ACTA ARITHMETICA
157.1 (2013)

Sárközy's theorem for $\mathcal{P}$-intersective polynomials

by

Alex Rice (Athens, GA)

1. Introduction

1.1. Background. 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 $\{1,\dots,N\}$. In the late 1970s, Sárközy and Furstenberg independently confirmed a conjecture of Lovász that any set of natural numbers of positive upper density necessarily contains two elements which differ by a perfect square. Furstenberg [2] used ergodic theory and obtained a purely qualitative result, proving the conjecture exactly as stated above. Sárközy, however, employed a Fourier-analytic density increment strategy, inspired by Roth's proof of the analogous conjecture for three-term arithmetic progressions [17], to prove the following quantitative strengthening.

Theorem A (Sárközy, [19]). *If $A \subseteq [1,N]$ and $n^2 \notin A-A$ for all $n \in \mathbb{N}$, then*

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

In this and the following theorems, $A - A$ denotes the difference set $\{a-a' : a,a' \in A\}$, the symbol $\ll$ denotes "less than a constant times", and we implicitly assume that N is large enough to make the right hand side of the inequalities defined and positive.

An extensive literature has been developed on improvements and extensions of Theorem A, for which the reader may refer to [15], [1], [21], [12], [13], [8], and [6]. In the same series of papers, Sárközy answered a similar question of Erdős concerning shifted primes.

2010 *Mathematics Subject Classification*: Primary 11B30.
Key words and phrases: arithmetic combinatorics, difference set, Sárközy's theorem, intersective polynomials.

DOI: 10.4064/aa157-1-4

THEOREM B (Sárközy, [20]). *If $A \subseteq [1,N]$ and $p-1 \notin A-A$ for all primes p, then*

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

The bound in Theorem B has been improved, first by Lucier [11] and later by Ruzsa and Sanders [18], who replaced (1) with $|A|/N \ll e^{-c(\log N)^{1/4}}$.

A natural generalization of Theorem A is the replacement of the squares with the image of a more general polynomial. However, to hope for such a result for a given polynomial $h \in \mathbb{Z}[x]$, it is clearly necessary that h has a root modulo q for every $q \in \mathbb{N}$, as otherwise there would be a set $q\mathbb{N}$ of positive density with no differences in the image of h. It follows from a theorem of Kamae and Mendès France [7] that this condition is also sufficient, in a qualitative sense, and in this case we say that h is an *intersective polynomial*. Examples of intersective polynomials include any polynomial with an integer root and any polynomial with two rational roots with coprime denominators. However, there are also intersective polynomials with no rational roots, for example $(x^3-19)(x^2+x+1)$. The best current bounds for this most general setting are essentially 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.

THEOREM C (Lucier, [12]). *Suppose $h \in \mathbb{Z}[x]$ is an intersective polynomial of degree $k \geq 2$ with positive leading term. If $A \subseteq [1,N]$ and $h(n) \notin A-A$ for all $n \in \mathbb{N}$ with $h(n) > 0$, then*

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

where the implied constant depends only on h.

In [16], the author made an extremely mild improvement to Theorem C, showing that one can in fact take $\mu = 1$. By the symmetry of difference sets, Theorem C and all the following theorems clearly imply the analogous results for the negative values of a polynomial with negative leading term.

Recently, Hamel, Lyall, and the author utilized Lucier's techniques in extending the best known bound on the size of a set with no square differences, due originally to Pintz, Steiger, and Szemerédi [15] and extended to k^{th} powers by Balog, Pelikan, Pintz, and Szemerédi [1], to all intersective polynomials of degree 2, which is to say quadratic polynomials which have two rational roots with coprime denominators.

THEOREM D (Hamel, Lyall, Rice, [6]). *Suppose $h \in \mathbb{Z}[x]$ is an intersective quadratic polynomial with positive leading term. If $A \subseteq [1,N]$ and*

$h(n) \notin A - A$ for all $n \in \mathbb{N}$ with $h(n) > 0$, then

$$|A|/N \ll (\log N)^{-c \log\log\log\log N}$$

for any $c < 1/\log 3$, where the implied constant depends only on h and c.

Some work has also been done to combine extensions of Theorem A with Theorem B. Li and Pan [10] established the following quantitative result.

THEOREM E (Li, Pan, [10]). *Suppose $h \in \mathbb{Z}[x]$ has positive leading term and $h(1) = 0$. If $A \subseteq [1, N]$ and $h(p) \notin A - A$ for all primes p with $h(p) > 0$, then*

$$|A|/N \ll 1/\log\log\log N.$$

Additionally, Lê and Li–Pan applied transference principles inspired by work of Green and Tao ([3], [4], [5]) to prove analogs of Theorems C and E, respectively, for dense subsets of the primes, which we denote by $\mathcal{P}$. We state the qualitative results below.

THEOREM F (Lê, [8]). *If $h \in \mathbb{Z}[x]$ is an intersective polynomial with positive leading term, $A \subseteq \mathcal{P}$, and*

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]|}{|\mathcal{P} \cap [1, N]|} > 0, \tag{2}$$

then there exist $a, a' \in A$ and $n \in \mathbb{N}$ with $a - a' = h(n) > 0$.

If $A \subseteq \mathcal{P}$ meets condition (2), we say that it has *positive relative upper density* in the primes.

THEOREM G (Li, Pan, [10]). *If $h \in \mathbb{Z}[x]$ has positive leading term, $h(1) = 0$, and $A \subseteq \mathcal{P}$ has positive relative upper density, then there exist $a, a' \in A$ and $p \in \mathcal{P}$ with $a - a' = h(p) > 0$.*

1.2. Main result of this paper. Just as there are intersective polynomials without integer roots, it seems natural to think that a result like Theorem E should hold for a larger class of polynomials.

A moment's consideration indicates that the correct analog to the intersective condition on a polynomial h when looking for differences of the form $h(p)$ is to insist that h not only has a root modulo q for every $q \in \mathbb{N}$, but has a root at a congruence class that admits infinitely many primes, leading to the following definition.

DEFINITION 1. A polynomial $h \in \mathbb{Z}[x]$ is called *$\mathcal{P}$-intersective* if, for every $q \in \mathbb{N}$, there exists $r \in \mathbb{Z}$ such that $(r, q) = 1$ and $q \mid h(r)$. Equivalently, for every $p \in \mathcal{P}$, there exists $z_p \in \mathbb{Z}_p$, where $\mathbb{Z}_p$ denotes the p-adic integers, such that $h(z_p) = 0$ and $z_p \not\equiv 0 \bmod p$.

REMARK. After the initial version of this paper was posted on arXiv server, the author learned that in [9], a survey on problems and results on

intersective sets, Thái Hoàng Lê independently posed the same question and arrived at the same notion, which he termed *intersective polynomials of the second kind*. Even later, the author learned that these polynomials were considered by Wierdl [23] in his thesis, where he called them *intersective along the primes*.

