has a symmetric matrix (with respect to the standard basis). Since  $M(\mathbf{h}) = M(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})$  is invariant under permutations of the  $\mathbf{h}_i$ , it suffices to establish the claim when  $i = d - 1$ . By (3.4) and linearity of differentiation, we have

$$(3.7) \quad \begin{aligned} M(\mathbf{h}_1, \dots, \mathbf{h}_{d-1}) &= (\partial^{\mathbf{e}_i}(\mathbf{h}_{d-1} \cdot \nabla f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]}))_{1 \leq i \leq n} \\ &= \mathbf{h}_{d-1} \cdot (\partial^{\mathbf{e}_i + \mathbf{e}_j} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \end{aligned}$$

Since  $\mathbf{e}_i + \mathbf{e}_j = \mathbf{e}_j + \mathbf{e}_i$ , the claim follows.

We can therefore apply the shrinking lemma (Lemma 2.2) to each block of variables  $\mathbf{h}_i$  in  $\mathbf{h}$  and conclude that for any  $\theta_1, \dots, \theta_{d-1} \in (0, 1]$  we have

$$\mathcal{M}_{H_1, \dots, H_{d-1}}(P^{-1}) \ll \frac{\mathcal{M}_{\theta_1 H_1, \dots, \theta_{d-1} H_{d-1}}(\theta_1 \cdots \theta_{d-1} P^{-1})}{(H_1^{-1} + \theta_1)^n \cdots (H_{d-1}^{-1} + \theta_{d-1})^n}.$$

By iterating (3.4), one can check that we have the formula

$$M(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})_j = \sum_{1 \leq i_1 \leq n} \cdots \sum_{1 \leq i_{d-1} \leq n} h_{1, i_1} \cdots h_{d-1, i_{d-1}} \partial^{\mathbf{e}_{i_1} + \cdots + \mathbf{e}_{i_{d-1}} + \mathbf{e}_j} f^{[d]}.$$

Hence there exists a constant  $C = O_{n,d}(1)$  such that if  $|\mathbf{h}_i| < \theta_i H_i$  for all  $i$ , then

$$|M(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})_j| < CH(\theta_1 H_1) \cdots (\theta_{d-1} H_{d-1}).$$

Let us choose  $\theta_1, \dots, \theta_{d-1} \in (0, 1]$  so that

$$(3.8) \quad \theta_1 \cdots \theta_{d-1} = \min\left\{1, \frac{1}{2\|q\alpha\|CHH_1 \cdots H_{d-1}}, \frac{P}{2q}, \max\left\{\frac{q}{CHH_1 \cdots H_{d-1}}, \|q\alpha\|P\right\}\right\}.$$

It follows that if  $|\mathbf{h}_i| < \theta_i H_i$  for all  $i$  and  $\|\alpha M(\mathbf{h})_j\| < \theta_1 \cdots \theta_{d-1} P^{-1}$  for all  $j$ , then

- (i)  $|M(\mathbf{h})_j| < \frac{1}{2}\|q\alpha\|^{-1}$ ,
- (ii)  $\|\alpha M(\mathbf{h})_j\| < \frac{1}{2}q^{-1}$ ,
- (iii)  $|M(\mathbf{h})_j| < q$  or  $\|\alpha M(\mathbf{h})_j\| < \|q\alpha\|$ .

Applying Lemma 2.3, we deduce that  $M(\mathbf{h})_j = 0$  for  $j = 1, \dots, n$ . Incorporating this into (3.6) we obtain the estimate

$$(3.9) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll (\log P)^n \frac{L_f(\theta_1 H_1, \dots, \theta_{d-1} H_{d-1})}{(1 + \theta_1 H_1)^n \cdots (1 + \theta_{d-1} H_{d-1})^n},$$

with  $\theta_1 \cdots \theta_{d-1}$  as in (3.8) and where

$$L_f(H_1, \dots, H_{d-1}) := \#\{\mathbf{h} \in \mathbb{Z}^{(d-1)n} : |\mathbf{h}_i| < H_i \text{ and } \nabla f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-1}}^{[1]} = \mathbf{0}\}.$$

We are therefore led to the estimation of  $L_f(H_1, \dots, H_{d-1})$ .

Using the notation (2.7), we may partition the set of  $\mathbf{h}_1 \in (-H_1, H_1)^n \cap \mathbb{Z}^n$  according to the value of  $\sigma_\infty(f_{\mathbf{h}_1}^{[d-1]})$ . Using this and the pigeon-hole principle, we deduce the existence of an integer  $s_1 \in [0, n]$ , a set  $\mathcal{H}_1 \subset (-H_1, H_1)^n \cap \mathbb{Z}^n$  and  $\mathbf{h}_1 \in \mathcal{H}_1$  such that both of the following hold:

- (i) For each  $\mathbf{h}'_1 \in \mathcal{H}_1$  we have  $\sigma_\infty(f_{\mathbf{h}'_1}^{[d-1]}) = s_1$ .
- (ii) One has  $L_f(H_1, \dots, H_{d-1}) \ll |\mathcal{H}_1| L_{f_{\mathbf{h}_1}}(H_2, \dots, H_{d-1})$ .

Iterating this process, we can find integers  $s_1, \dots, s_{d-2}$ , sets  $\mathcal{H}_1, \dots, \mathcal{H}_{d-2}$  and fixed elements  $\mathbf{h}_i \in \mathcal{H}_i$  such that

(i) For each  $\mathbf{h}'_i \in \mathcal{H}_i$  we have  $\sigma_\infty(f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}, \mathbf{h}'_i}^{[d-i]}) = s_i$ .

(ii) One has

$$(3.10) \quad L_f(H_1, \dots, H_{d-1}) \ll |\mathcal{H}_1| \cdots |\mathcal{H}_{d-2}| L_{f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}}(H_{d-1}).$$

By Euler's identity, the singular locus of  $f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]}$  is the set of  $\mathbf{x}$  such that

$$0 = \partial^{e_i} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]}(\mathbf{x}) = \sum_{j=1}^n (\partial^{e_i + e_j} f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]}) x_j \quad \text{for each } i \in \{1, \dots, n\}.$$

Hence by (3.7), the nullity of the linear map  $\mathbf{h}_{d-1} \mapsto M(\mathbf{h}_1, \dots, \mathbf{h}_{d-1})$  (over  $\mathbb{Q}$ ) coincides with the dimension

$$s_{d-2} = \sigma_\infty(f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}^{[2]}).$$

By elementary linear algebra (inducting on the nullity) one can show that for a given linear map  $T : \mathbb{C}^n \rightarrow \mathbb{C}^m$  and any  $\mathbf{y} \in \mathbb{C}^m$  we have the uniform estimate

$$\#\{\mathbf{x} \in (-P, P)^n \cap \mathbb{Z}^n : T\mathbf{x} = \mathbf{y}\} \ll_n P^{n - \text{rank}(T)}.$$

Hence for  $H_1 \geq 1$  we have

$$(3.11) \quad L_{f_{\mathbf{h}_1, \dots, \mathbf{h}_{d-2}}}(H_{d-1}) \ll H_{d-1}^{s_{d-2}}.$$

For each  $1 \leq i \leq d-2$ , define the algebraic variety

$$X_{i, \infty} := \{\mathbf{y} \in \mathbb{A}_{\mathbb{Q}}^n : \sigma_\infty(f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}, \mathbf{y}}^{[d-i]}) \geq s_i\}.$$

From (3.4) we have

$$f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}, \mathbf{y}}^{[d-i]} = \mathbf{y} \cdot \nabla f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}}^{[d-i+1]}.$$

We may therefore employ Lemma 2.4 with  $G = f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}}^{[d-i+1]}$  to deduce that  $X_{i, \infty}$  is an affine algebraic variety defined by  $O_{d,n}(1)$  equations of degree  $O_{d,n}(1)$ . Moreover, setting  $s_0 := \sigma$ , we see that the dimension of  $X_{i, \infty}$  is at most  $n - (s_i - s_{i-1})^+$ . Taking  $k_\infty = n - (s_i - s_{i-1})^+$  and  $\mathcal{P} = \{\infty\}$  in Lemma 2.5, we conclude that for  $H_i \geq 1$  we have

$$|\mathcal{H}_i| \ll H_i^{n - (s_i - s_{i-1})^+}.$$

Combining this estimate for  $|\mathcal{H}_i|$  with (3.10) and (3.11), we deduce that for  $H_i \geq 1$  there exist integers  $s_1, \dots, s_{d-2}$  such that

$$L_f(H_1, \dots, H_{d-1}) \ll H_1^{n - (s_1 - s_0)^+} \cdots H_{d-2}^{n - (s_{d-2} - s_{d-3})^+} H_{d-1}^{s_{d-2}}.$$

Set  $B_i := 1 + \theta_i H_i$ . Using this in (3.9), we see that for any  $H_1, \dots, H_{d-1}$  belonging to the interval  $[1, P]$  and any  $\theta_1, \dots, \theta_{d-1} \in (0, 1]$  whose product is equal to (3.8), we have

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll (\log P)^n B_1^{-(s_1 - s_0)^+} \cdots B_{d-2}^{-(s_{d-2} - s_{d-3})^+} B_{d-1}^{-(n - s_{d-2})}.$$

As we have no control over the values of the integers  $s_i$ , to proceed any further we must impose the condition that

$$\theta_1 H_1 = \theta_2 H_2 = \dots = \theta_{d-1} H_{d-1}.$$

Then since  $s_0 = \sigma$  we have

$$B_1^{-(s_1-s_0)^+} \dots B_{d-2}^{-(s_{d-2}-s_{d-3})^+} B_{d-1}^{-(n-s_{d-2})} \leq B_1^{-(n-\sigma)}.$$

Notice that  $B_1 \geq \theta_1 H_1$ . Moreover, writing  $\tilde{H} = H_1 \dots H_{d-1}$  it follows from (3.8) that

$$\theta_1 H_1 = (\theta_1 \dots \theta_{d-1} \tilde{H})^{\frac{1}{d-1}} \asymp \min \left\{ \tilde{H}, \frac{1}{\|q\alpha\|H}, \frac{P\tilde{H}}{q}, \max \left\{ \frac{q}{\tilde{H}}, \tilde{H} \|q\alpha\|P \right\} \right\}^{\frac{1}{d-1}}.$$

Thus we arrive at the estimate

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll (\log P)^n \max \left\{ \frac{1}{\tilde{H}}, \|q\alpha\|H, \frac{q}{P\tilde{H}}, \min \left\{ \frac{H}{q}, \frac{1}{\tilde{H}\|q\alpha\|P} \right\} \right\}^{\frac{n-\sigma}{d-1}}.$$

This bound is minimised by taking  $H_1 = \dots = H_{d-1} = P$ , which yields

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll (\log P)^n \max \left\{ \frac{1}{P^{d-1}}, \|q\alpha\|H, \frac{q}{P^d}, \min \left\{ \frac{H}{q}, \frac{1}{\|q\alpha\|P^d} \right\} \right\}^{\frac{n-\sigma}{d-1}},$$

which thereby completes the proof of the lemma. □

When the exponential sum  $S(\alpha, P)$  involves a cubic polynomial, one can get better estimates by applying Poisson summation instead of repeated applications of van der Corput differencing. This is one of the key innovations in Heath-Brown [7], for example, and it also played a critical role in [3].

