IMPROVED BOUNDS ON SÁRKÖZY'S THEOREM FOR QUADRATIC POLYNOMIALS

MARIAH HAMEL NEIL LYALL ALEX RICE

ABSTRACT. We extend the best known bound on the largest subset of $\{1, 2, \dots, N\}$ with no square differences to the largest possible class of quadratic polynomials.

1. INTRODUCTION

1.1. **Background and previous results.** Lovász conjectured verbally that any set of natural numbers of positive upper density[1] necessarily contains two distinct elements which differ by a perfect square. This conjecture was proven in the late 1970s independently by Sárközy and Furstenberg. Furstenberg used ergodic theory (see [3]) and obtained a purely qualitative result, proving the conjecture exactly as stated above. Sárközy, however, obtained a stronger, quantitative result by employing a Fourier analytic density increment strategy which utilized the Hardy-Littlewood circle method and was inspired by Roth's proof of the analogous conjecture for three-term arithmetic progressions.

Theorem A (Sárközy, [12])**.** *If $A \subseteq [1, N]$ and $n^2 \notin A - A$ for all $n \in \mathbb{N}$, then*

$$\frac{|A|}{N} \ll \Big(\frac{(\log\log N)^2}{\log N}\Big)^{1/3}. \tag{1}$$

Here we use $A - A$ to denote the difference set $\{a - a' : a, a' \in A\}$ and the symbol $\ll$ to denote "less than a constant times". The best known quantitative bound for the density of a subset $A \subseteq [1, N]$ with no square differences was obtained in [11] by Pintz, Steiger, and Szemerédi.

Theorem B (Pintz, Steiger, and Szemerédi, [11])**.** *If $A \subseteq [1, N]$ and $n^2 \notin A - A$ for all $n \in \mathbb{N}$, then*

$$\frac{|A|}{N} \ll (\log N)^{-c \log\log\log\log N}, \tag{2}$$

with $c = 1/12$.

A natural generalization of Sárközy's theorem is the replacement of the squares with the image of a more general integer polynomial. In [1], for example, Balog, Pelikán, Pintz, and Szemerédi modified the argument used in [11] to obtain the same bounds with squares replaced by perfect k^{th}-powers for an arbitrary fixed $k \in \mathbb{N}$. In fact, they improved the constant c in the exponent from $1/12$ to $1/4$.

However, it is not the case that an analogous result can be obtained for an arbitrary polynomial, even in a qualitative sense. Given a polynomial $f \in \mathbb{Z}[x]$, it is clearly necessary that f has a root modulo q for every $q \in \mathbb{N}$, as otherwise there would be a set of the form $q\mathbb{N}$ of positive density with no differences in the image of f. It follows from a theorem of Kamae and Mendès France in [5] that this condition is also sufficient, and in this case we say that f is an *intersective polynomial*. Equivalently, a polynomial is intersective if and only if it has a root in the p-adic integers for every prime p.

The first broad quantitative generalization of Theorem A beyond monomials was obtained in [13], in which Slijepčević showed triple logarithmic decay in the case of polynomials with an integer root. In [8], Lyall and Magyar obtained a stronger, single logarithmic bound in the integer root case as a corollary of a higher dimensional result. The best bounds for an arbitrary intersective polynomial are due to Lucier, who successfully adapted the density increment procedure by utilizing p-adic roots and allowing the polynomial to change at each step of the iteration.

2000 *Mathematics Subject Classification.* 11B30.

[1] A set $A \subseteq \mathbb{N}$ is said to have positive upper density if $\limsup_{N\to\infty} \frac{|A\cap[1,N]|}{N} > 0$, where $[1, N]$ denotes $\{1, 2, \dots, N\}$.

Theorem C (Lucier, [7]). *Suppose $f \in \mathbb{Z}[x]$ is an intersective polynomial of degree k. If $A \subseteq [1, N]$ and $f(n) \notin A - A$ for all $n \in \mathbb{N}$ with $f(n) \neq 0$, then*

$$\frac{|A|}{N} \ll \Big(\frac{(\log\log N)^{\mu}}{\log N}\Big)^{1/(k-1)}, \quad \mu = \begin{cases} 3 & \text{if } k = 2 \\ 2 & \text{if } k > 2 \end{cases},$$

where the implied constant depends only on f.

1.2. **Main result of this paper.** In this paper, we combine Lucier's modified density increment strategy with the methods of [11] and [1] in the special case of $k = 2$. We also improve the constant in the exponent from $1/4$ to $1/\log 3$, the natural limit of the method as remarked in [1], obtaining the following result.

Theorem 1. *Suppose $f \in \mathbb{Z}[x]$ is an intersective quadratic polynomial. If $A \subseteq [1, N]$ and $f(n) \notin A - A$ for all $n \in \mathbb{N}$ with $f(n) \neq 0$, then for any $\rho < 1/\log 3$,*

$$\frac{|A|}{N} \ll (\log N)^{-\rho \log\log\log\log N}, \tag{3}$$

where the implied constant depends only on f and ρ.

It is a pleasing consequence of Theorem 1 and the previous results of [11] and [1] that the primes, and even rather sparse subsets thereof, contain the desired arithmetic structure for any monomial or intersective quadratic based on density considerations alone. While the $1/\log N$ density barrier has not been broken for an arbitrary intersective polynomial, recent work of Lê in [6] uses Lucier's ideas together with Green's transference principle to show that for any intersective polynomial f, a subset of the primes of positive relative upper density is guaranteed to contain two distinct elements whose difference lies in the image of f.

1.3. **Remark on intersective quadratic polynomials.** It is worth pointing out that while the intersective condition can be somewhat mysterious and difficult to check for a general polynomial, this is not the case when restricted to degree 2.

Proposition 1. *A quadratic polynomial $f \in \mathbb{Z}[x]$ is intersective if and only if f has rational roots with coprime denominators. In other words,*

$$f(x) = a(\alpha x + \beta)(\gamma x + \lambda),\ a, \alpha, \beta, \gamma, \lambda \in \mathbb{Z},\ (\alpha, \beta) = (\gamma, \lambda) = (\alpha, \gamma) = 1.$$

It follows from Theorem 1 of [2] that an intersective polynomial with no rational roots must have degree at least five. Additionally, Proposition 1 can be directly shown rather easily by applying the quadratic formula over an appropriate field of p-adic numbers, and we include a short proof in Appendix B.1. While this characterization is not essential to the argument, it will allow us to greatly simplify some of Lucier's work. For example, we will avoid further discussion of p-adic numbers altogether.

Acknowledgement

The authors would like to acknowledge Julia Wolf, whose exposition in [14] of the sensitive and initially intimidating argument in [11] we found most helpful.

2. Preliminary Notation: The Fourier Transform and the Circle Method

In the classical case, if $A \subseteq [1, N]$ and $n^2 \in A - A$ with $n \in \mathbb{N}$, then it is clear that $1 \leq n^2 \leq N$. For technical reasons, it is useful to restrict to a slightly smaller collection of squares, for example those less than $N/2$. The same is true when generalizing to a polynomial, which by symmetry of difference sets we can assume has positive leading term, leading to the following definition.