Examples of $\mathcal{P}$-intersective polynomials include any polynomial with a root at 1 or -1, any polynomial with two rational roots a/b and c/d such that $(ab, cd) = 1$, and presumably lots more. The necessity of this condition is almost as clear as that of the original intersective condition. To exhibit this, suppose we have $h \in \mathbb{Z}[x]$ and $q \in \mathbb{N}$ such that the only roots of h modulo q share common factors with q. In particular, there are finitely many primes p such that $q \mid h(p)$. Letting $m = \max\{h(p)/q : p \in \mathcal{P},\ q \mid h(p)\}$ if such primes exist and $m = 0$ otherwise, we see that $q(m+1)\mathbb{N}$ is a set of positive upper density which contains no differences of the form $h(p)$.

Wierdl [23] observed in his thesis that one can again deduce the sufficiency of this condition, in a qualitative sense, from the aforementioned theorem of Kamae and Mendès France [7], and here we borrow heavily from [12], [18], [13], and [10] to establish the following quantitative result.

THEOREM 1. *Suppose $h \in \mathbb{Z}[x]$ is a $\mathcal{P}$-intersective polynomial of degree $k \geq 2$ with positive leading term. If $A \subseteq [1, N]$ and $h(p) \notin A - A$ for all $p \in \mathcal{P}$ with $h(p) > 0$, then*

$$|A|/N \ll (\log N)^{-c} \tag{3}$$

for any $c < 1/(2k-2)$, where the implied constant depends only on h and c.

In fact, with a few careful modifications one can sharpen (3) to

$$\frac{|A|}{N} \ll \left(\frac{(\log\log N)^2(\log\log\log N)^{2k}}{\log N}\right)^{1/(2k-2)},$$

but here we stick to the slightly less precise version for a more pleasing exposition.

1.3. Additional results. In addition to Theorem 1, one can conclude the following analogs of previous results from the estimates we obtain along the way.

THEOREM 2. *Suppose $h \in \mathbb{Z}[x]$ is a $\mathcal{P}$-intersective quadratic polynomial with positive leading term. If $A \subseteq [1, N]$ and $h(p) \notin A - A$ for all $p \in \mathcal{P}$ with $h(p) > 0$, then*

$$|A|/N \ll (\log N)^{-c\log\log\log\log N}$$

for any $c < 1/2\log 3$, where the implied constant depends only on h and c.

Just as in the traditional setting, the $\mathcal{P}$-intersective condition is greatly simplified when restricted to degree 2, as a quadratic polynomial is $\mathcal{P}$-intersective if and only if it has rational roots a/b and c/d with $(ab, cd) = 1$.

THEOREM 3. *If $h \in \mathbb{Z}[x]$ is a $\mathcal{P}$-intersective polynomial with positive leading term and $A \subseteq \mathcal{P}$ has positive relative upper density, then there exist $a, a' \in A$ and $p \in \mathcal{P}$ with $a - a' = h(p) > 0$.*

With the techniques and results of this paper, the modifications of the arguments in [6] and [8] required to obtain Theorems 2 and 3, respectively, are so minor that we do not provide the full details here. Alternatively, we discuss the required adaptations informally in Appendix B.

2. Preliminaries. To begin our effort to prove Theorem 1, we fix a $\mathcal{P}$-intersective polynomial h of degree $k \geq 2$ with positive leading term and an arbitrary $\epsilon > 0$, and we set $s = 2^k + 6$ and $K = 2^{10k}$. For our current purposes we only use that $s > 9$, but the choice also plays a role in our discussion of Theorems 2 and 3. We also fix a natural number N which, at the expense of the constant in Theorem 1, we are always free to insist is sufficiently large with respect to h and ϵ. For convenience, we take this as a perpetual hypothesis and abstain from including it further. We use the letters C and c to denote appropriately large or small positive constants, which will change from line to line and which we allow, along with any implied constants, to depend on h and ϵ. Further, we fix a set $A \subseteq [1, N]$ with $|A|/N = \delta > 0$ and set

$$Q(\delta) = \exp(C_0 \delta^{-(k+\epsilon-1)})$$

for a constant C_0.

2.1. Auxiliary polynomials and related definitions. We apply the modified density increment strategy described in [12], which allows for the polynomial to change at each stage of the iteration. The following definitions describe all of the polynomials that we could potentially encounter, as well as several other objects that will appear in the argument.

DEFINITION 2. For each $p \in \mathcal{P}$, we fix $z_p \in \mathbb{Z}_p$ with $h(z_p) = 0$ and $z_p \not\equiv 0 \mod p$. By reducing and applying the Chinese Remainder Theorem, the choices of z_p determine, for each natural number d, a unique integer $r_d \in (-d, 0]$, which consequently satisfies $d \mid h(r_d)$ and $(r_d, d) = 1$.

We define the function λ on $\mathbb{N}$ by letting $\lambda(p) = p^m$, where m is the multiplicity of z_p as a root of h in $\mathbb{Z}_p$, and then extending it to be completely multiplicative.

For each $d \in \mathbb{N}$, we define the *auxiliary polynomial* h_d by

$$h_d(x) = h(r_d + dx)/\lambda(d).$$

Lucier observed in Lemma 21 of [12] that each h_d has integer coefficients, and it is important to note that the leading coefficients grow with d at least as quickly, up to a constant, as the other coefficients.

Further, we let $\Lambda_d = \{x \in \mathbb{N} : r_d + dx \in \mathcal{P}\}$, and for $L \in \mathbb{N}$ we define

$$H_d = H_d(L) = \{x \in \mathbb{N} : 0 < h_d(x) < L/s\}$$

and

$$M_d = M_d(L) = (L/sb_d)^{1/k},$$

where b_d is the leading coefficient of h_d, noting that

$$|[1, M_d] \triangle H_d| = O(1), \tag{4}$$

where $\triangle$ denotes the symmetric difference. We also define a function ν_d on $\mathbb{Z}$ by

$$\nu_d(x) = \frac{\phi(d)}{d} \log(r_d + dx) 1_{\Lambda_d}(x),$$

where ϕ is the Euler totient function, and for a set $B \subseteq [1, L]$ we define

$$\mathcal{R}_d(B) = \mathcal{R}_d(B, L) = \sum_{\substack{x \in \mathbb{Z} \\ y \in H_d}} 1_B(x) 1_B(x + h_d(y)) \nu_d(y).$$

In the definitions above, L should always be replaced with the size of the appropriate ambient interval.

2.2. Counting primes in arithmetic progressions. For $X, a, q \in \mathbb{N}$, we define

$$\psi(X, a, q) = \sum_{\substack{p \in \mathcal{P} \cap [1, X] \\ p \equiv a \bmod q}} \log p.$$

The classical estimates on $\psi(X, a, q)$ come from the famous Siegel–Walfisz Theorem, which can be found for example in Corollary 11.19 of [14].

LEMMA 1 (Siegel–Walfisz Theorem). *If $q \leq (\log X)^D$ and $(a, q) = 1$, then*

$$\psi(X, a, q) = X/\phi(q) + O(Xe^{-c\sqrt{\log X}})$$

for some constant $c = c(D) > 0$.

