Forum of Mathematics, Sigma (2018), Vol. 6, e19, 89 pages
doi:10.1017/fms.2018.21

PRIME SOLUTIONS TO POLYNOMIAL EQUATIONS IN MANY VARIABLES AND DIFFERING DEGREES

SHUNTARO YAMAGISHI
Department of Mathematics and Statistics, Queen's University, Kingston, Ontario K7L 3N6, Canada;
email: sy46@queensu.ca

Received 20 February 2018; accepted 2 September 2018

Abstract

Let $\mathbf{f} = (f_1, \ldots, f_R)$ be a system of polynomials with integer coefficients in which the degrees need not all be the same. We provide sufficient conditions for which the system of equations $f_j(x_1, \ldots, x_n) = 0$ $(1 \leqslant j \leqslant R)$ satisfies a general local to global type statement, and has a solution where each coordinate is prime. In fact we obtain the asymptotic formula for number of such solutions, counted with a logarithmic weight, under these conditions. We prove the statement via the Hardy–Littlewood circle method. This is a generalization of the work of Cook and Magyar ['Diophantine equations in the primes', *Invent. Math.* **198** (2014), 701–737], where they obtained the result when the polynomials of $\mathbf{f}$ all have the same degree. Hitherto, results of this type for systems of polynomial equations involving different degrees have been restricted to the diagonal case.

2010 Mathematics Subject Classification: 11P32, 11P55 (primary); 11D45, 11D72 (secondary)

1. Introduction

Let $d \geqslant 1$, and let $\mathbf{f} = (\mathbf{f}_d, \ldots, \mathbf{f}_1)$ be a system of polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$, where $\mathbf{f}_\ell = (f_{\ell,1}, \ldots, f_{\ell,r_\ell})$ is the subsystem of degree ℓ polynomials of $\mathbf{f}$ $(1 \leqslant \ell \leqslant d)$. We are interested in finding prime solutions, which are solutions with each coordinate a prime number, to the equations

$$f_{\ell,r}(x_1, \ldots, x_n) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell). \tag{1.1}$$

Let us denote $V_{\mathbf{f},\mathbf{0}}(\mathbb{C})$ to be the affine variety in $\mathbb{C}^n$ defined by the equations (1.1).

Solving diophantine equations in primes is a fundamental problem in number theory. For example, the celebrated work of Green and Tao [10] on arithmetic progressions in primes can be phrased as the statement that given any $n \in \mathbb{N}$ the system of linear equations

$$x_{i+2} - x_{i+1} = x_{i+1} - x_i \quad (1 \leqslant i \leqslant n)$$

has a prime solution $(x_1, \ldots, x_{n+2}) = (p_1, \ldots, p_{n+2})$ where $p_1 < p_2 < \cdots < p_{n+2}$. The modern results on the large scale distribution of prime solutions on $V_{\mathbf{f},\mathbf{0}}(\mathbb{C})$ when $\mathbf{f}$ consists only of linear polynomials, for scenarios which do not reduce to a binary problem, is mostly summed up in the work of Green and Tao [12]. An example of a binary problem is bounding gaps between primes, an area which Maynard [19], Tao (see [19, page 385]), and Zhang [24] made significant progress in by building on the work of Goldston, Pintz, and Yıldırım [9]. In particular, it was shown in [19] that at least one of the equations

$$x_1 - x_2 = 2j \quad (1 \leqslant j \leqslant 300)$$

has infinitely many prime solutions. Another binary problem of significance is the Goldbach's conjecture, which states that the equation

$$x_1 + x_2 = N$$

has a prime solution for every even integer N greater than 2. It was proved by Vinogradov [21] that the equation

$$x_1 + x_2 + x_3 = N \tag{1.2}$$

has a prime solution for all sufficiently large odd $N \in \mathbb{N}$. The ternary Goldbach problem, which is the assertion that the equation (1.2) has a prime solution for all odd $N \in \mathbb{N}$ greater than or equal to 7, was solved by Helfgott in [13].

The examples given thus far had been for systems of linear equations. The scenario for systems involving higher degree polynomials is also complex, and has not been well-understood yet. Indeed, even the problem of whether a system of nonlinear polynomial equations has a solution over $\mathbb{Q}$ is 'one of considerable complexity' [4].

For solving nonlinear equations in primes, there are results due to Hua [14] for certain systems of homogeneous polynomials that are additive, for example on the system of the shape $x_1^j + \cdots + x_n^j = N_j$ $(1 \leqslant j \leqslant d)$ where $N_j \in \mathbb{N}$. Hua also has results on the Waring–Goldbach problem, which is regarding prime solutions of the equation $x_1^d + \cdots + x_n^d = N$ where $N \in \mathbb{N}$. These results were established via the Hardy–Littlewood circle method. We refer the reader to [16, 17] for recent progress on the Waring–Goldbach problem due to Kumchev and Wooley. There is also [5] by Chow regarding prime solutions of certain diagonal equations by a

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

transference principle approach. For the case of quadratic forms, there is a result due to Zhao [25].

The first result regarding prime solutions of general systems of nonlinear polynomials is contained in the breakthrough of Cook and Magyar [6], which we state in Theorem 1.1. Before we can state their result we need to introduce some notations. We also note that there is a discussion in [6] on this topic from the point of view of some recent results in sieve theory, which the list includes [2, 8, 18]. We refer the reader to [6] for more details on this discussion.

Let $\ell > 1$. Let $\mathbf{G} = (G_1, \dots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \dots, x_n]$. We define the *Birch singular locus* $V^*_{\mathbf{G}}$ to be the set of points in $\mathbb{C}^n$ given by

$$\operatorname{rank}\left(\frac{\partial G_r(\mathbf{x})}{\partial x_j}\right)_{\substack{1\leqslant r\leqslant r'\\ 1\leqslant j\leqslant n}} < r'. \tag{1.3}$$

Observe that this defines an affine variety in $\mathbb{C}^n$. We define the *Birch rank*, $\mathcal{B}_\ell(\mathbf{G})$, to be the codimension of $V^*_{\mathbf{G}}$. Given $\mathbf{g} = (g_1, \dots, g_{r'})$, a system of degree ℓ polynomials in $\mathbb{Q}[x_1, \dots, x_n]$, where G_r is the homogeneous degree ℓ portion of g_r $(1 \leqslant r \leqslant r')$, we extend the notion of the Birch rank to systems of degree ℓ polynomials by defining

$$\mathcal{B}_\ell(\mathbf{g}) := \mathcal{B}_\ell(\mathbf{G}).$$

When $\ell = 1$, following [6] we define $\mathcal{B}_1(\mathbf{g})$ to be the minimum number of nonzero coefficients in a nontrivial linear combination

$$\lambda_1 G_1 + \cdots + \lambda_{r'} G_{r'},$$

where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{r'}) \in \mathbb{Q}^{r'}\backslash\{\mathbf{0}\}$. Clearly $\mathcal{B}_1(\mathbf{g}) > 0$ if and only if the linear forms $G_1, \dots, G_{r'}$ are linearly independent over $\mathbb{Q}$. For any $\ell \geqslant 1$, if $r' = 0$ then we let $\mathcal{B}_\ell(\mathbf{g}) = +\infty$.

Let $\mathbf{F} = (\mathbf{F}_d, \dots, \mathbf{F}_1)$ be the system of homogeneous polynomials, where for each $1 \leqslant \ell \leqslant d$, $\mathbf{F}_\ell = (F_{\ell,1}, \dots, F_{\ell,r_\ell})$ and $F_{\ell,r}$ is the homogeneous degree ℓ portion of $f_{\ell,r}$ in (1.1). We let $V_{\mathbf{F},\mathbf{0}}(\mathbb{R})$ be the set of points in $\mathbb{R}^n$ satisfying

$$F_{\ell,r}(x_1, \dots, x_n) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell). \tag{1.4}$$

Let Λ be the von Mangoldt function, where $\Lambda(x)$ is $\log p$ if x is a power of prime p, and 0 otherwise. Given $\mathbf{x} = (x_1, \dots, x_n)$, we let

$$\Lambda(\mathbf{x}) = \Lambda(x_1) \cdots \Lambda(x_n). \tag{1.5}$$

We define the following quantity

$$\mathcal{M}_{\mathbf{f}}(X) := \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x})\mathbf{1}_{V_{\mathbf{f},\mathbf{0}}(\mathbb{C})}(\mathbf{x}), \tag{1.6}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

where $\mathbf{1}_{V_{\mathbf{f},\mathbf{0}}(\mathbb{C})}$ is the characteristic function of the set $V_{\mathbf{f},\mathbf{0}}(\mathbb{C})$. Thus the quantity $\mathcal{M}_{\mathbf{f}}(X)$ is the number of solutions, counted with a logarithmic weight, of the equations (1.1) in $[0, X]^n$ whose coordinates are all prime powers.

We may now phrase the main result of Cook and Magyar in [6], which is for the case when the polynomials of $\mathbf{f}$ in (1.1) all have the same degree.

THEOREM 1.1 [6, Theorem 1]. *Let* $\mathbf{f} = \mathbf{f}_d = (f_{d,1}, \ldots, f_{d,r_d})$ *be a system of degree* d *polynomials in* $\mathbb{Z}[x_1, \ldots, x_n]$*. Suppose* $\mathcal{B}_d(\mathbf{f})$ *is sufficiently large with respect to* d *and* r_d*. Then there exist* $\mathcal{C}(\mathbf{f})$*, a constant which depends only on* $\mathbf{f}$*, and* $c > 0$ *such that*

$$\mathcal{M}_{\mathbf{f}}(X) = \mathcal{C}(\mathbf{f}) X^{n-dr_d} + O\left(\frac{X^{n-dr_d}}{(\log X)^c}\right).$$

In this paper, we generalize Theorem 1.1 to handle systems of polynomials in which the degrees need not all be the same. The following is the main theorem of this paper.

THEOREM 1.2. *Let* $\mathbf{f} = (\mathbf{f}_d, \ldots, \mathbf{f}_1)$ *be a system of polynomials in* $\mathbb{Z}[x_1, \ldots, x_n]$*, where* $\mathbf{f}_\ell = (f_{\ell,1}, \ldots, f_{\ell,r_\ell})$ *is the subsystem of degree* ℓ *polynomials of* $\mathbf{f}$ $(1 \leqslant \ell \leqslant d)$*. For each* $1 \leqslant \ell \leqslant d$*, suppose* $\mathcal{B}_\ell(\mathbf{f}_\ell)$ *is sufficiently large with respect to* d *and* $r_d, \ldots, r_1$*. Then there exist* $\mathcal{C}(\mathbf{f})$*, a constant which depends only on* $\mathbf{f}$*, and* $c > 0$ *such that*

$$\mathcal{M}_{\mathbf{f}}(X) = \mathcal{C}(\mathbf{f}) X^{n-\sum_{\ell=1}^{d} \ell r_\ell} + O\left(\frac{X^{n-\sum_{\ell=1}^{d} \ell r_\ell}}{(\log X)^c}\right).$$

Note Theorem 1.2 recovers Theorem 1.1 when $r_1 = \cdots = r_{d-1} = 0$. We also prove in Section 7 that if the equations (1.1) has a nonsingular solution in $\mathbb{Z}_p^\times$, the units of p-adic integers, for each prime p, and $V_{\mathbf{F},\mathbf{0}}(\mathbb{R})$ has a nonsingular real point in $(0, 1)^n$, then

$$\mathcal{C}(\mathbf{f}) > 0.$$

We also present Theorem 8.1 in Section 8, where we obtain the asymptotic formula for the number of prime solutions, counted with a logarithmic weight, instead of solutions whose coordinates are all prime powers as in Theorem 1.2. Hitherto, the only examples in the literature of results of this type, for systems of polynomial equations involving different degrees, have been restricted to the diagonal case similar to the aforementioned result of Hua.

Theorems 1.1 and 1.2 are both obtained via the Hardy–Littlewood circle method. Circle method was pioneered by Hardy and Littlewood to give an asymptotic formula for the number of solutions to Waring's problem, and it has been quite effective at producing asymptotic formulas for the number of integer points of bounded height on varieties when the number of variables is sufficiently large. The results of this type on the distribution of integer points on varieties are provided by Birch [1] and Schmidt [20]. In [3], Browning and Heath-Brown succeeded in generalizing the seminal work of Birch [1], and showed 'how forms of unequal degrees can be handled in an efficient manner, so as to give the results in the spirit of Birch for arbitrary systems.' They obtained the following result.

THEOREM 1.3 [3, Theorem 1.2]. *Let* $\mathbf{f} = (\mathbf{f}_d, \ldots, \mathbf{f}_1)$ *be a system of polynomials in* $\mathbb{Z}[x_1, \ldots, x_n]$*, where* $\mathbf{f}_\ell = (f_{\ell,1}, \ldots, f_{\ell,r_\ell})$ *is the subsystem of degree* ℓ *polynomials of* $\mathbf{f}$ $(1 \leqslant \ell \leqslant d)$*. Let* $\mathcal{D}_j = r_1 + 2r_2 + \cdots + jr_j$ $(1 \leqslant j \leqslant d)$ *and* $\mathcal{D}_0 = 0$*. Let*

$$s_\ell = \sum_{k=\ell}^{d} \frac{2^{\ell-1}(\ell-1)r_\ell}{\mathcal{B}_\ell(\mathbf{f}_\ell)} \quad (1 \leqslant \ell \leqslant d).$$

Let $r_1 = 0$ *and* $r_d \geqslant 1$*, and suppose we have*

$$\mathcal{D}_\ell \left(\frac{2^{\ell-1}}{\mathcal{B}_\ell(\mathbf{f}_\ell)} + s_{\ell+1} \right) + s_{\ell+1} + \sum_{j=\ell+1}^{d} s_j r_j < 1$$

for $\ell = 0$ *and for every* ℓ *satisfying* $r_\ell > 0$*. Then there exist* $\mathcal{C}'(\mathbf{f})$*, a constant which depends only on* $\mathbf{f}$*, and* $\delta > 0$ *such that*

$$\sum_{\mathbf{x} \in [0,X]^n} \mathbf{1}_{V_{\mathbf{f},\mathbf{0}}(\mathbb{C})}(\mathbf{x}) = \mathcal{C}'(\mathbf{f}) X^{n - \sum_{\ell=1}^{d} \ell r_\ell} + O(X^{n - \sum_{\ell=1}^{d} \ell r_\ell - \delta}).$$

Note Theorem 1.3 recovers the main result of [1] when $r_1 = \cdots = r_{d-1} = 0$. The constant $\mathcal{C}'(\mathbf{f})$ is slightly different from $\mathcal{C}(\mathbf{f})$, and it is also positive assuming $\mathbf{f}$ satisfies suitable local conditions. (These are slightly different from the local conditions described in the paragraph after Theorem 1.2.) As stated in [3], 'Birch's original result needed the forms all to have the same degree, and there is a significant technical problem in extending the method to the general case.' It is required in Theorem 1.1 that the polynomials all have the same degree. As in the case for integer points, there are significant challenges to be overcome in generalizing the result on prime solutions of polynomial equations of equal degree to handle arbitrary systems.

We follow the main strategy of Cook and Magyar [6] to achieve Theorem 1.2. Let us briefly explain the idea of their approach using a very simple case.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let $\mathbf{y} = (x_2, \ldots, x_m)$ and $\mathbf{z} = (x_{m+1}, \ldots, x_n)$ for some $2 < m < n$ so that $\mathbf{x} = (x_1, \mathbf{y}, \mathbf{z})$. Suppose the system in consideration consists of one degree d homogeneous polynomial of the following shape

$$F(\mathbf{x}) = x_1^d + \mathfrak{F}_1(\mathbf{y}) + \mathfrak{F}_2(\mathbf{z}),$$

where $\mathfrak{F}_1(\mathbf{y})$ and $\mathfrak{F}_2(\mathbf{z})$ have sufficiently large Birch rank or h-invariant (defined in Section 2). By the orthogonality relation, we have (using the notation (1.6) with F in place of $\mathbf{f}$)

$$\begin{aligned}\mathcal{M}_F(X) &= \int_0^1 \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x}) e(\alpha F(\mathbf{x}))\, d\alpha \\ &= \int_{\mathfrak{M}(C)} \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x}) e(\alpha F(\mathbf{x}))\, d\alpha \\ &\quad + \int_{\mathfrak{m}(C)} \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x}) e(\alpha F(\mathbf{x}))\, d\alpha, \end{aligned} \tag{1.7}$$

where $\mathfrak{M}(C)$ is the *major arcs* and $\mathfrak{m}(C)$ is the *minor arcs* (defined in Section 4). By a fairly standard approach, the following estimate for the integral over the major arcs is obtained

$$\int_{\mathfrak{M}(C)} \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x}) e(\alpha F(\mathbf{x}))\, d\alpha = \mathcal{C}(F) X^{n-d} + O\left(\frac{X^{n-d}}{(\log X)^c}\right)$$

for some $c > 0$. The challenge is over the minor arcs. By the Cauchy–Schwartz inequality and the orthogonality relation, we have

$$\begin{aligned}&\left| \int_{\mathfrak{m}(C)} \sum_{\mathbf{x}\in[0,X]^n} \Lambda(\mathbf{x}) e(\alpha F(\mathbf{x}))\, d\alpha \right| \\ &\quad \leqslant N_1^{1/2} N_2^{1/2} (\log X)^{n-1} \sup_{\alpha\in\mathfrak{m}(C)} \left| \sum_{x_1\in[0,X]} \Lambda(x_1) e(\alpha x_1^d) \right|,\end{aligned}$$

where

$$N_1 = \#\{(\mathbf{y}, \mathbf{y}') \in [0, X]^{2(m-1)} : \mathfrak{F}_1(\mathbf{y}) = \mathfrak{F}_1(\mathbf{y}')\}$$

and

$$N_2 = \#\{(\mathbf{z}, \mathbf{z}') \in [0, X]^{2(n-m)} : \mathfrak{F}_2(\mathbf{z}) = \mathfrak{F}_2(\mathbf{z}')\}.$$

From our assumptions on $\mathfrak{F}_1(\mathbf{y})$ and $\mathfrak{F}_2(\mathbf{z})$, we have

$$N_1 \ll X^{2(m-1)-d} \quad \text{and} \quad N_2 \ll X^{2(n-m)-d}.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Then a method known as Weyl differencing is used to obtain the following

$$\sup_{\alpha\in\mathfrak{m}(C)}\left|\sum_{x_1\in[0,X]}\Lambda(x_1)e(\alpha x_1^d)\right|\ll\frac{X}{(\log X)^{c'}}$$

for some $c'>n-1$. Therefore, we obtain

$$\left|\int_{\mathfrak{m}(C)}\sum_{\mathbf{x}\in[0,X]^n}\Lambda(\mathbf{x})e(\alpha F(\mathbf{x}))\,d\alpha\right|\ll\frac{X^{n-d}}{(\log X)^{c'-n+1}},$$

and the asymptotic formula for $\mathcal{M}_F(X)$ holds. In this case, there were no terms of $F(\mathbf{x})$ involving variables from both $\mathbf{y}$ and $\mathbf{z}$, which makes the analysis very simple. Roughly speaking, the idea of [6] is that when such mixed terms are present, by considering certain level sets where the portions of the polynomials with these mixed terms are constant, the approach described above can be used. Also these level sets have to be chosen in a way that it does not introduce any unmanageable error terms. In order to generalize this argument of [6], we take the following steps. First, by an inductive argument we generalize [6, Proposition 2], which decomposes systems of degree d forms by a partition of variables, to handle systems involving forms of different degrees (Proposition 3.1). Next, using Gröbner basis theory we reduce the polynomials, without changing the solution set, to have suitable properties for the mentioned Weyl differencing argument to work. These properties become useful in the major arcs analysis as well. We note this reduction was not necessary in [6]. Here we also show that the reduction works well with the h-invariant. In our minor arcs analysis, we decompose the system of forms by Proposition 3.1 once and then we apply it again to one of the resulting systems of forms; the second application was not necessary in [6]. In [6], there are only the degree d forms in consideration, and the forms arising from the regularization process (Proposition 2.7) all have degrees strictly less than d. On the other hand, in our case the system in consideration has forms of degrees between 1 and d, and the forms arising from the regularization process have degrees strictly less than d. Thus, the forms in consideration and the forms arising from the regularization process are not as cleanly separated in our case. We overcome this challenge by the repeated application of Proposition 3.1. Finally, we treat the major arcs analysis in a similar manner as in [6]; however, the analysis becomes more complicated due to the fact that the system involves polynomials of differing degrees.

The organization of the rest of the paper is as follows. In Section 2, we collect some definitions and results related to the regularization process, which is an important part of the method in [6] and also of this paper. In Section 3, we prove

Proposition 3.1 to decompose systems of forms. We prepare the initial setup for the minor arcs analysis in Section 4; this is where we reduce the system of polynomials as mentioned above. We then obtain the desired minor arcs estimate in Section 5. In Section 6, we collect technical results that are necessary in obtaining our major arcs estimates in Section 7. Finally, we state our conclusions and further remarks in Section 8. We also have Appendix A, where we provide proofs for the results presented in Section 6. The work here is based on [20], and we chose to present these technical details at the end for an easier read of the paper.

Throughout the paper we do not distinguish between the two terms 'homogeneous polynomial' and 'form', and we will be using these terms interchangeably. By 'rational form' we mean it is a form with coefficients in $\mathbb{Q}$. By an affine variety we mean an algebraic set which is not necessarily irreducible. We use $\ll$ and $\gg$ to denote Vinogradov's well-known notation. We also use the notation $e(x)$ to denote $e^{2\pi i x}$. For $\mathbf{x} = (x_1, \ldots, x_n)$, the notation

$$\sum_{\mathbf{x} \in [0,X]^n}$$

means we are summing over all $\mathbf{x} \in \mathbb{Z}^n$ with $0 \leqslant x_i \leqslant X$ $(1 \leqslant i \leqslant n)$. For $q \in \mathbb{N}$, we use the numbers from $\{0, 1, \ldots, q-1\}$ to represent the residue classes of $\mathbb{Z}/q\mathbb{Z}$. Finally, given $\mathbf{x} = (x_1, \ldots, x_n)$ by $|\mathbf{x}|$ we mean $|\mathbf{x}| = n$ in Sections 3 and 5, and $|\mathbf{x}| = \max_{1 \leqslant i \leqslant n} |x_i|$ from Section 6 onwards. There should be no ambiguity since we are defining these notations as they come up.

2. Regularization lemmas

In this section, we collect results from [6, 20] related to regular systems (see Definition 2.3) and the regularization process. Given a system of rational forms $\mathbf{F}$, via the regularization process we obtain another system of forms which has at most the expected number of integer points, its number of forms is 'small', and partitions the level sets of $\mathbf{F}$. This was an important component of the method in [6] used to split the exponential sum in a controlled manner during the minor arcs estimate.

Let $\ell > 1$. Given a form $G \in \mathbb{Q}[x_1, \ldots, x_n]$ of degree ℓ, we define the *h-invariant* (also referred to as the *rational Schmidt rank* in [6]), $h_\ell(G)$, to be the least positive integer h such that G can be written identically as

$$G = \widetilde{U}_1 \widetilde{V}_1 + \cdots + \widetilde{U}_h \widetilde{V}_h, \tag{2.1}$$

where $\widetilde{U}_i$ and $\widetilde{V}_i$ are rational forms of positive degree $(1 \leqslant i \leqslant h)$. Let $\mathbf{G} = (G_1, \ldots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$. We generalize the definition of the h-invariant, and define the h-invariant of $\mathbf{G}$ to be

$$h_\ell(\mathbf{G}) = \min_{\boldsymbol{\mu}\in\mathbb{Q}^{r'}\setminus\{\mathbf{0}\}} h_\ell(\mu_1 G_1 + \cdots + \mu_{r'} G_{r'}). \tag{2.2}$$

Let $\mathbf{g} = (g_1, \ldots, g_{r'})$ be a system of degree ℓ polynomials in $\mathbb{Q}[x_1, \ldots, x_n]$. Let G_r be the homogeneous degree ℓ portion of g_r $(1 \leqslant r \leqslant r')$. We define

$$h_\ell(\mathbf{g}) := h_\ell(\mathbf{G}). \tag{2.3}$$

The h-invariant satisfies the following property.

LEMMA 2.1 [22, Lemma 2.2]. *Let $\ell > 1$ and let $\mathbf{G} = (G_1, \ldots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$. Suppose $h_\ell(\mathbf{G}) > 1$. Then for any $1 \leqslant i \leqslant n$, we have*

$$h_\ell(\mathbf{G}) - 1 \leqslant h_\ell(\mathbf{G}|_{x_i=0}) \leqslant h_\ell(\mathbf{G}),$$

where $\mathbf{G}|_{x_i=0} = (G_1|_{x_i=0}, \ldots, G_{r'}|_{x_i=0})$.

We have the following relation between the h-invariant and the Birch rank by combining [20, Lemma 16.1, (10.3), (10.5), (17.1)].

LEMMA 2.2. *Let $\ell > 1$ and let $\mathbf{G} = (G_1, \ldots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$. We have*

$$h_\ell(\mathbf{G}) \geqslant 2^{1-\ell} \mathcal{B}_\ell(\mathbf{G}). \tag{2.4}$$

DEFINITION 2.3. Let $d > 1$. Let $\mathbf{u} = (\mathbf{u}_d, \ldots, \mathbf{u}_1)$ be a system of polynomials in $\mathbb{Q}[x_1, \ldots, x_n]$, where $\mathbf{u}_\ell = (u_{\ell,1}, \ldots, u_{\ell,r_\ell})$ is the subsystem of degree ℓ polynomials of $\mathbf{u}$ $(1 \leqslant \ell \leqslant d)$. Let $D_{\mathbf{u}} = \sum_{\ell=1}^d \ell r_\ell$ and $R_{\mathbf{u}} = \sum_{\ell=1}^d r_\ell$. We denote $V_{\mathbf{u},\mathbf{0}}(\mathbb{Z})$ to be the set of solutions in $\mathbb{Z}^n$ to the equations

$$u_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell). \tag{2.5}$$

Let us denote the equations (2.5) by $\mathbf{u}(\mathbf{x}) = \mathbf{0}$. We say the system $\mathbf{u}$ is *regular* if

$$|V_{\mathbf{u},\mathbf{0}}(\mathbb{Z}) \cap [-X, X]^n| \ll X^{n-D_{\mathbf{u}}}.$$

Similarly as above we also define $V_{\mathbf{u},\mathbf{0}}(\mathbb{R})$ to be the set of solutions in $\mathbb{R}^n$ of the equations $\mathbf{u}(\mathbf{x}) = \mathbf{0}$. For a system of polynomials $\mathbf{u}$ as given in Definition 2.3, we let $\mathbf{U} = (\mathbf{U}_d, \ldots, \mathbf{U}_1)$ be the system of forms such that for each $1 \leqslant \ell \leqslant d$, we have $\mathbf{U}_\ell = \{U_{\ell,1}, \ldots, U_{\ell,r_\ell}\}$ where $U_{\ell,r}$ is the homogeneous degree ℓ portion of $u_{\ell,r}$ $(1 \leqslant r \leqslant r_\ell)$. The following theorem is one of the main results of [20] due to Schmidt.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

THEOREM 2.4 [20, Theorem II]. *Let $d > 1$. Let $\mathbf{u} = (\mathbf{u}_d, \ldots, \mathbf{u}_2)$ be a system of rational polynomials with notations as in Definition 2.3, and also let $\mathbf{U}_\ell$ be the system of homogeneous degree ℓ portions of $\mathbf{u}_\ell$ $(2 \leqslant \ell \leqslant d)$. If we have*

$$h_\ell(\mathbf{U}_\ell) \geqslant d2^{4\ell}(\ell!)r_\ell R_{\mathbf{u}} \quad (2 \leqslant \ell \leqslant d),$$

then the system $\mathbf{u}$ is regular.

Even though the statement of [20, Theorem II] is regarding systems of forms, the above Theorem 2.4, which is the inhomogeneous polynomials version, also holds by the explanation given in [20, Section 9] and 'Remark on inhomogeneous polynomials' in [20, page 262].

Let us denote

$$\rho_{d,\ell}(t) = d2^{4\ell}(\ell!)t^2 \quad (2 \leqslant \ell \leqslant d). \tag{2.6}$$

Then for each $2 \leqslant \ell \leqslant d$, $\rho_{d,\ell}(t)$ is an increasing function, and

$$\rho_{d,\ell}(R_{\mathbf{u}}) \geqslant d2^{4\ell}(\ell!)r_\ell R_{\mathbf{u}}.$$

Note Theorem 2.4 is regarding systems of polynomials which do not contain any linear polynomials. The following Corollary 2.5 is for systems that contain linear forms as well. We refer the reader to [6, Corollary 3] or [22, Corollary 3.3] for its proof.

COROLLARY 2.5. *Let $d > 1$. Let $\mathbf{u} = (\mathbf{u}_d, \ldots, \mathbf{u}_1)$ be a system of rational polynomials with notations as in Definition 2.3, and also let $\mathbf{U}_\ell$ be the system of homogeneous degree ℓ portions of $\mathbf{u}_\ell$ $(1 \leqslant \ell \leqslant d)$. Suppose $\mathbf{u}_1$ only contains linear forms, in other words $\mathbf{u}_1 = \mathbf{U}_1$, and that they are linearly independent over $\mathbb{Q}$. For each $2 \leqslant \ell \leqslant d$, let $\rho_{d,\ell}(\cdot)$ be as in (2.6). If we have*

$$h(\mathbf{U}_\ell) \geqslant \rho_{d,\ell}(R_{\mathbf{u}} - r_1) + r_1 \quad (2 \leqslant \ell \leqslant d),$$

then the system $\mathbf{u}$ is regular.

For $\mathbf{x} = (x_1, \ldots, x_n)$, by a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$ we mean that the set of variables of $\mathbf{y}$ and $\mathbf{z}$ partition $x_1, \ldots, x_n$. Let $\ell > 1$. Given $\mathbf{G} = (G_1, \ldots, G_{r'})$, a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$, and a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, we denote $\overline{\mathbf{G}}$ to be the system obtained by removing from $\mathbf{G}$ all forms which depend only on the $\mathbf{z}$ variables. Clearly if we have the trivial partition $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, where $\mathbf{z} = \emptyset$, then $\overline{\mathbf{G}} = \mathbf{G}$. Given a degree ℓ form $G(\mathbf{x})$ in $\mathbb{Q}[x_1, \ldots, x_n]$, we define the *$h$-invariant with respect to $\mathbf{z}$*, $h_\ell(G; \mathbf{z})$, to be the smallest number h_0

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

such that $G(\mathbf{x})$ can be expressed as

$$G(\mathbf{x}) = G(\mathbf{y}, \mathbf{z}) = \sum_{j=1}^{h_0} \widetilde{U}_j(\mathbf{y}, \mathbf{z}) \widetilde{V}_j(\mathbf{y}, \mathbf{z}) + W_0(\mathbf{z}),$$

where $\widetilde{U}_j$ and $\widetilde{V}_j$ are rational forms of positive degree ($1 \leqslant j \leqslant h_0$), and $W_0(\mathbf{z})$ is a rational form only in the $\mathbf{z}$ variables. We also define $h_\ell(\mathbf{G}; \mathbf{z})$ to be

$$h_\ell(\mathbf{G}; \mathbf{z}) = \min_{\boldsymbol{\lambda} \in \mathbb{Q}^{r'} \setminus \{\mathbf{0}\}} h_\ell(\lambda_1 G_1 + \cdots + \lambda_{r'} G_{r'}; \mathbf{z}).$$

If we have the trivial partition, then clearly we have $h_\ell(\mathbf{G}; \emptyset) = h_\ell(\mathbf{G})$. From this definition the following lemma holds.

LEMMA 2.6 [6, Lemma 2]. *Let $\ell > 1$. Let $\mathbf{G} = (G_1, \dots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \dots, x_n]$, and suppose we have a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$. Let $\mathbf{y}'$ be a set of variables with the same number of variables as $\mathbf{y}$. Then we have*

$$h_\ell(\mathbf{G}(\mathbf{y}, \mathbf{z}), \mathbf{G}(\mathbf{y}', \mathbf{z}); \mathbf{z}) = h_\ell(\mathbf{G}; \mathbf{z}),$$

where the left hand side denotes the h-invariant with respect to $\mathbf{z}$ of the system

$$(G_1(\mathbf{y}, \mathbf{z}), \dots, G_{r'}(\mathbf{y}, \mathbf{z}), G_1(\mathbf{y}', \mathbf{z}), \dots, G_{r'}(\mathbf{y}', \mathbf{z})).$$

In [6], the process in the following proposition is referred to as the regularization of systems. We will be utilizing this proposition in Section 5 to obtain the minor arcs estimate. Let us remark that results of this type had been obtained before for polynomials over finite fields [11, 15].

PROPOSITION 2.7 [6, Propositions 1 and 1′]. *Let $d > 1$, and let $\mathcal{F}$ be any collection of nondecreasing functions $\mathcal{F}_i : \mathbb{Z}_{\geqslant 0} \to \mathbb{Z}_{\geqslant 0}$ ($2 \leqslant i \leqslant d$). For a collection of nonnegative integers $r_1, \dots, r_d$, there exist constants*

$$C_1(r_1, \dots, r_d, \mathcal{F}), \dots, C_d(r_1, \dots, r_d, \mathcal{F})$$

such that the following holds.

Given a system of forms $\mathbf{U} = (\mathbf{U}_d, \dots, \mathbf{U}_1)$ in $\mathbb{Z}[x_1, \dots, x_n]$, where $\mathbf{U}_\ell = (U_{\ell,1}, \dots, U_{\ell,r_\ell})$ is the subsystem of degree ℓ forms of $\mathbf{U}$ ($1 \leqslant \ell \leqslant d$), and a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, there exists a system of forms $\mathcal{R}(\mathbf{U}) = (\mathcal{R}^{(d)}(\mathbf{U}), \dots, \mathcal{R}^{(1)}(\mathbf{U}))$ in $\mathbb{Q}[x_1, \dots, x_n]$, where $\mathcal{R}^{(\ell)}(\mathbf{U})$ is the subsystem of degree ℓ forms of $\mathcal{R}(\mathbf{U})$, satisfying the following. For each $1 \leqslant \ell \leqslant d$, let r'_ℓ be the number of forms in $\mathcal{R}^{(\ell)}(\mathbf{U})$, and let $R' = r'_1 + \cdots + r'_d$.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

(1) *Each form of the system* $\mathbf{U}$ *can be written as a rational polynomial expression in the forms of the system* $\mathcal{R}(\mathbf{U})$*. In particular, the level sets of* $\mathcal{R}(\mathbf{U})$ *partition those of* $\mathbf{U}$*.*

(2) *For each* $1 \leqslant \ell \leqslant d$*,* r'_ℓ *is at most* $C_\ell(r_1, \ldots, r_d, \boldsymbol{\mathcal{F}})$*.*

(3) *For each* $2 \leqslant \ell \leqslant d$*, we have* $h_\ell(\mathcal{R}^{(\ell)}(\mathbf{U})) \geqslant \mathcal{F}_\ell(R')$*. Moreover, the linear forms of* $\mathcal{R}^{(1)}(\mathbf{U})$ *are linearly independent over* $\mathbb{Q}$*.*

(4) *Let* $\overline{\mathcal{R}}^{(\ell)}(\mathbf{U})$ *be the system obtained by removing from* $\mathcal{R}^{(\ell)}(\mathbf{U})$ *all forms which depend only on the* $\mathbf{z}$ *variables* $(1 \leqslant \ell \leqslant d)$*. Then for each* $2 \leqslant \ell \leqslant d$*, we have* $h_\ell(\overline{\mathcal{R}}^{(\ell)}(\mathbf{U}); \mathbf{z}) \geqslant \mathcal{F}_\ell(R')$*. Furthermore, we may assume that the linear forms of* $\overline{\mathcal{R}}^{(1)}(\mathbf{U})$ *depend only on the* $\mathbf{y}$ *variables, and that they are linearly independent over* $\mathbb{Q}$*.*

We note that the last assertion in (4) regarding the linear forms of $\overline{\mathcal{R}}^{(1)}(\mathbf{U})$ is not stated in [6, Proposition 1′]. However, it is easy to deduce that this is indeed the case from [6, Proposition 1′] at the expense of possibly slightly larger constants $C_i(r_1, \ldots, r_d, \boldsymbol{\mathcal{F}})$ $(1 \leqslant i \leqslant d)$ than in [6, Proposition 1′]. We also note that with this assertion, it follows that every linear form of $\mathcal{R}^{(1)}(\mathbf{U})$ is either only in the $\mathbf{y}$ variables, or only in the $\mathbf{z}$ variables.