Definition 1. Given a nonconstant polynomial $f \in \mathbb{Z}[x]$ with positive leading term and $N \in \mathbb{N}$, we let

$$j = j(f) = \max\{x \in \mathbb{N} : f(x) \leq 0\},$$

taking $j = 0$ if the polynomial is strictly positive on $\mathbb{N}$, we let

$$M = M(f, N) = \min\{x \in \mathbb{N} : f(x) \geq N/3\},$$

and we define

$$S_f = S_f(N) = \{f(x) : x \in \mathbb{N}, j < x < M\}. \tag{4}$$

We identify subsets of the interval $[1, N]$ with subsets of the finite group $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$, on which we utilize the normalized discrete Fourier transform. Specifically, for a function $F : \mathbb{Z}_N \to \mathbb{C}$, we define $\widehat{F} : \mathbb{Z}_N \to \mathbb{C}$ by

$$\widehat{F}(t) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} F(x) e^{-2\pi i x t/N}.$$

We analyze the behavior of the Fourier transform using the Hardy-Littlewood circle method, decomposing the nonzero frequencies into two pieces: the roots of unity which are close to rationals with small denominator, and those which are not.

Definition 2. Given $N \in \mathbb{N}$, $\delta > 0$, $f \in \mathbb{Z}[x]$, and a parameter K, we let $\eta = c_0\delta$, where c_0 is an appropriately small constant depending only on f, we let M be as in Definition 1, and we define for each $q \in \mathbb{N}$ and $a \in [1, q]$

$$\mathbf{M}_{a,q}(K) = \left\{ t \in \mathbb{Z}_N \setminus \{0\} : \left| \frac{t}{N} - \frac{a}{q} \right| < \frac{K^2}{\eta^2 M^2} \right\} \quad \text{and} \quad \mathbf{M}_q(K) = \bigcup_{a \in [1,q]} \mathbf{M}_{a,q}(K).$$

We then define $\mathfrak{M}$, the *major arcs*, by

$$\mathfrak{M}(K) = \bigcup_{q=1}^{K^2\eta^{-2}} \mathbf{M}_q(K),$$

and $\mathfrak{m}(K)$, the *minor arcs*, by $\mathfrak{m}(K) = \mathbb{Z}_N \setminus (\mathfrak{M}(K) \cup \{0\})$. It is important to note that as long as $2K^3\eta^{-3} < M$, we have that $a/q \neq b/r$ implies $\mathbf{M}_{a,q} \cap \mathbf{M}_{b,r} = \emptyset$ whenever $q, r \leq K^2\eta^{-2}$.

Remark on notation. We note that the objects defined above certainly depend on N, δ, and f, despite their absence from the notation. These should always be replaced with the size of the appropriate ambient group, the density of the appropriate subset, and the polynomial in question, respectively. Also, the absence of the parameter K indicates $K = 1$. When considering a set A, we let $A(x)$ denote the characteristic function of A. We will use the letters C and c to denote appropriately large or small constants, which can change from line to line. We will occasionally use subscripts to indicate what these constants depend on, but they will always be independent of the main parameter N.

3. Overview of the argument

The underlying philosophy of many theorems of this flavor is that if a set of integers exhibits certain non-random phenomena, then these behaviors should be detectable in the Fourier transform of the characteristic function of the set. That information about the transform can then be used to obtain some structural information about the set, and eventually provide an upper bound on its size.

Specifically, in the case that $(A - A) \cap S_f = \emptyset$ for a set $A \subseteq [1, N]$ and a polynomial $f \in \mathbb{Z}[x]$ with positive leading term, this unexpected behavior leads to substantial L^2 mass of $\widehat{A}$ away from zero. After applying the Hardy-Littlewood circle method and decomposing the frequency space into major and minor arcs, one concludes that in fact there is substantial L^2 mass concentrated around rationals with small denominators. At this point, there are a number of possible paths one could to take in order to obtain the desired structural information.

The original method of Sárközy in [12], as well as that of Lucier in [7] and Lyall and Magyar in [8], is to use the pigeonhole principle to conclude that there is one single denominator q such that $\widehat{A}$ has L^2 concentration around rationals with denominator q. From this information, one can conclude that A has increased density on a long arithmetic progression with step size an appropriate multiple of q, for example q^2 in the classical case. By translating and scaling the intersection of A with this progression, one obtains a new subset A' of a slightly smaller interval with significantly greater density. In addition, if f is an intersective polynomial, A' inherits non-random behavior from the fact that $(A - A) \cap S_f = \emptyset$. In the case that f is a monomial, A' actually inherits the identical property, but more generally one sees that $(A' - A') \cap S_h = \emptyset$ for a potentially different intersective polynomial h obtained from f. One then shows that if the density of the original set A was too large, then this process could be iterated enough times for the density to surpass 1, hence obtaining a contradiction.

In [11], Pintz, Steiger, and Szemerédi first observed that pigeonholing to obtain a single denominator q is a potentially wasteful step. We follow their approach, specifically observing the following dichotomy:

> **Case 1.** There is a single denominator q such that $\widehat{A}$ has extremely high L^2 concentration, greater than yielded by the pigeonhole principle, around rationals with denominator q. This leads to a large density increment on a long arithmetic progression.
>
> **Case 2.** The L^2 mass of $\widehat{A}$ on the major arcs is spread over many denominators. In this case, an iteration procedure using the "combinatorics of rational numbers" is employed to build a large collection of frequencies at which $\widehat{A}$ is large, then Plancherel's identity is applied to obtain the bound on the density claimed in Theorem 1.

Philosophically, Case 1 provides more structural information about the original set A than Case 2 does, and could potentially lead to an even stronger result. The downside is that the density increment procedure yields a new set and potentially a new polynomial, while the iteration in Case 2 leaves these objects fixed. With theses cases in mind, we can now outline the argument, separated into two distinct phases.

> **Phase 1** (The Outer Iteration): Given a set A and an intersective quadratic polynomial f with $(A-A)\cap S_f = \emptyset$, we ask if the set falls into Case 1 or Case 2 described above. If it falls into Case 2, then we proceed to Phase 2.
>
> If it falls into Case 1, then the density increment procedure yields a new subset A_1 of a slightly smaller interval with significantly greater density, and an intersective quadratic f_1 with slightly larger coefficients and $(A_1 - A_1)\cap S_{f_1} = \emptyset$. We can then iterate this process as long as the resulting interval isn't too small and the coefficients of the resulting polynomial aren't too large, asking at each step if our set falls into Case 1 or Case 2.
>
> One can show that if the resulting sets never fall into Case 2, and the process iterates until the interval shrinks down or the coefficients grow to the limit, then the density of the original set A must have satisfied a bound stronger than the one purported in Theorem 1.

Contrapositively, we simply assume that the original density does not satisfy this stricter bound, and we conclude that one of the sets yielded by the density increment procedure must have fallen into Case 2. We call that set B and the corresponding polynomial h.

> **Phase 2** (The Inner Iteration): We now have a set B and an intersective quadratic h with $(B-B)\cap S_h = \emptyset$ which fall into Case 2, so we can adapt the strategy of [11] and [1].
>
> We use the usual approach to show that $(B-B)\cap S_h = \emptyset$ implies significant L^2 mass of $\widehat{B}$ on the major arcs, which we use to identify an initial set P of frequencies at which $\widehat{B}$ is large. Then, we prove that if you have a frequency close to a rational a/q at which $\widehat{B}$ is large, then there are lots of rationals b/r and frequencies close to $a/q + b/r$ at which $\widehat{B}$ is almost as large. This intuitively indicates that a set P of frequencies associated with large Fourier coefficients can be blown up to a much larger set P' of frequencies associated with nearly as large Fourier coefficients.
>
> The only obstruction to this intuition is the possibility that there are many pairs $(a/q, b/r)$ and $(a'/q', b'/r')$ with $a/q + b/r = a'/q' + b'/r'$. Observations made in [11] and [1] on the combinatorics of rational numbers demonstrate that this potentially harmful phenomenon can not occur terribly often.
>
> This process is applied as long as certain parameters are not too large, and the number of iterations is ultimately limited by the growth of the divisor function. Once the iteration is exhausted, we use the resulting set of large Fourier coefficients and Plancherel's Identity to get the upper bound on the density of B, which is by construction larger than the density of the original set A, claimed in Theorem 1.

