THE FURSTENBERG-SÁRKÖZY THEOREM FOR POLYNOMIALS IN ONE OR MORE PRIME VARIABLES

JOHN R. DOYLE ALEX RICE

Abstract. We establish upper bounds on the size of the largest subset of $\{1,2,\dots,N\}$ lacking nonzero differences of the form $h(p_1,\dots,p_\ell)$, where $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is a fixed polynomial satisfying appropriate conditions and $p_1,\dots,p_\ell$ are prime. The bounds are of the same type as the best-known analogs for unrestricted integer inputs, due to Bloom-Maynard and Arala for $\ell=1$, and to the authors for $\ell\geq 2$.

1. Introduction

Over the past half century, an array of results using a variety of methods have concerned the existence of certain differences within dense sets of integers. For $X\subseteq\mathbb{Z}$ and $N\in\mathbb{N}$, let

$$D(X,N)=\max\left\{|A| : A\subseteq[N],\ (A-A)\cap X\subseteq\{0\}\right\},$$

where $[N]=\{1,2,\dots,N\}$ and $A-A=\{a-b : a,b\in A\}$. In other words, $D(X,N)$ is the threshold such that a subset of $[N]$ with more than $D(X,N)$ elements necessarily contains two distinct elements that differ by an element of X, and in particular $D(X,N)=o(N)$ is equivalent to the statement that every set of natural numbers of positive upper density contains a nonzero difference in X.

Furstenberg [8] and Sárközy [31] independently established $D(S,N)=o(N)$, where S is the set of squares, answering a question of Lovász. While Furstenberg used ergodic methods to obtain a qualitative result, Sárközy used Fourier analysis to show $D(S,N)\leq N(\log N)^{-1/3+o(1)}$. In the same series of papers, Sárközy [32] established $D(\mathcal{P}-1,N)\leq N(\log\log N)^{-2+o(1)}$, where $\mathcal{P}$ is the set of primes, addressing a question of Erdős. A substantial literature has developed on refinements, generalizations, and alternative proofs of these results, see for example [9], [33], [19], [20], [28], [18], [17], [21], [26], [11], [36], [35], and [10].

More generally, for $\ell\in\mathbb{N}$ and $h\in\mathbb{Z}[x_1,\dots,x_\ell]$, the following conditions are immediately necessary for $D(h(\mathbb{Z}^\ell),N)=o(N)$ (resp. $D(h(\mathcal{P}^\ell),N)=o(N)$), to avoid counterexamples of the form $A=d\mathbb{N}$.

Definition 1.1. For $\ell\in\mathbb{N}$, a nonzero polynomial $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is *intersective* if for every $d\in\mathbb{N}$, there exists $\boldsymbol{r}=(r_1,\dots,r_\ell)\in\mathbb{Z}^\ell$ with $d\mid h(\boldsymbol{r})$. Equivalently, h is intersective if for every prime p, there exists $\boldsymbol{z}_p=((z_p)_1,\dots,(z_p)_\ell)\in\mathbb{Z}_p^\ell$ with $h(\boldsymbol{z}_p)=0$, where $\mathbb{Z}_p$ denotes the p-adic integers. Further, h is $\mathcal{P}$*-intersective* if $\boldsymbol{r}$ can always be chosen with $(r_i,d)=1$, or equivalently $\boldsymbol{z}_p$ can always be chosen with $(z_p)_i\not\equiv 0 \pmod p$, for all $1\leq i\leq\ell$.

When considering $\ell\geq 2$ variables and hoping to quantitatively improve on the univariate setting, examples like $h(x,y)=(x+y)^2$ force one to impose nonsingularity conditions, the core of which is defined below.

Definition 1.2. Suppose F is a field, $\ell\in\mathbb{N}$, and $g\in F[x_1,\dots,x_\ell]$ is a homogeneous polynomial. We say that g is *smooth* if the vanishing of g defines a smooth hypersurface in $\mathbb{P}^{\ell-1}$ (as opposed to $\mathbb{A}^\ell$). In other words, g is smooth if the system $g(\boldsymbol{x})=\frac{\partial g}{\partial x_1}(\boldsymbol{x})=\cdots=\frac{\partial g}{\partial x_\ell}(\boldsymbol{x})=0$ has no solution besides $x_1=\cdots=x_\ell=0$ in $\overline{F}^\ell$. For a general polynomial $h\in F[x_1,\dots,x_\ell]$ with $h=\sum_{i=0}^k h^i$, where h^i is homogeneous of degree i and $h^k\neq 0$, we say h is *Deligne* if the characteristic of F does not divide k and h^k is smooth. When considering polynomials with integer coefficients, we use the terms *smooth* and *Deligne* as defined above by embedding the coefficients in the field of rational numbers. For $\ell=1$, all nonconstant polynomials are Deligne.

Remark on notation. For the remainder of the paper, we take the notational convention that, for a polynomial h, h^i denotes the degree-i homogeneous part of h, as opposed to h raised to the i-th power.

For $k,\ell \geq 2$, let

$$\mu(k,\ell)=\begin{cases}[(k-1)^2+1]^{-1} & \text{if } \ell=2\\ 1/2 & \text{if } \ell\geq 3\end{cases},\quad \mu'(k,\ell)=\begin{cases}[2(k-1)^2+6]^{-1} & \text{if } \ell=2\\ 1/4 & \text{if } \ell\geq 3\end{cases}.$$

The current knowledge of upper bounds for $D(h(\mathbb{Z}^\ell),N)$ for $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is summarized in Theorem 1.3 below. The general $\ell=1$ case is due to Arala, combining efforts of Bloom and Maynard [3] and the second author [24], each of which built upon work of Pintz, Steiger, and Szemerédi [23]. The $\ell\geq 2$ case is due to the authors [7]. The term *strongly Deligne* denotes a large subclass of Deligne, intersective polynomials, the precise definition of which we delay to Section 2.1, and sufficient conditions for which we list in Theorem 1.4.

Theorem 1.3. *If $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is strongly Deligne of degree $k\geq 2$, then*

$$D(h(\mathbb{Z}^\ell),N)\ll_h \begin{cases}N(\log N)^{-c\log\log\log N} & \text{if } \ell=1\\ N\exp(-c(\log N)^{\mu(k,\ell)}) & \text{if } \ell\geq 2\end{cases},$$

where $c=c(h)>0$.

The fact that every intersective polynomial is strongly Deligne when $\ell=1$ follows from [19, Lemma 28], as discussed in Remark 2.3, while the remaining criteria below are established in [7].

Theorem 1.4. *An intersective, Deligne polynomial $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ of degree $k\geq 2$ is strongly Deligne if it meets any of the following conditions:*

(i) $\ell\neq 2$,
(ii) *There exist $a,b\in\mathbb{Z}$ such that $h(a,b)=0$ and the highest and lowest degree parts of $h(x+a,y+b)$ are smooth,*
(iii) *For an irreducible factorization $h=g_1\cdots g_n$ in $\overline{\mathbb{Z}}[x_1,\dots,x_\ell]$ and all but finitely many $p\in\mathcal{P}$, g_i has coefficients in $\mathbb{Z}_p$ for some $1\leq i\leq n$,*
(iv) *For all but finitely many p, there exists a p-adic integer root of h of multiplicity 1 or k. In particular, this condition holds when $k=2$.*

Our main results are as follows. Analogous to strongly Deligne, the term *$\mathcal{P}$-Deligne* refers to a large subclass of Deligne, $\mathcal{P}$-intersective polynomials, the precise definition of which is provided in Section 2.1.

Theorem 1.5. *If $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is $\mathcal{P}$-Deligne of degree $k\geq 2$, then*

$$D(h(\mathcal{P}^\ell),N)\ll_h \begin{cases}N(\log N)^{-c\log\log\log N} & \text{if } \ell=1\\ N\exp(-c(\log N)^{\mu'(k,\ell)}) & \text{if } \ell\geq 2\end{cases},$$

where $c=c(h)>0$.

As before, when $\ell=1$, all $\mathcal{P}$-intersective polynomials are $\mathcal{P}$-Deligne. The $\ell=1$ case of Theorem 1.5 improves upon the previous bound $D(h(\mathcal{P}),N)\leq N(\log N)^{-\frac{1}{2k-2}+o(1)}$ for $\mathcal{P}$-intersective $h\in\mathbb{Z}[x]$ of degree $k\geq 2$, due to the second author [26].

Remark 1.6. Under GRH, the $\ell\geq 2$ bounds in Theorem 1.5 holds with exponent $2\mu'(k,\ell)$. Further, unconditional results with exponent between $\mu(k,\ell)$ and $\mu'(k,\ell)$ may be possible using techniques from [36].

Theorem 1.7. *A $\mathcal{P}$-intersective, Deligne polynomial $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ of degree $k\geq 2$ is $\mathcal{P}$-Deligne if it meets any of the following conditions:*

(i) $\ell\neq 2$,
(ii) *There exist $a,b\in\{-1,1\}$ such that $h(a,b)=0$ and the highest and lowest degree parts of $h(x+a,y+b)$ are smooth,*
(iii) *Let $h=g_1\cdots g_n$ be an irreducible factorization in $\overline{\mathbb{Z}}[x_1,\dots,x_\ell]$. For all but finitely many $p\in\mathcal{P}$, there exists $1\leq i\leq n$ such that g_i has coefficients in $\mathbb{Z}_p$ and $x_j\nmid g_i$ for all $1\leq j\leq\ell$,*
(iv) *For all but finitely many p, there exists a p-adic integer root $\boldsymbol{z}_p$ of h, satisfying $(z_p)_i\not\equiv 0 \pmod p$ for $1\leq i\leq\ell$, of multiplicity 1 or k. In particular, this condition holds when $k=2$.*

Remark 1.8. The conclusions of Theorems 1.3 and 1.5 hold under a more general condition. It suffices that h can be written as $h(\boldsymbol{x}_1,\dots,\boldsymbol{x}_s)=h_1(\boldsymbol{x}_1)+\cdots+h_s(\boldsymbol{x}_s)$, where $h_i\in\mathbb{Z}[x_1,\dots,x_{\ell_i}]$ is strongly Deligne (resp. $\mathcal{P}$-Deligne) for $1\le i\le s$, and $\ell_1+\cdots+\ell_s=\ell$. In this case, the relevant exponential sums factor as products, reducing the problem to the treatment of strongly Deligne (resp. $\mathcal{P}$-Deligne) polynomials in fewer variables. In particular, this includes the diagonal case where $h(x_1,\dots,x_\ell)=h_1(x_1)+\cdots+h_\ell(x_\ell)$ with $h_1,\dots,h_\ell\in\mathbb{Z}[x]$ intersective (resp. $\mathcal{P}$-intersective), as treated in [24]. We stick to the conditions as stated in Theorem 1.5 for ease of exposition.

As discussed in Section 1.3 in [7], known lower bounds for $D(h(\mathbb{Z}^\ell),N)$ (which also bound $D(h(\mathcal{P}^\ell),N)$) take the form N^{1-c}, so are far removed from known upper bounds, and all are derived from a construction of Ruzsa [29] (see also [16], [37]). Another construction of Ruzsa [30] shows $D(\mathcal{P}-1,N)\gg N^{c/\log\log N}$, as compared with Green's [10] breakthrough upper bound $D(\mathcal{P}-1,N)\ll N^{1-c}$.

2. Preliminaries

2.1. **Auxiliary polynomials and $\mathcal{P}$-Deligne definition.**

We employ a standard density increment procedure, which takes as input a set lacking nonzero differences in the image of a polynomial, and produces a new, denser subset of a slightly smaller interval lacking nonzero differences in the image of a potentially modified polynomial. The following definition keeps track of the changes in the polynomial over the course of the iteration.

Definition 2.1. Suppose $h\in\mathbb{Z}[x_1,\dots,x_\ell]$ is an intersective polynomial and fix, for each prime p, $\boldsymbol{z}_p\in\mathbb{Z}_p^\ell$ with $h(\boldsymbol{z}_p)=0$. All objects defined below depend on this choice of p-adic integer roots, but we suppress that dependence in the subsequent notation.

By reducing modulo prime powers and applying the Chinese Remainder Theorem, the choice of roots determines, for each $d\in\mathbb{N}$, a unique $\boldsymbol{r}_d\in(-d,0]^\ell$ with $\boldsymbol{r}_d\equiv\boldsymbol{z}_p \bmod p^j$ for all prime powers $p^j\mid d$.

We define a completely multiplicative function λ (depending on h and $\{\boldsymbol{z}_p\}$) on $\mathbb{N}$ by letting $\lambda(p)=p^{m_p}$ for each prime p, where m_p is the multiplicity of $\boldsymbol{z}_p$ as a root of h, in other words

$$m_p=\min\left\{i_1+\cdots+i_\ell:\frac{\partial^{i_1+\cdots+i_\ell}h}{\partial x_1^{i_1}\cdots\partial x_\ell^{i_\ell}}(\boldsymbol{z}_p)\neq 0\right\}.$$

For each $d\in\mathbb{N}$, we define the *auxiliary polynomial*, $h_d\in\mathbb{Z}[x_1,\dots,x_\ell]$, by

$$h_d(\boldsymbol{x})=h(\boldsymbol{r}_d+d\boldsymbol{x})/\lambda(d).$$

Definition 2.2. We say that h is *strongly Deligne* if there exists a finite set of primes $X=X(h)$ and a choice $\{\boldsymbol{z}_p\}_{p\in\mathcal{P}}$ of p-adic integer roots of h such that the reduction of h_d modulo p is Deligne for all $p\notin X$ and all $d\in\mathbb{N}$. Analogously, we say h is *$\mathcal{P}$-Deligne* if such a choice $\{\boldsymbol{z}_p\}_{p\in\mathcal{P}}$ exists with $(z_p)_i\not\equiv 0 \pmod p$ for all p and all $1\le i\le\ell$. We note that all strongly Deligne polynomials are both Deligne and intersective, and all $\mathcal{P}$-Deligne polynomials are both Deligne and $\mathcal{P}$-intersective.