Ruzsa and Sanders [18] established asymptotics for $\psi(X, a, q)$ for certain moduli q beyond the limitations of Lemma 1 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. *If $Q(\delta) \leq e^{c_1\sqrt{\log N}}$ for a sufficiently small constant $c_1 = c_1(k) > 0$, then there exist $q_0 \leq Q(\delta)^{3K}$ and $\rho \in [1/2, 1)$ with $(1 - \rho)^{-1} \ll q_0$*

such that

$$\psi(X,a,q)=\frac{X}{\phi(q)}-\frac{\chi(a)X^{\rho}}{\phi(q)\rho}+O(Xe^{-30kK^2c_1\sqrt{\log X}}), \tag{5}$$

where χ is a Dirichlet character modulo q_0, provided $X\geq N^{1/10k}$, $q_0\,|\,q$, $(a,q)=1$, and $q\leq(q_0Q(\delta))^{3K}$.

Lemma 2 is a purpose-built special case of Proposition 4.7 of [18], which in the language of that paper can be deduced by considering the pair $(Q(\delta)^{10K^2},Q(\delta)^{3K})$, 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.

As remarked in the proof of Proposition 5.3 of [18], the asymptotic in Lemma 2 implies that under the hypotheses we have

$$\psi(X,a,q)\gg\frac{X}{\phi(q)}-\frac{X^{\rho}}{\phi(q)\rho}\geq(1-\rho)X/\phi(q)\gg\frac{X}{q_0\phi(q)}. \tag{6}$$

2.3. A uniform estimate on $\mathcal{R}_d$. We obtain Theorem 1 as a consequence of the following, stronger result, which says that the number of solutions to $a-a'=h(p)>0$ with $a,a'\in A$, $p\in\mathcal{P}$, has the correct order of magnitude. In addition, we obtain this estimate uniformly in d for a range of auxiliary polynomials h_d, which serves as the primary input required to apply the techniques of [8] and conclude Theorem 3.

THEOREM 4. *There exists a constant C depending only on h, ϵ, and C_0 such that*

$$\mathcal{R}_d(A)\geq\exp(-C\delta^{-(k+\epsilon-1)})NM_d$$

for all $d\leq\max\{\log N,Q(\delta)\}$, provided $\delta\geq C(\log N)^{-1/2(k+\epsilon-1)}$.

3. Main iteration lemma: deducing Theorem 4. We now make the assumption that

$$Q(\delta)\leq e^{c_1\sqrt{\log N}} \tag{7}$$

for a sufficiently small constant $c_1>0$, which is implied by the condition $\delta\geq C_0(\log N)^{-1/2(k+\epsilon-1)}/c_1$, and we fix ρ and q_0 yielded by Lemma 2. Also, we set $\gamma=k+\epsilon/2$ and for $d,L\in\mathbb{N}$ we define

$$\Psi_d=\Psi_d(L)=\phi(d)\psi(dM_d,r_d,d)/d,$$

noting by (6) that for appropriate d and L we have

$$\Psi_d\gg(1-\rho)M_d\gg M_d/q_0. \tag{8}$$

We deduce Theorem 4 from the following iteration lemma, which states that a set which is deficient in the desired arithmetic structure spawns a new, significantly denser subset of a slightly smaller interval with an inherited deficiency in the structure associated to an appropriate auxiliary polynomial.

LEMMA 3. *Suppose* $B \subseteq [1, L]$, $|B|/L = \sigma \geq \delta$, *and* $L \geq \sqrt{N}$. *If* $q_0 \mid d$, $d/q_0 \leq \max\{\log N, Q(\delta)\}^2$, *and*

$$\mathcal{R}_d(B) \leq \sigma^2 L \Psi_d/8,$$

then there exists $q \ll \sigma^{-\gamma}$ *and* $B' \subseteq [1, L']$ *with* $L' \gg \sigma^{\gamma(k+1)} L$, $\mathcal{R}_{qd}(B') \leq \mathcal{R}_d(B)$, *and*

$$|B'| \geq (\sigma + c\sigma^{\gamma})L'.$$

The following proposition exhibits the aforementioned inheritance of deficiency in arithmetic structure, and is essential to the deduction of Theorem 4 from Lemma 3 as well as the proof of Lemma 3 itself.

PROPOSITION 4. *If* $B \subseteq [1, L]$ *and* $B' \subseteq \{\ell \in [1, L'] : x + \ell\lambda(q) \in B\}$ *for some* $x \in \mathbb{Z}$, $q \in \mathbb{N}$, *and* $L' \leq L/\lambda(q)$, *then for any* $d \in \mathbb{N}$,

$$\mathcal{R}_{qd}(B') \leq \mathcal{R}_d(B).$$

Proof. Suppose $B \subseteq [1, L]$, $B' \subseteq \{\ell \in [1, L'] : x + \ell\lambda(q) \in B\}$, $L' \leq L/\lambda(q)$, and

$$L'/s > \ell - \ell' = h_{qd}(n) = \frac{h(r_{qd} + qdn)}{\lambda(q)\lambda(d)} > 0$$

for $\ell, \ell' \in B'$, $n \in \Lambda_{qd}$. Recalling that $r_{qd} \equiv r_d \bmod d$, there is an integer m such that $r_{qd} = r_d + md$, so

$$\ell - \ell' = \frac{h(r_d + d(m + qn))}{\lambda(q)\lambda(d)} = \frac{h_d(m + qn)}{\lambda(q)},$$

and therefore

$$0 < h_d(m+qn) = \lambda(q)\ell - \lambda(q)\ell' = (x+\lambda(q)\ell) - (x+\lambda(q)\ell') < \lambda(q)L'/s \leq L/s.$$

Moreover, we know that $r_d + d(m + qn) = r_{qd} + qdn \in \mathcal{P}$, so $m + qn \in \Lambda_d$, and the result follows. ∎

3.1. Proof of Theorem 4. On fixing $d \leq \max\{\log N, Q(\delta)\}$ and partitioning $[1, N]$ 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 $P = \{x + \ell\lambda(q_0) : 1 \leq \ell \leq N_0\}$ such that $N/2\lambda(q_0) \leq N_0 \leq N/\lambda(q_0)$ and $|A \cap P|/N_0 \geq \delta$.

This allows us to define $A_0 \subseteq [1, N_0]$ by

$$A_0 = \{\ell \in [1, N_0] : x + \ell\lambda(q_0) \in A\},$$

which consequently satisfies

$$|A_0|/N_0 = \delta_0 \geq \delta, \quad N_0 \geq N/Q(\delta)^{4kK}, \quad \mathcal{R}_{q_0 d}(A_0) \leq \mathcal{R}_d(A),$$

where the last inequality follows from Proposition 4.

