

A HARD-ANALYTIC PROOF OF “MOST” POLYNOMIAL WIENER-WINTNER THEOREMS FOR INFINITE MEASURE SPACES

BEN KRAUSE

ABSTRACT. We provide a new proof of “most” cases of the polynomial Wiener-Wintner theorem for σ -finite spaces, using hard-analytic methods. Specifically, we prove that whenever (X, μ, T) is a σ -finite measure-preserving system, and $f \in L^p(X)$, $1 \leq p < \infty$, there exists a co-null set $X_f \subset X$ so that for all $\omega \in X_f$

$$\frac{1}{N} \sum_{n \leq N} e^{2\pi i P(n)} f(T^n \omega)$$

converges for all polynomials P which are either linear, or vanish to degree 2 at the origin.

1. INTRODUCTION

The Wiener-Wintner ergodic theorem [21] is a classical generalization of Birkhoff’s Theorem [4]; here and throughout, by a *measure-preserving system*, (X, μ, T) , we mean a probability space, (X, μ) , equipped with a measure-preserving transformation $T : X \rightarrow X$, so that

$$\mu(T^{-1}E) = \mu(E) \text{ for all } E \subset X \text{ measurable;}$$

a σ -finite measure-preserving system is as above, but the underlying space (X, μ) is σ -finite.

Theorem 1.1 (Wiener-Wintner Ergodic Theorem). *Let (X, μ, T) be a measure-preserving system, and let $f \in L^1(X)$ be arbitrary. Then there exists a subset $X_f \subset X$ with $\mu(X_f) = 1$ so that for all $\omega \in X_f$*

$$\lim_N \frac{1}{N} \sum_{n \leq N} e^{2\pi i n \theta} f(T^n \omega)$$

exists for all $\theta \in [0, 1]$.

The significance of this theorem is that the limit holds for *all* θ , rather than just for almost every θ ; this weaker statement follows from applying Birkhoff’s Ergodic Theorem in the product setting. And, whenever (X, μ) is countably generated, one can choose a universal full-measure subset, $X_0 \subset X$ so that the conclusion of Theorem

1.1 holds for all $f \in L^1(X)$ whenever $\omega \in X_0$; this follows from the maximal ergodic theorem and a standard density argument.

At this point the Wiener-Wintner Theorem is well understood, and is known to accommodate more general weighted averages including polynomials [17], nilsequences [12], Hardy field functions [9], etc.; see [1, 2] for a fuller discussion. Moreover, this line of inquiry ultimately admits the profound extension to the *return times* setting [7].

In this note we provide a hard-analytic proof of “most” cases of the polynomial Wiener-Wintner Theorem which holds in the σ -finite setting; see [1] for an alternative proof of the case with linear modulations.

Theorem 1.2. *Let (X, μ, T) be a σ -finite measure-preserving system, and let $f \in L^p(X)$, $1 \leq p < \infty$ be arbitrary. Then there exists a co-null subset $X_f \subset X$ so that for all $\omega \in X_f$*

$$\lim_N \frac{1}{N} \sum_{n \leq N} e^{2\pi i P(n)} f(T^n \omega)$$

exists for all polynomials $P \in \mathbb{R}[\cdot]$ that are either linear, or vanish to degree two at the origin.

Standard arguments allow one to replace the character $t \mapsto e^{2\pi i t}$ with any Riemann integrable function $\phi : \mathbb{T} \rightarrow \mathbb{C}$ without adjusting X_f . And, we may remove the assumption that X is σ -finite by restricting to the $(\sigma$ -finite) σ -algebra generated by $\{f(T^n \cdot) : n \in \mathbb{Z}\}$ for each individual $f \in L^p(X)$, $1 \leq p < \infty$.

1.1. Proof Overview. In proving Theorem 1.2, by the Maximal Ergodic Theorem and a standard density argument, we may assume restrict to simple functions, and then by linearity to integrable indicators; with this in mind, we may restrict as well to lacunary times $N \in \{\lfloor \lambda^k \rfloor : k \in \mathbb{N}\}$ for $1 < \lambda \leq 2$. And, there is no harm in replacing the rough cut-off $\frac{1}{N} \mathbf{1}_{[1, N]}$ with a smooth bump function

$$\varphi_N(n) := \frac{1}{N} \varphi\left(\frac{n}{N}\right)$$

where

$$\|\varphi - \mathbf{1}_{[0, 1]}\|_{L^1(\mathbb{R})} \leq \epsilon_0$$

and

$$|\partial^\alpha \varphi| \lesssim_\alpha \epsilon_0^{-\alpha} \cdot \mathbf{1}_{[0, 1]}$$

for sufficiently many α ; see Subsection §1.2.1 below to recall the $\lesssim_\alpha$ notation. We will regard ϵ_0 as fixed throughout the remainder of the paper. Finally, we will restrict as we may to polynomials whose degree is bounded above by d .

Thus, if we define

$$\Phi_M^P f(\omega) := \Phi_M^{P;T} f(\omega) := \sum_m \varphi_M(m) e^{2\pi i P(m)} f(T^m \omega)$$

and for each $\omega \in X$, define the (truncated) *jump-counting* function

$$(1.3) \quad N_{\tau,H}^P f(\omega) := N_{\tau,H}^{P;T} f(\omega) := \\ := \sup \left\{ K : \text{there exists } M_0(\omega) < M_1(\omega) < \dots < M_K(\omega) \leq H : \right. \\ \left. |\Phi_{M_{k-1}(\omega)}^P f(\omega) - \Phi_{M_k(\omega)}^P f(\omega)| \gg \tau \right\}$$

where all times are of the form $\lfloor (1+\tau)^{\mathbb{N}} \rfloor$, the following elementary lemma will guide our analysis; see Subsection §1.2.1 below for a review of little-Oh notation.

Lemma 1.4. *Suppose that, for each integrable indicator f ,*

$$\mu(\{X : \sup_{P \in \mathcal{P}} N_{\tau,H}^P f(\omega) \geq L\}) = o_{L \rightarrow \infty; \tau, \|f\|_{L^2(X)}}(1),$$

independent of H ; then for any $f \in L^p(X)$, $1 \leq p < \infty$, there exists a co-null set $X_f \subset X$ so that for all $\omega \in X_f$

$$\lim_N \sum_n \varphi_N(n) e^{2\pi i P(n)} f(T^n \omega)$$

exists for all $P \in \mathcal{P}$.

Proof. As above, it suffices to prove convergence only for integrable indicators; let f be an arbitrary such indicator.

Since

$$P \mapsto N_{\tau,H}^P f(\omega)$$

is continuous for each ω, H, τ and bounded f , it suffices to restrict to polynomials with rational coefficients, so

$$\{X : \sup_{P \in \mathcal{P}} N_{\tau,H}^P f(\omega) \geq L\}$$

is measurable for each τ, H , and indicator function f . We proceed by contradiction. So, suppose that there exists a nontrivial set $E \subset X$ with $\mu(E) \gg \tau$ so that for all $\omega \in E$ there exists $P_\omega \in \mathcal{P}$ so that

$$\limsup_{N, M \rightarrow \infty} |\Phi_M^{P_\omega} f(\omega) - \Phi_N^{P_\omega} f(\omega)| \gg \tau;$$

there is no loss of generality in assuming that M, N lie in a $(1+\tau)$ -lacunary sequence.

