Journal of Fourier Analysis and Applications (2024) 30:61
https://doi.org/10.1007/s00041-024-10119-6

Vector-Valued Maximal Inequalities and Multiparameter Oscillation Inequalities for the Polynomial Ergodic Averages Along Multi-dimensional Subsets of Primes

Nathan Mehlhop[1]

Received: 1 June 2023 / Revised: 25 September 2024 / Accepted: 3 October 2024 /
Published online: 23 October 2024

Abstract

We prove uniform ℓ^2-valued maximal inequalities for polynomial ergodic averages and truncated singular operators of Cotlar type modeled over multidimensional subsets of primes. In the averages case, we combine this with earlier one-parameter oscillation estimates (Mehlhop and Słomian in Math Ann, 2023, https://doi.org/10.1007/s00208-023-02597-8) to prove corresponding multiparameter oscillation estimates. This provides a fuller quantitative description of the pointwise convergence of the mentioned averages and is a generalization of the polynomial Dunford–Zygmund ergodic theorem attributed to Bourgain (Mirek et al. in Rev Mat Iberoam 38:2249–2284, 2022).

Keywords Oscillation seminorm · Vector-valued inequality · Ergodic average along primes · Multiparameter average

Mathematics Subject Classification 37A30 (Primary) · 37A46 · 42B20

1 Introduction

1.1 Statement of Results

Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space endowed with a family of invertible commuting and measure preserving transformations $S_1, \ldots, S_d : X \to X$. Let Ω be a bounded convex open subset of $\mathbb{R}^k$ such that $B(0, c_\Omega) \subseteq \Omega \subseteq B(0, 1)$ for some $c_\Omega \in (0, 1)$, where $B(0, u)$ is the open Euclidean ball in $\mathbb{R}^k$ with radius $u > 0$ centered

Communicated by Stefan Steinerberger.

✉ Nathan Mehlhop
nmehlhop@lsu.edu

1 Department of Mathematics, Louisiana State University, Baton Rouge, LA 07802, USA

at $0 \in \mathbb{R}^k$. For any $t > 0$, we set

$$\Omega_t := \{x \in \mathbb{R}^k : t^{-1}x \in \Omega\}.$$

We consider a polynomial mapping

$$\mathcal{P} = (\mathcal{P}_1, \ldots, \mathcal{P}_d) \colon \mathbb{Z}^k \to \mathbb{Z}^d \tag{1.1}$$

where each $\mathcal{P}_j : \mathbb{Z}^k \to \mathbb{Z}$ is a polynomial of k variables with integer coefficients such that $\mathcal{P}_j(0) = 0$. Let $k', k'' \in \{0, 1, \ldots, k\}$ with $k = k' + k''$. For $f \in L^\infty(X, \mu)$, we define the associated ergodic averages by

$$\mathcal{A}_t^{\mathcal{P},k',k''} f(x) := \frac{1}{\vartheta_\Omega(t)} \sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} f\left(S_1^{\mathcal{P}_1(n,p)} \cdots S_d^{\mathcal{P}_d(n,p)} x\right) \mathbb{1}_{\Omega_t}(n, p) \left(\prod_{i=1}^{k''} \log |p_i|\right), x \in X, \tag{1.2}$$

where $\pm\mathbb{P}$ denotes the set of positive and negative prime numbers and

$$\vartheta_\Omega(t) := \sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} \mathbb{1}_{\Omega_t}(n, p) \left(\prod_{i=1}^{k''} \log |p_i|\right)$$

is the Chebyshev function. We also consider the Cotlar type ergodic averages (discrete singular integrals) given by

$$\mathcal{H}_t^{\mathcal{P},k',k''} f(x) := \sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} f\left(S_1^{\mathcal{P}_1(n,p)} \cdots S_d^{\mathcal{P}_d(n,p)} x\right) K(n, p) \mathbb{1}_{\Omega_t}(n, p) \left(\prod_{i=1}^{k''} \log |p_i|\right), x \in X, \tag{1.3}$$

where $K : \mathbb{R}^k \setminus \{0\} \to \mathbb{C}$ is a Calderón–Zygmund kernel satisfying the following conditions:

(1) The size condition: For every $x \in \mathbb{R}^k \setminus \{0\}$, we have

$$|K(x)| \lesssim |x|^{-k}. \tag{1.4}$$

(2) The cancellation condition: For every $0 < r < R < \infty$, we have

$$\int_{\Omega_R \setminus \Omega_r} K(y) \mathrm{d}y = 0. \tag{1.5}$$

(3) The Lipschitz continuity condition: For every $x, y \in \mathbb{R}^k \setminus \{0\}$ with $2|y| \leq |x|$, we have
$$|K(x) - K(x+y)| \lesssim |y||x|^{-(k+1)}. \tag{1.6}$$

In the case where $k' = k$ and $k'' = 0$, one may instead consider the Hölder continuity condition generalizing (1.6) (see Proposition 11 for discussion of this issue): For some $\sigma \in (0, 1]$ and for every $x, y \in \mathbb{R}^k \setminus \{0\}$ with $2|y| \leq |x|$, we have

$$|K(x) - K(x+y)| \lesssim |y|^{\sigma}|x|^{-(k+\sigma)}. \tag{1.7}$$

For a sequence of functions $(f_{\mathrm{i}})_{\mathrm{i}\in\mathbb{N}}$ with each $f_{\mathrm{i}} \in L^p(X, \mu)$, we define the $L^p(X; \ell^2)$ norm by

$$\|f_{\mathrm{i}}\|_{L^p(X;\ell^2)} = \left\| \left(\sum_{\mathrm{i}\in\mathbb{N}} |f_{\mathrm{i}}|^2 \right)^{1/2} \right\|_{L^p(X)} \tag{1.8}$$

and we say that $(f_{\mathrm{i}})_{\mathrm{i}\in\mathbb{N}} \in L^p(X; \ell^2)$ if $\|f_{\mathrm{i}}\|_{L^p(X;\ell^2)} < \infty$.

We can now state the main result of this paper.

Theorem 1 *Let $d, k \in \mathbb{N}$ and let $\mathcal{P}$ be a polynomial mapping as in (1.1). Let $k', k'' \in \{0, 1, \ldots, k\}$ with $k' + k'' = k$ and let $\mathcal{M}_t^{\mathcal{P},k',k''}$ be either $\mathcal{A}_t^{\mathcal{P},k',k''}$ or $\mathcal{H}_t^{\mathcal{P},k',k''}$. Then, for any $p \in (1, \infty)$, there is a constant $C_{p,d,k,\deg\mathcal{P}} > 0$ such that*

$$\left\| \sup_{t>0} \left| \mathcal{M}_t^{\mathcal{P},k',k''} f_{\mathrm{i}} \right| \right\|_{L^p(X;\ell^2)} \leq C_{p,d,k,\deg\mathcal{P}} \|f_{\mathrm{i}}\|_{L^p(X;\ell^2)} \tag{1.9}$$

for any $(f_{\mathrm{i}})_{\mathrm{i}\in\mathbb{N}} \in L^p(X; \ell^2)$. The constant $C_{p,d,k,\deg\mathcal{P}}$ is independent of the coefficients of the polynomial mapping $\mathcal{P}$.

In the proof of the above theorem, we use methods developed in [23, 27, 38] and very recently in [20, 21, 35]. We follow Bourgain's approach [5] to use the Calderón transference principle [7] which reduces the problem to the integer shift system (see Sect. 2.4) and then exploit the Hardy–Littlewood circle method to analyze the appropriate Fourier multipliers. The main tools used to handle the estimates for the multiplier operators are: an appropriate generalization of Weyl's inequality (Proposition 10); the Ionescu–Wainger multiplier theorem (see [13, 27] and [37]) combined with the Rademacher–Menshov inequality (see [23]) and standard multiplier approximations (Lemma 12); the Magyar–Stein–Wainger sampling principle [19] and [25]. Throughout, we also use the Marcinkiewicz-Zygmund inequality (Proposition 7) to extend scalar inequalities to their vector-valued analogues.

We recall the λ-jump counting function and the variation and oscillation seminorms, which give quantitative measures for pointwise convergence. We use the convention that a supremum taken over the empty set is zero. Let $\mathbb{I} \subseteq \mathbb{R}$ with $\#\mathbb{I} \geq 2$ and $f : \mathbb{I} \to \mathbb{C}$. For any $N \in \mathbb{N} \cup \{\infty\}$, we write $\mathfrak{S}_N(\mathbb{I})$ to denote the family of all strictly increasing sequences $(I_0, \ldots, I_N)$ of length $N + 1$ contained in $\mathbb{I}$.

Given $\lambda > 0$,the λ-*jump counting function* of f is defined by

$$N_\lambda(f(t) : t \in \mathbb{I}) := \sup\Big\{N \in \mathbb{N} \mid \exists_{\substack{t_0<\cdots<t_N \\ t_j \in \mathbb{I}}} : \min_{0\le j\le N-1} |f(t_{j+1}) - f(t_j)| \ge \lambda\Big\}.$$

Given $r \in [1, \infty)$, the *r-variation seminorm* V^r of a f is defined by

$$V^r(f(t) : t \in \mathbb{I}) := \sup_{\substack{t_0<\cdots<t_N \\ t_j \in \mathbb{I}}} \left(\sum_{j=0}^{N-1} |f(t_{j+1}) - f(t_j)|^r\right)^{1/r}.$$

Given $r \in [1, \infty)$, $N \in \mathbb{N} \cup \{\infty\}$, $I \in \mathfrak{S}_N(\mathbb{I})$, and $\mathbb{J} \subseteq \mathbb{I}$, the *$N$-truncated r-oscillation seminorm* of f is defined by

$$O^r_{I,N}(f(t) : t \in \mathbb{J}) := \left(\sum_{j=0}^{N-1} \sup_{\substack{I_j \le t < I_{j+1} \\ t \in \mathbb{J}}} |f(t) - f(I_j)|^r\right)^{1/r}.$$

There is no ambiguity if we instead take $I \in \mathfrak{S}_\infty(\mathbb{I})$ since only the first $N + 1$ terms of I are used.

Because of their preeminent role in multiparameter pointwise convergence problems, we also consider multiparameter analogues of the oscillation seminorms. Let $\mathbb{I} \subseteq \mathbb{R}^M$ with $\#\mathbb{I} \ge 2$ and $f : \mathbb{I} \to \mathbb{C}$. We now write $\mathfrak{S}_N(\mathbb{I})$ to denote the family of all sequences $(I_0, \ldots, I_N)$ of length $N + 1$ contained in $\mathbb{I}$ that are strictly increasing in every coordinate.

Given $r \in [1, \infty)$, $N \in \mathbb{N} \cup \{\infty\}$, $I \in \mathfrak{S}_N(\mathbb{I})$ (or even $I \in \mathfrak{S}_\infty(\mathbb{I})$), and $\mathbb{J} \subseteq \mathbb{I}$, the *$M$-parameter N-truncated r-oscillation seminorm* of f is defined by

$$O^r_{I,N}(f(t) : t \in \mathbb{J}) := \left(\sum_{j=0}^{N-1} \sup_{t \in \mathbb{B}[I_j] \cap \mathbb{J}} |f(t) - f(I_j)|^r\right)^{1/r},$$

where $\mathbb{B}[I_j] := [I^1_j, I^1_{j+1}) \times \ldots \times [I^M_j, I^M_{j+1})$ is a box determined by the element $I_j = (I^1_j, \ldots, I^M_j)$ of the sequence I.

For more information about these quantitative tools in the study of pointwise convergence problems, we refer to [24], see also [5, 17, 21, 24, 25, 34].