Remark 2.3. For $\ell=1$, all nonconstant polynomials are Deligne, so it follows from [19, Lemma 28] that all intersective polynomials are strongly Deligne and all $\mathcal{P}$-intersective polynomials are $\mathcal{P}$-Deligne, where X is the set of primes that could simultaneously divide all nonconstant coefficients of an auxiliary polynomial.

2.2. **Counting Primes in Arithmetic Progressions.**

For $x>0$ and $a,q\in\mathbb{N}$, we define

$$\psi(x,a,q)=\sum_{\substack{p\le x \text{ prime}\\ p\equiv a (\text{mod } q)}}\log p.$$

Classical estimates on $\psi(x,a,q)$ come from the Siegel-Walfisz Theorem, which can be found for example in Corollary 11.19 of [22]. Ruzsa and Sanders [28] established asymptotics for $\psi(x,a,q)$ for certain moduli q beyond the usual limitation $q\le(\log x)^C$ by exploiting a dichotomy based on exceptional zeros, or lack thereof, of Dirichlet L-functions. In particular, the following result follows from their work.

Lemma 2.4. *For any $Q, D > 0$, there exist $q_0 \leq Q^D$ and $\rho \in [1/2, 1)$ with $(1-\rho)^{-1} \ll q_0$ such that*

$$\psi(x,a,q) = \frac{x}{\varphi(q)} - \frac{\chi(a)x^\rho}{\varphi(q)\rho} + O\left(x \exp\left(-\frac{c\log x}{\sqrt{\log x} + D^2 \log Q}\right) D^2 \log Q\right), \tag{1}$$

where χ is a Dirichlet character modulo q_0, provided $q_0 \mid q$, $(a,q) = 1$, and $q \leq (q_0 Q)^D$.

Lemma 2.4 is a purpose-built special case of [28, Proposition 4.7], which in the language of that paper can be deduced by considering the pair (Q^{D^2+D}, Q^D), where q_0 is the modulus of the exceptional Dirichlet character if the pair is exceptional and $q_0 = 1$ if the pair is unexceptional.

It is a calculus exercise to verify that if $\epsilon \in [0, 1/2]$ and $x \geq 16$, then $1 - x^{-\epsilon}/(1-\epsilon) \geq \epsilon$, which implies that the main term in Lemma 2.4 satisfies

$$\Re\left(\frac{x}{\varphi(q)} - \frac{\chi(a)x^\rho}{\varphi(q)}\right) \geq (1-\rho)\frac{x}{\varphi(q)} \gg \frac{x}{q_0\varphi(q)}. \tag{2}$$

2.3. **Fourier analysis and the circle method on $\mathbb{Z}$.** We embed our finite sets in $\mathbb{Z}$, on which we utilize an unnormalized discrete Fourier transform. Specifically, for a function $F : \mathbb{Z} \to \mathbb{C}$ with finite support, we define $\widehat{F} : \mathbb{T} \to \mathbb{C}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, by

$$\widehat{F}(\alpha) = \sum_{x \in \mathbb{Z}} F(x) e(-x\alpha),$$

where $e(t) = e^{2\pi i t}$. Given $N \in \mathbb{N}$ and a set $A \subseteq [N]$ with $|A| = \delta N$, we examine the Fourier analytic behavior of A by considering the *balanced function*, f_A, defined by $f_A = 1_A - \delta 1_{[N]}$.

As is standard in the circle method, we decompose the frequency space into two pieces: the points of $\mathbb{T}$ that are close to rational numbers with small denominator, and the complement.

Definition 2.5. Given $\gamma > 0$ and $Q \geq 1$, we define, for each $a, q \in \mathbb{N}$ with $0 \leq a \leq q-1$,

$$\mathbf{M}_{a/q}(\gamma) = \left\{\alpha \in \mathbb{T} : \left|\alpha - \frac{a}{q}\right| < \gamma\right\}, \ \mathbf{M}_q(\gamma) = \bigcup_{(a,q)=1} \mathbf{M}_{a/q}(\gamma), \text{ and } \mathbf{M}'_q(\gamma) = \bigcup_{r|q} \mathbf{M}_r(\gamma) = \bigcup_{a=0}^{q-1} \mathbf{M}_{a/q}(\gamma).$$

We then define the *major arcs* by

$$\mathfrak{M}(\gamma, Q) = \bigcup_{q=1}^{Q} \mathbf{M}_q(\gamma),$$

and the *minor arcs* by $\mathfrak{m}(\gamma, Q) = \mathbb{T} \setminus \mathfrak{M}(\gamma, Q)$. We note that if $2\gamma Q^2 < 1$, then

$$\mathbf{M}_{a/q}(\gamma) \cap \mathbf{M}_{b/r}(\gamma) = \emptyset \tag{3}$$

whenever $a/q \neq b/r$ and $q, r \leq Q$.

2.4. **Inheritance proposition.** As previously noted, we defined auxiliary polynomials to keep track of an inherited lack of prescribed differences at each step of a density increment iteration. For our restriction to prime inputs, we define

$$\Lambda_d = \{\boldsymbol{n} \in \mathbb{Z}^\ell : \boldsymbol{r}_d + d\boldsymbol{n} \in \mathcal{P}^\ell\}.$$

The following proposition makes this inheritance precise.

Proposition 2.6. *If $h \in \mathbb{Z}[x_1, \ldots, x_\ell]$ is $\mathcal{P}$-intersective, $d, q \in \mathbb{N}$, $A \subseteq \mathbb{N}$, and $A' \subseteq \{a : x + \lambda(q)a \in A\}$ for some $x \in \mathbb{Z}$, then*

$$\{\lambda(q)x : x \in (A' - A') \cap h_{qd}(\Lambda_{qd})\} \subseteq (A - A) \cap h_d(\Lambda_d).$$

In particular, if $(A-A) \cap h_d(\Lambda_d) \subseteq \{0\}$, then $(A'-A') \cap h_{qd}(\Lambda_{qd}) \subseteq \{0\}$.

Proof. Suppose that $A \subseteq \mathbb{N}$, $A' \subseteq \{a : x + \lambda(q)a \in A\}$, and $t \in (A' - A') \cap h_{qd}(\Lambda_{qd})$. In other words,

$$t = a - a' = h_{qd}(\boldsymbol{n}) = h(\boldsymbol{r}_{qd} + qd\boldsymbol{n})/\lambda(qd)$$

for some $\boldsymbol{n} \in \mathbb{Z}^\ell$, $a, a' \in A'$, with $\boldsymbol{r}_{qd} + qd\boldsymbol{n} \in \mathcal{P}^\ell$. Clearly $\lambda(q)t = (x + \lambda(q)a) - (x + \lambda(q)a') \in A - A$, but we must also show $\lambda(q)t \in h_d(\Lambda_d)$. By construction, we have that $\boldsymbol{r}_{qd} \equiv \boldsymbol{r}_d \bmod d$, so there exists $\boldsymbol{s} \in \mathbb{Z}^\ell$ such that $\boldsymbol{r}_{qd} = \boldsymbol{r}_d + d\boldsymbol{s}$. Further, λ is completely multiplicative, and therefore

$$\lambda(q)t = \lambda(q)h_{qd}(\boldsymbol{n}) = h(\boldsymbol{r}_d + d(\boldsymbol{s} + q\boldsymbol{n}))/\lambda(d) = h_d(\boldsymbol{s} + q\boldsymbol{n}),$$

and $\boldsymbol{r}_d + d(\boldsymbol{s} + q\boldsymbol{n}) = \boldsymbol{r}_{qd} + qd\boldsymbol{n} \in \mathcal{P}^\ell$, so $\lambda(q)t \in h_d(\Lambda_d)$, completing the proof. □

2.5. **Input restriction.** In addition to restricting our inputs to (affine preimages of) the primes, we also sieve away roots of the gradient vector, as is done in [24] and [7], in order to apply Hensel's lemma when analyzing local exponential sums. For a $\mathcal{P}$-intersective polynomial $h \in \mathbb{Z}[x_1, \dots, x_\ell]$, and each prime p and $d \in \mathbb{N}$, we define $\varepsilon_d(p)$ to be 0 if $p \mid d$ and 1 otherwise. We define $\gamma_d(p)$ to be the smallest power such that ∇h_d does not vanish modulo $p^{\gamma_d(p)}$ on

$$J_d(p) = \{\boldsymbol{c} \in (\mathbb{Z}/p^{\gamma_d(p)}\mathbb{Z})^\ell : (r_d)_i + dc_i \not\equiv 0 \pmod{p} \text{ for all } 1 \le i \le \ell\}.$$

We let $j_d(p)$ denote the number of solutions to $\nabla h_d \equiv \boldsymbol{0} \pmod{p^{\gamma_d(p)}}$ on $J_d(p)$, and we let

$$w_d = \prod_{p \le Y} \left(1 - \frac{j_d(p)}{|J_d(p)|}\right) = \prod_{p \le Y} \left(1 - \frac{j_d(p)}{\big((p - \varepsilon_d(p))p^{\gamma_d(p)-1}\big)^\ell}\right) > 0.$$

Then, for $d \in \mathbb{N}$ and $Y > 0$ we define

$$W_d(Y) = \Big\{\boldsymbol{n} \in \Lambda_d : \nabla h_d(\boldsymbol{n}) \not\equiv \boldsymbol{0} \bmod p^{\gamma_d(p)} \text{ for all } p \le Y\Big\}.$$

To appropriately weight our doubly-restricted inputs, we define

$$\nu_d(\boldsymbol{n}) = w_d^{-1} \cdot \left(\frac{\varphi(d)}{d}\right)^\ell \cdot \prod_{i=1}^{\ell} \log((r_d)_i + dn_i) \cdot 1_{W_d(Y)}(\boldsymbol{n}) \cdot \begin{cases} h_d'(n) & \ell = 1 \\ 1 & \ell \ge 2 \end{cases}.$$

Further, for $x_1, \dots, x_\ell > 0$ and a collection of primes $p_1 < \dots < p_s$, we define

$$\mathcal{A}_{p_1 \cdots p_s} = \mathcal{A}_{p_1 \cdots p_s}(x_1, \dots, x_\ell) = \Big\{\boldsymbol{n} \in B : \boldsymbol{r}_d + d\boldsymbol{n} \in \mathcal{P}^\ell,\ \nabla h_d(\boldsymbol{n}) \equiv \boldsymbol{0} \bmod p_i^{\gamma(p_i)} \text{ for all } 1 \le i \le s\Big\},$$

where $B = [x_1] \times \dots \times [x_\ell]$. Finally, we let $b = b(h) = \max_{p \in \mathcal{P}} \gamma_d(p) \ll_h 1$ (see Section 5.2).

For the three lemmas that follow, we let Q, D, q_0, ρ, and χ be as in Lemma 2.4.

Lemma 2.7. *If $\ell \ge 2$, $x_1, \dots, x_\ell, Y > 0$, $q_0 \mid d$, $dY^{bt} \le (q_0 Q)^D$, and $t > 2C \log\log Y$ for a constant $C = C(h)$, then*

$$\sum_{\boldsymbol{n} \in B} \nu_d(\boldsymbol{n}) = \prod_{i=1}^{\ell} \left(x_i - \frac{\chi((r_d)_i) x_i^\rho}{\rho d^{1-\rho}}\right) + E_1 + E_2,$$

where $X = \max\{x_1, \dots, x_\ell\}$, $E_1 = O_h\left((dX)^\ell \left(\frac{C \log\log Y}{t}\right)^t\right)$, and

$$E_2 = O_h\left((dX)^\ell \log^\ell(dX) Y^{4bt} \exp\left(-\frac{c \log dx}{\sqrt{\log dx} + D^2 \log Q}\right) D^{2\ell} \log^\ell Q\right),$$

for a constant $c > 0$.

Proof. Suppose $\ell \ge 2$, and fix $x_1, \dots, x_\ell, Y > 0$ and $d \in \mathbb{N}$ with $q_0 \mid d$. Letting z denote the number of primes that are at most Y, we have by the Chinese Remainder Theorem and the inclusion-exclusion principle that

$$\sum_{\boldsymbol{n} \in B} \nu_d(\boldsymbol{n}) = w_d^{-1} \left(\frac{\varphi(d)}{d}\right)^\ell \sum_{s=0}^{z} (-1)^s \sum_{p_1 < \dots < p_s \le Y} \sum_{\mathcal{A}_{p_1 \cdots p_s}} \prod_{i=1}^{\ell} \log((r_d)_i + dn_i),$$

and moreover the true sum lies between any two consecutive truncated alternating sums in s.

