A VARIATION NORM CARLESON THEOREM

R. OBERLIN, A. SEEGER, T. TAO, C. THIELE, AND J. WRIGHT

1. Introduction

The partial (inverse) Fourier integral of a Schwartz function f on $\mathbb{R}$ is defined as

$$\mathcal{S}[f](\xi, x) = \int_{-\infty}^{\xi} \hat{f}(\xi')e^{2\pi i\xi' x}\, d\xi'$$

where $\hat{f}$ denotes the Fourier transform of f. The behaviour of the partial Fourier integrals as ξ tends to ∞ has been a subject of interest for a long time. The following uniform control is well known:

Theorem 1.1. *Suppose f is a Schwartz function and $1 < p < \infty$, then*

$$\| \sup_{\xi\in\mathbb{R}} |S[f](\xi, \cdot)| \|_{L^p(\mathbb{R})} \le C_p \|f\|_{L^p(\mathbb{R})}. \tag{1}$$

By a standard approximation argument it follows that $\mathcal{S}[f]$ may be meaningfully defined as a continuous function in ξ for almost every x whenever $f \in L^p$ and the a priori bound of the theorem continues to hold for such functions.

Theorem 1.1 is intimately related to almost everywhere convergence of partial Fourier sums for functions in $L^p[0,1]$. Via a transference principle [12], it is indeed equivalent to the celebrated theorem by Carleson [2] for $p = 2$ and the extension of Carleson's theorem by Hunt [9] for $1 < p < \infty$; see also [7],[15], and [8].

The main purpose of this paper is to sharpen Theorem 1.1 towards control of the variation norm in the parameter ξ. Thus we consider mixed L^p and V^r norms of the type:

$$\|\mathcal{S}[f]\|_{L^p(V^r)} = \left(\int_{\mathbb{R}} \sup_{\substack{K\in\mathbb{N} \\ \xi_0<\ldots<\xi_K\in\mathbb{R}}} \left(\sum_{k=1}^{K} |\mathcal{S}[f](\xi_k, x) - \mathcal{S}[f](\xi_{k-1}, x)|^r \right)^{\frac{p}{r}} dx \right)^{\frac{1}{p}}.$$

We will prove the following, where $r' = r/(r-1)$:

Theorem 1.2. *Suppose $r > 2$ and $r' < p < \infty$. Then*

$$\|\mathcal{S}[f]\|_{L^p(V^r)} \le C_{p,r} \|f\|_{L^p(\mathbb{R})}. \tag{2}$$

At the endpoint $r' = p$ we have the result:

Theorem 1.3. *Suppose $2 < r < \infty$ and $r' = p$. Then for all measurable functions f and sets F with $|f| \le 1_F$, we have*

$$\lambda^p |\{x : \|\mathcal{S}[f](\cdot, x)\|_{V^r_\xi} \ge \lambda\}| \le C_p |F|.$$

Note that if in the above definition of the mixed L^p and V^r norm we interchange the order between integration in the x variable and taking the supremum over the choices of K and the points ξ_0 to ξ_K so that these choices become independent of the variable x, then the estimates corresponding to Theorems 1.2 and 1.3 are weaker and follow by an inequality of Rubio de Francia [22], see also the proof [13] which is closer to the methods of this paper. As will be discussed in Section 2, the conditions on the exponents in Theorem 1.2 are sharp, and in the range of Lorentz norms no better than the stated weak-type estimate is possible in Theorem 1.3.

While the concept of r-variation norm is at least as old as Wiener's 1920s paper on quadratic variation [25], such norms and related oscillation norms have been pioneered by Bourgain [1] as a tool to prove convergence results for ergodic averages. Bourgain's simple motivation is that the variational estimate, rather than the weaker L^∞ estimate, allows him to prove pointwise convergence without previous knowledge that pointwise convergence holds for a dense subclass of functions. Such dense subclasses of functions, while usually available in the setting of analysis on Euclidean space, are less abundant in the ergodic theory setting. In Appendix D we demonstrate the use of Theorem 1.2 in the setting of Wiener-Wintner type theorems as developed in [14].

Additionally, we are motivated by the fact that variation norms are in certain situations more stable under nonlinear perturbation than supemum norms. For example one can deduce bounds for certain r-variational lengths of curves in Lie groups from the corresponding lengths of the "trace" of the curves in the corresponding Lie algebras, see Appendix C for definitions and details. What we have in mind is proving Carleson type theorems for nonlinear perturbations of the Fourier transform as discussed in [19], [20]. Unfortunately the naive approach fails and the ultimate goal remains unattained since we only know the correlation between lengths of the trace and the original curve for $r < 2$, while the variational Carleson theorem only holds for $r > 2$. Nonetheless, this method allows one to see that a variational version of the Christ-Kiselev theorem [4] follows from a variational Menshov-Paley-Zygmund theorem which we prove in Appendix B. The variational Carleson theorem can be viewed as an endpoint estimate in this theory.

The Carleson-Hunt theorem has previously been generalized by using other norms in place of the variation norm, see for example [14], [5], [6].

Our proof of Theorem 1.2 will follow the method of [15] as refined in [8]. In Section 3 we reduce the problem to that of bounding certain model operators which map f to linear combinations of wave-packets associated to collections of multitiles. In Section 5 we bound the model operators when the collection of multitiles is of a certain type called a tree; this bound is in terms of two quantities, energy and density, which are associated to the tree. These quantities are defined in Section 4 and an algorithm is given to decompose an arbitrary collection of multitiles into a union of trees with controlled energy and density. These ingredients are combined to complete the proof in Section 6. Finally, a variational estimate which is crucial for the proof of the model operator bound for trees is given in Appendix A.

The first author was partially supported by NSF VIGRE grant DMS 0502315. The second author was partially supported by NSF grant DMS 0652890. The third author was partially supported by NSF grant DMS 0701302. The fourth author was partially supported by a grant from the McArthur Foundation, and by NSF grant DMS 0649473. The fifth author was partially supported by an EPSRC grant.

2. Optimality of the exponents

In [11] it was shown that the condition $r > 2$ is necessary for the Fourier series analog of the bound (2) to hold; we begin by noting that similar considerations apply to the Fourier transform on the real line. For any integer k, consider the dyadic averaging operator

$$\mathbb{E}_k[f](x) = \frac{1}{|I_k(x)|}\int_{I_k(x)} f(y)\, dy$$

where $I_k(x)$ is the dyadic interval of length 2^k containing x. From arguments in [21] and [11] one sees that $\mathbb{E}$ is unbounded from $L^p \to L^p_x(V^2_k)$. Applying the square-function estimate from Appendix A it then follows that for $1 < p < \infty$, the operator $f \to f * \psi_k$ is unbounded from $L^p \to L^p(V^2_k)$, where $*$ denotes convolution, where ψ is a Schwartz function with $\hat{\psi}(\xi) = 1$ for $|\xi| \le 1$ and $\hat{\psi} = 0$ for $|\xi| > 2$, and where $\psi_k = 2^{-k}\psi(2^{-k}\cdot)$. Letting $\mathcal{S}_t[f](x) = \mathcal{S}[f](t,x) - \mathcal{S}[f](-t,x)$ one applies the standard estimates

$$\|\left(\sum_{k\in\mathbb{Z}} |\mathcal{S}_{2^{-k}}[g_k]|^2\right)^{1/2}\|_{L^p(\mathbb{R})} \le C_p \|\left(\sum_{k\in\mathbb{Z}} |g_k|^2\right)^{1/2}\|_{L^p(\mathbb{R})}$$

and

$$\|\left(\sum_{k\in\mathbb{Z}} |(\psi_k - \psi_{k+1}) * f|^2\right)^{1/2}\|_{L^p(\mathbb{R})} \le C_p \|f\|_{L^p(\mathbb{R})}$$

with $g_k = (\psi_k - \psi_{k+1}) * f$ to see that

$$\|\left(\sum_{k\in\mathbb{Z}} |\mathcal{S}_{2^{-k}}[f] - \psi_{k+1} * f|^2\right)^{1/2}\|_{L^p(\mathbb{R})} \le C_p \|f\|_{L^p(\mathbb{R})}$$

for $1 < p < \infty$. We thus have $f \to \mathcal{S}_{2^{-k}}[f](x)$ unbounded from $L^p \to L^p_x(V^2_k)$ and hence $\mathcal{S}$ is unbounded from $L^p \to L^p(V^2)$ for any p.

The necessity of the condition $p > r'$ is a consequence of the following argument. First note that, for $1 \le t \le 2$ we have

$$\mathcal{S}_t[\psi_{-1}](x) = \frac{\sin(tx)}{\pi x}.$$

For integers n let $t_{x,n} = \frac{\pi/2 + n\pi}{x}$ so that

$$\left|\frac{\sin(t_{x,n}x)}{x} - \frac{\sin(t_{x,n+1}x)}{x}\right| = \frac{2}{|x|}.$$

For each x let $E(x) = [1,2] \cap \{t_{x,n} : n \in \mathbb{Z}\}$ and note that for large $|x|$, the cardinality of $E(x)$ is $\ge C|x|$ and so

$$\|\mathcal{S}_t[\psi_{-1}](x)\|_{V^r_t} \ge C(1+|x|)^{-1/r'}.$$

It follows that $\|\mathcal{S}[\psi_{-1}]\|_{L^p_x(V^r_t)} = \infty$ for $p < r'$, and in fact the Lorentz norm $\|\mathcal{S}[\psi_{-1}]\|_{L^{r',s}_x(V^r_t)} = \infty$ for $s < \infty$.

3. The model operators

To start the proof of Theorems 1.2 and 1.3, we first linearize the variation norm. Fix K, measurable real valued functions $\xi_0(x) < \ldots < \xi_K(x)$, and measurable complex valued functions $a_1(x), \ldots, a_K(x)$ satisfying $|a_1(x)|^{r'} + \ldots + |a_K(x)|^{r'} = 1$. Letting

$$\mathcal{S}'[f](x) = \sum_{k=1}^{K} (\mathcal{S}[f](\xi_k(x), x) - \mathcal{S}[f](\xi_{k-1}(x), x))a_k(x)$$

Theorem 1.2 will follow by standard arguments from the estimate

$$\|\mathcal{S}'[f]\|_{L^p(\mathbb{R})} \leq C\|f\|_{L^p(\mathbb{R})} \tag{3}$$

where C is independent of K and the linearizing functions, and where f is any Schwartz function (an analogous statement holds for the endpoint $p = r'$ result, all such considerations for Theorem 1.3 will henceforth remain implicit).

Let $\mathcal{D} = \{[2^k m, 2^k(m+1)) : m, k \in \mathbb{Z}\}$ be the set of dyadic intervals. A tile will be any rectangle $I \times \omega$ where I, ω are dyadic intervals, and $|I||\omega| = 1/2$. We will write $\mathcal{S}'$ as the sum of wave packets adapted to tiles, and then decompose the operator into a finite sum of model operators by sorting the wave packets into a finite number of classes.