We now recall an abstract multiparameter oscillation result. For a linear operator $T : L^0(X) \to L^0(X)$, we denote by $|T|$ the sublinear maximal operator taken in the lattice sense defined by

$$|T|f(x) = \sup_{|g|\le|f|} |Tg(x)|, \quad x \in X, \quad f \in L^p(X).$$

Proposition 2 [24, Proposition 4.1] *Let $(X, \mathcal{B}(X), \mu)$ be a σ-finite measure space and let $\mathbb{I} \subseteq \mathbb{R}$ be such that $\#\mathbb{I} \geq 2$. Let $k \in \mathbb{N}_{\geq 2}$ and $p, r \in (1, \infty)$ be fixed. Let $(T_t)_{t \in \mathbb{I}^k}$ be a family of linear operators of the form*

$$T_t := T^1_{t_1} \cdots T^k_{t_k}, \qquad t = (t_1, \ldots, t_k) \in \mathbb{I}^k,$$

where $\{T^i_{t_i} : i \in [k],\ t_i \in \mathbb{I}\}$ is a family of commuting linear operators that are bounded on $L^p(X)$. If the set $\mathbb{I}$ is uncountable, then we also assume that $\mathbb{I} \ni t \mapsto T^i_t f$ is continuous μ-almost everywhere on X for every $f \in L^0(X)$ and $i \in [k]$. Further assume that, for every $i \in [k]$, we have

$$\sup_{J \in \mathbb{N}} \sup_{I \in \mathfrak{S}_J(\mathbb{I})} \left\| O^r_{I,J}(T^i_t f : t \in \mathbb{I}) \right\|_{L^p(X)} \lesssim_{p,r} \|f\|_{L^p(X)}, \qquad f \in L^p(X), \tag{1.10}$$

and

$$\left\| \left(\sum_{j \in \mathbb{Z}} \left(\sup_{t \in \mathbb{I}} |T^i_t| f_j \right)^r \right)^{1/r} \right\|_{L^p(X)} \lesssim_{p,r} \left\| \left(\sum_{j \in \mathbb{Z}} |f_j|^r \right)^{1/r} \right\|_{L^p(X)}, \qquad (f_j)_{j \in \mathbb{Z}} \in L^p(X; \ell^r(\mathbb{Z})). \tag{1.11}$$

Then we have the following multiparameter r-oscillation estimate:

$$\sup_{J \in \mathbb{N}} \sup_{I \in \mathfrak{S}_J(\mathbb{I}^k)} \left\| O^r_{I,J}(T_t f : t \in \mathbb{I}^k) \right\|_{L^p(X)} \lesssim \|f\|_{L^p(X)}, \qquad f \in L^p(X).$$

In the $\mathcal{M}^{\mathcal{P},k',k''}_t = \mathcal{A}^{\mathcal{P},k',k''}_t$ case, (1.9) gives us

$$\begin{aligned} \left\| \sup_{t>0} |\mathcal{A}^{\mathcal{P},k',k''}_t| f_{\mathrm{i}} \right\|_{L^p(X;\ell^2)} &= \left\| \sup_{t>0} \mathcal{A}^{\mathcal{P},k',k''}_t |f_{\mathrm{i}}| \right\|_{L^p(X;\ell^2)} \lesssim \| | f_{\mathrm{i}} | \|_{L^p(X;\ell^2)} \\ &= \|f_{\mathrm{i}}\|_{L^p(X;\ell^2)} \end{aligned} \tag{1.12}$$

which corresponds to condition (1.11) in the $r = 2$ case. We also recall the variation, jump, and one-parameter oscillation inequalities for $\mathcal{A}^{\mathcal{P},k',k''}_t$ and $\mathcal{H}^{\mathcal{P},k',k''}_t$.

Proposition 3 [38, Theorem C] [20, Theorem 1] *Let $d, k \geq 1$, $r \in (2, \infty)$, and let $\mathcal{P}$ be a polynomial mapping as in (1.1). Let $k', k'' \in \{0, 1, \ldots, k\}$ with $k' + k'' = k$ and let $\mathcal{M}^{\mathcal{P},k',k''}_t$ be either $\mathcal{A}^{\mathcal{P},k',k''}_t$ or $\mathcal{H}^{\mathcal{P},k',k''}_t$. Then, for any $p \in (1, \infty)$, there is a constant $C_{p,d,k,\deg \mathcal{P}} > 0$ such that*

$$\left\| V^r\big(\mathcal{M}^{\mathcal{P},k',k''}_t f : t > 0\big) \right\|_{L^p(X,\mu)} \leq \frac{r}{r-2} C_{p,d,k,\deg \mathcal{P}} \|f\|_{L^p(X,\mu)}, \tag{1.13}$$

$$\sup_{\lambda>0}\left\|\lambda N_\lambda\big(\mathcal{M}_t^{\mathcal{P},k',k''}f : t>0\big)^{1/2}\right\|_{L^p(X,\mu)} \le C_{p,d,k,\deg\mathcal{P}}\|f\|_{L^p(X,\mu)}, \tag{1.14}$$

$$\sup_{N\in\mathbb{N}}\sup_{I\in\mathfrak{S}_N(\mathbb{R}_+)}\left\|O_{I,N}^2\big(\mathcal{M}_t^{\mathcal{P},k',k''}f : t>0\big)\right\|_{L^p(X,\mu)} \le C_{p,d,k,\deg\mathcal{P}}\,\|f\|_{L^p(X,\mu)}\,, \tag{1.15}$$

for any $f \in L^p(X,\mu)$. The constant $C_{p,d,k,\deg\mathcal{P}}$ is independent of the coefficients of the polynomial mapping $\mathcal{P}$.

In particular, (1.15) corresponds to condition (1.10). As such, we have the following applications.

Corollary 4 *Let $M \in \mathbb{N}$ and let $(X,\mathcal{B},\mu)$ be a σ-finite measure space endowed with a family of invertible commuting and measure preserving transformations $S_1^1,\ldots,S_{d_1}^1,\ldots,S_1^M,\ldots,S_{d_M}^M : X \to X$. For each $j \in \{1,\ldots,M\}$, let Ω^j be a bounded convex open subset of $\mathbb{R}^{k_j}$ such that $B(0,c) \subseteq \Omega^j \subseteq B(0,1)$ for some $c \in (0,1)$, let $\mathcal{P}^j$ be a polynomial mapping*

$$\mathcal{P}^j = \big(\mathcal{P}_1^j,\ldots,\mathcal{P}_{d_j}^j\big)\colon \mathbb{Z}^{k_j} \to \mathbb{Z}^{d_j}$$

where each $\mathcal{P}_i^j : \mathbb{Z}^{k_j} \to \mathbb{Z}$ is a polynomial of k_j variables with integer coefficients such that $\mathcal{P}_i^j(0)=0$, and let $k_j', k_j'' \in \{0,\ldots,k_j\}$ with $k_j = k_j' + k_j''$. For $f \in L^\infty(X,\mu)$, we define the associated ergodic averages by

$$\mathcal{A}_t^{\mathcal{P}^j,k_j',k_j''}f(x) := \frac{1}{\vartheta_{\Omega^j}(t)}\sum_{(n,p)\in\mathbb{Z}^{k_j'}\times(\pm\mathbb{P})^{k_j''}} f\Big(\big(S_1^j\big)^{\mathcal{P}_1^j(n,p)}\cdots\big(S_{d_j}^j\big)^{\mathcal{P}_{d_j}^j(n,p)}x\Big)\mathbb{1}_{\Omega_t^j}(n,p)\left(\prod_{i=1}^{k_j''}\log|p_i|\right), \quad x\in X,$$

where

$$\vartheta_{\Omega^j}(t) := \sum_{(n,p)\in\mathbb{Z}^{k_j'}\times(\pm\mathbb{P})^{k_j''}} \mathbb{1}_{\Omega_t^j}(n,p)\left(\prod_{i=1}^{k_j''}\log|p_i|\right).$$

Letting $k = k_1+\ldots+k_M$, $k' = k_1'+\ldots+k_M'$, and $k'' = k_1''+\ldots+k_M''$, we let $(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}$ denote $(n_1,p_1,\ldots,n_M,p_M)\in\mathbb{Z}^{k_1'}\times(\pm\mathbb{P})^{k_1''}\times\ldots\times\mathbb{Z}^{k_M'}\times(\pm\mathbb{P})^{k_M''}\cong\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}$. For $f\in L^\infty(X,\mu)$ and $\vec{t}=(t_1,\ldots,t_M)\in\mathbb{R}_+^M$, we define the associated multiparameter ergodic averages by

$$\mathcal{A}_{\vec{t}}f(x) := \mathcal{A}_{t_1,\ldots,t_M}^{\mathcal{P}^1,\ldots,\mathcal{P}^M,k_1',k_1'',\ldots,k_M',k_M''}f(x) := \mathcal{A}_{t_1}^{\mathcal{P}^1,k_1',k_1''}\circ\cdots\circ\mathcal{A}_{t_M}^{\mathcal{P}^M,k_M',k_M''}f(x)$$
$$= \frac{1}{\vartheta(\vec{t})}\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} f\Big(\big(S_1^1\big)^{\mathcal{P}_1^1(n_1,p_1)}\cdots\big(S_{d_1}^1\big)^{\mathcal{P}_{d_1}^1(n_1,p_1)}\cdots\big(S_1^M\big)^{\mathcal{P}_1^M(n_M,p_M)}\cdots$$

$$\big(S_{d_M}^M\big)^{\mathcal{P}_{d_M}^M(n_M,p_M)}x\Big)$$
$$\times\, \mathbb{1}_{\Omega_{t_1}^1\times\ldots\times\Omega_{t_M}^M}(n,p)\left(\prod_{i=1}^{k''}\log|p_i|\right),\quad x\in X,$$

where

$$\vartheta(\vec{t}):=\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}}\mathbb{1}_{\Omega_{t_1}^1\times\ldots\times\Omega_{t_M}^M}(n,p)\left(\prod_{i=1}^{k''}\log|p_i|\right).$$

Let $p\in(1,\infty)$ and $f\in L^p(X,\mu)$. Then we have:

(i) (Mean ergodic theorem) the averages $\mathcal{A}_{\vec{t}}f$ converge in $L^p(X,\mu)$ norm as $\min\{t_1,\ldots,t_M\}\to\infty$;

(ii) (Pointwise ergodic theorem) the averages $\mathcal{A}_{\vec{t}}f$ converge pointwise μ-almost everywhere on X as $\min\{t_1,\ldots,t_M\}\to\infty$;

(iii) (Maximal ergodic theorem) the following maximal estimate holds, including with $p=\infty$:

$$\Big\|\sup_{\vec{t}\in\mathbb{R}_+^M}|\mathcal{A}_{\vec{t}}f|\Big\|_{L^p(X,\mu)}\lesssim_{d,k,p,M,\deg\mathcal{P}}\|f\|_{L^p(X,\mu)};\tag{1.16}$$

(iv) (Oscillation ergodic theorem) the following uniform oscillation inequality holds:

$$\sup_{N\in\mathbb{N}}\sup_{I\in\mathfrak{S}_N(\mathbb{R}_+^M)}\Big\|O_{I,N}^2\big(\mathcal{A}_{\vec{t}}f:\vec{t}\in\mathbb{R}_+^M\big)\Big\|_{L^p(X)}\lesssim_{d,k,p,M,\deg\mathcal{P}}\|f\|_{L^p(X)},$$
$$f\in L^p(X).\tag{1.17}$$

The implicit constants in (1.16) and (1.17) are independent of the coefficients of the polynomial mapping $\mathcal{P}$.

This generalizes the polynomial Dunford–Zygmund ergodic theorem due to Bourgain as we shall see in Sect. 1.3. We note that (i) follows from the dominated convergence theorem together with (ii) and (iii), and these each follow from (iv). Although Corollary 4 only requires the $\mathcal{M}_t^{\mathcal{P},k',k''}=\mathcal{A}_t^{\mathcal{P},k',k''}$ case of Theorem 1, we also prove the $\mathcal{M}_t^{\mathcal{P},k',k''}=\mathcal{H}_t^{\mathcal{P},k',k''}$ case for the sake of independent interest and to exhibit a unified approach that illustrates what common features of the operators are needed in the proof.