Given a fixed positive integer  $q$  we reserve the symbols  $b, c_1$  and  $c_2$  for the quantities

$$(3.12) \quad b := \prod_{\substack{p^e \parallel q \\ e \leq 2}} p^e, \quad c_1 := \prod_{\substack{p^e \parallel q \\ e > 2}} p^{\lfloor \frac{e}{2} \rfloor}, \quad c_2 := \prod_{\substack{p^e \parallel q \\ e > 2, 2 \nmid e}} p,$$

so that  $q = bc_1^2 c_2$ . Define the  $r$ -values of  $f$  with respect to  $q$  to be the numbers

$$(3.13) \quad r_i = r_i^{[d]}(f, q) := \prod_{\substack{p^e \parallel bd \\ \sigma_p(f^{[d]})=i}} p^e, \quad R_\xi = R_\xi^{[d]}(f, q) := \prod_{\xi \leq i \leq n} r_i^{i-\xi}.$$

Define the relative height of  $f$  with respect to  $P$  at scale  $d$  by

$$\text{Height}_{P,d}(f) := \text{Height}(P^{-d} f(P\mathbf{x})).$$

Notice that for  $P \geq 1$  and  $d \geq \deg(f)$  we have

$$\text{Height}(f^{[d]}) = \text{Height}_{P,d}(f^{[d]}) \leq \text{Height}_{P,d}(f) \leq \text{Height}(f).$$

Bearing this notation in mind, the following result is a reformulation of the key estimate in Browning and Heath-Brown [3].

**Lemma 3.4** (Cubic Poisson bound). *Suppose that  $\deg(f) \leq 3$  and  $H$  is such that*

$$\text{Height}_{P,3}(f) \leq H \leq P^{O(1)}.$$

*Then for  $\|q\alpha\| \leq P^{-1}$  and  $\zeta$  in the range  $\sigma_\infty(f^{[3]}) \leq \zeta \leq n$  we have*

$$S(\alpha, P) \ll R_\zeta^{[3]}(f, q)^{\frac{1}{2}} P^{n+\varepsilon} \left( \frac{\sqrt{q}}{P} + \sqrt{\|q\alpha\|PH} + \frac{H^{1/6} \min\{c_1, c_2 H\}^{\frac{1}{6}}}{b^{\frac{1}{2}} (c_1 c_2)^{\frac{1}{3}}} \right)^{n-\zeta}.$$

*Proof.* The statement of the lemma is trivial if  $q > P^2$  or if  $\deg(f) < 3$ , since then we are obliged to take  $\zeta = \sigma_\infty(f^{[3]}) = n$ . Suppose henceforth that  $q \leq P^2$  and  $\deg(f) = 3$ . Let

$$V := \max\{qP^{-1}, \sqrt{q\|q\alpha\|HP}\}$$

and

$$W := V + \min\{(c_1^2 c_2 H)^{\frac{1}{3}}, (c_1 V)^{\frac{1}{2}} + c_1^{\frac{5}{6}} H^{\frac{1}{6}}\}.$$

Then by [3, Proposition 2] we have

$$S(\alpha, P) \ll R_\zeta^{\frac{1}{2}} P^{n+\varepsilon} (Wq^{-\frac{1}{2}})^{n-\zeta}.$$

Using  $q = bc_1^2 c_2$ , we see that

$$Wq^{-\frac{1}{2}} \ll Vq^{-\frac{1}{2}} + \min\left\{\frac{H^{\frac{1}{3}}}{b^{\frac{1}{2}} c_1^{\frac{1}{3}} c_2^{\frac{1}{6}}}, \sqrt{\frac{V}{bc_1 c_2}}\right\} + \min\left\{\frac{H^{\frac{1}{3}}}{b^{\frac{1}{2}} c_1^{\frac{1}{3}} c_2^{\frac{1}{6}}}, \frac{H^{\frac{1}{6}}}{b^{\frac{1}{2}} c_1^{\frac{1}{6}} c_2^{\frac{1}{2}}}\right\},$$

with

$$Vq^{-\frac{1}{2}} \leq \frac{\sqrt{q}}{P} + \sqrt{\|q\alpha\|HP}.$$

It therefore suffices to establish that

$$(3.14) \quad \min\left\{\frac{H^{\frac{1}{3}}}{b^{\frac{1}{2}} c_1^{\frac{1}{3}} c_2^{\frac{1}{6}}}, \sqrt{\frac{V}{bc_1 c_2}}\right\} \ll Vq^{-\frac{1}{2}} + \min\left\{\frac{H^{\frac{1}{3}}}{b^{\frac{1}{2}} c_1^{\frac{1}{3}} c_2^{\frac{1}{6}}}, \frac{H^{\frac{1}{6}}}{b^{\frac{1}{2}} c_1^{\frac{1}{6}} c_2^{\frac{1}{3}}}\right\}.$$

If

$$\frac{H^{\frac{1}{3}}}{b^{\frac{1}{2}} c_1^{\frac{1}{3}} c_2^{\frac{1}{6}}} \leq \frac{H^{\frac{1}{6}}}{b^{\frac{1}{2}} c_1^{\frac{1}{6}} c_2^{\frac{1}{3}}},$$

then (3.14) follows immediately. We may therefore assume that the opposite inequality holds, or equivalently (after re-arrangement), that

$$(3.15) \quad Hc_2 > c_1.$$

In this case, (3.14) follows if we can prove that

$$\sqrt{\frac{V}{bc_1 c_2}} \leq \max\left\{\frac{V}{b^{\frac{1}{2}} c_1 c_2^{\frac{1}{2}}}, \frac{H^{\frac{1}{6}}}{b^{\frac{1}{2}} c_1^{\frac{1}{6}} c_2^{\frac{1}{3}}}\right\}.$$

By the trivial inequality  $\max\{X, Y\} \geq \sqrt{XY}$ , the right-hand side is at least

$$\left(\frac{V}{bc_1 c_2}\right)^{\frac{1}{2}} \left(\frac{Hc_2}{c_1}\right)^{\frac{1}{12}},$$

so that the desired condition now follows from (3.15).  $\square$

In Lemma 3.3 we gave a detailed account of how  $d - 1$  applications of van der Corput differencing can be used to transform the exponential sum  $S(\alpha, P)$  into a family of linear exponential sums associated to parameters  $H_1, \dots, H_{d-1} \in [1, P]$ , which we could ultimately estimate rather well. In the final stages of the argument we made the specialisation

$$H_1 = \dots = H_{d-1} = P,$$

which brought us to the usual Weyl estimate (as found in [4]). In the next result, we consider the effect of van der Corput differencing  $d - k$  times only, for any  $1 \leq k \leq d$ . Rather than ending up with something of the form (3.5), with  $r = d - k$ , it turns out that it will be more efficient to make a judicious application of the pigeon-hole principle at each differencing step separately, in order to control the singular locus of the underlying polynomial.

**Lemma 3.5** (van der Corput process). *Suppose that  $q \leq P^{O(1)}$  and  $\text{Height}(f) = O(1)$ . For each  $1 \leq Q \leq P$  and  $1 \leq k \leq d$ , there exists an exponential sum  $T(\alpha, P)$ , with underlying polynomial  $g$  of degree at most  $k$  and*

$$\text{Height}_{P,k}(g) \ll Q^{2-2^{k+1-d}},$$

such that for some  $\zeta \geq \max\{\sigma, \sigma_\infty(g^{[k]})\}$  we have

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-k}} \ll \frac{P^\varepsilon}{Q^{\zeta-\sigma} \sqrt{R_\zeta^{[k]}(g, q)}} \left| \frac{T(\alpha, P)}{P^n} \right|.$$

Moreover,  $T(\alpha, P)$  has weight in  $\mathcal{S}^+(\mathbf{c}')$  where  $\mathbf{c}'$  is the super-exponential sequence given by  $c' = c$  and  $c'_m = c_m 2^{(d-k)m}$ .

*Proof.* After a single iteration of van der Corput differencing, as in (3.2) and (3.3), we deduce that for any real  $H_1 \in [1, P]$  we have the bound

$$\left| \frac{S(\alpha, P)}{P^n} \right|^2 \ll \frac{1}{H_1^n} \sum_{-H_1 < \mathbf{h}_1 < H_1} \left| \frac{S_{\mathbf{h}_1}(\alpha, P)}{P^n} \right|,$$

where

$$S_{\mathbf{h}_1}(\alpha, P) := \sum_{\mathbf{x} \in \mathbb{Z}^n} \omega_{\mathbf{h}_1/P}(\mathbf{x}/P) e(\alpha(f_{\mathbf{h}_1}(\mathbf{x}))).$$

With  $b$  and  $c_2$  defined as in (3.12), let

$$\mathcal{V} := \{p : p \mid bc_2\} \cup \{\infty\}.$$

Notice that

$$(\sigma_v(f_{\mathbf{h}_1}^{[d-1]}))_{v \in \mathcal{V}} \in \prod_{v \in \mathcal{V}} \{0, 1, \dots, n\}.$$

Since  $|\mathcal{V}| \leq \omega(q) + 1$ , we see that there are at most  $O_{n,\varepsilon}(q^\varepsilon)$  choices for  $(\sigma_v(f_{\mathbf{h}_1}^{[d-1]}))_{v \in \mathcal{V}}$ . It follows from the pigeon-hole principle that there exists a tuple of integers  $\mathbf{s}_1$  and a set  $\mathcal{H}_1 \subset (-H_1, H_1)^n \cap \mathbb{Z}^n$  such that both of the following hold:

- (i) For each  $\mathbf{h}_1 \in \mathcal{H}_1$  we have  $(\sigma_v(f_{\mathbf{h}_1}^{[d-1]}))_{v \in \mathcal{V}} = \mathbf{s}_1$ .
- (ii) One has

$$\sum_{-H_1 < \mathbf{h}_1 < H_1} |S_{\mathbf{h}_1}(\alpha, P)| \ll q^\varepsilon \sum_{\mathbf{h}_1 \in \mathcal{H}_1} |S_{\mathbf{h}_1}(\alpha, P)|.$$