Note that for an ℓ-tuple of congruence classes $\boldsymbol{c}$ modulo $m = p_1^{\gamma_d(p_1)} \cdots p_s^{\gamma_d(p_s)}$ to contribute more than a single element to $\mathcal{A}_{p_1 \cdots p_s}(x)$, we must have $p_i \nmid (r_d)_j + dc_j$ for all $1 \leq i \leq s$ and all $1 \leq j \leq \ell$. We enumerate these congruence classes $\boldsymbol{c}_i$ for $1 \leq i \leq j_d(p_1) \cdots j_d(p_s)$, and letting $b = \max_{p \in \mathcal{P}} \gamma_d(p)$, we know by Lemma 2.8 that if $dY^{bt} \leq (q_0 Q)^D$, then

$$\begin{aligned}
&w_d^{-1} \left(\frac{\varphi(d)}{d}\right)^{\ell} \sum_{s=0}^{t} (-1)^s \sum_{p_1 < \cdots < p_s \leq Y} \left(\sum_{i=1}^{j_d(p_1) \cdots j_d(p_s)} \sum_{\substack{\boldsymbol{n} \in B \cap \Lambda_d \\ \boldsymbol{n} \equiv \boldsymbol{c}_i (\mathrm{mod}\ m)}} \prod_{i=1}^{\ell} \log((r_d)_i + dn_i) + O_h(Y^{bt} \log^{\ell}(dX)) \right) \\
=&w_d^{-1} \left(\frac{\varphi(d)}{d}\right)^{\ell} \sum_{s=0}^{t} (-1)^s \sum_{p_1 < \cdots < p_s \leq Y} \left(\sum_{i=1}^{j_d(p_1) \cdots j_d(p_s)} \prod_{i=1}^{\ell} \psi((r_d)_i + dx_i, (r_d)_i + dc_i, dm) + O_h(Y^{bt} \log^{\ell}(dX)) \right) \\
=&w_d^{-1} \left(\frac{\varphi(d)}{d}\right)^{\ell} \sum_{s=0}^{t} (-1)^s \sum_{p_1 < \cdots < p_s \leq Y} \left(\frac{j_d(p_1) \cdots j_d(p_s)}{(\varphi(dm))^{\ell}} \prod_{i=1}^{\ell} (dx_i - \chi((r_d)_i)(dx_i)^{\rho}/\rho) + C^s E_0 \right) \\
=&w_d^{-1} \sum_{s=0}^{t} (-1)^s \sum_{p_1 < \cdots < p_s \leq Y} \left(\prod_{i=1}^{\ell} \frac{j_d(p_i)}{\left((p - \epsilon_d(p_i))p^{\gamma_d(p_i)-1}\right)^{\ell}} \left(x_i - \frac{\chi((r_d)_i) x_i^{\rho}}{\rho d^{1-\rho}} \right) + C^s E_0 \right),
\end{aligned}$$

where the C^s term comes from the fact that $j_d(p) \ll_h 1$ (see Section 5.2), and

$$E_0 = O_h \left(Y^{bt} \log^{\ell}(dX) + (dX)^{\ell} \exp\left(-\frac{c \log dx}{\sqrt{\log dx} + D^2 \log Q} \right) D^{2\ell} \log^{\ell} Q \right),$$

which can be merged to

$$E_0 = O_h \left((dX)^{\ell} \log^{\ell}(dX) Y^{bt} \exp\left(-\frac{c \log dx}{\sqrt{\log dx} + D^2 \log Q} \right) D^{2\ell} \log^{\ell} Q \right). \tag{4}$$

Adding and subtracting the main term for $s > t$, and noting that

$$\sum_{s=0}^{z} (-1)^s \sum_{p_1 < \cdots < p_s \leq Y} \prod_{i=1}^{s} \frac{j_d(p_i)}{\left((p - \epsilon_d(p_i))p^{\gamma_d(p_i)-1}\right)^{\ell}} = w_d,$$

we have

$$\sum_{\boldsymbol{n} \in B} \nu_d(\boldsymbol{n}) = \prod_{i=1}^{\ell} \left(x_i - \frac{\chi((r_d)_i) x_i^{\rho}}{\rho d^{1-\rho}} \right) + E_1 + E_2, \tag{5}$$

where $E_1 = O\left((dX)^{\ell} \sum_{s=t+1}^{z} \sum_{p_1 < \cdots < p_s \leq Y} \prod_{i=1}^{s} \frac{j_d(p_i)}{\left((p-\epsilon_d(p_i))p^{\gamma_d(p_i)-1}\right)^{\ell}} \right)$ and $E_2 = \sum_{s=0}^{t} \binom{z}{s} C^s E_0$, and we note that $w^{-1} = O_h(1)$ since $\ell \geq 2$. First, we see

$$\begin{aligned}
E_1 &= O_h \left((dX)^{\ell} \sum_{s=t+1}^{z} \sum_{p_1 < \cdots < p_s \leq Y} \prod_{i=1}^{s} \frac{C}{p} \right) \\
&= O_h \left((dX)^{\ell} \sum_{s=t+1}^{z} \frac{1}{s!} \Big(\sum_{p \leq Y} \frac{C}{p} \Big)^s \right) \\
&= O_h \left((dX)^{\ell} \sum_{s=t+1}^{z} \frac{(C \log\log Y)^s}{s!} \right).
\end{aligned}$$

If $t > 2C \log\log Y$, then the sum is dominated by twice its first term, hence

$$E_1 = O_h \left((dX)^{\ell} \frac{(C \log\log Y)^t}{t!} \right) = O_h \left((dX)^{\ell} \Big(\frac{C \log\log Y}{t} \Big)^t \right), \tag{6}$$

where the last step uses that $t! \geq (t/e)^t$.

Further,

$$|E_2| = \left|\sum_{s=0}^{t} \binom{z}{s} C^s E_0\right| \leq |E_0| \sum_{s=0}^{t} \frac{(Cz)^s}{s!}.$$

This time the sum is dominated by twice its *last* term, so

$$E_2 = O(E_0(Cz)^t/t!) = O(E_0(Cz/t)^t). \tag{7}$$

Since $z \leq Y$, the result follows from (5), (4), (6), and (7). □

In particular, by the same reasoning as (2), the main term in Lemma 2.7 satisfies

$$\Re\left(\prod_{i=1}^{\ell}\left(x_i - \frac{\chi((r_d)_i)x_i^{\rho}}{\rho d^{1-\rho}}\right)\right) \gg x_1 \cdots x_\ell/q_0^\ell. \tag{8}$$

To account for the derivative weight when $\ell = 1$, we need the following estimate, which follows from Lemma 2.4 and partial summation.

Lemma 2.8. *If $x > 0$, $q_0 \mid q$, $(a, q) = 1$, and $q \leq (q_0Q)^D$, then*

$$\Psi_h(x, a, q) = \sum_{\substack{p \leq x \text{ prime} \\ p \equiv a (\text{mod } q)}} h'(p)\log p = \varphi(q)^{-1}\int_1^x h'(t)\left(1 - \chi(a)t^{\rho-1}\right) dt$$
$$+ O_h\left(x^k \exp\left(-\frac{c\log x}{\sqrt{\log x} + D^2\log Q}\right) D^2 \log Q\right).$$

Incorporating Lemma 2.8, we have the following analog of Lemma 2.7 when $\ell = 1$.

Lemma 2.9. *If $\ell = 1$, $x, Y > 0$, $q_0 \mid d$, $dY^{bt} \leq (q_0Q)^D$, and $t > 2C\log\log Y$ for a constant $C = C(h)$, then*

$$\sum_{n \leq x} \nu_d(n) = \int_1^x h_d'(t)\left(1 - \chi(r_d)(dt)^{\rho-1}\right) dt + E_1 + E_2,$$

where

$$E_1 = O_h\left(h_d(x)(\log Y)^k \left(\frac{C\log\log Y}{t}\right)^t\right),$$

and

$$E_2 = O_h\left((dx)^k \log(dx) Y^{4bt} \exp\left(-\frac{c\log dx}{\sqrt{\log dx} + D^2\log Q}\right) D^2\log Q\right),$$

for a constant $c > 0$.

3. Proof of Theorem 1.5

3.1. **Main iteration lemma and proof of Theorem 1.5.** For the remainder of the section, we fix a $\mathcal{P}$-Deligne polynomimal $h \in \mathbb{Z}[x_1, \ldots, x_\ell]$ of degree $k \geq 2$, and positive constants $C_0 = C_0(h)$ and $c_0 = c_0(h)$ that are appropriately large and small, respectively. We also fix $N \in \mathbb{N}$, and we let

$$\mathcal{Q} = \mathcal{Q}(N, h) = \exp(c_0\sqrt{\log N}), \quad \mathcal{Q}' = \mathcal{Q}'(N, h) = \exp(c_0(\log N)^{1/4}).$$

We apply Lemma 2.4 with $Q = \mathcal{Q}$ and $D = C_0$, letting $q_0 \leq \mathcal{Q}^{C_0}$, $\rho \in [1/2, 1)$ and the Dirichlet character χ be as in the conclusion of the that lemma. For a density $\delta \in (0, 1]$, we define

$$\theta(h, \delta) = \begin{cases} \exp\left(-C_0\frac{\log(2\delta^{-1})}{\log\log(3\delta^{-1})}\right) & \ell = 1 \\ \log^{-[(k-1)^2+1]}(2\delta^{-1}) & \ell = 2 \\ 1 & \ell \geq 3 \end{cases}.$$

We deduce Theorem 1.5 from the following iteration lemma, encapsulating the density increment strategy.

Lemma 3.1. *Suppose $A \subseteq [L]$ with $|A| = \delta L$ and $L \geq \sqrt{N}$. If $(A-A) \cap h_d(\Lambda_d) \subseteq \{0\}$, $C_0, \delta^{-1} \leq \mathcal{Q}'$, $q_0 \mid d$, and $d/q_0 \leq \mathcal{Q}$, then there exists $q \ll_h \delta^{-2}$ and $A' \subseteq [L']$ such that $L' \gg_h \delta^{4k} L$,*

$$|A'| \geq (1 + c\theta(h,\delta))\delta L',$$

where $c = c(h) > 0$, and

$$(A'-A') \cap h_{qd}(\Lambda_{qd}) \subseteq \{0\}.$$

Proof of Theorem 1.5. Throughout this proof, we let C and c denote sufficiently large or small positive constants, respectively, which we allow to change from line to line, but which depend only on h.

Suppose $A \subseteq [N]$ with $|A| = \delta N$ and

$$(A-A) \cap h(\mathcal{P}^\ell) \subseteq \{0\}.$$

Partitioning $[N]$ into arithmetic progressions of step size $\lambda(q_0)$ and length between $N/2\lambda(q_0)$ and $N/\lambda(q_0)$, the pigeonhole principle guarantees the existence of an arithmetic progression $P = \{x + a\lambda(q_0) : 1 \leq \ell \leq N_0\}$ such that $N/2\lambda(q_0) \leq N_0 \leq N/\lambda(q_0)$ and $|A \cap P|/N_0 = \delta_0 \geq \delta$. This allows us to define $A_0 \subseteq [N_0]$ by $A_0 = \{a \in [N_0] : x + a\lambda(q_0) \in A\}$. Setting $d_0 = q_0$, Lemma 3.1 yields, for each m, a set $A_m \subseteq [N_m]$ with $|A_m| = \delta_m N_m$ and $(A_m - A_m) \cap h_{d_m}(\Lambda_{d_m}) \subseteq \{0\}$. Further, we have that

$$N_m \geq c\delta^{4k} N_{m-1} \geq (c\delta)^{4km} N, \tag{9}$$

$$\delta_m \geq (1 + c\theta(h, \delta_{m-1}))\delta_{m-1}, \tag{10}$$

and

$$d_m \leq (c\delta)^{-2} d_{m-1} \leq (c\delta)^{-2m}, \tag{11}$$

as long as

$$C, \delta_m^{-1} \leq \mathcal{Q}', \qquad d_m/q_0 \leq \mathcal{Q}, \qquad \text{and} \qquad N_m \geq \sqrt{N}. \tag{12}$$

By (10), the density δ_m will exceed 1, and hence (12) must fail, for $m = M = M(h,\delta)$, where

$$M(h,\delta) = \begin{cases} C\log(2\delta^{-1}) & \text{if } \ell \geq 3 \\ C\log^{(k-1)^2+2}(2\delta^{-1}) & \text{if } \ell = 2 \\ \exp\left(C\frac{\log(2\delta^{-1})}{\log\log(3\delta^{-1})}\right) & \text{if } \ell = 1 \end{cases}.$$

However, for $\ell \geq 2$, by (9), (10), and (11), (12) holds for $m = M$ if

$$(c\delta)^{4kM} = \exp\left(C\log^{[2\mu'(k,\ell)]^{-1}}(2\delta^{-1})\right) \leq \mathcal{Q} = \exp(c\sqrt{\log N}). \tag{13}$$

Therefore, (13) must fail, or in other words $\delta \ll_h \exp(-c(\log N)^{\mu'(k,\ell)})$, as claimed.

Similarly, for $\ell = 1$, (12) holds for $m = M$ if

$$(c\delta)^{4kM} = \exp\left(\exp\left(C\frac{\log(2\delta^{-1})}{\log\log(3\delta^{-1})}\right)\right) \leq \mathcal{Q} = \exp(c\sqrt{\log N}). \tag{14}$$

Therefore, (14) must fail, which yields

$$\frac{\log(2\delta^{-1})}{\log\log(3\delta^{-1})} \geq c\log\log N,$$

and finally

$$\delta \ll_h (\log N)^{-c\log\log\log N},$$

completing the proof. □

3.2. L^2 **concentration and density increment lemmas.** As usual, we prove Lemma 3.1 by locating one small denominator q such that $\widehat{f_A}$ has L^2 concentration around rationals with denominator q, then invoke a standard lemma stating that L^2 concentration of $\widehat{f_A}$ implies the existence a long arithmetic progression on which A has increased density.

