The density of rational lines on hypersurfaces: a bihomogeneous perspective

Downloaded from: https://research.chalmers.se, 2026-11-16 20:08 UTC

Citation for the original published paper (version of record):
Brandes, J. (2021). The density of rational lines on hypersurfaces: a bihomogeneous perspective. Monatshefte für Mathematik, 195(2): 191-231. http://dx.doi.org/10.1007/s00605-021-01528-6

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Monatshefte für Mathematik
https://doi.org/10.1007/s00605-021-01528-6

The density of rational lines on hypersurfaces: a bihomogeneous perspective

Julia Brandes[1]

Received: 5 September 2020 / Accepted: 30 January 2021
© The Author(s) 2021

Abstract
Let F be a non-singular homogeneous polynomial of degree d in n variables. We give an asymptotic formula of the pairs of integer points $(\mathbf{x}, \mathbf{y})$ with $|\mathbf{x}| \leq X$ and $|\mathbf{y}| \leq Y$ which generate a line lying in the hypersurface defined by F, provided that $n > 2^{d-1}d^4(d+1)(d+2)$. In particular, by restricting to Zariski-open subsets we are able to avoid imposing any conditions on the relative sizes of X and Y.

Keywords Forms in many variables · Hardy–Littlewood method · Rational lines

Mathematics Subject Classification 11D72 · 11P55 · 11E76 · 14G05

1 Introduction

Questions concerning the number and distribution of rational points on hypersurfaces have long attracted the interest of both number theorists and algebraic geometers. Building on work by Davenport [15], Birch wrote an influential paper [6] in which he provided a method to prove the analytic Hasse principle and establish asymptotic formulæ for the number of integer points on projective hypersurfaces under moderate non-singularity conditions, provided that the dimension of the hypersurface is sufficiently large compared to its degree. In particular, suppose that $F \in \mathbb{Z}[x_1, \ldots, x_n]$ is a non-singular form of degree d defining a hypersurface $\mathcal{V}$, and write $N(X)$ for the number of points $\mathbf{x} \in \mathcal{V}(\mathbb{Z})$ with $|x_i| \leq X$ for $1 \leq i \leq n$. In this notation, Birch's main result [6, Theorem] states that whenever $n > 2^d(d-1)$, there exists a positive

Communicated by Alberto Minguez.

✉ Julia Brandes
brjulia@chalmers.se

[1] Department for Mathematical Sciences, University of Gothenburg and Chalmers University of Technology, 412 96 Göteborg, Sweden

Published online: 07 March 2021

real number v with the property that the number of integer points on $\mathcal{V}$ satisfies an asymptotic formula of the shape

$$N(X) = cX^{n-d} + O(X^{n-d-v}).$$

The constant c is non-negative and has an interpretation in terms of the density of $\mathbb{Q}_v$-points in $\mathcal{V}$ for all completions $\mathbb{Q}_v$ of $\mathbb{Q}$.

In the work at hand, we study a higher-dimensional generalisation of Birch's result. Denote by $N(X, Y)$ the number of points $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$ satisfying $|x_i| \leq X$ and $|y_i| \leq Y$ for $1 \leq i \leq n$, and having the property that

$$F(u\mathbf{x} + v\mathbf{y}) = 0 \qquad \text{identically in } u \text{ and } v. \tag{1.1}$$

This problem is related to that of counting rational lines contained in $\mathcal{V}$, in that it counts all possible sets of generating pairs $(\mathbf{x}, \mathbf{y})$ of suitably bounded height and with the property that the line spanned by $(\mathbf{x}, \mathbf{y})$ is fully contained in $\mathcal{V}$. Geometrically, it is known that the Fano scheme of lines on a generic hypersurface $\mathcal{V}$ of degree d has dimension $2n - d - 5$ whenever that number is positive (Langer [21]; see also the classical work by Altman and Kleiman [1]). More recent results by Harris, Mazur and Pandharipande [18], Beheshti [3–5] and others explore under what circumstances a similar statement holds for *all* smooth hypersurfaces. Unfortunately, these results within algebraic geometry do not contain much arithmetic information and thus are of limited use if one seeks to study rational lines.

When F is a cubic form, recent work of the author jointly with Dietmann [12] shows that (1.1) has non-trivial rational solutions whenever $n \geq 29$, but that there may not be any rational solutions when $n = 11$ or lower. For more general settings, (1.1) has been investigated in a series of papers by the present author [8–11]. We note at this point that, in order to strictly count lines, we would have to exclude those solutions of (1.1) where $\mathbf{x}$ and $\mathbf{y}$ are proportional. Fortunately, the contribution of such points is of a smaller order of magnitude than our eventual main term, so we do not lose any generality by omitting to explicitly exclude them.

A special role in problems of this flavour is played by certain points $\mathbf{y} \in \mathcal{V}$ that admit a disproportionate number of solutions $\mathbf{x} \in \mathcal{V}$ satisfying (1.1). Typically, the contribution arising from these solutions is counterbalanced by the relative sparsity of such points $\mathbf{y}$, but when Y is very small in comparison to X, such solutions might well dominate the overall count. It is therefore natural to exclude the solutions that arise from such special subvarieties. When $\mathcal{U} \subseteq \mathcal{V}$ is a Zariski-open subset, we denote by $N_{\mathcal{U}}(X, Y)$ the number of integral $\mathbf{x}, \mathbf{y} \in \mathcal{U}$ with $|x_i| \leq X$ and $|y_i| \leq Y$ for $1 \leq i \leq n$ that satisfy (1.1). We can now state the main result of this memoir.

Theorem 1.1 *Let $F \in \mathbb{Z}[x_1, \ldots, x_n]$ be a non-singular form of degree $d \geq 5$ defining a hypersurface $\mathcal{V}$. Let further*

$$n > 2^{d-1} d^4 (d+1)(d+2).$$

Then there exists a non-empty Zariski-open subset $\mathcal{U} \subseteq \mathcal{V}$ and a positive real number ν with the property that

$$N_{\mathcal{U}}(X, Y) = (XY)^{n-\frac{1}{2}d(d+1)}\chi_\infty \prod_{p \text{ prime}} \chi_p + O((XY)^{n-\frac{1}{2}d(d+1)-\nu}).$$

The Euler product converges absolutely, and its factors have an interpretation as the density of solutions of (1.1) *over the local fields $\mathbb{R}$ and $\mathbb{Q}_p$, respectively.*

Note that Theorem 1.1 is a slightly simplified version of what our methods yield; by a more thorough analysis it would be possible to obtain some improvements in the lower-order terms at the expense of a significantly more complicated expression, but no easy improvement of the order of growth $2^d d^6$ in our result. In particular, we do not expect our results to be competitive when d is small. For this reason, even though a modification of our approach would provide results for $d \in \{2, 3, 4\}$ also, we refrain from including the analysis of those cases as the expected results would likely be quite weak.

Clearly, the problem is symmetric in X and Y, so in our discussion we may assume without loss of generality that $Y \leq X$. In the special case when $Y = X$, the conclusion of Theorem 1.1 follows from [8, Theorem 1.1] under the more lenient condition that $n > 3 \cdot 2^d (d-1)(d+2)$, and subsequent work [9, Theorem 2.1] establishes a conclusion similar to that of Theorem 1.1 above under the additional condition that n should be large enough in terms of $\log X / \log Y$, which is acceptable if X is at most a bounded power of Y. The main new input in our present work is therefore our treatment of the situation when Y is vastly smaller than X. Unlike in our former work in [8,9], where we allowed the variables $\mathbf{x}, \mathbf{y}$ to vary independently, we pursue a slicing approach inspired by [23] in which we fix a point $\mathbf{y} \in \mathcal{U}(\mathbb{Z})$ and then investigate the number $N_{\mathbf{y}}(X; \mathcal{U})$ of points $\mathbf{x} \in \mathcal{U}(\mathbb{Z}) \cap [-X, X]^n$ for which (1.1) is satisfied with that particular value $\mathbf{y}$. We then have

$$N_{\mathcal{U}}(X, Y) = \sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \\ |\mathbf{y}| \leq Y}} N_{\mathbf{y}}(X; \mathcal{U}), \tag{1.2}$$

and we aim to establish bounds of the shape

$$N_{\mathbf{y}}(X; \mathcal{U}) = c_{\mathbf{y}} X^{n-\frac{1}{2}d(d+1)} + O_{\mathbf{y}}(X^{n-\frac{1}{2}d(d+1)-\nu})$$

for some constant $c_{\mathbf{y}}$ and some positive number ν.

For generic $\mathbf{y}$, the quantity $N_{\mathbf{y}}(X) = N_{\mathbf{y}}(X; \mathcal{V})$ can be understood by applying the methods of Browning and Heath-Brown [14] for systems of homogeneous equations with differing degrees, although we need to be careful to track the dependence on the coefficients as these will be polynomially dependent on $\mathbf{y}$. Unfortunately, this strategy breaks down if $\mathbf{y}$ fails to satisfy a certain second-order non-singularity condition. When

$H_{\mathbf{x}}$ denotes the Hessian of F at the point $\mathbf{x}$, we set

$$\mathcal{V}_{2,\rho}^* = \{\mathbf{x} \in \mathcal{V} : \operatorname{rank} H_{\mathbf{x}} \leq n - \rho\},$$

and let $\mathcal{V}_{2,\rho} = \mathcal{V} \backslash \mathcal{V}_{2,\rho}^*$. In particular, $\mathcal{V}_{2,\rho}$ is Zariski-open in $\mathcal{V}$ for all $1 \leq \rho \leq n$.

The set $\mathcal{V}_{2,1}^*$ is defined by the zero set of the simultaneous equations $F(\mathbf{x}) = 0$ and $\det H_{\mathbf{x}} = 0$, and is thus of codimension one inside $\mathcal{V}$. Here, the function $\Delta(\mathbf{x}) = \det H_{\mathbf{x}}$ is a form of degree $(d-2)n$ in n variables, and one might hope that, unless the variety defined by $\Delta(\mathbf{x}) = 0$ contains high-dimensional subvarieties of low degree, the set $\mathcal{V}_{2,1}^*(\mathbb{Z})$ might not be too large. In particular, in the eventuality that $\mathcal{V}_{2,1}^*(\mathbb{Z})$ should consist only of the origin for some function F, it would be permissible in Theorem 1.1 to take $\mathcal{U} = \mathcal{V} \backslash \{\mathbf{0}\}$. Unfortunately, our current understanding of the size of the set $\mathcal{V}_{2,\rho}^*$ is quite weak. Nonetheless, by bounding the number of integral points in $\mathcal{V}_{2,\rho}^*$ we are still able to establish asymptotic formulæ for our original counting function $N(X, Y)$ that extend the admissible range of Y compared to what had previously been known in [9, Theorem 2.1] without having to exclude an exceptional subvariety.

Theorem 1.2 *Let $F \in \mathbb{Z}[x_1, \ldots, x_n]$ be a non-singular form of degree $d \geq 5$, and suppose that $Y = X^{\psi}$, where $0 < \psi < (2d^4)^{-1}$. Furthermore, suppose that*

$$n > 2^{d-1} d^4 (d+1)(d+2) + \tfrac{1}{2} d(d-1)\psi^{-1}.$$

Then there exists a positive real number v with the property that

$$N(X, Y) = (XY)^{n - \frac{1}{2}d(d+1)} \chi_\infty \prod_{p \text{ prime}} \chi_p + O((XY)^{n-\frac{1}{2}d(d+1)} X^{-v}),$$

where the local factors are the same as in Theorem 1.1.

The reader may wonder how the lower bound on n compares with that which can be extracted from [9, Theorem 2.1]. In that result, the bound on the number of variables in the case when ψ is small can be written in terms of ψ as

$$n > 2^{d-2} d(d+1)(1 + \psi^{-1}).$$

It is clear that for $\psi \ll d^{-4}$ our new result is significantly stronger.

There are many questions of interest from a geometric point of view that use height functions different from the naive one. In particular, the conjectures of Manin and Manin–Peyre (see [2,17,22]) are phrased in terms of an anticanonical height function, which reflects the dimension and degree of the varieties under consideration. In bi- or multihomogeneous settings, this height forces the variables to lie in domains of a hyperbolic shape, which it is hard to capture from an analytic point of view. Fortunately, a very general method of transferring information between these settings has been provided by Blomer and Brüdern [7] in the context of proving Manin's conjecture for multihomogeneous diagonal equations. The same technology has later been used by Schindler [23] in her analysis of Manin's conjecture for bihomogeneous equations.

Here, the main criterion for the method of [7] to work is that one have a good understanding of the underlying counting function over lopsided boxes. The treatment of boxes with sides of comparable length is usually tractable by some minor modifications of the strategy employed for homogeneous boxes, so a main challenge lies in controlling the extreme cases when the boxes have very unequal sidelengths. In view of such potential future applications we state two by-products of our strategy which may be of independent interest and are simplified versions of Theorems 7.1 and 7.2 below, respectively.

Theorem 1.3 *Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be a non-singular form of degree $d \geq 5$ defining a hypersurface $\mathcal{V}$. Let further $\psi \in (0, 1/(2d^4)]$, and suppose that*

$$n \geq 2^d d(d^2 - 1) + \rho.$$

Then there exists a positive real number ν with the property that

$$N_{\mathbf{y}}(X) = X^{n-\frac{1}{2}d(d+1)}\mathfrak{S}_{\mathbf{y}}\mathfrak{J}_{\mathbf{y}} + O(X^{n-\frac{1}{2}d(d+1)-\nu})$$

uniformly for all $\mathbf{y} \in \mathcal{V}_{2,\rho}(\mathbb{Z})$ satisfying $|\mathbf{y}| \leq X^{\psi}$. Moreover, the local factors satisfy $0 \leq \mathfrak{S}_{\mathbf{y}} \ll_{\mathbf{y}} 1$ and $0 \leq \mathfrak{J}_{\mathbf{y}} \ll_{\mathbf{y}} 1$.

The set $\mathcal{V}^*_{2,\rho}$ is clearly algebraically defined for any ρ, and it is known (see e.g. [20, Lemma 2]) that $\dim \mathcal{V}^*_{2,\rho} \leq n - \rho$. Thus we have $N_{\mathbf{y}}(X; \mathcal{V}_{2,\rho}) = N_{\mathbf{y}}(X) + O(X^{n-\rho})$, and we see that when $\rho > \frac{1}{2}d(d+1)$, the anticipated main term exceeds any error that might arise if we replace $\mathcal{V}_{\rho,2}$ by $\mathcal{V}$ itself. This allows us to derive a bound on $N_{\mathcal{U}}(X, Y)$ from bounds on $N_{\mathbf{y}}(X)$.

Theorem 1.4 *Let F and $\mathcal{V}$ be as before with $d \geq 5$, and for some $\psi \in (0, 1/(2d^4)]$ set $Y = X^{\psi}$. Suppose that*

$$n \geq 2^d d(d^2 - 1)$$

and set $\mathcal{U} = \mathcal{V}_{2,\frac{1}{2}d(d+1)+1}$. Then there exists a real number $\nu > 0$ for which

$$N_{\mathcal{U}}(X, Y) = X^{n-\frac{1}{2}d(d+1)} \sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \\ |\mathbf{y}| \leq Y}} \mathfrak{S}_{\mathbf{y}}\mathfrak{J}_{\mathbf{y}} + O((XY)^{n-\frac{1}{2}d(d+1)-\nu}).$$

Notation Throughout the paper, the following notational conventions will be observed. Any statements containing the letter ε are asserted to hold for all sufficiently small values of ε, and we make no effort to track the precise 'value' of ε, which is consequently allowed to change from one line to the next. We will be liberal in our use of vector notation. In particular, equations and inequalities involving vectors should always be understood entrywise. In this spirit, we write $|\mathbf{x}| = \|\mathbf{x}\|_\infty = \max |x_i|$, as well as $(\mathbf{a}, b) = \gcd(a_1, \dots, a_n, b)$. For $\alpha \in \mathbb{R}$ we write $\|\alpha\| = \min_{z\in\mathbb{Z}} |\alpha - z|$. Finally, the implicit constants in the Landau and Vinogradov notations are allowed to depend on all parameters except X, Y and $\mathbf{y}$.

2 Preliminary manoevres

Let Φ denote the symmetric d-linear form associated to F, so that F can be written as $F(\mathbf{x}) = \Phi(\mathbf{x}, \dots, \mathbf{x})$. Then after expanding, the form F may be written as

$$F(u\mathbf{x} + v\mathbf{y}) = \sum_{j=0}^{d} \binom{d}{j} u^j v^{d-j} \Phi(\underbrace{\mathbf{x}, \dots, \mathbf{x}}_{j \text{ entries}}, \underbrace{\mathbf{y}, \dots, \mathbf{y}}_{d-j \text{ entries}}),$$

and our counting function $N_{\mathcal{U}}(X, Y)$ counts integer solutions $\mathbf{x}, \mathbf{y} \in \mathcal{U}$ to the system of equations

$$\Phi(\underbrace{\mathbf{x}, \dots, \mathbf{x}}_{j \text{ entries}}, \underbrace{\mathbf{y}, \dots, \mathbf{y}}_{d-j \text{ entries}}) = 0 \qquad (0 \le j \le d), \tag{2.1}$$

where $|x_i| \le X$ and $|y_i| \le Y$ for $1 \le i \le n$.

