

Diophantine Equations in Moderately Many Variables

OSCAR MARMON

1. Introduction

Let $f_1, \dots, f_r \in \mathbb{Z}[x_1, \dots, x_n]$ be polynomials of degrees d_1, \dots, d_r , respectively, and let \mathbf{f} denote the r -tuple of polynomials (f_1, \dots, f_r) . We are interested in upper bounds for the counting function

$$N(\mathbf{f}, B) := \#\{\mathbf{x} \in \mathbb{Z}^n; f_1(\mathbf{x}) = \dots = f_r(\mathbf{x}) = 0, |\mathbf{x}| \leq B\}.$$

(Here, and throughout the paper, $|\cdot|$ denotes the maximum norm $|\mathbf{x}| = \max\{|x_1|, \dots, |x_n|\}$.) If we assume that the polynomials f_i define a complete intersection in \mathbb{A}^n of dimension $n - r \geq 0$, then we have the well-known upper bound $N(\mathbf{f}, B) \ll B^{n-r}$, which we shall refer to as the trivial bound (cf. Lemma 2.5). Heuristic arguments suggest the bound $N(\mathbf{f}, B) \ll B^{n-D}$, where

$$D := \sum_{i=1}^r d_i,$$

at least as soon as $n > D$. In the special case where the polynomials f_i are homogeneous of the same degree d , a famous result by Birch establishes the heuristic upper bound and indeed an asymptotic formula, as soon as

$$n > s^* + 2^{d-1}(d-1)r(r+1). \tag{1}$$

Here, $s^* = s_{\mathbf{f}}^*$ is the dimension of the so-called Birch singular locus, the affine variety

$$\{\mathbf{x} \in \mathbb{A}^n \mid \text{rank } J(\mathbf{x}) < r\},$$

where $J(\mathbf{x})$ is the Jacobian matrix of size $r \times n$ with rows formed by the gradient vectors $\nabla f_i(\mathbf{x})$. (See also recent work by Dietmann [6] and, independently, Schindler [16], where s^* is replaced by an alternative quantity, sometimes leading to a stronger result.) Birch’s results have recently been extended to forms of differing degree by Browning and Heath-Brown [4].

Seeing as Birch’s theorem, like most results proven with the Hardy–Littlewood circle method, requires the number of variables to be rather large, we may ask if more modest upper bounds are still available for smaller values of n . At the far end of the spectrum, the *dimension growth conjecture* of Heath-Brown and Serre leads us to expect the bound $N(\mathbf{f}, B) \ll B^{n-\rho-1+\varepsilon}$ for an r -tuple of homogeneous polynomials \mathbf{f} defining an irreducible nonlinear variety of codimension $\rho \leq n - 2$ in \mathbb{P}^{n-1} . The determinant method has proved a useful tool in approaching this

Received February 27, 2015. Revision received May 23, 2016.

conjecture, and it has now been established in many cases; see [1, §3] for an overview. A full proof has recently been announced by Salberger.

We shall allow ourselves to call an r -tuple of polynomials \mathbf{f} as above a *system* of polynomials, and $\mathbf{d} := (d_1, \dots, d_r)$ its *multidegree*. For each polynomial f_i , we denote its leading form (homogeneous part of degree d_i) by F_i . The r -tuple $\mathbf{F} := (F_1, \dots, F_r)$ is then called the *system of leading forms* of \mathbf{f} . We associate with \mathbf{f} the projective variety $Z_{\mathbf{f}} \subset \mathbb{P}_{\mathbb{Q}}^{n-1}$ defined by the leading forms F_1, \dots, F_r . By a slight abuse of notation we write $\langle \mathbf{f} \rangle := \langle f_1, \dots, f_r \rangle$ for the ideal of $\mathbb{Z}[x_1, \dots, x_n]$ generated by the polynomials f_i . Finally, we define the *height* $\|\mathbf{f}\|$ to be the maximal absolute value of any coefficient appearing in one of the polynomials f_i in the system.

The results in this paper, which we are now ready to state, occupy a middle ground between the two types of bounds discussed above. We first state a simplified version of our main result.

THEOREM 1.1. *Suppose that $d_i \geq d \geq 4$ for all i and that $Z_{\mathbf{f}}$ is nonsingular of codimension r . Put*

$$\eta_{n,r,d} := \frac{2^{d-2}(d-1)r}{n + 2^{d-2}(d-1)r - 1}.$$

Then we have the estimate

$$N(\mathbf{f}, B) \ll_{n,\mathbf{d}} B^{n-rd(1-\eta_{n,r,d})} \log(\|\mathbf{F}\|)^{rd}$$

as soon as $n > 2^{d-2}(d-1)r$.

REMARK. For any ideal $I \subset \mathbb{Z}[x_1, \dots, x_n]$, one may define the *ideal of leading forms* I' consisting of the leading forms of all elements of I . If $I = \langle \mathbf{f} \rangle$ as above, then it may not always be the case that $I' = \langle \mathbf{F} \rangle$. However, it is well known that if the leading forms F_i cut out a subscheme of codimension r , then they do indeed generate the ideal I' of leading forms.

REMARK. In the excluded case where $\min\{d_i\} = 2$ or 3 , the estimate in Theorem 1.1 may be replaced by results of Luo [12, Thm. 2] or the author [13, Thm. 1.1], respectively.

In the case where the polynomials are truly of different degree, we can do better than Theorem 1.1. In order to state this more general result, we group the polynomials according to their degree. Thus, for each $2 \leq d \leq D$, where $D = \max_i \{d_i\}$, let $f_{d,1}, \dots, f_{d,r_d}$ be an enumeration of the polynomials of degree d among the f_i .

THEOREM 1.2. *Suppose that $D \geq 4$ and that $Z_{\mathbf{f}}$ is nonsingular of codimension r . Put*

$$\begin{aligned} \mathcal{D}' &:= \sum_{d=2}^{D-1} (d-1)r_d + Dr_D, \\ \Delta &:= \sum_{d=2}^{D-1} (d-2+2^{-d+1})r_d + (D-1)r_D, \\ \text{and } \eta &:= \frac{2^{D-2}\Delta}{n+2^{D-2}\Delta-1}. \end{aligned}$$

Then, provided that $n > 2^{D-2}\Delta$, we have the estimate

$$N(\mathbf{f}, B) \ll_{n, \mathbf{d}} B^{n-\mathcal{D}'(1-\eta)} \log(\|\mathbf{F}\|)^{\mathcal{D}}. \tag{2}$$

To compare the admissible range of n in the above theorems with (1), we consider the singular locus, in the usual sense, of the projective variety $Z_{\mathbf{f}}$, which is the variety

$$\text{Sing}(Z_{\mathbf{f}}) = \{\mathbf{x} \in \mathbb{P}_{\mathbb{Q}}^{n-1} \mid F_1(\mathbf{x}) = \dots = F_r(\mathbf{x}) = 0, \text{rank } J(\mathbf{x}) < r\}.$$

Here, $J(\mathbf{x})$ is the Jacobian matrix with rows $\nabla F_i(\mathbf{x})$. Putting $s = \dim(\text{Sing}(Z_{\mathbf{f}}))$, we have

$$s+1 \leq s^* \leq s+1+r$$

for $s^* = s_{\mathbf{f}}^*$. The nonsingularity assumption in our theorems translates to the condition that $s = -1$, but it would be an easy matter to derive more general results with $s+1$ playing a similar role as the quantity s^* in (1).

The proofs of Theorems 1.1 and 1.2 employ an iterated version of the multidimensional q -analogue of van der Corput differencing introduced by Heath-Brown [8] and are also built upon previous work by the author [13; 14].

REMARK 1.3. We observe that in the case where the polynomials are of equal degree, the exponent of B in Theorem 1.1 may be written as $n - \mu\mathcal{D}$, where $\mu \rightarrow 1$ as $n \rightarrow \infty$. Thus, our bound approaches the heuristic one asymptotically in this sense. The fact that the same is not true, in general, for the exponent in Theorem 1.2 indicates that our method is open to further improvement.

We shall now describe the strategy behind the proof of Theorems 1.1 and 1.2. Let us first introduce some notation.

