Algebra & Number Theory

Volume 18
2024
No. 8

Application of a polynomial sieve: beyond separation of variables

Dante Bonolis and Lillian B. Pierce

msp

ALGEBRA AND NUMBER THEORY 18:8 (2024)
https://doi.org/10.2140/ant.2024.18.1515

Application of a polynomial sieve: beyond separation of variables

Dante Bonolis and Lillian B. Pierce

Let a polynomial $f \in \mathbb{Z}[X_1, \dots, X_n]$ be given. The square sieve can provide an upper bound for the number of integral $\boldsymbol{x} \in [-B, B]^n$ such that $f(\boldsymbol{x})$ is a perfect square. Recently this has been generalized substantially: first to a power sieve, counting $\boldsymbol{x} \in [-B, B]^n$ for which $f(\boldsymbol{x}) = y^r$ is solvable for $y \in \mathbb{Z}$; then to a polynomial sieve, counting $\boldsymbol{x} \in [-B, B]^n$ for which $f(\boldsymbol{x}) = g(y)$ is solvable, for a given polynomial g. Formally, a polynomial sieve lemma can encompass the more general problem of counting $\boldsymbol{x} \in [-B, B]^n$ for which $F(y, \boldsymbol{x}) = 0$ is solvable, for a given polynomial F. Previous applications, however, have only succeeded in the case that $F(y, \boldsymbol{x})$ exhibits separation of variables, that is, $F(y, \boldsymbol{x})$ takes the form $f(\boldsymbol{x}) - g(y)$. In the present work, we present the first application of a polynomial sieve to count $\boldsymbol{x} \in [-B, B]^n$ such that $F(y, \boldsymbol{x}) = 0$ is solvable, in a case for which F does not exhibit separation of variables. Consequently, we obtain a new result toward a question of Serre, pertaining to counting points in thin sets.

1. Introduction

Fix an integer $m \geq 2$ and integers $d, e \geq 1$. Consider the polynomial

$$F(Y, \boldsymbol{X}) = Y^{md} + Y^{m(d-1)} f_1(\boldsymbol{X}) + \cdots + Y^m f_{d-1}(\boldsymbol{X}) + f_d(\boldsymbol{X}), \tag{1-1}$$

in which for each $1 \leq i \leq d$, $f_i \in \mathbb{Z}[X_1, \dots, X_n]$ is a form with $\deg f_i = m \cdot e \cdot i$. We are interested in counting

$$N(F, B) := \left|\{\boldsymbol{x} \in [-B, B]^n \cap \mathbb{Z}^n : \exists y \in \mathbb{Z} \text{ such that } F(y, \boldsymbol{x}) = 0\}\right|.$$

Trivially, $N(F, B) \ll B^n$; our main result proves a nontrivial upper bound. We assume in what follows that $f_d \not\equiv 0$, since otherwise $(0, \boldsymbol{X})$ is a solution to $F(Y, \boldsymbol{X}) = 0$ for all $\boldsymbol{X}$, and then $B^n \ll N(F, B) \ll B^n$. (Throughout, we use the convention that $A \ll_\kappa B$ if there exists a constant C, possibly depending on κ, such that $|A| \leq CB$.)

Theorem 1.1. *Fix $n \geq 3$. Fix integers $m \geq 2$ and $e, d \geq 1$. Let F be defined as in (1-1), with $f_d \not\equiv 0$. Suppose the weighted hypersurface $V(F(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \dots, 1)$ defined by $F(Y, \boldsymbol{X}) = 0$ is nonsingular over $\mathbb{C}$. Then*

$$N(F, B) \ll B^{n-1+\frac{1}{n+1}} (\log B)^{\frac{n}{n+1}}.$$

The implicit constant may depend on n, m, d, e, but is otherwise independent of F.

MSC2020: 11D45, 11D85, 11N36.
Keywords: thin sets, polynomial sieve.

© 2024 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

The main progress achieved in Theorem 1.1 is for $n \geq 4$, $e \geq 2, d \geq 2$. The requirement that $n \geq 3$ occurs since a key step, Proposition 5.2, is not true for $n = 2$ (see Remark 5.4). In any case, for $n = 2, 3$ the result of Theorem 1.1 is superseded by results of Broberg in [5], as described below in (1-14) and (1-15). When $e = 1$, the variety $V(F(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \ldots, 1)$ is unweighted, so that in the setting of Theorem 1.1, to bound $N(F, B)$ it is equivalent to count points $[Y : X_1 : \cdots : X_n]$ with $|Y|, |X_i| \ll B$ on a nonsingular projective hypersurface of degree at least 2 in $\mathbb{P}^n$. Then the result of Theorem 1.1 (in the stronger form $N(F, B) \ll_{m,d,n,\varepsilon} B^{n-1+\varepsilon}$) has already been obtained by work of Heath-Brown and Browning, appearing in [6; 9; 10; 26; 27], as summarized by Salberger in [42]. Finally, when $d = 1$, the result of Theorem 1.1 (aside from uniformity in the coefficients of F) follows from recent work of the first author in [2] (see Remark 3.2).

The condition $m \geq 2$ is applied in two ways: first, in the construction of certain sieve weights (see Section 1.2 and the proof of Lemma 1.2), and second, in Section 3.3 when we pass from the weighted variety to an unweighted variety. For illustration, we also describe how an alternative approach to the sieve lemma, conditional on GRH, can be devised when $m = 1$ (see Section 3.2 and Remark 1.3).

Bounding $N(F, B)$ relates to a question of Serre on counting integral points in thin sets. Let $\mathcal{V}$ denote the affine variety

$$\mathcal{V} = \{(Y, \boldsymbol{X}) \in \mathbb{A}^{n+1} : F(Y, \boldsymbol{X}) = 0\}, \tag{1-2}$$

and consider the projection

$$\pi : \mathcal{V} \to \mathbb{A}^n, \quad (y, \boldsymbol{x}) \mapsto \boldsymbol{x}. \tag{1-3}$$

Under the hypotheses of Theorem 1.1, the set $Z = \pi(\mathcal{V}(\mathbb{Q}))$ is a *thin set of type II* in $\mathbb{A}^n_{\mathbb{Q}}$, in the nomenclature of Serre. Serre has posed a general question that can be interpreted in our present setting as asking whether it is possible to prove that

$$N(F, B) \ll B^{n-1} (\log B)^c \tag{1-4}$$

for some c. Previous work by Broberg [5] nearly settled Serre's conjecture for thin sets of type II in $\mathbb{P}^{n-1}$ for $n = 2, 3$; see (1-14) and (1-15) below. For $n \geq 4$, Theorem 1.1 represents new progress toward resolving Serre's question for certain thin sets of type II. Note that as $n \to \infty$, the bound in Theorem 1.1 approaches a bound of the strength (1-4). We provide general background on Serre's question, and state precisely how Theorem 1.1 relates to previous literature on this question, in Section 1.1 and Section 1.2.

To prove Theorem 1.1, we develop an appropriate polynomial sieve lemma, and then bound each contribution to the sieve using analytic, algebraic, and geometric ideas. A novel feature of this work is that we do not assume that $F(Y, \boldsymbol{X})$ exhibits separation of variables: that is, when $d \geq 2$, $F(Y, \boldsymbol{X})$ of the form (1-1) cannot in general be written as $F(Y, \boldsymbol{X}) = g(Y) - G(\boldsymbol{X})$ for polynomials g, G. A formal polynomial sieve lemma has been formulated previously in a level of generality that does not require separation of variables; see [8; 13]. However, in those works it has so far only been applied to count points on a variety that does exhibit separation of variables. To our knowledge, Theorem 1.1 is the first application of a polynomial sieve to produce an upper bound for $N(F, B)$ in a case without separation of

variables. We state precisely how Theorem 1.1 relates to previous literature on so-called square, power, and polynomial sieves in Section 1.2.

A second strength of Theorem 1.1 is that the exponent in the upper bound for $N(F, B)$ is independent of e, where we recall that as a function of $\boldsymbol{X}$, F has highest degree $m \cdot e \cdot d$. For any given $\boldsymbol{x} \in [-B, B]^n$ such that $F(Y, \boldsymbol{x}) = 0$ is solvable, one observes that any solution y to $F(y, \boldsymbol{x}) = 0$ must satisfy $y \ll B^e$, and there can be at most md solutions y for the given $\boldsymbol{x}$ (or, equivalently, preimages under the projection π in (1-3)), since the coefficient of Y^{md} in $F(Y, \boldsymbol{X})$ is nonzero. Thus an alternative method to bound $N(F, B)$ (up to an implicit constant depending on md) would be to count all $(n+1)$-tuples $\{(y, \boldsymbol{x}) : y \ll B^e, x_i \ll B : F(y, \boldsymbol{x}) = 0\}$. Other potential methods might be sensitive to the role of e or size of d, m (see for example Remark 1.4), while in contrast both the method and the result of Theorem 1.1 do not depend on e (aside from a possible implicit constant).

Third, we note that the result of Theorem 1.1 is independent of the coefficients of F; the implicit constant depends only on F in terms of its degree. To accomplish this, we adapt a strategy of [27], also recently applied in a similar setting in [3], to show that either $N(F, B)$ is already acceptably small, or $\|F\| \ll B^{(mde)^{n+2}}$. In the latter case, we then show that any dependence on $\|F\|$ in the sieve method is at most logarithmic, which we show is allowable for the result in Theorem 1.1.

1.1. ***Context of Theorem 1.1 within the study of Serre's question on thin sets.*** Here we recall the notion of thin sets defined by Serre in [46, §9.1 p. 121] and [45, p. 19]. Let k be a field of characteristic zero and let V be an irreducible algebraic variety in $\mathbb{P}^n_k$ (respectively $\mathbb{A}^n_k$). A subset M of $V(k)$ is said to be a projective (respectively, affine) thin set of type I if there is a closed subset $W \subset V$, $W \neq V$, with $M \subset W(k)$ (i.e., M is not Zariski dense in V). A subset M of $V(k)$ is said to be a projective (respectively, affine) thin set of type II if there is an irreducible projective (respectively, affine) algebraic variety X with $\dim X = \dim V$, and a generically surjective morphism $\pi : X \to V$ of degree $d \geq 2$ with $M \subset \pi(X(k))$. Any thin set is a finite union of thin sets of type I and thin sets of type II. From now on we consider only $k = \mathbb{Q}$, although Serre's treatment considers any number field.

Given a thin set $M \subset \mathbb{A}^n_{\mathbb{Q}}$, define the counting function

$$M(B) := \left|\{\boldsymbol{x} \in M \cap \mathbb{Z}^n : \max_{1 \leq i \leq n} |x_i| \leq B\}\right|,$$

so that trivially $M(B) \ll B^n$ for all $B \geq 1$. A theorem of Cohen [16] (see also [46, Chapter 13, Theorem 1, p. 177]) shows that

$$M(B) \ll_M B^{n-1/2}(\log B)^{\gamma} \qquad \text{for some } \gamma < 1, \tag{1-5}$$

where $\ll_M$ denotes that the implicit constant can depend on the coefficients of the equations defining M. As Serre remarks, this bound is essentially optimal, since the thin set

$$M = \{\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n : x_1 \text{ is a square}\} \tag{1-6}$$

has $M(B) \gg B^{n-1/2}$. However, this M arises from a morphism that is singular; it is reasonable to

expect that the result can be improved under an appropriate nonsingularity assumption (such as in the setting of Theorem 1.1).

Now let $M \subset \mathbb{P}^{n-1}_{\mathbb{Q}}$ be a thin set in projective space. Define the height function $H(x)$ for $x = [x_1 : \cdots : x_n] \in \mathbb{P}^{n-1}_{\mathbb{Q}}$ such that $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ and $\gcd(x_1, \ldots, x_n) = 1$ by $H(x) = \max_{1 \le i \le n} |x_i|$. Define the associated counting function

$$M_H(B) = \{x \in M(\mathbb{Q}) : H(x) \le B\}$$

so that trivially $M_H(B) \ll B^n$. Serre deduces in [46, Chapter 13, Theorem 3] from an application of (1-5) that

$$M_H(B) \ll_M B^{n-1/2}(\log B)^{\gamma} \qquad \text{for some } \gamma < 1. \tag{1-7}$$

Serre raises a general question in [46, p. 178]: is it possible to prove that

$$M_H(B) \ll B^{n-1}(\log B)^{c} \tag{1-8}$$

for some c? (The set (1-6) is not an example of a thin set here because if $M = \{[x_1^2 : x_2 : \cdots : x_n]\} \subset \mathbb{P}^{n-1}_{\mathbb{Q}}$ then for any $x_1 \neq 0$,

$$[x_1 : x_2 : \cdots : x_n] = x_1[x_1 : x_2 : \cdots : x_n] = [x_1^2 : x_1x_2 : \cdots : x_1x_n] \in M,$$

so that $M \supset \mathbb{P}^{n-1}_{\mathbb{Q}}$.)

1.1.1. *Results for thin sets of type I.* If Z is an irreducible projective variety in $\mathbb{P}^{n-1}_{\mathbb{Q}}$ of degree $d \ge 2$, Serre deduces from (1-7) that $Z_H(B) \ll_Z B^{\dim Z + 1/2}(\log B)^{\gamma}$ for some $\gamma < 1$. Serre asks if it is possible to prove that $Z_H(B) \ll_Z B^{\dim Z}(\log B)^{c}$ for some c. (This question is raised in both [46, p. 178] and [45, p. 27]. Serre provides an example of a quadric for which a logarithmic factor necessarily arises. See also the question in the case of a hypersurface in Heath-Brown [24, p. 227], formally stated in both nonuniform and uniform versions as [27, Conjectures 1 and 2].) This is now called the *dimension growth conjecture* (in the terminology of [7]), and is often described as the statement that

$$Z_H(B) \ll_{Z,\varepsilon} B^{\dim Z + \varepsilon} \qquad \text{for every } \varepsilon > 0. \tag{1-9}$$

A refined version, credited to Heath-Brown and known as the *uniform dimension growth conjecture*, is the statement that

$$Z_H(B) \ll_{n,\deg Z,\varepsilon} B^{\dim Z + \varepsilon} \qquad \text{for every } \varepsilon > 0. \tag{1-10}$$

In the case that $Z \subset \mathbb{P}^{n-1}_{\mathbb{Q}}$ is a nonsingular projective hypersurface of degree $d \ge 2$, as mentioned before, combined works of Browning and Heath-Brown have proved (1-10) for all $n \ge 3$. More generally, Browning, Heath-Brown and Salberger proved (1-10) for all geometrically integral varieties of degree $d = 2$ and $d \ge 6$ (see [27] and [12], respectively). Recent work of Salberger has proved (1-9) in all remaining cases, and has even proved the uniform version (1-10) for $d \ge 4$ [43]. See [14] for a helpful survey, statements of open questions, and new progress such as an explicit bound $Z_H(B) \le C d^E B^{\dim Z}$

when $\deg Z = d \geq 5$, for a certain $C = C(n)$ and $E = E(n)$. The resolution of the dimension growth conjecture means that attention now turns to thin sets of type II, the subject of the present article.

1.1.2. *Results for thin sets of type II.* We turn to the case of thin sets of type II, our present focus. Given a finite cover $\phi : X \to \mathbb{P}^{n-1}$ over $\mathbb{Q}$ with $n \geq 2$, X irreducible and ϕ of degree at least 2, set

$$N_B(\phi) = |\{P \in X(\mathbb{Q}) : H(\phi(P)) \leq B\}| \tag{1-11}$$

for the standard height function above. Serre's question asks whether

$$N_B(\phi) \ll_{\phi,n} B^{n-1}(\log B)^c \qquad \text{for some } c, \tag{1-12}$$

or in a uniform version,

$$N_B(\phi) \ll_{\deg\phi,n} B^{n-1}(\log B)^c \qquad \text{for some } c. \tag{1-13}$$

For $n = 2, 3$ work of Broberg via the determinant method proves cases of Serre's conjecture up to the logarithmic factor [5]. Precisely, for $\phi : X \to \mathbb{P}^1$ of degree $r \geq 2$, Broberg proves

$$N_B(\phi) \ll_{\phi,\varepsilon} B^{2/r+\varepsilon} \qquad \text{for any } \varepsilon > 0. \tag{1-14}$$

For $\phi : X \to \mathbb{P}^2$ of degree r, Broberg proves

$$N_B(\phi) \ll_{\phi,\varepsilon} B^{2+\varepsilon} \text{ for } r \geq 3, \qquad N_B(\phi) \ll_{\phi,\varepsilon} B^{9/4+\varepsilon} \text{ for } r = 2, \text{ for any } \varepsilon > 0. \tag{1-15}$$

For $n \geq 4$, the question remains open whether one can achieve $N_B(\phi) \ll B^{n-1+\varepsilon}$ for all $\varepsilon > 0$, although we record some progress on this for specific types of ϕ in Section 1.2.

Now recall the setting of Theorem 1.1 in this paper, and the affine variety $\mathcal{V} \subset \mathbb{A}^{n+1}$ defined in (1-2) according to the polynomial $F(Y, \boldsymbol{X})$. Under the hypotheses of Theorem 1.1, we have:

(i) The variety $\mathcal{V}$ is irreducible (see Remark 3.3).

(ii) The projection π has degree $dm > 1$ since $m \geq 2$.

Thus $Z = \pi(\mathcal{V}(\mathbb{Q}))$ is a thin set of type II in $\mathbb{A}^n_{\mathbb{Q}}$, and in particular Cohen's result (1-5) implies that

$$Z(B) = N(F, B) \ll_F B^{n-1/2}(\log B)^{\gamma}, \tag{1-16}$$

following the same reasoning as [46, Chapter 13, Theorem 2, p. 178]. Or, interpreting the setting of Theorem 1.1 as counting points on a finite cover ϕ of $\mathbb{P}^{n-1}$ as in (1-11), this shows

$$N_B(\phi) \ll N(F, B) \ll_{\phi} B^{n-1/2}(\log B)^{\gamma}.$$

Our new work, Theorem 1.1, improves on (1-16) for each $n \geq 3$, for F of the form (1-1) with $V(F(Y, \boldsymbol{X}))$ nonsingular, and approaches a uniform bound of the strength (1-13) as $n \to \infty$.

1.2. ***Context of Theorem 1.1 within sieve methods.*** We now recall a few recent developments of sieve methods in the context of counting solutions to Diophantine equations, with a particular focus on progress toward Serre's conjecture for type II sets, as described above.

1.2.1. *Square sieve.* Let $f(\boldsymbol{X}) \in \mathbb{Z}[X_1, \dots, X_n]$ be a fixed polynomial. Let $\mathcal{B}$ be a "box," such as $[-B, B]^n$ or more generally $\prod_i [-B_i, B_i]$. In [25], Heath-Brown codified the square sieve to count the number of integral values $\boldsymbol{x} \in \mathcal{B}$ such that $f(\boldsymbol{x}) = y^2$ is solvable over $\mathbb{Z}$, building on a method of Hooley [31]. At its heart was a formal sieve lemma involving a character sum with Legendre symbols. Heath-Brown applied this in particular to improve the error term in an asymptotic for the number of consecutive square-free numbers in a range. In [40], Pierce developed a stronger version of the square sieve, with a sieving set comprised of products of two primes rather than primes; this effectively allows the underlying modulus to be larger relative to the box $\mathcal{B}$, by factoring the modulus and using the q-analogue of van der Corput differencing. Pierce applied this to prove a nontrivial upper bound for 3-torsion in class groups of quadratic fields [40]; Heath-Brown subsequently used this sieve method to prove there are finitely many imaginary quadratic fields having class group of exponent 5 [28]; Bonolis and Browning applied it to prove a uniform bound for counting rational points on hyperelliptic fibrations [3].

1.2.2. *Power sieve.* The square sieve has been generalized to a power sieve, in order to count integral values $\boldsymbol{x} \in \mathcal{B}$ with $f(\boldsymbol{x}) = y^r$ solvable, for a fixed $r \geq 2$. Recall the question of bounding $N_{\boldsymbol{B}}(\phi)$ as in (1-12). For any $n \geq 2$, in the special case that ϕ is a nonsingular cyclic cover of degree $r \geq 2$, Munshi observed this can be reduced to counting the number of integral values $\boldsymbol{x} \in [-B, B]^n$ with $F(x_1, \dots, x_n) = y^r$ solvable, for a nonsingular form F of degree mr for some $m \geq 1$. To bound this, Munshi developed a formal sieve lemma involving a character sum in terms of multiplicative Dirichlet characters [39]. Munshi applied it to prove that

$$\left|\{\boldsymbol{x} \in [-B, B]^n : F(\boldsymbol{x}) = y^r \text{ is solvable over } \mathbb{Z}\}\right| \ll B^{n-1+\frac{1}{n}} (\log B)^{\frac{n-1}{n}} \tag{1-17}$$

Consequently, this proved $N_{\boldsymbol{B}}(\phi) \ll B^{n-1+\frac{1}{n}} (\log B)^{\frac{n-1}{n}}$ for nonsingular cyclic covers. (See [2, Remark 1] for a note on the history of this result; the exponents stated here are slightly different from those presented in [39].)