For each k,

$$\mathcal{S}[f](\xi_k, x) - \mathcal{S}[f](\xi_{k-1}, x) = \int 1_{(\xi_{k-1},\xi_k)}(\xi)\hat{f}(\xi)e^{2\pi i\xi x}\, d\xi.$$

To suitably express the difference above as a sum of wave packets, we will first need to construct a partition of $1_{(\xi_{k-1},\xi_k)}$ adapted to certain dyadic intervals. The fact that (ξ_{k-1}, ξ_k) has two boundary points instead of the one from $(-\infty, \xi_k)$ will necessitate a slightly more involved discretization argument than that in [15].

For any $\xi < \xi'$, let $\mathbf{J}_{\xi,\xi'}$ be the set of maximal dyadic intervals J such that $J \subset (\xi, \xi')$ and $\operatorname{dist}(J, \xi), \operatorname{dist}(J, \xi') \geq |J|$. Let ν be a smooth function from $\mathbb{R}$ to $[0, 1]$ which vanishes on $(-\infty, -1/100]$ and is identically equal to 1 on $[1/100, \infty)$. Given an interval $J = [a, b)$ and $i \in \{-1, 0, 1\}$, define

$$\varphi_{J,i}(\xi) = \nu\left(\frac{\xi - a}{2^i|J|}\right) - \nu\left(\frac{\xi - b}{|J|}\right).$$

For each $J \in \mathbf{J}_{\xi,\xi'}$, one may check that there is a unique interval $J' \in \mathbf{J}_{\xi,\xi'}$ which lies strictly to the left of J and satisfies $\operatorname{dist}(J', J) = 0$, and one may check that J' has size $|J|/2, |J|$, or $2|J|$. We define $\varphi_J = \varphi_{J,i(J)}$ where $i(J)$ is chosen so that $|J'| = 2^{i(J)}|J|$. Then

$$1_{(\xi,\xi')} = \sum_{J\in\mathbf{J}_{\xi,\xi'}} \varphi_J.$$

We now write each multiplier φ_J as the sum of wave packets. For every tile $P = I \times J$, define $\phi_P(x) = \sqrt{|I|}\check{\sqrt{\varphi_J}}(x - c(I))$ where $c(I)$ denotes the center of the interval I and $\check{}$ denotes the inverse Fourier transform. For each J, we then have

$$\sum_{|I|=1/(2|J|)} \langle f, \phi_{I\times J}\rangle\, \widehat{\phi_{I\times J}} = \hat{f}\varphi_J.$$

This gives:

$$\mathcal{S}'[f](x) = \sum_{k=1}^{K}\left(\sum_{J\in\mathbf{J}_{\xi_{k-1}(x),\xi_k(x)}}\ \sum_{|I|=(1/(2|J|))} \langle f, \phi_{I\times J}\rangle\, \phi_{I\times J}\right) a_k(x).$$

The wave packets will be sorted into a finite number of classes, each well suited for further analysis. Sorting is accomplished by dividing every $\mathbf{J}_{\xi,\xi'}$ into a finite number of disjoint sets. These sets will be indexed by a fixed subset of $\{1, 2, 3\} \times \{1, 2, 3, 4\}^2 \times \{\text{left, right}\}$. Specifically, for each $(m, n, side) \in \{1, 2, 3, 4\}^2 \times \{\text{left, right}\}$, we define

- $\mathbf{J}_{\xi,\xi',(1,m,n,side)} = \{J \in \mathcal{D} : J \subset (\xi, \xi')$, ξ is in the interval $J-(m+1)|J|$, ξ' is in the interval $J + (n + 1)|J|$, and J is the *side*-child of its dyadic parent$\}$.

- $\mathbf{J}_{\xi,\xi',(2,m,n,side)} = \{J \in \mathcal{D} : J \subset (\xi,\xi'),\ \xi$ is in the interval $J - (m+1)|J|$, $\operatorname{dist}(\xi', J) \geq n|J|$, and J is the *side*-child of its dyadic parent$\}$.
- $\mathbf{J}_{\xi,\xi',(3,m,n,side)} = \{J \in \mathcal{D} : J \subset (\xi,\xi'),\ \operatorname{dist}(\xi, J) > m|J|,\ \xi'$ is in the interval $J + (n+1)|J|$, and J is the *side*-child of its dyadic parent$\}$.

We will choose $R \subset \{1,2,3\} \times \{1,2,3,4\}^2 \times \{\text{left, right}\}$ so that for each ξ, ξ', the collection $\{\mathbf{J}_{\xi,\xi',\rho}\}_{\rho \in R}$ is pairwise disjoint and $\mathbf{J}_{\xi,\xi'} = \cup_{\rho \in R} \mathbf{J}_{\xi,\xi',\rho}$. We will also assume that for each $\rho \in R$ there is an $i(\rho) \in \{-1,0,1\}$ such that $|J'| = 2^{i(\rho)}|J|$ for every $\xi < \xi'$, $J \in \mathbf{J}_{\xi,\xi',\rho}$ and $J' \in \mathbf{J}_{\xi,\xi'}$ with J' strictly to the left of J and $\operatorname{dist}(J,J') = 0$. One may check that these conditions are satisfied, say, for

$$\begin{aligned} R = \{&(1,2,1,\text{left}), (1,2,2,\text{left}), (1,3,1,\text{left}), \\ &(1,3,2,\text{left}), (2,1,1,\text{left}), (2,1,1,\text{right}), (2,2,1,\text{right}), \\ &(3,4,1,\text{left}), (3,3,1,\text{right}), (3,4,2,\text{left})\}. \end{aligned}$$

It now follows that

$$\mathcal{S}'[f] = \sum_{\rho \in R} \mathcal{S}^\rho[f]$$

where

$$\mathcal{S}^\rho[f](x) = \sum_{k=1}^{K} \left(\sum_{J \in \mathbf{J}_{\xi_{k-1}(x),\xi_k(x),\rho}} \ \sum_{|I| = 1/(2|J|)} \langle f, \phi_{I\times J}\rangle\, \phi_{I\times J} \right) a_k(x).$$

It will be convenient to rewrite each operator $\mathcal{S}^\rho$ in terms of multitiles. A multitile will be a subset of $\mathbb{R}^2$ of the form $I \times \omega$ where $I \in \mathcal{D}$ and where ω is the union of three intervals $\omega_l, \omega_u, \omega_h$. For each $\rho = (l,m,n,side) \in R$, we consider a set of ρ-multitiles which is parameterized by $\{(I,\omega_u) : I, \omega_u \in \mathcal{D}, |I||\omega_u| = 1/2,$ and ω_u is the *side*-child of its parent$\}$. Specifically, given $\omega_u = [a,b)$

- If $\rho = (1,m,n,side)$ then $\omega_l = \omega_u - (m+1)|\omega_u|$ and $\omega_h = \omega_u + (n+1)|\omega_u|$.
- If $\rho = (2,m,n,side)$ then $\omega_l = \omega_u - (m+1)|\omega_u|$ and $\omega_h = [a + (n+1)|\omega_u|, \infty)$
- If $\rho = (3,m,n,side)$ then $\omega_l = (-\infty, b - (m+1)|\omega_u|)$ and $\omega_h = \omega_u + (n+1)|\omega_u|$.

For every ρ-multitile P, let $a_P(x) = a_k(x)$ if k satisfies $1 \leq k \leq K$ and $\xi_{k-1}(x) \in \omega_l$ and $\xi_k \in \omega_h$ (such a k would clearly be unique), and $a_P(x) = 0$ if there is no such k. Then, using $\mathbf{P}_\rho$ to denote the set of ρ-multitiles, we have

$$\mathcal{S}^\rho[f](x) = \sum_{P \in \mathbf{P}_\rho} \langle f, \phi_P\rangle\, \phi_P(x) a_P(x)$$

where, for each ρ-multitile P, $\phi_P(x) = \sqrt{|I|}\sqrt{\check{\varphi}_{\omega_u, i(\rho)}}(x - c(I))$.

Inequality (3) and hence Theorem 1.2 will then follow after proving the bound

$$\|\mathcal{S}^\rho[f]\|_{L^p} \leq C\|f\|_{L^p} \tag{4}$$

for each $\rho \in R$. The argument for the case $\rho = (3,m,n,side)$ is analogous to that for the case $\rho = (2,m,n,side)$, so below we will assume $\rho = (2,m,n,side)$ in which case we say that ρ is a 2-index or $\rho = (1,m,n,side)$ in which case we say that ρ is a 1-index.

4. Energy and density

We want to prove

$$\| \sum_P \langle f, \phi_P \rangle \, \phi_P a_P \|_{L^p(\mathbb{R})} \leq C \|f\|_{L^p(\mathbb{R})}$$

where P ranges over an arbitrary finite collection of ρ-multitiles, ρ is a 1 or 2-index, and C does not depend on this collection or on the linearizing functions (which were used to define the functions a_P). By a standard limiting argument, this is sufficient to prove (4) and hence Theorem 1.2.

The wave packets ϕ_P are adapted to the multitiles P in the following sense. For each P, $\hat{\phi}_P$ is supported on the interval with the same center as ω_u and $\frac{11}{10}$ the diameter, which we denote $\frac{11}{10}\omega_u$. Fixing a large C and N and defining, for each I,

$$w_I(x) = C \frac{1}{|I|} \left(1 + \frac{|x - c(I)|}{|I|} \right)^{-N}$$

we have

$$\left| \frac{d^n}{dx^n} (e^{-2\pi i c(\omega_u) x} \phi_P)(x) \right| \leq C' |I|^{(1/2)-n} |w_I(x)| \tag{5}$$

for $n \geq 0$, where the constant above may depend on n.

Fix $1 \leq C_3 < C_2 < C_1$ such that for every multitile P, $\hat{\phi}_P$ is supported on $C_3\omega_u$, $C_2\omega_u \cap C_2\omega_l = \emptyset, C_2\omega_u \cap \omega_h = \emptyset, C_2\omega_l \subset C_1\omega_u, C_2\omega_u \subset C_1\omega_l$. One may check that the values $C_3 = 11/10, C_2 = 2$, and $C_1 = 12$ satisfy these properties.

Given a dyadic interval I_T and a point $\xi_T \in \mathbb{R}$, we say that a collection T of multitiles is a tree with top interval I_T and top frequency ξ_T if $I \subset I_T$ and $\omega_T \subset \omega_m$ for every $P \in T$ where ω_T is the interval $[\xi_T - (C_2-1)/(4|I_T|), \xi_T + (C_2-1)/(4|I_T|))$ and ω_m is the convex hull of $C_2\omega_u \cup C_2\omega_l$. A tree T will be said to be l-overlapping if for every $P \in T$, $\xi_T \in C_2\omega_l$; it will be said to be l-lacunary if for every $P \in T$, $\xi_T \notin C_2\omega_l$.