In this and the following sections we fix a value of $\mathbf{y}$ and consider (2.1) as a system of equations in $\mathbf{x}$ only. Eventually, we will have to consider only such choices for $\mathbf{y}$ that lie in a suitable Zariski-open subset $\mathcal{U}$. This allows us in particular to exclude the value $\mathbf{y} = \mathbf{0}$. For $1 \le j \le d$ we write $\Phi_{\mathbf{y}}^{(j)}(\mathbf{x})$ for the form having j entries $\mathbf{x}$ and $d - j$ entries $\mathbf{y}$. In this notation, $N_{\mathbf{y}}(X)$ denotes the number of points $\mathbf{x} \in \mathbb{Z}^n \cap [-X, X]^n$ satisfying

$$\Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) = 0 \qquad (1 \le j \le d). \tag{2.2}$$

The system (2.2) consists of forms of consecutive degrees $1, \dots, d$. Asymptotic formulæ for the number of solutions of such systems can be obtained by the machinery of Browning and Heath-Brown [14]. However, before embarking on that argument, it is convenient to eliminate one variable by solving the linear equation, so that all forms explicitly occurring in the system have degree two or higher. To this end, observe that the equation $\Phi_{\mathbf{y}}^{(1)}(\mathbf{x}) = 0$ can be expressed as

$$l_1(\mathbf{y})x_1 + \dots + l_n(\mathbf{y})x_n = 0, \tag{2.3}$$

where the coefficients $l_i = l_i(\mathbf{y})$ are given by $l_i(\mathbf{y}) = \frac{\partial}{\partial y_i} F(\mathbf{y})$, and are therefore polynomials of degree $d - 1$ in $\mathbf{y}$. Since F is non-singular by assumption, not all l_i can vanish simultaneously, and thus the set of $\mathbf{x} \in \mathbb{Z}^n$ satisfying (2.3) forms an $(n-1)$-dimensional lattice $\Lambda_{\mathbf{y}} \subseteq \mathbb{Z}^n$. Denote by $\mathfrak{A}_{\mathbf{y}}(X) \subseteq \mathbb{Z}^n$ the set of lattice points $\mathbf{x} \in \Lambda_{\mathbf{y}}$ for which $|\mathbf{x}| \le X$. Thus, we may equivalently consider the quantity $N_{\mathbf{y}}(X)$ to be given by the number of points $\mathbf{x} \in \mathfrak{A}_{\mathbf{y}}(X)$ satisfying the system of equations

$$\Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) = 0 \qquad (2 \le j \le d).$$

In order to understand the counting function $N_{\mathbf{y}}(X)$, we encode the summation conditions in exponential sums. Let $\boldsymbol{\alpha} = (\alpha_2, \dots, \alpha_d) \in [0,1)^{d-1}$, then $N_{\mathbf{y}}(X)$ is given by

$$N_{\mathbf{y}}(X) = \sum_{\mathbf{x} \in \mathfrak{A}_{\mathbf{y}}(X)} \int_{[0,1)^{d-1}} e\Big(\sum_{j=2}^{d} \alpha_j \Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) \Big) \, \mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^{d-1}} T_{\mathbf{y}}(\boldsymbol{\alpha}; X) \, \mathrm{d}\boldsymbol{\alpha}, \tag{2.4}$$

where we introduced the exponential sum

$$T_{\mathbf{y}}(\boldsymbol{\alpha}; P) = \sum_{\mathbf{x} \in \mathfrak{A}_{\mathbf{y}}(P)} e\Big(\sum_{j=2}^{d} \alpha_j \Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) \Big).$$

In our arguments below, we will omit the parameter P from the notation whenever there is no danger of confusion. In particular, we drop it in most cases when $P = X$, highlighting it only when we consider exponential sums of size different from X.

For simpler notation below, we write $s = n-1$. The following facts are all contained in [19, Lemma 1] or easy consequences thereof. By part (i) of that lemma, the lattice $\Lambda_{\mathbf{y}}$ has discriminant

$$d(\Lambda_{\mathbf{y}}) \ll |\mathbf{l}(\mathbf{y})| \ll |\mathbf{y}|^{d-1}. \tag{2.5}$$

Since we will require $\psi \le (2d)^{-2}$ and thus $d(\Lambda_{\mathbf{y}}) \ll Y^{d-1} \ll X^{1/d}$, it follows that $\operatorname{Card} \mathfrak{A}_{\mathbf{y}}(X) \asymp X^s / d(\Lambda_{\mathbf{y}})$. We denote the successive minima of the lattice $\Lambda_{\mathbf{y}}$ by $\mu_1 \le \dots \le \mu_s$, setting $\mu_s = \mu_{\max}$, and recall that $\mu_1 \cdots \mu_s \asymp d(\Lambda_{\mathbf{y}})$. Fix a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_s\} \subseteq \mathbb{R}^n$ of $\Lambda_{\mathbf{y}}$, which by part (iii) of the same lemma we are free to choose in such a way that when $\mathbf{x} \in \mathfrak{A}_{\mathbf{y}}(X)$ with $\mathbf{x} = \xi_1 \mathbf{b}_1 + \dots + \xi_s \mathbf{b}_s$, we have $\xi_i \ll X/|\mathbf{b}_i|$ for $1 \le i \le s$. We label the basis elements in increasing order, so that $|\mathbf{b}_1| \le |\mathbf{b}_2| \le \dots \le |\mathbf{b}_s|$. With $\mathcal{B}$ thusly chosen, it then is an immediate consequence of the second statement of part (iii) of the abovementioned lemma together with the definition of the successive minima that $|\mathbf{b}_i| \asymp \mu_i$ for $1 \le i \le s$.

Set

$$\mathfrak{B}_{\mathbf{y}}(X) = \prod_{i=1}^{s} [-cX/\mu_i, cX/\mu_i],$$

where $c \ll 1$ is chosen large enough so that the coordinate vector $\boldsymbol{\xi}$ of $\mathbf{x}$ lies in $\mathfrak{B}_{\mathbf{y}}(X)$ whenever $\mathbf{x} \in \mathfrak{A}_{\mathbf{y}}(X)$. Moreover, for $2 \le j \le d$ set $\Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi}) = \Phi_{\mathbf{y}}^{(j)}(\mathbf{x})$ and write

$$\phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi}) = \sum_{j=2}^{d} \alpha_j \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi}).$$

By an argument along the lines of that of Lemma 5.2 in [8] one sees that

$$T_{\mathbf{y}}(\boldsymbol{\alpha}) \ll X^{\varepsilon} U_{\mathbf{y}}(\boldsymbol{\alpha}), \tag{2.6}$$

where

$$U_{\mathbf{y}}(\boldsymbol{\alpha}) = \sup_{\boldsymbol{\eta}\in[0,1]^{d-1}} \Big| \sum_{\boldsymbol{\xi}\in\mathfrak{B}_{\mathbf{y}}(X)} e(\phi_{\mathbf{y}}(\boldsymbol{\alpha};\boldsymbol{\xi}) + \boldsymbol{\eta}\cdot\boldsymbol{\xi}) \Big|.$$

This exponential sum is related to that considered by Schindler and Sofos [24] in their treatment of forms in many variables over lopsided boxes. In comparison with their result, however, our argument is more sensitive to the degree of the lopsidedness of the box. Fortunately, the discriminant of our lattice is fairly small. Indeed, since our methods will break down when $\psi \gg 1/d^2$ (see (3.4) below), and for our theorems we require even $\psi \ll 1/d^4$, we find ourselves in a situation where the discriminant of our lattice satisfies the bound $d(\Lambda_{\mathbf{y}}) \ll X^{(d-1)\psi} \ll X^{O(1/d^3)}$.

3 Van der Corput differences

The discrete differencing operator ∂ is defined by its action on a polynomial F via the relation $\partial_{\mathbf{h}}F(\mathbf{x}) = F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x})$, and we write

$$\partial_{\mathbf{h}_i,\dots,\mathbf{h}_1}F(\mathbf{x}) = \partial_{\mathbf{h}_i}\cdots\partial_{\mathbf{h}_1}F(\mathbf{x})$$

for its i-fold iteration. This allows us to state our basic differencing lemma, which is fairly straightforward and essentially follows from [6, Lemma 2.1].

Lemma 3.1 *Let $1 \le i \le d-1$. Then one has*

$$|U_{\mathbf{y}}(\boldsymbol{\alpha})|^{2^i} \ll \Big(\frac{X^s}{d(\Lambda_{\mathbf{y}})}\Big)^{(2^i-i-1)} \sum_{\substack{\mathbf{h}_l\in\mathfrak{B}_{\mathbf{y}}(X)\\ 1\le l\le i}} \Big| \sum_{\boldsymbol{\xi}\in\mathfrak{C}} e\Big(\partial_{\mathbf{h}_i,\dots,\mathbf{h}_1}\phi_{\mathbf{y}}(\boldsymbol{\alpha};\boldsymbol{\xi})\Big)\Big|,$$

where the sets $\mathfrak{C} = \mathfrak{C}(\mathbf{h}_1,\dots,\mathbf{h}_i)$ are boxes (possibly empty) contained inside $\mathfrak{B}_{\mathbf{y}}(X)$.

Proof Upon recalling that $\mathrm{Card}\,\mathfrak{B}_{\mathbf{y}}(X) \asymp X^s/d(\Lambda_{\mathbf{y}})$, this is a straightforward reformulation of the standard Weyl differencing procedure as for instance in Davenport's monograph [16, Chapter 13]. □

At this stage, the usual procedure would be to apply Lemma 3.1 with $i = d-1$, so that the argument of the exponential function becomes linear, thus yielding either a non-trivial upper bound or good approximations to the coefficient α_d. In the situation at hand, however, this approach would lose all information connected to the forms $\Psi_{\mathbf{y}}^{(j)}$ with $j < d$. So instead we follow the approach by Browning and Heath-Brown [14] and replace the last Weyl differencing step by a suitable van der Corput step. For $2 \le j \le d$ we define functions $B_{\mathbf{y},m}^{(j)}$ for $1 \le m \le s$ via the relation

$$\Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi},\mathbf{h}_1,\dots,\mathbf{h}_{j-1}) = \sum_{m=1}^{s} \xi_m B_{\mathbf{y},m}^{(j)}(\mathbf{h}_1,\dots,\mathbf{h}_{j-1}). \tag{3.1}$$

Furthermore, let $\theta_2, \dots, \theta_d$ be parameters in the unit interval which will be fixed later, and define

$$v_j = (j-1)\theta_j \qquad \text{and} \qquad \omega_j = \sum_{i=j}^{d} v_i \qquad (2 \le j \le d). \tag{3.2}$$

Set further $D_j = \frac{1}{2}j(j+1)$ for $1 \le j \le d$, and for integers q_j with $2 \le j \le d$ put

$$Q_j = \prod_{i=j}^{d} q_i. \tag{3.3}$$

For notational reasons we write

$$D = D_d, \qquad D_0 = 0, \qquad \omega_{d+1} = 0 \qquad \text{and} \qquad Q_{d+1} = 1,$$

and we assume

$$\psi < 1/(2d^2) \tag{3.4}$$

throughout. For fixed $\theta_{j+1}, \dots, \theta_d$ set

$$R_j = X^{1-\omega_{j+1}} |\mathbf{y}|^{-D_{d-j}} \mu_{\max}^{-(d-j)} \tag{3.5}$$

and

$$\Upsilon_j = \sum_{\substack{\mathbf{h}_l \in \mathfrak{B}_{\mathbf{y}}(X) \\ 1 \le l \le j-2}} \sum_{\mathbf{w} \in \mathfrak{B}_{\mathbf{y}}(2R_j)} \prod_{m=1}^{s} \min\left\{ \frac{X}{\mu_m}, \left\| j! Q_{j+1} \alpha_j B_{\mathbf{y},m}^{(j)}(\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, \mathbf{w}) \right\|^{-1} \right\}. \tag{3.6}$$

We can now state one of our key iterative lemmas.

Lemma 3.2 *Let $j \in \{2, \dots, d\}$ be fixed. When $j < d$, suppose that $\theta_{j+1}, \dots, \theta_d$ are fixed in such a way that, observing (3.2), one has*

$$\omega_{j+1} + \psi(D_{d-j} + (d-1)(d-j)) < 1. \tag{3.7}$$

Suppose that for any i with $j < i \le d$ there exists a natural number q_i satisfying $q_i \ll X^{v_i} |\mathbf{y}|^{d-i} \mu_{\max}$ with the property that, in view of (3.3), one has

$$\| Q_i \alpha_i \| \ll X^{-i+\omega_i} |\mathbf{y}|^{D_{d-i}} \mu_{\max}^{d-i+1}.$$

Then we have the bound

$$|U_{\mathbf{y}}(\boldsymbol{\alpha})|^{2^{j-1}} \ll \left(\frac{X^s}{d(\Lambda_{\mathbf{y}})}\right)^{2^{j-1}-(j-1)} \left(\frac{R_j^s}{d(\Lambda_{\mathbf{y}})}\right)^{-1} \Upsilon_j.$$

Proof Suppose first that $j > 2$. In this case, applying Lemma 3.1 with $i = j-2$ followed by an application of Cauchy's inequality gives

$$|U_{\mathbf{y}}(\boldsymbol{\alpha})|^{2^{j-1}} \ll \left(\frac{X^s}{d(\Lambda_{\mathbf{y}})}\right)^{2^{j-1}-j} \sum_{\substack{\mathbf{h}_l \in \mathfrak{B}_{\mathbf{y}}(X) \\ 1 \le l \le j-2}} \left| \sum_{\boldsymbol{\xi} \in \mathfrak{C}_1} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi})) \right|^2 \tag{3.8}$$

for suitable boxes $\mathfrak{C}_1 = \mathfrak{C}_1(\mathbf{h}_1, \dots, \mathbf{h}_{j-2}) \subseteq \mathfrak{B}_{\mathbf{y}}(X)$. Let now $\mathbf{h}_1, \dots, \mathbf{h}_{j-2}$ be temporarily fixed, and observe that our hypothesis concerning the size of the q_i implies via (3.3) that $R_j Q_{j+1} \ll X$. Consequently, we have

$$\begin{aligned} &\frac{R_j^s}{d(\Lambda_{\mathbf{y}})} \sum_{\boldsymbol{\xi} \in \mathfrak{C}_1} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi})) \\ &\quad \ll \sum_{\mathbf{u} \in \mathfrak{B}_{\mathbf{y}}(R_j)} \sum_{\substack{\boldsymbol{\xi} \\ \boldsymbol{\xi} + Q_{j+1}\mathbf{u} \in \mathfrak{C}_1}} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{u})). \end{aligned} \tag{3.9}$$

We denote by $\mathfrak{C}_2$ the set of $\boldsymbol{\xi}$ for which $\boldsymbol{\xi} + Q_{j+1}\mathbf{u} \in \mathfrak{C}_1$ for some $\mathbf{u} \in \mathfrak{B}_{\mathbf{y}}(R_j)$; this box has cardinality $\operatorname{Card} \mathfrak{C}_2 \asymp X^s / d(\Lambda_{\mathbf{y}})$. Then with another application of Cauchy's inequality one obtains from (3.9) the bound

$$\begin{aligned} &\left(\frac{R_j^s}{d(\Lambda_{\mathbf{y}})}\right)^2 \left| \sum_{\boldsymbol{\xi} \in \mathfrak{C}_1} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi})) \right|^2 \\ &\quad \ll \frac{X^s}{d(\Lambda_{\mathbf{y}})} \sum_{\boldsymbol{\xi} \in \mathfrak{C}_2} \left| \sum_{\substack{\mathbf{u} \in \mathfrak{B}_{\mathbf{y}}(R_j) \\ \boldsymbol{\xi} + Q_{j+1}\mathbf{u} \in \mathfrak{C}_1}} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{u})) \right|^2 \\ &\quad \ll \frac{X^s}{d(\Lambda_{\mathbf{y}})} \sum_{\mathbf{u},\mathbf{v} \in \mathfrak{B}_{\mathbf{y}}(R_j)} \left| \sum_{\boldsymbol{\xi}} e\big(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}} (\phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{u}) - \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{v}))\big) \right|, \end{aligned}$$

where the inner sum runs over all $\boldsymbol{\xi}$ for which both $\boldsymbol{\xi} + Q_{j+1}\mathbf{u}$ and $\boldsymbol{\xi} + Q_{j+1}\mathbf{v}$ lie in $\mathfrak{C}_1$.

