THE LOCAL SOLUBILITY FOR HOMOGENEOUS POLYNOMIALS WITH RANDOM COEFFICIENTS OVER THIN SETS

HEEJONG LEE, SEUNGSU LEE, AND KISEOK YEON

Abstract. Let d and n be natural numbers greater or equal to 2. Let $\langle \boldsymbol{a}, \nu_{d,n}(\boldsymbol{x})\rangle \in \mathbb{Z}[\boldsymbol{x}]$ be a homogeneous polynomial in n variables of degree d with integer coefficients $\boldsymbol{a}$, where $\langle \cdot, \cdot \rangle$ denotes the inner product, and $\nu_{d,n} : \mathbb{R}^n \to \mathbb{R}^N$ denotes the Veronese embedding with $N = \binom{n+d-1}{d}$. Consider a variety $V_{\boldsymbol{a}}$ in $\mathbb{P}^{n-1}$, defined by $\langle \boldsymbol{a}, \nu_{d,n}(\boldsymbol{x})\rangle = 0$. In this paper, we examine a set of integer vectors $\boldsymbol{a} \in \mathbb{Z}^N$, defined by

$$\mathfrak{A}(A; P) = \{\boldsymbol{a} \in \mathbb{Z}^N : \ P(\boldsymbol{a}) = 0, \ \|\boldsymbol{a}\|_\infty \leq A\},$$

where $P \in \mathbb{Z}[\boldsymbol{x}]$ is a non-singular form in N variables of degree k with $2 \leq k \leq C(n, d)$ for some constant $C(n, d)$ depending at most on n and d. Suppose that $P(\boldsymbol{a}) = 0$ has a nontrivial integer solution. We confirm that the proportion of integer vectors $\boldsymbol{a} \in \mathbb{Z}^N$ in $\mathfrak{A}(A)$, whose associated equation $\langle \boldsymbol{a}, \nu_{d,n}(\boldsymbol{x})\rangle = 0$ is everywhere locally soluble, converges to a constant c_P as $A \to \infty$. Moreover, for each place v of $\mathbb{Q}$, if there exists a non-zero $\boldsymbol{b}_v \in \mathbb{Q}_v^N$ such that $P(\boldsymbol{b}_v) = 0$ and the variety $V_{\boldsymbol{b}_v}$ in $\mathbb{P}^{n-1}$ admits a smooth $\mathbb{Q}_v$-point, the constant c_P is positive.

Contents

1. Introduction

In this article, we study the p-adic and real solubility for projective varieties defined by forms with integer coefficients. Denote by $f(\boldsymbol{x}) \in \mathbb{Z}[x_1, x_2, \ldots, x_n]$ a homogeneous polynomial of degree d and let us write $V \subset \mathbb{P}^{n-1}$ for a projective variety defined by $f(\boldsymbol{x}) = 0$. One may ask if there exists $n_0 := n_0(d) \in \mathbb{N}$ such that whenever $n > n_0(d)$, the variety

2020 *Mathematics Subject Classification.* 11E76,14G25.
Key words and phrases. Projective variety, Local solubility.

V admits a $\mathbb{Q}_p$-point. The current knowledge is accessible to this type of question. In particular, it is known by Wooley [32] that it suffices to take $n_0 = d^{2^d}$ (see also [6], [15], [21]). As a different approach concerning the p-adic solubility for the variety V, the Ax–Kochen theorem [2] shows that a homogeneous polynomial $f(\boldsymbol{x}) \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree d with $n \geq d^2 + 1$ has a solution in $\mathbb{Q}_p^n \setminus \{\mathbf{0}\}$ for sufficiently large prime p with respect to d (see also [8]). It is also known that if $f(\boldsymbol{x}) \in \mathbb{Z}[x_1, x_2, \ldots, x_n]$ is an absolutely irreducible form of degree d over $\mathbb{F}_p$ with a prime p sufficiently large in terms of degree d, it follows by applying the Lang–Weil estimate (see [22] and [34, Theorem 3]) and the Hensel's lemma that the equation $f(\boldsymbol{x}) = 0$ has a solution in $\mathbb{Q}_p^n \setminus \{\mathbf{0}\}$. As for the real solubility for the variety V, one sees that if the degree d is odd, the variety V always admits a real point. For the degree d even, to the best of the authors' knowledge, it is not known how to determine, in general, whether the variety V has a real point or not.

One infers from the previous paragraph that if the number of variables n is not large enough, the main difficulties for verifying the p-adic solubility for the variety V occur from the case of small prime p. Nevertheless, thanks to the density lemma introduced in [25, Lemma 20], we are capable of obtaining some information about the p-adic solubility for the variety V even with small primes p. In order to describe this information, we temporarily pause and introduce some notation. Let n and d be natural numbers with $d \geq 2$. Let $N := N_{d,n} = \binom{n+d-1}{d}$. Let $\nu_{d,n} : \mathbb{R}^n \to \mathbb{R}^N$ denote the Veronese embedding, defined by listing all the monomials of degree d in n variables with the lexicographical ordering. Let $\langle \cdot, \cdot \rangle$ be the inner product on $\mathbb{A}^N$. We denote by $f_{\boldsymbol{a}}(\boldsymbol{x}) := \langle \boldsymbol{a}, \nu_{d,n}(\boldsymbol{x}) \rangle$ the homogeneous polynomial in n variables of degree d with coefficients $\boldsymbol{a} \in \mathbb{Z}^N$. Define $V \subset \mathbb{P}^{n-1} \times \mathbb{A}^N$ to be the subvariety given by $f_{\boldsymbol{a}}(\boldsymbol{x}) = 0$. Let $\pi : V \to \mathbb{A}^N$ be the projection onto the second factor. We denote by $V_{\boldsymbol{a}} := \pi^{-1}(\boldsymbol{a})$. For $A \in \mathbb{R}_{>0}$, we define

$$\mathfrak{A}(A) := \{\boldsymbol{a} \in \mathbb{Z}^N : \ \boldsymbol{a} \in [-A, A]^N\}. \tag{1.1}$$

The following quantity

$$\varrho_{d,n}^{\mathrm{loc}}(A) := \frac{\#\left\{\boldsymbol{a} \in \mathfrak{A}(A) : \ V_{\boldsymbol{a}}\left(\mathbb{R} \times \prod_{p \text{ prime}} \mathbb{Z}_p\right) \neq \emptyset\right\}}{\#\mathfrak{A}(A)}. \tag{1.2}$$

is the proportion of $\boldsymbol{a} \in \mathfrak{A}(A)$ such that $V_{\boldsymbol{a}}$ is *locally soluble*, i.e. that admit a real point and a p-adic point for all primes p. Thus, the behavior of $\varrho_{d,n}^{\mathrm{loc}}(A)$ provides the information about p-adic solubility of the varieties $V_{\boldsymbol{a}}$ for $\boldsymbol{a} \in \mathfrak{A}(A)$, even with small primes p. We also define c_∞ (resp. c_p) to be the measure of $\boldsymbol{a}$ in $[-1, 1]^N$ (resp. in $\mathbb{Z}_p^N$), whose associated variety $V_{\boldsymbol{a}}$ admits a real point (resp. a p-adic point). By using the density lemmas [25, Lemmas 20 and 21], Poonen and Voloch [26, Theorem 3.6] proved that whenever $n, d \geq 2$, one has

$$\lim_{A \to \infty} \varrho_{d,n}^{\mathrm{loc}}(A) = c, \tag{1.3}$$

where c is the product of c_∞ and c_p for all primes p.

In this paper, we investigate the proportion of $\boldsymbol{a}$ with locally soluble $V_{\boldsymbol{a}}$ in the type I thin set in the sense of Serre given by the vanishing locus of $P(\mathbf{t}) \in \mathbb{Z}[t_1, \dots, t_N]$. Here, we take P to be a non-singular form in N variables of degree $k \geq 2$. We define

$$\mathfrak{A}(A;P) := \{\boldsymbol{a} \in \mathfrak{A}(A) : P(\boldsymbol{a}) = 0\}.$$

Analogous to the quantity $\varrho^{\text{loc}}_{d,n}(A)$, we have the proportion of $\boldsymbol{a}$ in the thin set with locally soluble $V_{\boldsymbol{a}}$ defined as

$$\varrho^{P,\text{loc}}_{d,n}(A) := \frac{\#\left\{\boldsymbol{a} \in \mathfrak{A}(A;P) : V_{\boldsymbol{a}}\left(\mathbb{R} \times \prod_{p \text{ prime}} \mathbb{Z}_p\right) \neq \emptyset\right\}}{\#\mathfrak{A}(A;P)}. \tag{1.4}$$

Let μ_p denote the Haar measure on $\mathbb{Z}_p^N$ normalized to have total mass 1. For given measurable sets $S_p \subseteq \mathbb{Z}_p^N$ and $S_\infty \subseteq \mathbb{R}^N$ with the Haar measure μ_p and the Lebesgue measure, we define

$$\sigma_p(S_p) := \lim_{r\to\infty} p^{-r(N-1)} \# \left\{\boldsymbol{a} \pmod{p^r} \mid \boldsymbol{a} \in S_p \text{ and } P(\boldsymbol{a}) \equiv 0 \pmod{p^r}\right\}$$

and

$$\sigma_\infty(S_\infty) = \lim_{\eta\to 0+} (2\eta)^{-1} V_\infty(\eta),$$

where $V_\infty(\eta)$ is the volume of the subset of $\boldsymbol{y} \in S_\infty$ satisfying $|P(\boldsymbol{y})| < \eta$. Furthermore, for $d_1, d_2 \in \mathbb{N}$, we define $C_{n,d}(d_1, d_2)$ to be the rational number such that

$$C_{n,d}(d_1,d_2) = \frac{d!(n+d_1-1)!}{d_1!(n+d-1)!} + \frac{d!(n+d_2-1)!}{d_2!(n+d-1)!}.$$

Note that $C_{n,d}(d_1, d_2)$ belongs to $(0, 1)$ and maximizes at $(1, d-1)$ and $(d-1, 1)$; see Lemma 2.5. Our first main theorem shows that the proportion $\varrho^{P,\text{loc}}_{d,n}(A)$ converges to the product of "local proportions" as $A \to \infty$.

Theorem 1.1. *Suppose that $P(\mathbf{t}) \in \mathbb{Z}[t_1, \dots, t_N]$ is a non-singular form in N variables of degree k with $2 \leq k < (1 - C_{n,d}(1, d-1))N - 1$ and $(k-1)2^k < N$. Suppose that $P(\mathbf{t}) = 0$ has a nontrivial integer solution. For $n, d \geq 2$ such that*

$$(n,d) \neq (2,2), (2,3), (3,2), (3,3)$$

one has

$$\lim_{A\to\infty} \varrho^{P,loc}_{d,n}(A) = c_P,$$

where

$$c_P := \frac{\sigma_\infty(\pi(V(\mathbb{R})) \cap [-1,1]^N) \cdot \prod_p \sigma_p(\pi(V(\mathbb{Q}_p) \cap \mathbb{Z}_p^N)}{\sigma_\infty([-1,1]^N) \cdot \prod_p \sigma_p(\mathbb{Z}_p^N)}.$$

Remark 1. Since $P(\mathbf{t})$ is a non-singular form in N variables of degree k with $(k-1)2^k < N$ and $P(\mathbf{t}) = 0$ has a nontrivial integer solution, the classical argument (see [5] and [28, the proof of Theorem 1.3]) reveals that the quantity $\sigma_\infty([-1,1]^N) \cdot \prod_{p \text{ prime}} \sigma_p(\mathbb{Z}_p)$ is convergent and is bounded above and below by non-zero constants, respectively, depending on the polynomial $P(\mathbf{t})$. Then, on observing that

$$0 \leq \sigma_\infty(\pi(V(\mathbb{R})) \cap [-1,1]^N) \leq \sigma_\infty([-1,1]^N) \text{ and } 0 \leq \sigma_p(\pi(V(\mathbb{Q}_p) \cap \mathbb{Z}_p^N) \leq \sigma_p(\mathbb{Z}_p),$$

we infer that the infinite product $\sigma_\infty(\pi(V(\mathbb{R})) \cap [-1,1]^N) \cdot \prod_p \sigma_p(\pi(V(\mathbb{Q}_p) \cap \mathbb{Z}_p^N)$ in the numerator of c_P converges to a non-negative constant.