NOTATION. We define

$$Z_{\mathbf{f}} := \text{Proj } \mathbb{Z}[x_1, \dots, x_n] / \langle \mathbf{F} \rangle \subset \mathbb{P}_{\mathbb{Z}}^{n-1},$$

so that the variety $Z_{\mathbf{f}}$ featuring in our main results may be expressed as

$$Z_{\mathbf{f}} = Z_{\mathbf{f}} \otimes_{\mathbb{Z}} \mathbb{Q}.$$

For any prime p , reduction modulo p gives rise to the projective variety

$$Z_{\mathbf{f}, p} := Z_{\mathbf{f}} \otimes_{\mathbb{Z}} \mathbb{F}_p = \text{Proj } \mathbb{Z}[x_1, \dots, x_n] / \langle p, F_1, \dots, F_r \rangle \subset \mathbb{P}_{\mathbb{F}_p}^{n-1}.$$

Writing $Z_{\mathbf{f},\infty} := Z_{\mathbf{f}}$ and $\mathbb{F}_{\infty} := \mathbb{Q}$ for notational convenience, we now define

$$\rho_v(\mathbf{f}) := \text{codim}(Z_{\mathbf{f},v}, \mathbb{P}_{\mathbb{F}_v}^{n-1}) \quad \text{and} \quad s_v(\mathbf{f}) := \dim(\text{Sing } Z_{\mathbf{f},v}),$$

where either $v = p$ for a prime p , or $v = \infty$. The parameters $\rho_v(\mathbf{f})$ and $s_v(\mathbf{f})$ will appear frequently in our arguments. They clearly satisfy the inequalities

$$0 \leq \rho_v(\mathbf{f}) \leq r, \quad -1 \leq s_v(\mathbf{f}) \leq n - 1 - \rho_v(\mathbf{f}).$$

It is also clear that $\rho_p(\mathbf{f}) \leq \rho_{\infty}(\mathbf{f})$ for all p and that $\rho_p(\mathbf{f}) = \rho_{\infty}(\mathbf{f})$ for $p \gg_{\mathbf{f}} 1$. Conversely, we have $s_p(\mathbf{f}) \geq s_{\infty}(\mathbf{f})$ with equality for $p \gg_{\mathbf{f}} 1$.

Our approach for estimating $N(\mathbf{f}, B)$ involves replacing the system of equations $f_{d,1}(\mathbf{x}) = \dots = f_{d,r_d}(\mathbf{x}) = 0$ for each $2 \leq d \leq D$ with a system of congruences

$$f_{d,1}(\mathbf{x}) \equiv \dots \equiv f_{d,r_d}(\mathbf{x}) \equiv 0 \pmod{q_d} \tag{3}$$

for suitably chosen integers q_2, \dots, q_D . More precisely, let m be an integer between 0 and $D - 2$. Then, given a parameter $\xi \in [B^{1/2}, B]$ and a collection of pairwise different primes p_0, \dots, p_m , satisfying

$$p_1 \asymp \dots \asymp p_m \asymp \xi, \quad p_0 \asymp \xi^2, \quad \text{and} \tag{4}$$

$$(p_0 p_1 \dots p_m, D!) = 1, \tag{5}$$

we take

$$q_d := \prod_{i=0}^{\min\{m, d-2\}} p_{m-i}. \tag{6}$$

REMARK. The natural number m represents the number of differencing steps that will be performed on the polynomials of highest degree D . Each congruence to the modulus q_d will correspond, after one step in our van der Corput differencing process, to a family of congruences of degree $d - 1$ to the smaller modulus $p_m^{-1} q_d$. The numbers q_d are chosen in such a way that we only ever need to consider congruences involving polynomials of degree at least 2 (discarding, as we may, all congruences to the modulus 1).

To encode the congruence relations (3), we now put

$$\mathbf{q}_d := (q_d, \dots, q_d) \in \mathbb{N}^{r_d}$$

for $2 \leq d \leq D$ and let $\mathbf{q} \in \mathbb{N}^r$ be the vector

$$\mathbf{q} = (q_2, \dots, q_2, \dots, q_D, \dots, q_D) = (\mathbf{q}_2, \dots, \mathbf{q}_D). \tag{7}$$

Now, for a suitably chosen smooth weight function $W \in C_0^\infty(\mathbb{R}^n)$, we certainly have

$$N(\mathbf{f}, B) \leq \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap [-B, B]^n \\ \mathbf{q} | \mathbf{f}(\mathbf{x})}} 1 \ll \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ \mathbf{q} | \mathbf{f}(\mathbf{x})}} W(B^{-1} \mathbf{x}) =: N_W(\mathbf{f}, B, \mathbf{q}), \tag{8}$$

where the condition $\mathbf{q} | \mathbf{f}(\mathbf{x})$ should be interpreted to mean that the congruence systems (3) are simultaneously satisfied for $2 \leq d \leq D$.

In estimating the quantity $N_W(\mathbf{f}, B, \mathbf{q})$, we shall group the polynomials f_i into blocks in two different ways: we put

$$\mathbf{f}_d = (f_{d,1}, \dots, f_{d,r_d}) \in \mathbb{Z}[\mathbf{x}]^{r_d}$$

for each $2 \leq d \leq D$ and, furthermore,

$$\hat{\mathbf{f}}_i = (\mathbf{f}_{i+2}, \mathbf{f}_{i+3}, \dots, \mathbf{f}_D) \in \mathbb{Z}[\mathbf{x}]^{r_{i+2}+\dots+r_D}$$

for each $0 \leq i \leq m$. (Thus, each system $\hat{\mathbf{f}}_i$ is composed of all polynomials $f_{d,j}$ for which $p_i \mid q_d$, so that they will be subject to at least i differencing steps. In particular, $\hat{\mathbf{f}}_0 = \mathbf{f}$ and $\hat{\mathbf{f}}_m = (\mathbf{f}_{m+2}, \dots, \mathbf{f}_D)$.)

We are now ready to formulate the main technical result of the paper, giving an asymptotic formula for a weighted counting function $N_W(\mathbf{f}, B, \mathbf{q})$. The class of smooth weight functions $C_n(R, (\kappa_j))$ occurring in the statement will be defined in Section 2.

PROPOSITION 1.4. *Let $\xi \in [B^{1/2}, B]$ and suppose that p_0, \dots, p_m are primes satisfying (4) and (5). Suppose that $\rho_{p_i}(\mathbf{f}) = r$ for all $0 \leq i \leq m$ and put*

$$s := \max_{0 \leq i \leq m} s_{p_i}(\hat{\mathbf{f}}_i).$$

Then, for any smooth weight function $W \in C_n(R, (\kappa_j))$, we have the asymptotic formula

$$N_W(\mathbf{f}, B, \mathbf{q}) - \frac{\mathcal{B}}{q_2^{r_2} q_3^{r_3} \dots q_D^{r_D}} \ll B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-s-2)/2^m} + B^n \xi^{-r/2} \left(\frac{\xi}{B}\right)^{(n-s-1)/2}, \tag{9}$$

where

$$\mathcal{R} = \sum_{i=2}^{m+1} (1 - 2^{-i+1})r_i + \sum_{i=m+2}^D r_i$$

and

$$\mathcal{B} = \mathcal{B}_W := \sum_{\mathbf{x} \in \mathbb{Z}^n} W(B^{-1}\mathbf{x}).$$

The implied constant in (9) depends at most on the data

$$n, \mathbf{d}, m, R \text{ and } (\kappa_j).$$

The case $m = 0$ in Proposition 1.4 gives an asymptotic formula for the weighted number of solutions of height at most B to a system of congruences modulo a single prime $p_0 \in [B, B^2]$, with an error term depending on the dimension of the singular locus of the corresponding variety. Such a result appeared as Theorem 3.3 in [13], Heath-Brown [8] having treated the case $r = 1$. It may be viewed as an extension of Hooley’s estimate [10] for the number of points on a complete intersection over a finite field, which in turn generalized Deligne’s estimate [5, Thm. 8.1].

The case $m = 1$, where a single differencing step is performed, was treated in [13]. To deduce Theorem 1.1 from Proposition 1.4, we shall take $m = d - 2 \geq 2$, in this way performing the same number of differencings on each of the polynomials. For Theorem 1.2, we instead take $m = D - 2$, performing the maximum number of differencings on the polynomials of highest degree and gradually fewer for those of lower degree.

After collecting the necessary tools in Section 2, we devote Section 3 to the proof of Proposition 1.4. The deduction of Theorems 1.1 and 1.2 is carried out in Section 4.

2. Preliminaries

The differencing process that will be used in Section 3 gives rise to new polynomials

$$f_i^{\mathbf{y}}(\mathbf{x}) = f_i(\mathbf{x} + p_m \mathbf{y}) - f_i(\mathbf{x}).$$

If $\mathbf{y} \neq \mathbf{0}$, then the polynomial $f_i^{\mathbf{y}}$ has degree $d_i - 1$, and its leading form is

$$F_i^{\mathbf{y}}(\mathbf{x}) = p_m \mathbf{y} \cdot \nabla F_i(\mathbf{x}).$$

To apply Proposition 1.4 iteratively, we need to control the dimension of singular loci of varieties defined by collections of such forms. The following result, which appeared in previous work of the author [13, Lemma 2.9], addresses that problem. However, that version contained two errors, which we now take the opportunity to correct:

- The condition $p \nmid d_i$ in the hypotheses needs to be strengthened to $p \nmid d_i(d_i - 1)$ in order for the proof to be valid.
- In the proof of item (i), the identity $\mathbf{x} \cdot \nabla^2 G_i(\mathbf{x}) = \dots$ is both erroneous and superfluous, and one gets a valid proof by simply ignoring this statement.