1.2 Historical Background: One Parameter Problems

In 1931, Birkhoff [2] and von Neumann [30] proved that the averages

$$M_Nf(x):=\frac{1}{N}\sum_{n=1}^{N}f(S^nx)\tag{1.18}$$

converge pointwise μ-almost everywhere on X and in $L^p(X, \mu)$ norm respectively for any $f \in L^p(X, \mu)$, $p \in [1, \infty)$, as $N \to \infty$. In 1955, Cotlar [10] established the pointwise μ-almost everywhere convergence on X as $N \to \infty$ of the ergodic Hilbert transform given by

$$H_N f(x) := \sum_{1 \leq |n| \leq N} \frac{f(S^n x)}{n}$$

for any $f \in L^p(X, \mu)$. In 1968, Calderón [7] made an important observation (now called the Calderón transference principle) that some results in ergodic theory can be easily deduced from known results in harmonic analysis. Namely, the convergence of the Birkhoff averages M_N can be deduced from the boundedness of the Hardy–Littlewood maximal function, and the convergence of Cotlar's averages H_N follows from the boundedness of the maximal function for the truncated discrete Hilbert transform. As we will see ahead, this observation has had a huge impact in the study of convergence problems in ergodic theory.

We briefly sketch the classical approach of handling the problem of pointwise convergence. It consists of two steps:

(a) Establish L^p-boundedness for the corresponding maximal function.
(b) Find a dense class of functions in $L^p(X, \mu)$ for which the pointwise convergence holds.

In the case of Birkhoff's averages M_N, the Calderón transference principle allows one to deduce the estimate

$$\left\| \sup_{N \in \mathbb{N}} |M_N f| \right\|_{L^p(X,\mu)} \lesssim_p \|f\|_{L^p(X,\mu)}$$

for $p \in (1, \infty]$ from the estimate for the discrete Hardy–Littlewood maximal function (and we have a weak-type estimate for $p = 1$). In turn, estimates for the discrete Hardy–Littlewood maximal function follow easily from those for the continuous one. This establishes the first step (a). For the second step, one can use the idea of Riesz decomposition [32] to analyze the space $\mathbb{I}_S \oplus \mathbb{T}_S \subseteq L^2(X, \mu)$, where

$$\mathbb{I}_S := \{f \in L^2(X, \mu) : f \circ S = f\} \quad \text{and} \quad \mathbb{T}_S := \{h \circ S - h : h \in L^2(X, \mu) \cap L^\infty(X, \mu)\}.$$

We see that $M_N f = f$ for $f \in \mathbb{I}_S$ and, for $g = h \circ S - h \in \mathbb{T}_S$, we have

$$M_N g(x) = \frac{1}{N}\big(h\big(S^{N+1}x\big) - h(Sx)\big)$$

by telescoping. Consequently, we see that $M_N g \to 0$ as $N \to \infty$. This establishes μ-almost everywhere pointwise convergence of M_N on $\mathbb{I}_S \oplus \mathbb{T}_S$, which is dense in $L^2(X, \mu)$. Since $L^2(X, \mu)$ is dense in $L^p(X, \mu)$ for every $p \in [1, \infty)$, this establishes (b).

At the beginning of the 1980's, Bellow [1] and independently Furstenberg [12] posed the problem of the pointwise convergence of the averages along squares given by

$$T_N f(x) := \frac{1}{N}\sum_{n=1}^{N} f\big(S^{n^2}x\big).$$

Despite its similarity to Birkhoff's theorem, the problem of pointwise convergence of the T_N averages has a totally different nature from that of its linear counterpart, and the standard approach is insufficient in this case. For the first step, by the Calderón transference principle, it is enough to establish ℓ^p bounds for the maximal function given by

$$\sup_{N\in\mathbb{N}} \frac{1}{N}\sum_{n=1}^{N} f\big(x-n^2\big), \quad f \in \ell^p(\mathbb{Z}). \tag{1.19}$$

The ℓ^p estimate for the above maximal function does not follow directly from the continuous counterpart and requires completely new methods. However, a more serious problem arises in connection with the second step. The telescoping idea fails in the case of the averages $T_N g$ since the gap sizes $(n+1)^2 - n^2 = 2n+1$ are unbounded.

At the end of the 1980's, Bourgain established the pointwise convergence of the averages T_N in a series of groundbreaking articles [3–5]. By using the Hardy–Littlewood circle method from analytic number theory, he established ℓ^p-bounds for the maximal function (1.19), which establishes step (a). He then bypassed the problem of finding the requisite dense class of functions by using the oscillation seminorm. Bourgain [5] proved that, for any $\lambda > 1$ and any sequence of integers $I = (I_j : j \in \mathbb{N})$ with $I_{j+1} > 2I_j$ for all $j \in \mathbb{N}$, we have

$$\Big\| O^2_{I,N}(T_{\lambda^n} f : n \in \mathbb{N})\Big\|_{L^2(X,\mu)} \le C_{I,\lambda}(N)\,\|f\|_{L^2(X,\mu)}, \quad N \in \mathbb{N}, \tag{1.20}$$

for any $f \in L^2(X,\mu)$ with $\lim_{N\to\infty} N^{-1/2} C_{I,\lambda}(N) = 0$. Inequality (1.20) suffices to establish the pointwise convergence of the averaging operators $T_N f$ for any $f \in L^2(X,\mu)$ (see [24, Proposition 2.8] for why oscillation estimates give pointwise convergence and [6, Section 3.2] for proving convergence of $T_N f$ from that of $T_{\lambda^n} f$). Indeed, it can be thought of as the weakest possible quantitative form of pointwise convergence since one can derive (1.20) with $C_{I,\lambda}(N)$ at most $N^{1/2}$ from the ℓ^2 bound for the maximal function (1.19).

In the same series of papers, by similar methods, Bourgain established the pointwise convergence of the averages along primes

$$\frac{1}{|\mathbb{P}_N|}\sum_{n=1}^{N} f(S^n x)\mathbb{1}_{\mathbb{P}}(n)$$

for $f \in L^p(X,\mu)$ with $p > \frac{1}{2}(1+\sqrt{3})$. In the same year, Wierdl [39] extended Bourgain's result to $p \in (1,\infty)$.

The groundbreaking work of Bourgain led to work by many others in discrete harmonic analysis that proved various special cases of Proposition 3: [8, 13, 14, 16, 21–23, 27–29, 33, 35, 38, 40], see the historical background in [20] for a discussion of these papers and the techniques introduced along the way. One result we highlight here is that, in 2019, Trojan [38] proved the variation case of Proposition 3. A straightforward consequence is the μ-almost everywhere convergence of the averages $\mathcal{A}_t^{\mathcal{P},k',k''} f$ and $\mathcal{H}_t^{\mathcal{P},k',k''} f$. In this paper, we consider the same averages, but for a different purpose: by proving estimates in the vector-valued setting, we prove oscillation estimates for the multiparameter averages $\mathcal{A}_{\vec{t}} f$, and this again induces pointwise convergence results.

1.3 Historical Background: Multiparameter Problems

In 1951, Dunford [11] and independently Zygmund [41] showed that the two-step procedure can be applied in a multiparameter setting. Even for $S_1, \dots, S_d$ not necessarily commuting, the Dunford–Zygmund ergodic theorem states that the averages

$$A^{n_1,\dots,n_d}_{M_1,\dots,M_d;S_1,\dots,S_d} f(x) := \frac{1}{M_1 \cdots M_d} \sum_{n_1=1}^{M_1} \cdots \sum_{n_d=1}^{M_d} f\big(S_1^{n_1} \cdots S_d^{n_d} x\big), \quad x \in X,$$

converge almost everywhere on X and in $L^p(X)$ norm as $\min\{M_1, \dots, M_d\} \to \infty$ for every $f \in L^p(X)$, $p \in (1, \infty)$. Using the identity

$$A^{n_1,\dots,n_d}_{M_1,\dots,M_d;S_1,\dots,S_d} f = A^{n_1}_{M_1;S_1} \circ \cdots \circ A^{n_d}_{M_d;S_d} f,$$

the $L^p(X)$, $p \in (1, \infty]$, bounds for the strong maximal function $\sup_{M \in \mathbb{N}^d} |A^{n_1,\dots,n_d}_{M_1,\dots,M_d;S_1,\dots,S_d} f|$ follow by applying d times the corresponding $L^p(X)$ bounds for $\sup_{M \in \mathbb{N}} |A^n_{M;S} f|$. This establishes (a), and (b) can be established by a suitable adaptation of the telescoping argument to the multiparameter setting and an application of the classical Birkhoff ergodic theorem, see [31] for more details. We note that the operator $f \mapsto \sup_{M \in \mathbb{N}^d} |A^{n_1,\dots,n_d}_{M_1,\dots,M_d;S_1,\dots,S_d} f|$ is not of weak type $(1, 1)$ in general, so the pointwise convergence may fail if $p = 1$. A model example is $X = \mathbb{Z}^d$ with $S_j x = x - e_j$, $1 \le j \le d$, where e_j is the jth coordinate vector. It is well known that the weak type $(1, 1)$ estimate does not hold for the corresponding strong maximal operator, see [36, Section X.2.3].

After completing [3–5], Bourgain observed that the Dunford–Zygmund ergodic theorem can be extended to the polynomial setting at the expense of imposing that the measure-preserving transformations commute. Bourgain's result can be formulated as follows.

Proposition 5 (Polynomial Dunford–Zygmund ergodic theorem) [24, Theorem 1.25] *Let $M \in \mathbb{N}$, let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space endowed with a family of invertible commuting and measure preserving transformations $S_1, \dots, S_M : X \to X$, and consider a polynomial mapping*

$$\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_M) \colon \mathbb{Z}^M \to \mathbb{Z}^M$$

where each $\mathcal{P}_j : \mathbb{Z} \to \mathbb{Z}$ is a polynomial of one variable with integer coefficients such that $\mathcal{P}_j(0) = 0$. For $f \in L^\infty(X, \mu)$ and $t, t_1, \ldots, t_M \in \mathbb{N}$, we define the associated ergodic averages by

$$A_t^{\mathcal{P}_j} f(x) := \frac{1}{t} \sum_{n=1}^{t} f\left(S_1^{\mathcal{P}_j(n)} x\right), \quad x \in X,$$

and

$$A_{t_1,\ldots,t_M}^{\mathcal{P}_1,\ldots,\mathcal{P}_M} f(x) := A_{t_1}^{\mathcal{P}^1} \circ \cdots \circ A_{t_M}^{\mathcal{P}_M,} f(x)$$
$$= \frac{1}{t_1 \cdots t_M} \sum_{n_1=1}^{t_1} \cdots \sum_{n_M=1}^{t_M} f\left(S_1^{\mathcal{P}_1(n_1)} \cdots S_M^{\mathcal{P}_M(n_M)} x\right), \quad x \in X.$$

Let $p \in (1, \infty)$ and $f \in L^p(X, \mu)$. Then we have:

(i) (Mean ergodic theorem) the averages $A_{t_1,\ldots,t_M}^{\mathcal{P}_1,\ldots,\mathcal{P}_M} f$ converge in $L^p(X, \mu)$ norm as $\min\{t_1, \ldots, t_M\} \to \infty$;
(ii) (Pointwise ergodic theorem) the averages $A_{t_1,\ldots,t_M}^{\mathcal{P}_1,\ldots,\mathcal{P}_M} f$ converge pointwise μ-almost everywhere on X as $\min\{t_1, \ldots, t_M\} \to \infty$;
(iii) (Maximal ergodic theorem) the following maximal estimate holds, including with $p = \infty$:

$$\left\| \sup_{t \in \mathbb{N}^M} \left| A_{t_1,\ldots,t_M}^{\mathcal{P}_1,\ldots,\mathcal{P}_M} f \right| \right\|_{L^p(X,\mu)} \lesssim_{p,M,\deg \mathcal{P}} \|f\|_{L^p(X,\mu)};$$

(iv) (Oscillation ergodic theorem) the following uniform oscillation inequality holds:

$$\sup_{N \in \mathbb{N}} \sup_{I \in \mathfrak{S}_N(\mathbb{R}_+^M)} \left\| O_{I,N}^2 \left(A_{t_1,\ldots,t_M}^{\mathcal{P}_1,\ldots,\mathcal{P}_M} f : t \in \mathbb{N}^M \right) \right\|_{L^p(X)} \lesssim_{p,M,\deg \mathcal{P}} \|f\|_{L^p(X)},$$
$$f \in L^p(X),$$

with implicit constants independent of the coefficients of the polynomial mapping $\mathcal{P}$.