In [9, Corollary 1.6], Browning and Heath-Brown obtained the same constant for the case that P is a quadratic form of rank at least 5. The modicum computation reveals that when P is a non-singular quadratic form ($k = 2$), the conclusion of Theorem 1 holds for $N \geq 5$. Hence, we notice that the conclusion of Theorem 1.1 coincides with [9, Corollary 1.6] for non-singular quadratic forms. Furthermore, we emphasize that one could deal with P a quadratic form of rank at least 5 and obtain the same result in [9, Corollary 1.6], by using the argument described here. Additionally, we note that the argument in this paper seems plausible to be generalized for obtaining analogous results for thin sets defined by a system of non-singular forms.

Our second main theorem shows that under an additional condition on the polynomial P defining the thin set, the product of local proportions is strictly positive.

Theorem 1.2. *In addition to the setup of Theorem 1.1, suppose that for each place v of $\mathbb{Q}$, there exists $\boldsymbol{b}_v \in \mathbb{Q}_v^N \setminus \{\mathbf{0}\}$ such that the variety $V_{\boldsymbol{b}_v}$ in $\mathbb{P}_{\mathbb{Q}_v}^{n-1}$ admits a smooth $\mathbb{Q}_v$-point. Then, the constant c_P is positive.*

Notation

For a given vector $\boldsymbol{v} \in \mathbb{R}^N$, we write the i-th coordinate of $\boldsymbol{v}$ by $(\boldsymbol{v})_i$ or v_i. We use $\langle \cdot, \cdot \rangle$ for the inner product. We write $0 \leq \boldsymbol{x} \leq X$ or $\boldsymbol{x} \in [0, X]^s$ to abbreviate the condition $0 \leq x_1, \ldots, x_s \leq X$. For a prime p and vectors $\boldsymbol{v} \in \mathbb{R}^n$, we use $p^h \| \boldsymbol{v}$ when one has $p^h | v_i$ for all $1 \leq i \leq n$ but $p^{h+1} \nmid v_i$ for some $1 \leq i \leq n$. Throughout this paper, we use $\gg$ and $\ll$ to denote Vinogradov's well-known notation, and write $e(z)$ for $e^{2\pi i z}$. We use $A \asymp B$ when both $A \gg B$ and $A \ll B$ hold. We adopt the convention that when ϵ appears in a statement, then the statement holds for each $\epsilon > 0$, with implicit constants depending on ϵ.

Acknowledgement

The authors acknowledge support from NSF grant DMS-2001549 under the supervision of Trevor Wooley. The third author would like to thank James Cumberbatch and Dong-ha Kim for helpful discussion. Especially, the third author would like to thank Trevor Wooley for his constant encouragement to complete this work. Furthermore, the authors would like to thank Daniel Loughran and Nicholas Rome for letting us know many related works and providing very helpful comments on overall arguments in this paper.

2. Preliminaries

Throughout this section, we fix a non-singular form $P(\mathbf{t}) \in \mathbb{Z}[t_1, \dots, t_N]$ in N variables of degree $k \geq 2$. We also assume that $N > (k-1)2^k$. Let $\mathfrak{B}$ be a box in $[-1,1]^N \cap \mathbb{R}^N$. For given $A, B > 0$ and $\boldsymbol{r} \in \mathbb{Z}^N$ with $0 \leq \boldsymbol{r} \leq B-1$, we define

$$\mathcal{N}(A, \mathfrak{B}, B, \boldsymbol{r}, P) = \# \left\{ \boldsymbol{a} \in A\mathfrak{B} \cap \mathbb{Z}^N \middle|\ P(\boldsymbol{a}) = 0,\ \boldsymbol{a} \equiv \boldsymbol{r} \pmod{B} \right\}.$$

The first lemma in this section provides the asymptotic formula for

$$\mathcal{N}(A, \mathfrak{B}, B, \boldsymbol{r}, P) \text{ as } A \to \infty.$$

In advance of the statement of Lemma 2.1, we define $v_p(L)$ with $L \in \mathbb{Z}$ and p prime as the integer s such that $p^s \| L$.

Lemma 2.1. *Suppose that B is a natural number. Then, for a given $\boldsymbol{r} \in \mathbb{Z}^N$ with $0 \leq \boldsymbol{r} \leq B-1$ and for sufficiently large $A > 0$, there exists $\delta > 0$ such that*

$$\mathcal{N}(A, \mathfrak{B}, B, \boldsymbol{r}, P) = \prod_{p \nmid B} \sigma_p \cdot \prod_{p \mid B} \sigma_p^{B, \boldsymbol{r}} \cdot \sigma_\infty \cdot A^{N-k} + O(A^{N-k-\delta}),$$

where

$$\sigma_p := \lim_{l \to \infty} p^{-l(N-1)} \# \left\{ 1 \leq \boldsymbol{a} \leq p^l :\ P(\boldsymbol{a}) \equiv 0 \ (\mathrm{mod}\ p^l) \right\},$$

$$\sigma_p^{B, \boldsymbol{r}} := \lim_{l \to \infty} p^{-l(N-1)} \# \left\{ 1 \leq \boldsymbol{a} \leq p^l :\ P(\boldsymbol{a}) \equiv 0 \ (\mathrm{mod}\ p^l),\ \boldsymbol{a} \equiv \boldsymbol{r} \ (\mathrm{mod}\ p^{v_p(B)}) \right\}$$

and

$$\sigma_\infty := \sigma_\infty(\mathfrak{B}) = \lim_{\eta \to 0+} (2\eta)^{-1} V_\infty(\eta)$$

in which $V_\infty(\eta)$ is the volume of the subset of $\boldsymbol{y} \subseteq \mathfrak{B}$ satisfying $|P(\boldsymbol{y})| < \eta$. In particular, the implicit constant in $O(A^{N-k-\delta})$ depends on B and $P(\mathbf{t})$.

We record this lemma without proof because it is readily obtained by the previous results as follows. By repeating the argument of Birch in [5] replaced with the variables $\boldsymbol{x}$ imposed on the congruence condition $\boldsymbol{x} \equiv \boldsymbol{r} \pmod{B}$, we obtain an asymptotic formula for the number of integer solutions $\boldsymbol{x} \in [-A, A]^N$ of $P(\boldsymbol{x}) = 0$ with the congruence condition $\boldsymbol{x} \equiv \boldsymbol{r} \pmod{B}$ as $A \to \infty$. The main term of this asymptotic formula includes the product of p-adic densities and the singular integral. By applying the strategy proposed by Schmidt [29, 30] (see also a refined version [12, section 9]), the singular integral can be replaced by the real density σ_∞. Furthermore, we readily deduce by the coprimality between p and B that the product of p-adic densities in the main term becomes $\prod_{p \notin M} \sigma_p \cdot \prod_{p \in M} \sigma_p^{B, \boldsymbol{r}}$.

Thus, this yields the asymptotic formula for $N(A,\mathfrak{B},B,\boldsymbol{r},P)$ as desired in Lemma 2.1. For a detailed proof, see also [17, Lemma 4.4].

Lemma 2.2. *Suppose that A and Q are positive numbers with $Q \leq A$. Then, for a given $\boldsymbol{c} \in \mathbb{Z}^N$ with $1 \leq \boldsymbol{c} \leq Q$, we have*

$$\#\left\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N \middle|\ P(\boldsymbol{a}) = 0 \text{ and } \boldsymbol{a} \equiv \boldsymbol{c} \ (\mathrm{mod}\ Q)\right\} \ll (A/Q)^{N-k},$$

where the implicit constant may depend on $P(\mathbf{t})$.

Proof. By orthogonality, we have

$$\begin{aligned} &\#\left\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N \middle|\ P(\boldsymbol{a}) = 0 \text{ and } \boldsymbol{a} \equiv \boldsymbol{c} \ (\mathrm{mod}\ Q)\right\} \\ &= \int_0^1 \sum_{-A \leq Q\boldsymbol{y}+\boldsymbol{c} \leq A} e(\alpha P(Q\boldsymbol{y}+\boldsymbol{c}))d\alpha. \end{aligned} \tag{2.1}$$

By change of variable $\alpha = \beta/Q^k$, the last expression is seen to be

$$\begin{aligned} &Q^{-k} \int_0^{Q^k} \sum_{-A \leq Q\boldsymbol{y}+\boldsymbol{c} \leq A} e(\beta P(\boldsymbol{y}+\boldsymbol{c}/Q))d\beta \\ &\ll \sup_{\substack{1 \leq l \leq Q^k \\ l \in \mathbb{N}}} \int_{l-1}^{l} \sum_{-A \leq Q\boldsymbol{y}+\boldsymbol{c} \leq A} e(\beta P(\boldsymbol{y}+\boldsymbol{c}/Q))d\beta \\ &= \sup_{\substack{1 \leq l \leq Q^k \\ l \in \mathbb{N}}} \int_{l-1}^{l} \sum_{-A \leq Q\boldsymbol{y}+\boldsymbol{c} \leq A} e(\beta (P(\boldsymbol{y}) + g(\boldsymbol{y})))d\beta, \end{aligned} \tag{2.2}$$

where $g \in \mathbb{Q}[\boldsymbol{y}]$ is a polynomial of degree at most $k-1$.

On writing that $S(\beta) = \sum_{-A \leq Q\boldsymbol{y}+\boldsymbol{c} \leq A} e(\beta(P(\boldsymbol{y}) + g(\boldsymbol{y})))$, we claim that whenever $N > (k-1)2^k$, one has

$$\int_{l-1}^{l} S(\beta)d\beta \ll (A/Q)^{N-k}, \tag{2.3}$$

uniformly in l. Suppose that the inequality (2.3) holds. Then, on substituting (2.3) into the last expression of (2.2), it follows from (2.1) and (2.2) that we complete the proof of Lemma 2.2.

It remains to verify the inequality (2.3). For a given $l \in \mathbb{Z}$, we define the major arcs $\mathfrak{M}^l(H)$ by

$$\mathfrak{M}^l(H) = \bigcup_{\substack{0 \leq a \leq q \leq H \\ (q,a)=1}} \mathfrak{M}^l(H,q,a),$$

where

$$\mathfrak{M}^l(H,q,a) = \left\{\beta \in [l-1,l) :\ \left|\beta - (l-1) - \frac{a}{q}\right| \leq \frac{H}{q(A/Q)^k}\right\}.$$

Furthermore, we define the minor arcs $\mathfrak{m}^l := [l-1,l) \setminus \mathfrak{M}^l$.