We then iteratively apply Lemma 3, which yields, for each m, a set $A_m \subseteq [1, N_m]$ with $|A_m| = \delta_m N_m$ and

$$\mathcal{R}_{d_m}(A_m) \le \mathcal{R}_d(A) \tag{9}$$

satisfying

$$N_m \ge (c\delta)^{Cm} N_0, \quad \delta_m \ge \delta_{m-1} + c\delta_{m-1}^{\gamma}, \quad q_0 \mid d_m, \quad d_m/q_0 \le (c\delta)^{-Cm} d \tag{10}$$

as long as

$$N_m \ge \sqrt{N}, \quad d_m/q_0 \le \max\{\log N, Q(\delta)\}^2, \tag{11}$$

and

$$\mathcal{R}_{d_m}(A_m) \le \delta_m^2 N_m \Psi_{d_m}/8. \tag{12}$$

By (10), we see that the density δ_m would surpass 1, and hence (11) or (12) must fail, with

$$m = C\delta^{-(\gamma-1)}. \tag{13}$$

However, if C_0 is sufficently large then (13) implies $(c\delta)^{-Cm} \le Q(\delta)$, hence $N_m \ge N/Q(\delta) \ge \sqrt{N}$ and $d_m/q_0 \le Q(\delta)d \le \max\{\log N, Q(\delta)\}^2$, so (11) holds. Further, we see by (8) and (10) that

$$\delta_m^2 N_m \Psi_{d_m} \ge (c\delta)^{Cm} N_0 M_d/q_0 \ge \exp(-C\delta^{-(k+\epsilon-1)}) N M_d,$$

so if $\mathcal{R}_d(A) \le \exp(-C\delta^{-(k+\epsilon-1)}) N M_d$ for a sufficiently large constant C, then by (9) we deduce that (12) also holds. This yields a contradiction, and the theorem follows. ∎

4. Density increment strategy: deducing Lemma 3

4.1. Fourier analysis on $\mathbb{Z}$. We embed our finite sets in $\mathbb{Z}$, on which we utilize the 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 $\mathbb{T}$ denotes the circle parameterized by the interval $[0, 1]$ with 0 and 1 identified, by

$$\widehat{F}(\alpha) = \sum_{x \in \mathbb{Z}} F(x) e^{-2\pi i x \alpha}.$$

Given $L \in \mathbb{N}$ and a set $B \subseteq [1, L]$ with $|B| = \sigma L$, we examine the Fourier-analytic behavior of B by considering the *balance function*, f_B, defined by

$$f_B = 1_B - \sigma 1_{[1,L]}.$$

4.2. The circle method. We analyze the behavior of $\widehat{f_B}$ using the Hardy–Littlewood circle method, decomposing the frequency space into two pieces: the points on the circle which are close to rationals with small denominator, and those which are not.

Definition 3. Given $L \in \mathbb{N}$ and $\eta > 0$, we define, for each $q \in \mathbb{N}$ and $a \in [1, q]$,

$$\mathbf{M}_{a/q} = \mathbf{M}_{a/q}(L, \eta) = \left\{ \alpha \in \mathbb{T} : \left| \alpha - \frac{a}{q} \right| < \frac{1}{\eta^{\gamma} L} \right\} \quad \text{and} \quad \mathbf{M}_q = \bigcup_{(a,q)=1} \mathbf{M}_{a/q}.$$

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

$$\mathfrak{M} = \bigcup_{q=1}^{\eta^{-\gamma}} \mathbf{M}_q,$$

and $\mathfrak{m}$, the *minor arcs*, by

$$\mathfrak{m} = \mathbb{T} \setminus \mathfrak{M}.$$

4.3. L^2 concentration and density increment lemmas. As usual, the philosophy behind the argument is that a deficiency in the desired arithmetic structure from a set B represents nonrandom behavior, which should be detected in the Fourier-analytic behavior of B. Specifically, we follow the approach of Lyall and Magyar [13] to locate one small denominator q such that $\widehat{f_B}$ has L^2 concentration around rationals with denominator q, then use that information to find a long arithmetic progression on which B has increased density.

Lemma 5 (L^2 concentration). *Suppose $B \subseteq [1, L]$, $|B|/L = \sigma \geq \delta$, and $L \geq \sqrt{N}$, and let $\eta = c_2\sigma$ for a sufficiently small constant $c_2 > 0$. Suppose further that*

$$q_0 \,|\, d, \qquad d/q_0 \leq \max\{\log N, Q(\delta)\}^2, \qquad \mathcal{R}_d(B) \leq \sigma^2 L \Psi_d/8.$$

If $|B \cap (L/9, 8L/9)| \geq 3\sigma L/4$, then there exists $q \leq \eta^{-\gamma}$ such that

$$\int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 \, d\alpha \gg \sigma^{\gamma+1} L.$$

We now invoke a variation of the usual L^2 density increment. Specifically, we quote a result which follows from Proposition 7.2 of [18].

Lemma 6 (Density increment). *Suppose $B \subseteq [1, L]$ with $|B| = \sigma L$ and let $\eta = c_2\sigma$. If*

$$\int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2 \, d\alpha \geq \omega \sigma^2 L,$$

then there exists an arithmetic progression

$$P = \{x + \ell\lambda(q) : 1 \leq \ell \leq L'\}$$

with

$$L/\lambda(q) \geq L' \gg \min\{\eta^{\gamma}, \omega\sigma\} L/\lambda(q) \quad \text{and} \quad |A \cap P|/L' \geq \sigma + \omega\sigma/4.$$

4.4. Proof of Lemma 3. Suppose $B \subset [1, L]$ meets all the hypotheses of the lemma.

If $|B \cap (L/9, 8L/9)| < 3\sigma L/4$, then

$$\max\{|B \cap [1, L/9]|, |B \cap [8L/9]|\} \geq \sigma L/8.$$

In other words, B has density at least $9\sigma/8$ on one of these intervals.

Otherwise, Lemmas 5 and 6 apply, so in either case there exists $q \leq \eta^{-\gamma}$ and an arithmetic progression $P = \{x + \ell\lambda(q) : 1 \leq \ell \leq L'\}$ with

$$L/\lambda(q) \geq L' \gg \sigma^{\gamma} L/\lambda(q)) \gg \sigma^{\gamma(k+1)} L \quad \text{and} \quad |B \cap P|/L' \geq \sigma + c\sigma^{\gamma}.$$

This allows us to define a new set $B' \subset [1, L']$ by

$$B' = \{\ell \in [1, L'] : x + \ell\lambda(q) \in B\},$$

which by Proposition 4 satisfies $\mathcal{R}_{qd}(B') \leq \mathcal{R}_d(B)$, as required. ∎

4.5. Proof of Lemma 5. Suppose $B \subseteq [1, L]$, $|B|/L = \sigma \geq \delta$, and $L \geq \sqrt{N}$. Let $\eta = c_2\sigma$, and suppose further that $q_0 \,|\, d$ and $d/q_0 \leq \max\{\log N, Q(\delta)\}^2$. For the remainder of the proof, we keep this d fixed and omit it from the notations H_d, M_d, ν_d, $\mathcal{R}_d$, and Ψ_d defined in Sections 2 and 3.