3. Decomposition of forms

In this section, we decompose a system of forms into two parts in a way that both parts have large Birch rank. Let $d, n \in \mathbb{N}$, and let $\mathbf{F}$ be a system of forms in $\mathbb{Q}[x_1, \ldots, x_n]$ of degrees less than or equal to d. We use a slightly different notation in this section compared to the previous sections in order to make the argument as clear as possible. We denote $\mathbf{F} = (\mathbf{F}^{(d)}, \ldots, \mathbf{F}^{(1)})$, where $\mathbf{F}^{(\ell)}$ is the subsystem of degree ℓ forms of $\mathbf{F}$ $(1 \leqslant \ell \leqslant d)$. For each $1 \leqslant \ell \leqslant d$, we denote the elements of $\mathbf{F}^{(\ell)}$ by

$$\mathbf{F}^{(\ell)} = (F_1^{(\ell)}, \ldots, F_{r_\ell}^{(\ell)}),$$

where r_ℓ is the number of forms in $\mathbf{F}^{(\ell)}$. Suppose we have a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$. We define $\mathbf{F}_{\mathbf{y},\mathbf{z}}^{(\ell)}(\mathbf{y}, \mathbf{z})$ to be the following system of forms

$$\mathbf{F}_{\mathbf{y},\mathbf{z}}^{(\ell)}(\mathbf{y}, \mathbf{z}) = (F_1^{(\ell)}(\mathbf{y}, \mathbf{z}) - F_1^{(\ell)}(\mathbf{0}, \mathbf{z}), \ldots, F_{r_\ell}^{(\ell)}(\mathbf{y}, \mathbf{z}) - F_{r_\ell}^{(\ell)}(\mathbf{0}, \mathbf{z})). \quad (3.1)$$

Note for each $1 \leqslant r \leqslant r_\ell$, we have

$$F_r^{(\ell)}(\mathbf{y}, \mathbf{z}) - F_r^{(\ell)}(\mathbf{0}, \mathbf{z}) = F_r^{(\ell)}(\mathbf{y}, \mathbf{0}) + (F_r^{(\ell)}(\mathbf{y}, \mathbf{z}) - F_r^{(\ell)}(\mathbf{y}, \mathbf{0}) - F_r^{(\ell)}(\mathbf{0}, \mathbf{z})),$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

and every monomial with nonzero coefficient in $(F_r^{(\ell)}(\mathbf{y},\mathbf{z}) - F_r^{(\ell)}(\mathbf{y},\mathbf{0}) - F_r^{(\ell)}(\mathbf{0},\mathbf{z}))$ involves both the $\mathbf{y}$ variables and the $\mathbf{z}$ variables, in other words it cannot be in terms of only the $\mathbf{y}$ variables or only the $\mathbf{z}$ variables. For each $1 \leqslant \ell \leqslant d$, we also define

$$\mathbf{F}_{\mathbf{z}}^{(\ell)}(\mathbf{z}) = (F_1^{(\ell)}(\mathbf{0},\mathbf{z}), \ldots, F_{r_\ell}^{(\ell)}(\mathbf{0},\mathbf{z})). \tag{3.2}$$

It should be clear from the context which partition of variables is being used when the notations (3.1) and (3.2) come up in this section. We now give an example of how these notations may be used. Let us consider $\mathbf{F}^{(\ell)}$ with $\ell > 1$. Suppose we have partitions of variables $\mathbf{x} = (\mathbf{v},\mathbf{z})$ and $\mathbf{z} = (\mathbf{y},\mathbf{z}')$, and let us denote $\mathbf{x} = (\mathbf{v},(\mathbf{y},\mathbf{z}'))$. From the first partition of variables, we have $\mathbf{F}_{\mathbf{v},\mathbf{z}}^{(\ell)}(\mathbf{v},\mathbf{z})$ and $\mathbf{F}_{\mathbf{z}}^{(\ell)}(\mathbf{z})$ as above. Since $\mathbf{F}_{\mathbf{z}}^{(\ell)}(\mathbf{z})$ is in terms of the $\mathbf{z}$ variables, we can consider (3.1) and (3.2) of this system with respect to the partition $\mathbf{z} = (\mathbf{y},\mathbf{z}')$. We then have

$$\begin{aligned}(\mathbf{F}_{\mathbf{z}}^{(\ell)})_{\mathbf{y},\mathbf{z}'}(\mathbf{y},\mathbf{z}') \\ &= (F_1^{(\ell)}(\mathbf{0},(\mathbf{y},\mathbf{z}')) - F_1^{(\ell)}(\mathbf{0},(\mathbf{0},\mathbf{z}')), \ldots, F_{r_\ell}^{(\ell)}(\mathbf{0},(\mathbf{y},\mathbf{z}')) - F_{r_\ell}^{(\ell)}(\mathbf{0},(\mathbf{0},\mathbf{z}')))\end{aligned}$$

and

$$(\mathbf{F}_{\mathbf{z}}^{(\ell)})_{\mathbf{z}'}(\mathbf{z}') = (F_1^{(\ell)}(\mathbf{0},(\mathbf{0},\mathbf{z}')), \ldots, F_{r_\ell}^{(\ell)}(\mathbf{0},(\mathbf{0},\mathbf{z}'))).$$

Given a set of variables $\mathbf{y}$, we denote $|\mathbf{y}|$ to be the number of variables of $\mathbf{y}$. The following proposition for a system of forms of equal degree is proved in [6]. We will be using this proposition as the base case in induction to generalize it in Proposition 3.4.

PROPOSITION 3.1 [6, Proposition 2]. *Let C_1 and C_2 be some positive integers. Let $d \geqslant 1$ and $\mathbf{F} = \mathbf{F}^{(d)} = (F_1^{(d)}, \ldots, F_{r_d}^{(d)})$ be a system of degree d forms in $\mathbb{Q}[x_1, \ldots, x_n]$, where $\mathcal{B}_d(\mathbf{F})$ is sufficiently large with respect to C_1, C_2, r_d and d. Then there exists a partition of variables $\mathbf{x} = (\mathbf{y},\mathbf{z})$ such that*

$$|\mathbf{y}| \leqslant C_1 r_d,$$
$$\mathcal{B}_d(\mathbf{F}_{\mathbf{y},\mathbf{z}}(\mathbf{y},\mathbf{z})) \geqslant C_1, \quad \textit{and} \quad \mathcal{B}_d(\mathbf{F}_{\mathbf{z}}(\mathbf{z})) \geqslant C_2.$$

The following lemma and its corollary are also proved in [6].

LEMMA 3.2 [6, Lemma 3]. *Let $\ell \geqslant 1$ and let $\mathbf{G} = (G_1, \ldots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$. Given any $1 \leqslant j \leqslant n$, we have*

$$\mathcal{B}_\ell(\mathbf{G}) \geqslant \mathcal{B}_\ell(\mathbf{G}|_{x_j=0}) \geqslant \mathcal{B}_\ell(\mathbf{G}) - r' - 1, \tag{3.3}$$

where $\mathbf{G}|_{x_j=0} = (G_1|_{x_j=0}, \ldots, G_{r'}|_{x_j=0})$. When $\ell = 1$, we in fact have

$$\mathcal{B}_1(\mathbf{G}) \geqslant \mathcal{B}_1(\mathbf{G}|_{x_j=0}) \geqslant \mathcal{B}_1(\mathbf{G}) - 1.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Proof. The lower bounds are provided in [6, Lemma 3], so we only provide the argument for the upper bounds here. We remark that due to a minor oversight the lower bound is stated to be $\mathcal{B}_\ell(\mathbf{G}) - r'$ in [6, Lemma 3] instead of the lower bound given in (3.3). However, by following through their argument it can be seen that in fact (3.3) is the correct lower bound. We begin by considering the case $\ell > 1$. Let us denote

$$\mathbf{G}' = (G'_1, \ldots, G'_{r'}),$$

where $G'_r = G_r|_{x_j=0}$ for each $1 \leqslant r \leqslant r'$. It follows from the definition of the Birch singular locus given in (1.3) that

$$V^*_{\mathbf{G}} \cap \{\mathbf{x} \in \mathbb{C}^n : x_j = 0\} \subseteq V^*_{\mathbf{G}'}.$$

Since the dimension of $V^*_{\mathbf{G}} \cap \{\mathbf{x} \in \mathbb{C}^n : x_j = 0\}$ is either $\dim(V^*_{\mathbf{G}}) - 1$ or $\dim(V^*_{\mathbf{G}})$, we have

$$\dim(V^*_{\mathbf{G}}) - 1 \leqslant \dim(V^*_{\mathbf{G}'}),$$

and equivalently,

$$\mathcal{B}_\ell(\mathbf{G}) = n - \dim(V^*_{\mathbf{G}}) \geqslant n - 1 - \dim(V^*_{\mathbf{G}'}) = \mathcal{B}_\ell(\mathbf{G}|_{x_j=0}).$$

For the case $\ell = 1$, it follows immediately from the definition. □

COROLLARY 3.3 [6, Corollary 4]. *Let $\ell \geqslant 1$ and let $\mathbf{G} = (G_1, \ldots, G_{r'})$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \ldots, x_n]$. If $\mathcal{H}$ is an affine linear space of codimension m, then the restriction of $\mathbf{G}$ to $\mathcal{H}$ has Birch rank at least $(\mathcal{B}_\ell(\mathbf{G}) - m(r' + 1))$. When $\ell = 1$, we in fact have that it is at least $(\mathcal{B}_1(\mathbf{G}) - m)$.*

We obtain the following technical result for a system of forms that is more general than in Proposition 3.1.

PROPOSITION 3.4. *Let $d, n \in \mathbb{N}$. Let $\boldsymbol{F} = (\boldsymbol{F}^{(d)}, \ldots, \boldsymbol{F}^{(1)})$ be a system of forms in $\mathbb{Q}[x_1, \ldots, x_n]$, where $\boldsymbol{F}^{(i)} = (F^{(i)}_1, \ldots, F^{(i)}_{r_i})$ is the subsystem of degree i forms of $\boldsymbol{F}$ $(1 \leqslant i \leqslant d)$. Let $C_{i,1}, C_{i,2}$ $(1 \leqslant i \leqslant d)$ be positive integers. For each $1 \leqslant i \leqslant d$, suppose $\mathcal{B}_i(\mathbf{F}^{(i)})$ is sufficiently large with respect to $C_{1,1}, \ldots, C_{d,1}, C_{1,2}, \ldots, C_{d,2}$, $r_d, \ldots, r_1$, and d. Then there exists a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$ such that*

$$|\mathbf{y}| \leqslant \sum_{i=1}^{d} C_{i,1} r_i,$$

and for each $1 \leqslant i \leqslant d$, we have

$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{y},\mathbf{z}}(\mathbf{y}, \mathbf{z})) \geqslant C_{i,1} - (r_i + 1) \sum_{\ell=1}^{i-1} C_{\ell,1} r_\ell,$$

and

$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{z}}(\mathbf{z})) \geqslant C_{i,2} - (r_i + 1)\sum_{\ell=1}^{i-1} C_{\ell,1} r_\ell.$$

Proof. We prove by induction the following statement: Given $2 \leqslant j \leqslant d$, there exists a partition of variables $\mathbf{x} = (\mathbf{v}_j, \mathbf{z}_j)$, where $\mathbf{v}_j = (\mathbf{y}_d, \ldots, \mathbf{y}_j)$, such that for each $j \leqslant i \leqslant d$ we have

$$|\mathbf{y}_i| \leqslant C_{i,1} r_i,$$
$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)) \geqslant C_{i,1} - (r_i + 1)\sum_{\ell=j}^{i-1} C_{\ell,1} r_\ell,$$

and

$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{z}_j}(\mathbf{z}_j)) \geqslant C_{i,2} - (r_i + 1)\sum_{\ell=j}^{i-1} C_{\ell,1} r_\ell.$$

We begin with the base case $j = d$. We know from Proposition 3.1 that there exists a partition of variables $\mathbf{x} = (\mathbf{y}_d, \mathbf{z}_d)$ such that

$$|\mathbf{y}_d| \leqslant C_{d,1} r_d,$$
$$\mathcal{B}_d(\mathbf{F}^{(d)}_{\mathbf{y}_d,\mathbf{z}_d}(\mathbf{y}_d, \mathbf{z}_d)) \geqslant C_{d,1}, \quad \text{and} \quad \mathcal{B}_d(\mathbf{F}^{(d)}_{\mathbf{z}_d}(\mathbf{z}_d)) \geqslant C_{d,2}.$$

This concludes our base case.

Suppose the statement holds for $j + 1$, in other words there exists a partition of variables $\mathbf{x} = (\mathbf{v}_{j+1}, \mathbf{z}_{j+1})$, where $\mathbf{v}_{j+1} = (\mathbf{y}_d, \ldots, \mathbf{y}_{j+1})$, such that for each $j + 1 \leqslant i \leqslant d$ we have

$$|\mathbf{y}_i| \leqslant C_{i,1} r_i,$$
$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1})) \geqslant C_{i,1} - (r_i + 1)\sum_{\ell=j+1}^{i-1} C_{\ell,1} r_\ell, \tag{3.4}$$

and

$$\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}}(\mathbf{z}_{j+1})) \geqslant C_{i,2} - (r_i + 1)\sum_{\ell=j+1}^{i-1} C_{\ell,1} r_\ell. \tag{3.5}$$

First we observe that by Lemma 3.2 the following holds

$$\begin{aligned}\mathcal{B}_j(\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}}(\mathbf{z}_{j+1})) = \mathcal{B}_j(\mathbf{F}^{(j)}(\mathbf{x})|_{\mathbf{v_{j+1}=0}}) &\geqslant \mathcal{B}_j(\mathbf{F}^{(j)}(\mathbf{x})) - (r_j + 1)|\mathbf{v}_{j+1}| \\ &\geqslant \mathcal{B}_j(\mathbf{F}^{(j)}(\mathbf{x})) - (r_j + 1)\sum_{i=j+1}^{d} C_{i,1} r_i.\end{aligned}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Since $\mathcal{B}_j(\mathbf{F}^{(j)}(\mathbf{x}))$ is sufficiently large with respect to $d, r_j, \ldots, r_d, C_{j,1}, \ldots, C_{d,1}$, and $C_{j,2}$, we obtain from Proposition 3.1 a partition of variables $\mathbf{z}_{j+1} = (\mathbf{y}_j, \mathbf{z}_j)$ such that

$$|\mathbf{y}_j| \leqslant C_{j,1} r_j, \tag{3.6}$$

$$\mathcal{B}_j((\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j)) \geqslant C_{j,1}, \tag{3.7}$$

and

$$\mathcal{B}_j((\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{z}_j}(\mathbf{z}_j)) \geqslant C_{j,2}. \tag{3.8}$$

We denote $\mathbf{v}_j = (\mathbf{v}_{j+1}, \mathbf{y}_j) = (\mathbf{y}_d, \ldots, \mathbf{y}_j)$, and consider the partition of variables $\mathbf{x} = (\mathbf{v}_j, \mathbf{z}_j)$.

Since $\mathbf{v}_{j+1} \subseteq \mathbf{v}_j$, we have

$$F_r{}^{(j)}(\mathbf{x})|_{\mathbf{v}_j=\mathbf{0}} = (F_r{}^{(j)}(\mathbf{x})|_{\mathbf{v}_{j+1}=\mathbf{0}})|_{\mathbf{v}_j=\mathbf{0}} \quad (1 \leqslant r \leqslant r_j)$$

and consequently,

$$\mathbf{F}^{(j)}_{\mathbf{z}_j}(\mathbf{z}_j) = (\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{z}_j}(\mathbf{z}_j). \tag{3.9}$$

Therefore, we obtain by (3.8) that

$$\mathcal{B}_j(\mathbf{F}^{(j)}_{\mathbf{z}_j}(\mathbf{z}_j)) \geqslant C_{j,2}. \tag{3.10}$$

We have the following two decompositions for $\mathbf{F}^{(j)}(\mathbf{x})$,

$$\begin{aligned} &\mathbf{F}^{(j)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j) + \mathbf{F}^{(j)}_{\mathbf{z}_j}(\mathbf{z}_j) \\ &\quad = \mathbf{F}^{(j)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1}) + (\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j) + (\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{z}_j}(\mathbf{z}_j), \end{aligned}$$

where the first decomposition is via the partition $\mathbf{x} = (\mathbf{v}_j, \mathbf{z}_j)$, and the second via the partitions $\mathbf{x} = (\mathbf{v}_{j+1}, \mathbf{z}_{j+1})$ and $\mathbf{z}_{j+1} = (\mathbf{y}_j, \mathbf{z}_j)$. It follows from (3.9) that

$$\mathbf{F}^{(j)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j) = \mathbf{F}^{(j)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1}) + (\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j). \tag{3.11}$$

Since $\mathbf{v}_{j+1} \cap (\mathbf{y}_j, \mathbf{z}_j) = \emptyset$, we obtain from (3.11)

$$\begin{aligned} (\mathbf{F}^{(j)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j))|_{\mathbf{v}_{j+1}=\mathbf{0}} &= ((\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j))|_{\mathbf{v}_{j+1}=\mathbf{0}} \\ &= (\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j). \end{aligned} \tag{3.12}$$

Also see the sentence after (3.1) for the explanation of $\mathbf{F}^{(j)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1})|_{\mathbf{v}_{j+1}=\mathbf{0}} = \mathbf{0}$. Consequently, we obtain from Lemma 3.2, (3.12), and (3.7),

$$\begin{aligned} \mathcal{B}_j(\mathbf{F}^{(j)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)) &\geqslant \mathcal{B}_j(\mathbf{F}^{(j)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)|_{\mathbf{v}_{j+1}=\mathbf{0}}) \\ &= \mathcal{B}_j((\mathbf{F}^{(j)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j)) \\ &\geqslant C_{j,1}. \end{aligned} \tag{3.13}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let $j+1 \leqslant i \leqslant d$. Recall we have partitions $\mathbf{x} = (\mathbf{v}_{j+1}, \mathbf{z}_{j+1})$, $\mathbf{x} = (\mathbf{v}_j, \mathbf{z}_j)$, and $\mathbf{z}_{j+1} = (\mathbf{y}_j, \mathbf{z}_j)$. Since $\mathbf{v}_j = (\mathbf{v}_{j+1}, \mathbf{y}_j)$, it follows that

$$(F_r^{(i)}(\mathbf{x})|_{\mathbf{v}_{j+1}=\mathbf{0}})|_{\mathbf{y}_j=\mathbf{0}} = F_r^{(i)}(\mathbf{x})|_{\mathbf{v}_j=\mathbf{0}} \quad (1 \leqslant r \leqslant r_i)$$

and consequently,

$$(\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}})_{\mathbf{z}_j}(\mathbf{z}_j) = \mathbf{F}^{(i)}_{\mathbf{z}_{j+1}}(\mathbf{z}_{j+1})|_{\mathbf{y}_j=\mathbf{0}} = \mathbf{F}^{(i)}_{\mathbf{z}_j}(\mathbf{z}_j). \tag{3.14}$$

Therefore, we obtain from (3.14), Lemma 3.2, (3.5), and (3.6),

$$\begin{aligned}\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{z}_j}(\mathbf{z}_j)) &\geqslant \mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}}(\mathbf{z}_{j+1})) - (r_i+1)|\mathbf{y}_j| \\ &\geqslant \left(C_{i,2} - (r_i+1)\sum_{\ell=j+1}^{i-1} C_{\ell,1} r_\ell\right) - (r_i+1)|\mathbf{y}_j| \\ &\geqslant C_{i,2} - (r_i+1)\sum_{\ell=j}^{i-1} C_{\ell,1} r_\ell.\end{aligned}$$

Also we have the following two decompositions for $\mathbf{F}^{(i)}(\mathbf{x})$,

$$\begin{aligned}&\mathbf{F}^{(i)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j) + \mathbf{F}^{(i)}_{\mathbf{z}_j}(\mathbf{z}_j) \\ &\quad = \mathbf{F}^{(i)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1}) + (\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j) + (\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}})_{\mathbf{z}_j}(\mathbf{z}_j),\end{aligned} \tag{3.15}$$

where the first decomposition is via the partition $\mathbf{x} = (\mathbf{v}_j, \mathbf{z}_j)$, and the second via the partitions $\mathbf{x} = (\mathbf{v}_{j+1}, \mathbf{z}_{j+1})$ and $\mathbf{z}_{j+1} = (\mathbf{y}_j, \mathbf{z}_j)$. Therefore, it follows from (3.14) and (3.15) that

$$\begin{aligned}&\mathbf{F}^{(i)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)|_{\mathbf{y}_j=\mathbf{0}} \\ &\quad = (\mathbf{F}^{(i)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1}) + (\mathbf{F}^{(i)}_{\mathbf{z}_{j+1}})_{\mathbf{y}_j,\mathbf{z}_j}(\mathbf{y}_j, \mathbf{z}_j))|_{\mathbf{y}_j=\mathbf{0}} \\ &\quad = \mathbf{F}^{(i)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1})|_{\mathbf{y}_j=\mathbf{0}}.\end{aligned} \tag{3.16}$$

Consequently, we have by Lemma 3.2, (3.16), (3.4), and (3.6),

$$\begin{aligned}\mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)) &\geqslant \mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{v}_j,\mathbf{z}_j}(\mathbf{v}_j, \mathbf{z}_j)|_{\mathbf{y}_j=\mathbf{0}}) \\ &\geqslant \mathcal{B}_i(\mathbf{F}^{(i)}_{\mathbf{v}_{j+1},\mathbf{z}_{j+1}}(\mathbf{v}_{j+1}, \mathbf{z}_{j+1})) - (r_i+1)|\mathbf{y}_j| \\ &\geqslant \left(C_{i,1} - (r_i+1)\sum_{\ell=j+1}^{i-1} C_{\ell,1} r_\ell\right) - (r_i+1)C_{j,1} r_j \\ &\geqslant C_{i,1} - (r_i+1)\sum_{\ell=j}^{i-1} C_{\ell,1} r_\ell.\end{aligned} \tag{3.17}$$

Hence, from (3.6), (3.10), (3.13), (3.15), and (3.17), we see that we have completed induction. From the $j = 2$ case with $\mathbf{v}_2 = (\mathbf{y}_d, \mathbf{y}_{d-1}, \ldots, \mathbf{y}_2)$ and $\mathbf{z}_2$, we can proceed in the exact same manner as above to deal with the linear forms even though the definition of the Birch rank, $\mathcal{B}_1$, is slightly different than that for the higher degrees. We can do so because Proposition 3.1 is still applicable with $\mathcal{B}_1$ for systems of linear forms. By letting the resulting variables $\mathbf{v}_1 = (\mathbf{y}_d, \mathbf{y}_{d-1}, \ldots, \mathbf{y}_2, \mathbf{y}_1)$ and $\mathbf{z}_1$ be $\mathbf{y}$ and $\mathbf{z}$, respectively, we complete the proof of Proposition 3.4. □

4. Initial setup to prove Theorem 1.2

Let $\mathbf{f} = (\mathbf{f}_d, \ldots, \mathbf{f}_1)$ be a system of polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$, where $\mathbf{f}_\ell = (f_{\ell,1}, \ldots, f_{\ell,r_\ell})$ is the subsystem of degree ℓ polynomials of $\mathbf{f}$ $(1 \leqslant \ell \leqslant d)$. We let $\mathbf{F} = (\mathbf{F}_d, \ldots, \mathbf{F}_1)$ be the system of forms such that for each $1 \leqslant \ell \leqslant d$, we have $\mathbf{F}_\ell = (F_{\ell,1}, \ldots, F_{\ell,r_\ell})$ where $F_{\ell,r}$ is the homogeneous degree ℓ portion of $f_{\ell,r}$ $(1 \leqslant r \leqslant r_\ell)$. Recall in Theorem 1.2 we consider the system of equations

$$f_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \tag{4.1}$$

where for each $1 \leqslant \ell \leqslant d$, $\mathcal{B}_\ell(\mathbf{f}_\ell)$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. Also recall we denote the integer solutions of these equations by $V_{\mathbf{f},\mathbf{0}}(\mathbb{Z})$. In order to prove Theorem 1.2, we begin by simplifying the polynomials in (4.1) to satisfy more properties suitable for our purposes without changing the solution set $V_{\mathbf{f},\mathbf{0}}(\mathbb{Z})$.

By reducing the polynomials in (4.1) without changing the solution set, we transform system (4.1) into the following system:

$$f_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \tag{4.2}$$

where for $2 \leqslant \ell \leqslant d$,

$$f_{\ell,r}(\mathbf{x}) = c_{\ell,r}\mathbf{w}^{\mathbf{j}_{\ell,r}} + \chi_{\ell,r}(\mathbf{x}) + \widetilde{f}_{\ell,r}(\mathbf{x}) \quad (1 \leqslant r \leqslant r_\ell)$$

and

$$f_{1,r}(\mathbf{x}) = c_{1,r}\mathbf{w}^{\mathbf{j}_{1,r}} + \widetilde{f}_{1,r}(\mathbf{x}) \quad (1 \leqslant r \leqslant r_1)$$

with the following properties. Here $\mathbf{w}$ is a subset of the variables $\mathbf{x} = (x_1, \ldots, x_n)$, and for each $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ is a degree ℓ monomial in $\mathbf{w}$.

(1) For each $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, we have $c_{\ell,r} \in \mathbb{Z}\backslash\{0\}$, and $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ is the leading monomial of $f_{\ell,r}(\mathbf{x})$ with respect to a graded lexicographic ordering.

(2) The monomials $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ are distinct, and given $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ is not divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell',r'}}$ $(1 \leqslant \ell' < \ell, 1 \leqslant r' \leqslant r_{\ell'})$.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

(3) For each $2 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, the polynomial $\chi_{\ell,r}(\mathbf{x})$ has degree less than or equal to ℓ with coefficients in $\mathbb{Z}$. Also $\chi_{\ell,r}(\mathbf{x})$ does not contain any monomial divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell',r'}}$ $(1 \leqslant \ell' \leqslant \ell, 1 \leqslant r' \leqslant r_{\ell'})$.

(4) For each $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, the polynomial $\widetilde{f}_{\ell,r}(\mathbf{x})$ has degree ℓ with coefficients in $\mathbb{Z}$. Also $\widetilde{f}_{\ell,r}(\mathbf{x})$ does not contain any monomial divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell',r'}}$ $(1 \leqslant \ell' \leqslant \ell, 1 \leqslant r' \leqslant r_{\ell'})$.

(5) For each $2 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$, we have

$$h_\ell(\chi_{\ell,r}) \leqslant C_0'',$$

where C_0'' is a constant depending only on d and $r_d, \dots, r_1$.

(6) For each $1 \leqslant \ell \leqslant d$, $\mathcal{B}_\ell(\{\widetilde{f}_{\ell,r} : 1 \leqslant r \leqslant r_\ell\})$ is sufficiently large with respect to d and $r_d, \dots, r_1$.

(7) For each $2 \leqslant \ell \leqslant d$, $h_\ell(\mathbf{f}_\ell)$ is sufficiently large with respect to d and $r_d, \dots, r_1$, and $\mathcal{B}_1(\mathbf{f}_1)$ is sufficiently large with respect to d and $r_d, \dots, r_1$.

These conditions system (4.2) satisfies become crucial during our minor arcs estimate. Before we describe this reduction process, first we note basic properties of the Birch rank which will be utilized in this section. Let $\ell \geqslant 1$ and $\mathbf{G} = \{G_1, \dots, G_{r''}\}$ be a system of degree ℓ forms in $\mathbb{Q}[x_1, \dots, x_n]$. Let $\kappa_1, \dots, \kappa_{r''} \in \mathbb{Q}\backslash\{0\}$. Then it follows from the definition of the Birch rank that

$$\mathcal{B}_\ell(\{\kappa_r G_r : 1 \leqslant r \leqslant r''\}) = \mathcal{B}_\ell(\mathbf{G}).$$

Let $\kappa \in \mathbb{Q}$ and $1 \leqslant i, j \leqslant r''$. Let $G_r' = G_r$ if $r \neq i$, and $G_i' = G_i + \kappa G_j$. It also follows from the definition of the Birch rank that

$$\mathcal{B}_\ell(\{G_r' : 1 \leqslant r \leqslant r''\}) = \mathcal{B}_\ell(\mathbf{G}).$$

We now transform system (4.1) into system (4.2) beginning with $\ell = 1$. By considering the reduced row echelon form of the matrix formed by the coefficients of $F_{1,1}, \dots, F_{1,r_1}$, and relabeling the variables if necessary, we reduce the linear polynomials in (4.1) without changing the solution set to be of the shape

$$f_{1,r}(\mathbf{x}) = x_{n-r+1} + \widetilde{f}_{1,r}(x_1, \dots, x_{n-r_1}) \quad (1 \leqslant r \leqslant r_1),$$

where $\widetilde{f}_{1,r}(x_1, \dots, x_{n-r_1})$ is a linear polynomial in variables $x_1, \dots, x_{n-r_1}$ with rational coefficients. Then by substituting $x_{n-r+1} = -\widetilde{f}_{1,r}(x_1, \dots, x_{n-r_1})$ into each equation in (4.1) with $\ell > 1$, we may further reduce without changing the solution set such that for $\ell > 1$ the polynomials $f_{\ell,r}$ do not involve any of the variables

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$x_{n-r_1+1}, \ldots, x_n$. Let us label $w_r = x_{n-r+1}$ $(1 \leqslant r \leqslant r_1)$. By multiplying each of the resulting equation by an integer constant if necessary, we replace system (4.1) with the following system of equations

$$f_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \tag{4.3}$$

where

$$f_{\ell,r}(\mathbf{x}) \in \mathbb{Z}[x_1, \ldots, x_{n-r_1}] \quad (1 < \ell \leqslant d, 1 \leqslant r \leqslant r_\ell),$$

and for each $1 \leqslant r \leqslant r_1$, we have

$$f_{1,r}(\mathbf{x}) = c_{1,r} w_r + \widetilde{f}_{1,r}(x_1, \ldots, x_{n-r_1})$$

with $c_{1,r} \in \mathbb{Z}\backslash\{0\}$ and $\widetilde{f}_{1,r}(x_1, \ldots, x_{n-r_1}) \in \mathbb{Z}[x_1, \ldots, x_{n-r_1}]$. From the definition of the Birch rank, we have that $\mathcal{B}_1(\mathbf{f}_1)$ remains the same under these changes. Therefore, it follows by Lemma 3.2 that $\mathcal{B}_1(\{\widetilde{f}_{1,r} : 1 \leqslant r \leqslant r_1\})$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. For $1 < \ell \leqslant d$, we can deduce from Corollary 3.3 that we still have $\mathcal{B}_\ell(\mathbf{f}_\ell)$ sufficiently large with respect to d and $r_d, \ldots, r_1$. Let us put a graded lexicographic ordering on the monomials formed by $x_1, \ldots, x_{n-r_1}$, $w_1, \ldots, w_{r_1}$ such that w_r is the leading coefficient of $f_{1,r}$. By denoting $\mathbf{w}^{\mathbf{j}_{1,r}} = w_r$ $(1 \leqslant r \leqslant r_1)$, we see that the linear polynomials in (4.3) satisfy the conditions of system (4.2). We note that for us this graded lexicographic ordering is essentially on the monomials formed by $x_1, \ldots, x_{n-r_1}$ as $w_1, \ldots, w_r$ do not appear in $f_{\ell,r}$ with $\ell > 1$.

Let us denote $B_2 := \mathcal{B}_2(\mathbf{f}_2)$ for the Birch rank of $\mathbf{f}_2$ in (4.3). We consider $\mathbf{F}_2 = (F_{2,1}, \ldots, F_{2,r_2})$, the system of homogeneous degree 2 portions of $\mathbf{f}_2$. From each $F_{2,1}, \ldots, F_{2,r_2}$, we collect the coefficient of the monomial $x_{i_1}x_{i_2}$ and turn it into a vector in $\mathbb{Q}^{r_2}$. We do this for every degree 2 monomial. We then form a matrix by putting these vectors in columns from left to right in the decreasing order of the degree 2 monomials. Since $B_2 = \mathcal{B}_2(\mathbf{F}_2) > 0$, this matrix has full rank. We row reduce this matrix, and we denote the r_2 monomials where the leading 1's occur to be $\mathbf{w}^{\mathbf{j}_{2,r}}$ $(1 \leqslant r \leqslant r_2)$, and label the distinct variables involved in these r_2 monomials to be $\mathbf{w}_2 = (w_{r_1+1}, \ldots, w_{r_1+K_2})$. Clearly we have $K_2 \leqslant 2r_2$.

From the row reduction operations done on the coefficient matrix of $\mathbf{F}_2$, without changing the solution set we can reduce $\mathbf{f}_2$ to

$$f_{2,r}(\mathbf{x}) = c_{2,r}\mathbf{w}^{\mathbf{j}_{2,r}} + \widetilde{f}_{2,r}(\mathbf{x}) \quad (1 \leqslant r \leqslant r_2), \tag{4.4}$$

where $\mathbf{w}^{\mathbf{j}_{2,r}}$ is the leading monomial of $f_{2,r}(\mathbf{x})$, with respect to the graded lexicographic ordering, and none of the monomials of $\widetilde{f}_{2,r}(\mathbf{x})$ is divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell',r'}}$ $(1 \leqslant \ell' \leqslant 2, 1 \leqslant r \leqslant r_{\ell'})$. We have that $c_{2,r}\mathbf{w}^{\mathbf{j}_{2,r}} + \widetilde{f}_{2,r}(\mathbf{x})$ is a $\mathbb{Q}$-linear combination of $f_{2,1}, \ldots, f_{2,r_2}$ in (4.3), where the $\mathbb{Q}$-linear combination comes from the row reduction operations applied on the coefficient matrix described above. Thus by the basic properties of the Birch rank, it follows that

$$\mathcal{B}_2(\{c_{2,r}\mathbf{w}^{\mathbf{j}_{2,r}} + \widetilde{f}_{2,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_2\}) = B_2.$$

It then follows from (2.4) that the h-invariant of $\mathbf{f}_2$ in (4.4) satisfies

$$h_2(\mathbf{f}_2) \geqslant 2^{1-2}B_2,$$

and hence $h_2(\mathbf{f}_2)$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. We also have by Lemma 3.2,

$$\begin{aligned}\mathcal{B}_2(\{\widetilde{f}_{2,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_2\}) &\geqslant \mathcal{B}_2(\{\widetilde{f}_{2,r}(\mathbf{x})|_{\mathbf{w}_2=\mathbf{0}} : 1 \leqslant r \leqslant r_2\}) \\ &= \mathcal{B}_2(\{(c_{2,r}\mathbf{w}^{\mathbf{j}_{2,r}} + \widetilde{f}_{2,r}(\mathbf{x}))|_{\mathbf{w}_2=\mathbf{0}} : 1 \leqslant r \leqslant r_2\}) \\ &\geqslant \mathcal{B}_2(\{c_{2,r}\mathbf{w}^{\mathbf{j}_{2,r}} + \widetilde{f}_{2,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_2\}) - (r_2+1)K_2 \\ &= B_2 - (r_2+1)K_2.\end{aligned}$$

Thus $\mathcal{B}_2(\{\widetilde{f}_{2,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_2\})$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. It is also clear that $\mathbf{w}^{\mathbf{j}_{2,r}}$ is not divisible by any one of $\mathbf{w}^{\mathbf{j}_{1,r'}} = w_{r'}$ $(1 \leqslant r' \leqslant r_1)$. Therefore, we have obtained that we can reduce the degree 2 polynomials of system (4.3) without changing the solution set to satisfy the conditions of system (4.2) with $\chi_{2,r}(\mathbf{x})$ being the zero polynomial $(1 \leqslant r \leqslant r_2)$.

Using the $\ell = 2$ case as the base case, we prove our statement by induction. Let $\ell_0 \geqslant 3$. Suppose we have reduced the polynomials $\mathbf{f}_\ell$ in (4.3) for each $2 \leqslant \ell \leqslant \ell_0 - 1$, without changing the solution set, to satisfy the conditions of (4.2). First we take the distinct variables involved in the monomials $\mathbf{w}^{\mathbf{j}_{3,r}}$ $(1 \leqslant r \leqslant r_3)$ that have not yet appeared in $\mathbf{w}_2$, and label them as $w_{r_1+K_2+1}, \ldots, w_{r_1+K_2+K_3}$. Clearly we have $K_3 \leqslant 3r_3$. We adjoin these variables to $\mathbf{w}_2$, and let $\mathbf{w}_3 = (w_{r_1+1}, \ldots, w_{r_1+K_2+K_3})$. Then we take the distinct variables involved in the monomials $\mathbf{w}^{\mathbf{j}_{4,r}}$ $(1 \leqslant r \leqslant r_4)$ that have not yet appeared in $\mathbf{w}_3$, and label them as $w_{r_1+K_2+K_3+1}, \ldots, w_{r_1+K_2+K_3+K_4}$. Clearly we have $K_4 \leqslant 4r_4$. We adjoin these variables to $\mathbf{w}_3$, and let $\mathbf{w}_4 = (w_{r_1+1}, \ldots, w_{r_1+K_2+K_3+K_4})$. We continue in this manner until we obtain

$$\mathbf{w}_{\ell_0-1} = (w_{r_1+1}, \ldots, w_{r_1+K_2+\cdots+K_{\ell_0-1}}),$$

where $K_j \leqslant jr_j$ $(2 \leqslant j \leqslant \ell_0 - 1)$.

For each $1 \leqslant r \leqslant r_{\ell_0}$, we let

$$f_{\ell_0,r}(\mathbf{x}) = \chi''_{\ell_0,r}(\mathbf{x}) + f''_{\ell_0,r}(\mathbf{x}), \tag{4.5}$$

where every monomial of $\chi''_{\ell_0,r}(\mathbf{x})$ is divisible by one of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(2 \leqslant \ell < \ell_0, 1 \leqslant r \leqslant r_\ell)$, and none of the monomials of $f''_{\ell_0,r}(\mathbf{x})$ is divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(2 \leqslant \ell < \ell_0, 1 \leqslant r \leqslant r_\ell)$. Since

$$f_{\ell_0,r}(\mathbf{x})|_{\mathbf{w}_{\ell_0-1}=\mathbf{0}} = f''_{\ell_0,r}(\mathbf{x})|_{\mathbf{w}_{\ell_0-1}=\mathbf{0}} \quad (1 \leqslant r \leqslant r_{\ell_0}),$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

we have by Lemma 3.2 that

$$\begin{aligned}\mathcal{B}_{\ell_0}(\{f''_{\ell_0,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_{\ell_0}\}) &\geqslant \mathcal{B}_{\ell_0}(\{f''_{\ell_0,r}(\mathbf{x})|_{\mathbf{w}_{\ell_0-1}=\mathbf{0}} : 1 \leqslant r \leqslant r_{\ell_0}\}) \\ &= \mathcal{B}_{\ell_0}(\{f_{\ell_0,r}(\mathbf{x})|_{\mathbf{w}_{\ell_0-1}=\mathbf{0}} : 1 \leqslant r \leqslant r_{\ell_0}\}) \\ &\geqslant \mathcal{B}_{\ell_0}(\mathbf{f}_{\ell_0}) - (r_{\ell_0}+1)(K_2+\cdots+K_{\ell_0-1}).\end{aligned}$$

Consequently, we have that $\mathcal{B}_{\ell_0}(\{f''_{\ell_0,r} : 1 \leqslant r \leqslant r_{\ell_0}\})$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. Also it follows by basic facts about reduction in Gröbner basis theory that we may write

$$\chi''_{\ell_0,r}(\mathbf{x}) = \chi'_{\ell_0,r}(\mathbf{x}) + \sum_{2\leqslant \ell' < \ell_0} \sum_{1 \leqslant r' \leqslant r_{\ell'}} \zeta_{\ell_0,r:\ell',r'}(\mathbf{x}) f_{\ell',r}(\mathbf{x}), \tag{4.6}$$

where $\chi'_{\ell_0,r}(\mathbf{x})$ is a polynomial which does not contain any monomial divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(2 \leqslant \ell < \ell_0, 1 \leqslant r \leqslant r_\ell)$. Furthermore, $\chi'_{\ell_0,r}(\mathbf{x})$ is a polynomial of degree less than or equal to ℓ_0, and $\zeta_{\ell_0,r:\ell',r'}(\mathbf{x})$ is a polynomial of degree less than or equal to $\ell_0 - \ell'$. We obtain by the definition of the h-invariant that

$$\begin{aligned}h_{\ell_0}(\chi'_{\ell_0,r}) &\leqslant h_{\ell_0}(\chi''_{\ell_0,r}(\mathbf{x})) + h_{\ell_0}\left(\sum_{2\leqslant \ell' < \ell_0} \sum_{1 \leqslant r' \leqslant r_{\ell'}} \zeta_{\ell_0,r:\ell',r'}(\mathbf{x}) f_{\ell',r}(\mathbf{x})\right) \\ &\leqslant \sum_{\ell=2}^{\ell_0-1} r_\ell + \sum_{\ell=2}^{\ell_0-1} r_\ell \end{aligned} \tag{4.7}$$

for each $1 \leqslant r \leqslant r_{\ell_0}$. Also, via (4.6) we can reduce $\mathbf{f}_{\ell_0}$ of (4.5) without changing the solution set, and assume it is of the shape

$$f_{\ell_0,r}(\mathbf{x}) = \chi'_{\ell_0,r}(\mathbf{x}) + f''_{\ell_0,r}(\mathbf{x}) \quad (1 \leqslant r \leqslant r_{\ell_0}), \tag{4.8}$$

where none of the monomials of $f''_{\ell_0,r}(\mathbf{x})$ or $\chi'_{\ell_0,r}(\mathbf{x})$ is divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(2 \leqslant \ell < \ell_0, 1 \leqslant r \leqslant r_\ell)$.