LEMMA 2.1. *Let $G_1, \dots, G_r \in \mathbb{Z}[x_1, \dots, x_n]$ be homogeneous polynomials of degrees d_1, \dots, d_r , respectively, let p be a prime such that $p \nmid d_i(d_i - 1)$ for all $i = 1, \dots, r$, and suppose that $\rho_p(\mathbf{G}) = r$ and $s_p(\mathbf{G}) = -1$. Define the closed subschemes $S \subseteq \mathbb{P}_{\mathbb{F}_p}^{n-1} \times \mathbb{P}_{\mathbb{F}_p}^{n-1}$ and $S_{\mathbf{y}} \subseteq \mathbb{P}_{\mathbb{F}_p}^{n-1}$ for each $\mathbf{y} \in \mathbb{P}_{\mathbb{F}_p}^{n-1}$ by setting*

$$S = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{P}_{\mathbb{F}_p}^{n-1} \times \mathbb{P}_{\mathbb{F}_p}^{n-1}; \mathbf{y} \cdot \nabla G_i(\mathbf{x}) = 0, i = 1, \dots, r, \\ \text{rank}(\mathbf{y} \cdot \nabla^2 G_i(\mathbf{x}))_{1 \leq i \leq r} < r\}$$

and

$$S_{\mathbf{y}} = \{\mathbf{x} \in \mathbb{P}_{\mathbb{F}_p}^{n-1}; \mathbf{y} \cdot \nabla G_i(\mathbf{x}) = 0, i = 1, \dots, r, \\ \text{rank}(\mathbf{y} \cdot \nabla^2 G_i(\mathbf{x}))_{1 \leq i \leq r} < r\}.$$

For $s = -1, 0, 1, \dots, n - 1$, let $T_s = \{\mathbf{y} \in \mathbb{P}_{\mathbb{F}_p}^{n-1}; \dim S_{\mathbf{y}} \geq s\}$. Then T_s is also Zariski closed, and we have the following estimates:

- (i) $\dim S \leq n - 2$;
- (ii) $\dim T_s \leq n - s - 2$;

(iii) If $T_s^{(1)}, T_s^{(2)}, \dots$ are the irreducible components of T_s , then

$$\sum_j \deg(T_s^{(j)}) = O_{n,d}(1).$$

REMARK 2.2. In particular, it follows that the set of \mathbf{y} such that $\mathbf{y} \cdot G_i(\mathbf{x})$ vanishes identically in \mathbf{x} for some i is (projectively) empty. Indeed, in this case, we would have $S_{\mathbf{y}} = \mathbb{P}^{n-1}$, and by (ii), we have $T_{n-1} = \emptyset$.

In our arguments, we use smooth weight functions, for which we introduce the following notation, partly following [9].

DEFINITION (Smooth weights). For any function $W \in C_0^\infty(\mathbb{R}^n)$, we define its radius $\text{Rad}(W)$ to be the smallest real number R such that $\text{supp}(W) \subset [-R, R]^n$. Furthermore, we put

$$\kappa_j(W) := \max_{j_1 + \dots + j_n = j} \max_{\mathbf{x} \in \mathbb{R}^n} \left| \frac{\partial^j}{\partial x_1^{j_1} \dots \partial x_n^{j_n}} W(\mathbf{x}) \right|$$

for any integer $j \geq 0$. In particular, $\kappa_0(W) = \max_{\mathbf{x}} |W(\mathbf{x})|$. Using this notation, given a positive real number R and a sequence of positive real numbers $(\kappa_j)_{j=0}^\infty$, we let $C_n(R, (\kappa_j))$ be the set of all functions $W \in C_0^\infty(\mathbb{R}^n)$ such that

$$\text{Rad}(W) \leq R \quad \text{and} \quad \kappa_j(W) \leq \kappa_j \quad \text{for all } j \in \mathbb{Z}_{\geq 0}.$$

In our next result, we list some elementary properties of the classes $C_n(R, (\kappa_j))$.

PROPOSITION 2.3. Suppose that $W \in C_n(R, (\kappa_j))$. Then we have the following properties:

- (i) For any $\mathbf{u} \in \mathbb{R}^n$, the translation $\tau_{\mathbf{u}}W$ defined by $\mathbf{x} \mapsto W(\mathbf{x} + \mathbf{u})$ satisfies $\tau_{\mathbf{u}}W \in C_n(R + |\mathbf{u}|, (\kappa_j))$.
- (ii) For any $x \in \mathbb{R}$, the restriction $r_xW : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ defined by $\mathbf{x} \mapsto W(\mathbf{x}, x)$ satisfies $r_xW \in C_{n-1}(R, (\kappa_j))$.
- (iii) If $W' \in C_n(R', (\kappa'_j))$, then $WW' \in C_n(R'', (\kappa''_j))$, where

$$R'' = \min\{R, R'\} \quad \text{and} \quad \kappa''_j \ll_n \sum_{j_1 + j_2 = j} \kappa_{j_1} \kappa'_{j_2}.$$

- (iv) For any matrix $M = (m_{ij}) \in \text{GL}_n(\mathbb{R})$, the composition $W \circ M$ satisfies $W \circ M \in C_n(R', (\kappa'_j))$, where

$$R' \ll_{n, \|M^{-1}\|} R \quad \text{and} \quad \kappa'_j \ll_{n, \|M\|} \kappa_j,$$

with $\|M\| := \max |m_{ij}|$.

Proof. Properties (i) and (ii) follow directly from the given definition, whereas (iii) and (iv) are easy consequences of the Leibniz formula and the chain rule, respectively. □

Our next result is a reformulation of [14, Lemma 3.1] in the notation introduced. It is a consequence of the Poisson summation formula and plays a key role in the proof of Proposition 1.4.

LEMMA 2.4. *Let $W \in \mathcal{C}(R, (\kappa_j))$, and let a and B be real numbers with $B \geq 1$ and $1 \leq a \leq B$. Then we have*

$$\begin{aligned} & \sum_{\mathbf{x} \in \mathbb{Z}^n} W\left(\frac{1}{B}\mathbf{x}\right) \sum_{\mathbf{y} \in \mathbb{Z}^n} W\left(\frac{1}{B}(\mathbf{x} + a\mathbf{y})\right) \\ &= a^{-n} \left(\sum_{\mathbf{x} \in \mathbb{Z}^n} W\left(\frac{1}{B}\mathbf{x}\right) \right)^2 + O_{n,R,N}(\kappa_0 \kappa_N B^{2n-N} a^{-n+N}) \\ & \quad + O_{n,R,N}(\kappa_N^2 B^{2(n-N)} a^{-n+N}) \end{aligned}$$

for any $N \in \mathbb{Z}_{\geq 0}$.

We may view this result as counting, with smooth weights, the number of points of height at most B in the lattice $\{(\mathbf{x}, \mathbf{x} + a\mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in \mathbb{Z}^n\} \subseteq \mathbb{R}^{2n}$. This lattice has determinant a^n , explaining the main term in the asymptotic formula, and the smooth weight allows for a strong error term.

Given any integer q , we may define the congruence counting function

$$N(\mathbf{f}, B, q) := \#\{\mathbf{x} \in \mathbb{Z}^n; f_1(\mathbf{x}) \equiv \dots \equiv f_r(\mathbf{x}) \equiv 0 \pmod{q}, |\mathbf{x}| \leq B\}.$$

The following result records “trivial” estimates for the counting functions $N(\mathbf{f}, B)$ and $N(\mathbf{f}, B, p)$ for any prime p .

LEMMA 2.5. *Given $\mathbf{f} \in \mathbb{Z}[x_1, \dots, x_n]^r$, we have the upper bound*

$$N(\mathbf{f}, B) \ll_{n,\mathbf{d}} B^{n-\rho_\infty(\mathbf{f})}. \tag{10}$$

Furthermore, for any prime p , we have

$$N(\mathbf{f}, B, p) \ll_{n,\mathbf{d}} B^{n-\rho_p(\mathbf{f})} + B^n p^{-\rho_p(\mathbf{f})}. \tag{11}$$

In particular,

$$\#\mathcal{Z}_{\mathbf{f},p}(\mathbb{F}_p) \ll_{n,\mathbf{d}} p^{n-\rho_p(\mathbf{f})}. \tag{12}$$

Such bounds are certainly used frequently, with or without proof, in the literature. In particular, a well-known reference for the bound (12) is the classic paper by Lang and Weil [11]. We allow ourselves to give a self-contained proof of all of these statements, thereby reproducing parts of the proof of [3, Lemma 4].

Proof of Lemma 2.5. Let us first assume that $B \leq p$ and prove the bound

$$N(\mathbf{f}, B, p) \ll_{n,\mathbf{d}} B^{n-\rho_p(\mathbf{f})}, \tag{13}$$

which will establish (11) in this case. For any closed subvariety $V \subset \mathbb{A}_{\mathbb{F}_p}^n$, let $V(B)$ be the set of points in $V(\mathbb{F}_p)$ having a representative \mathbf{x} in \mathbb{Z}^n with $|\mathbf{x}| \leq B$. Furthermore, let $\delta(V)$ be the sum of the degrees of the irreducible components

of V . We shall prove that

$$\#V(B) \ll_{\delta(V)} B^{\dim(V)}. \tag{14}$$

Since $\delta(Z_{\mathbf{f},p}) \leq \prod d_i$ by Bézout’s theorem [7, Ex. 8.4.6], this will imply the desired bound (13). We proceed by induction on the dimension $m := \dim(V)$. If $m = 0$, then it is clear that $\#V(B) \leq \delta(V)$. Assume now that $m \geq 1$. We may further assume that V is irreducible since it has at most $\delta(V)$ irreducible components, each of dimension at most m . By irreducibility we may choose $i \in \{1, \dots, n\}$ such that the hyperplane H_b defined by $x_i = b$ intersects V properly for any $b \in \mathbb{F}_p$ and write

$$\#V(B) \leq \sum_{b \in \mathbb{Z}; |b| \leq B} \#(V \cap H_b)(B).$$

Now we have $\delta(V \cap H_b) \leq \delta(V) = \deg(V)$, so our induction hypothesis implies that $\#(V \cap H_b)(B) \ll_{\delta(V)} B^{m-1}$ for each b in the sum, and the desired bound (14) follows.