Since  $q \leq P^{O(1)}$ , we deduce that there exists an  $\mathbf{h}_1 \in \mathcal{H}_1$  satisfying

$$\left| \frac{S(\alpha, P)}{P^n} \right|^2 \ll P^\varepsilon \frac{|\mathcal{H}_1|}{H_1^n} \left| \frac{S_{\mathbf{h}_1}(\alpha, P)}{P^n} \right|.$$

Next, let  $r := d - k$ . Applying the van der Corput differencing process to  $S_{\mathbf{h}_1}(\alpha, P)$  and iterating, we deduce that for any  $1 \leq H_1, \dots, H_r \leq P$  there exist sets  $\mathcal{H}_i \subset (-H_i, H_i)^n \cap \mathbb{Z}^n$  and elements  $\mathbf{h}_i \in \mathcal{H}_i$  such that

$$(3.16) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^{2^r} \ll P^\varepsilon \left( \frac{|\mathcal{H}_1|}{H_1^n} \right)^{2^{r-1}} \left( \frac{|\mathcal{H}_2|}{H_2^n} \right)^{2^{r-2}} \dots \left( \frac{|\mathcal{H}_r|}{H_r^n} \right) \left| \frac{S_{\mathbf{h}_1, \dots, \mathbf{h}_r}(\alpha, P)}{P^n} \right|.$$

Moreover, there exist tuples of integers  $\mathbf{s}_i = (s_{i,v})_{v \in \mathcal{V}}$  such that for any  $\mathbf{h}'_i \in \mathcal{H}_i$  we have

$$(3.17) \quad (\sigma_v(f_{\mathbf{h}_1, \dots, \mathbf{h}_{i-1}, \mathbf{h}'_i}^{[d-i]}))_{v \in \mathcal{V}} = \mathbf{s}_i.$$

For notational convenience, let us define

$$\mathbf{s}_0 := (\sigma_v(f^{[d]}))_{v \in \mathcal{V}}.$$

For any prime  $p$  we have  $s_{0,p} \geq s_{0,\infty}$ , with strict inequality for only finitely many primes. It follows that there exists a constant  $C = O_f(1)$  such that  $C \geq d$  and for any  $p > C$  we have  $s_{0,p} = s_{0,\infty}$ . Set

$$\mathcal{V}_C := \{v \in \mathcal{V} : v > C\}.$$

For each  $v \in \mathcal{V}_C$  define the sets

$$X_v := \{\mathbf{y} \in \mathbb{A}_{\mathbb{F}_v}^n : \sigma_v(f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}, \mathbf{y}}^{[d-r]}) \geq s_{r,v}\}.$$

Notice that

$$f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}, \mathbf{y}}^{[d-r]} = \mathbf{y} \cdot \nabla f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}^{[d-(r-1)]}$$

by (3.4). We may therefore employ Lemma 2.4 with  $G = f_{\mathbf{h}_1, \dots, \mathbf{h}_{r-1}}^{[d-(r-1)]}$ , and deduce that  $X_v$  is an affine algebraic variety defined by  $O_{d,n}(1)$  equations of degree  $O_{d,n}(1)$ . Moreover, the dimension of  $X_v$  is at most  $n - (s_{r,v} - s_{r-1,v})^+$ . Taking  $\mathcal{P} = \mathcal{V}_C$  and  $k_v = n - (s_{r,v} - s_{r-1,v})^+$  in Lemma 2.5, we conclude that

$$\frac{|\mathcal{H}_r|}{H_r^n} \ll q^\varepsilon \sum_{v \in \mathcal{V}_C} H_r^{-(s_{r,v} - s_{r-1,v})^+} \prod_{\substack{p \in \mathcal{V}_C \\ s_{r,p} - s_{r-1,p} \geq \\ (s_{r,v} - s_{r-1,v})^+}} p^{-(s_{r,p} - s_{r-1,p}) + (s_{r,v} - s_{r-1,v})^+}.$$

By the pigeon-hole principle, we see that there exists a  $v \in \mathcal{V}_C$  such that on setting  $s_r := s_{r,v}$  and  $t_r := s_{r-1,v}$  we have

$$(3.18) \quad \frac{|\mathcal{H}_r|}{H_r^n} \ll q^{2\varepsilon} H_r^{-(s_r - t_r)^+} \prod_{\substack{p \in \mathcal{V}_C \\ s_{r,p} - s_{r-1,p} \geq \\ (s_r - t_r)^+}} p^{-(s_{r,p} - s_{r-1,p}) + (s_r - t_r)^+}.$$

Next, define the set

$$\mathcal{V}_C^{(r-1)} := \{\mu \in \mathcal{V}_C : s_{r-1,\mu} \geq t_r\}.$$

Repeating the argument leading to (3.18), we deduce that there is a  $\mu \in \mathcal{V}_C^{(r-1)}$  such that on setting  $s_{r-1} := s_{r-1,\mu}$  and  $t_{r-1} := s_{r-2,\mu}$ , we have

$$\frac{|\mathcal{H}_{r-1}|}{H_{r-1}^n} \ll q^\varepsilon H_{r-1}^{-(s_{r-1}-t_{r-1})^+} \prod_{\substack{p \in \mathcal{V}_C^{(r-1)} \\ s_{r-1,p}-s_{r-2,p} \geq \\ (s_{r-1}-t_{r-1})^+}} p^{-(s_{r-1,p}-s_{r-2,p})+(s_{r-1}-t_{r-1})^+}.$$

Let us write  $\mathcal{V}_C^{(r)}$  for  $\mathcal{V}_C$  and  $t_{r+1} := 0$ . Iterating the above process, we obtain integers  $s_i$  and  $t_i$  for  $1 \leq i \leq r$ , with  $s_i \geq t_{i+1}$ , such that on setting

$$\mathcal{V}_C^{(i)} = \{v \in \mathcal{V}_C : s_{i,v} \geq t_{i+1}\},$$

we have the bound

$$(3.19) \quad \frac{|\mathcal{H}_i|}{H_i^n} \ll q^\varepsilon H_i^{-(s_i-t_i)^+} \prod_{\substack{p \in \mathcal{V}_C^{(i)} \\ s_{i,p}-s_{i-1,p} \geq \\ (s_i-t_i)^+}} p^{-(s_{i,p}-s_{i-1,p})+(s_i-t_i)^+}.$$

Let us set  $\zeta_0 := \sigma$  and for  $i \geq 1$  set

$$\begin{aligned} \zeta_i &:= (s_i - t_i)^+ + \zeta_{i-1} \\ &= (s_i - t_i)^+ + (s_{i-1} - t_{i-1})^+ + \cdots + (s_1 - t_1)^+ + \sigma. \end{aligned}$$

Notice that  $s_{i,v} \geq s_{i,\infty}$  for all  $i$  and  $v$ . Also  $t_1 = s_{0,v}$  for some  $v \in \mathcal{V}_C$ , and by our choice of  $C$  this means that  $t_1 = \sigma$ . We claim that

$$(3.20) \quad \zeta_j \geq s_j \geq \max\{t_{j+1}, s_{j,\infty}\} \quad \text{for } 0 \leq j \leq r.$$

The second inequality follows since  $s_j \geq t_{j+1}$  for each  $j \leq r$ . To see the first inequality we argue by induction on  $j$ , the case  $j = 0$  being trivial. For  $j > 0$  we need to show that  $\zeta_j = (s_j - t_j)^+ + \zeta_{j-1} \geq s_j$ . Now the induction hypothesis implies that  $\zeta_{j-1} \geq s_{j-1}$ . If  $s_j \geq t_j$ , then  $\zeta_j \geq s_j - t_j + s_{j-1} \geq s_j$ . If, on the other hand,  $s_j < t_j$ , then  $\zeta_j \geq s_{j-1} \geq t_j > s_j$ . This therefore establishes (3.20).

The estimate (3.19) now becomes

$$\frac{|\mathcal{H}_i|}{H_i^n} \ll q^\varepsilon H_i^{-(\zeta_i - \zeta_{i-1})} \prod_{\substack{p \in \mathcal{V}_C^{(i)} \\ s_{i,p}-s_{i-1,p} \geq \\ \zeta_i - \zeta_{i-1}}} p^{-(s_{i,p}-\zeta_i)+(\zeta_i - \zeta_{i-1})}.$$

An expression of the form  $\prod_{p \in \mathcal{P}} p^{e_p}$  is minimised by taking  $\mathcal{P} = \{p : e_p < 0\}$  and maximised by taking  $\mathcal{P} = \{p : e_p > 0\}$ . Therefore

$$\begin{aligned} \prod_{\substack{p \in \mathcal{V}_C^{(i)} \\ s_{i,p}-s_{i-1,p} \geq \\ \zeta_i - \zeta_{i-1}}} p^{-(s_{i,p}-\zeta_i)+(\zeta_i - \zeta_{i-1})} &\leq \prod_{\substack{p \in \mathcal{V}_C^{(i)} \\ s_{i,p} \geq \zeta_i}} p^{-(s_{i,p}-\zeta_i)+(\zeta_i - \zeta_{i-1})} \\ &\leq \prod_{\substack{p \in \mathcal{V}_C \\ s_{i,p} \geq \max\{\zeta_i, t_{i+1}\}}} p^{-(s_{i,p}-\zeta_i)} \prod_{\substack{p \in \mathcal{V}_C \\ s_{i-1,p} \geq \zeta_{i-1}}} p^{(s_{i-1,p}-\zeta_{i-1})}. \end{aligned}$$

By (3.20) we have  $\max\{\zeta_i, t_{i+1}\} = \zeta_i$ . Re-setting  $\varepsilon$  and using the estimate  $q \leq P^{O(1)}$ , we see that

$$\frac{|\mathcal{H}_i|}{H_i^n} \ll P^\varepsilon H_i^{-(\zeta_i - \zeta_{i-1})} \prod_{\substack{p \in \mathcal{V}_C \\ s_{i,p} > \zeta_i}} p^{-(s_{i,p} - \zeta_i)} \prod_{\substack{p \in \mathcal{V}_C \\ s_{i-1,p} > \zeta_{i-1}}} p^{(s_{i-1,p} - \zeta_{i-1})}.$$

The above process produces a sequence  $\zeta_r \geq \dots \geq \zeta_1 \geq \zeta_0 = \sigma$  such that, on setting

$$L_i := \prod_{\substack{p \in \mathcal{V}_C \\ s_{i,p} > \zeta_i}} p^{s_{i,p} - \zeta_i},$$

we have

$$\frac{|\mathcal{H}_i|}{H_i^n} \ll P^\varepsilon H_i^{-(\zeta_i - \zeta_{i-1})} \frac{L_{i-1}}{L_i}.$$