Lemma 3.2. *Suppose $A \subseteq [L]$ with $|A| = \delta L$, $L \geq \sqrt{N}$, $\eta = c_0\delta$, and $\gamma = \eta^{-2k}/L$. Further suppose $(A-A)\cap h_d(\Lambda_d) \subseteq \{0\}$, $C_0, \delta^{-1} \leq \mathcal{Q}'$, $q_0 \mid d$, and $d/q_0 \leq \mathcal{Q}$. If $|A\cap (L/9, 8L/9)| \geq 3\delta L/4$, then there exists $q \leq \eta^{-2}$ such that*

$$\int_{\mathbf{M}'_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \gg_h \theta(h,\delta)\delta^2 L.$$

Lemma 3.1 follows from Lemma 3.2 and the following standard L^2 density increment lemma.

Lemma 3.3 (Lemma 2.3 in [25], see also [19], [28]). *Suppose $A \subseteq [L]$ with $|A| = \delta L$. If $0 < \theta \leq 1$, $q \in \mathbb{N}$, $\gamma > 0$, and*

$$\int_{\mathbf{M}'_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \geq \theta\delta^2 L,$$

then there exists an arithmetic progression $P = \{x + aq : 1 \leq a \leq L'\}$ with $qL' \gg \min\{\theta L, \gamma^{-1}\}$ and $|A \cap P| \geq (1+\theta/32)\delta L'$.

The deduction of Lemma 3.1 from Lemmas 3.2 and 3.3 is standard, and in particular is identical to the analogous deduction in [7, Section 4].

3.3. **Proof of Lemma 3.2.** Before delving into the proof of Lemma 3.2, we take the opportunity to define some relevant sets and quantities, depending on our $\mathcal{P}$-Deligne polynomial $h \in \mathbb{Z}[x_1,\dots,x_\ell]$, scaling parameter d, a parameter $Y > 0$, and the size L of the ambient interval.

We define W_d, w_d, γ_d, j_d, and ν_d in terms of h as in Section 2.5. We let $M = \left(\frac{L}{9J}\right)^{1/k}$, where J is the sum of the absolute value of all the coefficients of h_d, and hence $h_d([M]^\ell) \subseteq [-L/9, L/9]$. We let $Z = \{\boldsymbol{n} \in \mathbb{Z}^\ell : h_d(\boldsymbol{n}) = 0\}$, and we let $H = \big([M]^\ell \cap W_d(Y)\big) \setminus Z$. We note that the hypotheses $\mathcal{Q}' \geq C_0$ and $L \geq \sqrt{N}$ allow us to assume at any point that $\mathcal{Q}'$, $\mathcal{Q}$, and L are all sufficiently large with respect to h. Under this assumption, it follows from Lemmas 2.7 and 2.9 (with $t = c\sqrt{\log M}/\log Y$), and the estimate

$$|Z \cap [M]^\ell| \ll_h M^{\ell-1}, \tag{15}$$

that for $\ell \geq 2$ we have

$$T = \sum_{\boldsymbol{n}\in H} \nu_d(\boldsymbol{n}) \gg_h \left|\prod_{i=1}^{\ell} (M - \chi((r_d)_i)M^\rho/\rho)\right| \gg (M/q_0)^\ell,$$

and for $\ell = 1$ we have

$$T = \sum_{n\in H} \nu_d(n) \gg_h \int_1^M h'_d(x)\left(1 - \chi(r_d)(dx)^{\rho-1}\right)dx \gg L/q_0.$$

Proof of Lemma 3.2 for $\ell \geq 2$. Suppose $\ell \geq 2$, $A \subseteq [L]$ with $|A| = \delta L$, $(A-A)\cap h_d(\Lambda_d) \subseteq \{0\}$, $C_0, \delta^{-1} \leq \mathcal{Q}'$, $q_0 \mid d$, and $d/q_0 \leq \mathcal{Q}$. Further, let $\eta = c_0\delta$, $Q = \eta^{-2}$, and $Y = \eta^{-2k}$. As $h_d(H) \subseteq [-L/9, L/9]\setminus\{0\}$, we have

$$\begin{aligned}\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} f_A(x)f_A(x+h_d(\boldsymbol{n}))\nu_d(\boldsymbol{n}) &= \sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} 1_A(x)1_A(x+h_d(\boldsymbol{n}))\nu_d(\boldsymbol{n}) - \delta\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} 1_A(x)1_{[L]}(x+h_d(\boldsymbol{n}))\nu_d(\boldsymbol{n}) \\ &\quad - \delta\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} 1_A(x+h_d(\boldsymbol{n}))1_{[L]}(x)\nu_d(\boldsymbol{n}) + \delta^2\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} 1_{[L]}(x)1_{[L]}(x+h_d(\boldsymbol{n}))\nu_d(\boldsymbol{n}) \\ &\leq \Big(\delta^2 L - 2\delta|A\cap(L/9,8L/9)|\Big)T.\end{aligned}$$

Therefore, if $|A \cap (L/9, 8L/9)| \geq 3\delta L/4$, we have

$$\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} f_A(x)f_A(x+h_d(\boldsymbol{n})) \leq -\delta^2 LT/2. \tag{16}$$

We see from (15) and orthogonality of characters that

$$\sum_{\substack{x\in\mathbb{Z}\\ \boldsymbol{n}\in H}} f_A(x)f_A(x+h_d(\boldsymbol{n})) = \int_0^1 |\widehat{f_A}(\alpha)|^2 S(\alpha)d\alpha + O_h(LM^{\ell-1}\log^\ell(dM)), \tag{17}$$

where

$$S(\alpha) = \sum_{\boldsymbol{n}\in[M]^\ell\cap W_d(Y)} \nu_d(\boldsymbol{n})e(h_d(\boldsymbol{n})\alpha).$$

Combining (16) and (17), we have

$$\int_0^1 |\widehat{f_A}(\alpha)|^2|S(\alpha)|d\alpha \geq \delta^2 LT/4. \tag{18}$$

Letting $\gamma = \eta^{-2k}/L$, we deduce in Section 5 that for $\alpha \in \mathbf{M}_q(\gamma),\ q \leq Q$, we have

$$|S(\alpha)| \ll_h \begin{cases} (q/\varphi(q))^C((k-1)^2+2)^{\omega(q)}T/q & \ell = 2 \\ C^{\omega(q)}T/q^{3/2} & \ell \geq 3 \end{cases}, \tag{19}$$

where $C = C(h)$, while for $\alpha \in \mathfrak{m}(\gamma, Q)$ we have

$$|S(\alpha)| \leq \delta T/8. \tag{20}$$

From (20) and Plancherel's Identity, we have

$$\int_{\mathfrak{m}(\gamma,Q)} |\widehat{f_A}(\alpha)|^2|S(\alpha)|d\alpha \leq \delta^2 LT/8,$$

which together with (18) yields

$$\int_{\mathfrak{M}(\gamma,Q)} |\widehat{f_A}(\alpha)|^2|S(\alpha)|d\alpha \geq \delta^2 LT/8. \tag{21}$$

From (19) and (21) , we have

$$\sum_{q=1}^{Q}(q/\varphi(q))^C((k-1)^2+2)^{\omega(q)}q^{-1}\int_{\mathbf{M}_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \gg_h \delta^2 L \tag{22}$$

when $\ell = 2$ and

$$\sum_{q=1}^{Q} C^{\omega(q)}q^{-3/2}\int_{\mathbf{M}_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \gg_h \delta^2 L \tag{23}$$

when $\ell \geq 3$. For $\ell = 2$, the function $b(q) = (q/\varphi(q))^C((k-1)^2+2)^{\omega(q)}$ satisfies $b(qr) \geq b(r)$, and we make use of the following proposition, which is based on a trick that originated in [28].

Proposition 3.4 (Proposition 5.6, [24])**.** *For any $\gamma, Q > 0$ satisfying $2\gamma Q^2 < 1$ and any function $b : \mathbb{N} \to [0,\infty)$ satisfying $b(qr) \geq b(r)$ for all $q, r \in \mathbb{N}$, we have*

$$\max_{q\leq Q}\int_{\mathbf{M}'_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \geq Q\Big(2\sum_{q=1}^{Q} b(q)\Big)^{-1}\sum_{r=1}^{Q}\frac{b(r)}{r}\int_{\mathbf{M}_r(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha.$$

Because b is multiplicative, $b(p^v) = ((k-1)^2+2)(1+1/(p-1))^C \ll_k 1$ for all prime powers p^v, and

$$\sum_{q=1}^{Q}\frac{b(q)}{q} \leq \prod_{p\leq Q}\left(1+\frac{b(p)}{p}+\frac{b(p)}{p^2}+\cdots\right) = \prod_{p\leq Q}\left(1+\frac{(k-1)^2+2}{p}+O_k(1/p^2)\right) \ll_k \log^{(k-1)^2+2} Q,$$

it follows from [12, Theorem 01] that

$$\sum_{q=1}^{Q} b(q) \ll_k Q \log^{(k-1)^2+1} Q,$$

and the lemma for $\ell = 2$ follows from (22) and Proposition 3.4. For $\ell \geq 3$, since $C^{\omega(q)} \ll_{h,\epsilon} q^{\epsilon}$ for every $\epsilon > 0$, the sum $\sum_{q=1}^{\infty} C^{\omega(q)} q^{-3/2}$ is convergent, and hence (23) immediately yields

$$\max_{q \leq Q} \int_{\mathbf{M}_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \gg_h \delta^2 L.$$

Since $\mathbf{M}_q(\gamma) \subseteq \mathbf{M}'_q(\gamma)$, this establishes the lemma for $\ell \geq 3$. □

The major contribution of [3] was to replace an iterative estimate on repeated sumsets of rational numbers, developed in [23], with a single higher additive energy estimate. The definitions and results that we import from [3] are as follows, after which we prove Lemma 3.2 in the case $\ell = 1$.

Definition 3.5. For $m \geq 1$, $\mathcal{B} \subseteq \mathbb{T}$, and $\varepsilon > 0$, we define

$$E_{2m}(\mathcal{B}) = |\{b_1, \dots, b_{2m} \in \mathcal{B} : b_1 + \cdots + b_m = b_{m+1} + \cdots + b_{2m}\}|$$

and

$$E_{2m}(\mathcal{B}, \varepsilon) = |\{b_1, \dots, b_{2m} \in \mathcal{B} : \|b_1 + \cdots + b_m - b_{m+1} - \cdots - b_{2m}\|_{\mathbb{T}} \leq \varepsilon\}|.$$

Lemma 3.6 (Theorem 2, [3])**.** *Suppose $m \geq 2$, $Q \geq 4$ and $\mathcal{B} \subseteq \{a/q \in \mathbb{T} : q \leq Q\}$, and for each q let $\mathcal{B}_q$ denote the elements of $\mathcal{B}$ of reduced denominator q. If $|\mathcal{B}_q| \leq n$ for all q, then*

$$E_{2m}(\mathcal{B}) \leq (Qn)^m \log^{C^m} Q,$$

where $C > 0$ is an absolute constant.

Lemma 3.7 (Lemma 7, [3])**.** *Suppose $\varepsilon > 0$, $A \subseteq [L]$ with $|A| = \delta L$ and $\mathcal{B} \subseteq \mathbb{T}$. Then, for each $m \geq 1$,*

$$\sum_{\alpha \in \mathcal{B}} |\widehat{1_A}(\alpha)| \ll \delta^{1-1/2m} L E_{2m}(\mathcal{B}, (2L)^{-1})^{1/2m}.$$

Proof of Lemma 3.2 for $\ell = 1$. Here we follow closely the methods of Lemmas 5 and 6 in [3]. Suppose $\ell = 1$, $A \subseteq [L]$ with $|A| = \delta L$, $(A - A) \cap h_d(\Lambda_d) \subseteq \{0\}$, $C_0, \delta^{-1} \leq \mathcal{Q}'$, $q_0 \mid d$, and $d/q_0 \leq \mathcal{Q}$. Further, let $\eta = c_0 \delta$, let $Q = \eta^{-2}$, and let $Y = \eta^{-2k}$. Similar to the beginning of the proof in the $\ell \geq 2$ case, but simpler because we use only one balanced function instead of two, we have that if $|A \cap (L/9, 8L/9)| \geq 3\delta L/4$, then

$$\sum_{\substack{x \in \mathbb{Z} \\ n \in H}} f_A(x) f_A(x + h_d(n)) \nu_d(n) = \int_0^1 \widehat{f_A}(\alpha) \overline{\widehat{1_A}(\alpha)} S(\alpha) d\alpha + O_h(L(dM)^{k-1} \log(dM)) \leq -3\delta^2 LT/4,$$

where $S(\alpha)$ is defined as before, hence

$$\int_0^1 |\widehat{f_A}(\alpha)||\widehat{1_A}(\alpha)||S(\alpha)| d\alpha \geq \delta^2 LT/2.$$

Our deduction of (20) in Section 5 still applies when $\ell = 1$, so as before, with Cauchy-Schwarz in place of Plancherel, we have

$$\int_{\mathfrak{M}(\gamma, Q)} |\widehat{f_A}(\alpha)||\widehat{1_A}(\alpha)||S(\alpha)| d\alpha = \sum_{q=1}^{Q} \sum_{(a,q)=1} \int_{\mathbf{M}_q(\gamma)} |\widehat{f_A}(\alpha)||\widehat{1_A}(\alpha)||S(\alpha)| d\alpha \geq \delta^2 LT/4.$$