We now make the change of variables $\boldsymbol{\xi}' = \boldsymbol{\xi} + Q_{j+1}\mathbf{v}$ and $\mathbf{w} = \mathbf{u} - \mathbf{v}$, so that

$$\phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{u}) - \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi} + Q_{j+1}\mathbf{v}) = \partial_{Q_{j+1}\mathbf{w}} \phi_{\mathbf{y}}(\boldsymbol{\alpha}; \boldsymbol{\xi}').$$

Thus, upon summing trivially over $\mathbf{v}$, we have shown that

$$\begin{aligned}&\left|\sum_{\boldsymbol{\xi}\in\mathfrak{C}_1} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2}}\phi_{\mathbf{y}}(\boldsymbol{\alpha};\boldsymbol{\xi}))\right|^2\\ &\ll \left(\frac{X}{R_j}\right)^s \sum_{\mathbf{w}\in\mathfrak{B}_{\mathbf{y}}(2R_j)} \sup_{\mathfrak{C}\subseteq\mathfrak{B}_{\mathbf{y}}(X)} \left|\sum_{\boldsymbol{\xi}\in\mathfrak{C}} e(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2},Q_{j+1}\mathbf{w}}\phi_{\mathbf{y}}(\boldsymbol{\alpha};\boldsymbol{\xi}))\right|,\end{aligned}$$

where the supremum is over all coordinate-aligned boxes $\mathfrak{C}$ inside $\mathfrak{B}_{\mathbf{y}}(X)$. Thus, upon combining this bound with (3.8), it follows that the exponential sum can be bounded above via

$$|U_{\mathbf{y}}(\boldsymbol{\alpha})|^{2^{j-1}} \ll \left(\frac{X^s}{d(\Lambda_{\mathbf{y}})}\right)^{2^{j-1}-(j-1)} \left(\frac{R_j^s}{d(\Lambda_{\mathbf{y}})}\right)^{-1} \mathcal{W}_j,$$

where

$$\mathcal{W}_j = \sum_{\substack{\mathbf{h}_l\in\mathfrak{B}_{\mathbf{y}}(X)\\ 1\le l\le j-2}} \sum_{\mathbf{w}\in\mathfrak{B}_{\mathbf{y}}(2R_j)} \sup_{\mathfrak{C}\subseteq\mathfrak{B}_{\mathbf{y}}(X)} \left|\sum_{\boldsymbol{\xi}\in\mathfrak{C}} e\left(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2},Q_{j+1}\mathbf{w}} \sum_{i=2}^{d} \alpha_i \Psi_{\mathbf{y}}^{(i)}(\boldsymbol{\xi})\right)\right|. \tag{3.10}$$

An analogous bound is also derived easily in the omitted case when $j = 2$ upon interpreting the empty sum over $\mathbf{h}_l$ and the concomitant differences as void, and noting that the phase factor in $U_{\mathbf{y}}(\boldsymbol{\alpha})$ disappears in the van der Corput step.

The size of the innermost exponential sum in (3.10) is dominated by the term corresponding to $i = j$. In fact, observe that after $j-1$ differences taken only the terms $\Psi_{\mathbf{y}}^{(i)}(\boldsymbol{\xi})$ with $i \ge j$ occur explicitly in the argument of the exponential, and due to the last Q_{j+1}-van der Corput step all of these contain a factor Q_{j+1}. Hence whenever $j < d$ and $1 \le l \le s$ one has

$$\begin{aligned}&\frac{\partial}{\partial\xi_l} e\left(\partial_{\mathbf{h}_1,\dots,\mathbf{h}_{j-2},Q_{j+1}\mathbf{w}} \sum_{i=j+1}^{d} \alpha_i \Psi_{\mathbf{y}}^{(i)}(\boldsymbol{\xi})\right)\\ &\ll \sum_{i=j+1}^{d} \|Q_{j+1}\alpha_i\| X^{i-2} R_j |\mathbf{y}|^{d-i} \mu_l\\ &\ll \sum_{i=j+1}^{d} \left|\frac{Q_{j+1}}{Q_i}\right| \|Q_i\alpha_i\| X^{i-1-\omega_{j+1}} |\mathbf{y}|^{-D_{d-j}+d-i} \mu_{\max}^{-(d-j)} \mu_l\\ &\ll X^{-1}\mu_l,\end{aligned}$$

where in the last step we used the hypotheses of the lemma. Upon iterating this procedure, one confirms for any subset $\{l_1, \dots, l_k\} \subseteq \{1, \dots, s\}$ that

$$\frac{\partial^k}{\partial \xi_{l_1} \cdots \partial \xi_{l_k}} e\bigg(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \sum_{i=j+1}^{d} \alpha_i \Psi_{\mathbf{y}}^{(i)}(\boldsymbol{\xi}) \bigg) \ll X^{-k} \mu_{l_1} \cdots \mu_{l_k}.$$

Suppose that $\mathfrak{C} = \prod_i [C_i, C_i']$, recalling that $\mathfrak{C} \subseteq \mathfrak{B}$ forces $\max\{|C_i|, |C_i'|\} \ll X/\mu_i$ for $1 \leq i \leq s$. Thus, it follows from multidimensional partial summation that

$$\begin{aligned}
&\sum_{\boldsymbol{\xi} \in \mathfrak{C}} e\bigg(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \sum_{i=j}^{d} \alpha_i \Psi_{\mathbf{y}}^{(i)}(\boldsymbol{\xi}) \bigg) \\
&\quad \ll \bigg| \sum_{\boldsymbol{\xi} \in \mathfrak{C}} e(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \alpha_j \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})) \bigg| \\
&\qquad + \sum_{l=1}^{s} \frac{\mu_l}{X} \int_{C_l}^{C_l'} \bigg| \sum_{\substack{\boldsymbol{\xi} \in \mathfrak{C} \\ \xi_l \leq t}} e(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \alpha_j \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})) \bigg| \, \mathrm{d}t \\
&\qquad + \dots + \frac{\mu_1 \cdots \mu_s}{X^s} \int_{\mathfrak{C}} \bigg| \sum_{\substack{\boldsymbol{\xi} \in \mathfrak{C} \\ \xi_l \leq t_l \ (1 \leq l \leq s)}} e(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \alpha_j \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})) \bigg| \, \mathrm{d}\mathbf{t} \\
&\quad \ll \sup_{\mathfrak{C}' \subseteq \mathfrak{C}} \bigg| \sum_{\boldsymbol{\xi} \in \mathfrak{C}'} e(\partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \alpha_j \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})) \bigg|,
\end{aligned}$$

where the supremum is over all coordinate-aligned boxes $\mathfrak{C}' \subseteq \mathfrak{C}$. Thus, we discern that the dominant contribution arises indeed from the term of degree j, so that

$$\mathcal{W}_j \ll \sum_{\substack{\mathbf{h}_l \in \mathfrak{B}_{\mathbf{y}}(X) \\ 1 \leq l \leq j-2}} \sum_{\mathbf{w} \in \mathfrak{B}_{\mathbf{y}}(2R_j)} \sup_{\mathfrak{C} \subseteq \mathfrak{B}_{\mathbf{y}}(X)} \bigg| \sum_{\boldsymbol{\xi} \in \mathfrak{C}} e\Big(\alpha_j \partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi}) \Big) \bigg|.$$

The argument of the exponential is now linear in $\boldsymbol{\xi}$. Since $\mathfrak{C} \subseteq \mathfrak{B}_{\mathbf{y}}(X)$ is a box oriented along the coordinate axes, upon recalling the definition (3.1) the standard estimate on linear exponential sums yields the bound

$$\begin{aligned}
&\bigg| \sum_{\boldsymbol{\xi} \in \mathfrak{C}} e\Big(\alpha_j \partial_{\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, Q_{j+1}\mathbf{w}} \Psi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi}) \Big) \bigg| \\
&\qquad \ll \prod_{m=1}^{s} \min \bigg\{ \frac{X}{\mu_m}, \Big\| j! Q_{j+1} \alpha_j B_{\mathbf{y},m}^{(j)}(\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, \mathbf{w}) \Big\|^{-1} \bigg\}.
\end{aligned}$$

Thus we have shown that $\mathcal{W}_j \ll \Upsilon_j$ and the proof of the lemma is complete. □

4 Geometry of numbers and a nonsingularity condition

The next step is to estimate the quantity Υ_j. For positive real numbers U, V, W set

$$N_{j,\mathbf{y}}(U, V; W) = \mathrm{Card}\Big\{\mathbf{h}_1, \dots, \mathbf{h}_{j-2} \in \mathfrak{B}_{\mathbf{y}}(U), \mathbf{z} \in \mathfrak{B}_{\mathbf{y}}(V), \\ \Big\| j! Q_{j+1} \alpha_j B^{(j)}_{\mathbf{y},m}(\mathbf{h}_1, \dots, \mathbf{h}_{j-2}, \mathbf{z}) \Big\| < \frac{\mu_m}{W} \quad (1 \le m \le s)\Big\}. \tag{4.1}$$

In this notation, standard arguments similar to those in the proof of [16, Lemma 13.2] show that for any fixed $\theta_{j+1}, \dots, \theta_d$ one has

$$\Upsilon_j \ll \left(\frac{X^s}{d(\Lambda_{\mathbf{y}})}\right)^{1+\varepsilon} N_{j,\mathbf{y}}(X, R_j; X). \tag{4.2}$$

Our next goal is to bound the size of $N_{j,\mathbf{y}}(X, R_j; X)$. For this purpose we need a generalisation of Davenport's lemma on the geometry of numbers (see [16, Lemma 12.6]). Let $A_{k,m} > 1$ be real numbers for $1 \le k \le j-1$, $1 \le m \le s$, and write

$$\mathscr{A}_k = \prod_{m=1}^{s} [-A_{k,m}, A_{k,m}] \qquad (1 \le k \le j-1).$$

Let further $0 < Z_k \le 1$ for $1 \le k \le j-1$. For any l with $1 \le l \le j-1$ write $\mathcal{R}_l(Z)$ for the number of $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1} \in \mathbb{Z}^s$ such that $\boldsymbol{\xi}_k \in Z_k \mathscr{A}_k$ for all $1 \le k \le j-1$ with $k \ne l$ and $\boldsymbol{\xi}_l \in Z\mathscr{A}_l$, having the property that

$$\Big\| j! Q_{j+1} \alpha_j B^{(j)}_{\mathbf{y},m}(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1}) \Big\| \le Z A_{l,m}^{-1} \qquad (1 \le m \le s).$$

In this notation, Schindler and Sofos [24] give the following variant of Davenport's result.

Lemma 4.1 (Lemma 2.4 in [24]). *Fix $Z_1, \dots, Z_{j-1} \in (0, 1]$ and l with $1 \le l \le j-1$. For any Z, Z' in the range $0 < Z' \le Z \le 1$ one has*

$$\mathcal{R}_l(Z) \ll (Z/Z')^s \mathcal{R}_l(Z').$$

Suppose that $\theta_{j+1}, \dots, \theta_d$ are fixed in such a way that (3.7) is satisfied. For any θ_j satisfying

$$0 < \theta_j \le 1 - \omega_{j+1} - \psi(D_{d-j} + (d-1)(d-j)) \tag{4.3}$$

and all $1 \le m \le s$ we set

$$A_{k,m} = (X/\mu_m) X^{\frac{(1-\theta_j)(k-1)}{2}}, \quad Z_k = X^{-\frac{(1-\theta_j)(k-1)}{2}}, \quad Z'_k = X^{-\frac{(1-\theta_j)(k+1)}{2}}$$

for $1 \le k \le j-2$, and

$$A_{j-1,m} = \frac{(R_j X)^{1/2}}{\mu_m} X^{\frac{(1-\theta_j)(j-2)}{2}}, \qquad Z_{j-1} = \left(\frac{R_j}{X}\right)^{1/2} X^{-\frac{(1-\theta_j)(j-2)}{2}},$$
$$Z'_{j-1} = \left(\frac{X}{R_j}\right)^{1/2} X^{-\frac{(1-\theta_j)j}{2}}.$$

Thus $0 < Z'_k < Z_k \le 1$ for all k, and one has

$$A_{k,m} Z_k = X/\mu_m, \qquad A_{k,m} Z'_k = X^{\theta_j}/\mu_m, \qquad Z_k/Z'_k = X^{1-\theta_j},$$
$$Z'_k/A_{k,m} = \mu_m X^{-1-(1-\theta_j)k} = Z_{k+1}/A_{k+1,m}$$

for $1 \le k \le j-2$, and

$$A_{j-1,m} Z_{j-1} = R_j/\mu_m, \qquad A_{j-1,m} Z'_{j-1} = X^{\theta_j}/\mu_m, \qquad Z_{j-1}/Z'_{j-1} = R_j X^{-\theta_j},$$
$$Z_{j-1}/A_{j-1,m} = \mu_m X^{-1-(1-\theta_j)(j-2)}, \qquad Z'_{j-1}/A_{j-1,m} = \frac{\mu_m}{R_j} X^{-(j-1)(1-\theta_j)}.$$

Note here that (4.3) implies via (3.5) and (2.5) that $R_j > X^{\theta_j}$. Applying Lemma 4.1 consecutively for the indices $k = 1, \dots, j-1$ shows that

$$N_{j,\mathbf{y}}(X, R_j; X) \ll X^{(j-2)(1-\theta_j)s} (R_j X^{-\theta_j})^s N_{j,\mathbf{y}}(X^{\theta_j}, X^{\theta_j}; X^{(j-1)(1-\theta_j)} R_j),$$

and hence we infer from (4.2) that

$$\Upsilon_j \ll \frac{X^{(j-1)(1-\theta_j)s+\varepsilon} R_j^s}{d(\Lambda_{\mathbf{y}})} N_{j,\mathbf{y}}(X^{\theta_j}, X^{\theta_j}; X^{(j-1)(1-\theta_j)} R_j). \tag{4.4}$$

If we now make the assumption that $|T_{\mathbf{y}}(\boldsymbol{\alpha})| \gg (X^s/d(\Lambda_{\mathbf{y}})) X^{-k_j\theta_j}$ for some $k_j > 0$ and some θ_j satisfying (4.3), we obtain from Lemma 3.2 together with (2.6) and (4.4) the bound

$$N_{j,\mathbf{y}}(X^{\theta_j}, X^{\theta_j}; X^{-(j-1)(1-\theta_j)} R_j) \gg \left(\frac{X^{\theta_j s}}{d(\Lambda_{\mathbf{y}})}\right)^{j-1} X^{-2^{j-1}k_j\theta_j - \varepsilon}.$$

The diophantine approximation condition that is implicit in (4.1) is satisfied either if the functions $B^{(j)}_{\mathbf{y},m}$ $(1 \le m \le s)$ tend to vanish for geometric reasons, or if α_j has a good approximation in the rational numbers. Suppose that $j! B^{(j)}_{\mathbf{y},m}(\mathbf{h}_1, \dots, \mathbf{h}_{j-1})$ is non-zero for some m and some choice of $\mathbf{h}_1, \dots, \mathbf{h}_{j-1}$ counted by $N_{j,\mathbf{y}}(X^{\theta_j}, X^{\theta_j}; X^{(j-1)(1-\theta_j)} R_j)$, and denote its absolute value by q_j. Then $q_j \ll X^{\nu_j} |\mathbf{y}|^{d-j} \mu_{\max}$, and the approximation condition implied by the definition (4.1) takes the shape

$$\|\alpha_j q_j Q_{j+1}\| \ll \mu_{\max} X^{(j-1)(\theta_j-1)} R_j^{-1} \ll X^{-j+\omega_j} |\mathbf{y}|^{D_{d-j}} \mu_{\max}^{d-j+1}.$$

We summarise the conclusions of our arguments in a lemma.

Lemma 4.2 *Let $j \in \{2, \dots, d\}$ be fixed. Recalling (3.2), when $j < d$ assume that $\theta_{j+1}, \dots, \theta_d$ are such that (3.7) is satisfied. Suppose that for any i with $j < i \leq d$ there are positive integers $q_i \ll X^{\nu_i} |\mathbf{y}|^{d-i} \mu_{\max}$ with the property that, in view of (3.3), one has*

$$\| Q_i \alpha_i \| \ll X^{-i+\omega_i} |\mathbf{y}|^{D_{d-i}} \mu_{\max}^{d-i+1}.$$

Finally, take $k_j > 0$ and $\theta_j > 0$ to be parameters, where θ_j satisfies (4.3). For any $\boldsymbol{\alpha} \in [0,1)^{d-1}$ one of the following holds.

(A) The exponential sum is bounded by

$$|T_{\mathbf{y}}(\boldsymbol{\alpha})| \ll \frac{X^s}{d(\Lambda_{\mathbf{y}})} X^{-k_j\theta_j+\varepsilon}.$$

(B) There exist integers a_j and q_j satisfying $1 \leq q_j \ll X^{\nu_j} |\mathbf{y}|^{d-j} \mu_{\max}$ as well as $0 \leq a_j \leq Q_j$ such that

$$|Q_j \alpha_j - a_j| \ll X^{-j+\omega_j} |\mathbf{y}|^{D_{d-j}} \mu_{\max}^{d-j+1}.$$

(C) The number of $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1} \in \mathfrak{B}_{\mathbf{y}}(X^{\theta_j})$ for which $B^{(j)}_{\mathbf{y},m}(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1}) = 0$ for $1 \leq m \leq s$ is at least of order $(X^{\theta_j s}/d(\Lambda_{\mathbf{y}}))^{j-1} X^{-2^{j-1}k_j\theta_j-\varepsilon}$.

