1-1-2020

Averages along the Primes: Improving and Sparse Bounds

Rui Han
Georgia Institute of Technology

Ben Krause
Princeton University

Michael T. Lacey
Georgia Institute of Technology

Fan Yang
Georgia Institute of Technology

Recommended Citation

Han, R., Krause, B., Lacey, M., & Yang, F. (2020). Averages along the Primes: Improving and Sparse Bounds. *Concrete Operators, 7* (1), 45-54. https://doi.org/10.1515/conop-2020-0003

DE GRUYTER Concr. Oper. 2020; 7:45–54

Research Article Open Access

Rui Han, Ben Krause, Michael T. Lacey*, and Fan Yang

Averages Along the Primes: Improving and Sparse Bounds

https://doi.org/10.1515/conop-2020-0003
Received September 6, 2019; accepted December 9, 2019

Abstract: Consider averages along the prime integers $\mathbb{P}$ given by

$$\mathcal{A}_N f(x) = N^{-1} \sum_{p \in \mathbb{P}\,:\, p \le N} (\log p) f(x-p).$$

These averages satisfy a uniform scale-free ℓ^p-improving estimate. For all $1 < p < 2$, there is a constant C_p so that for all integer N and functions f supported on $[0, N]$, there holds

$$N^{-1/p'} \|\mathcal{A}_N f\|_{\ell^{p'}} \le C_p N^{-1/p} \|f\|_{\ell^p}.$$

The maximal function $\mathcal{A}^\star f = \sup_N |\mathcal{A}_N f|$ satisfies (p, p) sparse bounds for all $1 < p < 2$. The latter are the natural variants of the scale-free bounds. As a corollary, $\mathcal{A}^\star$ is bounded on $\ell^p(w)$, for all weights w in the Muckenhoupt A_p class. No prior weighted inequalities for $\mathcal{A}^\star$ were known.

Keywords: primes, circle method, improving, sparse bounds, maximal function

MSC: Primary: 42A45, 42B25 Secondary: 11L05

1 Introduction

Let $\mathbb{P} = \{3, 5, 7, \dots, \}$ be the odd primes and define the logarithmically weighted averages along the primes by

$$\mathcal{A}_N f(x) = N^{-1} \sum_{p \in \mathbb{P}\,:\, p \le N} (\log p) f(x-p),$$

We prove scale-free ℓ^p improving bounds for these averages, and sparse bounds for the associated maximal function

$$\mathcal{A}^\star f = \sup_N |\mathcal{A}_N f|. \tag{1.1}$$

For a function f on $\mathbb{Z}$, and an interval $I \subset \mathbb{Z}$, define

$$\langle f \rangle_{I,p} := \left(\frac{1}{|I|} \sum_{x \in I} |f(x)|^p \right)^{1/p} \tag{1.2}$$

to be the normalized ℓ^p norm on I. Throughout the paper, if $I = [a, b] \cap \mathbb{Z}$, with $a, b \in \mathbb{Z}$, is an interval on $\mathbb{Z}$, let $2I = [2a - b - 1, b] \cap \mathbb{Z}$ be the doubled interval (on the left-hand side), let $3I = [2a - b - 1, 2b - a + 1]$ be the tripled interval which has the same center as I.

We prove that the averages along the primes improve integrability, uniformly over all scales.

Rui Han: Georgia Institute of Technology, E-mail: rui.han@math.gatech.edu
Ben Krause: Princeton University, E-mail: bkrause@princeton.edu
***Corresponding Author: Michael T. Lacey:** Georgia Institute of Technology, E-mail: lacey@math.gatech.edu
Fan Yang: Georgia Institute of Technology, E-mail: fan.yang@math.gatech.edu

Theorem 1.1. *For $1 < p < \infty$, there is a constant C_p so that for all integers N, and interval I of length N, there holds for all functions f,*

$$\langle \mathcal{A}_N f \rangle_{I,p'} \le C_p \langle f \rangle_{2I,p}, \tag{1.3}$$

where $p' = \frac{p}{p-1}$.

We turn to the sparse inequalities. They are the natural extensions of the ℓ^p improving inequalities above for the maximal function (1.1). We say that a sublinear operator B has sparse type (r, s), for $1 < r, s < \infty$ if there is a constant C so that for all finitely supported functions f, g there are a sparse collection of intervals $\mathcal{S}$ so that

$$|(Bf, g)| \le C \sum_{I \in \mathcal{S}} \langle f \rangle_{2I,r} \langle g \rangle_{I,s} |I|, \tag{1.4}$$

where (f, g) is the standard inner product on $\ell^2(\mathbb{Z})$. A collection of intervals $\mathcal{S}$ is said to be *sparse* if there are subsets $E_I \subset I$ for $I \in \mathcal{S}$ which are pairwise disjoint, and satisfy $|E_I| > \frac{1}{10}|I|$.