We will show in Section 5 that

$$\int_{\mathbf{M}_{a/q}} |S(\alpha)|^2 d\alpha \ll_h C^{\omega(q)} \frac{T^2}{qL}, \tag{24}$$

for $q \leq Q$ and $(a,q)=1$, where $C=C(h)$, so again applying Cauchy-Schwarz we have

$$\sum_{q=1}^{Q} \sum_{(a,q)=1} C^{\omega(q)} q^{-1/2} \sup_{\alpha \in \mathbf{M}_{a/q}} |\widehat{1_A}(\alpha)| \left(\int_{\mathbf{M}_{a/q}} |\widehat{f_A}(\alpha)|^2 d\alpha \right)^{1/2} \gg_h \delta^2 L^{3/2}.$$

Let

$$R = \left\{ a/q \in \mathbb{T} : q \leq Q, \int_{\mathbf{M}_{a/q}(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \leq L/Q^5 \right\},$$

so $|R| \leq Q^2$ and

$$\sum_{a/q \in R} C^{\omega(q)} q^{-1/2} \sup_{\alpha \in \mathbf{M}_{a/q}} |\widehat{1_A}(\alpha)| \left(\int_{\mathbf{M}_{a/q}} |\widehat{f_A}(\alpha)|^2 d\alpha \right)^{1/2} \ll_h Q^2 \delta L^{3/2}/Q^{5/2} = \delta L^{3/2}/Q^{1/2} = \delta\eta L^{3/2},$$

hence

$$\sum_{q=1}^{Q} \sum_{a/q \notin R} C^{\omega(q)} q^{-1/2} \sup_{\alpha \in \mathbf{M}_{a/q}} |\widehat{1_A}(\alpha)| \left(\int_{\mathbf{M}_{a/q}} |\widehat{f_A}(\alpha)|^2 d\alpha \right)^{1/2} \gg_h \delta^2 L^{3/2}. \tag{25}$$

Further, because the measure of $\mathbf{M}_{a/q}$ is Q^k/L and $|\widehat{f_A}(\alpha)| \ll \delta L$, we know

$$\int_{\mathbf{M}_{a,q}(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \ll Q^k \delta^2 L.$$

Dyadically pigeonholing in both q and the integral value, there exist $1 \leq Q' \leq Q$, $Q^{-k} \leq K \leq Q^3$, and $\mathcal{B} \subseteq \{a/q \in \mathbb{T} : q \leq Q, a/q \notin R\}$ such that all reduced denominators in $\mathcal{B}$ are between Q' and $2Q'$,

$$\delta^2 L/K^2 \leq \int_{\mathbf{M}_{a,q}(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha \leq 2\delta^2 L/K^2 \quad \text{for all } a/q \in \mathcal{B}, \tag{26}$$

and

$$\sum_{a/q \in \mathcal{B}} C^{\omega(q)} q^{-1/2} \sup_{\alpha \in \mathbf{M}_{a/q}} |\widehat{1_A}(\alpha)| \left(\int_{\mathbf{M}_{a/q}} |\widehat{f_A}(\alpha)|^2 d\alpha \right)^{1/2} \gg_h \delta^2 L^{3/2}/\log^2 Q. \tag{27}$$

Letting $\alpha_{a/q}$ denote the point in $\mathbf{M}_{a/q}$ on which $|\widehat{1_A}|$ attains its maximum, substituting (26) into (27) gives

$$\sum_{a/q \in \mathcal{B}} |\widehat{1_A}(\alpha_{a/q})| \gg_h \frac{\delta\sqrt{Q'}KL}{\tau \log^2 Q}, \tag{28}$$

where $\tau = \max_{q \leq Q} C^{\omega(q)} \leq \exp(C\log(2\delta^{-1})/\log\log(3\delta^{-1}))$. Let

$$\theta = (\delta^2 L)^{-1} \max_{q \leq Q} \int_{\mathbf{M}_q(\gamma)} |\widehat{f_A}(\alpha)|^2 d\alpha.$$

For fixed q, we let $\mathcal{B}_q$ denote the elements of $\mathcal{B}$ with reduced denominator exactly q, and (26) yields

$$|\mathcal{B}_q|\delta^2 L/K^2 \leq \sum_{a/q \in \mathcal{B}_q} \int_{\mathbf{M}_{a/q}} |\widehat{f_A}(\alpha)|^2 d\alpha \leq \theta\delta^2 L,$$

so in particular

$$|\mathcal{B}_q| \leq \theta K^2. \tag{29}$$

Letting $m = 2\lceil \log\log(3\delta^{-1}) \rceil$, (28) and the pigeonhole principle ensure the existence of $\mathcal{B}' \subseteq \mathcal{B}$, contained in an interval of length $(8m)^{-1}$ satisfying

$$\sum_{a/q \in \mathcal{B}'} |\widehat{1_A}(\alpha_{a/q})| \gg_h \frac{\delta\sqrt{Q'}KL}{m\tau \log^2 Q}. \tag{30}$$

Letting $\Gamma = \{\alpha_{a/q} : a/q \in \mathcal{B}'\}$, (30) and Lemma 3.7 yield

$$\frac{\delta\sqrt{Q'}KL}{\tau m \log^2 Q} \ll_h \delta^{1-1/2m} L E_{2m}(\Gamma, (2L)^{-1})^{1/2m}. \tag{31}$$

However, the elements inside the norm in the definition of $E_{2m}(\Gamma, (2L)^{-1})$ are all rationals of denominator at most $(Q')^{2m}$, and since $(Q')^{2m} = \mathcal{Q}^{O_h(\log\log L)} \leq L$, such a rational can only be less than $(2L)^{-1}$ in absolute value if it is 0, meaning $E_{2m}(\Gamma, (2L)^{-1}) = E_{2m}(\Gamma)$. Combining with (31), (29), and Lemma 3.6, we have

$$(Q'\theta K^2)^m \log^{C^m} Q \geq E_{2m}(\Gamma) \gg_h \delta \left(\frac{\sqrt{Q'}K}{\tau m \log^2 Q} \right)^{2m},$$

which rearranges to

$$\theta \geq \frac{\delta^{1/m}}{m^2 \tau^2 \log^{C^m} Q}.$$

Since $Q \ll_h \delta^{-2}$ and $m = 2\lceil \log\log(3\delta^{-1})\rceil$, this yields the desired lower bound

$$\theta \gg_h \exp\left(-C \frac{\log(2\delta^{-1})}{\log\log(3\delta^{-1})} \right).$$

□

4. Criteria for $\mathcal{P}$-Deligne Polynomials

In this section we establish the sufficient conditions for $\mathcal{P}$-Deligne polynomials enumerated in Theorem 1.7. Most of the statements we make here are analogous to certain statements in Sections 2 and 5 of [7], and in those cases we simply mention the corresponding statement in [7] and that the proof is essentially the same.

We begin with two geometric lemmas. The first is a straightforward consequence of the point-counting estimates for varieties over finite fields due to Lang and Weil; a short proof is provided in [7, Lemma 5.2]. We use V^{ns} to denote the nonsingular points on a variety V.

Lemma 4.1. *Let k, ℓ, m, and r be positive integers, and let q be a prime power. Let V be a (reduced) closed subvariety of $\mathbb{P}^\ell$, defined over $\mathbb{F}_q$, of degree k and dimension r. Let $m \geq 1$ be the number of geometrically irreducible components of V which are defined over $\mathbb{F}_q$. Then*

$$|V(\mathbb{F}_q)|,\ |V^{\mathrm{ns}}(\mathbb{F}_q)| = mq^r + O_{k,\ell,r}(q^{r-1/2}). \tag{32}$$

Moreover, the same is true if we replace V with a closed subvariety $W \subseteq \mathbb{A}^\ell$.

The second geometric lemma is a slight variation on [7, Lemma 5.3].

Lemma 4.2. *Let $V \subseteq \mathbb{P}^\ell$ be a variety (reduced, but not necessarily irreducible) of dimension $r \geq 1$ defined over $\mathbb{Z}$, let V^{ns} be the nonsingular locus of V, and let V_0^{ns} be the Zariski open subset of V^{ns} obtained by imposing the conditions $x_i \neq 0$ for all $0 \leq i \leq \ell$. If p is sufficiently large (with respect to V), the following are equivalent:*

(a) $V_0^{\mathrm{ns}}(\mathbb{F}_p) \neq \emptyset$.
(b) $V_0^{\mathrm{ns}}(\mathbb{Z}_p) \neq \emptyset$.
(c) At least one of the geometric components of V is defined over $\mathbb{Z}_p$ and is not contained in a coordinate hyperplane $\{x_i = 0\}$.

Proof. The proof is the same as for [7, Lemma 5.3], with one additional observation: in the context of showing that (c) implies (a), if Z is an irreducible component of V defined over $\mathbb{Z}_p$ not contained in a coordinate hyperplane, then the number of elements of $Z(\mathbb{F}_p)$ with at least one coordinate 0 is at most $O_{k,\ell,r}(p^{r-1})$ by Lemma 4.1; applying Lemma 4.1 again, there are plenty of nonsingular points leftover, provided p is sufficiently large. □

In the remainder of this section we prove Theorem 1.7. We will use the following definition, which is modified from [7, Definition 2.7].

Definition 4.3. For $\ell \in \mathbb{N}$ and $h \in \mathbb{Z}[x_1, \dots, x_\ell]$, we say that h is *smoothly $\mathcal{P}$-intersective* if there exists a choice $\{\boldsymbol{z}_p\}_{p\in\mathcal{P}}$ of p-adic integer roots of h such that $(\boldsymbol{z}_p)_i \not\equiv 0 \pmod p$ for all $1 \leq i \leq \ell$ and all p, and $m_p = 1$ for all but finitely many p.

Now, part (ii) of Theorem 1.7 is proven with an argument which is identical to that of [7, Proposition 2.5]. The next proposition is item (iv) from Theorem 1.7, and we note that the proof is essentially the same as for [7, Proposition 2.8], but using Lemma 4.2 in place of [7, Lemma 5.3].

Proposition 4.4. *Suppose $\ell \geq 2$ and $h \in \mathbb{Z}[x_1, \dots, x_\ell]$ is Deligne and $\mathcal{P}$-intersective with* $\deg(h) = k \geq 2$. *If there exists a choice $\{\boldsymbol{z}_p\}_{p\in\mathcal{P}}$ of p-adic integer roots of h satisfying $(z_p)_i \not\equiv 0 \pmod p$ for all $1 \leq i \leq \ell$ and all p, and $m_p \in \{1, k\}$ for all but finitely many p, then h is $\mathcal{P}$-Deligne. In particular, if $k = 2$ or h is smoothly $\mathcal{P}$-intersective, then h is $\mathcal{P}$-Deligne.*

It remains to show parts (i) and (iii) of Theorem 1.7, both of which are a consequence of the following sufficient condition for smooth $\mathcal{P}$-intersectivity, analogous to [7, Corollary 5.4]. Note that the following proposition is precisely part (iii) from Theorem 1.7.

Proposition 4.5. *Suppose $\ell \geq 2$ and $h \in \mathbb{Z}[x_1, \dots, x_\ell]$ is Deligne and $\mathcal{P}$-intersective, and let $h = g_1 \cdots g_n$ be an irreducible factorization of h in $\overline{\mathbb{Z}}[x_1, \dots, x_\ell]$. If, for all but finitely many $p \in \mathcal{P}$, there exists $1 \leq i \leq n$ such that g_i has coefficients in $\mathbb{Z}_p$ and $x_j \nmid g_i$ for all $1 \leq j \leq \ell$, then h is smoothly $\mathcal{P}$-intersective, hence $\mathcal{P}$-Deligne.*

Proof. Suppose $\ell \geq 2$ and $h \in \mathbb{Z}[x_1, \dots, x_\ell]$ satisfies the hypotheses of the proposition. Lemma 4.2 gives us exactly what we need for smooth $\mathcal{P}$-intersectivity—a nonsingular $\mathbb{Z}_p$ point, with all coordinates nonzero modulo p, for all but finitely many p—except in the following pathological scenario: for infinitely many primes p, every irreducible component of the variety $\{h = 0\}$ defined over $\mathbb{Z}_p$ is in fact a coordinate hyperplane. □

For $\ell \geq 3$, a reducible hypersurface in $\mathbb{P}^{\ell-1}$ must be singular; thus, a Deligne polynomial in at least three variables must be geometrically irreducible. We therefore obtain the following consequence of Proposition 4.5, proving item (i) from—and thus completing the proof of—Theorem 1.7.