Taking $B = p$ in (13), we immediately infer the bound (12). Furthermore, we clearly obtain a proof of (10) by replacing \mathbb{F}_p by \mathbb{Q} in the proof of (13).

Suppose next that $B > p$. For each $\mathbf{u} \in \mathbb{F}_p^n$, the number of points \mathbf{x} counted by $N(\mathbf{f}, B, p)$ whose reduction (mod p) is \mathbf{u} is at most $O_n((B/p)^n)$. By (12) we conclude that $N(\mathbf{f}, B, p) \ll_{n,\mathbf{d}} B^n p^{-\rho_p(\mathbf{f})}$, thus establishing the remaining case of (11). □

Our final result in this section is a variant of Bertini’s theorem, formulated in an “effective” way. Here, for $\mathbf{a} \in \mathbb{Z}^n$, we use the notation $H_{\mathbf{a}} \subset \mathbb{P}^{n-1}$ for the hyperplane defined by the equation $\mathbf{a} \cdot \mathbf{x} = 0$.

LEMMA 2.6. *Let $\Pi = \{p_1, \dots, p_N\}$ be a finite set of primes, and suppose that $\rho_p(\mathbf{f}) = r$ for each $p \in \Pi$. Then, provided that $\sum_{i=1}^N \frac{1}{p_i} \ll_{n,\mathbf{d}} 1$, there is a primitive integer vector $\mathbf{a} \in \mathbb{Z}^n$, with $\mathbf{a} \ll_{n,N,\mathbf{d}} 1$, such that, for each $p \in \Pi$, we have the properties*

- (i) $\dim(\text{Sing}(Z_{\mathbf{f},p} \cap H_{\mathbf{a}})) = \max\{-1, \dim(\text{Sing } Z_{\mathbf{f},p}) - 1\}$,
- (ii) $\dim(Z_{\mathbf{f},p} \cap H_{\mathbf{a}}) = \dim Z_{\mathbf{f},p} - 1$.

Proof. It follows from [13, Lemma 2.8] that there exist forms $\Phi_i \in \mathbb{Z}[\xi_1, \dots, \xi_n]$ of degree $O_{n,\mathbf{d}}(1)$, not identically divisible by p_i , such that properties (i) and (ii) hold for $p = p_i$ as soon as $p_i \nmid \Phi_i(\mathbf{a})$. For any $A \in \mathbb{N}$, let

$$\mathcal{M}(A) := \{\mathbf{a} \in \mathbb{Z}^n \cap [-A, A]^n; \gcd(a_1, \dots, a_n) = 1\}.$$

We note that

$$\#\mathcal{M}(A) = \frac{2^n}{\zeta(n)} A^n + O(A^{n-1} \log(A)) > A^n$$

as soon as A exceeds some explicit constant $C_1(n)$. Furthermore, we put

$$\mathcal{N}_i(A) := \{\mathbf{a} \in \mathcal{M}(A); p_i \mid \Phi_i(\mathbf{a})\}$$

for $i = 1, \dots, n$ and $\mathcal{N}(A) := \bigcup_i \mathcal{N}_i(A)$. Then by Lemma 2.5 there is a constant $C_2(n, \mathbf{d})$ such that

$$\#\mathcal{N}_i(A) \leq C_2(n, \mathbf{d})(A^{n-1} + A^n p_i^{-1})$$

for all i . This implies that

$$\#\mathcal{N}(A) \leq C_2(n, \mathbf{d}) \left(N A^{n-1} + A^n \sum_{i=1}^N \frac{1}{p_i} \right),$$

so if

$$\sum_{i=1}^N \frac{1}{p_i} \leq \frac{1}{2C_2(n, \mathbf{d})} \quad \text{and} \quad A \geq \max \left\{ C_1(n), \frac{2N}{C_2(n, \mathbf{d})} \right\},$$

then we get $\#\mathcal{N}(A) \leq A^n < \#\mathcal{M}(A)$, and in particular $\mathcal{N}(A) \neq \mathcal{M}(A)$, establishing the existence of an integer vector \mathbf{a} with the desired properties. \square

3. Proof of Proposition 1.4

We shall argue by induction over m . The inductive base $m = 0$ follows from [13, Thm. 3.3]. Indeed, we then have $\mathcal{R} = r$ and $q_d = p_0$ for all d . In this case, [13, Eq. (5)] with $q = p_0 \times \xi^2$ reads

$$N_W(\mathbf{f}, B, \mathbf{q}) - \frac{\mathcal{B}}{q^r} \ll B^{s+2} \xi^{n-r-s-2} = B^n \xi^{-r} \left(\frac{\xi}{B} \right)^{n-s-2},$$

as required for (9).

For $m \geq 1$, we shall use induction over s , the inductive base being the case $s = -1$. Thus, suppose that $s_{p_i}(\mathbf{f}_i) = -1$ for all $0 \leq i \leq m$. The quantity we wish to estimate is

$$V_W(\mathbf{f}, B, \mathbf{q}) := N_W(\mathbf{f}, B, \mathbf{q}) - \frac{\mathcal{B}}{Q},$$

where $Q := q_2^{r_2} q_3^{r_3} \cdots q_D^{r_D}$. We shall do this by carrying out a p_m -van der Corput differencing step, reducing all the moduli involved by a factor p_m . Our first step is to split the sum into congruence classes modulo p_m , yielding

$$V_W(\mathbf{f}, B, \mathbf{q}) = \sum_{\substack{\mathbf{u} \pmod{p_m} \\ p_m | \mathbf{f}(\mathbf{u})}} \sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{q} | \mathbf{f}(\mathbf{x})}} W(B^{-1}\mathbf{x}) - Q^{-1}\mathcal{B}, \tag{15}$$

where $\tilde{\mathbf{f}} := \hat{\mathbf{f}}_1 = (\mathbf{f}_3, \dots, \mathbf{f}_D)$. Note that if $\mathbf{x} \equiv \mathbf{u} \pmod{p_m}$, where $p_m | \mathbf{f}(\mathbf{u})$, then for each $3 \leq d \leq D$, the condition $q_d | \mathbf{f}_d(\mathbf{x})$ is equivalent to $\tilde{q}_d | \mathbf{f}_d(\mathbf{x})$, where $\tilde{q}_d := p_m^{-1} q_d$. Therefore, the ‘‘expected value’’ of the inner sum

$$\sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{q} | \mathbf{f}(\mathbf{x})}} W(B^{-1}\mathbf{x}) =: \Upsilon(\mathbf{u}),$$

say, is $p_m^{-n} \tilde{Q}^{-1} \mathcal{B}$, where

$$\tilde{Q} := \tilde{q}_3^{r_3} \cdots \tilde{q}_D^{r_D} = p_m^{-r} Q.$$

Writing $\tilde{\mathbf{q}} := (\tilde{q}_3, \dots, \tilde{q}_3, \dots, \tilde{q}_D, \dots, \tilde{q}_D) \in \mathbb{N}^{r_3} \times \dots \times \mathbb{N}^{r_D}$ and

$$S := \sum_{\substack{\mathbf{u} \pmod{p_m} \\ p_m | \mathbf{f}(\mathbf{u})}} \{\Upsilon(\mathbf{u}) - p_m^{-n} \tilde{Q}^{-1} \mathcal{B}\},$$

we obtain

$$V_W(\mathbf{f}, B, \mathbf{q}) = S + p_m^{-n+r} Q^{-1} \mathcal{B} \sum_{\substack{\mathbf{u} \pmod{p_m} \\ p_m | \mathbf{f}(\mathbf{u})}} 1 - Q^{-1} \mathcal{B}.$$

We have assumed that $\rho_{p_m}(\mathbf{f}) = r$ and $s_{p_m}(\mathbf{f}) = -1$, so we have

$$\sum_{\substack{\mathbf{u} \pmod{p_m} \\ p_m | \mathbf{f}(\mathbf{u})}} 1 = p_m^{n-r} + O(p_m^{(n-r+1)/2})$$

by Hooley’s extension [10] of Deligne’s theorem (see [13, Lemma 3.2]). Observing that $Q \asymp \xi^\kappa$, where

$$\kappa := \sum_{i=2}^{m+1} (i-1)r_i + (m+2) \sum_{i=m+2}^D r_i,$$

we conclude that

$$V_W(\mathbf{f}, B, \mathbf{q}) = S + O(B^n \xi^{-\kappa - (n-r-1)/2}). \tag{16}$$

The error term here is clearly admissible for (9). Indeed, since $\kappa - \frac{r}{2} \geq \mathcal{R}$ and $\xi^{-1} \leq \xi/B \leq 1$, we have

$$\begin{aligned} B^n \xi^{-\kappa - (n-r-1)/2} &< B^n \xi^{-\mathcal{R} - (n-1)/2} \leq B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2} \\ &\leq B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2^m}, \end{aligned}$$

as required.