So, for all $\omega \in E$, there exists an infinite sequence $M_0(\omega) < M_1(\omega) < \dots$ so that

$$|\Phi_{M_{k-1}(\omega)}^{P_\omega} f(\omega) - \Phi_{M_k(\omega)}^{P_\omega} f(\omega)| \gg \tau, \quad k \geq 1;$$

possibly after replacing E with a subset, $E' \subset E$ of measure $\gg \tau$, we may assume that for all $\omega \in E'$, there exists a subsequence

$$M_0(\omega) < M_1(\omega) < \dots < M_K(\omega) \leq H$$

with $K \geq L \gg_\tau 1$ so that the above oscillatory estimate holds; while H may depend on τ, L , our estimates will be independent of this precise choice of H . In particular, for each such ω ,

$$\sup_{P \in \mathcal{P}} N_{\tau, H}^P f(\omega) \geq L,$$

so that

$$\tau \ll \mu(E') \leq \mu(\{X : \sup_{P \in \mathcal{P}} N_{\tau, H}^P f(\omega) \geq L\}) = o_{L \rightarrow \infty; \tau, \|f\|_{L^2(X)}}(1),$$

for the desired contradiction. $\square$

We will apply Lemma 1.4 by dominating the pertaining jump-counting function (1.3) by its *variational* counterpart,

$$(1.5) \quad V_{\mathcal{P}}^r f := \sup_{P \in \mathcal{P}} \sup \left(\sum_i |\Phi_{M_k}^P f - \Phi_{M_{k-1}}^P f|^r \right)^{1/r}, \quad 2 < r < \infty$$

where the inner supremum is over finite increasing subsequences, as we may bound

$$\sup_{P \in \mathcal{P}} \tau N_{\tau, H}^P f(\omega)^{1/r} \leq V_{\mathcal{P}}^r f(\omega)$$

for each τ, H . In particular, we will be interested in proving the norm estimates

$$\|V_{\mathcal{P}}^r f\|_{L^2(X)} \lesssim_{r, \epsilon_0, \lambda, d} \|f\|_{L^2(X)}$$

whenever the sequence of times in the definition of $V_{\mathcal{P}}^r$ live inside a lacunary sequence, $\{M_k\} \subset \{\lfloor \lambda^n \rfloor\}$, $1 < \lambda \leq 2$; by Calderón's transference principle, [8], we may conduct our analysis on the integer lattice, and focus on estimating

$$(1.6) \quad \mathcal{V}_{\mathcal{P}}^r f := \sup_{P \in \mathcal{P}} \sup \left(\sum_i |A_{M_k}^P f - A_{M_{k-1}}^P f|^r \right)^{1/r}, \quad 2 < r < \infty$$

on $\ell^2(\mathbb{Z})$, where

$$A_M^P f(x) := \sum_n \varphi_M(n) e^{2\pi i P(n)} f(x - n) : \mathbb{Z} \rightarrow \mathbb{C}$$

are polynomially modulated discrete convolution operators.

At this point, the analysis splits into two cases: in the first case, when $\mathcal{P}$ consists of linear polynomials, (1.6) encodes a modulation invariance, namely

$$\mathcal{V}_{\mathcal{P}}^r f \equiv \mathcal{V}_{\mathcal{P}}^r f_\theta, \quad f_\theta(n) := e^{2\pi i n \theta} f(n)$$

which necessitates an approach deriving from time frequency analysis. Fortunately, the relevant analysis was developed in deep work of Oberlin-Seeger-Tao-Thiele-Wright

[20], in the context of quantitative convergence of Fourier series. We make no comments on their work, which goes by way of wave packet analysis, but simply import and appropriately transfer the estimate

$$\begin{aligned} & \left\| \sup_{\theta} \sup_k \left(\sum_k \left| \int_{-r_k}^{r_k} f(x-t) e^{2\pi i \theta t} \frac{dt}{2r_k} - \int_{-r_{k-1}}^{r_{k-1}} f(x-t) e^{2\pi i \theta t} \frac{dt}{2r_{k-1}} \right|^r \right)^{1/r} \right\|_{L^2(\mathbb{R})} \\ & \lesssim_r \|f\|_{L^2(\mathbb{R})}, \quad r > 2, \end{aligned}$$

see [20, Remark D.4]; the inner supremum is over finite increasing subsequences.

The second case we address is when $\mathcal{P}$ is given by

$$(1.7) \quad \mathcal{P}_d := \{P \in \mathbb{R}[\cdot] : P(0) = P'(0) = 0, \deg(P) \leq d\},$$

polynomials of bounded degree which vanish to degree two at the origin. In this context, no modulation invariance presents, and orthogonality methods become more relevant; on the other hand, the lack of linearity in the class of $\mathcal{P}_d$ eliminates the connection to the real-variable problem. We build off prior work [15] to prove

$$\|\mathcal{V}_{\mathcal{P}}^r f\|_{\ell^2(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda, d} \left(\frac{r}{r-2}\right)^2 \|f\|_{\ell^2(\mathbb{Z})},$$

which in fact implies

$$\|\mathcal{V}_{\mathcal{P}}^r f\|_{\ell^p(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda, d} \left(\frac{r}{r-2}\right)^2 \|f\|_{\ell^p(\mathbb{Z})}, \quad r > \max\{p, p'\},$$

by interpolating against a special case of [15]. While the arguments of [15] lived at the interface of analytic number theory and harmonic analysis, deriving from Bourgain's celebrated work on the polynomial ergodic theorems [5–7], our work here is purely analytic, as the number theoretic input of [15] can be imported directly; our main ingredient is a “variable coefficient” multi-frequency variation estimate, see [11] for similar work.

Finally, we remark that addressing the case of general polynomial modulations using these methods is out of range of current techniques, as the polynomial modulation invariance whenever $\mathcal{P}$ consist of all real variable polynomials of degree $\leq d$,

$$\mathcal{V}_{\mathcal{P}}^r f \equiv \mathcal{V}_{\mathcal{P}}^r f_P, \quad f_P(n) := e^{2\pi i P(n)} f(n), \quad \deg(P) \leq d,$$

introduces a degree of complexity similar to that found in addressing the pointwise convergence of the multiple ergodic averages

$$\frac{1}{N} \sum_{n \leq N} T^n f_1 \cdots T^{n(d+1)} f_{d+1}, \quad f_i \in L^\infty(X);$$

we hope to address this issue in future work.

The structure of the paper is as follows:

In §2, we quickly address the case of linear polynomials;

In §3, we turn to the case of oscillatory polynomials, and complete the proof of Theorem 1.2 by developing the necessary variational theory.