Theorem 1.2. *The maximal operator $\mathcal{A}^\star$ is of sparse type (r, s), for all $1 < r, s < 2$.*

This statement is much stronger than just asserting that $\mathcal{A}^\star$ is bounded on ℓ^p, for all $1 < p < \infty$. It implies for instance these weighted inequalities, which match the classical result of Muckenhoupt for the ordinary maximal function. (Although the quantitative estimates of the norm will not match.)

Corollary 1.3. *For any $1 < p < \infty$, and any weight w in the Muckenhoupt class A_p, we have that $\mathcal{A}^\star$ is a bounded operator on $\ell^p(w)$.*

We remark that for the simple averages along the primes, one can check that for non-negative f

$$\sup_N \frac{\log N}{N} \sum_{p \in \mathbb{P} \,:\, p \le N} f(x - p) \lesssim \mathcal{A}^\star f.$$

Therefore, the sparse bounds hold for the maximal function on the left. Our argument for the fixed scale inequalities (1.3) requires the logarithmic averages.

Following Bourgain's work on arithmetic ergodic theorems 1], Wierdl 20] showed that $\mathcal{A}^\star$ is bounded on ℓ^p for all $1 < p < \infty$. At the time, this was the first arithmetic example for which this fact was known for all $1 < p < 2$. Bourgain's work 3] gave a comprehensive approach to the ℓ^p theory of arithmetic averages. The subject continues to be under development, with important contributions by 8, 15, 16]. We point to the work of Mirek-Trojan and Trojan 17, 18] also focused on the primes. The methods therein are different from those of this paper.

Our subject, developing the ℓ^p-improving properties and sparse bounds started with 4], and continued in 6, 14]. It now encompasses the discrete spherical maximal operators 9, 10, 12], as well as the square integers 5].

We use the High Low Method 5, 7, 11]. This depends upon efficient use of ℓ^2-methods, followed by a fine analysis of certain ℓ^1-type expressions. The latter are frequently the most intricate part. In this argument, they depend upon a relatively accessible property of Ramanujan sums, Lemma 3.4. Our argument is new, even if one is only interested in the $\ell^p \to \ell^p$ bounds for $\mathcal{A}^\star$.

2 Preliminaries

Throughout, let $\phi(q)$ be the Euler totient function, let $\mu(q)$ be the Möbius function. The following estimate for $\phi(q)$ is well known:

$$\phi(q) \gtrsim_\varepsilon q^{1-\varepsilon}. \tag{2.1}$$

We count primes in the standard logarithmic fashion. Put

$$\vartheta(N) = \sum_{p\in\mathbb{P}\,:\,p\le N} \log P. \tag{2.2}$$

By the prime number theorem

$$\left|\frac{\vartheta(N)-N}{N}\right| \le Ce^{-c\sqrt{\log N}}, \tag{2.3}$$

holds for some constant $c, C > 0$. This obviously implies $\vartheta(N) \sim N$.

We now redefine the averaging operators A_N, by setting

$$\mathcal{A}_N f(x) = \vartheta(N)^{-1} \sum_{p\in\mathbb{P}\,:\,p\le N} (\log p) f(x-p) \tag{2.4}$$

As this is a positive operator, there is no harm in this new definition.

The Fourier transform of a measure σ on $\mathbb{Z}$ is given by

$$\widehat{\sigma}(\xi) = \sum_{x\in\mathbb{Z}} \sigma(x)e(x\xi),$$

where $e(\zeta) = e^{2\pi i\zeta}$ throughout. The inverse Fourier transform is denoted $\check{\eta}$. Occasionally, we may also denote the Fourier transform by $\mathcal{F}$, and inverse Fourier transform by $\mathcal{F}^{-1}$.

We further set $e_q(\zeta) = e^{2\pi i\zeta/q}$. Recall that Ramanujan sums are defined by

$$c_q(n) = \sum_{a\in\mathbb{A}_q} e_q(an/q), \tag{2.5}$$

where $\mathbb{A}_q = \{1 \le a < q \ : \ (a,q) = 1\}$ is the multiplicative group associated to q. Define by convention that $c_1(n) \equiv 1$.

A finer property of Ramanujan sums is recalled in Lemma 3.4 below.

