

IMPROVING AND MAXIMAL INEQUALITIES FOR PRIMES IN PROGRESSIONS

CHRISTINA GIANNITSI, MICHAEL T. LACEY, HAMED MOUSAVI,
AND YAGHOUB RAHIMI

ABSTRACT. Assume that $y < N$ are integers, and that $(b, y) = 1$. Define an average along the primes in a progression of spacing y , given by integer $(b, y) = 1$.

$$A_{N,y,b} := \frac{\phi(y)}{N} \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \Lambda(n) f(x - n)$$

Above, Λ is the von Mangoldt function and ϕ is the totient function. We establish improving and maximal inequalities for these averages. These bounds are uniform in the choice of progression. For instance, for $1 < r < \infty$ there is an integer $N_{y,r}$ so that for $(b, y) = 1$, we have

$$\| \sup_{N > N_{y,r}} |A_{N,y,b} f| \|_r \ll \|f\|_r.$$

The implied constant is only a function of r . The uniformity over progressions imposes several novel elements on the proof.

CONTENTS

1. Introduction	1
2. Preliminaries	3
3. Approximation	8
4. Properties of the High and Low Parts	15
4.1. Control of the Low Part	15
4.2. Properties of the High Term	17
5. Proof of the Main Inequalities	19
5.1. The Maximal Function Estimates	19
5.2. Fixed Scale Estimates	20
References	20

1. INTRODUCTION

We study averages over primes in arithmetic progressions, establishing inequalities with constants independent of the choice of progression. As far as we know, these are new. And the underlying proof entails some new complications, as compared to known results and their proofs. The averages we are concerned with are

MTL: The author is a 2020 Simons Fellow.

Research of all authors is supported in part by grant from the US National Science Foundation, DMS-1949206.

defined as follows. For $b, y \in \mathbb{N}$ and $y \leq N$, and function $f : \mathbb{Z} \rightarrow \mathbb{R}$, define

$$A_{N,y,b}f := \frac{\phi(y)}{N} \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \Lambda(n) f(x - n),$$

where Λ is the von Mangoldt function and ϕ is the Euler totient function. This is the average of f along the primes in the arithmetic progression $\{n : n \equiv b \pmod{y}\}$. We are only interested in the case of $(b, y) = 1$ of course, hence we use the totient function $\phi(y)$ above.

As our first result, we establish ℓ^r improving type inequalities.

Theorem 1.1. *For $r \in (1, 2)$, there exists $C_r > 0$, so that for all integers y , there is a $N_{r,y} > 0$ such that for all $N > N_{r,y}$ and compactly supported function f ,*

$$(1.2) \quad \max_{(b,y)=1} \|A_{N,y,b}f\|_{\ell^{r'}} \leq C_r \left(\frac{y}{N}\right)^{\frac{1}{r} - \frac{1}{r'}} \|f\|_{\ell^r}.$$

Above, we set $N_{r,y} = e^{C_{r,\delta} y^\delta}$, for any $\delta > 0$, and $C_{r,\delta}$ sufficiently large.

The right hand side of (1.2) is the correct scale factor for the inequality to hold uniformly in N . And, it is sharpest when f is assumed to be supported on a progression of spacing y . It is natural to suppose that N is sufficiently large, as a function of y . For the average over all primes, this inequality was established in [7], with study of the endpoint case in [11]. The novelty here is the uniformity in choice of arithmetic progression.

We also study the maximal inequality.

Theorem 1.3. *For $1 < r < \infty$, there is a constant C_r so that for all integers y , there is a $N_{r,y} > 0$ so that*

$$(1.4) \quad \left\| \sup_{N > N_{r,y}} |A_{N,y,b}f| \right\|_{\ell^r} \leq C_r \|f\|_{\ell^r}$$

The inequality above is uniform in y and $(b, y) = 1$.

We prove these theorems using the Siegel-Walfisz Theorem, and methods that are common to the study of these averages and their ℓ^r improving and sparse bounds. The bounds from the Siegel-Walfisz Theorem are ineffective. So, our bounds are also ineffective.

The uniformity over the progressions introduces important differences with prior papers studying averages over the primes. We describe them here.

The Hardy-Littlewood circle method is key. The decomposition of the Fourier transform of the averages leads to two competing sets of properties. The first, is the *height* of rational points in the circle. This property was identified by Bourgain [5], and refined by Ionescu and Wainger [8]. Its role is well understood.

This height is, for our purposes, dictated by the size of Gauss sums associated to the rational. Most commonly, this height is given by the denominator of the rational point in its lowest terms. In our setting, *these are decoupled*. Rational points whose denominator divides y all have Gauss sums of magnitude one. Specializing the discussion to the primes, the Gauss sum associated with rational a/q in lowest terms, is $\mu(q)/\phi(q)$. In our setting, the Gauss sums are given by a Ramanujan type sum along a progression. These are evaluated in Lemma 2.4. And, the height of a/q is given by $\text{lcm}(q, y)/y$. In particular, there are more than y rational points of height one.

This is a novel feature, and once identified, only adds a little extra difficulty to the proof of the improving inequality. In particular, the formulation of the Fourier multiplier approximation theorem, Theorem 3.14, is different from standard statements of this type. For the maximal inequality, however, one cannot use the standard approach. The latter approach uses the Bourgain Multifrequency Maximal Inequality [5]. It has a bound that is logarithmic in the number of rationals of a given height. And so, we cannot appeal to it. We use a different inequality at this point. See Lemma 4.12. Also note that the large number of points of height one would complicate applications of the Ionescu-Wainger theory, in seeking ℓ^p estimates. But we do not need to confront them, due to our approach to the improving inequalities.

The improving inequalities require a second property, call it a *Ramanujan height*. It depends upon subtle cancellation and size conditions on certain Ramanujan's sums. Again, there is a complication in evaluating these sums, and we need a progression variant of a familiar identity due to Cohen, see Lemma 2.7. Applying this identity is not so straightforward. An inverse Fourier transform calculation, easy in the case of the full sequence of primes, becomes much more involved. See Lemma 4.9.

In addition, one needs to know that Ramanujan's sums are typically of size one. This is quantified in a famous inequality due to Bourgain, stated in Lemma 2.10. Again, we need a progression version, stated in Lemma 2.11.

Bourgain [2, 3] initiated the study of these discrete averages, with the ℓ^2 result for the square integers being an important breakthrough. The first example of an arithmetic sequence for which the full ℓ^p inequalities were known is Wierdl's result for the primes [15]. See Mirek and Trojan [13] for a discussion of this proof. Averages along the primes, and closely related objects, have been studied by many, including variational results by [14], thin subsets of the primes [12], Carleson type theorems [6], and endpoint type results [11, 14]. This is the only paper we are aware of that discusses the uniformity over progressions.

The remainder of the paper begins with §2, where some notation and standard facts are collected. This section also has the crucial progression variants of some standard facts about Ramanujan's sums. These facts are probably known, but we could not find relevant sources to cite, so we include complete proofs for these facts. The remaining sections develop the tools along standard lines, while addressing the complications from the decoupling of the size of the Gauss sum at rational a/q and q mentioned above. The circle method is used to build approximation to the multipliers in §3. There are differences in the standard approaches here, accounting for the fact that the different role that height plays in this argument. See Definition 3.9. The following section §4 develops the properties of the High and Low decomposition of the multipliers. These definitions are not completely standard. The analysis of the Low part depends very much on the progression versions of the Ramanujan multipliers. The Bourgain Multifrequency Maximal Inequality cannot be used for the High part. The concluding section §5 is standard in nature.

2. PRELIMINARIES

For quantities a and b , we write $a \ll b$ if $|a| \leq Cb$ for some constant $C > 0$. We write $a \ll_p b$ if they implied constant depends on p .

For a function f on the integers, $\widehat{f}$ or $\mathcal{F}f$ denotes the discrete time Fourier transform of f , defined as

$$\mathcal{F}f(\theta) = \sum_{x \in \mathbb{Z}} f(x) e^{-2\pi i x \theta},$$

and $\check{f}$ or $\mathcal{F}^{-1}$ the inverse discrete time Fourier transform,

$$\mathcal{F}^{-1}f(x) = \int_0^1 \widehat{f}(\theta) e^{2\pi i x \theta} d\theta.$$

Finally, let $e(x) := e^{2\pi i x}$.

Let Ψ denote the Chebyshev function, which counts the primes in a progression.

$$\Psi(N, y, b) = \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \Lambda(n).$$

The fundamental estimate on it is given here, requiring that the average be sufficiently large, depending upon y and the ℓ^r index of the inequalities.