Since $h_d(H) \subseteq [1, L/9)$, we see that

$$\begin{aligned}
&\sum_{\substack{x\in\mathbb{Z}\\ y\in H}} f_B(x) f_B(x + h_d(y))\nu(y) \\
&= \sum_{\substack{x\in\mathbb{Z}\\ y\in H}} 1_B(x) 1_B(x + h_d(y))\nu(y) - \sigma \sum_{\substack{x\in\mathbb{Z}\\ y\in H}} 1_B(x) 1_{[1,L]}(x + h_d(y))\nu(y) \\
&\quad - \sigma \sum_{\substack{x\in\mathbb{Z}\\ y\in H}} 1_{[1,L]}(x - h_d(y)) 1_B(x)\nu(y) + \sigma^2 \sum_{\substack{x\in\mathbb{Z}\\ y\in H}} 1_{[1,L]}(x) 1_{[1,L]}(x + h_d(y))\nu(y) \\
&\leq \mathcal{R}(B) + \big(\sigma^2 L - \sigma\big(|B \cap [1, 8L/9)| + |B \cap (L/9, L]|\big)\big) \sum_{y\in H} \nu(y).
\end{aligned}$$

By (4) we have

$$\sum_{y\in H} \nu(y) = \Psi + O(\log L),$$

so if $|B \cap (L/9, 8L/9)| \geq 3\sigma L/4$ and $\mathcal{R}(B) \leq \sigma^2 L\Psi/8$, then

$$\sum_{\substack{x\in\mathbb{Z}\\ y\in H}} f_B(x) f_B(x + h_d(y))\nu(y) \leq -\sigma^2 L\Psi/8. \tag{14}$$

One can easily check using (4) and orthogonality of characters that

$$\sum_{\substack{x\in\mathbb{Z}\\ y\in H}} f_B(x) f_B(x + h_d(y))\nu(y) = \int_0^1 |\widehat{f_B}(\alpha)|^2 S_M(\alpha)\, d\alpha + O(L \log L), \tag{15}$$

where

$$S_X(\alpha) = \sum_{x=1}^{X} \nu(x) e^{2\pi i h_d(x)\alpha}.$$

Combining (14) and (15), we have

$$\int_0^1 |\widehat{f_B}(\alpha)|^2 |S_M(\alpha)|\, d\alpha \geq \sigma^2 L\Psi/16. \tag{16}$$

It follows from Lemma 2, an observation of Lucier on auxiliary polynomials, and Theorem 4.1 of [10] that

$$|S_M(\alpha)| \ll q^{-1/\gamma}\Psi \quad \text{for all } \alpha \in \mathbf{M}_q \subset \mathfrak{M}, \tag{17}$$

and

$$|S_M(\alpha)| \leq C\eta\Psi \leq \sigma\Psi/32 \quad \text{for all } \alpha \in \mathfrak{m}, \tag{18}$$

provided we choose $c_2 < 1/32C$. We discuss these estimates in more detail in Section 5.

From (18) and Plancherel's identity, we have

$$\int_{\mathfrak{m}} |\widehat{f_B}(\alpha)|^2 |S_M(\alpha)|\, d\alpha \leq \sigma^2 L\Psi/32,$$

which together with (16) yields

$$\int_{\mathfrak{M}} |\widehat{f_B}(\alpha)|^2 |S_M(\alpha)|\, d\alpha \geq \sigma^2 L\Psi/32. \tag{19}$$

By (17) and (19), we have

$$\sigma^2 L \ll \Big(\sum_{q=1}^{\eta^{-\gamma}} q^{-1/\gamma}\Big) \max_{q\leq\eta^{-\gamma}} \int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2\, d\alpha \ll \sigma^{-\gamma+1} \max_{q\leq\eta^{-\gamma}} \int_{\mathbf{M}_q} |\widehat{f_B}(\alpha)|^2\, d\alpha,$$

and the lemma follows. ∎

5. Major and minor arc estimates: proof of (17) and (18). We remain in the setting of the proof of Lemma 5, recalling all hypotheses and notation defined there. We first state some required estimates, which we use to deduce (17) and (18); we include the necessary proofs in Appendix A.

Lemma 7. *If $Q(\delta) \geq \log N$ and $\alpha = a/q + \beta$ with $q \leq (q_0 Q(\delta)^2)^K$, $(a,q) = 1$, and $|\beta| < (q_0 Q(\delta)^2)^K/L$, then*

$$S_M(\alpha) = \frac{\phi(d)}{\phi(qd)} G(a,q) \int_1^M (1 - \chi(r_d)(dx)^{\rho-1}) e^{2\pi i h_d(x)\beta}\, dx + O(Me^{-5K^2 c_1\sqrt{\log N}}),$$

where

$$G(a,q) = \sum_{\substack{\ell=0 \\ (r_d+d\ell,q)=1}}^{q-1} e^{2\pi i h_d(\ell)a/q}.$$

LEMMA 8. *If $Q(\delta) \le \log N$ and $\alpha = a/q + \beta$ with $q \le (q_0(\log N)^2)^K$, $(a,q)=1$, and $|\beta| < (q_0(\log N)^2)^K/L$, then*

$$S_M(\alpha) = \frac{\phi(d)}{\phi(qd)} G(a,q) \int_1^M e^{2\pi i h_d(x)\beta}\,dx + O(Me^{-c\sqrt{\log N}}).$$

In Appendix A, we exhibit how Lemma 7 follows from Lemma 2, and Lemma 8 follows from Lemma 1 in an analogous, more standard way.

LEMMA 9. *Suppose $g(x) = a_0 + a_1 x + \cdots + a_k x^k \in \mathbb{Z}[x]$. If $W, b \in \mathbb{Z}$, $q \in \mathbb{N}$ and $(a,q)=1$, then*

$$\Big| \sum_{\substack{\ell=0 \\ (W\ell+b,q)=1}}^{q-1} e^{2\pi i g(\ell)a/q} \Big| \ll \big(\gcd(\mathrm{cont}(g), q_1)\gcd(a_k, q_2)\big)^{1/k} q^{1-1/k}, \tag{20}$$

where $q = q_1 q_2$, q_2 is the maximal divisor of q which is coprime to W, and

$$\mathrm{cont}(g) := \gcd(a_1, \ldots, a_k).$$

The statement of Lemma 9 indicates that we could lose control of the sum $G(a,q)$ if the coefficients of the auxiliary polynomials h_d share larger and larger common factors. The following observation of Lucier ensures that this does not occur.

LEMMA 10 (Lemma 28 in [12]). *For every $d \in \mathbb{N}$,*

$$\mathrm{cont}(h_d) \le |\Delta(h)|^{(k-1)/2}\mathrm{cont}(h),$$

where $\Delta(h) = a^{2k-2}\prod_{i\ne j}(\alpha_i - \alpha_j)^{e_i e_j}$ if h factors over the complex numbers as $a(x-\alpha_1)^{e_1}\cdots(x-\alpha_r)^{e_r}$ with all the α_i's distinct.

While the statement of Lemma 10 is pleasingly precise, we only use that $\mathrm{cont}(h_d)$ is uniformly bounded in terms of the original polynomial h.