For simplicity, we put $L = A/Q$. Let δ be a positive number with $\delta < 2^{-1-k}$. Define $I \in \mathbb{N}$ to be the minimum number satisfying $2^I L^\delta > (1/4)L^{k/2}$. We notice here that $I = O_{k,\delta}(\log L)$. By using those major and minor arcs dissections of $[l-1, l)$, defined in the previous paragraph, we deduce that

$$\int_{l-1}^{l} S(\beta)d\beta \ll S_1 + S_2 + S_3, \tag{2.4}$$

where

$$\begin{aligned} S_1 &= \int_{\mathfrak{M}^l(L^\delta)} |S(\beta)|d\beta \\ S_2 &= \sum_{i=1}^{I} \int_{\mathfrak{M}^l(2^i L^\delta)\setminus \mathfrak{M}^l(2^{i-1}L^\delta)} |S(\beta)|d\beta \\ S_3 &= \int_{\mathfrak{m}^l((1/4)L^{k/2})} |S(\beta)|d\beta. \end{aligned}$$

We shall show that whenever $N > (k-1)2^k$, each of S_1, S_2 and S_3 is $O(L^{N-k})$, which delivers the desired bound (2.3). We first obtain the upper bound for S_3. On noting that

$$\begin{aligned} S(\beta) &= \sum_{-A \le Q\boldsymbol{y}+\boldsymbol{c} \le A} e(\beta(P(\boldsymbol{y}) + g(\boldsymbol{y}))) \\ &= \sum_{-A \le Q\boldsymbol{y}+\boldsymbol{c} \le A} e((\beta - (l-1))P(\boldsymbol{y}) + \beta g(\boldsymbol{y})), \end{aligned}$$

and recalling that $P(\mathbf{t})$ is a non-singular form, we infer by [23, Lemma 3.6] with $R = 1$, $\alpha_1 = \beta - (l-1)$, $f_1(\boldsymbol{x}) = P(\boldsymbol{y})$, $G(\boldsymbol{x}) = \beta g(\boldsymbol{y}), \mathrm{dim} V_{f_1}^* = 0, \kappa = N$ that

$$\sup_{\beta \in \mathfrak{m}^l((1/4)L^{k/2})} |S(\beta)| \ll L^{N - kN/((k-1)2^k) + \epsilon}.$$

Hence, whenever $N > (k-1)2^k$, one has

$$S_3 \ll \sup_{\beta \in \mathfrak{m}^l((1/4)L^{k/2})} |S(\beta)| \cdot \int_{l-1}^{l} 1 d\beta \ll L^{N-k}. \tag{2.5}$$

Next, we derive the upper bound for S_2. Note that $\mathrm{mes}(\mathfrak{M}^l(H)) \ll H^2 L^{-k}$ and note again by [23, Lemma 3.6] that whenever $N > (k-1)2^k$ one has

$$\begin{aligned} \sup_{\beta \in \mathfrak{m}^l(2^{i-1}L^\delta)} |S(\beta)| &\ll L^N (2^{i-1}L^\delta)^{-N/((k-1)2^{k-1}) + \epsilon} \\ &\ll L^N (2^{i-1}L^\delta)^{-2+\epsilon} \cdot (2^{i-1}L^\delta)^{-1/((k-1)2^{k-1})}. \end{aligned}$$

Hence, one has

$$\int_{\mathfrak{M}^l(2^iL^\delta)\setminus\mathfrak{M}^l(2^{i-1}L^\delta)} |S(\beta)|d\beta \ll \text{mes}(\mathfrak{M}^l(2^iL^\delta)) \cdot \sup_{\beta\in\mathfrak{m}^l(2^{i-1}L^\delta)} |S(\beta)|$$
$$\ll L^{N-k} \cdot (2^iL^\delta)^{\epsilon-1/((k-1)2^{k-1})}$$
$$\ll L^{N-k-\eta},$$

for all $i = 1, 2, \ldots, I$ and with some $\eta > 0$. Therefore, we find by $I = O(\log L)$ that $S_2 = O(L^{N-k})$.

Lastly, we deduce the upper bound for S_1. For this, it is convenient to define differencing operators Δ_1 by

$$\Delta_1(P(\boldsymbol{x}); \boldsymbol{h}) = P(\boldsymbol{x} + \boldsymbol{h}) - P(\boldsymbol{x}),$$

and so we define Δ_j for $j \geq 2$ recursively by means of the relations

$$\Delta_j(P(\boldsymbol{x}); \boldsymbol{h}_1, \ldots, \boldsymbol{h}_j) = \Delta_1(\Delta_{j-1}(P(\boldsymbol{x}); \boldsymbol{h}_1, \ldots, \boldsymbol{h}_{j-1}); \boldsymbol{h}_j). \tag{2.6}$$

For given variables $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k)} \in \mathbb{Z}^N$, we define

$$\begin{aligned}&\phi_1(\boldsymbol{y}^{(k-1)}, \boldsymbol{y}^{(k)})\\ &:= \phi_1(\boldsymbol{y}^{(k-1)}, \boldsymbol{y}^{(k)}; \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)})\\ &= \Delta_{k-2}(P(\boldsymbol{y}^{(k-1)}); \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)}) - \Delta_{k-2}(P(\boldsymbol{y}^{(k)}); \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)}).\end{aligned}$$

Then, by applying the Cauchy-Schwarz inequality together with the classical Weyl differencing argument, we deduce that

$$S(\beta)^{2^{k-1}} \ll L^{(2^{k-1}-k)N} \sum_{\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(k-2)}} \sum_{\boldsymbol{y}^{(k-1)},\boldsymbol{y}^{(k)}\in\mathcal{D}L} e(\beta\phi_1(\boldsymbol{y}^{(k-1)}, \boldsymbol{y}^{(k)})), \tag{2.7}$$

where the variables $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)}$ run over $[-2L, 2L]^N$ and

$$\mathcal{D} := \mathcal{D}(\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)})$$

is a box in $[-2, 2]^N$ suitably defined through the classical Weyl differencing argument in terms of $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-2)}$.

We denote by $\mathcal{Y}(\beta; \mathcal{D})$ the exponential sum over $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k)}$ on the right hand side of (2.7). Obviously, we have

$$\mathcal{Y}(\beta; \mathcal{D}) \ll \sup_{\mathcal{B}\subseteq[-1,1]^N} |\mathcal{Y}(\beta; \mathcal{B})|, \tag{2.8}$$

where $\mathcal{B}$ is over boxes in $[-2, 2]^N$. For $\beta \in \mathfrak{M}^l(L^\delta, q, a)$ and for a given box $\mathcal{B} \subseteq [-2, 2]^N$, we derive the upper bound for $\mathcal{Y}(\beta; \mathcal{B})$. To do this, we make use of the classical argument of Birch [5, Section 5]. Hence, we shall be brief in some steps.

Put $\alpha = \beta - (l - 1) - a/q$. Write $\boldsymbol{y}^{(i)} = q\boldsymbol{x}^{(i)} + \boldsymbol{z}^{(i)}$ $(i = 1, \ldots, k)$, where $1 \leq \boldsymbol{z}^{(i)} \leq q$ and $\boldsymbol{x}^{(i)}$ runs over boxes so that $\|\boldsymbol{y}\|_\infty \leq 2L$ with $i = 1, \ldots, k-2$,

and $\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)}$ run over boxes so that $\boldsymbol{y}^{(k-1)}, \boldsymbol{y}^{(k)} \in \mathcal{B}L$. Then, we have
(2.9)
$$\mathcal{Y}(\beta;\mathcal{B}) = \sum_{1\leq \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k)}\leq q} e\Big(\frac{a}{q}\phi_1(\boldsymbol{z}^{(k-1)}, \boldsymbol{z}^{(k)}; \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k-2)})\Big)\mathcal{T}(\alpha, \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k)}),$$
where
$$\mathcal{T}(\alpha, \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k)}) = \sum_{\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(k)}} e(\alpha\phi_2(\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)})),$$
in which
$$\begin{aligned}
&\phi_2(\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)})\\
&:= \phi_2(\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)}; \boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(k-2)})\\
&= \phi_1(q\boldsymbol{x}^{(k-1)}+\boldsymbol{z}^{(k-1)}, q\boldsymbol{x}^{(k)}+\boldsymbol{z}^{(k)}; q\boldsymbol{x}^{(1)}+\boldsymbol{z}^{(1)},\ldots,q\boldsymbol{x}^{(k-2)}+\boldsymbol{z}^{(k-2)}).
\end{aligned}$$

Note that $\sup_{\boldsymbol{\gamma}\in[0,1]^N} |\boldsymbol{\gamma}\cdot\nabla e(\alpha\phi_2(\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)}))| \ll qL^{k-1}|\alpha|$, since $\phi_2(\boldsymbol{x}^{(k-1)}, \boldsymbol{x}^{(k)}))$ is a polynomial of degree k in $\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(k)}$. Hence, if we define
$$I(\alpha;\mathcal{B}) = \int_{[-2,2]^{(k-2)N}}\int_{\mathcal{B}^2} e(\alpha\phi_1(\boldsymbol{\eta}^{(k-1)}, \boldsymbol{\eta}^{(k)}; \boldsymbol{\eta}^{(1)}\ldots,\boldsymbol{\eta}^{(k-2)}))d\boldsymbol{\eta},$$
where $d\boldsymbol{\eta} = d\boldsymbol{\eta}^{(k-1)}d\boldsymbol{\eta}^{(k)}d\boldsymbol{\eta}^{(1)}\cdots d\boldsymbol{\eta}^{(k-2)}$, it follows by the classical argument in [5, Section 5] using multi-dimensional mean value theorem that
$$\mathcal{T}(\alpha, \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k)}) - (L/q)^{kN}I(L^k\alpha) \ll E_1+E_2, \tag{2.10}$$
where
$$\begin{aligned}
E_1 &= q(L/q)^{kN}L^{k-1}|\alpha|\\
E_2 &= (L/q)^{kN-1}.
\end{aligned}$$

Write
$$S(q,a) = \sum_{1\leq \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k)}\leq q} e\Big(\frac{a}{q}\phi_1(\boldsymbol{z}^{(k-1)}, \boldsymbol{z}^{(k)}; \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k-2)})\Big). \tag{2.11}$$
Then, on substituting (2.10) into (2.9), we obtain that
$$\mathcal{Y}(\beta;\mathcal{B}) = q^{-kN}S(q,a)I(L^k\alpha;\mathcal{B})L^{kN} + O(E(\beta)), \tag{2.12}$$
where $E(\beta) = q^{kN}(E_1+E_2)$. By the definition E_1 and E_2 together with the fact that $\beta\in\mathfrak{M}^l(L^\delta,q,a)$, we deduce that
$$E(\beta) \ll L^{kN-1+\delta}. \tag{2.13}$$

Meanwhile, it follows from (2.7) together with (2.8) that
$$S_1 \ll L^{(1-2^{1-k}k)N}\int_{\mathfrak{M}^l(L^\delta)} \sup_{\mathcal{B}\subseteq[-1,1]^N} |\mathcal{Y}(\beta;\mathcal{B})|^{2^{1-k}}d\beta. \tag{2.14}$$

Let us write
$$\mathfrak{S}(L^\delta) = \sum_{1\leq q\leq L^\delta}\sum_{\substack{1\leq a\leq q\\ (q,a)=1}} (q^{-kN}S(q,a))^{2^{1-k}}. \tag{2.15}$$

Then, on noting that $\mathrm{mes}(\mathfrak{M}^l(L^\delta)) \ll L^{2\delta-k}$ and substituting (2.13) into (2.12) and that into (2.14), we find that
(2.16)
$$S_1 \ll L^{N-k} \cdot \mathfrak{S}(L^\delta) \cdot \int_{|\alpha_1| \le \frac{L^\delta}{q}} \sup_{\mathcal{B} \subseteq [-1,1]^N} |I(\alpha_1;\mathcal{B})|^{2^{1-k}} d\alpha_1 + L^{N-k+2\delta-(1-\delta)2^{1-k}},$$

where we have assumed that $\sup_{\mathcal{B}\subseteq[-1,1]^N} |\mathcal{Y}(\beta;\mathcal{B})|^{2^{1-k}}$ and $\sup_{\mathcal{B}\subseteq[-1,1]^N} |I(L^k\alpha;\mathcal{B})|^{2^{1-k}}$ is a measurable function and we have used a change of variable $\alpha_1 = L^k\alpha$. We will prove the measurability for these functions at the end of the proof of this lemma.