Theorem 2.1. *[Siegel-Walfisz Theorem] Let $J > 1$ be an integer. This holds for all $x > 1$, $y \leq (\log N)^J$ and $b \pmod{y}$.*

$$\left| \Psi(x, y, b) - \frac{x}{\phi(y)} \right| \leq C_J \cdot x \exp(-c_J \sqrt{\log x}),$$

where the constants C_J and c_J depend only on J .

Throughout, we denote $\mathbb{A}_q = \{a \in \mathbb{Z}/q\mathbb{Z} : (a, q) = 1\}$, so that $|\mathbb{A}_q| = \phi(q)$, the totient function. This lower bound on the totient function is well known. For all $0 < \epsilon < 1$, we have

$$\phi(q) \gg q^{1-\epsilon}.$$

We also make use of the major and minor arc decomposition. For integers $q, s \geq 1$ consider the following sets

$$\mathcal{R}_s = \left\{ \frac{a}{q} \in [0, 1) : a \in \mathbb{A}_q, \ 2^{s-1} \leq q < 2^s \right\}$$

For $0 < \varepsilon \leq 1/4$ and $\frac{a}{q} \in \mathcal{R}_s$, with $s \leq j\varepsilon$, we define the j -th major arc at a/q as

$$\mathfrak{M}_j(a/q) := \left(\frac{a}{q} - 2^{(\varepsilon-2)j}, \frac{a}{q} + 2^{(\varepsilon-2)j} \right),$$

which are disjoint for ε small enough. The j -th major arcs are given by $\mathfrak{M}_j := \bigcup_{\frac{a}{q} \in \mathcal{R}_s} \mathfrak{M}_j(\frac{a}{q})$. We define the j -th minor arcs $\mathfrak{m}_j$ as the complement of $\mathfrak{M}_j$.

We turn to exponential and Ramanujan's sums. Define Ramanujan's sums by

$$(2.2) \quad \tau_q(x) = \sum_{a \in \mathbb{A}_q} e(ax/q).$$

Cancellative properties of the Ramanujan's sums are very important for us, and expressed in different ways. The first of these is

$$\tau_q(x) = \mu(q) \quad (q, x) = 1.$$

Above, μ is the Möbius function, the multiplicative function with $\mu(p) = -1$ for all primes p , that vanishes on integers that are not square free. A second example of the cancellative properties is

$$\sum_{d|r} \tau_d(x) = \begin{cases} r & \text{if } r|x \\ 0 & \text{otherwise.} \end{cases}$$

We mention the next cancellation property known as Cohen's identity

$$(2.3) \quad \sum_{\substack{r < q \\ (r, q) = 1}} \tau_q(x + r) = \mu(q) \tau_q(-x).$$

Their relationship to the prime numbers are well known. In this study, we will need these properties, as well as certain progression versions of them.

Firstly, we examine Ramanujan's sum restricted to a progression. This formula must be known, but we were not able to find it in the literature.

Lemma 2.4. *Let $q, y, b \in \mathbb{N}$, with $g = \gcd(q, y)$, $(a, q) = 1$, $(b, g) = 1$. If $(g, q/g) = 1$, let $1 - g\bar{g} = \frac{a}{g}t$, where $\bar{g}$ is the multiplicative inverse of $g \bmod q/g$. Then,*

$$(2.5) \quad \sum_{\substack{r \in \mathbb{A}_q \\ r \equiv b \pmod{g}}} e\left(\frac{ra}{q}\right) = \begin{cases} 0 & \text{if } 1 < g < q \text{ and } (g, \frac{q}{g}) > 1 \\ \mu\left(\frac{q}{g}\right) e\left(\frac{abt}{g}\right) & \text{if } 1 \leq g < q \text{ and } (g, \frac{q}{g}) = 1 \\ e\left(\frac{ab}{q}\right) & \text{if } g = q. \end{cases}$$

Proof. The case of $g = 1, q$ are elementary, and we leave them to the reader. Below, we will assume that g is a proper divisor of q . Let u be a divisor of q . We have

$$(2.6) \quad \sum_{j=0}^{\frac{q}{u}-1} e(a(b + ju)/q) = \begin{cases} 0 & u < q \\ e(ab/q) & u = q \end{cases}$$

In the case of $u < q$, note that since $a \in \mathbb{A}_q$, we also have $a \in \mathbb{A}_{q/u}$, hence $j \rightarrow aj$ is a permutation on $\mathbb{Z}/(q/u)\mathbb{Z}$. And, if $u = q$, there is only a single term in the summation, so there is no cancellation.

For a set $B \subset \mathbb{Z}_q$, set

$$S(B) = \sum_{s \in B} e(sa/q).$$

We need to evaluate the term $S(A)$, where $A = \{r \in \mathbb{A}_q : r \equiv b \pmod{g}\}$. To do so, we use the Inclusion-Exclusion Principle to write $S(A)$ as a sum of progressions, as in (2.6).

Consider the set $T_u = \{b + ju : 0 \leq j < q/u\}$, and note that (2.6) is essentially an estimate of $S(T_g)$. Now, suppose g is a proper divisor of q . Then, for all prime factors p of g , we have

$$b + jg \equiv b \not\equiv 0 \pmod{p},$$

since $(b, g) = 1$. That is, if $r \in T_g \setminus A$, it must be divided by a prime factor of q that does not divide g .

Let U_g be all square free proper divisors of q that are relatively prime to g . If $U_g = \emptyset$, that means that g and q are powers of the same prime p . Therefore, $A = T_g$, since for every $p|q$, we conclude that $p|g$, and $b + jg \equiv b \not\equiv 0 \pmod{p}$. So $S(A) = S(T_g)$ and our desired estimate follows from (2.6).

On the other hand, if $U_g \neq \emptyset$, consider $u \in U_g$, and let r denote an integer $r = b + jg \in T_g$. We have

$$b + jg \equiv 0 \pmod{u} \quad \text{iff} \quad j \equiv -b\bar{g} \pmod{u}.$$

This holds since $(g, u) = 1$. Set $\beta_u = -b\bar{g} \pmod{u}$, (we may have $b \equiv 0 \pmod{u}$ for some $u \in U_g$.) and $\beta_g = b$. Let

$$R_u = \{rg + b \in \mathbb{Z}_q : r \equiv \beta_u \pmod{u}\},$$

and notice that we can then write A as

$$A = T_g \setminus \bigcup_{u \in U_g} R_u.$$

We can now utilize the Inclusion-Exclusion Principle. Let $\omega(n)$ be the number of distinct prime factors of n . Then

$$S(A) = S(T_g) + \sum_{u \in U} (-1)^{\omega(u)} S(R_u).$$

The equation above implies that the desired sum in (2.5) is a linear combination of other sums that can be expressed in the form of (2.6), and can be therefore estimated. Additionally, (2.6) forces a lot of the sums above to be zero. Specifically, all of them are zero except for when $gu = q$. In that case, the progression consists of a single term. This forces q/g to be square free, since U_g consists of square free integers. The corresponding coefficient from the Inclusion-Exclusion Principle is $(-1)^{\omega(q/g)} = \mu(q/g)$, which means that

$$S(A) = \mu(q/g) S(T_{q/g})$$

Recalling the definition of $\beta_{q/g}$, we see that $\beta_{q/g} = -b\bar{g} = \frac{-b}{g}(\frac{q}{g}t - 1)$. So $r = b + g\bar{g} = b + \frac{q}{g}t - b = \frac{q}{g}t$. The result follows from (2.6). $\square$

Secondly, we need a progression version of Cohen's identity (2.3).

Lemma 2.7. *We have for $g = \gcd(y, q)$*

$$(2.8) \quad \sum_{\substack{t \in A_q \\ t \equiv b \pmod{g}}} \tau_q(x + t) = \begin{cases} 0 & (g, \frac{q}{g}) > 1 \\ \mu(q/g) \tau_{q/g}(x) \tau_g(x + b) & (g, \frac{q}{g}) = 1 \end{cases}$$

Remark 2.9. Note that if $g = 1$, Lemma 2.7 reduces to the usual Cohen's identity. It is expected, because the progression on y and on q become independent. Also if $g = q$, we will get only the term $t = b$ from the sum in the left hand side of (2.8). This term is equal $\tau_q(x + b)$, which happens to be the right hand side.