We then consider $\mathbf{F}_{\ell_0} = (F_{\ell_0,1}, \ldots, F_{\ell_0,r_{\ell_0}})$ where $F_{\ell_0,r}$ is the homogeneous degree ℓ_0 portion of $f_{\ell_0,r}$ in (4.8). From each $F_{\ell_0,1}, \ldots, F_{\ell_0,r_{\ell_0}}$, we collect the coefficient of the monomial $x_{i_1} \cdots x_{i_{\ell_0}}$ and turn it into a vector in $\mathbb{Q}^{r_{\ell_0}}$. We do this for every degree ℓ_0 monomial. We then form a matrix by putting these vectors in columns from left to right in the decreasing order of the degree ℓ_0 monomials. From the definition of the h-invariant, we can deduce

$$h_{\ell_0}(\mathbf{F}_{\ell_0}) + \sum_{r=1}^{r_{\ell_0}} h_{\ell_0}(\chi'_{\ell_0,r}(\mathbf{x})) \geqslant h_{\ell_0}(\{f''_{\ell_0,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_{\ell_0}\}).$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Consequently, we obtain from (2.4) and (4.7) that

$$h_{\ell_0}(\mathbf{F}_{\ell_0}) \geqslant 2^{1-\ell_0}\mathcal{B}_{\ell_0}(\{f''_{\ell_0,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_{\ell_0}\}) - 2r_{\ell_0}\sum_{\ell=2}^{\ell_0-1} r_\ell.$$

Thus it follows that $h_{\ell_0}(\mathbf{F}_{\ell_0})$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. In particular, since we have $h_{\ell_0}(\mathbf{F}_{\ell_0}) > 0$, the coefficient matrix of $\mathbf{F}_{\ell_0}$ above has full rank. We row reduce this matrix, and we denote the r_{ℓ_0} monomials where the leading 1's occur to be $\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ $(1 \leqslant r \leqslant r_{\ell_0})$. We then take the distinct variables involved in these r_{ℓ_0} monomials that have not yet appeared in $\mathbf{w}_{\ell_0-1}$, and label them as

$$w_{r_1+K_2+\cdots+K_{\ell_0-1}+1}, \ldots, w_{r_1+K_2+\cdots+K_{\ell_0-1}+K_{\ell_0}}.$$

Clearly we have $K_{\ell_0} \leqslant \ell_0 r_{\ell_0}$. We adjoin these variables to $\mathbf{w}_{\ell_0-1}$, and let $\mathbf{w}_{\ell_0} = (w_{r_1+1}, \ldots, w_{r_1+K_2+\cdots+K_{\ell_0}})$.

From the row reduction operations done on the coefficient matrix, without changing the solution set we can reduce $\mathbf{f}_{\ell_0}$ to

$$f_{\ell_0,r}(\mathbf{x}) = c_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}} + \chi_{\ell_0,r}(\mathbf{x}) + \widetilde{f}_{\ell_0,r}(\mathbf{x}) \quad (1 \leqslant r \leqslant r_{\ell_0}), \tag{4.9}$$

where $\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ is the leading monomial of $f_{\ell_0,r}$, with respect to the graded lexicographic ordering, and none of the monomials of $\widetilde{f}_{\ell_0,r}(\mathbf{x})$ or $\chi_{\ell_0,r}(\mathbf{x})$ is divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell',r'}}$ $(1 \leqslant \ell' \leqslant \ell_0, 1 \leqslant r' \leqslant r_\ell)$. Also $\chi_{\ell_0,r}(\mathbf{x}) + c^{(1)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ is a $\mathbb{Q}$-linear combination of $\chi'_{\ell_0,1}, \ldots, \chi'_{\ell_0,r_{\ell_0}}$, and similarly $\widetilde{f}_{\ell_0,r}(\mathbf{x}) + c^{(2)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ is a $\mathbb{Q}$-linear combination of $f''_{\ell_0,1}, \ldots, f''_{\ell_0,r_{\ell_0}}$ for some appropriate rational coefficients $c^{(1)}_{\ell_0,r}$ and $c^{(2)}_{\ell_0,r}$, where $c^{(1)}_{\ell_0,r} + c^{(2)}_{\ell_0,r} = c_{\ell_0,r}$. It then follows by the definition of the h-invariant and (4.7) that

$$h(\chi_{\ell_0,r}) \leqslant 1 + \sum_{r=1}^{r_{\ell_0}} h_{\ell_0}(\chi'_{\ell_0,r}) \leqslant 1 + 2r_{\ell_0}\sum_{\ell=2}^{\ell_0-1} r_\ell \quad (1 \leqslant r \leqslant r_{\ell_0}).$$

We obtained $\widetilde{f}_{\ell_0,r}(\mathbf{x}) + c^{(2)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ as a $\mathbb{Q}$-linear combination of $f''_{\ell_0,1}, \ldots, f''_{\ell_0,r_{\ell_0}}$, where the $\mathbb{Q}$-linear combination came from the row reduction operations applied to the coefficient matrix of $\mathbf{F}_{\ell_0}$. Thus by the basic properties of the Birch rank, it follows that

$$\mathcal{B}_{\ell_0}(\{\widetilde{f}_{\ell_0,r}(\mathbf{x}) + c^{(2)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}} : 1 \leqslant r \leqslant r_{\ell_0}\}) = \mathcal{B}_{\ell_0}(\{f''_{\ell_0,r} : 1 \leqslant r \leqslant r_{\ell_0}\}).$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Therefore, we obtain by Lemma 3.2 that

$$\begin{aligned}
&\mathcal{B}_{\ell_0}(\{\widetilde{f}_{\ell_0,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_{\ell_0}\}) \\
&\quad \geqslant \mathcal{B}_{\ell_0}(\{\widetilde{f}_{\ell_0,r}(\mathbf{x})|_{\mathbf{w}_{\ell_0}=\mathbf{0}} : 1 \leqslant r \leqslant r_{\ell_0}\}) \\
&\quad = \mathcal{B}_{\ell_0}(\{(\widetilde{f}_{\ell_0,r}(\mathbf{x}) + c^{(2)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}})|_{\mathbf{w}_{\ell_0}=\mathbf{0}} : 1 \leqslant r \leqslant r_{\ell_0}\}) \\
&\quad \geqslant \mathcal{B}_{\ell_0}(\{\widetilde{f}_{\ell_0,r}(\mathbf{x}) + c^{(2)}_{\ell_0,r}\mathbf{w}^{\mathbf{j}_{\ell_0,r}} : 1 \leqslant r \leqslant r_{\ell_0}\}) - (r_{\ell_0}+1)(K_2 + \cdots + K_{\ell_0}) \\
&\quad = \mathcal{B}_{\ell_0}(\{f''_{\ell_0,r} : 1 \leqslant r \leqslant r_{\ell_0}\}) - (r_{\ell_0}+1)(K_2 + \cdots + K_{\ell_0}).
\end{aligned}$$

Thus we have that $\mathcal{B}_{\ell_0}(\{\widetilde{f}_{\ell_0,r}(\mathbf{x}) : 1 \leqslant r \leqslant r_{\ell_0}\})$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. It then follows by a similar argument given above, to show the h-invariant of $\mathbf{f}_{\ell_0}$ in (4.8) is sufficiently large, that the h-invariant of $\mathbf{f}_{\ell_0}$ in (4.9) is sufficiently large with respect to d and $r_d, \ldots, r_1$. Finally, we also have by the construction that for each $1 \leqslant r \leqslant r_{\ell_0}$, the monomial $\mathbf{w}^{\mathbf{j}_{\ell_0,r}}$ is not divisible by any one of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(1 \leqslant \ell < \ell_0, 1 \leqslant r \leqslant r_\ell)$. Thus we have completed induction. Therefore, we obtain that we can transform system (4.1) into system (4.2) without changing the solution set. Let us adjoin $\mathbf{w}_d$ to $(w_1, \ldots, w_{r_1})$ and denote the resulting set of variables to be $\mathbf{w} = (w_1, \ldots, w_K)$, where

$$K = r_1 + K_2 + \cdots + K_d \leqslant \sum_{\ell=1}^{d} \ell r_\ell. \tag{4.10}$$

We also add that if there are any ℓ with $r_\ell = 0$, we simply skip these cases in the above argument.

Let $\alpha_{\ell,r} \in \mathbb{R}$ $(1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$, and consider

$$\sum_{1 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} f_{\ell,r}(\mathbf{x})$$

as a polynomial in $x_1, \ldots, x_n$ with real coefficients, where $\mathbf{f}$ is the system of polynomials in (4.2). Given any $1 \leqslant \ell \leqslant d$ and $1 \leqslant r \leqslant r_\ell$, it follows from the construction that the coefficient of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ of the above polynomial is $c_{\ell,r}\alpha_{\ell,r}$. Let $\mathbf{x} = (\mathbf{w}, \mathbf{x}')$. Let us also fix $\mathbf{x}' = \mathbf{x}'_0 \in \mathbb{Z}^{n-K}$. It is clear that if we consider

$$\sum_{1 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} f_{\ell,r}(\mathbf{w}, \mathbf{x}'_0) \tag{4.11}$$

as a polynomial in $\mathbf{w}$ with real coefficients, then given $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$ we still have that the coefficient of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ of this polynomial is $c_{\ell,r}\alpha_{\ell,r}$. Furthermore, this polynomial does not contain any monomial divisible by $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ other than itself.

We set $R = r_d + \cdots + r_1$. Let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_d, \ldots, \boldsymbol{\alpha}_1) \in \mathbb{R}^R$ where $\boldsymbol{\alpha}_\ell = (\alpha_{\ell,1}, \ldots, \alpha_{\ell,r_\ell}) \in \mathbb{R}^{r_\ell}$ $(1 \leqslant \ell \leqslant d)$. Similarly, we denote $\mathbf{a} = (\mathbf{a}_d, \ldots, \mathbf{a}_1) \in (\mathbb{Z}/q\mathbb{Z})^R$, where $q \in \mathbb{N}$ and $\mathbf{a}_\ell = (a_{\ell,1}, \ldots, a_{\ell,r_\ell}) \in (\mathbb{Z}/q\mathbb{Z})^{r_\ell}$ $(1 \leqslant \ell \leqslant d)$. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\|\beta\|$ denote the distance from $\beta \in \mathbb{R}$ to the nearest integer, which induces a metric on $\mathbb{T}$ via $d(\alpha, \beta) = \|\alpha - \beta\|$. For a given value of $C > 0$ and an integer $1 \leqslant q \leqslant (\log X)^C$, we define

$$\mathfrak{M}_{\mathbf{a},q}(C) = \left\{ \boldsymbol{\alpha} \in \mathbb{T}^R : \max_{1 \leqslant r \leqslant r_\ell} \|\alpha_{\ell,r} - a_{\ell,r}/q\| \leqslant X^{-\ell}(\log X)^C \ (1 \leqslant \ell \leqslant d) \right\}$$

for each $\mathbf{a} \in (\mathbb{Z}/q\mathbb{Z})^R$ with $\gcd(\mathbf{a}, q) = 1$ (by which we mean that the greatest common divisor of the numbers $a_{d,1}, \ldots, a_{1,r_1}$ and q is 1). These arcs are disjoint for X sufficiently large.

We define the *major arcs* to be

$$\mathfrak{M}(C) = \bigcup_{q \leqslant (\log X)^C} \bigcup_{\substack{\mathbf{a} \in (\mathbb{Z}/q\mathbb{Z})^R \\ \gcd(\mathbf{a},q)=1}} \mathfrak{M}_{\mathbf{a},q}(C),$$

and define the *minor arcs* to be

$$\mathfrak{m}(C) = \mathbb{T}^R \backslash \mathfrak{M}(C).$$

In other words, the major arcs is a collection of elements in $\mathbb{T}^R$ that can be simultaneously 'well approximated' by rational numbers of the same denominator q, where q is 'small'.

For a system of polynomials $\mathbf{f}$, we define

$$T(\mathbf{f}; \boldsymbol{\alpha}) := \sum_{\mathbf{x} \in [0,X]^n} \Lambda(\mathbf{x}) e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{x}) \cdot \alpha_{\ell,r} \right), \tag{4.12}$$

where we defined $\Lambda(\mathbf{x})$ in (1.5). By the orthogonality relation, we have

$$\begin{aligned} \mathcal{M}_{\mathbf{f}}(X) &= \sum_{\mathbf{x} \in [0,X]^n} \Lambda(\mathbf{x}) \mathbf{1}_{V_{\mathbf{f},\mathbf{0}}(\mathbb{C})}(\mathbf{x}) \\ &= \int_{\mathbb{T}} \cdots \int_{\mathbb{T}} T(\mathbf{f}; \boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha} \\ &= \int_{\mathfrak{M}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha} + \int_{\mathfrak{m}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha}. \end{aligned} \tag{4.13}$$

For the system of polynomials $\mathbf{f}$ in (4.2), we prove the following results on the minor arcs and the major arcs.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

PROPOSITION 4.1. *Let* $\mathbf{f}$ *be the polynomials in* (4.2). *Given any* $c > 0$, *for sufficiently large* $C > 0$ *we have*

$$\int_{\mathfrak{m}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, d\boldsymbol{\alpha} \ll \frac{X^{n-\sum_{\ell=1}^{d} \ell r_\ell}}{(\log X)^c}.$$

PROPOSITION 4.2. *Let* $\mathbf{f}$ *be the polynomials in* (4.2). *Given any* $c > 0$, *for sufficiently large* $C > 0$ *we have*

$$\int_{\mathfrak{M}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \mathcal{C}(\mathbf{f}) X^{n-\sum_{\ell=1}^{d} \ell r_\ell} + O\left(\frac{X^{n-\sum_{\ell=1}^{d} \ell r_\ell}}{(\log X)^c}\right),$$

where $\mathcal{C}(\mathbf{f})$ *is a constant that depends only on* $\mathbf{f}$.

We prove Proposition 4.1 in Section 5, and Proposition 4.2 in Section 7.

5. Hardy–Littlewood circle method: minor arcs

Proof of Proposition 4.1. We consider the system of polynomials $\mathbf{f}$ in (4.2) constructed in the previous section, which satisfies all the conditions described below (4.2). Recall we denote $\mathbf{w} = (w_1, \ldots, w_K)$, where $K \leqslant dR$ and $R = \sum_{\ell=1}^{d} r_\ell$. We let $\widetilde{\mathbf{F}} = (\widetilde{\mathbf{F}}_d, \ldots, \widetilde{\mathbf{F}}_1)$ be the system of forms such that for each $1 \leqslant \ell \leqslant d$, $\widetilde{\mathbf{F}}_\ell = (\widetilde{F}_{\ell,1}, \ldots, \widetilde{F}_{\ell,r_\ell})$ and $\widetilde{F}_{\ell,r}$ is the homogeneous degree ℓ portion of $\widetilde{f}_{\ell,r}$ $(1 \leqslant r \leqslant r_\ell)$. For each $1 \leqslant \ell \leqslant d$, we know that $\mathcal{B}_\ell(\widetilde{\mathbf{F}}_\ell)$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. Thus we apply Proposition 3.4 to the system

$$(\widetilde{\mathbf{F}}_d|_{\mathbf{w}=\mathbf{0}}, \ldots, \widetilde{\mathbf{F}}_1|_{\mathbf{w}=\mathbf{0}}),$$

and denote the partition of variables of $\mathbf{x}\backslash\mathbf{w}$ we obtain by $(\mathbf{y}, \mathbf{z})$ so that $\mathbf{x} = (\mathbf{w}, \mathbf{y}, \mathbf{z})$. Let

$$\widetilde{Q}_{\ell,r}(\mathbf{y}, \mathbf{z}) = \widetilde{F}_{\ell,r}(\mathbf{0}, \mathbf{y}, \mathbf{z}) - \widetilde{F}_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell).$$

Then the partition of variables $\mathbf{x} = (\mathbf{w}, \mathbf{y}, \mathbf{z})$ satisfies

$$|\mathbf{y}| = M \leqslant \sum_{\ell=1}^{d} r_\ell C_{\ell,1}^{\bullet}, \tag{5.1}$$

and also

$$\mathcal{B}_\ell(\{\widetilde{Q}_{\ell,r}(\mathbf{y}, \mathbf{z}) : 1 \leqslant r \leqslant r_\ell\}) \geqslant C_{\ell,1}^{\bullet} - r_\ell \sum_{j=1}^{\ell-1} C_{j,1}^{\bullet} r_j \quad (2 \leqslant \ell \leqslant d), \tag{5.2}$$

$$\mathcal{B}_1(\{\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0}) : 1 \leqslant r \leqslant r_1\}) \geqslant C_{1,1}^{\bullet}, \tag{5.3}$$

and

$$\mathcal{B}_\ell(\{\widetilde{F}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) : 1 \leqslant r \leqslant r_\ell\}) \geqslant C^\bullet_{\ell,2} - r_\ell \sum_{j=1}^{\ell-1} C^\bullet_{j,1} r_j \quad (1 \leqslant \ell \leqslant d), \qquad (5.4)$$

where $C^\bullet_{\ell,1}$ and $C^\bullet_{\ell,2}$ $(1 \leqslant \ell \leqslant d)$ are positive integer constants depending only on d and $r_d, \ldots, r_1$ to be chosen later. In particular, we will make sure that the right hand side of (5.4) for $2 \leqslant \ell \leqslant d$ is sufficiently large with respect to d and $r_d, \ldots, r_1$. For notational convenience, we label $\mathbf{y} = (y_1, \ldots, y_M)$ and $\mathbf{z} = (z_1, \ldots, z_{n-M-K})$.

We then apply Proposition 3.4 (with $C_{1,1} = C_{1,2} = C_{2,2} = \cdots = C_{d,2} = 1$) to the system of forms $(\widetilde{\mathbf{F}}_d(\mathbf{0},\mathbf{0},\mathbf{z}), \ldots, \widetilde{\mathbf{F}}_2(\mathbf{0},\mathbf{0},\mathbf{z}))$, where $\widetilde{\mathbf{F}}_\ell(\mathbf{0},\mathbf{0},\mathbf{z}) = (\widetilde{F}_{\ell,1}(\mathbf{0},\mathbf{0},\mathbf{z}), \ldots, \widetilde{F}_{\ell,r_\ell}(\mathbf{0},\mathbf{0},\mathbf{z}))$ for each $2 \leqslant \ell \leqslant d$. Let the partition of variables we obtain to be $\mathbf{z} = (\mathbf{a}, \mathbf{b})$, which satisfies

$$|\mathbf{a}| = M' \leqslant \sum_{\ell=2}^{d} r_\ell C^\star_{\ell,1}, \qquad (5.5)$$

and

$$\mathcal{B}_\ell(\{\widetilde{P}_{\ell,r}(\mathbf{a},\mathbf{b}) : 1 \leqslant r \leqslant r_\ell\}) \geqslant C^\star_{\ell,1} - r_\ell \sum_{j=2}^{\ell-1} C^\star_{j,1} r_j \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell), \qquad (5.6)$$

where

$$\widetilde{P}_{\ell,r}(\mathbf{a},\mathbf{b}) = \widetilde{F}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - \widetilde{F}_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b})) \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell).$$

Note we are only mentioning the statement (5.6) for $2 \leqslant \ell < d$, because we will not be needing it for the case $\ell = d$. Recall from (4.2) we have for $2 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$,

$$f_{\ell,r}(\mathbf{x}) = c_{\ell,r}\mathbf{w}^{\mathbf{j}_{\ell,r}} + \chi_{\ell,r}(\mathbf{x}) + \widetilde{f}_{\ell,r}(\mathbf{x}),$$

where

$$h_\ell(\chi_{\ell,r}) \leqslant C''_0$$

for some constant C''_0 dependent only on d and $r_d, \ldots, r_1$. Let $\chi^{(\ell)}_{\ell,r}(\mathbf{x})$ denote the homogeneous degree ℓ portion of $\chi_{\ell,r}(\mathbf{x})$. Then it is easy to deduce from the definition of the h-invariant that the quantities

$$h_\ell(\chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z})),\ h_\ell(\chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b}))),$$
$$h_\ell(\chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}))$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

are all bounded by $2C_0''$. We then let

$$Q_{\ell,r}(\mathbf{y},\mathbf{z}) = \widetilde{Q}_{\ell,r}(\mathbf{y},\mathbf{z}) + \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) \tag{5.7}$$

and

$$P_{\ell,r}(\mathbf{a},\mathbf{b}) = \widetilde{P}_{\ell,r}(\mathbf{a},\mathbf{b}) + \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - \chi^{(\ell)}_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b})). \tag{5.8}$$

We remark that from the definition it follows that $Q_{\ell,r}(\mathbf{y},\mathbf{z})$ is precisely the homogeneous degree ℓ portion of the polynomial $f_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - f_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z})$. Clearly every monomial of $Q_{\ell,r}(\mathbf{y},\mathbf{z})$ with nonzero coefficient contains at least one of the $\mathbf{y}$ variables, and hence $h_\ell(Q_{\ell,r}(\mathbf{y},\mathbf{z})) \leqslant |\mathbf{y}|$. Similarly, $P_{\ell,r}(\mathbf{a},\mathbf{b})$ is precisely the homogeneous degree ℓ portion of the polynomial $f_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - f_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b}))$. Clearly every monomial of $P_{\ell,r}(\mathbf{a},\mathbf{b})$ with nonzero coefficient contains at least one of the $\mathbf{a}$ variables, and hence $h_\ell(P_{\ell,r}(\mathbf{a},\mathbf{b})) \leqslant |\mathbf{a}|$.

We obtain the following three inequalities from (5.2), (5.6), and (5.4), respectively, by applying (2.4) and the definition of the h-invariant with the comment before (5.7),

$$\begin{aligned} &h_\ell(\{Q_{\ell,r}(\mathbf{y},\mathbf{z}) : 1 \leqslant r \leqslant r_\ell\}) \\ &\quad \geqslant 2^{1-\ell}\left(C^\bullet_{\ell,1} - r_\ell \sum_{j=1}^{\ell-1} C^\bullet_{j,1} r_j\right) - 2r_\ell C_0'' \quad (2 \leqslant \ell \leqslant d), \end{aligned} \tag{5.9}$$

$$\begin{aligned} &h_\ell(\{P_{\ell,r}(\mathbf{a},\mathbf{b}) : 1 \leqslant r \leqslant r_\ell\}) \\ &\quad \geqslant 2^{1-\ell}\left(C^\star_{\ell,1} - r_\ell \sum_{j=2}^{\ell-1} C^\star_{j,1} r_j\right) - 2r_\ell C_0'' \quad (2 \leqslant \ell < d), \end{aligned} \tag{5.10}$$

and

$$\begin{aligned} &h_d(\{\widetilde{F}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \chi^{(d)}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) : 1 \leqslant r \leqslant r_d\}) \\ &\quad \geqslant 2^{1-d}\left(C^\bullet_{d,2} - r_d \sum_{j=1}^{d-1} C^\bullet_{j,1} r_j\right) - 2r_d C_0''. \end{aligned} \tag{5.11}$$

It is clear from the definition that the homogeneous degree d portion of $f_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z})$ is precisely $\widetilde{F}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \chi^{(d)}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z})$ for each $1 \leqslant r \leqslant r_d$. Thus we have

$$F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) = \widetilde{F}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \chi^{(d)}_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) \quad (1 \leqslant r \leqslant r_d),$$

and we also let

$$\mathbf{F}_d(\mathbf{0},\mathbf{0},\mathbf{z}) = (F_{d,1}(\mathbf{0},\mathbf{0},\mathbf{z}), \ldots, F_{d,r_d}(\mathbf{0},\mathbf{0},\mathbf{z})). \tag{5.12}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

With the notations we have defined so far, we have

$$F_{d,r}(\mathbf{0},\mathbf{y},\mathbf{z}) = Q_{d,r}(\mathbf{y},\mathbf{z}) + F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) \quad (1 \leqslant r \leqslant r_d),$$

and for each $2 \leqslant \ell \leqslant d$,

$$F_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) = Q_{\ell,r}(\mathbf{y},\mathbf{z}) + P_{\ell,r}(\mathbf{a},\mathbf{b}) + F_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b})) \quad (1 \leqslant r \leqslant r_\ell),$$

where

$$F_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) = P_{\ell,r}(\mathbf{a},\mathbf{b}) + F_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b})) \quad (1 \leqslant r \leqslant r_\ell).$$

Let $2 \leqslant \ell \leqslant d$. For each $1 \leqslant r \leqslant r_\ell$, the partition of variables $\mathbf{x} = (\mathbf{w},\mathbf{y},\mathbf{z})$ gives the decomposition of the following shape

$$\begin{aligned} f_{\ell,r}(\mathbf{w},\mathbf{y},\mathbf{z}) &= f_{\ell,r}(\mathbf{w},\mathbf{0},\mathbf{0}) + \sum_{j=1}^{\ell-1} \sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant K} \left(\sum_{k=1}^{\ell-j} \Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y},\mathbf{z}) \right) w_{i_1}\cdots w_{i_j} \\ &\quad + \sum_{j=1}^{\ell-1} \sum_{1\leqslant t_1\leqslant\cdots\leqslant t_j\leqslant M} \left(\sum_{k=0}^{\ell-j} \Psi^{(k)}_{\ell,r:t_1,\ldots,t_j}(\mathbf{z}) \right) y_{t_1}\cdots y_{t_j} + F_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{0}) \\ &\quad + F_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \sum_{k=1}^{\ell-1} G^{(k)}_{\ell,r}(\mathbf{z}), \end{aligned} \tag{5.13}$$

which we describe below. We note that $\Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y},\mathbf{z})$ and $\Psi^{(k)}_{\ell,r:t_1,\ldots,t_j}(\mathbf{z})$ are forms of degree k. The above decomposition establishes the following. The term

$$f_{\ell,r}(\mathbf{w},\mathbf{0},\mathbf{0}) + \sum_{j=1}^{\ell-1} \sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant K} \left(\sum_{k=1}^{\ell-j} \Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y},\mathbf{z}) \right) w_{i_1}\cdots w_{i_j}$$

consists of all the monomials of $f_{\ell,r}(\mathbf{x})$ which involve any variables of $\mathbf{w}$, and also the constant term. The term

$$\sum_{j=1}^{\ell-1} \sum_{1\leqslant t_1\leqslant\cdots\leqslant t_j\leqslant M} \left(\sum_{k=0}^{\ell-j} \Psi^{(k)}_{\ell,r:t_1,\ldots,t_j}(\mathbf{z}) \right) y_{t_1}\cdots y_{t_j} + F_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{0}) \tag{5.14}$$

consists of all the monomials of $f_{\ell,r}(\mathbf{x})$ which involve any variables of $\mathbf{y}$ and do not involve any of the $\mathbf{w}$ variables. In other words, it is precisely $f_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - f_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z})$. Finally, we have the terms which only involve the $\mathbf{z}$ variables

$$F_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \sum_{k=1}^{\ell-1} G^{(k)}_{\ell,r}(\mathbf{z}),$$

where $G^{(k)}_{\ell,r}(\mathbf{z})$ is the homogeneous degree k portion of $f_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z})$.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We denote

$$\Phi = \{\Phi^{(k)}_{\ell,r:i_1,\ldots,i_j} : 2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell, 1 \leqslant j \leqslant \ell - 1,\\ 1 \leqslant i_1 \leqslant \cdots \leqslant i_j \leqslant K, 1 \leqslant k \leqslant \ell - j\}.$$

Note every form of Φ has degree strictly less than d, and involves only the $\mathbf{y}$ variables and the $\mathbf{z}$ variables. We shall use the notation $|\Phi|$ to denote the number of forms in Φ, and other instances of notation of this type should be interpreted in a similar manner. Clearly we have

$$|\Phi| \leqslant \sum_{\ell=2}^{d} r_\ell \ell^2 K^\ell \leqslant \sum_{\ell=2}^{d} R\ell^2 (dR)^\ell \leqslant d^{d+3} R^{d+1}.$$

Recall the function $\rho_{d,\ell}$ defined in (2.6). We apply Proposition 2.7 to the system Φ with respect to the partition of variables $(\mathbf{y}, \mathbf{z})$ and the functions $\boldsymbol{\mathcal{F}} = \{\mathcal{F}_2, \ldots, \mathcal{F}_{d-1}\}$, where

$$\mathcal{F}_i(t) = \rho_{d,d}(2R + 2t) + 2t + 4r_1\\ + 2R(dR(R^2+1)^{d-2} 2^d (\rho_{d,d}(2R+2t) + 2t + 4r_1 + 2RC''_0)\\ + dR^3(R^2+1)^{d-2}(2t+1)),$$

for each $2 \leqslant i \leqslant d-1$, and obtain $\mathcal{R}(\Phi) = (\mathcal{R}^{(d-1)}(\Phi), \ldots, \mathcal{R}^{(1)}(\Phi))$. For each $1 \leqslant s \leqslant d-1$,

$$\mathcal{R}^{(s)}(\Phi) = \{A_i^{(s)} : 1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Phi)|\}$$

is precisely all the degree s forms of $\mathcal{R}(\Phi)$. For each form $A_i^{(s)} \in \mathcal{R}^{(s)}(\Phi)$ $(1 \leqslant s \leqslant d-1, 1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Phi)|)$, we write

$$A_i^{(s)}(\mathbf{y}, \mathbf{z}) = \sum_{k=0}^{s} \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant M} \widetilde{\Psi}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{z}) y_{i_1} \cdots y_{i_k}, \tag{5.15}$$

where each $\widetilde{\Psi}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{z})$ is a form of degree $s-k$. Thus each form $A_i^{(s)}$ introduces at most $(s+1)M^s \leqslant dM^d$ forms in $\mathbf{z}$. Also for each $1 \leqslant s \leqslant d-1$, we denote $\overline{\mathcal{R}}^{(s)}(\Phi)$ to be the system obtained by removing from $\mathcal{R}^{(s)}(\Phi)$ all forms which depend only on the $\mathbf{z}$ variables. Let $\overline{\mathcal{R}}(\Phi) = (\overline{\mathcal{R}}^{(d-1)}(\Phi), \ldots, \overline{\mathcal{R}}^{(1)}(\Phi))$, $R_2 = \sum_{s=1}^{d-1} |\overline{\mathcal{R}}^{(s)}(\Phi)|$, and $D_2 = \sum_{s=1}^{d-1} s\, |\overline{\mathcal{R}}^{(s)}(\Phi)|$. By relabeling if necessary, for each $1 \leqslant s \leqslant d-1$ we denote the elements of $\overline{\mathcal{R}}^{(s)}(\Phi)$ by

$$\overline{\mathcal{R}}^{(s)}(\Phi) = \{A_i^{(s)} : 1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(s)}(\Phi)|\}. \tag{5.16}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let

$$\begin{aligned}\Psi = \{&F_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b})) : 2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell\}\\ &\cup\{G^{(k)}_{\ell,r}(\mathbf{z}) : 2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell, 1 \leqslant k \leqslant \ell - 1\}\\ &\cup\{\Psi^{(k)}_{\ell,r:t_1,\dots,t_j}(\mathbf{z}) : 2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell, 1 \leqslant j \leqslant \ell - 1,\\ &\quad 1 \leqslant t_1 \leqslant \dots \leqslant t_j \leqslant M, 1 \leqslant k \leqslant \ell - j\}\\ &\cup\{\widetilde{\Psi}^{(s-k)}_{s,i:i_1,\dots,i_k}(\mathbf{z}) : 1 \leqslant s \leqslant d-1, 1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Phi)|, 0 \leqslant k < s,\\ &\quad 1 \leqslant i_1 \leqslant \dots \leqslant i_k \leqslant M\}.\end{aligned}$$

In other words, Ψ is the collection of all $G^{(k)}_{\ell,r}(\mathbf{z})$, $\Psi^{(k)}_{\ell,r:t_1,\dots,t_j}(\mathbf{z})$, $\widetilde{\Psi}^{(s-k)}_{s,i:i_1,\dots,i_k}(\mathbf{z})$ except the constants, and all $F_{\ell,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b}))$ but not $F_{d,r}(\mathbf{0},\mathbf{0},(\mathbf{0},\mathbf{b}))$. In particular, every form of Ψ has degree strictly less than d. We can see that

$$|\Psi| \leqslant R + dR + Rd^2M^d + |\mathcal{R}(\Phi)|dM^d.$$

Furthermore, every form of Ψ is only in terms of the $\mathbf{z}$ variables.

We let $\mathcal{R}(\Psi) = (\mathcal{R}^{(d-1)}(\Psi), \dots, \mathcal{R}^{(1)}(\Psi))$ be a regularization of Ψ with respect to the functions $\boldsymbol{\mathcal{F}}' = \{\mathcal{F}'_2, \dots, \mathcal{F}'_{d-1}\}$, where

$$\begin{aligned}\mathcal{F}'_i(t) = \rho_{d,d}&(2R + 2t) + 2t + 4r_1\\ &+ 2R(dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R + 2t) + 2t + 4r_1 + 2RC''_0))\end{aligned}$$

for each $2 \leqslant i \leqslant d-1$. For each $1 \leqslant s \leqslant d-1$,

$$\mathcal{R}^{(s)}(\Psi) = \{V^{(s)}_i : 1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Psi)|\}$$

is precisely all the degree s forms of $\mathcal{R}(\Psi)$. Let $R_1 = \sum_{s=1}^{d-1} |\mathcal{R}^{(s)}(\Psi)|$ and $D_1 = \sum_{s=1}^{d-1} s|\mathcal{R}^{(s)}(\Psi)|$.

Let $\Phi^{(j)}$ denote the degree j forms of Φ. It follows from Proposition 2.7 that each $|\mathcal{R}^{(i)}(\Phi)|$ is bounded by some constant dependent only on $\boldsymbol{\mathcal{F}}$ and $|\Phi^{(d-1)}|, \dots, |\Phi^{(1)}|$. Thus we see that $|\mathcal{R}(\Phi)|$ and R_2 are bounded by a constant dependent only on d and $r_d, \dots, r_1$. It also follows from Proposition 2.7 that each $|\mathcal{R}^{(i)}(\Psi)|$ is bounded by some constant dependent only on $\boldsymbol{\mathcal{F}}'$, d, R, M, and $|\mathcal{R}(\Phi)|$. Thus we obtain that R_1 is bounded by a constant dependent only on M, d and $r_d, \dots, r_1$.

We first set $C^{\bullet}_{1,1} = 2R_2 + 1$. We now set the values of $C^{\bullet}_{\ell,1}$ $(2 \leqslant \ell \leqslant d)$ such that they satisfy

$$2^{1-\ell}\left(C^{\bullet}_{\ell,1} - r_\ell \sum_{j=1}^{\ell-1} C^{\bullet}_{j,1} r_j\right) - 2r_\ell C''_0 \geqslant \rho_{d,\ell}(2R + 2R_2) + 2R_2 + 4r_1. \qquad (5.17)$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Since (5.17) is equivalent to

$$C^{\bullet}_{\ell,1} \geqslant 2^{\ell-1}(\rho_{d,\ell}(2R+2R_2)+2R_2+4r_1+2r_\ell C''_0)+r_\ell\sum_{j=1}^{\ell-1}C^{\bullet}_{j,1}r_j,$$

we can also make sure $C^{\bullet}_{\ell,1}$ satisfies the additional constraint

$$C^{\bullet}_{\ell,1} \leqslant 2^{d}(\rho_{d,d}(2R+2R_2)+2R_2+4r_1+2RC''_0)+R^2\sum_{j=1}^{\ell-1}C^{\bullet}_{j,1}.$$

It is then not difficult to show by induction that

$$\begin{aligned} C^{\bullet}_{\ell,1} &\leqslant (R^2+1)^{\ell-2}2^{d}(\rho_{d,d}(2R+2R_2)+2R_2+4r_1+2RC''_0)\\ &\quad + R^2(R^2+1)^{\ell-2}C^{\bullet}_{1,1}. \end{aligned}$$

In particular, $C^{\bullet}_{\ell,1}$ is bounded by a constant dependent only on d and $r_d,\ldots,r_1$. Therefore, it follows from (5.1) that

$$\begin{aligned} M &\leqslant R\sum_{\ell=1}^{d}C^{\bullet}_{\ell,1}\\ &\leqslant dR(R^2+1)^{d-2}2^{d}(\rho_{d,d}(2R+2R_2)+2R_2+4r_1+2RC''_0)\\ &\quad + dR^3(R^2+1)^{d-2}(2R_2+1). \end{aligned} \tag{5.18}$$

Thus it follows that R_1 is bounded by a constant dependent only on d and $r_d,\ldots,r_1$.