We split our arbitrary finite collection of multitiles into a bounded number of subcollections (i.e. henceforth all multitiles will be assumed to belong to a fixed subcollection) to obtain the following two separation properties.

$$\text{If } P, P' \text{ satisfy } |\omega_u'| < |\omega_u|, \text{ then } |\omega_u'| \leq \frac{C_2 - C_3}{2C_1} |\omega_u|. \tag{6}$$

$$\text{If } P, P' \text{ satisfy } C_1\omega_u \cap C_1\omega_u' \neq \emptyset \text{ and } |\omega_u| = |\omega_u'| \text{ then } \omega_u = \omega_u'. \tag{7}$$

From (6), it follows that if $P, P' \in T$, T is a l-lacunary tree, and $|\omega_u'| < |\omega_u|$ then $C_3\omega_l \cap C_3\omega_l' = \emptyset$, and that if $P, P' \in T$, T is an l-overlapping tree, and $|\omega_u'| < |\omega_u|$ then $C_3\omega_u \cap C_3\omega_u' = \emptyset$. From (7), it follows that if $P, P' \in T$, T is a tree, and $|\omega_u| = |\omega_u'|$, then $I \cap I' = \emptyset$.

Given any collection of multitiles $\mathbf{P}$, we define

$$\text{energy}(\mathbf{P}) = \sup_T \sqrt{\frac{1}{|I_T|} \sum_{P \in T} | \langle f, \phi_P \rangle |^2}$$

where the sup ranges over all l-overlapping trees $T \subset \mathbf{P}$. We set

$$\text{density}(\mathbf{P}) =$$

$$= \sup_T \left(\frac{1}{|I_T|} \int_E (1 + |x - c(I_T)|/|I_T|)^{-4} \sum_{k=1}^{K} |a_k(x)|^{r'} 1_{\omega_T}(\xi_{k-1}(x))\, dx \right)^{1/r'}$$

where the sup is over all non-empty trees $T \subset \mathbf{P}$, and where $E \subset \mathbb{R}$ is a fixed set which will be chosen later .

The following proposition allows one to decompose an arbitrary collection of multitiles into the union of trees, where the trees are divided into collections $\mathbf{T}_j$ with the energy of trees from $\mathbf{T}_j$ bounded by 2^{-j}. The control over energy is balanced by an L^q bound for the functions $N_{j,l} = \sum_{T \in \mathbf{T}_j} 1_{2^l I_T}$. In contrast to [15] and [8], it is necessary here to consider $q > 1$ and $l > 0$ in order to effectively use the tree estimate Proposition 5.1 with $q > 1$. The bound (11) permits one to make further decompositions to take advantage of large $|F|$ in the L^q bound for the $N_{j,l}$ while maintaining compatibility with bounds for trees with a fixed density obtained from Proposition 4.2.

Proposition 4.1. *Let* $\mathbf{P}$ *be a collection of multitiles with energy bounded above by* e*, and let* $|f|$ *be bounded above by* 1_F*. Then, there is a collection of trees* $\mathbf{T}$ *such that*

$$\sum_{T \in \mathbf{T}} |I_T| \leq C e^{-2} |F| \tag{8}$$

and

$$\text{energy}\,(\mathbf{P} \setminus \cup_{T \in \mathbf{T}} T) \leq e/2.$$

and such that, for every integer $l \geq 0$*,*

$$\| \sum_{T \in \mathbf{T}} 1_{2^l I_T} \|_{BMO} \leq C 2^{2l} e^{-2}. \tag{9}$$

Furthermore, if for some collection of trees $\mathbf{T}'$*,*

$$\mathbf{P} = \bigcup_{T' \in \mathbf{T}'} T' \tag{10}$$

then

$$\sum_{T \in \mathbf{T}} |I_T| \leq C \sum_{T' \in \mathbf{T}'} |I_{T'}|. \tag{11}$$

Above, and subsequently, $\| \cdot \|_{BMO}$ denotes the dyadic BMO norm.

Proof. Without loss of generality, assume that $e > 0$. We select trees through an iterative procedure. Suppose that trees S_k, T_k have been chosen for $k = 1, \ldots, j$. Set

$$\mathbf{P}_j = \mathbf{P} \setminus \bigcup_{k=1}^{j} T_k$$

If energy$(\mathbf{P}_j) \leq e/2$ then we terminate the procedure, set $\mathbf{T} = \{T_k\}_{1 \leq k \leq j}$ and $n = j$. Otherwise, we may find an l-overlapping tree $S \subset \mathbf{P}_j$ such that

$$\frac{1}{|I_S|} \sum_{P \in S} | \langle f, \phi_P \rangle |^2 \geq e^2/4. \tag{12}$$

Choose such a tree S_{j+1} with $\xi_{S_{j+1}}$ maximal in the sense that for any l-overlapping tree S satisfying (12) with $\xi_S > \xi_{S_{j+1}}$ we have that $(S_{j+1}, \xi_S, I_{S_{j+1}})$ is an l-overlapping tree. Let T_{j+1} be the maximal, with respect to inclusion, tree with top data $(\xi_{S_{j+1}}, I_{S_{j+1}})$. This process will eventually stop since each T_j is nonempty and $\mathbf{P}$ is finite.

To verify (8) it suffices to show

$$\left(\frac{e^2}{|F|}\sum_{j=1}^{n}|I_{S_j}|\right)^2 \leq C\frac{e^2}{|F|}\sum_{j=1}^{n}|I_{S_j}|.$$

Since the S_j satisfy (12), we have

$$\left(\frac{e^2}{|F|}\sum_{j=1}^{n}|I_{S_j}|\right)^2 \leq \left(4\sum_{j=1}^{n}\sum_{P\in S_j}|\langle\frac{f}{|F|^{1/2}},\phi_P\rangle|^2\right)^2$$

Since $\|\frac{f}{|F|^{1/2}}\|_{L^2} \leq 1$, the right hand side above is

$$\leq 16\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{P\in S_j}\sum_{P'\in S_k}|\langle\frac{f}{|F|^{1/2}},\phi_P\rangle||\langle\frac{f}{|F|^{1/2}},\phi_{P'}\rangle||\left\langle\phi_P,\phi_{P'}\right\rangle|.$$

By symmetry, it remains, for (8), to show that

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{P\in S_j}\sum_{P'\in S_k:|I'|=|I|}|\langle f,\phi_P\rangle||\langle f,\phi_{P'}\rangle||\left\langle\phi_P,\phi_{P'}\right\rangle| \leq Ce^2\sum_{j=1}^{n}|I_{S_j}| \tag{13}$$

and

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{P\in S_j}\sum_{P'\in S_k:|I'|<|I|}|\langle f,\phi_P\rangle||\langle f,\phi_{P'}\rangle||\left\langle\phi_P,\phi_{P'}\right\rangle| \leq Ce^2\sum_{j=1}^{n}|I_{S_j}|. \tag{14}$$

In both cases, we will use the estimate

$$|\left\langle\phi_P,\phi_{P'}\right\rangle| \leq C\left(\frac{|I|}{|I'|}\right)^{1/2}\left\langle w_I, 1_{I'}\right\rangle. \tag{15}$$

which holds whenever $|I'| \leq |I|$.

Estimating the product of two terms by the square of their maximum, we see that the left side of (13) is

$$\leq 2\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{P\in S_j}\sum_{P'\in S_k:|I'|=|I|}|\langle f,\phi_P\rangle|^2|\left\langle\phi_P,\phi_{P'}\right\rangle|.$$

Recall that $\langle\phi_P,\phi'_P\rangle = 0$ unless $C_3\omega_u \cap C_3\omega'_u \neq \emptyset$. Thus, by (7), (15) and the fact that the S_k are pairwise disjoint, we have that the display above is

$$\leq 2\sum_{j=1}^{n}\sum_{P\in S_j}|\langle f,\phi_P\rangle|^2\sum_{I':|I'|=|I|}\langle w_I, 1_{I'}\rangle \leq 2C\sum_{j=1}^{n}\sum_{P\in S_j}|\langle f,\phi_P\rangle|^2.$$

Since the energy of $\mathbf{P}$ is bounded above by e, the right side above is

$$\leq 2C\sum_{j=1}^{n}e^2|I_{S_j}|$$

which finishes the proof of (13).

Applying Cauchy-Schwarz, we see that the left side of (14) is

$$\leq \sum_{j=1}^{n} \left(\sum_{P \in S_j} |\langle f, \phi_P \rangle|^2 \right)^{1/2} \left(\sum_{P \in S_j} \left(\sum_{k=1}^{n} \sum_{P' \in S_k : |I'| < |I|} |\langle f, \phi_{P'} \rangle| \, | \langle \phi_P, \phi_{P'} \rangle | \right)^2 \right)^{1/2}.$$

Twice using the fact that the energy of $\mathbf{P}$ is bounded by e, we see that the display above is

$$\leq e^2 \sum_{j=1}^{n} |I_{S_j}|^{1/2} \left(\sum_{P \in S_j} \left(\sum_{k=1}^{n} \sum_{P' \in S_k : |I'| < |I|} | \left\langle \phi_P, |I_{P'}|^{1/2} \phi_{P'} \right\rangle | \right)^2 \right)^{1/2}.$$

Thus, to prove (14) it remains to show that, for each j,

$$\sum_{P \in S_j} \left(\sum_{k=1}^{n} \sum_{P' \in S_k : |I'| < |I|} | \left\langle \phi_P, |I_{P'}|^{1/2} \phi_{P'} \right\rangle | \right)^2 \leq C |I_{S_j}|.$$

Again, we only have $| \left\langle \phi_P, |I_{P'}|^{1/2} \phi_{P'} \right\rangle |$ nonzero when $C_3 \omega_u \cap C_3 \omega'_u \neq \emptyset$ which can only happen if $\sup C_3 \omega_u \in C_3 \omega'_u$ or $\inf C_3 \omega_u \in C_3 \omega'_u$. Applying (15), we thus see that the left side of above is

$$\leq 2 \sum_{P \in S_j} \left(\sum_{k=1}^{n} \sum_{\substack{P' \in S_k : |I'| < |I| \\ \sup C_3 \omega_u \in C_3 \omega'_u}} |I_P|^{1/2} \langle w_I, 1_{I'} \rangle \right)^2$$

$$+ 2 \sum_{P \in S_j} \left(\sum_{k=1}^{n} \sum_{\substack{P' \in S_k : |I'| < |I| \\ \inf C_3 \omega_u \in C_3 \omega'_u}} |I_P|^{1/2} \langle w_I, 1_{I'} \rangle \right)^2.$$

Suppose $P \in S_k$, $P' \in S_{k'}$, $P \neq P'$, $|I'| \leq |I|$ and $C_3 \omega_u \cap C_3 \omega'_u \neq \emptyset$. If $|I'| = |I|$, then from (7) it follows that $\omega_u = \omega'_u$ and hence, since $P \neq P'$, we have $I \cap I' = \emptyset$. If $|I'| < |I|$, then from (6) it follows that $\xi_{S_k} > \xi_{S_{k'}}$ and $\xi_{S_k} \notin C_2 \omega_{l'}$, and hence $k < k'$. But $\omega_{S_k} \subset \omega'_m$ and $P' \notin T_k$, so $I' \cap I_{S_k} = \emptyset$. We conclude that each of the two terms above is

$$\leq 2 \sum_{P \in S_j} |I_P| \left\langle w_I, 1_{\mathbb{R} \setminus I_{S_j}} \right\rangle^2 \leq 2 \sum_{l : 2^l \leq |I_{S_j}|} 2^l \sum_{P \in S_j : |I| = 2^l} \left\langle w_I, 1_{\mathbb{R} \setminus I_{S_j}} \right\rangle \tag{16}$$

One may check that for each l

$$\sum_{P \in S_j : |I| = 2^l} \left\langle w_I, 1_{\mathbb{R} \setminus I_{S_j}} \right\rangle \leq C$$

and so the right side of (16) is $\leq C |I_{S_j}|$, which finishes the proof of (14) and thus (8).