1.2. **Notation.** We use

$$e(t) := e^{2\pi it}$$

throughout to denote the complex exponential, and let

$$M_{\text{HL}}f(x) := \sup_{N \geq 0} \frac{1}{2N+1} \sum_{|n| \leq N} |f(x-n)|$$

denote the discrete Hardy-Littelwood maximal function.

For a sequence of scalars, $\{a_N\}$, we define the r -variation

$$\mathcal{V}^r(a_N : N) := \sup \left(\sum_k |a_{N_k} - a_{N_{k-1}}|^r \right)^{1/r},$$

where the supremum is over all finite increasing subsequences. We will only be interested below in the case where $2 < r < \infty$.

For functions $\{f_N\} : X \rightarrow \mathbb{C}$, we define

$$\mathcal{V}^r(f_N : N)(x) := \mathcal{V}^r(f_N(x) : N)$$

pointwise.

1.2.1. *Asymptotic Notation.* We will make use of the modified Vinogradov notation. We use $X \lesssim Y$ or $Y \gtrsim X$ to denote the estimate $X \leq CY$ for an absolute constant C and $X, Y \geq 0$. If we need C to depend on a parameter, we shall indicate this by subscripts, thus for instance $X \lesssim_p Y$ denotes the estimate $X \leq C_p Y$ for some C_p depending on p . We use $X \approx Y$ as shorthand for $Y \lesssim X \lesssim Y$. We use the notation $X \ll Y$ or $Y \gg X$ to denote that the implicit constant in the $\lesssim$ notation is extremely large, and analogously $X \ll_p Y$ and $Y \gg_p X$.

We also make use of big-Oh and little-Oh notation: we let $O(Y)$ denote a quantity that is $\lesssim Y$, and similarly $O_p(Y)$ will denote a quantity that is $\lesssim_p Y$; we let $o_{t \rightarrow a}(Y)$ denote a quantity whose quotient with Y tends to zero as $t \rightarrow a$ (possibly ∞), and $o_{t \rightarrow a; p}(Y)$ denote a quantity whose quotient with Y tends to zero as $t \rightarrow a$ at a rate depending on p .

2. LINEAR MODULATIONS

We organize our analysis in this section by introducing the following definition:

Definition 2.8. A function $\chi : \mathbb{R} \rightarrow \mathbb{C}$ is said to satisfy r -Variational Carleson with constant C if the following inequality holds, independent over all choices of (finitely many) measurable functions $\{R_i\}, \theta$:

$$\begin{aligned} & \left\| \left(\sum_i \left| \int \left(\widehat{\chi}(R_i(x)(\beta - \theta(x))) - \widehat{\chi}(R_{i+1}(x)(\beta - \theta(x))) \right) \widehat{f}(\beta) e(\beta x) \right|^r \right)^{1/r} \right\|_{L^2(\mathbb{R})} \\ & \leq C \frac{r}{r-2} \|f\|_{L^2(\mathbb{R})}. \end{aligned}$$

With this in mind, Remark D.4 of [20] can be stated as follows.

Proposition 2.9 (Variational Carleson, Averaging Formulation). $\mathbf{1}_{[-1/2, 1/2]}$ satisfies r -Variational Carleson with constant $O(1)$.

Our task, therefore, is to prove the following proposition.

Proposition 2.10. Suppose that φ is smooth, and satisfies

$$\|\partial^\alpha \varphi\|_\infty \leq \epsilon^{-\alpha} \cdot \mathbf{1}_{[0,1]}$$

for sufficiently many α , and that all times $M \in \lfloor \lambda^\mathbb{N} \rfloor \subset \mathbb{N}$, $\lambda > 1$. Then

$$\left\| \sup_{\theta} \mathcal{V}^r \left(\sum_m \varphi_M(m) g(a+m) e(m\theta) : M \right) \right\|_{\ell^2(\mathbb{Z})} \lesssim \epsilon^{-O(1)} \frac{r}{r-2} \frac{\lambda}{\lambda-1} \|f\|_{\ell^2(\mathbb{Z})}.$$

The proof of Proposition 2.10 will derive from Proposition 2.9 by way of a transference argument. We begin with a convexity result.

Lemma 2.11. Suppose that $\varphi \in \mathcal{C}^1(\mathbb{R}) \cap \mathcal{C}_0(\mathbb{R})$. Then φ satisfies r -Variational Carleson with constant $O(C_\varphi)$, where

$$C_\varphi := \|x\varphi'(x)\|_{L^1(\mathbb{R})}.$$

Proof. By the fundamental theorem of calculus we may express

$$\phi(t) = - \int \frac{1}{s} \mathbf{1}_{[0,s]}(t) s\phi'(s) ds,$$

and thus whenever ϕ is even

$$\phi(t) = - \int \frac{1}{2s} \mathbf{1}_{[-s,s]}(t) s\phi'(s) ds,$$

so

$$\begin{aligned} & \int \left(\widehat{\phi}(R_i(x)(\beta - \theta(x))) - \widehat{\phi}(R_{i+1}(x)(\beta - \theta(x))) \right) \widehat{f}(\beta) e(\beta x) d\beta \\ &= - \int \left(\int \left(\widehat{\chi}(sR_i(x)(\beta - \theta(x))) - \widehat{\chi}(sR_{i+1}(x)(\beta - \theta(x))) \right) \widehat{f}(\beta) e(\beta x) \right) s\phi'(s) ds, \end{aligned}$$

where $\chi = \frac{1}{2} \mathbf{1}_{[-1,1]}$, from which the result follows. $\square$

Since, for the purpose of convergence, we are only interested in lacunary sequences, i.e. sequences of the form

$$\{M_i\} \subset \lfloor \lambda^\mathbb{N} \rfloor, \quad \lambda > 1,$$

we introduce the following definition.

Definition 2.12. A function $\chi : \mathbb{R} \rightarrow \mathbb{C}$ is said to satisfy lacunary r -Variational Carleson with constant C if the following inequality holds, independent over all choices of (finitely many) measurable functions $\{R_i\} \subset \lfloor \lambda^{\mathbb{N}} \rfloor, \theta$:

$$\begin{aligned} & \left\| \left(\sum_i \left| \int (\widehat{\chi}(R_i(x)(\beta - \theta(x))) - \widehat{\chi}(R_{i+1}(x)(\beta - \theta(x)))) \widehat{f}(\beta) e(\beta x) \right|^r \right)^{1/r} \right\|_{L^2(\mathbb{R})} \\ & \leq C \frac{\lambda}{\lambda - 1} \frac{r}{r - 2} \|f\|_{L^2(\mathbb{R})}. \end{aligned}$$

We next introduce a quantity that will allow us to pass from one function which satisfies lacunary r -Variational Carleson to another:

Given two real-variable functions φ, χ , define

$$\begin{aligned} A[\varphi, \chi] := & \sup_{\{R\}, 0 < s \leq S < \infty} \left| \sum_R \int_{s \leq |t| \leq S} \varphi(t/R) - \chi(t/R) \frac{dt}{R} \right| \\ & + \sup_{\{R\}, t \neq 0} (|t|/R \cdot \sum_R |\varphi(t/R) - \chi(t/R)|) + \sup_{R, t \neq 0} (|t|^2/R^2 \cdot \sum_R |\varphi'(t/R) - \chi'(t/R)|) \end{aligned}$$

where the supremum is over every 2-lacunary sequence of times, $\{R\} \subset \mathbb{N}$.