Proposition 5(i)–(iii) is attributed to Bourgain, though it was never published (see [24] for a proof and additional historical notes), and Proposition 5(iv) with linear polynomials $\mathcal{P}_1(t) = \ldots = \mathcal{P}_M(t) = t$ was established in [15]. Corollary 4 is a significant generalization of Proposition 5 (one may check that the proof is easily adaptable to sums taken over $\mathbb{N}$ instead of $\mathbb{Z}$) in that it allows for averages taken over primes and over more general polynomial orbits. Indeed, from the point of view of the permitted polynomial orbits, Corollary 4 is the most one can extend Proposition 5 without having to go beyond averaging operators that can be written as the composition of one-parameter averaging operators. For comparison, proving the analogue of Proposition

5 for averages of the form

$$\frac{1}{t_1 \cdots t_M} \sum_{n_1=1}^{t_1} \cdots \sum_{n_M=1}^{t_M} f\left(S_1^{\mathcal{P}_1(n_1,\ldots,n_M)} \cdots S_M^{\mathcal{P}_M(n_1,\ldots,n_M)} x\right)$$

is a central open problem in modern ergodic theory that can be seen as a multiparameter variant of the Bellow and Furstenberg problem (cf. [24, Conjecture 1.29], and see [6] for some recent progress).

2 Notation and Necessary Tools

2.1 Basic Notation

We denote $\mathbb{N}:=\{1, 2, \ldots\}$, $\mathbb{N}_0:=\{0, 1, 2, \ldots\}$, and $\mathbb{R}_+:=(0, \infty)$. For $d \in \mathbb{N}$, the sets $\mathbb{Z}^d$, $\mathbb{R}^d$, $\mathbb{C}^d$, and $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d \equiv [-1/2, 1/2)^d$ have the standard meanings. For each $N \in \mathbb{N}$, we set

$$\mathbb{N}_N:=\{1, \ldots, N\}, \quad \mathbb{P}_N:=\mathbb{P} \cap \{1, \ldots, N\}.$$

For any $x \in \mathbb{R}$, we set

$$\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \le x\}.$$

For $u \in \mathbb{N}$, we define the set

$$2^{u\mathbb{N}}:=\{2^{un} : n \in \mathbb{N}\}.$$

For two non-negative numbers A and B, we write $A \lesssim B$ to indicate that $A \le CB$ for some $C > 0$ that may change from line to line, and we may write $\lesssim_\delta$ if the implicit constant depends on δ.

We denote the standard inner product on $\mathbb{R}^d$ by $x \cdot \xi$. Moreover, for any $x \in \mathbb{R}^d$, we denote the ℓ^2-norm and the maximum norm respectively by

$$|x|:=|x|_2:=\sqrt{x \cdot x} \quad \text{and} \quad |x|_\infty:= \max_{1\le k\le d} |x_k|.$$

For a multi-index $\gamma = (\gamma_1, \ldots, \gamma_k) \in \mathbb{N}_0^k$, we abuse the notation to write $|\gamma|:=\gamma_1 + \cdots + \gamma_k$. No confusion should arise since all multi-indices will be denoted by γ.

2.2 Rademacher–Menshov Inequality

We recall a basic numerical inequality. A variational version of this inequality was proven by Lewko–Lewko [18, Lemma 13], see also [26, Lemma 2.5, p. 534].

Proposition 6 *For any $k, m \in \mathbb{N}$ with $k < 2^m$ and any sequence of complex numbers $(a_n : n \in \mathbb{N})$, we have*

$$\sup_{k \le n \le 2^m} |a_n| \le |a_k| + \sqrt{2} \sum_{i=1}^{s} \left(\sum_j \left| a_{u^i_{j+1}} - a_{u^i_j} \right|^2 \right)^{1/2}, \tag{2.1}$$

where each $[u^i_j, u^i_{j+1})$ is a dyadic interval contained in $[k, 2^m]$ of the form $[j2^i, (j+1)2^i)$ for some $0 \le i \le m$ and $0 \le j \le 2^{m-i} - 1$.

2.3 Marcinkiewicz–Zygmund Inequality

We recall a result extending the Marcinkiewicz–Zygmund inequality to the Hilbert space setting. Let $(T_m : m \in \mathbb{N}_0)$ be a family of bounded linear operators, $T_m : L^p(X) \to L^p(X)$. For each $\omega \in [0, 1]$, we define

$$T^\omega = \sum_{m \in \mathbb{N}_0} \epsilon_m(\omega) T_m$$

where $(\epsilon_m : m \in \mathbb{N}_0)$ is the sequence of Rademacher functions on $[0, 1]$.

Proposition 7 [22, Lemma 2.1] *Let $p \in (0, \infty)$. Suppose there is a constant $C_p > 0$ such that, for all $\omega \in [0, 1]$ and $f \in L^p(X)$, we have*

$$\|T^\omega f\|_{L^p(X)} \le C_p \|f\|_{L^p(X)},$$

then there is a constant C such that

$$\left\| \left(\sum_{m \in \mathbb{N}_0} |T_m f_{\mathfrak{i}}|^2 \right)^{1/2} \right\|_{L^p(X;\ell^2)} \le C C_p \|f_{\mathfrak{i}}\|_{L^p(X;\ell^2)} \tag{2.2}$$

for every sequence of functions $(f_{\mathfrak{i}})_{\mathfrak{i} \in \mathbb{N}}$ in $L^p(X; \ell^2)$. Moreover, if $T_m \equiv 0$ for all $m \in \mathbb{N}$, then (2.2) recovers the Marcinkiewicz–Zygmund inequality

$$\|T_0 f_{\mathfrak{i}}\|_{L^p(X;\ell^2)} \le C C_p \|f_{\mathfrak{i}}\|_{L^p(X;\ell^2)}. \tag{2.3}$$

2.4 Reductions: Calderón Transference and Lifting

By the Calderón transference principle [7], we may restrict attention to the model dynamical system of $\mathbb{Z}^d$ equipped with the counting measure and the shift operators $S_j : \mathbb{Z}^d \to \mathbb{Z}^d$ given by $S_j(x_1, \ldots, x_d) := (x_1, \ldots, x_j - 1, \ldots, x_d)$. We denote the corresponding averaging operators by

$$A_t^{\mathcal{P},k',k''} f(x) = \frac{1}{\vartheta_\Omega(t)} \sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} f\big(x - \mathcal{P}(n,p)\big)\mathbb{1}_{\Omega_t}(n,p)\left(\prod_{j=1}^{k''} \log|p_j|\right)$$

and

$$H_t^{\mathcal{P},k',k''} f(x) = \sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} f\big(x - \mathcal{P}(n,p)\big)K(n,p)\mathbb{1}_{\Omega_t}(n,p)\left(\prod_{j=1}^{k''} \log|p_j|\right).$$

Moreover, by a standard lifting argument, it suffices to prove Theorem 1 for a canonical case of the polynomial mapping $\mathcal{P}$. Let $\mathcal{P}$ be a polynomial mapping as in (1.1). We define

$$\deg \mathcal{P} := \max\{\deg \mathcal{P}_j : 1 \le j \le d\}$$

and consider the set of multi-indices

$$\Gamma := \big\{\gamma \in \mathbb{N}_0^k \setminus \{0\} : 0 < |\gamma| \le \deg \mathcal{P}\big\}$$

equipped with the lexicographic order. We define the *canonical polynomial mapping* by

$$\mathbb{R}^k \ni x = (x_1, \ldots, x_k) \mapsto \mathcal{Q}(x) := (x^\gamma : \gamma \in \Gamma) \in \mathbb{R}^\Gamma, \tag{2.4}$$

where $x^\gamma = x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_k^{\gamma_k}$. By invoking the lifting procedure described in [22, Lemma 2.2] (see also [36, Chapter XI]), the following implies Theorem 1.

Theorem 8 *Let $k \in \mathbb{N}$, and let $k', k'' \in \{0, 1, \ldots, k\}$ with $k' + k'' = k$. Let $M_t^{k',k''}$ be either $A_t^{\mathcal{Q},k',k''}$ or $H_t^{\mathcal{Q},k',k''}$. For any $p \in (1, \infty)$, there is a constant $C_{p,k,|\Gamma|} > 0$ such that*

$$\left\| \sup_{t>0} \left| M_t^{k',k''} f_{\mathfrak{i}} \right| \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \le C_{p,k,|\Gamma|} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}. \tag{2.5}$$

2.5 Fourier Transform and Ionescu–Wainger Multiplier Theorem

Let $\mathbb{G} = \mathbb{R}^d$ or $\mathbb{G} = \mathbb{Z}^d$ and let $\mathbb{G}^*$ denote the dual group of $\mathbb{G}$. For every $z \in \mathbb{C}$, we set $\boldsymbol{e}(z) := e^{2\pi \boldsymbol{i} z}$, where $\boldsymbol{i}^2 = -1$. Let $\mathcal{F}_\mathbb{G}$ denote the Fourier transform on $\mathbb{G}$ defined for any $f \in L^1(\mathbb{G})$ by

$$\mathcal{F}_\mathbb{G} f(\xi) := \int_\mathbb{G} f(x)\boldsymbol{e}(x \cdot \xi)\mathrm{d}\mu(x), \quad \xi \in \mathbb{G}^*,$$

where μ is the usual Haar measure on $\mathbb{G}$. For any bounded function $\mathfrak{m} \colon \mathbb{G}^* \to \mathbb{C}$, we define the corresponding Fourier multiplier operator by

$$T_\mathbb{G}[\mathfrak{m}]f(x) := \int_{\mathbb{G}^*} \boldsymbol{e}(-\xi \cdot x)\mathfrak{m}(\xi)\mathcal{F}_\mathbb{G} f(\xi)\mathrm{d}\xi, \quad x \in \mathbb{G}. \tag{2.6}$$

Here, we assume that $f : \mathbb{G} \to \mathbb{C}$ is a compactly supported function on $\mathbb{G}$ (and smooth if $\mathbb{G} = \mathbb{R}^d$) or any other function for which (2.6) makes sense.

An indispensable tool in the proof of Theorem 8 is the vector-valued Ionescu–Wainger multiplier theorem from [27, Section 2] with an improvement by Tao [37].

Theorem 9 *For every $\varrho > 0$, there exists a family $(P_{\leq N})_{N\in\mathbb{N}}$ of subsets of $\mathbb{N}$ such that:*

(i) $\mathbb{N}_N \subseteq P_{\leq N} \subseteq \mathbb{N}_{\max\{N, e^{N^\varrho}\}}$.
(ii) If $N_1 \leq N_2$, then $P_{\leq N_1} \subseteq P_{\leq N_2}$.
(iii) If $q \in P_{\leq N}$, then all factors of q also lie in $P_{\leq N}$.
(iv) $\operatorname{lcm}(P_N) \leq 3^N$.