Hence

$$\left(\frac{|\mathcal{H}_1|}{H_1^n}\right)^{2^{r-1}} \left(\frac{|\mathcal{H}_2|}{H_2^n}\right)^{2^{r-2}} \dots \left(\frac{|\mathcal{H}_r|}{H_r^n}\right) \ll \frac{P^\varepsilon L_0^{2^{r-1}} L_r^{-1}}{L_1^{2^{r-2}} L_2^{2^{r-3}} \dots L_{r-1}} \prod_{i=1}^r H_i^{(\zeta_{i-1} - \zeta_i) 2^{r-i}}.$$

For each  $v \in \mathcal{V}_C$  we have  $s_{0,v} = s_{0,\infty} = \sigma = \zeta_0$ , so that  $L_0 = 1$ . Also for all  $1 \leq i \leq r-1$  we have  $L_i \geq 1$ . Therefore the left-hand side is

$$\ll P^\varepsilon H_1^{(\zeta_0 - \zeta_1) 2^{r-1}} H_2^{(\zeta_1 - \zeta_2) 2^{r-2}} \dots H_r^{(\zeta_{r-1} - \zeta_r)} \prod_{\substack{p \in \mathcal{V}_C \\ s_{r,p} > \zeta_r}} p^{-(s_{r,p} - \zeta_r)}.$$

Let us take

$$(3.21) \quad H_j = Q^{\frac{1}{2^{r-j}}} \quad \text{for } 1 \leq j \leq r.$$

Then we deduce that

$$(3.22) \quad \left(\frac{|\mathcal{H}_1|}{H_1^n}\right)^{2^{r-1}} \left(\frac{|\mathcal{H}_2|}{H_2^n}\right)^{2^{r-2}} \dots \left(\frac{|\mathcal{H}_r|}{H_r^n}\right) \ll P^\varepsilon Q^{\zeta_0 - \zeta_r} \prod_{\substack{p \in \mathcal{V}_C \\ s_{r,p} > \zeta_r}} p^{-(s_{r,p} - \zeta_r)}.$$

Let  $g = f_{\mathbf{h}_1, \dots, \mathbf{h}_r}$  with each  $\mathbf{h}_i$  the fixed element of  $\mathcal{H}_i$  determined by the van der Corput process (3.16). Put  $\zeta := \zeta_r$  and  $s := s_{r,\infty}$ , so  $\zeta \geq s$  by (3.20). Recall that  $r = d - k$  so that by (3.17) we have  $\sigma_\nu(g^{[k]}) = s_{r,\nu}$ . In the notation (3.13) it therefore follows that

$$r_\zeta := r_\zeta^{[k]}(g, q) = \prod_{\substack{p^e \parallel bd \\ s_{r,p} = \zeta}} p^e.$$

Set

$$\rho_\zeta := \prod_{\substack{p \mid bd, p > C \\ s_{r,p} = \zeta}} p.$$

We note that

$$\rho_\zeta \gg r_\zeta^{\frac{1}{2}}.$$

One can then re-write (3.22) using our above notation to conclude that there exists an integer  $\xi \geq s$  such that

$$\begin{aligned} \left(\frac{|\mathcal{H}_1|}{H_1^n}\right)^{2^{r-1}} \left(\frac{|\mathcal{H}_2|}{H_2^n}\right)^{2^{r-2}} \cdots \left(\frac{|\mathcal{H}_r|}{H_r^n}\right) &\ll P^\varepsilon Q^{\sigma-\xi} (\rho_{\xi+1} \rho_{\xi+2}^2 \cdots \rho_n^{n-\xi})^{-1} \\ &\ll P^\varepsilon Q^{\sigma-\xi} (r_{\xi+1} r_{\xi+2}^2 \cdots r_n^{n-\xi})^{-\frac{1}{2}} \\ &= P^\varepsilon Q^{\sigma-\xi} R_\xi^{[k]}(g, q)^{-\frac{1}{2}}. \end{aligned}$$

To complete the proof of Lemma 3.5 it remains to establish that  $g$  satisfies the bound

$$\text{Height}_{P,k}(g) \ll Q^{2-2^{k+1-d}}.$$

Taylor's formula implies that

$$f_{\mathbf{h}}(\mathbf{x}) = f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \sum_{|\mathbf{m}|>0} \frac{\mathbf{h}^{\mathbf{m}}}{\mathbf{m}!} \partial^{\mathbf{m}} f(\mathbf{x}).$$

Hence there exist forms  $F_l$  and  $G_l$ , each of degree  $l$  with  $\text{Height}(F_l) \ll \text{Height}(f) = O(1)$  and  $\text{Height}(G_l) \ll 1$  such that

$$f_{\mathbf{h}}(\mathbf{x}) = \sum_{l=0}^{d-1} G_{d-l}(\mathbf{h}) F_l(\mathbf{x}).$$

Supposing that  $|\mathbf{h}| \leq H \leq P$  we have

$$\text{Height}_{P,d-1}(f_{\mathbf{h}}) \ll \sum_{l=0}^{d-1} P^{l-d+1} |G_{d-l}(\mathbf{h})| \text{Height}(F_l) \ll H.$$

Hence if  $H_1, \dots, H_r$  are chosen as in (3.21) it follows from induction that

$$\text{Height}_{P,d-r}(f_{\mathbf{h}_1, \dots, \mathbf{h}_r}) \ll H_1 \cdots H_r = Q^{\sum_{i=0}^{r-1} 2^{-i}},$$

as required.  $\square$

We are now ready to reveal our two main estimates for the exponential sum  $S(\alpha, P)$ . The first of these is Proposition 3.6. It is based on applying van der Corput differencing  $d - k$  times (Lemma 3.5) before applying the Weyl bound to the resulting exponential sum with underlying polynomial of degree at most  $k$  (Lemma 3.3). The second result is Proposition 3.8. This is proved using  $d - 3$  applications of van der Corput differencing (Lemma 3.5) together with an application of the bound for cubic exponential sums obtained via Poisson summation (Lemma 3.4).

**Proposition 3.6** (van der Corput + Weyl). *For  $1 \leq k \leq d$ , let  $B_k := (k - 1)2^{d-1}$  and let  $V_k := 2^{d+1-k} - 2$ . Suppose that  $\text{Height}(f) = O(1)$ . Then*

$$(3.23) \quad \begin{aligned} S(\alpha, P) &\ll P^{n+\varepsilon} \left( P^{-2^{1-d}} + \|q\alpha\|^{\frac{1}{B_k+V_k}} + (qP^{-k})^{\frac{1}{B_k}} \right. \\ &\quad \left. + \min\left\{ q^{-\frac{1}{B_k+V_k}}, (\|q\alpha\| P^k)^{-\frac{1}{B_k}} \right\} \right)^{n-\sigma}. \end{aligned}$$

**Remark 3.7.** For a fixed value of  $k$  one can show that the estimate in Proposition 3.6 yields the bound  $S(\alpha, P) \ll P^{n-d-\Omega(1)}$  for 'generic'  $\alpha$  in the minor arcs, provided that

$$n - \sigma > \left(1 - \frac{1}{k} \left(1 - \frac{1}{2^{k-1}}\right)\right) d 2^d - \frac{2d}{k}.$$

The optimal choice of  $k$  in (3.23) therefore appears to be  $k = 2$ , which gives  $n - \sigma > \frac{3}{4} d 2^d - d$ . Note that this inequality becomes  $n - \sigma > (d - 1) 2^d$  when  $k = d$ .

*Proof of Proposition 3.6.* Let  $1 \leq Q \leq P$ . Applying Lemma 3.5 we obtain

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll \frac{P^\varepsilon}{Q^{(\zeta-\sigma)2^{k-1}}} \left| \frac{T(\alpha, P)}{P^n} \right|^{2^{k-1}},$$

for some exponential sum  $T(\alpha, P)$  with underlying polynomial  $g$  satisfying the conclusions of the lemma. In particular,

$$\zeta \geq \max\{\sigma, \sigma_\infty(g^{[k]})\}.$$

Setting  $\tau := 2 - 2^{k+1-d}$ ,  $\theta := \sigma_\infty(g^{[k]})$  and applying Lemma 3.3, we see that

$$|P^{-n} T(\alpha, P)|^{2^{k-1}} \ll P^\varepsilon \mathfrak{E}^{\frac{n-\theta}{k-1}},$$

where

$$\mathfrak{E} := P^{1-k} + \|q\alpha\| Q^\tau + q P^{-k} + \min\{Q^\tau q^{-1}, (\|q\alpha\| P^k)^{-1}\}.$$

Since  $\zeta \geq \theta$ , we have

$$|P^{-n} T(\alpha, P)|^{2^{k-1}} \ll P^\varepsilon \mathfrak{E}^{\frac{n-\zeta}{k-1}}$$

(this is obvious when  $\mathfrak{E} \leq 1$  and follows from the trivial estimate  $T(\alpha, P) \ll P^n$  otherwise). We thus obtain

$$\begin{aligned} (3.24) \quad \left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} &\ll \frac{P^\varepsilon}{Q^{(n-\sigma)2^{k-1}}} (1 + Q^{(k-1)2^{k-1}} \mathfrak{E})^{\frac{n-\zeta}{k-1}} \\ &\leq \frac{P^\varepsilon}{Q^{(n-\sigma)2^{k-1}}} (1 + Q^{(k-1)2^{k-1}} \mathfrak{E})^{\frac{n-\sigma}{k-1}} \\ &= P^\varepsilon (Q^{-(k-1)2^{k-1}} + \mathfrak{E})^{\frac{n-\sigma}{k-1}}. \end{aligned}$$

Let us take  $Q$  such that

$$Q^{(k-1)2^{k-1}} = \min\left\{P^{k-1}, \|q\alpha\|^{-\frac{(k-1)2^{k-1}}{\tau+(k-1)2^{k-1}}}, P^k q^{-1}, \max\left\{q^{\frac{(k-1)2^{k-1}}{\tau+(k-1)2^{k-1}}}, \|q\alpha\| P^k\right\}\right\}.$$

We may assume that  $q \leq P^k$ , since the result is trivial otherwise. Using this assumption, one can check that  $1 \leq Q \leq P$ , so that our choice of  $Q$  is indeed valid. Moreover, with this choice, the  $Q^{-(k-1)2^{k-1}}$  term dominates in (3.24). Hence

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-1}} \ll P^\varepsilon Q^{-(n-\sigma)2^{k-1}}$$