Corollary 4.6. *If $\ell \geq 3$ and $h \in \mathbb{Z}[x_1, \dots, x_\ell]$ is Deligne and $\mathcal{P}$-intersective, then h is smoothly $\mathcal{P}$-intersective, hence $\mathcal{P}$-Deligne.*

5. Exponential Sum Estimates

In this section, we import all objects and parameters from Section 3.3, introduced prior to and during the proof of Lemma 3.2. Further, for $q \in \mathbb{N}$ we let

$$W_{d,q}(Y) = \Big\{ \boldsymbol{n} \in \mathbb{Z}^\ell : ((r_d)_i + dn_i, q) = 1 \text{ for all } 1 \leq i \leq \ell,\ \nabla h_d(\boldsymbol{n}) \not\equiv \boldsymbol{0} \bmod p^{\gamma_d(p)} \text{ for all } p \leq Y, p^{\gamma_d(p)} \mid q \Big\},$$

and for $\boldsymbol{s} \in [q]^\ell$ and a prime p with $p^{\gamma_d(p)} \nmid q$, we let $j_{d,q,\boldsymbol{s}}(p)$ denote the number of ℓ-tuples of congruence classes $\boldsymbol{c}$ modulo $p^{\gamma_d(p)}$ satisfying $\nabla h_d(\boldsymbol{c}) \equiv \boldsymbol{0} \pmod{p^{\gamma_d(p)}}$, $p \nmid (r_d)_i + dc_i$ for all $1 \leq i \leq \ell$, and $\boldsymbol{c} \equiv \boldsymbol{s} \pmod{p^{\mathrm{ord}_p(q)}}$. We let $\epsilon_{d,q}(p) = 0$ if $p \mid dq$ and 1 otherwise. Finally, we let

$$w_{d,q}(\boldsymbol{s}) = \prod_{\substack{p \leq Y \\ p^{\gamma_d(p)} \nmid q}} \left(1 - \frac{j_{d,q,\boldsymbol{s}}(p)}{((p - \epsilon_{d,q}(p))p^{\gamma_d(p) - \mathrm{ord}_p(q) - 1})^\ell} \right).$$

5.1. **Major arc estimate.** A minor adaptation of the proof of Proposition 2.7, with $t = c\sqrt{\log M}/\log Y$, gives the following estimates. For the remainder of the section, we let E denote an error term of the form $M^\ell \exp(-c\sqrt{\log M})$ for $\ell \geq 2$, and $L\exp(-c\sqrt{\log M})$ for $\ell = 1$, for a constant $c = c(h) > 0$, noting that E can absorb terms of the form $\mathcal{Q}^{O_h(1)}$.

Lemma 5.1. *Suppose $\ell \geq 2$. For $0 < x_1, \dots, x_\ell \leq M$, $q \leq \mathcal{Q}^{O_h(1)}$ and $\boldsymbol{s} \in [q]^\ell$, we have*

$$\sum_{\substack{\boldsymbol{n}\in B\\ \boldsymbol{n}\equiv \boldsymbol{s} (\mathrm{mod}\ q)}} \nu_d(\boldsymbol{n}) = \left(\frac{\varphi(d)}{\varphi(qd)}\right)^\ell \frac{w_{d,q}(\boldsymbol{s})}{w_d} \prod_{i=1}^{\ell} \left(x_i - \frac{\chi((r_d)_i)x_i^\rho}{\rho d^{1-\rho}}\right) 1_{W_{d,q}(Y)}(\boldsymbol{s}) + O(E).$$

Lemma 5.2. *Suppose $\ell = 1$. For $0 < x \leq M$, $q \leq \mathcal{Q}^{O_h(1)}$ and $s \in [q]$, we have*

$$\sum_{\substack{n\in B\\ n\equiv s (\mathrm{mod}\ q)}} \nu_d(n) = \frac{\varphi(d)}{\varphi(qd)} \frac{w_{d,q}(s)}{w_d} \left(\int_1^x h_d'(t)\left(1 - \chi(r_d)(dt)^{\rho-1}\right) dt\right) 1_{W_{d,q}(Y)}(s) + O(E).$$

We use Lemmas 5.1 and 5.2, together with partial summation, to get asymptotic formulas for $S(\alpha)$ near rationals with small denominator.

Lemma 5.3 (Multivariable Partial Summation, Lemma 7.4 [7])**.** *Suppose $\ell \in \mathbb{N}$ and $a : \mathbb{N}^\ell \to \mathbb{C}$. Suppose further that $b : \mathbb{R}^\ell \to \mathbb{C}$ is C^ℓ. For any $X \geq 1$, we have*

$$\sum_{\boldsymbol{n}\in[X]^\ell} a(\boldsymbol{n})b(\boldsymbol{n}) = A(X,\dots,X)b(X,\dots,X) + \sum_{i=1}^{\ell}(-1)^i \sum_{1\leq j_1<\cdots<j_i\leq \ell} \int_{[1,X]^i} A(\star)\frac{\partial^i b}{\partial x_{j_1}\cdots\partial x_{j_i}}(\star)\ dx_{j_1}\cdots dx_{j_i},$$

where

$$A(x_1,\dots,x_\ell) = \sum_{\boldsymbol{n}\in[x_1]\times\cdots\times[x_\ell]} a(\boldsymbol{n})$$

and $\star = (X, \dots, x_{j_1}, \dots, x_{j_i}, \dots, X)$, with $x_{j_1}, \dots, x_{j_i}$ plugged into coordinate positions $j_1, \dots, j_i$ and all other coordinates evaluated at X.

For the following two lemmas, let J be the sum of the absolute value of all coefficients of h_d.

Lemma 5.4. *Suppose $\ell \geq 2$. If $a, q \in \mathbb{N}$, $q \leq \mathcal{Q}^{O_h(1)}$, and $\alpha = a/q + \beta$, then*

$$S(\alpha) = \left(\frac{\varphi(d)}{\varphi(qd)}\right)^\ell w_d^{-1}\mathcal{G}(a,q)\int_{[1,M]^\ell}\prod_{i=1}^{\ell}\left(1-\chi((r_d)_i)(dx_i)^{\rho-1}\right) e(h_d(\boldsymbol{x})\beta)d\boldsymbol{x} + O_h\left(E(1+JM^k|\beta|)^\ell\right),$$

where

$$\mathcal{G}(a,q) = \sum_{\boldsymbol{s}\in[q]^\ell\cap W_{d,q}(Y)} w_{d,q}(\boldsymbol{s})e(h_d(\boldsymbol{s})a/q).$$

Proof. We begin by noting that for any $a, q \in \mathbb{N}$ and $0 \leq x_1, \dots, x_\ell \leq M$, letting $B = [x_1] \times \cdots \times [x_\ell]$, we have

$$\begin{aligned}\mathcal{S}(x_1,\dots,x_\ell) :&= \sum_{\boldsymbol{n}\in B}\nu_d(\boldsymbol{n})e(h_d(\boldsymbol{n})a/q)\\ &= \sum_{\boldsymbol{s}\in[q]^\ell} e(h_d(\boldsymbol{s})a/q) \sum_{\substack{\boldsymbol{n}\in B\\ \boldsymbol{n}\equiv\boldsymbol{s}\ (\mathrm{mod}\ q)}}\nu_d(\boldsymbol{n}).\end{aligned}$$

Lemma 5.1 then gives

$$\mathcal{S}(x_1,\dots,x_\ell) = \left(\frac{\varphi(d)}{\varphi(qd)}\right)^{\ell} w_d^{-1} \prod_{i=1}^{\ell} \left(x_i - \frac{\chi((r_d)_i)x_i^{\rho}}{\rho d^{1-\rho}}\right) \sum_{\boldsymbol{s}\in[q]^\ell\cap W_{d,q}(Y)} w_{d,q}(\boldsymbol{s})e(h_d(\boldsymbol{s})a/q) + O(E). \tag{33}$$

Letting $b(\boldsymbol{n}) = e(h_d(\boldsymbol{n})\beta)$, we now decompose our sum as

$$S(\alpha) = \sum_{\boldsymbol{n}\in[M]^\ell} \nu_d(\boldsymbol{n})e(h_d(\boldsymbol{n})a/q)b(\boldsymbol{n})$$

and apply Lemma 5.3, yielding

$$\begin{aligned} S(\alpha) &= \mathcal{S}(M,\dots,M)b(M,\dots,M) \\ &\quad + \sum_{m=1}^{\ell}(-1)^m \sum_{1\le j_1<\cdots<j_m\le\ell} \int_{[1,X]^m} \mathcal{S}(\star)\frac{\partial^m b}{\partial x_{j_1}\cdots\partial x_{j_m}}(\star)\ dx_{j_1}\cdots dx_{j_m}, \end{aligned}$$

where $\star$ is as in Lemma 5.3. Substituting (33) gives the main term

$$\begin{aligned} &\left(\frac{\varphi(d)}{\varphi(qd)}\right)^{\ell} w_d^{-1} \sum_{\boldsymbol{s}\in[q]^\ell\cap W_{d,q}(Y)} w_{d,q}(\boldsymbol{s})e(h_d(\boldsymbol{s})a/q)\Big(\prod_{1\le i\le\ell}\left(M - \frac{\chi((r_d)_i)M^{\rho}}{\rho d^{1-\rho}}\right) b(M,\dots,M) \\ &+ \sum_{m=1}^{\ell}(-1)^m \sum_{1\le j_1<\cdots<j_m\le\ell} \prod_{\substack{1\le i\le\ell \\ i\ne j_1,\dots,j_m}} \left(M - \frac{\chi((r_d)_i)M^{\rho}}{\rho d^{1-\rho}}\right) \\ &\cdot \int_{[1,M]^m} \prod_{1\le i\le m}\left(x_i - \frac{\chi((r_d)_{j_i})x_i^{\rho}}{\rho d^{1-\rho}}\right)\frac{\partial^m b}{\partial x_{j_1}\cdots\partial x_{j_m}}(\star)\ dx_{j_1}\cdots dx_{j_m}\Big). \end{aligned}$$

By iteratively applying integration by parts, this equals

$$\left(\frac{\varphi(d)}{\varphi(qd)}\right)^{\ell} w_d^{-1} \sum_{\boldsymbol{s}\in[q]^\ell\cap W_{d,q}(Y)} w_{d,q}(\boldsymbol{s})e(h_d(\boldsymbol{s})a/q) \int_{[1,M]^\ell} \prod_{i=1}^{\ell}\left(1-\chi((r_d)_i)(dx)^{\rho-1}\right)e(h_d(\boldsymbol{x})\beta)d\boldsymbol{x},$$

as desired. The error that results from our substitution of (33) consists of 2^ℓ terms of order

$$O_h\left(E(1+JM^k|\beta|)^\ell\right),$$

completing the proof. □

Analogous to the deduction of Lemma 5.4 from Lemma 5.1, the following asymptotic formula for $\ell = 1$ follows from Lemma 5.2.

Lemma 5.5. *Suppose $\ell = 1$. If $a, q\in\mathbb{N}$, $q\le \mathcal{Q}^{O_h(1)}$, and $\alpha = a/q+\beta$, then*

$$S(\alpha) = \frac{\varphi(d)}{\varphi(qd)} w_d^{-1}\mathcal{G}(a,q)\int_1^M h_d'(x)\left(1-\chi(r_d)(dx)^{\rho-1}\right)e(h_d(x)\beta)dx + O_h\left(E(1+JM^k|\beta|)\right).$$

5.2. **Common divisors and gradient vanishing.** Here we collect facts assuring that several quantities depending a priori on h_d can actually be bounded in terms of the original polynomial h, which we use implicitly in the remainder of Section 5.

Definition 5.6. For $g(\boldsymbol{x}) = \sum_{|\boldsymbol{i}|\le k} a_{\boldsymbol{i}}\boldsymbol{x}^{\boldsymbol{i}} \in \mathbb{Z}[x_1,\dots,x_\ell]$, we define $\mathrm{cont}(g) = \gcd(\{a_{\boldsymbol{i}} : |\boldsymbol{i}|>0\})$.

Proposition 5.7 (Proposition 3.6, [7])**.** *If $g(\boldsymbol{x}) = \sum_{|\boldsymbol{i}|\le k} a_{\boldsymbol{i}}\boldsymbol{x}^{\boldsymbol{i}} \in \mathbb{Z}[x_1,\dots,x_\ell]$ is identically zero modulo $q\in\mathbb{N}$, then $q \mid k!\gcd(\{a_{\boldsymbol{i}}\})$. In particular, if ∇g is identically $\mathbf{0}$ modulo q, then $q \ll_k \mathrm{cont}(g)$.*

Proposition 5.8 (Proposition 6.5, [7])**.** *If $g\in\mathbb{Z}[x_1,\dots,x_\ell]$ is strongly Deligne, then $\mathrm{cont}(g_d)\ll_g 1$.*

Lemma 5.9 (Corollary 7.3, [7])**.** *If $g\in\mathbb{F}_q[x_1,\dots,x_\ell]$ is Deligne of degree $k\ge1$, then*

$$|\{\boldsymbol{x}\in\mathbb{F}_q^\ell : \nabla g(\boldsymbol{x}) = \mathbf{0}| \ll_{k,\ell} 1.$$

5.3. **Local cancellation.** The primary purpose for defining the Deligne condition is the following estimate on multivariate exponential sums over finite fields, due to Deligne in his proof of the Weil conjectures.

Lemma 5.10 (Theorem 8.4, [6]). *Suppose $\ell \in \mathbb{N}$ and $p \in \mathcal{P}$. If $g \in \mathbb{F}_p[x_1, \dots, x_\ell]$ is Deligne, then*

$$\left| \sum_{\boldsymbol{x} \in \mathbb{F}_p^\ell} e(g(\boldsymbol{x})/p) \right| \leq (\deg(g) - 1)^\ell p^{\ell/2}.$$

As first done in [24], to effectively utilize 5.10, we must reduce to the case of prime moduli, which we do with a multivariable version of Hensel's Lemma that follows from [5, Theorem 1.1].

Lemma 5.11 (Multivariable Hensel's Lemma). *Suppose $\ell \in \mathbb{N}$, $g \in \mathbb{Z}[x_1, \dots, x_\ell]$, p is prime, $\boldsymbol{n} \in \mathbb{Z}^\ell$, and $\gamma, v \in \mathbb{N}$ with $v \geq 2\gamma - 1$. If $g(\boldsymbol{n}) \equiv 0 \pmod{p^{2\gamma-1}}$ and $\nabla g(\boldsymbol{n}) \not\equiv \boldsymbol{0} \pmod{p^\gamma}$, then there exists $\boldsymbol{m} \in \mathbb{Z}^\ell$ with $g(\boldsymbol{m}) \equiv 0 \pmod{p^v}$.*

Equipped with Lemmas 5.10 and 5.11, we establish the following estimate on the local exponential sums that appear in our asymptotic formulas.