COROLLARY 11. *If $(a,q)=1$, then*

$$|G(a,q)| \ll q^{1-1/k},$$

where the implied constant depends only on h.

5.1. Proof of (17). We treat the case of $Q(\delta) \ge \log N$ using Lemma 7, and the other case follows in a similar, slightly simpler fashion from Lemma 8. Since $\eta^{-\gamma} < Q(\delta)$, the hypotheses of Lemma 7 are certainly satisfied whenever $\alpha \in \mathbf{M}_q$ with $q \le \eta^{-\gamma}$. Therefore, Lemma 7 and Corollary 11,

combined with the bound

$$\Big|\int_1^M (1-\chi(r_d)(dx)^{\rho-1})e^{2\pi i h_d(x)\beta}\,dx\Big| \le M-\chi(r_d)(dM)^{\rho}/d\rho \ll \Psi \tag{21}$$

from Lemma 2 and the well-known facts $\phi(qd)\ge\phi(q)\phi(d)$ and

$$\phi(q)\ge c_\mu q^{1-\mu} \quad \text{for any } \mu>0,$$

yield

$$|S_M(\alpha)| \ll q^{-1/\gamma}\Psi + O(Me^{-5K^2c_1\sqrt{\log N}}).$$

Finally, the lower bound

$$\Psi \gg M/q_0 \ge Me^{-3Kc_1\sqrt{\log N}} \tag{22}$$

given by (8) and (7) ensures that the error term is negligible, and the estimate follows. ∎

For our minor arc estimate we need the following analog of Weyl's inequality, due to Li and Pan, which generalizes work of Vinogradov.

LEMMA 12. *Suppose that* $g(x)=a_0+a_1x+\cdots+a_kx^k\in\mathbb{Z}[x]$ *with* $a_k>0$, $D,W\in\mathbb{N}$, *and* $b\in\mathbb{Z}$. *If* $U\ge\log D$, $a_k\gg|a_{k-1}|+\cdots+|a_0|$, *and* $W,|b|,a_k\le U^k$, *then*

$$\sum_{\substack{x=1\\ Wx+b\in\mathcal{P}}}^{D}\log(Wx+b)e^{2\pi i g(x)\alpha} \ll \frac{D}{U}+U^CD^{1-c}$$

for some constants $C=C(k)$ *and* $c=c(k)>0$, *provided*

$$|\alpha-a/q|<q^{-2} \quad \text{for some } U^K\le q\le g(D)/U^K \text{ and } (a,q)=1.$$

Lemma 12 is a rougher, only nominally generalized version of Theorem 4.1 of [10]. That result restricts to the case where U is a power of $\log D$, and provides a more precise bound in place of K, but the main achievement of the theorem is that one can take U to be that small. Larger values of U, and hence stricter conditions on q, actually make the proof, which can be found in the appendix of that paper, slightly easier. Specifically, one can observe that the precise condition on q is not utilized until Lemmas 4.11 and 4.12, and adaptations of those lemmas are sufficient to adapt the proof of Theorem 4.1.

5.2. Proof of (18). Again, we only treat the case $Q(\delta)\ge\log N$. For a fixed $\alpha\in\mathfrak{m}$, by the pigeonhole principle there exist

$$1\le q\le L/(q_0Q(\delta)^2)^K$$

and $(a,q)=1$ with

$$|\alpha-a/q|<(q_0Q(\delta)^2)^K/(qL).$$

If $\eta^{-\gamma} \le q \le (q_0 Q(\delta)^2)^K$, then α meets the hypotheses of Lemma 7, and by reasoning identically to the proof of (17) we have

$$|S_M(\alpha)| \ll q^{-1/\gamma}\Psi \ll \eta\Psi.$$

If $(q_0 Q(\delta)^2)^K \le q \le L/(q_0 Q(\delta)^2)^K$, then we can apply Lemma 12 with $U = q_0 Q(\delta)^2$, along with (8) and the fact that $\eta > Q(\delta)^{-1}$, to conclude

$$|S_M(\alpha)| \ll \frac{M}{q_0 Q(\delta)^2} \ll \frac{\Psi}{Q(\delta)^2} < \eta\Psi,$$

as required.

If $1 \le q \le \eta^{-\gamma}$, then, letting $\beta = \alpha - a/q$, it must be the case that

$$|\beta| > \eta^{-\gamma}/L, \tag{23}$$

as otherwise we would have $\alpha \in \mathfrak{M}$. By Lemma 7 it suffices to show

$$\Big| \int_1^M (1 - \chi(r_d)(dx)^{\rho-1}) e^{2\pi i h_d(x)\beta}\, dx \Big| \ll \eta\Psi.$$

From (23) and Lemma 2.8 of [22], for any $x > 1$ we have

$$\Big| \int_1^x e^{2\pi i h_d(y)\beta}\, dy \Big| \ll (b_d|\beta|)^{-1/k} \ll \eta M,$$

hence by integration by parts, Lemma 2, and (8) we see that

$$\Big| \int_1^M (1 - \chi(r_d)(dx)^{\rho-1}) e^{2\pi i h_d(x)\beta}\, dx \Big|$$
$$\ll \eta(M - \chi(r_d)(dM)^\rho/d) \le \eta\bigg(M - \frac{\chi(r_d)(dM)^\rho}{d\rho} + 2(1-\rho)M \bigg) \ll \eta\Psi,$$

and the estimate is complete. ∎

Appendix A. Exponential sum estimates: proof of Lemma 7, Lemma 9, and Corollary 11

A.1. Proof of Lemma 7. Fixing $q \le (q_0 Q(\delta)^2)^K$ and $(a,q) = 1$, we first investigate the values of $S_X(a/q)$ for $X \ge N^{1/10k}$. We see that

$$S_X(a/q) = \sum_{x=1}^{X} \nu(x) e^{2\pi i h_d(x) a/q} = \sum_{\ell=o}^{q-1} e^{2\pi i h_d(\ell) a/q} \sum_{\substack{x=1 \\ x \equiv \ell \bmod q}}^{X} \nu(x), \tag{24}$$

and we note that

$$\sum_{\substack{x=1 \\ x \equiv \ell \bmod q}}^{X} \nu(x) = \phi(d)\psi(dX + r_d, r_d + d\ell, qd)/d. \tag{25}$$

Since $(r_d, d) = 1$, we see that $(r_d + d\ell, qd) = 1$ if and only if $(r_d + d\ell, q) = 1$. Therefore, if $(r_d+d\ell, q) > 1$, we have $\psi(dX+r_d, r_d+d\ell, qd) \le \log(dX+r_d) \ll \log X$, whereas if $(r_d + d\ell, q) = 1$, then (25) and Lemma 2 show that

$$\sum_{\substack{x=1\\ x\equiv \ell \bmod q}}^{X} \nu(x) = \frac{\phi(d)}{\phi(qd)}(X - \chi(r_d)(dX)^\rho/d\rho) + O(Xe^{-30kK^2c_1\sqrt{\log X}}). \tag{26}$$

Combining (24) and (26), we have

$$S_X(a/q) = \frac{\phi(d)}{\phi(qd)}G(a,q)(X - \chi(r_d)(dX)^\rho/d\rho) + O(qXe^{-30kK^2c_1\sqrt{\log X}})$$

for all $X \ge N^{1/10k}$. In particular, since $q \le e^{5K^2c_1\sqrt{\log N}}$ and $M \gg N^{1/4k}$, we can apply trivial bounds for small values of X and conclude

$$S_X(a/q) = \frac{\phi(d)}{\phi(qd)}G(a,q)(X - \chi(r_d)(dX)^\rho/d\rho) + O(Me^{-10K^2c_1\sqrt{\log N}}) \tag{27}$$

for all $X \le M$.