We then set the values for $C^{\star}_{\ell,1}$ $(2\leqslant \ell\leqslant d)$ to satisfy

$$2^{1-\ell}\left(C^{\star}_{\ell,1}-r_\ell\sum_{j=2}^{\ell-1}C^{\star}_{j,1}r_j\right)-2r_\ell C''_0 \geqslant \rho_{d,\ell}(2R+2R_1)+2R_1+4r_1. \tag{5.19}$$

By a similar argument as for the $C^{\bullet}_{\ell,1}$ above, we can also make sure that $C^{\star}_{\ell,1}$ satisfies

$$C^{\star}_{\ell,1} \leqslant (R^2+1)^{\ell-2}2^{d}(\rho_{d,d}(2R+2R_1)+2R_1+4r_1+2RC''_0),$$

and it follows from (5.5) that

$$M' \leqslant dR(R^2+1)^{d-2}2^{d}(\rho_{d,d}(2R+2R_1)+2R_1+4r_1+2RC''_0). \tag{5.20}$$

In particular, each $C^{\star}_{\ell,1}$ and M' are bounded by constants dependent only on d and $r_d,\ldots,r_1$. Let us set $C^{\bullet}_{1,2}=2R_1+1$. We then make sure that for each $2\leqslant\ell\leqslant d$,

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$C^{\bullet}_{\ell,2}$ is sufficiently large with respect to $C^{\star}_{2,1}, \ldots, C^{\star}_{d,1}, C^{\bullet}_{1,1}, \ldots, C^{\bullet}_{d-1,1}, r_d, \ldots, r_1$, and d, and also that $C^{\bullet}_{d,2}$ satisfies

$$2^{1-d}\left(C^{\bullet}_{d,2} - r_d \sum_{j=1}^{d-1} C^{\bullet}_{j,1} r_j\right) - 2r_d C''_0 \geqslant \rho_{d,d}(2R + 2R_1) + 2R_1 + 4r_1. \quad (5.21)$$

We note that the three inequalities (5.17), (5.19) and (5.21) provide lower bounds for the h-invariants in (5.9), (5.10), and (5.11), respectively.

We now decompose the linear polynomials. From Proposition 2.7, we know that every linear form of $\mathcal{R}^{(1)}(\Phi)$ is either only in the $\mathbf{y}$ variables, or only in the $\mathbf{z}$ variables. First we consider the linear forms of $\overline{\mathcal{R}}^{(1)}(\Phi) = \{A_i^{(1)}(\mathbf{y}) : 1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(1)}(\Phi)|\}$, which we know to be linearly independent over $\mathbb{Q}$ and involve only the $\mathbf{y}$ variables. By considering their linear combinations, we may assume without loss of generality that these linear forms are of the shape

$$A_i^{(1)}(\mathbf{y}) = y_i + A'_i(y_{|\overline{\mathcal{R}}^{(1)}(\Phi)|+1}, \ldots, y_M) \quad (1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(1)}(\Phi)|),$$

where $A'_i(y_{|\overline{\mathcal{R}}^{(1)}(\Phi)|+1}, \ldots, y_M)$ is a linear form in the variables $y_{|\overline{\mathcal{R}}^{(1)}(\Phi)|+1}, \ldots, y_M$ with coefficients in $\mathbb{Q}$. By (5.3) and Lemma 3.2, we have

$$\mathcal{B}_1(\{\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0})|_{y_i=0\ (1\leqslant i\leqslant|\overline{\mathcal{R}}^{(1)}(\Phi)|)} : 1 \leqslant r \leqslant r_1\}) \geqslant C^{\bullet}_{1,1} - |\overline{\mathcal{R}}^{(1)}(\Phi)| \geqslant R_2 + 1 > 0.$$

Therefore, we can find r_1 variables from $y_{|\overline{\mathcal{R}}^{(1)}(\Phi)|+1}, \ldots, y_M$ such that the $r_1 \times r_1$ matrix, where the rth row consists of the coefficients of $\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0})$ of these r_1 variables, is invertible. Let us denote these variables by $\widetilde{y}_1, \ldots, \widetilde{y}_{r_1}$, and let $\widetilde{\mathbf{y}} = (\widetilde{y}_1, \ldots, \widetilde{y}_{r_1})$. We can then write

$$\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0}) = g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} + \widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0})|_{\widetilde{\mathbf{y}}=\mathbf{0}},$$

where $g_{r,1}, \ldots, g_{r,r_1} \in \mathbb{Z}$. Let $\mathcal{R}^{(1)}_+(\Phi)$ be a maximal linearly independent (over $\mathbb{Q}$) subset of

$$\mathcal{R}^{(1)}(\Phi) \cup \{\widetilde{F}_{1,1}(\mathbf{0}, \mathbf{y}, \mathbf{0})|_{\widetilde{\mathbf{y}}=\mathbf{0}}, \ldots, \widetilde{F}_{1,r_1}(\mathbf{0}, \mathbf{y}, \mathbf{0})|_{\widetilde{\mathbf{y}}=\mathbf{0}}\}.$$

The important thing to note is that by our construction, we have that the set of linear forms

$$\{g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} : 1 \leqslant r \leqslant r_1\} \cup \overline{\mathcal{R}}^{(1)}_+(\Phi)$$

is linearly independent over $\mathbb{Q}$. Here $\overline{\mathcal{R}}^{(1)}_+(\Phi)$ is the set of forms obtained by removing from $\mathcal{R}^{(1)}_+(\Phi)$ all forms that depend only on the $\mathbf{z}$ variables. Note we also have $|\overline{\mathcal{R}}^{(1)}_+(\Phi)| \leqslant |\overline{\mathcal{R}}^{(1)}(\Phi)| + r_1$.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We also decompose the linear forms $\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})$ in a similar manner. First we consider the linear forms of $\mathcal{R}^{(1)}(\Psi) = \{V_i^{(1)}(\mathbf{z}) : 1 \leqslant i \leqslant |\mathcal{R}^{(1)}(\Psi)|\}$, which we know to be linearly independent over $\mathbb{Q}$ and involve only the $\mathbf{z}$ variables. By considering their linear combinations, we may assume without loss of generality that these linear forms are of the shape

$$V_i^{(1)}(\mathbf{z}) = z_i + V_i'(z_{|\mathcal{R}^{(1)}(\Psi)|+1}, \ldots, z_{n-M-K}) \quad (1 \leqslant i \leqslant |\mathcal{R}^{(1)}(\Psi)|),$$

where $V_i'(z_{|\mathcal{R}^{(1)}(\Psi)|+1}, \ldots, z_{n-M-K})$ is a linear form in the variables $z_{|\mathcal{R}^{(1)}(\Psi)|+1}, \ldots, z_{n-M-K}$ with coefficients in $\mathbb{Q}$. By (5.4) and Lemma 3.2, we have

$$\mathcal{B}_1(\{\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})|_{z_i=0\ (1\leqslant i\leqslant|\mathcal{R}^{(1)}(\Psi)|)} : 1 \leqslant r \leqslant r_1\}) \geqslant C^{\bullet}_{1,2} - |\mathcal{R}^{(1)}(\Psi)| \geqslant R_1 + 1 > 0.$$

Therefore, we can find r_1 variables from $z_{|\mathcal{R}^{(1)}(\Psi)|+1}, \ldots, z_{n-M-K}$ such that the $r_1 \times r_1$ matrix, where the rth row consists of the coefficients of $\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})$ of these r_1 variables, is invertible. Let us denote these variables by $\widetilde{z}_1, \ldots, \widetilde{z}_{r_1}$, and let $\widetilde{\mathbf{z}} = (\widetilde{z}_1, \ldots, \widetilde{z}_{r_1})$. We can then write

$$\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) = g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} + \widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})|_{\widetilde{\mathbf{z}}=\mathbf{0}},$$

where $g'_{r,1}, \ldots, g'_{r,r_1} \in \mathbb{Z}$. Let $\mathcal{R}^{(1)}_{+}(\Psi)$ be a maximal linearly independent (over $\mathbb{Q}$) subset of

$$\mathcal{R}^{(1)}(\Psi) \cup \{\widetilde{F}_{1,1}(\mathbf{0}, \mathbf{0}, \mathbf{z})|_{\widetilde{\mathbf{z}}=\mathbf{0}}, \ldots, \widetilde{F}_{1,r_1}(\mathbf{0}, \mathbf{0}, \mathbf{z})|_{\widetilde{\mathbf{z}}=\mathbf{0}}\}.$$

The important thing to note is that by our construction, we have that the set of linear forms

$$\{g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} : 1 \leqslant r \leqslant r_1\} \cup \mathcal{R}^{(1)}_{+}(\Psi)$$

is linearly independent over $\mathbb{Q}$. We also have that $|\mathcal{R}^{(1)}_{+}(\Psi)| \leqslant |\mathcal{R}^{(1)}(\Psi)| + r_1$.

We replace $\mathcal{R}^{(1)}(\Phi)$ of $\mathcal{R}(\Phi)$ with $\mathcal{R}^{(1)}_{+}(\Phi)$ and refer to the resulting set of forms as $\mathcal{R}_{+}(\Phi)$. It follows easily from the construction that the linear forms of $\mathcal{R}^{(1)}_{+}(\Phi)$ are either only in the $\mathbf{y}$ variables, or only in the $\mathbf{z}$ variables. We denote

$$\mathcal{R}_{+}(\Phi) = (\mathcal{R}^{(d-1)}(\Phi), \ldots, \mathcal{R}^{(2)}(\Phi), \mathcal{R}^{(1)}_{+}(\Phi)),$$

and by abusing notation slightly let

$$\mathcal{R}^{(1)}_{+}(\Phi) = \{A_i^{(1)} : 1 \leqslant i \leqslant |\mathcal{R}^{(1)}_{+}(\Phi)|\} \quad \text{and}$$
$$\overline{\mathcal{R}}^{(1)}_{+}(\Phi) = \{A_i^{(1)}(\mathbf{y}) : 1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(1)}_{+}(\Phi)|\}.$$

We then define $\overline{\mathcal{R}}_+(\Phi)$, R_2', and D_2' for $\mathcal{R}_+(\Phi)$ in an analogous manner as $\overline{\mathcal{R}}(\Phi)$, R_2, and D_2 for $\mathcal{R}(\Phi)$, respectively. Similarly, we replace $\mathcal{R}^{(1)}(\Psi)$ of $\mathcal{R}(\Psi)$ with $\mathcal{R}_+^{(1)}(\Psi)$ and refer to the resulting set of forms as $\mathcal{R}_+(\Psi)$. We denote

$$\mathcal{R}_+(\Psi) = (\mathcal{R}^{(d-1)}(\Psi), \ldots, \mathcal{R}^{(2)}(\Psi), \mathcal{R}_+^{(1)}(\Psi)),$$

and by abusing notation slightly let

$$\mathcal{R}_+^{(1)}(\Psi) = \{V_i^{(1)}(\mathbf{z}) : 1 \leqslant i \leqslant |\mathcal{R}_+^{(1)}(\Psi)|\}.$$

We also define R_1' and D_1' for $\mathcal{R}_+(\Psi)$ in an analogous manner as R_1 and D_1 for $\mathcal{R}(\Psi)$, respectively. It then follows that we have $R_2' \leqslant R_2 + r_1$ and $R_1' \leqslant R_1 + r_1$.

For each $\mathbf{H} \in \mathbb{Z}^{R_1'}$, we define the following set

$$Z(\mathbf{H}) = \{\mathbf{z} \in [0, X]^{n-M-K} \cap \mathbb{Z}^{n-M-K} : \mathcal{R}_+(\Psi)(\mathbf{z}) = \mathbf{H}\}.$$

By $\mathcal{R}_+(\Psi)(\mathbf{z}) = \mathbf{H}$, we mean that $V_i^{(s)}(\mathbf{z}) = H_{s,i}$, where $H_{s,i}$ is the corresponding term of $\mathbf{H}$, for every $V_i^{(s)} \in \mathcal{R}_+(\Psi)$. Other instances of notation of this type should be interpreted in a similar manner. By Proposition 2.7, we know that each of the polynomials $F_{\ell,r}(\mathbf{0}, \mathbf{0}, (\mathbf{0}, \mathbf{b}))$ $(2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell)$ and $G_{\ell,r}^{(k)}(\mathbf{z})$ in (5.13) can be expressed as a rational polynomial in the forms of $\mathcal{R}_+(\Psi)$. Let us denote

$$F_{\ell,r}(\mathbf{0}, \mathbf{0}, (\mathbf{0}, \mathbf{b})) = c_{\ell,r}^{\sharp}(\mathcal{R}_+(\Psi)),$$

and

$$G_{\ell,r}^{(k)}(\mathbf{z}) = c_{\ell,r:k}^{\flat}(\mathcal{R}_+(\Psi)),$$

where $c_{\ell,r}^{\sharp}$ and $c_{\ell,r:k}^{\flat}$ are rational polynomials in R_1' variables. Therefore, for any $\mathbf{z}_0 = (\mathbf{a}_0, \mathbf{b}_0) \in Z(\mathbf{H})$ we have

$$F_{\ell,r}(\mathbf{0}, \mathbf{0}, (\mathbf{0}, \mathbf{b}_0)) = c_{\ell,r}^{\sharp}(\mathbf{H}) \quad \text{and} \quad G_{\ell,r}^{(k)}(\mathbf{z}_0) = c_{\ell,r:k}^{\flat}(\mathbf{H}).$$

We also know that $\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})|_{\widetilde{\mathbf{z}}=\mathbf{0}}$ is constant on $Z(\mathbf{H})$, and we denote this constant value by $c_{1,r}^{\sharp}(\mathbf{H})$.

Similarly, we know that each of the polynomials $\Psi_{\ell,r:t_1,\ldots,t_j}^{(k)}(\mathbf{z})$ in (5.13) can be expressed as a rational polynomial in the forms of $\mathcal{R}_+(\Psi)$. Let us denote

$$\Psi_{\ell,r:t_1,\ldots,t_j}^{(k)}(\mathbf{z}) = \hat{c}_{\ell,r:t_1,\ldots,t_j}^{(k)}(\mathcal{R}_+(\Psi)), \tag{5.22}$$

where $\hat{c}_{\ell,r:t_1,\ldots,t_j}^{(k)}$ is a rational polynomial in R_1' variables. Therefore, for any $\mathbf{z}_0 \in Z(\mathbf{H})$ we have

$$\Psi_{\ell,r:t_1,\ldots,t_j}^{(k)}(\mathbf{z}_0) = \hat{c}_{\ell,r:t_1,\ldots,t_j}^{(k)}(\mathbf{H}).$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Since each of the forms $\widetilde{\Psi}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{z})$ in (5.15) can be expressed as a rational polynomial in the forms of $\mathcal{R}_+(\Psi)$, let us denote

$$\widetilde{\Psi}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{z}) = \widetilde{c}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathcal{R}_+(\Psi)),$$

where each $\widetilde{c}^{(s-k)}_{s,i:i_1,\ldots,i_k}$ is a rational polynomial in R_1' variables. Therefore, for each $A_i^{(s)}$ with $1 < s \leqslant d-1$ and $1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Phi)|$, we can write

$$A_i^{(s)}(\mathbf{y}, \mathbf{z}) = \sum_{k=0}^{s} \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant M} \widetilde{c}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathcal{R}_+(\Psi)) y_{i_1} \cdots y_{i_k}. \tag{5.23}$$

Consequently, we can define the following polynomial for each $1 < s \leqslant d-1$ and $1 \leqslant i \leqslant |\mathcal{R}^{(s)}(\Phi)|$,

$$A_i^{(s)}(\mathbf{y}, Z(\mathbf{H})) = \sum_{k=0}^{s} \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant M} \widetilde{c}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{H}) y_{i_1} \cdots y_{i_k}, \tag{5.24}$$

so that given any $\mathbf{z}_0 \in Z(H)$ we have

$$A_i^{(s)}(\mathbf{y}, \mathbf{z}_0) = A_i^{(s)}(\mathbf{y}, Z(\mathbf{H})).$$

We then define

$$\overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, Z(\mathbf{H})) = \{A_i^{(s)}(\mathbf{y}, Z(\mathbf{H})) : 2 \leqslant s \leqslant d-1, 1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(s)}(\Phi)|\} \cup \overline{\mathcal{R}}_+^{(1)}(\Phi),$$

which is a system consisting of R_2' polynomials (with possible repetitions).

For each $\mathbf{G} \in \mathbb{Z}^{R_2'}$, we let

$$Y(\mathbf{G}; \mathbf{H}) = \{\mathbf{y} \in [0, X]^M \cap \mathbb{Z}^M : \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, Z(\mathbf{H})) = \mathbf{G}\}.$$

It follows from the definition of $\overline{\mathcal{R}}_+^{(1)}(\Phi)$ that for each $1 \leqslant r \leqslant r_1$ the polynomial $\widetilde{F}_{1,r}(\mathbf{0}, \mathbf{y}, \mathbf{0})|_{\widetilde{\mathbf{y}}=\mathbf{0}}$ is constant on $Y(\mathbf{G}; \mathbf{H})$, and we denote this constant value by $c'_{1,r}(\mathbf{G}, \mathbf{H})$.

Recall Φ is the collection of all $\Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y}, \mathbf{z})$ in (5.13), and that each $\Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y}, \mathbf{z})$ can be expressed as a rational polynomial in the forms of $\mathcal{R}_+(\Phi)$. It follows from our definition that the forms of $\mathcal{R}_+(\Phi)$ which depend only on the $\mathbf{z}$ variables are constant on $Z(\mathbf{H})$, and the remaining forms, which are precisely the forms of $\overline{\mathcal{R}}_+(\Phi)$, are constant on $Y(\mathbf{G}; \mathbf{H}) \times Z(\mathbf{H})$. Thus each $\Phi^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{y}, \mathbf{z})$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

is constant on $Y(\mathbf{G};\mathbf{H})\times Z(\mathbf{H})$, and we denote this constant value by $c^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{G},\mathbf{H})$. Let $2\leqslant \ell < d$ and $1\leqslant r\leqslant r_\ell$. Therefore, for any choice of $\mathbf{z}=(\mathbf{a},\mathbf{b})\in Z(\mathbf{H})$ and $\mathbf{y}\in Y(\mathbf{G};\mathbf{H})$, the polynomial $f_{\ell,r}(\mathbf{x})$ takes the following shape

$$\begin{aligned} f_{\ell,r}(\mathbf{w},\mathbf{y},\mathbf{z}) &= f_{\ell,r}(\mathbf{w},\mathbf{0},\mathbf{0}) + \sum_{j=1}^{\ell-1}\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant K}\left(\sum_{k=1}^{\ell-j} c^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{G},\mathbf{H})\right) w_{i_1}\cdots w_{i_j} \\ &\quad + \sum_{j=1}^{\ell-1}\sum_{1\leqslant t_1\leqslant\cdots\leqslant t_j\leqslant M}\left(\sum_{k=0}^{\ell-j} \hat{c}^{(k)}_{\ell,r:t_1,\ldots,t_j}(\mathbf{H})\right) y_{t_1}\cdots y_{t_j} + F_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{0}) \\ &\quad + P_{\ell,r}(\mathbf{a},\mathbf{b}) + c^{\sharp}_{\ell,r}(\mathbf{H}) + \sum_{k=1}^{\ell-1} c^{\flat}_{\ell,r:k}(\mathbf{H}). \end{aligned} \tag{5.25}$$

When $\ell=d$, we replace the term $P_{\ell,r}(\mathbf{a},\mathbf{b})+c^{\sharp}_{\ell,r}(\mathbf{H})$ in (5.25) with $F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z})$. Similarly, when $\ell=1$ and $1\leqslant r\leqslant r_1$, for any choice of $\mathbf{z}=(\mathbf{a},\mathbf{b})\in Z(\mathbf{H})$ and $\mathbf{y}\in Y(\mathbf{G};\mathbf{H})$, the polynomial $f_{1,r}(\mathbf{x})$ takes the following shape

$$\begin{aligned} f_{1,r}(\mathbf{x}) = c_{1,r}w_r + \widetilde{f}_{1,r}(\mathbf{w},\mathbf{0},\mathbf{0}) + (g_{r,1}\widetilde{y}_1+\cdots+g_{r,r_1}\widetilde{y}_{r_1}) \\ + c'_{1,r}(\mathbf{G},\mathbf{H}) + (g'_{r,1}\widetilde{z}_1+\cdots+g'_{r,r_1}\widetilde{z}_{r_1}) + c^{\sharp}_{1,r}(\mathbf{H}) \end{aligned}$$

where $\widetilde{f}_{1,r}$ is defined in (4.2).

For each $2\leqslant \ell\leqslant d$, $1\leqslant r\leqslant r_\ell$, we label

$$\begin{aligned} \mathfrak{C}_{\ell,r}(\mathbf{w},\mathbf{G},\mathbf{H}) = f_{\ell,r}(\mathbf{w},\mathbf{0},\mathbf{0}) \\ + \sum_{j=1}^{\ell-1}\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant K}\left(\sum_{k=1}^{\ell-j} c^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{G},\mathbf{H})\right) w_{i_1}\cdots w_{i_j}, \end{aligned}$$

and

$$\mathfrak{U}_{\ell,r}(\mathbf{y},\mathbf{H}) = \sum_{j=1}^{\ell-1}\sum_{1\leqslant t_1\leqslant\cdots\leqslant t_j\leqslant M}\left(\sum_{k=0}^{\ell-j} \hat{c}^{(k)}_{\ell,r:t_1,\ldots,t_j}(\mathbf{H})\right) y_{t_1}\cdots y_{t_j} + F_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{0}). \tag{5.26}$$

We let

$$\mathfrak{X}_{\ell,r}(\mathbf{a},\mathbf{b},\mathbf{H}) = P_{\ell,r}(\mathbf{a},\mathbf{b}) + c^{\sharp}_{\ell,r}(\mathbf{H}) + \sum_{k=1}^{\ell-1} c^{\flat}_{\ell,r:k}(\mathbf{H}) \quad (2\leqslant \ell < d,\ 1\leqslant r\leqslant r_\ell),$$

and also

$$\mathfrak{X}_{d,r}(\mathbf{a},\mathbf{b},\mathbf{H}) = F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) + \sum_{k=1}^{d-1} c^{\flat}_{d,r:k}(\mathbf{H}) \quad (1\leqslant r\leqslant r_d).$$

Then for $\mathbf{z} = (\mathbf{a}, \mathbf{b}) \in Z(\mathbf{H})$ and $\mathbf{y} \in Y(\mathbf{G}; \mathbf{H})$, we have

$$f_{\ell,r}(\mathbf{w}, \mathbf{y}, \mathbf{z}) = \mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}) + \mathfrak{U}_{\ell,r}(\mathbf{y}, \mathbf{H}) + \mathfrak{X}_{\ell,r}(\mathbf{a}, \mathbf{b}, \mathbf{H})$$
$$(2 \leqslant \ell \leqslant d,\ 1 \leqslant r \leqslant r_\ell).$$

We define the following three exponential sums,

$$S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) = \sum_{\mathbf{w} \in [0,X]^K} \Lambda(\mathbf{w}) e\Bigg(\sum_{1 \leqslant r \leqslant r_1} \alpha_{1,r}(c_{1,r} w_r + \widetilde{f}_{1,r}(\mathbf{w}, \mathbf{0}, \mathbf{0})) + \sum_{2 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot \mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}) \Bigg),$$

$$S_1(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) = \sum_{\mathbf{y} \in Y(\mathbf{G};\mathbf{H})} \Lambda(\mathbf{y}) e\Bigg(\sum_{1 \leqslant r \leqslant r_1} \alpha_{1,r}(g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} + c'_{1,r}(\mathbf{G}, \mathbf{H})) + \sum_{2 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot \mathfrak{U}_{\ell,r}(\mathbf{y}, \mathbf{H}) \Bigg),$$

and

$$S_2(\boldsymbol{\alpha}, \mathbf{H}) = \sum_{\mathbf{z} = (\mathbf{a},\mathbf{b}) \in Z(\mathbf{H})} \Lambda(\mathbf{z}) e\Bigg(\sum_{1 \leqslant r \leqslant r_1} \alpha_{1,r}(g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} + c^{\sharp}_{1,r}(\mathbf{H})) + \sum_{2 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot \mathfrak{X}_{\ell,r}(\mathbf{a}, \mathbf{b}, \mathbf{H}) \Bigg).$$

Let

$$\mathcal{L}_1(X) = \{\mathbf{H} \in \mathbb{Z}^{R'_1} : Z(\mathbf{H}) \neq \emptyset\},$$

and for each $\mathbf{H} \in \mathcal{L}_1(X)$, let

$$\mathcal{L}_2(X; \mathbf{H}) = \{\mathbf{G} \in \mathbb{Z}^{R'_2} : Y(\mathbf{G}, \mathbf{H}) \neq \emptyset\}.$$

We have the following bounds on the cardinalities of these sets,

$$|\mathcal{L}_1(X)| \ll X^{D'_1} \quad \text{and} \quad |\mathcal{L}_2(X; \mathbf{H})| \ll X^{D'_2}.$$

It is not difficult to deduce the first inequality. The implicit constant in the second inequality is independent of $\mathbf{H}$, and to see this we note that given $A_i^{(s)}$ with $1 < s \leqslant d-1$ and $1 \leqslant i \leqslant |\overline{\mathcal{R}}^{(s)}(\Phi)|$, we have

$$|A_i^{(s)}(\mathbf{y}, \mathbf{z})| = \left| \sum_{k=0}^{s} \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant M} \widetilde{\Psi}^{(s-k)}_{s,i:i_1,\ldots,i_k}(\mathbf{z}) y_{i_1} \cdots y_{i_k} \right| \ll X^s$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

for any $(\mathbf{y}, \mathbf{z}) \in [0, X]^{n-K} \cap \mathbb{Z}^{n-K}$, and similarly for the linear forms of $\overline{\mathcal{R}}_{+}^{(1)}(\Phi)$. Therefore, we obtain by applying the Cauchy–Schwarz inequality

$$
\begin{aligned}
&\left| \int_{\mathfrak{m}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha} \right|^2 \\
&\leqslant \Bigg| \sum_{\mathbf{H} \in \mathcal{L}_1(X)} \sum_{\mathbf{G} \in \mathcal{L}_2(X; \mathbf{H})} \int_{\mathfrak{m}(C)} \sum_{\substack{\mathbf{w} \in [0,X]^K \\ \mathbf{y} \in Y(\mathbf{G}; \mathbf{H}) \\ \mathbf{z} = (\mathbf{a}, \mathbf{b}) \in Z(\mathbf{H})}} \Lambda(\mathbf{w}) \Lambda(\mathbf{y}) \Lambda(\mathbf{z}) \\
&\quad \cdot e \Bigg(\sum_{1 \leqslant r \leqslant r_1} \alpha_{1,r} (c_{1,r} w_r + \widetilde{f}_{1,r}(\mathbf{w}, \mathbf{0}, \mathbf{0}) + (g_{r,1} \widetilde{y}_1 + \cdots + g_{r,r_1} \widetilde{y}_{r_1}) \\
&\quad + c'_{1,r}(\mathbf{G}, \mathbf{H}) \\
&\quad + (g'_{r,1} \widetilde{z}_1 + \cdots + g'_{r,r_1} \widetilde{z}_{r_1}) + c^{\sharp}_{1,r}(\mathbf{H})) \Bigg) \\
&\quad \cdot e \left(\sum_{2 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot (\mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}) + \mathfrak{U}_{\ell,r}(\mathbf{y}, \mathbf{H}) + \mathfrak{X}_{\ell,r}(\mathbf{a}, \mathbf{b}, \mathbf{H})) \right) \mathbf{d}\boldsymbol{\alpha} \Bigg|^2 \\
&\ll X^{D'_1 + D'_2} \sum_{\mathbf{H} \in \mathcal{L}_1(X)} \sum_{\mathbf{G} \in \mathcal{L}_2(X; \mathbf{H})} \left| \int_{\mathfrak{m}(C)} S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) S_1(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) S_2(\boldsymbol{\alpha}, \mathbf{H})\, \mathbf{d}\boldsymbol{\alpha} \right|^2 \\
&\ll X^{D'_1 + D'_2} \left(\sup_{\substack{\mathbf{H} \in \mathcal{L}_1(X) \\ \mathbf{G} \in \mathcal{L}_2(X; \mathbf{H})}} \sup_{\boldsymbol{\alpha} \in \mathfrak{m}(C)} |S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H})|^2 \right) \\
&\quad \cdot \sum_{\mathbf{H} \in \mathcal{L}_1(X)} \sum_{\mathbf{G} \in \mathcal{L}_2(X; \mathbf{H})} \| S_1(\cdot, \mathbf{G}, \mathbf{H}) \|_2^2 \| S_2(\cdot, \mathbf{H}) \|_2^2, \qquad (5.27)
\end{aligned}
$$

where $\| \cdot \|_2$ denotes the L^2-norm on $[0, 1]^R$.

By the orthogonality relation, it follows that

$$
\| S_1(\cdot, \mathbf{G}, \mathbf{H}) \|_2^2 \| S_2(\cdot, \mathbf{H}) \|_2^2 \leqslant (\log X)^{2n-2K} \mathcal{N}_1(\mathbf{G}; \mathbf{H}) \mathcal{N}_2(\mathbf{H}),
$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

where

$$\begin{aligned}\mathcal{N}_1(\mathbf{G};\mathbf{H}) = |\{(\mathbf{y},\mathbf{y}') \in Y(\mathbf{G};\mathbf{H}) \times Y(\mathbf{G};\mathbf{H}) : &\ \mathfrak{U}_{\ell,r}(\mathbf{y},\mathbf{H}) \\ &= \mathfrak{U}_{\ell,r}(\mathbf{y}',\mathbf{H}) \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \\ &\quad g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} = g_{r,1}\widetilde{y}'_1 + \cdots + g_{r,r_1}\widetilde{y}'_{r_1}\ (1 \leqslant r \leqslant r_1)\}|,\end{aligned}$$

and with $\mathbf{z} = (\mathbf{a},\mathbf{b})$ and $\mathbf{z}' = (\mathbf{a}',\mathbf{b}')$,

$$\begin{aligned}\mathcal{N}_2(\mathbf{H}) = |\{(\mathbf{z},\mathbf{z}') \in Z(\mathbf{H}) \times Z(\mathbf{H}) : &\ \mathfrak{X}_{\ell,r}(\mathbf{a},\mathbf{b},\mathbf{H}) \\ &= \mathfrak{X}_{\ell,r}(\mathbf{a}',\mathbf{b}',\mathbf{H}) \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \\ &\quad g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} = g'_{r,1}\widetilde{z}'_1 + \cdots + g'_{r,r_1}\widetilde{z}'_{r_1}\ (1 \leqslant r \leqslant r_1)\}|.\end{aligned}$$

Here $\widetilde{y}'_i$ $(1 \leqslant i \leqslant r_1)$ are r_1 of the $\mathbf{y}'$ variables in the exact same way $\widetilde{y}_i$ $(1 \leqslant i \leqslant r_1)$ are r_1 of the $\mathbf{y}$ variables. Similarly, $\widetilde{z}'_i$ $(1 \leqslant i \leqslant r_1)$ are r_1 of the $\mathbf{z}'$ variables in the exact same way $\widetilde{z}_i$ $(1 \leqslant i \leqslant r_1)$ are r_1 of the $\mathbf{z}$ variables. Other instances of notation of this type should be interpreted in a similar manner.

With these notations, we may further bound (5.27) as follows

$$\begin{aligned}&\left| \int_{\mathfrak{m}(C)} T(\mathbf{f};\boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha} \right|^2 \\ &\quad \ll (\log X)^{2n-2K} X^{D'_1+D'_2} \left(\sup_{\substack{\mathbf{H}\in\mathcal{L}_1(X) \\ \mathbf{G}\in\mathcal{L}_2(X;\mathbf{H})}} \sup_{\boldsymbol{\alpha}\in\mathfrak{m}(C)} |S_0(\boldsymbol{\alpha},\mathbf{G},\mathbf{H})|^2 \right) \mathcal{W}, \quad (5.28)\end{aligned}$$

where

$$\mathcal{W} = \sum_{\mathbf{H}\in\mathcal{L}_1(X)} \sum_{\mathbf{G}\in\mathcal{L}_2(X;\mathbf{H})} \mathcal{N}_1(\mathbf{G};\mathbf{H})\mathcal{N}_2(\mathbf{H}).$$

We can express $\mathcal{W}$ as the number of solutions $\mathbf{y},\mathbf{y}' \in [0,X]^M \cap \mathbb{Z}^M$ and $\mathbf{z} = (\mathbf{a},\mathbf{b}), \mathbf{z}' = (\mathbf{a}',\mathbf{b}') \in [0,X]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ of the system

$$\begin{aligned}\mathcal{R}_+(\Psi)(\mathbf{z}) &= \mathcal{R}_+(\Psi)(\mathbf{z}') = \mathbf{H} \qquad (5.29)\\ \overline{\mathcal{R}}_+(\Phi)(\mathbf{y},Z(\mathbf{H})) &= \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}',Z(\mathbf{H})) = \mathbf{G} \\ \mathfrak{U}_{\ell,r}(\mathbf{y},\mathbf{H}) &= \mathfrak{U}_{\ell,r}(\mathbf{y}',\mathbf{H}) \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell) \\ g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} &= g_{r,1}\widetilde{y}'_1 + \cdots + g_{r,r_1}\widetilde{y}'_{r_1} \quad (1 \leqslant r \leqslant r_1) \\ \mathfrak{X}_{\ell,r}(\mathbf{a},\mathbf{b},\mathbf{H}) &= \mathfrak{X}_{\ell,r}(\mathbf{a}',\mathbf{b}',\mathbf{H}) \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell) \\ g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} &= g'_{r,1}\widetilde{z}'_1 + \cdots + g'_{r,r_1}\widetilde{z}'_{r_1} \quad (1 \leqslant r \leqslant r_1)\end{aligned}$$

for any $\mathbf{H} \in \mathcal{L}_1(X)$ and $\mathbf{G} \in \mathcal{L}_2(X;\mathbf{H})$. By $\mathcal{R}_+(\Psi)(\mathbf{z}) = \mathcal{R}_+(\Psi)(\mathbf{z}') = \mathbf{H}$, we mean that $V_i^{(s)}(\mathbf{z}) = V_i^{(s)}(\mathbf{z}') = H_{s,i}$, where $H_{s,i}$ is the corresponding term of $\mathbf{H}$, for

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

every $V_i^{(s)} \in \mathcal{R}_+(\Psi)$. The second set of equations in (5.29) should be interpreted in a similar manner.

We know that the system of polynomials $\overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, Z(\mathbf{H}))$ is identical to $\overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, \mathbf{z}_0)$ for any choice of $\mathbf{z}_0 \in Z(\mathbf{H})$. Similarly, it follows from (5.14), (5.22), and (5.26) that the polynomial $\mathfrak{U}_{\ell,r}(\mathbf{y}, \mathbf{H})$ is identical to $f_{\ell,r}(\mathbf{0}, \mathbf{y}, \mathbf{z}_0) - f_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}_0)$ for any choice of $\mathbf{z}_0 \in Z(\mathbf{H})$. Furthermore, for $2 \leqslant \ell < d$ we know that each term of $\mathfrak{X}_{\ell,r}(\mathbf{a}, \mathbf{b}, \mathbf{H})$ except for $P_{\ell,r}(\mathbf{a}, \mathbf{b})$ is constant on $\mathbf{z} = (\mathbf{a}, \mathbf{b}) \in Z(\mathbf{H})$. Therefore, since $\mathcal{R}_+(\Psi)(\mathbf{z}) = \mathbf{H}$ implies $\mathbf{z} \in Z(\mathbf{H})$, we can rearrange the system (5.29) and deduce that $\mathcal{W}$ is the number of solutions $\mathbf{y}, \mathbf{y}' \in [0, X]^M \cap \mathbb{Z}^M$ and $\mathbf{z}, \mathbf{z}' \in [0, X]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ of the following system

$$
\begin{aligned}
\mathcal{R}_+(\Psi)(\mathbf{z}) &= \mathcal{R}_+(\Psi)(\mathbf{z}') \\
\overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, \mathbf{z}) &= \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}', \mathbf{z}) \\
f_{\ell,r}(\mathbf{0}, \mathbf{y}, \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) &= f_{\ell,r}(\mathbf{0}, \mathbf{y}', \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) \\
&\quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell) \\
g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} &= g_{r,1}\widetilde{y}'_1 + \cdots + g_{r,r_1}\widetilde{y}'_{r_1} \quad (1 \leqslant r \leqslant r_1) \\
F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) &= F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}') \quad (1 \leqslant r \leqslant r_d) \\
P_{\ell,r}(\mathbf{a}, \mathbf{b}) &= P_{\ell,r}(\mathbf{a}', \mathbf{b}') \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell) \\
g'_{r,1}\widetilde{z}_1 + \cdots + g'_{r,r_1}\widetilde{z}_{r_1} &= g'_{r,1}\widetilde{z}'_1 + \cdots + g'_{r,r_1}\widetilde{z}'_{r_1} \quad (1 \leqslant r \leqslant r_1).
\end{aligned}
\tag{5.30}
$$

Our result then follows from the following two claims.

CLAIM 1. *Given any $c > 0$, for sufficiently large $C > 0$ we have*

$$
\sup_{\substack{\mathbf{H} \in \mathcal{L}_1(X) \\ \mathbf{G} \in \mathcal{L}_2(X;\mathbf{H})}} \sup_{\boldsymbol{\alpha} \in \mathfrak{m}(C)} |S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H})| \ll \frac{X^K}{(\log X)^c}.
$$

CLAIM 2. *We have the following bound on $\mathcal{W}$,*

$$
\mathcal{W} \ll X^{2n-2K-2\sum_{\ell=1}^d \ell r_\ell - D'_1 - D'_2}.
$$

Let $c > 0$. By substituting the bounds from the two claims above into (5.28), we obtain that for sufficiently large $C > 0$ we have

$$
\int_{\mathfrak{m}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, \mathbf{d}\boldsymbol{\alpha} \ll \frac{X^{n-\sum_{\ell=1}^d \ell r_\ell}}{(\log X)^c},
$$

which is the bound in the statement of this proposition. Therefore, we are only left to prove Claims 1 and 2 to establish our proposition. We now present the proof of

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Claim 2. Claim 1 is obtained via Weyl differencing, which is a technique based on the Cauchy–Schwarz inequality, and we prove it in Section 5.3 after the proof of Claim 2.