As the endgame, we shall use [23, Lemma 3.4] to obtain the upper bound the functions $\mathfrak{S}(L^\delta)$ and $I(\alpha_1 : \mathcal{B})$. Recall the definition of $\mathcal{Y}(\beta;\mathcal{B})$. By applying the Weyl's differencing argument, we see that

$$\mathcal{Y}(\beta;\mathcal{B}) \le \sum_{\boldsymbol{y}^{(1)},\ldots,\boldsymbol{y}^{(k-1)}} \Big| \sum_{\boldsymbol{y}^{(k)} \in \mathcal{B}_1 L} e(\beta\Delta_{k-1}(P(\boldsymbol{y}^{(k)}); \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(k-1)})) \Big|, \tag{2.17}$$

where $\mathcal{B}_1 \subseteq [-2,2]^N$ is a box suitably defined through the classical Weyl differencing argument in terms of $\boldsymbol{y}^{(k-1)}$. By [23, Lemma 3.4] with $P = L, R = 1, \alpha_1 = \beta - (l-1), f_1 = P(\boldsymbol{y}), \dim V_{f_1}^* = 0$ and $\kappa = N$, we infer that whenever $\beta \notin \mathfrak{M}^l(H)$ with $H \le L^{k-1}$, one has

$$\mathcal{Y}(\beta;\mathcal{B}) = \mathcal{Y}(\beta - (l-1);\mathcal{B}) \ll L^{kN+\epsilon} H^{-N/(k-1)}. \tag{2.18}$$

Meanwhile, we note that if we temporarily assume that $|\beta - (l-1)| < (1/2)L^{-k/2}$, it follows that $\mathfrak{M}^l(|\beta-(l-1)|L^k, q, a)$ are disjoint over $0 \le a \le q \le |\beta-(l-1)|L^k$ with $(q,a) = 1$, and β is on the edge of $\mathfrak{M}^l(|\beta-(l-1)|L^k, 1, 0)$. Thus, one sees that $\beta \notin \mathfrak{M}^l(H)$ with $H = |\beta-(l-1)|L^{k-\epsilon}$ for any $\epsilon > 0$. Hence, it follows by (2.18) that whenever $|\beta-(l-1)| < (1/2)L^{-k/2}$, one has

$$\mathcal{Y}(\beta;\mathcal{B}) \ll L^{kN+\epsilon}(|\beta-(l-1)|L^k)^{-N/(k-1)}. \tag{2.19}$$

First, we claim that for any $\alpha_1 > 0$ one has

$$I(\alpha_1;\mathcal{B}) \ll \min\{1, |\alpha_1|^{-N/(k-1)+\epsilon}\}. \tag{2.20}$$

The argument for this is based on [5, Lemma 5.2]. The estimate $I(\alpha_1;\mathcal{B}) \ll 1$ is trivial. To obtain the second bound, we may assume that $|\alpha_1| > 1$. Taking $a = 0$ and $q = 1$ in (2.12), we deduce from the definition of E_1 and E_2 that

$$\mathcal{Y}(\beta;\mathcal{B}) = I(L^k\alpha;\mathcal{B})L^{kN} + O((|\alpha|L^k+1)L^{kN-1}), \tag{2.21}$$

with $\alpha = \beta-(l-1)$. Then, on writing $L^k\alpha = \alpha_1$, it follows by (2.21) together with (2.19) that whenever $1 < |\alpha_1| < (1/2)L^{k/2}$, one has

$$I(\alpha_1;\mathcal{B}) \ll |\alpha_1|^{-N/(k-1)}L^\epsilon + |\alpha_1|L^{-1}. \tag{2.22}$$

On observing that $I(\alpha_1;\mathcal{B})$ does not depend on L, by taking $L = |\alpha_1|^{1+N/(k-1)}$, the inequality (2.22) delivers (2.20).

Recall the definition (2.11) of $S(q,a)$. Next, we claim that whenever $(q,a)=1$, one has

$$S(q,a) \ll q^{kN-N/(k-1)+\epsilon}. \tag{2.23}$$

By applying the Weyl's differencing argument, we see that

$$S(q,a) \leq \sum_{1\leq \boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k-1)}\leq q} \Big| \sum_{1\leq \boldsymbol{z}^{(k)}\leq q} e\big(\frac{a}{q}\Delta_{k-1}(P(\boldsymbol{z}^{(k)});\boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(k-1)}\big)\Big|. \tag{2.24}$$

Meanwhile, note that for any $\epsilon>0$, one cannot find $q'\in\mathbb{N}, a'\in\mathbb{Z}$ with $(q',a')=1$ and $1\leq q'\leq q^{1-\epsilon}$ satisfying

$$|q'a-qa'| \leq q^{1-\epsilon}q^{1-k} \tag{2.25}$$

Hence, by applying again [23, Lemma 3.4] with $P=q,\ R=1,\ \alpha_1=a/q,\ f_1=P(\boldsymbol{y}),\ \dim V_{f_1}^*=0,\ X=q^{(1-\epsilon)/(k-1)}$ and $\kappa=N$, we infer that whenever $(q,a)=1$, one sees that the inequality (2.23) holds.

On substituting (2.23) into (2.15), we see that whenever $N>(k-1)2^k$

$$\mathfrak{S}(L^{\delta}) \ll 1. \tag{2.26}$$

Hence, on substituting (2.20) into (2.16), one infers that $S_1=O(L^{N-k})$, since $0<\delta<2^{-1-k}$. We have shown thus far that $S_1+S_2+S_3=O(L^{N-k})$, which establishes (2.3) by (3.4).

We complete the proof of this lemma by confirming that the functions $g_1(\beta):=\sup_{\mathcal{B}\subseteq[-1,1]^N}|\mathcal{Y}(\beta;\mathcal{B})|^{2^{1-k}}$ and $g_2(\alpha):=\sup_{\mathcal{B}\subseteq[-1,1]^N}|I(L^k\alpha;\mathcal{B})|^{2^{1-k}}$ are measurable. It is enough to show that for any given $k>0$, the set $\{\alpha\in\mathbb{R}: g_i(\alpha)>k\}$ $(i=1,2)$ is open. Let $g_2(\alpha_0)>k$. Then, there exists $\mathcal{B}\subseteq[-1,1]^N$ such that $|I(L^k\alpha;\mathcal{B})|^{2^{1-k}}>k$. On recalling the definition of $I(\cdot;\mathcal{B})$, one sees that there exists $\epsilon>0$ such that whenever $|\alpha-\alpha_0|<\epsilon$ we have $|I(L^k\alpha;\mathcal{B})|^{2^{1-k}}>k$, which proves that the set $\{\alpha\in\mathbb{R}: g_2(\alpha)>k\}$ is open as desired. By applying the same argument, we see that $g_1(\beta)$ is a measurable function.

□

The following lemma provides an upper asymptotic estimate for the number of integer points in the variety cut by the polynomial $P(\mathbf{t})\in\mathbb{Z}[\mathbf{t}]$, that reduce modulo p, for some sufficiently large $p>M$, to an $\mathbb{F}_p$-point of an another given variety $Y\subset\mathbb{A}^n$ defined over $\mathbb{Z}$.

Lemma 2.3. *Let $\mathfrak{B}$ be a compact region in $\mathbb{R}^N$ having a finite measure, and let Y be any closed subscheme of $\mathbb{A}^N_{\mathbb{Z}}$ of codimension $r\geq 1$. Let A and M be positive real numbers. Suppose that $r-1>k$. Then, there exists*

$A_0 := A_0(P(\mathbf{t})) \in \mathbb{R}_{>0}$ *such that whenever* $A > A_0$*, we have*

(2.27)
$$\#\left\{\boldsymbol{a} \in A\mathfrak{B} \cap \mathbb{Z}^N \,\middle|\, \begin{array}{l} (i)\ \boldsymbol{a}\ (mod\ p) \in Y(\mathbb{F}_p)\ for\ some\ prime\ p > M \\ (ii)\ P(\boldsymbol{a}) = 0 \end{array}\right\} \ll \frac{A^{N-k}}{M^{r-k-1}\log M} + A^{N-r+1},$$

where the implicit constant may depend on $\mathfrak{B}$ *and* Y*.*

In [4, Theorem 3.3], Bhargava provided an upper asymptotic estimate that

$$\#\left\{\boldsymbol{a} \in A\mathfrak{B} \cap \mathbb{Z}^N \,\middle|\, \boldsymbol{a}\ (\mathrm{mod}\ p) \in Y(\mathbb{F}_p) \text{ for some prime } p > M\right\} \ll \frac{A^N}{M^{r-1}\log M} + A^{N-r+1}. \tag{2.28}$$

Furthermore, as alluded in [4, Remark 3.4], the bound in (2.28) can be achieved for suitable choices of Y. Thus, this bound is essentially optimal.

In the proof of Lemma 2.3, we mainly adopt the argument in [4, Theorem 3.3] though, we freely admit that the bound in (2.27) seems not optimal order of magnitude of the bound. Especially, one finds that the second term is trivially obtained by that in (2.28). We are independently interested in a sharper bound in (2.27) and expect that one might be able to improve this bound. Nevertheless, since the strength of the upper asymptotic estimate in Lemma 2.3 is enough for our purpose, we do not put our effort into optimizing this upper bound in this paper.