Furthermore, for every $p \in (1, \infty)$, there exists $0 < C_{p,\varrho,|\Gamma|} < \infty$ such that, for every $N \in \mathbb{N}$, the following holds:

Let $0 < \varepsilon_N \leq e^{-N^{2\varrho}}$ and let $\boldsymbol{Q}:=[-1/2, 1/2)^\Gamma$ be a unit cube. Let $\mathfrak{m}\colon \mathbb{R}^\Gamma \to L(H_0, H_1)$ be a measurable function supported on $\varepsilon_N \boldsymbol{Q}$ taking values in $L(H_0, H_1)$, the space of bounded linear operators between separable Hilbert spaces H_0 and H_1. Let $0 \leq \boldsymbol{A}_p \leq \infty$ denote the smallest constant such that

$$\left\| T_{\mathbb{R}^\Gamma}[\mathfrak{m}] f \right\|_{L^p(\mathbb{R}^\Gamma; H_1)} \leq \boldsymbol{A}_p \| f \|_{L^p(\mathbb{R}^\Gamma; H_0)}$$

for every function $f \in L^2(\mathbb{R}^\Gamma; H_0) \cap L^p(\mathbb{R}^\Gamma; H_0)$. Then, the multiplier

$$\Delta_N(\xi) := \sum_{b \in \Sigma_{\leq N}} \mathfrak{m}(\xi - b),$$

where $\Sigma_{\leq N}$ is defined by

$$\Sigma_{\leq N} := \left\{ \frac{a}{q} \in \mathbb{Q}^\Gamma \cap \mathbb{T}^\Gamma : q \in P_{\leq N} \text{ and } \gcd(a, q) = 1 \right\},$$

satisfies

$$\left\| T_{\mathbb{Z}^\Gamma}[\Delta_N] f \right\|_{\ell^p(\mathbb{Z}^\Gamma; H_1)} \leq C_{p,\varrho,|\Gamma|} (\log N) \boldsymbol{A}_p \| f \|_{\ell^p(\mathbb{Z}^\Gamma; H_0)} \tag{2.7}$$

for every $f \in \ell^p(\mathbb{Z}^\Gamma; H_0)$, (cf. [37, Theorem 1.4] which removes the factor of $\log N$ *in the inequality (2.7)).*

2.6 Exponential Sums

In this section, we present some general results concerning the behavior of exponential sums. The following proposition is an enhancement of the variant of Weyl's inequality due to Trojan [38, Theorem 2] that allows us to estimate exponential sums related to a possibly non-differentiable function ϕ, (cf. [27, Theorem A.1]).

Proposition 10 (Weyl's inequality) [20, Proposition 6] *Let $\alpha > 0$, $k \in \mathbb{N}$, and let $\Gamma \subset \mathbb{N}^k \setminus \{0\}$ be a nonempty finite set. Let $\Omega' \subseteq \Omega \subseteq B(0, N) \subset \mathbb{R}^k$ be convex sets and let $\phi\colon \Omega \cap \mathbb{Z}^k \to \mathbb{C}$. There is $\beta_\alpha > 0$ such that, for any $\beta > \beta_\alpha$, if there is a multi-index $\gamma_0 \in \Gamma$ with*

$$\left|\xi_{\gamma_0} - \frac{a}{q}\right| \le \frac{1}{q^2}$$

for some coprime integers a and q with $1 \le a \le q$ and $(\log N)^\beta \le q \le N^{|\gamma_0|}(\log N)^{-\beta}$, then

$$\left|\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} \boldsymbol{e}(\xi \cdot \mathcal{Q}(n, p))\phi(n, p)\mathbb{1}_{\Omega\setminus\Omega'}(n, p)\right| \lesssim N^k \log(N)^{-\alpha} \|\phi\|_{L^\infty(\Omega\setminus\Omega')} + N^k \sup_{\substack{|x-y|\le N(\log N)^{-\alpha}\\ x,y\in\Omega\setminus\Omega'}} |\phi(x) - \phi(y)|.$$

The implicit constant is independent of the function ϕ, the variable ξ, the sets Ω, Ω', and the numbers a, q, and N.

The next result is a generalization of [38, Proposition 4.1] and [38, Proposition 4.2] in the spirit of [27, Proposition 4.18]. For $q \in \mathbb{N}$ and $a \in \mathbb{N}_q^\Gamma$ with $\gcd(a, q) = 1$, the *Gaussian sum* related to the polynomial mapping $\mathcal{Q}$ is given by

$$G(a/q):=\frac{1}{q^{k'}}\frac{1}{\varphi(q)^{k''}}\sum_{x\in\mathbb{N}_q^{k'}}\sum_{y\in A_q^{k''}} \boldsymbol{e}((a/q)\cdot\mathcal{Q}(x, y)), \tag{2.8}$$

where $A_q:=\{a \in \mathbb{N}_q : \gcd(a, q) = 1\}$ and φ is Euler's totient function. There is $\delta > 0$ such that

$$\left|G(a/q)\right| \lesssim q^{-\delta}, \tag{2.9}$$

according to [38, Theorem 3].

Proposition 11 [20, Lemma 7] *Let $N \in \mathbb{N}$ and let $\Omega \subseteq B(0, N) \subset \mathbb{R}^k$ be a convex set or a Boolean combination of finitely many convex sets. Let $\mathcal{K}\colon \mathbb{R}^k \to \mathbb{C}$ be a function supported in Ω with $\mathcal{K}|_\Omega$ continuous. Then, for each $\beta > 0$, there is a constant $c = c_\beta > 0$ such that, for any $q \in \mathbb{N}$ with $1 \le q \le (\log N)^\beta$, $a \in A_q$, and $\xi = a/q + \theta \in \mathbb{R}^\Gamma$, we have*

$$\Bigg|\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}} \boldsymbol{e}\big(\xi \cdot \mathcal{Q}(n, p)\big)\mathcal{K}(n, p)\left(\prod_{i=1}^{k''}\log|p_i|\right) - G(a/q)\int_\Omega \boldsymbol{e}\big((\xi - a/q)\cdot\mathcal{Q}(t)\big)\mathcal{K}(t)\mathrm{d}t\Bigg|$$

$$\lesssim \big[N^{k-1}\|\mathcal{K}\|_{L^\infty(\Omega)}\Big(1+\sum_{\gamma\in\Gamma}|\theta_\gamma|N^{|\gamma|}\Big)+N^k\sup_{\substack{x,y\in\Omega\\|x-y|\le q\sqrt{k}}}|\mathcal{K}(x)-\mathcal{K}(y)|\big]N$$
$$\exp\big(-c\sqrt{\log N}\big).$$

The implied constant is independent of N, a, q, ξ and the kernel $\mathcal{K}$.

We remark that, in the case where $k'=k$ and $k''=0$, the factor of $N\exp\big(-c\sqrt{\log N}\big)$ can be omitted because the proof will no longer apply the Siegel–Walfisz theorem (cf. [27, Proposition 4.18]). Then the weaker kernel condition (1.7) suffices to apply this proposition for Property 3 ahead.

2.7 Multipliers for the Averaging Operators

For a function $f:\mathbb{Z}^\Gamma\to\mathbb{C}$ with finite support, we have

$$A_t^{\mathcal{Q},k',k''}f(x)=T_{\mathbb{Z}^\Gamma}[\mathfrak{m}_t]f(x)\quad\text{and}\quad H_t^{\mathcal{Q},k',k''}f(x)=T_{\mathbb{Z}^\Gamma}[\mathfrak{n}_t]f(x)$$

for the discrete Fourier multipliers

$$\mathfrak{m}_t(\xi):=\frac{1}{\vartheta_\Omega(t)}\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}}\boldsymbol{e}\big(\xi\cdot\mathcal{Q}(n,p)\big)\mathbb{1}_{\Omega_t}(n,p)\left(\prod_{i=1}^{k''}\log|p_i|\right),\quad \xi\in\mathbb{T}^\Gamma,$$

and

$$\mathfrak{n}_t(\xi):=\sum_{(n,p)\in\mathbb{Z}^{k'}\times(\pm\mathbb{P})^{k''}}\boldsymbol{e}(\xi\cdot\mathcal{Q}(n,p))K(n,p)\mathbb{1}_{\Omega_t}(n,p)\left(\prod_{i=1}^{k''}\log|p_i|\right),\quad \xi\in\mathbb{T}^\Gamma.$$

Their continuous counterparts are given by

$$\Phi_t(\xi):=\frac{1}{|\Omega_t|}\int_{\Omega_t}\boldsymbol{e}(\xi\cdot\mathcal{Q}(t))\mathrm{d}t\quad\text{and}\quad\Psi_t(\xi):=\text{p.v.}\int_{\Omega_t}\boldsymbol{e}(\xi\cdot\mathcal{Q}(t))K(t)\mathrm{d}t$$

respectively. To present a unified approach, we write $M_t^{k',k''}$, $\mathfrak{y}_t$, and Θ_t to represent either $A_t^{\mathcal{Q},k',k''}$, $\mathfrak{m}_t$, and Φ_t or $H_t^{\mathcal{Q},k',k''}$, $\mathfrak{n}_t$, and Ψ_t respectively. We now present the key properties of our multiplier operators that will be used in the proof of Theorem 8. Let $N_n:=\lfloor 2^{n^\tau}\rfloor$ for $n\in\mathbb{N}$ and some $\tau\in(0,1]$ adjusted later.

Property 1 For each $\alpha>0$, there is $\beta_\alpha>0$ such that, for any $\beta>\beta_\alpha$ and $n\in\mathbb{N}$, if there is a multi-index $\gamma_0\in\Gamma$ with

$$\left|\xi_{\gamma_0}-\frac{a}{q}\right|\le\frac{1}{q^2}$$

for some coprime integers a and q with $1 \le a \le q$ and $(\log N_n)^\beta \le q \le N_n^{|\gamma_0|}(\log N_n)^{-\beta}$, then

$$|(\mathfrak{y}_{N_n} - \mathfrak{y}_{N_{n-1}})(\xi)| \lesssim C(\log N_n)^{-\alpha}.$$

This follows from Proposition 10 with $\phi(x) \equiv (\vartheta_\Omega(N_n))^{-1}$ for the $\mathfrak{y}_t = \mathfrak{m}_t$ case and with $\phi(x) = K(x)$ for the $\mathfrak{y}_t = \mathfrak{n}_t$ case, noting the size condition (1.4) and the continuity condition (1.6).

Property 2 Let A be the $|\Gamma| \times |\Gamma|$ diagonal matrix with

$$(Av)_\gamma = |\gamma| v_\gamma. \tag{2.10}$$

For any $t > 0$, we set $t^A v := \left(t^{|\gamma|} v_\gamma : \gamma \in \Gamma\right)$. Then

$$\left|\Theta_{N_n}(\xi) - \Theta_{N_{n-1}}(\xi)\right| \lesssim \min\left\{|N_n^A \xi|_\infty, |N_n^A \xi|_\infty^{-1/|\Gamma|}\right\}, \quad \text{for each } n \in \mathbb{N}.$$

In the $\Theta_t = \Phi_t$ case, this follows from the mean value theorem and the standard van der Corput lemma. In the $\Theta_t = \Psi_t$ case, this follows from the cancellation condition (1.5) and [26, Proposition B.2] (see [26, p. 21] for details).

Property 3 For each $\alpha > 0$, $n \in \mathbb{N}$, and $\xi \in \mathbb{T}^\Gamma$ satisfying

$$\left|\xi_\gamma - \frac{a_\gamma}{q}\right| \le N_n^{-|\gamma|} L \qquad \text{for all } \gamma \in \Gamma$$

with $1 \le q \le L$, $a \in A_q^\Gamma$, and $1 \le L \le \exp\left(c\sqrt{\log N_n}\right)(\log N_n)^{-\alpha}$, we have

$$\mathfrak{y}_{N_n}(\xi) - \mathfrak{y}_{N_{n-1}}(\xi) = G(a/q)\left(\Theta_{N_n}(\xi - a/q) - \Theta_{N_{n-1}}(\xi - a/q)\right) + \mathcal{O}\left((\log N_n)^{-\alpha}\right),$$

for some constant $c > 0$ which is independent of n, ξ, a and q.

In the $\mathfrak{y}_t = \mathfrak{m}_t$, $\Theta_t = \Phi_t$ case, this is [38, Property 6]. In the $\mathfrak{y}_t = \mathfrak{n}_t$, $\Theta_t = \Psi_t$ case, this follows from Property 1 alongside Proposition 11 with $\Omega := \Omega_{N_n} \setminus \Omega_{N_{n-1}}$ and $\mathcal{K}(n, p) := K(n, p)\mathbb{1}_\Omega$, noting the size condition (1.4) and the continuity condition (1.6). For details see [38, Lemmas 3 and 5].