3 Approximating Multipliers

We define the approximating multipliers. Let $\mathbf{1}_{[-1/8,1/8]} \le \eta \le \mathbf{1}_{[-1/4,1/4]}$ be a Schwartz function. For an integer s, let $\eta_s(\xi) = \eta(8^s\xi)$. Define the Fourier transform of the usual averages by

$$\widehat{\gamma_N} = \frac{1}{N}\sum_{n=1}^{N} \widehat{\delta_n}. \tag{3.1}$$

The building blocks of the approximating multipliers are

$$\begin{aligned} \widehat{L_{1,N}}(\xi) &= \widehat{\gamma_N}(\xi)\eta_1(\xi) \\ \widehat{L_{q,N}}(\xi) &= \frac{\mu(q)}{\phi(q)} \sum_{a\in\mathbb{A}_q} \widehat{\gamma_N}\cdot\eta_s(\xi - a/q), \qquad 2^s \le q < 2^{s+1},\ \ s\ge 1. \end{aligned} \tag{3.2}$$

Throughout, q and s have the relationship above, although this will be suppressed in the notation. (This is a useful convention in the application of the multi-frequency maximal function inequality in the proof of the sparse bounds, see (5.7).)

Theorem 3.1. *Let $A, N > 10$ be integers. If $K \lesssim (\log N)^A$, there holds*

$$\widehat{\mathcal{A}_N} = \sum_{1\le q\le K} \widehat{L_{q,N}} + r_{A,N,K}, \tag{3.3}$$

where $\|r_{A,N,K}\|_{L^\infty} \lesssim_A K^{-1+1/A}$.

This is a consequence of standard facts in the number theory literature, and is very similar to how these facts are used in 20]. We recall them here.

Lemma 3.2. *For positive integers B, there is an integer N_B so that for all $N > N_B$*

1. *If $|\xi| < \frac{(\log N)^B}{N}$, then*
$$\widehat{\mathcal{A}_N}(\xi) = \widehat{\gamma_N}(\xi) + O(e^{-c\sqrt{\log N}}). \tag{3.4}$$
2. *If $|\xi - a/q| < \frac{(\log N)^B}{N}$ for $(a, q) = 1$ and $1 < q < (\log N)^B$, then*
$$\widehat{\mathcal{A}_N}(\xi) = \frac{\mu(q)}{\phi(q)}\widehat{\gamma_N}(\xi - \frac{a}{q}) + O(e^{-c\sqrt{\log N}}). \tag{3.5}$$
3. *If ξ does not meet any of the hypotheses of the prior two conditions, then*
$$\widehat{\mathcal{A}_N}(\xi) = O\left((\log N)^{4-\frac{B}{2}}\right). \tag{3.6}$$
4. *The following holds for $|\xi| \le 1/2$*
$$|\widehat{\gamma_N}(\xi)| \lesssim \min\{1, (N|\xi|)^{-1}\}. \tag{3.7}$$

The points (1), (2) and (3) above are in 19, Lemma 3.1 & Thm. 3.1], while the last point is well known.

Proof of Theorem 3.1. We note that by construction, the multipliers $\{\widehat{L_{q,N}} \;:\; 2^s \le q < 2^{s+1}\}$ are supported on disjoint intervals around the rationals a/q, with $a \in \mathbb{A}_q$, and $2^s \le q < 2^{s+1}$. From this, it follows from (2.1) that
$$\Big\| \sum_{2^s \le q < 2^{s+1}} \widehat{L_{q,N}} \Big\|_{L^\infty} \le \max_{2^s \le q < 2^{s+1}} \phi(q)^{-1} \lesssim 2^{-s(1-1/A)}. \tag{3.8}$$
Above, A is the integer in Theorem 3.1.

It suffices to argue that for $B = 2A + 8$
$$\widehat{\mathcal{A}_N} = \sum_{1 \le q \le (\log N)^B} \widehat{L_{q,N}} + O(\log N)^{-A}, \tag{3.9}$$
because we can use (3.8) to complete the proof of (3.3).