For (9), we need to show that for each dyadic interval J, we have

$$\frac{1}{|J|} \int_J | \sum_{T \in \mathbf{T}} 1_{2^l I_T}(x) - \frac{1}{|J|} \int_J \sum_{T \in \mathbf{T}} 1_{2^l I_T}(y) \, dy | \, dx \leq C 2^{2l} e^{-2}.$$

To this end, it will suffice to show that

$$\sum_{\substack{T\in\mathbf{T}\\ J\cap 2^l I_T\neq\emptyset, J}} |I_T| \leq Ce^{-2}2^l|J|. \tag{17}$$

Let $\tilde{\mathbf{T}} = \{T \in \mathbf{T} : I_T \subset 2^{l+1}J, |I_T| \leq |J|\}$ and note that if $T \in \mathbf{T}$ with $2^l I_T \cap J \neq \emptyset, J$ then $T \in \tilde{\mathbf{T}}$. Write $f = f' + f''$ where $|f'| \leq 1_{2^{l+5}J}$ and $|f''| \leq 1_{\mathbb{R}\setminus 2^{l+5}J}$.

We will write $\tilde{\mathbf{T}}$ as the union of collections of trees $\mathbf{T}' \cup \mathbf{T}^0 \cup \mathbf{T}^1 \cup \ldots$ each of which will have certain properties related to the energy. For each tree $T \in \tilde{\mathbf{T}}$ there is an l-overlapping tree S chosen in the algorithm above with $I_S = I_T$ and

$$\frac{1}{|I_S|}\sum_{P\in S} |\langle f, \phi_P\rangle|^2 \geq e^2/4. \tag{18}$$

Let $\mathbf{T}^0 = \{T \in \tilde{\mathbf{T}} : \frac{1}{|I_S|}\sum_{P\in S} |\langle f'', \phi_P\rangle|^2 \geq e^2/16\}$. For $j \geq 1$, define

$$\mathbf{T}^j = \{T \in \tilde{\mathbf{T}} : \sup_{\substack{S'\subset S\\ |I_{S'}|=2^{-j}|I_T|}} \frac{1}{|I_{S'}|}\sum_{P\in S'} |\langle f'', \phi_P\rangle|^2 \geq e^2/16\}$$

where, for each T, the sup above is taken over all l-overlapping trees S' with $S' \subset S$. We then let $\mathbf{T}' = \{T \in \tilde{\mathbf{T}} \setminus (\mathbf{T}^0 \cup \mathbf{T}^1 \cup \ldots)\}$.

For each j, we have

$$\sum_{T\in\mathbf{T}^j} |I_T| \leq C \sum_{T\in\mathbf{T}^j} 2^j e^{-2} \sum_{P\in S:|I|\leq 2^{-j}|J|} |\langle f'', \phi_P\rangle|^2.$$

Since the S above are pairwise disjoint, the right hand side is

$$\leq C2^j e^{-2}\sum_{k\geq j}\sum_{\substack{P:|I|=2^{-k}|J|\\ I\subset 2^{l+1}J}} |\langle f'', \phi_P\rangle|^2. \tag{19}$$

Fixing k, we apply Minkowski's inequality to obtain

$$\sum_{\substack{P:|I|=2^{-k}|J|\\ I\subset 2^{l+1}J}} |\langle f'', \phi_P\rangle|^2 \leq \left(\sum_{\substack{K:|K|=2^{1-k}|J|\\ K\cap 2^{l+2}J=\emptyset}} \left(\sum_{\substack{P:|I|=2^{-k}|J|\\ I\subset 2^{l+1}J}} |\langle 1_K f'', \phi_P\rangle|^2\right)^{1/2}\right)^2$$

where above, we sum over dyadic intervals K and use the fact that f'' is supported on $\mathbb{R}\setminus 2^{l+5}J$. Since $\phi_{P'} = ce^{2\pi i(c(\omega'_u)-c(\omega_u))\cdot}\phi_P$ when $I = I'$, we may use orthogonality and the fact that $|f''| \leq 1$ to see that the right side above is

$$\leq C2^{-k}|J|\left(\sum_{\substack{K:|K|=2^{1-k}|J|\\ K\cap 2^{l+2}J=\emptyset}} \left(\sum_{\substack{I:|I|=2^{-k}|J|\\ I\subset 2^{l+1}J}} \|1_K\phi_{P_I}\|_{L^2}^2\right)^{1/2}\right)^2$$

where for each I, P_I is any multitile with time interval I. Using (5) gives

$$\|1_K\phi_{P_I}\|_{L^2}^2 \leq C(1 + \mathrm{dist}(K, I)/|I|)^{-N},$$

and so we see that the display above is

$$\leq C2^{-k(N-2)}|J|$$

Summing over k and j, we conclude that

$$\sum_{T\in\cup_j \mathbf{T}_j} |I_T| \le Ce^{-2}|J|.$$

Thus, to prove (17), it suffices to show

$$\sum_{T\in\mathbf{T}'} |I_T| \le Ce^{-2}2^l|J|.$$

Let $T\in\mathbf{T}'$ and let S' be any l-overlapping tree contained in S satisfying $|I_{S'}|\le |I_S|$. Since the energy of $\mathbf{P}$ is bounded by e and since T is not in any $\mathbf{T}^j$, we have

$$\frac{1}{|S'|}\sum_{P\in S'}|\langle f',\phi_P\rangle|^2 \le 2\frac{1}{|S'|}\sum_{P\in S'}|\langle f,\phi_P\rangle|^2 + 2\frac{1}{|S'|}\sum_{P\in S'}|\langle f'',\phi_P\rangle|^2 \le Ce^2.$$

From (12) and the fact that $T\notin\mathbf{T}^0$, we have

$$\frac{1}{|S|}\sum_{P\in S}|\langle f',\phi_P\rangle|^2 \ge e^2/8 - e^2/16 = e^2/16.$$

By the same reasoning as in the proof of (8), we thus have

$$\sum_{T\in\mathbf{T}'}|I_T| \le Ce^{-2}\|f'\|_2^2 \le C'e^{-2}2^l|J|.$$

Moving on to (11), for each $T\in\mathbf{T}$, let S be the corresponding l-overlapping tree from the selection algorithm above and recall

$$\sum_{T\in\mathbf{T}}|I_T|(e/2)^2 \le \sum_{P\in\bigcup_{T\in\mathbf{T}}S}|\langle f,w_u\rangle|^2.$$

Since $\mathbf{P}=\bigcup_{T'\in\mathbf{T}'}T'$, the right side above is

$$\le \sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \xi_{T'}\in C_2\omega_l}}|\langle f,w_u\rangle|^2 + \sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \xi_{T'}\ge\inf C_3\omega_u}}|\langle f,w_u\rangle|^2$$
$$+\sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \sup C_2\omega_l\le\xi_{T'}<\inf C_3\omega_u}}|\langle f,w_u\rangle|^2.$$

Since $\mathbf{P}$ has energy bounded by e

$$\sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \xi_{T'}\in C_2\omega_l}}|\langle f,w_u\rangle|^2 \le \sum_{T'\in\mathbf{T}'}e^2|I_{T'}|.$$

Since the rectangles $\{I\times[\inf C_3\omega_u,\sup C_2\omega_u) : P\in\bigcup_{T\in\mathbf{T}}S\}$ are pairwise disjoint, we apply the energy bound again to see that

$$\sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \xi_{T'}\ge\inf C_3\omega_u}}|\langle f,w_u\rangle|^2 \le \sum_{T'\in\mathbf{T}'}\sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}}S\\ \xi_{T'}\ge\inf C_3\omega_u}}e^2|I| \le \sum_{T'\in\mathbf{T}'}e^2|I_{T'}|.$$

Now, suppose $P\in T'\cap S$, $\tilde{P}\in T'\cap\tilde{S}$ where $T,\tilde{T}\in\mathbf{T}$ and

$$\xi_{T'}\in[\sup C_2\omega_l,\inf C_3\omega_u)\cap[\sup C_2\tilde{\omega}_l,\inf C_3\tilde{\omega}_u),$$

and suppose $I\subset\tilde{I}$ and $P\ne\tilde{P}$. From (7) we have $I\subsetneq\tilde{I}$. We also have $\inf C_2\tilde{\omega}_l<\sup C_2\omega_l$ since otherwise it would follow that $\tilde{S}$ was selected prior to S and hence

$P \in \tilde{T}$ which is impossible. From (6), we have $\inf C_2\tilde{\omega}_l \geq \sup C_3\omega_l$ and so P is in the maximal l-overlapping tree with top data $(\tilde{I}, \inf C_2\tilde{\omega}_l)$.

For each $T' \in \mathbf{T}'$ let T'' be the collection of multitiles $P \in T' \cap \bigcup_{T\in\mathbf{T}} S$ with $\xi_{T'} \in [\sup C_2\omega_l, \inf C_3\omega_u)$ and I maximal among such multitiles. Then

$$\sum_{T'\in\mathbf{T}'} \sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}} S \\ \sup C_2\omega_l \leq \xi_{T'} < \inf C_3\omega_u}} |\langle f, w_u\rangle|^2 \leq \sum_{T'\in\mathbf{T}'} \sum_{P''\in T''} \sum_{\substack{P\in T'\cap\bigcup_{T\in\mathbf{T}} S \\ \sup C_2\omega_l \leq \xi_{T'} < \inf C_3\omega_u \\ I\subset I''}} |\langle f, w_u\rangle|^2$$

Considering the discussion in the preceding paragraph, we may apply the energy bound to see that the right side above is

$$\leq \sum_{T'\in\mathbf{T}'} \sum_{P''\in T''} 2e^2|I''| \leq \sum_{T'\in\mathbf{T}'} 2e^2|I_{T'}|.$$

We thus obtain (11). □

The proposition below is for use in tandem with Proposition 4.1.