Proof of Lemma 2.3. We can and do assume that Y is irreducible. Otherwise, we can take its irreducible components and add up the equation (2.27) to deduce general cases. Since Y has codimension r, there exists $f_1, \ldots, f_r \in \mathbb{Z}[t_1, \ldots, t_N]$ such that the vanishing locus $V(f_1, \ldots, f_r)$ contains an irreducible component of codimension r containing Y. Indeed, we can assume that Y equals to the irreducible component, as they have the same underlying reduced subscheme, and we only consider $\mathbb{Z}$ or $\mathbb{F}_p$-points of them. By [4, Lemma 3.1], the number of $\boldsymbol{a} \in A\mathfrak{B} \cap Y(\mathbb{Z})$ is $\ll A^{N-r}$. (Note that $A\mathfrak{B} \cap Y(\mathbb{Z})$ equals to $A\mathfrak{B} \cap \mathbb{Z}^N \cap Y(\mathbb{R})$) Thus, it suffices to find an upper bound of the size of the set by

$$\#\left\{\boldsymbol{a} \in A\mathfrak{B} \cap \mathbb{Z}^N \,\middle|\, \begin{array}{l} (i)\ P(\boldsymbol{a}) = 0 \\ (ii)\ \boldsymbol{a}\ (\mathrm{mod}\ p) \in Y(\mathbb{F}_p) \text{ for some prime } p > M \\ (iii)\ \boldsymbol{a} \notin Y(\mathbb{Z}) \end{array}\right\} \ll \frac{A^{N-k}}{M^{r-k-1}\log M} + A^{N-r+1}.$$

We may assume that $r > k+1$, since $r - 1 > k$ from the hypothesis in the statement of Lemma 2.3. We shall find an upper bound of a slightly larger set by

$$\#\left\{(\boldsymbol{a},p)\middle|\begin{array}{l}(i)\ \boldsymbol{a}\in A\mathfrak{B}\cap\mathbb{Z}^N \text{ and } P(\boldsymbol{a})=0\\(ii)\ p>M \text{ a prime and } \boldsymbol{a}\ (\mathrm{mod}\ p)\in Y(\mathbb{F}_p)\\(iii)\ \boldsymbol{a}\notin Y(\mathbb{Z})\end{array}\right\} \ll \frac{A^{N-k}}{M^{r-k-1}\log M}+A^{N-r+1}. \tag{2.29}$$

First, we count pairs $(\boldsymbol{a},p)$ on the left hand side of (2.29) for each prime p satisfying $p\leq A$; such primes arise only when $A>M$. Meanwhile, by Lemma 2.2 we find that for a given $\boldsymbol{c}\in[1,p]^N$, the number of integer solutions $\boldsymbol{a}\in[-A,A]^N$ of $P(\boldsymbol{a})=0$ with the congruence condition $\boldsymbol{a}\equiv\boldsymbol{c}\ (\mathrm{mod}\ p)$ is $O((A/p)^{N-k})$. Then, since $\#Y(\mathbb{F}_p)=O(p^{N-r})$, we see that the number of $\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N$ such that $\boldsymbol{a}\ (\mathrm{mod}\ p)$ is in $Y(\mathbb{F}_p)$ is $O(p^{N-r})\cdot O((A/p)^{N-k})=O(A^{N-k}/p^{r-k})$. Thus the total number of desired pairs $(\boldsymbol{a},p)$ with $p\leq A$ is at most

$$\#\left\{(\boldsymbol{a},p)\middle|\begin{array}{l}(i)\ \boldsymbol{a}\in A\mathfrak{B}\cap\mathbb{Z}^N \text{ and } P(\boldsymbol{a})=0\\(ii)\ A\geq p>M \text{ a prime and } \boldsymbol{a}\ (\mathrm{mod}\ p)\in Y(\mathbb{F}_p)\\(iii)\ \boldsymbol{a}\notin Y(\mathbb{Z})\end{array}\right\} \ll \sum_{M<p\leq A} O\left(\frac{A^{N-k}}{p^{r-k}}\right)=O\left(\frac{A^{N-k}}{M^{r-k-1}\log M}\right). \tag{2.30}$$

Next, we count pairs $(\boldsymbol{a},p)$ with $p>A$. It follows from (see the equation (17) in the proof of [4, Theorem 3.3]) that

$$\#\left\{(\boldsymbol{a},p)\middle|\begin{array}{l}(i)\ \boldsymbol{a}\in A\mathfrak{B}\cap\mathbb{Z}^N \text{ and } P(\boldsymbol{a})=0\\(ii)\ p>A \text{ a prime and } \boldsymbol{a}\ (\mathrm{mod}\ p)\in Y(\mathbb{F}_p)\\(iii)\ \boldsymbol{a}\notin Y(\mathbb{Z})\end{array}\right\} \ll \#\left\{(\boldsymbol{a},p)\middle|\begin{array}{l}(i)\ \boldsymbol{a}\in A\mathfrak{B}\cap\mathbb{Z}^N\\(ii)\ p>A \text{ a prime and } \boldsymbol{a}\ (\mathrm{mod}\ p)\in Y(\mathbb{F}_p)\\(iii)\ \boldsymbol{a}\notin Y(\mathbb{Z})\end{array}\right\} \ll A^{N-r+1}. \tag{2.31}$$

Therefore, we find by (2.30) and (2.31) that the inequality (2.29) holds. Hence, we complete the proof of Lemma 2.3. □

Next, we will prove that most of $f_{\boldsymbol{a}}(\boldsymbol{x})$ are irreducible. Let Y be a subset of $\mathbb{A}^N_{\mathbb{Z}}$ defined to be

$$Y:=\left\{\boldsymbol{a}\in\mathbb{A}^N_{\mathbb{Z}}\,\middle|\, f_{\boldsymbol{a}}(\boldsymbol{x}) \text{ is reducible over } \mathbb{C}\right\}. \tag{2.32}$$

Our goal here is to show that Y is, in fact, an algebraic variety and that the codimension of Y is strictly greater than a constant depending on n and d. Here, we fix $r:=\mathrm{codim}_{\mathbb{A}^N_{\mathbb{Z}}} Y$. To show the claim, we record the following two useful lemmas:

Lemma 2.4. *For any integers $n \geq 3$ and $d \geq 3$,*

$$\frac{(n+d)(n+d-1)}{n-1} < \binom{n+d}{d}$$

Proof. We do this by induction on n. When $n = 3$, we have

$$\begin{aligned} 3 &< 3+d-2 \\ \frac{3(3+d)(3+d-1)}{2} &< \frac{(3+d)(3+d-1)(3+d-2)}{2} \\ \frac{(3+d)(3+d-1)}{2} &< \frac{(3+d)(3+d-1)(3+d-2)}{3 \cdot 2} = \binom{n+d}{d} \end{aligned}$$

Suppose that the lemma is true for n. Then,

$$\begin{aligned} \frac{(n+d+1)(n+d)}{n} &= \frac{(n+d+1)(n+d)(n+d-1)(n-1)}{n(n-1)(n+d-1)} \\ &< \frac{(n+d)!}{n!d!} \cdot \frac{(n+d+1)(n-1)}{n(n+d-1)} \\ &= \frac{(n+d+1)!}{(n+1)!d!} \cdot \frac{(n+1)(n-1)}{n(n+d-1)} \\ &< \frac{(n+d+1)!}{(n+1)!d!} \end{aligned} \tag{2.33}$$

The last inequality follows from $\frac{(n-1)(n+1)}{n(n+d-1)} < 1$. □

Lemma 2.5. *Let $n \geq 3$ and $d \geq 3$ be integers. Suppose that d_1 and d_2 are natural numbers with $d = d_1 + d_2$. Let $C_{n,d}(d_1, d_2)$ be a rational number defined as*

$$C_{n,d}(d_1, d_2) = \frac{d!(n+d_1-1)!}{d_1!(n+d-1)!} + \frac{d!(n+d_2-1)!}{d_2!(n+d-1)!}$$

Then, for given n and d, the quantity $C_{n,d}$ attains the maximum value when $(d_1, d_2) = (1, d-1)$ or $(d_1, d_2) = (d-1, 1)$. Furthermore, its maximum is strictly less than 1.

Proof. Without loss of generality, we assume that $d_1 \leq d_2$. We shall first show that one has

$$C_{n,d}(d_1, d_2) \leq C_{n,d}(d_1 - 1, d_2 + 1). \tag{2.34}$$

In order to verify the inequality (2.34), we observe that whenever $n \geq 3$ one has

$$(n-1) \cdot \frac{(n+d_1-2)!}{d_1!} \leq (n-1) \cdot \frac{(n+d_2-1)!}{(d_2+1)!}.$$

Equivalently, this is seen to be

$$\frac{(n+d_1-1)!}{d_1!} - \frac{(n+d_1-2)!}{(d_1-1)!} \leq \frac{(n+d_2)!}{(d_2+1)!} - \frac{(n+d_2-1)!}{d_2!},$$

and thus

$$\frac{(n+d_1-1)!}{d_1!}+\frac{(n+d_2-1)!}{d_2!}\le\frac{(n+d_2)!}{(d_2+1)!}+\frac{(n+d_1-2)!}{(d_1-1)!}. \tag{2.35}$$

Therefore, we find from (2.35) that

$$\frac{(n+d-1)!}{d!}C_{n,d}(d_1,d_2)\le\frac{(n+d-1)!}{d!}C_{n,d}(d_1-1,d_2+1).$$

This confirms the inequality (2.34). Then, by applying (2.34) iteratively, we conclude that the quantity $C_{n,d}$ attains the maximum value when $(d_1,d_2)=(1,d-1)$ or $(d_1,d_2)=(d-1,1)$.

Next, we deduce by applying Lemma 2.4 that

$$\begin{aligned} C_{n,d}(1,d-1) &= \frac{d!n!}{(n+d-1)!}+\frac{d!(n+d-2)!}{(d-1)!(n+d-1)!}\\ &= \frac{d!n!}{(n+d-1)!}+\frac{d}{n+d-1}\\ &= \frac{n+d}{\binom{n+d}{d}}+\frac{d}{n+d-1}\\ &< \frac{(n-1)(n+d)}{(n+d)(n+d-1)}+\frac{d}{n+d-1}\\ &= \frac{n-1}{n+d-1}+\frac{d}{n+d-1}=1. \end{aligned} \tag{2.36}$$

Therefore, this completes the proof of Lemma 2.5. ☐

Proposition 2.6. *Recall the definition of $Y\subseteq\mathbb{A}^N_{\mathbb{Z}}$ in (2.32). Then, Y is an affine variety. Further, let r be the codimension of Y. Then, one has $r>(1-C_{n,d}(1,d-1))N-1$.*

Proof. Let $\langle \boldsymbol{a}, v_{d,n}(\boldsymbol{x})\rangle$ be a reducible polynomial. Then, we write $\langle \boldsymbol{a}, v_{d,n}(\boldsymbol{x})\rangle = f^{(1)}(\boldsymbol{x})f^{(2)}(\boldsymbol{x})$. Here, $f^{(1)}(\boldsymbol{x})$ and $f^{(2)}(\boldsymbol{x})$ are homogeneous whose degrees are strictly less than d. Let $d_i=\deg(f^{(i)}(\boldsymbol{x}))$ and $t=\binom{n+d_1-1}{n-1}$. Then,

$$\langle \boldsymbol{a}, v_{d,n}(\boldsymbol{x})\rangle = \underbrace{(u_1x_1^{d_1}+\cdots u_tx_n^{d_1})}_{=f^{(1)}(\boldsymbol{x})}\underbrace{(u_{t+1}x_1^{d_2}+\cdots+u_{M(d_1,d_2)}x_n^{d_2})}_{=f^{(2)}(\boldsymbol{x})} \tag{2.37}$$

Let $Y_{(d_1,d_2)}\subset Y$ be a subset where $\langle \boldsymbol{a}, v_{d,n}(\boldsymbol{x})\rangle$ seperates as (2.37). Comparing the coefficients in (2.37) for both sides, we attain a_i's as a polynomial of $u_1,\dots,u_M$. Now, let us write $a_i=g_i(u_1,\dots,u_{M(d_1,d_2)})$. Consider a map $\varphi_{(d_1,d_2)}:\mathbb{Z}[t_1,\dots,t_N]\to\mathbb{Z}[u_1,\dots,u_{M(d_1,d_2)}]$ sending $t_i\mapsto g_i(u_1,\dots,u_{M(d_1,d_2)})$. Then, by the construction, $Y_{(d_1,d_2)}=V(\ker\varphi_{(d_1,d_2)})$, and so $Y=\bigcup Y_{(d_1,d_2)}$ is an affine variety.