Lemma 2.13. Suppose that χ satisfies lacunary r -Variational Carleson constant C . Then φ satisfies lacunary r -Variational Carleson with constant $C + O(A[\varphi, \chi])$.

Proof. Set $\psi := \varphi - \chi$; it suffices to bound

$$\left\| \sup_{\theta} \left(\sum_R \left| \int \psi(R(\beta - \theta)) \widehat{f}(\beta) e(\beta x) d\beta \right|^2 \right)^{1/2} \right\|_{L^2(\mathbb{R})} \lesssim A[\varphi, \chi] \|f\|_{L^2(\mathbb{R})}$$

for any 2-lacunary sequence $\{R_i\}$. The result follows from randomization and Carleson's Theorem for Calderón-Zygmund Kernels, a result implicitly contained in e.g. [10], see also [3] or [18] for the explicit result along with a significant strengthening. $\square$

We now begin transferring to the discrete context.

Lemma 2.14. Suppose that φ satisfies lacunary r -variational Carleson with constant C , and has Fourier support inside $\{|\xi| \leq \epsilon^{-1}\}$, and that all times $M \in \lfloor \lambda^{\mathbb{N}} \rfloor \subset \mathbb{N}$. Then

$$\left\| \sup_{\theta} \mathcal{V}^r \left(\sum_m \varphi_M(m) g(a + m) e(m\theta) : M \right) \right\|_{\ell^2(\mathbb{Z})} \lesssim C \epsilon^{-1} \frac{r}{r - 2} \frac{\lambda}{\lambda - 1} \|g\|_{\ell^2(\mathbb{Z})}.$$

Proof. By convexity and lacunarity we can assume that $M \geq 2^{\epsilon^{-1}}$. If we bound

$$\mathcal{V}^r\left(\sum_m \varphi_M(m)g(a+m)e(m\theta) : M\right) \leq \sum_{0 \leq l \leq 4} \sup_{l/5 \leq \theta < (l+1)/5} \mathcal{V}^r\left(\sum_m \varphi_M(m)g(a+m)e(m\theta) : M\right)$$

then it suffices to each summand in l individually; by monotone convergence, it suffices to estimate the ℓ^2 norms of

$$\sup_{\theta \in [l/5, (l+1)/5] \cap \frac{1}{T} \cdot \mathbb{Z}} \mathcal{V}^r\left(\sum_m \varphi_M(m)g(a+m)e(m\theta) : M \leq T\right),$$

independent of T ; the result then follows from Magyar-Stein-Wainger transference [19, Proposition 2.1] and the hypothesis. $\square$

The proof of Proposition 2.10 now follows from a standard dyadic decomposition.

Proof of Proposition 2.10. By the Riemann-Lebesgue lemma, φ is smooth and satisfies

$$|\widehat{\varphi}(\xi)| \lesssim_A (1 + \epsilon|\xi|)^{-A}$$

for sufficiently large A ; set

$$\varphi^E(t) := \frac{\varphi(t) + \varphi(-t)}{2},$$

which satisfies the same Fourier decay, and decompose

$$\varphi := \varphi_0 + \sum_{j \geq 1} \varphi_j, \quad \varphi^E = \varphi_0^E + \sum_{j \geq 1} \varphi_j^E$$

where

$$\widehat{\varphi_j}(\xi) := \widehat{\varphi}(\xi) \cdot \psi(\xi^2/2^j),$$

and φ_0 is defined by subtraction, and similarly for φ_j^E , where

$$\sum_k \psi(\xi/2^k) = \mathbf{1}_{\xi \neq 0}$$

and ψ is a smooth approximation to $\mathbf{1}_{|\xi| \approx 1}$.

Note that φ_0^E satisfies r -Variational Carleson with constant $O(\epsilon^{-O(1)})$, and similarly

$$A[\varphi_0, \varphi_0^E] \lesssim \epsilon^{-O(1)},$$

so φ_0 satisfies the lacunary r -Variational Carleson with constant $O(\epsilon^{-O(1)})$. And, for each $j \geq 1$, $\widehat{\varphi_j}, \widehat{\varphi_j^E}$ are supported in $\{|\xi| \lesssim \epsilon^{-2}2^j\}$, and that

$$C_{\varphi_j^E}, A[\varphi_j, \varphi_j^E] \lesssim \epsilon^{10}2^{-10j}$$

(say), so the result follows from the triangle inequality. $\square$

This concludes our work on linear modulations; for the remainder of the paper, all polynomials will be assumed to vanish to degree two at the origin, see (1.7).

3. POLYNOMIAL WIENER WINTNER

It will be convenient to adopt a singular integral perspective. Thus, fix a constant of lacunarity $1 < \lambda \leq 2$, and set

$$\psi(t) := \varphi(t) - \lambda^{-1}\varphi(t/\lambda), \quad \psi_k(t) := \lambda^{-k}\psi(\lambda^{-1}t),$$

so that $\{\psi_k\}$ are mean-zero, supported in $\{\lambda^k/10 \leq |x| \leq 10\lambda^k\}$ and satisfy

$$\frac{\|\psi_k\|_{L^\infty(\mathbb{R})}}{\lambda^k} + \frac{\|\psi'_k\|_{L^\infty(\mathbb{R})}}{\lambda^{2k}} \leq C < \infty$$

uniformly in k ; consolidate

$$\Psi_k(t) := \sum_{1 \leq j \leq k} \psi_j(t), \quad \Psi_k^s(t) := \sum_{2^{s/A_0} \leq j \leq k} \psi_j(t).$$

By telescoping appropriately and applying Calderón's transference principle as above, our task is to prove the following proposition.

Proposition 3.15. *The following estimate holds:*

$$\|\mathcal{V}_d^r f\|_{\ell^2(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda, d} \left(\frac{r}{r-2}\right)^2 \|f\|_{\ell^2(\mathbb{Z})},$$

where

$$\mathcal{V}_d^r f(x) := \sup_{P \in \mathcal{P}_d} \sup \left(\sum_i \left| \sum_n (\Psi_{k_i}(n) - \Psi_{k_{i-1}}(n)) e(P(n)) f(x-n) \right|^r \right)^{1/r}$$

where the inner supremum is over all finite increasing subsequence $\{k_i\}$.