Proof. The sum in question is

$$\sum_{\substack{t \in \mathbb{A}_q \\ t \equiv b \pmod{g}}} \sum_{r \in \mathbb{A}_q} e\left(\frac{(x+t)r}{q}\right) = \sum_{r \in \mathbb{A}_q} e\left(\frac{rx}{q}\right) \sum_{\substack{t \in \mathbb{A}_q \\ t \equiv b \pmod{g}}} e\left(\frac{tr}{q}\right)$$

By Lemma 2.4, the inner-most sum on the right hand-side is zero, when $\gcd(g, \frac{q}{g}) > 1$. Continuing with the assumption that $\gcd(g, \frac{q}{g}) = 1$, the sum above is equal to

$$\sum_{r \in A_q} e\left(\frac{r(x + bsq/g)}{q}\right) = \mu(q/g) \tau_q(x + bsq/g)$$

where $1 - g\bar{g} = \frac{q}{g}s$ if $g < q$ and $s = 1$ if $g = q$. Ramanujan's sums are multiplicative, leading to

$$\begin{aligned} \sum_{\substack{t \in A_q \\ t \equiv b \pmod{g}}} \tau_q(x+t) &= \mu(q/g) \tau_{q/g}(x + bsq/g) \tau_g(x + b(1 - g\bar{g})) \\ &= \mu(q/g) \tau_{q/g}(x) \tau_g(x + b). \end{aligned}$$

The last equality follows from the periodicity of Ramanujan's sum. $\square$

A final property of Ramanujan's sums is a fundamental inequality due to Bourgain [4]. It implies that typical values of $\tau_q(n)$ are approximately 1, on average.

Lemma 2.10. *Given integer k and $\epsilon > 0$, we have for all integers $M > yQ^k$*

$$\left[\frac{1}{M} \sum_{n \leq M} \left[\sum_{q \leq Q} |\tau_q(n)| \right]^k \right]^{1/k} \ll Q^{1+\epsilon}$$

The implied constant depends only on ϵ .

We need a progression version of this inequality.

Lemma 2.11. *Given integer t and $\epsilon > 0$, and integers b, y , with $(b, y) = 1$ we have for all integers $M > yQ^t$*

$$\left[\frac{y}{M} \sum_{\substack{n \leq M \\ n \equiv b \pmod{y}}} \left[\sum_{\substack{q \leq Q \\ (q, y) = 1}} |\tau_q(n)| \right]^t \right]^{1/t} \ll Q^{1+\epsilon}$$

The implied constant depends only on ϵ .

Notice that the length of the average is required to grow with t . That the constant is independent of t is not recorded as such in the literature, but follows from a modification of the proof in [9]. The ϵ dependence is traced to an inequality for the divisor function.

Proof. We follow the proof from [9, §3]. Firstly, we have $|\tau_q(n)| \leq (q, n)$. Secondly, for $\vec{q} \in [1, Q]^t$, let $\mathcal{L}(\vec{q})$ be the least common multiple of $q_1, \dots, q_t$. We assume throughout that all q_j are relatively prime to y . The map $m \mapsto \prod_{j=1}^t \tau_{q_j}(my + b)$ is periodic with period $\mathcal{L}(\vec{q})$. The condition $M > yQ^t$ then implies that for any $\vec{q} \in [1, Q]^t$,

$$\frac{y}{M} \sum_{\substack{n \leq M \\ n \equiv b \pmod{y}}} \prod_{j=1}^t (q_j, n) \ll \frac{1}{\mathcal{L}(\vec{q})} \sum_{n \leq \mathcal{L}(\vec{q})} \prod_{j=1}^t (q_j, ny + b).$$

On the right, we have dropped the modularity assumption on n .

Thirdly, we have, uniformly in $\vec{q} \in [1, Q]^t$, subject to the condition that q_j are coprime to y ,

$$\sum_{n \leq \mathcal{L}(\vec{q})} \prod_{j=1}^t (q_j, ny + b) \ll Q^{t+\epsilon}.$$

We establish this here. Due to the multiplicative structure of the estimate above, it suffices to consider this case. Consider the inequality below for prime $p \nmid y$, and integers t and $x_1 \geq x_2 \geq \dots \geq x_t$.

$$\sum_{n \leq p^{x_1}} \prod_{j=1}^t (p^{x_j}, ny + b) \ll p^{x_1 t + \epsilon}.$$

To see this, note that

$$\sum_{n \leq p^{x_1}} \prod_{j=1}^t (p^{x_j}, ny + b) \leq \prod_{j=1}^t p^{x_1 - x_j} \sum_{n \leq p^{x_j}} (p^{x_j}, ny + b)$$

For $w_j \leq x_j$, if $p^{w_j} \parallel n_i y + b$ for $i = 1, 2$ and $n_1 \neq n_2 \leq p^{x_j}$. Then, $p^{w_j} \mid (n_1 - n_2)$, since $p \nmid y$. That is, there are at most $p^{x_j - w_j}$ values of n such that $p^{w_j} \parallel ny + b$. It follows that

$$\begin{aligned} \sum_{n \leq p^{x_1}} \prod_{j=1}^t (p^{x_j}, ny + b) &\leq \prod_{j=1}^t p^{x_1 - x_j} \sum_{w_j \leq x_j} p^{x_j - w_j} p^{w_j} \\ &\ll p^{x_1 t} \prod_{j=1}^t \sum_{w_j \leq x_j} 1 \ll p^{x_1 t} \prod_{j=1}^t x_j \ll p^{x_1 t + \epsilon}. \end{aligned}$$

Fourth, we have the bound

$$\sum_{\vec{q} \in [1, Q]^t} \frac{1}{\mathcal{L}(\vec{q})} \ll Q^\epsilon.$$

Pulling together the different estimates gives us this chain of inequalities, which completes the proof.

$$\begin{aligned} \frac{y}{M} \sum_{\substack{n \leq M \\ n \equiv b \pmod{y}}} \left[\sum_{\substack{q \leq Q \\ (q, y) = 1}} |\tau_q(n)| \right]^t &\ll \frac{y}{M} \sum_{\substack{\vec{q} \in [1, Q]^t \\ (y, \mathcal{L}(\vec{q})) = 1}} \sum_{\substack{n \leq M \\ n \equiv b \pmod{y}}} \prod_{j=1}^t (q_j, yn + b) \\ &\ll \sum_{\substack{\vec{q} \in [1, Q]^t \\ (y, \mathcal{L}(\vec{q})) = 1}} \frac{1}{\mathcal{L}(\vec{q})} \sum_{n \leq \mathcal{L}(\vec{q})} \prod_{j=1}^t (q_j, yn + b) \\ &\ll \sum_{\vec{q} \in [1, Q]^t} \frac{Q^{t+\epsilon}}{\mathcal{L}(\vec{q})} \ll Q^{t+2\epsilon}. \end{aligned}$$

□

3. APPROXIMATION

Our strategy of proving the desired results consists of firstly approximating our kernel by another multiplier. We opt to do that on the Fourier side, and obtain an error that is easily controlled. This is established in Theorem 3.14. The next step is to take a closer look at the approximating multiplier and split it into two pieces, one that is well behaved on the time domain, and one that is well-behaved in the frequency domain. We call these pieces the Low and High parts and they

are thoroughly discussed in the next section. The principal result of this section is to prove Theorem 3.14, the approximation result for

$$\widehat{A}_{N,y,b}(\theta) = \frac{\phi(y)}{N} \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \Lambda(n) e(-n\theta).$$

This is the Fourier transform of our averaging kernel. The standard average over the integers from 1 to N is a multiplier with kernel

$$\widehat{M}_N(\theta) = \frac{1}{N} \sum_{n < N} e(-n\theta).$$

The progression version of the average over the integers congruent to $b \pmod{y}$, and less than N is denoted by $M_{N,y,b}$. As a Fourier multiplier, its kernel is

$$\begin{aligned} \widehat{M}_{N,y,b}(\theta) &= \frac{y}{N} \sum_{\substack{n \leq N \\ n \equiv b \pmod{y}}} e(-n\theta) \\ &= e(-b\theta) \frac{y}{N} \sum_{n \leq \frac{N-b}{y}} e(-ny\theta) \\ &= e(-b\theta) \widehat{M}_{\frac{N-b}{y}}(y\theta). \end{aligned}$$

We record an elementary relation between these two definitions.

$$(3.1) \quad \widehat{M}_{N,y,b}(\theta) = \widehat{M}_{\frac{N-b}{y}}(y\theta)(1 + O(b|\theta|))$$

Also note that because of the relative sizes of b and y , we always have $\frac{b}{y} < 1$. This means that there can only be at most one integer $n_0 \in [\frac{N-b}{y}, \frac{N}{y})$. Therefore

$$(3.2) \quad \widehat{M}_{\frac{N-b}{y}}(\theta) = \begin{cases} \frac{N}{N-b} \widehat{M}_{N/y}(\theta), & \text{if } [\frac{N-b}{y}, \frac{N}{y}) \cap \mathbb{Z} = \emptyset \\ \frac{N}{N-b} \widehat{M}_{N/y}(\theta) - \frac{ye(-n_0\theta)}{N-b}, & \text{if } [\frac{N-b}{y}, \frac{N}{y}) \cap \mathbb{Z} = \{n_0\} \end{cases}$$

Let $\|x\|$ denote the distance of real number x from its nearest integer. For the complete average, the estimate below is elementary.

$$\widehat{M}_N(\theta) = \frac{1}{N} \sum_{n \leq N} e(-n\theta) \ll \min \left(1, \frac{1}{N\|\theta\|} \right)$$

The progression version of this inequality is

$$\widehat{M}_{N,y,b}(\theta) \ll \min \left\{ 1, \frac{y}{N\|y\theta\|} \right\}.$$

Our primary focus is on the multiplier $\widehat{A}_{N,y,b}$. The first step in approximating it is taken here, where we focus our attention around the origin.