Now, we prove $r>(1-C_{n,d}(1,d-1))N-1$. We will instead find the upper bound of the dimension of Y. Let $M=\max M(d_1,d_2)$. Since $\mathbb{Z}[t_1,\dots,t_N]/\ker\varphi_{(d_1,d_2)}$ injects into $\mathbb{Z}[u_1,\dots,u_{M(d_1,d_2)}]$, $\dim Y$ is less than or equal to M. Hence, it suffices to show $M\le N-(1-C_{n,d}(1,d-1))N$.

Note that we have $M(d_1,d_2) = C_{d,n}(d_1,d_2)N \le N - (1 - C_{n,d}(1,d-1))N$. Indeed, we have

$$\begin{aligned} M(d_1,d_2) &= \binom{n+d_1-1}{n-1} + \binom{n+d_2-1}{n-1} \\ &= \frac{(n+d_1-1)!}{(n-1)!d_1!} + \frac{(n+d_2-1)!}{(n-1)!d_2!} \\ &= \frac{(n+d-1)!}{d!(n-1)!} \underbrace{\left(\frac{d!(n+d_1-1)!}{d_1!(n+d-1)!} + \frac{d!(n+d_2-1)!}{d_2!(n+d-1)!}\right)}_{C_{n,d}(d_1,d_2)} \\ &= C_{n,d}(d_1,d_2)N \le N - (1 - C_{n,d}(1,d-1))N \end{aligned}$$

The latter inequality is by Lemma 2.5. Now, since $\dim Y \le M \le N - (1 - C_{n,d}(1,d-1))N$, we have $(1 - C_{n,d}(1,d-1))N - 1 < N - \dim Y = r$, which is desired. □

Remark 2. Using the identity of Lemma 2.4 and Lemma 2.5, it is not difficult to show that $(1 - C_{n,d}(1,d-1))N - 1$ is a positive integer. In fact, if $d \ge 4$, the constant is strictly greater than 2.

Here, in order to justify the positivity, we need to show

$$1 - \frac{1}{N} > C_{n,d}(1,d-1),$$

or equivalently,

$$1 - C_{n,d}(1,d-1) > \frac{1}{N} \tag{2.38}$$

First of all, $N = \binom{n+d-1}{d}$ and by Lemma 2.4, we have

$$\frac{1}{N} < \frac{n-2}{(n+d-1)(n+d-2)} \tag{2.39}$$

(Note: In order to make it strictly greater than 2, i.e., $(1-C_{n,d}(1,d-1))N - 1 > 2$, it suffices to show $1 - C_{n,d}(1,d-1) > \frac{3}{N}$. i.e., we just need to vary the factor of (2.39)). Hence, it suffices to bound the left-hand side of (2.38) below by the right-hand side of (2.39). Using the representations in (2.36),

$$\begin{aligned} 1 - C_{n,d}(1,d-1) &= \left(\frac{n-1}{n+d-1} + \frac{d}{n+d-1}\right) - \left(\frac{n+d}{\binom{n+d}{d}} + \frac{d}{n+d-1}\right) \\ &= \frac{n-1}{n+d-1}\left(1 - \frac{\frac{(n+d)(n+d-1)}{n-1}}{\binom{n+d}{d}}\right) \\ &\overset{(*)}{>} \frac{n-1}{n+d-1}\left(1 - \frac{n(n-2)}{(n-1)(n+d-2)}\right) \\ &= \frac{(n-1)(d-1)+1}{(n+d-1)(n+d-2)} > \frac{n-2}{(n+d-1)(n+d-2)} \end{aligned}$$

The last inequality is due to the fact that $d \geq 3$ (Here, if $d \geq 4$, the last inequality gets us $1 - C_{n,d}(1, d-1) > \frac{3}{N}$). The inequality $(*)$ is true since

$$\frac{(n+d)(n+d-1)}{n-1} < \frac{(n+d)!}{n!d!} \cdot \frac{n(n-2)}{(n-1)(n+d-2)}$$

by applying (2.33) for $n-1$.

3. Proof of Theorem 1.1 and Theorem 1.2

In this section, we provide the proofs of Theorem 1.1 and 1.2. To ease the notations, we denote by

$$\begin{aligned} T_\infty &:= \pi(V(\mathbb{R})) \cap [-1,1]^N \\ T_p &:= \pi(V(\mathbb{Q}_p)) \cap \mathbb{Z}_p^N \ \forall p \text{ prime.} \end{aligned}$$

We begin this section with a lemma on the measurability of the sets T_∞ and T_p.

Lemma 3.1. *The sets $T_\infty \subseteq \mathbb{R}^N$ and $T_p \subseteq \mathbb{Z}_p^N$ are measurable with the Lebesgue measure and the Haar measure μ_p, respectively.*

Proof. This follows from a version of Tarski–Seidenberg–Macintyre theorem which implies that T_∞ and T_p are semialgebraic sets (see [24, Theorem 3]). □

In advance of the proofs of Theorem 1.1 and Theorem 1.2, we provide some definitions and observations. We consider the natural map

$$\begin{aligned} \Phi^A: \ [-A,A]^N \cap \mathbb{Z}^N &\to [-1,1]^N \times \prod_{p \text{ prime}} \mathbb{Z}_p^N \\ \boldsymbol{a} &\mapsto \left(\frac{\boldsymbol{a}}{A}, \boldsymbol{a}, \dots, \boldsymbol{a}, \dots\right). \end{aligned}$$

We sometimes use a different order of primes in the product $\prod_{p \text{ prime}} \mathbb{Z}_p$, for notational convenience.

Furthermore, for a given subset $\mathcal{U}$ of $[-1,1]^N \times \prod_{p \text{ prime}} \mathbb{Z}_p^N$ and a given polynomial $P(\mathbf{t}) \in \mathbb{Z}[t_1, \dots, t_N]$, we define a quantity $\boldsymbol{d}(\mathcal{U}, A; P)$ by

$$\boldsymbol{d}(\mathcal{U}, A; P) := \frac{\#\left\{\boldsymbol{a} \in (\Phi^A)^{-1}(\mathcal{U}) \middle|\ P(\boldsymbol{a}) = 0\right\}}{\#\left\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N \middle|\ P(\boldsymbol{a}) = 0\right\}}. \tag{3.1}$$

We say that a subset of $\mathbb{Z}_p$ is an open interval if it has the form $\{x \in \mathbb{Z}_p |\ |x-a|_p \leq b\}$ for some $a \in \mathbb{Z}_p$ and $b \in \mathbb{R}$. Furthermore, by an open box I_p in $\mathbb{Z}_p^N$ with a given prime p, we mean a Cartesian product of open intervals. Suppose that $\mathfrak{B}$ be an open box in $[-1,1]^N \cap \mathbb{R}^N$. For a given set $\mathfrak{p}$ of prime numbers, it follows that

$$\begin{aligned}&\boldsymbol{d}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N, A; P\bigg)\\&=\frac{\#\bigg\{\boldsymbol{a}\in(\Phi^A)^{-1}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N\bigg)\bigg|\ P(\boldsymbol{a})=0\bigg\}}{\#\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N\middle|\ P(\boldsymbol{a})=0\right\}}.\end{aligned}$$

We observe that the set $(\Phi^A)^{-1}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N\bigg)$ can be viewed by a set of integers in $A\mathfrak{B}\cap\mathbb{Z}^N$ satisfying certain congruence conditions associated with the radius and the center of the open intervals defining I_p. Thus, we infer from the Chinese remainder theorem that there exists $B\in\mathbb{Z}$ whose prime divisors are in $\mathfrak{p}$, and $\boldsymbol{r}\in\mathbb{Z}^N$ such that

$$\boldsymbol{d}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N, A; P\bigg)=\frac{\#\bigg\{\boldsymbol{a}\in A\mathfrak{B}\cap\mathbb{Z}^N\bigg|\ P(\boldsymbol{a})=0,\ \boldsymbol{a}\equiv\boldsymbol{r}\bmod B\bigg\}}{\#\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N\middle|\ P(\boldsymbol{a})=0\right\}}.$$

Then, by applying Lemma 2.1, we obtain

$$\begin{aligned}&\boldsymbol{d}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N, A; P\bigg)\\&=\frac{\frac{1}{\zeta(N-k)}\prod_{p\notin\mathfrak{p}}\sigma_p\cdot\prod_{p\in\mathfrak{p}}\sigma_p^{B,\boldsymbol{r}}\cdot\sigma_\infty(\mathfrak{B})+O(A^{-\delta})}{\frac{1}{\zeta(N-k)}\prod_p\sigma_p\cdot\sigma_\infty([-1,1]^N)+O(A^{-\delta})}.\end{aligned}\tag{3.2}$$

For a given measurable set $S_p\subseteq\mathbb{Z}_p^N$, recall the definition of $\sigma_p(S_p)$ in the preamble to the statement of Theorem 1.1. We observe that $\sigma_p^{B,\boldsymbol{r}}=\sigma_p(I_p)$ and $\sigma_p=\sigma_p(\mathbb{Z}_p)$. Therefore, we find from (3.2) that

$$\lim_{A\to\infty}\boldsymbol{d}\bigg(\mathfrak{B}\times\prod_{p\in\mathfrak{p}}I_p\times\prod_{p\notin\mathfrak{p}}\mathbb{Z}_p^N, A; P\bigg)=\frac{\prod_{p\in\mathfrak{p}}\sigma_p(I_p)\cdot\sigma_\infty(\mathfrak{B})}{\prod_{p\in\mathfrak{p}}\sigma_p(\mathbb{Z}_p)\cdot\sigma_\infty([-1,1]^N)}\tag{3.3}$$

Proof of Theorem 1.1. Recall that

$$\begin{aligned}T_\infty&=\left\{\boldsymbol{a}\in[-1,1]^N\cap\mathbb{R}^N\middle|\ \exists\ \boldsymbol{x}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}\text{ such that }f_{\boldsymbol{a}}(\boldsymbol{x})=0\right\}\\T_p&=\left\{\boldsymbol{a}\in\mathbb{Z}_p^N\middle|\ \exists\ \boldsymbol{x}\in\mathbb{Z}_p^n\setminus\{\boldsymbol{0}\}\text{ such that }f_{\boldsymbol{a}}(\boldsymbol{x})=0\right\}.\end{aligned}$$

On recalling the definition (3.1) of $\boldsymbol{d}(\cdot,A;P)$, we infer that

$$\varrho_{d,n}^{P,\mathrm{loc}}(A)=\boldsymbol{d}\bigg(T_\infty\times\prod_{p\text{ prime}}T_p, A; P\bigg).$$

Thus, it suffices to show that

$$\lim_{A\to\infty}\boldsymbol{d}\bigg(T_\infty\times\prod_{p\text{ prime}}T_p, A; P\bigg)=c_P.\tag{3.4}$$

In order to verify that the equality (3.4) holds, we introduce

$$\boldsymbol{d}(A,M)=\boldsymbol{d}\biggl(T_\infty\times\prod_{p<M}T_p\times\prod_{p\geq M}\mathbb{Z}_p,A;P\biggr).$$