4. Reduction of Theorem 1 to Three Lemmas

To make the strategy outlined in Section 3 precise, we fix an intersective quadratic f and a set $A \subseteq [1, N]$ with $|A| = \delta N$ and $f(n) \notin A - A$ for all $n \in \mathbb{N}$ with $f(n) \neq 0$. By the symmetry of difference sets, we can assume without loss of generality that f has positive leading term, and we see in particular that $(A - A) \cap S_f = \emptyset$.

We also fix an arbitrary $\epsilon > 0$ and set $Q = (\log N)^{\epsilon \log\log\log N}$, and we will prove Theorem 1 with $\rho = (1 - 20\epsilon)/\log 3$. At any point we are free to insist that N is sufficiently large with respect to f and ϵ, as this will only affect the implied constant in (3), so for convenience we will take these to be perpetual implicit hypotheses and abstain from including them further. From this point on, we will allow all of our constants to depend on f and ϵ, even if this is not specifically indicated with subscripts.

4.1. **Three Key Lemmas.** We first reduce Theorem 1 to three key lemmas, the first of which corresponds to Phase 1 outlined above, and the last two of which correspond to Phase 2.

Lemma 1. *If*

$$\delta \geq e^{-(\log N)^{\epsilon/2}/1000}, \tag{5}$$

then there exists a set $B \subseteq [1, N']$ satisfying $N' \geq N^{.99}$, $|B|/N' = \sigma \geq \delta$, and $(B - B) \cap S_h(N') = \emptyset$, where

$$h(x) = \begin{cases} ax^2, & \text{if } f(x) = a(x-b)^2 \\ f(r + dx)/d, & \text{else} \end{cases},$$

$d \leq N^{.01}$, and $r \in (-d, 0]$ is a root of f modulo d. Further, the set B satisfies

$$\max_{q \leq Q} \sum_{t \in \mathbf{M}_q(Q)} |\widehat{B}(t)|^2 \leq \sigma^2 (\log N)^{-1+\epsilon}. \tag{6}$$

We note that $e^{-(\log N)^{\epsilon/2}/1000} \leq C_K (\log N)^{-K \log\log\log\log N}$ for any K. In particular, if hypothesis (5) is not satisfied, then Theorem 1 is already more than true. For the set B produced by Lemma 1, it must be the case that either $|B \cap [1, N'/2]| \geq \sigma N'/2$ or $|B \cap [N'/2, N']| \geq \sigma N'/2$. We assume without loss of generality that the former holds and set $B_1 = B \cap [1, N'/2]$.

The next lemma corresponds to the identification of an initial collection of large Fourier coefficients discussed at the beginning of Phase 2.

Lemma 2. *If B and h are as in the conclusion of Lemma 1, then there exist natural numbers $U \ll \sigma^{-5/2}$ and $V \ll \sigma^{-2}$ and a set of large frequencies*

$$P \subseteq \left\{ t \in \bigcup_{q=V/2}^{V} \bigcup_{(a,q)=1} \mathbf{M}_{a,q} : \frac{\sigma}{U} \leq \min\left\{ |\widehat{B}(t)|, |\widehat{B_1}(t)| \right\} \leq \frac{2\sigma}{U} \right\}$$

satisfying

$$\frac{|P|}{U^2} \gg \frac{V^{1/2}}{(\log(\sigma^{-1}))^3},$$

and

$$|P \cap \mathbf{M}_{a,q}| \leq 1 \quad \text{for all} \quad \mathbf{M}_{a,q} \subseteq \mathfrak{M}.$$

Lemma 2 provides a starting point for the iteration scheme described in the following lemma, in which a set of large Fourier coefficients from distinct major arcs is blown up in such a way that the relative growth of the size of the set is much greater than the relative loss of pointwise mass.

Lemma 3. *Suppose B and h are as in the conclusion of Lemma 1, and suppose $\sigma \geq Q^{-1/6}$. Given $U, V, K \in \mathbb{N}$ with $\max\{U, V, K\} \leq Q^{1/6}$ and a set*

$$P \subseteq \left\{ t \in \bigcup_{q=1}^{V} \mathbf{M}_q(K) : |\widehat{B_1}(t)| \geq \frac{\sigma}{U} \right\}$$

satisfying

$$|P \cap \mathbf{M}_{a,q}(K)| \leq 1 \quad \textit{for all} \quad \mathbf{M}_{a,q}(K) \subseteq \mathfrak{M}(K), \tag{7}$$

there exist $U', V', K' \in \mathbb{N}$ *with* $\max\{U', V', K'\} \ll (\max\{U, V, K\})^3 \sigma^{-3}$ *and a set of large frequencies*

$$P' \subseteq \left\{ t \in \bigcup_{q=1}^{V'} \mathbf{M}_q(K') : |\widehat{B_1}(t)| \geq \frac{\sigma}{U'} \right\} \tag{8}$$

satisfying

$$|P' \cap \mathbf{M}_{a,q}(K')| \leq 1 \quad \textit{for all} \quad \mathbf{M}_{a,q}(K') \subseteq \mathfrak{M}(K'), \tag{9}$$

and

$$\frac{|P'|}{(U')^2} \geq \frac{|P|}{U^2} (\log N)^{1-18\epsilon}. \tag{10}$$

4.2. **Proof that Lemmas 1, 2, and 3 imply Theorem 1.** In order to prove Theorem 1, we can assume that

$$\delta \geq (\log N)^{-\log\log\log\log N}.$$

Therefore, Lemma 1 produces a set B of density $\sigma \geq \delta$ with the stipulated properties, and Lemma 2 produces an initial set of frequencies P_0 with parameters U_0, V_0, K_0 such that

$$\max\{U_0, V_0, K_0\} \leq (\log N)^{3 \log\log\log\log N}$$

and

$$\frac{|P_0|}{U_0^2} \gg (\log\log\log\log N \log\log N)^{-3}.$$

Lemma 3 then yields, for each n, a set P_n with parameters U_n, V_n, K_n such that

$$\max\{U_n, V_n, K_n\} \leq (\log N)^{3^{n+2} \log\log\log\log N}$$

and

$$\frac{1}{\sigma} \geq \frac{|P_n|}{U_n^2} \geq (\log N)^{n(1-19\epsilon)},$$

where the left-hand inequality comes from Plancherel's Identity, as long as $\max\{U_n, V_n, K_n\} \leq Q^{1/6}$. This holds with $n = (1-\epsilon)\log\log\log\log N / \log 3$, and Theorem 1 follows. □

5. The Outer Iteration: Proof of Lemma 1

Recall that we have fixed $A \subseteq [1, N]$, an intersective quadratic $f \in \mathbb{Z}[x]$, and $\epsilon > 0$. As previously mentioned, we will apply the modified density increment strategy described in [7], which allows for the polynomial to change at each stage of the iteration. The following definition completely describes all of the polynomials that we could potentially encounter.