Lemma 3.3. *For all $J > 1$, there is a $0 < c < 1$ so that for $|\theta| < \frac{\log^J(N)}{N}$, and $y < \log^J(N)$, there holds for all $b \in \mathbb{A}_y$*

$$\widehat{A}_{N,y,b}(\theta) - \widehat{M}_{N/y}(y\theta) \ll \exp(-c\sqrt{\log N}).$$

Proof. We establish the closely related inequality

$$(3.4) \quad \widehat{A}_{N,y,b}(\theta) - \widehat{M}_{N,y,b}(\theta) \ll \exp(-c\sqrt{\log N}).$$

Then appeal to (3.1) and (3.2) to see the Lemma as written.

The left hand-side of (3.4) equals

$$\frac{\phi(y)}{N} \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \left[\Lambda(n) - \frac{y}{\phi(y)} \right] e(-n\theta).$$

We will use a trivial bound for $n \leq \sqrt{N}$. Apply the Siegel-Walfisz Theorem 2.1 and Abel summation to see that

$$\begin{aligned} & \frac{\phi(y)}{N} \sum_{\substack{n < N \\ n \equiv b \pmod{y}}} \left[\Lambda(n) - \frac{y}{\phi(y)} \right] e(-n\theta) \\ &= \frac{\phi(y)}{N} (\Psi(N, y, b) - N/\phi(y)) e(N\theta) \\ & \quad - \frac{\phi(y)}{N} (\Psi(\sqrt{N}, y, b) - \sqrt{N}/\phi(y)) e(\sqrt{N}\theta) \\ (3.5) \quad & \quad - 2\pi i \frac{\phi(y)}{N} \theta \int_{\sqrt{N}}^N [\Psi(t, y, b) - t/\phi(y)] e(-\theta t) dt + O(\sqrt{N}). \end{aligned}$$

Each term is at most $\exp(-c_J \sqrt{\log N})$. The integral is the one that uses the information on θ . We have

$$(3.5) \ll \frac{(\log N)^J}{N} \cdot N \exp(-c_J \sqrt{\log N}).$$

This is enough to finish the proof. $\square$

The approximation result on a so-called major arc is below. Recall that from their definition this concerns points in neighborhoods around rationals whose denominators have controlled magnitudes. The statement introduces the parameters $\ell = \text{lcm}(y, q)$ and $g = \text{gcd}(y, q)$ which play an important role in what follows. One should also note that the Gauss sum in (3.7) depends upon these parameters, and has itself a complicated expression. Nevertheless, it is explicitly evaluated in Lemma 2.4.

Lemma 3.6. *For all $J > 1$, there is a $0 < c < 1$ so that the following holds. For $y, q < \log^J(N)$, set $\ell = \ell_q = \text{lcm}(y, q)$, and $g = \text{gcd}(y, q)$. With $(a, q) = 1$, suppose that $|\xi - \frac{a}{q}| < \frac{\log^J(N)}{N}$. We have the inequality below.*

$$\widehat{A}_{N,y,b}(\xi) = \Upsilon(q, a) \widehat{M}_{N/\ell}(\ell(\xi - \frac{a}{q})) + O(\exp(-c_J \sqrt{\log N})),$$

where

$$(3.7) \quad \Upsilon(a, q) = \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r \in \mathbb{A}_q \\ r \equiv b \pmod{g}}} e(-ra/q).$$

Proof. The sum defining $\widehat{A}_{N,y,b}(\xi)$ is divided into residue classes mod q . Consider the conditions

$$n \equiv b \pmod{y}, \quad n \equiv r \pmod{g}.$$

If $g := \text{gcd}(y, q)$ and $b \not\equiv r \pmod{g}$, there is no solution. Otherwise, the conditions above are equivalent to $n \equiv \beta_r \pmod{\ell}$, where $\ell = \text{lcm}(q, y)$, for some choice of β_r . The choice of β_r can be made more explicit using a generalized Chinese Remainder Theorem, but that is not necessary for our purposes.

We will write $\xi = \frac{a}{q} + \theta$, where $|\theta| < \frac{\log^J(N)}{N}$. Observe that

$$\begin{aligned} \widehat{A}_{N,y,b}(\xi) &= \frac{\phi(y)}{N} \sum_{\substack{r < q \\ r \equiv b \pmod{g}}} \sum_{\substack{n < N \\ n \equiv b \pmod{y} \\ n \equiv r \pmod{q}}} \Lambda(n) e(-n(a/q + \theta)) \\ &= \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r < q \\ r \equiv b \pmod{g}}} e(-ra/q) \cdot \frac{\phi(\ell)}{N} \sum_{\substack{n < N \\ n \equiv \beta_r \pmod{\ell}}} \Lambda(n) e(-n\theta). \end{aligned}$$

Without loss of generality assume that n is a prime. If $\gcd(r, q) > 1$, then $n|q$. It gives the contribution of at most

$$\frac{\phi(y)q}{N} \sum_{n|q} \log(n) \ll \frac{\phi(y)q^{\epsilon+1}}{N}.$$

So we conclude that

$$\begin{aligned} \widehat{A}_{N,y,b}(\xi) &= \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r \in \mathbb{A}_q \\ r \equiv b \pmod{g}}} e(-ra/q) \cdot \frac{\phi(\ell)}{N} \sum_{\substack{n < N \\ n \equiv \beta_r \pmod{\ell}}} \Lambda(n) e(-n\theta) \\ &\quad + O(\exp(-c_J \sqrt{\log N})). \end{aligned}$$

By our hypotheses, Lemma 3.3 applies to the inner most sum, for each $r \in \mathbb{A}_q$ (with a different choice of J , that is larger by a square). It follows that

$$\widehat{A}_{N,y,b}(a/q + \theta) = \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r \in \mathbb{A}_q \\ r \equiv b \pmod{g}}} e\left(\frac{ra}{q}\right) \widehat{M}_{N/\ell}(\ell\theta) + O(\exp(-c_J \sqrt{\log N})).$$

That completes our proof. $\square$

In (3.7), the sum is a progression restricted Ramanujan's sum as in Lemma 2.4. Applying the latter, we have

Lemma 3.8. *We have this equality for $\Upsilon(a, q)$, defined in (3.7).*

$$\Upsilon(a, q) = \begin{cases} 0 & \text{if } 1 < g < q \text{ and } (g, \frac{a}{q}) > 1 \\ \frac{\phi(y)}{\phi(\ell)} \mu(q/g) e(-abt/g) & \text{if } 1 \leq g < q \text{ and } (g, \frac{a}{q}) = 1 \\ e(-ba/q) & \text{if } g = q. \end{cases}$$

This formula has implications for how the proof should be organized. Typically, one expects the Gauss sum at rational a/q to decay at a rate dictated by q . That is not the case here.

- (1) If $q \mid y$, then $g = q$, and $\Upsilon(a, q) = e(-ba/q)$. That is, there is no decay in the height of the Gauss sum. This is reflection of the fact our sum is restricted to a progression.
- (2) If $1 < g = \gcd(q, y) < q$, and $(g, q/g) = 1$, there is some decay in the Gauss sum, but only at the rate of g/q .
- (3) If $(q, y) = 1$, then $|\Upsilon(a, q)| = \phi(q)^{-1}$. These rational points act as if there is no progression.