In [29] Heath-Brown and Pierce have strengthened the power sieve, by using a sieving set comprised of products of primes, generalizing the approach of [40]. They used this method to prove that for any polynomial $f(\boldsymbol{X}) \in \mathbb{Z}[X_1, \dots, X_n]$ of degree $d \geq 3$ with nonsingular leading form, and for any $r \geq 2$,

$$|\{\boldsymbol{x} \in [-B, B]^n : f(\boldsymbol{x}) = y^r \text{ is solvable over } \mathbb{Z}\}| \ll \begin{cases} B^{n-1+\frac{n(8-n)+4}{6n+4}} (\log B)^2, & 2 \leq n \leq 8, \\ B^{n-1+\frac{1}{2n+10}} (\log B)^2, & n = 9, \\ B^{n-1-\frac{n-10}{2n+10}} (\log B)^2, & n \geq 10. \end{cases} \tag{1-18}$$

This proves Serre's conjecture (1-12) for $N_{\boldsymbol{B}}(\phi)$, for all nonsingular cyclic covers, for $n \geq 10$. Indeed, the bound achieved is even smaller than the general conjecture, which is reasonable due to the imposed nonsingularity assumption.

Independently, Brandes also developed a power sieve in [4], applied to counting sums and differences of power-free numbers.

1.2.3. *Polynomial sieve: with separation of variables.* The next significant generalization addressed counting $\boldsymbol{x} \in \mathcal{B}$ for which $g(y) = f(\boldsymbol{x})$ is solvable, for appropriate polynomials g, f. Here, a quite general framework for a polynomial sieve lemma was developed by Browning in [8]. Specifically, in that work, Browning applied the polynomial sieve lemma to count x_1, x_2 such that $g(y) = f(x_1, x_2)$ is solvable, for particular functions f, g, that enabled an application showing the sparsity of like sums of a quartic polynomial of one variable.

Bonolis [2] further developed a polynomial sieve lemma with a character sum involving trace functions. Applying this, he proved that for any polynomial $g \in \mathbb{Z}[Y]$ of degree $r \geq 2$, and any irreducible form $F \in \mathbb{Z}[X_1, \dots, X_n]$ of degree $e \geq 2$ such that the projective hypersurface $V(F)$ defined by $F = 0$ is nonsingular over $\mathbb{C}$, then

$$\left|\{\boldsymbol{x} \in [-B, B]^n : F(\boldsymbol{x}) = g(y) \text{ is solvable over } \mathbb{Z}\}\right| \ll B^{n-1+\frac{1}{n+1}} (\log B)^{\frac{n}{n+1}}. \tag{1-19}$$

(This improves (1-17) and recovers the result initially stated in [39]; see [2, Remark 1].) This can also be seen as an improvement on Cohen's theorem (1-16) for a special type of thin set (defined as the image of $\mathcal{V} = \{(y, \boldsymbol{x}) \in \mathbb{A}^{n+1} : F(\boldsymbol{x}) - g(y) = 0\}$ under $(y, \boldsymbol{x}) \mapsto \boldsymbol{x}$, under the assumption that $V(F)$ defines a nonsingular projective hypersurface). The special case of our Theorem 1.1 when $d = 1$ follows from [2, Theorem 1.1]; see Remark 3.2.

Notably, the method employed in [2] to prove (1-19) was the first to demonstrate nontrivial averaging over pairs of primes in the sieving set, and exploiting such a strategy is central to the strength of our main theorem. We explain explicitly the advantage of such averaging in equations (1-25) and (1-26), below. For now, we simply state abstractly that any polynomial sieve method tests the solvability of the desired equation modulo p for primes in a chosen sieving set $\mathcal{P}$. The outcome of applying a sieve lemma (such as Lemma 1.2 below) is that one must bound from above an expression roughly of the form $|\mathcal{P}|^{-2} \sum_{p \neq q \in \mathcal{P}} T(p, q)$, where $T(p, q)$ studies the solvability of the desired equation modulo pairs $p \neq q \in \mathcal{P}$. Previous to [2], papers applying any type of polynomial sieve produced an upper bound for $|T(p, q)|$ that was uniform over p, q and then summed trivially over $p \neq q \in \mathcal{P}$. Instead, averaging nontrivially over p, q exploits the fact that $T(p, q)$ is typically smaller than its worst (largest) upper bound.

Most recently, a geometric generalization of Browning's polynomial sieve lemma has been developed over function fields by Bucur, Cojocaru, Lalín and the second author in [13]. They pose an analogue of Serre's question (1-8) in that setting (also raised by Browning and Vishe [11]), and apply a polynomial sieve to prove a bound of analogous strength to (1-19), in the special case of nonsingular cyclic covers in a function field setting. It remains an interesting open question to achieve a stronger bound such as (1-18), or to prove results for finite covers that are noncyclic, in such a function field setting.

1.2.4. *Polynomial sieve: without separation of variables.* So far we have mentioned applications of a sieve lemma to count solutions to $G(Y, \boldsymbol{X}) = 0$ when G separates variables as $G(Y, \boldsymbol{X}) = g(Y) - f(\boldsymbol{X})$ for some polynomials g, f. More generally, it is reasonable to ask — and this is a motivation for the

present paper — whether an appropriate polynomial sieve can be employed to count solutions to equations of the form $G(Y, \boldsymbol{X}) = 0$ where $G(Y, \boldsymbol{X}) \in \mathbb{Z}[Y, X_1, \dots, X_n]$ is a polynomial of degree D of the form

$$G(Y, \boldsymbol{X}) = Y^D + Y^{D-1} f_1(\boldsymbol{X}) + \cdots + Y f_{D-1}(\boldsymbol{X}) + f_D(\boldsymbol{X}), \tag{1-20}$$

where each f_i is a form of degree $i \cdot e$, and we assume that the weighted hypersurface $V(G(Y, \boldsymbol{X}))) \subset \mathbb{P}(e, 1, \dots, 1)$ defined by $G(Y, \boldsymbol{X}) = 0$ is nonsingular. Define

$$N(G, B) := \left|\{\boldsymbol{x} \in [-B, B]^n : \exists y \in \mathbb{Z} \text{ such that } G(y, \boldsymbol{x}) = 0\}\right|.$$

Under the assumption $f_D \not\equiv 0$, the aim is to improve on the trivial bound $N(G, B) \ll B^n$. To be clear, the formal sieve lemmas appearing in [8; 13] include this level of generality, but have only been applied to prove a bound for $N(G, B)$ when separation of variables occurs. In this paper we accomplish the first application of the polynomial sieve without assuming separation of variables, but under the additional assumption that the degree D of $G(Y, \boldsymbol{X})$ defined in (1-20) factors as $D = md$ for some $m \geq 2$, and all powers of Y that appear are divisible by m. (To see why this restriction is useful, see the proof of Lemma 1.2; for an alternative approach when $m = 1$, conditional on GRH, see Remark 1.3 and Section 3.2.)

The strength of our approach hinges on a particular formulation of the polynomial sieve, given in Lemma 1.2. It is worthwhile to compare our formulation with the polynomial sieve presented in [8, Theorem 1.1]. In [8, Theorem 1.1], the sieve weight system, adapted to counting solutions to (1-20), is defined as follows:

$$w_{p,\mathrm{Bro}}(\boldsymbol{k}) = \alpha + (\nu_p(\boldsymbol{k}) - 1)(D - \nu_p(\boldsymbol{k})),$$

in which $\nu_p(\boldsymbol{k}) = |\{y \in \mathbb{F}_p : G(y, \boldsymbol{k}) = 0 \in \mathbb{F}_p\}|$. (These weights are then applied in an inequality analogous to (3-1) below, to derive a sieve lemma.) Consequently, if $G(Y, \boldsymbol{k}) = 0$ is solvable over $\mathbb{Z}$, the conditions $1 \leq \nu_p(\boldsymbol{k}) \leq D$ and $\alpha > 0$ guarantee that $w_{p,\mathrm{Bro}}(\boldsymbol{k}) > 0$ *for any* p. In our approach, we consider simpler weights:

$$w_p(\boldsymbol{k}) = \nu_p(\boldsymbol{k}) - 1.$$

Thus, in our situation, if $G(Y, \boldsymbol{k}) = 0$ is solvable over $\mathbb{Z}$, we can only conclude that $w_p(\boldsymbol{k}) \geq 0$. However, it is still possible to establish that $w_p(\boldsymbol{k}) > 0$ for *a positive proportion of primes*, which suffices for our application. (Precisely, we obtain $\omega_p(\boldsymbol{k}) > 0$ for those $p \equiv 1 \pmod{m}$ where $m \geq 2$; see (3-2) in the proof of Lemma 1.2.)

The simplicity of our weight system turns out to be crucial for bounding the terms that appear in the polynomial sieve lemma. In the setting of the polynomial $F(Y, \boldsymbol{X})$ as in (1-1), our main task will be to prove square root cancellation for the sum

$$\sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(z^e, \boldsymbol{a}) = 0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u} \rangle),$$

for generic $\boldsymbol{a} \in \mathbb{F}_p^n$, which can be accomplished by exploiting the smoothness of the variety $V(F(Z^e, \boldsymbol{X}))$. On the other hand, if we were to adopt [8, Theorem 1.1], the presence of the factor $(\nu_p(\boldsymbol{k}))^2$ would lead to the exponential sum

$$\sum_{\substack{(z_1,z_2,\boldsymbol{a})\in\mathbb{F}_p^{n+2}\\ F(z_1^e,\boldsymbol{a})=0\\ F(z_2^e,\boldsymbol{a})=0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle),$$

which is more challenging to handle, due to the highly singular nature of the variety $V(F(Z_1^e, \boldsymbol{X})) \cap V(F(Z_2^e, \boldsymbol{X}))$.

1.3. ***Overview of the method.*** We now provide an overview of our method, highlighting four key aspects of our strategy. To prove a nontrivial upper bound for $N(F, B)$ via a sieve, we introduce a smooth nonnegative function $W : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ defined by $W(\boldsymbol{x}) = w(\boldsymbol{x}/B)$, where w is an infinitely differentiable, compactly supported function that is $\equiv 1$ on $[-1, 1]^n$, and supported in $[-2, 2]^n$. Define the smoothed counting function

$$\mathcal{S}(F, B) := \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^n\\ F(y,\boldsymbol{k})=0 \text{ solvable}}} W(\boldsymbol{k}), \tag{1-21}$$

which sums over $\boldsymbol{k} \in \mathbb{Z}^n$ such that there exists $y \in \mathbb{Z}$ with $F(y, \boldsymbol{k}) = 0$. By construction

$$N(F, B) \leq \mathcal{S}(F, B),$$

and we may focus on proving a nontrivial upper bound for $\mathcal{S}(F, B)$. We employ the following sieve lemma, which we prove in Section 3.1. Here and throughout, given a polynomial f, we let $\|f\|$ denote the maximum absolute value of any coefficient of f.

Lemma 1.2 (polynomial sieve lemma). *Let $e, d \geq 1$ and $m \geq 2$ be integers. Consider the polynomial*

$$F(Y, \boldsymbol{X}) = Y^{md} + Y^{m(d-1)} f_1(\boldsymbol{X}) + \cdots + Y^m f_{d-1}(\boldsymbol{X}) + f_d(\boldsymbol{X}),$$

under the assumption that $f_d \not\equiv 0$, and that $\deg f_i = m \cdot e \cdot i$ *for each $1 \leq i \leq d$.*

Let $B \geq 1$ and define a smooth weight W supported in $[-2B, 2B]^n$ and $\equiv 1$ on $[-B, B]^n$, as above. Let $\mathcal{P} \subset \{p \equiv 1 \bmod m\}$ be a finite set of primes $p \in [Q, 2Q]$, with cardinality P. Suppose that $Q = B^\kappa$ for some fixed $0 < \kappa \leq 1$ and that $P \gg Q/\log Q$. Suppose also that

$$P \gg_{m,e,d} \max\{\log \|f_d\|, \log B\}. \tag{1-22}$$

For each $\boldsymbol{k} \in \mathbb{Z}^n$ and $p \in \mathcal{P}$ define

$$\nu_p(\boldsymbol{k}) = |\{y \in \mathbb{F}_p : F(y, \boldsymbol{k}) = 0 \ (mod\ p)\}|.$$

Then

$$\mathcal{S}(F, B) \ll_{m,e,d} \sum_{\boldsymbol{k}: f_d(\boldsymbol{k})=0} W(\boldsymbol{k}) + \frac{1}{P}\sum_{\boldsymbol{k}} W(\boldsymbol{k}) + \frac{1}{P^2}\sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \left|\sum_{\boldsymbol{k}} W(\boldsymbol{k})(\nu_p(\boldsymbol{k})-1)(\nu_q(\boldsymbol{k})-1)\right|.$$

Remark 1.3. We observe that the same lemma holds for $m = 1$, conditional on GRH, with (1-22) replaced by $Q \gg_{m,e,d} \max\{(\log \|F\|)^{\alpha_0}, (\log B)^{\alpha_0}\}$ for some $\alpha_0 > 2$. For the sake of illustration, we demonstrate this in Section 3.2, although we do not apply such a conditional result in this paper.

We now point out four key aspects of our method for applying this sieve lemma to prove Theorem 1.1. First, for all $\boldsymbol{k}$ and for all primes p, $\nu_p(\boldsymbol{k}) \le md$; this is because Y^{md} has coefficient 1 in $F(Y, \boldsymbol{X})$, so that for all values of $\boldsymbol{k}$, $F(Y, \boldsymbol{k})$ is of degree md as a polynomial in Y. On the other hand, in the proof of the lemma, we use the assumption that each prime in the sieving set has $p \equiv 1 \pmod{m}$ in order to provide a lower bound $\nu_p(\boldsymbol{k}) - 1 \ge m - 1 > 0$ for many $\boldsymbol{k}$, motivating our requirement that $m \ge 2$. This is the first novelty of our method for dealing with a case in which the variables $Y, \boldsymbol{X}$ are not "separated."

For each pair of primes $p \ne q \in \mathcal{P}$, the sieve lemma leads us to study

$$T(p, q) := \sum_{\boldsymbol{k} \in \mathbb{Z}^n} W(\boldsymbol{k})(\nu_p(\boldsymbol{k}) - 1)(\nu_q(\boldsymbol{k}) - 1). \tag{1-23}$$

After an application of the Poisson summation formula, we see that

$$T(p, q) = \left(\frac{1}{pq}\right)^n \sum_{\boldsymbol{u} \in \mathbb{Z}^n} \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq),$$

where

$$g(\boldsymbol{u}, pq) := \sum_{\boldsymbol{a} \pmod{pq}} (\nu_p(\boldsymbol{a}) - 1)(\nu_q(\boldsymbol{a}) - 1) e_{pq}(\langle \boldsymbol{a}, \boldsymbol{u} \rangle). \tag{1-24}$$

Here we write each coordinate of $\boldsymbol{a}$ in terms of its residue class modulo pq, and $e_{pq}(t) = e^{2\pi i t/pq}$. After showing that $g(\boldsymbol{u}, pq)$ satisfies a multiplicativity relation, we can focus on the case of prime modulus, and study

$$g(\boldsymbol{u}, p) := \sum_{\boldsymbol{a} \in \mathbb{F}_p^n} (\nu_p(\boldsymbol{a}) - 1) e_p(\langle \boldsymbol{a}, \boldsymbol{u} \rangle).$$

We show that the main task to bound $g(\boldsymbol{u}, p)$ is to bound the exponential sum

$$\sum_{\substack{(y, \boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(y, \boldsymbol{a}) = 0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u} \rangle).$$

Here we highlight a second aspect: the fact that the polynomial $F(Y, \boldsymbol{X})$ is not homogeneous motivates a more sophisticated approach to bounding this sum (see Remark 4.6). Given a polynomial H, let $V(H)$ denote the corresponding variety $\{H = 0\}$, and let $\langle \boldsymbol{X}, \boldsymbol{U} \rangle = \sum_i X_i U_i$. Roughly speaking, for each prime p we divide $\boldsymbol{u} \in \mathbb{Z}^n$ into three cases: a *type zero* case when $\boldsymbol{u} \equiv 0 \pmod{p}$, a *good* case when $V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle)$ is not tangent to $V(F(Y, \boldsymbol{X}))$ over $\overline{\mathbb{F}}_p$, and finally a *bad* case in which $V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle)$ is tangent to $V(F(Y, \boldsymbol{X}))$ over $\overline{\mathbb{F}}_p$. (More precisely, we reformulate this in terms of varieties in unweighted projective space.) In the type zero case, we can only show that $g(\boldsymbol{0}, p) \ll p^{n-1/2}$, but such cases are sparse. In the

remaining two cases, we apply a version of the Weil bound to $g(\boldsymbol{u}, p)$, obtaining $g(\boldsymbol{u}, p) \ll p^{n/2}$ if $\boldsymbol{u}$ is good and $g(\boldsymbol{u}, p) \ll p^{n/2+1/2}$ if $\boldsymbol{u}$ is bad (Proposition 4.2).

A third crucial aspect arises when we assemble this information efficiently inside the third term on the right-hand side of the sieve lemma, namely

$$\frac{1}{P^2} \sum_{p \neq q \in \mathcal{P}} |T(p,q)| \ll \frac{1}{P^2 Q^{2n}} \sum_{p \neq q \in \mathcal{P}} \sum_{\boldsymbol{u} \in \mathbb{Z}^n} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq) \right|. \tag{1-25}$$

In many earlier applications of the power sieve or polynomial sieve to count solutions to Diophantine equations, the strategy has been to bound $|T(p,q)|$ uniformly over $p \neq q$ and simply sum trivially over $p \neq q$. However, recent work of the first author demonstrated how to take advantage of nontrivial averaging over the sum of $p \neq q \in \mathcal{P}$; see [2]. In this paper, we also average nontrivially over $p \neq q$ and this contributes to the strength of our main theorem.

In order to average nontrivially over $p \neq q \in \mathcal{P}$, we quantify the fact that there cannot be many triples $\boldsymbol{u}, p, q$ for which $\boldsymbol{u}$ is simultaneously bad for both p and q. Roughly speaking, we characterize the dual variety of the original hypersurface $V(F(Y, \boldsymbol{X}))$ according to an irreducible polynomial $G(U_Y, U_1, \dots, U_n)$, and observe that $G(0, \boldsymbol{u}) \neq 0$ precisely when the hyperplane $V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle)$ is not tangent to $V(F(Y, \boldsymbol{X}))$ over $\mathbb{C}$. Then we reverse the order of summation in the right-hand side of (1-25), writing it as

$$\frac{1}{P^2 Q^{2n}} \sum_{\boldsymbol{u} \in \mathbb{Z}^n} \sum_{p \neq q \in \mathcal{P}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq) \right|. \tag{1-26}$$

The sum over $\boldsymbol{u}$ can be split into case (a) where $G(0, \boldsymbol{u}) \neq 0$ and case (b) where $G(0, \boldsymbol{u}) = 0$. In case (a), we show $\boldsymbol{u}$ is bad modulo p and q only if p and q divide the (nonzero) value of a certain resultant polynomial; thus there can only be very few such p, q.

A fourth key aspect arises in case (b), for which $\boldsymbol{u}$ is bad for all primes (since the value of the resultant is zero). To compensate, we show that there are not too many $\boldsymbol{u}$ for which $G(0, \boldsymbol{u}) = 0$. This step is one of the significant novelties of the paper. It requires understanding not the variety $V(G(U_Y, \boldsymbol{U}))$ but $V(G(U_Y, \boldsymbol{U})) \cap V(U_Y)$, the intersection with the hyperplane $U_Y = 0$. To tackle this, we show that any polynomial divisor of $G(0, \boldsymbol{U})$ has degree at least 2 (Proposition 5.2), so that we can apply strong bounds of Heath-Brown [27] and Pila [41] to count solutions to $G(0, \boldsymbol{u}) = 0$ (see (5-18)). To prove the key result in Proposition 5.2, we employ a geometric argument to show that given a nonsingular projective hypersurface X and a projective line ℓ not contained in X, the generic hyperplane containing ℓ is not tangent to X. This statement, proved in Section 6 via a strategy suggested by Per Salberger, is critical to the method and the ultimate strength of Theorem 1.1.