We note that the intervals of ξ that appear in the conditions 1 and 2 of Lemma 3.2 are pairwise disjoint. Let us assume that ξ meets the condition 2, so $|\xi - a/q| < \frac{(\log N)^B}{N}$ for $(a, q) = 1$ and $1 < q < (\log N)^B$. To prove (3.3) in this case, we need to see that,
$$\eta_s(\xi - b/q) = \begin{cases} 0, & \text{if } \mathbb{A}_q \ni b \ne a \\ 1, & \text{if } b = a \end{cases}$$
Hence
$$\widehat{L_{q,N}}(\xi) = \frac{\mu(q)}{\phi(q)}\widehat{\gamma_N}(\xi - a/q),$$
and furthermore by (3.5),
$$|\widehat{\mathcal{A}_N}(\xi) - \widehat{L_{q,N}}(\xi)| \le e^{-c\sqrt{\log N}}. \tag{3.10}$$
We also need to see that all the other $L_{q',N}(\xi)$ are small. Indeed, for $1 < q' \ne q \le (\log N)^B$, and $a' \in \mathbb{A}_{q'}$, we have $|\xi - a'/q'| \ge \frac{(\log N)^B}{N}$. Hence, by (3.7), we have
$$|\widehat{L_{q',N}}(\xi)| \lesssim \phi(q')^{-1}(\log N)^{-B}.$$
Similarly, we have $|\xi| \ge \frac{(\log N)^B}{N}$, hence
$$|\widehat{L_{1,N}}(\xi)| \lesssim (\log N)^{-B}.$$

Summing the estimates for $\widehat{L_{q',N}}$ over $1 \le q' \ne q \le (\log N)^B$ and using (2.1), we have

$$\sum_{1 \le q' \ne q \le (\log N)^B} |\widehat{L_{q',N}}(\xi)| \lesssim (\log N)^{-A}. \tag{3.11}$$

Putting (3.10) and (3.11) together, we have verified (3.9) in this case. If ξ meets condition 1 of Lemma 3.2, the proof is completely analogous.

We now assume that ξ does not meet the first or second condition of Lemma 3.2. Then, (3.6) holds. And, similar to (3.11), we have

$$\sum_{1 \le q \le (\log N)^B} |\widehat{L_{q,N}}(\xi)| \lesssim \sum_{1 \le q \le (\log N)^B} \phi(q)^{-1} (\log N)^{-B} \lesssim (\log N)^{-B+1}.$$

Combining (3.6) with (3.11), we have completed the proof of (3.9). □

The building blocks of the approximating multipliers have explicit inverse Fourier transforms.

Lemma 3.3. *With the notation of* (3.2), *there holds*

$$L_{q,N}(x) = \frac{\mu(q)}{\phi(q)} c_q(-x) \cdot \gamma_N \star \check{\eta}_s(x) \tag{3.12}$$

Proof. For $q \ge 2$, compute

$$\begin{aligned} L_{q,N}(x) &= \int_{\mathbb{T}} \widehat{L_{q,N}}(\xi) e(-x\xi)\, d\xi \\ &= \frac{\mu(q)}{\phi(q)} \sum_{a \in \mathbb{A}_q} \int_{\mathbb{T}} \widehat{\gamma_N} \cdot \eta_s(\xi - a/q) e(-x\xi)\, d\xi \\ &= \frac{\mu(q)}{\phi(q)} \gamma_N \star \check{\eta}_s(x) \sum_{a \in \mathbb{A}_q} e_q(-ax) = \frac{\mu(q)}{\phi(q)} c_q(-x) \cdot \gamma_N \star \check{\eta}_s(x), \end{aligned}$$

where we are using the notation of Ramanujan sums (2.5). Above $\check{\eta}_s$ is understood as $\widecheck{\eta_{s,\mathrm{per}}}$, where $\eta_{s,\mathrm{per}}$ is the 1-periodic extension of η_s. For $q = 1$, (3.12) holds since $c_1(x) \equiv 1$. □

The term on the right in (3.12) includes an average γ_N. It also includes a Ramanujan sum term. One should note that $c_q(0) = \phi(q)$, but this is far from typical behavior. This crude estimate shows that for most x, $c_q(x)$ is about one.

Lemma 3.4. *For any $\epsilon > 0$, and integer $k > 1$, uniformly in $M > Q^k$, there holds*

$$\Big[\frac{1}{M} \sum_{|x|<M} \Big|\sum_{q=1}^{Q} \frac{c_q(x)}{\phi(q)}\Big|^k\Big]^{1/k} \lesssim Q^{\epsilon}. \tag{3.13}$$

The implied constant depends upon k and ε.

Sketch of Proof. We will not give a complete proof. It follows from work of Bourgain 2, (3.43), page 126] that we have, under the assumptions above, that for any integer P,

$$\left[\frac{1}{M} \sum_{|x|<M} \Big| \sum_{P \le q < 2P} |c_q(x)| \Big|^k \right]^{1/k} \lesssim P^{1+\epsilon/2}, \qquad M > P^k.$$

This is given a stand-alone proof in 12, Lemma 3.13]. Using the well known lower bound $\phi(q) \gtrsim q^{1-\epsilon/4}$, we see that

$$\frac{1}{M} \sum_{|x|<M} \Big| \sum_{P \le q < 2P} \frac{|c_q(x)|}{\phi(q)} \Big|^k \lesssim P^{3k\varepsilon/4}. \tag{3.14}$$