Remark 3.16. *By interpolating this proposition with the $r = \infty$ version addressed in [15], the following norm estimates present:*

$$\|\mathcal{V}_d^r f\|_{\ell^p(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda, d, p} \left(\frac{r}{r-2}\right)^2 \|f\|_{\ell^p(\mathbb{Z})}, \quad r > \max\{p, p'\}.$$

3.1. Preliminaries. Let $A_0 \in \mathbb{N}$ be large but fixed, and for $P \in \mathcal{P}_d$, abbreviate

$$P_{\vec{\lambda}}(t) := \sum_{j=2}^d \lambda_j t^j.$$

For $P \in \mathbb{R}[\cdot]$, we use

$$N_{2^k}(P)$$

to denote the *coefficient norm* of P at scale 2^k , see [15, Definition 3.1]. We will use $\vec{\lambda}$ to denote elements of $\mathbb{T}^{d-1}$, $\vec{\lambda} = (\lambda_2, \dots, \lambda_d)$, and similarly use the notation

$$\vec{A} := (A_2, \dots, A_d, Q)$$

where $\vec{A} \in \mathbb{Z}^{d-1}$ so that in particular $(\vec{A}, Q) = 1$ means that $(A_2, \dots, A_d, Q) = 1$.

We apply the decomposition of [15, Proposition 4.2]; specifically, we have the following lemma.

Lemma 3.17. *The following pointwise bound holds:*

$$\mathcal{V}_d^r f \leq \sum_{s \geq 1} \mathcal{A}_s f + \mathcal{E}f + O(M_{HL}f) + \mathcal{H}^r f,$$

where $\mathcal{E}$ is exactly as in [15, Proposition 4.2], with bounded $\ell^2(\mathbb{Z})$ operator norm, $\mathcal{H}^r f$ is the r -variation of a truncated singular integral (whose ℓ^2 -operator norm is $O(\frac{r}{r-2})$), and

$$\mathcal{A}_s f(x) := \mathcal{V}^r \left(\sum_{k=2^{s/A_0}}^{k_0} \sum_m \psi_k(m) e(P_{\vec{\lambda}(x)}(m)) f(x-m) \cdot \mathbf{1}_{n: N_{2^k}(P_{\vec{\lambda}}(n))=2^s}(x) : k_0 \geq 2^{s/A_0} \right).$$

In particular, to prove Proposition 3.15 it suffices to prove that

$$\|\mathcal{A}_s f\|_{\ell^2(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda, d} \left(\frac{r}{r-2} \right)^2 2^{-cs} \|f\|_{\ell^2(\mathbb{Z})} \text{ for some } c > 0.$$

Next, by [15, Proposition 5.6], we may replace $\mathcal{A}_s f$ with

$$\mathcal{V}_{s,d}^r f := \sup_{\vec{\lambda}} \mathcal{V}^r((L_{J,\vec{\lambda}}^s)^\vee * f : J)$$

where

$$\{L_{J,\vec{\lambda}}^s\}$$

are Fourier multipliers, defined below:

$$L_{J,\vec{\lambda}}^s(\beta) = \sum_{(\vec{A}, Q)=1: 2^{s-1} \leq Q < 2^s} \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \Phi_{J,\vec{\lambda}-\vec{A}/Q}^*(\beta - B/Q) \chi_s(\beta - B/Q) \\ \times \mathbf{1}_{\|\lambda_j - A_j/Q\|_{\mathbb{T}} \leq 2^{-10s-10}, 2 \leq j \leq d}$$

where

$$S(\vec{A}/Q, B/Q) := \frac{1}{Q} \sum_{r \leq Q} e\left(-\frac{A_2 r^2 + \dots + A_d r^d + rB}{Q}\right)$$

are complete Weyl sums;

$$\Phi_{J,\vec{\lambda}}^\vee(x) = \Psi_J^s(x) e(-P_{\vec{\lambda}}(x)) \cdot \mathbf{1}_{\|\lambda_k\|_{\mathbb{T}} \leq J^{A_0} 2^{-kJ}, 2 \leq k \leq d}$$

are polynomially modulated truncated singular kernels, and

$$\mathbf{1}_{|\beta| \leq 2^{-2s/10A_0}} \leq \chi_s \leq \mathbf{1}_{|\beta| \leq 2^{1-2s/10A_0}}$$

is a smooth cut-off.

The key property of the $\{L_{J,\vec{\lambda}}^s\}$ that we will use is encoded in the behavior of the Weyl sums, which satisfy

$$\sup_{(\vec{A}, B, Q)=1} |S(\vec{A}/Q, B/Q)| \leq Q^{-cd}$$

for some absolute $c_d > 0$ by standard Weyl sum estimates, see e.g. [14, Appendix B], but also stronger orthogonality estimates, see [15, Lemma 6.3]:

Lemma 3.18. *There exists an absolute $c = c_d > 0$ so that the following maximal estimate holds:*

$$\left\| \sup_{(\vec{A}, Q)=1, 2^{s-1} \leq Q < 2^s} \left| \int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) \right| \right\|_{\ell^2(\mathbb{Z})} \lesssim 2^{-cs} \|f\|_{\ell^2(\mathbb{Z})}.$$

With this quantitative input, we can effectively estimate $\mathcal{V}_{s,d}^r f$; the proof of the below estimate will occupy the remainder of the paper.

Proposition 3.19. *There exists an absolute constant $0 < c = c_d < 1$ so that following bound holds:*

$$\|\mathcal{V}_{s,d}^r f\|_{\ell^2(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda} \left(\frac{r}{r-2}\right)^2 s^2 2^{-cs} \|f\|_{\ell^2(\mathbb{Z})}.$$

We let

$$I_s = [-2^{-2^{s/10}A_0}, 2^{-2^{s/10}A_0}]$$

denote the moral support of χ_s ; and we below we let $c = c_d > 0$ denote small constants, depending only on the degree, d ; and so, we will often suppress the implicit dependence on d in the estimates below.

With this in mind, the following two corollaries immediately present:

The first is a sequence-space estimate.

Corollary 3.20. *There exists $c > 0$ so that whenever $|I| \geq |I_s|$,*

$$\left\| \sup_{(\vec{A}, Q)=1, 2^{s-1} \leq Q < 2^s} \left| \sum_{B \leq Q} c_{B/Q} S(\vec{A}/Q, B/Q) e(B/Qx) \right| \right\|_{\ell^2(I)} \lesssim 2^{-cs} |I|^{1/2} \|c_{B/Q}\|_{\ell^2(\mathbb{Z})}.$$

Proof. Let $\mathbf{1}_I \leq |V_I|$ be a Schwartz function with spatial scale I whose Fourier transform is supported in a $|I|^{-1}$ neighborhood of 0, and restricting the supremum to

$$(\vec{A}, Q) = 1, 2^{s-1} \leq Q < 2^s,$$

bound

$$\begin{aligned}
& \left\| \sup_{\vec{A}/Q} \left| \sum_{B \leq Q} c_{B/Q} S(\vec{A}/Q, B/Q) e(B/Qx) \right| \right\|_{\ell^2(I)} \\
& \leq \left\| \sup_{\vec{A}/Q} \left| \sum_{B \leq Q} c_{B/Q} V_I(x) S(\vec{A}/Q, B/Q) e(B/Qx) \right| \right\|_{\ell^2(\mathbb{Z})} \\
& = \left\| \sup_{\vec{A}/Q} \left| \int e(\beta x) \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \widehat{V}_I(\beta - B/Q) \left(\sum_{B \leq Q} c_{B/Q} \chi_s(\beta - B/Q) \right) d\beta \right| \right\|_{\ell^2(\mathbb{Z})} \\
& = \left\| \sup_{\vec{A}/Q} \left| \int e(\beta x) \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \chi_s(\beta - B/Q) \left(\sum_{B \leq Q} c_{B/Q} \widehat{V}_I(\beta - B/Q) \right) d\beta \right| \right\|_{\ell^2(\mathbb{Z})} \\
& \lesssim 2^{-cs} \|c_{B/Q}\|_{\ell^2(\mathbb{Z})} |I|^{1/2}.
\end{aligned}$$

□

We record the following simple consequence of convexity in the below lemma.