Applying Cauchy’s inequality to the sum defining S and using (12) with $\rho_{p_m}(\mathbf{f}) = r$, we get

$$S \ll p_m^{(n-r)/2} \Sigma^{1/2},$$

and thus

$$V_W(\mathbf{f}, B, \mathbf{q}) \ll p_m^{(n-r)/2} \Sigma^{1/2} + B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2^m},$$

where

$$\Sigma := \sum_{\mathbf{u} \pmod{p_m}} \{\Upsilon(\mathbf{u}) - p_m^{-n} \tilde{Q}^{-1} \mathcal{B}\}^2$$

may be viewed as the “variance” of the quantity $\Upsilon(\mathbf{u})$ introduced before. To facilitate the analysis of this sum, we add some extra terms to it, obtaining

$$\Sigma \leq \sum_{3 \leq d \leq D} \sum_{\mathbf{a}_d \in (\mathbb{Z}/\tilde{q}_d \mathbb{Z})^{r_d}} \sum_{\mathbf{u} \pmod{p_m}} T_{\mathbf{u}}(\mathbf{a}_3, \dots, \mathbf{a}_D)^2, \tag{17}$$

where

$$T_{\mathbf{u}}(\mathbf{a}_3, \dots, \mathbf{a}_D) := \sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{f}_d(\mathbf{x}) \equiv \mathbf{a}_d \pmod{\tilde{q}_d}, 3 \leq d \leq D}} W(B^{-1}\mathbf{x}) - p_m^{-n} \tilde{Q}^{-1} \mathcal{B}.$$

In particular, $T_{\mathbf{u}}(\mathbf{0}, \dots, \mathbf{0}) = \Upsilon(\mathbf{u}) - p_m^{-n} \tilde{Q}^{-1} \mathcal{B}$.

REMARK 3.1. As seen in [14], one can obtain better bounds by circumventing this step, at the cost of more complicated geometric considerations. The current approach, however, seems better suited for an inductive argument with an arbitrary number of differencing steps.

When summing the expression

$$\begin{aligned} T_{\mathbf{u}}(\mathbf{a}_3, \dots, \mathbf{a}_D)^2 &= \left\{ \sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{f}_d(\mathbf{x}) \equiv \mathbf{a}_d \pmod{\tilde{q}_d}, 3 \leq d \leq D}} W(B^{-1}\mathbf{x}) \right\}^2 \\ &\quad - 2p_m^{-n} \tilde{Q}^{-1} \mathcal{B} \left\{ \sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{f}_d(\mathbf{x}) \equiv \mathbf{a}_d \pmod{\tilde{q}_d}, 3 \leq d \leq D}} W(B^{-1}\mathbf{x}) \right\} \\ &\quad + p_m^{-2n} \tilde{Q}^{-2} \mathcal{B}^2 \end{aligned}$$

over $\mathbf{a}_3, \dots, \mathbf{a}_D$ and \mathbf{u} as in (17), we see that the contribution from the last term precisely cancels the contribution from one of the cross terms, and thus the right-hand side of (17) equals

$$\sum_{3 \leq d \leq D} \sum_{\mathbf{a}_d \pmod{\tilde{q}_d}} \sum_{\mathbf{u} \pmod{p_m}} \left\{ \sum_{\substack{\mathbf{x} \equiv \mathbf{u} \pmod{p_m} \\ \mathbf{f}_d(\mathbf{x}) \equiv \mathbf{a}_d \pmod{\tilde{q}_d}, 3 \leq d \leq D}} W(B^{-1}\mathbf{x}) \right\}^2 - p_m^{-n} \tilde{Q}^{-1} \mathcal{B}^2.$$

Expanding the square, we express the leftmost sum as

$$\begin{aligned} &\sum_{\mathbf{x} \in \mathbb{Z}^n} \sum_{\substack{\mathbf{x}' \equiv \mathbf{x} \pmod{p_m} \\ \mathbf{f}_d(\mathbf{x}') \equiv \mathbf{f}_d(\mathbf{x}) \pmod{\tilde{q}_d}, 3 \leq d \leq D}} W(B^{-1}\mathbf{x}) W(B^{-1}\mathbf{x}') \\ &= \sum_{\mathbf{y} \in \mathbb{Z}^n} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ \tilde{q}_d | \mathbf{f}_d(\mathbf{x} + p_m \mathbf{y}) - \mathbf{f}_d(\mathbf{x}), 3 \leq d \leq D}} W(B^{-1}(\mathbf{x} + p_m \mathbf{y})) W(B^{-1}\mathbf{x}). \end{aligned}$$

Now we define, for each $\mathbf{y} \in \mathbb{Z}^n$ and each polynomial f_i , the differenced polynomial $f_i^{\mathbf{y}}$ by setting

$$f_i^{\mathbf{y}}(\mathbf{x}) = f_i(\mathbf{x} + p_m \mathbf{y}) - f_i(\mathbf{x}).$$

Accordingly, we put $\mathbf{f}_d^{\mathbf{y}} := (f_{d,1}^{\mathbf{y}}, \dots, f_{d,r_d}^{\mathbf{y}})$ for each $3 \leq d \leq D$, $\hat{\mathbf{f}}_i^{\mathbf{y}} := (\mathbf{f}_{m-i+2}^{\mathbf{y}}, \dots, \mathbf{f}_D^{\mathbf{y}})$ for each $0 \leq i \leq m - 1$, and $\hat{\mathbf{f}}^{\mathbf{y}} := \hat{\mathbf{f}}_3^{\mathbf{y}}$. Similarly, we define new weight functions $W_{\mathbf{y}}$ by

$$W_{\mathbf{y}}(\mathbf{t}) := W(\mathbf{t}) W(\mathbf{t} + B^{-1} p_m \mathbf{y}).$$

By Proposition 2.3 we have $W_{\mathbf{y}} \in \mathcal{C}(R, (\kappa'_j))$, where κ'_j may be bounded in terms of the numbers $\kappa_0, \dots, \kappa_j$. In this notation, we conclude from the previous considerations that

$$\Sigma \leq \sum_{\mathbf{y} \in \mathbb{Z}^n} N_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) - p_m^{-n} \tilde{Q}^{-1} B^2. \tag{18}$$

In fact, by the definition of $W_{\mathbf{y}}$ the sum runs only over $\mathbf{y} \ll B/p_m$, a fact we shall soon use.

If we denote the leading form of each polynomial $f_i^{\mathbf{y}}$ by $F_i^{\mathbf{y}}$, then we observe that

$$F_i^{\mathbf{y}} = p_m \mathbf{y} \cdot \nabla F_i,$$

unless the right-hand side vanishes identically in \mathbf{x} . But by Remark 2.2 this happens only if $\mathbf{y} = \mathbf{0}$. In particular, for $\mathbf{y} \neq \mathbf{0}$, $f_i^{\mathbf{y}}$ is a polynomial of degree $d_i - 1$. We shall shortly invoke our induction hypothesis, with \mathbf{f} replaced by $\tilde{\mathbf{f}}^{\mathbf{y}}$ and m replaced by $m - 1$, to obtain an estimate for

$$V_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) = N_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) - \tilde{Q}^{-1} \sum_{\mathbf{x} \in \mathbb{Z}^n} W_{\mathbf{y}}(B^{-1} \mathbf{x}) \tag{19}$$

for each $\mathbf{y} \neq \mathbf{0}$. We then need to show that the last term in (19), when summed over $\mathbf{y} \in \mathbb{Z}^n$, approximately cancels the last term in (17). To this end, note that

$$\begin{aligned} & \sum_{\mathbf{y} \in \mathbb{Z}^n} N_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) - \sum_{\mathbf{y} \in \mathbb{Z}^n} V_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) \\ &= \tilde{Q}^{-1} \sum_{\mathbf{y} \in \mathbb{Z}^n} \sum_{\mathbf{x} \in \mathbb{Z}^n} W(B^{-1}(\mathbf{x} + p_m \mathbf{y})) W(B^{-1} \mathbf{x}), \end{aligned}$$

which by Lemma 2.4 is

$$= p_m^{-n} \tilde{Q}^{-1} B^2 + O_N(p_m^{-n} \tilde{Q}^{-1} B^{2n} (p_m/B)^N)$$

for any $N \in \mathbb{N}$. The rightmost term here is

$$\ll p_m^{-n+r} B^{2n} \xi^{-\kappa} \left(\frac{\xi}{B}\right)^N,$$

and thus its contribution to S is

$$\ll B^n \xi^{-\kappa/2} \left(\frac{\xi}{B}\right)^N \leq B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^N,$$

since $\kappa \geq 2\mathcal{R}$. This is certainly admissible if N is chosen large enough. Thus, we obtain

$$\begin{aligned} V_W(\mathbf{f}, B, \mathbf{q}) &\ll p_m^{(n-r)/2} \left(\sum_{\mathbf{y} \ll B/p_m} V_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) \right)^{1/2} \\ &\quad + O\left(B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2^m} \right). \end{aligned} \tag{20}$$