From (5.30), we can write

$$\mathcal{W} = \sum_{\mathbf{z}=(\mathbf{a},\mathbf{b})\in[0,X]^{n-M-K}} \mathcal{W}_1'(\mathbf{z}) \cdot \mathcal{W}_2'(\mathbf{z}),$$

where $\mathcal{W}_1'(\mathbf{z})$ is the number of solutions $\mathbf{y}, \mathbf{y}' \in [0, X]^M \cap \mathbb{Z}^M$ to the system

$$\begin{aligned} \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}, \mathbf{z}) &= \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}', \mathbf{z}),\\ f_{\ell,r}(\mathbf{0}, \mathbf{y}, \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) &= f_{\ell,r}(\mathbf{0}, \mathbf{y}', \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{0}, \mathbf{z})\\ &\quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)\\ g_{r,1}\widetilde{y}_1 + \cdots + g_{r,r_1}\widetilde{y}_{r_1} &= g_{r,1}\widetilde{y}_1' + \cdots + g_{r,r_1}\widetilde{y}_{r_1}' \quad (1 \leqslant r \leqslant r_1), \end{aligned}$$

and $\mathcal{W}_2'(\mathbf{z})$ is the number of solutions $\mathbf{z}' = (\mathbf{a}', \mathbf{b}') \in [0, X]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ to the system

$$\begin{aligned} \mathcal{R}_+(\Psi)(\mathbf{z}) &= \mathcal{R}_+(\Psi)(\mathbf{z}')\\ F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) &= F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}') \quad (1 \leqslant r \leqslant r_d)\\ P_{\ell,r}(\mathbf{a}, \mathbf{b}) &= P_{\ell,r}(\mathbf{a}', \mathbf{b}') \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell)\\ g_{r,1}'\widetilde{z}_1 + \cdots + g_{r,r_1}'\widetilde{z}_{r_1} &= g_{r,1}'\widetilde{z}_1' + \cdots + g_{r,r_1}'\widetilde{z}_{r_1}' \quad (1 \leqslant r \leqslant r_1). \end{aligned}$$

Define $\mathcal{W}_i := \sum_{\mathbf{z}} \mathcal{W}_i'(\mathbf{z})^2$ $(i = 1, 2)$ so that we have $\mathcal{W}^2 \leqslant \mathcal{W}_1\mathcal{W}_2$ by the Cauchy–Schwarz inequality. We estimate $\mathcal{W}_1$ and $\mathcal{W}_2$ in Sections 5.1 and 5.2, respectively. In Section 5.1, we prove $\mathcal{W}_1 \ll X^{n+3M-K-2\sum_{\ell=1}^d \ell r_\ell - 2D_2'}$, and in Section 5.2 we prove $\mathcal{W}_2 \ll X^{3(n-M-K)-2\sum_{\ell=1}^d \ell r_\ell - 2D_1'}$. Combining these bounds for $\mathcal{W}_1$ and $\mathcal{W}_2$, we obtain

$$\mathcal{W} \leqslant \mathcal{W}_1^{1/2}\mathcal{W}_2^{1/2} \ll X^{2n-2K-2\sum_{\ell=1}^d \ell r_\ell - D_1' - D_2'},$$

which proves Claim 2.

5.1. Estimate for $\mathcal{W}_1$.

We first estimate $\mathcal{W}_1$, which we can deduce to be the number of solutions $\mathbf{y}, \mathbf{y}', \mathbf{v}, \mathbf{v}' \in [0, X]^M \cap \mathbb{Z}^M$ and $\mathbf{z} \in [0, X]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ satisfying the equations

$$\begin{aligned} f_{\ell,r}(\mathbf{0}, \mathbf{y}, \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{y}', \mathbf{z}) &= 0 \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell) \qquad (5.31)\\ f_{\ell,r}(\mathbf{0}, \mathbf{v}, \mathbf{z}) - f_{\ell,r}(\mathbf{0}, \mathbf{v}', \mathbf{z}) &= 0 \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)\\ \sum_{i=1}^{r_1} g_{r,i}\widetilde{y}_i - \sum_{i=1}^{r_1} g_{r,i}\widetilde{y}_i' &= 0 \quad (1 \leqslant r \leqslant r_1) \end{aligned}$$

$$\sum_{i=1}^{r_1} g_{r,i}\widetilde{v}_i - \sum_{i=1}^{r_1} g_{r,i}\widetilde{v}'_i = 0 \quad (1 \leqslant r \leqslant r_1)$$
$$\overline{\mathcal{R}}_+(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}',\mathbf{z}) = \mathbf{0}$$
$$\overline{\mathcal{R}}_+(\Phi)(\mathbf{v},\mathbf{z}) - \overline{\mathcal{R}}_+(\Phi)(\mathbf{v}',\mathbf{z}) = \mathbf{0}.$$

Let $\overline{\mathcal{R}}_+^{(i)}(\Phi)$ denote the degree i forms of $\overline{\mathcal{R}}_+(\Phi)$ $(1 \leqslant i < d)$. By $\overline{\mathcal{R}}_+(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+(\Phi)(\mathbf{y}',\mathbf{z})$, we mean the system of forms where its degree i forms are

$$\overline{\mathcal{R}}_+^{(i)}(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+^{(i)}(\Phi)(\mathbf{y}',\mathbf{z}) = \{A_j^{(i)}(\mathbf{y},\mathbf{z}) - A_j^{(i)}(\mathbf{y}',\mathbf{z}) : 1 \leqslant j \leqslant |\overline{\mathcal{R}}_+^{(i)}(\Phi)|\},$$

for each $1 \leqslant i \leqslant d-1$. Recall we have $\overline{\mathcal{R}}_+^{(i)}(\Phi) = \overline{\mathcal{R}}^{(i)}(\Phi)$ for $2 \leqslant i \leqslant d-1$. We also define

$$\overline{\mathcal{R}}_+(\Phi)(\mathbf{v},\mathbf{z}) - \overline{\mathcal{R}}_+(\Phi)(\mathbf{v}',\mathbf{z})$$

in a similar manner.

We consider the h-invariant of the system of polynomials on the left hand side of (5.31), and show that it is a regular system. Recall we defined $Q_{\ell,r}(\mathbf{y},\mathbf{z})$ in (5.7) and also remarked that it is the homogeneous degree ℓ portion of the polynomial $f_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - f_{\ell,r}(\mathbf{0},\mathbf{0},\mathbf{z})$. Therefore, the homogeneous degree ℓ portion of the polynomial $f_{\ell,r}(\mathbf{0},\mathbf{y},\mathbf{z}) - f_{\ell,r}(\mathbf{0},\mathbf{y}',\mathbf{z})$ is precisely $Q_{\ell,r}(\mathbf{y},\mathbf{z}) - Q_{\ell,r}(\mathbf{y}',\mathbf{z})$. Thus the homogeneous degree d portions of the degree d polynomials of the system (5.31) are $Q_{d,r}(\mathbf{y},\mathbf{z}) - Q_{d,r}(\mathbf{y}',\mathbf{z})$, $Q_{d,r}(\mathbf{v},\mathbf{z}) - Q_{d,r}(\mathbf{v}',\mathbf{z})$ $(1 \leqslant r \leqslant r_d)$. We let h_d be the h-invariant of these degree d forms. Suppose for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{Q}^{r_d}$, not both $\mathbf{0}$, we have

$$\sum_{r=1}^{r_d} \lambda_r \cdot (Q_{d,r}(\mathbf{y},\mathbf{z}) - Q_{d,r}(\mathbf{y}',\mathbf{z})) + \mu_r \cdot (Q_{d,r}(\mathbf{v},\mathbf{z}) - Q_{d,r}(\mathbf{v}',\mathbf{z}))$$
$$= \sum_{j=1}^{h_d} \widetilde{U}_j \cdot \widetilde{V}_j, \tag{5.32}$$

where $\widetilde{U}_j = \widetilde{U}_j(\mathbf{y},\mathbf{y}',\mathbf{v},\mathbf{v}',\mathbf{z})$ and $\widetilde{V}_j = \widetilde{V}_j(\mathbf{y},\mathbf{y}',\mathbf{v},\mathbf{v}',\mathbf{z})$ are rational forms of positive degree $(1 \leqslant j \leqslant h_d)$. Without loss of generality, suppose $\boldsymbol{\lambda} \neq \mathbf{0}$. If we set $\mathbf{v} = \mathbf{v}' = \mathbf{y}' = \mathbf{0}$, then the above equation (5.32) becomes

$$\sum_{r=1}^{r_d} \lambda_r \cdot Q_{d,r}(\mathbf{y},\mathbf{z}) = \sum_{j=1}^{h_d} \widetilde{U}_j(\mathbf{y},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{z}) \cdot \widetilde{V}_j(\mathbf{y},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{z}).$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Therefore, we obtain from (5.9) and (5.17) that

$$\begin{aligned} h_d &\geqslant h_d(\{Q_{d,r}(\mathbf{y},\mathbf{z}) : 1 \leqslant r \leqslant r_d\}) \\ &\geqslant \rho_{d,d}(2R + 2R_2) + 2R_2 + 4r_1 \\ &\geqslant \rho_{d,d}(2R + 2R_2' - 2|\overline{\mathcal{R}}_+^{(1)}(\Phi)| - 2r_1) + 2|\overline{\mathcal{R}}_+^{(1)}(\Phi)| + 2r_1. \end{aligned}$$

We now estimate the h-invariant of the degree ℓ polynomials of the system (5.31) for each $2 \leqslant \ell \leqslant d-1$. The homogeneous degree ℓ portion of the degree ℓ polynomials of the system (5.31) is precisely $Q_{\ell,r}(\mathbf{y},\mathbf{z}) - Q_{\ell,r}(\mathbf{y}',\mathbf{z})$, $Q_{\ell,r}(\mathbf{v},\mathbf{z}) - Q_{\ell,r}(\mathbf{v}',\mathbf{z})$ $(1 \leqslant r \leqslant r_\ell)$, and the forms of $\overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y}',\mathbf{z})$ and $\overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{v},\mathbf{z}) - \overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{v}',\mathbf{z})$. We let h_ℓ be the h-invariant of these degree ℓ forms. Suppose for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{Q}^{r_\ell}$ and $\boldsymbol{\gamma}, \boldsymbol{\gamma}' \in \mathbb{Q}^{|\overline{\mathcal{R}}^{(\ell)}(\Phi)|}$, not all zero vectors, we have

$$\begin{aligned} &\sum_{r=1}^{r_\ell} \lambda_r \cdot (Q_{\ell,r}(\mathbf{y},\mathbf{z}) - Q_{\ell,r}(\mathbf{y}',\mathbf{z})) + \mu_r \cdot (Q_{\ell,r}(\mathbf{v},\mathbf{z}) - Q_{\ell,r}(\mathbf{v}',\mathbf{z})) \\ &\qquad + \sum_{1 \leqslant j \leqslant |\overline{\mathcal{R}}^{(\ell)}(\Phi)|} \gamma_j (A_j^{(\ell)}(\mathbf{y},\mathbf{z}) - A_j^{(\ell)}(\mathbf{y}',\mathbf{z})) + \gamma_j'(A_j^{(\ell)}(\mathbf{v},\mathbf{z}) - A_j^{(\ell)}(\mathbf{v}',\mathbf{z})) \\ &\quad = \sum_{j=1}^{h_\ell} \widetilde{U}_j \cdot \widetilde{V}_j, \end{aligned} \tag{5.33}$$

where $\widetilde{U}_j = \widetilde{U}_j(\mathbf{y},\mathbf{y}',\mathbf{v},\mathbf{v}',\mathbf{z})$ and $\widetilde{V}_j = \widetilde{V}_j(\mathbf{y},\mathbf{y}',\mathbf{v},\mathbf{v}',\mathbf{z})$ are rational forms of positive degree $(1 \leqslant j \leqslant h_\ell)$. We must consider two cases, $\boldsymbol{\gamma} = \boldsymbol{\gamma}' = \mathbf{0}$ and at least one of $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$ not being a zero vector. If $\boldsymbol{\gamma} = \boldsymbol{\gamma}' = \mathbf{0}$, then at least one of $\boldsymbol{\lambda}$ or $\boldsymbol{\mu}$ is not a zero vector. Without loss of generality, suppose $\boldsymbol{\lambda} \neq \mathbf{0}$. Then by setting $\mathbf{v} = \mathbf{v}' = \mathbf{y}' = \mathbf{0}$, we have

$$\sum_{r=1}^{r_\ell} \lambda_r Q_{\ell,r}(\mathbf{y},\mathbf{z}) = \sum_{j=1}^{h_\ell} \widetilde{U}_j(\mathbf{y},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{z}) \cdot \widetilde{V}_j(\mathbf{y},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{z}).$$

Consequently, we obtain from (5.9) and (5.17) that

$$h_\ell \geqslant h_\ell(\{Q_{\ell,r}(\mathbf{y},\mathbf{z}) : 1 \leqslant r \leqslant r_\ell\}) \geqslant \rho_{d,\ell}(2R + 2R_2) + 2R_2 + 4r_1.$$

For the second case, suppose without loss of generality that $\boldsymbol{\gamma} \neq \mathbf{0}$. First we set $\mathbf{v} = \mathbf{v}' = \mathbf{0}$ and simplify the equation (5.33) to

$$\sum_{r=1}^{r_\ell} \lambda_r \cdot (Q_{\ell,r}(\mathbf{y},\mathbf{z}) - Q_{\ell,r}(\mathbf{y}',\mathbf{z})) + \sum_{1 \leqslant j \leqslant |\overline{\mathcal{R}}^{(\ell)}(\Phi)|} \gamma_j (A_j^{(\ell)}(\mathbf{y},\mathbf{z}) - A_j^{(\ell)}(\mathbf{y}',\mathbf{z}))$$
$$= \sum_{j=1}^{h_\ell} \widetilde{U}_j(\mathbf{y},\mathbf{y}',\mathbf{0},\mathbf{0},\mathbf{z}) \cdot \widetilde{V}_j(\mathbf{y},\mathbf{y}',\mathbf{0},\mathbf{0},\mathbf{z}). \tag{5.34}$$

Recall every monomial of $Q_{\ell,r}(\mathbf{y},\mathbf{z})$ contains at least one of the $\mathbf{y}$ variables. Thus it follows from the definition of the h-invariant, (5.1), and (5.18) that

$$\begin{aligned} &h_\ell(Q_{\ell,r}(\mathbf{y},\mathbf{z})) \\ &\quad \leqslant M \\ &\quad \leqslant dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R+2R_2)+2R_2+4r_1+2RC_0'') \\ &\qquad + dR^3(R^2+1)^{d-2}(2R_2+1). \end{aligned}$$

Therefore, by moving the term $\sum_{r=1}^{r_\ell} \lambda_r \cdot (Q_{\ell,r}(\mathbf{y},\mathbf{z}) - Q_{\ell,r}(\mathbf{y}',\mathbf{z}))$ to the right hand side of the equation (5.34), we obtain via Lemma 2.6 and (4) of Proposition 2.7 that

$$\begin{aligned} &h_\ell + 2r_\ell M \\ &\quad \geqslant h_\ell(\overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y}',\mathbf{z})) \\ &\quad \geqslant h_\ell(\overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y}',\mathbf{z}); \mathbf{z}) \\ &\quad \geqslant h_\ell(\overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}), \overline{\mathcal{R}}_+^{(\ell)}(\Phi)(\mathbf{y}',\mathbf{z}); \mathbf{z}) \\ &\quad = h_\ell(\overline{\mathcal{R}}^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}), \overline{\mathcal{R}}^{(\ell)}(\Phi)(\mathbf{y}',\mathbf{z}); \mathbf{z}) \\ &\quad = h_\ell(\overline{\mathcal{R}}^{(\ell)}(\Phi)(\mathbf{y},\mathbf{z}); \mathbf{z}) \\ &\quad \geqslant \mathcal{F}_\ell(R_2) \\ &\quad = \rho_{d,d}(2R+2R_2)+2R_2+4r_1 \\ &\qquad + 2R(dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R+2R_2)+2R_2+4r_1+2RC_0'') \\ &\qquad + dR^3(R^2+1)^{d-2}(2R_2+1)). \end{aligned} \tag{5.35}$$

Thus it follows that

$$h_\ell \geqslant \rho_{d,d}(2R+2R_2)+2R_2+4r_1 \geqslant \rho_{d,\ell}(2R+2R_2)+2R_2+4r_1.$$

Therefore, in either case we obtain

$$\begin{aligned} h_\ell \geqslant \rho_{d,\ell}(2R+2R_2)+2R_2+4r_1 \geqslant \rho_{d,\ell}(&2R+2R_2'-2|\overline{\mathcal{R}}_+^{(1)}(\Phi)|-2r_1) \\ &+2|\overline{\mathcal{R}}_+^{(1)}(\Phi)|+2r_1. \end{aligned}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Finally, we also have to show that the linear forms of the system (5.31) are linearly independent over $\mathbb{Q}$. Recall the linear forms of

$$\left\{ \sum_{i=1}^{r_1} g_{r,i} \widetilde{y}_i : 1 \leqslant r \leqslant r_1 \right\} \bigcup \overline{\mathcal{R}}_+^{(1)}(\Phi)(\mathbf{y}, \mathbf{z})$$

are linearly independent over $\mathbb{Q}$, and by construction they are only in the $\mathbf{y}$ variables. It is then a basic exercise in linear algebra to verify that the linear forms of

$$\begin{aligned} &\overline{\mathcal{R}}_+^{(1)}(\Phi)(\mathbf{y}, \mathbf{z}) - \overline{\mathcal{R}}_+^{(1)}(\Phi)(\mathbf{y}', \mathbf{z}) \bigcup \overline{\mathcal{R}}_+^{(1)}(\Phi)(\mathbf{v}, \mathbf{z}) - \overline{\mathcal{R}}_+^{(1)}(\Phi)(\mathbf{v}', \mathbf{z}) \\ &\bigcup \left\{ \sum_{i=1}^{r_1} g_{r,i} \widetilde{y}_i - \sum_{i=1}^{r_1} g_{r,i} \widetilde{y}'_i : 1 \leqslant r \leqslant r_1 \right\} \\ &\bigcup \left\{ \sum_{i=1}^{r_1} g_{r,i} \widetilde{v}_i - \sum_{i=1}^{r_1} g_{r,i} \widetilde{v}'_i : 1 \leqslant r \leqslant r_1 \right\} \end{aligned} \tag{5.36}$$

are linearly independent over $\mathbb{Q}$.

Therefore, we obtain from Corollary 2.5 that

$$\mathcal{W}_1 \ll X^{n+3M-K-2\sum_{\ell=1}^{d} \ell r_\ell - 2D_2'}.$$

5.2. Estimate for $\mathcal{W}_2$. We now estimate $\mathcal{W}_2$, which we can deduce to be the number of solutions $\mathbf{z}, \mathbf{z}', \mathbf{z}'' \in [0, X]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ satisfying the equations

$$\begin{aligned} F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) - F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}') &= 0 \quad (1 \leqslant r \leqslant r_d) \\ F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}) - F_{d,r}(\mathbf{0}, \mathbf{0}, \mathbf{z}'') &= 0 \quad (1 \leqslant r \leqslant r_d) \\ P_{\ell,r}(\mathbf{a}, \mathbf{b}) - P_{\ell,r}(\mathbf{a}', \mathbf{b}') &= 0 \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell) \\ P_{\ell,r}(\mathbf{a}, \mathbf{b}) - P_{\ell,r}(\mathbf{a}'', \mathbf{b}'') &= 0 \quad (2 \leqslant \ell < d, 1 \leqslant r \leqslant r_\ell) \\ \sum_{i=1}^{r_1} g'_{r,i} \widetilde{z}_i - \sum_{i=1}^{r_1} g'_{r,i} \widetilde{z}'_i &= 0 \quad (1 \leqslant r \leqslant r_1) \\ \sum_{i=1}^{r_1} g'_{r,i} \widetilde{z}_i - \sum_{i=1}^{r_1} g'_{r,i} \widetilde{z}''_i &= 0 \quad (1 \leqslant r \leqslant r_1) \\ \mathcal{R}_+(\Psi)(\mathbf{z}) - \mathcal{R}_+(\Psi)(\mathbf{z}') &= \mathbf{0} \\ \mathcal{R}_+(\Psi)(\mathbf{z}) - \mathcal{R}_+(\Psi)(\mathbf{z}'') &= \mathbf{0}, \end{aligned} \tag{5.37}$$

where $\mathbf{z} = (\mathbf{a}, \mathbf{b})$, $\mathbf{z}' = (\mathbf{a}', \mathbf{b}')$, and $\mathbf{z}'' = (\mathbf{a}'', \mathbf{b}'')$. Here the notations $\mathcal{R}_+(\Psi)(\mathbf{z}) - \mathcal{R}_+(\Psi)(\mathbf{z}')$ and $\mathcal{R}_+(\Psi)(\mathbf{z}) - \mathcal{R}_+(\Psi)(\mathbf{z}'')$ should be interpreted in a similar manner as in Section 5.1.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We consider the h-invariant of the system of forms on the left hand side of (5.37), and show that it is a regular system. The degree d forms of the system (5.37) are precisely $F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}')$ and $F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}'')$ $(1 \leqslant r \leqslant r_d)$, and we let h_d be the h-invariant of these degree d forms. Suppose for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{Q}^{r_d}$, not both $\mathbf{0}$, we have

$$\sum_{r=1}^{r_d} \lambda_r \cdot (F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}')) + \sum_{r=1}^{r_d} \mu_r \cdot (F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) - F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}''))$$
$$= \sum_{j=1}^{h_d} \widetilde{U}_j \cdot \widetilde{V}_j, \tag{5.38}$$

where $\widetilde{U}_j = \widetilde{U}_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ and $\widetilde{V}_j = \widetilde{V}_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ are rational forms of positive degree $(1 \leqslant j \leqslant h_d)$. We consider two cases, $(\boldsymbol{\lambda} + \boldsymbol{\mu}) \neq \mathbf{0}$ and $(\boldsymbol{\lambda} + \boldsymbol{\mu}) = \mathbf{0}$. Suppose $(\boldsymbol{\lambda} + \boldsymbol{\mu}) \neq \mathbf{0}$. If we set $\mathbf{z}' = \mathbf{z}'' = \mathbf{0}$, then the above equation (5.38) becomes

$$\sum_{r=1}^{r_d} (\lambda_r + \mu_r) \cdot F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}) = \sum_{j=1}^{h_d} \widetilde{U}_j(\mathbf{z},\mathbf{0},\mathbf{0}) \cdot \widetilde{V}_j(\mathbf{z},\mathbf{0},\mathbf{0}).$$

Thus we obtain

$$h_d \geqslant h_d(\mathbf{F}_d(\mathbf{0},\mathbf{0},\mathbf{z})).$$

On the other hand, suppose $(\boldsymbol{\lambda} + \boldsymbol{\mu}) = \mathbf{0}$, then the above equation (5.38) simplifies to

$$-\sum_{r=1}^{r_d} \lambda_r \cdot (F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}') - F_{d,r}(\mathbf{0},\mathbf{0},\mathbf{z}'')) = \sum_{j=1}^{h_d} \widetilde{U}_j \cdot \widetilde{V}_j.$$

From this equation, by setting $\mathbf{z}'' = \mathbf{0}$ we obtain

$$h_d \geqslant h_d(\mathbf{F}_d(\mathbf{0},\mathbf{0},\mathbf{z})).$$

Therefore, in either case we obtain from (5.11), (5.12), and (5.21) that

$$\begin{aligned} h_d &\geqslant h_d(\mathbf{F}_d(\mathbf{0},\mathbf{0},\mathbf{z})) \\ &\geqslant \rho_{d,d}(2R + 2R_1) + 2R_1 + 4r_1 \\ &\geqslant \rho_{d,d}(2R + 2R_1' - 2|\mathcal{R}_+^{(1)}(\Psi)| - 2r_1) + 2|\mathcal{R}_+^{(1)}(\Psi)| + 2r_1. \end{aligned}$$

We now estimate the h-invariant of the degree ℓ forms of the system (5.37) for each $2 \leqslant \ell < d$. Recall we have $\mathcal{R}_+(\Psi) = (\mathcal{R}^{(d-1)}(\Psi), \ldots, \mathcal{R}^{(2)}(\Psi), \mathcal{R}_+^{(1)}(\Psi))$. The degree ℓ forms of the system (5.37) are precisely $P_{\ell,r}(\mathbf{a},\mathbf{b}) - P_{\ell,r}(\mathbf{a}',\mathbf{b}')$,

$P_{\ell,r}(\mathbf{a},\mathbf{b}) - P_{\ell,r}(\mathbf{a}'',\mathbf{b}'')$ $(1 \leqslant r \leqslant r_\ell)$, and the forms of $\mathcal{R}^{(\ell)}(\Psi)(\mathbf{z}) - \mathcal{R}^{(\ell)}(\Psi)(\mathbf{z}')$ and $\mathcal{R}^{(\ell)}(\Psi)(\mathbf{z}) - \mathcal{R}^{(\ell)}(\Psi)(\mathbf{z}'')$. We let h_ℓ be the h-invariant of these degree ℓ forms. Then for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{Q}^{r_\ell}$ and $\boldsymbol{\gamma}, \boldsymbol{\gamma}' \in \mathbb{Q}^{|\mathcal{R}^{(\ell)}(\Psi)|}$, not all zero vectors, we have

$$\begin{aligned}
&\sum_{r=1}^{r_\ell} \lambda_r (P_{\ell,r}(\mathbf{a},\mathbf{b}) - P_{\ell,r}(\mathbf{a}',\mathbf{b}') + \mu_r (P_{\ell,r}(\mathbf{a},\mathbf{b}) - P_{\ell,r}(\mathbf{a}'',\mathbf{b}''))\\
&\quad + \sum_{j=1}^{|\mathcal{R}^{(\ell)}(\Psi)|} \gamma_j \cdot (V_j^{(\ell)}(\mathbf{z}) - V_j^{(\ell)}(\mathbf{z}')) + \gamma_j' \cdot (V_j^{(\ell)}(\mathbf{z}) - V_j^{(\ell)}(\mathbf{z}''))\\
&= \sum_{j=1}^{h_\ell} \widetilde{U}_j \cdot \widetilde{V}_j,
\end{aligned} \tag{5.39}$$

where $\widetilde{U}_j = \widetilde{U}_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ and $\widetilde{V}_j = \widetilde{V}_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ are rational forms of positive degree $(1 \leqslant j \leqslant h_\ell)$. We consider two cases, $\boldsymbol{\gamma} = \boldsymbol{\gamma}' = \mathbf{0}$ and at least one of $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$ is not a zero vector.

First we suppose that $\boldsymbol{\gamma} = \boldsymbol{\gamma}' = \mathbf{0}$. In this case, at least one of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ is not a zero vector. Without loss of generality, suppose $\boldsymbol{\lambda} \neq \mathbf{0}$. Then by setting $\mathbf{z} = \mathbf{z}''$ and $\mathbf{z}' = \mathbf{0}$, the equation (5.39) becomes

$$\sum_{1 \leqslant r \leqslant r_\ell} \lambda_r P_{\ell,r}(\mathbf{a},\mathbf{b}) = \sum_{j=1}^{h_\ell} \widetilde{U}_j(\mathbf{z},\mathbf{0},\mathbf{z}) \cdot \widetilde{V}_j(\mathbf{z},\mathbf{0},\mathbf{z}).$$

Therefore, it follows from (5.10) and (5.19) that

$$\begin{aligned}
h_\ell &\geqslant h_\ell(\{P_{\ell,r}(\mathbf{a},\mathbf{b}) : 1 \leqslant r \leqslant r_\ell\})\\
&\geqslant \rho_{d,\ell}(2R + 2R_1) + 2R_1 + 4r_1\\
&\geqslant \rho_{d,\ell}(2R + 2R_1' - 2|\mathcal{R}_+^{(1)}(\Psi)| - 2r_1) + 2|\mathcal{R}_+^{(1)}(\Psi)| + 2r_1.
\end{aligned}$$

Next we suppose at least one of $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$ is not a zero vector. Without loss of generality, suppose $\boldsymbol{\gamma} \neq \mathbf{0}$. We consider two further subcases, $(\boldsymbol{\gamma} + \boldsymbol{\gamma}') \neq \mathbf{0}$ and $(\boldsymbol{\gamma} + \boldsymbol{\gamma}') = \mathbf{0}$.

Suppose $(\boldsymbol{\gamma} + \boldsymbol{\gamma}') \neq \mathbf{0}$. In this case, we set $\mathbf{z}' = \mathbf{z}'' = \mathbf{0}$, and the equation (5.39) simplifies to

$$\begin{aligned}
&\sum_{1 \leqslant r \leqslant r_\ell} (\lambda_r + \mu_r) P_{\ell,r}(\mathbf{a},\mathbf{b}) + \sum_{j=1}^{|\mathcal{R}^{(\ell)}(\Psi)|} (\gamma_j + \gamma_j') \cdot V_j^{(\ell)}(\mathbf{z})\\
&= \sum_{j=1}^{h_\ell} \widetilde{U}_j(\mathbf{z},\mathbf{0},\mathbf{0}) \cdot \widetilde{V}_j(\mathbf{z},\mathbf{0},\mathbf{0}).
\end{aligned} \tag{5.40}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Recall every monomial of $P_{\ell,r}(\mathbf{a},\mathbf{b})$ contains at least one of the **a** variables. Thus it follows from the definition of the h-invariant, (5.5), and (5.20) that

$$h_\ell(P_{\ell,r}(\mathbf{a},\mathbf{b})) \leqslant M' \leqslant dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R+2R_1)+2R_1+4r_1+2RC_0'').$$

Therefore, by moving the term $\sum_{1\leqslant r\leqslant r_\ell}(\lambda_r+\mu_r)P_{\ell,r}(\mathbf{a},\mathbf{b})$ to the right hand side of the equation (5.40), we obtain via (3) of Proposition 2.7 that

$$\begin{aligned} h_\ell + M'r_\ell &\geqslant h_\ell(\mathcal{R}^{(\ell)}(\Psi)) \\ &\geqslant \mathcal{F}_\ell'(R_1) \\ &= \rho_{d,d}(2R+2R_1)+2R_1+4r_1 \\ &\quad + 2R(dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R+2R_1)+2R_1+4r_1+2RC_0'')). \end{aligned}$$

Thus we obtain

$$\begin{aligned} h_\ell &\geqslant \rho_{d,d}(2R+2R_1)+2R_1+4r_1 \\ &\geqslant \rho_{d,\ell}(2R+2R_1'-2|\mathcal{R}_+^{(1)}(\Psi)|-2r_1)+2|\mathcal{R}_+^{(1)}(\Psi)|+2r_1. \end{aligned} \tag{5.41}$$

On the other hand, we now suppose $(\boldsymbol{\gamma}+\boldsymbol{\gamma}')=\mathbf{0}$. By setting $\mathbf{z}=\mathbf{z}''=\mathbf{0}$, the equation (5.39) simplifies to

$$-\sum_{r=1}^{r_\ell}\lambda_r\cdot P_{\ell,r}(\mathbf{a}',\mathbf{b}') - \sum_{j=1}^{|\mathcal{R}^{(\ell)}(\Psi)|}\gamma_j\cdot V_j^{(\ell)}(\mathbf{z}') = \sum_{j=1}^{h_\ell}\widetilde{U}_j(\mathbf{0},\mathbf{z}',\mathbf{0})\cdot\widetilde{V}_j(\mathbf{0},\mathbf{z}',\mathbf{0}).$$

Then by a similar argument as above, we have

$$\begin{aligned} h_\ell + M'r_\ell &\geqslant h_\ell(\mathcal{R}^{(\ell)}(\Psi)) \\ &\geqslant \mathcal{F}_\ell'(R_1) \\ &= \rho_{d,d}(2R+2R_1)+2R_1+4r_1 \\ &\quad + 2R(dR(R^2+1)^{d-2}2^d(\rho_{d,d}(2R+2R_1)+2R_1+4r_1+2RC_0'')). \end{aligned}$$

Therefore, we also obtain

$$\begin{aligned} h_\ell &\geqslant \rho_{d,d}(2R+2R_1)+2R_1+4r_1 \\ &\geqslant \rho_{d,\ell}(2R+2R_1'-2|\mathcal{R}_+^{(1)}(\Psi)|-2r_1)+2|\mathcal{R}_+^{(1)}(\Psi)|+2r_1 \end{aligned} \tag{5.42}$$

in this case.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We also have to show that the linear forms of the system (5.37),

$$\{\mathcal{R}_+^{(1)}(\Psi)(\mathbf{z}) - \mathcal{R}_+^{(1)}(\Psi)(\mathbf{z}')\} \bigcup \{\mathcal{R}_+^{(1)}(\Psi)(\mathbf{z}) - \mathcal{R}_+^{(1)}(\Psi)(\mathbf{z}'')\}$$
$$\bigcup \left\{ \sum_{i=1}^{r_1} g'_{r,i}\widetilde{z}_i - \sum_{i=1}^{r_1} g'_{r,i}\widetilde{z}'_i : 1 \leqslant r \leqslant r_1 \right\}$$
$$\bigcup \left\{ \sum_{i=1}^{r_1} g'_{r,i}\widetilde{z}_i - \sum_{i=1}^{r_1} g'_{r,i}\widetilde{z}''_i : 1 \leqslant r \leqslant r_1 \right\}, \tag{5.43}$$

are linearly independent over $\mathbb{Q}$. Recall the linear forms of

$$\mathcal{R}_+^{(1)}(\Psi)(\mathbf{z}) \bigcup \left\{ \sum_{i=1}^{r_1} g'_{r,i}\widetilde{z}_i : 1 \leqslant r \leqslant r_1 \right\}$$

are linearly independent over $\mathbb{Q}$. Using this fact, the verification of linear independence over $\mathbb{Q}$ of the system of linear forms (5.43) is a basic exercise in linear algebra.

Therefore, we obtain by Corollary 2.5 that

$$\mathcal{W}_2 \ll X^{3(n-M-K)-2\sum_{\ell=1}^{d} \ell r_\ell - 2D'_1}.$$

5.3. Proof of Claim 1. Recall we defined

$$S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) = \sum_{\mathbf{w}\in[0,X]^K} \Lambda(\mathbf{w}) e\left(\sum_{1\leqslant r\leqslant r_1} \alpha_{1,r}(c_{1,r}\mathbf{w}^{\mathbf{j}_{1,r}} + \widetilde{f}_{1,r}(\mathbf{w}, \mathbf{0}, \mathbf{0})) \right.$$
$$\left. + \sum_{2\leqslant \ell\leqslant d} \sum_{1\leqslant r\leqslant r_\ell} \alpha_{\ell,r} \cdot \mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}) \right), \tag{5.44}$$

where

$$\mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}) = f_{\ell,r}(\mathbf{w}, \mathbf{0}, \mathbf{0})$$
$$+ \sum_{j=1}^{\ell-1} \sum_{1\leqslant i_1\leqslant\cdots\leqslant i_j\leqslant K} \left(\sum_{k=1}^{\ell-j} c^{(k)}_{\ell,r:i_1,\ldots,i_j}(\mathbf{G}, \mathbf{H}) \right) w_{i_1}\cdots w_{i_j}.$$

Also recall we defined the monomials $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ $(1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$ in (4.2). If we consider the expression in the exponent of (5.44),

$$\sum_{1\leqslant r\leqslant r_1} \alpha_{1,r}(c_{1,r}\mathbf{w}^{\mathbf{j}_{1,r}} + \widetilde{f}_{1,r}(\mathbf{w}, \mathbf{0}, \mathbf{0})) + \sum_{2\leqslant \ell\leqslant d} \sum_{1\leqslant r\leqslant r_\ell} \alpha_{\ell,r} \cdot \mathfrak{C}_{\ell,r}(\mathbf{w}, \mathbf{G}, \mathbf{H}),$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

as a polynomial in $\mathbf{w}$ with real coefficients, then it follows from the discussion after (4.11) that the coefficient of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ of this polynomial is $c_{\ell,r}\alpha_{\ell,r}$. Furthermore, this polynomial does not contain any monomial divisible by $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ other than itself.

We need to introduce few definitions and lemmas before we can begin with the proof of Claim 1. Let $1 \leqslant \ell \leqslant d$, $q \in \mathbb{N}$, and $a_\ell \in \mathbb{Z}/q\mathbb{Z}$. For $q > 1$ we define

$$\mathfrak{N}^{(\ell)}_{a_\ell,q}(C_0) = \{\xi_\ell \in [0,1) : |\xi_\ell - a_\ell/q| \leqslant (\log X)^{C_0} X^{-\ell}\},$$

and when $q = 1$ we let

$$\mathfrak{N}^{(\ell)}_{0,1}(C_0) = \{\xi_\ell \in [0,1) : \min\{|\xi_\ell|, |\xi_\ell - 1|\} \leqslant (\log X)^{C_0} X^{-\ell}\}.$$

We set

$$\mathfrak{N}(C_0) = \bigcup_{q \leqslant (\log X)^{C_0}} \ \bigcup_{\substack{\gcd(a_d,\ldots,a_1,q)=1 \\ a_d,\ldots,a_1 \in \mathbb{Z}/q\mathbb{Z}}} \mathfrak{N}^{(d)}_{a_d,q}(C_0) \times \cdots \times \mathfrak{N}^{(1)}_{a_1,q}(C_0),$$

and denote

$$\mathfrak{n}(C_0) = [0,1)^d \backslash \mathfrak{N}(C_0).$$

Let $\mathbb{U}_q$ be the group of units in $\mathbb{Z}/q\mathbb{Z}$. When $q = 1$ we let $\mathbb{U}_1 = \{0\}$. Let us also denote

$$\mathfrak{n}^{(\ell)}(C_0) = [0,1) \Bigg\backslash \left(\bigcup_{q \leqslant (\log X)^{C_0}} \ \bigcup_{a_\ell \in \mathbb{U}_q} \mathfrak{N}^{(\ell)}_{a_\ell,q}(C_0) \right).$$

Suppose $\boldsymbol{\xi} = (\xi_d, \ldots, \xi_1) \in [0,1)^d$ satisfies $\xi_\ell \in \mathfrak{n}^{(\ell)}(C_0)$ for some $1 \leqslant \ell \leqslant d$. Then it is clear that $\boldsymbol{\xi} \in \mathfrak{n}(C_0)$.

We have the following lemma which is a special case of [**14**, Ch. VI, Section 1, Theorem 10].

LEMMA 5.1 [**14**, Ch. VI, Section 1, Theorem 10]. *Let* $\ell \geqslant 1$, $\alpha_{\ell-1}, \ldots, \alpha_1, \alpha_0 \in \mathbb{R}$, *and* $\gcd(a,q) = 1$ *with* $(\log X)^\sigma < q \leqslant X^\ell (\log X)^{-\sigma}$. *Suppose we have* $\sigma_0 > 0$ *such that* $\sigma \geqslant 2^{6\ell}(\sigma_0 + 1)$. *Then we have*

$$\sum_{\substack{p \leqslant X \\ p \text{ prime}}} e\left(\frac{a}{q} p^\ell + \alpha_{\ell-1} p^{\ell-1} + \cdots + \alpha_1 p + \alpha_0\right) \ll \frac{X}{(\log X)^{\sigma_0}},$$

where the implicit constant depends only on ℓ.