In particular, there are more than $\phi(y)$ rational points $\frac{a}{q}$ with $|\Upsilon(a, q)| \simeq 1$. And, our estimates should be independent of y . This situation is rather different from most of the literature on this type of subject. This next definition is used to keep track of the relationship between the rational point and the value of the Gauss sum.

Definition 3.9. Define the *height (with respect to y)* of a rational a/q with $(a, q) = 1$, or an integer q to be

$$h_y(a/q) = h_y(q) = \begin{cases} 0 & \text{if } 1 < g < q \text{ and } (g, \frac{q}{g}) > 1 \\ \ell/y & \text{otherwise} \end{cases}$$

Here, and throughout, $g = \gcd(y, q)$ and $\ell = \text{lcm}(y, q)$. In particular, we have for any $\epsilon > 0$,

$$(3.10) \quad |\Upsilon(a, q)| \ll h_y(q)^{-1+\epsilon}, \quad \text{whenever } h_y(q) > 0.$$

We chose to refer to this height as the *Ramanujan height*. The “traditional” notion of height, as that term is frequently used in the related literature, is dictated, essentially, by the magnitude of the denominator. For our study, this is not good enough, as it does not take into consideration the restriction to a progression. There is again the dependence on the denominator q , which is indicated by the existence of the least common multiple in the formula, however notice that the part of q that actually contributes is the part that is co-prime with y . And the same applies to y as well.

Proof of (3.10). From (3.7), if $h_y(a/q) = 0$, then $\Upsilon(q, a)$ is also zero. Otherwise

$$|\Upsilon(a, q)| = \frac{\phi(y)}{\phi(\ell)} = \frac{\phi(y)}{\phi(yq/g)}$$

If $g = q$, the expression above is 1, so that (3.10) trivially holds. If $1 \leq g < q$, we have $\phi(yq/g) \geq \phi(y)\phi(q/g)$, so that (3.10) follows in this case as well. $\square$

It is important to observe that there are a potentially large number of rational points of a given height r . The exact number is

$$\begin{aligned} (3.11) \quad \#\{a/q : h_y(q) = r\} &= \sum_{\substack{q : \ell/y=r \\ (g,r)=1}} \phi(q) \\ &= \sum_{\substack{g|y \\ (g,r)=1}} \phi(gr) \\ &= \phi(r) \sum_{g|\frac{y}{(y,r)}} \phi(g) = \phi(r) \frac{y}{(y,r)}. \end{aligned}$$

Approaches to different aspects of this question are then limited by these bounds.

This notation is needed for the statement of our principal approximation result. For $0 \leq \xi < 1$, let $\ell := \text{lcm}(y, q)$ and

$$(3.12) \quad \widehat{L}_{N,y}^{a,q}(\xi) := \Upsilon(q, a) \widehat{M}_{N/\ell}(\ell(\xi - a/q)) \widehat{\eta}_{\ell^2}(\xi - a/q),$$

where η is a non-negative Schwartz function such that $\mathbf{1}_{[-1/16, 1/16]} \leq \eta \leq \mathbf{1}_{[-1/4, 1/4]}$, and

$$(3.13) \quad \widehat{\eta}_t(\xi) := \widehat{\eta}(t\xi).$$

One should not fail to note that *the cutoff function η above is scaled by ℓ^2* , to ensure that the major arcs remain disjoint, meaning the support of the multipliers is disjoint as well. The importance of this cutoff will also come into play in the next section when discussing the Low Part, as it allows us to use a multiplicative property of the spatial domain that is important for our estimates.

Theorem 3.14. *We have the estimate below, uniformly in $0 \leq \xi < 1$, uniformly for $b \in \mathbb{A}_y$,*

$$(3.15) \quad \widehat{A}_{N,y,b} = \sum_{q < N^{-1/10}} \sum_{a \in \mathbb{A}_q} \widehat{L}_{N,y}^{a,q}(\xi) + \widehat{E}_{N,y}(\xi),$$

where

$$(3.16) \quad \|\widehat{E}_{N,y}(\cdot)\|_{L^\infty} \ll \exp(-c' \sqrt{\log N}).$$

for a positive constant c' that depends on y .

Proof. Let $y^8 \simeq R = 2^r \simeq e^{c\sqrt{\log(N)}}$ for a sufficiently small choice of $c > 0$. Fix a choice of $\xi \in \mathbb{T}$. Using Dirichlet's Approximation Theorem, we can choose $0 \leq a < q < N^{1/10}$ with $a \in \mathbb{A}_q$, so that $|\xi - \frac{a}{q}| < \frac{1}{qN^{1/10}}$. The proof will be organized around the relative sizes of q and R .

We have this estimate for 'large' major arcs. For each fixed $s > r/2$, the quantity $\Upsilon(q, a)$ is at most $y2^{-s}$, which has small contribution. We also bound $\widehat{M}_{N/\ell} \ll 1$. So, using Lemma 2.4 we have

$$\begin{aligned} \sum_{s \geq r/2} \sum_{\frac{a}{q} \in \mathcal{R}_s} \widehat{L}_{N,y}^{a,q}(\xi) &\ll \sum_{s \geq r/2} \max_{2^{s-1} \leq q < 2^s} \left| \Upsilon(q, a) \widehat{M}_{N/\ell}(\ell(\xi - \frac{a}{q})) \right| \\ &\ll \sum_{s \geq r/2} \max_{2^{s-1} \leq q < 2^s} \frac{\phi(y)}{\phi(q)} \ll R^\epsilon \log(N) y R^{-1/2} \\ &\ll R^{-1/4}. \end{aligned}$$

The implication of this estimate is that we need not concern ourselves with this part of what will end up being the High term of our decomposition.

The remaining analysis is split according to the relative sizes of q and R . In the case of $q \geq R$, concerning the function $\widehat{A}_{N,y,b}$ we are in the setting of classical estimates of Vinogradov. The particular result we apply to in this setting is the main result of Balog and Perelli [1]. It gives us

$$\begin{aligned} \widehat{A}_{N,y,b}(\xi) &\ll \frac{y}{N} (NR^{-1/2} + \sqrt{RN} + R^{3/14} N^{5/7}) (\log N)^{18} \\ &\ll y R^{-1/2} (\log N)^{18} \ll R^{-1/3}, \end{aligned}$$

under our assumptions on R and y .

We establish a corresponding estimate for (the remaining part of) the High and Low terms. This will establish (3.16). Assume that $1 \leq s < r/2$. We know that $\ell'(\xi - \frac{a'}{q'})$ should be less than $1/4$ so that $\eta_{q'}(\xi - \frac{a'}{q'}) \neq 0$. So $\|\ell'(\xi - \frac{a'}{q'})\| = \ell'(\xi - \frac{a'}{q'})$. For $R < q < N^{1/10}$, one must note that for $\frac{a}{q} \neq \frac{a'}{q'} \in \mathcal{R}_s$ we have

$$\left| \xi - \frac{a'}{q'} \right| > \frac{1}{qq'} - \left| \xi - \frac{a}{q} \right| > \frac{1}{q'N^{1/10}} - \frac{1}{N^{1/10}q}$$

$$\begin{aligned}
&> \frac{1}{N^{1/10}}(2^{-s} - \frac{1}{R}) \\
&> \frac{1}{N^{1/10}}2^{-s-1}.
\end{aligned}$$

This implies that, for $\ell' = \text{lcm}(y, q')$, we get the upper bound

$$M_{N/\ell'}(\ell'(\xi - \frac{a'}{q'})) \ll \frac{\ell'}{N\|\ell'(\xi - \frac{a'}{q'})\|} \ll \frac{1}{N^{1/2}}.$$

To complete the proof, we now consider the following three cases, dictated by the relative sizes of q and R , as well as s and r .

Case 1: So if $q > R$, let $\ell' = \text{lcm}(q', y)$

$$\begin{aligned}
&\sum_{s \leq r/2} \sum_{\frac{a}{q} \neq \frac{a'}{q'} \in \mathcal{R}_s} \widehat{L}_{N,y}^{a',q'}(\xi) \\
&\ll \sum_{s < r/2} \sum_{\frac{a}{q} \neq \frac{a'}{q'} \in \mathcal{R}_s} \frac{\phi(y)}{\phi(\ell')} \left| \widehat{M}_{N/\ell'}(\ell'(\xi - \frac{a'}{q'})) \right| \eta_{q'}\left(\xi - \frac{a'}{q'}\right) \\
&\ll \sum_{s < r/2} \max_{q' \sim 2^s} \left| \widehat{M}_{N/\ell'}(\ell'(\xi - \frac{a'}{q'})) \right|.
\end{aligned}$$

There is a uniform bound, in s , on the number of a'/q' that contributes above. So

$$\sum_{s < r/2} \sum_{\frac{a}{q} \neq \frac{a'}{q'} \in \mathcal{R}_s} \widehat{L}_{N,y}^{a',q'}(\xi) \ll \sum_{s < r/2} \frac{1}{N^{1/2}} \simeq \frac{r}{N^{1/2}} \ll R^{-1}.$$

We conclude that both $\widehat{A}_N$ and the High and Low terms are small if $q > R$. This concludes the proof of (3.16) in this case.