Definition 3. For each $d \in \mathbb{N}$, we fix an integer $r_d \in (-d, 0]$ such that $f(r_d) \equiv 0 \bmod d$ and $r_d \equiv r_s \bmod s$ whenever $s \mid d$, and we define the *auxiliary polynomials* $f_d \in \mathbb{Z}[x]$ by

$$f_d(x) = \begin{cases} ax^2, & \text{if } f(x) = a(x-b)^2 \\ f(r_d + dx)/d, & \text{else} \end{cases}.$$

One can find a collection of roots with the property stipulated in Definition 3 in the following way. If $f(x) = a(\alpha x + \beta)(\gamma x + \lambda)$ with $(\alpha, \gamma) = 1$, partition the primes into $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$, with $p \nmid \alpha$ for all $p \in \mathcal{P}_1$ and $p \nmid \gamma$ for all $p \in \mathcal{P}_2$.

For each $d \in \mathbb{N}$, write $d = p_1^{a_1} \cdots p_k^{a_k} s_1^{j_1} \cdots s_\ell^{j_\ell}$ with $p_i \in \mathcal{P}_1$ and $s_i \in \mathcal{P}_2$. By the Chinese Remainder Theorem, there is a unique integer $r_d \in (-d, 0]$ such that $r_d \equiv -\beta\alpha^{-1} \mod p_i^{a_i}$ and $r_d \equiv -\lambda\gamma^{-1} \mod s_n^{j_n}$ for all $1 \leq i \leq k$ and $1 \leq n \leq \ell$, and we see that this choice of r_d meets the purported condition.

One can easily show from the characterization in Proposition 1 and the construction of the roots r_d that each of these auxiliary polynomials are themselves intersective quadratics. It is of potential concern to us that the coefficients of the polynomials could grow out of control with d. In particular, the required exponential

sum estimates would be irreparably damaged if the coefficients of the auxiliary polynomials shared larger and larger common factors. To address this issue, we define for a polynomial $g(x) = a_o + a_1 x + \cdots + a_k x^k$,

$$\text{cont}(g) = \gcd(a_1, \ldots, a_k). \tag{11}$$

In the usual sense of content, this would be the content of $g(x) - g(0)$, although in the case of an intersective polynomial it agrees with the content of g itself. The following proposition assures us that the scenario we fear does not occur.

Proposition 2 (Lemma 28 in [7]). *If $f(x) = (\alpha x + \beta)(\gamma x + \lambda)$ with $\alpha, \beta, \gamma, \lambda \in \mathbb{Z}$ and f does not have a double root, then for any $d \in \mathbb{N}$,*

$$\text{cont}(f_d) \leq |\alpha\lambda - \beta\gamma|.$$

We note that in the case $f(x) = a(x-b)^2$ excluded by the hypotheses, we trivially have $\text{cont}(f_d) = a$ for all d. Proposition 2 is simply a special case of Lemma 28 of [7], in which Lucier shows an analogous, highly nontrivial result for a general intersective polynomial which may or may not have rational roots. Again, the degree 2 case is considerably simpler, and we include an elementary proof of Proposition 2 in Appendix B.2.

This observation allows us to define, for each polynomial f_d, the constant $c_0 = c_0(f_d)$ from Definition 2 in terms of only the original polynomial f. Namely, we fix

$$c_0 = \begin{cases} c/a, & \text{if } f(x) = a(x-b)^2 \\ c/|\alpha\lambda - \beta\gamma|, & \text{if } f(x) = (\alpha x + \beta)(\gamma x + \lambda) \neq a(x-b)^2 \end{cases}, \tag{12}$$

where c is an appropriately small absolute constant.

We now invoke the usual density increment lemma which states that L^2-concentration of the Fourier transform leads to increased density on a long progression, with the added observation that the resulting subset of a smaller interval contains no differences in the image of an appropriate auxiliary polynomial. The particular statement below is a special case of Lemma 20 in [7], while the additional observation is made in Lemma 31 of the same paper.

Lemma 4 (Density Increment). *Suppose $A_k \subseteq [1, N_k]$ with $|A_k| = \delta_k N_k$, and let $\eta_k = c_0 \delta_k$ as usual. If*

$$\delta_k \geq \frac{(2\pi)^{1/6} Q}{c_0 (N_k)^{1/6}} \tag{13}$$

and

$$\sum_{t \in \mathbf{M}_q(Q)} |\widehat{A}_k(t)|^2 \geq \delta_k^2 (\log N)^{-1+\epsilon} \tag{14}$$

for some $q \leq \eta_k^{-2} Q^2$, then there exists an arithmetic progression

$$\Lambda = \{x + \lambda(q)\ell : 1 \leq \ell \leq N_{k+1}\} \subseteq [1, N_k],$$

where

$$\lambda(q) = \begin{cases} q^2, & \text{if } f(x) = a(x-b)^2 \\ q, & \text{else} \end{cases},$$

satisfying

$$N_{k+1} \geq c \delta_k^7 Q^{-6} (\log N)^{-1+\epsilon} N_k$$

and

$$|A_k \cap \Lambda| / N_{k+1} = \delta_{k+1} \geq \delta_k + \delta_k / (8 (\log N)^{1-\epsilon}). \tag{15}$$

Further, if $(A_k - A_k) \cap S_{f_{d_k}} = \emptyset$, then the set $A_{k+1} \subseteq [1, N_{k+1}]$ defined by

$$A_k \cap \Lambda = \{x + \lambda(q)\ell : \ell \in A_{k+1}\}$$

satisfies $|A_{k+1}| = \delta_{k+1} N_{k+1}$ and

$$(A_{k+1} - A_{k+1}) \cap S_{f_{d_{k+1}}} = \emptyset,$$

where $d_{k+1} = q d_k$.

5.1. **Proof of Lemma 1.** Setting A_0=A, and therefore $\delta_0 = \delta$, $N_0 = N$, and $d_0 = 1$, we input this initial set into Lemma 4 and allow it to iterate as long as conditions (13) and (14) are satisfied.

We see that by (5) and (15), the density δ_k will exceed 1 after

$$16\log(\delta^{-1})(\log N)^{1-\epsilon} \le (\log N)^{1-\epsilon/2}$$

steps, so one of those conditions must fail for some $k \le (\log N)^{1-\epsilon/2}$.

However, (5) guarantees that

$$N_{k+1} \ge \frac{c_f e^{-(\log N)^{\epsilon/2}/140}}{Q^6 \log N} N_k \ge e^{-(\log N)^{\epsilon/2}/100} N_k,$$

so for any $k \le (\log N)^{1-\epsilon/2}$, we have

$$N_k \ge (e^{-(\log N)^{\epsilon/2}/100})^{(\log N)^{1-\epsilon/2}} N = N^{.99}.$$

Therefore, since $\delta_k \ge \delta$ for all k, we see that (5) guarantees that (13) is comfortably satisfied for all $k \le (\log N)^{1-\epsilon/2}$. We can now conclude that (14) must fail for some $k \le (\log N)^{1-\epsilon/2}$, so we set $B = A_k$, $N' = N_k$, $\sigma = \delta_k$, and $h = f_{d_k}$, and we need only show the purported bound on d_k. To this end, we see that (5) implies

$$d_k \le (C_f e^{(\log N)^{\epsilon/2}/500} Q^2)^{(\log N)^{1-\epsilon/2}} \le (e^{(\log N)^{\epsilon/2}/100})^{(\log N)^{1-\epsilon/2}} = N^{.01},$$

as required. □

6. The Inner Iteration: Proof of Lemmas 2 and 3

Let B and h be as in the conclusion of Lemma 1.

6.1. **Proof of Lemma 2.** Recall that $S_h = S_h(N') = \{h(x) : x \in \mathbb{N}, j < x < M\}$, where j and M are defined in Definition 1. One can see from the definition of the polynomial h, the lower bound on N', and the upper bound on the integer d that $j \le C_f$ while $M \ge c_f N^{.49}$. In particular, we have for large N that $j < M/2$.