One can now check that the result follows with the appropriate exponents.  $\square$

**Proposition 3.8** (van der Corput + cubic Poisson). *Suppose that  $q \leq P^2$ ,  $\|q\alpha\| \leq P^{-1}$  and  $\text{Height}(f) = O(1)$ . Then*

$$S(\alpha, P) \ll P^{n+\varepsilon} \left( (qP^{-2})^{\frac{1}{2d-2}} + (\|q\alpha\|P)^{\frac{1}{2d-1-2}} + \eta_q^{-1} \right)^{n-\sigma},$$

where

$$(3.25) \quad \eta_q := \max \left\{ (b^3 c_1 c_2^2)^{\frac{1}{2d-2}}, (b^3 c_1^2 c_2)^{\frac{1}{5 \cdot 2d-2-4}} \right\}.$$

*Proof.* Proceeding as before, we employ Lemma 3.5 to van der Corput difference down to a cubic exponential sum, which we then estimate with Lemma 3.4. We thereby deduce that for any  $1 \leq Q \leq P$  there exists a  $\zeta \geq \sigma$  such that

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-3}} \ll P^\varepsilon Q^{-(n-\sigma)} \left( 1 + \frac{Q\sqrt{q}}{P} + Q\sqrt{\|q\alpha\|PQ^\tau} + \frac{Q^{1+\frac{\tau}{6}}}{b^{\frac{1}{2}}(c_1 c_2)^{\frac{1}{3}}} \min\{c_1, c_2 Q^\tau\}^{\frac{1}{6}} \right)^{n-\zeta},$$

where  $\tau := 2 - 2^{4-d}$ . Absorbing  $Q^{-(n-\sigma)}$  into the brackets we see that the right-hand side is at most

$$P^\varepsilon \left( Q^{-1} + \frac{\sqrt{q}}{P} + \sqrt{\|q\alpha\|PQ^\tau} + \frac{Q^{\frac{\tau}{6}}}{b^{\frac{1}{2}}(c_1 c_2)^{\frac{1}{3}}} \min\{c_1, c_2 Q^\tau\}^{\frac{1}{6}} \right)^{n-\sigma}.$$

Let us take

$$Q := \min \left\{ \frac{P}{\sqrt{q}}, (\|q\alpha\|P)^{-\frac{1}{2+\tau}}, \max \left\{ (b^3 c_1 c_2^2)^{\frac{1}{6+\tau}}, (b^3 c_1^2 c_2)^{\frac{1}{6+2\tau}} \right\} \right\}.$$

Since  $q \leq P^2$  and  $\|q\alpha\| \leq P^{-1}$ , we have  $1 \leq Q \leq P$ . One can also check that

$$Q^{-1} \geq \max \left\{ \frac{\sqrt{q}}{P}, \sqrt{\|q\alpha\|PQ^\tau}, \frac{Q^{\frac{\tau}{6}}}{b^{\frac{1}{2}}(c_1 c_2)^{\frac{1}{3}}} \min\{c_1, c_2 Q^\tau\}^{\frac{1}{6}} \right\},$$

whence the  $Q^{-1}$  term dominates and we find that

$$\left| \frac{S(\alpha, P)}{P^n} \right|^{2^{d-3}} \ll P^\varepsilon \left( \frac{\sqrt{q}}{P} + (\|q\alpha\|P)^{\frac{1}{2+\tau}} + \min \left\{ (b^3 c_1 c_2^2)^{-\frac{1}{6+\tau}}, (b^3 c_1^2 c_2)^{-\frac{1}{6+2\tau}} \right\} \right)^{n-\sigma}.$$

The desired result easily follows.  $\square$

#### 4. The minor arc bound

It follows from the definition (2.3) of the major arcs that if  $\alpha \in \mathfrak{m}$ , then for any  $q \in \mathbb{N}$  either

$$q > P^\Delta \quad \text{or} \quad \|q\alpha\| > P^{\Delta-d}.$$

Our objective in this section is to establish the following estimate for the minor arc contribution, which clearly suffices for (2.5).

**Lemma 4.1.** *Suppose that*

$$n - \sigma \begin{cases} > \frac{3}{4}d2^d - 2d, & \text{if } 3 \leq d \leq 9, \\ \geq (d - \frac{1}{2}\sqrt{d})2^d, & \text{if } d \geq 10. \end{cases}$$

Then

$$\int_{\mathfrak{m}} |S(\alpha, P)| d\alpha \ll P^{n-d-\Omega(1)},$$

where the  $\Omega(1)$  term depends at most on  $d$  and  $\Delta$ .

The work in this section will involve a number of quantities that are defined in terms of  $d$  and  $k \in \{1, \dots, d\}$  and it is convenient to record them here for ease of reference. We put

$$B_k = (k-1)2^{d-1}, \quad V_k = 2^{d+1-k} - 2$$

and

$$(4.1) \quad \xi := \frac{5 - 2^{4-d}}{3 - 2^{3-d}}, \quad \gamma := \frac{1}{3 \cdot 2^{d-2} - 2},$$

together with

$$(4.2) \quad A_d := 7 \cdot 2^{d-4} - \frac{5}{4}.$$

In particular,  $\xi \leq \frac{5}{3}$ . It will also be convenient to define

$$(4.3) \quad C(\alpha, q) := (\|q\alpha\|P)^{\frac{1}{2^{d-1}-2}} + (qP^{-2})^{\frac{1}{2^{d-2}}} + \eta_q^{-1}$$

and

$$(4.4) \quad W_k(\alpha, q) := P^{-2^{1-d}} + \|q\alpha\|^{\frac{1}{B_k+V_k}} + (qP^{-k})^{\frac{1}{B_k}} \\ + \min\{q^{-\frac{1}{B_k+V_k}}, (\|q\alpha\|P^k)^{-\frac{1}{B_k}}\}.$$

Next, we define the minimum

$$M(\alpha, q) := \min\{C(\alpha, q), W_3(\alpha, q), \dots, W_d(\alpha, q)\}.$$

In the following result we estimate the minor arc contribution in terms of these quantities.

**Lemma 4.2.** *There exists an index  $i \in \{1, 2\}$  such that*

$$\int_{\mathfrak{m}} |S(\alpha, P)| d\alpha \ll P^{n-d-\Omega(1)} + P^{n+\varepsilon} \sum_{P^{\Delta_i} \leq q \leq P^\xi} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} M(\alpha, q)^{n-\sigma} d\alpha,$$

where

$$\Delta_1 := \Delta, \quad \Delta_2 := 0, \quad d_1 := d + 2, \quad d_2 := d - \Delta.$$

*Proof.* Let

$$\mathcal{B}(q, Q) := \{\alpha \in \mathbb{T} : \|q\alpha\| \leq Q^{-1} \text{ and } \alpha \text{ is primitive to } q\}.$$

Then by Dirichlet's theorem on Diophantine approximation and the definition of the minor arcs, we have

$$\mathfrak{m} \subset \left( \bigcup_{P^{\Delta} \leq q \leq P^\xi} \mathcal{B}(q, P^\xi) \right) \cup \left( \bigcup_{q \leq P^\xi} \mathcal{B}(q, P^\xi) \setminus \mathcal{B}(q, P^{d-\Delta}) \right).$$

For  $\alpha \in \mathcal{B}(q, P^\xi)$  it follows from Propositions 3.6 and 3.8 that

$$S(\alpha, P) \ll P^{n+\varepsilon} M(\alpha, q)^{n-\sigma}.$$

Hence by sub-additivity of integration and dropping the primitivity condition on  $q$  and  $\alpha$ , we deduce that

$$\begin{aligned} \int_{\mathfrak{m}} |S(\alpha, P)| d\alpha &\ll P^{n+\varepsilon} \sum_{P^\Delta \leq q \leq P^\xi} \int_{\|q\alpha\| \leq P^{-\xi}} M(\alpha, q)^{n-\sigma} d\alpha \\ &+ P^{n+\varepsilon} \sum_{q \leq P^\xi} \int_{P^{\Delta-d} \leq \|q\alpha\| \leq P^{-\xi}} M(\alpha, q)^{n-\sigma} d\alpha \end{aligned}$$

for any  $\varepsilon > 0$ . The second term above is of the required form. In the first term, we need to remove the possibility that  $\|q\alpha\| \leq P^{-d-2}$ . But the trivial bound  $M(\alpha, q) \leq 1$  yields

$$\sum_{P^\Delta \leq q \leq P^\xi} \int_{\|q\alpha\| \leq P^{-d-2}} M(\alpha, q)^{n-\sigma} d\alpha \ll P^{\xi-d-2}.$$

Thus the result follows with the  $\Omega(1)$  term equal to any positive real strictly less than  $\frac{1}{3}$ .  $\square$

Our aim is to show that when  $n - \sigma > \frac{3}{4}d2^d - 2d$ , then  $M(\alpha, q)^{n-\sigma}$  contributes at most  $P^{-d-\Omega(1)}$  once integrated over the range of  $(\alpha, q)$  afforded by Lemma 4.2. We emphasise that for both  $i \in \{1, 2\}$  we have

$$\sum_{P^{\Delta_i} \leq q \leq P^\xi} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} d\alpha \ll 1,$$

so that a pointwise bound of the form  $M(\alpha, q)^{n-\sigma} \ll P^{-d-\Omega(1)}$  suffices for our purposes. To this end, we first utilise the bound  $M(\alpha, q) \leq C(\alpha, q)$ , and show that for  $\alpha$  and  $q$  in the range in question, all terms of  $C(\alpha, q)$  are negligible bar possibly the  $\eta_q$  term.

**Lemma 4.3.** *Let  $\alpha \in \mathbb{T}$  and  $q \in \mathbb{N}$  satisfy*

$$q \leq P^\xi \quad \text{and} \quad \|q\alpha\| \leq P^{-\xi},$$

*with  $\xi$  as in (4.1). Then provided that  $n - \sigma > \frac{3}{4}d2^d - 2d$  we have*

$$(4.5) \quad M(\alpha, q)^{n-\sigma} \ll P^{-d-\Omega(1)} + \eta_q^{-(n-\sigma)}.$$

*Proof.* This follows from the inequality  $M(\alpha, q) \leq C(\alpha, q)$  and by substituting the bounds on  $q$  and  $\|q\alpha\|$  into (4.3).  $\square$

Next, recall the definition (4.1) of  $\gamma$ . When  $\eta_q > P^\gamma$  and  $n - \sigma > \frac{3}{4}d2^d - 2d$ , then (4.5) implies that  $M(\alpha, q)^{n-\sigma} \ll P^{-d-\Omega(1)}$ . It follows that we can restrict our estimation of  $M(\alpha, q)$  to those  $q$  satisfying  $\eta_q \leq P^\gamma$ . It is helpful to have an estimate for the number of  $q$  which lie in this range. This is provided by the following lemma.