Lemma 5.12. *If $q \in \mathbb{N}$ has prime factorization $q = p_1^{v_1} \cdots p_r^{v_r}$ with $p_1 < \cdots < p_t \leq Y < p_{t+1} < \cdots < p_r$, and $(a, q) = 1$, then*

$$\begin{aligned} |\mathcal{G}(a,q)| \leq C_1 &\prod_{\substack{p \leq Y \\ p \nmid q}} \left(1 - \frac{j_d(p)}{((p - \epsilon_d(p))p^{\gamma_d(p)-1})^\ell}\right) \\ &\cdot \prod_{i=1}^{t} \Big((k-1)^\ell p_i^{\ell/2} + \mathbb{1}_{\ell=1} + (2p-1)\mathbb{1}_{\ell=2} + [(k-1)p_i^{\ell-3/2} + ((k-2)\ell + 2^\ell)p_i^{\ell-2}]\mathbb{1}_{\ell \geq 3} + j_d(p_i)\Big) \\ &\cdot \prod_{i=t+1}^{r} C_2(v_i+1)^\ell p_i^{v_i(\ell - 1/k)}, \end{aligned}$$

where $C_1 = C_1(h)$ and $C_2 = C_2(h)$. Further, $\mathcal{G}(a,q) = 0$ if $v_i \geq 2\gamma_d(p_i)$ for some $1 \leq i \leq t$.

Proof. Factor $q = p_1^{v_1} \cdots p_r^{v_r}$ as in the lemma. By the Chinese Remainder Theorem, we have

$$\mathcal{G}(a,q) = \prod_{\substack{p \leq Y \\ p \nmid q}} \left(1 - \frac{j_d(p)}{((p - \epsilon_d(p))p^{\gamma_d(p)-1})^\ell}\right) \prod_{m=1}^{r} \tilde{\mathcal{G}}(a, p_m^{v_m}),$$

where

$$\tilde{\mathcal{G}}(a, p_m^{v_m}) = \sum_{\boldsymbol{s} \in [p_m^{v_m}]^\ell \cap W_{d, p_m^{v_m}}(Y)} e(ah_d(\boldsymbol{s})/p_m^{v_m}) \cdot \begin{cases} c(\boldsymbol{s}) & \text{if } m \leq t,\ v_m < \gamma_d(p_m) \\ 1 & \text{else} \end{cases},$$

and $|c(\boldsymbol{s})| \leq 1$.

Suppose $p^v = p_m^{v_m}$ with $\gamma_d(p) > 1$ and $v < 2\gamma_d(p)$. Since $p^{2\gamma_d(p)-1} \leq p^{3(\gamma_d(p)-1)}$, we can trivially bound the contributions from all such p^v by the cube of the product of prime powers $p^{\gamma_d(p)}$ for which $\gamma_d(p) > 1$, which is $O_h(1)$, and absorb them into $C_1 = C_1(h)$.

Next suppose $p^v = p_m^{v_m}$ with $p \leq Y$ and $v = \gamma_d(p) = 1$. If h_d is not Deligne modulo p, absorb $\tilde{\mathcal{G}}(a,p)$ into C_1. Otherwise, recalling that $j_d(p)$ is the number of zeros of ∇h_d modulo p on $J_d(p)$ we have for $p \nmid a$ that

$$\left| \sum_{\boldsymbol{s} \in [p]^\ell \cap W_{d,p}(Y)} e(ah_d(\boldsymbol{s})/p) \right| \leq \left| \sum_{\substack{\boldsymbol{s} \in [p]^\ell \\ p \nmid (r_d)_i + ds_i \text{ for all } 1 \leq i \leq \ell}} e(ah_d(\boldsymbol{s})/p) \right| + j_d(p). \tag{34}$$

If $p \mid d$, the sum on the right hand side is complete and bounded by $(k-1)^\ell p^{\ell/2}$ by Theorem 5.10. If $p \nmid d$, let $m_i = -d^{-1}(r_d)_i \pmod p$ for $1 \le i \le \ell$, so the sum on the right hand side above is

$$\sum_{j=0}^{\ell}(-1)^j \sum_{1\le i_1<\cdots<i_j\le \ell} \sum_{\substack{\boldsymbol{s}\in[p]^\ell \\ s_{i_n}=m_{i_n} \text{ for all } 1\le n\le j}} e(ah_d(\boldsymbol{s})/p). \tag{35}$$

For $j=0$, we have a complete sum, which is bounded by $(k-1)^\ell p^{\ell/2}$ by Theorem 5.10. If $\ell=1$, we are removing only a single term, while if $\ell=2$ we are removing $2p-1$ terms.

Now suppose $\ell \ge 3$. For $j=1$, by geometric irreducibility, we know that h_d is nonconstant modulo p on each of the hyperplanes $x_i = m_i$. Let $g_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_\ell) = h_d(x_1,\dots,x_{i-1},m_i,x_{i+1},\dots,x_\ell)$ for each $1 \le i \le \ell$. Since we cannot guarantee that g_i is a Deligne polynomial in $\ell-1$ variables, we only use that it is nonconstant and exploit cancellation in a single variable. For each $\tilde{\boldsymbol{s}} \in [p]^{\ell-2}$, we let $g_{i,\tilde{\boldsymbol{s}}}(x) = g_i(x,\tilde{\boldsymbol{s}})$ for $i>1$ and $g_{1,\tilde{\boldsymbol{s}}}(x) = g_1(\tilde{\boldsymbol{s}},x)$. Then, we have

$$\begin{aligned}\left|\sum_{\substack{\boldsymbol{s}\in[p]^\ell \\ s_i=m_i}} e(ah_d(\boldsymbol{s})/p)\right| &= \left|\sum_{\boldsymbol{s}\in[p]^{\ell-1}} e(ag_i(\boldsymbol{s})/p)\right| \\ &= \left|\sum_{\tilde{\boldsymbol{s}}\in[p]^{\ell-2}}\sum_{x\in[p]} e(ag_{i,\tilde{\boldsymbol{s}}}(x)/p)\right| \\ &\le (k-1)p^{1/2}\sum_{\tilde{\boldsymbol{s}}\in[p]^{\ell-2}} \gcd(\operatorname{cont}(g_{i,\tilde{\boldsymbol{s}}}),p)^{1/2}.\end{aligned}$$

Consider a coefficient of g_i not divisible by p, with a positive exponent on x_1 (or x_ℓ if $i=1$, and if no such coefficient exists relabel the coordinates). Then, $\gcd(\operatorname{cont}(g_{i,\tilde{\boldsymbol{s}}}),p) = p$ only if p divides the product of the coordinates of $\tilde{\boldsymbol{s}}$, which occurs for fewer than $\ell p^{\ell-3}$ choices. Putting everything together, we have

$$\left|\sum_{\substack{\boldsymbol{s}\in[p]^\ell \\ s_i=m_i}} e(ah_d(\boldsymbol{s})/p)\right| \le (k-1)(p^{\ell-3/2}+\ell p^{\ell-2}),$$

and hence by (34) and (35), trivially bounding the terms with $j>1$, we know

$$\left|\sum_{\boldsymbol{s}\in[p]^\ell\cap W_{d,p}(Y)} e(ah_d(\boldsymbol{s})/p)\right| \le (k-1)^\ell p^{\ell/2} + (k-1)p^{\ell-3/2} + ((k-2)\ell+2^\ell)p^{\ell-2} + j_d(p).$$

Now suppose that $p^v = p_m^{v_m}$ with $p \le Y$ and $v \ge 2\gamma_d(p)$, and let $w = 2\gamma_d(p)-1$. If $\boldsymbol{s} \in [p^v]^\ell$ and $\tilde{\boldsymbol{s}}$ is the reduced residue class of $\boldsymbol{s}$ modulo p^w, then $h_d(\boldsymbol{s}) \equiv p^w t + h_d(\tilde{\boldsymbol{s}}) \pmod{p^v}$ for some $0 \le t \le p^{v-w}-1$. Conversely, if $\tilde{\boldsymbol{s}} \in [p^w]^\ell$ with $\nabla h_d(\tilde{\boldsymbol{s}}) \not\equiv \boldsymbol{0} \pmod{p^{\gamma_d(p)}}$, then for every $0 \le t \le p^{v-w}-1$, Lemma 5.11 applied to the polynomial $h_d(\boldsymbol{x}) - (p^w t + h_d(\tilde{\boldsymbol{s}}))$ yields $\boldsymbol{s} \in [p^v]^\ell$ with $h_d(\boldsymbol{s}) \equiv p^w t + h_d(\tilde{\boldsymbol{s}}) \pmod{p^v}$.

In other words, the map F on $\mathbb{Z}/p^{v-w}\mathbb{Z}$ defined by $h_d(p^wt+\tilde{\boldsymbol{s}}) \equiv p^wF(t)+h_d(\tilde{\boldsymbol{s}}) \pmod{p^v}$ is a bijection. In particular,

$$\begin{aligned}\sum_{\boldsymbol{s}\in[p^v]^\ell\cap W_{d,p^v}(Y)} e(ah_d(\boldsymbol{s})/p^v) &= \sum_{\substack{\tilde{\boldsymbol{s}}\in[p^w]^\ell \\ \nabla h_d(\tilde{\boldsymbol{s}})\not\equiv\boldsymbol{0} \pmod{p^{\gamma_d(p)}} \\ p\nmid(r_d)_i+d\tilde{s}_i \text{ for } 1\le i\le\ell}} \sum_{t=0}^{p^{v-w}-1} e(ah_d(p^wt+\tilde{\boldsymbol{s}})/p^v) \\ &= \sum_{\substack{\tilde{\boldsymbol{s}}\in[p^w]^\ell \\ \nabla h_d(\tilde{\boldsymbol{s}})\not\equiv\boldsymbol{0} \pmod{p^{\gamma_d(p)}} \\ p\nmid(r_d)_i+d\tilde{s}_i \text{ for } 1\le i\le\ell}} \sum_{t=0}^{p^{v-w}-1} e\left(a\left(p^wt+h_d(\tilde{\boldsymbol{s}})\right)/p^v\right) \\ &= 0,\end{aligned}$$

where the last equality is the fact that the sum in t runs over the full collection of p^{v-w}-th roots of unity.

Finally, suppose $p^v = p_m^{v_m}$ with $p > Y$, so there is no longer a gradient nonvanishing condition. Similar to the $j=1$ case of (35), we only exploit cancellation in a single variable.

To this end, for $\tilde{\boldsymbol{s}} = (s_2,\dots,s_\ell) \in [p^v]^{\ell-1}$, we define $g_{\tilde{\boldsymbol{s}}}$ by $g_{\tilde{\boldsymbol{s}}}(x) = h_d(x,\tilde{\boldsymbol{s}})$, and we have

$$\begin{aligned}\left|\sum_{\substack{\boldsymbol{s}\in[p^v]^\ell \\ p\nmid(r_d)_i+d\tilde{s}_i \text{ for } 1\le i\le\ell}} e(ah_d(\boldsymbol{s})/p^v)\right| &\le \sum_{\tilde{\boldsymbol{s}}\in[p^v]^{\ell-1}} \left|\sum_{\substack{x\in[p^v] \\ p\nmid(r_d)_1+dx}} e(ag_{\tilde{\boldsymbol{s}}}(x)/p^v)\right| \\ &= \sum_{\tilde{\boldsymbol{s}}\in[p^v]^{\ell-1}} \left|\sum_{x\in[p^v]} e(ag_{\tilde{\boldsymbol{s}}}(x)/p^v) - \sum_{y\in[p^{v-1}]} e(a\overline{g_{\tilde{\boldsymbol{s}}}}(y)/p^{v-1})\right|,\end{aligned}$$

where $m \equiv -d^{-1}(r_d)_1 \pmod{p}$, $\overline{g_{\tilde{\boldsymbol{s}}}}(y) = (g_{\tilde{\boldsymbol{s}}}(m+py) - g_{\tilde{\boldsymbol{s}}}(m))/p$, and the second inner sum is only present when $p \nmid d$. By the standard single-variable complete sum estimate (see [4] for example), the first inner sum is bounded by $p^{v(1-1/k)}\gcd(\mathrm{cont}(g_{\tilde{\boldsymbol{s}}}),p^v)^{1/k}$. Further, the second inner sum is bounded by $p^{(v-1)(1-1/k)}\gcd(\mathrm{cont}(\overline{g_{\tilde{\boldsymbol{s}}}}),p^{v-1})^{1/k}$, and since this term is only present when $p\nmid d$, we know in this case that $\gcd(\mathrm{cont}(\overline{g_{\tilde{\boldsymbol{s}}}}),p^{v-1}) \ll_h p^{k-1}$. In any case, we have

$$\left|\sum_{\substack{\boldsymbol{s}\in[p^v]^\ell \\ p\nmid(r_d)_i+d\tilde{s}_i \text{ for } 1\le i\le\ell}} e(ah_d(\boldsymbol{s})/p^v)\right| \ll_h p^{v(1-1/k)} \sum_{\tilde{s}\in[p^v]^{\ell-1}} \gcd(\mathrm{cont}(g_{\tilde{\boldsymbol{s}}}),p^v)^{1/k}.$$