Since $(B - B) \cap S_h = \emptyset$ and $S_h \subset [1, N'/2)$, we see that there are no solutions to

$$x - y \equiv z \pmod{N'}, \quad x \in B, y \in B_1, z \in S_h.$$

This implies

$$\frac{1}{N'M^2} \sum_{\substack{x \in \mathbb{Z}_{N'} \\ j \le y \le M}} B(x)B_1(x - h(y))y = \sum_{t \in \mathbb{Z}_{N'}} \widehat{B}(t)\overline{\widehat{B_1}(t)}W(t) = 0,$$

where

$$W(t) = \frac{1}{M^2} \sum_{x=j}^{M} x e^{2\pi i h(x) t/N}.$$

This immediately yields

$$\sum_{t \in \mathbb{Z}_{N'} \setminus \{0\}} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ge \Big| \sum_{t \in \mathbb{Z}_{N'} \setminus \{0\}} \widehat{B}(t)\overline{\widehat{B_1}(t)}W(t) \Big| = \widehat{B}(0)\widehat{B_1}(0)W(0) \ge \sigma^2/8. \tag{16}$$

since $W(0) \ge 1/4$. It follows from traditional Weyl sum estimates, Lemma 11 in [7], and Proposition 2 that for some absolute constant C,

$$|W(t)| \le C(\mathrm{cont}(h))^{1/2}\eta \le \sigma/16 \quad \text{for all } t \in \mathfrak{m}, \tag{17}$$

provided we chose an appropriately small absolute constant when defining $c_0 = \eta/\sigma$ in (12), and

$$|W(t)| \le C(\mathrm{cont}(h)/q)^{1/2} \min\{1, (N'|t/N' - a/q|)^{-1}\} \tag{18}$$

$$\le C_f\, q^{-1/2} \min\{1, (N'|t/N' - a/q|)^{-1}\} \tag{19}$$

provided $t \in \mathbf{M}_{a/q} \subseteq \mathfrak{M}$ and $(a, q) = 1$. We discuss estimates (17) and (18) in more detail in Appendix A.

We have by (17), Cauchy-Schwarz, Plancherel's Identity that

$$\sum_{t\in\mathfrak{m}} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \le \sigma^2/16,$$

which together with (16) yields

$$\sum_{t\in\mathfrak{M}} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ge \sigma^2/16. \tag{20}$$

We now wish to assert that we can ignore those frequencies in the major arcs at which the transform of B or B_1 is particularly small. In order to make this precise, we first need to invoke a weighted version of a well-known estimate on the higher moments of Weyl sums. Specifically, we have that

$$\sum_{t\in\mathbb{Z}_{N'}} |W(t)|^6 \le C_f, \tag{21}$$

which can be seen by adapting the proof of Proposition 3.3 in [9] and applying Proposition 2, and we provide a proof in Appendix A. Choosing a constant $0 < c_1 < (64C_f^{1/6})^{-3}$, where C_f comes from (21), we define

$$X := \Big\{t \in \mathfrak{M} : \min\Big\{|\widehat{B}(t)|, |\widehat{B_1}(t)|\Big\} \le c_1\sigma^{7/2}\Big\} \quad \text{and} \quad Y := \mathfrak{M} \setminus X. \tag{22}$$

Using Hölder's Inequality to exploit the sixth moment estimate on W, followed by Plancherel's Identity, we see that

$$\begin{aligned}
\sum_{t\in X} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| &\le \Big(\sum_{t\in X} |\widehat{B}(t)|^{6/5}|\widehat{B_1}(t)|^{6/5}\Big)^{5/6}\Big(\sum_{t\in\mathbb{Z}_{N'}} |W(t)|^6\Big)^{1/6} \\
&\le c_1^{1/3}\sigma^{7/6}\Big(\sum_{t\in\mathbb{Z}_{N'}} \max\Big\{|\widehat{B}(t)|^2, |\widehat{B_1}(t)|^2\Big\}\Big)^{5/6}\cdot C_f^{1/6} \\
&\le \frac{\sigma^{7/6}}{64}\Big(\sum_{t\in\mathbb{Z}_{N'}} |\widehat{B}(t)|^2 + |\widehat{B_1}(t)|^2\Big)^{5/6} \le \sigma^2/32,
\end{aligned}$$

and hence by (20) we have

$$\sum_{t\in Y} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ge \sigma^2/32. \tag{23}$$

For $i, j \in \mathbb{N}$, we define

$$\mathcal{R}_{i,j} = \Big\{a/q : 2^i - 1 \le q \le 2^i,\ (a,q) = 1,\ \frac{\sigma}{2^j} \le \max_{t\in\mathbf{M}_{a/q}} \Big\{\min\Big\{|\widehat{B}(t)|, |\widehat{B_1}(t)|\Big\}\Big\} \le \frac{\sigma}{2^{j-1}}\Big\}.$$

We see that we have

$$\sum_{a/q\in\mathcal{R}_{i,j}} \sum_{t\in\mathbf{M}_{a/q}} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ll |\mathcal{R}_{i,j}|\frac{\sigma^2}{2^{2j}} \max_{a/q\in\mathcal{R}_{i,j}} \sum_{t\in\mathbf{M}_{a/q}} |W(t)|. \tag{24}$$

It follows from (18) that

$$\sum_{t\in\mathbf{M}_{a/q}} |W(t)| \ll q^{-1/2}\log(\sigma^{-1}),$$

hence by (24) we have

$$\sum_{a/q\in\mathcal{R}_{i,j}} \sum_{t\in\mathbf{M}_{a/q}} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ll |\mathcal{R}_{i,j}|\frac{\sigma^2}{2^{2j}}2^{-i/2}\log(\sigma^{-1}). \tag{25}$$

By our definitions, the sets $\mathcal{R}_{i,j}$ exhaust Y by taking $1 \le 2^i \le (c_0\sigma)^{-2}$ and $1 \le 2^j \le c_1^{-1}\sigma^{-5/2}$, a total search space of size $\ll (\log(\sigma^{-1}))^2$. Therefore, by (23) and (25) there exist $i, j \ll \log(\sigma^{-1})$ such that

$$\frac{\sigma^2}{(\log(\sigma^{-1}))^2} \ll |\mathcal{R}_{i,j}|\frac{\sigma^2}{2^{2j}}2^{-i/2}\log(\sigma^{-1}).$$

Setting $V = 2^i$, $U = 2^j$, and forming P by taking one element from each of the pairwise disjoint major arcs specified by $\mathcal{R}_{i,j}$, the result follows. □

6.2. **Proof of Lemma 3.** Suppose $\sigma \geq Q^{-1/6}$ and we have a set P with parameters U, V, K as specified in Lemma 3, and fix an element $s \in P$ with corresponding rational a/q.