Now suppose $\alpha = a/q + \beta$ with $|\beta| < (q_0Q(\delta)^2)^K/L$. By (27) and successive applications of summation and integration by parts, we see

$$\begin{aligned}
S_M(\alpha) &= \sum_{x=1}^{M} \nu(x)e^{2\pi i h_d(x)a/q}e^{2\pi i h_d(x)\beta}\\
&= S_M(a/q)e^{2\pi i h_d(M)\beta} - \int_1^M S_x(a/q)2\pi i\beta h_d'(x)e^{2\pi i h_d(M)\beta}\,dx\\
&= S_M(a/q)e^{2\pi i h_d(M)\beta}\\
&\quad - \frac{\phi(d)}{\phi(qd)}G(a,q)\int_1^M \big(x - \chi(r_d)(dx)^\rho/d\rho\big)2\pi i\beta h_d'(x)e^{2\pi i h_d(x)\beta}\,dx\\
&\quad + O((1+\beta L)Me^{-10K^2c_1\sqrt{\log N}})\\
&= \frac{\phi(d)}{\phi(qd)}G(a,q)\int_1^M (1 - \chi(r_d)(dx)^{\rho-1})e^{2\pi i h_d(x)\beta}\,dx\\
&\quad + O(Me^{-5K^2c_1\sqrt{\log N}}),
\end{aligned}$$

and the asymptotic is established. ∎

A.2. Proof of Lemma 9. Fix g, W, b, a, q as in Lemma 9. We primarily make use of the well-known complete Gauss sum estimate

$$\Big|\sum_{\ell=0}^{q-1} e^{2\pi i g(\ell)a/q}\Big| \ll \gcd(\mathrm{cont}(g), q)^{1/k}q^{1-1/k}, \tag{28}$$

which can be found for example in Lemma 6 of [12]. As is often the case with this type of sum, we can simplify our argument by taking advantage of multiplicativity. Specifically, it is not difficult to show that if $q = q_1q_2$ with $(q_1, q_2) = 1$, then

$$\sum_{\substack{\ell=0\\(W\ell+b,q)=1}}^{q-1} e^{2\pi i g(\ell)a/q} = \Big(\sum_{\substack{\ell_1=0\\(W\ell_1+b,q_1)=1}}^{q_1-1} e^{2\pi i g(\ell_1)a_1/q_1}\Big)\Big(\sum_{\substack{\ell_2=0\\(W\ell_2+b,q_2)=1}}^{q_2-1} e^{2\pi i g(\ell_2)a_2/q_2}\Big),$$

where $a/q = a_1/q_1 + a_2/q_2$, so we can assume without loss of generality that $q = p^j$ for some $p \in \mathcal{P}$, $j \in \mathbb{N}$.

If $p \mid W$ and $p \mid b$, then $W\ell+b$ is never coprime to p^j, so the sum is clearly zero. If $p \mid W$ and $p \nmid b$, then $W\ell + b$ is always coprime to p^j, so the sum is complete and the result follows from (28). If $p \nmid W$, then $p \mid W\ell + b$ if and only if $\ell \equiv -bW^{-1} \bmod p$. Therefore,

$$\sum_{\substack{\ell=0\\p\nmid W\ell+b}}^{p^j-1} e^{2\pi i g(\ell)a/p^j} = \sum_{\ell=0}^{p^j-1} e^{2\pi i g(\ell)a/p^j} - \sum_{r=0}^{p^{j-1}-1} e^{2\pi i g(pr+m)a/p^j}, \tag{29}$$

where $m \equiv -bW^{-1} \bmod p$, and by (28) we need only obtain the estimate for the second sum.

Setting

$$\tilde{g}(r) = \frac{g(pr+m) - g(m)}{p},$$

we see that $\tilde{g}$ is a polynomial with integer coefficients and leading coefficient a_kp^{k-1}. In particular,

$$(\mathrm{cont}(\tilde{g}), p^{j-1}) \le p^{k-1}(a_k, p^{j-1}).$$

Therefore, by (28) we have

$$\Big|\sum_{r=0}^{p^{j-1}-1} e^{2\pi i g(pr+m)a/p^j}\Big| = \Big|\sum_{r=0}^{p^{j-1}-1} e^{2\pi i (g(pr+m)-g(m))a/p^j}\Big| = \Big|\sum_{r=0}^{p^{j-1}-1} e^{2\pi i \tilde{g}(r)a/p^{j-1}}\Big|$$
$$\ll (p^{k-1}(a_k, p^{j-1}))^{1/k} p^{(j-1)(1-1/k)} \le (a_k, p^j)^{1/k} p^{j(1-1/k)},$$

as required. ∎

A.3. Proof of Corollary 11. From its definition, we see that the leading coefficient of h_d is $d^k b/\lambda(d)$, where b is the leading coefficient of h. Given $q \in \mathbb{N}$ and $(a, q) = 1$, we write $q = q_1q_2$, where q_2 is the maximal divisor of q which is coprime to d. In particular,

$$(d^k b/\lambda(d), q_2) \le b. \tag{30}$$

Therefore, by Lemmas 9 and 10 and by (30) we have

$$|G(a,q)| = \Big| \sum_{\substack{\ell=0 \\ (r_d+d\ell,q)=1}}^{q-1} e^{2\pi i h_d(\ell)a/q} \Big| \ll \big((\mathrm{cont}(h_d), q_1)b\big)^{1/k} q^{1-1/k} \ll q^{1-1/k},$$

and all the required estimates are established. ∎

Appendix B. Theorems 2 and 3: an informal discussion. Using the result of Theorem 4, one can almost immediately conclude Theorem 3 by replicating the transference principle argument used in [8] to obtain Theorem F from a uniform version of Theorem C. Similarly, using weighted analogs of the major and minor arc estimates from this paper, one can almost immediately conclude Theorem 2 by reproducing the method of [6] used to prove Theorem D. In each instance, there are a few issues that arise and are addressed in this section, which is best read in conjunction with those two papers. First, we recall that for the arguments in [6] and [8], it is convenient, if not necessary, to do analysis with a discrete frequency domain, that is, to embed subsets of $[1, N]$ into the finite group $\mathbb{Z}/N\mathbb{Z}$ as opposed to the integers.