Our next goal is to interpret the third case geometrically. Write $\mathcal{M}_j(\mathbf{y})$ for the variety containing all $(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1}) \in \mathbb{A}_{\mathbb{C}}^{(j-1)s}$ that satisfy

$$B^{(j)}_{\mathbf{y},m}(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{j-1}) = 0 \quad (1 \leq m \leq s).$$

It is clear (for instance from Theorem 3.1 in [13]) that for any positive real number Z one has

$$\mathrm{Card}\left\{(\mathbf{h}_1, \dots, \mathbf{h}_{j-1}) \in \mathbb{Z}^{(j-1)s} \cap \mathcal{M}_j(\mathbf{y}) : |\mathbf{h}_i| \leq Z \ (1 \leq i \leq j-1)\right\} \ll Z^{\dim \mathcal{M}_j(\mathbf{y})}.$$

As in the work of Schindler and Sofos [24] we cover the domain $(\mathfrak{B}_{\mathbf{y}}(X^{\theta_j}))^{j-1}$ by at most $O(\mu_{\max}^{s(j-1)}/d(\Lambda_{\mathbf{y}})^{j-1})$ translates of the box $[-X^{\theta_j}/\mu_{\max}, X^{\theta_j}/\mu_{\max}]^{s(j-1)}$. In particular, since $d(\Lambda_{\mathbf{y}}) \asymp \mu_1 \cdots \mu_s \ll \mu_{\max}^s$, the number of boxes in the covering is positive. Suppose now that

$$\psi < \varpi k_j \theta_j \tag{4.5}$$

for all j and a suitably small parameter ϖ, so that $\mu_{\max} \ll X^{(d-1)\varpi k_j \theta_j}$. Since [13, Theorem 3.1] allows for translations, we infer that

$$\begin{aligned}\operatorname{Card}\Big\{(\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_{j-1}) \in (\mathfrak{B}_{\mathbf{y}}(X^{\theta_j}))^{j-1} \cap \mathcal{M}_j(\mathbf{y})\Big\}&\\ \ll \left(\frac{\mu_{\max}^s}{d(\Lambda_{\mathbf{y}})}\right)^{j-1}\left(\frac{X^{\theta_j}}{\mu_{\max}}\right)^{\dim \mathcal{M}_j(\mathbf{y})}&\\ \ll \left(\frac{X^{(d-1)\varpi k_j\theta_j s}}{d(\Lambda_{\mathbf{y}})}\right)^{j-1}\left(X^{\theta_j(1-(d-1)\varpi k_j)}\right)^{\dim \mathcal{M}_j(\mathbf{y})},&\end{aligned}$$

where in the second step we used that $(j-1)s - \dim \mathcal{M}_j(\mathbf{y}) \geq 0$ trivially by the definition of $\mathcal{M}_j(\mathbf{y})$ as a variety inside $\mathbb{A}^{(j-1)s}$. We thus discern that whenever we are in case (C) of Lemma 4.2, we must have the bound

$$\left(\frac{X^{\theta_j s}}{d(\Lambda_{\mathbf{y}})}\right)^{j-1} X^{-2^{j-1}k_j\theta_j-\varepsilon} \ll \left(\frac{X^{(d-1)\varpi k_j\theta_j s}}{d(\Lambda_{\mathbf{y}})}\right)^{j-1}\left(X^{\theta_j(1-(d-1)\varpi k_j)}\right)^{\dim \mathcal{M}_j(\mathbf{y})},$$

which simplifies to

$$(X^{\theta_j(1-(d-1)\varpi k_j)})^{(j-1)s-\dim \mathcal{M}_j(\mathbf{y})} \ll X^{2^{j-1}k_j\theta_j-\varepsilon}.$$

It follows that for any j, the case (C) of Lemma 4.2 is excluded when

$$(j-1)s - \dim \mathcal{M}_j(\mathbf{y}) > \frac{2^{j-1}k_j}{1-(d-1)\varpi k_j}. \tag{4.6}$$

We thus want to choose our parameters in such a way that (4.6) holds for all $2 \leq j \leq d$.

We begin by observing that $\mathcal{M}_{j-1}(\mathbf{y})$ is obtained from $\mathcal{M}_j(\mathbf{y})$ by intersecting with the s hyperplanes defined by $\mathbf{h}_{j-1} = \mathbf{y}$. This gives the inequality $\dim \mathcal{M}_{j-1}(\mathbf{y}) \geq \dim \mathcal{M}_j(\mathbf{y}) - s$ for all j with $3 \leq j \leq d$, and upon solving the recursion we deduce that

$$\dim \mathcal{M}_j(\mathbf{y}) \leq (j-2)s + \dim \mathcal{M}_2(\mathbf{y}) \qquad (2 \leq j \leq d). \tag{4.7}$$

It thus suffices to understand the set $\mathcal{M}_2(\mathbf{y})$.

Lemma 4.3 *Let $\mathbf{y} \in \mathcal{V}$. We have $\mathcal{M}_2(\mathbf{y}) = \langle \ker H_{\mathbf{y}}, \mathbf{y}\rangle$ and thus*

$$\dim \mathcal{M}_2(\mathbf{y}) \leq \dim \ker H_{\mathbf{y}} + 1.$$

Proof It follows from the definition of $\Psi^{(2)}$ that $\mathcal{M}_2(\mathbf{y})$ is given by the set of all $\mathbf{h} \in \mathbb{A}^n_{\mathbb{C}}$ satisfying $\Phi^{(1)}_{\mathbf{y}}(\mathbf{h}) = 0$ and $(H_{\mathbf{y}}\mathbf{h})\cdot\mathbf{x} = 0$ for all $\mathbf{x}$ having $\Phi^{(1)}_{\mathbf{y}}(\mathbf{x}) = 0$. In particular, $\mathbf{h}$ has to be such that $(H_{\mathbf{y}}\mathbf{h})\cdot\mathbf{x} = 0$ whenever $(H_{\mathbf{y}}\mathbf{y})\cdot\mathbf{x} = 0$. This is clearly satisfied if $\mathbf{h} \in \ker H_{\mathbf{y}}$, as then the first equation holds trivially. On the other hand, if

$\mathbf{h} \notin \ker H_{\mathbf{y}}$, both equations define hyperplanes which coincide precisely if the vectors $H_{\mathbf{y}}\mathbf{h}$ and $H_{\mathbf{y}}\mathbf{y}$ are proportional, or in other words, $\mathbf{h} - \alpha\mathbf{y} \in \ker H_{\mathbf{y}}$ for some scalar α. Rewriting gives $\mathbf{h} \in \langle \ker H_{\mathbf{y}}, \mathbf{y}\rangle$, and the statement follows. □

We now quantify the set of points $\mathbf{y}$ for which $\ker H_{\mathbf{y}}$ is large. For a natural number ρ set

$$\mathcal{A}(\rho) = \{\mathbf{y} \in \mathbb{A}^n_{\mathbb{C}} : \dim \ker H_{\mathbf{y}} \leq \rho - 1\}$$

and

$$\mathcal{B}(\rho) = \{\mathbf{y} \in \mathbb{A}^n_{\mathbb{C}} : \dim \ker H_{\mathbf{y}} \geq \rho\},$$

so that the sets $\mathcal{A}(\rho)$ and $\mathcal{B}(\rho)$ are complementary. Observe also that with this definition we have $\mathcal{V}^*_{2,\rho} = \mathcal{B}(\rho) \cap \mathcal{V}$ and $\mathcal{V}_{2,\rho} = \mathcal{A}(\rho) \cap \mathcal{V}$. Suppose that $\mathbf{y} \in \mathcal{A}(\rho)$ for some natural number ρ. It then follows from (4.6) and (4.7) via Lemma 4.3 that case (C) of Lemma 4.2 is excluded whenever the inequalities

$$s - \rho > \frac{2^{j-1}k_j}{1 - (d-1)\varpi k_j} \qquad (2 \leq j \leq d) \tag{4.8}$$

are satisfied.

To conclude the section, we record the bound

$$\mathrm{Card}\{\mathbf{y} \in \mathcal{V}^*_{2,\rho}(\mathbb{Z}) : |\mathbf{y}| \leq Y\} \leq \mathrm{Card}\{\mathbf{y} \in \mathcal{B}(\rho) \cap \mathbb{Z}^n : |\mathbf{y}| \leq Y\} \ll Y^{n-\rho}, \tag{4.9}$$

which follows from the argument of [20, Lemma 2] via Theorem 3.1 in [13].

5 Major and minor arcs

Lemma 4.2 is designed to inductively define a partition into major and minor arcs for the entries α_j of $\boldsymbol{\alpha}$ as j runs from d to 2. The size of the major arcs obtained in this way is controlled by the parameters θ_j and k_j which it is now our job to choose optimally. Throughout this section and the next we will assume that $\mathbf{y} \in \mathcal{A}(\rho)$ for some parameter ρ. Also, we will work on the assumption that (4.8) is satisfied, so that the singular case in Lemma 4.2 is excluded.

Given an index j and parameters $\theta_j, \dots, \theta_d \in (0, 1]$, we define the major arcs $\mathfrak{M}_{\mathbf{y}}(X; \theta_j, \dots, \theta_d)$ to be the set of all $\boldsymbol{\alpha} \in [0, 1)^{d-1}$ for which there exist integers $q_j, \dots, q_d$ and $a_j, \dots, a_d$ having the property that for all $i \in \{j, j+1, \dots, d\}$ one has

$$\begin{aligned} 1 \leq q_i \leq c_j X^{\nu_i} |\mathbf{y}|^{d-i} \mu_{\max}, \qquad 0 \leq a_i \leq Q_i, \\ |\alpha_i Q_i - a_i| \leq c_j X^{-i+\omega_i} |\mathbf{y}|^{D_{d-i}} \mu_{\max}^{d-i+1} \end{aligned} \tag{5.1}$$

for some suitable constant c_j. Here, we implicitly used the notation (3.2) and (3.3). Let

$$\mathfrak{m}_{\mathbf{y}}(X;\theta_j,\dots,\theta_d)=[0,1)^{d-1}\backslash\mathfrak{M}_{\mathbf{y}}(X;\theta_j,\dots,\theta_d)$$

be the corresponding minor arcs. One checks that the major arcs are disjoint as soon as X is sufficiently large and

$$\omega_j<j/2-\psi(D_{d-j}+(d-1)(d-j+1)).$$

The definition of the major arcs as given above is iterative in nature in that the approximation of α_j involves the denominators q_i for all $i>j$, and this reflects the fact that our work of the previous section generates an approximation for α_j only in the case when all α_i with $i>j$ have already been approximated. In a sense, therefore, the major arcs $\mathfrak{M}_{\mathbf{y}}(X;\theta_j,\dots,\theta_d)$ are only defined inside the set $\mathfrak{M}_{\mathbf{y}}(X;\theta_{j+1},\dots,\theta_d)$.

At this point, we observe that the size of $T_{\mathbf{y}}(\boldsymbol{\alpha})$ is well defined for any particular $\boldsymbol{\alpha}$. In the light of Lemma 4.2, this means that we lose nothing by making the choice

$$k_j\theta_j=k_i\theta_i\quad(2\le i,j\le d).\tag{5.2}$$

With this assumption, as a consequence of our nested definition of the major arcs we have

$$|T_{\mathbf{y}}(\boldsymbol{\alpha})|\ll\frac{X^s}{d(\Lambda_{\mathbf{y}})}X^{-k_j\theta_j+\varepsilon}\quad\text{whenever }\boldsymbol{\alpha}\in\mathfrak{m}_{\mathbf{y}}(X,\theta_j,\dots,\theta_d),$$

where the ε absorbs any possible dependence on the constants c_j. As the convention (5.2) renders much of the information in our above notation superfluous, we put

$$\mathfrak{M}_{\mathbf{y}}^{(j)}(X;\theta_j)=\mathfrak{M}_{\mathbf{y}}(X;\theta_j,(k_j/k_{j+1})\theta_j,\dots,(k_j/k_d)\theta_j),$$

and we adopt an analogous convention for the minor arcs.

It is useful to make the definition

$$\Omega_j=\sum_{i=j}^{d}\omega_i=\sum_{i=j}^{d}(i-j+1)\nu_i\qquad(2\le j\le d).$$

Write further

$$\sigma_j=\sum_{i=j}^{d}\frac{(i-1)}{k_i}\quad\text{and}\quad\Sigma_j=\sum_{i=j}^{d}\sigma_j=\sum_{i=j}^{d}\frac{(i-j+1)(i-1)}{k_i},\tag{5.3}$$

then (5.2) implies that

$$\omega_j=\sigma_jk_j\theta_j\quad\text{and}\quad\Omega_j=\Sigma_jk_j\theta_j.\tag{5.4}$$

When there is no danger of confusion, we will employ the convention that

$$\Sigma = \Sigma_2, \qquad \sigma = \sigma_2, \qquad \omega = \omega_2.$$

Also, define

$$\Delta_j = \sum_{i=j}^{d} D_{d-i} = \tfrac{1}{6}(d-j)(d-j+1)(d-j+2),$$

noting that $\Delta_d = 0$ and

$$\Delta_j \le \Delta_2 = \tfrac{1}{6}d(d^2-3d+2) \quad \text{for all } j.$$

We then have the following simple lemma.

Lemma 5.1 *For any j with $2 \le j \le d$ the volume of the multi-dimensional major arcs is bounded by*

$$\operatorname{vol} \mathfrak{M}_{\mathbf{y}}(X; \theta_j, \dots, \theta_d) \ll X^{-(D-D_{j-1})+\Omega_j+\omega_j} |\mathbf{y}|^{\Delta_j + D_{d-j}} \mu_{\max}^{D_{d-j+2}-1}.$$

Proof Recall the notation (3.2) and (3.3). The condition (5.1) implies that

$$\begin{aligned} &\operatorname{vol} \mathfrak{M}_{\mathbf{y}}(X; \theta_j, \dots, \theta_d) \\ &\ll \sum_{q_j=1}^{c_j X^{\nu_j} |\mathbf{y}|^{d-j} \mu_{\max}} \sum_{a_j=0}^{Q_j} \left(\frac{X^{-j+\omega_j} |\mathbf{y}|^{D_{d-j}} \mu_{\max}^{d-j+1}}{Q_j} \right) \times \dots \\ &\quad \times \sum_{q_d=1}^{c_j X^{\nu_d} \mu_{\max}} \sum_{a_d=0}^{Q_d} \left(\frac{X^{-d+\omega_d} \mu_{\max}}{Q_d} \right). \end{aligned}$$

We can perform the summations over all a_j. After that, the sums disentangle, and we obtain

$$\begin{aligned} \operatorname{vol} \mathfrak{M}_{\mathbf{y}}(X; \theta_j, \dots, \theta_d) &\ll \prod_{i=j}^{d} \left(\sum_{q_i=1}^{c_i X^{\nu_i} |\mathbf{y}|^{d-i} \mu_{\max}} X^{-i+\omega_i} |\mathbf{y}|^{D_{d-i}} \mu_{\max}^{d-i+1} \right) \\ &\ll \prod_{i=j}^{d} X^{-i+\omega_i+\nu_i} |\mathbf{y}|^{D_{d-i}+d-i} \mu_{\max}^{d-i+2} \\ &\ll X^{-(D-D_{j-1})+\Omega_j+\omega_j} |\mathbf{y}|^{\Delta_j+D_{d-j}} \mu_{\max}^{D_{d-j+2}-1} \end{aligned}$$

as claimed. □

Our next task is to analyse under which conditions the contribution of the minor arcs is under control. We first consider the one-dimensional minor arcs $\mathfrak{m}_{\mathbf{y}}^{(d)}(X; \theta_d)$.

Lemma 5.2 *For any choice of positive parameters $\theta_d \in (0, 1]$, k_d and δ_d suppose that*

$$k_d > D - 1 + \delta_d \tag{5.5}$$

and

$$(1 - \Sigma_d - \sigma_d)k_d\theta_d > D_{d-1} - 1 + \delta_d. \tag{5.6}$$

Then for some $\nu > 0$ we have the bound

$$\int_{\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d)} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s-(D-1)-\delta_d-\nu}\mu_{\max}.$$

Proof Let θ_d be given. We can find a sequence $\theta_d^{(i)}$ with the property

$$1 = \theta_d^{(0)} > \theta_d^{(1)} > \ldots > \theta_d^{(M)} = \theta_d > 0$$

and subject to the condition

$$\big(\theta_d^{(i-1)} - \theta_d^{(i)}\big)k_d < (1 - \Sigma_d - \sigma_d)k_d\theta_d - (D_{d-1} - 1) - \delta_d \qquad (1 \le i \le M). \tag{5.7}$$

Thanks to (5.6), this is always possible with $M = O(1)$. We now infer from Lemma 4.2 and (5.5) that

$$\int_{\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \ll \sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})| \ll \frac{X^s}{d(\Lambda_{\mathbf{y}})}X^{-k_d+\varepsilon} \ll X^{s-(D-1)-\delta_d-\nu},$$

provided that ν is small enough in terms of the other parameters. Further, if we write

$$\mathfrak{m}_{\mathbf{y},i}^{(d)} = \mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d^{(i)}) \cap \mathfrak{M}_{\mathbf{y}}^{(d)}(X;\theta_d^{(i-1)}) \qquad (1 \le i \le M),$$

one obtains via Lemma 5.1, (5.4) and Lemma 4.2 that

$$\begin{aligned}\int_{\mathfrak{m}_{\mathbf{y},i}^{(d)}} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} &\ll \operatorname{vol}\mathfrak{M}_{\mathbf{y}}^{(d)}(X;\theta_d^{(i-1)}) \sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d^{(i)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})| \\ &\ll \frac{X^s}{d(\Lambda_{\mathbf{y}})}X^{-(D-D_{d-1})+(\Sigma_d+\sigma_d)k_d\theta_d^{(i-1)}-k_d\theta_d^{(i)}+\varepsilon}\mu_{\max}^2,\end{aligned}$$

and (5.7) ensures that in the exponent one has for every $i = 1, \ldots, M$ the relation

$$\begin{aligned}-k_d\theta_d^{(i)} + (\Sigma_d + \sigma_d)k_d\theta_d^{(i-1)} &\le \big(\theta_d^{(i-1)} - \theta_d^{(i)}\big)k_d - \big(1 - (\Sigma_d + \sigma_d)\big)k_d\theta_d \\ &< -(D_{d-1} - 1) - \delta_d - \nu\end{aligned}$$

for some sufficiently small $\nu > 0$. Since $\mu_{\max} \ll d(\Lambda_{\mathbf{y}})$ and

$$\int_{\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d)} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} = \int_{\mathfrak{m}_{\mathbf{y}}^{(d)}(X;\theta_d^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} + \sum_{i=1}^{M} \int_{\mathfrak{m}_{\mathbf{y},i}^{(d)}} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha}$$

with $M = O(1)$, this completes the proof. □

We now employ an iterative argument in order to control the contribution from the nested sets of minor arcs. Fix some j in the range $2 \le j \le d-1$, and suppose that the contribution arising from the sets $\mathfrak{m}_{\mathbf{y}}^{(i)}(X;\theta_i)$ is already bounded for all $i > j$ and some suitable parameter θ_{j+1}^*, where the θ_i with $i > j+1$ are determined by θ_{j+1}^* via (5.2).