Remark 1.4. It would be interesting to consider bounding $N(F, B)$, in the setting of Theorem 1.1, by other methods. As mentioned earlier, one approach is to count all $(n+1)$-tuples $\{(y, \boldsymbol{x}) \in \mathbb{Z}^{n+1} : y \ll B^e, x_i \ll B : F(y, \boldsymbol{x}) = 0\}$, for example, by applying the determinant method. Since the range of y depends on e, such a direct approach is likely to produce a bound for $N(F, B)$ with an exponent depending on e. Alternatively, one could fix $x_2, \dots, x_n$ (with $\approx B^{n-1}$ such choices) and consider the resulting

equation as a projective curve in variables y, x_1. Supposing that the resulting curve is generically of degree dme, an application of Bombieri-Pila [1] could count (y, x_1) in the square $[-B^e, B^e]^2$. This could ultimately lead to a total bound of the form $N(F, B) \ll B^{n-1} \cdot B^{e/dme+\varepsilon} = B^{n-1+1/dm+\varepsilon}$. This putative outcome appears independent of e, but the method has overcounted x_1 in the range B^e; nevertheless, such an approach could be advantageous for large d, m.

1.4. ***Notation.*** We use $e_q(t) = e^{2\pi i t/q}$. We denote $\boldsymbol{X} = (X_1, \dots, X_n)$, $\boldsymbol{U} = (U_1, \dots, U_n)$. Moreover, for two vectors $\boldsymbol{s} = (s_1, \dots s_n)$, $\boldsymbol{t} = (t_1, \dots, t_n)$, we define $\langle \boldsymbol{s}, \boldsymbol{t} \rangle = \sum_{i=1}^n s_i t_i$. We let $\|F\|$ denote the absolute value of the maximum coefficient in a polynomial $F \in \mathbb{Z}[X_1, \dots, X_n]$; similarly $\|\boldsymbol{X}\| = \max_{1 \le i \le n} |X_i|$ for $\boldsymbol{X} \in \mathbb{Z}^n$.

2. Reduction to remove dependence on $\|F\|$

Recall that Theorem 1.1 states that the upper bound for $N(F, B)$ is only dependent on the degree of F, and not on the coefficients of F. In fact, the sieve methods we apply prove an upper bound for $N(F, B)$ that *can* depend on $\|F\|$. In this section we show by alternative methods that we may assume that $\|F\| \ll B^{(mde)^{n+2}}$. The method does not rely on assuming $m \ge 2$ in (1-1), and so without any additional trouble we may work more generally in the setting of (1-20).

Lemma 2.1. *Let $V(G(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \dots, 1)$ be defined by*

$$G(Y, \boldsymbol{X}) = Y^D + Y^{D-1} f_1(\boldsymbol{X}) + \cdots + Y f_{D-1}(\boldsymbol{X}) + f_D(\boldsymbol{X})$$

with each f_i a form of $\deg f_i = i \cdot e$, *for fixed $D, e \ge 1$ and $n \ge 1$. Assume that $f_D \not\equiv 0$ and the weighted hypersurface $V(G(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \dots, 1)$ is absolutely irreducible. Then either*

$$\|G\| \ll B^{(De)^{n+2}},$$

or $N(G, B) \ll_{n,D,e} B^{n-1}$.

Remark 2.2. Under the hypotheses of Theorem 1.1, for F as in (1-1), $V(F(Y, \boldsymbol{X}))$ is absolutely irreducible (following similar reasoning to Remark 3.3). As a result of this lemma, we can obtain the bound claimed in Theorem 1.1 as long as all later dependence on $\|F\|$ is at most logarithmic in $\|F\|$, which we track as the argument proceeds.

Proof. The method of proof follows [27, Theorem 4], or the recent similar result [3, Lemma 2.1]. Fix $n, D, e \ge 1$. We start by considering the set of monomials

$$\mathcal{E} := \left\{ Y^{d_Y} X_1^{d_1} \cdots X_n^{d_n} : \ d_Y e + \sum_{i=1}^n d_i = De \right\},$$

in which the degrees $d_Y, d_1, \dots, d_n$ vary over all nonnegative integers satisfying $d_Y e + \sum d_i = De$. It is easy to see that $|\mathcal{E}| \le (De)^{n+1}$.

Let $B \geq 1$ be fixed. Let $\boldsymbol{v}$ denote coordinates $(y, x_1, \ldots, x_n)$ and let $\{\boldsymbol{v}_1, \ldots \boldsymbol{v}_N\}$ enumerate the set of points that are solutions to $G(Y, \boldsymbol{X}) = 0$, with each of the last n coordinates of $\boldsymbol{v}_j$ lying in $[-B, B]$. Note that these count each $\boldsymbol{X} \in [-B, B]^n$ for which $G(Y, \boldsymbol{X})$ is solvable at least once, so that $N(G, B) \leq N \leq D \cdot N(G, B)$. (For the upper bound, we recall that the coefficient of Y^D in $G(Y, \boldsymbol{X})$ is nonzero, so that any given $\boldsymbol{X}$ can correspond to at most D such Y.) Then, we construct the $N \times |\mathcal{E}|$ matrix

$$\boldsymbol{C} = (\boldsymbol{v}_i^{\boldsymbol{e}})_{\substack{1 \leq i \leq N \\ \boldsymbol{e} \in \mathcal{E}}}.$$

Notice that rank $\boldsymbol{C} \leq |\mathcal{E}| - 1$, since the vector $\boldsymbol{a} \in \mathbb{Z}^{|\mathcal{E}|} \setminus \{0\}$ whose entries correspond to the coefficients of $G(Y, \boldsymbol{X})$ is such that $\boldsymbol{C}\boldsymbol{a} = \boldsymbol{0}$. Moreover, $\boldsymbol{a}$ is primitive since the coefficient associated to Y^D is 1. Now the strategy is to find another nonzero vector $\boldsymbol{b}$ in the nullspace of $\boldsymbol{C}$ and show that if $\boldsymbol{b}$ is in the span of $\boldsymbol{a}$ then $\|G\|$ is small, and if $\boldsymbol{b}$ is not in the span of $\boldsymbol{a}$ then we have an improved count for $N(G, B)$. We may assume henceforward that $|\mathcal{E}| \leq N$, since otherwise we already have the upper bound $N(G, B) \leq N \leq |\mathcal{E}| \leq (De)^{n+1}$, which suffices for the lemma.

If rank $\boldsymbol{C} \leq |\mathcal{E}| - 2$, then the nullspace has dimension at least 2, and we can take $\boldsymbol{b} \in \mathbb{Z}^{|\mathcal{E}|}$ to be any element in the nullspace that is not in the span of $\boldsymbol{a}$. Let $H(Y, \boldsymbol{X})$ be the polynomial defined by the coefficients corresponding to the vector $\boldsymbol{b}$ and consider the polynomial $R(\boldsymbol{X}) = \mathrm{Res}(G(Y, \boldsymbol{X}), H(Y, \boldsymbol{X}))$, which is a polynomial in $\boldsymbol{X}$ of degree $\ll_{D,e,n} 1$. (See, e.g., [21, Ch 12], which we apply to take the resultant of two polynomials in the variable Y, whose coefficients are determined by $\boldsymbol{X}$.) We claim that $R(\boldsymbol{X}) \not\equiv 0$: indeed, if $R(\boldsymbol{X}) \equiv 0$, then G and H would share an irreducible component. Since $G(Y, \boldsymbol{X}) = 0$ is irreducible, and $\deg H \leq De = \deg G$, it would follow that G is a constant multiple of H, but this is not possible since we are assuming that $\boldsymbol{a}$ and $\boldsymbol{b}$ are not proportional. Thus $R(\boldsymbol{X}) \not\equiv 0$. Moreover, observe that for any $\boldsymbol{x} \in \mathbb{Z}^n$

$$R(\boldsymbol{x}) = 0 \iff G(Y, \boldsymbol{x}) \text{ and } H(Y, \boldsymbol{x}) \text{ have a common root.}$$

Note that any $\boldsymbol{x}$ such that $G(y, \boldsymbol{x}) = 0$ is solvable contributes at least one row to the matrix $\boldsymbol{C}$; each such row also corresponds to a solution to $H(y, \boldsymbol{x}) = 0$. Thus it follows that

$$\begin{aligned} N(G, B) &= |\{\boldsymbol{x} \in [-B, B]^n : \exists y \in \mathbb{Z} \text{ such that } G(y, \boldsymbol{x}) = H(y, \boldsymbol{x}) = 0\}| \\ &\leq |\{\boldsymbol{x} \in [-B, B]^n : \ R(\boldsymbol{x}) = 0\}| \\ &\ll_{n,D,e,} B^{n-1}, \end{aligned}$$

with an implicit constant independent of the coefficients of R, via an application of a trivial counting bound for the nonzero polynomial R. (This bound is sometimes called the Schwartz-Zippel bound, and a proof can be found in [27, Theorem 1]; we remark that although in that context the polynomial under consideration is absolutely irreducible, the method of proof only requires that it is not identically zero.)

The remaining case is when rank $\boldsymbol{C} = |\mathcal{E}| - 1$, so that all $|\mathcal{E}| \times |\mathcal{E}|$ minors vanish, but at least one $(|\mathcal{E}| - 1) \times (|\mathcal{E}| - 1)$ minor does not; we claim there is a nonzero $\boldsymbol{b} \in \mathbb{Z}^{|\mathcal{E}|}$ in the nullspace of $\boldsymbol{C}$ such that

$|\boldsymbol{b}| = O(B^{De|\mathcal{E}|}) = O(B^{(De)^{n+2}})$. If so, then since $\boldsymbol{a}$ is primitive (and $\boldsymbol{b}$ must be proportional to $\boldsymbol{a}$) it follows that $|\boldsymbol{a}| \le |\boldsymbol{b}| \ll B^{(De)^{n+2}}$. This shows that $\|G\| \ll B^{(De)^{n+2}}$ as claimed.

An appropriate $\boldsymbol{b}$ can be constructed with entries that are $(|\mathcal{E}|-1)\times(|\mathcal{E}|-1)$ minors, so that the size estimate $|\boldsymbol{b}| = O(B^{De|\mathcal{E}|})$ follows from the fact that each entry of $\boldsymbol{C}$ is $O(B^{De})$. For completeness, we sketch this construction. Without loss of generality, we can let $\boldsymbol{C}'$ denote the top $|\mathcal{E}|\times|\mathcal{E}|$ submatrix in $\boldsymbol{C}$, and assume that the minor $\boldsymbol{C}'_{1,1}$ (obtained by omitting the first row and first column of $\boldsymbol{C}'$) is nonzero. Define a vector $\boldsymbol{b}$ as follows: for each $1 \le j \le |\mathcal{E}|$, define the entry b_j to be the $(1, j)$-th cofactor of $\mathbf{C}'$; in particular $b_1 \ne 0$ so $\boldsymbol{b}$ is nonzero, and $|\boldsymbol{b}| = O(B^{De(|\mathcal{E}|-1)}) = O(B^{De|\mathcal{E}|})$. We now show that $\boldsymbol{b}$ is in the nullspace of $\boldsymbol{C}$. Let $\boldsymbol{r}_i$ denote the i-th row of $\boldsymbol{C}$; then for each $1 \le i \le N$,

$$\boldsymbol{r}_i \cdot \boldsymbol{b} = \det \begin{pmatrix} \boldsymbol{r}_i \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_{|\mathcal{E}|} \end{pmatrix} = 0. \tag{2-1}$$

Indeed, for $i = 1$ or $i > |\mathcal{E}|$, up to sign, $\boldsymbol{r}_i \cdot \boldsymbol{b}$ is an $|\mathcal{E}|\times|\mathcal{E}|$ minor of $\boldsymbol{C}$, and all such minors vanish since $\operatorname{rank}\boldsymbol{C} < |\mathcal{E}|$. For $2 \le i \le |\mathcal{E}|$, the matrix (2-1) has two identical rows. Thus $\boldsymbol{C}\boldsymbol{b} = \boldsymbol{0}$. □

3. Preliminaries on the sieve lemma

In this section we gather together two preliminary steps: first, we prove the sieve inequality in Lemma 1.2; for $m = 1$ we provide an alternative proof, conditional on GRH. Second, we formulate an equivalent nonsingularity condition in unweighted projective space. We also make preliminary remarks on the sieving set.

3.1. ***Proof of the polynomial sieve lemma.*** To prove Lemma 1.2, observe that

$$\mathcal{S}(F, B) = \sum_{\boldsymbol{k}: f_d(\boldsymbol{k})=0} W(\boldsymbol{k}) + \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^n: \\ f_d(\boldsymbol{k})\ne 0 \\ F(y,\boldsymbol{k})=0 \text{ solvable}}} W(\boldsymbol{k}),$$

since within the first term, $y = 0$ is always a solution to $F(y, \boldsymbol{k}) = 0$. We consider the weighted sum

$$\sum_{\boldsymbol{k}: f_d(\boldsymbol{k})\ne 0} W(\boldsymbol{k}) \Bigg(\sum_{p\in\mathcal{P}} (\nu_p(\boldsymbol{k}) - 1) \Bigg)^2. \tag{3-1}$$

Fix $\boldsymbol{k}$ such that $f_d(\boldsymbol{k}) \ne 0$ and the polynomial $F(Y, \boldsymbol{k})$ is solvable over $\mathbb{Z}$, so that there exists $y_0 \in \mathbb{Z}$ such that $F(y_0, \boldsymbol{k}) = 0$. For any $p \in \mathcal{P}$ such that $p \nmid f_d(\boldsymbol{k})$, then $y_0 \not\equiv 0 \bmod p$. Then since $p \equiv 1 \bmod m$, and due to the structure of F in (1-1), we have that $\{y_0, \gamma_p y_0, \dots, \gamma_p^{m-1} y_0\}$ are distinct solutions of $F(Y, \boldsymbol{k}) \equiv 0 \pmod p$, where $\gamma_p^m \equiv 1 \bmod p$ and γ_p is a primitive m-th root of unity in $\mathbb{F}_p$. In particular, for such p, $\nu_p(\boldsymbol{k}) \ge m$. Consequently, for each $\boldsymbol{k}$ such that $f_d(\boldsymbol{k}) \ne 0$ and $F(Y, \boldsymbol{k})$ is solvable, we have

that

$$\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1)\geq(m-1)\sum_{p\in\mathcal{P},p\nmid f_d(\boldsymbol{k})}1\gg_m P-\sum_{p\in\mathcal{P},p\mid f_d(\boldsymbol{k})}1\geq(1/2)P, \tag{3-2}$$

as long as $P\gg_{m,e,d}\max\{\log\|f_d\|,\log B\}$. The last step follows since the number $\omega(f_d(\boldsymbol{k}))$ of distinct prime divisors of $f_d(\boldsymbol{k})\neq 0$ is at most

$$\begin{aligned}\omega(f_d(\boldsymbol{k}))&\ll\log(f_d(\boldsymbol{k}))/\log\log(f_d(\boldsymbol{k}))\\&\ll\log(\|f_d\|B^{dem})\\&\ll_{m,e,d}\log\|f_d\|+\log B.\end{aligned}$$

Thus the last inequality in (3-2) holds as long as

$$P\gg_{m,e,d}\max\{\log\|f_d\|,\log B\}, \tag{3-3}$$

leading to the corresponding hypothesis in the lemma.

From (3-2) and the nonnegativity of the weight W, we see that

$$P^2\sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^n:\\ f_d(\boldsymbol{k})\neq 0\\ F(y,\boldsymbol{k})=0\text{ solvable}}}W(\boldsymbol{k})\ll\sum_{\boldsymbol{k}:f_d(\boldsymbol{k})\neq 0}W(\boldsymbol{k})\Big(\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1)\Big)^2\leq\sum_{\boldsymbol{k}}W(\boldsymbol{k})\Big(\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1)\Big)^2.$$

Opening the square on the right-hand side, the contribution from $p=q\in\mathcal{P}$ is

$$\sum_{p\in\mathcal{P}}\sum_{\boldsymbol{k}}W(\boldsymbol{k})(\nu_p(\boldsymbol{k})-1)^2\ll_{m,d}P\sum_{\boldsymbol{k}}W(\boldsymbol{k}),$$

since $\nu_p(\boldsymbol{k})\leq md$ for all $\boldsymbol{k}$, as previously mentioned. The contribution from $p\neq q\in\mathcal{P}$ is bounded in absolute value by

$$\sum_{p\neq q\in\mathcal{P}}\Big|\sum_{\boldsymbol{k}}W(\boldsymbol{k})(\nu_p(\boldsymbol{k})-1)(\nu_q(\boldsymbol{k})-1)\Big|.$$

Assembling all these terms, we see that Lemma 1.2 is proved.

Remark 3.1. When we apply Lemma 1.2 to prove Theorem 1.1, we can assume that $\|f_d\|\leq\|F\|\ll B^{(mde)^{n+2}}$, by Lemma 2.1. This will allow us to verify that (3-3) holds for our choice of sieving set, as we will verify in Section 7 when we choose Q in (7-4).

3.2. ***Alternative proof when m = 1, conditional on GRH.*** Recall from Section 1.2.4 the general problem of counting $\boldsymbol{x}\in[-B,B]^n$ such that $G(y,\boldsymbol{x})=0$ is solvable in $\mathbb{Z}$, with $G(Y,\boldsymbol{X})$ of degree D as in (1-20). In our main work in this paper, we assume that $D=md$ with $m\geq 2$ and that G is a polynomial in Y^m. This additional structure allowed us to choose a sieving set $\mathcal{P}\subset[Q,2Q]$ of primes $p\equiv 1\pmod m$, so that all the m-th roots of unity are present in $\mathbb{F}_p$, for each $p\in\mathcal{P}$. With this property, we could define sieve weights that exhibit an appropriate lower bound in the form (3-2) for most $\boldsymbol{k}$ in the support of $W(\boldsymbol{k})$ and a positive proportion of primes.

Nevertheless, we can proceed by a different argument to develop a sieve lemma to bound the number of $\boldsymbol{x} \in [-B, B]^n$ such that $G(y, \boldsymbol{x}) = 0$ is solvable over $\mathbb{Z}$, with no condition on the degree D; that is, to prove a version of Lemma 1.2 in the case $m = 1$. As a first step, we naturally try to introduce a system of weights, according to a fixed set of primes. Let us take $\mathcal{P} = \{Q \le p \le 2Q : p \text{ prime}\}$ for some parameter Q to be chosen optimally with respect to B. In particular, by the prime number theorem, $|\mathcal{P}| \gg Q(\log Q)^{-1}$ for all $Q \gg 1$. Fix $\boldsymbol{k} \in \mathbb{Z}^n$. For each prime $p \in \mathcal{P}$, set

$$\nu_p(\boldsymbol{k}) = |\{y \in \mathbb{F}_p : G(y, \boldsymbol{k}) = 0 \pmod p\}|.$$

Since $G(y, \boldsymbol{k})$ contains the term y^D, it is not the zero polynomial in y, and $\nu_p(\boldsymbol{k}) \le D$. Consider, as in the proof of Lemma 1.2 above, the weighted sum

$$\sum_{\boldsymbol{k} : f_D(\boldsymbol{k}) \neq 0} W(\boldsymbol{k}) \Big(\sum_{p \in \mathcal{P}} (\nu_p(\boldsymbol{k}) - 1) \Big)^2. \tag{3-4}$$

In order to deduce a sieve lemma, we need a lower bound for the arithmetic weight (the squared term), for those $\boldsymbol{k}$ for which $f_D(\boldsymbol{k}) \neq 0$ and $G(Y, \boldsymbol{k}) = 0$ is solvable over $\mathbb{Z}$.