From this lemma we can obtain the following, which is essentially a special case of [**14**, Ch. X, Section 5, Lemma 10.8].

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

LEMMA 5.2 [14, Ch. X, Section 5, Lemma 10.8]. *Suppose* $\ell \geqslant 1$ *and* $\alpha_\ell, \ldots, \alpha_1 \in \mathbb{R}$. *Let*

$$T_1(\alpha_\ell, \ldots, \alpha_1) = \sum_{x \in [0, X]} \Lambda(x) e(\alpha_\ell x^\ell + \cdots + \alpha_1 x).$$

Given any $c_0 > 0$*, for sufficiently large* $C_0 > 0$ *we have*

$$|T_1(\alpha_\ell, \ldots, \alpha_1)| \ll \frac{X}{(\log X)^{c_0}}$$

for any $\alpha_\ell, \ldots, \alpha_1 \in \mathbb{R}$ *with* $\alpha_\ell \in \mathfrak{n}^{(\ell)}(C_0)$. *Here the implicit constant depends only on* ℓ.

Proof. By Dirichlet's theorem on diophantine approximation, there exist $a, q \in \mathbb{Z}$ such that $\gcd(a, q) = 1$, $1 \leqslant q \leqslant X^\ell (\log X)^{-C_0}$, and

$$|q\alpha_\ell - a| < \frac{(\log X)^{C_0}}{X^\ell}. \tag{5.45}$$

Since we have

$$\left| \alpha_\ell - \frac{a}{q} \right| < \frac{(\log X)^{C_0}}{q X^\ell} \leqslant \frac{(\log X)^{C_0}}{X^\ell}, \tag{5.46}$$

it follows from the definition of $\mathfrak{n}^{(\ell)}(C_0)$ that $q > (\log X)^{C_0}$. Let $\beta_\ell = \alpha_\ell - a/q$. Then we obtain from (5.45) that

$$|\beta_\ell| = \left| \alpha_\ell - \frac{a}{q} \right| < \frac{(\log X)^{C_0}}{q X^\ell} \leqslant \frac{1}{X^\ell}.$$

We now have the setup to apply Lemma 5.1. Let us define

$$T_0(\alpha_\ell, \ldots, \alpha_1) = \sum_{\substack{1 \leqslant p \leqslant X \\ p \text{ prime}}} e(\alpha_\ell p^\ell + \cdots + \alpha_1 p).$$

By following the argument in the proof of [14, Ch. X, Section 5, Lemma 10.8], we obtain that given any $c_0 > 0$, for $C_0 > 0$ sufficiently large we have

$$|T_0(\alpha_\ell, \ldots, \alpha_1)| \ll \frac{X}{(\log X)^{c_0}},$$

where the implicit constant depends only on ℓ. From here we obtain via partial summation the required bound on $T_1(\alpha_\ell, \ldots, \alpha_1)$. □

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Recall $\|\alpha\|$ is the distance from $\alpha \in \mathbb{R}$ to the closest integer. The following is a special case of [20, Lemma 14.1].

LEMMA 5.3 [20, Lemma 14.1]. *Suppose* $\lambda \in \mathbb{R}$, $A > 1$, *and* $Z > 0$. *Let* $\mathcal{N}(Z)$ *be the number of integers* v *such that*

$$|v| \leqslant ZA \quad \textit{and} \quad \|\lambda v\| \leqslant ZA^{-1}. \tag{5.47}$$

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

$$\mathcal{N}(Z_1) \gg (Z_1/Z_2)\mathcal{N}(Z_2),$$

where the implicit constant is an absolute constant.

We now begin with the proof of Claim 1. Let M_0 be the diagonal $R \times R$ matrix where its diagonal entries from the top left corner to the right bottom corner are $c_{d,1}, c_{d,2}, \ldots, c_{d,r_d}, c_{d-1,1}, c_{d-1,2}, \ldots, c_{d-1,r_{d-1}}, \ldots, c_{1,1}, c_{1,2}, \ldots, c_{1,r_1}$ in this order. Clearly M_0 is an invertible matrix. Let $\gamma_{\ell,r} = \alpha_{\ell,r}c_{\ell,r}$. Consider the polynomial in the exponent of (5.44) as a polynomial in the $\mathbf{w}$ variables. Then we know that the coefficient of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ of this polynomial is $\gamma_{\ell,r}$. We also have

$$M_0 \cdot \begin{bmatrix} \alpha_{d,1} \\ \vdots \\ \alpha_{1,r_1} \end{bmatrix} = \begin{bmatrix} \gamma_{d,1} \\ \vdots \\ \gamma_{1,r_1} \end{bmatrix} \in \mathbb{R}^R.$$

Suppose $\boldsymbol{\gamma} \in \mathfrak{M}(C')$ for some $C' > 0$, then there exist $\mathbf{a} \in \mathbb{Z}^R$ and $q \in \mathbb{N}$ such that $\gcd(\mathbf{a}, q) = 1$, $0 < q \leqslant (\log X)^{C'}$, and $|\gamma_{\ell,r} - a_{\ell,r}/q| \leqslant (\log X)^{C'}/X^{\ell}$ $(1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell)$. Let us denote

$$M_0^{-1} \cdot \begin{bmatrix} a_{d,1}/q \\ \vdots \\ a_{1,r_1}/q \end{bmatrix} = \begin{bmatrix} a'_{d,1}/q' \\ \vdots \\ a'_{1,r_1}/q' \end{bmatrix} \quad \text{and} \quad M_0^{-1} \cdot \begin{bmatrix} \gamma_{d,1} - a_{d,1}/q \\ \vdots \\ \gamma_{1,r_1} - a_{1,r_1}/q \end{bmatrix} = \begin{bmatrix} \beta'_{d,1} \\ \vdots \\ \beta'_{1,r_1} \end{bmatrix}.$$

It is easy to deduce that

$$q' \leqslant (\log X)^{C'+1} \quad \text{and} \quad |\beta'_{\ell,r}| \leqslant \frac{(\log X)^{C'+1}}{X^{\ell}} \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$$

for X sufficiently large with respect to $c_{d,1}, \ldots, c_{1,r_1}$. Since $\alpha_{\ell,r} = a'_{\ell,r}/q' + \beta'_{\ell,r}$, we see that $\boldsymbol{\alpha} \in \mathfrak{M}(C'+1)$. Now since $\boldsymbol{\alpha} \in \mathfrak{m}(C)$, it follows from this argument that $\boldsymbol{\gamma} \in \mathfrak{m}(C-1)$. Then there exist ℓ and r such that $\gamma_{\ell,r} \in \mathfrak{n}^{(\ell)}(C'')$, where

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$C'' = (C-1)/R$, by the following reason. Suppose $\gamma_{\ell,r} \notin \mathfrak{n}^{(\ell)}(C'')$ $(1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell)$. Then for each $1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$ there exist $q_{\ell,r} \in \mathbb{N}$ and $a_{\ell,r} \in \mathbb{Z}$ such that

$$q_{\ell,r} \leqslant (\log X)^{C''} \quad \text{and} \quad |\gamma_{\ell,r} - a_{\ell,r}/q_{\ell,r}| \leqslant \frac{(\log X)^{C''}}{X^\ell}.$$

By taking q to be the appropriate factor of the lowest common multiple of $q_{d,1}$, $\ldots, q_{1,r_1}$, we see that $\boldsymbol{\gamma} \in \mathfrak{M}(C-1)$, which is a contradiction.

Throughout the remainder of this section, we fix ℓ and r to be such that $\gamma_{\ell,r} \in \mathfrak{n}^{(\ell)}(C'')$. Following [6], we consider two cases depending on $\mathbf{w}^{\mathbf{j}_{\ell,r}}$: Case 1 is when $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ contains only one distinct variable, and Case 2 is when it has more than one distinct variable.

Case 1: Without loss of generality, suppose $\mathbf{w}^{\mathbf{j}_{\ell,r}} = w_1^\ell$. We may bound $S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H})$ as follows

$$S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) \leqslant (\log X)^{K-1} \cdot \sum_{w_K \in [0,X]} \ldots \sum_{w_2 \in [0,X]} \left| \sum_{w_1 \in [0,X]} \Lambda(w_1) e(\gamma_{\ell,r} w_1^\ell + \tau(w_1, w_2, \ldots, w_K, \mathbf{G}, \mathbf{H})) \right|, \quad (5.48)$$

where $\tau(w_1, w_2, \ldots, w_K, \mathbf{G}, \mathbf{H})$ has degree strictly less than ℓ as a polynomial in w_1 with coefficients possibly dependent on $w_2, \ldots, w_K, \mathbf{G}, \mathbf{H}$. This follows from the fact that the coefficient of w_1^ℓ of the polynomial in the exponent of (5.44) is $\gamma_{\ell,r}$, and that there are no other monomials divisible by w_1^ℓ.

Therefore, since $\gamma_{\ell,r} \in \mathfrak{n}^{(\ell)}(C'')$ we may apply Lemma 5.2 with $c_0 = c + K - 1$ to the inner sum of (5.48) and obtain

$$S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H}) \ll (\log X)^{K-1} X^{K-1} \frac{X}{(\log X)^{c+K-1}} = \frac{X^K}{(\log X)^c}.$$

Case 2: We have that $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ contains at least two distinct variables. In particular, we must have $\ell > 1$. By relabeling if necessary, let $\mathbf{w}^{\mathbf{j}_{\ell,r}} = w_1^{j_1} \ldots w_k^{j_k}$ where $j_1, \ldots, j_k > 0$. We know that the coefficient of $\mathbf{w}^{\mathbf{j}_{\ell,r}}$ of the polynomial in the exponent of (5.44) is $\gamma_{\ell,r}$. In this case, we may bound $S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H})$ as follows

$$|S_0(\boldsymbol{\alpha}, \mathbf{G}, \mathbf{H})| \leqslant (\log X)^{K-k} \cdot \sum_{w_K \in [0,X]} \cdots \sum_{w_{k+1} \in [0,X]} |S(w_1, \ldots, w_K, \mathbf{G}, \mathbf{H})|, \quad (5.49)$$

where

$$S(w_1, \ldots, w_K, \mathbf{G}, \mathbf{H}) = \sum_{w_1 \in [0,X]} \cdots \sum_{w_k \in [0,X]} \Lambda(w_1) \cdots \Lambda(w_k) e(\gamma_{\ell,r} \mathbf{w}^{\mathbf{j}_{\ell,r}} + \Theta(w_1, \ldots, w_k)),$$

and $\Theta(w_1, \ldots, w_k) = \Theta(w_1, \ldots, w_k : w_{k+1}, \ldots, w_K, \mathbf{G}, \mathbf{H})$ is a polynomial in $w_1, \ldots, w_k$ with coefficients possibly dependent on $w_{k+1}, \ldots, w_K, \mathbf{G}, \mathbf{H}$. By construction, we also know that this polynomial does not have any monomial divisible by $\mathbf{w}^{\mathbf{j}_{\ell,r}}$.

We now apply Weyl differencing ℓ times, where we apply it j_i times to the variable w_i for each $1 \leqslant i \leqslant k$. The point is that with this process every monomial of $\gamma_{\ell,r} \mathbf{w}^{\mathbf{j}_{\ell,r}} + \Theta(w_1, \ldots, w_k)$ for which at least one of w_i has degree strictly less than j_i will vanish, in particular every monomial of $\Theta(w_1, \ldots, w_k)$ will vanish. Let $\widetilde{c} = j_1! \ldots j_k!$. As a result, we obtain

$$|S(w_1, \ldots, w_K, \mathbf{G}, \mathbf{H})|^{2^\ell} \ll (\log X)^{k2^\ell} X^{k2^\ell - \ell} \sum_{\substack{v_i \in [-X,X] \\ 1 \leqslant i \leqslant \ell-1}} \min\{X, \|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\|^{-1}\}. \quad (5.50)$$

Since this is a standard application of Weyl differencing, and also similar to the argument in [6, pages 725–726], we leave the details to the reader.

Let

$$\mathcal{A}_X := \left\{ (v_1, \ldots, v_{\ell-1}) \in [-X, X]^{\ell-1} \cap \mathbb{Z}^{\ell-1} : \|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\| \leqslant \frac{1}{X} \right\}.$$

For any $1 \leqslant X' < X$, we define the set

$$\mathcal{A}_{X,X'} := \left\{ (v_1, \ldots, v_{\ell-1}) \in [-X/X', X/X']^{\ell-1} \cap \mathbb{Z}^{\ell-1} : \|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\| \leqslant \frac{1}{X(X')^{\ell-1}} \right\}.$$

By applying Lemma 5.3 successively in the variables $v_1, \ldots, v_{\ell-1}$, we obtain

$$|\mathcal{A}_X| \ll (X')^{\ell-1} |\mathcal{A}_{X,X'}|.$$

Let $X' = X(\log X)^{-C''/d}$. Suppose there exists $(v_1, \ldots, v_{\ell-1}) \in \mathcal{A}_{X,X'}$ such that $(v_1 \cdots v_{\ell-1}) \neq 0$. Then we have

$$|\widetilde{c} v_1 \cdots v_{\ell-1}| \leqslant (\log X)^{C''} \quad \text{and} \quad \|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\| \leqslant \frac{(\log X)^{C''}}{X^\ell}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

for X sufficiently large with respect to ℓ, and this contradicts the fact that $\gamma_{\ell,r} \in \mathfrak{n}^{(\ell)}(C'')$. Thus at least one of $v_1, \ldots, v_{\ell-1}$ must be 0. Therefore, we have

$$|\mathcal{A}_{X,X'}| \ll (\log X)^{(\ell-2)C''/d},$$

and consequently,

$$|\mathcal{A}_X| \ll \left(\frac{X}{(\log X)^{C''/d}}\right)^{\ell-1} |\mathcal{A}_{X,X'}| \ll \frac{X^{\ell-1}}{(\log X)^{C''/d}}. \tag{5.51}$$

We now proceed in a similar manner as in [20, Lemma 13.2]. First let us deal with the case $\ell > 2$. Let $N_0(v_1', \ldots, v_{\ell-2}')$ be the number of points $v_{\ell-1} \in [-X, X] \cap \mathbb{Z}$ such that $(v_1', \ldots, v_{\ell-2}', v_{\ell-1}) \in \mathcal{A}_X$. Then we have

$$|\mathcal{A}_X| = \sum_{v_1 \in [-X,X]} \cdots \sum_{v_{\ell-2} \in [-X,X]} N_0(v_1, \ldots, v_{\ell-2}), \tag{5.52}$$

and let $N_0 = |\mathcal{A}_X|$ when $\ell = 2$.

Let us write $\{\alpha\}$ for the fractional part of a real number α, that is, $\{\alpha\} = \alpha - \max_{\substack{z \leqslant \alpha \\ z \in \mathbb{Z}}} z$. Then for any set of integers $v_1, \ldots, v_{\ell-2}$, and $a \in \mathbb{Z}$ with $0 \leqslant a < X$, the inequality

$$\frac{a}{X} \leqslant \{\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\} < \frac{a+1}{X} \tag{5.53}$$

cannot hold for more than $N_0(v_1, \ldots, v_{\ell-2})$ integer points $v_{\ell-1}$ lying inside $[-X, X]$ for the following reason. Suppose this is indeed the case, and let $v_{\ell-1} \in [-X, X]$ be one integer which satisfies (5.53). If $v_{\ell-1}$ and $v_{\ell-1}'$ are two distinct points that satisfy (5.53), then we have

$$\|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-2}(v_{\ell-1} - v_{\ell-1}')\| < \frac{1}{X}$$

and $(v_{\ell-1} - v_{\ell-1}') \in [-X, X] \cap \mathbb{Z}$. Consequently, we have $(v_1, \ldots, v_{\ell-2}, v_{\ell-1} - v_{\ell-1}') \in \mathcal{A}_X$ from which we can obtain contradiction. Therefore, we obtain the following inequalities

$$\begin{aligned} &\sum_{v_{\ell-1} \in [-X,X]} \min(X, \|\widetilde{c}\gamma_{\ell,r} v_1 \cdots v_{\ell-1}\|^{-1}) \\ &\ll N_0(v_1, \ldots, v_{\ell-2}) \sum_{0 \leqslant a \leqslant X} \min\left(X, \max\left(\frac{X}{a}, \frac{X}{|X-a-1|}\right)\right) \\ &\ll N_0(v_1, \ldots, v_{\ell-2}) X \log X. \end{aligned} \tag{5.54}$$

Thus via (5.51), (5.52), and (5.54), we have the following bound for (5.50),

$$\begin{aligned}
&|S(w_1,\ldots,w_K,\mathbf{G},\mathbf{H})|^{2^\ell}\\
&\leqslant (\log X)^{2^\ell k} X^{2^\ell k-\ell} \sum_{v_1\in[-X,X]}\cdots\sum_{v_{\ell-1}\in[-X,X]} \min(X,\|\widetilde{c}\gamma_{\ell,r}v_1\cdots v_{\ell-1}\|^{-1})\\
&\ll (\log X)^{2^\ell k} X^{2^\ell k-\ell} \sum_{v_1\in[-X,X]}\cdots\sum_{v_{\ell-2}\in[-X,X]} N_0(v_1,\ldots,v_{\ell-2})X\log X\\
&= (\log X)^{2^\ell k} X^{2^\ell k-\ell}|\mathcal{A}_X|X\log X\\
&\leqslant X^{2^\ell k}(\log X)^{2^\ell k+1-C''/d},
\end{aligned}$$

and hence

$$|S(w_1,\ldots,w_K,\mathbf{G},\mathbf{H})| \ll X^k(\log X)^{k+2^{-\ell}(1-C''/d)}.$$

Therefore, we obtain from (5.49) that

$$\begin{aligned}
|S_0(\boldsymbol{\alpha},\mathbf{G},\mathbf{H})| &\ll (\log X)^{K-k}\cdot \sum_{w_K\in[0,X]}\cdots\sum_{w_{k+1}\in[0,X]} X^k(\log X)^{k+2^{-\ell}(1-C''/d)}\\
&\ll (\log X)^K X^K(\log X)^{2^{-\ell}(1-C''/d)}.
\end{aligned}$$

The case $\ell=2$ can be dealt with in a similar and more simple manner. Recall from above $C''=(C-1)/R$ and $K\leqslant dR$. Thus we make sure C is sufficiently large with respect to d and R. This completes the proof of Claim 1, and hence the proof of Proposition 4.1 as well. □

6. Technical estimates

In this section, we collect results related to Weyl differencing that are necessary in obtaining estimates for the singular integral and the singular series defined in (7.4) and (7.19), respectively.

Let us denote $\mathfrak{B}_0=[0,1]^n$. Let $\boldsymbol{\alpha}=(\boldsymbol{\alpha}_d,\ldots,\boldsymbol{\alpha}_1)\in\mathbb{R}^R$, where $R=r_1+\cdots+r_d$ and $\boldsymbol{\alpha}_\ell=(\alpha_{\ell,1},\ldots,\alpha_{\ell,r_\ell})\in\mathbb{R}^{r_\ell}$ $(1\leqslant\ell\leqslant d)$. We define

$$\|\boldsymbol{\alpha}\| = \max_{\substack{1\leqslant\ell\leqslant d\\1\leqslant r\leqslant r_\ell}}\|\alpha_{\ell,r}\| \quad\text{and}\quad |\boldsymbol{\alpha}| = \max_{\substack{1\leqslant\ell\leqslant d\\1\leqslant r\leqslant r_\ell}}|\alpha_{\ell,r}|.$$

Let $\mathbf{u}=(\mathbf{u}_d,\ldots,\mathbf{u}_1)$ be a system of polynomials in $\mathbb{Q}[x_1,\ldots,x_n]$, where $\mathbf{u}_\ell=(u_{\ell,1},\ldots,u_{\ell,r_\ell})$ is the subsystem of degree ℓ polynomials of $\mathbf{u}$ $(1\leqslant\ell\leqslant d)$. We let $\mathbf{U}=(\mathbf{U}_d,\ldots,\mathbf{U}_1)$ be the system of forms, where for each $1\leqslant\ell\leqslant d$, $\mathbf{U}_\ell=(U_{\ell,1},\ldots,U_{\ell,r_\ell})$ and $U_{\ell,r}$ is the homogeneous degree ℓ portion of $u_{\ell,r}$ $(1\leqslant r\leqslant r_\ell)$.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We define the following exponential sum associated to $\mathbf{u}$,

$$S(\boldsymbol{\alpha}) = S(\mathbf{u}, \mathfrak{B}_0; \boldsymbol{\alpha}) := \sum_{\mathbf{x} \in P\mathfrak{B}_0 \cap \mathbb{Z}^n} e\left(\sum_{1 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot u_{\ell,r}(\mathbf{x}) \right). \tag{6.1}$$

Let $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,n})$ for $i \geqslant 1$, and let

$$\Gamma_{\ell, U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_\ell) = \sum_{t_1=0}^{1} \cdots \sum_{t_\ell=0}^{1} (-1)^{t_1+\cdots+t_\ell} U_{\ell,r}(t_1\mathbf{x}_1 + \cdots + t_\ell \mathbf{x}_\ell).$$

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors of $\mathbb{C}^n$. Let $1 < \ell \leqslant d$ and $r_\ell > 0$. We define $\mathbb{M}_\ell = \mathbb{M}_\ell(\mathbf{U}_\ell)$ to be the set of $(\ell-1)$-tuples $(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}) \in (\mathbb{C}^n)^{\ell-1}$ for which the matrix

$$[m_{ri}] = [\Gamma_{\ell, U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i)] \quad (1 \leqslant r \leqslant r_\ell, 1 \leqslant i \leqslant n) \tag{6.2}$$

has rank strictly less than r_ℓ. For $P_0 > 0$, we denote $z_{P_0}(\mathbb{M}_\ell)$ to be the number of integer points $(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1})$ on $\mathbb{M}_\ell$ such that

$$\max_{1 \leqslant i \leqslant \ell-1} \max_{1 \leqslant j \leqslant n} |x_{i,j}| \leqslant P_0.$$

We define $g_\ell(\mathbf{U}_\ell)$ to be the largest real number such that

$$z_P(\mathbb{M}_\ell) \ll P^{n(\ell-1) - g_\ell(\mathbf{U}_\ell) + \varepsilon} \tag{6.3}$$

holds for each $\varepsilon > 0$. It was proved in [20, page 280, Corollary] that

$$h_\ell(\mathbf{U}_\ell) < \frac{\ell!}{(\log 2)^\ell} (g_\ell(\mathbf{U}_\ell) + (\ell-1) r_\ell (r_\ell - 1)). \tag{6.4}$$

Let

$$\gamma_\ell = \frac{2^{\ell-1}(\ell-1) r_\ell}{g_\ell(\mathbf{U}_\ell)}$$

when $r_\ell > 0$ and $g_\ell(\mathbf{U}_\ell) > 0$. We let $\gamma_\ell = 0$ if $r_\ell = 0$, and let $\gamma_\ell = +\infty$ if $r_\ell > 0$ and $g_\ell(\mathbf{U}_\ell) = 0$. For ℓ with $r_\ell > 0$, we also define

$$\gamma'_\ell = \frac{2^{\ell-1}}{g_\ell(\mathbf{U}_\ell)} = \frac{\gamma_\ell}{(\ell-1) r_\ell}. \tag{6.5}$$

We need the following lemma to obtain estimates on the singular integral. Let

$$\mathcal{I}(\mathfrak{B}_0, \boldsymbol{\tau}) = \int_{\mathbf{v} \in \mathfrak{B}_0} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} \tau_{\ell,r} \cdot U_{\ell,r}(\mathbf{v}) \right) \mathbf{d}\mathbf{v}.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

LEMMA 6.1 [23, Lemma 2.7]. *Suppose* $\mathbf{u}$ *has coefficients in* $\mathbb{Z}$, *and that* $\mathcal{B}_1(\mathbf{u}_1)$ *is sufficiently large with respect to* $r_d, \ldots, r_1$, *and* d. *Furthermore, suppose* $\gamma_2, \ldots, \gamma_d$ *are sufficiently small with respect to* $r_d, \ldots, r_1$, *and* d. *Then we have*

$$\mathcal{I}(\mathfrak{B}_0, \boldsymbol{\tau}) \ll \min(1, |\boldsymbol{\tau}|^{-R-1}), \tag{6.6}$$

where the implicit constant depends at most on $d, r_d, \ldots, r_1$, *and* $\mathbf{U}$.

We refer the reader to [23] for a proof of this lemma. The proof in [23] is similar to that of [20, Lemma 8.1], which is for systems without linear polynomials. However, due to the presence of linear polynomials it requires some justification not available in [20].

We also need to deal with certain situations where the coefficients of $\mathbf{u}$ may depend on P (but not the coefficients of $\mathbf{U}$). There are essentially two different scenarios we have to consider, first of which we refer to as the following.

Condition ($\star'$): The polynomials of $\mathbf{u}$ have coefficients in $\mathbb{Z}$, and the coefficients of $\mathbf{U}$ do not depend on P. However, for each $u_{\ell,r}(\mathbf{x})$ ($1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell$) the coefficients of its monomials whose degrees are strictly less than ℓ may depend on P.

We have the following result when $\mathbf{u}$ satisfies Condition ($\star'$).

COROLLARY 6.2. *Suppose* $\mathbf{u}$ *satisfies Condition* ($\star'$). *Let* $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in* (6.1). *Suppose* $\varepsilon' > 0$ *is sufficiently small and* $Q > 0$ *satisfies*

$$Q\gamma_d' < 1.$$

Then one of the following alternatives must hold:

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{Q\gamma_d+\varepsilon'} \quad \text{and} \quad \|n_0\boldsymbol{\alpha}_d\| \ll P^{-d+Q\gamma_d+\varepsilon'}.$$

The implicit constants depend only on $n, d, r_d, \varepsilon', Q$, *and* $\mathbf{U}_d$.

Next we present the result in our second scenario for when the coefficients of $\mathbf{u}$ may depend on P. Let $u_{\ell,r}^{(j)}(\mathbf{x})$ be the homogeneous degree j portion of the polynomial $u_{\ell,r}(\mathbf{x})$. In the following corollary, for $j < \ell$ the coefficients of $u_{\ell,r}^{(j)}(\mathbf{x})$ may be in $\mathbb{Q}$ and also depend on P, but in a controlled manner. On the other hand, the coefficients of $U_{\ell,r}(\mathbf{x})$ do not depend on P.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

COROLLARY 6.3. *Suppose* $\mathbf{u}$ *has coefficients in* $\mathbb{Q}$*, and further suppose* $\mathbf{U}$ *has coefficients in* $\mathbb{Z}$*. Let* $Q > 0$ *and* $\varepsilon > 0$*. Let* $2 \leqslant \ell \leqslant d$ *with* $r_\ell > 0$*. If* $\ell = d$*, then let* $\theta = 0$ *and* $q = 1$*. On the other hand, if* $2 \leqslant \ell < d$*, then suppose* $0 \leqslant \theta < 1/4$ *and that there is* $q \in \mathbb{N}$ *with*

$$q \leqslant P^{\theta}, \quad q\boldsymbol{\alpha}_j \in \mathbb{Z}^{r_j} \ (\ell < j \leqslant d),$$

and

$$q\alpha_{\ell',r} u^{(j)}_{\ell',r}(\mathbf{x}) \in \mathbb{Z}[x_1, \ldots, x_n]$$

for every $\ell < \ell' \leqslant d, 0 \leqslant j < \ell', 1 \leqslant r \leqslant r_{\ell'}$*. Let* $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in* (6.1)*. Suppose*

$$4\theta + Q\gamma'_\ell < 1.$$

Then one of the following alternatives must hold:

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{Q\gamma_\ell+\varepsilon} \quad \text{and} \quad \|n_0 q \boldsymbol{\alpha}_\ell\| \ll P^{-\ell+4\theta+Q\gamma_\ell+\varepsilon}.$$

The implicit constants depend at most on $n, d, r_d, \ldots, r_1, Q, \varepsilon$*, and* $\mathbf{U}$*.*

We present the details of the proofs of Corollaries 6.2 and 6.3 in Appendix A.

7. Hardy–Littlewood circle method: major arcs

For $\mathbf{x} = (x_1, \ldots, x_n)$, let us denote $\widehat{\mathbf{x}} = (x_1, \ldots, x_{n-r_1})$. In this section, we consider the system of equations

$$f_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell), \tag{7.1}$$

where we assume $\mathbf{f}$ is of the shape

$$f_{\ell,r}(\mathbf{x}) = f_{\ell,r}(\widehat{\mathbf{x}}) \in \mathbb{Z}[x_1, \ldots, x_{n-r_1}] \quad (2 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell),$$

and

$$f_{1,r}(\mathbf{x}) = c_{1,r} x_{n-r_1+r} + \widetilde{f}_{1,r}(\widehat{\mathbf{x}}) \quad (1 \leqslant r \leqslant r_1),$$

where $c_{1,r} \in \mathbb{Z}\backslash\{0\}$ and $\widetilde{f}_{1,r}(\widehat{\mathbf{x}}) \in \mathbb{Z}[x_1, \ldots, x_{n-r_1}]$. We further assume $\mathbf{f}$ satisfies the following: $h_d(\mathbf{f}_d), \ldots, h_2(\mathbf{f}_2)$, and $\mathcal{B}_1(\mathbf{f}_1)$ are all sufficiently large with respect to d and $r_d, \ldots, r_1$. Clearly systems with these assumptions include the reduced system $\mathbf{f}$ in (4.2) (see property (7) which was obtained using Lemma 2.2). Let us

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

remark that in contrast to the major arcs analysis in [6], we have conditions on the h-invariant instead of the Birch rank, and these conditions on the h-invariant required in this section are 'comparable' to those in [20]. We also denote $F_{\ell,r}$ to be the homogeneous degree ℓ portion of $f_{\ell,r}$ $(1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$, and let $\mathbf{F}_\ell = (F_{\ell,1}, \ldots, F_{\ell,r_\ell})$ $(1 \leqslant \ell \leqslant d)$.

Let $\mathfrak{B}_0 = [0, 1]^n \subseteq \mathbb{R}^n$. Given $\mathbf{b} \in (\mathbb{Z}/q\mathbb{Z})^n$, we define

$$\boldsymbol{\psi}_{\mathbf{b}}(\mathbf{t}) = \psi_{b_1}(t_1) \cdots \psi_{b_n}(t_n),$$

where

$$\psi_{b_j}(t_j) = \sum_{\substack{0 \leqslant v \leqslant t_j \\ v \equiv b_j \, (\mathrm{mod}\, q)}} \Lambda(v).$$

We use the notation $\mathbf{x} \equiv \mathbf{b}$ (mod q) to mean $x_i \equiv b_i$ (mod q) for each $1 \leqslant i \leqslant n$. Suppose for $\boldsymbol{\alpha} \in [0, 1)^R$, we have $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$ where $\mathbf{a} \in (\mathbb{Z}/q\mathbb{Z})^R$. Then we have

$$\begin{aligned}
&T(\mathbf{f}; \boldsymbol{\alpha}) \\
&= \sum_{\mathbf{x} \in [0,X]^n} \Lambda(\mathbf{x}) e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} \alpha_{\ell,r} \cdot f_{\ell,r}(\mathbf{x}) \right) \\
&= \sum_{\mathbf{b} \in (\mathbb{Z}/q\mathbb{Z})^n} \sum_{\substack{\mathbf{x} \in [0,X]^n \\ \mathbf{x} \equiv \mathbf{b} (\mathrm{mod}\, q)}} \Lambda(\mathbf{x}) e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} a_{\ell,r} \cdot f_{\ell,r}(\mathbf{b})/q \right) \\
&\quad \cdot e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} \beta_{\ell,r} \cdot f_{\ell,r}(\mathbf{x}) \right) \\
&= \sum_{\mathbf{b} \in (\mathbb{Z}/q\mathbb{Z})^n} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} a_{\ell,r} \cdot f_{\ell,r}(\mathbf{b})/q \right) \\
&\quad \cdot \int_{\mathbf{t} \in X\mathfrak{B}_0} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} \beta_{\ell,r} \cdot f_{\ell,r}(\mathbf{t}) \right) \mathbf{d}\boldsymbol{\psi}_{\mathbf{b}}(\mathbf{t}),
\end{aligned} \tag{7.2}$$

where $\mathbf{d}\boldsymbol{\psi}_{\mathbf{b}}(\mathbf{t})$ denotes the product measure $d\psi_{b_1}(t_1) \times \cdots \times d\psi_{b_n}(t_n)$.

Let ϕ be Euler's totient function. For a positive integer q, recall we put $\mathbb{U}_q$ for the group of units in $\mathbb{Z}/q\mathbb{Z}$. Lemma 7.1 follows immediately from the proof of [6, Lemma 6] as the proof does not depend on the fact that the polynomials of the system all have the same degree.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

LEMMA 7.1. *Let $c > 0$, $C > 0$, $q \leqslant (\log X)^C$, and $\mathbf{b} \in (\mathbb{Z}/q\mathbb{Z})^n$. Suppose $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta} \in \mathfrak{M}_{\mathbf{a},q}(C)$. Then we have*

$$\int_{\mathbf{t}\in X\mathfrak{B}_0} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \beta_{\ell,r}\cdot f_{\ell,r}(\mathbf{t})\right) \mathbf{d}\boldsymbol{\psi}_{\mathbf{b}}(\mathbf{t})$$
$$= \mathbf{1}_{\mathbf{b}\in(\mathbb{U}_q)^n}\frac{1}{\phi(q)^n}\int_{\mathbf{v}\in X\mathfrak{B}_0} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \beta_{\ell,r}\cdot f_{\ell,r}(\mathbf{v})\right)\mathbf{dv} + O(X^n/(\log X)^c),$$

where $\mathbf{1}_{\mathbf{b}\in(\mathbb{U}_q)^n}$ is 1 if $\mathbf{b} \in (\mathbb{U}_q)^n$ and 0 otherwise.

Let $\varepsilon > 0$. We simplify the above integral by a change of variable as follows

$$\begin{aligned}\int_{\mathbf{v}\in X\mathfrak{B}_0} & e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \beta_{\ell,r}\cdot f_{\ell,r}(\mathbf{v})\right)\mathbf{dv}\\ &= \int_{\mathbf{v}\in X\mathfrak{B}_0} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \beta_{\ell,r}\cdot F_{\ell,r}(\mathbf{v})\right)\mathbf{dv} + O(X^{n-1+\varepsilon})\\ &= X^n\mathcal{I}(\mathfrak{B}_0,\boldsymbol{\beta}') + O(X^{n-1+\varepsilon}),\end{aligned} \tag{7.3}$$

where

$$\beta'_{\ell,r} = X^\ell \beta_{\ell,r} \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell),$$

and

$$\mathcal{I}(\mathfrak{B}_0,\boldsymbol{\tau}) = \int_{\mathbf{v}\in\mathfrak{B}_0} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \tau_{\ell,r}\cdot F_{\ell,r}(\mathbf{v})\right)\mathbf{dv}.$$

We define

$$J(L) = \int_{|\boldsymbol{\tau}|\leqslant L} \mathcal{I}(\mathfrak{B}_0,\boldsymbol{\tau})\,\mathbf{d}\boldsymbol{\tau}.$$

By our assumptions on $\mathbf{f}$ and (6.4), we know we can apply Lemma 6.1 and obtain $\mathcal{I}(\mathfrak{B}_0,\boldsymbol{\tau}) \ll \min(1, |\boldsymbol{\tau}|^{-R-1})$. With this estimate, it is an easy exercise to show that

$$\mu(\infty) = \int_{\boldsymbol{\tau}\in\mathbb{R}^R} \mathcal{I}(\mathfrak{B}_0,\boldsymbol{\tau})\,\mathbf{d}\boldsymbol{\tau}, \tag{7.4}$$

which is called the *singular integral*, exists, and that

$$|\mu(\infty) - J(L)| \ll L^{-1}. \tag{7.5}$$

We note that $\mu(\infty)$ is the same as what is defined in [3, (2.3)], and we have

$$\mu(\infty) > 0 \tag{7.6}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

provided that the system of equations

$$F_{\ell,r}(\mathbf{x}) = 0 \quad (1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$$

has a nonsingular real solution in $(0, 1)^n$. The argument used to show this fact is standard and we refer the reader to see for example [7, Ch. 16], or the explanation in [3].

We define the following sums:

$$\mathcal{S}_{\mathbf{a},q} = \sum_{\mathbf{k}\in(\mathbb{U}_q)^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k}) \cdot a_{\ell,r}/q\right), \tag{7.7}$$

$$B(q) = \sum_{\substack{\gcd(\mathbf{a},q)=1 \\ \mathbf{a}\in(\mathbb{Z}/q\mathbb{Z})^R}} \frac{1}{\phi(q)^n}\mathcal{S}_{\mathbf{a},q},$$

and

$$\mathfrak{S}(X) = \sum_{q\leqslant(\log X)^C} B(q). \tag{7.8}$$

By combining Lemma 7.1 with the definitions given above, we have the following.

LEMMA 7.2 [6, Lemma 8]. *Given any $c > 0$, $C > 0$, and $q \leqslant (\log X)^C$, we have*

$$\int_{\mathfrak{M}_{\mathbf{a},q}(C)} T(\mathbf{f};\boldsymbol{\alpha})\,\mathbf{d}\boldsymbol{\alpha} = \frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{\phi(q)^n}\mathcal{S}_{\mathbf{a},q}J((\log X)^C) + O\left(\frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log X)^c}\right).$$

Therefore, we obtain the following estimate as a consequence of the definition of the major arcs, (7.5), and Lemma 7.2.

LEMMA 7.3. *Given any $c > 0$ and $C > 0$, we have*

$$\int_{\mathfrak{M}(C)} T(\mathbf{f};\boldsymbol{\alpha})\,\mathbf{d}\boldsymbol{\alpha} = \mathfrak{S}(X)\mu(\infty)X^{n-\sum_{\ell=1}^{d}\ell r_\ell} + O\left(\mathfrak{S}(X)\frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log X)^C} + \frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log X)^c}\right).$$

We still have to deal with the term $\mathfrak{S}(X)$, and this is done in the following section.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

7.1. Singular series. In order to estimate the term $\mathfrak{S}(X)$, we begin by obtaining estimates for the exponential sum $\mathcal{S}_{\boldsymbol{a},q}$ defined in (7.7). We define $g_\ell(\mathbf{F}_\ell)$ as in (6.3). It then follows from (6.4) that

$$h_\ell(\mathbf{F}_\ell) < (\log 2)^{-\ell} \cdot \ell! \cdot (g_\ell(\mathbf{F}_\ell) + (\ell-1)r_\ell(r_\ell - 1))$$

for $2 \leqslant \ell \leqslant d$ with $r_\ell > 0$. From this inequality, for $2 \leqslant \ell \leqslant d$ with $r_\ell > 0$ we see that we can assume $g_\ell(\mathbf{F}_\ell)$ to be sufficiently large with respect to d and $r_d, \ldots, r_1$.