Lemma 5.3 *Fix an index j with $2 \le j \le d-1$. Suppose that the parameters k_i with $j+1 \le i \le d$ as well as θ_{j+1}^* are given in accordance with (3.7). For some $\delta_{j+1} \ge 0$ assume that*

$$(1 - \Sigma_{j+1} - \sigma_{j+1})k_{j+1}\theta_{j+1}^* > D_j - 1 + \delta_{j+1}. \tag{5.8}$$

Furthermore, for non-negative parameters δ_j and k_j suppose that θ_j satisfies (4.3) as well as the inequalities

$$0 < \theta_j < \theta_j^{(0)} = \frac{k_{j+1}}{k_j}\theta_{j+1}^* \tag{5.9}$$

and

$$(1 - (\Sigma_j + \sigma_j))k_j\theta_j > D_{j-1} - 1 + \delta_j. \tag{5.10}$$

Then the j-th minor arcs contribution is bounded by

$$\int_{\mathfrak{m}^{(j)}(X;\theta_j)} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s-(D-1)} \sum_{i=j}^{d} X^{-\delta_i-\nu}|\mathbf{y}|^{\Delta_i+D_{d-i}}\mu_{\max}^{D_{d-i+2}-2},$$

where ν is some suitably small real number.

Proof Observe first that with our notation in (5.9) we have the decomposition

$$\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)}) = \mathfrak{m}_{\mathbf{y}}^{(j+1)}(X;\theta_{j+1}^*) \cup \left(\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)}) \cap \mathfrak{M}_{\mathbf{y}}^{(j+1)}(X;\theta_{j+1}^*)\right).$$

Suppose that the lemma has been established for j replaced by $j+1$, and recall (2.5) and (5.4). We infer from the inductive hypothesis and Lemmata 5.1 and 4.2 that

$$\begin{aligned}
&\int_{\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \\
&\ll \int_{\mathfrak{m}_{\mathbf{y}}^{(j+1)}(X;\theta_{j+1}^*)} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} + \operatorname{vol}\mathfrak{M}_{\mathbf{y}}^{(j+1)}(X;\theta_{j+1}^*) \sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})| \\
&\ll \sum_{i=j+1}^{d} X^{s-(D-1)-\delta_i-\nu}|\mathbf{y}|^{\Delta_i+D_{d-i}}\mu_{\max}^{D_{d-i+2}-2} \\
&\quad + \frac{X^s}{d(\Lambda_{\mathbf{y}})} X^{-(D-D_j)+(\Sigma_{j+1}+\sigma_{j+1})k_{j+1}\theta_{j+1}^*-k_j\theta_j^{(0)}+\varepsilon}|\mathbf{y}|^{\Delta_{j+1}+D_{d-j-1}}\mu_{\max}^{D_{d-j+1}-1}.
\end{aligned}$$

Recall (5.9). Thus the above bound implies via (5.8) and the relation $\mu_{\max} \ll d(\Lambda_{\mathbf{y}})$ that

$$\int_{\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} \ll X^{s-(D-1)} \sum_{i=j+1}^{d} X^{-\delta_i-\nu}|\mathbf{y}|^{\Delta_i+D_{d-i}}\mu_{\max}^{D_{d-i+2}-2},$$

provided that ν is small enough in terms of the other parameters.

Let now θ_j be given according to (5.9) and (5.10). We can find a sequence $\theta_j^{(i)}$ satisfying

$$\theta_j^{(0)} > \theta_j^{(1)} > \ldots > \theta_j^{(M)} = \theta_j > 0,$$

and subject to the condition

$$\big(\theta_j^{(i-1)} - \theta_j^{(i)}\big)k_j < (1-(\Sigma_j+\sigma_j))k_j\theta_j - (D_{j-1}-1) - \delta_j \qquad (1 \le i \le M). \tag{5.11}$$

This is always possible with $M = O(1)$. For $i \ge 1$ set

$$\mathfrak{m}_{\mathbf{y},i}^{(j)} = \mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(i)}) \cap \mathfrak{M}_{\mathbf{y}}^{(j)}(X;\theta_j^{(i-1)}).$$

Then one deduces from Lemma 5.1, Lemma 4.2, (5.4) and (5.11) that

$$\begin{aligned}
\int_{\mathfrak{m}_{\mathbf{y},i}^{(j)}} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} &\ll \operatorname{vol}\mathfrak{M}_{\mathbf{y}}^{(j)}(X;\theta_j^{(i-1)}) \sup_{\boldsymbol{\alpha}\in\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(i)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})| \\
&\ll \frac{X^s}{d(\Lambda_{\mathbf{y}})} X^{-(D-D_{j-1})+(\Sigma_j+\sigma_j)k_j\theta_j^{(i-1)}-k_j\theta_j^{(i)}+\varepsilon}|\mathbf{y}|^{\Delta_j+D_{d-j}}\mu_{\max}^{D_{d-i+2}-1} \\
&\ll X^{s-(D-1)-\delta_j-\nu}|\mathbf{y}|^{\Delta_j+D_{d-j}}\mu_{\max}^{D_{d-j+2}-2}
\end{aligned}$$

for each $i \geq 1$, and thus altogether

$$\int_{\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j)} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} = \int_{\mathfrak{m}_{\mathbf{y}}^{(j)}(X;\theta_j^{(0)})} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha} + \sum_{i=1}^{M} \int_{\mathfrak{m}_{\mathbf{y},i}^{(j)}} |T_{\mathbf{y}}(\boldsymbol{\alpha})|\,\mathrm{d}\boldsymbol{\alpha}$$
$$\ll MX^{s-(D-1)} \sum_{i=j}^{d} X^{-\delta_i-\nu} |\mathbf{y}|^{\Delta_i+D_{d-i}} \mu_{\max}^{D_{d-i+2}-2}.$$

Since $M = O(1)$, this completes the proof. □

We may now apply first Lemma 5.2 and then Lemma 5.3 successively to each of the θ_j. Thus, for the initial step we need to ensure that the condition (5.5) is satisfied, and after that we have to satisfy the requirements described by (5.10) for all $2 \leq j \leq d$. On the other hand, we have to be careful to ensure that in each iteration we can take θ_{j+1}^* small enough for Lemma 4.2 to be applicable within Lemma 5.3. The crucial requirement here is for the bound (4.3) to be satisfied for $\theta_j^{(0)}$ for all j with $2 \leq j \leq d-1$. Using (5.4) and our convention (5.2), the bound of (4.3) can be re-written in the form

$$(\sigma_j + 1/k_{j-1})k_j\theta_j < 1 - (D_{d-j+1} + (d-1)(d-j+1))\psi \qquad (3 \leq j \leq d). \tag{5.12}$$

For $3 \leq j \leq d$, the condition (5.12) is compatible with the hypotheses (5.6) and (5.10) of Lemmata 5.2 and 5.3, respectively, only if

$$\left(\sigma_j + k_{j-1}^{-1}\right) \frac{D_{j-1} - 1 + \delta_j}{1 - (D_{d-j+1} + (d-1)(d-j+1))\psi} + \Sigma_j + \sigma_j < 1. \tag{5.13}$$

At the same time, a comparison of (5.12) with (4.5) shows that we also require

$$\sigma_j + \frac{1}{k_{j-1}} < \frac{(1 - (D_{d-j+1} + (d-1)(d-j+1))\psi)\varpi}{\psi}. \tag{5.14}$$

Meanwhile, when $j = 2$, the bound of (5.12) does not apply, and we only have the constraints stemming from (4.5) and (5.10), which can be rewritten as

$$k_2\theta_2 > \max\left\{\frac{\delta_2}{1 - (\Sigma_2 + \sigma_2)}, \psi\varpi^{-1}\right\}. \tag{5.15}$$

We will attend to this bound later, but in the meanwhile we remark that regardless of the specific values $\theta_2 > 0$ and $\delta_2 \geq 0$, it implies that we must have $\Sigma_2 + \sigma_2 < 1$. Summarising, we obtain the following intermediate result.

Proposition 5.4 *Assume* (4.8). *Suppose that* (5.13) *and* (5.14) *are satisfied for all* $j \geq 3$, *and that furthermore* (5.5) *and* (5.15) *hold. Then for some* $\nu > 0$ *we have*

$$N_{\mathbf{y}}(X) = \int_{\mathfrak{M}_{\mathbf{y}}^{(2)}(X;\theta)} T_{\mathbf{y}}(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} + O\left(X^{s-(D-1)-\nu}\sum_{j=2}^{d} X^{-\delta_j}|\mathbf{y}|^{\Delta_j+D_{d-j}}\mu_{\max}^{D_{d-j+2}-2}\right). \tag{5.16}$$

Proof This follows from (2.4) upon applying Lemmata 5.2 and 5.3, and the discussion preceding the statement of the proposition. □

6 Understanding the main term

In order to show that the main term of (5.16) is indeed of the expected shape, it is necessary for the approximations of all components of $\boldsymbol{\alpha}$ to have the same denominator. Recall that we wrote $\omega = \omega_2$, and set

$$q = Q_2 \quad \text{and} \quad b_j = (q/Q_j)a_j \qquad (2 \leq j \leq d).$$

For some positive constant c set $W = cX^{\omega}|\mathbf{y}|^{D_{d-2}+(d-1)^2}$, where ω is as obtained in Proposition 5.4. Our final set $\mathfrak{P}_{\mathbf{y}}(X;\omega)$ of major arcs is now the set of all $\boldsymbol{\alpha}$ with an approximation of the shape

$$1 \leq q \leq W \quad \text{and} \quad |\alpha_j - b_j/q| \leq X^{-j}W \quad (2 \leq j \leq d). \tag{6.1}$$

Recall (2.5). When c is sufficiently large, the set $\mathfrak{P}_{\mathbf{y}}(X;\omega)$ is slightly larger than $\mathfrak{M}_{\mathbf{y}}^{(2)}(X;\theta)$, so the corresponding minor arcs $\mathfrak{p}_{\mathbf{y}}(X;\omega) = [0,1)^{d-1}\setminus\mathfrak{P}_{\mathbf{y}}(X;\omega)$ are contained in $\mathfrak{m}_{\mathbf{y}}^{(2)}(X;\theta)$. In the statement of Proposition 5.4, we may therefore replace the major arcs $\mathfrak{M}_{\mathbf{y}}^{(2)}(X;\theta)$ by the larger set $\mathfrak{P}_{\mathbf{y}}(X;\omega)$.

Let $\mathcal{L}_{\mathbf{y}}$ denote the s-dimensional subspace of $\mathbb{R}^n$ containing $\Lambda_{\mathbf{y}}$. Furthermore, we define $\mathcal{L}_{\mathbf{y}}(X) = \mathcal{L}_{\mathbf{y}} \cap [-X, X]^n$, and we let $\Lambda_{\mathbf{y}}(q)$ denote the set of residue classes modulo q of lattice points $\mathbf{x} \in \Lambda_{\mathbf{y}}$. Also, set

$$\vartheta_{\mathbf{y}}(\boldsymbol{\alpha};\mathbf{x}) = \sum_{j=2}^{d} \alpha_j \Phi_{\mathbf{y}}^{(j)}(\mathbf{x})$$

for the analogue of $\phi_{\mathbf{y}}$ in terms of the original variables $\mathbf{x} \in \Lambda_{\mathbf{y}}$. In this notation, we can now define

$$S_{\mathbf{y}}(q,\mathbf{a}) = \sum_{\mathbf{x}\in\Lambda_{\mathbf{y}}(q)} e(\vartheta_{\mathbf{y}}(\mathbf{a}/q;\mathbf{x})) \quad \text{and} \quad v_{\mathbf{y}}(\boldsymbol{\beta},X) = \int_{\mathcal{L}_{\mathbf{y}}(X)} e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi}. \tag{6.2}$$

These functions allow us to approximate the exponential sum $T_{\mathbf{y}}(\boldsymbol{\alpha})$ on the major arcs.

Lemma 6.1 *Suppose that* $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$ *with* $q \le X^{1-\psi(d-1)}$. *We have*

$$\left| T_{\mathbf{y}}(\boldsymbol{\alpha}) - \frac{S_{\mathbf{y}}(q,\mathbf{a})}{q^s}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},X)}{d(\Lambda_{\mathbf{y}})}\right| \ll X^{s-1}q\left(1+\frac{1}{d(\Lambda_{\mathbf{y}})}\sum_{j=2}^{d}|\beta_j|X^j|\mathbf{y}|^{d-j}\right). \quad (6.3)$$

Proof This is essentially standard, but due to our specific setting over a lattice we prefer to provide a full proof. Sorting the terms into arithmetic progressions modulo q, we find that

$$T_{\mathbf{y}}(\boldsymbol{\alpha}) = \sum_{\mathbf{z}\in\Lambda_{\mathbf{y}}(q)} e(\vartheta_{\mathbf{y}}(\mathbf{a}/q;\mathbf{z})) \sum_{\substack{\mathbf{w}\in\Lambda_{\mathbf{y}}\\ q\mathbf{w}+\mathbf{z}\in\mathfrak{A}_{\mathbf{y}}(X)}} e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};q\mathbf{w}+\mathbf{z})),$$

and hence

$$\left| T_{\mathbf{y}}(\boldsymbol{\alpha}) - \frac{S_{\mathbf{y}}(q,\mathbf{a})}{q^s}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},X)}{d(\Lambda_{\mathbf{y}})}\right| \ll \sum_{\mathbf{z}\in\Lambda_{\mathbf{y}}(q)} e(\vartheta_{\mathbf{y}}(\mathbf{a}/q;\mathbf{z}))H(q,\mathbf{z},\boldsymbol{\beta}),$$

where

$$H(q,\mathbf{z},\boldsymbol{\beta}) = \sum_{\substack{\mathbf{w}\in\Lambda_{\mathbf{y}}\\ q\mathbf{w}+\mathbf{z}\in\mathfrak{A}_{\mathbf{y}}(X)}} e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};q\mathbf{w}+\mathbf{z})) - \frac{1}{q^s d(\Lambda_{\mathbf{y}})}\int_{\boldsymbol{\xi}\in\mathcal{L}_{\mathbf{y}}(X)} e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi}.$$

Denote the fundamental domain of $\Lambda_{\mathbf{y}}$ by $\mathcal{F}$, and for $\mathbf{w}\in\Lambda_{\mathbf{y}}$ write $\mathcal{F}(\mathbf{w}) = \mathbf{w}+\mathcal{F}$ for the fundamental domain located at $\mathbf{w}$. Moreover, we write $\mathcal{F}_{q,\mathbf{z}}(\mathbf{w}) = q(\mathbf{w}+\mathcal{F})+\mathbf{z}$ for the domain, stretched by a factor q, that is located at $q\mathbf{w}+\mathbf{z}$. We want to replace $H(q,\mathbf{z},\boldsymbol{\beta})$ by the related quantity

$$H^*(q,\mathbf{z},\boldsymbol{\beta}) = \sum_{\substack{\mathbf{w}\in\Lambda_{\mathbf{y}}\\ q\mathbf{w}+\mathbf{z}\in\mathfrak{A}_{\mathbf{y}}(X)}} \left\{ e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};q\mathbf{w}+\mathbf{z})) - \frac{1}{q^s d(\Lambda_{\mathbf{y}})}\int_{\mathcal{F}_{q,\mathbf{z}}(\mathbf{w})} e(\vartheta_{\mathbf{y}}(\boldsymbol{\alpha};\boldsymbol{\xi}))\,\mathrm{d}\boldsymbol{\xi}\right\}.$$

Clearly, we have $\operatorname{vol}\mathcal{F} = d(\Lambda_{\mathbf{y}})$ and $\operatorname{vol}\mathcal{F}_{q,\mathbf{z}}(\mathbf{w}) = q^s d(\Lambda_{\mathbf{y}})$. Thus, $\mathcal{L}_{\mathbf{y}}(X)$ may be covered by $O(X^s/(q^s d(\Lambda_{\mathbf{y}})))$ domains $\mathcal{F}_{q,\mathbf{z}}(\mathbf{w})$ as $\mathbf{w}$ varies over $\Lambda_{\mathbf{y}}$, and the boundary intersects at most roughly $(X/q)^{s-1}\mu_{\max}/d(\Lambda_{\mathbf{y}}) \ll (X/q)^{s-1}$ of these. Thus, the defect is of size at most $O(X^{s-1}qd(\Lambda_{\mathbf{y}}))$. With this information, we find upon partitioning the integrating domain that $H(q,\mathbf{z},\boldsymbol{\beta}) - H^*(q,\mathbf{z},\boldsymbol{\beta}) \ll (X/q)^{s-1}$, and thus

$$\left| T_{\mathbf{y}}(\boldsymbol{\alpha}) - \frac{S_{\mathbf{y}}(q,\mathbf{a})}{q^s}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},X)}{d(\Lambda_{\mathbf{y}})}\right| \ll \sum_{\mathbf{z}\in\Lambda_{\mathbf{y}}(q)} e(\vartheta_{\mathbf{y}}(\mathbf{a}/q;\mathbf{z}))H^*(q,\mathbf{z},\boldsymbol{\beta}) + O(X^{s-1}q).$$