**Lemma 4.4.** *Let  $R \geq 1$ . Then*

$$\#\{q \in \mathbb{N} : \eta_q \leq R\} \ll R^{A_d},$$

*where  $A_d$  is given by (4.2).*



*Proof.* Recall the definition of  $b$ ,  $c_1$  and  $c_2$  given in (3.12). Setting

$$e := \prod_{p^5 \parallel q} p,$$

we have that  $e \mid c_2$  and that  $\frac{c_1}{c_2 e}$  is a square-full positive integer. Since

$$\eta_q \geq (b^3 c_1^2 c_2)^{\frac{1}{5 \cdot 2^{d-2-4}}},$$

we see that  $\eta_q \leq R$  implies that

$$\frac{c_1}{c_2 e} \leq \frac{R^{5 \cdot 2^{d-3}-2}}{b^{\frac{3}{2}} c_2^{\frac{3}{2}} e}.$$

As a square-full number is a product of a square and a cube, the hyperbola method shows that the number of square-full integers less than or equal to  $X$  is at most  $3X^{\frac{1}{2}}$ . Thus for a fixed choice of  $b$  and  $c_2$ , the number of choices of  $c_1$  for which  $\eta_q \leq R$  is at most

$$3 \sum_{e \mid c_2} \frac{R^{5 \cdot 2^{d-4}-1}}{b^{\frac{3}{4}} c_2^{\frac{3}{4}} e^{\frac{1}{2}}}.$$

Using the fact that  $\eta_q \geq (b^3 c_1 c_2^2)^{\frac{1}{2^{d-2}}}$  we see that  $\eta_q \leq R$  implies that for fixed  $b$  we have

$$c_2 \leq \frac{R^{2^{d-1}-1}}{b^{\frac{3}{2}}}.$$

Hence the number of choices for  $q$  for which  $\eta_q \leq R$  is of order

$$\begin{aligned} \sum_b \sum_{c_2 \leq \frac{R^{2^{d-1}-1}}{R b^{\frac{3}{2}}}} \sum_{e \mid c_2} \frac{R^{5 \cdot 2^{d-4}-1}}{b^{\frac{3}{4}} c_2^{\frac{3}{4}} e^{\frac{1}{2}}} &\ll R^{5 \cdot 2^{d-4}-1} \sum_b b^{-\frac{3}{4}} \sum_{ef \leq \frac{R^{2^{d-1}-1}}{R b^{\frac{3}{2}}}} f^{-\frac{3}{4}} e^{-\frac{5}{4}} \\ &\ll R^{5 \cdot 2^{d-4}-1} \sum_b b^{-\frac{3}{4}} \left( \frac{R^{2^{d-1}-1}}{R b^{3/2}} \right)^{\frac{1}{4}} \\ &\ll R^{5 \cdot 2^{d-4}-1+2^{d-3}-\frac{1}{4}}. \end{aligned}$$

This completes the proof of the lemma.  $\square$

We are now in a position to use information coming from  $W_k(\alpha, q)$  in our definition of  $M(\alpha, q)$ . We first show that both the  $P^{-2^{1-d}}$  and  $qP^{-k}$  terms appearing in  $W_k(\alpha, q)$  are negligible when  $\eta_q \leq P^\nu$  and  $n - \sigma > \frac{3}{4}d2^d - 2d$ . To this end, define

$$(4.6) \quad W_k^*(\alpha, q) = \max \left\{ \|q\alpha\|^{\frac{1}{B_k + \nu_k}}, \min \{ \eta_q^{-1}, (\|q\alpha\| P^k)^{-\frac{1}{B_k}} \} \right\}$$

and

$$M^*(\alpha, q) := \min \{ W_3^*(\alpha, q), \dots, W_d^*(\alpha, q) \}.$$

Then we have the following result.

**Lemma 4.5.** *Assume that  $n - \sigma > \frac{3}{4}d2^d - 2d$  and let  $i \in \{1, 2\}$ . Then*

$$\begin{aligned} & \sum_{P^{\Delta_i} \leq q \leq P^\xi} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} M(\alpha, q)^{n-\sigma} d\alpha \\ & \ll P^{-d-\Omega(1)} + \sum_{P^{\gamma_i} \leq \eta_q \leq P^\nu} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} M^*(\alpha, q)^{n-\sigma} d\alpha, \end{aligned}$$

where

$$\gamma_i = \frac{\Delta_i}{5 \cdot 2^{d-2} - 4}.$$

*Proof.* If  $n - \sigma > \frac{3}{4}d2^d - 2d$  and  $d \geq 3$ , then  $\frac{n-\sigma}{2^{d-1}} > d$ , so that

$$P^{-2^{1-d}(n-\sigma)} \ll P^{-d-\Omega(1)}.$$

This deals with the  $P^{-2^{1-d}}$  term.

For the second term, we note from (3.25) that

$$\eta_q \geq q^{\frac{1}{5 \cdot 2^{d-2} - 4}}.$$

We may also assume that  $\eta_q \leq P^\nu$  by the remark after Lemma 4.3. Combining this with (4.4) and (4.5) one sees that the lemma follows if for each  $3 \leq k \leq d$  we have

$$\sum_{P^{\gamma_i} \leq \eta_q \leq P^\nu} P^{-\xi} (qP^{-k})^{\frac{n-\sigma}{B_k}} \ll P^{-d-\Omega(1)}.$$

Here we have used the fact that the set  $\{\alpha \in \mathbb{T} : \|q\alpha\| \leq P^{-\xi}\}$  has measure  $2P^{-\xi}$ . Using Lemma 4.4 and the bound  $q \leq P^\xi$  we have

$$\begin{aligned} P^{-\xi} \sum_{\eta_q \leq P^\nu} (qP^{-k})^{\frac{n-\sigma}{B_k}} & \leq P^{\xi(\frac{n-\sigma}{B_k}-1) - \frac{k(n-\sigma)}{B_k}} \sum_{\eta_q \leq P^\nu} 1 \\ & \leq P^{\xi(\frac{n-\sigma}{B_k}-1) - \frac{k(n-\sigma)}{B_k} + \gamma A_d}. \end{aligned}$$

One can check from the definitions (4.1) and (4.2) that  $\gamma A_d \leq \frac{7}{12}$  and  $\xi \leq \frac{5}{3}$ . Incorporating this together with the fact that  $\frac{n-\sigma}{B_k} - 1 \geq 0$ , we obtain

$$P^{\xi(\frac{n-\sigma}{B_k}-1) - \frac{k(n-\sigma)}{B_k} + \gamma A_d} \leq P^{-(k-\frac{5}{3})\frac{n-\sigma}{B_k} - \frac{13}{12}}.$$

To complete the proof of the lemma, we require that

$$\frac{(n-\sigma)(k-\frac{5}{3})}{B_k} + \frac{13}{12} > d$$

for  $k \geq 3$ . This is equivalent to

$$n - \sigma > \frac{1}{2} \left( \frac{k-1}{k-\frac{5}{3}} \right) (d - \frac{13}{12}) 2^d = \frac{1}{2} \left( 1 + \frac{2}{3k-5} \right) (d - \frac{13}{12}) 2^d.$$

But

$$\frac{2}{3k-5} \leq \frac{1}{2} \quad (\text{as } k \geq 3)$$

and the required bound is thus implied by the assumption that  $n - \sigma > \frac{3}{4}d2^d - 2d$ , together with the estimate  $\frac{3 \cdot 13}{4 \cdot 12} 2^d \geq 2d$ , which is valid for  $d \geq 3$ .  $\square$

In view of Lemma 4.5, it remains to analyse the contribution from the term  $M^*(\alpha, q)$ . The following result represents a key step in our argument and provides us with a concrete condition under which the bulk of this contribution is satisfactory.

**Lemma 4.6.** *Let  $k \geq 3$  and suppose that for each  $\ell \in \{k, \dots, d\}$  we have*

$$(4.7) \quad n - \sigma > A_d + \frac{d}{\ell} B_\ell + \left(\frac{d}{\ell} - 1\right)(B_{\ell+1} + V_{\ell+1}).$$

Then

$$M^*(\alpha, q) \ll E(\alpha, q) + \|q\alpha\|^{\frac{1}{B_k + V_k}},$$

where  $E(\alpha, q)$  is a function satisfying

$$(4.8) \quad \sum_{P^{\gamma i} \leq \eta_q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} E(\alpha, q)^{n-\sigma} d\alpha \ll P^{-d-\Omega(1)} \quad (i \in \{1, 2\}).$$

*Proof.* We proceed by inducting downwards on  $k$ , starting with  $k = d$ . By the definition (4.6) of  $W_d^*(\alpha, q)$ , it suffices to prove that

$$\sum_{P^{\gamma i} \leq \eta_q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} \min\left\{\eta_q^{-1}, (\|q\alpha\| P^d)^{-\frac{1}{B_d}}\right\}^{n-\sigma} d\alpha \ll P^{-d-\Omega(1)}.$$

Let us denote by  $I_d$  the left-hand side of this desired estimate. Since

$$[P^{\gamma i}, P^\gamma] \times [P^{-d_i}, P^{-\xi}] \subset [1, P] \times [P^{-d-2}, 1],$$

we can partition the ranges of  $\|q\alpha\|$  and  $\eta_q$  occurring in  $I_d$  into  $O(\log^2 P)$  dyadic rectangles. Thus, by the pigeon-hole principle, there exist  $\delta \gg_{d, \Delta} 1$ ,  $R \leq P^\gamma$  and  $t \leq P^{-\xi}$  such that

$$R \geq P^\delta \quad \text{or} \quad t \geq P^{\delta-d},$$

and

$$I_d \ll (\log P)^2 \sum_{R \leq \eta_q \leq 2R} \int_{t \leq \|q\alpha\| \leq 2t} \min\left\{\eta_q^{-1}, (\|q\alpha\| P^d)^{-\frac{1}{B_d}}\right\}^{n-\sigma} d\alpha.$$

Using Lemma 4.4, we find that

$$\begin{aligned} I_d &\ll (\log P)^2 2t \min\left\{R^{-1}, (tP^d)^{-\frac{1}{B_d}}\right\}^{n-\sigma} \sum_{R \leq \eta_q \leq 2R} 1 \\ &\ll (\log P)^2 t R^{A_d} \min\left\{R^{-1}, (tP^d)^{-\frac{1}{B_d}}\right\}^{n-\sigma}. \end{aligned}$$

Now for any  $a, b, X, Y \geq 0$  we have the inequality  $\min\{X, Y\}^{a+b} \leq X^a Y^b$ . Hence to establish the base case, it suffices to find  $a, b \geq 0$  with  $a + b = n - \sigma$  such that

$$R^{A_d - a} t^{1 - \frac{b}{B_d}} P^{-\frac{db}{B_d}} \ll P^{-d-\Omega(1)}.$$

Notice that when  $\ell = d$ , condition (4.7) reduces to  $n - \sigma > A_d + B_d$ . There are two cases to consider. Suppose first that  $R \geq P^\delta$ . In this case, we can take  $a = A_d + \Omega(1)$  and  $b = B_d$ , which gives

$$R^{A_d - a} t^{1 - \frac{b}{B_d}} P^{-\frac{db}{B_d}} \leq P^{-\Omega(\delta) - d}.$$

Alternatively, we suppose that  $t \geq P^{\delta-d}$ . But in this case we take  $a = A_d$  and  $b = B_d + \Omega(1)$ , giving

$$R^{A_d-a} t^{1-\frac{b}{B_d}} P^{-\frac{db}{B_d}} \leq P^{-\Omega(\frac{\delta}{B_d})-d}.$$

The base case then follows.