We will begin by obtaining a fact similar to (16), but with a modulation in the original sum on the space side. Specifically, because $(B-B) \cap S_h = \emptyset$, we have

$$\frac{1}{N'M^2} \sum_{\substack{x \in \mathbb{Z}_{N'} \\ j \leq y \leq M}} B(x) B_1(x-h(y)) e^{2\pi i (x-h(y))s/N} y = \sum_{t \in \mathbb{Z}_{N'}} \widehat{B}(t) \overline{\widehat{B_1}(s+t)} W(t) = 0,$$

which immediately implies

$$\sum_{t \in \mathbb{Z}_{N'} \setminus \{0\}} |\widehat{B}(t)||\widehat{B_1}(s+t)||W(t)| \geq \Big| \sum_{t \in \mathbb{Z}_{N'} \setminus \{0\}} \widehat{B}(t) \overline{\widehat{B_1}(s+t)} W(t) \Big| = \widehat{B}(0)\widehat{B_1}(s) W(0) \geq \sigma^2/(4U), \tag{26}$$

since $|\widehat{B_1}(s)| \geq \sigma/U$. Analogous to (17) and (18), we have from traditional estimates that

$$|W(t)| \leq \sigma/(8U) \text{ for all } t \in \mathfrak{m}(U) \tag{27}$$

and

$$|W(t)| \leq C_f r^{-1/2} \min\left\{1, (N'|t/N' - b/r|)^{-1}\right\} \quad \text{if} \quad t \in \mathbf{M}_{b,r}(U) \subseteq \mathfrak{M}(U),\ (b,r) = 1. \tag{28}$$

By (27), Cauchy-Schwarz, and Plancherel, we have

$$\sum_{t \in \mathfrak{m}(U)} |\widehat{B}(t)||\widehat{B_1}(s+t)||W(t)| \leq \sigma^2/(8U),$$

which together with (26) yields

$$\sum_{t \in \mathfrak{M}(U)} |\widehat{B}(t)||\widehat{B_1}(s+t)||W(t)| \geq \sigma^2/(8U). \tag{29}$$

This time around, we would like to ignore the frequencies at which either $|\widehat{B}(t)|$ or $|\widehat{B_1}(s+t)|$ is small, so choosing a constant $0 < c_1 < (32C_f^{1/6})^{-3}$, where C_f comes from (21), we define

$$X(U) := \left\{ t \in \mathfrak{M}(U) : \min\left\{ |\widehat{B}(t)|, |\widehat{B_1}(s+t)| \right\} \leq \frac{c_1 \sigma^{7/2}}{U^3} \right\} \quad \text{and} \quad Y(U) := \mathfrak{M}(U) \setminus X(U).$$

By (21), Hölder's Inequality, and Plancherel, we see

$$\begin{aligned}
\sum_{t \in X(U)} |\widehat{B}(t)||\widehat{B_1}(s+t)||W(t)| &\leq \Big(\sum_{t \in X(U)} |\widehat{B}(t)|^{6/5} |\widehat{B_1}(s+t)|^{6/5} \Big)^{5/6} \Big(\sum_{t \in \mathbb{Z}_{N'}} |W(t)|^6 \Big)^{1/6} \\
&\leq \frac{c_1^{1/3} \sigma^{7/6}}{U} \Big(\sum_{t \in \mathbb{Z}_{N'}} \max\left\{ |\widehat{B}(t)|^2, |\widehat{B_1}(s+t)|^2 \right\} \Big)^{5/6} \cdot C_f^{1/6} \\
&\leq \frac{\sigma^{7/6}}{32U} \Big(\sum_{t \in \mathbb{Z}_{N'}} |\widehat{B}(t)|^2 + |\widehat{B_1}(s+t)|^2 \Big)^{5/6} \\
&\leq \sigma^2/(16U),
\end{aligned}$$

which by (29) yields

$$\sum_{t \in Y(U)} |\widehat{B}(t)||\widehat{B_1}(s+t)||W(t)| \geq \sigma^2/(16U). \tag{30}$$

For $i, j, k \in \mathbb{N}$, we define

$$\mathcal{R}_{i,j,k} = \left\{ b/r : 2^i - 1 \leq r \leq 2^{i+1},\ \frac{\sigma}{2^j} \leq \max_{t \in \mathbf{M}_{b/r}} |\widehat{B}(t)| \leq \frac{\sigma}{2^{j-1}},\ \frac{\sigma}{2^k} \leq \max_{t \in \mathbf{M}_{b/r}} |\widehat{B_1}(s+t)| \leq \frac{\sigma}{2^{k-1}} \right\}.$$

Analogous to (25), we have by (28) and (27) that

$$\sum_{b/r \in \mathcal{R}_{i,j,k}} \sum_{t \in \mathbf{M}_{b/r}(U)} |\widehat{B}(t)||\widehat{B_1}(t)||W(t)| \ll |\mathcal{R}_{i,j,k}| \frac{\sigma^2}{2^j 2^k} 2^{-i/2} \left(\log(\sigma^{-1}) + \log U \right). \tag{31}$$

By our definitions, the sets $\mathcal{R}_{i,j,k}$ exhaust $Y(U)$ by taking $1 \le 2^i \le (c_0\sigma)^{-2}U^2$ and $1 \le 2^j, 2^k \le c_1^{-1}\sigma^{-5/2}U^3$, a total search space of $\ll (\log(\sigma^{-1})+\log U)^3$. Therefore, by (30) and (31) there exists $i,j,k \ll \log(\sigma^{-1})+\log U$ such that

$$\frac{\sigma^2}{U(\log(\sigma^{-1})+\log U)^3} \ll |\mathcal{R}_{i,j,k}|\frac{\sigma^2}{2^j 2^k}2^{-i/2}\log(\sigma^{-1}).$$

In other words, we can set $V_s = 2^i$, $W_s = 2^j$, and $U_s = 2^k$ and take one element from each of the major arcs specified by $\mathcal{R}_{i,j,k}$ to form a set

$$P_s \subseteq \left\{ t \in \bigcup_{r=V_s/2}^{V_s} \bigcup_{(b,r)=1} \mathbf{M}_{b,r}(U) : \frac{\sigma}{W_s} \le |\widehat{B}(t)| \le \frac{2\sigma}{W_s},\ \frac{\sigma}{U_s} \le |\widehat{B}(s+t)| \le \frac{2\sigma}{U_s} \right\}$$

which satisfies

$$|P_s| \gg \frac{U_s W_s V_s^{1/2}}{U(\log(\sigma^{-1})+\log U)^4}$$

and

$$|P_s \cap \mathbf{M}_{b,r}(U)| \le 1 \ \text{ for all } \ \mathbf{M}_{b,r}(U) \subseteq \mathfrak{M}(U). \tag{32}$$

We now note that there will be some subset $\tilde{P} \subseteq P$ with

$$|\tilde{P}| \gg |P|/(\log(\sigma^{-1})+\log U)^3 \tag{33}$$

for which the triple U_s, W_s, V_s is the same. We call those common parameters $\tilde{U}, \tilde{W}$ and $\tilde{V}$, respectively, and we can now foreshadow by asserting that the claimed parameters in the conclusion of Lemma 3 will be $U' = \tilde{U}$, $V' = \tilde{V}V$, and $K' = K + U$, which do satisfy the purported bound.

Let

$$\mathcal{R} = \left\{ \frac{a}{q} + \frac{b}{r} : \frac{a}{q} \text{ corresponds to } s \in \tilde{P},\ \frac{b}{r} \text{ corresponds to } t \in P_s \right\}.$$

By taking one frequency $s+t$ from each of the major arcs specified by $\mathcal{R}$, we can form our set P', which immediately satisfies conditions (8) and (9) from the conclusion of Lemma 3, but the crucial condition (10) on $|P'|$, which by construction is equal to $|\mathcal{R}|$, remains to be shown. To this end, we invoke the work on the combinatorics of rational numbers found in [11] and [1].