Case 2: If $q < R$ and $s < r/2$ and $a/q \neq a'/q'$, then ξ and a'/q' are far apart. Namely,

$$\left| \xi - \frac{a'}{q'} \right| > \frac{1}{qq'} - \left| \xi - \frac{a}{q} \right| > \frac{1}{Rq'} - \frac{1}{N^{1/10}} > \frac{1}{2Rq'}.$$

This implies that

$$M_{N/\ell'}(\ell'(\xi - \frac{a'}{q'})) \ll \frac{\ell'}{N\|\ell'(\xi - \frac{a'}{q'})\|} \ll \frac{2^s R}{N}$$

Using this, we then have

$$\begin{aligned}
&\sum_{s < r/2} \sum_{\substack{\frac{a'}{q'} \in \mathcal{R}_s \\ \frac{a'}{q'} \neq \frac{a}{q}}} \widehat{L}_{N,y}^{a',q'}(\xi) \\
&\ll \sum_{s < r/2} \sum_{\frac{a'}{q'} \in \mathcal{R}_s} \frac{\phi(y)}{\phi(\ell')} \left| \widehat{M}_{N/\ell'}(\ell'(\xi - \frac{a'}{q'})) \right| \eta_s\left(\xi - \frac{a'}{q'}\right) \\
&\ll \sum_{s < r/2} \frac{2^s R}{N} \ll \frac{R^2}{N} < R^{-1}.
\end{aligned}$$

Case 3: Finally, when $q < R$ and $a/q = a'/q'$, Lemma 3.6 immediately implies that

$$\widehat{A}_{N,y,b}(\xi) = \widehat{L}_{N,y}^{a,q}(\xi) + O(R^{-1}),$$

and our proof is complete. $\square$

4. PROPERTIES OF THE HIGH AND LOW PARTS

We are now ready to define the High and Low part. Crucially, the definitions use the Ramanujan height, Definition 3.9. For an integer $Q < N^{1/10}$, with Q a power of 2, we set

$$(4.1) \quad \text{Hi}_{N,y,Q} = \sum_{s : Q \leq h_y(q) \leq N^{1/10}} \sum_{\frac{a}{q} \in \mathcal{R}_s} L_{N,y}^{a,q},$$

$$(4.2) \quad \text{Lo}_{N,y,Q} = \sum_{q : h_y(q) < Q} \sum_{\frac{a}{q} \in \mathcal{R}_s} L_{N,y}^{a,q}.$$

The terms $L_{N,y}^{a,q}$ are defined in (3.12). Again, the division into High and Low parts is done via the height function $h_y(q)$. The norm inequalities for these terms are as follows.

Lemma 4.3. *For all $\epsilon > 0$, $1 < r < 2$ and $1 \leq Q < (\log N)^{Cr'}$, and finite sets $F \subset \mathbb{Z}$ supported in $[0, N]$, we have that there exists $N_{r,y} > 0$ so that for $N > N_{r,y}$, r , the High term satisfies*

$$(4.4) \quad \|\text{Hi}_{N,y,Q} * \mathbf{1}_F\|_{\ell^2} \ll Q^{-1+\epsilon} |F|^{1/2},$$

$$(4.5) \quad \left\| \sup_{N=2^n > N_{2,y}} \|\text{Hi}_{N,y,Q} * \mathbf{1}_F\|_{\ell^2} \right\|_{\ell^2} \ll Q^{-1+\epsilon} |F|^{1/2},$$

and the Low term satisfies

$$(4.6) \quad \|\text{Lo}_{N,y,Q} * \mathbf{1}_F\|_{\ell^\infty} \ll Q^\epsilon (y/N)^{1/r} |F|^{1/r},$$

$$(4.7) \quad \left\| \sup_{N=2^n > N_{r,y}} \|\text{Lo}_{N,y,Q} * \mathbf{1}_F\|_{\ell^r} \right\|_{\ell^r} \ll Q^\epsilon |F|^{1/r}.$$

The power of y/N is needed to keep the estimate scale free. The constant $N_{r,y}$ is the same as in Theorem 1.3. The maximal inequalities (4.5) and (4.7) are $\ell^r \rightarrow \ell^r$, so the power of y/N is not needed. (And, they are sharpest when F is restricted to a progression of spacing y .)

4.1. Control of the Low Part. The estimates for the Low part are more challenging, so we address them first. Define

$$\widehat{\Phi}_{N,q}(\xi) = M_{N/\ell,\ell}(\ell\xi) \widehat{\eta}_{\ell^2}(\xi).$$

We record the elementary inequality for $\Phi_{N,q}$.

Proposition 4.8. *We have the estimate*

$$\Phi_{N,q}(x) \ll \eta_N(x).$$

Proof. Recall from (3.13) that η is a non-negative Schwartz function with $\mathbf{1}_{[-1/16, 1/16]} \leq \eta \leq \mathbf{1}_{[-1/4, 1/4]}$. The function η_{ℓ^2} then has spatial scale $\ell^2 \leq \sqrt{N}$, while $M_{N/\ell,\ell}$ is an average of length N/ℓ , along a progression of spacing ℓ . Then, the conclusion above is clear. $\square$

We invert the Fourier transform of the Low term. Experts will recognize that this step is typically routine, leading directly to Ramanujan's sums. In this instance, the proof is notably more complicated.

Lemma 4.9. *With the notation of (4.2), we have*

$$(4.10) \quad \text{Lo}_{N,y,Q}(x) \leq y \mathbf{1}_{y|x-b}(x) \sum_{\substack{q' < Q \\ (q',y)=1}} \frac{|\tau_{q'}(x)|}{\phi(q')} \eta_N(x).$$

Here, $\ell = \ell_q = \text{lcm}(q, y)$, and $\tau_q(x)$ is the Ramanujan function from (2.2).

Proof. For any q , we can calculate as follows.

$$\begin{aligned} & \mathcal{F}^{-1} \sum_{a \in \mathbb{A}_q} L_{N,y}^{a,q}(\xi - \frac{a}{q}) \\ &= \sum_{\substack{r \in A_q \\ r \equiv b \pmod{g}}} e(-ar/q) \sum_{a \in \mathbb{A}_q} \int_{\mathbb{T}} \widehat{M}_N(\xi - a/q) \eta_s(\xi - a/q) \frac{\phi(y)}{\phi(\ell)} e(x\xi) d\xi \\ &= \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r \in A_q \\ r \equiv b \pmod{g}}} \sum_{a \in \mathbb{A}_q} e(a(x-r)/q) \int_{\mathbb{T}} \widehat{M}_N(\theta) \eta_s(\theta) e(x\theta) d\theta \\ &= \Phi_{N,q}(x) \frac{\phi(y)}{\phi(\ell)} \sum_{\substack{r \in A_q \\ r \equiv b \pmod{g}}} \tau_q(x-r) \\ &= \Phi_{N,q}(x) \frac{\phi(y)}{\phi(\ell)} \mu(q/g) \mathbf{1}_{(g, \frac{q}{g})=1} \tau_{q/g}(x) \tau_g(x-b). \end{aligned}$$

A change of variables allows us to pull the sum over $a \in \mathbb{A}_q$ outside the integral. And, we use Lemma 2.7 in the last line.