Lemma 3.21. *There exists $c > 0$ so that whenever $\|\phi\|_{L^1(\mathbb{R})} \leq 1$ and*

$$m_{\vec{\mu}} := \phi e(\widehat{P_{\vec{\mu}}(\cdot)}),$$

the following estimate holds:

$$\left\| \sup_{(\vec{A}, Q)=1, 2^{s-1} \leq Q < 2^s, \vec{\mu}} \left| \int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) (m_{\vec{\mu}} \chi_s)(\beta - B/Q) \widehat{f}(\beta) e(\beta x) \right| \right\|_{\ell^2(\mathbb{Z})} \lesssim 2^{-cs} \|f\|_{\ell^2(\mathbb{Z})}.$$

The proof of Proposition 3.19 will largely derive from the below proposition concerning the following less oscillatory variation operators:

$$\mathcal{V}_s^r f(x) := \sup_{(\vec{A}, Q)=1, 2^{s-1} \leq Q < 2^s} \mathcal{V}^r \left(\int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \widehat{\Psi}_J^s(\beta - B/Q) \chi_s(\beta - B/Q) \widehat{f}(\beta) e(\beta x) d\beta : J \right).$$

Proposition 3.22. *The following estimate holds for some $c > 0$:*

$$\|\mathcal{V}_s^r f\|_{\ell^2(\mathbb{Z})} \lesssim_{\epsilon_0, \lambda} \left(\frac{r}{r-2} \right)^2 s^2 2^{-cs} \|f\|_{\ell^2(\mathbb{Z})}.$$

To prove Proposition 3.22, we linearize our supremum via measurable functions $\vec{A}, Q : \mathbb{Z} \rightarrow \mathbb{Z}^{d-1} \times \mathbb{Z}_{\geq 1}$ and express

$$\mathcal{V}_s^r f(x)$$

$$\equiv \mathcal{V}^r \left(\sum_{B \leq Q(x)} S(\vec{A}(x)/Q(x), B/Q(x)) \widehat{\Psi}_J^s(\beta - B/Q(x)) \chi_s(\beta - B/Q(x)) \hat{f}(\beta) e(\beta x) d\beta : J \right).$$

We emphasize that $J \geq 2^{s/A_0}$ and all scales involved in the definition of Ψ_J^s are similarly $\geq 2^{s/A_0}$, and that our estimates below will be independent of this particular choice of linearization. With this in mind, for notational we will suppress the superscript $\Psi_J^s \rightarrow \Psi_J$, and will similarly suppress the dependence on ϵ_0, λ .

Proof of Proposition 3.22. Set $F_{B/Q}(x) := \chi_s * (\text{Mod}_{-B/Q} f)(x)$ so that we may express the foregoing as variation with respect to the operator

$$\sum_{B \leq Q, 2^{s-1} \leq Q < 2^s} e(B/Qx) \Psi_j * F_{B/Q}(x) \cdot S(\vec{A}(x)/Q(x), B/Q) \mathbf{1}_{B/Q \in \{B/Q(x) : B \leq Q(x)\}}.$$

Note that whenever $|I| = |I_s|$ is an interval, for any $x_I, y_I \in I$ we may Taylor expand

$$\begin{aligned} & \sum_{B \leq Q, 2^{s-1} \leq Q < 2^s} e(B/Qx) \Psi_j * F_{B/Q}(x) S(\vec{A}(x)/Q(x), B/Q) \mathbf{1}_{B/Q \in \{B/Q(x)\}} \\ &= \sum_{B \leq Q, 2^{s-1} \leq Q < 2^s} e(B/Qx) \Psi_j * F_{B/Q}(x_I) S(\vec{A}(x)/Q(x), B/Q) \mathbf{1}_{B/Q \in \{B/Q(x)\}} + O(2^{-10s} M_{\text{HL}} f(y_I)) \end{aligned}$$

since all scales involved satisfy $j \geq 2^{s/A_0}$.

With x_I fixed, we apply metric chaining to the set

$$X(x_I) := \{(\Psi_j * F_{B/Q}(x_I))_{B \leq Q, 2^{s-1} \leq Q < 2^s} : j \geq 2^{s/A_0}\}.$$

In particular, let

$$\vec{N}_\lambda(x_I)$$

denote the $\ell^2(B/Q)$ -jump-counting function associated to the set $X(x_I)$, namely

$$\begin{aligned} \vec{N}_\lambda(x_I) &:= \sup \{K : \text{there exists } k_0 < k_1 < \dots < k_K \text{ so that} \\ &\quad \|\Psi_{k_i} * F_{B/Q}(x_I) - \Psi_{k_{i-1}} * F_{B/Q}(x_I)\|_{\ell^2(B/Q)} \geq \lambda\}, \end{aligned}$$

let

$$\mathcal{F}_s f(x) := \sup_{j \geq 2^{s/A_0}} \left(\sum_{B \leq Q, 2^{s-1} \leq Q < 2^s} |\Psi_j * F_{B/Q}(x)|^2 \right)^{1/2}$$

and

$$\mathcal{V}^r f(x) := \sup \left(\sum_i \|\Psi_{j_i} * F_{B/Q}(x_I) - \Psi_{j_{i+1}} * F_{B/Q}(x)\|_{\ell^2(B/Q)}^r \right)^{1/r}$$

where the supremum runs over all finite increasing subsequences. Note that

$$\|\mathcal{F}_s f\|_{\ell^2(\mathbb{Z})} + \frac{r-2}{r} \|\mathcal{V}^r f\|_{\ell^2(\mathbb{Z})} \lesssim \|f\|_{\ell^2(\mathbb{Z})}$$

by standard singular integral estimates, see [13, Theorem 1.2], and Magyar-Stein-Wainger Transference.