Here is one approach. Let $\boldsymbol{k}$ be fixed, with $f_D(\boldsymbol{k}) \neq 0$ and $G(Y, \boldsymbol{k}) = 0$ solvable over $\mathbb{Z}$, and $\boldsymbol{k}$ in the support of W. Then $G(Y, \boldsymbol{k}) = (Y - y_0)\tilde{g}_{\boldsymbol{k}}(Y)$ for some $y_0 \in \mathbb{Z} \setminus \{0\}$ and some (monic) $\tilde{g}_{\boldsymbol{k}}(Y) \in \mathbb{Z}[Y]$ of degree $D - 1$. For each such $\boldsymbol{k}$, we can obtain a suitable lower bound for the arithmetic weight in (3-4) as long as for a positive proportion of $p \in \mathcal{P}$, $\tilde{g}_{\boldsymbol{k}}$ has a root over $\mathbb{F}_p$. Let $g_{\boldsymbol{k}}$ be an irreducible factor of $\tilde{g}_{\boldsymbol{k}}$. Let $F_{\boldsymbol{k}}$ denote the splitting field of $g_{\boldsymbol{k}}$ over $\mathbb{Q}$, say $F_{\boldsymbol{k}} = \mathbb{Q}(\alpha_{\boldsymbol{k}})$. Since $g_{\boldsymbol{k}}$ is irreducible, then it is the minimal polynomial of $\alpha_{\boldsymbol{k}}$ in $\mathbb{Z}[Y]$, and it is separable (since we are working over characteristic zero), and the splitting field is Galois over $\mathbb{Q}$. By Dedekind's theorem, for all $p \nmid [\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]$, $g_{\boldsymbol{k}}$ splits completely over $\mathbb{F}_p$ precisely when $(p) = p\mathcal{O}_{F_{\boldsymbol{k}}}$ splits completely in $F_{\boldsymbol{k}}$; see, e.g., [37, Theorem 27, p. 79]. Then

$$\sum_{p \in \mathcal{P}} (\nu_p(\boldsymbol{k}) - 1) = \sum_{p \in \mathcal{P}} \big|\{y \in \mathbb{F}_p : \tilde{g}_{\boldsymbol{k}}(y) = 0\}\big| \ge \sum_{p \in \mathcal{P}} \big|\{y \in \mathbb{F}_p : g_{\boldsymbol{k}}(y) = 0\}\big|.$$

If $g_{\boldsymbol{k}}$ is linear in $\mathbb{Z}[Y]$, this sum is of size $|\mathcal{P}|$, which suffices. If $\deg g_{\boldsymbol{k}} \ge 2$, we continue to argue that

$$\begin{aligned} \sum_{p \in \mathcal{P}} (\nu_p(\boldsymbol{k}) - 1) &\ge \deg(g_{\boldsymbol{k}}) \big|\{p \in \mathcal{P} : g_{\boldsymbol{k}}(Y) \text{ completely split over } \mathbb{F}_p\}\big| \\ &\ge \big|\{p \in \mathcal{P} : p\mathcal{O}_{F_{\boldsymbol{k}}} \text{ splits completely in } F_{\boldsymbol{k}}\}\big| - \big|\{p \in \mathcal{P} : p \mid [\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]\}\big|. \end{aligned} \tag{3-5}$$

Let

$$\pi_{\boldsymbol{k}}(Q) = \big|\{p \le Q : p\mathcal{O}_{F_{\boldsymbol{k}}} \text{ splits completely in } F_{\boldsymbol{k}}\}\big|$$

and $N(\boldsymbol{k}) = |\{p \mid [\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]\}|$. The Chebotarev density theorem, in the unconditional form of [34, Theorem 1.3], shows that

$$\left| \pi_{\boldsymbol{k}}(Q) - \frac{1}{|G_{\boldsymbol{k}}|} \frac{Q}{\log Q} \right| = \frac{1}{|G_{\boldsymbol{k}}|} \frac{Q^{\beta_0}}{\log Q^{\beta_0}} + O_{D,A}(Q(\log Q)^{-A}) \tag{3-6}$$

for every $A \geq 2$, as long as $Q \geq \exp(10 \deg F_{\boldsymbol{k}} (\log |D(F_{\boldsymbol{k}})|)^2)$. Here $G_{\boldsymbol{k}}$ is the Galois group $\mathrm{Gal}(F_{\boldsymbol{k}}/\mathbb{Q})$, $D(F_{\boldsymbol{k}})$ is the discriminant of the splitting field $F_{\boldsymbol{k}}/\mathbb{Q}$, and $\deg F_{\boldsymbol{k}} = \deg |F_{\boldsymbol{k}}/\mathbb{Q}|$ is the degree of the extension. The implicit constant in the error term depends only on A and $\deg F_{\boldsymbol{k}} = |G_{\boldsymbol{k}}| \leq (D-1)!$. The real number $1/2 < \beta_0 < 1$, if it exists, is the (real, simple) exceptional zero of the associated Dedekind zeta function $\zeta_{F_{\boldsymbol{k}}}$; if no exceptional zero exists, that term does not appear in the result.

In particular, under the assumption of GRH for $\zeta_{F_{\boldsymbol{k}}}$, Lagarias and Odlyzko's Theorem 1.1 in [34] (in the refined form of Serre [44, Theorem 4]) shows that for any $Q > 2$, the entire right-hand side of (3-6) may be replaced by

$$O(|G_{\boldsymbol{k}}|^{-1} Q^{1/2} \log(|D(F_{\boldsymbol{k}})| Q^{\deg F_{\boldsymbol{k}}})) = O_D(Q^{1/2} \log Q) + O_D(Q^{1/2} \log |D(F_{\boldsymbol{k}})|),$$

in which the implied constant is absolute and effectively computable. There exists a constant $Q_0(D)$ depending only on D such that the first term is $\leq \frac{1}{4} \frac{1}{(D-1)!} Q(\log Q)^{-1}$ for all $Q \geq Q_0(D)$. The second term is also $\leq \frac{1}{4} \frac{1}{(D-1)!} Q(\log Q)^{-1}$ if for example $Q \geq Q_1(D)(\log D(F_{\boldsymbol{k}}))^{\alpha_0}$ for a constant $Q_1(D)$ and some fixed $\alpha_0 > 2$. This shows that under GRH, for all $Q \gg_D (\log D(F_{\boldsymbol{k}}))^{\alpha_0}$ some fixed $\alpha_0 > 2$,

$$\pi_{\boldsymbol{k}}(Q) - \pi_{\boldsymbol{k}}(Q/2) \gg_D Q/\log Q \gg_D |\mathcal{P}|. \tag{3-7}$$

Two tasks remain in order to complete a lower bound for (3-5): (i) to bound $D(F_{\boldsymbol{k}})$ from above, so that the lower bound $Q \gg_D (\log D(F_{\boldsymbol{k}}))^{\alpha_0}$ can be made uniform over $\boldsymbol{k}$, and (ii) to count

$$N(\boldsymbol{k}) = |\{p | [\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]\}| = \omega([\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]) \ll \log[\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]/\log\log[\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]].$$

We note the relation

$$D(F_{\boldsymbol{k}})[\mathcal{O}_{F_{\boldsymbol{k}}} : \mathbb{Z}[\alpha_{\boldsymbol{k}}]]^2 = \mathrm{Disc}(g_{\boldsymbol{k}}), \tag{3-8}$$

which holds by [38, Remark 2.25 and equation (8) on p. 38]. (Since $g_{\boldsymbol{k}}$ was assumed to be irreducible and we are in characteristic zero, then $g_{\boldsymbol{k}}$ is separable and $\mathrm{Disc}(g_{\boldsymbol{k}}) \neq 0$.) Thus for both remaining tasks, it suffices to bound $\mathrm{Disc}(g_{\boldsymbol{k}})$ from above, since by (3-8) both

$$N(\boldsymbol{k}) \ll \log \mathrm{Disc}\,(g_{\boldsymbol{k}}), \qquad \log D(F_{\boldsymbol{k}}) \leq \log \mathrm{Disc}\,(g_{\boldsymbol{k}}).$$

Now $\mathrm{Disc}\,(g_{\boldsymbol{k}})$ (the resultant of $g_{\boldsymbol{k}}(Y)$ and $g'_{\boldsymbol{k}}(Y)$, as defined in [21, Chapter 13, Proposition 1.1]) is a polynomial in the coefficients of $g_{\boldsymbol{k}}$ with degree bounded in terms of D. The coefficients of $g_{\boldsymbol{k}}$ are polynomials in $\boldsymbol{k}$ and the coefficients of $G(Y, \boldsymbol{X})$ with degree at most D. Since we only consider $\boldsymbol{k}$ in the support of W, $|\boldsymbol{k}| \ll B$, and the coefficients of $g_{\boldsymbol{k}}$ are $\ll \|G\| B^D$. Thus

$$\log \mathrm{Disc}\,(g_{\boldsymbol{k}}) \ll_D \log \|G\| + \log B.$$

In combination with (3-7), we can conclude in (3-5) that for some constant C_D,

$$\sum_{p \in \mathcal{P}} (\nu_p(\boldsymbol{k}) - 1) \gg_D Q/\log Q - C_D(\log \|G\| + \log B),$$

for all $Q \geq C_D' \max\{(\log \|G\|)^{\alpha_0}, (\log B)^{\alpha_0}\}$ for some $\alpha_0 > 2$. By taking C_D' sufficiently large, we achieve $\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1) \gg |\mathcal{P}| = P$. This shows that, conditional on GRH,

$$P^2 \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^n:\\ f_D(\boldsymbol{k})\neq 0\\ G(y,\boldsymbol{k})=0 \text{ solvable}}} W(\boldsymbol{k}) \ll \sum_{\boldsymbol{k}: f_D(\boldsymbol{k})\neq 0} W(\boldsymbol{k})\Big(\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1)\Big)^2 \leq \sum_{\boldsymbol{k}} W(\boldsymbol{k})\Big(\sum_{p\in\mathcal{P}}(\nu_p(\boldsymbol{k})-1)\Big)^2.$$

From here, the remainder of the proof used above for Lemma 1.2 can be repeated, and this completes the proof of the claim in Remark 1.3.

3.3. ***Associated variety in unweighted projective space.*** It is a hypothesis of Theorem 1.1 that the weighted hypersurface $V(F(Y,\boldsymbol{X})) \subset \mathbb{P}(e,1,\dots,1)$, defined by $F(Y,\boldsymbol{X})=0$, is nonsingular over $\mathbb{C}$. It is convenient to relate $V(F(Y,\boldsymbol{X}))$ to a variety in unweighted projective space. We claim that for

$$F(Y,\boldsymbol{X}) = Y^{dm} + Y^{(d-1)m} f_1(\boldsymbol{X}) + \cdots + f_d(\boldsymbol{X}),$$

then $V(F(Y,\boldsymbol{X})) \subset \mathbb{P}(e,1,\dots,1)$ is nonsingular if and only if $V(F(Z^e,\boldsymbol{X})) \subset \mathbb{P}^n$ is nonsingular. Here, we again apply the assumption $m \geq 2$. Indeed the weighted projective variety is nonsingular if and only if the only solution of

$$\begin{cases} F(Y,\boldsymbol{X}) = 0, \\ \dfrac{\partial F}{\partial Y}(Y,\boldsymbol{X}) = \sum\limits_{i=0}^{d-1} f_i(\boldsymbol{X})\cdot m(d-i)Y^{m(d-i)-1} = 0, \\ \dfrac{\partial F}{\partial X_j}(Y,\boldsymbol{X}) = 0, \quad j=1,\dots,n, \end{cases} \tag{3-9}$$

on $\mathbb{A}^{n+1}$ is the point $P=\mathbf{0}$. (By convention we set $f_0(\boldsymbol{X})=1$.) Similarly, the projective variety $V(F(Z^e,\boldsymbol{X}))$ is nonsingular if and only if the only solution of

$$\begin{cases} F(Z^e,\boldsymbol{X}) = 0, \\ \dfrac{\partial F}{\partial Z}(Z^e,\boldsymbol{X}) = \sum\limits_{i=0}^{d-1} f_i(\boldsymbol{X})\cdot me(d-i)Z^{em(d-i)-1} = 0, \\ \dfrac{\partial F}{\partial X_j}(Z^e,\boldsymbol{X}) = 0, \quad j=1,\dots,n, \end{cases} \tag{3-10}$$

on $\mathbb{A}^{n+1}$ is the point $P=\mathbf{0}$. Moreover, note that

$$\begin{aligned} \frac{\partial F}{\partial Y}(Y,\boldsymbol{X}) &= mY^{m-1}\sum_{i=0}^{d-1} f_i(\boldsymbol{X})(d-i)Y^{m(d-i-1)}, \\ \frac{\partial F}{\partial Z}(Z^e,\boldsymbol{X}) &= emZ^{em-1}\sum_{i=0}^{d-1} f_i(\boldsymbol{X})(d-i)Z^{em(d-i-1)}. \end{aligned} \tag{3-11}$$

We will momentarily use this to confirm that if $m \geq 2$, a nonzero solution (say $P = (y, \boldsymbol{x}) \in \mathbb{A}^{n+1}$) to (3-9) exists if and only if a solution (namely $Q = (y^{1/e}, \boldsymbol{x}) \in \mathbb{A}^{n+1}$) to (3-10) exists.

To clarify the role of the assumption $m \geq 2$, let us briefly make a general observation. In general, let a polynomial $G(Y, \boldsymbol{X})$ be given as in (1-20) and assume $V(G(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \ldots, 1)$ is nonsingular; we may assume $e \geq 2$ (since otherwise the variety is already unweighted). Then we claim $V(G(Z^e, \boldsymbol{X}))$ is nonsingular (as a projective variety) if and only if $V(G(Y, \boldsymbol{X})) \cap V(Y)$ is nonsingular (as a weighted projective variety). By the chain rule,

$$\frac{\partial G}{\partial Z}(Z^e, \boldsymbol{X}) = eZ^{e-1}\left(\frac{\partial G}{\partial Y}\right)(Z^e, \boldsymbol{X}).$$

Observe that

$$\begin{aligned}\operatorname{Sing}(V(G(Z^e, \boldsymbol{X}))) &= \{(z, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{Z,\boldsymbol{X}} G(z^e, \boldsymbol{x}) = \boldsymbol{0}\} \\ &= \{(0, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{\boldsymbol{X}} G(0, \boldsymbol{x}) = \boldsymbol{0}\} \cup \{(z, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{Y,\boldsymbol{X}} G(z^e, \boldsymbol{x}) = \boldsymbol{0}\} \\ &= \{(0, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{\boldsymbol{X}} G(0, \boldsymbol{x}) = \boldsymbol{0}\} \cup \varnothing \end{aligned} \tag{3-12}$$

under the assumption that $V(G(Y, \boldsymbol{X}))$ is nonsingular. On the other hand, by the Jacobian criterion,

$$\operatorname{Sing}(V(G(Y, \boldsymbol{X})) \cap V(Y)) = \{(0, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{\boldsymbol{X}} G(0, \boldsymbol{x}) = \boldsymbol{0}\}.$$

(Here we have used that $G(0, \boldsymbol{X})$ is itself homogeneous in $\boldsymbol{X}$, so that $\nabla_{\boldsymbol{X}} G(0, \boldsymbol{X}) = 0$ implies $G(0, \boldsymbol{X}) = 0$ by Euler's identity.) Since the singular sets are identical, this proves the claim.

Let us apply this in our case with G taken to be the polynomial $F(Y, \boldsymbol{X})$, with $V(F(Y, \boldsymbol{X}))$ assumed to be nonsingular. We consider whether there are any $(0, \boldsymbol{x}) \in \mathbb{P}^n$ such that $\nabla_{\boldsymbol{X}} F(0, \boldsymbol{x}) = 0$. Supposing such $(0, \boldsymbol{x})$ exists, it must be the case that $(\partial F/\partial Y)(0, \boldsymbol{x}) \neq 0$, since otherwise $(0, \boldsymbol{x})$ would be a singular point on $V(F(Y, \boldsymbol{X}))$. If $m \geq 2$, then due to the leading factor Y^{m-1} in (3-11), any point $(0, \boldsymbol{x}) \in \mathbb{P}^n$ must lead to $(\partial F/\partial Y)(0, \boldsymbol{x}) = 0$. Consequently there can be no such $(0, \boldsymbol{x})$, and $\operatorname{Sing}(V(F(Y, \boldsymbol{X})) \cap V(Y))$ must be empty. Hence by the general argument above, so is $\operatorname{Sing}(V(F(Z^e, \boldsymbol{X})))$. In conclusion, if $m \geq 2$, $V(F(Y, \boldsymbol{X}))$ being nonsingular implies $V(F(Z^e, \boldsymbol{X}))$ is nonsingular.

However if $m = 1$, there is no leading factor of Y in (3-11), and indeed at $(0, \boldsymbol{x})$, (3-11) evaluates to $f_{d-1}(\boldsymbol{x})$. Thus points $(0, \boldsymbol{x})$ for which $f_{d-1}(\boldsymbol{x}) \neq 0$ and $\nabla_{\boldsymbol{X}} F(0, \boldsymbol{x}) = 0$ can lead to singular points on $V(F(Y, \boldsymbol{X})) \cap V(Y)$ and hence to singular points on $F(F(Z^e, \boldsymbol{X}))$. (Nevertheless, there cannot be too many singular points, as we will observe in (4-1) below that the singular locus has at most dimension 0.)

In the other direction, suppose that $V(F(Z^e, \boldsymbol{X}))$ is nonsingular, so that as computed in (3-12),

$$\operatorname{Sing}(V(F(Z^e, \boldsymbol{X}))) = \{(0, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{\boldsymbol{X}} F(0, \boldsymbol{x}) = \boldsymbol{0}\} \cup \{(z, \boldsymbol{x}) \in \mathbb{P}^n : \nabla_{Y,\boldsymbol{X}} F(z^e, \boldsymbol{x}) = \boldsymbol{0}\}$$

is empty. If there were a point $(y, \boldsymbol{x})$ in $\operatorname{Sing}(V(Y, \boldsymbol{X}))$ then if $y = 0$ this would produce an element in the first set on the right-hand side, while if $y \neq 0$ then taking $z = y^{1/e}$ (working over $\mathbb{C}$) would produce a point in the second set on the right-hand side. Thus $V(F(Y, \boldsymbol{X}))$ must be nonsingular (and here we did not need to apply $m \geq 2$).

Remark 3.2. In the special case that $d = 1$, then $F(Y, \boldsymbol{X}) = Y^m + f_1(\boldsymbol{X})$. Thus $V(F(Y, \boldsymbol{X})) \subset \mathbb{P}(e, 1, \dots, 1)$ is nonsingular if and only if $V(Z^{em} + f_1(\boldsymbol{X})) \subset \mathbb{P}^n$ is nonsingular, with $f_1 \not\equiv 0$ homogeneous of degree em. This occurs if and only if $V(f_1(\boldsymbol{X})) \subset \mathbb{P}^{n-1}$ is nonsingular; in this special case, the problem we consider falls in the scope of the work in [2, Theorem 1.1], which proves this case of Theorem 1.1. Our method of proof works regardless, so we allow $d = 1$ as we continue.

Remark 3.3. Recall the affine hypersurface $\mathcal{V} \subset \mathbb{A}^{n+1}_{\mathbb{C}}$ defined in (1-2) according to the polynomial $F(Y, \boldsymbol{X})$. We note that $\mathcal{V}$ is irreducible under the conditions of Theorem 1.1. Suppose it is reducible, so that $F(Y, \boldsymbol{X}) = G(Y, \boldsymbol{X})H(Y, \boldsymbol{X})$ for some nonconstant polynomials. Then $F(Z^e, \boldsymbol{X}) = G(Z^e, \boldsymbol{X})H(Z^e, \boldsymbol{X})$ so that the projective variety $V(F(Z^e, \boldsymbol{X}))$ is reducible. Consequently, by [13, Lemma 11.1], $V(F(Z^e, \boldsymbol{X}))$ is singular, which is a contradiction because by the discussion above, $V(F(Y, \boldsymbol{X}))$ is nonsingular if and only if $V(F(Z^e, \boldsymbol{X}))$ is nonsingular.

3.4. ***Initial considerations of the sieving set.*** We suppose that $Q = B^\kappa$ for some $0 < \kappa \le 1$ to be chosen later (see (7-4)). We will choose a sieving set

$$\mathcal{P} \subset [Q, 2Q]$$

comprised of primes with certain properties. In the special case that $(e, m) = 1$, it is sensible to restrict our attention to a set $\mathcal{P}_0$ of primes in $[Q, 2Q]$ such that

(i) $p \equiv 1 \pmod{m}$ (recalling $m \ge 2$) and

(ii) $p \equiv 2 \bmod e$, and

(iii) the reduction of $V(F(Y, \boldsymbol{X}))$ as a weighted variety over $\overline{\mathbb{F}}_p$ is nonsingular.

The first criterion (i) we have used in the proof of the sieve lemma (Lemma 1.2). The second criterion (ii) ensures that $(e, p-1) = 1$ so that every $y \in \mathbb{F}_p$ satisfies $y = z^e$ for some $z \in \mathbb{F}_p$. Then for each $p \in \mathcal{P}$, we can simply consider the reduction $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\mathbb{F}_p}$ in place of the weighted variety, so that (iii) is equivalent to

(iii′) the reduction of $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is nonsingular.

By the Chinese remainder theorem and the Siegel–Walfisz theorem on primes in arithmetic progressions, under the assumption that $(e, m) = 1$, there are $\gg_{m,e} Q/\log Q$ primes that satisfy (i) and (ii) in any dyadic region $[Q, 2Q]$, for all Q sufficiently large. We could then choose the sieving set $\mathcal{P}_0$ to be the subset of such primes for which (iii′) holds; the remaining task is to show there are sufficiently few primes that violate (iii′).