Finally, let integer m_0 be such that $2^{m_0} \le Q < 2^{m_0+1}$. We have

$$\begin{aligned}\frac{1}{M}\sum_{|x|<M}\Big|\sum_{q=1}^{Q}\frac{|c_q(x)|}{\phi(q)}\Big|^k \le &\frac{1}{M}\sum_{|x|<M}\Big|\sum_{m=0}^{m_0}\sum_{2^m\le q<2^{m+1}}\frac{|c_q(x)|}{\phi(q)}\Big|^k\\ \le &(m_0+1)^{k-1}\sum_{m=0}^{m_0}\frac{1}{M}\sum_{|x|<M}\Big|\sum_{2^m\le q<2^{m+1}}\frac{|c_q(x)|}{\phi(q)}\Big|^k\\ \lesssim &(\log Q)^{k-1}\sum_{m=0}^{m_0}2^{3mk\varepsilon/4}\\ \lesssim &Q^{k\varepsilon},\end{aligned}$$

where we used (3.14). This proves the claimed result. □

4 Fixed Scale

The fixed scale result has fewer complications than the sparse bound. We show that for any $1 < p < 2$, there holds

$$N^{-1}(\mathcal{A}_N f, g) \le C_p \langle f\rangle_{2E,p}\langle g\rangle_{E,p}, \tag{4.1}$$

where E is an interval of length N, and the inequality is independent of N. Since the condition is open with respect to p, it suffices to consider the case of $p' \in \mathbb{N}$, with $f = \mathbf{1}_F$ supported on $2E$ and $g = \mathbf{1}_G$ supported on E. We trivially have

$$N^{-1}(\mathcal{A}_N f, g) \lesssim \log N \cdot \langle f\rangle_{2E,1}\langle g\rangle_{E,1}$$

so that we conclude (4.1) if

$$(\log N)(\langle f\rangle_{2E,1}\langle g\rangle_{E,1})^{1/p'} \le 1. \tag{4.2}$$

We assume that this fails, thus

$$\min\{\langle f\rangle_{2E,1}, \langle g\rangle_{E,1}\} > (\log N)^{-p'}. \tag{4.3}$$

Now, we prove this auxiliary estimate–the High Low estimate. For constants $1 \le J \le (\log N)^{p'}$, we can write $\mathcal{A}_N f = H + L$ where

$$\langle H\rangle_{E,2} \lesssim J^{-1+\frac{1}{p'}}\langle f\rangle_{2E,1}^{1/2} \tag{4.4}$$

$$\langle L\rangle_{E,\infty} \lesssim J^{1/p'}\langle f\rangle_{2E,1}^{1/p}. \tag{4.5}$$

The implied constants depend upon p. The term H is the High term, and it satisfies a quantified ℓ^2 estimate, while L satisfies something close to the $\ell^1 \to \ell^\infty$ endpoint. It consists of the 'low frequency' terms.

From this, it follows that

$$N^{-1}(\mathcal{A}_N f, g) \lesssim J^{-1+\frac{1}{p'}}(\langle f\rangle_{2E,1}\langle g\rangle_{E,1})^{1/2} + J^{1/p'}\langle f\rangle_{2E,1}^{1/p}\langle g\rangle_{E,1}.$$

The two sides are equal provided that

$$J \simeq \langle f\rangle_{2E,1}^{1/2-1/p}\langle g\rangle_{E,1}^{-1/2}. \tag{4.6}$$

By our lower bound on $\langle f\rangle_{2E,1}$ and $\langle g\rangle_{E,1}$ from (4.3), this is an allowed choice of J. And, then (4.1) follows.

It remains to prove (4.4) and (4.5). Apply our decomposition of the averaging operator (3.3) with $A = p'$ and $K = J$. With the notation from (3.3), set $H = \mathcal{F}^{-1}(r_{N,A,J}\widehat{f})$. The ℓ^2 estimate (4.4) on H follows from the L^∞ bound on $r_{N,A,J}$. Turning to (4.5), the estimate for L, from (3.12), we have

$$\begin{aligned}\Big|\sum_{q=1}^{J}L_{q,N} * f(x)\Big| \le &\sum_{q=1}^{J}\left(\frac{|c_q(\cdot)|}{\phi(q)}|\gamma_N * \check{\eta}_s(\cdot)|\right) * f(x)\\ \le &\sum_{y=1}^{N}\sum_{q=1}^{J}\frac{|c_q(y)|}{\phi(q)}|\gamma_N * \check{\eta}_s(y)|\,f(x-y).\end{aligned}$$