2.8 Parameters Discussion

Let $p \in (1, \infty)$ be fixed and let $\chi \in (0, 1/10)$. Fix τ with $0 < \tau < 1 - \min(2, p)^{-1}$ and let $N_n := \lfloor 2^{n^\tau} \rfloor$ for $n \in \mathbb{N}$. If $p \in (1, 2)$, fix p_0 such that $1 < p_0 < p$. If instead $p \in (2, \infty)$, fix $p_0 > p$. If $p = 2$, the discussion is moot since all the interpolation arguments in the article become unnecessary. We choose ρ with

$$\rho > \frac{1}{\tau} \frac{pp_0 - 2p}{2p_0 - 2p}$$

so that interpolation of the estimates

$$\|T\|_{\ell^2} \lesssim n^{-\rho\tau} \quad \text{and} \quad \|T\|_{\ell^{p_0}} \lesssim 1$$

yields

$$\|T\|_{\ell^p} \lesssim n^{-(1+\varepsilon)} \text{ for some } \varepsilon > 0.$$

Property 1 gives us a corresponding β_ρ. We fix a choice of $\beta > \beta_\rho$ and then fix a choice of $u \in \mathbb{N}$ with $u > |\Gamma|\beta$. We also have the value of δ coming from the Gaussian sum estimate (2.9). With these fixed, we choose the value of ϱ in Theorem 9 to be

$$\varrho := \min\left(\frac{\chi}{10u}, \frac{\delta}{8\tau}\right).$$

3 Proof of Theorem 8

By the monotone convergence theorem and standard density arguments, it is enough to prove that

$$\left\| \sup_{t\in\mathbb{I}} \left|M_t^{k',k''} f_{\mathrm{i}}\right| \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim_{p,k,|\Gamma|} \|f_{\mathrm{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

holds for every finite subset $\mathbb{I} \subset \mathbb{R}_+$ with the implicit constant independent of the set $\mathbb{I}$. For any $t_0 \in \mathbb{I}$, we have

$$\sup_{t\in\mathbb{I}} \left|M_t^{k',k''} f_{\mathrm{i}}\right| \le \sup_{t\in\mathbb{I}} \left|\left(M_t^{k',k''} - M_{t_0}^{k',k''}\right) f_{\mathrm{i}}\right| + \left|M_{t_0}^{k',k''} f_{\mathrm{i}}\right|,$$

so,

$$\left\| \sup_{t\in\mathbb{I}} \left|M_t^{k',k''} f_{\mathrm{i}}\right| \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim \mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(M_t^{k',k''} f_{\mathrm{i}} : t \in \mathbb{I}\right) + \left\|M_{\min\mathbb{I}}^{k',k''} f_{\mathrm{i}}\right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

Let $\mathbb{I} \subset \mathbb{R}_+$ with $\#\mathbb{I} < \infty$, let E be either of $\mathbb{R}^d$ or $\mathbb{Z}^d$ with the usual measures, and let $(f_{\mathrm{i},t} : \mathrm{i} \in \mathbb{N}) \in L^p(E, \ell^2)$ for all $t \in \mathbb{I}$. We define

$$\mathcal{S}^p_E(f_{\mathrm{i},t} : t \in \mathbb{I}) := \left\| \sup_{t\in\mathbb{I}} |f_{\mathrm{i},t} - f_{\mathrm{i},\min\mathbb{I}}| \right\|_{L^p(E;\ell^2)} = \left\| \left(\sum_{\mathrm{i}=1}^\infty \sup_{t\in\mathbb{I}} |f_{\mathrm{i},t} - f_{\mathrm{i},\min\mathbb{I}}|^2 \right)^{1/2} \right\|_{L^p(E)}.$$

As we shall see, working with this rather than the usual maximal function is just a technical adaptation to be more similar to the variation, jump, and oscillation quantities that have been studied before. With this notation established, it suffices to show that

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}(M_t^{k',k''} f_{\mathfrak{i}} : t \in \mathbb{I}) \lesssim \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

We start by splitting (cf. [17, Lemma 1.3], [23, Lemma 8.1]) into long and short suprema along the subexponential sequence N_n. Letting $\mathbb{I}_n := [N_n, N_{n+1}) \cap \mathbb{I}$, we have

$$\begin{aligned}\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(M_t^{k',k''} f_{\mathfrak{i}} : t \in \mathbb{I}\right) &\lesssim \mathcal{S}^p_{\mathbb{Z}^\Gamma}(T_{\mathbb{Z}^\Gamma}[\mathfrak{y}_{N_n}] f_{\mathfrak{i}} : n \in \mathbb{N}_0) \\ &+ \left\| \left(\sum_{n\in\mathbb{N}_0} \sup_{t\in\mathbb{I}_n} \left| (M_t^{k',k''} - M_{\min \mathbb{I}_n}^{k',k''}) f_{\mathfrak{i}} \right|^2 \right)^{1/2} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.\end{aligned}$$

3.1 Short Suprema

Let $s_{n,0} < s_{n,1} < \ldots < s_{s,J(n)}$ be the increasing enumeration of $[N_n, N_{n+1}] \cap \mathbb{I}$ and let $r = \min(2, p)$. Monotonicity of ℓ^p norms, Minkowski's inequality, and the triangle inequality give

$$\begin{aligned}&\left\| \left(\sum_{n\in\mathbb{N}_0} \sup_{t\in\mathbb{I}_n} \left| (M_t^{k',k''} - M_{\min \mathbb{I}_n}^{k',k''}) f_{\mathfrak{i}} \right|^2 \right)^{1/2} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \\ &\le \left(\sum_{n\in\mathbb{N}_0} \left(\sum_{j=1}^{J(n)} \left\| (M_{s_{n,j}}^{k',k''} - M_{s_{n,j-1}}^{k',k''}) f_{\mathfrak{i}} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \right)^r \right)^{1/r}. \qquad (3.1)\end{aligned}$$

Since

$$\left\| (M_{s_{n,j}}^{k',k''} - M_{s_{n,j-1}}^{k',k''}) f \right\|_{\ell^p(\mathbb{Z}^\Gamma)} \le \| \check{\mathfrak{y}}_{s_{n,j}} - \check{\mathfrak{y}}_{s_{n,j-1}} \|_{\ell^1(\mathbb{Z}^\Gamma)} \|f\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

by Young's convolution inequality, (2.3) gives

$$\left\| (M_{s_{n,j}}^{k',k''} - M_{s_{n,j-1}}^{k',k''}) f_{\mathfrak{i}} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \le \| \check{\mathfrak{y}}_{s_{n,j}} - \check{\mathfrak{y}}_{s_{n,j-1}} \|_{\ell^1(\mathbb{Z}^\Gamma)} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

Therefore, we control the right hand side of (3.1) by

$$\begin{aligned}&\left(\sum_{n\in\mathbb{N}_0} \left(\left\| \sum_{j=1}^{J(n)} |\check{\mathfrak{y}}_{s_{n,j}} - \check{\mathfrak{y}}_{s_{n,j-1}}| \right\|_{\ell^1(\mathbb{Z}^\Gamma)} \right)^r \right)^{1/r} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \\ &\lesssim \left(\sum_{n\in\mathbb{N}_0} (n^{-r(1-\tau)} \right)^{1/r} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \\ &\lesssim \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.\end{aligned}$$

The last estimates follow from [20, Eq. 4.2] with $f = \delta_0$ and the discussion thereafter.

3.2 Long Suprema and the Circle Method

Let $\eta\colon \mathbb{R}^\Gamma \to [0,1]$ be a smooth function with

$$\eta(x) = \begin{cases} 1 & \text{if } |x|_\infty \le \frac{1}{32|\Gamma|}, \\ 0 & \text{if } |x|_\infty \ge \frac{1}{16|\Gamma|}. \end{cases}$$

For $N \in \mathbb{R}_+$, we define the scaling notation

$$\eta_N(\xi) := \eta\big(2^{N\cdot A - N^\chi \cdot \mathrm{Id}}\xi\big)$$

where A is the matrix given in (2.10) and Id is the $|\Gamma| \times |\Gamma|$ identity matrix. For dyadic integers $s \in 2^{u\mathbb{N}}$, we define the *annuli sets of fractions* by

$$\Sigma_s := \begin{cases} \Sigma_{\le s} & \text{if } s = 2^u, \\ \Sigma_{\le s} \setminus \Sigma_{\le s/2^u} & \text{if } s > 2^u, \end{cases} \tag{3.2}$$

where the $\Sigma_{\le \cdot}$ are the sets of Ionescu–Wainger fractions as in Theorem 9. For $t \ge 2^u$, we set $F(t) := \max\{s \in 2^{u\mathbb{N}} : s \le t\}$. We define

$$\Xi_{\le j^{\tau u}}(\xi) := \sum_{a/q \in \Sigma_{\le F(j^{\tau u})}} \eta_{j^\tau}(\xi - a/q)$$

and, for $s \in 2^{u\mathbb{N}}$, we define the *annuli functions*

$$\Xi_j^s(\xi) := \sum_{a/q \in \Sigma_s} \eta_{j^\tau}(\xi - a/q). \tag{3.3}$$

By (3.2), we have the telescoping property

$$\Xi_{\le j^{\tau u}} = \sum_{\substack{s \in 2^{u\mathbb{N}} \\ s \le j^{\tau u}}} \Xi_j^s.$$

Note that $\eta_{j^\tau}(\xi)$ satisfies the hypothesis about the support for $\mathfrak{m}$ in Theorem 9 since $\frac{1}{8|\Gamma|} 2^{-j^\tau + j^{\tau\chi}} \le e^{-j^{2\tau u \varrho}}$ provided that $\varrho \le \chi/(10u)$. Using the $\Xi_{\le j^{\tau u}}$ functions, we bound the long suprema by

$$\begin{aligned} &\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{j=1}^n T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j} - \mathfrak{y}_{N_{j-1}})\Xi_{\le j^{\tau u}}]f_{\mathfrak{i}} : n \in \mathbb{N}\right) \\ &\quad + \mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{j=1}^n T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j} - \mathfrak{y}_{N_{j-1}})(1 - \Xi_{\le j^{\tau u}})]f_{\mathfrak{i}} : n \in \mathbb{N}\right). \end{aligned}$$

These terms correspond to major and minor arcs respectively.

3.3 Minor Arcs

Monotonicity of ℓ^p norms, telescoping, and the triangle inequality give

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{j=1}^{n} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})(1-\Xi_{\leq j^{\tau u}})]f_{\mathfrak{i}} : n\in\mathbb{N}\right)$$
$$\leq \sum_{n=1}^{\infty}\left\|T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n}-\mathfrak{y}_{N_{n-1}})(1-\Xi_{\leq n^{\tau u}})]f_{\mathfrak{i}}\right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

It then suffices to show that

$$\left\|T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n}-\mathfrak{y}_{N_{n-1}})(1-\Xi_{\leq n^{\tau u}})]f_{\mathfrak{i}}\right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim n^{-(1+\varepsilon)}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

for some $\varepsilon > 0$. This follows from

$$\left\|T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n}-\mathfrak{y}_{N_{n-1}})(1-\Xi_{\leq n^{\tau u}})]f\right\|_{\ell^p(\mathbb{Z}^\Gamma)} \lesssim n^{-(1+\varepsilon)}\|f\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

by (2.3). This uses Property 1 and follows from the proof of [38, Eqs. (5.8), (5.9)] with only small changes due to our differing scaling in the definition of $\eta_N(\xi)$. We omit the details.