Suppose $a_{\boldsymbol{i}} = a_{i_1,\dots,i_\ell}$ with $0 < |\boldsymbol{i}| \le k$ is a coefficient of h_d, corresponding to $x_1^{i_1}\cdots x_\ell^{i_\ell}$, that is not divisible by p. Further, assume that $i_1 > 0$, as if $i_1 = 0$ then we could just relabel our coordinates. In this case, for each $0 \le w \le v$, $\gcd(\mathrm{cont}(g_{\tilde{\boldsymbol{s}}}),p^v) = p^w$ only if $p^w \mid s_2^{i_2}\cdots s_\ell^{i_\ell}$, so in particular $p^{\lceil w/k\rceil} \mid s_2\cdots s_\ell$, which occurs for fewer than $(w+1)^{\ell-1}p^{v(\ell-1)-w/k}$ choices of $\tilde{\boldsymbol{s}}$. In particular,

$$\begin{aligned}\sum_{\tilde{s}\in[p^v]^{\ell-1}} \gcd(\mathrm{cont}(g_{\tilde{\boldsymbol{s}}}),p^v)^{1/k} &\le \sum_{w=0}^{v}(w+1)^{\ell-1}p^{v(\ell-1)-w/k}p^{w/k} \\ &\le (v+1)^\ell p^{v(\ell-1)}.\end{aligned}$$

The resulting bound on the exponential sum modulo p^v is a constant depending on h times

$$p^{v(1-1/k)}(v+1)^\ell p^{v(\ell-1)} = (v+1)^\ell p^{v(\ell-1/k)},$$

as required. Having accounted for all prime power divisors of q, the proof is complete. □

Corollary 5.13. *If $(a,q)=1$, then*

$$|\mathcal{G}(a,q)| \ll_h \prod_{\substack{p\leq Y\\ p\nmid q}} \left(1-\frac{j_d(p)}{((p-\epsilon_d(p))p^{\gamma_d(p)-1})^{\ell}}\right)\cdot \begin{cases} C^{\omega(q)}q^{1/2} & \ell=1,\ q\leq Y \\ (q/\varphi(q))^{C}\left((k-1)^2+2\right)^{\omega(q)}q & \ell=2,\ q\leq Y \\ C^{\omega(q)}q^{\ell-3/2} & \ell\geq 3,\ q\leq Y \\ C^{\omega(q)}\tau(q)^{\ell}q^{\ell-1/k} & q>Y \end{cases},$$

where $\tau(q)=\sum_{m|q}1$ and $C=C(h)$.

5.4. **Minor arc estimate.** When α is not close to a rational with small denominator, we use the following variant of Weyl's inequality, which is a version of [17, Theorem 4.1]. For $k\in\mathbb{N}$, let $K=2^{10k}$.

Lemma 5.14 (Lemma 12, [26])**.** *Suppose $g(x)=a_0+a_1x+\cdots+a_kx^k\in\mathbb{Z}[x]$ with $a_k>0$, $X,d\in\mathbb{N}$, and $r\in\mathbb{Z}$. If $U\geq\log X$, $a_k\gg|a_{k-1}|+\cdots+|a_0|$, and $d,|r|,a_k\leq U^k$, then*

$$\sum_{\substack{x\leq X\\ r+dx\in\mathcal{P}}}\log(r+dx)e(g(x)\alpha)\ll\frac{X}{U}+U^BX^{1-c}$$

for $B=B(k)$ and $c=c(k)>0$, provided $|\alpha-a/q|<q^{-2}$ for some $U^K\leq q\leq g(X)/U^K$ and $(a,q)=1$.

Using Lemma 5.14 to exploit cancellation in only one variable, combined with the techniques of the proof of [7, Lemma 7.12] to account for the sieve, yields the following.

Corollary 5.15. *If $C\geq1$ and $|\alpha-a/q|<q^{-2}$, $(a,q)=1$, for some $\mathcal{Q}^C\leq q\leq M^k/\mathcal{Q}^C$, then*

$$|S(\alpha)|\ll_h\mathcal{Q}^{-C/K}M^{\ell}+\mathcal{Q}^BM^{\ell-c},$$

where $B=B(C,k)$ and $c=c(k)>0$.

5.5. **Proof of** (19) **and** (20) **for** $\ell\geq2$**.** Fixing $\alpha\in\mathbb{T}$, and letting $C=C(h)$ be a sufficiently large constant, the pigeonhole principle guarantees the existence of $1\leq q\leq M^k/\mathcal{Q}^C$ and $(a,q)=1$ such that

$$\left|\alpha-\frac{a}{q}\right|<\frac{\mathcal{Q}^C}{qM^k}.$$

Letting $\beta=\alpha-a/q$, if $q\leq\mathcal{Q}^C$, we have by Lemma 5.4 that

$$S(\alpha)=\left(\frac{\varphi(d)}{\varphi(qd)}\right)^{\ell}w_d^{-1}\mathcal{G}(a,q)\int_{[1,M]^{\ell}}\prod_{i=1}^{\ell}\left(1-\chi((r_d)_i)(dx)^{\rho-1}\right)e(h_d(\boldsymbol{x})\beta)d\boldsymbol{x}+O_h(E), \tag{36}$$

so in particular

$$|S(\alpha)|\ll_h\left(\frac{\varphi(d)}{\varphi(qd)}\right)^{\ell}w_d^{-1}|\mathcal{G}(a,q)|T,$$

which combines with Corollary 5.13 to yield (19) if $q\leq Q$ and $|\beta|<\gamma$ and (20) if $Q\leq q\leq\mathcal{Q}^C$. Further, assuming x_1 appears in a degree k term of h_d (if not, relabel coordinates), if $q\leq\mathcal{Q}^C$ and $|\beta|\geq\gamma$, standard van der Corput estimates (see for example Lemma 2.8 in [34]) give

$$\left|\int_1^M e(h_d(\boldsymbol{x})\beta)dx_1\right|\ll_h|J\beta|^{-1/k}\ll\eta M.$$

Integration by parts yields

$$\left|\int_1^M(1-\chi((r_d)_1)(dx_1)^{\rho-1})e(h_d(\boldsymbol{x})\beta)dx_1\right|\ll\eta\left(M-\frac{\chi((r_d)_1)(dM)^{\rho}}{d\rho}+2(1-\rho)M\right),$$

which combines with (36) to yield (20). Finally, again exploiting cancellation in only a single variable, (20) holds by Corollary 5.15 if $\mathcal{Q}^C\leq q\leq M^k/\mathcal{Q}^C$. □

The proof of (20) for $\ell=1$ is similar to the corresponding proof above, with partial summation when appropriate to account for the derivative weight. What requires some final attention, however, is the estimate (24) for the L^2 mass of $S(\alpha)$ over a full major arc when $\ell=1$.

5.6. **Proof of** (24) **for** $\ell = 1$**.** Suppose $q \leq Q$ and $\alpha = a/q + \beta$ with $(a,q) = 1$ and $|\beta| < \gamma$. By Lemma 5.5, we have

$$S(\alpha) = \frac{\varphi(d)}{\varphi(qd)} w_d^{-1} \mathcal{G}(a,q) \int_1^M h_d'(x) \left(1 - \chi(r_d)(dx)^{\rho-1}\right) e(h_d(x)\beta) dx + O_h\left(E\right).$$

Further, we see that

$$\left| \int_1^M h_d'(x) e(h_d(x)\beta) \right| = \left| \int_{h_d(1)}^{h_d(M)} e(y\beta) dy \right| \ll |\beta|^{-1}.$$

Integration by parts then yields

$$|S(\alpha)| \ll_h \frac{\varphi(d)}{\varphi(qd)} w_d^{-1} |\mathcal{G}(a,q)| T \min\{1, (L|\beta|)^{-1}\}.$$

Applying Corollary 5.13 and splitting the integral

$$\int_{|\beta|<\gamma} |S(a/q+\beta)|^2 d\beta$$

into the regions $|\beta| < L^{-1}$ and $|\beta| \geq L^{-1}$ gives the required estimate. □

Acknowledgements: The first author's research was partially supported by NSF grant DMS-2302394. The second author's research was partially supported by an AMS-Simons research grant for PUI faculty.

References

[1] N. Arala, *A maximal extension of the Bloom-Maynard bound for sets with no square differences*, preprint (2023), `arXiv:2303.03345`.

[2] A. Balog, J. Pelikán, J. Pintz, E. Szemerédi, *Difference sets without κ-th powers*, Acta. Math. Hungar. 65 (2) (1994), 165-187.

[3] T. Bloom, J. Maynard, *A new upper bound for sets with no square differences*, preprint (2020), `arxiv:2011.13266`.

[4] J.R. Chen, *On Professor Hua's estimate of exponential sums*, Sci. Sinica 20 (1977), 711-719.

[5] K. Conrad, *A multivariable Hensel's lemma*, https://kconrad.math.uconn.edu/blurbs/gradnumthy/multivarhensel.pdf

[6] P. Deligne, *La conjecture de Weil I*, Pub. Math. I.H.E.S. 43 (1974), 273-307.

[7] J. R. Doyle, A. Rice, *Multivariate polynomial values in difference sets*, Discete Analysis, 2021:11, 46pp.

[8] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math, 71 (1977), 204-256.

[9] B. Green, *On arithmetic structures in dense sets of integers*, Duke Math. Jour. 114 (2002) no.2, 215-238.

[10] B. Green, *On Sárözy's theorem for shifted primes*, J. Amer. Math. Soc.(2023), https://doi.org/10.1090/jams/1036.

[11] B. Green, T. Tao, T. Ziegler, *A Fourier-free proof of the Furstenberg-Sárközy theorem*, https://terrytao.wordpress.com/2013/02/28/a-fourier-free-proof-of-the-furstenberg-sarkozy-theorem/.

[12] R. Hall, G. Tenanbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90, Cambridge University Press, 1990.

[13] M. Hamel, N. Lyall, A. Rice, *Improved bounds on Sárközy's theorem for quadratic polynomials*, Int. Math. Res. Not. no. 8 (2013), 1761-1782

[14] T. Kamae, M. Mendès France, *van der Corput's difference theorem*, Israel J. Math. 31, no. 3-4, (1978), pp. 335-342.

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

[16] M. Lewko, *An improved lower bound related to the Sárközy-Furstenberg Theorem*, Electron. J. Combin. 22 (2015), No. 32, 1-6.

[17] H.-Z. Li, H. Pan, *Difference sets and polynomials of prime variables*, Acta. Arith. 138, no. 1 (2009), 25-52.

[18] N. Lyall, À. Magyar, *Polynomial configurations in difference sets*, J. Number Theory 129 (2009), 439-450.

[19] J. Lucier, *Intersective sets given by a polynomial*, Acta Arith. 123 (2006), 57-95.

[20] J. Lucier, *Difference sets and shifted primes*, Acta. Math. Hungar. 120 (2008), 79-102.

[21] N. Lyall, *A new proof of Sárközy's theorem*, Proc. Amer. Math. Soc. 141 (2013), 2253-2264.

[22] H. L. Montgomery, R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge Studies in Advanced Mathematics 97, 2007.

[23] J. Pintz, W. L. Steiger, E. Szemerédi, *On sets of natural numbers whose difference set contains no squares*, J. London Math. Soc. 37 (1988), 219-231.

[24] A. Rice, *A maximal extension of the best-known bounds for the Furstenberg-Sárközy Theorem*, Acta Arith. 187 (2019), 1-41.

[25] A. Rice, *Improvements and extensions of two theorems of Sárközy*, Ph.D. thesis, University of Georgia, 2012. http://alexricemath.com/wp-content/uploads/2013/06/AlexThesis.pdf.

[26] A. Rice, *Sárközy's theorem for $\mathcal{P}$-intersective polynomials*, Acta Arith. 157 (2013), no. 1, 69-89.

[27] A. Rice, *Binary quadratic forms in difference sets*, Combinatorial and Additive Number Theory III, Springer Proc. of Math. and Stat., vol. 297 (2020), 175-196.
[28] I. Ruzsa, T. Sanders, *Difference sets and the primes*, Acta. Arith. 131, no. 3 (2008), 281-301.
[29] I. Ruzsa, *Difference sets without squares*, Period. Math. Hungar. 15 (1984), 205-209.
[30] I. Ruzsa, *On measures on intersectivity*, Acta Math. Hungar. 43(3-4) (1984), 335-340.
[31] A. Sárközy, *On difference sets of sequences of integers I*, Acta. Math. Hungar. 31(1-2) (1978), 125-149.
[32] A. Sárközy, *On difference sets of sequences of integers III*, Acta. Math. Hungar. 31(3-4) (1978), 355-386.
[33] S. Slijepčević, *A polynomial Sárközy-Furstenberg theorem with upper bounds*, Acta Math. Hungar. 98 (2003), 275-280.
[34] R. C. Vaughan, *The Hardy-Littlewood method*, Cambridge University Press, Second Edition, 1997.
[35] M. Wang, *A quantitative bound on Furstenberg-Sárközy patterns with shifted prime power common differences in primes*, preprint (2021), arXiv:2102.11441.
[36] R. Wang, *On a theorem of Sárközy for difference sets and shifted primes*, Journal of Number Theory, Volume 211 (2020), 220-234.
[37] K. Younis, *Lower bounds in the polynomial Szemerédi theorem*, preprint (2019), arXiv:1908.06058.

Department of Mathematics, Oklahoma State University, Stillwater, OK 74075

Email address: john.r.doyle@okstate.edu

Department of Mathematics, Millsaps College, Jackson, MS 39210

Email address: riceaj@millsaps.edu