Here note that

$$|\gamma_N \star \check{\eta}_s(y)| \le \frac{1}{N}\|\check{\eta}_s\|_{\ell^1} \lesssim \frac{1}{N}. \tag{4.7}$$

Hence

$$\begin{aligned}\Big|\sum_{q=1}^{J} L_{q,N} \star f(x)\Big| &\le \frac{1}{N}\sum_{y=1}^{2N}\sum_{q=1}^{J}\frac{|c_q(y)|}{\phi(q)} f(x-y)\\ &\lesssim \Big[\frac{1}{2N}\sum_{y=1}^{2N}\Big|\sum_{q=1}^{J}\frac{|c_q(y)|}{\phi(q)}\Big|^{p'}\Big]^{1/p'}\Big[\frac{1}{2N}\sum_{y=1}^{2N} f(x-y)\Big]^{1/p}\\ &\lesssim J^{1/p'}\langle f\rangle_{2E,1}^{1/p}.\end{aligned}$$

Above, we have appealed to Hölder inequality and (3.13), with appropriate choice of parameters. Note this (3.13) only applies for $N > N_p$, for a choice of N_p that is only depending on p. After that, we simplify the expression, since f is an indicator set. This completes the proof.

5 Sparse Bound

We prove the sparse bound in Theorem 1.2. The sparse bound is stronger for smaller choices of (r, s), and so it suffices to prove the (p, p) sparse bound for all $1 < p < 2$. Again, by openness of the condition we are proving, it suffices to restrict attention to functions f, g that are indicator sets.

The sparse bound is proved by recursion, which depends upon the following definition. Let E be an interval of length 2^{n_0}. Let $f = \mathbf{1}_F$ be supported on $2E$, and $g = \mathbf{1}_G$ be supported on E. Let $\tau : E \to \{2^n : 1 \le n \le n_0\}$ be a choice of stopping time. We say that τ is *admissible* if for any interval $I \subset E$ such that $\langle f\rangle_{3I,1} > 100\langle f\rangle_{2E,1}$, there holds

$$\inf_{x\in I}\tau(x) > |I|. \tag{5.1}$$

We will have direct recourse to this at the end of the proof of the Lemma below.

Lemma 5.1. *For all admissible stopping times, and* $1 < p < 2$*, there holds*

$$(\mathcal{A}_\tau f, g) \lesssim (\langle f\rangle_{2E,1}\langle g\rangle_{E,1})^{1/p}|E|. \tag{5.2}$$

It is a routine argument to see that this implies the sparse bound as written in Theorem 1.2, see 5, Lemma 2.8] or 11, Lemma 2.1]. We prove the Lemma with the auxiliary High Low construction. For integers $J = 2^j$, we write $\mathcal{A}_\tau f \le H + L$ where

$$\langle H\rangle_{E,2} \lesssim J^{-1+1/p'}\langle f\rangle_{2E,1}^{1/2}, \tag{5.3}$$

$$\langle L\rangle_{E,\infty} \lesssim J^{1/p'}\langle f\rangle_{2E,1}^{1/p}. \tag{5.4}$$

The conclusion of (5.2) is very similar to the earlier argument in (4.6), and we omit the details.

We proceed with the construction of the High and Low terms. We begin with the trivial bound, following from admissibility,

$$\langle \mathcal{A}_\tau f\rangle_{E,\infty} \lesssim \sup_x(\log\tau(x))\langle f\rangle_{2E,1}. \tag{5.5}$$

On the set $B = \{\log\tau(x) \le D_p J^{1/p'}\}$, we see that (5.4) holds. Here, D_p is a constant that depends only on p, which we specify in the discussion of the Low term below. We proceed under the assumption that the set B is empty. Hence the following holds on E:

$$\tau(x) \ge D_p J^{1/p'}. \tag{5.6}$$

We are then concerned with averages $A_{2^n} f$, where $n \ge D_p J^{1/p'}$. Let

$$m := (p' + 1)\lfloor \log_2 n\rfloor.$$

Hence $(n/2)^{p'+1} < 2^m \le n^{p'+1}$. Apply the decomposition (3.3) with $N = 2^n$, $A = p' + 1$ and $K = 2^m$. Then, we have

$$\widehat{\mathcal{A}_{2^n}} = \sum_{q=1}^{2^m} \widehat{L_{2^n,q}} + \rho_{2^n},$$

where $\|\rho_{2^n}\|_\infty \lesssim n^{-p'}$. Our first contribution to the term H is $H_1 = |\widetilde{\rho_\tau} \star f|$. Note that by a familiar square function argument,

$$\begin{aligned}\|H_1\|_2^2 &\le \sum_{n \ge D_p J^{1/p'}} \|\widetilde{\rho_{2^n}} \star f\|_2^2 \\ &\lesssim \|f\|_2^2 \sum_{n \ge D_p J^{1/p'}} n^{-2p'} \lesssim J^{-2+1/p'} \|f\|_2^2.\end{aligned}$$

This satisfies the requirement in (5.3).