Rewriting

$$H^*(q,\mathbf{z},\boldsymbol{\beta}) = \sum_{\substack{\mathbf{w}\in\Lambda_{\mathbf{y}}\\ q\mathbf{w}+\mathbf{z}\in\mathfrak{A}_{\mathbf{y}}(X)}} \frac{1}{d(\Lambda_{\mathbf{y}})}\int_{\mathcal{F}(\mathbf{w})} e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};q\mathbf{w}+\mathbf{z})) - e(\vartheta_{\mathbf{y}}(\boldsymbol{\beta};q\boldsymbol{\xi}+\mathbf{z}))\,\mathrm{d}\boldsymbol{\xi}$$

puts us into a position where we can apply the mean value theorem, whereupon we see that

$$\begin{aligned} H^*(q,\mathbf{z},\boldsymbol{\beta}) &\ll \sum_{\substack{\mathbf{w}\in\Lambda_{\mathbf{y}}\\ q\mathbf{w}+\mathbf{z}\in\mathfrak{A}_{\mathbf{y}}(X)}} q\sum_{j=2}^{d}|\beta_j|X^{j-1}|\mathbf{y}|^{d-j}\\ &\ll \left(\frac{X^s}{q^s d(\Lambda_{\mathbf{y}})}+1\right) q\sum_{j=2}^{d}|\beta_j|X^{j-1}|\mathbf{y}|^{d-j}. \end{aligned}$$

The desired bound follows now upon applying the trivial bound $S_{\mathbf{y}}(q,\mathbf{a}) \ll q^s$. □

In particular, when $\boldsymbol{\alpha} \in \mathfrak{P}_{\mathbf{y}}(X;\omega)$, inserting the conditions (6.1) into (6.3) shows that

$$\left| T_{\mathbf{y}}(\boldsymbol{\alpha}) - \frac{S_{\mathbf{y}}(q,\mathbf{a})}{q^s}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},X)}{d(\Lambda_{\mathbf{y}})}\right| \ll X^{s-1}W^2.$$

Since

$$\mathrm{vol}\,\mathfrak{P}_{\mathbf{y}}(X;\omega) \ll \sum_{q=1}^{W}\prod_{j=2}^{d} qX^{-j}W \ll X^{-(D-1)}W^{2d-1},$$

it follows that

$$\begin{aligned} \int_{\mathfrak{P}_{\mathbf{y}}(X;\omega)} T_{\mathbf{y}}(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} &= \sum_{q=1}^{W}\sum_{\substack{\mathbf{a}=0\\(\mathbf{a},q)=1}}^{q-1}\frac{S_{\mathbf{y}}(q,\mathbf{a})}{q^s}\int_{\substack{|\beta_j|\le X^{-j}W\\(2\le j\le d)}}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},X)}{d(\Lambda_{\mathbf{y}})}\,\mathrm{d}\boldsymbol{\beta}\\ &\quad + O\left(X^{s-D}W^{2d+1}\right). \end{aligned} \tag{6.4}$$

As usual, the growth rate of the main term in the asymptotic formula comes from the contribution of $v_{\mathbf{y}}(\boldsymbol{\beta},X)$. Setting $\gamma_j = X^j\beta_j$ for $2 \le j \le d$, the identity

$$v_{\mathbf{y}}(\boldsymbol{\beta},X) = X^s v_{\mathbf{y}}(\boldsymbol{\gamma},1) \tag{6.5}$$

follows from (6.2) by applying integration by parts, and in the same manner one finds further that

$$\int_{\substack{|\beta_j|\le X^{-j}W\\(2\le j\le d)}} v_{\mathbf{y}}(\boldsymbol{\beta},X)\,\mathrm{d}\boldsymbol{\beta} = X^{s-(D-1)}\int_{|\boldsymbol{\beta}|\le W} v_{\mathbf{y}}(\boldsymbol{\beta},1)\,\mathrm{d}\boldsymbol{\beta}.$$

Let

$$\mathfrak{J}_{\mathbf{y}}(W)=\int_{[-W,W]^{d-1}}\frac{v_{\mathbf{y}}(\boldsymbol{\beta},1)}{d(\Lambda_{\mathbf{y}})}\,\mathrm{d}\boldsymbol{\beta}\qquad\text{and}\qquad \mathfrak{S}_{\mathbf{y}}(W)=\sum_{q=1}^{W}q^{-s}\sum_{\substack{\mathbf{a}=0\\(\mathbf{a},q)=1}}^{q-1}S_{\mathbf{y}}(q,\mathbf{a}),$$

then we can rewrite (6.4) in the shape

$$\int_{\mathfrak{P}_{\mathbf{y}}(X;\omega)} T_{\mathbf{y}}(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = X^{s-D+1}\mathfrak{S}_{\mathbf{y}}(W)\mathfrak{J}_{\mathbf{y}}(W)+O\left(X^{s-D}W^{2d+1}\right). \tag{6.6}$$

In order to understand the main term in (6.6), we extend the truncated singular integral $\mathfrak{J}_{\mathbf{y}}(W)$ and the truncated singular series $\mathfrak{S}_{\mathbf{y}}(W)$ to infinity by taking the limits $X\to\infty$ in both expressions. In our analysis of these limits, the notations $\boldsymbol{\beta}_j=(\beta_j,\dots,\beta_d)$ and $\mathbf{a}_j=(a_j,\dots,a_d)$ $(2\le j\le d)$ will prove useful.

We start by considering the singular integral.

Lemma 6.2 *We have*

$$|v_{\mathbf{y}}(\boldsymbol{\beta},1)|\ll \min_{2\le j\le d}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j}(1+|\boldsymbol{\beta}_j|)^{-1/\sigma_j+\varepsilon}.$$

Proof Fix j with $2\le j\le d$. For $|\boldsymbol{\beta}_j|\le 1$ the claim is trivial, so we may assume that $|\boldsymbol{\beta}_j|>1$. Choose $P=|\boldsymbol{\beta}|^A$ for some large parameter A to be fixed later, and write $\boldsymbol{\gamma}=(P^{-2}\beta_2,\dots,P^{-d}\beta_d)$ and $\boldsymbol{\gamma}_j=(\gamma_j,\dots,\gamma_d)$. Recalling (5.4), we fix θ_j such that

$$\max_{j\le i\le d}\frac{|\boldsymbol{\beta}_i|}{c_j P^{\omega_i}|\mathbf{y}|^{D_{d-i}}\mu_{\max}^{d-i+1}}=1,$$

so that

$$P^{-k_j\theta_j}\ll|\boldsymbol{\beta}_j|^{-1/\sigma_j}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j}. \tag{6.7}$$

With this choice, $\boldsymbol{\gamma}_j$ lies in the major arcs $\mathfrak{M}_{\mathbf{y}}^{(j)}(P;\theta_j)$. Clearly, the major arcs are disjoint when A is sufficiently large, so $\boldsymbol{\gamma}_j$ is best approximated by $q=1$ and $\mathbf{a}_j=\mathbf{0}$. We therefore have from Lemma 6.1 and (6.5) that

$$|v_{\mathbf{y}}(\boldsymbol{\beta},1)|\ll\left(\frac{P^s}{d(\Lambda_{\mathbf{y}})}\right)^{-1}|T_{\mathbf{y}}(\boldsymbol{\gamma};P)|+P^{-1}|\boldsymbol{\beta}|. \tag{6.8}$$

On the other hand, $\boldsymbol{\gamma}$ lies just on the boundary of the major arcs and thus by continuity the minor arcs bound continues to apply. Consequently, we obtain from Lemma 4.2 and (6.7) the complementary estimate

$$|T_{\mathbf{y}}(\boldsymbol{\gamma}; P)| \ll \left(\frac{P^s}{d(\Lambda_{\mathbf{y}})}\right) P^{-k_j\theta_j+\varepsilon} \ll \left(\frac{P^s}{d(\Lambda_{\mathbf{y}})}\right) P^{\varepsilon}|\boldsymbol{\beta}_j|^{-1/\sigma_j}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j}.$$

Inserting this into (6.8) leads to

$$|v_{\mathbf{y}}(\boldsymbol{\beta}, 1)| \ll P^{\varepsilon}|\boldsymbol{\beta}_j|^{-1/\sigma_j+\varepsilon}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j} + P^{-1}|\boldsymbol{\beta}|,$$

and upon recalling that $P = |\boldsymbol{\beta}|^A \geq |\boldsymbol{\beta}_j|^A$, this reproduces the desired estimate whenever A is sufficiently large. □

It follows from Lemma 6.2 that for any tuple $\lambda_2, \ldots, \lambda_d \in [0, 1]$ satisfying the relation $\lambda_2 + \ldots + \lambda_d = 1$ we have

$$\begin{aligned}&\int_{\substack{\boldsymbol{\beta}\in\mathbb{R}^{d-1}\\ |\boldsymbol{\beta}|>W}} \frac{|v_{\mathbf{y}}(\boldsymbol{\beta}, 1)|}{d(\Lambda_{\mathbf{y}})}\,\mathrm{d}\boldsymbol{\beta}\\ &\quad\ll \frac{1}{d(\Lambda_{\mathbf{y}})}\int_{\substack{\boldsymbol{\beta}\in\mathbb{R}^{d-1}\\ |\boldsymbol{\beta}|>W}} \prod_{j=2}^{d}\Big(|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j}(1+|\boldsymbol{\beta}|)^{-1/\sigma_j+\varepsilon}\Big)^{\lambda_j}\,\mathrm{d}\boldsymbol{\beta}.\end{aligned}$$

The set of all $\boldsymbol{\beta} \in \mathbb{R}^{d-1}$ having $|\boldsymbol{\beta}| = r$ has volume $O(r^{d-2})$. Recalling that we have $\mu_{\max} \ll d(\Lambda_{\mathbf{y}})$, it follows that the above integral is bounded by

$$\int_{\substack{\boldsymbol{\beta}\in\mathbb{R}^{d-1}\\ |\boldsymbol{\beta}|>W}} \frac{|v_{\mathbf{y}}(\boldsymbol{\beta}, 1)|}{d(\Lambda_{\mathbf{y}})}\,\mathrm{d}\boldsymbol{\beta} \ll |\mathbf{y}|^{\kappa_1}\mu_{\max}^{-1+\kappa_2}\int_{r>W}(1+r)^{-\kappa_3+d-2+\varepsilon}\,\mathrm{d}r,$$

where

$$\kappa_1 = \sum_{j=2}^{d}\frac{D_{d-j}\lambda_j}{\sigma_j}, \qquad \kappa_2 = \sum_{j=2}^{d}\frac{(d-j+1)\lambda_j}{\sigma_j}, \qquad \kappa_3 = \sum_{j=2}^{d}\frac{\lambda_j}{\sigma_j}.$$

The integral converges if we can pick $\lambda_2, \ldots, \lambda_d$ in such a way that $\kappa_3 > d - 1$. We take $\lambda_j = \sigma_j$ for $j \geq 3$, so that $\lambda_2 = 1 - \Sigma_3 = \sigma_2 + (1 - \Sigma_2)$. With this choice, the desired inequality $\kappa_3 > d - 1$ is satisfied if $\Sigma < 1$, and we have

$$\kappa_3 = d - 1 + \frac{1-\Sigma_2}{\sigma_2} = d + \frac{1-\Sigma-\sigma}{\sigma}.$$

Moreover, using these values in our expression for κ_1 and κ_2 we obtain

$$\kappa_1 = \Delta_2 + D_{d-2} + D_{d-2}\frac{1-\sigma-\Sigma}{\sigma} \qquad \text{and} \qquad \kappa_2 = D - 1 + (d-1)\frac{1-\sigma-\Sigma}{\sigma}.$$

Upon referring to (2.5), this allows us to conclude that

$$\begin{aligned}\mathfrak{J}_{\mathbf{y}} - \mathfrak{J}_{\mathbf{y}}(W) &\ll |\mathbf{y}|^{\Delta_2+D_{d-2}+(d-1)(D-2)+(D_{d-2}+(d-1)^2)\frac{1-\sigma-\Sigma}{\sigma}} W^{-1-\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon} \\ &\ll |\mathbf{y}|^{\frac{1}{3}(2d^3-11d+9)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}} W^{-1-\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon}, \end{aligned} \tag{6.9}$$

and we have the bound

$$\mathfrak{J}_{\mathbf{y}}(W) \ll |\mathbf{y}|^{\frac{1}{3}(2d^3-11d+9)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}} \tag{6.10}$$

uniformly in W.

The next step is to complete the truncated singular series.

Lemma 6.3 *The terms of the singular series are bounded by*

$$|q^{-s} S_{\mathbf{y}}(q, \mathbf{a})| \ll \min_{2\le j\le d} q^{\varepsilon} \left(\frac{q}{(q, \mathbf{a}_j)}\right)^{-1/\sigma_j} |\mathbf{y}|^{D_{d-j}/\sigma_j} \mu_{\max}^{(d-j+1)/\sigma_j}.$$

Proof For $q = 1$ the estimate is trivial, so we may suppose that $q > 1$. Fix $P = q^A$ for some large A to be determined later. For any j with $2 \le j \le d$ fix θ_j such that

$$\max_{j\le i\le d} \frac{q/(q, \mathbf{a}_i)}{c_j^{d-i} P^{\omega_i} |\mathbf{y}|^{D_{d-i}} \mu_{\max}^{d-i+1}} = 1,$$

so that in particular

$$P^{-k_j\theta_j} \ll \left(\frac{q}{(q, \mathbf{a}_j)}\right)^{-1/\sigma_j} |\mathbf{y}|^{D_{d-j}/\sigma_j} \mu_{\max}^{(d-j+1)/\sigma_j} \tag{6.11}$$

and $\mathbf{a}_j/q \in \mathfrak{M}_{\mathbf{y}}^{(j)}(P; \theta_j)$. Note that by taking A sufficiently large we may ensure that the major arcs $\mathfrak{M}_{\mathbf{y}}^{(j)}(P; \theta_j)$ are disjoint, so $\mathbf{a}_j/q$ is best approximated by itself. Applying Lemma 6.1 and (6.5) with $\boldsymbol{\beta} = \mathbf{0}$ and observing that $v_{\mathbf{y}}(\mathbf{0}, 1) \asymp 1$, it follows that

$$q^{-s} S_{\mathbf{y}}(q, \mathbf{a}) \ll \left(\frac{P^s}{d(\Lambda_{\mathbf{y}})}\right)^{-1} |T_{\mathbf{y}}(q^{-1}\mathbf{a}; P)| + P^{-1}q. \tag{6.12}$$

At the same time, $\mathbf{a}_j/q$ can be viewed as lying just on the boundary of the major arcs in the q-aspect. As before, this implies that Lemma 4.2 and (6.11) furnish the additional minor arcs bound

$$\begin{aligned}|T_{\mathbf{y}}(q^{-1}\mathbf{a}; P)| &\ll d(\Lambda_{\mathbf{y}})^{-1} P^{s-k_j\theta_j+\varepsilon} \\ &\ll \frac{P^{s+\varepsilon}}{d(\Lambda_{\mathbf{y}})} \left(\frac{q}{(q, \mathbf{a}_j)}\right)^{-1/\sigma_j} |\mathbf{y}|^{D_{d-j}/\sigma_j} \mu_{\max}^{(d-j+1)/\sigma_j}, \end{aligned}$$

and on substituting this into (6.12) we discern that

$$q^{-s}S_{\mathbf{y}}(q,\mathbf{a}) \ll P^{\varepsilon}\left(\frac{q}{(q,\mathbf{a}_j)}\right)^{-1/\sigma_j}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j} + P^{-1}q \qquad (2 \le j \le d).$$

Recalling that $P = q^A$, it is clear that for A sufficiently large the first term dominates. □

Lemma 6.3 implies that the singular series may be extended to infinity. Let $\tau_2, \dots, \tau_d$ be natural numbers with the property that $\tau_j | \tau_{j+1}$ for $2 \le j \le d-1$ and $\tau_d | q$. For any j the number of choices of $\mathbf{a} \pmod q$ satisfying $(q, \mathbf{a}_j) = \tau_j$ is $O(q^{d-1}/\tau_j^{d-j+1})$. It thus follows that we have

$$\begin{aligned}
&\sum_{q=1}^{W}\sum_{\substack{\mathbf{a}=0\\(\mathbf{a},q)=1}}^{q-1} q^{-s}|S_{\mathbf{y}}(q,\mathbf{a})| \\
&\ll \sum_{q=1}^{W}\sum_{\tau_2|\dots|\tau_d|q}\min_{2\le j\le d} q^{j-2+\varepsilon}\left(\frac{q}{\tau_j}\right)^{d-j+1-1/\sigma_j}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j} \\
&\ll \sum_{q=1}^{W} q^{d-1+\varepsilon}\prod_{j=2}^{d}\left(q^{-1/\sigma_j}|\mathbf{y}|^{D_{d-j}/\sigma_j}\mu_{\max}^{(d-j+1)/\sigma_j}\right)^{\lambda_j}
\end{aligned}$$

for any choice of $\lambda_2, \dots, \lambda_d \in [0,1]$ with $\lambda_2 + \dots + \lambda_d = 1$. Just like in the treatment of the singular integral, we can take $\lambda_j = \sigma_j$ for $3 \le j \le d$, and $\lambda_2 = 1 - \Sigma_3$. This choice yields the bound

$$\begin{aligned}
\mathfrak{S}_{\mathbf{y}} - \mathfrak{S}_{\mathbf{y}}(W) &\ll |\mathbf{y}|^{\Delta_2 + D_{d-2} + (d-1)(D-1) + \frac{1-\sigma-\Sigma}{\sigma}(D_{d-2}+(d-1)^2)}\sum_{q\ge W} q^{-1-\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon} \\
&\ll |\mathbf{y}|^{\frac{2}{3}(d^3-4d+3)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}} W^{-\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon} \qquad (6.13)
\end{aligned}$$

whenever we have $\Sigma + \sigma < 1$. Again, we recall that this last inequality is satisfied as a consequence of the more stringent condition (5.15). In particular, we have the bound

$$\mathfrak{S}_{\mathbf{y}}(W) \ll |\mathbf{y}|^{\frac{2}{3}(d^3-4d+3)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}}, \qquad (6.14)$$

which holds uniformly in W.