Recall from Section 3.3 that $V(F(Y, \boldsymbol{X}))$ is nonsingular over $\mathbb{C}$ (as a weighted projective variety) if and only if $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n$ is nonsingular over $\mathbb{C}$. Thus under the hypothesis of Theorem 1.1, the latter is nonsingular, and consequently there are no nontrivial simultaneous solutions of the system (3-10), and thus the resultant

$$r := \operatorname{Res}\left(F, \frac{\partial F}{\partial Z}, \frac{\partial F}{\partial X_1}, \dots, \frac{\partial F}{\partial X_n}\right)$$

of those $n+2$ polynomials in $n+1$ variables is a nonzero integer. Moreover, by [21, Chapter 13, Proposition 1.1], r is a polynomial in the coefficients of F with degree bounded in terms of m, e, d. By [15, Section IV], the reduction $V_p(F(Z^e, \boldsymbol{X}))$ of $V(F(Z^e, \boldsymbol{X}))$ modulo p is singular precisely when $p|r$, which can only occur for at most $\omega(r)$ primes, where

$$\omega(r) \ll \log r / \log\log r \ll_{m,e,d} \log \|F\|. \tag{3-13}$$

(Notice that the argument in this paragraph made no assumption on the relative primality of e and m.)

In particular, if $(e, m) = 1$, then as long as Q is sufficiently large, say $Q \gg_{m,e,d} (\log \|F\|)^{1+\delta_0}$ for any fixed $\delta_0 > 0$ or even $Q \gg_{m,e,d} (\log \|F\|)(\log\log \|F\|)$, we can conclude that $|\mathcal{P}_0| \gg_{m,e,d} Q/\log Q$. After we choose Q to be a certain power of B (see (7-4)), this will only require a lower bound on B that is on the order of a power of $\log \|F\|$, which we will see can be accommodated by the bound on the right-hand side of our claim in Theorem 1.1.

These remarks all apply in the case that $(e, m) = 1$. However, we can also argue more generally without this assumption, as we demonstrate in the next section, by working not with $V(F(Z^e, \boldsymbol{X}))$ as above, but with a finite collection of varieties W_i, defined according to $F(\gamma^i z^e, \boldsymbol{X}) = 0$ in $\mathbb{F}_p$, for a certain primitive root $\gamma \in \mathbb{F}_p^\times$ (see Lemma 4.3). Thus we postpone our definition of the sieving set, in general, until the end of the next section.

4. Estimates for exponential sums

In this section we apply the Weil bound to prove an upper bound for the exponential sum $g(\boldsymbol{u}, p)$ (see (1-24)) in the case that $\boldsymbol{u}$ is each of three types: type zero, good, or bad modulo p (Definition 4.1). At the end, in Section 4.2 we then define the sieving set $\mathcal{P}$.

We note the multiplicativity condition

$$g(\boldsymbol{u}, pq) := \sum_{\boldsymbol{a} \bmod pq} (\nu_p(\boldsymbol{a}) - 1)(\nu_q(\boldsymbol{a}) - 1) e_{pq}(\langle \boldsymbol{a}, \boldsymbol{u} \rangle) = g(\bar{q}\boldsymbol{u}, p) g(\bar{p}\boldsymbol{u}, q),$$

where $q\bar{q} \equiv 1 \bmod p$, and $p\bar{p} \equiv 1 \bmod q$. This leads us to study the key exponential sums with prime modulus:

$$g(\boldsymbol{u}, p) := \sum_{\boldsymbol{a} \in \mathbb{F}_p^n} (\nu_p(\boldsymbol{a}) - 1) e_p(\langle \boldsymbol{a}, \boldsymbol{u} \rangle).$$

Let p be a fixed prime of good reduction for $F(Z^e, \boldsymbol{X})$, so that $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is a nonsingular projective hypersurface. For any point $P \in V(F(Z^e, \boldsymbol{X}))$, let $T_P \subseteq \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ denote the projective tangent space to $V(F(Z^e, \boldsymbol{X}))$ at P. A linear space L is tangent to $V(F(Z^e, \boldsymbol{X}))$ at P if $T_P \subseteq L$; if L is a hyperplane, this is equivalent to P being a singular point of $V(F(Z^e, \boldsymbol{X})) \cap L$ (see [20, p. 57]).

Given $\boldsymbol{u} \in \mathbb{Z}^n$ with $\boldsymbol{u} \not\equiv \boldsymbol{0} \pmod p$, if $V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is not tangent to $V(F(Z^e, \boldsymbol{X}))$ at any point (i.e., they intersect transversely), we simply say $V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle)$ is not tangent to $V(F(Z^e, \boldsymbol{X}))$; otherwise, we will say they are tangent (and as we will discuss below in (4-1), there are at most finitely many points at which they are tangent).

Using this terminology, we will classify $\boldsymbol{u} \in \mathbb{Z}^n$ in terms of three cases:

Definition 4.1. For $\boldsymbol{u} \in \mathbb{Z}^n$ and $p \in \mathcal{P}$ we say that:

(i) $\boldsymbol{u}$ is of type zero mod p if $\boldsymbol{u} \equiv \boldsymbol{0} \pmod{p}$,

(ii) $\boldsymbol{u}$ is good mod p if $\boldsymbol{u} \not\equiv \boldsymbol{0} \pmod{p}$ and $V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is not tangent to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$,

(iii) $\boldsymbol{u}$ is bad mod p if $\boldsymbol{u} \not\equiv \boldsymbol{0} \pmod{p}$, and $V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is tangent to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$.

(The fact that we define these types in relation to $V(F(Z^e, \boldsymbol{X}))$, is justified by Lemma 4.4, below.) The main result of this section is the following:

Proposition 4.2. *Assume that $p > 2$ is a prime of good reduction for $F(Z^e, \boldsymbol{X})$, that is $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is nonsingular.*

(i) *If $\boldsymbol{u}$ is type zero modulo p then $g(\boldsymbol{u}, p) \ll p^{n-1/2}$;*

(ii) *If $\boldsymbol{u}$ is good modulo p then $g(\boldsymbol{u}, p) \ll p^{n/2}$;*

(iii) *If $\boldsymbol{u}$ is bad modulo p then $g(\boldsymbol{u}, p) \ll p^{(n+1)/2}$.*

The implied constants can depend on n, m, e, d, but are independent of $\|F\|, \boldsymbol{u}, p$.

In a final step of the proof, we will apply the property that if $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n$ is nonsingular, any hyperplane L has

$$\dim\{P \in V(F(Z^e, \boldsymbol{X})) : T_P \subseteq L\} = \dim(\mathrm{Sing}(V(F(Z^e, \boldsymbol{X})) \cap L)) \leq 0. \tag{4-1}$$

Here, by $\dim(\mathrm{Sing}(V))$ we mean the dimension of the singular locus of a variety $V \subset \mathbb{P}^n$. We will apply this in (4-3) over $\overline{\mathbb{F}}_p$ for p a prime of good reduction for $F(Z^e, \boldsymbol{X})$. The result (4-1) is a special case of Zak's theorem on tangencies as in [20, Theorem 7.1, Remark 7.5], valid over any algebraically closed field, or [33, Lemma 3], valid over any perfect field. More simply, in our setting (4-1) can be shown directly, and we do so in Remark 4.5.

As preparation for proving Proposition 4.2, we transform $g(\boldsymbol{u}, p)$ into an exponential sum over solutions to $F(y, \boldsymbol{a}) = 0$ by writing

$$\begin{aligned} g(\boldsymbol{u}, p) &= \sum_{\boldsymbol{a} \in \mathbb{F}_p^n} \nu_p(\boldsymbol{a}) e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle) - \sum_{\boldsymbol{a} \in \mathbb{F}_p^n} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle) \\ &= -\delta_{\boldsymbol{u}=\boldsymbol{0}} \cdot p^n + \sum_{\boldsymbol{a} \in \mathbb{F}_p^n} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle) \sum_{\substack{y \in \mathbb{F}_p \\ F(y, \boldsymbol{a}) = 0}} 1 \\ &= -\delta_{\boldsymbol{u}=\boldsymbol{0}} \cdot p^n + \sum_{\substack{(y, \boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(y, \boldsymbol{a}) = 0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle), \end{aligned}$$

where $\delta_{\boldsymbol{u}=\mathbf{0}} = 1$ if $\boldsymbol{u} \equiv \mathbf{0} \pmod{p}$ and is 0 otherwise. The task now is to estimate the sum

$$g(\boldsymbol{u}, p) + \delta_{\boldsymbol{u}=\mathbf{0}} \cdot p^n = \sum_{\substack{(y,\boldsymbol{a})\in\mathbb{F}_p^{n+1} \\ F(y,\boldsymbol{a})=0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle).$$

A barrier to doing this efficiently is that the polynomial $F(Y, \boldsymbol{X})$ is not homogeneous (see Remark 4.6). Recall the definition of $F(Y, \boldsymbol{X})$ in (1-1), and recall the integer $e \geq 1$ fixed in that definition. As a first step, we prove:

Lemma 4.3. *Fix a prime $p > 2$. Let $f = (e, p-1)$, and let $\gamma \in \mathbb{F}_p^\times$ be a primitive f-th root of unity. Then*

$$\sum_{(y,\boldsymbol{a})\in W} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle) = \frac{1}{f} \sum_{i=0}^{f-1} \sum_{(z,\boldsymbol{a})\in W_i} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle),$$

where

$$W = \{(y, \boldsymbol{a}) \in \mathbb{F}_p^{n+1} : F(y, \boldsymbol{a}) = 0\},$$
$$W_i = \{(z, \boldsymbol{a}) \in \mathbb{F}_p^{n+1} : F(\gamma^i z^e, \boldsymbol{a}) = 0\}, \quad \textit{for } i = 0, \dots, f-1.$$

(This lemma replaces the remarks in Section 3.4 that applied in the special case $(e, p-1) = 1$.)

Proof. We start by claiming that for any $y \in \mathbb{F}_p^\times$ there exists an unique $i \in \{0, \dots, f-1\}$ and some $z \in \mathbb{F}_p^\times$ such that $y = \gamma^i z^e$: we write $e = \ell k$ where

$$(\ell, q) = 1 \text{ for any } q \,|\, (p-1), \qquad k = \frac{e}{\ell}.$$

Note that then $f|k$ and also there exists some integer N such that $k|(f^N)$. Since γ is a generator for the group $\mathbb{F}_p^\times/\mathbb{F}_p^{\times f}$, then for any $y \in \mathbb{F}_p^\times$ there exists an unique $i \in \{0, \dots, f-1\}$ and $z_1 \in \mathbb{F}_p^\times$ such that $y = \gamma^i z_1^f$. On the other hand, we can apply the same principle to z_1, finding an unique $j \in \{0, \dots, f-1\}$ and $z_2 \in \mathbb{F}_p^\times$ such that $z_1 = \gamma^j z_2^f$. Thus, $y = \gamma^i z_1^f = \gamma^i (\gamma^j z_2^f)^f = \gamma^i z_2^{f^2}$. Iterating this process N times, we can find $z_N \in \mathbb{F}_p^\times$ such that $y = \gamma^i z_N^{f^N}$ with $k|f^N$. Then, $y = \gamma^i (z_N^{f^N/k})^k$. On the other hand, since $(\ell, p-1) = 1$, we have that $z_N^{f^N/k} = z^\ell$ for some $z \in \mathbb{F}_p^\times$, so that $y = \gamma^i z^{\ell k} = \gamma^i z^e$ and this proves the claim. Moreover, note that once we have obtained z such that $y = \gamma^i z^e$ then we can multiply z by any f-th root of unity, so that there are f such values z.

Next, for any $i \in \{0, \dots, f-1\}$ we can consider the map

$$\varphi_i : W_i \longrightarrow W \quad (z, \boldsymbol{a}) \mapsto (\gamma^i z^e, \boldsymbol{a}).$$

From this, we deduce that if $(y, \boldsymbol{a})$ is in the image of φ_i then

$$|\varphi_i^{-1}(y, \boldsymbol{a})| = \begin{cases} f & \text{if } y \neq 0, \\ 1 & \text{if } y = 0. \end{cases}$$

On the other hand, if $(0, \boldsymbol{a}) \in W$, then $(0, \boldsymbol{a}) \in W_i$ for each of $i = 0, \dots, f-1$. The result follows. $\square$

When we apply Lemma 4.3 it will be convenient to treat all cases analogously as i varies; to do so we will employ the following lemma.

Lemma 4.4. *Fix $e \geq 1$ and recall $F(Y, \boldsymbol{X})$ from (1-1). Let p be a prime, and let $\boldsymbol{u} \in \overline{\mathbb{F}}_p^n$. Then for any $\alpha \in \overline{\mathbb{F}}_p^\times$ the variety $V(F(\alpha Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is isomorphic to $V(F(Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$. In particular, for $\boldsymbol{u} = \boldsymbol{0}$, we conclude $V(F(\alpha Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is isomorphic to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$.*

Proof. Let $\beta \in \overline{\mathbb{F}}_p^\times$ be such that $\beta^e = \alpha$. Then the change of variables $(Z, \boldsymbol{X}) \mapsto (\beta Z, \boldsymbol{X})$ induces an isomorphism between $V(F(Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle)$ and $V(F(\alpha Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u}\rangle)$. □

4.1. *Proof of Proposition 4.2.* We are now ready to prove our main result of this section, Proposition 4.2. In the following, we denote $f = (e, p-1)$. An application of Lemma 4.3 leads to

$$g(\boldsymbol{u}, p) = -\delta_{\boldsymbol{u}=\boldsymbol{0}} p^n + \frac{1}{f} \sum_{i=0}^{f-1} \sum_{(z,\boldsymbol{a}) \in W_i} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle). \tag{4-2}$$

4.1.1. *Type zero case.* Assume $\boldsymbol{u} \equiv \boldsymbol{0} \pmod p$. The right-hand side of (4-2) becomes

$$g(\boldsymbol{0}, p) = -p^n + \frac{1}{f} \sum_{i=0}^{f-1} \sum_{(z,\boldsymbol{a}) \in W_i} 1 = -p^n + \frac{1}{f} \sum_{i=0}^{f-1} |W_i|.$$

By definition, for any $i = 0, \dots, f-1$ the set W_i is the set of the $\mathbb{F}_p$-points on the affine variety $V(F(\gamma^i Z^e, \boldsymbol{X})) \subset \mathbb{A}^{n+1}_{\mathbb{F}_p}$. By hypothesis, p is of good reduction for $V(F(Z^e, \boldsymbol{X}))$, so $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is nonsingular. Then by Lemma 4.4, we have that $V(F(\gamma^i Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is a nonsingular variety for each $i = 0, \dots, f-1$ (and in particular is absolutely irreducible over $\overline{\mathbb{F}}_p$), and certainly $V(F(\gamma^i Z^e, \boldsymbol{X}))$ is defined over $\mathbb{F}_p$. Thus the Lang-Weil bound [35] implies that (counting projectively)

$$|V(F(\gamma^i Z^e, \boldsymbol{X}))(\mathbb{F}_p)| = p^{n-1} + O_{m,e,d}(p^{n-1-1/2}) \qquad \text{for each } i = 0, \dots, f-1,$$

so that $|W_i| = p^n + O_{m,e,d,}(p^{n-1/2})$ for each $i = 0, \dots, f-1$. Thus we may conclude that $g(\boldsymbol{0}, p) \ll p^{n-1/2}$.

4.1.2. *Good/bad case.* Assume $\boldsymbol{u} \neq \boldsymbol{0} \pmod p$; we may initially argue the good and the bad cases together. The right hand side of (4-2) becomes

$$g(\boldsymbol{u}, p) = \frac{1}{f} \sum_{i=0}^{f-1} \sum_{(z,\boldsymbol{a}) \in W_i} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle).$$

In either the good or the bad case, it suffices to estimate each sum

$$g_i(\boldsymbol{u}, p) = \sum_{(z,\boldsymbol{a}) \in W_i} e_p(\langle \boldsymbol{a}, \boldsymbol{u}\rangle), \qquad \text{for } i = 0, .., f-1.$$

First we prove that for any $\alpha \in \mathbb{F}_p^\times$, $g_i(\boldsymbol{u}, p) = g_i(\alpha \boldsymbol{u}, p)$. Indeed

$$\begin{aligned} g_i(\alpha \boldsymbol{u}, p) &= \sum_{(z,\boldsymbol{a}) \in W_i} e_p(\langle \boldsymbol{a}, \alpha \boldsymbol{u} \rangle) = \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} e_p(\langle \boldsymbol{a}, \alpha \boldsymbol{u} \rangle) \\ &= \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} e_p(\langle \alpha \boldsymbol{a}, \boldsymbol{u} \rangle) = \sum_{\substack{(t,\boldsymbol{b}) \in \mathbb{F}_p^{n+1} \\ \overline{\alpha}^{med} F(\gamma^i t^e, \boldsymbol{b}) = 0}} e_p(\langle \boldsymbol{b}, \boldsymbol{u} \rangle) \\ &= \sum_{\substack{(t,\boldsymbol{b}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i t^e, \boldsymbol{b}) = 0}} e_p(\langle \boldsymbol{b}, \boldsymbol{u} \rangle) = g_i(\boldsymbol{u}, p), \end{aligned}$$

where in the fourth step we use the change of variables $(z, \boldsymbol{a}) = (\overline{\alpha} t, \overline{\alpha} \boldsymbol{b})$, for $\alpha \overline{\alpha} \equiv 1 \pmod p$. Hence

$$\begin{aligned} (p-1) g_i(\boldsymbol{u}, p) &= \sum_{\alpha \in \mathbb{F}_p^\times} g_i(\alpha \boldsymbol{u}, p) \\ &= \sum_{\alpha \in \mathbb{F}_p^\times} \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} e_p(\langle \boldsymbol{a}, \alpha \boldsymbol{u} \rangle) \\ &= \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} \sum_{\alpha \in \mathbb{F}_p^\times} e_p(\alpha \langle \boldsymbol{a}, \boldsymbol{u} \rangle) = \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} \sum_{\alpha \in \mathbb{F}_p} e_p(\alpha \langle \boldsymbol{a}, \boldsymbol{u} \rangle) - \sum_{\substack{(z,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(\gamma^i z^e, \boldsymbol{a}) = 0}} 1 \\ &= p(p-1) \left| (V(F(\gamma^i Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle))(\mathbb{F}_p) \right| - (p-1) \left| V(F(\gamma^i Z^e, \boldsymbol{X})(\mathbb{F}_p) \right| + (p-1), \end{aligned}$$

where in the last step we have passed to counting points over $\mathbb{F}_p$ in the projective sense. Applying [32, Appendix by N. Katz, Theorem 1], we have that

$$|V(F(\gamma^i Z^e, \boldsymbol{X}))(\mathbb{F}_p)| = \sum_{j=0}^{n-1} p^j + O_{n,m,e,d}(p^{\frac{n+\delta_i}{2}}),$$

$$|(V(F(\gamma^i Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle))(\mathbb{F}_p)| = \sum_{j=0}^{n-2} p^j + O_{n,m,e,d}(p^{\frac{n-1+\delta_{i,\boldsymbol{u}}}{2}}),$$

where $\delta_i = \dim(\mathrm{Sing}(V(F(\gamma^i Z^e, \boldsymbol{X}))))$ and $\delta_{i,\boldsymbol{u}} = \dim(\mathrm{Sing}(V(F(\gamma^i Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle)))$.

On the other hand, Lemma 4.4 implies that $\delta_i = \delta_0$ and $\delta_{i,\boldsymbol{u}} = \delta_{0,\boldsymbol{u}}$ for each i. Moreover, $\delta_0 = -1$ since we are assuming that p is of good reduction for $V(F(Z^e, \boldsymbol{X}))$. Thus, we obtain

$$g_i(\boldsymbol{u}, p) = O(p^{\frac{n+1+\delta_{0,\boldsymbol{u}}}{2}}), \tag{4-3}$$

with an implicit constant depending only on n, m, e, d. Finally, by (4-1),

$$\delta_{0,\boldsymbol{u}} = \begin{cases} 0 & \text{if } V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle) \text{ is tangent to } V(F(Z^e, \boldsymbol{X})), \\ -1 & \text{otherwise,} \end{cases}$$

and this completes the proof of the good and bad cases in Proposition 4.2.