One sees by applying the triangle inequality that

$$\lim_{A\to\infty}\biggl|\boldsymbol{d}\biggl(T_\infty\times\prod_{p\ \text{prime}}T_p,A;P\biggr)-c_P\biggr|\leq\lim_{A\to\infty}|d_1(A,M)|+\lim_{A\to\infty}|d_2(A,M)|, \tag{3.5}$$

where

$$d_1(A,M)=\boldsymbol{d}\biggl(T_\infty\times\prod_{p\ \text{prime}}T_p,A;P\biggr)-\boldsymbol{d}(A,M)$$

and

$$d_2(A,M)=\boldsymbol{d}(A,M)-c_P.$$

First, we analyze the quantity

$$\lim_{A\to\infty}|d_1(A,M)|.$$

We readily see from the definition of $\boldsymbol{d}(A,M)$ that

$$|d_1(A,M)|=\boldsymbol{d}\biggl(T_\infty\times\prod_{p<M}T_p\times\biggl(\prod_{p\geq M}T_p\biggr)^c,A;P\biggr).$$

Furthermore, we find that

$$\begin{aligned}&|d_1(A,M)|\\&\leq\boldsymbol{d}\biggl([-1,1]^N\times\prod_{p<M}\mathbb{Z}_p\times\biggl(\prod_{p\geq M}T_p\biggr)^c,A;P\biggr)\\&\leq\frac{\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N\,\middle|\,\begin{array}{l}(i)\ \exists\ p>M\ \text{s.t.}\ f_{\boldsymbol{a}}(\boldsymbol{x})=0\ \text{has no solution in}\ \mathbb{Z}_p^n\\(ii)\ P(\boldsymbol{a})=0\end{array}\right\}}{\#\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N|\ P(\boldsymbol{a})=0\right\}}.\end{aligned} \tag{3.6}$$

Meanwhile, for sufficiently large prime p, whenever $f_{\boldsymbol{a}}(\boldsymbol{x})$ is irreducible over $\overline{\mathbb{F}}_p$, the Lang-Weil estimate [22] (see also [34, Theorem 3]) ensures the existence of a smooth point $\boldsymbol{x}\in(\mathbb{F}_p)^n$ satisfying $f_{\boldsymbol{a}}(\boldsymbol{x})=0$. Then, by Hensel's lemma, we have a point $\boldsymbol{x}$ in $\mathbb{Q}_p$ satisfying $f_{\boldsymbol{a}}(\boldsymbol{x})=0$. Therefore, we conclude from (3.6) that for sufficiently large $M>0$, one has

$$\begin{aligned}&|d_1(A,M)|\\&\leq\frac{\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N\,\middle|\,\begin{array}{l}(i)\ \exists\ p>M\ \text{s.t.}\ f_{\boldsymbol{a}}(\boldsymbol{x})\ \text{is reducible over}\ \overline{\mathbb{F}}_p\\(ii)\ P(\boldsymbol{a})=0\end{array}\right\}}{\#\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N|\ P(\boldsymbol{a})=0\right\}}.\end{aligned} \tag{3.7}$$

We shall apply Lemma 2.3 with Y defined in (2.32). With this Y in mind, we find from (3.7) that

(3.8)

$$|d_1(A,M)| \leq \frac{\#\left\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N \,\middle|\, \begin{array}{l} (i)\ \boldsymbol{a} \pmod p \in Y(\mathbb{F}_p) \text{ for some prime } p > M \\ (ii)\ P(\boldsymbol{a}) = 0 \end{array}\right\}}{\#\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N |\ P(\boldsymbol{a}) = 0\}}.$$

Note by the classical argument (see [5] and [28, the proof of Theorem 1.3]) that the fact that $P(\mathbf{t}) = 0$ has a nontrivial integer solution implies that $\sigma_\infty([-1,1]^N) \cdot \prod_p \sigma_p(\mathbb{Z}_p) \asymp 1$. Then, one infers by Lemma 2.1 that

$$\#\left\{\boldsymbol{a} \in [-A,A]^N \cap \mathbb{Z}^N \middle|\ P(\boldsymbol{a}) = 0\right\} \asymp A^{N-k}.$$

Proposition 2.6 together with the hypothesis in the statement of Theorem 1.1 that $k < \lfloor (1 - C_{n,d}(1,d-1))N \rfloor$ reveals that the codimension r of Y is strictly greater than $k+1$. Hence, we find by applying Lemma 2.3 that

$$|d_1(A,M)| \ll \frac{1}{M^{r-k-1}\log M} + A^{k-r+1}.$$

Therefore, we obtain

$$\lim_{A\to\infty} |d_1(A,M)| \ll \frac{1}{M^{r-k-1}\log M}. \tag{3.9}$$

Next, we turn to estimate $\lim_{A\to\infty} |d_2(A,M)|$. For simplicity, we temporarily write

$$\lim_{A\to\infty} |d_2(A,M)| = \varphi(M). \tag{3.10}$$

One infers by (3.3) together with Lemma 3.1 that

$$d_2(A,M) = \frac{\prod_{p\leq M} \sigma_p(T_p) \cdot \sigma_\infty(T_\infty)}{\prod_{p\in\mathfrak{p}} \sigma_p(\mathbb{Z}_p) \cdot \sigma_\infty([-1,1]^N)},$$

and thus by the definition of c_P, we discern that

$$\varphi(M) \to 0, \tag{3.11}$$

as $M \to \infty$.

Hence, we conclude from (3.5), (3.9) and (3.10) that

$$\lim_{A\to\infty} \left| \boldsymbol{d}\Big(T_\infty \times \prod_{p \text{ prime}} T_p, A; P\Big) - c_P \right| \ll \frac{1}{M^{r-k-1}\log M} + \varphi(M).$$

By letting $M \to \infty$, we see from (3.11) that

$$\lim_{A\to\infty} \left| \boldsymbol{d}\Big(T_\infty \times \prod_{p \text{ prime}} T_p, A; P\Big) - c_P \right| = 0,$$

which gives (3.4). This completes the proof of Theorem 1.1. $\square$

Remark 3. By making use of Lemma 2.3, it seems possible to apply [19, Theorem 2.5] in order to obtain the conclusion of Theorem 1.1. However, in order to make this paper self-contained and for readers who are interested in an explicit way for particular thin sets we are dealing with, we record the procedure in full.

For the proof of Theorem 1.2, we require a proposition that plays an important role in guaranteeing the positiveness of c_P. In order to describe this proposition, it is convenient to define $S_\infty := \pi(V(\mathbb{R}))$. Also, for given points $\boldsymbol{b}_p \in \mathbb{Z}_p^N$ and $\boldsymbol{b}_\infty \in \mathbb{R}^N$, we define

$$B_\infty(\boldsymbol{b}_\infty, \eta) = \left\{\boldsymbol{a} \in \mathbb{R}^N \middle| \, |a_i - (\boldsymbol{b}_\infty)_i| < \eta \text{ for } 1 \le i \le N\right\}$$
$$B_p(\boldsymbol{b}_p, \eta) = \left\{\boldsymbol{a} \in \mathbb{Z}_p^N \middle| \, |a_i - (\boldsymbol{b}_p)_i|_p < \eta \text{ for } 1 \le i \le N\right\},$$

for every prime p.

Proposition 3.2. *Suppose that there exists $\boldsymbol{b}_p \in \mathbb{Z}_p^N \setminus \{\boldsymbol{0}\}$ (resp. $\boldsymbol{b}_\infty \in \mathbb{R}^N \setminus \{0\}$) such that the variety $V_{\boldsymbol{b}_p}$ in $\mathbb{P}^{n-1}_{\mathbb{Q}_p}$ (resp. $V_{\boldsymbol{b}_\infty}$ in $\mathbb{P}^{n-1}_{\mathbb{R}}$) admits a smooth $\mathbb{Q}_p$-point (resp. $\mathbb{R}$-point) for each prime p. Then, for each prime p, there exists positive numbers η_p and η_∞ less than or equal to 1 such that*

$$B_p(\boldsymbol{b}_p, \eta_p) \subseteq T_p$$
$$B_\infty(\boldsymbol{b}_\infty, \eta_\infty) \subseteq S_\infty.$$

Proof. By the existence of a smooth $\mathbb{Q}_p$-point in $V_{\boldsymbol{b}_p}$, there is an open neighborhood of V containing the smooth $\mathbb{Q}_p$-point such that the restriction of π (after base change to $\mathbb{Q}_p$) to the neighborhood is a smooth map into a neighborhood of $\mathbb{A}^N$ containing $\boldsymbol{b}_p$. By [14, Theorem 10.5.1], this map induces a topologically open map between $\mathbb{Q}_p$-points, and thus the image contains an open ball $B_p(\boldsymbol{b}_p, \eta_p)$ for some η_p. The proof for the real place is identical. □

Proof of Theorem 1.2. Recall the natural map

$$\Phi^A: \ [-A, A]^N \cap \mathbb{Z}^N \to [-1,1]^N \times \prod_{p \text{ prime}} \mathbb{Z}_p^N$$
$$\boldsymbol{a} \mapsto \left(\frac{\boldsymbol{a}}{A}, \boldsymbol{a}, \dots, \boldsymbol{a}, \dots\right).$$

Furthermore, we recall the definitions of $S_\infty := \pi(V(\mathbb{R}))$ and $T_p := \pi(V(\mathbb{Q}_p)) \cap \mathbb{Z}_p^N$ for all primes p.

Recall from the hypothesis in the statement of Theorem 1.2 that for each place v of $\mathbb{Q}$, there exists a non-zero $\boldsymbol{b}_v \in \mathbb{Q}_v^N$ such that $P(\boldsymbol{b}_v) = 0$ and the variety $V_{\boldsymbol{b}_v}$ in $\mathbb{P}^{n-1}$ admits a smooth $\mathbb{Q}_v$-point. For $v = p$ (p prime), we assume that this $\boldsymbol{b}_p$ is a $\mathbb{Z}_p$-point. Otherwise, by multiplying powers of p, we can make it a $\mathbb{Z}_p$-point not changing the variety $V_{\boldsymbol{b}_p}$. One sees by applying Proposition 3.2 that for each prime p there exist positive numbers η_p and η_∞ less than 1 such that $B_p(\boldsymbol{b}_p, \eta_p) \subseteq T_p$ and $B_\infty(\boldsymbol{b}_\infty, \eta_\infty) \subseteq S_\infty$. We choose a sufficiently large number $C := C(\boldsymbol{b}_\infty) > 0$ such that $B_\infty(\boldsymbol{b}_\infty/C, \eta_\infty/C) \subseteq$