Next, we turn to the induction step. Let us assume that  $3 \leq k < d$ . By the induction hypothesis we have

$$M^*(\alpha, q) \ll E(\alpha, q) + \|q\alpha\|^{\frac{1}{B_{k+1}+V_{k+1}}},$$

where  $E(\alpha, q)$  satisfies (4.8). Combining this with the definition of  $W_k^*(\alpha, q)$ , it suffices to prove that

$$\sum_{P^{\gamma_i} \leq \eta q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} \min\left\{\eta_q^{-1}, (\|q\alpha\| P^k)^{-\frac{1}{B_k}}, \|q\alpha\|^{\frac{1}{B_{k+1}+V_{k+1}}}\right\}^{n-\sigma} d\alpha$$

is  $O(P^{-d-\Omega(1)})$ . As before, breaking into dyadic rectangles, we see that it suffices to find  $a, b, c \geq 0$  with  $a + b + c = n - \sigma$  such that for  $R \leq P^\gamma$  and  $t \leq P^{-\xi}$  we have

$$R^{A_d-a} t^{1-\frac{b}{B_k} + \frac{c}{B_{k+1}+V_{k+1}}} P^{-\frac{kb}{B_k}} \ll P^{-d-\Omega(1)}.$$

Recall our assumption (4.7). Taking

$$a = A_d, \quad b = \frac{dB_k}{k} + \Omega_d(1), \quad c = \left(\frac{d}{k} - 1\right)(B_{k+1} + V_{k+1}),$$

one readily verifies that the desired estimate holds.  $\square$

The size of the term  $\|q\alpha\|^{\frac{1}{B_k+V_k}}$  in Lemma 4.6 is smallest when  $k$  is minimal. When the condition (4.7) holds for every  $3 \leq \ell \leq d$ , the following result shows that the contribution from the term  $\|q\alpha\|^{\frac{1}{B_3+V_3}}$  is satisfactory under the assumptions of Lemma 4.1.

**Lemma 4.7.** *Suppose that  $n - \sigma > \frac{3}{4}d2^d - 2d$  and  $d \geq 3$ . Then for  $i \in \{1, 2\}$  we have*

$$\sum_{P^{\gamma_i} \leq \eta q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} \min\left\{\eta_q^{-1}, \|q\alpha\|^{\frac{1}{B_3+V_3}}\right\}^{n-\sigma} d\alpha \ll P^{-d-\Omega(1)}.$$

*Proof.* An easy calculation reveals that

$$\frac{3}{4}d2^d - 2d \geq A_d + \left(\frac{d}{\xi} - 1\right)(B_3 + V_3)$$

for  $d \geq 3$ . Hence

$$(4.9) \quad n - \sigma > A_d + \left(\frac{d}{\xi} - 1\right)(B_3 + V_3),$$

under the assumptions of the lemma. As before, we split into dyadic intervals and deduce that it suffices to find  $a, b \geq 0$  with  $a + b = n - \sigma$  such that

$$R^{A_d-a} t^{1+\frac{b}{B_3+V_3}} \ll P^{-d-\Omega(1)},$$

where  $R \leq P^\gamma$  and  $t \leq P^{-\xi}$ . Taking  $a = A_d$ , inequality (4.9) ensures that

$$b = \left(\frac{d}{\xi} - 1\right)(B_3 + V_3) + \Omega_d(1).$$

Hence

$$R^{A_d-a} t^{1+\frac{b}{B_3+V_3}} \leq P^{-\xi(1+\frac{b}{B_3+V_3})} \leq P^{-d-\Omega_d(1)},$$

as required.  $\square$

We now have everything in place to establish Lemma 4.1. Suppose first that  $3 \leq d \leq 9$  and  $n - \sigma > \frac{3}{4}d2^d - 2d$ . Combining Lemma 4.2 and Lemma 4.5, there exists an index  $i \in \{1, 2\}$  such that

$$\int_{\mathfrak{m}} |S(\alpha, P)| d\alpha \ll P^{n-d-\Omega(1)} + P^{n+\varepsilon} \sum_{P^{\gamma_i} \leq \eta_q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} M^*(\alpha, q)^{n-\sigma} d\alpha.$$

One can check (most expediently by computer) that when  $3 \leq d \leq 9$  and  $3 \leq k \leq d$ , we have

$$\frac{3}{4}d2^d - 2d + 1 > A_d + \frac{d}{k}B_k + \left(\frac{d}{k} - 1\right)(B_{k+1} + V_{k+1}).$$

The hypotheses of Lemma 4.6 are therefore satisfied with  $k = 3$ , so that

$$\begin{aligned} & \sum_{P^{\gamma_i} \leq \eta_q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} M^*(\alpha, q)^{n-\sigma} d\alpha \\ & \ll \sum_{P^{\gamma_i} \leq \eta_q \leq P^\gamma} \int_{P^{-d_i} \leq \|q\alpha\| \leq P^{-\xi}} \min\left\{\eta_q^{-1}, \|q\alpha\|^{\frac{1}{B_3+V_3}}\right\}^{n-\sigma} d\alpha. \end{aligned}$$

But the latter quantity is  $O(P^{-d-\Omega(1)})$  by Lemma 4.7, which therefore concludes the proof of Lemma 4.1 when  $3 \leq d \leq 9$ .

Suppose now that  $d \geq 10$  and  $n - \sigma \geq (d - \frac{1}{2}\sqrt{d})2^d$ . As above it suffices to show that for all  $k$  in the range  $3 \leq k \leq d$  we have

$$d2^d - \sqrt{d}2^{d-1} > A_d + \frac{d}{k}B_k + \left(\frac{d}{k} - 1\right)(B_{k+1} + V_{k+1}).$$

Substituting the definitions of  $A_d$ ,  $B_k$  and  $V_k$  into the right-hand side, we see that it is strictly less than

$$\begin{aligned} d2^d - \left(\frac{d}{2}\left(\frac{1}{k} - \frac{1}{k2^{k-1}}\right) + \frac{1}{2}\left(k - \frac{7}{8}\right)\right)2^d &< d2^d - \left(\frac{d}{2}\left(\frac{1}{2k}\right) + \frac{1}{2}\left(\frac{k}{2}\right)\right)2^d \\ &= d2^d - \frac{\sqrt{d}}{4}\left(\frac{\sqrt{d}}{k} + \frac{k}{\sqrt{d}}\right)2^d \\ &\leq d2^d - \sqrt{d} \cdot 2^{d-1}. \end{aligned}$$

This therefore concludes the proof of Lemma 4.1.

**Remark 4.8.** It is somewhat disappointing that we are unable to do better when  $d \geq 10$ . However, one easily confirms that the maximum of the conditions (4.7) is asymptotically at least  $d2^d$ . Thus one cannot obtain the required minor arc bound with roughly  $\frac{3}{4}d2^d$  variables.

## 5. The major arc asymptotic

The purpose of this section is to establish (2.4) under suitable hypotheses on the form  $F$  and on the parameter  $\Delta$  occurring in the definition (2.3) of  $\mathfrak{M}$ .

According to our local solubility hypothesis in Theorems 1.1 and 1.2 we may assume that the system (1.2) has a solution over the reals. Thus there exists a vector  $\mathbf{x}_0 \in \mathbb{R}^n$  such that  $F(\mathbf{x}_0) = 0$  and  $\nabla F(\mathbf{x}_0) \neq \mathbf{0}$ . This vector is to be considered fixed once and for all in what

follows. It will be convenient to work with a weight function that forces us to count points lying very close to  $\mathbf{x}_0$ . For any  $\delta \in (0, 1]$ , we define the function  $\omega : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  by

$$\omega(\mathbf{x}) := w(\delta^{-1} \|\mathbf{x} - \mathbf{x}_0\|_2),$$

where  $\|\mathbf{y}\|_2 = \sqrt{y_1^2 + \dots + y_n^2}$  and

$$w(x) := \begin{cases} e^{-\frac{1}{1-x^2}}, & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1. \end{cases}$$

We will require  $\delta$  to be sufficiently small, with  $1 \ll \delta \leq 1$ . It is clear that  $\omega$  belongs to the class  $\mathfrak{S}^+(\mathbf{c})$  from Definition 3.1 for a suitable infinite tuple  $\mathbf{c} = (c, c_0, c_1, \dots)$  depending on  $\delta$  and  $\mathbf{x}_0$ .

Let us define

$$S_q(a) := \sum_{\mathbf{y} \pmod{q}} e_q(aF(\mathbf{y}))$$

for  $a \in \mathbb{Z}$  such that  $(a, q) = 1$ , together with the associated truncated singular series

$$(5.1) \quad \mathfrak{S}(R) = \sum_{q \leq R} \frac{1}{q^n} \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} S_q(a)$$

for any  $R > 1$ . We put  $\mathfrak{S} = \lim_{R \rightarrow \infty} \mathfrak{S}(R)$ , whenever this limit exists. Next, let

$$(5.2) \quad \mathfrak{I}(R) = \int_{-R}^R \int_{\mathbb{R}^n} \omega(\mathbf{x}) e(\gamma F(\mathbf{x})) d\mathbf{x} d\gamma$$

for any  $R > 0$ . We put  $\mathfrak{I} = \lim_{R \rightarrow \infty} \mathfrak{I}(R)$ , whenever the limit exists. The main aim of this section is to establish the following result.