We continue with the construction of H. The second contribution is

$$\begin{aligned}H_2 &= \sup_n \Big| \sum_{2^j < q \le 2^m} L_{2^n,q} \star f \Big| \\ &\le \sum_{k=j}^{m} \sup_n \Big| \sum_{2^k < q \le 2^{k+1}} L_{2^n,q} \star f \Big|\end{aligned}$$

The point of this last line is that the inequality below is a direct consequence of Bourgain's multi-frequency maximal inequality, and the bound $\phi(q) \gtrsim_p q^{1-1/(2p')}$:

$$\Big\| \sup_n \Big| \sum_{2^k \le q < 2^{k+1}} L_{2^n,q} \star f \Big| \Big\|_{\ell^2} \lesssim k \max_{2^k \le q < 2^{k+1}} \frac{1}{\phi(q)} \cdot \|f\|_{\ell^2} \lesssim 2^{-k(1-1/p')} \|f\|_{\ell^2}. \tag{5.7}$$

Summing this estimate over $k \ge j$ completes the analysis of the High term.

Remark 5.8. One of the main results of Bourgain 3] is the multi-frequency maximal inequality, a key aspect of discrete Harmonic Analysis. In the form that we have used it in (5.7), see for instance 13, Prop. 5.11].

The term that remains is the Low term below. We appeal to (3.12), to see that

$$\Big| \sum_{q=1}^{J} L_{q,\tau} \star f(x) \Big| \le \sum_{q=1}^{J} \left(\frac{|c_q(\cdot)|}{\phi(q)} |\gamma_\tau \star \check{\eta}_s(\cdot)| \right) \star f(x) \tag{5.9}$$

We need the following simple Lemma concerning $\gamma_\tau \star \check{\eta}_s$.

Lemma 5.2. *We have*

$$|\gamma_\tau \star \check{\eta}_s(y)| \lesssim \begin{cases} \tau^{-1} & \text{if } |y| \le 4\tau \\ 2^{-2k}\,\tau^{-1} & \text{if } |y| \in (2^k\tau, 2^{k+1}\tau], \ \text{for } k \ge 2. \end{cases}$$

Proof. The proof for the case $|y| \le 4\tau$ follows from (4.7). Now, assume $|y| \in (2^k\tau, 2^{k+1}\tau]$ for $k \ge 2$. We have

$$|\gamma_\tau \star \check{\eta}_s(y)| \le \frac{1}{8^s\tau} \sum_{z=1}^{\tau} |\check{\eta}(\frac{y-z}{8^s})| \lesssim \frac{1}{8^s\tau} \sum_{z=1}^{\tau} \left(1 + (\frac{y-z}{8^s})^2\right)^{-1} \lesssim \frac{8^s}{2^{2k}\tau^2} \lesssim \frac{1}{2^{2k}\tau}.$$

In the last inequality we used $8^s \le q^3 \le J^3 < 2^{D_p J^{1/p'}} < \tau(x)$, due to (5.6) with a proper choice of D_p. □

Plugging the estimates in Lemma 5.2 into (5.9), we have

$$
\begin{aligned}
\Big|\sum_{q=1}^{J} L_{q,\tau} \star f(x)\Big| \lesssim & \frac{1}{\tau} \sum_{|y|\le 4\tau} \sum_{q=1}^{J} \frac{|c_q(y)|}{\phi(q)} f(x-y) \\
& + \sum_{k=2}^{\infty} \frac{1}{2^{2k}\tau} \sum_{|y|\le 2^{k+1}\tau} \sum_{q=1}^{J} \frac{|c_q(y)|}{\phi(q)} f(x-y) \\
\lesssim & \Big[\sum_{y=1}^{\tau} \frac{1}{\tau} \Big|\sum_{q=1}^{J} \frac{|c_q(y)|}{\phi(q)}\Big|^{p'}\Big]^{\frac{1}{p'}} \Big[\frac{1}{\tau}\sum_{y=1}^{\tau} f(x-y)\Big]^{\frac{1}{p}} \\
& + \sum_{k=2}^{\infty} \frac{1}{2^k} \Big[\sum_{y=1}^{2^{k+1}\tau} \frac{1}{2^{k+1}\tau} \Big|\sum_{q=1}^{J} \frac{|c_q(y)|}{\phi(q)}\Big|^{p'}\Big]^{\frac{1}{p'}} \Big[\frac{1}{2^{k+1}\tau} \sum_{y=1}^{2^{k+1}\tau} f(x-y)\Big]^{\frac{1}{p}} \\
\lesssim & J^{1/p'} \langle f \rangle_{2E,1}^{1/p}.
\end{aligned}
$$