Now, for each v so that

$$2^{-v} \leq \text{diam}(X(x_I)) \leq 2\mathcal{F}_s f(x_I),$$

define $\Lambda_v(x_I)$ to be a collection of times t so that

$$X(x_I) \subset \bigcup_{t \in \Lambda_v(x_I)} \{(b_{B/Q})_{B/Q} : \|b_{B/Q} - \Psi_t * F_{B/Q}(x_I)\|_{\ell^2(B/Q)} \leq 2^{-v}\},$$

and for each $t \in \Lambda_v(x_I)$ define the *parent* of t , $\varrho(t) \in \Lambda_{v-1}(x_I)$ to be the minimal time so that

$$\begin{aligned} & \{(b_{B/Q})_{B/Q} : \|b_{B/Q} - \Psi_t * F_{B/Q}(x_I)\|_{\ell^2(B/Q)} \leq 2^{-v}\} \\ & \cap \{(b_{B/Q})_{B/Q} : \|b_{B/Q} - \Psi_{\varrho(t)} * F_{B/Q}(x_I)\|_{\ell^2(B/Q)} \leq 2^{1-v}\} \neq \emptyset. \end{aligned}$$

Set

$$\nu_t := \Psi_t - \Psi_{\varrho(t)}.$$

Then for $x, y \in I$

$$\begin{aligned} \mathcal{V}^r f(x) & \leq \sum_{2^{-v} \leq 2\mathcal{F}_s(x_I)} \mathcal{V}^r \left(\sum_{B \leq Q(x)} e(B/Q(x)x) S(\vec{A}(x)/Q(x), B/Q(x)) \nu_t * F_{B/Q(x)}(x_I) : t \in \Lambda_v(x_I) \right) \\ & \quad + 2^{-10s} M_{\text{HL}} f(y) \\ & \leq \sum_{2^{-v} \leq 2\mathcal{F}_s(x_I)} \left(\sum_{t \in \Lambda_v(x_I)} \left| \sum_{B \leq Q(x)} e(B/Q(x)x) S(\vec{A}(x)/Q(x), B/Q(x)) \nu_t * F_{B/Q(x)}(x_I) \right|^r \right)^{1/r} \\ & \quad + 2^{-10s} M_{\text{HL}} f(y). \end{aligned}$$

The argument now concludes as in the proof of [16, Theorem 1.4], with the key estimate being

$$\begin{aligned} & \left\| \left(\sum_{t \in \Lambda_v(x_I)} \left| \sum_{B \leq Q(x)} e(B/Q(x)x) S(\vec{A}(x)/Q(x), B/Q(x)) \nu_t * F_{B/Q(x)}(x_I) \right|^r \right)^{1/r} \right\|_{\ell^2(I)} \\ & \lesssim 2^{-cs} 2^{-v} |I|^{1/2} \min\{2^s \vec{N}_{2^{-v}}(x_I)^{1/r}, \vec{N}_{2^{-v}}(x_I)^{1/2}\}; \end{aligned}$$

the first estimate follows Cauchy-Schwartz and pointwise considerations, noting that

$$|\{B \leq Q \leq 2^s\}| \approx 2^{2s},$$

while the second estimate follows from replacing the ℓ^r sum with the stronger ℓ^2 sum, and then applying Corollary 3.20. $\square$

With this in hand, we can quickly prove Proposition 3.19.

The Proof of Proposition 3.19. We will use the *polynomial coefficient norm*,

$$\|P\| := \sum_{j \geq 1} |\lambda_j|, \quad P(t) := \sum_{j \geq 0} \lambda_j t^j \in \mathbb{R}[\cdot]$$

to organize our analysis.

Let A_1 be a large constant, and notice that for each $\vec{\mu}$, note that there are only $O(d^2 A_1)$ many scales j so that there exists $2 \leq k \neq k' \leq d$

$$2^{-A_1} \leq \frac{|\mu_k| 2^{jk}}{|\mu_{k'}| 2^{jk'}} \leq 2^{A_1};$$

collect these scales in the set $\mathcal{J}_0(\vec{\mu})$.

For each $\vec{\mu} \in [2^{-10s}]^{d-1}$, we partition our set of scales into three sets: let

$$\begin{aligned} \mathcal{J}_{\leq}(\vec{\mu}) &:= \{j \notin \mathcal{J}_0(\vec{\mu}) : \|P_{\vec{\mu}}(2^j \cdot)\| \leq 2^{-A_1 s}\} \\ \mathcal{J}_{\approx}(\vec{\mu}) &:= \mathcal{J}_0(\vec{\mu}) \cup \{j : 2^{-A_1 s} \leq \|P_{\vec{\mu}}(2^j \cdot)\| \leq 2^{A_1 s}\} \quad \text{and} \\ \mathcal{J}_{\geq}(\vec{\mu}) &:= \{j \notin \mathcal{J}_0(\vec{\mu}) : \|P_{\vec{\mu}}(2^j \cdot)\| \geq 2^{A_1 s}\}. \end{aligned}$$

Note that for each $j \notin \mathcal{J}_0(\vec{\mu})$, there exists a unique $k = k(j)$ so that

$$\|P_{\vec{\mu}}(2^j \cdot)\| \approx |\mu_k| 2^{kj},$$

so in particular

$$\sup_{\vec{\mu}} |\mathcal{J}_{\approx}(\vec{\mu})| \lesssim_d s$$

and $\mathcal{J}_{\leq}(\vec{\mu}), \mathcal{J}_{\geq}(\vec{\mu})$ have at most $O_d(1)$ many connected components. More to the point, if we let

$$\mathcal{J}_l(\vec{\mu}) := \{j \notin \mathcal{J}_0(\vec{\mu}) : \|P_{\vec{\mu}}(2^j \cdot)\| \approx 2^l\}$$

then

$$\sup_{\vec{\mu}, l \geq -A_1 s} |\mathcal{J}_l(\vec{\mu})| \lesssim_d 1.$$

With this in mind, set

$$\phi_{j,\vec{\mu}}(t) := \psi_j(t)e(P_{\vec{\mu}}(t)),$$

and, abbreviating

$$\sup_{(\vec{A},Q)=1, 2^{s-1} \leq Q < 2^s} \longrightarrow \sup_{\vec{A}/Q}$$

bound

$$\mathcal{V}_{s,d}^r f(x)$$

$$\begin{aligned} &\leq \mathcal{V}_s^r f(x) + \sup_{\vec{A}/Q, \vec{\mu}} \sum_{B \leq Q} \sum_n |f(x-n)| \left(\sum_{j \in \mathcal{J}_{\leq}(\vec{\mu})} \int |\chi_s^\vee(n-t)| |\psi_j(t)| |e(P_{\vec{\mu}}(t)) - 1| dt \right) \\ &+ \sup_{\vec{A}/Q, \vec{\mu}} \sum_{j \in \mathcal{J}_{\approx}(\vec{\mu})} \left| \int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \phi_{j,\vec{\mu}}(\beta - B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) d\beta \right| \\ &+ \sum_{l \geq A_1 s} \sup_{\vec{A}/Q, \vec{\mu}} \sum_{j \in \mathcal{J}_l(\vec{\mu})} \left| \int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \phi_{j,\vec{\mu}}(\beta - B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) d\beta \right| \\ &\leq \mathcal{V}_s^r f(x) + 2^{-10s} M_{\text{HL}} f(x) \\ &+ \sup_{\vec{A}/Q, \vec{\mu}} \sum_{j \in \mathcal{J}_{\approx}(\vec{\mu})} \left| \int \sum_{B \leq Q} S(\vec{A}/Q, B/Q) \phi_{j,\vec{\mu}}(\beta - B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) d\beta \right| \\ &+ \sum_{\vec{A}/Q} \sum_{B \leq Q} \sum_{l \geq A_1 s} \sup_{\vec{\mu}} \sum_{j \in \mathcal{J}_l(\vec{\mu})} \left| \int \phi_{j,\vec{\mu}}(\beta - B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) d\beta \right|; \end{aligned}$$

in the first step we have expanded

$$\psi_j(t)e(P_{\vec{\mu}}(t)) = \psi_j(t) + \psi_j(t)(e(P_{\vec{\mu}}(t)) - 1).$$