**Lemma 5.1.** *Assume that  $n - \sigma > \frac{3}{4}(d - 1)2^d$ . Then the singular series  $\mathfrak{S}$  and the singular integral  $\mathfrak{I}$  are absolutely convergent. Moreover, if we choose  $\Delta = \frac{1}{6}$ , then*

$$\int_{\mathfrak{M}} S(\alpha) d\alpha = \mathfrak{S} \mathfrak{I} P^{n-d} + O(P^{n-d-\Omega(1)}).$$

Here the leading constant  $\mathfrak{S} \mathfrak{I}$  is a product of local densities and, in the usual way, one has  $\mathfrak{S} \mathfrak{I} > 0$  under the local solubility assumptions of Theorems 1.1 and 1.2. Once taken in conjunction with Lemma 4.1, the proof of our main results will therefore stand once Lemma 5.1 is verified.

Our treatment of Lemma 5.1 is standard and closely follows the argument of Birch in [1, Section 5], as revisited in [3, Section 10]. Thus we shall allow ourselves to be brief. For  $q \leq P^\Delta$  and  $a \in \mathbb{Z}$  coprime to  $q$ , let us put

$$\alpha = \frac{a}{q} + \theta$$

for any  $\alpha \in \mathfrak{M}_{a,q}$ . To begin with, the argument of [1, Lemma 5.1] (cf. [3, equation (10.5)]) easily gives

$$(5.3) \quad S(\alpha, P) = q^{-n} P^n S_q(a) I(\theta P^d) + O(q|\theta| P^{n+d-1} + qP^{n-1}),$$

where  $S_q(a)$  is given above and we put

$$I(\gamma) = \int_{\mathbb{R}^n} \omega(\mathbf{x}) e(\gamma F(\mathbf{x})) d\mathbf{x}$$

for any  $\gamma \in \mathbb{R}$ . Recalling that  $|\theta| \leq P^{-d+\Delta}$  and  $q \leq P^\Delta$  on the major arcs, this implies that

$$S(\alpha, P) = q^{-n} P^n S_q(a) I(\theta P^d) + O(P^{n-1+2\Delta}).$$

Noting that the major arcs have measure  $O(P^{-d+3\Delta})$ , it now follows that

$$(5.4) \quad \int_{\mathfrak{M}} S(\alpha, P) d\alpha = P^{n-d} \mathfrak{C}(P^\Delta) \mathfrak{F}(P^\Delta) + O(P^{n-d-1+5\Delta}),$$

where  $\mathfrak{C}(P^\Delta)$  is given by (5.1), and  $\mathfrak{F}(P^\Delta)$  is given by (5.2).

Next, we claim that

$$(5.5) \quad I(\gamma) \ll \min\left\{1, |\gamma|^{-\frac{n-\sigma}{(d-1)2^{d-1}} + \varepsilon}\right\}.$$

The argument for this is based on [1, Lemma 5.2] (cf. [3, Lemma 24]). The estimate  $I(\gamma) \ll 1$  is trivial. In proving the second estimate we may clearly assume that  $|\gamma| > 1$ . Taking  $a = 0$  and  $q = 1$  in (5.3), we deduce that

$$S(\alpha, P) = P^n I(\alpha P^d) + O((|\alpha| P^d + 1) P^{n-1})$$

for any  $P \geq 1$ . On the other hand, assuming that  $|\alpha| < P^{-\frac{d}{2}}$ , Lemma 3.3 gives

$$S(\alpha, P) \ll P^{n+\varepsilon} (|\alpha| P^d)^{-\frac{n-\sigma}{(d-1)2^{d-1}}}.$$

Writing  $\alpha P^d = \gamma$ , we may combine these estimates to obtain

$$I(\gamma) \ll |\gamma|^{-\frac{n-\sigma}{(d-1)2^{d-1}}} P^\varepsilon + |\gamma| P^{-1}$$

when  $|\gamma| < P^{\frac{d}{2}}$ . Finally, we observe that  $I(\gamma)$  is independent of  $P$ . Thus we are free to choose

$$P = |\gamma|^{1 + \frac{n-\sigma}{(d-1)2^{d-1}}},$$

which thereby establishes (5.5).

Suppose that  $n - \sigma > \frac{1}{2}(d - 1)2^d$ . It now follows from (5.5) that

$$\mathfrak{F} - \mathfrak{F}(R) = \int_{|\gamma| \geq R} I(\gamma) d\gamma \ll \int_R^\infty \min\left\{1, \gamma^{-\frac{n-\sigma}{(d-1)2^{d-1}} + \varepsilon}\right\} d\gamma \ll R^{1 - \frac{n-\sigma}{(d-1)2^{d-1}} + \varepsilon}.$$

This shows that  $\mathfrak{F}$  is absolutely convergent for  $n - \sigma > \frac{1}{2}(d - 1)2^d$ , which is more than enough for Lemma 5.1.

Next, we need to show that

$$(5.6) \quad \mathfrak{C} - \mathfrak{C}(R) \ll R^{-\eta},$$

for some  $\eta > 0$ , provided that  $n - \sigma > \frac{3}{4}(d - 1)2^d$ . Assuming this to be the case for the moment

and observing that  $\mathfrak{S}(P^\Delta) \ll 1$ , it follows from (5.4) that

$$\begin{aligned} \int_{\mathfrak{M}} S(\alpha, P) d\alpha &= \mathfrak{S} P^{n-d} \mathfrak{S}(P^\Delta) + O(P^{n-d-1+5\Delta} + P^{n-d-\Delta\eta}) \\ &= \mathfrak{S} P^{n-d} + O(P^{n-d-1+5\Delta} + P^{n-d-\Delta\eta} + P^{n-d-\Delta(\frac{n-\sigma}{(d-1)2^{d-1}}-1)+\varepsilon}). \end{aligned}$$

We therefore obtain the statement of Lemma 5.1 by choosing  $\Delta = \frac{1}{6}$  and taking  $\varepsilon > 0$  to be sufficiently small.

Turning finally to the proof of (5.6), we put

$$A(q) = q^{-n} \sum_{\substack{a \pmod{q} \\ (a,q)=1}} S_q(a).$$

This is a multiplicative function of  $q$ . It follows from Birch [1, Lemma 5.4] that

$$(5.7) \quad A(p^k) \ll p^{k(1-\frac{n-\sigma}{(d-1)2^{d-1}})+\varepsilon}.$$

When  $k = 1$ , we instead call upon the estimate

$$A(p) \ll p^{1-\frac{n-\sigma}{2}},$$

as recorded in [3, Lemma 7]. This is established by induction on  $\sigma$ , the inductive base  $\sigma = 0$  being taken care of by Deligne's estimate [5].

We now establish estimate (5.6) under the assumption that  $n - \sigma > \frac{3}{4}(d - 1)2^d$ . Let us write  $q = uv$ , where  $u$  is the square-free part of  $q$ . Then, by multiplicativity, we have

$$A(u) \ll u^{1-\frac{n-\sigma}{2}+\varepsilon}.$$

Once combined with (5.7), this shows there exists a real number  $\eta > 0$  such that

$$\begin{aligned} |\mathfrak{S} - \mathfrak{S}(R)| &\ll \sum_{q=uv>R} u^{1-\frac{n-\sigma}{2}+\varepsilon} v^{1-\frac{n-\sigma}{(d-1)2^{d-1}}+\varepsilon} \\ &\ll \sum_{q=uv>R} u^{-2-\eta} v^{-\frac{1}{2}-2\eta} \\ &\ll R^{-\eta} \sum_{u,v=1}^{\infty} u^{-2} v^{-\frac{1}{2}-\eta}, \end{aligned}$$

provided that  $\varepsilon > 0$  is taken to be sufficiently small. Since the number of square-full integers  $v \in (V, 2V]$  is  $O(V^{\frac{1}{2}})$ , the sum over  $v$  is seen to be convergent, as is the sum over  $u$ . The bound recorded in (5.6) therefore follows.

### References

- [1] *B. J. Birch*, Forms in many variables, Proc. R. Soc. Lond. Ser. A Math. Phys. Sci. **265** (1961/62), 245–263.
- [2] *J. Brandes*, Forms representing forms and linear spaces on hypersurfaces, Proc. Lond. Math. Soc. (3) **108** (2014), 809–835.
- [3] *T. D. Browning* and *D. R. Heath-Brown*, Rational points on quartic hypersurfaces, J. reine angew. Math. **629** (2009), 37–88.



- [4] *H. Davenport*, Analytic methods for Diophantine equations and Diophantine inequalities, 2nd ed., edited by T.D. Browning, Cambridge University Press, Cambridge 2005.
- [5] *P. Deligne*, La conjecture de Weil. I, Publ. Math. Inst. Hautes Études Sci. **43** (1974), 273–307.
- [6] *M. Hanselmann*, Rational points on quartic hypersurfaces, Ph.D. thesis, Ludwig Maximilians Universität München, Munich 2012.
- [7] *D. R. Heath-Brown*, Cubic forms in 10 variables, Proc. Lond. Math. Soc. (3) **47** (1983), 225–257.
- [8] *D. R. Heath-Brown*, Cubic forms in 14 variables, Invent. Math. **170** (2007), 199–230.
- [9] *D. R. Heath-Brown*, Zeros of  $p$ -adic forms, Proc. Lond. Math. Soc. (3) **100** (2010), 560–584.
- [10] *C. Hooley*, On nonary cubic forms, J. reine angew. Math. **386** (1988), 32–98.
- [11] *S.-L. A. Lee*, Birch's theorem in function fields, Ph.D. thesis, University of Bristol, Bristol 2013.
- [12] *D. B. Leep* and *C. C. Yeomans*, Quintic forms over  $p$ -adic fields, J. Number Theory **57** (1996), 231–241.
- [13] *D. Schindler*, Bihomogeneous forms in many variables, J. Théorie Nombres Bordeaux **26** (2014), 483–506.
- [14] *C. M. Skinner*, Rational points on nonsingular cubic hypersurfaces, Duke Math. J. **75** (1994), 409–466.
- [15] *C. M. Skinner*, Forms over number fields and weak approximation, Compos. Math. **106** (1997), 11–29.
- [16] *J. Zahid*, Zeros of  $p$ -adic forms, J. Number Theory **129** (2009), 2439–2456.

---

T. D. Browning, School of Mathematics, University of Bristol, Bristol, BS8 1TW, UK  
e-mail: t.d.browning@bristol.ac.uk

S. M. Prendiville, School of Mathematics, University of Bristol, Bristol, BS8 1TW, UK  
e-mail: sean.prendiville@bristol.ac.uk

Eingegangen 4. Juni 2014, in revidierter Fassung 21. Oktober 2014