Lemma 5 (Lemma CR in [1])**.**

$$|\mathcal{R}| \ge \frac{|\tilde{P}|(\min_{s\in\tilde{P}}|P_s|)^2}{\tilde{V}\tau^8(1+\log V)},$$

where

$$L = \max_{r\le\tilde{V}} \left| \left\{ \frac{b}{r} : \frac{b}{r} \text{ corresponds to } t \in \bigcup_{s\in\tilde{P}} P_s \right\} \right|,$$

$\tau(q)$ *is the divisor function and* $\tau = \max_{q\le V\tilde{V}}\tau(q)$.

It is a well-known fact of the divisor function that $\tau(n) \le n^{1/\log\log n}$ for large n, and since $V\tilde{V} \le Q$, we have that $\tau \le (\log N)^{2\epsilon}$. We also have from (6) that

$$\sigma^2(\log N)^{-1+\epsilon} \ge \max_{r\le Q} \sum_{t\in\mathbf{M}_r(Q)} |\widehat{B}(t)|^2 \ge \max_{r\le\tilde{V}} \sum_{t\in\mathbf{M}_r(U)} |\widehat{B}(t)|^2 \ge \frac{\sigma^2}{\tilde{W}^2}L,$$

and hence

$$L \le \tilde{W}^2(\log N)^{-1+\epsilon}. \tag{34}$$

Combining the estimates on τ and L with (32), (33), and Lemma 5, we have

$$|P'| \gg \frac{|P|}{(\log\sigma^{-1}+\log U)^3}\frac{\tilde{U}^2\tilde{W}^2\tilde{V}}{U^2(\log(\sigma^{-1})+\log U)^8}\frac{(\log N)^{1-\epsilon}}{\tilde{V}\tilde{W}^2(\log N)^{16\epsilon}} \ge \tilde{U}^2\frac{|P|}{U^2}(\log N)^{1-18\epsilon}.$$

Recalling that we set $U' = \tilde{U}$, the lemma follows. □

Appendix A. Exponential Sum Estimates: Proofs of (17), (18), (21), (27), and (28)

Throughout Appendix A we write $h(x) = \alpha x^2 + \beta x + \gamma$. We note that (18) is just a special case of (28) with $U = 1$, and the same can be said about (17) and (27), respectively, with just a slight difference in requirement for the absolute constant defining c_0 in (12). In either case, it suffices to establish the latter versions, for which we invoke some classical Weyl sum estimates.

Lemma 6. *If $t \in \mathbb{Z}_{N'}$ and $t/N' = a/q + \lambda$ with $q \leq M^{0.1}$, $(a,q) = 1$, and $|\lambda| < M^{-1.9}$, then*

$$W(t) = \frac{1}{qM^2} S(a,q) \int_1^M x e^{2\pi i h(x)\lambda} dx + O(M^{-0.8}),$$

where

$$S(a,q) = \sum_{r=0}^{q-1} e^{2\pi i h(r) a/q}.$$

Lemma 7. *If $t \in \mathbb{Z}_{N'}$, $|t/N' - a/q| < 1/q^2$, and $(a,q) = 1$, then*

$$|W(t)| \ll \log M (\alpha/q + \alpha/M + q/M^2)^{1/2}$$

Lemma 6 is a weighted version of the traditional major arc asymptotic for Weyl sums, and in particular is a special case of Lemma 11 in [7]. Lemma 7 follows from the standard Weyl Inequality for quadratic polynomials (see [10] for example) and summation by parts.

We will also need some facts about the polynomial h which follow from its construction as an auxiliary polynomial of f in Lemma 1. Specifically, the bounds on d and N' in Lemma 1 and the definition of M tell us that $\max\{\alpha, |\beta|, |\gamma|\} \leq C_f d \leq C_f N^{.01}$, hence

$$M \geq c_f N^{.49} \tag{35}$$

and

$$\max\{\alpha, |\beta|, |\gamma|\} < M^{.03}. \tag{36}$$

Also, $\alpha \geq c_f(|\beta| + |\gamma|)$, and therefore

$$\alpha M^2 \geq N'/4. \tag{37}$$

Finally, by (5) and (35), we have that

$$\eta > M^{-.01}, \tag{38}$$

and we are ready to establish the purported estimates.

A.1. **Proof of (28).** By (38), we see that for $U \leq Q^{1/6}$ the hypotheses of Lemma 6 are comfortably satisfied whenever $t \in \mathbf{M}_{a/q}(U) \subset \mathfrak{M}(U)$ with $(a,q) = 1$. We will in fact show that (28) holds whenever the hypotheses of Lemma 6 hold, even if t is not in the major arcs as we defined them, and we will need this when proving (27).

For the Gauss sum $S(a,q)$, we use the well known estimate

$$|S(a,q)| \leq C(\mathrm{cont}(h) q)^{1/2}, \tag{39}$$

which, for example, is a special case of Lemma 6 in [7]. Combining (39) and (37) with Lemma 6, it suffices to show

$$\Big| \int_1^M x e^{2\pi i h(x)\lambda} dx \Big| \leq \min\left\{ M^2, (\alpha|\lambda|)^{-1} \right\}. \tag{40}$$

The first of the two implicit inequalities is trivial, and in particular holds with $\lambda = 0$. For $\lambda \neq 0$, we first ignore the lower order terms in the polynomial by observing

$$\Big| \int_1^M x e^{2\pi i h(x)\lambda} - x e^{2\pi i \alpha x^2 \lambda} dx \Big| \leq M \int_1^M |1 - e^{2\pi i \beta x \lambda}| dx \leq 2\pi M^3 |\beta||\lambda| \leq M^{-0.7}/|\lambda|,$$

since $|\beta| < M^{.03}$ by (36) and $|\lambda| < M^{-1.9}$ by assumption.

For the main term, we change variables ($y := \alpha x^2$) to see

$$\int_1^M x e^{2\pi i \alpha x^2 \lambda} dx = \frac{1}{2\alpha} \int_\alpha^{\alpha M^2} e^{2\pi i y \lambda} dy = \frac{1}{2\alpha} \Big(\frac{e^{2\pi i \alpha M^2} - e^{2\pi i \alpha}}{2\pi i \lambda} \Big),$$

which in absolute value is clearly at most $(4\alpha|\lambda|)^{-1}$, and the estimate follows. □

A.2. **Proof of (27).** Fixing $t \in \mathfrak{m}(U)$, we have by the pigeonhole principle that there exist $1 \le q \le M^{1.9}$ and $(a,q) = 1$ with $|t/N' - a/q| < 1/(qM^{1.9})$. If $U^2\eta^{-2} \le q \le M^{0.1}$, then (39) and Lemma 6, with the trivial bound on the integral, immediately yield the desired estimate. If $M^{0.1} \le q \le M^{1.9}$, then (36) and Lemma 7 imply

$$|W(t)| \le M^{-.03},$$

which by (38) is much stronger than the required estimate. If $1 \le q \le U^2\eta^{-2}$, then it must be the case that

$$|t/N' - a/q| \ge U^2/(\eta^2 M^2), \tag{41}$$

as otherwise we would have $t \in \mathfrak{M}(U)$. It then follows from Lemma 6, (40), and (41) that

$$|W(t)| \le C\eta^2/U^2,$$

which is again stronger than the requirement. □

A.3. **Proof of (21).** We first note that

$$\begin{aligned}\sum_{t \in \mathbb{Z}_{N'}} |W(t)|^6 &= \frac{1}{M^{12}} \sum_{j < x_1, \dots, x_6 < M} x_1 \cdots x_6 \sum_{t \in \mathbb{Z}_{N'}} e^{2\pi i (h(x_1)+h(x_2)+h(x_3)-h(x_4)-h(x_5)-h(x_6))t/N'} \\ &\le \frac{N'}{M^6} \cdot \#\left\{(x_1, \dots, x_6) : j < x_i < M, h(x_1)+h(x_2)+h(x_3) \equiv h(x_4)+h(x_5)+h(x_6) \pmod{N'}\right\}.\end{aligned}$$