Remark 4.5. This remark justifies (4-1). Let $V = V(H(\boldsymbol{X})) \subset \mathbb{P}^n$ be a nonsingular hypersurface and $L = V(\langle \boldsymbol{a}, \boldsymbol{X} \rangle)$ be a hyperplane. We may suppose without loss of generality that $a_1 \neq 0$. By the Jacobian criterion, $\mathrm{Sing}(V \cap L)$ is the set of points on the intersection $V \cap L$ for which the $(n+1) \times 2$ matrix with columns ∇H and $\boldsymbol{a}$ has rank 1. Consequently, $\mathrm{Sing}(V \cap L) \subset W$ where

$$W = V \cap V(g_2) \cap \cdots \cap V(g_n),$$

in which for each $i = 2, \ldots, n$,

$$g_i(\boldsymbol{X}) = a_1 \frac{\partial H}{\partial X_i}(\boldsymbol{X}) - a_i \frac{\partial H}{\partial X_1}(\boldsymbol{X}).$$

On the other hand, $W \cap V(\partial H/\partial X_1) = \mathrm{Sing}(V) = \varnothing$ under the hypothesis that V is nonsingular. Consequently, $\dim W \leq 0$, implying $\dim(\mathrm{Sing}(V \cap L)) \leq 0$, as desired.

Remark 4.6. It is worth remarking what we have gained from the arguments in this section. Briefly, suppose $\boldsymbol{u} \not\equiv 0 \pmod{p}$ and consider

$$g(\boldsymbol{u}, p) = \sum_{\substack{(y,\boldsymbol{a}) \in \mathbb{F}_p^{n+1} \\ F(y,\boldsymbol{a})=0}} e_p(\langle \boldsymbol{a}, \boldsymbol{u} \rangle).$$

To work directly with this sum rather than passing through the dissection into the components W_i as we did above, we would first need to homogenize the polynomial $F(Y, \boldsymbol{x})$, say defining a homogeneous polynomial

$$\tilde{F}(T, Y, \boldsymbol{X}) = T^{md(e-1)} Y^{md} + \cdots + T^{m(e-1)} Y^m f_{d-1}(\boldsymbol{X}) + f_d(\boldsymbol{X}).$$

(Here we suppose that $e \geq 2$ for this example.) Then observe that $[1 : 0 : \cdots : 0]$ is a singular point on $V(\tilde{F}(T, Y, \boldsymbol{X})) \subset \mathbb{P}^{n+1}$. Consequently, if one proceeded to estimate $g(\boldsymbol{u}, p)$, roughly analogous to the approach in (4-3), by counting points on the complete intersection described by

$$V(\tilde{F}(T, Y, \boldsymbol{X})) \cap V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle) \cap V(T = 1),$$

the role of $\delta_{0,\boldsymbol{u}}$ in the exponent is now played by a dimension that is always at least 0, ultimately leading to a result that is larger by a factor of $p^{1/2}$ than the results we obtain in Proposition 4.2.

4.2. ***Choice of the sieving set.*** We can now continue the discussion initiated in Section 3.4, and choose the sieving set. We suppose that $Q = B^\kappa$ for some $1/2 \leq \kappa \leq 1$ to be chosen later (see (7-4)). We choose the sieving set

$$\mathcal{P} \subset [Q, 2Q]$$

comprised of all primes in this range such that (i) $p \equiv 1 \pmod{m}$ (recalling $m \geq 2$), and (iii′) the reduction $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ is nonsingular.

By the Siegel–Walfisz theorem on primes in arithmetic progressions, there are $\gg_m Q/\log Q$ primes such that $p \equiv 1 \pmod{m}$ in any dyadic region $[Q, 2Q]$, for all $Q \gg_m 1$ sufficiently large, which we assume is a condition met henceforward. We recall from (3-13) that at most $O_{m,e,d}(\log \|F\|)$ primes fail (iii′). We henceforward assume that

$$Q \gg_{m,e,d} (\log \|F\|)(\log\log \|F\|) \tag{4-4}$$

for an appropriately large implied constant, so that consequently

$$P = |\mathcal{P}| \gg_m Q/\log Q - C_{m,e,d}(\log \|F\|) \gg_{m,e,d} Q/\log Q. \tag{4-5}$$

When we finally choose Q as a power of B, (4-4) will impose a lower bound on B; we defer this to (7-4).

5. Estimating the main sieve term: the bad-bad case

This section is the technical heart of the paper. We show how to bound the most difficult contribution to the sieve, which occurs when $\boldsymbol{u}$ is bad with respect to two primes $p \neq q \in \mathcal{P}$. (We reserve the treatment of all other cases, when $\boldsymbol{u}$ is either type zero, or good with respect to at least one of these primes, to Section 7; these remaining cases are significantly easier.)

We recall from the sieve lemma, Lemma 1.2, that $\mathcal{S}(F, B)$ is bounded above by a sum of three terms. The first two terms can be bounded simply:

$$\sum_{\boldsymbol{k}: f_d(\boldsymbol{k})=0} W(\boldsymbol{k}) + \frac{1}{P}\sum_{\boldsymbol{k}} W(\boldsymbol{k}) \ll B^{n-1} + B^n P^{-1}. \tag{5-1}$$

Here the first term follows from the Schwartz-Zippel trivial bound $\ll_{n,e,d} B^{n-1}$ for the number of zeroes of f_d with $\boldsymbol{k} \in \operatorname{supp}(W)$, since $f_d \not\equiv 0$ (see, e.g., [27, Theorem 1], which as mentioned before has a method of proof that applies even if f_d is not absolutely irreducible). We will call the remaining, third, term on the right-hand side of the sieve lemma the main sieve term.

Now we are ready to estimate the main sieve term, which after an application of Poisson summation inside the definition (1-23) of $T(p, q)$ is

$$\begin{aligned}\frac{1}{P^2}\sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} |T(p,q)| &= \frac{1}{P^2}\sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \left(\frac{1}{pq}\right)^n \left|\sum_{\boldsymbol{u}} \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \\ &\ll \frac{1}{P^2 Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\boldsymbol{u}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right|.\end{aligned} \tag{5-2}$$

We will apply Proposition 4.2 to bound $g(\boldsymbol{u}, pq)$, according to the "type" of $\boldsymbol{u}$ modulo p and q, respectively; this leads to cases we can abbreviate as zero-zero, zero-good, zero-bad, good-good, good-bad, and bad-bad. Unsurprisingly, the greatest difficulty is to bound the contribution of the bad-bad case, and we focus on this first, returning to the other cases in Section 7.

Recall that W is a nonnegative function with $W(\boldsymbol{u}) = w(\boldsymbol{u}/B)$ for an infinitely differentiable, nonnegative function w that is $\equiv 1$ on $[-1, 1]$ and vanishes outside of $[-2, 2]$. Thus $\hat{W}(\boldsymbol{u}) = B^n \hat{w}(B\boldsymbol{u})$ and $\hat{w}(\boldsymbol{u})$ has rapid decay in $\boldsymbol{u}$, so that

$$|\hat{W}(\boldsymbol{u})| \ll B^n \prod_{i=1}^{n} (1 + |u_i|B)^{-M} \tag{5-3}$$

for any $M \geq 1$; we will for example specify a lower bound on M at (5-22) and can certainly always assume $M \geq 2n$. In particular, we will later apply the fact that for any $B, L \geq 1$,

$$\sum_{\boldsymbol{u} \in \mathbb{Z}^n} |\hat{W}(\boldsymbol{u}/L)| \ll \max\{B^n, L^n\}. \tag{5-4}$$

5.1. ***The dual variety.*** To consider any bad case, it is useful to consider certain facts about the dual variety. Recall that $m \geq 2$ and $d, e \geq 1$, and

$$F(Y, \boldsymbol{X}) = Y^{md} + Y^{m(d-1)} f_1(\boldsymbol{X}) + \cdots + f_d(\boldsymbol{X}), \tag{5-5}$$

in which for each $1 \leq i \leq d$, f_i is a polynomial in $\mathbb{Z}[X_1, \ldots, X_n]$ with $\deg f_i = m \cdot e \cdot i$. By hypothesis, the variety defined by $F(Y, \boldsymbol{X}) = 0$ in weighted projective space, denoted $V(F(Y, \boldsymbol{X})) \subset \mathbb{P}_{\mathbb{C}}(e, 1, \ldots, 1)$, is nonsingular. Recall from Section 3.3 that $V(F(Y, \boldsymbol{X})) \subset \mathbb{P}_{\mathbb{C}}(e, 1, \ldots, 1)$ is nonsingular if and only if $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\mathbb{C}}$ is nonsingular. The dual variety $V^* = V(F(Z^e, \boldsymbol{X}))^* \subset \mathbb{P}^n_{\mathbb{C}}$ of a hypersurface is a hypersurface. We denote by

$$G(U_Y, U_1, \ldots, U_n) \tag{5-6}$$

the irreducible homogeneous polynomial such that $V(G) = V^*$ (see, e.g., [13, Proposition 11.2, Appendix]). Recall that $\deg F(Z^e, \boldsymbol{X}) = mde$; by [19, Proposition 2.9],

$$\deg G = mde(mde - 1)^{n-1} \geq 2.$$

In our analysis of the bad-bad case in Section 5.2, our strategy is to divide our analysis depending on whether $\boldsymbol{u}$ has the property $G(0, \boldsymbol{u}) \neq 0$ or $G(0, \boldsymbol{u}) = 0$. In the first case, we now show via an explicit constructive argument that

$$|\{p : \boldsymbol{u} \text{ is bad modulo } p\}| \ll_{n,m,e,d} \log(\|F\|\|\boldsymbol{u}\|). \tag{5-7}$$

Let us prove this. A given $\boldsymbol{u}$ has the property $G(0, \boldsymbol{u}) \neq 0$ if and only if the hyperplane $V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle) \subset \mathbb{P}^n_{\mathbb{C}}$ is not tangent to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\mathbb{C}}$; that is, if and only if for any $[z : \boldsymbol{x}] \in V(F(Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle)$, the matrix

$$\begin{pmatrix} \frac{\partial F}{\partial Z}(z^e, \boldsymbol{x}) & 0 \\ \frac{\partial F}{\partial X_1}(z^e, \boldsymbol{x}) & u_1 \\ \vdots & \\ \frac{\partial F}{\partial X_n}(z^e, \boldsymbol{x}) & u_n \end{pmatrix} \tag{5-8}$$

has maximal rank (i.e., at least one 2×2 minor is nonvanishing). Now define $n+2$ polynomials in $Z, X_1, \ldots, X_n$, with integral coefficients (depending on $\boldsymbol{u}$) as follows: set

$$H_{0,\boldsymbol{u}}(Z, \boldsymbol{X}) = F(Z^e, \boldsymbol{X}), \qquad H_{n+1,\boldsymbol{u}}(Z, \boldsymbol{X}) = \langle \boldsymbol{X}, \boldsymbol{u} \rangle,$$

and for $1 \le i \le n$ set

$$H_{i,\boldsymbol{u}}(Z, \boldsymbol{X}) = \begin{cases} \det \begin{pmatrix} \frac{\partial F}{\partial Z}(z^e, \boldsymbol{x}) & 0 \\ \frac{\partial F}{\partial X_1}(z^e, \boldsymbol{x}) & u_1 \end{pmatrix} & \text{for } i = 1, \\ \det \begin{pmatrix} \frac{\partial F}{\partial X_{i-1}}(z^e, \boldsymbol{x}) & u_{i-1} \\ \frac{\partial F}{\partial X_i}(z^e, \boldsymbol{x}) & u_i \end{pmatrix} & \text{for } 2 \le i \le n. \end{cases}$$

Then define the resultant (see [21, Chapter 13])

$$R(\boldsymbol{u}) = \operatorname{Res}(H_{0,\boldsymbol{u}}, H_{1,\boldsymbol{u}}, \ldots, H_{n+1,\boldsymbol{u}}). \tag{5-9}$$

The following are all equivalent:

(1) $\boldsymbol{u}$ has the property that $V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle)$ is tangent to $V(F(Z^e, \boldsymbol{X}))$.

(2) For some $[z : \boldsymbol{x}] \in V(F(Z^e, \boldsymbol{X})) \cap V(\langle \boldsymbol{X}, \boldsymbol{u} \rangle)$, (5-8) has rank < 2.

(3) The polynomials $H_{i,\boldsymbol{u}}(Z, \boldsymbol{X})$ (for $0 \le i \le n+1$) share a common (nonzero) root.

(4) $R(\boldsymbol{u}) = 0$.

Now we consider the analogues of these statements for each p. Fix a prime p. For a polynomial $L \in \mathbb{Z}[\boldsymbol{U}]$, let $\overline{L}$ denote its reduction modulo p. By definition, $\boldsymbol{u}$ is bad modulo p precisely when $\overline{H}_{i,\boldsymbol{u}}$ (for $0 \le i \le n+1$) have a common nontrivial root modulo p, that is if and only if $p | \operatorname{Res}(\overline{H}_{0,\boldsymbol{u}}, \ldots, \overline{H}_{n+1,\boldsymbol{u}})$. By [15, Section IV], as a polynomial in $\boldsymbol{U}$,

$$\operatorname{Res}(\overline{H}_{0,\boldsymbol{U}}, \ldots, \overline{H}_{n+1,\boldsymbol{U}}) = \overline{R}(\boldsymbol{U}),$$

where R is defined as in (5-9). (That is, the resultant of the reductions modulo p is the reduction modulo p of the resultant.) Thus for each $\boldsymbol{u}$ such that $G(0, \boldsymbol{u}) \neq 0$ so that $R(\boldsymbol{u}) \neq 0$, we can conclude that

$$\left| \{ p : \boldsymbol{u} \text{ is bad modulo } p \} \right| = \omega(\operatorname{Res}(H_{0,\boldsymbol{u}}, \ldots, H_{n+1,\boldsymbol{u}})),$$

where $\omega(r)$ indicates the number of distinct prime divisors of an integer r; we recall in particular that $\omega(r) \ll (\log r)/(\log\log r)$. By [21, Chapter 13, Proposition 1.1], the resultant is a homogeneous polynomial in the coefficients of the forms $H_{0,\boldsymbol{u}}, \ldots, H_{n+1,\boldsymbol{u}}$ (with degree bounded in terms of n, m, e, d). Thus, for every value of $\boldsymbol{u}$ such that $G(0, \boldsymbol{u}) \neq 0$ so that $\operatorname{Res}(H_{0,\boldsymbol{u}}, \ldots, H_{n+1,\boldsymbol{u}})$ is a nonzero integer,

$$\omega(\operatorname{Res}(H_{0,\boldsymbol{u}}, \ldots, H_{n+1,\boldsymbol{u}})) \ll_{n,m,e,d} \log(\|F\| \|\boldsymbol{u}\|). \tag{5-10}$$

Finally, if $G(0, \boldsymbol{u}) = 0$, then the hyperplane $V(\langle \boldsymbol{u}, \boldsymbol{X} \rangle) \subset \mathbb{P}^n_{\mathbb{C}}$ is tangent to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n_{\mathbb{C}}$ so that (5-8) has rank 1 over $\mathbb{C}$; consequently $\boldsymbol{u}$ is bad for all primes p. Thus in this latter case, we will instead focus on showing there are sufficiently few solutions to $G(0, \boldsymbol{u}) = 0$.

Remark 5.1. It is a common occurrence that one requires the fact that there are "quite few" primes of bad reduction for a variety of the form $\mathcal{V} \cap \{u_0 X_0 + \cdots u_n X_n = 0\}$ for some variety $\mathcal{V}$ and parameter $(u_0, u_1, \ldots, u_n)$ of interest, in this case $V(G)$ with G describing the dual of F, and $u_0 = 0$. The fact that our result (5-7) depends only logarithmically on $\|F\|$ is important for our ultimate deduction that the implicit constant in Theorem 1.1 is independent of $\|F\|$; see the application in Section 5.2.1. This motivated the explicit argument we gave above. Alternatively, we thank Per Salberger for pointing out that the useful references [17, pp. 95–98] and [18] also provide similar constructions leading to explicit results of the form (5-10) and hence (5-7). We remark that if we did not require logarithmic dependence on $\|F\|$, one could apply a result such as [13, Proposition 11.5(3), Appendix] to conclude immediately that for all sufficiently large primes (in an inexplicit sense), $\boldsymbol{u}$ is bad modulo p precisely when $p|G(0, \boldsymbol{u})$ (so that $|\{p : \boldsymbol{u} \text{ is bad modulo p}\}| \ll_G \log \|\boldsymbol{u}\|$ when $G(0, \boldsymbol{u}) \neq 0$), but with dependence on G and hence on F that has not been made explicit, and so does not immediately suffice for our application.

5.2. ***Bad-bad case.*** We use the above facts to control the contribution of the bad-bad case to the sieve, which by Proposition 4.2 is bounded by

$$\frac{1}{P^2 Q^{2n}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q}} \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq) \right| \ll \frac{Q^{n+1}}{P^2 Q^{2n}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q}} \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) \right|. \tag{5-11}$$

We start by exchanging the order of summation between $\boldsymbol{u}$ and the primes p, q, and then splitting the sum as

$$\sum_{\boldsymbol{u} \in \mathbb{Z}^n} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) \right| = \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ G(0,\boldsymbol{u})=0}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} + \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ G(0,\boldsymbol{u}) \neq 0}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}}.$$

In this section, we will prove that the contribution from $G(0, \boldsymbol{u}) \neq 0$ is

$$\sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ G(0,\boldsymbol{u}) \neq 0}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) \right| \ll_{n,m,e,d} Q^{2n} (\log B)^2. \tag{5-12}$$

On the other hand, we will prove that the contribution from $G(0, \boldsymbol{u}) = 0$ is

$$\sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ G(0,\boldsymbol{u})=0}} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) \right| \ll_\varepsilon P^2 \left(Q^{2n} B^{-\alpha(M-1)} + B^n \left(\frac{Q^2}{B^{1-\alpha}} \right)^{n-2+\frac{1}{3}+\varepsilon} \right), \tag{5-13}$$

for a small $0 < \alpha < 1$ of our choice, and any $\varepsilon > 0$. Once we have proved these two inequalities, we will wrap up the contribution of the bad-bad case in Section 5.2.3.

5.2.1. *The case* $G(0,\boldsymbol{u}) \neq 0$. Proving (5-12) is quite simple; by the decay (5-3) for $\hat{W}$ and the bound (5-10) for counting p, q,

$$\sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})\neq 0}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q\\ \boldsymbol{u}\text{ bad mod }p\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right)\right| \ll B^n \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})\neq 0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} \omega(R(\boldsymbol{u}))^2$$

$$\ll B^n \sum_{\boldsymbol{u}\in\mathbb{Z}^n} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} (\log(\|F\|\|\boldsymbol{u}\|))^2$$

$$\ll_{n,m,e,d} Q^{2n}(\log B)^2.$$

Here we have used the fact that $Q = B^\kappa$ with $1/2 \le \kappa \le 1$ (so that $Q^{2n} \gg B^n$), and the fact from Lemma 2.1 that in the only case we need to consider, $\log\|F\| \ll_{m,e,d} \log B$. This proves (5-12) with an implied constant independent of $\|F\|$.

5.2.2. *The case* $G(0,\boldsymbol{u}) = 0$. Proving (5-13) is a key novel aspect of our proof. Note that if $G(0,\boldsymbol{u}) = 0$, then $\boldsymbol{u}$ is bad mod p for all $p \in \mathcal{P}$. Then

$$\sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q\\ \boldsymbol{u}\text{ bad mod }p\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right)\right| \ll B^n P^2 \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M}. \tag{5-14}$$

Let $0 < \alpha < 1$ be a parameter to be chosen later and consider the cube

$$\mathcal{C}_\alpha = [-Q^2/B^{1-\alpha}, Q^2/B^{1-\alpha}]^n \subset \mathbb{R}^n.$$

This is slightly larger than the "essential support" of the sum over $\boldsymbol{u}$, so that outside this box we can exploit decay more efficiently. We will ultimately prove that

$$\sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} \ll_\varepsilon Q^{2n}B^{-n}B^{-\alpha(M-1)} + \left(\frac{Q^2}{B^{1-\alpha}}\right)^{n-2+1/3+\varepsilon}, \tag{5-15}$$

for any $\varepsilon > 0$. We split the sum as

$$\sum_{\substack{\boldsymbol{u}\in\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} + \sum_{\substack{\boldsymbol{u}\notin\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M}. \tag{5-16}$$

In the second sum in (5-16), we can exploit decay:

$$\sum_{\substack{\boldsymbol{u}\notin\mathcal{C}_\alpha\\ G(0,\boldsymbol{u})=0}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} \ll \sum_{j=1}^{n} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0\\ |u_j|>Q^2/B^{1-\alpha}}} \prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M} \ll \left(\frac{Q^2}{B}\right)^n \frac{1}{B^{\alpha(M-1)}}.$$

The contribution of these $\boldsymbol{u}$ to (5-14) is thus $\ll Q^{2n}P^2B^{-\alpha(M-1)}$ for $0<\alpha<1$ and any $M\geq 2n$; this contributes the first term in (5-13).