The first term is governed by Proposition 3.22, the sum over $\mathcal{J}_{\approx}(\vec{\mu})$ contribute $O_d(s)$ many scales, so by Lemma 3.21 the total contribution from these scales is $O(s2^{-cs})$. It remains only to address the contribution from the scales in $\mathcal{J}_l(\vec{\mu})$; by conceding a constant factor, we can assume for that each $\vec{\mu}$, there exists a unique such scale, call it $j_l(\vec{\mu})$. By modulating, it suffices to estimate

$$\left\| \sup_{\vec{\mu}} \left| \int (\chi_s * f)(x-t) \psi_{j_l(\vec{\mu})}(t) e(P_{\vec{\mu}}(t)) dt \right| \right\|_{\ell^2(\mathbb{Z})};$$

by Magyar-Stein-Wainger transference and TT^* , see [14, §14], there exists an absolute constant $c = c_d > 0$ so that, uniformly in $l \geq 1$,

$$\left\| \sup_{\vec{\mu}} \left| \int (\chi_s * f)(x - t) \psi_{j_l(\vec{\mu})}(t) e(P_{\vec{\mu}}(t)) dt \right| \right\|_{\ell^2(\mathbb{Z})} \lesssim 2^{-cl} \|f\|_{\ell^2(\mathbb{Z})},$$

so

$$\begin{aligned} & \left\| \sum_{\vec{A}/Q} \sum_{B \leq Q} \sum_{l \geq A_1 s} \sup_{\vec{\mu}} \sum_{j \in \mathcal{J}_l(\vec{\mu})} \left| \int \psi_j e(\widehat{P_{\vec{\mu}}(\cdot)})(\beta - B/Q) \chi_s(\beta - B/Q) \hat{f}(\beta) e(\beta x) d\beta \right| \right\|_{\ell^2(\mathbb{Z})} \\ & \lesssim 2^{sd} \sum_{l \geq A_1 s} 2^{-cl} \|f\|_{\ell^2(\mathbb{Z})} \leq 2^{-s} \|f\|_{\ell^2(\mathbb{Z})}, \end{aligned}$$

provided A_1 is chosen sufficiently large. Putting everything together, we have bounded

$$\|\mathcal{V}_{s,d}^r f\|_{\ell^2(\mathbb{Z})} \lesssim \left(\left(\frac{r}{r-2} \right)^2 s^2 2^{-cs} + s 2^{-cs} + 2^{-s} \right) \|f\|_{\ell^2(\mathbb{Z})},$$

as desired. $\square$

REFERENCES

- [1] I. Assani. *Wiener Wintner ergodic theorems*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003, pages 2
- [2] I. Assani. *Wiener-Wintner ergodic theorem, in brief*. Notices Amer. Math. Soc. 69 (2022), no. 2, 198–209. pages 2
- [3] L. Becker; F. van Doorn; A. Jamneshan; R. Srivastava; C. Thiele. *Carleson Operators on Doubling Metric Measure Spaces*. Preprint, <https://arxiv.org/abs/2405.06423> pages 8
- [4] G. Birkhoff. *Proof of the ergodic theorem*. Proc Natl Acad Sci USA 17 (12): 656–660 (1931) pages 1
- [5] J. Bourgain. *On the maximal ergodic theorem for certain subsets of the positive integers*. Israel J. Math. 61 (1988), 39–72. pages 5
- [6] J. Bourgain. *On the pointwise ergodic theorem on L^p for arithmetic sets*. Israel J. Math. 61 (1988), no. 1, 73–84. pages 5
- [7] J. Bourgain. *Pointwise ergodic theorems for arithmetic sets*. Inst. Hautes Études Sci. Publ. Math. (69):5–45, 1989. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. pages 2, 5
- [8] A. Calderón. *Ergodic theory and translation invariant operators*. Proc. Nat. Acad. Sci., USA 59 (1968), 349–353 pages 4
- [9] T. Eisner; B. Krause. *(Uniform) convergence of twisted ergodic averages*. Ergodic Theory Dynam. Systems 36 (2016), no. 7, 2172–2202. pages 2
- [10] C. Fefferman. *Pointwise convergence of Fourier series*. Ann. of Math. (2) 98 (1973), 551–571. pages 8
- [11] S. Guo. P. Zorin-Kranich. *A variable coefficient multi-frequency lemma* Preprint, <https://arxiv.org/abs/2010.00812> pages 5
- [12] B. Host; B. Kra. *Uniformity seminorms on ℓ^∞ and applications*. J. Anal. Math. 108 (2009), 219–276. pages 2

- [13] R. Jones; A. Seeger; J. Wright. *Strong variational and jump inequalities in harmonic analysis*. Trans. Amer. Math. Soc. 360 (2008), no. 12, 6711–6742. pages 15
- [14] B. Krause. *Discrete Analogues in Harmonic Analysis: Bourgain, Stein, and Beyond*. Graduate Studies in Mathematics, 224. American Mathematical Society, Providence, RI, 2023 pages 12, 18
- [15] B. Krause. *Discrete analogues in harmonic analysis: a theorem of Stein-Wainger*. J. Funct. Anal. 287 (2024), no. 5, Paper No. 110498, 49 pp. pages 5, 10, 11, 12
- [16] B. Krause. *Multi-Frequency Oscillation Estimates Arising in Pointwise Ergodic Theory*. Preprint, <https://arxiv.org/abs/2502.12887> pages 16
- [17] E. Lesigne. *Spectre quasi-discret et théorème ergodique de Wiener-Wintner pour les polynômes*. Ergod. Th. & Dynam. Sys., 13 (1993), 767–78 pages 2
- [18] V. Lie. *The polynomial Carleson operator*. Ann. of Math. (2) 192 (2020), no. 1, 47–163. pages 8
- [19] A. Magyar; E. Stein; S. Wainger. *Discrete analogues in harmonic analysis: spherical averages*. Ann. of Math. (2) 155 (2002), no. 1, 189–208. pages 9
- [20] R. Oberlin; A. Seeger; T. Tao; C. Thiele; J. Wright. *A variation norm Carleson theorem*. J. Eur. Math. Soc. (JEMS) 14 (2012), no. 2, 421–464. pages 5, 7
- [21] N. Wiener; A. Wintner. *Harmonic analysis and ergodic theory*. Amer. J. Math., 63, (1941), 415–42 pages 1

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRISTOL, BEACON HOUSE, QUEENS RD,
BRISTOL BS8 1QU

Email address: `ben.krause@bristol.ac.uk`