Take $q' = h_y(q) = \ell/y = q/g$. As $(g, q/g) = 1$. Note that $\Phi_{N,q}$ is just a function of ℓ . It means that $\Phi_{N,q}$ does not depend on g and only depends on q' and y . We have

$$\text{Lo}_{N,y,Q}(x) = \sum_{q: h_y(q) < Q} \Phi_{N,q}(x) \frac{\phi(y)}{\phi(\ell)} \mu(q/g) \mathbf{1}_{(g, \frac{q}{g})=1} \tau_{q/g}(x) \tau_g(x-b).$$

Observe that $\gcd(q/g, g) = 1$, if and only if $\gcd(q/g, y) = 1$. This obvious but important property makes the condition $\mathbf{1}_{(q/g, g)=1}$ independent of g , and only depends on q', y . So

$$\begin{aligned} \text{Lo}_{N,y,Q}(x) &= \sum_{q' < Q} \sum_{\substack{g|y \\ (g, q')=1}} \Phi_{N,q'}(x) \frac{\phi(y)}{\phi(q')\phi(y)} \mu(q') \tau_{q'}(x) \tau_g(x-b) \\ &= \sum_{\substack{q' < Q \\ (y, q')=1}} \Phi_{N,q'}(x) \frac{\phi(y)}{\phi(q')\phi(y)} \mu(q') \tau_{q'}(x) \sum_{g|y} \tau_g(x-b). \end{aligned}$$

Next we use well-known Ramanujan's sum property

$$\sum_{q|r} \tau_q(n) = \begin{cases} r & r|n \\ 0 & \text{otherwise.} \end{cases}$$

Applying this property gives us

$$\begin{aligned} \text{Lo}_{N,y,Q}(x) &= \sum_{\substack{q' < Q \\ (y,q')=1}} \Phi_{N,q}(x) \frac{\mu(q')}{\phi(q')} \tau_{q'}(x) y \mathbf{1}_{y|x-b} \\ &\ll (y \mathbf{1}_{y|x-b}) \sum_{\substack{q' < Q \\ (q',y)=1}} \frac{|\tau_{q'}(x)|}{\phi(q')} \eta_N(x). \end{aligned}$$

Hence we have the result. $\square$

We address the fixed scale estimate (4.6) here. We appeal to details in this proof to prove the maximal estimate (4.7).

Proof of (4.6). We of course use (4.10), together with Hölder's inequality. That gives us

$$\begin{aligned} \text{Lo}_{N,y,Q} * \mathbf{1}_F(x) &= \sum_z \mathbf{1}_F(z) \text{Lo}_{N,y,Q}(x-z) \\ (4.11) \quad &\leq A_N(x) B_N(x), \end{aligned}$$

where

$$\begin{aligned} A_N(x)^{r'} &= y \sum_{x \equiv z+b \pmod y} \left[\sum_{\substack{q' < Q \\ (q',y)=1}} \frac{|\tau_{q'}(x-z)|}{\phi(q')} \right]^{r'} \eta_N(x-z) \\ B_N(x)^r &= y \sum_{x \equiv z+b \pmod y} \mathbf{1}_F(z) \eta_N(x-z). \end{aligned}$$

We have treated $y\eta_N(x-\cdot)$ as a measure, in our inequality above. The second term satisfies $\|B_N\|_\infty \ll [(y/N)|F|]^{1/r}$. The first term is controlled by Lemma 2.10. Recalling the familiar lower bound on the totient function $\phi(q) \gg q^{1-\epsilon}$, we see that $\|A_N\|_\infty \ll Q^\epsilon$. That completes the proof. $\square$

Proof of (4.7). We take advantage of (4.11) again, along with the fact that we always have $\|A_N\|_\infty \ll Q^\epsilon$ and

$$\left\| \sup_{N > N_{y,r}} B_N \right\|_{\ell^s} \ll |F|^{1/s}, \quad 1 < r < s < 2,$$

by the usual maximal function estimates. $\square$

4.2. Properties of the High Term. The first inequality is the fixed scale ℓ^2 estimate.

Proof of (4.4). This is entirely elementary. By Plancherel, it suffices to estimate

$$\begin{aligned} \|\widehat{\text{Hi}}_{N,y,Q}\|_{L^\infty} &\leq \sum_{s: 2^{s-1} \leq N^{1/10}} \left\| \sum_{\substack{q: 2^s \leq q < 2^{s+1} \\ h_y(q) > Q}} \sum_{a \in \mathbb{A}_q} \widehat{L}_{N,y}^{a,q} \right\|_{L^\infty} \\ &\ll \sum_{s: 2^{s+1} > Q} 2^{-s(1-\epsilon)} \\ &\ll Q^{-1+\epsilon}. \end{aligned}$$

We have taken care to define the functions $\{\widehat{L}_{N,y}^{a,q} : 2^s \leq q < 2^{s+1}\}$ so that they have disjoint support. That is done by inserting $\eta(\ell^2(\xi - a/q))$ into the definition of $L_{N,y}^{a,q}$ in (3.12). And the L^∞ norm of $\widehat{L}_{N,y}^{a,q}$ is at most $\phi(y)/\phi(\ell) \leq h_y(q)^{-1+\epsilon}$. $\square$

For the maximal function estimate (4.5), it is typical to apply the Bourgain Multifrequency Maximal Inequality from [5]. Also, in the typical setting, the height of the rationals and the number of rationals are coupled. In the current setting, this is no longer true. Following this path would result in an estimate that is logarithmic in y , because of the estimate (3.11).

Instead, we recall an inequality from [10, Lemma 2.1]. It requires the multi-frequency base points to share a common denominator, and the averages be over scales large relative to the common denominator. The constant in the maximal inequality is then independent of the number of base points.

Lemma 4.12. *Let $r_1, \dots, r_J$ be distinct rational points in $\mathbb{T}$, with common denominator $D < 2^d$. Then, we have*

$$(4.13) \quad \left\| \sup_{n > 2d} \left| \mathcal{F}^{-1} \left\{ \sum_{j=1}^J \widehat{\eta}(2^n(\theta - r_j)) \widehat{f}(\theta) \right\} (x) \right| \right\|_2 \ll \|f\|_2.$$

In our application of this lemma, the number of distinct rational points a/q with $2^s < h_y(q/q) \leq 2^{s+1}$ is at most $Cy^2 2^{2s}$.

Proof of (4.5). In the definition of the High term, we fix s with $2^s > Q/2$, and consider the maximal function formed over the kernels

$$\Gamma_{N,s} = \sum_{q : 2^s \leq h_y(q) < 2^{s+1}} \sum_{a \in \mathbb{A}_q} L_{N,y}^{a,q}.$$

The sum above is over at most $Cy^2 2^{2s}$ rational points. A denominator is qq' , where g divides y and $q' < 2^{s+1}$. Their common denominator is then at most $Cy^2 2^{2s}$. This means that we can apply (4.13) for the supremum over $N = 2^n > y^2 2^{2s}$.

Recall that we only consider $N > N_{y,2} = Cy^{2C}$, for a large absolute constant C . For values of $N_{y,2} \leq 2^n = N < Cy^2 2^{2s}$, turn to the fixed scale case, namely (4.4), to conclude that

$$\left\| \sup_{n : N_{y,2} \leq 2^n = N < Cy^2 2^{2s}} |\Gamma_{N,y} * f| \right\|_2 \ll 2^{-s(1-\epsilon)} \|f\|_2.$$

For the remaining supremum, the definition of $\Gamma_{N,s}$ needs a slight adjustment in order to apply (4.13). Define

$$\mathcal{F} \tilde{\Gamma}_{N,s} = \sum_{q : 2^s \leq h_y(q) < 2^{s+1}} \sum_{a \in \mathbb{A}_q} \Upsilon(q, a) \widehat{M}_N(\xi - a/q) \widehat{\eta}_{2^{2s}}(\xi - a/q).$$

Here, we have modified the definition of $L_{N,y}^{a,q}$ in (3.12) by replacing the average $\widehat{M}_{N/\ell}(\ell(\xi - a/q))$ by $\widehat{M}_N(\xi - a/q)$ and $\widehat{\eta}_{\ell^2}(\xi - a/q)$ by $\widehat{\eta}_{2^{2s}}(\xi - a/q)$. With this definition, by a square function argument, we have

$$\begin{aligned} \left\| \sup_{n : N = 2^n > Cy^2 2^{2s}} |(\Gamma_{N,s} - \tilde{\Gamma}_{N,s}) * f| \right\|_2^2 &\leq \sum_{n : N = 2^n > Cy^2 2^{2s}} \left\| (\Gamma_{N,s} - \tilde{\Gamma}_{N,s}) * f \right\|_2^2 \\ &\ll 2^{-2s(1-\epsilon)} \|f\|_2^2. \end{aligned}$$

And then, we have a direct application of (4.13) to control the supremum below.

$$\left\| \sup_{n : N=2^n > Cy2^{2s}} |\tilde{\Gamma}_{N,s} * f| \right\|_2 \ll 2^{-s(1-\epsilon)} \|f\|_2.$$

We conclude (4.5) by summing over s such that $2^s > Q$. $\square$

5. PROOF OF THE MAIN INEQUALITIES

5.1. The Maximal Function Estimates. We prove (1.4). To do so, it suffices to suppose that the function f on $\mathbb{Z}$ is the indicator of a set F . Indeed, we will prove a weak-type estimate for the maximal function. We need only consider the weak-type estimate at heights $0 < \lambda < 1$.