It remains to deal with the first sum appearing on the right-hand side of (5-16), summing over $\boldsymbol{u}\in\mathcal{C}_\alpha$ such that $G(0,\boldsymbol{u})=0$. Here we show that there are few solutions to $G(0,\boldsymbol{u})=0$. Recall the definition of the form G from Section 5.1. Consider $V(G(0,\boldsymbol{U}))\subset\mathbb{P}^{n-1}$ defined by $G(0,\boldsymbol{U})=0$ as a function of $\boldsymbol{U}$. (First notice that $G(0,\boldsymbol{U})$ is not identically zero; indeed, if it were then we would conclude that $\{U_Y=0\}\subset\{G(U_Y,U_1,\dots,U_n)=0\}$. Recalling that $G(U_Y,\boldsymbol{U})$ is irreducible, both these projective varieties have dimension $n-1$ so that in fact we must have $\{G=0\}=\{U_Y=0\}$. But this is impossible, since G has degree >1.) Thus $V(G(0,\boldsymbol{U}))\subset\mathbb{P}^{n-1}_{\mathbb{C}}$ is a projective variety of dimension $n-2$ and $\deg G(0,\boldsymbol{U})=\deg G(U_Y,\boldsymbol{U})\geq 2$. Moreover, let us decompose $G(0,\boldsymbol{U})$ into irreducible components, i.e., by writing

$$G(0,\boldsymbol{U})=\prod_{\ell=1}^{L}G_\ell(\boldsymbol{U}), \tag{5-17}$$

where $G_\ell(\boldsymbol{U})$ is an irreducible polynomial for each $\ell\leq L$ (and $L\ll_{n,m,e,d}1$). Set $d_\ell:=\deg G_\ell$. We have

$$\sum_{\substack{\boldsymbol{u}\in\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}}\prod_{i=1}^{n}\left(1+\frac{B|u_i|}{Q^2}\right)^{-M}\leq\sum_{\substack{\boldsymbol{u}\in\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G(0,\boldsymbol{u})=0}}1\leq\sum_{\ell=1}^{L}\sum_{\substack{\boldsymbol{u}\in\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G_\ell(\boldsymbol{u})=0}}1.$$

In the next section, we shall prove:

Proposition 5.2. *Let $n\geq 3$. For the homogeneous polynomial $G(U_Y,U_1,\dots,U_n)\in\mathbb{C}[U_Y,U_1,\dots,U_n]$ defined in (5-6), $G(0,U_1,\dots,U_n)$ contains no linear factor, that is, we cannot write $G(0,\boldsymbol{U})=L(\boldsymbol{U})\tilde{H}(\boldsymbol{U})$ for any linear form $L(\boldsymbol{U})\in\mathbb{C}[U_1,\dots,U_n]$.*

Remark 5.3. As a consequence of Proposition 5.2, $G(0,U_1,\dots,U_n)$ contains no factor in one or two variables. For suppose that in the notation of (5-17) some factor $G_\ell(\boldsymbol{U})$ (after an appropriate $GL_n(\mathbb{C})$ change of variables) can be written as a polynomial $g_1(U_1)$ or $g_2(U_1,U_2)$. Then $g_1(U_1)$ is a monomial, hence a product of linear factors, contradicting the proposition. Alternatively, any form $g_2(U_1,U_2)$ factors over $\mathbb{C}$ into homogeneous linear factors in U_1,U_2, as a consequence of the fundamental theorem of algebra applied to $g_2(1,t)\in\mathbb{C}[t]$, followed by noting $g_2(U_1,U_2)=U_1^{\deg g_2}g_2(1,U_2/U_1)$. This again would contradict the proposition. (Since the statement of Proposition 5.2 is false if $n=2$, see Remark 5.4 for an alternative approach for $n=2$.)

The crucial point is that Proposition 5.2 implies that for each $\ell=1,\dots,L$ the degree $d_\ell\geq 2$ (and G_ℓ depends on at least 3 variables). By [27, Theorem 2] and [41, Theorem A], we have, for any $\varepsilon>0$,

$$\sum_{\substack{\boldsymbol{u}\in\mathcal{C}_\alpha\cap\mathbb{Z}^n\\ G_\ell(\boldsymbol{u})=0}}1\ll_\varepsilon\begin{cases}(Q^2/B^{1-\alpha})^{n-2+\varepsilon} & \text{if } d_\ell=2,\\ (Q^2/B^{1-\alpha})^{n-2+\frac{1}{d_\ell}+\varepsilon} & \text{if } d_\ell>2.\end{cases} \tag{5-18}$$

Within these results, the implied constant is independent of $\|F\|$ in each case. In particular, we may

conclude that for each $\ell = 1, \dots, L$,

$$\sum_{\substack{\boldsymbol{u} \in \mathcal{C}_\alpha \cap \mathbb{Z}^n \\ G_\ell(0,\boldsymbol{u})=0}} 1 \ll_\varepsilon \left(\frac{Q^2}{B^{1-\alpha}} \right)^{n-2+\frac{1}{3}+\varepsilon}.$$

Thus the total contribution of these terms to (5-14) is

$$\ll_\varepsilon B^n P^2 \left(\frac{Q^2}{B^{1-\alpha}} \right)^{n-2+\frac{1}{3}+\varepsilon}.$$

This contributes the second term in (5-13), and hence (5-13) is proved.

5.2.3. *Conclusion of the bad-bad sieve term.* From (5-12) and (5-13) we conclude that the total contribution of the bad-bad case (5-11) to the sieve is

$$\frac{Q^{n+1}}{P^2 Q^{2n}} \left(Q^{2n} (\log B)^2 + Q^{2n} P^2 B^{-\alpha(M-1)} + B^n P^2 \left(\frac{Q^2}{B^{1-\alpha}} \right)^{n-2+\frac{1}{3}+\varepsilon} \right)$$
$$\ll_{\varepsilon'} Q^n \left(QP^{-2} (\log B)^2 + QB^{-\alpha(M-1)} + \left(\frac{B^{\frac{5}{3}+g(\alpha)+\varepsilon'}}{Q^{\frac{7}{3}+\varepsilon'}} \right) \right), \quad (5\text{-}19)$$

where $g(\alpha) = \alpha\big(n - \frac{5}{3} + \varepsilon'\big)$, for any $\varepsilon' > 0$. To simplify the third term above, henceforward we assume $Q = B^\kappa$ with

$$\tfrac{3}{4} \le \kappa \le 1. \quad (5\text{-}20)$$

Then the above is

$$\ll_{\varepsilon'} Q^n (QP^{-2} (\log B)^2 + QB^{-\alpha(M-1)} + B^{-\frac{1}{12}+g(\alpha)+\varepsilon'}), \quad (5\text{-}21)$$

for any $\varepsilon' > 0$. In the first term on the right-hand side, we observe by (4-5) that $P \gg Q/\log Q$ so that

$$QP^{-2} (\log B)^2 \ll Q^{-1} (\log B)^4 \ll B^{-3/4} (\log B)^4.$$

In the second term, we can choose $\alpha = \frac{1}{24}(n - \frac{5}{3} + \varepsilon')^{-1}$ so $g(\alpha) = \frac{1}{24}$, and set $M \ge \max\{2n, \alpha^{-1} + 1\}$. Regarding the third term, so far this is true for any $\varepsilon' > 0$; let us take $\varepsilon' = 1/100$, say. We conclude that

$$\frac{Q^{n+1}}{P^2 Q^{2n}} \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ G(0,\boldsymbol{u})=0}} \sum_{\substack{p,q \in \mathcal{P} \\ p \ne q \\ \boldsymbol{u} \text{ bad mod } p \\ \boldsymbol{u} \text{ bad mod } q}} \left| \hat{W}\left(\frac{\boldsymbol{u}}{pq} \right) \right| \ll Q^n \big(B^{-3/4} (\log B)^4 + QB^{-1} + B^{-\frac{1}{24}+\frac{1}{100}} \big) \ll Q^n, \quad (5\text{-}22)$$

since $B \ge Q$. The implied constant is independent of $\|F\|$. (Here we could even obtain a term that is $o(Q^n)$, but this will not change our main theorem, since the good-good contribution to the sieve is $O(Q^n)$.) This completes the treatment of the bad-bad contribution to the sieve, except for the proof of Proposition 5.2, which we provide in the next section. Then in Section 7 we show that the contributions

of all the other types to the sieve are also dominated by $\ll Q^n$, and then conclude the proof of our main theorem.

Remark 5.4 (the case $n = 2$). The method of this paper applies for $n = 2$ up until Proposition 5.2; arguing as in Remark 5.3 shows that $G(0, U_1, U_2)$ factors over $\mathbb{C}$ into homogeneous linear factors in U_1, U_2, so that proposition is false for $n = 2$. Thus in the nomenclature of (5-17), each degree $d_\ell = 1$, and the estimate (5-18) is replaced by $(Q^2/B^{1-\alpha})^{n-1}$. Thus (5-19) is replaced by

$$Q^n\big(QP^{-2}(\log B)^2 + QB^{-\alpha(M-1)} + B^{(n-1)\alpha+1}Q^{-1}\big) \ll Q^{n+1},$$

upon taking $\alpha = 0$ and using $Q \gg B^{1/2}$. Ultimately, arguing in this way for $n = 2$ leads to the choice $Q = B^{1/2}(\log B)^{1/2}$ and the outcome $S(F, B) \ll B^{n-1+1/2}(\log B)^{1/2}$, which is essentially no better than (1-16), aside from the fact that we can remove the dependence on $\|F\|$ in the implicit constant. In any case, Broberg's results (1-14) and (1-15) supersede the outcome of the methods of this paper for $n = 2, 3$.

6. Proof of Proposition 5.2

In this section we prove the critical Proposition 5.2 that allows us to deduce all factors in $G(0, \boldsymbol{U})$ have at least degree 2, so that we can apply the nontrivial bounds of Heath-Brown and Pila in (5-18). We thank Per Salberger for suggesting the following strategy to prove the proposition.

Let $n \geq 3$. Suppose to the contrary that $G(0, \boldsymbol{U})$ contains a linear factor, that is,

$$G(0, \boldsymbol{U}) = L(\boldsymbol{U})\tilde{H}(\boldsymbol{U}) \tag{6-1}$$

for some linear form L. Then by a linear change of variables we can reduce to the case in which we may assume that $L(\boldsymbol{U}) = U_1$, and conclude that

$$G(0, \boldsymbol{U}) = U_1 H(\boldsymbol{U})$$

for some homogeneous polynomial H. Then *any* point $(0, 0, u_2, \dots, u_n) \in \{U_Y = U_1 = 0\} \subset \mathbb{P}^n$ satisfies $G(0, \boldsymbol{U}) = 0$ and thus defines a tangent hyperplane to $V(F(Z^e, \boldsymbol{X})) \subset \mathbb{P}^n$, given by

$$u_2 X_2 + \cdots + u_n X_n = 0.$$

In particular, for all $[u_2 : \cdots : u_n] \in \mathbb{P}^{n-2}$, this hyperplane contains the line ℓ given by $X_2 = \cdots = X_n = 0$ in $\mathbb{P}^n$. We note that this line ℓ is not contained in $V(F(Z^e, \boldsymbol{X}))$, since for example in the coordinates $[U_Y : U_1 : U_2 : \cdots : U_n]$ we see that the point $[1 : 0 : 0 : \cdots : 0] \in \ell$ but $[1 : 0 : 0 : \cdots : 0] \notin V$, since in the definition of F the coefficient of Z^{mde} is 1. Thus under the assumption (6-1) we have shown that the generic hyperplane through ℓ is tangent to $V(F(Z^e, \boldsymbol{X}))$. We will see this is impossible, and our assumption (6-1) is false (so that Proposition 5.2 is verified), by the following proposition.

Proposition 6.1. *Let $n \geq 3$. Let $X \subset \mathbb{P}^n$ be a nonsingular hypersurface and let ℓ be a line not contained in X. Then the generic hyperplane in $\mathbb{P}^n$ containing ℓ is not tangent to X.*

Let X be given as in the proposition. Without loss of generality we can make a change of coordinates so that

$$\ell = \{X_2 = \cdots = X_n = 0\}.$$

Let $F \in \mathbb{C}[X_0, X_1, \ldots, X_n]$ be such that $X = \{F = 0\}$, and let D denote the degree of F. Our strategy is to construct the blow-up of X along the zero-dimensional subvariety $Z \subset X$, where we define

$$Z = \ell \cap X \subset \mathbb{P}^n.$$

Under the hypothesis that ℓ is not contained in X, then $\deg Z \leq D$. We also define the open set

$$U := X \setminus Z.$$

To prove the proposition, we first notice that we can parametrize the hyperplanes containing ℓ in $\mathbb{P}^n$ by points in $\mathbb{P}^{n-2}$ using the map

$$\mathbb{P}^{n-2} \to \{H \subset \mathbb{P}^n : \deg H = 1,\ \ell \subset H\}, \quad [v_2 : \cdots : v_n] \mapsto \{v_2 X_2 + \cdots + v_n X_n = 0\}.$$

Thus, it will suffice to show that there exists an open set $V \subset \mathbb{P}^{n-2}$ such that for all $\boldsymbol{v} = [v_2 : \cdots : v_n] \in V$,

$$X \cap \{v_2 X_2 + \cdots + v_n X_n = 0\}$$

is smooth, so that in particular the hyperplane $\{v_2 X_2 + \cdots + v_n X_n = 0\} \subset \mathbb{P}^n$ is not tangent to X. We will prove this in two steps, first focusing on the intersection of the hyperplane with the open set $U = X \setminus Z$, and then focusing on the intersection of the hyperplane with the finite set of points in Z. In agreement with the citations we apply in what follows, from now on we will use the terminology "regular" for a scheme instead of "smooth." For a nonsingular hypersurface such as X, these notions are identical by the Jacobian criterion [36, Chapter 4, Theorem 2.19 and Example 2.10]; more generally, the notions are equivalent for any algebraic variety over a perfect field, and in particular over $\mathbb{C}$ [36, Chapter 4, Corollary 3.33].

Define a rational map $\varphi : X \dashrightarrow \mathbb{P}^{n-2}$ given by

$$\varphi : [X_0 : X_1 : X_2 : \cdots : X_n] \mapsto [X_2 : \cdots : X_n].$$

This is a regular map on U. We claim that there exists a projective variety $\tilde{Y}$ and two morphisms $\pi : \tilde{Y} \to X$, and $\tilde{\varphi} : \tilde{Y} \to \mathbb{P}^{n-2}$ such that:

(i) The diagram

$$\begin{tikzcd} \tilde{Y} \arrow[d, "\pi"'] \arrow[dr, "\tilde{\varphi}"] & \\ X \arrow[r, dashed, "\varphi"] & \mathbb{P}^{n-2} \end{tikzcd}$$

is commutative.

(ii) The morphism π restricts to an isomorphism $\pi : \pi^{-1}(U) \to U$.

(iii) The projective variety $\tilde{Y}$ is regular.

Let us assume this claim for now and see how to conclude the proof of the proposition. Since $\tilde{Y}$ is regular, we can apply Kleiman's Bertini theorem [23, Chapter III, Corollary 10.9] to the morphism $\tilde{\varphi} : \tilde{Y} \to \mathbb{P}^{n-2}$, and deduce that given a generic hyperplane $H \subset \mathbb{P}^{n-2}$, $\tilde{\varphi}^{-1}(H) \subseteq \tilde{Y}$ is regular. Let us fix one of these generic hyperplanes, and call it

$$H = \{u_2 X_2 + \cdots + u_n X_n = 0\} \subset \mathbb{P}^{n-2}.$$

By the choice of H, $\tilde{\varphi}^{-1}(H) \cap \pi^{-1}(U)$ is nonsingular. Recall that π is an isomorphism when restricted to the open set $\pi^{-1}(U)$. Thus we also learn that

$$\begin{aligned}\pi(\tilde{\varphi}^{-1}(H) \cap \pi^{-1}(U)) = \pi(\tilde{\varphi}^{-1}(H)) \cap U = \varphi^{-1}(H) \cap U \\ = \{[x_0 : x_1 : x_2 : \cdots : x_n] \in U : u_2 x_2 + \cdots + u_n x_n = 0\}\end{aligned}$$

is regular. Since such H are generic in $\mathbb{P}^{n-2}$, we conclude that there is an open set $V_1 \subset \mathbb{P}^{n-2}$ such that for all $\boldsymbol{v} = [v_2 : \cdots : v_n] \in V_1$, the intersection

$$U \cap \{v_2 X_2 + \cdots + v_n X_n = 0\}$$

is regular.

Let us next focus on the intersection of the hyperplane with the set Z. For any $P \in Z$, a hyperplane $\{v_2 X_2 + \cdots + v_n X_n = 0\}$ with $[v_2 : \cdots : v_n] \in \mathbb{P}^{n-2}$ is tangent to X at P if the Jacobian matrix at P,

$$J_{\boldsymbol{v}}(P) = \begin{pmatrix} \frac{\partial F}{\partial X_0}(P) & 0 \\ \frac{\partial F}{\partial X_1}(P) & 0 \\ \frac{\partial F}{\partial X_2}(P) & v_2 \\ \vdots & \vdots \\ \frac{\partial F}{\partial X_n}(P) & v_n \end{pmatrix},$$

has rank ≤ 1. From this it is clear that if either $\frac{\partial F}{\partial X_0}(P) \neq 0$ or $\frac{\partial F}{\partial X_1}(P) \neq 0$ then rank $J_{\boldsymbol{v}}(P) = 2$ for any $\boldsymbol{v} \in \mathbb{P}^{n-2}$. On the other hand, if $\frac{\partial F}{\partial X_0}(P) = \frac{\partial F}{\partial X_1}(P) = 0$ then $\operatorname{rank}_{\boldsymbol{v}}(P) \leq 1$ if and only if $\boldsymbol{v} = \left[\frac{\partial F}{\partial X_2}(P) : \cdots : \frac{\partial F}{\partial X_n}(P)\right]$ since we are assuming that X is a nonsingular hypersurface. For each $P \in Z$ we define

$$C_P = \begin{cases} \left\{\left[\frac{\partial F}{\partial X_2}(P) : \cdots : \frac{\partial F}{\partial X_n}(P)\right]\right\} & \text{if } \frac{\partial F}{\partial X_0}(P) = \frac{\partial F}{\partial X_1}(P) = 0, \\ \varnothing & \text{otherwise.} \end{cases}$$

If we define $V_P = \mathbb{P}^{n-2} \setminus C_P$, it follows that for any $\boldsymbol{v} \in V_P$ the intersection

$$X \cap \{v_2 X_2 + \cdots + v_n X_n = 0\}$$

is regular at P.

Finally consider the set

$$V = V_1 \cap \bigcap_{P \in Z} V_P.$$

Since $\deg Z \leq D$, then V is a nonempty open subset of $\mathbb{P}^{n-2}$. For each $\boldsymbol{v} \in V$, the hyperplane $v_2 x_2 + \cdots + v_n x_n = 0$ contains ℓ, and

$$\{v_2 X_2 + \cdots + v_n X_n = 0\} \cap (U \cup Z) = \{v_2 X_2 + \cdots + v_n X_n = 0\} \cap X$$

is regular, or equivalently, nonsingular; thus $\{v_2 X_2 + \cdots + v_n X_n = 0\}$ is not tangent to X. This completes the proof of Proposition 6.1, except for the proof of properties (i), (ii), and (iii) in the claim.

We now prove the claim of properties (i), (ii) and (iii). From the rational map $\varphi : X \dashrightarrow \mathbb{P}^{n-2}$ given by

$$\varphi : [X_0 : X_1 : X_2 : \cdots : X_n] \mapsto [X_2 : \cdots : X_n],$$

we consider the graph $\Gamma = \Gamma_\varphi$ of the map φ,

$$\Gamma = \{(\boldsymbol{x}, \varphi(\boldsymbol{x})) : \boldsymbol{x} \in U\} \subset X \times \mathbb{P}^{n-2}.$$

Define the Zariski closure $\tilde{X} = \overline{\Gamma} \subset X \times \mathbb{P}^{n-2}$. Define the projection map $\pi' : \tilde{X} \to X$ acting by $(\boldsymbol{x}, \varphi(\boldsymbol{x})) \to (\boldsymbol{x})$. Then the blow-up is $\tilde{X}$ along with a morphism φ' such that

$$\begin{tikzcd} \tilde{X} \arrow[d, "\pi'"'] \arrow[dr, "\varphi'"] & \\ X \arrow[r, dashed, "\varphi"] & \mathbb{P}^{n-2} \end{tikzcd}$$

is a commutative diagram (see, e.g., [22, Chapter 7, p. 82]). Moreover, from the definition of the blow-up it follows that π' restricts to an isomorphism $\pi' : (\pi')^{-1}(U) \to U$, i.e., $\tilde{X}$ satisfies properties (i) and (ii), but it might be singular. To resolve this, we apply Hironaka's resolution of singularities: as a consequence of [30, Theorem 1] (see also [30, p. 112]), there is a projective variety $\tilde{Y}$ and a morphism $f : \tilde{Y} \to \tilde{X}$ such that f is an isomorphism when restricted to the inverse image $f^{-1}(V)$ of the open set V of the regular points of $\tilde{X}$, and such that $\tilde{Y}$ is regular. Then the claim follows by taking $\pi = \pi' \circ f$, $\tilde{\varphi} = \varphi' \circ f$ and observing that $(\pi')^{-1}(U) \subset V$.