We use Hölder's inequality in ℓ^p-$\ell^{p'}$, and use (3.13) above to gain the factor of $J^{1/p'}$. Recall that (3.13) holds in this setting, since we assumed (5.6). Thus, $J^{p'} < 2^{D_p J^{1/p'}} < \tau(x)$, for appropriate choice of constant D_p. Note that admissibility of τ, namely the condition (5.1), gives us the estimate in terms of $\langle f \rangle_{2E,1}$. This completes the proof of (5.4), and completes the proof of the sparse bound.

References

[1] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. **61** (1988), no. 1, 39–72. MR937581

[2] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations*, Geom. Funct. Anal. **3** (1993), no. 2, 107–156. MR1209299

[3] 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

[4] Amalia Culiuc, Robert Kesler, and Michael T. Lacey, *Sparse bounds for the discrete cubic Hilbert transform*, Anal. PDE **12** (2019), no. 5, 1259–1272. MR3892403

[5] Rui Han, Michael T Lacey, and Fan Yang, *Averages along the Square Integers: ℓ^p improving and Sparse Inequalities*, arXiv e-prints (2019Jul), arXiv:1907.05734, available at https://arxiv.org/abs/1907.05734.

[6] K. Hughes, *ℓ^p-improving for discrete spherical averages*, ArXiv e-prints (April 2018), available at 1804.09260.

[7] Alexandru D. Ionescu, *An endpoint estimate for the discrete spherical maximal function*, Proc. Amer. Math. Soc. **132** (2004), no. 5, 1411–1417. MR2053347

[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] R. Kesler, *$\ell^p(\mathbb{Z}^d)$-Improving Properties and Sparse Bounds for Discrete Spherical Maximal Averages*, ArXiv e-prints (May 2018), available at https://arxiv.org/abs/1805.09925.

[10] R. Kesler, *$\ell^p(\mathbb{Z}^d)$-Improving Properties and Sparse Bounds for Discrete Spherical Maximal Means, Revisited*, ArXiv e-prints (September 2018), available at https://arxiv.org/abs/1809.06468.

[11] Robert Kesler, Michael T. Lacey, and Darío Mena, *Sparse bounds for the discrete spherical maximal functions*, Pure Appl. Anal. **2** (2020), no. 1, 75–92, DOI 10.2140/paa.2020.2.75. MR4041278

[12] Robert Kesler, Michael T. Lacey, and Darío Mena Arias, *Lacunary discrete spherical maximal functions*, New York J. Math. **25** (2019), 541–557. MR3982252

[13] Ben Krause, *Discrete Analogoues in Harmonic Analysis: Maximally Monomially Modulated Singular Integrals Related to Carleson's Theorem*, arXiv e-prints (2018Mar), arXiv:1803.09431, available at 1803.09431.

[14] Michael T. Lacey and Robert Kesler, *ℓ^p-improving inequalities for Discrete Spherical Averages*, arXiv e-prints (April 2018), arXiv:1804.09845, available at 1804.09845.

[15] Mariusz Mirek, Elias M. Stein, and Bartosz Trojan, *$\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: maximal functions and vector-valued estimates*, J. Funct. Anal. **277** (2019), no. 8, 2471–2521, DOI 10.1016/j.jfa.2018.10.020. MR3990725

[16] Mariusz Mirek, Elias M. Stein, and Bartosz Trojan, *$\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: variational estimates*, Invent. Math. **209** (2017), no. 3, 665–748. MR3681393

[17] Mariusz Mirek and Bartosz Trojan, *Cotlar's ergodic theorem along the prime numbers*, J. Fourier Anal. Appl. **21** (2015), no. 4, 822–848. MR3370012

[18] Bartosz Trojan, *Endpoint estimates for the maximal function over prime numbers*, arXiv e-prints (2019Jul), arXiv:1907.04753, available at 1907.04753.

[19] R. C. Vaughan, *The Hardy-Littlewood method*, Cambridge Tracts in Mathematics, vol. 80, Cambridge University Press, Cambridge-New York, 1981. MR628618

[20] Máté Wierdl, *Pointwise ergodic theorem along the prime numbers*, Israel J. Math. **64** (1988), no. 3, 315–336 (1989). MR995574