For each $\mathbf{y} \in \mathbb{Z}^n$, we define

$$\rho(\mathbf{y}) := \min_{0 \leq i \leq m-1} \rho_{p_i}(\tilde{\mathbf{f}}^{\mathbf{y}}).$$

The number of equations in our new system $\tilde{\mathbf{F}}^{\mathbf{y}} = \mathbf{0}$ is $\tilde{r} := r - r_2$, so we have $0 \leq \rho(\mathbf{y}) \leq \tilde{r}$. Estimating the sum

$$\sum_{\mathbf{y} \ll B/p_m} V_{W_{\mathbf{y}}}(\tilde{\mathbf{F}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}), \tag{21}$$

we shall first consider the contribution from those \mathbf{y} for which the system $\tilde{\mathbf{F}}^{\mathbf{y}}$ still defines a complete intersection, that is, for which $\rho(\mathbf{y}) = \tilde{r}$. In this case, our induction hypothesis applies. Putting

$$s(\mathbf{y}) := \max_{0 \leq i \leq m-1} s_{p_i}(\hat{\mathbf{F}}_i^{\mathbf{y}}),$$

we have

$$V_{W_{\mathbf{y}}}(\tilde{\mathbf{F}}^{\mathbf{y}}, B, \tilde{\mathbf{q}}) \ll B^n \xi^{-\tilde{\mathcal{R}}} \left(\frac{\xi}{B} \right)^{(n-s(\mathbf{y})-2)/2^{m-1}},$$

where

$$\tilde{\mathcal{R}} = \sum_{i=2}^m (1 - 2^{-i+1})r_{i+1} + \sum_{i=m+1}^{D-1} r_{i+1}.$$

We shall apply Lemma 2.1 to estimate how often $s(\mathbf{y})$ attains a given value. Thus, suppose that $s(\mathbf{y}) = t$, where $-1 \leq t \leq n - 1$. (Here we ignore the fact that our assumption on $\rho(\mathbf{y})$ would allow us to use a smaller upper bound for $s(\mathbf{y})$ since the bounds we use are such that large values of t make a negligible contribution.) Let us fix a prime p_i such that $s_{p_i}(\hat{\mathbf{F}}_i^{\mathbf{y}}) = t$. Invoking our condition (5) and recalling that $s_{p_i}(\hat{\mathbf{F}}_i) = -1$ by assumption, we shall then apply Lemma 2.1 to the forms $F_{d,j}$, where $i + 2 \leq d \leq D$. If we denote by $V_{\mathbf{y}} \subset \mathbb{P}_{\mathbb{F}_{p_i}}^{n-1}$ the closed subvariety defined by

$$F_{d,j}^{\mathbf{y}} = 0, \quad 1 \leq j \leq r_d, i + 2 \leq d \leq D,$$

then our assumption means that $\dim(\text{Sing}(V_{\mathbf{y}})) = t$. However, since $\rho_{p_i}(\tilde{\mathbf{F}}^{\mathbf{y}}) = \tilde{r}$ by assumption and thus necessarily $\rho_{p_i}(\hat{\mathbf{F}}_i^{\mathbf{y}}) = r_{i+2} + \dots + r_D$, the singular locus of $V_{\mathbf{y}}$ is, by the Jacobian criterion, precisely the algebraic set $S_{\mathbf{y}}$ defined in Lemma 2.1. Hence, \mathbf{y} lies in (the affine cone over) the algebraic set T_s . By Lemmas 2.1 and 2.5 we conclude that there are at most $O((B/p_m)^{n-t-1})$ choices of \mathbf{y} for which $s(\mathbf{y}) = t$. Thus, the contribution to (21) from \mathbf{y} such that $\rho(\mathbf{y}) = \tilde{r}$ is at most

$$\begin{aligned} &\ll B^n \xi^{-\tilde{\mathcal{R}}} \left(\frac{\xi}{B} \right)^{(n-1)/2^{m-1}} \sum_{t'=0}^n \sum_{\substack{\mathbf{y} \ll B/p_m \\ s(\mathbf{y})=t'-1}} \left(\frac{B}{\xi} \right)^{t'/2^{m-1}} \\ &\ll B^n \xi^{-\tilde{\mathcal{R}}} \left(\frac{\xi}{B} \right)^{(n-1)/2^{m-1}} \sum_{t'=0}^n \left(\frac{B}{\xi} \right)^{n-t'(1-1/2^{m-1})} \\ &\ll B^{2n} \xi^{-n-\tilde{\mathcal{R}}} \left(\frac{\xi}{B} \right)^{(n-1)/2^{m-1}}. \end{aligned}$$

Then the contribution to the right-hand side of (20) from these values of \mathbf{y} is

$$\begin{aligned} &\ll p_m^{(n-r)/2} B^n \xi^{-(n+\tilde{\mathcal{R}})/2} \left(\frac{\xi}{B}\right)^{(n-1)/2^m} \\ &\ll B^n \xi^{-(r+\tilde{\mathcal{R}})/2} \left(\frac{\xi}{B}\right)^{(n-1)/2^m} \\ &= B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2^m}, \end{aligned}$$

as required.

Now we treat those values of \mathbf{y} for which $\rho(\mathbf{y}) < \tilde{r}$. In this case, the induction hypothesis does not apply, but we may apply a more elementary bound. If $\rho(\mathbf{y}) = \rho$, then there is some p_i such that $\rho_{p_i}(\tilde{\mathbf{f}}^{\mathbf{y}}) = \rho$, so a crude upper bound for $N_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}})$, and hence for $V_{W_{\mathbf{y}}}(\tilde{\mathbf{f}}^{\mathbf{y}}, B, \tilde{\mathbf{q}})$, is given by $B^n \xi^{-\rho}$, by Lemma 2.5. Furthermore, it is easy to see [13, Lemma 4.2] that in this case we actually have $\mathbf{y} \in T_{n-\rho-1}$. Again, by Lemmas 2.1 and 2.5 we conclude that the contribution to (21) from values of \mathbf{y} with $\rho(\mathbf{y}) < \tilde{r}$ (including $\mathbf{y} = \mathbf{0}$) is at most

$$\begin{aligned} &\ll \sum_{\rho=0}^{\tilde{r}-1} \sum_{\substack{\mathbf{y} \ll B/p_m \\ \rho(\mathbf{y})=\rho}} B^n \xi^{-\rho} \ll \sum_{\rho=0}^{\tilde{r}-1} \left(\frac{B}{\xi}\right)^{\rho} B^n \xi^{-\rho} \\ &= B^n \sum_{\rho=0}^{\tilde{r}-1} \left(\frac{B}{\xi^2}\right)^{\rho} \ll B^n. \end{aligned}$$

This yields a contribution to the right-hand side of (20) of size at most

$$\ll B^n \xi^{-r/2} \left(\frac{\xi}{B}\right)^{n/2},$$

corresponding to the second error term in (9). This concludes the treatment of the case $s = -1$.

Next, assume that $s \geq 0$. By Lemma 2.6 there is a primitive integer vector $\mathbf{a} \ll_{n,D} 1$ such that

$$\rho_{p_i}(\mathbf{f}|_{\mathbf{a}}) = r + 1 \quad \text{and} \quad s_{p_i}(\hat{\mathbf{f}}_i|_{\mathbf{a}}) = s - 1 \tag{22}$$

for all $0 \leq i \leq m$, where $\mathbf{f}|_{\mathbf{a}} := (f_1, \dots, f_r, a_1 x_1 + \dots + a_n x_n)$, and analogously for $\hat{\mathbf{f}}_i|_{\mathbf{a}}$. We may then find a unimodular matrix $M \in \text{GL}_n(\mathbb{Z})$, all of whose entries are $O_{n,D}(1)$, such that $M^T \mathbf{a} = \mathbf{e}_n$ and write

$$\begin{aligned} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ \mathbf{q}|\mathbf{f}(\mathbf{x})}} W(B^{-1}(\mathbf{x})) &= \sum_{b \ll B} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ \mathbf{q}|\mathbf{f}(\mathbf{x}), \mathbf{a} \cdot \mathbf{x} = b}} W(B^{-1} \mathbf{x}) \\ &= \sum_{b \ll B} \sum_{\substack{\mathbf{x}' \in \mathbb{Z}^n \\ \mathbf{q}|\mathbf{f}(M\mathbf{x}'), x'_n = b}} W(B^{-1} M\mathbf{x}') \\ &= \sum_{b \ll B} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^{n-1} \\ \mathbf{q}|\mathbf{g}_b(\mathbf{x})}} \tilde{W}_{B^{-1}b}(B^{-1} \mathbf{x}). \end{aligned}$$

Here we have put $\mathbf{g}_b = (g_{b,1}, \dots, g_{b,r})$, where $g_{b,i}$, for each $1 \leq i \leq r$, is the polynomial in $n - 1$ variables given by $g_{b,i}(\mathbf{x}) := f_i(M(\mathbf{x}, b))$, and we have defined $\tilde{W}_u : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ by $\tilde{W}_u(\mathbf{t}) := W(M(\mathbf{t}, u))$. Moreover, we have

$$\sum_{b \ll B} \sum_{\mathbf{x} \in \mathbb{Z}^{n-1}} \tilde{W}_{B^{-1}b}(B^{-1}\mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{Z}^n} W(B^{-1}M\mathbf{x}) = \sum_{\mathbf{x} \in \mathbb{Z}^n} W(B^{-1}\mathbf{x})$$

since M is unimodular.