Proposition 4.2. *Let $\mathbf{P}$ be a collection of multitiles and $d > 0$. Then, there is a collection of trees $\mathbf{T}$ such that*

$$\sum_{T\in\mathbf{T}} |I_T| \leq Cd^{-r'}|E| \tag{20}$$

and such that

$$\text{density}\,(\mathbf{P} \setminus \cup_{T\in\mathbf{T}} T) \leq d/2.$$

Proof. We select trees through an iterative procedure. Suppose that trees T_j, T_j^+, T_j^- have been chosen for $j = 1, \ldots, k$. Let

$$\mathbf{P}_k = \mathbf{P} \setminus \bigcup_{j=1}^{k} T_j \cup T_j^+ \cup T_j^-.$$

If density$(\mathbf{P}_k) \leq d/2$ then we terminate the procedure and set

$$\mathbf{T} = \{T_1, T_1^+, T_1^-, \ldots, T_k, T_k^+, T_k^-\}.$$

Otherwise, we may find a nonempty tree $T \subset \mathbf{P}_k$ such that

$$\frac{1}{|I_T|}\int_E (1 + |x - c(I_T)|/|I_T|)^{-4} \sum_{k:\xi_{k-1}(x)\in\omega_T} |a_k(x)|^{r'}\, dx > (d/2)^{r'}. \tag{21}$$

Choose $T_{k+1} \subset \mathbf{P}_k$ so that $|I_{T_{k+1}}|$ is maximal among all nonempty trees contained in $\mathbf{P}_k$ which satisfy (21), and so that T_{k+1} is the maximal, with respect to inclusion, tree contained in $\mathbf{P}_k$ with top data $(I_{T_{k+1}}, \xi_{T_{k+1}})$. Let $T_{k+1}^+ \subset \mathbf{P}_k$ be the maximal tree contained in $\mathbf{P}_k$ with top data $(I_{T_{k+1}}, \xi_{T_{k+1}} + (C_2-1)/(2|I_{T_{k+1}}|))$ and $T_{k+1}^- \subset \mathbf{P}_k$ be the maximal tree contained in $\mathbf{P}_k$ with top data $(I_{T_{k+1}}, \xi_{T_{k+1}} - (C_2-1)/(2|I_{T_{k+1}}|))$. Since each T_j is nonempty and $\mathbf{P}$ is finite, this process will eventually stop.

To prove (20), it will suffice to verify

$$\sum_j |I_{T_j}| \leq Cd^{-r'}|E|. \tag{22}$$

To this end, we first observe that the tiles $I_{T_j} \times \omega_{T_j}$ are pairwise disjoint. Indeed, suppose that $(I_{T_j} \times \omega_{T_j}) \cap (I_{T_{j'}} \times \omega_{T_{j'}}) \neq \emptyset$ and $j < j'$. Then, by the first maximality condition, we have $|I_{T_j}| \geq |I_{T_{j'}}|$ and so $I_{T_{j'}} \subset I_{T_j}$ and $|\omega_{T_j}| \leq |\omega_{T_{j'}}|$. From the latter inequality, it follows that for every $P \in T_{j'}$, either $\omega_{T_j} \subset \omega_m$, $\omega_{T_j^+} \subset \omega_m$, or $\omega_{T_j^-} \subset \omega_m$. Thus, $T_{j'} \subset T_j \cup T_j^+ \cup T_j^-$ which contradicts the selection algorithm.

Breaking the integral up into pieces and applying a pigeonhole argument, it follows from (21) that for each j there is a positive integer l_j such that

$$|I_{T_j}| \leq C 2^{-3l_j} d^{-r'} \int_{E \cap 2^{l_j} I_{T_j}} \sum_{k : \xi_{k-1}(x) \in \omega_{T_j}} |a_k(x)|^{r'} \, dx. \tag{23}$$

For each l we let $\mathbf{T}^{(l)} = \{T_j : l_j = l\}$ and choose elements of $\mathbf{T}^{(l)}$: $T_1^{(l)}, T_2^{(l)}, \ldots$ and subsets of $\mathbf{T}^{(l)}$: $\mathbf{T}_1^{(l)}, \mathbf{T}_2^{(l)}, \ldots$ as follows. Suppose $T_j^{(l)}$ and $\mathbf{T}_j^{(l)}$ have been chosen for $j = 1, \ldots, k$. If $\mathbf{T}^{(l)} \setminus \bigcup_{j=1}^k \mathbf{T}_j^{(l)}$ is empty, then terminate the selection procedure. Otherwise, let $T_{k+1}^{(l)}$ be an element of $\mathbf{T}^{(l)} \setminus \bigcup_{j=1}^k \mathbf{T}_j^{(l)}$ with $|I_{T_{k+1}^{(l)}}|$ maximal, and let

$$\mathbf{T}_{k+1}^{(l)} = \{T \in \mathbf{T}^{(l)} \setminus \bigcup_{j=1}^{k} \mathbf{T}_j^{(l)} : (2^l I_T \times \omega_T) \cap (2^l I_{T_{k+1}^{(l)}} \times \omega_{T_{k+1}^{(l)}}) \neq \emptyset\}.$$

By construction, $\mathbf{T}^{(l)} = \bigcup_j \mathbf{T}_j^{(l)}$ and so

$$\sum_{T \in \mathbf{T}^{(l)}} |I_T| \leq \sum_j \sum_{T \in \mathbf{T}_j^{(l)}} |I_T|. \tag{24}$$

Using the fact that the tiles $I_{T_j} \times \omega_{T_j}$ are pairwise disjoint, and (twice) the fact that $|I_T| \leq |I_{T_j^{(l)}}|$ for every $T \in \mathbf{T}_j^{(l)}$, we see that for each j

$$\sum_{T \in \mathbf{T}_j^{(l)}} |I_T| \leq C 2^l |I_{T_j^{(l)}}|.$$

From (23), we thus see that the right side of (24) is

$$\leq C 2^{-2l} d^{-r'} \int_E \sum_j 1_{2^l I_{T_j^{(l)}}}(x) \sum_{k : \xi_{k-1}(x) \in \omega_{T_j^{(l)}}} |a_k(x)|^{r'} \, dx.$$

Since each $\sum_{k=1}^K |a_k(x)|^{r'} \leq 1$ and the tiles $2^l I_{T_j^{(l)}} \times \omega_{T_j^{(l)}}$ are pairwise disjoint, the display above is

$$\leq C 2^{-2l} d^{-r'} |E|.$$

Summing over l, we thus obtain (22). □

5. The tree estimate

The following bound allows us to estimate the model operator in the special case where the collection of multitiles is a tree. The bound will be applied in Section 6 with $q = r'$ and $q = 1$.

Proposition 5.1. *Let T be a tree with energy bounded above by e and density bounded above by d. Then, for each $1 \leq q \leq 2$*

$$\| \sum_{P \in T} \langle f, \phi_P \rangle \phi_P a_P 1_E \|_{L^q} \leq C e d^{\min(1, r'/q)} |I_T|^{1/q}. \tag{25}$$

Furthermore, for $l \geq 0$ we have

$$\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(\mathbb{R}\setminus 2^l I_T)} \leq C2^{-l(N-10)} e d^{\min(1,r'/q)} |I_T|^{1/q}. \tag{26}$$

The bounds above also hold for $2 < q < \infty$, but we omit the proof for this range of exponents since it requires an additional L^p estimate for $\sum_{P\in T}\langle f,\phi_P\rangle\phi_P$, and is not required for our purposes.

Proof. Let $\mathbf{J}$ be the collection of dyadic intervals J which are maximal with respect to the property that $I \not\subset 3J$ for every $P \in T$.

Our first goal is to prove

$$\|\sum_{P\in T:|I|\leq C''|J|}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(J)} \leq Ced^{\min(1,r'/q)}|J|^{1/q}(1+\operatorname{dist}(I_T,J)/|I_T|)^{-(N-6)} \tag{27}$$

for each $J \in \mathbf{J}$, where $C'' \geq 1$ is a constant to be determined later. By Hölder's inequality, we may assume that $q \geq r'$. Fix $P \in T$ with $|I| \leq C''|J|$. From the energy bound, we have

$$\|\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(J)} \leq Ce(1+\operatorname{dist}(I,J)/|I|)^{-N}\|a_P 1_E\|_{L^q(J)}. \tag{28}$$

From the density bound applied to $\approx 1/(C_2-1)$ nonempty trees, each with top time interval I, we obtain

$$\frac{1}{|I|}\int_E (1+|x-c(I)|/|I|)^{-4}\sum_{k:\xi_{k-1}(x)\in\omega_l}|a_k(x)|^{r'}\,dx \leq Cd^{r'}.$$

Since $I \not\subset 3J$, it follows that $1+|x-y|/|I| \leq C(1+\operatorname{dist}(I,J)/|I|)$ for every $x \in J$ and $y \in I$. Thus

$$\|a_P 1_E\|^q_{L^q(J)} \leq \|a_P 1_E\|^{r'}_{L^{r'}(J)} \leq C(1+\operatorname{dist}(I,J)/|I|)^4|I|d^{r'}$$

where, above, we use the fact that $|a_P| \leq 1$. Since $|I| \leq C''|J|$ the right side above is

$$\leq C(1+\operatorname{dist}(I,J)/|I|)^4|J|d^{r'}.$$

and so the right side of (28) is

$$\leq Ced^{r'/q}|J|^{1/q}(1+\operatorname{dist}(I,J)/|I|)^{-(N-4)}.$$

Summing this estimate and using the fact that T is a tree, we have

$$\|\sum_{P\in T:|I|=2^{-k}|J|}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(J)} \leq C2^{-k}(1+\operatorname{dist}(I_T,J)/|I_T|)^{-(N-6)}ed^{r'/q}|J|^{1/q}$$

and summing over k gives (27).

Using the maximality of each J, we see that if $l \geq 4$ and $J\cap(\mathbb{R}\setminus 2^l I_T) \neq \emptyset$ then $\operatorname{dist}(I_T,J) \geq |J|/2$ and $|J| \geq 2^{l-3}|I_T|$. It thus follows from (27) that

$$\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(J)}$$
$$\leq C(|I_T|/|J|)(\operatorname{dist}(I_T,J)/|J|)^{-2}ed^{\min(1,r'/q)}|I_T|^{1/q}2^{-l(N-10)}$$

whenever $J\cap(\mathbb{R}\setminus 2^l I_T) \neq \emptyset$. Summing over all J, we thus obtain (26) for $l \geq 4$.

It remains to prove

$$\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(16 I_T)} \leq Ced^{\min(1,r'/q)}|I_T|^{1/q}, \tag{29}$$

and, again, we may assume that $q \geq r'$. The first step will be to demonstrate

$$\int_{J\cap E} \sum_{k:\xi_{k-1}(x)\in\omega_J} |a_k(x)|^{r'}\, dx \leq C d^{r'} |J| \tag{30}$$

where $\omega_J = \bigcup_{P\in T:|I|\geq C''|J|} \omega_l$.

We will say that an l-overlapping tree T is l^--overlapping if for every $P \in T$, $\xi_T \leq \inf \omega_l$. We will say that an l-overlapping tree T is l^+-overlapping if for every $P \in T$, $\xi_T > \inf \omega_l$. For the remainder of the proof, we assume without loss of generality that T is either l^+-overlapping, l^--overlapping, or l-lacunary.

By the maximality of J there is a multitile $P \in T$ with $I \subset 3\tilde{J}$ where $\tilde{J}$ is the dyadic double of J. This implies that there is a dyadic interval J' with $|J| \leq |J'| \leq 4|J|$ and $\mathrm{dist}(J, J') \leq |J|$ and $I \subset J'$.

If T is l^+-overlapping then $T' = (\{P\}, \xi_T, J')$ is a tree. For every $P' \in T$ with $|I'| \geq C''|J|$, we have $\omega_l' \subset [\xi_T - C_1/(2C''|J|), \xi_T + C_1/(2C''|J|))$. Thus, by choosing $C'' \geq 8C_1/(C_2-1)$, we have $\omega_J \subset \omega_{T'}$.