3.4 Introduction to Major Arcs

Using the annuli multipliers (3.3) and the triangle inequality, we bound the major arcs term by

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{j=1}^{n}\sum_{\substack{s\in 2^{u\mathbb{N}}\\ s\leq j^{\tau u}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j]f_{\mathfrak{i}} : n\in\mathbb{N}\right)$$
$$\leq \sum_{s\in 2^{u\mathbb{N}}}\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\leq j\leq n\\ j\geq s^{1/(\tau u)}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j]f_{\mathfrak{i}} : n\geq s^{1/\tau u}\right).$$

It then suffices to show for large $s \in 2^{u\mathbb{N}}$ that

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\le j\le n\\ j\ge s^{1/(\tau u)}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j]f_{\mathfrak{i}} : n \ge s^{1/\tau u}\right) \lesssim s^{-\varepsilon}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \tag{3.4}$$

for some $\varepsilon > 0$ since $\sum_{s\in 2^{u\mathbb{N}}} s^{-\varepsilon} < \infty$. Let $\kappa_s := s^{2\lfloor \varrho \rfloor}$. By splitting the left hand side of (3.4) at $n \approx 2^{\kappa_s}$ into small and large scales, it suffices to prove that

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\le j\le n\\ j\ge s^{1/(\tau u)}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j]f_{\mathfrak{i}} : n^\tau \in \left[s^{1/u}, 2^{\kappa_s+1}\right]\right) \lesssim s^{-\varepsilon}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \tag{3.5}$$

and

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\le j\le n\\ j\ge 2^{\kappa_s/\tau}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j]f_{\mathfrak{i}} : n^\tau > 2^{\kappa_s}\right) \lesssim s^{-\varepsilon}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}. \tag{3.6}$$

For the small scales (3.5), we will use the Rademacher–Menshov inequality (2.1) and Theorem 9. For the large scales (3.6), we will use the Magyar–Stein–Wainger sampling principle from [19, Proposition 2.1] and its counterpart for the jump inequality from [25, Theorem 1.7]. We first recall an approximation lemma to replace our discrete multipliers with continuous counterparts. Let

$$v^s_j(\xi) := \sum_{a/q\in\Sigma_s} G(a/q)\big(\Theta_{N_j}-\Theta_{N_{j-1}}\big)(\xi - a/q)\eta_{j^\tau}(\xi - a/q) \tag{3.7}$$

and

$$\Lambda^s_j(\xi) := \sum_{a/q\in\Sigma_s} \big(\Theta_{N_j}-\Theta_{N_{j-1}}\big)(\xi - a/q)\eta_{j^\tau}(\xi - a/q). \tag{3.8}$$

Lemma 12 [20, Lemma 8] *Let $M \in \mathbb{N}$, $\alpha' > 0$, and $S_M := \lfloor 2^{M^\tau - 3M^{\tau\chi}} \rfloor$. For $j \in \mathbb{N}$ with $s^{1/(\tau u)} \le j$ and $M \le j \le 2M$, we have*

$$\|(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j - v^s_j\|_{\ell^\infty(\mathbb{T}^\Gamma)} \lesssim j^{-\alpha'\tau} \tag{3.9}$$

and

$$\|(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j - \Lambda^s_j \mathfrak{m}_{S_M}\|_{\ell^\infty(\mathbb{T}^\Gamma)} \lesssim j^{-\alpha'\tau}. \tag{3.10}$$

3.5 Small Scales

Splitting $[s^{1/u}, 2^{\kappa_s+1}]$ into dyadic intervals and preparing via the triangle inequality to use (3.10), we bound the left hand side of (3.5) by

$$\overbrace{\sum_{M\in 2^{\mathbb{N}}\cap[s^{1/u},2^{\kappa_s}]} \mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\le j\le n\\ j\ge s^{1/(\tau u)}}} T_{\mathbb{Z}^\Gamma}[\Lambda^s_j \mathfrak{m}_{S_M}]f_{\mathfrak{i}} : n^\tau \in [M,2M]\right)}^{\text{Main Term 1}}$$
$$+\underbrace{\sum_{M\in 2^{\mathbb{N}}\cap[s^{1/u},2^{\kappa_s}]} \mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(\sum_{\substack{1\le j\le n\\ j\ge s^{1/(\tau u)}}} T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_j}-\mathfrak{y}_{N_{j-1}})\Xi^s_j - \Lambda^s_j \mathfrak{m}_{S_M}]f_{\mathfrak{i}} : n^\tau \in [M,2M]\right)}_{\text{Error Term 1}}.$$

For Error Term 1, it will suffice to show that

$$\left\| T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n}-\mathfrak{y}_{N_{n-1}})\Xi^s_n - \Lambda^s_n \mathfrak{m}_{S_M}]f_{\mathfrak{i}}\right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim n^{-(1+\varepsilon')}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

for some $\varepsilon' > 0$ since we would then bound it by

$$\sum_{n\ge s^{1/(\tau u)}} n^{-(1+\varepsilon')}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim s^{-\varepsilon'/(\tau u)}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim s^{-\varepsilon}\|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

This follows from

$$\left\| T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n}-\mathfrak{y}_{N_{n-1}})\Xi^s_n - \Lambda^s_n \mathfrak{m}_{S_M}]f\right\|_{\ell^p(\mathbb{Z}^\Gamma)} \lesssim n^{-(1+\varepsilon')}\|f\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

by (2.3), and that is [20, Eq. 4.12].

For Main Term 1, we apply the Rademacher–Menshov inequality (2.1) to bound it by

$$\sum_{M\in 2^{\mathbb{N}}\cap[s^{1/u},2^{\kappa_s}]} \sum_{i=0}^{\log_2(2M)} \left\| \left(\sum_j \left| \sum_{k\in I^M_{i,j}} T_{\mathbb{Z}^\Gamma}[\Lambda^s_k \mathfrak{m}_{S_M}]f_{\mathfrak{i}}\right|^2\right)^{1/2}\right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)},$$

where j is taken over $j \ge 0$ such that $I^M_{i,j} := [j2^i, (j+1)2^i] \cap [M^{1/\tau}, (2M)^{1/\tau}] \ne \emptyset$. Let $\tilde{\eta}_N(\xi) := \eta_N(\xi/2)$. Then $\tilde{\eta}_N \eta_{k^\tau} = \eta_{k^\tau}$ for $k^\tau \ge N$ due to the nesting supports. This lets us write

$$\Lambda^s_k \mathfrak{m}_{S_M} = \Lambda^s_k \mathfrak{m}_{S_M} \sum_{a/q\in\Sigma_s} \tilde{\eta}_M(\xi - a/q) =: \Lambda^s_k \mathfrak{m}_{S_M} \tilde{\Xi}^s_{M^{1/\tau}}$$

for $k \in I_{i,j}^M$ since then $k \geq M^{1/\tau}$.

By (2.2), it suffices to get an appropriate estimate for

$$\left\| \sum_j \sum_{k \in I_{i,j}^M} \epsilon_j(\omega) T_{\mathbb{Z}^\Gamma}[\Lambda_k^s \mathfrak{m}_{S_M}] f \right\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

for any Rademacher sequence $\epsilon = (\epsilon_j(\omega))$ with $\epsilon_j(\omega) \in \{-1, 1\}$ and for every $\omega \in [0, 1]$.

We get the appropriate bound on $\ell^p(\mathbb{Z}^\Gamma)$ by the Ionescu–Wainger theorem and the bound for the continuous analogue

$$\left\| \sum_j \sum_{k \in I_{i,j}^M} \epsilon_j(\omega) T_{\mathbb{Z}^\Gamma}[(\Theta_{N_k} - \Theta_{N_{k-1}}) \eta_k^\tau] g \right\|_{L^p(\mathbb{R}^\Gamma)} \lesssim \|g\|_{L^p(\mathbb{R}^\Gamma)}$$

with a bound independent of the Rademacher sequence ϵ, see [36, Chapter XI] or [9]. Therefore,

$$\left\| \sum_j \sum_{k \in I_{i,j}^M} \epsilon_j(\omega) T_{\mathbb{Z}^\Gamma}[\Lambda_k^s \mathfrak{m}_{S_M}] f \right\|_{\ell^{p_0}(\mathbb{Z}^\Gamma)} \lesssim \|T_{\mathbb{Z}^\Gamma}[\mathfrak{m}_{S_M}] f\|_{\ell^{p_0}(\mathbb{Z}^\Gamma)} \lesssim \|f\|_{\ell^{p_0}(\mathbb{Z}^\Gamma)} \tag{3.11}$$

using the uniform ℓ^p-boundedness of the averaging operators.

We get an improved bound on ℓ^2. To do this, we use that

$$\|\mathfrak{m}_{S_M} \tilde{\Xi}_{M^{1/\tau}}^s\|_{\ell^\infty(\mathbb{T}^\Gamma)} \lesssim s^{-\delta}$$

for $M \in 2^{\mathbb{N}} \cap [s^{1/u}, 2^{\kappa_s}]$, see [20, Section 4.5]. Then

$$\left\| \sum_j \sum_{k \in I_{i,j}^M} \epsilon_j(\omega) T_{\mathbb{Z}^\Gamma}[\Lambda_k^s \mathfrak{m}_{S_M}] f \right\|_{\ell^2(\mathbb{Z}^\Gamma)} \lesssim \left\| T_{\mathbb{Z}^\Gamma}\left[\mathfrak{m}_{S_M} \tilde{\Xi}_{M_{1/\tau}}^s\right] f \right\|_{\ell^2(\mathbb{Z}^\Gamma)}$$
$$\lesssim s^{-\delta} \|f\|_{\ell^2(\mathbb{Z}^\Gamma)}. \tag{3.12}$$

Interpolation of (3.11) with (3.12) then gives that

$$\left\| \sum_j \sum_{k \in I_{i,j}^M} \epsilon_j(\omega) T_{\mathbb{Z}^\Gamma}[\Lambda_k^s \mathfrak{m}_{S_M}] f \right\|_{\ell^p(\mathbb{Z}^\Gamma)} \lesssim s^{-8\varrho} \|f\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

since $8\varrho \le \delta/(\rho\tau)$. We then apply (2.2)

$$\left\| \left(\sum_j \left| \sum_{k \in I_{i,j}^M} T_{\mathbb{Z}^\Gamma}[\Lambda_k^s \mathfrak{m}_{S_M}] f_{\mathfrak{i}} \right|^2 \right)^{1/2} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

Thus, we may dominate Main Term 1 by

$$\sum_{M \in 2^{\mathbb{N}} \cap [s^{1/u}, 2^{\kappa_s}]} \sum_{i=0}^{\log_2(2M)} s^{-8\varrho} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim \kappa_s^2 s^{-8\varrho} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$
$$\lesssim s^{-4\varrho} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

since $\kappa_s \le s^{2\varrho}$, concluding the proof of (3.5).

3.6 Large Scales

We bound the left hand side of (3.6) by

$$\overbrace{\mathcal{S}_{\mathbb{Z}^\Gamma}^p \left(\sum_{\substack{1 \le j \le n \\ j \ge 2^{\kappa_s/\tau}}} T_{\mathbb{Z}^\Gamma}[v_j^s] f_{\mathfrak{i}} : n^\tau > 2^{\kappa_s} \right)}^{\text{Main Term 2}} + \overbrace{\sum_{n \ge 2^{\kappa_s/\tau}} \left\| T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n} - \mathfrak{y}_{N_{n-1}}) \Xi_n^s - v_n^s] f_{\mathfrak{i}} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}}^{\text{Error Term 2}}.$$

For Error Term 2, it will suffice to show that

$$\left\| T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n} - \mathfrak{y}_{N_{n-1}}) \Xi_n^s - v_n^s] f_{\mathfrak{i}} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim e^{(|\Gamma|+1)s^\varrho} n^{-(1+\varepsilon')} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$

for some $\varepsilon' > 0$ since we would then bound it by

$$e^{(|\Gamma|+1)s^\varrho} \sum_{n \ge 2^{\kappa_s/\tau}} n^{-(1+\varepsilon')} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim e^{(|\Gamma|+1)s^\varrho} 2^{-s^{2\varrho}\varepsilon'/\tau} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}$$
$$\lesssim s^{-\varepsilon} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)}.$$

This follows from

$$\left\| T_{\mathbb{Z}^\Gamma}[(\mathfrak{y}_{N_n} - \mathfrak{y}_{N_{n-1}}) \Xi_n^s - v_n^s] f \right\|_{\ell^p(\mathbb{Z}^\Gamma)} \lesssim e^{(|\Gamma|+1)s^\varrho} n^{-(1+\varepsilon')} \|f\|_{\ell^p(\mathbb{Z}^\Gamma)},$$

by (2.3), and that is [20, Eq. 4.15].