7. Concluding arguments

In Section 5 we proved that the contribution of the bad-bad terms to the sieve is $\ll Q^n$. We now turn to analyzing the contributions of the other types, as defined in Definition 4.1. We will treat these in three sections; in each case we apply the relevant bound for $|g(\boldsymbol{u}, pq)|$ from Proposition 4.2 and the bound (5-4) for $\hat{W}$. Once we have treated these cases, we proceed in Section 7.4 to choose the parameter Q, and conclude the proof of Theorem 1.1.

7.1. ***Zero-type cases.*** We first consider any case in which $\boldsymbol{u}$ is zero-type modulo p, divided into cases according to whether $\boldsymbol{u}$ is zero-type, good, or bad modulo q. The contribution of the first case (upon

setting $\boldsymbol{u} = pq\boldsymbol{v}$ and applying (5-4)) is

$$\frac{1}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ zero mod }p\\ \boldsymbol{u}\text{ zero mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \ll \frac{Q^{2n-1}}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\boldsymbol{v}\in\mathbb{Z}^n} \left|\hat{W}(\boldsymbol{v})\right| \ll B^nQ^{-1}.$$

The contribution of the second case (upon setting $\boldsymbol{u} = p\boldsymbol{v}$, applying (5-4) with $L = Q < B$) is

$$\frac{1}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ zero mod }p\\ \boldsymbol{u}\text{ good mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \ll \frac{Q^{n-1/2}Q^{n/2}P^2}{P^2Q^{2n}} \sum_{\boldsymbol{v}\in\mathbb{Z}^n} \left|\hat{W}\left(\frac{\boldsymbol{v}}{Q}\right)\right| \ll B^nQ^{-n/2-1/2}.$$

The contribution of the third case (upon setting $\boldsymbol{u} = p\boldsymbol{v}$, applying (5-4) with $L = Q < B$) is

$$\frac{1}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ zero mod }p\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \ll \frac{Q^{n-1/2}Q^{n/2+1/2}P^2}{P^2Q^{2n}} \sum_{\boldsymbol{v}\in\mathbb{Z}^n} \left|\hat{W}\left(\frac{\boldsymbol{v}}{Q}\right)\right| \ll B^nQ^{-n/2}.$$

As long as $n \geq 2$, all these cases contribute at most $\ll B^nQ^{-1}$ to the sieve, which is acceptable.

7.2. ***Good-good case.*** The contribution to the sieve from the good-good case is:

$$\frac{1}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ good mod }p\\ \boldsymbol{u}\text{ good mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \ll \frac{Q^nP^2}{P^2Q^{2n}} \sum_{\boldsymbol{u}\in\mathbb{Z}^n} \left|\hat{W}\left(\frac{\boldsymbol{u}}{Q^2}\right)\right| \ll Q^n,$$

after applying (5-4) with $L = Q^2 > B$, since under the assumption (5-20), $\kappa \geq 1/2$.

7.3. ***Good-bad case.*** The contribution to the sieve from the good-bad case is

$$\frac{1}{P^2Q^{2n}} \sum_{\substack{p,q\in\mathcal{P}\\ p\neq q}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ good mod }p\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right) g(\boldsymbol{u}, pq)\right| \ll \frac{Q^{n+1/2}}{P^2Q^{2n}} \sum_{p\in\mathcal{P}} \sum_{q\neq p\in\mathcal{P}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right)\right|. \tag{7-1}$$

Here we proceed by imitating the key step from Section 5 for the bad-bad case, and sum over q before summing over $\boldsymbol{u}$. We again define $G(U_Y, \boldsymbol{U})$ as in (5-6), and let $R(\boldsymbol{u})$ denote the resultant (5-9), so that

$$\sum_{p\in\mathcal{P}} \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})\neq 0}} \sum_{\substack{q\neq p\in\mathcal{P}\\ \boldsymbol{u}\text{ bad mod }q}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{pq}\right)\right| \ll P \sum_{\substack{\boldsymbol{u}\in\mathbb{Z}^n\\ G(0,\boldsymbol{u})\neq 0}} \left|\hat{W}\left(\frac{\boldsymbol{u}}{Q^2}\right)\right| \omega(R(\boldsymbol{u})) \ll_{n,m,e,d} PQ^{2n}\log B,$$

with an implied constant independent of $\|F\|$ (in the first case of Lemma 2.1), by arguing as in the proof of (5-12).

Notice that in the good-bad case, we do not need to consider a possible contribution from those $\boldsymbol{u}$ for which $G(0, \boldsymbol{u}) = 0$: when $G(0, \boldsymbol{u}) = 0$, then all q have the property that $\boldsymbol{u}$ is bad for q, whereas by

definition in the good-bad case, $\boldsymbol{u}$ is good for at least one prime. In total, the contribution to the sieve from the good-bad case is thus

$$\frac{Q^{n+1/2}}{P^2 Q^{2n}} \cdot PQ^{2n}(\log B) \ll Q^{n+1/2} P^{-1}(\log B) \ll Q^n,$$

since $Q = B^\kappa$ for some $1/2 \le \kappa \le 1$ and under our acting assumption (4-4), by (4-5), $P \gg Q/\log Q$. Thus we can conclude that the total contribution of the good-bad case (7-1) of the sieve is $\ll Q^n$, with an implied constant independent of $\|F\|$ (in the first case of Lemma 2.1).

7.4. ***Final conclusion of the sieve and choice of parameters.*** We now assemble all the terms of the main sieve term in (5-2): we can conclude that

$$\frac{1}{P^2} \sum_{\substack{p,q \in \mathcal{P} \\ p \neq q}} |T(p,q)| \ll B^n Q^{-1} + Q^n. \tag{7-2}$$

The first term is from all zero-type cases, and the last term includes the good-good, good-bad, and bad-bad cases. We apply this in the sieve lemma, along with the bound (5-1) for the two simple terms in the sieve, to conclude that (in the first case of Lemma 2.1) our counting function admits the bound

$$\mathcal{S}(F,B) \ll_{n,m,e,d} (B^{n-1} + B^n P^{-1} + B^n Q^{-1} + Q^n) \ll (B^n P^{-1} + Q^n). \tag{7-3}$$

Choose

$$Q = B^{n/(n+1)} (\log B)^{1/(n+1)}. \tag{7-4}$$

The requirement (5-20) is met for all $n \ge 3$. (If $n = 2$, then this argument leads to the choice $Q \approx B^{2/3}$, which does not suffice to prove sufficient decay in the bad-bad case; see Remark 5.4.) Recall from (4-4) and (4-5) that

$$P = |\mathcal{P}| \gg_{m,e,d} Q(\log Q)^{-1} \gg_{n,m,e,d} B^{\frac{n}{n+1}} (\log B)^{-\frac{n}{n+1}}$$

as long as

$$Q \gg_{m,e,d} (\log \|F\|)(\log\log \|F\|). \tag{7-5}$$

Recall also that we require $P \gg_{m,e,d} \max\{\log \|f_d\|, \log B\}$ in Lemma 1.2. Certainly the first condition is satisfied under the assumption (7-5). The second condition is satisfied for Q as in (7-4) for all $B \gg_n 1$.

To meet the requirement (7-5) for Q as chosen in (7-4), it suffices to require that

$$B \gg_{m,e,d} (\log \|F\| \log\log \|F\|)^{\frac{n+1}{n}}.$$

For such B, the conclusion of the sieve process in (7-3) shows that

$$\mathcal{S}(F,B) \ll_{n,m,e,d} B^{n-1+\frac{1}{n+1}} (\log B)^{\frac{n}{n+1}},$$

where the implicit constant is independent of $\|F\|$. This suffices for Theorem 1.1. Finally, for all $B \ll_{m,e,d} (\log\|F\|\log\log\|F\|)^{\frac{n+1}{n}}$, we apply the trivial bound

$$\mathcal{S}(F,B) \ll_n B^n \ll_{n,m,e,d} (\log\|F\|\log\log\|F\|)^{n+1} \ll (\log\|F\|)^{n+2}$$
$$\ll_{n,m,d,e} (\log B)^{n+2} \ll_n B^{n-1+\frac{1}{n+1}}(\log B)^{\frac{n}{n+1}}.$$

Here we applied the fact from Lemma 2.1 that in the case it remains to prove Theorem 1.1, $\|F\| \ll B^{(mde)^{n+2}}$ so that $\log\|F\| \ll_{n,m,d,e} \log B$. This completes the proof of Theorem 1.1.

Acknowledgements

The authors thank T. Browning for suggesting the application of the polynomial sieve to smooth coverings and for useful discussions, and J. Lyczak for many helpful remarks. In addition, the authors thank P. Salberger for suggesting a strategy to prove Proposition 5.2, and both Salberger and an anonymous referee for helpful remarks on an earlier version of this manuscript. The authors credit ChatGPT with expository edits to the last two paragraphs of Section 1.2.4.

Bonolis has been supported by FWF grant P 32428-N35. Pierce has been partially supported by NSF DMS-2200470 and NSF CAREER grant DMS-1652173, a Sloan Research Fellowship, and a Joan and Joseph Birman Fellowship. The authors thank the Hausdorff Center for Mathematics for hosting a productive collaboration visit and the RTG DMS-2231514. Pierce thanks HCM for hosting a visit as a Bonn Research Chair.

References

[1] E. Bombieri and J. Pila, "The number of integral points on arcs and ovals", *Duke Math. J.* **59**:2 (1989), 337–357. MR Zbl

[2] D. Bonolis, "A polynomial sieve and sums of Deligne type", *Int. Math. Res. Not.* **2021**:2 (2021), 1096–1137. MR Zbl

[3] D. Bonolis and T. Browning, "Uniform bounds for rational points on hyperelliptic fibrations", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **24**:1 (2023), 173–204. Correction in **24**:4 (2023), 2501–2504. MR Zbl

[4] J. Brandes, "Sums and differences of power-free numbers", *Acta Arith.* **169**:2 (2015), 169–180. MR Zbl

[5] N. Broberg, "Rational points on finite covers of $\mathbb{P}^1$ and $\mathbb{P}^2$", *J. Number Theory* **101**:1 (2003), 195–207. MR Zbl

[6] T. D. Browning, "A note on the distribution of rational points in threefolds", *Quart. J. Math.* **54** (2003), 33–39.

[7] T. D. Browning, *Quantitative arithmetic of projective varieties*, Progr. Math. **277**, Birkhäuser, Basel, 2009. MR Zbl

[8] T. D. Browning, "The polynomial sieve and equal sums of like polynomials", *Int. Math. Res. Not.* **2015**:7 (2015), 1987–2019. MR Zbl

[9] T. D. Browning and D. R. Heath-Brown, "The density of rational points on non-singular hypersurfaces, I", *Bull. Lond. Math. Soc.* **38**:3 (2006), 401–410. MR Zbl

[10] T. D. Browning and D. R. Heath-Brown, "The density of rational points on non-singular hypersurfaces, II", *Proc. Lond. Math. Soc.* (3) **93**:2 (2006), 273–303. MR Zbl

[11] T. D. Browning and P. Vishe, "Rational points on cubic hypersurfaces over $\mathbb{F}_q(t)$", *Geom. Funct. Anal.* **25**:3 (2015), 671–732. MR Zbl

[12] T. D. Browning, D. R. Heath-Brown, and P. Salberger, "Counting rational points on algebraic varieties", *Duke Math. J.* **132**:3 (2006), 545–578. MR Zbl

[13] A. Bucur, A. C. Cojocaru, M. N. Lalín, and L. B. Pierce, "Geometric generalizations of the square sieve, with an application to cyclic covers", *Mathematika* **69**:1 (2023), 106–154. MR Zbl

[14] W. Castryck, R. Cluckers, P. Dittmann, and K. H. Nguyen, "The dimension growth conjecture, polynomial in the degree and without logarithmic factors", *Algebra Number Theory* **14**:8 (2020), 2261–2294. MR Zbl

[15] M. Chardin, "The resultant via a Koszul complex", pp. 29–39 in *Computational algebraic geometry* (Nice, Italy, 1992), edited by F. Eyssette and A. Galligo, Progr. Math. **109**, Birkhäuser, Boston, MA, 1993. MR Zbl

[16] S. D. Cohen, "The distribution of Galois groups and Hilbert's irreducibility theorem", *Proc. Lond. Math. Soc.* (3) **43**:2 (1981), 227–250. MR Zbl

[17] D. A. Cox, J. Little, and D. O'Shea, *Using algebraic geometry*, 2nd ed., Grad. Texts in Math. **185**, Springer, 2005. MR Zbl

[18] M. Demazure, "Résultant, discriminant", *Enseign. Math.* (2) **58**:3-4 (2012), 333–373. MR Zbl

[19] D. Eisenbud and J. Harris, 3264 *and all that: a second course in algebraic geometry*, Cambridge Univ. Press, 2016. MR Zbl

[20] W. Fulton and R. Lazarsfeld, "Connectivity and its applications in algebraic geometry", pp. 26–92 in *Algebraic geometry* (Chicago, IL, 1980), edited by A. Libgober and P. Wagreich, Lecture Notes in Math. **862**, Springer, 1981. MR Zbl

[21] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Birkhäuser, Boston, MA, 1994. MR Zbl

[22] J. Harris, *Algebraic geometry: a first course*, Grad. Texts in Math. **133**, Springer, 1992. MR Zbl

[23] R. Hartshorne, *Algebraic geometry*, Grad. Texts in Math. **52**, Springer, 1977. MR Zbl

[24] D. R. Heath-Brown, "Cubic forms in ten variables", *Proc. Lond. Math. Soc.* (3) **47**:2 (1983), 225–257. MR Zbl

[25] D. R. Heath-Brown, "The square sieve and consecutive square-free numbers", *Math. Ann.* **266**:3 (1984), 251–259. MR Zbl

[26] D. R. Heath-Brown, "The density of rational points on nonsingular hypersurfaces", *Proc. Indian Acad. Sci. Math. Sci.* **104**:1 (1994), 13–29. MR Zbl

[27] D. R. Heath-Brown, "The density of rational points on curves and surfaces", *Ann. of Math.* (2) **155**:2 (2002), 553–595. MR Zbl

[28] D. R. Heath-Brown, "Imaginary quadratic fields with class group exponent 5", *Forum Math.* **20**:2 (2008), 275–283. MR Zbl

[29] D. R. Heath-Brown and L. B. Pierce, "Counting rational points on smooth cyclic covers", *J. Number Theory* **132**:8 (2012), 1741–1757. MR Zbl

[30] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero, I", *Ann. of Math.* (2) **79**:1 (1964), 109–203. MR Zbl

[31] C. Hooley, "On the representations of a number as the sum of four cubes, I", *Proc. Lond. Math. Soc.* (3) **36**:1 (1978), 117–140. MR Zbl

[32] C. Hooley, "On the number of points on a complete intersection over a finite field", *J. Number Theory* **38**:3 (1991), 338–358. MR Zbl

[33] N. M. Katz, "Estimates for 'singular' exponential sums", *Int. Math. Res. Not.* **1999**:16 (1999), 875–899. MR Zbl

[34] J. C. Lagarias and A. M. Odlyzko, "Effective versions of the Chebotarev density theorem", pp. 409–464 in *Algebraic number fields*: *L-functions and Galois properties* (Durham, 1975), edited by A. Fröhlich, Academic Press, London, 1977. MR Zbl

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

[36] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Grad. Texts in Math. **6**, Oxford Univ. Press, 2002. MR Zbl

[37] D. A. Marcus, *Number fields*, Springer, 1977. MR Zbl

[38] J. S. Milne, "Algebraic number theory", preprint, version 3.08, 2020, available at https://www.jmilne.org/math/CourseNotes/ant.html.

[39] R. Munshi, "Density of rational points on cyclic covers of $\mathbb{P}^n$", *J. Théor. Nombres Bordeaux* **21**:2 (2009), 335–341. MR Zbl

[40] L. B. Pierce, "A bound for the 3-part of class numbers of quadratic fields by means of the square sieve", *Forum Math.* **18**:4 (2006), 677–698. MR Zbl

[41] J. Pila, "Density of integral and rational points on varieties", pp. 183–187 in *Columbia University Number Theory Seminar* (New York, 1992), Astérisque **228**, Soc. Math. France, Paris, 1995. MR Zbl

[42] P. Salberger, "On the density of rational and integral points on algebraic varieties", *J. Reine Angew. Math.* **606** (2007), 123–147. MR Zbl

[43] P. Salberger, "Counting rational points on projective varieties", *Proc. Lond. Math. Soc.* (3) **126**:4 (2023), 1092–1133. MR Zbl

[44] J.-P. Serre, "Quelques applications du théorème de densité de Chebotarev", *Inst. Hautes Études Sci. Publ. Math.* **54** (1981), 323–401. MR Zbl

[45] J.-P. Serre, *Topics in Galois theory*, Res. Notes in Math. **1**, Jones & Bartlett, Boston, MA, 1992. MR Zbl

[46] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, 3rd ed., Vieweg & Sohn, Braunschweig, Germany, 1997. MR Zbl

Communicated by Philippe Michel
Received 2022-09-06 Revised 2023-06-19 Accepted 2023-10-31

dante.bonolis@duke.edu *Mathematics Department, Duke University, Durham, NC 27708, United States*

pierce@math.duke.edu *Mathematics Department, Duke University, Durham, NC 27708, United States*

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Antoine Chambert-Loir
Université Paris-Diderot
France

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Jason P. Bell	University of Waterloo, Canada
Bhargav Bhatt	University of Michigan, USA
Frank Calegari	University of Chicago, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Saclay, France
Brian D. Conrad	Stanford University, USA
Samit Dasgupta	Duke University, USA
Hélène Esnault	Freie Universität Berlin, Germany
Gavril Farkas	Humboldt Universität zu Berlin, Germany
Sergey Fomin	University of Michigan, USA
Edward Frenkel	University of California, Berkeley, USA
Wee Teck Gan	National University of Singapore
Andrew Granville	Université de Montréal, Canada
Ben J. Green	University of Oxford, UK
Christopher Hacon	University of Utah, USA
Roger Heath-Brown	Oxford University, UK
János Kollár	Princeton University, USA
Michael J. Larsen	Indiana University Bloomington, USA
Philippe Michel	École Polytechnique Fédérale de Lausanne
Martin Olsson	University of California, Berkeley, USA
Irena Peeva	Cornell University, USA
Jonathan Pila	University of Oxford, UK
Anand Pillay	University of Notre Dame, USA
Bjorn Poonen	Massachusetts Institute of Technology, USA
Victor Reiner	University of Minnesota, USA
Peter Sarnak	Princeton University, USA
Michael Singer	North Carolina State University, USA
Vasudevan Srinivas	SUNY Buffalo, USA
Shunsuke Takagi	University of Tokyo, Japan
Pham Huu Tiep	Rutgers University, USA
Ravi Vakil	Stanford University, USA
Akshay Venkatesh	Institute for Advanced Study, USA
Melanie Matchett Wood	Harvard University, USA
Shou-Wu Zhang	Princeton University, USA

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2024 is US $525/year for the electronic version, and $770/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2024 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 18 No. 8 2024

The strong maximal rank conjecture and moduli spaces of curves 1403
FU LIU, BRIAN OSSERMAN, MONTSERRAT TEIXIDOR I BIGAS and NAIZHEN ZHANG

Unramifiedness of weight 1 Hilbert Hecke algebras 1465
SHAUNAK V. DEO, MLADEN DIMITROV and GABOR WIESE

Failure of the local-global principle for isotropy of quadratic forms over function fields 1497
ASHER AUEL and V. SURESH

Application of a polynomial sieve: beyond separation of variables 1515
DANTE BONOLIS and LILLIAN B. PIERCE

Functorial embedded resolution via weighted blowings up 1557
DAN ABRAMOVICH, MICHAEL TEMKIN and JAROSŁAW WŁODARCZYK