It follows that

$$\begin{aligned} V_W(\mathbf{f}, B, \mathbf{q}) &= \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \\ \mathbf{q}\mathbf{f}(\mathbf{x})}} W(B^{-1}(\mathbf{x})) - Q^{-1} \sum_{\mathbf{x} \in \mathbb{Z}^n} W(B^{-1}(\mathbf{x})) \\ &= \sum_{b \ll B} V_{\tilde{W}_{B^{-1}b}}(\mathbf{g}_b, B, \mathbf{q}). \end{aligned} \tag{23}$$

Note that the polynomials $\mathbf{g}_{b,i}$ have the same leading forms $G_i(\mathbf{x}) := F_i(M(\mathbf{x}, 0))$ for each $b \in \mathbb{Z}$ (indeed, by (22), G_i cannot vanish identically) and that the subscheme $V(G_1, \dots, G_r) \subset \mathbb{P}_{\mathbb{Z}}^{n-2}$ is isomorphic to the subscheme $V(F_1, \dots, F_r, \mathbf{a} \cdot \mathbf{x}) \subset \mathbb{P}_{\mathbb{Z}}^{n-1}$. Thus, we may conclude by (22) that $\rho_{p_i}(\mathbf{g}_b) = r$ and $s_{p_i}(\mathbf{g}_b) = s - 1$ for all $b \in \mathbb{Z}$ and all $0 \leq i \leq m$. Moreover, we have $\tilde{W}_{B^{-1}b} \in \mathcal{C}(R', (\kappa'_j))$, where $R' \ll R$ and $\kappa'_j \ll \kappa_j$, by Proposition 2.3. Consequently, we may invoke our induction hypothesis to obtain the bound

$$\begin{aligned} V_{\tilde{W}_{B^{-1}b}}(\mathbf{g}_b, B, \mathbf{q}) &\ll B^{n-1} \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-s-2)/2^m} \\ &\quad + B^{n-1} \xi^{-r/2} \left(\frac{\xi}{B}\right)^{(n-s-1)/2}. \end{aligned}$$

Summing over $b \ll B$, by (23) we get

$$V_W(\mathbf{f}, B, \mathbf{q}) \ll B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-s-2)/2^m} + B^n \xi^{-r/2} \left(\frac{\xi}{B}\right)^{(n-s-1)/2},$$

as desired.

4. Proof of the Main Results

Suppose that we are given a system \mathbf{f} of polynomials with $\rho_\infty(\mathbf{f}) = r$ and $s_\infty(\mathbf{f}) = -1$. We want to apply Proposition 1.4, so our first step will be to find an alternative system $\mathbf{g} = (g_1, \dots, g_r)$, generating the same ideal (\mathbf{f}) , such that $s_\infty(\mathbf{g}_i) = -1$ for each $i = 0, \dots, D - 2$. The following result, which is a slight modification of [2, Lemma 3.1], shows that this is possible.

LEMMA 4.1. *Let $\mathbf{f} = (f_1, \dots, f_r)$ be a system of polynomials in $\mathbb{Z}[\mathbf{x}]$ with*

$$\deg(f_1) \leq \dots \leq \deg(f_r),$$

and let $I = \langle \mathbf{f} \rangle$ be the ideal in $\mathbb{Z}[\mathbf{x}]$ that it generates. Suppose that the leading forms F_1, \dots, F_r define a nonsingular complete intersection of codimension r in $\mathbb{P}_{\mathbb{Q}}^{n-1}$. Then there is another system $\mathbf{g} = (g_1, \dots, g_r)$, with $\deg(g_i) = \deg(f_i)$ for

all i , such that the corresponding leading forms G_1, \dots, G_r have the following property:

$$\text{For each } 0 \leq j \leq r, \text{ the subset } G_{r-j+1}, \dots, G_r \text{ defines} \tag{24}$$

$$\text{a nonsingular complete intersection of codimension } j.$$

Furthermore, $\|\mathbf{g}\|$ may be bounded above by a polynomial in $\|\mathbf{f}\|$ whose height and degree are bounded in terms of n and \mathbf{d} .

Proof. Let us write I' for the ideal generated by the leading forms F_i . In the proof of [2, Lemma 3.1], a new system of generators G_i satisfying property (24) is found by taking

$$G_i = \sum_{\substack{1 \leq j \leq i \\ d_i = d_j}} \lambda_{ij} F_j + \sum_{\substack{1 \leq j \leq i \\ d_i > d_j}} \sum_{1 \leq k \leq n} \lambda_{ijk} x_k^{d_i - d_j} F_j \tag{25}$$

for suitable integers $\lambda_{ij}, \lambda_{ijk}$. It is not hard to see that we may take $\lambda_{ii} = 1$ for all i , so that the G_i indeed generate the ideal I' as a $\mathbb{Z}[\mathbf{x}]$ -module. It follows immediately by construction that the polynomials

$$g_i = \sum_{\substack{1 \leq j \leq i \\ d_i = d_j}} \lambda_{ij} f_j + \sum_{\substack{1 \leq j \leq i \\ d_i > d_j}} \sum_{1 \leq k \leq n} \lambda_{ijk} x_k^{d_i - d_j} f_j$$

then generate the ideal I .

It remains to verify the height bound. To this end, we need to redo the previous argument in a more abstract setting. Let \mathbb{P}_{d_i} , for $i = 1, \dots, r$, be the projective space $\mathbb{P}_{\mathbb{Z}}^{N_i}$ parameterizing degree d_i hypersurfaces in $\mathbb{P}_{\mathbb{Z}}^{n-1}$. By abuse of notation, we will write F_i for the element of \mathbb{P}_{d_i} representing the hypersurface $F_i = 0$, and $\mathbf{F} = (F_1, \dots, F_r)$ for an element of $\mathbb{P}_{d_1} \times \dots \times \mathbb{P}_{d_r}$. Moreover, for each $i = 1, \dots, r$, we consider the parameter spaces $\mathbf{\Lambda}_i = \mathbb{P}_{\mathbb{Z}}^{M_i - 1}$, where M_i is the number of unknowns $\lambda_{ij}, \lambda_{ijk}$ occurring in the right-hand side of (25) for a fixed i , and denote a general element of $\mathbf{\Lambda}_i$ by λ_i . Let $V_{\mathbf{F}, \lambda_i}$ denote the hypersurface in \mathbb{P}^{n-1} defined by the form G_i in (25). It is then a standard fact that the locus of $(\mathbf{F}, \lambda_1, \dots, \lambda_r)$ such that the intersection $V_{\mathbf{F}, \lambda_r} \cap \dots \cap V_{\mathbf{F}, \lambda_{r-j+1}}$ is proper and smooth for each $j = 1, \dots, r$ is an open subset \mathcal{U} of $\mathbb{P}_{d_1} \times \dots \times \mathbb{P}_{d_r} \times \mathbf{\Lambda}_1 \times \dots \times \mathbf{\Lambda}_r$. Its complement, being a closed subset, is thus defined by the simultaneous vanishing of a finite collection of multihomogeneous forms $\mathcal{F}_i(\mathbf{F}; \lambda_1; \dots; \lambda_r)$ with coefficients in \mathbb{Z} .

For the specific r -tuple \mathbf{F} under consideration, we have already verified the nonemptiness of $\pi^{-1}(\mathbf{F})$, where π denotes the projection from \mathcal{U} to $\mathbb{P}_{d_1} \times \dots \times \mathbb{P}_{d_r}$. In particular, at least one of these forms $\mathcal{G}(\lambda_1; \dots; \lambda_r) := \mathcal{F}_i(\mathbf{F}; \lambda_1; \dots; \lambda_r)$, say, does not vanish identically in $\lambda_1, \dots, \lambda_r$. Clearly, we have the bounds

$$\|\mathcal{G}\| \ll_{n, \mathbf{d}} \|\mathbf{F}\|^{O_{n, \mathbf{d}}(1)} \quad \text{and} \quad \deg(\mathcal{G}) \ll_{n, \mathbf{d}} 1.$$

As an easy consequence of Lemma 2.5, we may then find a solution $\lambda_1, \dots, \lambda_r$ to $\mathcal{G}(\lambda_1; \dots; \lambda_r) \neq 0$ satisfying $|\lambda_i| \ll_{n, \mathbf{d}} \|\mathbf{F}\|^{O_{n, \mathbf{d}}(1)}$. This in turn gives the desired bound on $\|\mathbf{g}\|$. □

By Lemma 4.1 we may now assume that $s_\infty(\hat{\mathbf{f}}_i) = -1$ for each $i = 0, \dots, D - 2$. Indeed, the possible increase in $\|\mathbf{f}\|$ is absorbed by the implied constant in (2).