For Main Term 2, we define

$$w^s(\xi) := \sum_{a/q \in \Sigma_s} G(a/q)\tilde{\eta}_{2^{\kappa_s}}(\xi - a/q), \quad \Pi^s(\xi) := \sum_{a/q \in \Sigma_s} \tilde{\eta}_{2^{\kappa_s}}(\xi - a/q),$$

and

$$\omega_n^s(\xi) := \sum_{2^{\kappa_s/\tau} \le j \le n} (\Theta_{N_j} - \Theta_{N_{j-1}})(\xi)\eta_{j^\tau}(\xi).$$

Let $Q_s := \mathrm{lcm}(q : a/q \in \Sigma_s)$. By property *(iv)* from Theorem 9, we have $Q_s \le 3^s$. The function ω_n^s is supported on $[-\frac{1}{4Q_s}, \frac{1}{4Q_s}]$ for large $s \in 2^{u\mathbb{N}}$ since, on the support of $\eta_{2^{\kappa_s}}$, we have $|\xi_\gamma| \le 2^{-2^{-\kappa_s} + 2^{\kappa_s \chi}} \le (4Q_s)^{-1}$ for all $\gamma \in \Gamma$ and large s. We also have

$$\sum_{2^{\kappa_s/\tau} \le j \le n} v_j^s(\xi) = w^s(\xi) \sum_{b \in \mathbb{Z}^\Gamma} \omega_n^s(\xi - b/Q_s).$$

Therefore, it suffices to prove

$$\mathcal{S}^p_{\mathbb{Z}^\Gamma}\left(T_{\mathbb{Z}^\Gamma}\left[\sum_{b \in \mathbb{Z}^\Gamma} \omega_n^s(\cdot - b/Q_s)\right] f_{\mathfrak{i}} : n^\tau > 2^{\kappa_s}\right) \lesssim \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \tag{3.13}$$

and

$$\left\| T_{\mathbb{Z}^\Gamma}[w^s] f_{\mathfrak{i}} \right\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \lesssim s^{-\varepsilon} \|f_{\mathfrak{i}}\|_{\ell^p(\mathbb{Z}^\Gamma;\ell^2)} \tag{3.14}$$

for some $\varepsilon > 0$. (3.14) follows from

$$\left\| T_{\mathbb{Z}^\Gamma}[w^s] f \right\|_{\ell^p(\mathbb{Z}^\Gamma)} \lesssim s^{-\varepsilon} \|f\|_{\ell^p(\mathbb{Z}^\Gamma)}$$

by (2.3), and that is [20, Eq. 4.17].

By the Magyar–Stein–Wainger sampling principle [19, Proposition 2.1] for the supremum seminorm, (3.13) follows from

$$\mathcal{S}^p_{\mathbb{R}^\Gamma}\left(T_{\mathbb{R}^\Gamma}[\omega_n^s] f_{\mathfrak{i}} : n^\tau > 2^{\kappa_s}\right) \lesssim \|f_{\mathfrak{i}}\|_{L^p(\mathbb{R}^\Gamma;\ell^2)}. \tag{3.15}$$

To prove (3.15), we use that the ω_n^s functions are almost telescoping. We define

$$\Delta_n^s(\xi) := \sum_{2^{\kappa_s/\tau} \le j \le n} (\Theta_{N_j} - \Theta_{N_{j-1}})(\xi) = (\Theta_{N_n} - \Theta_{N_{2^{\kappa_s/\tau}-1}})(\xi).$$

Then (3.15) follows from

$$\mathcal{S}^p_{\mathbb{R}^\Gamma}(T_{\mathbb{R}^\Gamma}[\Delta_n^s] f_{\mathfrak{i}} : n^\tau > 2^{\kappa_s}) \lesssim \|f_{\mathfrak{i}}\|_{L^p(\mathbb{R}^\Gamma;\ell^2)} \tag{3.16}$$

since the error term is bounded by

$$\sum_{n>2^{\kappa s/\tau}} \left\| T_{\mathbb{R}^\Gamma}[(\Theta_{N_n} - \Theta_{N_{n-1}})(\eta_{n^\tau} - 1)] f_{\mathfrak{i}} \right\|_{L^p(\mathbb{R}^\Gamma;\ell^2)} \lesssim \|f_{\mathfrak{i}}\|_{L^p(\mathbb{R}^\Gamma;\ell^2)}.$$

This last estimate follows from

$$\sum_{n>2^{\kappa s/\tau}} \left\| T_{\mathbb{R}^\Gamma}[(\Theta_{N_n} - \Theta_{N_{n-1}})(\eta_{n^\tau} - 1)] f \right\|_{L^p(\mathbb{R}^\Gamma)} \lesssim \|f\|_{L^p(\mathbb{R}^\Gamma)}$$

by (2.3), and this follows from Property 2 and interpolation. (3.16) itself follows from

$$\mathcal{S}^p_{\mathbb{R}^\Gamma}(T_{\mathbb{R}^\Gamma}[\Theta_t] f_{\mathfrak{i}} : t > 0) \lesssim \|f_{\mathfrak{i}}\|_{L^p(\mathbb{R}^\Gamma;\ell^2)},$$

and this follows from [22, Appendix A].

Funding The author was partially supported by the National Science Foundation (NSF) Grant DMS-2154712.

References

1. Bellow, A.: Measure theory Oberwolfach 1981. In: Kölzow, D., Maharam-Stone, D. (eds.) Proceedings of the Conference Held at Oberwolfach, June 21-27, 1981. Volume 945 of Lecture Notes in Mathematics. Section: Two Problems Submitted by A. Bellow, pp. 429–431. Springer-Verlag, Berlin, Heidelberg (1982)
2. Birkhoff, G.: Proof of the ergodic theorem. Proc. Natl. Acad. Sci. U.S.A. **17**(12), 656–660 (1931)
3. Bourgain, J.: On the maximal ergodic theorem for certain subsets of the integers. Isr. J. Math. **61**, 39–72 (1988)
4. Bourgain, J.: On the pointwise ergodic theorem on L^p for arithmetic sets. Isr. J. Math. **61**, 73–84 (1988)
5. Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, H. Furstenberg, Y. Katznelson, and D.S. Ornstein. Inst. Hautes Etudes Sci. Publ. Math. **69**, 5–45 (1989)
6. Bourgain, J., Mirek, M., Stein, E.M., Wright, J.: On a multi-parameter variant of the Bellow–Furstenberg problem. Preprint (2022)
7. Calderón, A.: Ergodic theory and translation invariant operators. Proc. Natl. Acad. Sci. U.S.A. **59**, 349–353 (1968)
8. Campbell, J.T., Jones, R.L., Reinhold, K., Wierdl, M.: Oscillation and variation for the Hilbert transform. Duke Math. J. **105**(1), 59–83 (2000)
9. Duoandikoetxea, J., Rubio de Francia, J.L.: Maximal and singular integral operators via Fourier transform estimates. Invent. Math. **84**(3), 541–561 (1986)
10. Cotlar, M.: A unified theory of Hilbert transforms and ergodic theorems. Rev. Mat. Cuyana **1**, 105–167 (1955)
11. Dunford, N.: An individual ergodic theorem for non-commutative transformations. Acta Sci. Math. Szeged **14**, 1–4 (1951)

12. Furstenberg, H.: Problems Session, Conference on Ergodic Theory and Applications. University of New Hampshire, Durham (1982)
13. Ionescu, A.D., Wainger, S.: L^p boundedness of discrete singular Radon transforms. J. Am. Math. Soc. **19**(2), 357–383 (2005)
14. Jones, R.L., Kaufman, R., Rosenblatt, J.M., Wierdl, M.: Oscillation in ergodic theory. Ergodic Theory Dyn. Syst. **18**(4), 889–935 (1998)
15. Jones, R.L., Rosenblatt, J.M., Wierdl, M.: Oscillation inequalities for rectangles. Proc. Am. Math. Soc. **129**(5), 1349–1358 (2001)
16. Jones, R.L., Rosenblatt, R., Wierdl, M.: Oscillation in ergodic theory: higher dimensional results. Isr. J. Math. **135**, 1–27 (2003)
17. Jones, R.L., Seeger, A., Wright, J.: Strong variational and jump inequalities in harmonic analysis. Trans. Am. Math. Soc. **360**(12), 6711–6742 (2008)
18. Lewko, A., Lewko, M.: Estimates for the square variation of partial sums of Fourier series and their rearrangements. J. Funct. Anal. **262**(6), 2561–2607 (2012)
19. Magyar, A., Stein, E.M., Wainger, S.: Discrete analogues in harmonic analysis: spherical averages. Ann. Math. **155**, 189–208 (2002)
20. Mehlhop, N., Słomian, W.: Oscillation and jump inequalities for the polynomial ergodic averages along multi-dimensional subsets of primes. Math. Ann. (2023). https://doi.org/10.1007/s00208-023-02597-8
21. Mirek, M., Słomian, W., Szarek, T.Z.: Some remarks on oscillation inequalities. Ergodic Theory Dyn. Syst. (2023). https://doi.org/10.1017/etds.2022.77
22. Mirek, M., Stein, E.M., Trojan, B.: $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon types I: maximal functions and vector-valued estimates. J. Funct. Anal. **277**, 2471–2892 (2019)
23. Mirek, M., Stein, E.M., Trojan, B.: $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: variational estimates. Invent. Math. **209**(3), 665–748 (2017)
24. Mirek, M., Szarek, T.Z., Wright, J.: Oscillation inequalities in ergodic theory and analysis: one-parameter and multi-parameter perspective. Rev. Mat. Iberoam. **38**, 2249–2284 (2022)
25. Mirek, M., Stein, E.M., Zorin-Kranich, P.: Jump inequalities via real interpolation. Math. Ann. **376**, 797–819 (2020)
26. Mirek, M., Stein, E.M., Zorin-Kranich, P.: A bootstrapping approach to jump inequalities and their applications. Anal. PDE **13**(2), 527–558 (2020)
27. Mirek, M., Stein, E.M., Zorin-Kranich, P.: Jump inequalities for translation-invariant operators of Radon type on $\mathbb{Z}^d$. Adv. Math. **365**, 107065 (2020)
28. Mirek, M., Trojan, B.: Cotlar's ergodic theorem along the prime numbers. J. Fourier Anal. Appl. **21**, 822–848 (2015)
29. Mirek, M., Trojan, B.: Discrete maximal functions in higher dimensions and applications to ergodic theory. Am. J. Math. **138**(6), 1495–1532 (2016)
30. von Neumann, J.: Proof of the quasi-ergodic hypothesis. Proc. Natl. Acad. Sci. U.S.A. **18**, 70–82 (1932)
31. Nevo, A.: Chapter 13—pointwise ergodic theorems for actions of groups. In: Katok, A., Hasselblatt, B. (eds.) Handbook of Dynamical Systems, vol. 1B. Elsevier Science, New York (2005)
32. Riesz, F.: Some mean ergodic theorems. J. Lond. Math. Soc. **13**, 274–278 (1938)
33. Rosenblatt, J., Wierdl, M.: Pointwise ergodic theorems via harmonic analysis. In: Proceedings of the Conference on Ergodic Theory (Alexandria, Egypt, 1993). London Mathematical Society Lecture Notes, vol. 205, pp. 3–151 (1995)
34. Słomian, W.: Bootstrap methods in bounding discrete Radon operators. J. Funct. Anal. **283**(9), 109650 (2022)
35. Słomian, W.: Oscillation estimates for truncated singular Radon operators. J. Fourier Anal. Appl. (2023). https://doi.org/10.1007/s00041-022-09986-8
36. Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
37. Tao, T.: The Ionescu-Wainger multiplier theorem and the adeles. Mathematika **67**, 647–677 (2021)
38. Trojan, B.: Variational estimates for discrete operators modeled on multi-dimensional polynomial subsets of primes. Math. Ann. **374**, 1597–1656 (2019)
39. Wierdl, M.: Pointwise ergodic theorem along the prime numbers. Isr. J. Math. **64**, 315–336 (1988)
40. Zorin-Kranich, P.: Variation estimates for averages along primes and polynomials. J. Funct. Anal. **268**(1), 210–238 (2015)

41. Zygmund, A.: An individual ergodic theorem for non-commutative transformations. Acta Sci. Math. Szeged **14**, 103–110 (1951)