Let

$$Q = 1 + \max\left\{\frac{1 + R(800d^3+2)}{800d^3+1}, \frac{R+1}{1 - 1/(800d^3+1)}\right\}.$$

With our assumptions in this section, Q satisfies the following,

$$4\left(\gamma_2 Q + \gamma_3 Q + \cdots + \gamma_d Q + \frac{1}{800d}\right) < \frac{1}{100d},$$

$$Q \cdot r_\ell(\ell-1) \cdot 2^{\ell-1}\left(\frac{(\log 2)^\ell (h_\ell(\mathbf{F}_\ell) - (800d^3+1)Q)}{\ell!} - (\ell-1)r_\ell(r_\ell-1)\right)^{-1}$$
$$< \frac{1}{1600d^3+2} \quad (2 \leqslant \ell \leqslant d), \tag{7.9}$$

and

$$0 < Q < \frac{d-1}{d(r_1+1)}(\gamma_2 + 4\gamma_3 + \cdots + 4^{d-2}\gamma_d)^{-1}, \tag{7.10}$$

where γ_ℓ is defined (with respect to $\mathbf{F}_\ell$ here) after (6.4). We fix this value of Q throughout the remainder of this section. Also since $\mathcal{B}_1(\mathbf{F}_1)$ is sufficiently large with respect to d and $r_d, \ldots, r_1$, we have $\mathcal{B}_1(\mathbf{F}_1) > Q$.

We consider two cases depending on $\mathbf{a}$ to bound $\mathcal{S}_{\boldsymbol{a},q}$ when q is a prime power. These cases are treated separately in Lemmas 7.4 and 7.5.

LEMMA 7.4. *Let p be a prime and let $q = p^t$, $t \in \mathbb{N}$. Let $\mathbf{a} = (\mathbf{a}_d, \ldots, \mathbf{a}_1) \in (\mathbb{Z}/q\mathbb{Z})^R$ with $\gcd(\mathbf{a}, q) = 1$. Furthermore, suppose there exists $\ell \in \{2, \ldots, d\}$ such that $\gcd(\mathbf{a}_\ell, q) = 1$. Then we have the following bounds*

$$\mathcal{S}_{\boldsymbol{a},q} \ll \begin{cases} q^{n-Q} & \text{if } t \leqslant 800d^3+1, \\ p^Q q^{n-Q} & \text{if } t > 800d^3+1, \end{cases}$$

where the implicit constants are independent of p.

Proof. We consider the two cases $t \leqslant 800d^3+1$ and $t > 800d^3+1$ separately. We begin with the case $t \leqslant 800d^3+1$. In this case, we apply the inclusion–exclusion principle to $\mathcal{S}_{\mathbf{a},q}$. As a result, we obtain

$$\begin{aligned}
\mathcal{S}_{\mathbf{a},q} &= \sum_{\mathbf{k}\in(\mathbb{U}_q)^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\\
&= \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n}\prod_{i=1}^{n}\left(1-\sum_{v_i\in\mathbb{Z}/p^{t-1}\mathbb{Z}}\mathbf{1}_{k_i=pv_i}\right) e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\\
&= \sum_{I\subseteq\{1,2,\ldots,n\}}(-1)^{|I|}\sum_{\mathbf{v}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}}\sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n}\mathfrak{F}_I(\mathbf{k};\mathbf{v})e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right),
\end{aligned}\tag{7.11}$$

where

$$\mathbf{1}_{k_i=pv_i} = \begin{cases} 1 & \text{if } k_i = pv_i,\\ 0 & \text{if } k_i \neq pv_i,\end{cases}$$

and

$$\mathfrak{F}_I(\mathbf{k};\mathbf{v}) = \prod_{i\in I}\mathbf{1}_{k_i=pv_i}$$

for $\mathbf{v}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}$. In other words, $\mathfrak{F}_I(\mathbf{k};\mathbf{v})$ is the characteristic function of the set $H_{I,\mathbf{v}} = \{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n : k_i = pv_i\ (i\in I)\}$. We now bound the summand in the final expression of (7.11) by further considering two cases, $|I|\geqslant tQ$ and $|I|<tQ$. In the first case $|I|\geqslant tQ$, we use the following trivial estimate

$$\begin{aligned}
\left|\sum_{\mathbf{v}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}}\sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n}\mathfrak{F}_I(\mathbf{k};\mathbf{v})e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\right| &\leqslant p^{(t-1)|I|}(p^t)^{n-|I|}\\
&= q^{n-|I|/t}\\
&\leqslant q^{n-Q}.
\end{aligned}$$

On the other hand, suppose $|I|<tQ$. Let us label $\mathbf{s}=(s_1,\ldots,s_{n-|I|})$ to be the remaining variables of $\mathbf{x}$ after setting $x_i=0$ for each $i\in I$. For each $1\leqslant\ell\leqslant d$, $1\leqslant r\leqslant r_\ell$, let

$$\mathfrak{g}_{\ell,r}(\mathbf{s}) = f_{\ell,r}(\mathbf{x})|_{x_i=pv_i\,(i\in I)},$$

or equivalently the polynomial $\mathfrak{g}_{\ell,r}(\mathbf{s})$ is obtained by substituting $x_i=pv_i$ $(i\in I)$ to the polynomial $f_{\ell,r}(\mathbf{x})$. Thus $\mathfrak{g}_{\ell,r}(\mathbf{s})$ is a polynomial in $n-|I|$ variables whose coefficients may depend on p. With these notations we have

$$\begin{aligned}\sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} \mathfrak{F}_I(\mathbf{k};\mathbf{v}) e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right) &= \sum_{\mathbf{s}\in(\mathbb{Z}/q\mathbb{Z})^{n-|I|}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s})\cdot a_{\ell,r}/q\right)\\ &= \sum_{\mathbf{s}\in[0,q-1]^{n-|I|}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s})\cdot a_{\ell,r}/q\right).\end{aligned}$$

We can also deduce easily that the homogeneous degree ℓ portion of the polynomial $\mathfrak{g}_{\ell,r}(\mathbf{s})$, which we denote $\mathfrak{G}_{\ell,r}(\mathbf{s})$, is obtained by substituting $x_i = 0$ $(i \in I)$ to $F_{\ell,r}(\mathbf{x})$. Hence, we have

$$\mathfrak{G}_{\ell,r}(\mathbf{s}) = F_{\ell,r}(\mathbf{x})|_{x_i=0\ (i\in I)},$$

and in particular, it is independent of p. Thus the system of polynomials $\mathfrak{g}_{\ell,r}(\mathbf{s})$ $(1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$ satisfies Condition $(\star')$. It also follows by Lemma 2.1 that

$$h_\ell(\{\mathfrak{G}_{\ell,r} : 1 \leqslant r \leqslant r_\ell\}) \geqslant h_\ell(\mathbf{F}_\ell) - |I| > h_\ell(\mathbf{F}_\ell) - (800d^3+1)Q \quad (2 \leqslant \ell \leqslant d).$$

By our choice of Q, namely (7.9), and from (6.4), we have

$$Q\gamma'_\ell \leqslant Q\gamma_\ell < \frac{1}{1600d^3+2} < 1 \quad (2 \leqslant \ell \leqslant d),$$

where γ'_ℓ and γ_ℓ are defined with respect to $\{\mathfrak{G}_{\ell,r} : 1 \leqslant r \leqslant r_\ell\}$ here.

Take $\varepsilon > 0$ sufficiently small. Let us suppose that p and t are sufficiently large with respect to the coefficients of $\mathbf{F}$, $n, d, r_d, \ldots, r_1, \varepsilon$, and Q, which implies that q is sufficiently large with respect to the coefficients of $\mathfrak{G}_{\ell,r_\ell}(\mathbf{s})$ $(1 \leqslant \ell \leqslant d$, $1 \leqslant r \leqslant r_\ell)$. Suppose we have

$$\sum_{\mathbf{s}\in[0,q-1]^{n-|I|}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s})\cdot a_{\ell,r}/q\right) > (q-1)^{n-|I|-Q}. \tag{7.12}$$

Then by Corollary 6.2 there must exist $n_0 \in \mathbb{N}$ such that

$$n_0 \ll q^{Q\gamma_d+\varepsilon} \quad \text{and} \quad \|n_0(\mathbf{a}_d/q)\| \ll q^{-d+Q\gamma_d+\varepsilon}.$$

Since p and t are sufficiently large, we have $n_0 < q^{1/(1600d^3+2)}$ and $\|n_0(\mathbf{a}_d/q)\| < q^{-d+1/(1600d^3+2)}$, because $Q\gamma_d + \varepsilon < 1/(1600d^3+2)$. Then it follows that

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$n_0 < p^{t/(1600d^3+2)} < p$. Suppose now that not all entries of $n_0\mathbf{a}_d$ are divisible by q. In this case, we obtain

$$\frac{1}{q} \leqslant \|n_0(\mathbf{a}_d/q)\| < \frac{1}{q^{d-1/(1600d^3+2)}},$$

which is a contradiction. Thus all of the entries of $n_0\mathbf{a}_d$ must be divisible by $q = p^t$ and since $\gcd(n_0, p) = 1$, it follows that all of the entries of $\mathbf{a}_d$ must be divisible by q. Therefore, we can simplify the exponential sum in consideration since $e(m) = 1$ when $m \in \mathbb{Z}$, and the inequality (7.12) becomes

$$\sum_{\mathbf{s}\in[0,q-1]^{n-|I|}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s}) \cdot a_{\ell,r}/q\right) = \sum_{\mathbf{s}\in[0,q-1]^{n-|I|}} e\left(\sum_{\ell=1}^{d-1}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s}) \cdot a_{\ell,r}/q\right) > (q-1)^{n-|I|-Q}. \tag{7.13}$$

We may repeat the argument as above, because we are now dealing with the system of polynomials $\mathfrak{g}_{\ell,r}(\mathbf{s})$ $(1 \leqslant \ell \leqslant d-1, 1 \leqslant r \leqslant r_\ell)$. Again by Corollary 6.2 there must exist $n_0' \in \mathbb{N}$ such that

$$n_0' \ll q^{Q\gamma_{d-1}+\varepsilon} \quad \text{and} \quad \|n_0'(\mathbf{a}_{d-1}/q)\| \ll q^{-(d-1)+Q\gamma_{d-1}+\varepsilon}.$$

Since p and t are sufficiently large, we have $n_0' < q^{1/(1600d^3+2)}$ and $\|n_0'(\mathbf{a}_{d-1}/q)\| < q^{-(d-1)+1/(1600d^3+2)}$, because $Q\gamma_{d-1} + \varepsilon < 1/(1600d^3 + 2)$. Then it follows that $n_0' < p^{t/(1600d^3+2)} < p$. Suppose now that not all entries of $n_0'\mathbf{a}_{d-1}$ are divisible by q. In this case, we obtain

$$\frac{1}{q} \leqslant \|n_0'(\mathbf{a}_{d-1}/q)\| < \frac{1}{q^{(d-1)-1/(1600d^3+2)}},$$

which is a contradiction. Thus all of the entries of $n_0'\mathbf{a}_{d-1}$ must be divisible by $q = p^t$ and since $\gcd(n_0', p) = 1$, it follows that all of the entries of $\mathbf{a}_{d-1}$ must be divisible by q. It is then clear that we can repeat the argument and keep reducing until we obtain that all of the entries of $\mathbf{a}_\ell$ must be divisible by q for each $2 \leqslant \ell \leqslant d$. We remark that if there exists ℓ' with $r_{\ell'} = 0$, then we simply skip the case $\ell = \ell'$ during this process. Thus we have $\gcd(\mathbf{a}_\ell, q) > 1$ for each $2 \leqslant \ell \leqslant d$, which is a contradiction. As a result we must have

$$\sum_{\mathbf{s}\in[0,q-1]^{n-|I|}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \mathfrak{g}_{\ell,r}(\mathbf{s}) \cdot a_{\ell,r}/q\right) \ll q^{n-|I|-Q}.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Thus we obtain

$$\begin{aligned}\sum_{\mathbf{v}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}}\ \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n}\mathfrak{F}_I(\mathbf{k};\mathbf{v})e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\\ \ll (p^{t-1})^{|I|}q^{n-|I|-Q}\\ \leqslant q^{n-Q}.\end{aligned}$$

Consequently, by combining the estimates for the two cases $|I|\geqslant tQ$ and $|I|<tQ$, we obtain

$$\mathcal{S}_{\mathbf{a},q}\ll q^{n-Q}$$

when $t\leqslant 800d^2+1$.

We now consider the case $t>800d^3+1$. By the definition of $\mathcal{S}_{\mathbf{a},q}$, we have

$$\begin{aligned}\mathcal{S}_{\mathbf{a},q}&=\sum_{\mathbf{k}\in(\mathbb{U}_q)^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\\ &=\sum_{\mathbf{k}\in(\mathbb{U}_p)^n}\ \sum_{\mathbf{y}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k}+p\mathbf{y})\cdot a_{\ell,r}/q\right)\\ &=\sum_{\mathbf{k}\in(\mathbb{U}_p)^n}\ \sum_{\mathbf{y}\in[0,p^{t-1}-1]^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k}+p\mathbf{y})\cdot a_{\ell,r}/q\right).\end{aligned}\tag{7.14}$$

For each fixed $\mathbf{k}\in(\mathbb{U}_p)^n$, we have

$$f_{\ell,r}(\mathbf{k}+p\mathbf{y})=p^\ell F_{\ell,r}(\mathbf{y})+\omega_{\ell,r:p,\mathbf{k}}(\mathbf{y})\quad(1\leqslant\ell\leqslant d,\ 1\leqslant r\leqslant r_\ell),$$

where $\omega_{\ell,r:p,\mathbf{k}}(\mathbf{y})$ is a polynomial in $\mathbf{y}$ of degree at most $\ell-1$ and its coefficients are integers which may depend on p and $\mathbf{k}$. We let

$$u_{\ell,r}(\mathbf{y})=F_{\ell,r}(\mathbf{y})+\frac{1}{p^\ell}\omega_{\ell,r:p,\mathbf{k}}(\mathbf{y})\quad(1\leqslant\ell\leqslant d,\ 1\leqslant r\leqslant r_\ell).$$

We can then express the inner sum of the last expression of (7.14) as

$$\sum_{\mathbf{y}\in[0,p^{t-1}-1]^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} u_{\ell,r}(\mathbf{y})\cdot a_{\ell,r}/(q/p^\ell)\right).$$

We have that $\mathbf{u}$ has coefficients in $\mathbb{Q}$, and $\mathbf{U}$ has coefficients in $\mathbb{Z}$. Let $\alpha_{\ell,r}=a_{\ell,r}/p^{t-\ell}$ $(1\leqslant\ell\leqslant d,\ 1\leqslant r\leqslant r_\ell)$, and $P=(p^{t-1}-1)$.

Recall we have set Q to satisfy

$$4\left(\gamma_2Q+\gamma_3Q+\cdots+\gamma_dQ+\frac{1}{800d}\right)<\frac{1}{100d},$$

where γ_ℓ is defined with respect to $\mathbf{F}_\ell$ here. Suppose we have

$$\left|\sum_{\mathbf{y}\in[0,P]^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} u_{\ell,r}(\mathbf{y})\cdot\alpha_{\ell,r}\right)\right| > P^{n-Q}.$$

Then by Corollary 6.3, there must exist $n_d \in \mathbb{N}$ such that

$$n_d \ll P^{Q\gamma_d+\varepsilon} \quad \text{and} \quad \|n_d\boldsymbol{\alpha}_d\| \ll P^{-d+Q\gamma_d+\varepsilon}.$$

For $P = p^{t-1} - 1$ sufficiently large, we have

$$n_d < p^{(t-1)/(100d)} \quad \text{and} \quad \|n_d(\mathbf{a}_d/p^{t-d})\| < p^{(t-1)(-d+1/(100d))},$$

because $Q\gamma_d + \varepsilon < 1/(100d)$. Suppose now that not all entries of $n_d\mathbf{a}_d$ are divisible by p^{t-d}. In this case, we obtain

$$\frac{1}{p^{t-d}} \leqslant \|n_d(\mathbf{a}_d/p^{t-d})\| < \frac{1}{p^{(t-1)(d-1/(100d))}},$$

which is a contradiction. Thus all of the entries of $n_d\mathbf{a}_d$ must be divisible by p^{t-d}. In particular,

$$n_d\boldsymbol{\alpha}_d = n_d(\mathbf{a}_d/p^{t-d}) \in \mathbb{Z}^{r_d},$$

and we can assume without loss of generality that n_d is a power of p satisfying $n_d < p^{(t-1)/(100d)}$. Since $t-d > (t-1)/(100d)$, it follows that every entry of $\mathbf{a}_d$ is divisible by p.

From the inequality $t > 800d^3 + 1$, we have $p^d < p^{(t-1)/(800d^2)}$. Thus we have

$$(n_d p^d)\alpha_{d,r}\frac{1}{p^d}\omega_{d,r:p,\mathbf{k}}(\mathbf{y}) \in \mathbb{Z}[y_1,\ldots,y_n] \quad (1 \leqslant r \leqslant r_d),$$

and

$$n_d p^d \leqslant P^{Q\gamma_d+2\varepsilon}P^{1/(800d^2)}.$$

With this setup, we can apply Corollary 6.3 again with $\ell = d-1$ and

$$\theta = Q\gamma_d + \frac{1}{800d^2} + \varepsilon_d < \frac{1}{100d} < \frac{1}{4},$$

where $\varepsilon_d > 0$ is sufficiently small, and deduce that there must exist $n_{d-1} \in \mathbb{N}$ such that

$$n_{d-1} \ll P^{Q\gamma_{d-1}+\varepsilon} \quad \text{and} \quad \|n_{d-1}n_d p^d\boldsymbol{\alpha}_{d-1}\| \ll P^{-(d-1)+4\theta+Q\gamma_{d-1}+\varepsilon}.$$

For $P = p^{t-1} - 1$ sufficiently large, we have

$$n_{d-1} < p^{(t-1)/(100d)} \quad \text{and} \quad \|n_{d-1}n_d p^d(\mathbf{a}_{d-1}/p^{t-(d-1)})\| < p^{(t-1)(-(d-1)+1/(100d))},$$

because

$$Q\gamma_{d-1} + 4\theta + \varepsilon = Q\gamma_{d-1} + 4\left(Q\gamma_d + \frac{1}{800d^2} + \varepsilon_d\right) + \varepsilon < \frac{1}{100d}.$$

Suppose now that not all entries of $(n_{d-1}n_d\ p^d\mathbf{a}_{d-1})$ are divisible by $p^{t-(d-1)}$. In this case, we obtain

$$\frac{1}{p^{t-(d-1)}} \leqslant \|n_{d-1}n_d p^d(\mathbf{a}_{d-1}/p^{t-(d-1)})\| < \frac{1}{p^{(t-1)((d-1)-1/(100d))}},$$

which is a contradiction. Thus all of the entries of $(n_{d-1}n_d p^d\mathbf{a}_{d-1})$ must be divisible by $p^{t-(d-1)}$. In particular,

$$n_{d-1}n_d p^d\boldsymbol{\alpha}_{d-1} = n_{d-1}n_d p^d(\mathbf{a}_{d-1}/p^{t-(d-1)}) \in \mathbb{Z}^{r_{d-1}},$$

and we can assume without loss of generality that n_{d-1} is a power of p satisfying $n_{d-1} < p^{(t-1)/(100d)}$. Since $t-(d-1) > 2(t-1)/(100d)+d$, it follows that every entry of $\mathbf{a}_{d-1}$ is divisible by p.

We have

$$(n_{d-1}n_d p^{d+(d-1)})\alpha_{d,r}\frac{1}{p^d}\omega_{d,r:p,\mathbf{k}}(\mathbf{y}) \in \mathbb{Z}[y_1,\dots,y_n] \quad (1 \leqslant r \leqslant r_d),$$
$$(n_{d-1}n_d p^{d+(d-1)})\alpha_{d-1,r}\frac{1}{p^{d-1}}\omega_{d-1,r:p,\mathbf{k}}(\mathbf{y}) \in \mathbb{Z}[y_1,\dots,y_n] \quad (1 \leqslant r \leqslant r_{d-1}),$$

and

$$n_{d-1}n_d p^{d+(d-1)} \leqslant P^{Q\gamma_{d-1}+2\varepsilon}P^{Q\gamma_d}P^{2/(800d^2)}.$$

With this setup, we can apply Corollary 6.3 again with $\ell = d-2$ and

$$\begin{aligned}\theta &= Q\gamma_{d-1} + \varepsilon_{d-1} + Q\gamma_d + \frac{2}{800d^2}\\ &< Q\gamma_{d-1} + \varepsilon_{d-1} + Q\gamma_d + \frac{d}{800d^2}\\ &< \frac{1}{100d}\\ &< \frac{1}{4},\end{aligned} \tag{7.15}$$

where $\varepsilon_{d-1} > 0$ is sufficiently small. At this point it is clear that we can repeat the process, in fact we continue in this manner until $\ell = 2$. We remark that if there exists ℓ' with $r_{\ell'} = 0$, then we simply skip the case $\ell = \ell'$ during this process. As a result, we obtain that every entry of $\mathbf{a}_d, \ldots, \mathbf{a}_2$ is divisible by p. Thus we have $\gcd(\mathbf{a}_\ell, q) > 1$ for each $2 \leqslant \ell \leqslant d$, which is a contradiction. Therefore, we obtain

$$\begin{aligned}\left| \sum_{\mathbf{y} \in [0, p^{t-1}-1]^n} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k} + p\mathbf{y}) \cdot a_{\ell,r}/q \right) \right| &= \left| \sum_{\mathbf{y} \in [0, P]^n} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} u_{\ell,r}(\mathbf{y}) \cdot \alpha_{\ell,r} \right) \right| \\ &\ll P^{n-Q} \\ &\ll (p^{t-1})^{n-Q}.\end{aligned}$$

Thus we can bound (7.14) as follows

$$\begin{aligned}\mathcal{S}_{\mathbf{a},q} &\leqslant \sum_{\mathbf{k} \in \mathbb{U}_p^n} \left| \sum_{\mathbf{y} \in [0, p^{t-1}-1]^n} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k} + p\mathbf{y}) \cdot a_{\ell,r}/q \right) \right| \\ &\ll p^n (p^{t-1})^{n-Q} \\ &= p^Q q^{n-Q}. \qquad \square\end{aligned}$$

LEMMA 7.5. *Let p be a prime and let $q = p^t$, $t \in \mathbb{N}$. Let $\mathbf{a} = (\mathbf{a}_d, \ldots, \mathbf{a}_1) \in (\mathbb{Z}/q\mathbb{Z})^R$ with $\gcd(\mathbf{a}, q) = 1$. Furthermore, suppose $\gcd(\mathbf{a}_\ell, q) > 1$ for $2 \leqslant \ell \leqslant d$, and $\gcd(\mathbf{a}_1, q) = 1$. Then we have*

$$\mathcal{S}_{\mathbf{a},q} \ll \begin{cases} q^{n-Q} & \text{if } t \leqslant 800d^3 + 1, \\ p^Q q^{n-Q} & \text{if } t > 800d^3 + 1, \end{cases}$$

where the implicit constants depend only on n and the coefficients of $\mathbf{F}_1$, and in particular they are independent of p.

Proof. First we consider the case $t > 1$. Since $\gcd(\mathbf{a}_1, q) = 1$, there exists $1 \leqslant r' \leqslant r_1$ such that $\gcd(a_{1,r'}, p) = 1$. By our assumption on $\mathbf{f}$, we have

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$$\begin{aligned}\mathcal{S}_{\mathbf{a},q} &= \sum_{\mathbf{k}\in(\mathbb{U}_q)^n} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/q\right)\\ &= \left(\prod_{1\leqslant r\leqslant r_1}\sum_{k_{n-r_1+r}\in\mathbb{U}_q} e(c_{1,r}k_{n-r_1+r}\cdot a_{1,r}/q)\right)\sum_{\widehat{\mathbf{k}}\in(\mathbb{U}_q)^{n-r_1}} e\left(\sum_{1\leqslant r\leqslant r_1}\widetilde{f}_{1,r}(\widehat{\mathbf{k}})\cdot a_{1,r}/q\right.\\ &\quad\left.+\sum_{\ell=2}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\widehat{\mathbf{k}})\cdot a_{\ell,r}/q\right).\end{aligned}$$

If $\gcd(p, c_{1,r'}) = 1$, then $\gcd(a_{1,r'}\cdot c_{1,r'},\ q) = 1$. Consequently, we have

$$\sum_{k_{n-r_1+r'}\in\mathbb{U}_q} e((a_{1,r'}\cdot c_{1,r'})k_{n-r_1+r'}/q) = \sum_{k\in\mathbb{U}_q} e(k/q) = 0,$$

because $t > 1$. In this case, it follows that

$$\mathcal{S}_{\mathbf{a},q} = 0.$$

Otherwise, we have $p|c_{1,r'}$. Let $c_{1,r'} = p^{i_0}m_0$ where $p\nmid m_0$. By a similar argument, we have

$$\sum_{k_{n-r_1+r'}\in\mathbb{U}_q} e((a_{1,r'}\cdot c_{1,r'})k_{n-r_1+r'}/q) = \begin{cases}0 & \text{if } t > i_0+1,\\ -p^{i_0} & \text{if } t = i_0+1,\\ \phi(q) & \text{if } t\leqslant i_0.\end{cases}$$

Therefore, for all but finite possibilities (depending only on $c_{1,1},\ldots,c_{1,r_1}$) of q we always have $\mathcal{S}_{\mathbf{a},q} = 0$. Thus for $t > 1$ we see that we can obtain the bounds in the statement of the lemma with the implicit constant depending only on $c_{1,1}$, $\ldots, c_{1,r_1}$.

In the case $t = 1$, since $e(m) = 1$ for $m\in\mathbb{Z}$, we have by our hypothesis that

$$\mathcal{S}_{\mathbf{a},p} = \sum_{\mathbf{k}\in(\mathbb{U}_p)^n} e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right).$$

We bound this sum in a similar manner as in Lemma 7.4. We apply the inclusion–exclusion principle and obtain

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$$\begin{aligned}
&\sum_{\mathbf{k}\in(\mathbb{U}_p)^n} e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right)\\
&\quad=\sum_{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n}\prod_{i=1}^{n}(1-\mathbf{1}_{k_i=0})\,e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right)\\
&\quad=\sum_{I\subseteq\{1,2,\dots,n\}}(-1)^{|I|}\sum_{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n}\mathfrak{F}_I(\mathbf{k})e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right), \qquad (7.16)
\end{aligned}$$

where

$$\mathbf{1}_{k_i=0}=\begin{cases}1 & \text{if } k_i=0,\\ 0 & \text{if } k_i\neq 0,\end{cases}$$

and

$$\mathfrak{F}_I(\mathbf{k})=\prod_{i\in I}\mathbf{1}_{k_i=0}.$$

In other words, $\mathfrak{F}_I(\mathbf{k})$ is the characteristic function of the set $H_I=\{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n : k_i=0\ (i\in I)\}$. We now bound the summand in the final expression of (7.16) by further considering two cases, $|I|\geqslant Q$ and $|I|<Q$. In the first case $|I|\geqslant Q$, we use the following trivial estimate

$$\left|\sum_{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n}\mathfrak{F}_I(\mathbf{k})e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right)\right|\leqslant p^{n-|I|}\leqslant p^{n-Q}.$$

On the other hand, suppose $|I|<Q$. Let us label $\mathbf{s}=(s_1,\dots,s_{n-|I|})$ to be the remaining variables of $\mathbf{x}$ after setting $x_i=0$ for each $i\in I$. For each $1\leqslant r\leqslant r_1$, let

$$\mathfrak{g}_{1,r}(\mathbf{s})=f_{1,r}(\mathbf{x})|_{x_i=0\ (i\in I)},$$

or equivalently the polynomial $\mathfrak{g}_{1,r}(\mathbf{s})$ is obtained by substituting $x_i=0\ (i\in I)$ to the polynomial $f_{1,r}(\mathbf{x})$. Thus $\mathfrak{g}_{1,r}(\mathbf{s})$ is a polynomial in $n-|I|$ variables. Let us denote

$$\mathfrak{g}_{1,r}(\mathbf{s})=\sum_{i=1}^{n-|I|}c'_{r,i}s_i+c'_{r,0}\quad (1\leqslant r\leqslant r_1).$$

With these notations, we have

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$$\begin{aligned}
&\sum_{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n} \mathfrak{F}_I(\mathbf{k}) e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right)\\
&= \sum_{\mathbf{s}\in[0,p-1]^{n-|I|}} e\left(\sum_{r=1}^{r_1} \mathfrak{g}_{1,r}(\mathbf{s})\cdot a_{1,r}/p\right)\\
&= e\left(\sum_{r=1}^{r_1} c'_{r,0}\, a_{1,r}/p\right) \sum_{\mathbf{s}\in[0,p-1]^{n-|I|}} e\left(\sum_{r=1}^{r_1}\sum_{i=1}^{n-|I|} c'_{r,i} s_i \cdot a_{1,r}/p\right)\\
&= e\left(\sum_{r=1}^{r_1} c'_{r,0}\, a_{1,r}/p\right) \sum_{\mathbf{s}\in[0,p-1]^{n-|I|}} e\left(\sum_{i=1}^{n-|I|} \frac{s_i}{p}\left(\sum_{r=1}^{r_1} c'_{r,i}\cdot a_{1,r}\right)\right). \qquad (7.17)
\end{aligned}$$

We can also deduce easily that the homogeneous linear portion of the polynomial $\mathfrak{g}_{1,r}(\mathbf{s})$, which we denote $\mathfrak{G}_{1,r}(\mathbf{s}) = \sum_{i=1}^{n-|I|} c'_{r,i}\, s_i$, is obtained by substituting $x_i = 0$ $(i \in I)$ to $F_{1,r}(\mathbf{x})$. Hence, we have

$$\mathfrak{G}_{1,r}(\mathbf{s}) = F_{1,r}(\mathbf{x})|_{x_i=0\ (i\in I)}.$$

It then follows by Lemma 3.2 that

$$\mathcal{B}_1(\{\mathfrak{G}_{1,r} : 1 \leqslant r \leqslant r_1\}) \geqslant \mathcal{B}_1(\mathbf{F}_1) - |I| > \mathcal{B}_1(\mathbf{F}_1) - Q > 0.$$

In particular, it follows that $\mathfrak{G}_{1,1}(\mathbf{s}), \ldots, \mathfrak{G}_{1,r_1}(\mathbf{s})$ are linearly independent over $\mathbb{Q}$. Thus for p sufficiently large with respect to the coefficients of $\mathbf{F}_1$, the coefficient matrix of $\mathfrak{G}_{1,1}(\mathbf{s}), \ldots, \mathfrak{G}_{1,r_1}(\mathbf{s})$ has full rank modulo p. Therefore, it follows that if

$$\sum_{r=1}^{r_1} c'_{r,i}\cdot a_{1,r} \equiv 0 \pmod{p}$$

for each $1 \leqslant i \leqslant n-|I|$, then it must be that $a_{1,1} \equiv \cdots \equiv a_{1,r_1} \equiv 0 \pmod{p}$. Since we have $\gcd(\mathbf{a}_1, p) = 1$, this is a contradiction. Thus without loss of generality suppose

$$\varsigma = \sum_{r=1}^{r_1} c'_{r,1}\cdot a_{1,r} \not\equiv 0 \pmod{p}.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Then equation (7.17) becomes

$$\begin{aligned}
&\left|\sum_{\mathbf{k}\in(\mathbb{Z}/p\mathbb{Z})^n} \mathfrak{F}_I(\mathbf{k}) e\left(\sum_{r=1}^{r_1} f_{1,r}(\mathbf{k})\cdot a_{1,r}/p\right)\right| \\
&= \left|\left(\sum_{0\leqslant s_1\leqslant p-1} e(\varsigma s_1/p)\right) \sum_{\substack{0\leqslant s_i\leqslant p-1\\ 2\leqslant i\leqslant n-|I|}} e\left(\sum_{i=2}^{n-|I|} \frac{s_i}{p}\left(\sum_{r=1}^{r_1} c'_{r,i}\cdot a_{1,r}\right)\right)\right| \\
&= 0,
\end{aligned}$$

because

$$\sum_{0\leqslant s_1\leqslant p-1} e(\varsigma s_1/p) = \sum_{0\leqslant s_1\leqslant p-1} e(s_1/p) = 0.$$

Therefore, by combining the estimates for the two cases $|I|\geqslant Q$ and $|I|<Q$, we obtain

$$\mathcal{S}_{\mathbf{a},p} \ll p^{n-Q},$$

where the implicit constant depends only on n and the coefficients of $\mathbf{F}_1$. □

By a similar argument as in [14, Ch. VIII, Section 2, Lemma 8.1], one can show that $B(q)$ is a multiplicative function of q. We leave the proof of the following lemma as a basic exercise involving the Chinese remainder theorem and manipulating summations.

LEMMA 7.6. *Suppose $q, q'\in\mathbb{N}$ and $\gcd(q,q')=1$. Then we have*

$$B(qq') = B(q)B(q').$$

Recall we defined the term $\mathfrak{S}(N)$ in (7.8). For each prime p, we define

$$\mu(p) = 1 + \sum_{t=1}^{\infty} B(p^t), \tag{7.18}$$

which converges absolutely under our assumptions on $\mathbf{f}$. Furthermore, under our assumptions on $\mathbf{f}$ the following limit exists

$$\mathfrak{S}(\infty) := \lim_{N\to\infty} \mathfrak{S}(N) = \prod_{p \text{ prime}} \mu(p), \tag{7.19}$$

which is called the *singular series*. We prove these statements in the following Lemma 7.7.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

LEMMA 7.7. *There exists* $\delta_1 > 0$ *such that for each prime* p*, we have*

$$\mu(p) = 1 + O(p^{-1-\delta_1}),$$

where the implicit constant is independent of p*. Furthermore, we have*

$$|\mathfrak{S}(N) - \mathfrak{S}(\infty)| \ll (\log N)^{-C\delta_2}$$

for some $\delta_2 > 0$.

Therefore, the limit in (7.19) exists, and the product in (7.19) converges. We leave the details that these two quantities are equal to the reader.

Proof. Let $\varepsilon_0 > 0$ be sufficiently small and let $\widetilde{Q} = Q - \varepsilon_0$. It follows from the definition of Q that $\widetilde{Q}$ satisfies

$$\widetilde{Q} > \frac{1 + R(800d^3 + 2)}{800d^3 + 1} \quad \text{and} \quad \widetilde{Q} > \frac{R+1}{1 - \frac{1}{800d^3+1}} > R + 1.$$

We substitute $Q = \widetilde{Q} + \varepsilon_0$ into the bounds in Lemmas 7.4 and 7.5. It is then clear that we may assume the implicit constants in Lemmas 7.4 and 7.5 are 1 for p sufficiently large with the cost of using $\widetilde{Q}$ in place of Q. For any $t \in \mathbb{N}$, we know that $\phi(p^t) = p^t(1 - 1/p) \geqslant \frac{1}{2}p^t$. Therefore, by considering the two cases as in the statements of Lemmas 7.4 and 7.5, we obtain

$$\begin{aligned}
&|\mu(p) - 1| \\
&\leqslant \sum_{1 \leqslant t \leqslant 800d^3+1} \left| \sum_{\substack{\gcd(\mathbf{a}, p^t)=1 \\ \mathbf{a} \in (\mathbb{Z}/p^t\mathbb{Z})^R}} \frac{1}{\phi(p^t)^n} \mathcal{S}_{\mathbf{a}, p^t} \right| + \sum_{t > 800d^3+1} \left| \sum_{\substack{\gcd(\mathbf{a}, p^t)=1 \\ \mathbf{a} \in (\mathbb{Z}/p^t\mathbb{Z})^R}} \frac{1}{\phi(p^t)^n} \mathcal{S}_{\mathbf{a}, p^t} \right| \\
&\ll \sum_{1 \leqslant t \leqslant 800d^3+1} p^{tR} p^{-nt} p^{nt - t\widetilde{Q}} + \sum_{t > 800d^3+1} p^{tR} p^{-nt} p^{\widetilde{Q} + nt - t\widetilde{Q}} \\
&\ll p^{R - \widetilde{Q}} + p^{\widetilde{Q}} p^{-(800d^3+2)(\widetilde{Q} - R)} \\
&\ll p^{-1-\delta_1},
\end{aligned}$$

for some $\delta_1 > 0$. We note that the implicit constants in $\ll$ are independent of p here.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let $q = p_1^{t_1}\cdots p_v^{t_v}$ be the prime factorization of $q \in \mathbb{N}$. Without loss of generality, suppose we have $t_j \leqslant 800d^3+1$ $(1 \leqslant j \leqslant v_0)$ and $t_j > 800d^3+1$ $(v_0 < j \leqslant v)$. By a similar calculation as above and the multiplicativity of $B(q)$, it follows that

$$\begin{aligned}
B(q) &= B(p_1^{t_1})\cdots B(p_v^{t_v}) \\
&\ll \left(\prod_{j=1}^{v_0} p_j^{t_j R} p_j^{-nt_j} p_j^{t_j(n-\widetilde{Q})}\right)\cdot\left(\prod_{j=v_0+1}^{v} p_j^{t_j R} p_j^{-nt_j} p_j^{\widetilde{Q}} p_j^{t_j(n-\widetilde{Q})}\right) \\
&= q^{R-\widetilde{Q}}\cdot\left(\prod_{j=v_0+1}^{v} p_j^{\widetilde{Q}}\right) \\
&\leqslant q^{R-\widetilde{Q}}\cdot q^{\widetilde{Q}/(800d^3+1)} \\
&\leqslant q^{-1-\delta_2},
\end{aligned}$$

for some $\delta_2 > 0$. We note that the implicit constant in $\ll$ is independent of q here, because the implicit constants in Lemmas 7.4 and 7.5 are 1 for p sufficiently large as mentioned above. Therefore, we obtain

$$\begin{aligned}
|\mathfrak{S}(N) - \mathfrak{S}(\infty)| &\leqslant \sum_{q>(\log N)^C} |B(q)| \\
&\ll \sum_{q>(\log N)^C} q^{-1-\delta_2} \\
&\ll (\log N)^{-C\delta_2}.
\end{aligned}$$

□

Let $\nu_t(p)$ denote the number of solutions $\mathbf{x} \in (\mathbb{U}_{p^t})^n$ to the congruence relations

$$f_{\ell,r}(\mathbf{x}) \equiv 0 \pmod{p^t}\ (1 \leqslant \ell \leqslant d,\ 1 \leqslant r \leqslant r_\ell). \tag{7.20}$$

Then using the fact that

$$\sum_{a\in\mathbb{Z}/p^t\mathbb{Z}} e\left(m\cdot a/p^t\right) = \begin{cases} p^t & \text{if } p^t|m, \\ 0 & \text{otherwise,}\end{cases}$$

we deduce

$$\begin{aligned}
&1+\sum_{j=1}^{t} B(p^j) \\
&\quad = 1+\sum_{j=1}^{t}\frac{1}{\phi(p^j)^n}\sum_{\mathbf{k}\in(\mathbb{U}_{p^j})^n}\ \sum_{\substack{\gcd(\mathbf{a},p^j)=1 \\ \mathbf{a}\in(\mathbb{Z}/p^j\mathbb{Z})^R}} e\left(\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k})\cdot a_{\ell,r}/p^j\right)
\end{aligned}$$

$$= \frac{1}{\phi(p^t)^n} \sum_{\mathbf{k} \in (\mathbb{U}_{p^t})^n} \sum_{\mathbf{a} \in (\mathbb{Z}/p^t\mathbb{Z})^R} e\left(\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} f_{\ell,r}(\mathbf{k}) \cdot a_{\ell,r}/p^t \right)$$

$$= \frac{p^{tR}}{\phi(p^t)^n} \nu_t(p).$$

Therefore, under our assumptions on $\mathbf{f}$ we obtain

$$\mu(p) = \lim_{t \to \infty} \frac{p^{tR} \nu_t(p)}{\phi(p^t)^n}.$$

We can then deduce by an application of Hensel's lemma that

$$\mu(p) > 0,$$

if the system (7.1) has a nonsingular solution in $\mathbb{Z}_p^\times$, the units of p-adic integers. The details are left to the reader. From this it follows in combination with Lemma 7.7 that if the system (7.1) has a nonsingular solution in $\mathbb{Z}_p^\times$ for every prime p, then

$$\mathfrak{S}(\infty) = \prod_{p \text{ prime}} \mu(p) > 0. \tag{7.21}$$

By combining Lemmas 7.3 and 7.7, we obtain the following.