Fix $1 < r < 2$, and let $\epsilon = \frac{r-1}{4}$. Below, N will always be a power of 2. We trivially have

$$\sum_{N=2^n < 2^{\lambda^{-r+1}}} \|A_{N,y,b} \mathbf{1}_F\|_1 \ll \lambda^{-r+1} |F|.$$

So, we can restrict attention to $N > 2^{\lambda^{-r+1}}$. Importing this condition allows us to take advantage of the maximal inequalities (4.5) and (4.7), which means we can allow Q to be as large as

$$Q \leq (\log 2^{\lambda^{-r+1}})^{Cr'} = \lambda^{Cr}.$$

Take $Q \simeq \lambda^{-1+r/2}$. We will show that for $N_0 = \max\{N_{y,r}, 2^{\lambda^{-r+1}}\}$,

$$(5.1) \quad \begin{aligned} |\{ \sup_{N > N_0} A_{N,y,b} \mathbf{1}_F > \lambda \}| &\ll \{Q^\epsilon \lambda^{-r} + Q^{-2+\epsilon} \lambda^{-2}\} |F| \\ &\ll \lambda^{-r+\epsilon(1-r/2)} |F| \\ &\ll \lambda^{-r+\epsilon} |F|. \end{aligned}$$

This proves the restricted weak type estimate $\ell^{s,1} \rightarrow \ell^{s,\infty}$, where $s = r - \epsilon$. As r decreases to one, so does s . We deduce the restricted weak type inequality for all $1 < r < 2$. Interpolation completes the argument.

Recall our approximation (3.15). Use the value of Q above in the definition of the High and Low terms in (4.1) and (4.2), respectively. Then,

$$A_{N,y} = \text{Lo}_{N,y,Q} + \text{Hi}_{N,y,Q} + E_{N,y,Q}.$$

Then, by (4.7), we have

$$|\{ \sup_{N > N_0} |\text{Lo}_{N,y,Q} * \mathbf{1}_F| > \lambda/3 \}| \ll Q^\epsilon \lambda^{-r} |F|.$$

This is the first half of (5.1). The estimate below matches the second half of (5.1), and it follows from (4.5).

$$|\{ \sup_{N > N_0} |\text{Hi}_{N,y,Q} * \mathbf{1}_F| > \lambda/3 \}| \ll Q^{-2+\epsilon} \lambda^{-2} |F|.$$

Last of all, recalling (3.16), we have

$$\begin{aligned} |\{ \sup_{N > N_0} |E_{N,y} * \mathbf{1}_F| > \lambda/3 \}| &\ll \lambda^{-2} \sum_{N > N_0} \|E_{N,y} * \mathbf{1}_F\|_2^2 \\ &\ll \lambda^{-2} \exp(-c' \lambda^{-(r-1)/2}) |F|. \end{aligned}$$

This is better than the second half of (5.1). So, it completes the proof.

5.2. Fixed Scale Estimates. We prove the estimate (1.1). By duality, that estimate is the same as

$$(5.2) \quad \frac{y}{N} \langle A_{N,y,b} \mathbf{1}_F, g \rangle \ll \left(\frac{y}{N} |F| \right)^{1/r} \left(\frac{y}{N} |G| \right)^{1/r},$$

where F and G are subsets of an interval E of length N .

Observe that trivially

$$\frac{y}{N} \langle A_{N,y,b} \mathbf{1}_F, g \rangle \ll \log N \frac{y}{N} |F| \cdot \frac{y}{N} |G|.$$

This implies that the inequality (5.2) is true, unless

$$(5.3) \quad \frac{y^2}{N^2} |F| \cdot |G| \ll (\log N)^{-r'}.$$

So it suffices to only study this case.

We take $N > N_{r,y}$, and $0 < \epsilon < \frac{r-1}{100}$ small, and apply the High/Low decomposition with parameter Q to be determined later. Using the estimates (4.4) and (4.6), we have

$$\begin{aligned} \frac{y}{N} \langle \text{Hi}_{N,y,Q} * \mathbf{1}_F, g \rangle &\ll Q^{-1+\epsilon} \left(\frac{y}{N} |F| \right)^{1/2} \left(\frac{y}{N} |G| \right)^{1/2}, \\ \frac{y}{N} \langle \text{Lo}_{N,y,Q} * \mathbf{1}_F, g \rangle &\ll Q^\epsilon \left(\frac{y}{N} |F| \right)^{1/r} \frac{y}{N} |G|. \end{aligned}$$

Optimize over Q so that the right hand sides above are approximately equal. We obtain

$$Q^{1+2\epsilon} \simeq \left(\frac{y}{N} |F| \right)^{\frac{1}{2}-\frac{1}{r}} \left(\frac{y}{N} |G| \right)^{-\frac{1}{2}}.$$

By (5.3), this is an allowed choice for us.

So our estimate becomes

$$\frac{y}{N} \langle A_{N,y,b} \mathbf{1}_F, g \rangle \ll \left(\frac{y}{N} |F| \right)^{\frac{1}{r}+\epsilon'} \left(\frac{y}{N} |G| \right)^{\frac{1}{r}+\epsilon'}.$$

Above $\epsilon' < c_r \epsilon$. Thus, we see that (5.2) holds for $1 < r < 2$.

Remark 5.4. The estimate above could be improved to a sparse bound for the maximal function. However, the notion of a sparse bound would have to be refined to one that is adapted to progressions. Not having a ready application of such a result, we do not pursue the details herein.

REFERENCES

- [1] Antal Balog and Alberto Perelli, *Exponential sums over primes in an arithmetic progression*, Proc. Amer. Math. Soc. **93** (1985), no. 4, 578–582. MR776182
- [2] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. **61** (1988), no. 1, 39–72. MR937581
- [3] ———, *On the pointwise ergodic theorem on L^p for arithmetic sets*, Israel J. Math. **61** (1988), no. 1, 73–84. MR937582
- [4] ———, *Fourier transform restriction phenomena for certain lattice subsets and applications to non-linear evolution equations. Part I Schrödinger equations*, Geom. Funct. Anal. **3**, No. 2 (1993), 107–156 (en). MR1209299
- [5] Jean Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 5–45. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. MR1019960
- [6] Laura Cladek, Kevin Henriot, Ben Krause, Izabella Łaba, and Malabika Pramanik, *A discrete Carleson theorem along the primes with a restricted supremum*, Math. Z. **289** (2018), no. 3–4, 1033–1057. MR3830238

- [7] Rui Han, Ben Krause, Michael T. Lacey, and Fan Yang, *Averages along the primes: Improving and sparse bounds*, Concrete Operators **7** (01 Jan. 2020), no. 1, 45–54.
- [8] Alexandru D. Ionescu and Stephen Wainger, *L^p boundedness of discrete singular Radon transforms*, J. Amer. Math. Soc. **19** (2006), no. 2, 357–383. MR2188130
- [9] Robert Kesler, Micheal Lacey, and Dario Mena, *Lacunary discrete spherical maximal functions*, New York Journal of Mathematics **25** (2019), 541–557.
- [10] Michael Lacey, *On an inequality due to Bourgain*, Illinois J. Math. **41** (1997), no. 2, 231–236. MR1441675
- [11] Michael T. Lacey, Hamed Mousavi, and Yaghoub Rahimi, *Endpoint ℓ^r improving estimates for prime averages*, Math Research Letters, to appear (2021), available at [2101.10401](#).
- [12] Mariusz Mirek, *$\ell^p(\mathbb{Z})$ -boundedness of discrete maximal functions along thin subsets of primes and pointwise ergodic theorems*, Math. Z. **279** (2015), no. 1-2, 27–59. MR3299842
- [13] Mariusz Mirek and Bartosz Trojan, *Cotlar’s ergodic theorem along the prime numbers*, J. Fourier Anal. Appl. **21** (2015), no. 4, 822–848. MR3370012
- [14] Bartosz Trojan, *Endpoint estimates for the maximal function over prime numbers*, J. Fourier Anal. Appl. **25** (2019), no. 6, 3123–3153. MR4029173
- [15] Máté Wierdl, *Pointwise ergodic theorem along the prime numbers*, Israel J. Math. **64** (1988), no. 3, 315–336 (1989). MR995574

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
Email address: cgiannitsi3@math.gatech.edu

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
Email address: lacey@math.gatech.edu

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
Email address: hmousavi6@gatech.edu

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332, USA
Email address: yaghoub.rahimi@gatech.edu