If T is l^--overlapping then $T' = (\{P\}, \xi_T + (C_2-1)/(4|J|), J')$ is a tree. Using the fact that T is l^--overlapping, we see that $\omega_l' \subset [\xi_T, \xi_T + C_2/(2C''|J|))$ for every $P' \in T$ with $|I'| \geq C''|J|$. Thus, by choosing $C'' \geq 4C_2/(C_2-1)$, we have $\omega_J \subset \omega_{T'}$.

If T is l-lacunary then $T' = (\{P\}, \xi_T - (C_2-1)/(4|J|))$ is a tree. Using the fact that T is l-lacunary, we see that $\omega_l' \subset [\xi_T - C_1/(2C''|J|), \xi_T)$ for every $P' \in T$ with $|I'| \geq C''|J|$. Thus, by choosing $C'' \geq 4C_1/(C_2-1)$, we have $\omega_J \subset \omega_{T'}$.

In any of the three cases, the density bound gives

$$\frac{1}{|J'|}\int_E (1+|x-c(J')|/|J'|)^{-4} \sum_{k:\xi_{k-1}(x)\in\omega_{T'}} |a_k(x)|^{r'}\, dx \leq d^{r'}$$

and hence (30).

We now show that if T is l-lacunary then (29) follows from (30). We start by observing that for each x there is at most one integer m and at most one integer k such that there exists a $P \in T$ with $|I| = 2^m$, $\xi_{k-1} \in \omega_l$, and $\xi_k \in \omega_h$. Indeed suppose such a P exists, and $P' \in T$ with $|I'| > |I|$. Since T is l-lacunary, we have $\inf(\omega_l') > \sup(\omega_l)$ by (6), and so $\xi_{k-1} < \inf(\omega_l')$. We also have $\xi_k > \sup(\omega_T) > \sup(\omega_l')$ since $C_2\omega_u \cap \omega_h = \emptyset$ and $\xi_T \notin C_2\omega_l'$. It follows that there is no k' with $\xi_{k'-1} \in \omega_l'$.

We thus have

$$\|\sum_{P\in T:|I|\geq C''|J|} \langle f, \phi_P\rangle \phi_P a_P 1_E\|^q_{L^q(J)}$$

$$\leq \int_{J\cap E} \left(|a(x)| \sum_{P\in T:|I|=2^{m(x)}} |\langle f,\phi_P\rangle \phi_P(x)| \right)^q dx.$$

where, $a(x) = a_k(x)$ if there exists an $m(x)$ as in the previous paragraph with $2^{m(x)} \geq C''|J|$, and $a(x) = 0$ otherwise. From the energy bound and the bound for $|\phi_P|$, the right side above is

$$\leq \int_{J\cap E} \left(|a(x)| \sum_{P\in T:|I|=2^{m(x)}} e(1+|x-c(I)|/|I|)^{-N} \right)^q dx.$$

Noting that $\sum_{P\in T:|I|=2^{m(x)}}(1+|x-c(I)|/|I|)^{-N}\leq C$, we see that the display above is

$$\leq Ce^q\int_{J\cap E}|a(x)|^q\,dx$$

and by our choice of $a(x)$, the display above is

$$\leq Ce^q\int_{J\cap E}\left(\sum_{k:\xi_{k-1}(x)\in\omega_J}|a_k(x)|^{r'}\right)^{q/r'}dx$$

Using (30) and the fact that $\sum|a_k(x)|^{r'}\leq 1$, the display above is

$$\leq Ce^q d^{r'}|J|.$$

Summing over J gives (29).

It remains to consider the case when T is l-overlapping. For each J, we have

$$\|\sum_{P\in T:|I|\geq C''|J|}\langle f,\phi_P\rangle\phi_P a_P 1_E\|^q_{L^q(J)}\leq\int_{J\cap E}\left(\sum_{k:\xi_{k-1}(x)\in\omega_J}|a_k(x)|^{r'}\right)^{q/r'}\times \left(\sum_{k:\xi_{k-1}(x)\in\omega_J}\left|\sum_{P\in T:|I|\geq C''|J|,\xi_{k-1}(x)\in\omega_l,\xi_k(x)\in\omega_h}\langle f,\phi_P\rangle\phi_P(x)\right|^r\right)^{q/r}dx \tag{31}$$

By breaking up T into a bounded number of subtrees, we may assume without loss of generality that for each $P\in T$, $\xi_T\in\omega_l+j|\omega_l|$ for some integer j with $|j|\leq C_2$. We will show that, for any $\xi_{k-1}<\xi_k$, there exist integers $l_1\leq l_2$ with $2^{l_1}\geq|J|$ such that

$$\sum_{P\in T:|I|\geq C''|J|,\xi_{k-1}\in\omega_l,\xi_k\in\omega_h}\langle f,\phi_P\rangle\phi_P=(e^{2\pi i\xi_T\cdot}(\psi_{l_1}-\psi_{l_2}))*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P \tag{32}$$

where $\psi_l=2^{-l}\psi(2^{-l}\cdot)$, and ψ is any Schwartz function with $\hat{\psi}(\xi)=1$ for $|\xi|\leq C_1+C_3$ and $\hat{\psi}(\xi)=0$ for $|\xi|\geq 2C_1$. From (6) we have, for each l such that $2^l=|I|$ for some multitile P,

$$(e^{2\pi i\xi_T\cdot}\psi_l)*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P=\sum_{P\in T:|I|\geq 2^l}\langle f,\phi_P\rangle\phi_P.$$

Thus, to prove (32) it will suffice to show that there exist integers l_1 and l_2 such that

$$\{P\in T:|I|\geq C''|J|,\xi_{k-1}\in\omega_l,\xi_k\in\omega_h\}=\{P\in T:2^{l_1}\leq|I|\leq 2^{l_2}\}. \tag{33}$$

Again using (6), we see that for $P,P'\in T$ with $|I|<|I'|$ we have $\inf\omega_h'<\inf\omega_h$, and if we are in the setting of ρ-multitiles where ρ is a 1-index, we have the stronger inequality $\sup\omega_h'<\inf\omega_h$. Thus, (33) will follow after finding l_1 and l_2 with

$$\{P\in T:\xi_{k-1}\in\omega_l\}=\{P\in T:2^{l_1}\leq|I|\leq 2^{l_2}\}. \tag{34}$$

The equation above follows when $|j|>1$ from the fact that $\omega_l\cap\omega_l'=\emptyset$ if $P,P'\in T$ and $|I|<|I'|$; it follows when $j=0$ from the fact that the intervals $\{\omega_l:P\in T\}$ are nested. Finally, when $j=\pm1$ it follows from the property that if $P,P',P''\in T$, $|I|,\leq|I'|\leq|I''|$ and $\omega_l\cap\omega_l''\neq\emptyset$ then $\omega_l''\subset\omega_l'\subset\omega_l$.

Using (32), we have

$$\left(\sum_{k:\xi_{k-1}(x)\in\omega_J}\left|\sum_{P\in T:|I|\geq C''|J|,\xi_{k-1}(x)\in\omega_l,\xi_k(x)\in\omega_h}\langle f,\phi_P\rangle\phi_P(x)\right|^r\right)^{1/r}$$
$$\leq \|(e^{2\pi i\xi_T\cdot}\psi_k)*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P(x)\|_{V_k^r(\mathbb{Z}^+ + \log_2(|J|))}.$$

For $\log_2(|J|)\leq k_1<k_2$, we have

$$(e^{2\pi i\xi_T\cdot}(\psi_{k_1}-\psi_{k_2}))*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P$$
$$= (e^{2\pi i\xi_T\cdot}\psi_{C+\log_2(|J|)})*(e^{2\pi i\xi_T\cdot}(\psi_{k_1}-\psi_{k_2}))*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P$$

and so, for $x\in J$

$$\|(e^{2\pi i\xi_T\cdot}\psi_k)*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P(x)\|_{V_k^r(\mathbb{Z}^+ + \log_2(|J|))}$$
$$\leq C\sup_{x\in J}\sup_{R\geq|J|}\frac{2}{|R|}\int_{x-R}^{x+R}\|(e^{2\pi i\xi_T\cdot}\psi_k)*\sum_{P\in T}\langle f,\phi_P\rangle\phi_P(y)\|_{V_k^r(\mathbb{Z}^+ + \log_2(|J|))}\,dy.$$

Denoting the right side of the inequality above by M_J, we see that the right side of (31) is

$$\leq M_J^q d^{r'}|J|\leq d^{r'}\int_J\mathcal{M}[\|\psi_k*(e^{-2\pi i\xi_T\cdot}\sum_{P\in T}\langle f,\phi_P\rangle\phi_P)\|_{V_k^r}](x)^q\,dx$$

where $\mathcal{M}$ is the Hardy-Littlewood maximal operator. Summing over J gives

$$\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P a_P 1_E\|_{L^q(16I_T)}^q$$
$$\leq Ce^q d^{r'}|I_T| + d^{r'}\|\mathcal{M}[\|\psi_k*(e^{-2\pi i\xi_T\cdot}\sum_{P\in T}\langle f,\phi_P\rangle\phi_P)\|_{V_k^r}](x)\|_{L_x^q(16I_T)}^q.$$

Since $q\leq 2$, it follows from Hölder's inequality that the right side above is

$$\leq Ce^q d^{r'}|I_T| + Cd^{r'}|I_T|^{(2-q)/2}\|\mathcal{M}[\|\psi_k*(e^{-2\pi i\xi_T\cdot}\sum_{P\in T}\langle f,\phi_P\rangle\phi_P)\|_{V_k^r}](x)\|_{L_x^2(16I_T)}^q.$$

Applying the variation estimate (44) from Appendix A with $p=2$ and the L^2 estimate for $\mathcal{M}$ one sees that the display above is

$$\leq Ce^q d^{r'}|I_T| + Cd^{r'}|I_T|^{(2-q)/2}\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P\|_{L^2}^q.$$

To finish the proof, it only remains to see that $\|\sum_{P\in T}\langle f,\phi_P\rangle\phi_P\|_{L^2}^2\leq Ce^2|I_T|$. The left side of this inequality is

$$\leq\sum_{P\in T}\sum_{P'\in T}|\langle f,\phi_P\rangle||\langle f,\phi'_P\rangle||\langle\phi_P,\phi_{P'}\rangle|$$
$$\leq 2\sum_{P\in T}|\langle f,\phi_P\rangle|^2\sum_{P'\in T}|\langle\phi_P,\phi_{P'}\rangle|.$$

Since T is an l-overlapping tree, we have $\langle \phi_P, \phi_{P'} \rangle$ unless $|I| = |I'|$, in which case, we have $| \langle \phi_P, \phi_{P'} \rangle | \leq C(1 + \text{dist}(I, I')/|I|)^{-N}$. It follows that the right side above is

$$\leq C \sum_{P \in T} | \langle f, \phi_P \rangle |^2 \leq Ce^2 |I_T|.$$

□

6. Main argument

To prove Theorem 1.2, it will suffice by interpolation and monotonicity of the V^r norms to prove the restricted weak type estimate

$$|\{| \sum_{P \in \mathbf{P}} \langle f, \phi_P \rangle \phi_P a_P| > \lambda\}| \leq C \frac{|F|}{\lambda^p}$$

where $\mathbf{P}$ is a finite collection of multitiles as in Section 4, $F \subset \mathbb{R}$, $|f| \leq 1_F$, $\lambda > 0$, $2 < r < \infty$, and $r' \leq p < (1/2 - 1/r)^{-1}$.