We can now complete the singular series and integral. Here, from (6.9), (6.10), (6.13) and (6.14) and upon inserting our value $W = cX^{\omega}|\mathbf{y}|^{D_{d-2}+(d-1)^2}$, we find that

$$\begin{aligned}|\mathfrak{J}_{\mathbf{y}}\mathfrak{S}_{\mathbf{y}} - \mathfrak{J}_{\mathbf{y}}(W)\mathfrak{S}_{\mathbf{y}}(W)| &\ll |\mathbf{y}|^{\frac{1}{3}(4d^3-19d+15)+(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}} W^{-\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon} \\ &\ll X^{-\omega\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon}|\mathbf{y}|^{\frac{1}{3}(4d^3-19d+15)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}}.\end{aligned} \tag{6.15}$$

It remains to collect our estimates.

Proposition 6.4 *Make the assumption* (3.4) *and suppose that the conditions* (5.5), (5.13), (5.14) *and* (5.15) *are satisfied. Moreover, assume* (4.8). *In this case we have the asymptotic formula*

$$N_{\mathbf{y}}(X) = X^{n-D}\left(\mathfrak{S}_{\mathbf{y}}\mathfrak{J}_{\mathbf{y}} + O(E(\mathbf{y},\theta))\right),$$

where

$$\begin{aligned}E(\mathbf{y},\theta) = &\sum_{j=2}^{d} X^{-\delta_j-\nu}|\mathbf{y}|^{\Delta_j+D_{d-j}+(D_{d-j+1}+d-j)(d-1)} \\ &+ X^{-1+(2d+1)\omega}|\mathbf{y}|^{\frac{1}{2}(6d^3-11d^2+d+4)} \\ &+ X^{-\omega\frac{1-\sigma-\Sigma}{\sigma}+\varepsilon}|\mathbf{y}|^{\frac{1}{3}(4d^3-19d+15)+\frac{1}{2}(3d^2-7d+4)\frac{1-\sigma-\Sigma}{\sigma}}.\end{aligned} \tag{6.16}$$

Proof Recall that we had $n = s+1$. The statement now follows from Proposition 5.4 together with (6.6) and (6.15). □

Before concluding the section, we remark that the singular series and integral can be expressed in terms of solution densities of the system (2.2) over the real and p-adic numbers. Indeed, since under the hypotheses of the proposition the singular series is absolutely convergent, by standard arguments it can be written as an absolutely convergent Euler product $\mathfrak{S}_{\mathbf{y}} = \prod_p \chi_p$, where

$$\begin{aligned}\chi_p &= \sum_{h=0}^{\infty} p^{-hs} \sum_{\substack{\mathbf{a}=1 \\ (\mathbf{a},p)=1}}^{p^h} S_{\mathbf{y}}(p^h,\mathbf{a}) \\ &= \lim_{H\to\infty} p^{H(D-1-s)}\#\{\mathbf{x}\in\Lambda_{\mathbf{y}}(p^H) : \Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) \equiv 0 \pmod{p^H} \text{ for } 2\le j\le d\}.\end{aligned}$$

Upon recalling that $\Lambda_{\mathbf{y}}(q)$ denotes the set of all $\mathbf{x}\in(\mathbb{Z}/q\mathbb{Z})^n$ that satisfy the congruence $\Phi_{\mathbf{y}}^{(1)}(\mathbf{x}) \equiv 0 \pmod{q}$, we see that the above can be re-written as

$$\chi_p = \lim_{H\to\infty} p^{H(D-n)}\#\{\mathbf{x}\in(\mathbb{Z}/p^H\mathbb{Z})^n : \Phi_{\mathbf{y}}^{(j)}(\mathbf{x}) \equiv 0 \pmod{p^H} \text{ for } 1\le j\le d\}.$$

Thus, each factor χ_p reflects the solution density of (2.2) in $\mathbb{Q}_p$.

For the singular integral we proceed in a similar manner. Recall that $\Phi_{\mathbf{y}}^{(1)}$ is an invertible linear transformation. Consider the manifold

$$M(h) = \{\boldsymbol{\xi} \in [-1, 1]^n : \Phi_{\mathbf{y}}^{(1)}(\boldsymbol{\xi}) = h\}$$

with associated measure μ, normalised such that $\mu(M(0)) = d(\Lambda_{\mathbf{y}})^{-1}$. Let now

$$g(\boldsymbol{\xi}) = \int_{\mathbb{R}^{d-1}} e\left(\sum_{j=2}^{d} \eta_j \Phi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})\right) \mathrm{d}\boldsymbol{\eta} \qquad \text{and} \qquad f(h) = \int_{M(h)} g(\boldsymbol{\xi})\, \mathrm{d}\mu(\boldsymbol{\xi}),$$

so that $f(0) = \mathfrak{J}_{\mathbf{y}}$. The inverse Fourier transform of f is given by

$$\mathcal{F}^{-1} f(\alpha) = \int_{[-1,1]^n} g(\boldsymbol{\xi}) e(\alpha \Phi_{\mathbf{y}}^{(1)}(\boldsymbol{\xi}))\, \mathrm{d}\boldsymbol{\xi},$$

and upon taking the (regular) Fourier transform it follows from the Fourier inversion formula that

$$f(N) = \int_{\mathbb{R}} \int_{[-1,1]^n} g(\boldsymbol{\xi}) e(\alpha(\Phi_{\mathbf{y}}^{(1)}(\boldsymbol{\xi}) - N))\, \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\alpha.$$

Thus we conclude that

$$\mathfrak{J}_{\mathbf{y}} = f(0) = \int_{[-1,1]^n} \int_{\mathbb{R}^d} e\left(\sum_{j=1}^{d} \eta_j \Phi_{\mathbf{y}}^{(j)}(\boldsymbol{\xi})\right) \mathrm{d}\boldsymbol{\eta}\, \mathrm{d}\boldsymbol{\xi}.$$

One can now show by standard arguments (for instance Lemma 2 and §11 in [25]) that this expression indeed describes the solution density of (2.2) over the real unit hypercube.

7 Endgame

The quantities σ_j and Σ_j can be expressed in terms of s itself. It is a straightforward exercise to confirm the identities

$$\sum_{n=1}^{N} n2^n = 2^{N+1}(N-1) + 2 \qquad \text{and} \qquad \sum_{n=1}^{N} n^2 2^n = 2^{N+1}(N^2 - 2N + 3) - 6. \tag{7.1}$$

Note that (4.8) transforms into

$$\frac{1}{k_j} > \frac{2^{j-1}}{s-\rho} + (d-1)\varpi.$$

Using this within (5.3), an application of (7.1) produces the bounds

$$\sigma_j > \frac{2^d(d-2) - 2^{j-1}(j-3)}{s-\rho} + \varpi(d-1)\frac{d(d-1) - (j-1)(j-2)}{2}$$

and

$$\Sigma_j > \frac{2^d(d^2 - 2d + 2 - j(d-2)) + 2^{j-1}(j-5)}{s-\rho} + \varpi\frac{(d-1)(d-j+1)(d-j+2)(2d+j-3)}{6},$$

which we require to hold for all indices j in our range $2 \le j \le d$. For the sake of simplicity we replace all these bounds by

$$\begin{aligned} &\frac{1}{k_j} > \frac{2^{d-1}}{s-\rho} + (d-1)\varpi, \qquad \sigma_j > \frac{2^d(d-1)}{s-\rho} + \frac{\varpi d(d-1)^2}{2}, \\ &\Sigma_j > \frac{2^d(d^2-4d+6)}{s-\rho} + \frac{\varpi(2d-1)d(d-1)^2}{6}. \end{aligned} \tag{7.2}$$

This allows us to state a first result.

Theorem 7.1 *Let $F \in \mathbb{Z}[x_1, \dots, x_n]$ be a non-singular form of degree $d \ge 5$ defining a hypersurface $\mathcal{V}$. Let further $\psi > 0$ be a parameter satisfying*

$$\psi^{-1} > d^4 + \tfrac{3}{2}d^3 - \tfrac{11}{2}d^2 + d + 2, \tag{7.3}$$

and set

$$n_1(\psi) = \frac{2^{d-1}\left(d^3 + \frac{1}{2}d^2 - \frac{11}{2}d + 10 - \psi p_6(d)\right)}{1 - (d^4 + \frac{3}{2}d^3 - \frac{11}{2}d^2 + d + 2)\psi},$$

where $p_6(d) = \frac{1}{12}(50d^6 - 171d^5 + 88d^4 + 517d^3 - 732d^2 + 8d - 120)$. For some integer $\rho \in [1, n]$ suppose that $n - \rho > n_1(\psi)$. Then there exists a real positive number ν with the property that

$$N_{\mathbf{y}}(X) = X^{n-\frac{1}{2}d(d+1)}\mathfrak{S}_{\mathbf{y}}\mathfrak{J}_{\mathbf{y}} + O(X^{n-\frac{1}{2}d(d+1)-\nu})$$

uniformly for all $\mathbf{y} \in \mathcal{V}_{2,\rho}(\mathbb{Z})$ satisfying $|\mathbf{y}| \le X^{\psi}$, and the factors $\mathfrak{S}_{\mathbf{y}}$ and $\mathfrak{J}_{\mathbf{y}}$ satisfy $0 \le \mathfrak{S}_{\mathbf{y}} \ll_{\mathbf{y}} 1$ and $0 \le \mathfrak{J}_{\mathbf{y}} \ll_{\mathbf{y}} 1$.

Proof Our main task here is to bound the error terms given by (6.16) in the conclusion of Proposition 6.4, while at the same time ensuring that the hypotheses of said proposition are satisfied. In order to control the first term in (6.16) we choose

$$\delta_j = \psi(\Delta_j + D_{d-j} + (D_{d-j+2} - 2)(d-1))$$

for $2 \leq j \leq d$. Thus, we have $\delta_d = (d-1)\psi$. With this choice, and recalling (3.4), the bound in (5.5) is certainly majorised by $k_d > D$. In a similar manner, upon taking into account the uniform bounds (7.2) as well as the relations $D_j \leq D$ and

$$\delta_j \leq \delta_2 = \tfrac{1}{3}(2d^3 - 11d + 9)\psi$$

for all j, a modicum of computation reveals that for all ψ satisfying (7.3) one has

$$\frac{D_{j-1} - 1 + \delta_j}{1 - (D_{d-j+1} + (d-1)(d-j+1))\psi} \leq \frac{d(d-1)}{2},$$

and hence the condition (5.13) may be simplified to

$$\tfrac{1}{2}d(d-1)\left(\sigma_j + k_{j-1}^{-1}\right) + \Sigma_j + \sigma_j < 1. \tag{7.4}$$

Upon inserting (7.2), we see that the conditions (5.5), (5.13) (as simplified to (7.4)) and (5.14) of Proposition 6.4 are satisfied whenever

$$s - \rho > \max\{a_0(\varpi), a_1(\varpi), a_2(\varpi, \psi)\},$$

where

$$\begin{aligned} a_0(\varpi) &= \frac{2^{d-2}d(d+1)}{1 - \frac{1}{2}d(d^2-1)\varpi}, \\ a_1(\varpi) &= \frac{2^{d-2}(2d^3 + d^2 - 11d + 20)}{1 - \frac{1}{12}(d-1)^2 d(3d^2 + d + 10)\varpi}, \\ a_2(\varpi, \psi) &= \frac{2^{d-1}(2d-1)\psi}{\varpi(1 - \frac{1}{2}d(d^2 + d - 2)\psi)}. \end{aligned}$$

For this to be defined, we require in particular that

$$\varpi^{-1} > \tfrac{1}{12}(d-1)^2 d(3d^2 + d + 10), \tag{7.5}$$

which we will assume henceforth.

Meanwhile, to control the second and third term in (6.16) we require that

$$\frac{\frac{1}{3}(4d^3 - 19d + 15)\psi}{1 - \Sigma - \sigma} + \frac{\frac{1}{2}(3d^2 - 7d + 4)\psi}{\sigma} < k_2\theta_2 < \frac{1 - \frac{1}{2}(6d^3 - 11d^2 + d + 4)\psi}{(2d+1)\sigma}, \tag{7.6}$$

while simultaneously the bound (5.15) should be satisfied. Upon re-writing, we see that the interval in (7.6) is non-empty when

$$\left(1 + \frac{\frac{1}{3}(8d^4 + 4d^3 - 38d^2 + 11d + 15)\psi}{1 - (6d^3 - 11d^2 + d + 4)\psi}\right)\sigma + \Sigma < 1. \tag{7.7}$$

When ψ satisfies (7.3) one can show for $d \geq 5$ that

$$\frac{\frac{1}{3}(8d^4 + 4d^3 - 38d^2 + 11d + 15)\psi}{1 - (6d^3 - 11d^2 + d + 4)\psi} \leq 8,$$

and hence (7.7) may be simplified to $9\sigma + \Sigma < 1$. In combination with (7.2) this delivers the bound $s - \rho > b_1(\varpi)$ where

$$b_1(\varpi) = \frac{2^d(d^2 + 5d - 3)}{1 - \frac{1}{3}d(d-1)^2(d+13)\varpi}.$$

In order to handle the bound (5.15) one confirms that $\delta_2/(1 - \sigma - \Sigma)$ is smaller than the first term on the left hand side of (7.6), and hence (5.15) is compatible with the right hand side of (7.6) if the inequality

$$\psi\varpi^{-1} < \frac{1 - \frac{1}{2}(6d^3 - 11d^2 + d + 4)\psi}{(2d+1)\sigma}$$

is satisfied. Re-arranging yields

$$\frac{\psi(2d+1)\varpi^{-1}}{1 - \frac{1}{2}(6d^3 - 11d^2 + d + 4)\psi}\sigma < 1,$$

which upon inserting (7.2) delivers the bound $s - \rho > b_2(\varpi, \psi)$ where

$$b_2(\varpi, \psi) = \frac{2^d(d-1)(2d+1)\psi}{\varpi(1 - (d^4 + \frac{3}{2}d^3 - \frac{11}{2}d^2 + d + 2)\psi)}.$$

Thus, altogether we have shown that the conclusion of the theorem follows if for some suitable value of ϖ one has

$$s - \rho > \max\{a_0(\varpi), a_1(\varpi), a_2(\varpi, \psi), b_1(\varpi), b_2(\varpi, \psi)\}.$$

We see that $b_2(\varpi, \psi) > a_2(\varpi, \psi)$ for all admissible values of ψ and ϖ. In a similar manner, when $d \geq 5$ we have the inequalities $a_1(\varpi) \geq \max\{a_0(\varpi), b_1(\varpi)\}$ for all admissible values of ϖ. One can compute (for instance with the help of a computer algebra programme) that $a_1(\varpi) = b_2(\varpi, \psi)$ when $\varpi = \varpi_0(\psi)$, where

$$\varpi_0(\psi) = \frac{(d-1)(2d+1)\psi}{d^3 + \frac{1}{2}d^2 - \frac{11}{2}d + 10 - \psi\, p_6(d)},$$

where $p_6(d)$ is as in the statement of the theorem. The quantity $\varpi_0(\psi)$ is increasing in ψ, and a final computation confirms that it is admissible within (7.5) for all values of ψ satisfying (7.3). Thus, for any given value of ψ within the admissible range the bound $s - \rho > b_2(\psi, \varpi_0(\psi))$ dominates overall. Setting $n_1(\psi) = b_2(\psi, \varpi_0(\psi))$ concludes the proof of Theorem 7.1.

Theorem 1.3 is a simplification of Theorem 7.1. Indeed, upon choosing $\psi = \psi_1$ with $\psi_1^{-1} = 2d^4$ we find that

$$\begin{aligned} n_1(\psi_1) &= \frac{2^d(24d^7 - 38d^6 + 39d^5 + 152d^4 - 517d^3 + 732d^2 - 8d - 240)}{24d^4 - 36d^3 + 132d^2 - 24d - 48} \\ &< 2^d d(d^2 - 1) \end{aligned}$$

for all admissible values of d. Since the function $n_1(\psi)$ is increasing in ψ, this bound is sufficient for all $\psi < \psi_1$ also. This completes the proof of Theorem 1.3.