$[-1,1]^N$. Furthermore, on observing the relation that $f_{\boldsymbol{b}_\infty/C}(\boldsymbol{x}) = (1/C)\cdot f_{\boldsymbol{b}_\infty}(\boldsymbol{x})$, we infer that $B_\infty(\boldsymbol{b}_\infty/C, \eta_\infty/C) \subseteq S_\infty$. Therefore, on noting that

$$T_\infty = S_\infty \cap [-1,1]^N.$$

one deduces that

$$B_\infty(\boldsymbol{b}_\infty/C, \eta/C) \subseteq T_\infty.$$

Meanwhile, it follows by Theorem 1.1 that

$$\lim_{A\to\infty} \varrho_{d,n}^{P,\mathrm{loc}}(A) = c_P,$$

and thus, we find that

$$\lim_{A\to\infty} \varrho_{d,n}^{P,\mathrm{loc}}(A) \geq \frac{\prod_{p<M}\sigma_p(B_p(\boldsymbol{b}_p,\eta_p))\cdot\prod_{p\geq M}\sigma_p(T_p)\cdot\sigma_\infty(B_\infty(\boldsymbol{b}_\infty/C,\eta_\infty/C))}{\prod_p \sigma_p(\mathbb{Z}_p)\cdot\sigma_\infty([-1,1]^N)}, \tag{3.12}$$

for any $M>0$. Then, it suffices to show that the right-hand side in (3.12) is greater than 0. We shall prove this by showing that there exists $M>0$ such that

$$\frac{\prod_{p<M}\sigma_p(B_p(\boldsymbol{b}_p,\eta_p))\cdot\sigma_\infty(B_\infty(\boldsymbol{b}_\infty/C,\eta_\infty/C))}{\prod_{p<M}\sigma_p(\mathbb{Z}_p)\cdot\sigma_\infty([-1,1]^N)} > 0 \tag{3.13}$$

and

$$\frac{\prod_{p\geq M}\sigma_p(T_p)}{\prod_{p\geq M}\sigma_p(\mathbb{Z}_p)} > 0. \tag{3.14}$$

First, we shall show that the inequality (3.13) holds. For a given $B\in\mathbb{N}$ and $\boldsymbol{r}\in\mathbb{Z}^N$, we recall the definition of $\sigma_p^{B,\boldsymbol{r}}$ in the statement of Lemma 2.1. Note that there exists $B\in\mathbb{N}\cup\{0\}$ such that

$$\prod_{p<M}\sigma_p(B_p(\boldsymbol{b}_p,\eta_p)) = \prod_{p<M}\sigma_p^{B,\boldsymbol{b}_p}. \tag{3.15}$$

Furthermore, the p-adic densities $\sigma_p(\mathbb{Z}_p)$, $\sigma_p^{B,\boldsymbol{b}_p}$ and the real densities $\sigma_\infty([-1,1]^N)$, $\sigma_\infty(B_\infty(\boldsymbol{b}_\infty/C,\eta_\infty/C))$ are greater than 0 by the application of the Hensel's lemma and the implicit function theorem (see [28, the proof of Theorem 1.3] or [10, Lemma 5.7]). Therefore, one sees from (3.15) that the inequality (3.13) holds for any $M>0$.

Next, we shall show that (3.14) holds. For any M_1 with $M_1>M$, we find from (3.3) with $[M,M_1)$ and $[-1,1]^N$ in place of $\mathfrak{p}$ and $\mathfrak{B}$ that

$$\lim_{A\to\infty} \boldsymbol{d}\Big([-1,1]^N \times \prod_{p\in[M,M_1)} T_p \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big) = \frac{\prod_{p\in[M,M_1)}\sigma_p(T_p)}{\prod_{p\in[M,M_1)}\sigma_p(\mathbb{Z}_p)}. \tag{3.16}$$

Meanwhile, we see that

$$\begin{aligned}&\boldsymbol{d}\Big([-1,1]^N \times \Big(\prod_{p\in[M,M_1)} T_p\Big)^c \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big)\\ &\leq \frac{\left\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N \,\middle|\, \begin{array}{l}(i)\ \exists\, p>M \text{ s.t. } f_{\boldsymbol{a}}(\boldsymbol{x})=0 \text{ has no solution in } \mathbb{Z}_p^n\\ (ii)\ P(\boldsymbol{a})=0\end{array}\right\}}{\#\{\boldsymbol{a}\in[-A,A]^N\cap\mathbb{Z}^N|\ P(\boldsymbol{a})=0\}}.\end{aligned}$$

Then, it follows by the same argument leading from (3.6) to (3.9) that

(3.17)

$$\lim_{A\to\infty}\boldsymbol{d}\Big([-1,1]^N \times \Big(\prod_{p\in[M,M_1)} T_p\Big)^c \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big) \ll \frac{1}{M^{r-k-1}\log M}.$$

Thus, on noting that

$$\begin{aligned}&\boldsymbol{d}\Big([-1,1]^N \times \prod_{p\in[M,M_1)} T_p \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big)\\ &= 1-\boldsymbol{d}\Big([-1,1]^N \times \Big(\prod_{p\in[M,M_1)} T_p\Big)^c \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big),\end{aligned}$$

we find from (3.17) that

$$\lim_{A\to\infty}\boldsymbol{d}\Big([-1,1]^N \times \prod_{p\in[M,M_1)} T_p \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big) > 1/2, \tag{3.18}$$

for sufficiently large $M>0$. Therefore, it follows from (3.16) and (3.18) that for sufficiently large $M>0$, one has

$$\begin{aligned}&\frac{\prod_{p\geq M}\sigma_p(T_p)}{\prod_{p\geq M}\sigma_p(\mathbb{Z}_p)}\\ &= \lim_{M_1\to\infty}\frac{\prod_{p\in[M,M_1)}\sigma_p(T_p)}{\prod_{p\in[M,M_1)}\sigma_p(\mathbb{Z}_p)}\\ &= \lim_{M_1\to\infty}\lim_{A\to\infty}\boldsymbol{d}\Big([-1,1]^N \times \prod_{p\in[M,M_1)} T_p \times \prod_{p\notin[M,M_1)} \mathbb{Z}_p^N, A; P\Big) > 1/2.\end{aligned}$$

Hence, the inequality (3.14) holds. By using the inequalities (3.13) and (3.14), one finds that

$$\frac{\prod_{p<M}\sigma_p(B_p(\boldsymbol{b}_p,\eta_p))\cdot\prod_{p\geq M}\sigma_p(T_p)\cdot\sigma_\infty(B_\infty(\boldsymbol{b}_\infty/C,\eta_\infty/C))}{\prod_p\sigma_p(\mathbb{Z}_p)\cdot\sigma_\infty([-1,1]^N)} > 0,$$

and thus we conclude from (3.12) that

$$\lim_{A\to\infty}\varrho_{d,n}^{P,\mathrm{loc}}(A) > 0.$$

□

References

[1] E. Artin, The collected papers of Emil Artin, (Addison–Wesley, Reading, MA, 1965).

[2] J. Ax and S. Kochen, Diophantine problems over local fields. I, Amer. J. Math., 87 (1965), 605–630.

[3] M. Bhargava, A positive proportion of plane cubics fail the Hasse principle, arXiv:1402.1131v1 (2014).

[4] M. Bhargava, The geometric sieve and the density of squarefree values of invariant polynomials, arXiv:1402.0031v1 (2014).

[5] B. J. Birch, Forms in many variables, Proc. Roy. Soc. London Ser. A 265 (1961/62), 245–263

[6] R. Brauer, A note on systems of homogeneous algebraic equations, Bull. Amer. Math. Soc., 51 (1945), 749–755.

[7] M. Bright, T. Browning, D. Loughran, Failures of weak approximation in families. Compositio Math. 152(7) (2016), 1435-1475.

[8] S.S. Brown, Bounds on transfer principles for algebraically closed and complete discretely valued fields, Mem. Amer. Math. Soc., 15 (1978), no. 204, iv+92pp.

[9] T. Browning and R. Heath-Brown. The geometric sieve for quadrics, Forum Mathematicum, vol. 33, no. 1, 2021, pp. 147-165.

[10] T. D. Browning, P. Le Boudec, W. Sawin, The Hasse principle for random fano hypersurfaces, Ann. of Math. (2) 197 (3), 1115-1203.

[11] J. Brüdern and R. Dietmann, Random Diophantine equations, I, Adv. Math. 256 (2014), 18-45.

[12] J. Brüdern and T. D. Wooley, An instance where the major and minor arc integrals meet, Bull. London Math. Soc. 51 (2019), no. 6, 1113–1128.

[13] J.-L. Colliot-Thélène, Points rationnels sur les fibrations, Higher dimensional varieties and rational points (Budapest, 2001), Bolyai Soc. Math. Stud., vol. 12, Springer, Berlin, 2003, pp. 171–221.

[14] J-L. Colliot-Thélène and A. N. Skorobogatov, The Brauer-Grothendieck group. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 71. Springer, Cham, (2021)

[15] H. Davenport and D. J. Lewis, Homogeneous additive equations, Proc. Roy. Soc. Ser. A, 274 (1963), 443–460.

[16] T. Ekedahl, An infinite version of the Chinese remainder theorem, Comment. Math. Univ. St. Paul. 40 (1991), 53–59.

[17] D. El-Baz, D. Loughran, E. Sofos, Multivariate normal distribution for integral points on varieties, Trans. Amer. Math. Soc. 375 (2022), 3089-3128.

[18] D. R. Heath-Brown, Artin's Conjecture on Zeros of p-Adic Forms, Proceedings of the International Congress of Mathematicians, Volume II (2010), 249–257.

[19] Z. Huang, Equidistribution of rational points and the geometric sieve for toric varieties, arXiv:2111.01509.

[20] F.-V. Kuhlmann. Maps on ultrametric spaces, Hensel's lemma, and differential equations over valued fields. Communications in Algebra, 39(5):1730– 1776, 2011.

[21] D. B. Leep and W. M. Schmidt, Systems of homogeneous equations, Invent. Math., 71 (1983), 539–549.

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

[23] J.Liu and L. Zhao, On forms in prime variables, arXiv:2105.12956 (2021).

[24] K. N. Ponomaryov, Semialgebraic sets and variants of the Tarski–Seidenberg–Macintyre theorem. Algebra and Logic 34 (1995), 182–191.

[25] B. Poonen, M. Stoll, The Cassels-Tate pairing on polarized abelian varieties. Ann. of Math. (2) 150 (1999), 1109-1149.

[26] B. Poonen and J. F. Voloch, Random Diophantine equations, Arithmetic of higher-dimensional algebraic varieties (Palo Alto, CA, 2002), Progr. Math., vol. 226, Birkhäuser Boston, Boston, MA, 2004, pp. 175–184.

[27] S. L. Rydin Myerson, Systems of many forms. DPhil thesis, University of Oxford (2016).

[28] S. L. Rydin Myerson, Quadratic forms and systems of forms in many variables. Invent. math. 213 (2018), 205–235.

[29] W. M. Schmidt, Simultaneous rational zeros of quadratic forms, in: Seminar on Number Theory (Paris 1980-81), Progr. Math. 22, Birkhäuser, Boston, MA, 1982, 281–307.

[30] W. M. Schmidt, The density of integer points on homogeneous varieties, Acta Math. 154 (1985), no. 3–4, 243–296

[31] G. Terjanian, Un contre-exemple à une conjecture d'Artin, C. R. Acad. Sci. Paris Sér. A-B, 262 (1966), A612.

[32] T.D. Wooley, On the local solubility of Diophantine systems, Compositio Math., 111 (1998), 149–165.

[33] K. Yeon, The Hasse principle for homogeneous polynomials with random coefficients over thin sets, arXiv:2305.08035.

[34] J. Zahid, Nonsingular points on hypersurfaces over $\mathbb{F}_q$, J. Math. Sci. (N.Y.) 171, 731–735 (2010).

(H. Lee) Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA
Email address: lee4878@purdue.edu

(S. Lee) Department of Mathematics, University of Michigan, 4839 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, USA
Email address: starlee@umich.edu

(K. Yeon) Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA
Email address: kyeon@purdue.edu