This is equivalent to proving that, for every $E \subset \mathbb{R}$,

$$\left| \left\{ x \in E : | \sum_{P \in \mathbf{P}} \langle f, \phi_P \rangle \phi_P a_P| > C \left(\frac{|F|}{|E|} \right)^{1/p} \right\} \right| \leq |E|/2. \tag{35}$$

After possibly rescaling, we assume that $1 \leq |E| \leq 2$. It will suffice, by Chebyshev's inequality to show

$$\|1_E \sum_{P \in \mathbf{P}} \langle f, \phi_P \rangle \phi_P a_P\|_{L^1(\mathbb{R} \setminus G)} \leq C|F|^{1/p} \tag{36}$$

for some exceptional set G with $|G| \leq 1/4$.

The density of $\mathbf{P}$ (which will henceforth be defined with respect to the set E above) is clearly bounded above by a universal constant. Let T be any l-overlapping tree. Writing $f = f' + f''$ where $f' = 1_{3I_T} f$ and $f'' = f - f'$, it follows from arguments in the proof of (8) that

$$\sum_{P \in T} | \langle f', \phi_P \rangle |^2 \leq C\|f'\|_{L^2}^2 \leq C|I_T|.$$

Furthermore, since $|f''| \leq 1_{\mathbb{R} \setminus 3I_T}$, we have the estimate

$$| \langle f'', \phi_P \rangle | \leq C|I|^{1/2}(1 + \text{dist}(I, \mathbb{R} \setminus 3I_T)/|I|)^{-(N-1)} \leq C|I|^{1/2}(|I|/|I_T|)^{N-1}$$

Summing the inequality above, we obtain

$$\sum_{P \in T} | \langle f'', \phi_P \rangle |^2 \leq C|I_T|$$

and so the energy of $\mathbf{P}$ with respect to f is bounded above by a universal constant.

We first consider the case when $|F| > 1$. Repeatedly applying Propositions 4.1 and 4.2 we write $\mathbf{P}$ as the disjoint union

$$\mathbf{P} = \bigcup_{j \geq 0} \bigcup_{T \in \mathbf{T}_j} T$$

where each $\mathbf{T}_j$ is a collection of trees T each of which have energy bounded by $C2^{-j/2}|F|^{1/2}$, density bounded by $C2^{-j/r'}$, and satisfy

$$\sum_{T \in \mathbf{T}_j} |I_T| \leq 2^j.$$

For each j we apply Proposition 4.1 again, this time using (9) and (11) to write

$$\bigcup_{T\in\mathbf{T}_j} T = \bigcup_{k\geq 0}\bigcup_{T\in\mathbf{T}_{j,k}} T$$

where each tree $T\in\mathbf{T}_{j,k}$ has energy bounded by $C2^{-(j+k)/2}|F|^{1/2}$, density bounded by $C2^{-j/r'}$, and satisfies

$$\sum_{T\in\mathbf{T}_{j,k}} |I_T| \leq C2^j. \tag{37}$$

and for every $l\geq 0$

$$\|\sum_{T\in\mathbf{T}_{j,k}} 1_{2^l I_T}\|_{BMO} \leq C2^{2l}2^{j+k}|F|^{-1} \tag{38}$$

From (37), (38), and a standard technique involving the sharp maximal function, it follows that for $1\leq q<\infty$

$$\|\sum_{T\in\mathbf{T}_{j,k}} 1_{2^l I_T}\|_q \leq C2^{j+k+2l}|F|^{-1/q'}.$$

Let $\epsilon>0$ be small and $C'>0$ be large, depending on p,q,r. For each j,k,l define

$$G_{j,k,l} = \{\sum_{T\in\mathbf{T}_{j,k}} 1_{2^l I_T} \geq C'|F|^{-1/q'}2^{(1+\epsilon)(j+k+2l)}\}.$$

By Chebyshev's inequality, we have

$$|G_{j,k,l}| \leq c'2^{-\epsilon(j+k+2l)},$$

so setting $G=\bigcup_{j,k,l\geq 0} G_{j,k,l}$ we have $|G|\leq 1/4$.

Applying Minkowski's inequality gives

$$\begin{aligned}\|1_E&\sum_{P\in\mathbf{P}}\langle f,\phi_P\rangle\,\phi_P a_P\|_{L^1(\mathbb{R}\setminus G)}\\ &\leq \sum_{j,k\geq 0}\Bigg(\|1_E\sum_{T\in\mathbf{T}_{j,k}} 1_{I_T}\sum_{P\in T}\langle f,\phi_P\rangle\,\phi_P a_P\|_{L^1(\mathbb{R}\setminus G_{j,k,0})}\\ &\qquad + \sum_{l\geq 1}\|1_E\sum_{T\in\mathbf{T}_{j,k}} 1_{2^l I_T\setminus 2^{l-1}I_T}\sum_{P\in T}\langle f,\phi_P\rangle\,\phi_P a_P\|_{L^1(\mathbb{R}\setminus G_{j,k,l})}\Bigg).\end{aligned}$$

From Hölder's inequality, Fubini's theorem, and the definition of $G_{j,k,l}$, it follows that the right side above is $\leq C(S_1+S_2)$ where

$$S_1 = \sum_{j,k\geq 0}|F|^{-1/(q'r)}2^{(1+\epsilon)(j+k)/r}\left(\sum_{T\in\mathbf{T}_{j,k}}\|1_E\sum_{P\in T}\langle f,\phi_P\rangle\,\phi_P a_P\|^{r'}_{L^{r'}(\mathbb{R})}\right)^{1/r'}$$

and

$$S_2 = \sum_{\substack{j,k\geq 0\\ l\geq 1}}|F|^{-1/(q'r)}2^{(1+\epsilon)(j+k+2l)/r}\left(\sum_{T\in\mathbf{T}_{j,k}}\|1_E\sum_{P\in T}\langle f,\phi_P\rangle\,\phi_P a_P\|^{r'}_{L^{r'}(\mathbb{R}\setminus 2^{l-1}I_T)}\right)^{1/r'}$$

Applying Proposition 5.1 with the energy and density bounds for trees $T \in \mathbf{T}_{j,k}$, we see that

$$S_2 \leq C \sum_{\substack{j,k\geq 0\\ l\geq 1}} |F|^{-1/(q'r)} 2^{(1+\epsilon)(j+k+2l)/r} 2^{-l(N-10)} 2^{-(j+k)/2} |F|^{1/2} 2^{-j/r'} \left(\sum_{T\in\mathbf{T}_{j,k}} |I_T| \right)^{1/r'}$$
$$\leq C \sum_{\substack{j,k\geq 0\\ l\geq 1}} 2^{(j+k)((1+\epsilon)(2/r)-1)/2} 2^{-l(N-14)} |F|^{1/2-1/(q'r)}$$

Choosing ϵ small enough and q large enough so that $(1+\epsilon)(2/r)-1<0$ and $1/2-1/(q'r)<1/p$ we have $S_2 \leq C|F|^{1/p}$. We similarly obtain $S_1 \leq C|F|^{1/p}$, thus giving (36).

We will finish by proving (36) for $|F|\leq 1$. Here, we let $G = \{\mathcal{M}[1_F] > C''|F|\}$ where $\mathcal{M}$ is the Hardy-Littlewood maximal operator and C'' is chosen large enough so that the weak-type 1-1 estimate for $\mathcal{M}$ guarantees $|G| \leq 1/4$. From the proposition below, which is a special case of an estimate from [8] (we will provide a proof for convenience), and the fact that $p \geq r'$, it will remain to show that

$$\|1_E \sum_{P\in\mathbf{P}'} \langle f,\phi_P\rangle \phi_P a_P\|_{L^1(\mathbb{R}\setminus G)} \leq C|F|^{1/p}. \tag{39}$$

where $\mathbf{P}' = \{P\in\mathbf{P} : I \not\subset G\}$.

Proposition 6.1. *Let* $\mathbf{P}$ *be a finite set of multitiles, and let* $\lambda>0$, $F\subset\mathbb{R}$, *and* $|f|\leq 1_F$. *Then*

$$\|\sum_{P\in\mathbf{P}:I\subset\Omega} \langle f,\phi_P\rangle \phi_P a_P\|_{L^1(\mathbb{R}\setminus\Omega)} \leq C\frac{|F|}{\lambda^{1/r}} \tag{40}$$

where $\Omega = \{\mathcal{M}[1_F] > \lambda\}$.

Finally, it follows from the proposition below, the proof of which may be found on page 12 of [24] or as a special case of a lemma from [8], that the energy of $\mathbf{P}'$ is bounded above by $C|F|$.

Proposition 6.2. *Let* T *be an* l*-overlapping tree. Then*

$$\frac{1}{|I_T|} \sum_{P\in T: I\not\subset\Omega_D} |\langle f,\phi_P\rangle|^2 \leq C\lambda^2$$

where $\Omega_D = \{\mathcal{M}_D[1_F] > \lambda\}$ *and* $\mathcal{M}_D$ *is the maximal dyadic average operator.*

Repeatedly applying Propositions 4.1 and 4.2 we write $\mathbf{P}'$ as the disjoint union

$$\mathbf{P}' = \bigcup_{j\geq 0} \bigcup_{T\in\mathbf{T}_j} T$$

where each $\mathbf{T}_j$ is a collection of trees T each of which have energy bounded by $C2^{-j/2}|F|^{1/2}$, density bounded by $C2^{-j/r'}$, and satisfy

$$\sum_{T\in\mathbf{T}_j} |I_T| \leq 2^j.$$

We then have

$$\|1_E \sum_{P\in\mathbf{P}'} \langle f,\phi_P\rangle \phi_P a_P\|_{L^1} \leq \sum_{j\geq 0}\sum_{T\in\mathbf{T}_j} \|1_E \sum_{P\in T} \langle f,\phi_P\rangle \phi_P a_P\|_{L^1}.$$

Applying Proposition 5.1, we see that the right side above is

$$\leq C\sum_{j\geq 0}\sum_{T\in\mathbf{T}_j}\min(2^{-j/2}|F|^{1/2},|F|)2^{-j/r'}|I_T|\leq C\sum_{j\geq 0}2^{j/r}\min(2^{-j/2}|F|^{1/2},|F|).$$

Summing over j, we see that the right side above is $\leq C|F|^{1/r'}$. This finishes the proof, since $p\geq r'$.