In order to obtain an estimate for $N_{\mathcal{U}}(X, X^{\psi})$ and thus complete the proof of Theorems 1.4 and 1.2, we need to sum over all values of $\mathbf{y} \in \mathcal{U}(\mathbb{Z})$ satisfying $|\mathbf{y}| \leq X^{\psi}$ and $F(\mathbf{y}) = 0$.

Theorem 7.2 *Let $F \in \mathbb{Z}[x_1, \ldots, x_n]$ be a non-singular form of degree $d \geq 5$ defining a hypersurface $\mathcal{V}$. Let further $\psi > 0$ be a parameter satisfying*

$$\psi^{-1} > d^4 + \tfrac{3}{2}d^3 - 5d^2 + \tfrac{1}{2}d + 2. \tag{7.8}$$

Set

$$n_2(\psi) = \frac{2^{d-1}\left(d^3 + \frac{1}{2}d^2 - \frac{11}{2}d + 10 - q_6(d)\psi\right)}{1 - (d^4 + \frac{1}{2}d^3 - \frac{5}{2}d^2 - 2d + 2)\psi}$$

where $q_6(d) = \frac{1}{12}(50d^6 - 165d^5 + 85d^4 + 481d^3 - 639d^2 - 52d + 240)$. For some integer ρ in the range $\frac{1}{2}d(d+1) + 1 < \rho < n$ suppose that $n - \rho > n_2(\psi)$. Then there exists a positive real number ν for which we have the asymptotic formula

$$N_{\mathcal{V}_{2,\rho}}(X, Y) = X^{n-D} \sum_{\substack{\mathbf{y} \in \mathcal{V}_{2,\rho}(\mathbb{Z}) \\ |\mathbf{y}| \leq Y}} \mathfrak{S}_{\mathbf{y}} \mathfrak{J}_{\mathbf{y}} + O((XY)^{n-D} X^{-\nu}), \tag{7.9}$$

and the factors satisfy $0 \leq \mathfrak{S}_{\mathbf{y}} \ll_{\mathbf{y}} 1$ and $0 \leq \mathfrak{J}_{\mathbf{y}} \ll_{\mathbf{y}} 1$.

Proof Recall from Birch's theorem [6] that for $n > 2^d(d-1)$ the number of points $\mathbf{z} \in \mathbb{Z}^n$ with $|\mathbf{z}| \leq Z$ and $F(\mathbf{z}) = 0$ is given by $N(Z) \ll Z^{n-d}$. Upon combining (1.2) and Proposition 6.4, we find that

$$N_{\mathcal{U}}(X, X^{\psi}) = X^{n-D} \sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \cap \mathcal{A}(\rho) \\ |\mathbf{y}| \leq X^{\psi} \\ F(\mathbf{y})=0}} \mathfrak{S}_{\mathbf{y}} \mathfrak{J}_{\mathbf{y}} + O\left(E_{\mathcal{A}}(\psi) + E_{\mathcal{B}}(\psi) + E_{\mathcal{U}}(\psi)\right), \tag{7.10}$$

where

$$E_{\mathcal{A}}(\psi) = X^{n-D} \sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \cap \mathcal{A}(\rho) \\ |\mathbf{y}| \leq X^{\psi} \\ F(\mathbf{y})=0}} E(\mathbf{y}, \theta), \qquad E_{\mathcal{B}}(\psi) = \sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \cap \mathcal{B}(\rho) \\ 0 < |\mathbf{y}| \leq X^{\psi} \\ F(\mathbf{y})=0}} \sum_{\substack{\mathbf{x} \in \mathcal{U}(\mathbb{Z}) \\ |\mathbf{x}| \leq X \\ F(\mathbf{x})=0}} 1$$

and

$$E_{\mathcal{U}}(\psi) = \sum_{\substack{|\mathbf{y}| \leq X^{\psi} \\ F(\mathbf{y})=0}} \sum_{\substack{\mathbf{x} \in \mathcal{V}(\mathbb{Z}) \setminus \mathcal{U}(\mathbb{Z}) \\ |\mathbf{x}| \leq X}} 1 \ll X^{\dim \mathcal{V} \setminus \mathcal{U} + \psi(n-d)}.$$

The choice $\mathcal{U} = \mathcal{A}(\rho) \cap \mathcal{V} = \mathcal{V}_{2,\rho}$ entails that $\dim \mathcal{V} \setminus \mathcal{V}_{2,\rho} = \dim \mathcal{V}^*_{2,\rho} \leq n - \rho$, and we conclude that the error $E_{\mathcal{U}}(\psi)$ is acceptable within (7.10) if $\rho > D + \psi(D-d)$. In particular, it follows from (3.4) that the choice $\rho = D + 1$ is permissible. Clearly, with this choice of $\mathcal{U}$ the set $\mathcal{B}(\rho) \cap \mathcal{U}$ is empty and we can disregard the error term $E_{\mathcal{B}}(\psi)$. Thus, it suffices to bound the error $E_{\mathcal{A}}(\psi)$. We have

$$E_{\mathcal{A}}(\psi) \ll X^{n-D} N(X^{\psi}) \sup_{|\mathbf{y}| \leq X^{\psi}} E(\mathbf{y}, \theta) \ll (X^{1+\psi})^{n-D}(U_1 + U_2 + U_3),$$

where

$$U_1 = \sum_{j=2}^{d} X^{-\delta_j + \psi(\Delta_j + D_{d-j} + (D_{d-j+2}-2)(d-1) + D - d) - \nu},$$

$$U_2 = X^{-1 + (2d+1)\omega + (3d^3 - 5d^2 + 2)\psi},$$

$$U_3 = X^{\frac{1-\sigma-\Sigma}{\sigma}(-\omega + \frac{1}{2}(3d^2 - 7d + 4)\psi) + \frac{1}{6}(8d^3 + 3d^2 - 41d + 30)\psi}.$$

Assuming that

$$\delta_j = \psi(\Delta_j + D_{d-j} + (D_{d-j+2} - 2)(d-1) + D - d) \qquad (2 \leq j \leq d),$$

the exponent in the first term is negative. With this choice we have $\delta_d = (D-1)\psi$ and

$$\delta_j \le \delta_2 = \tfrac{1}{6}(4d^3 + 3d^2 - 25d + 18)\psi.$$

As before, this choice allows us to simplify the conditions (5.5) and (5.13), and we see that they and (5.14) are satisfied whenever $s - \rho > \max\{a_0(\varpi), a_1(\varpi), a_2(\varpi, \psi)\}$, with the same values as in the proof of Theorem 7.1.

Meanwhile, the error terms U_2 and U_3 are acceptable if we can choose θ_2 such that

$$\begin{aligned} &\frac{\frac{1}{6}(8d^3 + 3d^2 - 41d + 30)\psi}{1 - \sigma - \Sigma} + \frac{\frac{1}{2}(3d^2 - 7d + 4)\psi}{\sigma} \\ &\quad < k_2\theta_2 < \frac{1 - (3d^3 - 5d^2 + 2)\psi}{(2d+1)\sigma}, \end{aligned} \tag{7.11}$$

and this interval can be seen to be non-empty if (7.8) is satisfied and further

$$\left(1 + \frac{\frac{1}{6}(16d^4 + 14d^3 - 79d^2 + 19d + 30)\psi}{1 - \frac{1}{2}(12d^3 - 21d^2 + d + 8)\psi}\right)\sigma + \Sigma < 1. \tag{7.12}$$

When ψ satisfies (7.8) one can show for $d \ge 5$ that

$$\frac{\frac{1}{6}(16d^4 + 14d^3 - 79d^2 + 19d + 30)\psi}{1 - \frac{1}{2}(12d^3 - 21d^2 + d + 8)\psi} \le \frac{25}{3},$$

and hence (7.12) can be simplified to $\frac{28}{3}\sigma + \Sigma < 1$. Upon recalling (7.2) this gives $s - \rho > \beta_1(\varpi)$, where

$$\beta_1(\varpi) = \frac{2^d(d^2 + \frac{16}{3}d - \frac{10}{3})}{1 - \frac{1}{6}d(d-1)^2(2d+27)\varpi}.$$

It remains to compare the right hand side of (7.11) with the bound of (5.15). As before, with our choice of δ_2 we find that the first term in the maximum in (5.15) is bounded above by the left hand side of (7.11). Thus, it suffices to ensure that the interval

$$\psi\varpi^{-1} < k_2\theta_2 < \frac{1 - (3d^3 - 5d^2 + 2)\psi}{(2d+1)\sigma}$$

is non-empty. Such is the case when

$$\frac{(2d+1)\psi}{\varpi(1 - (3d^3 - 5d^2 + 2)\psi)}\sigma < 1,$$

and on inserting (7.2) we obtain the bound $s - \rho > \beta_2(\varpi, \psi)$ where

$$\beta_2(\varpi, \psi) = \frac{2^d(d-1)(2d+1)\psi}{\varpi(1 - (d^4 + \frac{3}{2}d^3 - 5d^2 + \frac{1}{2}d + 2)\psi)}.$$

When $d \geq 5$ one checks by a modicum of computation that $\beta_2(\varpi, \psi) \geq a_2(\varpi, \psi)$ and that $a_1(\varpi)$ exceeds both $\beta_1(\varpi)$ and $a_0(\varpi)$ in the appropriate ranges for ϖ and ψ. Just as before, we see that $a_1(\varpi) = \beta_2(\varpi, \psi)$ when $\varpi = \varpi_1(\psi)$, where

$$\varpi_1(\psi) = \frac{2(1+2d)(d-1)\psi}{d^3 + \frac{1}{2}d^2 - \frac{11}{2}d + 10 - q_6(d)\psi},$$

and $q_6(d)$ is the polynomial defined in the statement of the theorem. This is in accordance with (7.5), so that just as before we obtain our final bound $s - \rho > n_2(\psi)$ where we put $n_2(\psi) = \beta_2(\varpi_1(\psi), \psi)$. This completes the proof of the theorem. □

As before, one can show that $n_2(\psi)$ is increasing in ψ, and by taking $\psi = \psi_1$ with $\psi_1^{-1} = 2d^4$ we see after some calculations that

$$\begin{aligned} n_2(\psi_1) &= \frac{2^d(24d^7 - 38d^6 + 33d^5 + 155d^4 - 481d^3 + 639d^2 + 52d - 240)}{24d^4 - 36d^3 + 120d^2 - 12d - 48} \\ &\leq 2^d d(d^2-1) - \tfrac{1}{2}d(d+1) - 1. \end{aligned}$$

The conclusion of Theorem 1.4 now follows upon choosing $\rho = \frac{1}{2}d(d+1) + 1$.

It thus remains to evaluate the sum over the singular integral and singular series. This task can be absolved swiftly by invoking Theorem 2.1 in [9] and imitating arguments from [23, Section 8]. For fixed Y we set $\psi_0 = (d^3(d + \frac{3}{2}) - 1)^{-1}$ and $X_0 = Y^{1/\psi_0}$. Now assume that

$$n - \rho > 2^{d-1}d(d+1)(1 + \psi_0^{-1}). \tag{7.13}$$

Then by [9, Theorem 2.1] we have the alternative asymptotic formula

$$N(X_0, Y) = (X_0 Y)^{n-D}\chi_\infty \prod_{p \text{ prime}} \chi_p + O((X_0 Y)^{n-D} Y^{-\nu}).$$

On the other hand, one can check that the condition in (7.13) is stricter than the hypothesis of Theorem 7.2, so we may compare this bound with (7.9) and deduce that

$$\sum_{\substack{\mathbf{y} \in \mathcal{U}(\mathbb{Z}) \\ |\mathbf{y}| \leq Y \\ F(\mathbf{y}) = 0}} \mathfrak{S}_{\mathbf{y}} \mathfrak{J}_{\mathbf{y}} = Y^{n-D}\chi_\infty \prod_{p \text{ prime}} \chi_p + O(Y^{n-D-\nu}). \tag{7.14}$$

Note in particular that (7.14) does not depend on X_0 any longer. Thus, if (7.13) is satisfied, we are able to replace the sum over the singular series and integral in

Theorem 1.4 by a product of local densities as in (7.14). This establishes Theorem 1.1 for all $\psi \leq \psi_0$, while for $\psi_0 \leq \psi \leq 1$ the corresponding result follows from Theorem 2.1 in [9]. Finally, we recall that we need $\rho \geq \frac{1}{2}d(d+1)+1$ and note that

$$2^{d-1}d^4(d+1)(d+\tfrac{3}{2})+\tfrac{1}{2}d(d+1)+1 \leq 2^{d-1}d^4(d+1)(d+2)$$

for all admissible values d. This completes the proof of Theorem 1.1.

In order to complete the proof of our final result in Theorem 1.2, we note that in this case $\mathcal{U} = \mathcal{V}\backslash\{\mathbf{0}\}$. Thus, the error $E_{\mathcal{U}}(\psi) \ll X^{\psi(n-d)}$ is under control, and it remains to understand the error arising from any singular set $\mathcal{B}(\rho)$. From (4.9) we infer that $E_{\mathcal{B}}(\psi) \ll X^{n-d}X^{\psi(n-\rho)}$, which is acceptable within (7.10) if $\rho > D+d(d-1)/(2\psi)$. Picking ρ minimal in this way, we can now proceed precisely as in the proof of Theorem 1.1.

Acknowledgements During the production of this memoir, the author was supported by Starting Grant No. 2017-05110 from the Swedish Research Council (Vetenskapsrådet). The author is also grateful to Tim Browning and Per Salberger for valuable conversations around the topic of this paper, as well as useful comments from an anonymous referee.

Funding Open access funding provided by University of Gothenburg.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Altman, A.B., Kleiman, S.L.: Foundations of the theory of Fano schemes. Compositio Math. **34**(1), 3–47 (1977)
2. Batyrev, V.V., Tschinkel, Y.: Manin's conjecture for toric varieties. J. Algebraic Geom. **7**(1), 15–53 (1998)
3. Beheshti, R.: Linear subvarieties of hypersurfaces. Int. Math. Res. Not. **49**, 3055–3063 (2005)
4. Beheshti, R.: Lines on projective hypersurfaces. J. Reine Angew. Math. **592**, 1–21 (2006)
5. Beheshti, R.: Hypersurfaces with too many rational curves. Math. Ann. **360**(3–4), 753–768 (2014)
6. Birch, B.J.: Forms in many variables. Proc. Roy. Soc. Ser. A **265**, 245–263 (1962)
7. Blomer, V., Brüdern, J.: Counting in hyperbolic spikes: the diophantine analysis of multihomogeneous diagonal equations. J. Reine Angew. Math. **737**, 255–300 (2018)
8. Brandes, J.: Forms representing forms and linear spaces on hypersurfaces. Proc. Lond. Math. Soc. **108**, 809–835 (2014)
9. Brandes, J.: Forms representing forms: the definite case. J. Lond. Math. Soc. **92**, 393–410 (2015)
10. Brandes, J.: Linear spaces on hypersurfaces over number fields. Michigan Math. J. **66**, 769–784 (2017)
11. Brandes, J.: On the number of linear spaces on hypersurfaces with a prescribed discriminant. Math. Z. **289**(3–4), 803–827 (2018)
12. Brandes, J., Dietmann, R.: Rational lines on cubic hypersurfaces, Math. Proc. Camb. Philos. Soc., to appear
13. Browning, T.D.: Quantitative Arithmetic of Projective Varieties, Progress in Mathematics, vol. 277. Birkhäuser Verlag, Basel (2009)

14. Browning, T.D., Heath-Brown, D.H.: Forms in many variables and differing degrees. J. Eur. Math. Soc. **9**, 357–394 (2017)
15. Davenport, H.: Cubic forms in thirty-two variables. Philos. Trans. R. Soc. Ser. A **251**, 193–232 (1959)
16. Davenport, H.: Analytic Methods for Diophantine Equations and Diophantine inequalities, 2nd edn. Cambridge University Press, Cambridge (2005)
17. Franke, J., Manin, Y.I., Tschinkel, Y.: Rational points of bounded height on Fano varieties. Invent. Math. **95**, 421–435 (1989)
18. Harris, J., Mazur, B., Pandharipande, R.: Hypersurfaces of low degree. Duke Math. J. **95**, 125–160 (1998)
19. Heath-Brown, D.R.: Rational points on curves and surfaces. Ann. Math. **155**(2), 553–598 (2002)
20. Heath-Brown, D.R.: Cubic forms in ten variables. Proc. Lond. Math. Soc. **47**, 225–257 (1983)
21. Langer, A.: Fano schemes of linear spaces on hypersurfaces. Manuscripta Math. **93**, 21–28 (1997)
22. Peyre, E.: Hauteurs et mesures de Tamgawa sur les variétés de Fano. Duke Math. J. **79**, 101–218 (1995)
23. Schindler, D.: Manin's conjecture for certain biprojective hypersurfaces. J. Reine Angew. Math. **714**, 209–250 (2016)
24. Schindler, D., Sofos, E.: Sarnak's saturation problem for complete intersections. Mathematika **65**, 1–56 (2019)
25. Schmidt, W.M.: Simultaneous rational zeros of quadratic forms. In: Seminar on Number Theory, Paris 1980–81 (Paris, 1980/1981), pp. 281–307, Progr. Math., **22**, Birkhäuser, Boston, Mass., (1982)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.