B.1. Higher moments of Weyl sums. To adapt the methods of [6] and [8], we need analogous estimates on higher moments of weighted and unweighted exponential sums over polynomials in primes. Specifically, if we borrow some notation from Section 2 and define

$$T(\alpha) = \frac{1}{\Psi_d} \sum_{x \in H_d} \nu_d(x) e^{2\pi i h_d(x)\alpha}, \qquad W(\alpha) = \frac{M_d}{\Psi_d N} \sum_{x \in H_d} \nu_d(x) h_d'(x) e^{2\pi i h_d(x)\alpha},$$

then it is straightforward to apply the major and minor arc estimates from this paper, weighted analogs thereof, and higher moment estimates on standard Weyl sums (see [24] for example) to conclude under appropriate conditions that

$$\sum_{t \in \mathbb{Z}/N\mathbb{Z}} |T(t/N)|^s = N \int_0^1 |T(\alpha)|^s \, d\alpha \ll 1,$$

$$\sum_{t \in \mathbb{Z}/N\mathbb{Z}} |W(t/N)|^s = N \int_0^1 |W(\alpha)|^s \, d\alpha \ll 1.$$

It is with these estimates in mind that we chose $s = 2^k + 6$, although something much smaller would suffice, and the above equalities follow from the dependence on s in the definition of H_d, as the relevant mod N congruences imply equality.

B.2. Applying Lemma 2 to Theorem 2. Because the method of [6] does not involve estimating the number of solutions to the desired equation, it suffices for the proof of Theorem 2 to use a simplified form of Lemma 2 in which $Q(\delta)$ is replaced with $e^{c_1\sqrt{\log N}}$ throughout. In order to obtain a usable analog to Lemma 1 of [6], we need to initially pass to a subprogression of step size $\lambda(q_0)$ and replace the condition $d \leq N^{.01}$ with $d \leq e^{c\sqrt{\log N}}$ for a sufficiently small constant c. This requires us to replace the L^2 concentration upper bound $\sigma^2(\log N)^{-1+\epsilon}$ with $\sigma^2(\log N)^{-1/2+\epsilon}$, which is the reason for the factor of 2 discrepancy between Theorem D and Theorem 2.

B.3. "Square root cancellation" in Theorem 2. The proof of Theorem D intimately uses the fact that for a quadratic polynomial, the normalized, weighted Weyl sum has "square root cancellation" on the major arcs. In our setting, we can apply weighted analogs of Lemmas 7 and 8 to conclude under appropriate conditions that if t/N is close to a rational a/q with $(a,q) = 1$, then

$$\begin{aligned} W(t/N) &\ll \frac{q^{1/2}}{\phi(q)} \min\{1, (N|t/N - a/q|)^{-1}\} \\ &\ll q^{-1/2} \log\log q \min\{1, (N|t/N - a/q|)^{-1}\}, \end{aligned}$$

where W is as in Section B.1 and the last inequality is a standard estimate on ϕ. While this is not quite as good as the estimate used in the proof of Lemma 2 of [6], the error of $\log\log q$ can easily be absorbed with other negligible terms as at the end of that proof (in fact $\log q$ would be fine as well). For a more detailed proof of Theorem 2, see [16].

B.4. Rephrasing Theorem 4 to deduce Theorem 3. Theorem 4 implies the following, less precise statement, which uses notation defined in Section 2 and is ready-made for applying a transference principle.

THEOREM 5. *If $h \in \mathbb{Z}[x]$ is a $\mathcal{P}$-intersective polynomial and $F : \mathbb{Z}/N\mathbb{Z} \to [0,1]$ with*

$$\frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} F(x) \geq \delta > 0,$$

then there exist constants $c(h,\delta) > 0$ and $N_0(h,\delta)$ such that

$$\frac{1}{NM_d} \sum_{\substack{x \in \mathbb{Z}/N\mathbb{Z} \\ y \in H_d}} F(x)F(x + h_d(y))\nu_d(y) \geq c(h,\delta)$$

provided $d \leq \log N$ and $N \geq N_0(h,\delta)$.

Once armed with Theorem 5 and the unweighted higher moment estimate from Section B.1, we derive Theorem 3 in the identical fashion that Theorem F follows from a uniform version of Theorem C, as in [8].

Acknowledgements. The author would like to thank his thesis advisor, Neil Lyall.

References

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

[2] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math. 31 (1977), 204–256.

[3] B. Green, *Roth's theorem in the primes*, Ann. of Math. 161 (2005), 1609–1636.

[4] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. 167 (2008), 481–547.

[5] B. Green and T. Tao, *Restriction theory of the Selberg sieve, with applications*, J. Théor. Nombres Bordeaux 18 (2006), 147–182.

[6] M. Hamel, N. Lyall and A. Rice, *Improved bounds on Sárközy's theorem for quadratic polynomials*, Int. Math. Res. Notices, to appear; arXiv:1111.5786v1.

[7] T. Kamae and M. Mendès France, *Van der Corput's difference theorem*, Israel J. Math. 31 (1978), 335–342.

[8] T. H. Lê, *Intersective polynomials and the primes*, J. Number Theory 130 (2010), 1705–1717.

[9] T. H. Lê, *Problems and results on intersective sets*, preprint.

[10] H.-Z. Li and H. Pan, *Difference sets and polynomials of prime variables*, Acta Arith. 138 (2009), 25–52.

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

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

[13] N. Lyall and Á. Magyar, *Polynomial configurations in difference sets*, J. Number Theory 129 (2009), 439–450.

[14] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge Stud. Adv. Math. 97, Cambridge Univ. Press, 2007.

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

[16] A. Rice, *Improvements and extensions of two theorems of Sárközy*, Ph.D. thesis, Univ. of Georgia, submitted July 2012.

[17] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104–109.

[18] I. Z. Ruzsa and T. Sanders, *Difference sets and the primes*, Acta Arith. 131 (2008), 281–301.

[19] A. Sárközy, *On difference sets of sequences of integers I*, Acta Math. Acad. Sci. Hungar. 31 (1978), 125–149.

[20] A. Sárközy, *On difference sets of sequences of integers III*, Acta Math. Acad. Sci. Hungar. 31 (1978), 355–386.

[21] S. Slijepčević, *A polynomial Sárközy–Furstenberg theorem with upper bounds*, Acta Math. Hungar. 98 (2003), 111–128.

[22] R. C. Vaughan, *The Hardy–Littlewood Method*, 2nd ed., Cambridge Univ. Press, 1997.

[23] M. Wierdl, *Almost everywhere convergence and recurrence along subsequences in ergodic theory*, Ph.D. thesis, Ohio State Univ., 1989.

[24] T. D. Wooley, *Some remarks on Vinogradov's mean value theorem and Tarry's problem*, Monatsh. Math. 122 (1996), 265–273.

Alex Rice
Department of Mathematics
University of Georgia
Athens, GA 30602, U.S.A.
E-mail: alex.rice@bucknell.edu

Current address:
Department of Mathematics
Bucknell University
Lewisburg, PA 17837, U.S.A.
E-mail: alex.rice@bucknell.edu

Received on 28.3.2012
and in revised form on 15.7.2012 (7017)