By definition of j and M, both sides of the congruence above lie in $[1, N')$, so congruence modulo N' implies equality. Noting this fact, we have

$$\sum_{t \in \mathbb{Z}_{N'}} |W(t)|^6 \le \frac{N'}{M^6} \cdot J(\alpha, \beta, M)$$

where

$$J(\alpha,\beta,M) = \#\{(x_1,\dots,x_6) : 1 \le x_i \le M, \alpha(x_1^2+x_2^2+x_3^2-x_4^2-x_5^2-x_6^2) = \beta(x_1+x_2+x_3-x_4-x_5-x_6)\}.$$

By Proposition 2, we know that $(\alpha,\beta) = \mathrm{cont}(h) \le C_f$, so by (37) it suffices to show under the assumption $(\alpha,\beta) = 1$ that

$$J(\alpha,\beta,M) \ll M^4/\alpha.$$

Examining the equation

$$\alpha(x_1^2+x_2^2+x_3^2-x_4^2-x_5^2-x_6^2) = \beta(x_1+x_2+x_3-x_4-x_5-x_6), \tag{42}$$

we see that the right hand side must be divisible by α, so if $(\alpha,\beta) = 1$, it must be the case that α divides $x_1+x_2+x_3-x_4-x_5-x_6$. Since this expression takes values in $(-3M, 3M)$, there are at most $6M/\alpha + 1 \le 7M/\alpha$ choices for its value, where the last inequality follows from (36). Also, a chosen value for this expression determines the value of $x_1^2+x_2^2+x_3^2-x_4^2-x_5^2-x_6^2$ required to satisfy (42).

Now we invoke a special case of the solution to Tarry's problem, which says that for any fixed $s, t \in \mathbb{Z}$, the number of solutions to the system

$$\begin{aligned}x_1+x_2+x_3-x_4-x_5-x_6 &= s \\ x_1^2+x_2^2+x_3^2-x_4^2-x_5^2-x_6^2 &= t\end{aligned}$$

with $1 \le x_i \le M$ is at most CM^3. Discussions of this fact and Tarry's problem in general can be found in [4] and [15]. Putting the pieces together, we have that for $(\alpha,\beta) = 1$,

$$J(\alpha,\beta,M) \ll \frac{M}{\alpha} \cdot M^3 = \frac{M^4}{\alpha},$$

and the result follows. □

Appendix B. Proofs of Propositions 1 and 2

B.1. **Proof of Proposition 1.** First we recall that a polynomial is intersective if and only if it has a root in the p-adic integers for every prime p.

Suppose $f(x) = ax^2 + bx + c \in \mathbb{Z}[x]$ has no rational roots, hence $b^2 - 4ac$ is not a perfect square, and let p be any prime such that $\mathrm{ord}_p(b^2 - 4ac)$, the exponent of p in the prime factorization of $b^2 - 4ac$, is odd. Letting $\mathbb{Q}_p$ denote the field of p-adic numbers, we have that $b^2 - 4ac$ is not a square in $\mathbb{Q}_p$. Therefore, by the quadratic formula, f has no roots in $\mathbb{Q}_p$, hence no p-adic integer roots, so f is not an intersective polynomial.

Now suppose that $f(x) = a(\alpha x + \beta)(\gamma x + \lambda)$ with $a, \alpha, \beta, \gamma, \lambda \in \mathbb{Z}$ and $(\alpha, \beta) = (\gamma, \lambda) = 1$. If p is a prime that divides both α and γ, then we see that f has no root modulo p^k whenever $p^k \nmid a$, hence f is not an intersective polynomial.

Conversely, if $(\alpha, \gamma) = 1$, we see that $-\beta/\alpha$ is a p-adic integer root of f whenever $p \nmid \alpha$, and $-\lambda/\gamma$ is a p-adic integer root of f whenever $p \nmid \gamma$. Since at least one of these divisibility conditions holds for every prime p, f is an intersective polynomial. □

B.2. **Proof of Proposition 2.** Recall that at this stage we have a fixed intersective quadratic $f \in \mathbb{Z}[x]$.

We can assume $f(x) = ax^2 + bx + c = (\alpha x + \beta)(\gamma x + \lambda)$ with $(a, b, c) = (\alpha, \beta) = (\gamma, \lambda) = 1$, since $\mathrm{cont}(f_d)$ and the expression $|\alpha\lambda - \beta\gamma|$ both behave predictably under scaling of f. Further, since f is intersective, we have that $(a, b) = (\alpha, \gamma) = 1$.

In this case, we have $f_d(x) = dax^2 + (2ar_d + b)x + f(r_d)/d$, so

$$\mathrm{cont}(f_d) = (da, 2ar_d + b) = (d, 2ar_d + b),$$

where the last equality holds because $(a, b) = 1$ implies $(a, 2ar_d + b) = 1$.

Now suppose that a prime power p^k divides both d and $2ar_d + b = f'(r_d) = \alpha(\gamma r_d + \lambda) + \gamma(\alpha r_d + \beta)$. Because $p \mid d$, and by the construction of the root r_d described following Definition 3, it is either the case that $p^k \mid \alpha r_d + \beta$ or $p^k \mid \gamma r_d + \lambda$. We will assume the former without loss of generality, so in particular $p \nmid \alpha$.

We then see that $p^k \mid \alpha r_d + \beta$ and $p^k \mid f'(r_d)$ implies $p^k \mid \alpha(\gamma r_d + \lambda)$, and since $p \nmid \alpha$, it must be the case that $p^k \mid \gamma r_d + \lambda$. In summary, we have that

$$r_d \equiv -\beta/\alpha \equiv -\lambda/\gamma \bmod p^k.$$

In particular $p^k \mid \alpha\lambda - \beta\gamma$, and the result follows. □

References

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

[2] D. Berend, Y. Bilu, *Polynomials with roots modulo every integer*, Proc. Amer. Math. Soc. 124 (1996), pp. 1663-1671.

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

[4] L. K. Hua, *Additive theory of prime numbers* , American Mathematical Society, Providence, RI 1965.

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

[6] T. H. Lê, *Intersective polynomials and the primes*, J. Number Theory 130 no. 8 (2010), pp. 1705-1717.

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

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

[9] N. Lyall, Á. Magyar, *Simultaneous polynomial recurrence*, Bull. Lond. Math. Soc. 43 (2011), no. 4, 765785

[10] H. L. Montgomery *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathematics, 84.

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

[12] A. Sárközy, *On difference sets of sequences of integers I*, Acta. Math. Hungar. 31 (1-2) (1978), pp. 125-149.

[13] S. Slijepčević, *A polynomial Sárközy-Furstenberg theorem with upper bounds*, Acta Math. Hungar. 98 (2003), pp. 275-280

[14] J. Wolf, *Arithmetic structures in sets of integers*, Ph.D. thesis, University of Cambridge, submitted December 2007

[15] T. Wooley, *Some remarks on Vinogradov's mean value theorem and Tarry's problem*, Monatsh. Math. 122, no. 3 (1996), pp. 265-273.

Département de Mathématiques et de Statistique, Université de Montréal, CP 6128, Centre-ville, Montréal, QC H3C 3J7, Canada

E-mail address: mhamel@dms.umontreal.ca

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: lyall@math.uga.edu

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: arice@math.uga.edu