LEMMA 4.2. *Suppose that*

$$\rho_\infty(\mathbf{f}) = r \quad \text{and} \quad s_\infty(\hat{\mathbf{f}}_0) = \dots = s_\infty(\hat{\mathbf{f}}_{D-2}) = -1.$$

Then, provided that $\xi \gg \log \|\mathbf{F}\|$, there exist m distinct primes p_1, \dots, p_m with $p_j \asymp \xi$ such that

$$\rho_{p_j}(\mathbf{f}) = r \quad \text{and} \quad s_{p_j}(\hat{\mathbf{f}}_0) = \dots = s_{p_j}(\hat{\mathbf{f}}_{D-2}) = -1 \tag{26}$$

for all $j = 1, \dots, m$. The implied constants depend only on n, \mathbf{d} , and m .

Proof. By an argument similar to that in the proof of Lemma 4.1 the set of $\mathbf{F} = (F_1, \dots, F_r)$ such that either $\rho_\infty(\mathbf{F}) < r$ or $s_\infty(\hat{\mathbf{F}}_i) > -1$ for some i is a closed subset of $\mathbb{P}_{d_1} \times \dots \times \mathbb{P}_{d_r}$, defined by a collection of multihomogeneous forms. By assumption, at least one of these, say $\mathcal{G}(\mathbf{F})$, does not vanish when \mathbf{F} is taken to be the system of leading forms of the system \mathbf{f} in the hypotheses. Furthermore, any prime p violating condition (26) has to divide the integer $\mathcal{G}(\mathbf{F})$.

It is a straightforward consequence of Chebyshev’s theorem that for any natural number m , there is a constant C depending only on m such that for any natural number $A \geq 3$, the interval $(\log A, C \log A]$ contains at least m distinct primes not dividing A . From this the lemma follows by taking $A = \mathcal{G}(\mathbf{F})$ and observing that

$$\log \mathcal{G}(\mathbf{F}) \ll_{n, \mathbf{d}} \log \|F\|. \quad \square$$

We shall derive the estimate in Theorem 1.2 by applying Proposition 1.4 with $m = D - 2$. A suitable smooth weight function is given by

$$W(\mathbf{t}) = \prod_{i=1}^n w(t_i/2), \quad \text{where } w(t) = \begin{cases} \exp(-1/(1-t^2)), & |t| < 1, \\ 0, & |t| \geq 1. \end{cases}$$

The asymptotic formula (9) in particular yields an upper bound for $N_W(\mathbf{f}, B, q)$. Assuming that $s = -1$, the optimal such upper bound will be obtained when

$$\xi = B^{1-\eta}, \tag{27}$$

where η is the quantity defined in Theorem 1.2. Provided that $n > 2^{D-2}\Delta$, this implies that $\xi \geq B^{1/2}$, as required in Proposition 1.4. Suppose first that

$$B \gg (\log \|\mathbf{F}\|)^{1/(1-\eta)}.$$

In view of Lemma 4.2, we may then indeed apply Proposition 1.4, and the primes p_0, \dots, p_{D-2} may be chosen so that $s = -1$. We obtain

$$N(\mathbf{f}, B) \ll B^n \xi^{-D'} + B^n \xi^{-\mathcal{R}} \left(\frac{\xi}{B}\right)^{(n-1)/2^{D-2}} + B^n \xi^{-r/2} \left(\frac{\xi}{B}\right)^{n/2}.$$

In (27), we have already set ξ to make the first two terms on the right-hand side equal. Verifying that the third term is then of negligible size as soon as $D \geq 4$, we arrive at the bound

$$N(\mathbf{f}, B) \ll B^{n-D'(1-\eta)},$$

thus establishing the bound in Theorem 1.2 in this case. Here, the implied constant depends only on n and \mathbf{d} .

In the complementary case where $B \ll (\log \|\mathbf{F}\|)^{1/(1-\eta)}$, we may, for example, use the trivial bound from Lemma 2.5 to obtain

$$\begin{aligned} N(\mathbf{f}, B) &\ll B^{n-r} \ll B^{n-\mathcal{D}'(1-\eta)} (\log \|\mathbf{F}\|)^{\mathcal{D}'-r/(1-\eta)} \\ &\leq B^{n-\mathcal{D}'(1-\eta)} (\log \|\mathbf{F}\|)^{\mathcal{D}'}, \end{aligned}$$

as desired. This completes the proof of Theorem 1.2.

To prove Theorem 1.1, one instead takes $m = d - 2 \geq 2$ in the application of Proposition 1.4. More generally, for certain given specifications of the data n, \mathbf{d} , the optimal upper bound for $N(\mathbf{f}, B)$ may be attained by choosing some other value of $m \in \{0, \dots, D - 2\}$. Thus, there is a whole range of intermediate results between Theorem 1.1 and Theorem 1.2, a precise formulation of which is left to the reader.

5. Concluding Remarks

Continuing the discussion from Remark 1.3, it would be desirable to have a result where the exponent of B approaches the heuristic one also for systems of truly differing degree. In fact, the expression $n - \mathcal{D}'$, which is the asymptotic value of the exponent of B in Theorem 1.2 as $n \rightarrow \infty$, would be the expected one for a system where all polynomials of degree $d \leq D - 1$ were replaced by polynomials of degree $d - 1$. This is because our setup at the moment needs to exclude polynomials of degree ≤ 2 in each differencing step.

One could also imagine using ideas from [14] to improve our results. That paper considers the case where $r = 1$ and $m = 2$. Using the ideas of Salberger [15], an alternative differencing procedure is used, which allows one to keep the original polynomial in play in each step (see Remark 3.1). This produces varieties of gradually increasing codimension, giving stronger bounds. But it is not obvious how to turn the arguments in [14] into an iterative approach, nor how to generalize the geometric considerations playing the role of Lemma 2.1 in [14] to the case $r > 1$.

ACKNOWLEDGMENTS. I wish to thank the anonymous referee for valuable remarks and Jörg Brüdern for helpful discussions concerning this paper.

References

- [1] T. D. Browning, *Quantitative arithmetic of projective varieties*, Progr. Math., 277, Birkhäuser, Basel, 2009, 160 pp.
- [2] T. D. Browning, R. Dietmann, and D. R. Heath-Brown, *Rational points on intersections of cubic and quadric hypersurfaces*, J. Inst. Math. Jussieu 14 (2015), no. 4, 703–749.
- [3] T. D. Browning and D. R. Heath-Brown, *Rational points on quartic hypersurfaces*, J. Reine Angew. Math. 629 (2009), 37–88.

- [4] ———, *Forms in many variables and differing degrees*, J. Eur. Math. Soc. (2014, to appear), [arXiv:1403.5937](https://arxiv.org/abs/1403.5937).
- [5] P. Deligne, *La conjecture de Weil. I*, Publ. Math. Inst. Hautes Études Sci. 43 (1974), 273–307.
- [6] R. Dietmann, *Weyl's inequality and systems of forms*, Q. J. Math. 66 (2015), no. 1, 97–110.
- [7] W. Fulton, *Intersection theory*, Ergeb. Math. Grenzgeb. (3) [Results in Mathematics and Related Areas (3)], 2, Springer-Verlag, Berlin, 1984.
- [8] D. R. Heath-Brown, *The density of rational points on nonsingular hypersurfaces*, Proc. Indian Acad. Sci. Math. Sci. 104 (1994), no. 1, 13–29.
- [9] ———, *A new form of the circle method, and its application to quadratic forms*, J. Reine Angew. Math. 481 (1996), 149–206.
- [10] C. Hooley, *On the number of points on a complete intersection over a finite field*, J. Number Theory 38 (1991), no. 3, 338–358.
- [11] S. Lang and A. Weil, *Number of points of varieties in finite fields*, Amer. J. Math. 76 (1954), 819–827.
- [12] W. Luo, *Rational points on complete intersections over \mathbf{F}_p* , Int. Math. Res. Not. IMRN 16 (1999), 901–907.
- [13] O. Marmon, *The density of integral points on complete intersections*, Q. J. Math. 59 (2008), no. 1, 29–53, with an appendix by Per Salberger.
- [14] ———, *The density of integral points on hypersurfaces of degree at least four*, Acta Arith. 141 (2010), no. 3, 211–240.
- [15] P. Salberger, *Integral points on hypersurfaces of degree at least three*, Unpublished.
- [16] D. Schindler, *A variant of Weyl's inequality for systems of forms and applications*, Advances in the theory of numbers, Fields Inst. Commun., 77, pp. 207–218, Fields Inst. Res. Math. Sci, Toronto, ON, 2015.

Department of Mathematical
Sciences
University of Copenhagen
Universitetsparken 5
2100 København Ø
Denmark

marmon@math.ku.dk