PROPOSITION 7.8. *Let* $\mathbf{f}$ *be the polynomials in* (7.1). *Given any* $c > 0$, *for sufficiently large* $C > 0$ *we have*

$$\int_{\mathfrak{M}(C)} T(\mathbf{f}; \boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \mathfrak{S}(\infty)\mu(\infty) X^{n - \sum_{\ell=1}^{d} \ell r_\ell} + O\left(\frac{X^{n - \sum_{\ell=1}^{d} \ell r_\ell}}{(\log X)^c} \right).$$

We note that this proposition contains Proposition 4.2 as a special case with

$$\mathcal{C}(\mathbf{f}) = \mathfrak{S}(\infty)\mu(\infty). \tag{7.22}$$

8. Conclusions and further remarks

Let us refer to the polynomials in (4.1) as $\mathbf{f}$, and the polynomials in (4.2) as $\mathfrak{f}$ in this section. We let $\mathbf{F}$ and $\mathfrak{F}$ be the systems of the highest degree homogeneous portions of $\mathbf{f}$ and $\mathfrak{f}$, respectively.

As a consequence of Propositions 4.1 and 4.2, we obtain the following asymptotic formula for the system of equations (4.2). We have that given any $c > 0$, there exists $C > 0$ such that

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

$$\begin{aligned}\mathcal{M}_{\mathfrak{f}}(X) &= \int_{\mathbb{T}^R} T(\mathfrak{f};\boldsymbol{\alpha})\,\mathbf{d}\boldsymbol{\alpha}\\ &= \int_{\mathfrak{M}(C)} T(\mathfrak{f};\boldsymbol{\alpha})\,\mathbf{d}\boldsymbol{\alpha} + \int_{\mathfrak{m}(C)} T(\mathfrak{f};\boldsymbol{\alpha})\,\mathbf{d}\boldsymbol{\alpha}\\ &= \mathcal{C}(\mathfrak{f})X^{n-\sum_{\ell=1}^{d}\ell r_\ell} + O\left(\frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log X)^c}\right),\end{aligned} \tag{8.1}$$

which proves Theorem 1.2 for $\mathfrak{f}$.

Recall from Section 4 that transforming the system $\mathbf{f}$ into $\mathfrak{f}$ does not affect its solution set, in other words $V_{\mathbf{f},\mathbf{0}}(\mathbb{Z}) = V_{\mathfrak{f},\mathbf{0}}(\mathbb{Z})$. Therefore, we in fact have

$$\mathcal{M}_{\mathbf{f}}(X) = \mathcal{M}_{\mathfrak{f}}(X) = \mathcal{C}(\mathfrak{f})X^{n-\sum_{\ell=1}^{d}\ell r_\ell} + O\left(\frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log N)^c}\right).$$

Since $\mathcal{C}(\mathfrak{f})$ is a constant dependent only on $\mathfrak{f}$, in turn it follows that it is a constant which depends only on $\mathbf{f}$. Thus by setting $\mathcal{C}(\mathbf{f}) = \mathcal{C}(\mathfrak{f})$, we have obtained Theorem 1.2.

We also remark that if $V_{\mathbf{F},\mathbf{0}}(\mathbb{R})$ has a nonsingular real point in $(0, 1)^n$, then so does $V_{\mathfrak{F},\mathbf{0}}(\mathbb{R})$, and if the system of equations (4.1) has a nonsingular solution in $\mathbb{Z}_p^\times$ for every prime p, then so does the system (4.2). Under these conditions, it follows from (7.6), (7.21), and (7.22) that $\mathcal{C}(\mathbf{f}) = \mathcal{C}(\mathfrak{f}) > 0$. We leave the details here to the reader.

Finally, we followed [6] and used the von Mangoldt function Λ as our weight for the exponential sum. Consequently, $\mathcal{M}_{\mathbf{f}}(X)$ counts the number of solutions, with a logarithmic weight, to the equations $\mathbf{f} = \mathbf{0}$ whose coordinates are all prime powers. Let $\mathbf{1}_{\mathcal{P}}$ denote the characteristic function of the set of prime numbers. For $\mathbf{x} = (x_1, \ldots, x_n)$, we let $\mathbf{1}_{\mathcal{P}}(\mathbf{x}) = \mathbf{1}_{\mathcal{P}}(x_1)\cdots\mathbf{1}_{\mathcal{P}}(x_n)$ and $\log(\mathbf{x}) = \log(x_1)\cdots\log(x_n)$. Let us define

$$\mathcal{M}'_{\mathbf{f}}(X) := \sum_{\mathbf{x}\in[0,X]^n} \log(\mathbf{x})\mathbf{1}_{\mathcal{P}}(\mathbf{x})\mathbf{1}_{V_{\mathbf{f},\mathbf{0}}(\mathbb{C})}(\mathbf{x})$$

with the convention that $\log(\mathbf{x})\mathbf{1}_{\mathcal{P}}(\mathbf{x}) = 0$ if $x_i = 0$ for some $1 \leqslant i \leqslant n$. The quantity $\mathcal{M}'_{\mathbf{f}}(X)$ counts the number of prime solutions, with a logarithmic weight, to the equations $\mathbf{f} = \mathbf{0}$. We record the following result for $\mathcal{M}'_{\mathbf{f}}(X)$.

THEOREM 8.1. *Under the same hypotheses as in Theorem 1.2, we have*

$$\mathcal{M}'_{\mathbf{f}}(X) = \mathcal{C}(\mathbf{f})X^{n-\sum_{\ell=1}^{d}\ell r_\ell} + O\left(\frac{X^{n-\sum_{\ell=1}^{d}\ell r_\ell}}{(\log N)^c}\right),$$

where $\mathcal{C}(\mathbf{f})$ is the same constant as in the statement of Theorem 1.2.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

We can obtain this asymptotic formula by changing the weight from $\Lambda(\mathbf{x})$ to $\log(\mathbf{x})\mathbf{1}_{\mathcal{P}}(\mathbf{x})$ in the proof of Theorem 1.2. Since the resulting changes in the proof are minimal, we leave the details to the reader.

Acknowledgements

The author would like to thank James Maynard and Kannan Soundararajan for their helpful advices, and the anonymous referee for her/his careful reading and helpful suggestions. The author would also like to thank the following people for helpful conversations and/or encouragement while working on this paper: Matthew Beckett, Arunabha Biswas, Tim Browning, Francesco Cellarosi, Robert Krone, Jamie Mingo, Abdol-Reza Mansouri, M. Ram Murty, Mike Roth, Trevor Wooley, Stanley Yao Xiao, and Serdar Yüksel.

Appendix A. Proofs of the results in Section 6

In this appendix, we provide proof for the results presented in Section 6. Let us denote $\mathfrak{B}_1 = [-1, 1]^n$ and $\mathfrak{B}_0 = [0, 1]^n$. Let $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{x}_j = (x_{j,1}, \dots, x_{j,n})$ for $j \geqslant 1$. Given a function $G(\mathbf{x})$, we define

$$\Gamma_{\ell,G}(\mathbf{x}_1, \dots, \mathbf{x}_\ell) = \sum_{t_1=0}^{1} \cdots \sum_{t_\ell=0}^{1} (-1)^{t_1+\cdots+t_\ell} G(t_1\mathbf{x}_1 + \cdots + t_\ell \mathbf{x}_\ell).$$

Then it follows that $\Gamma_{\ell,G}$ is symmetric in its ℓ arguments, and that $\Gamma_{\ell,G}(\mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}, \mathbf{0}) = 0$ [20, Section 11]. It is clear from the definition that if $G'(\mathbf{x})$ is another function, then $\Gamma_{\ell,G} + \Gamma_{\ell,G'} = \Gamma_{\ell,G+G'}$. We also have that if G is a form of degree d and $\ell > d > 0$, then $\Gamma_{\ell,G} = 0$ [20, Lemma 11.2].

For $\alpha \in \mathbb{R}$, let $\|\alpha\|$ denote the distance from α to the closest integer. Let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_d, \dots, \boldsymbol{\alpha}_1) \in \mathbb{R}^R$, where $R = r_1 + \cdots + r_d$ and $\boldsymbol{\alpha}_\ell = (\alpha_{\ell,1}, \dots, \alpha_{\ell,r_\ell}) \in \mathbb{R}^{r_\ell}$ $(1 \leqslant \ell \leqslant d)$. We define

$$\|\boldsymbol{\alpha}\| = \max_{\substack{1\leqslant \ell \leqslant d \\ 1 \leqslant r \leqslant r_\ell}} \|\alpha_{\ell,r}\| \quad \text{and} \quad |\boldsymbol{\alpha}| = \max_{\substack{1\leqslant \ell \leqslant d \\ 1 \leqslant r \leqslant r_\ell}} |\alpha_{\ell,r}|.$$

We have the following standard results related to Weyl differencing.

Lemma A.1 [20, Lemma 13.1]. *Suppose $G(\mathbf{x}) = G^{(0)} + G^{(1)}(\mathbf{x}) + \cdots + G^{(d)}(\mathbf{x})$, where $G^{(j)}$ is a form of degree j with real coefficients $(1 \leqslant j \leqslant d)$ and $G^{(0)} \in \mathbb{R}$. Let $P > 1$, and put*

$$S' = S'(G, P, \mathfrak{B}_0) := \sum_{\mathbf{x} \in P\mathfrak{B}_0 \cap \mathbb{Z}^n} e(G(\mathbf{x})).$$

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *be the standard basis vectors of* $\mathbb{R}^n$. *Let* $\varepsilon > 0$ *and* $2 \leqslant \ell \leqslant d$. *If* $\ell = d$, *then let* $\theta = 0$ *and* $q = 1$. *On the other hand, if* $2 \leqslant \ell < d$, *then suppose* $0 \leqslant \theta < 1/4$ *and that there is* $q \in \mathbb{N}$ *with*

$$q \leqslant P^{\theta} \quad \text{and} \quad \|qG^{(j)}\| \leqslant cP^{\theta-j} \quad (\ell < j \leqslant d).$$

Then we have

$$|S'|^{2^{\ell-1}} \ll P^{(2^{\ell-1}-\ell+2\theta)n+\varepsilon} \sum \left(\prod_{i=1}^{n} \min(P^{1-2\theta}, \|q\Gamma_{\ell,G^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i)\|^{-1}) \right),$$

where the sum $\sum$ *is over* $(\ell-1)$*-tuples of integer points* $\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}$ *in* $P\mathfrak{B}_1$, *and the implicit constant in* $\ll$ *depends only on* n, d, c, *and* ε.

We remark that the term c which appears in the statement of Lemma A.1 is not present in the statement of [20, Lemma 13.1]. However, it can be seen from the proof of [20, Lemma 13.1] that this change does not affect the result, or see the explanation given in [20, page 275, line 5].

LEMMA A.2 [20, Lemma 14.2]. *Make all the assumptions of Lemma A.1. Suppose further that*

$$|S'| \geqslant P^{n-Q}$$

where $Q > 0$. *Let* $\eta > 0$ *and* $\eta + 4\theta \leqslant 1$. *Then the number* $N(\eta)$ *of integral* $(\ell - 1)$*-tuples*

$$\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1} \in P^{\eta}\mathfrak{B}_1$$

with

$$\|q\Gamma_{\ell,G^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i)\| < P^{-\ell+4\theta+(\ell-1)\eta} \quad (i = 1, \ldots, n)$$

satisfies

$$N(\eta) \gg P^{n(\ell-1)\eta-2^{\ell-1}Q-\varepsilon},$$

where the implicit constant in $\gg$ *depends only on* n, d, c, η, *and* ε.

Let $\mathbf{u} = (\mathbf{u}_d, \ldots, \mathbf{u}_1)$ be a system of polynomials in $\mathbb{Q}[x_1, \ldots, x_n]$, where $\mathbf{u}_\ell = (u_{\ell,1}, \ldots, u_{\ell,r_\ell})$ is the subsystem of degree ℓ polynomials of $\mathbf{u}$ $(1 \leqslant \ell \leqslant d)$. We let $\mathbf{U} = (\mathbf{U}_d, \ldots, \mathbf{U}_1)$ be the system of forms, where for each $1 \leqslant \ell \leqslant d$, $\mathbf{U}_\ell = (U_{\ell,1}, \ldots, U_{\ell,r_\ell})$ and $U_{\ell,r}$ is the homogeneous degree ℓ portion of $u_{\ell,r}$ $(1 \leqslant r \leqslant r_\ell)$. We define the following exponential sum associated to $\mathbf{u}$,

$$S(\boldsymbol{\alpha}) = S(\mathbf{u}, \mathfrak{B}_0; \boldsymbol{\alpha}) := \sum_{\mathbf{x} \in P\mathfrak{B}_0 \cap \mathbb{Z}^n} e\left(\sum_{1 \leqslant \ell \leqslant d} \sum_{1 \leqslant r \leqslant r_\ell} \alpha_{\ell,r} \cdot u_{\ell,r}(\mathbf{x}) \right). \tag{A.1}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors of $\mathbb{C}^n$. Let $1 < \ell \leqslant d$. We define $\mathbb{M}_\ell = \mathbb{M}_\ell(\mathbf{U}_\ell)$ to be the set of $(\ell-1)$-tuples $(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}) \in (\mathbb{C}^n)^{\ell-1}$ for which the matrix

$$[m_{ri}] = [\Gamma_{\ell, U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i)] \quad (1 \leqslant r \leqslant r_\ell, 1 \leqslant i \leqslant n) \tag{A.2}$$

has rank strictly less than r_ℓ. For $P_0 > 0$, we denote $z_{P_0}(\mathbb{M}_\ell)$ to be the number of integer points $(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1})$ on $\mathbb{M}_\ell$ such that

$$\max_{1 \leqslant i \leqslant \ell-1} \max_{1 \leqslant j \leqslant n} |x_{i,j}| \leqslant P_0.$$

Given a degree ℓ polynomial

$$u(\mathbf{x}) = \sum_{\substack{i_j \in \mathbb{N} \cup \{0\} (1 \leqslant j \leqslant n) \\ 0 \leqslant i_1 + \cdots + i_n \leqslant \ell}} A_{i_1, \ldots, i_n} x_1^{i_1} \ldots x_n^{i_n}$$

with real coefficients, we denote

$$|u| = \max_{\substack{i_j \in \mathbb{N} \cup \{0\} (1 \leqslant j \leqslant n) \\ 0 \leqslant i_1 + \cdots + i_n \leqslant \ell}} |A_{i_1, \ldots, i_n}| \quad \text{and} \quad \|u\| = \max_{\substack{i_j \in \mathbb{N} \cup \{0\} (1 \leqslant j \leqslant n) \\ 0 \leqslant i_1 + \cdots + i_n \leqslant \ell}} \|A_{i_1, \ldots, i_n}\|.$$

LEMMA A.3 [20, Lemma 11.3]. *Suppose $U(\mathbf{x})$ is a form of degree ℓ. Then we have*

$$\|\Gamma_{\ell,U}\| \leqslant 2^\ell \ell^\ell \, \|U\|.$$

By a similar proof as in [20, Lemma 11.3], we can also show that for a degree ℓ form $U(\mathbf{x})$ the following holds

$$|\Gamma_{\ell,U}| \leqslant 2^\ell \ell^\ell |U|. \tag{A.3}$$

Let $1 < \ell \leqslant d$ and $r_\ell > 0$. We define $g_\ell(\mathbf{U}_\ell)$ to be the largest real number such that

$$z_P(\mathbb{M}_\ell) \ll P^{n(\ell-1) - g_\ell(\mathbf{U}_\ell) + \varepsilon} \tag{A.4}$$

holds for each $\varepsilon > 0$. It was proved in [20, page 280, Corollary] that

$$h_\ell(\mathbf{U}_\ell) < \frac{\ell!}{(\log 2)^\ell} (g_\ell(\mathbf{U}_\ell) + (\ell-1) r_\ell (r_\ell - 1)). \tag{A.5}$$

Let

$$\gamma_\ell = \frac{2^{\ell-1}(\ell-1) r_\ell}{g_\ell(\mathbf{U}_\ell)}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

when $r_\ell > 0$ and $g_\ell(\mathbf{U}_\ell) > 0$. We let $\gamma_\ell = 0$ if $r_\ell = 0$, and let $\gamma_\ell = +\infty$ if $r_\ell > 0$ and $g_\ell(\mathbf{U}_\ell) = 0$. For ℓ with $r_\ell > 0$, we also define

$$\gamma'_\ell = \frac{2^{\ell-1}}{g_\ell(\mathbf{U}_\ell)} = \frac{\gamma_\ell}{(\ell-1)r_\ell}. \tag{A.6}$$

We have to deal with the cases when the coefficients of $\mathbf{u}$ may depend on P (but not the coefficients of $\mathbf{U}$). There are essentially two different scenarios we have to consider, the first of which we refer to as follows.

Condition ($\star'$): The polynomials of $\mathbf{u}$ have coefficients in $\mathbb{Z}$, and the coefficients of $\mathbf{U}$ do not depend on P. However, given $u_{\ell,r}(\mathbf{x})$ $(1 \leqslant \ell \leqslant d, 1 \leqslant r \leqslant r_\ell)$ the coefficients of its monomials whose degrees are strictly less than ℓ may depend on P.

The following lemma is essentially [20, Lemma 15.1]. The point here is that if we are only considering the case $\ell = d$, then the implicit constants may depend on $\mathbf{U}_d$ but not on $\mathbf{u}$. (Note for the case $\ell < d$ the implicit constants may depend on $\mathbf{u}$, see [23, Lemma 2.2].)

LEMMA A.4 [20, Lemma 15.1]. *Suppose* $\mathbf{u}$ *satisfies Condition* ($\star'$). *Let* $Q > 0$, $\varepsilon > 0$, *and let* P *be sufficiently large with respect to* d *and* $r_d, \dots, r_1$. *Let* $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in* (A.1). *Given* $0 < \eta \leqslant 1$, *one of the following three alternatives must hold:*

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{r_d(d-1)\eta} \quad \textit{and} \quad \|n_0 \boldsymbol{\alpha}_d\| \ll P^{-d+r_d(d-1)\eta}.$$

(iii) $z_{P_0}(\mathbb{M}_\ell) \gg P_0^{(d-1)n - 2^{d-1}(Q/\eta) - \varepsilon}$ *holds with* $P_0 = P^\eta$.

The implicit constants depend at most on $n, d, r_d, \eta, \varepsilon$, *and* $\mathbf{U}_d$.

Proof. We have $\boldsymbol{\alpha} \in \mathbb{R}^R$. Let us denote

$$\sum_{\ell=1}^{d} \sum_{r=1}^{r_\ell} \alpha_{\ell,r} u_{\ell,r}(\mathbf{x}) = G^{(0)} + G^{(1)}(\mathbf{x}) + \cdots + G^{(d)}(\mathbf{x}),$$

where $G^{(j)}$ is a form of degree j $(1 \leqslant j \leqslant d)$ and $G^{(0)} \in \mathbb{R}$. Then it is clear that $G^{(d)}(\mathbf{x}) = \sum_{r=1}^{r_d} \alpha_{d,r} U_{d,r}(\mathbf{x})$, and it depends on $\mathbf{U}_d$ only, and not on $\mathbf{u}$. With this observation, by following through the proof of [20, Lemma 15.1] for the case $\ell = d$ while keeping track of the constant dependency, we obtain the result. □

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

From Lemma A.4, we obtain the following corollary in a similar manner as in [20, page 276, Corollary].

COROLLARY A.5 [20, page 276, Corollary]. *Suppose* $\mathbf{u}$ *satisfies Condition* $(\star')$. *Let* $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in* (A.1). *Suppose* $\varepsilon' > 0$ *is sufficiently small and* $Q > 0$ *satisfies*

$$Q\gamma_d' < 1.$$

Then one of the following two alternatives must hold:

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{Q\gamma_d+\varepsilon'} \quad \text{and} \quad \|n_0\boldsymbol{\alpha}_d\| \ll P^{-d+Q\gamma_d+\varepsilon'}.$$

The implicit constants depend at most on $n, d, r_d, \varepsilon', Q$, *and* $\mathbf{U}_d$.

Now we move on to our next scenario of when the coefficients of $\mathbf{u}$ may depend on P. Let $u_{\ell,r}^{(j)}(\mathbf{x})$ be the homogeneous degree j portion of the polynomial $u_{\ell,r}(\mathbf{x})$. In the following lemma, for $j < \ell$ the coefficients of $u_{\ell,r}^{(j)}(\mathbf{x})$ may be in $\mathbb{Q}$ and also depend on P, but in a controlled manner. On the other hand, the coefficients of $U_{\ell,r}(\mathbf{x})$ do not depend on P. We also note the implicit constants may depend on $\mathbf{U}$ but not on $\mathbf{u}$.

LEMMA A.6 [20, Lemma 15.1]. *Suppose* $\mathbf{u}$ *has coefficients in* $\mathbb{Q}$, *and further suppose* $\mathbf{U}$ *has coefficients in* $\mathbb{Z}$. *Let* $Q > 0$ *and* $\varepsilon > 0$. *Let* $2 \leqslant \ell \leqslant d$ *with* $r_\ell > 0$. *If* $\ell = d$, *then let* $\theta = 0$ *and* $q = 1$. *On the other hand, if* $2 \leqslant \ell < d$, *then suppose* $0 \leqslant \theta < 1/4$ *and that there is* $q \in \mathbb{N}$ *with*

$$q \leqslant P^\theta, \quad q\boldsymbol{\alpha}_{\ell'} \in \mathbb{Z}^{r_{\ell'}} \quad (\ell < \ell' \leqslant d),$$

and

$$q\alpha_{j,r}u_{j,r}^{(\ell')}(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$$

for every $\ell < j \leqslant d, 0 \leqslant \ell' < j, 1 \leqslant r \leqslant r_j$.

Let $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in (A.1). Given* $\eta > 0$ *with* $\eta + 4\theta \leqslant 1$, *one of the following three alternatives must hold:*

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{r_\ell(\ell-1)\eta} \quad \text{and} \quad \|qn_0\boldsymbol{\alpha}_\ell\| \ll P^{-\ell+4\theta+r_\ell(\ell-1)\eta}.$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

(iii) $z_{P_0}(\mathbb{M}_\ell) \gg P_0^{(\ell-1)n-2^{\ell-1}(Q/\eta)-\varepsilon}$ *holds with* $P_0 = P^\eta$.

The implicit constants depend at most on $n, d, r_d, \ldots, r_1, \eta, \varepsilon$, *and* $\mathbf{U}$.

Proof. We have $\boldsymbol{\alpha} \in \mathbb{R}^R$. Let us denote

$$\sum_{\ell=1}^{d}\sum_{r=1}^{r_\ell} \alpha_{\ell,r} u_{\ell,r}(\mathbf{x}) = G^{(0)} + G^{(1)}(\mathbf{x}) + \cdots + G^{(d)}(\mathbf{x}),$$

where $G^{(\ell')}$ is a form of degree ℓ' $(1 \leqslant \ell' \leqslant d)$ and $G^{(0)} \in \mathbb{R}$. Then it is clear that $G^{(d)}(\mathbf{x}) = \sum_{r=1}^{r_d} \alpha_{d,r} U_{d,r}(\mathbf{x})$. Recall we denote $u_{j,r}^{(\ell')}(\mathbf{x})$ to be the homogeneous degree ℓ' portion of the polynomial $u_{j,r}(\mathbf{x})$. Then we have

$$G^{(\ell')}(\mathbf{x}) = \sum_{r=1}^{r_{\ell'}} \alpha_{\ell',r} U_{\ell',r}(\mathbf{x}) + \sum_{j=\ell'+1}^{d} \sum_{r=1}^{r_j} \alpha_{j,r} u_{j,r}^{(\ell')}(\mathbf{x}) \quad (1 \leqslant \ell' < d).$$

If $\ell < d$, then it is clear from our hypothesis that we have

$$\| q G^{(\ell')} \| = 0 \leqslant P^{\theta-\ell'}$$

for each $\ell < \ell' \leqslant d$.

Suppose the alternative (i) fails. In this case, we may apply Lemma A.2 and obtain that the number $N(\eta)$ of integral $(\ell-1)$-tuples $\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}$ in $P^\eta \mathfrak{B}_1$ with

$$\| q \Gamma_{\ell, G^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \| < P^{-\ell+4\theta+(\ell-1)\eta} \quad (i = 1, \ldots, n) \tag{A.7}$$

satisfies

$$N(\eta) \gg P_0^{n(\ell-1)-2^{\ell-1}(Q/\eta)-\varepsilon},$$

where $P_0 = P^\eta$, and the implicit constant in $\gg$ depends only on n, d, η, and ε. We have

$$\begin{aligned} &\| q \Gamma_{\ell, G^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \| \\ &\quad = \left\| \sum_{r=1}^{r_\ell} q \alpha_{\ell,r} \Gamma_{\ell, U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) + \sum_{j=\ell+1}^{d} \sum_{r=1}^{r_j} q \alpha_{j,r} \Gamma_{\ell, u_{j,r}^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \right\| \\ &\quad = \left\| \sum_{r=1}^{r_\ell} q \alpha_{\ell,r} \Gamma_{\ell, U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \right\|, \end{aligned}$$

because

$$q \alpha_{j,r} \Gamma_{\ell, u_{j,r}^{(\ell)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \in \mathbb{Z} \tag{A.8}$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

for each $\ell < j \leqslant d$, $1 \leqslant r \leqslant r_j$. Thus we see that (A.7) implies

$$\left\| \sum_{r=1}^{r_\ell} q\alpha_{\ell,r} \Gamma_{\ell,U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \right\| < P^{-\ell+4\theta+(\ell-1)\eta} \quad (i = 1, \ldots, n). \tag{A.9}$$

Given $\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}$ as above, we form a matrix

$$[m_{ri}]_{\mathbf{x}_1,\ldots,\mathbf{x}_{\ell-1}},$$

where its entries are

$$m_{ri} = \Gamma_{\ell,U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \quad (1 \leqslant r \leqslant r_\ell, 1 \leqslant i \leqslant n).$$

Now if this matrix $[m_{ri}]_{\mathbf{x}_1,\ldots,\mathbf{x}_{\ell-1}}$ has rank strictly less than r_ℓ for each of the $(\ell-1)$-tuples counted by $N(\eta)$, then by the definition of $z_{P_0}(\mathbb{M}_\ell)$ we have

$$z_{P_0}(\mathbb{M}_\ell) \geqslant N(\eta) \gg P_0^{n(\ell-1)-2^{\ell-1}(Q/\eta)-\varepsilon},$$

where the implicit constant in $\gg$ depends only on n, d, η, and ε. Thus we have the alternative (iii) in this case. Hence, we may suppose that at least one of these matrices, which we denote by $[m_{ri}]$, has rank r_ℓ. Without loss of generality, suppose the submatrix M_0 formed by taking the first r_ℓ columns of $[m_{ri}]$ has rank r_ℓ.

It follows from the definition of $\Gamma_{\ell,U_{\ell,r}}$ that every monomial occurring in $\Gamma_{\ell,U_{\ell,r}}(\mathbf{z}_1, \ldots, \mathbf{z}_\ell)$ has some component of $\mathbf{z}_i = (z_{i,1}, \ldots, z_{i,n})$ as a factor for each $1 \leqslant i \leqslant \ell$ [20, Proof of Lemma 11.2]. Recall we also have

$$|\Gamma_{\ell,U_{\ell,r}}| \leqslant 2^\ell \ell^\ell |U_{\ell,r}|$$

from (A.3). Therefore, we have

$$m_{ri} = \Gamma_{\ell,U_{\ell,r}}(\mathbf{x}_1, \ldots, \mathbf{x}_{\ell-1}, \mathbf{e}_i) \ll P_0^{\ell-1},$$

and also

$$n_0 := \det(M_0) \ll P_0^{r_\ell(\ell-1)} = P^{r_\ell(\ell-1)\eta},$$

where the implicit constants in $\ll$ depend only on n, ℓ, r_ℓ, and $\mathbf{U}_\ell$. Hence, from (A.9) we may write

$$q \sum_{r=1}^{r_\ell} \alpha_{\ell,r} m_{ri} = c_i + \beta_i' \quad (1 \leqslant i \leqslant n),$$

where c_i are integers and β_i' are real numbers satisfying

$$|\beta_i'| < P^{-\ell+4\theta+(\ell-1)\eta} \quad (1 \leqslant i \leqslant n).$$

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

Let $v_1, \ldots, v_{r_\ell}$ be the solution to the system of linear equations

$$\sum_{r=1}^{r_\ell} v_r m_{ri} = n_0 c_i \quad (1 \leqslant i \leqslant r_\ell). \tag{A.10}$$

Then we have

$$\sum_{r=1}^{r_\ell} (q n_0 \alpha_{\ell,r} - v_r) m_{ri} = n_0 \beta_i' \quad (1 \leqslant i \leqslant r_\ell). \tag{A.11}$$

By applying Cramér's rule to (A.10), it follows that $v_r \in \mathbb{Z}$ $(1 \leqslant r \leqslant r_\ell)$. Also by applying Cramér's rule to (A.11), we obtain

$$\|q n_0 \alpha_{\ell,r}\| \leqslant |q n_0 \alpha_{\ell,r} - v_r| \ll P_0^{(\ell-1)(r_\ell-1)} P^{-\ell+4\theta+(\ell-1)\eta} = P^{-\ell+4\theta+r_\ell(\ell-1)\eta}, \tag{A.12}$$

where the implicit constant in $\ll$ depends only on n, ℓ, r_ℓ, and $\mathbf{U}_\ell$. This completes the proof of Lemma A.6. □

We then have the following corollary.

COROLLARY A.7 [20, page 276, Corollary]. *Suppose* $\mathbf{u}$ *has coefficients in* $\mathbb{Q}$, *and further suppose* $\mathbf{U}$ *has coefficients in* $\mathbb{Z}$. *Let* $Q > 0$ *and* $\varepsilon > 0$. *Let* $2 \leqslant \ell \leqslant d$ *with* $r_\ell > 0$. *If* $\ell = d$, *then let* $\theta = 0$ *and* $q = 1$. *On the other hand, if* $2 \leqslant \ell < d$, *then suppose* $0 \leqslant \theta < 1/4$ *and that there is* $q \in \mathbb{N}$ *with*

$$q \leqslant P^\theta, \quad q\boldsymbol{\alpha}_j \in \mathbb{Z}^{r_j} \quad (\ell < j \leqslant d),$$

and

$$q \alpha_{\ell',r} u_{\ell',r}^{(j)}(\mathbf{x}) \in \mathbb{Z}[x_1, \ldots, x_n]$$

for every $\ell < \ell' \leqslant d, 0 \leqslant j < \ell', 1 \leqslant r \leqslant r_{\ell'}$.

Let $S(\boldsymbol{\alpha})$ *be the sum associated to* $\mathbf{u}$ *as in* (A.1). *Suppose*

$$4\theta + Q\gamma_\ell' < 1.$$

Then one of the following two alternatives must hold:

(i) $|S(\boldsymbol{\alpha})| \leqslant P^{n-Q}$.

(ii) *There exists* $n_0 \in \mathbb{N}$ *such that*

$$n_0 \ll P^{Q\gamma_\ell+\varepsilon} \quad \textit{and} \quad \|n_0 q \boldsymbol{\alpha}_\ell\| \ll P^{-\ell+4\theta+Q\gamma_\ell+\varepsilon}.$$

The implicit constants depend at most on $n, d, r_d, \ldots, r_1, Q, \varepsilon$, *and* $\mathbf{U}$.

Proof. The proof is similar to that of [20, page 276, Corollary]. If we have

$$2^{\ell-1}Q/\eta < g_\ell(\mathbf{U}_\ell),$$

then it is clear that the alternative (iii) of Lemma A.6 cannot occur for P sufficiently large with respect to $n, d, r_d, \ldots, r_1, \eta, \varepsilon$, and $\mathbf{U}$. In particular, this is the case with $\eta = Q\gamma'_\ell + \varepsilon'$ where $\varepsilon' > 0$ is sufficiently small. Note we also have

$$\eta + 4\theta < 1,$$

given $4\theta + Q\gamma'_\ell < 1$. □

References

[1] B. J. Birch, 'Forms in many variables', *Proc. R. Soc. Lond. Ser. A* **265** (1961/62), 245–263.

[2] J. Bourgain, A. Gamburd and P. Sarnak, 'Affine linear sieve, expanders, and sum-product', *Invent. Math.* **179**(3) (2010), 559–644.

[3] T. D. Browning and D. R. Heath-Brown, 'Forms in many variables and differing degrees', *J. Eur. Math. Soc. (JEMS)* **9** (2017), 357–394.

[4] J. Brüdern, R. Dietmann, J. Liu and T. D. Wooley, 'A Birch–Goldbach theorem', *Arch. Math. (Basel)* **94**(1) (2010), 53–58.

[5] S. Chow, 'Roth-Waring-Goldbach', *Int. Math. Res. Not. IMRN* **2018**(8) (2018), 2341–2374.

[6] B. Cook and Á. Magyar, 'Diophantine equations in the primes', *Invent. Math.* **198** (2014), 701–737.

[7] H. Davenport, *Analytic Methods for Diophantine Equations and Diopantine Inequalities*, 2nd edn (Cambridge University Press, Cambridge, 2005).

[8] W. Duke, Z. Rudnick and P. Sarnak, 'Density of integer points on affine homogeneous varieties', *Duke Math. J.* **71**(1) (1993), 143–179.

[9] D. A. Goldston, J. Pintz and C. Y. Yıldırım, 'Primes in tuples. I', *Ann. of Math. (2)* **170**(2) (2009), 819–862.

[10] B. Green and T. Tao, 'The primes contain arbitrarily long arithmetic progressions', *Ann. of Math. (2)* **167**(2) (2008), 481–547.

[11] B. Green and T. Tao, 'The distribution of polynomials over finite fields, with applications to the Gowers norms', *Contrib. Discrete Math.* **4**(2) (2009), 1–36.

[12] B. Green and T. Tao, 'Linear equations in primes', *Ann. of Math. (2)* **171**(3) (2010), 1753–1850.

[13] H. A. Helfgott, 'The ternary Goldbach problem', *Ann. Math. Stud.*, to appear.

[14] L. K. Hua, *Additive Theory of Prime Numbers*, Translations of Mathematical Monographs, 13 (American Mathematical Society, Providence, RI, 1965).

[15] T. Kaufman and S. Lovett, 'Worst case to average case reductions for polynomials', in *49th Annual IEEE Symposium on Foundations of Computer Science* (2008), 166–175. https://www.computer.org/csdl/proceedings/focs/2008/3436/00/3436a166-abs.html.

[16] A. V. Kumchev and T. D. Wooley, 'On the Waring-Goldbach problem for eighth and higher powers', *J. Lond. Math. Soc. (2)* **93**(3) (2016), 811–824.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press

[17] A. V. Kumchev and T. D. Wooley, 'On the Waring-Goldbach problem for seventh and higher powers', *Monatsh. Math.* **183**(2) (2017), 303–310.

[18] J. Liu and P. Sarnak, 'Integral points on quadrics in three variables whose coordinates have few prime factors', *Israel J. Math.* **178** (2010), 393–426.

[19] J. Maynard, 'Small gaps between primes', *Ann. of Math. (2)* **181**(1) (2015), 383–413.

[20] W. M. Schmidt, 'The density of integer points on homogeneous varieties', *Acta Math.* **154**(3–4) (1985), 243–296.

[21] I. M. Vinogradov, 'Representation of an odd number as a sum of three primes', *Dokl. Akad. Nauk. SSSR* **15** (1937), 291–294.

[22] S. Y. Xiao and S. Yamagishi, 'Zeroes of polynomials in many variables with prime inputs'. Preprint, 2015, arXiv:1512.01258.

[23] S. Yamagishi, 'An exponential sum estimate for systems with linear polynomials', *J. Théor. Nombres Bordeaux*, to appear.

[24] Y. Zhang, 'Bounded gaps between primes', *Ann. of Math. (2)* **179**(3) (2014), 1121–1174.

[25] L. Zhao, 'The quadratic form in nine prime variables', *Nagoya Math. J.* **223**(1) (2016), 21–65.

https://doi.org/10.1017/fms.2018.21 Published online by Cambridge University Press