Proof of Proposition 6.1. Fix l and let $I_l\subset\Omega$ be a dyadic interval satisfying

$$2^l I_l\subset\Omega \text{ and } 2^{l+1}I_l\not\subset\Omega. \tag{41}$$

We consider

$$\left(\sum_{P:I=I_l}|\langle f,\phi_P\rangle|^2\right)^{1/2}.$$

Applying Minkowski's inequality, the display above is

$$\leq\left(\sum_{P:I=I_l}|\langle 1_{4I_l}f,\phi_P\rangle|^2\right)^{1/2}+\sum_{j=2}^{\infty}\left(\sum_{P:I=I_l}|\langle 1_{2^{j+1}I_l\setminus 2^jI_l}f,\phi_P\rangle|^2\right)^{1/2}$$

Using orthogonality, the display above is

$$\leq(4|I_l|)^{1/2}\|1_{4I_l}f\phi_{P_0}\|_{L^2}+\sum_{j=2}^{\infty}(2^{j+1}|I_l|)^{1/2}\|1_{2^{j+1}I_l\setminus 2^jI_l}f\phi_{P_0}\|_{L^2}$$

where P_0 is any multitile with $I=I_l$. Applying the bounds (5) and $|f|\leq 1_F$, we see that the display above is

$$\leq C|F\cap 4I_l|^{1/2}+\sum_{j=2}^{\infty}C2^{-j(N-1)}|F\cap 2^{j+1}I_l|^{1/2}.$$

Since $2^{l+1}I_l\not\subset\Omega$, we have $|F\cap 2^{j+1}I_l|\leq C2^{\max(l,j)}|I_l|\lambda$ for each j. Thus, the display above is

$$\leq C(2^l\lambda|I_l|)^{1/2}.$$

Similarly,

$$\sup_{P:I=I_l}|\langle f,\phi_P\rangle|\leq C2^l\lambda|I_l|^{1/2}$$

and so, by interpolation,

$$\left(\sum_{P:I=I_l}|\langle f,\phi_P\rangle|^r\right)^{1/r}\leq(2^l\lambda)^{1/r'}|I_l|^{1/2} \tag{42}$$

whenever $2\leq r\leq\infty$. For each ξ,I_l there is at most one $P\in\mathbf{P}$ with $\xi\in\omega_l$ and $I=I_l$. Thus, using the fact that, for each x, $\sum_{k=1}^{K}|a_k(x)|^{r'}\leq 1$, we see that

$$\|\sum_{P\in\mathbf{P}:I=I_l}\langle f,\phi_P\rangle\,\phi_P a_P\|_{L^1(\mathbb{R}\setminus\Omega)}\leq C(2^l\lambda)^{1/r'}|I_l|^{1/2}\|\phi_{P_0}\|_{L^1(\mathbb{R}\setminus\Omega)}$$

where P_0 is any multitile with $I_0=I_l$. Using the fact that $2^lI_l\subset\Omega$, it follows that the right side above is

$$\leq C2^{-l(N-2)}\lambda^{1/r'}|I_l|.$$

For $l \geq 0$ let $\mathcal{I}_l$ be the set of all dyadic intervals satisfying (41). If $I \subset \mathcal{I}_l$ then for each $j > 0$ there are at most 2 intervals $I' \in \mathcal{I}_l$ with $I' \subset I$ and $|I'| = 2^{-j}|I|$. By considering the collection of maximal dyadic intervals in $\mathcal{I}_l$, one sees that

$$\sum_{I \in \mathcal{I}_l} |I| \leq C|\Omega|$$

Thus,

$$\| \sum_{P \in \mathbf{P} : I \in \mathcal{I}_l} \langle f, \phi_P \rangle \, \phi_P a_P \|_{L^1(\mathbb{R} \setminus \Omega)} \leq C 2^{-l(N-2)} \lambda^{1/r'} |\Omega|.$$

Summing over l and applying the weak-type 1-1 estimate for $\mathcal{M}$ then gives (40). □

A. Variational estimates for averages

The purpose of this appendix is to give the bound (44), which may be considered as a lacunary-"smooth cutoff" version of the main result Theorem 1.2 (a non-smooth version follows from the smooth version by the square function argument in Section 2). Although this estimate seems to be well-known, we provide a proof for the convenience of the reader. We will follow a method from [10], see also the references therein.

For any integer k, we consider the dyadic averaging operator

$$\mathbb{E}_k[f](x) = \frac{1}{|I_k(x)|} \int_{I_k(x)} f(y) \, dy$$

where $I_k(x)$ is the dyadic interval of length 2^k containing x.

It is a special case of Lépingle's inequality [16] (of which alternative proofs, using Doob's jump inequality, may be found for example in [1],[6]) that

$$\|\mathbb{E}_k[f](x)\|_{L^p_x(V^r_k)} \leq C_{p,r} \|f\|_{L^p} \tag{43}$$

whenever $1 < p < \infty$ and $r > 2$, where

$$\|g\|_{V^r} = \sup_{N, k_1 < k_2 < \ldots < k_N} \left(\sum_{j=1}^{N-1} |g(k_{j+1}) - g(k_j)|^r \right)^{1/r}.$$

Let ψ be a Schwartz function on $\mathbb{R}$ with $\int \psi = 1$, and for each k let $\psi_k = 2^{-k}\psi(2^{-k}\cdot)$. Our aim is to see that the bound

$$\|\psi_k * f(x)\|_{L^p_x(V^r_k)} \leq C'_{p,r} \|f\|_{L^p} \tag{44}$$

follows from (43) whenever $1 < p < \infty$ and $r > 2$. Letting

$$\mathbb{S}[f](x) = \left(\sum_{k=-\infty}^{\infty} |\psi_k * f(x) - \mathbb{E}_k[f](x)|^2 \right)^{1/2},$$

it will suffice to show that

$$\|\mathbb{S}[f]\|_{L^p} \leq C_p \|f\|_{L^p} \tag{45}$$

holds whenever $1 < p < \infty$.

We let

$$\mathbb{D}_k[f] = \mathbb{E}_{k-1}[f] - \mathbb{E}_k[f] = \sum_{|I| = 2^k} \langle f, h_I \rangle \, h_I,$$

where, for every dyadic interval I, $h_I = (1_{[\inf(I), c(I))} - 1_{[c(I), \sup(I))})/|I|^{1/2}$ is the L^2 normalized Haar function associated to I.

The case $p = 2$ of (45), which is the case used in the proof of Theorems 1.2 and 1.3, will follow from

Lemma A.1. *Suppose ψ is a Schwartz function with $\int \psi = 1$. Then for every $f \in L^2$*

$$\|\psi_k * \mathbb{D}_j[f] - \mathbb{E}_k[\mathbb{D}_j[f]]\|_{L^2} \leq C 2^{-|j-k|/4} \|\mathbb{D}_j[f]\|_{L^2} \tag{46}$$

where C may depend on ψ.

Indeed

$$\begin{aligned}\|\mathbb{S}[f]\|_{L^2} &= \|\psi_k * f(x) - \mathbb{E}_k[f](x)\|_{l^2_k(L^2_x)} \\ &= \|\psi_k * (\sum_{j=-\infty}^{\infty} \mathbb{D}_j[f])(x) - \mathbb{E}_k(\sum_{j=-\infty}^{\infty} \mathbb{D}_j[f])(x)\|_{l^2_k(L^2_x)}\end{aligned}$$

where the second equation follows from the fact that the Haar functions are a complete orthonormal system in L^2. After applying (46), the right side above is

$$\begin{aligned}&\leq C\|\sum_j 2^{-|j-k|/4} \|\mathbb{D}_j[f]\|_{L^2}\|_{l^2_k} \\ &\leq C\|2^{-|j-k|/8} \mathbb{D}_j[f](x)\|_{l^2_k(l^2_j(L^2_x))} \\ &\leq C\|\mathbb{D}_j[f](x)\|_{l^2_j(L^2(x))} \\ &= \|f\|_{L^2}.\end{aligned}$$

Proof of lemma A.1. First, suppose $k \geq j$. Then, for every $x \in \mathbb{R}$

$$\psi_k * \mathbb{D}_j[f](x) = \sum_{|I|=2^j} \langle f, h_I \rangle \int_I (\psi_k(x-y) - \psi_k(x - c(I)))h_I(y)\, dy.$$

Applying the triangle inequality and mean value theorem, the absolute value of the right side above is

$$\leq \sum_{|I|=2^j} |\langle f, h_I \rangle| 2^{j-k} \int_I \sup_{y \in I} 2^k |\psi_k'(x-y)||I|^{-1/2}\, dy.$$

Since ψ is a Schwartz function, we have $2^k|\psi_k'(x-y)| \leq C2^{-k}(1 + 2^{-k}|x-y|)^{-2}$. Thus, the display above is

$$\begin{aligned}&\leq \sum_{|I|=2^j} |\langle f, h_I \rangle| 2^{j-k} \int_I C 2^{-k}(1 + 2^{-k}|x-y|)^{-2}|I|^{-1/2}\, dy \\ &= C2^{j-k}2^{-k}(1 + 2^{-k}|\cdot|)^{-2} * |\mathbb{D}_j[f]|.\end{aligned}$$

Since $k \geq j$, we have $\mathbb{E}_k[\mathbb{D}_j[f]] = 0$, and we thus obtain (46) from Young's inequality.

For the case $k < j$, we write $\psi_k = \psi_k^{(0)} + \psi_k^{(1)}$ where $\psi_k^{(0)} = \psi_k 1_{[-2^{(j+k-2)/2}, 2^{(j+k-2)/2}]}$. Since ψ is a Schwartz function, $|\psi_k(x)| \leq C2^{-k}(1 + 2^{-k}|x|)^{-2}$ and so

$$\int |\psi_k^{(1)}| \leq C2^{-|j-k|/2}.$$

Since $\int \psi_k = 1$ and $\mathbb{E}_k[\mathbb{D}_j[f]] = \mathbb{D}_j[f]$, we have

$$\psi_k * \mathbb{D}_j[f](x) - \mathbb{E}_k[\mathbb{D}_j[f]](x) = \int \psi_k(y) \sum_{|I|=2^j} \langle f, h_I \rangle (h_I(x-y) - h_I(x))\, dy.$$

Since $h_I(x-y) = h_I(x)$ unless $x-y$ and y are in different dyadic intervals of length 2^j, we have

$$\int \psi_k^{(0)}(y)(\sum_{|I|=2^j} \langle f, h_I\rangle (h_I(x-y) - h_I(x))\ dy$$

supported on $\cup_{m\in\mathbb{Z}}(m2^j - 2^{(j+k-2)/2}, m2^j + 2^{(j+k-2)/2})$. Using an L^1 estimate of $\psi_k^{(0)}$ and, again using its support property, we see that

$$\left| 1_{(m2^j-2^{(j+k-2)/2}, m2^j+2^{(j+k-2)/2})}(x) \int \psi_k^{(0)}(y) \sum_{|I|=2^j} \langle f, h_I\rangle (h_I(x-y) - h_I(x))\ dy \right|$$
$$\leq C(|\left\langle f, h_{[m2^j,(m+1)2^j)}\right\rangle| + |\left\langle f, h_{[(m-1)2^j, m2^j)}\right\rangle|)2^{-j/2}.$$

Thus

$$\| \int \psi_k^{(0)}(y) \sum_{|I|=2^j} \langle f, h_I\rangle (h_I(x-y) - h_I(x))\ dy\|_{L^2(x)} \leq C2^{-|j-k|/4}\|\mathbb{D}_j[f]\|_{L^2}.$$

From the L^1 estimate of $\psi_k^{(1)}$, we have

$$\begin{aligned} &\| \int \psi_k^{(1)}(y) \sum_{|I|=2^j} \langle f, h_I\rangle (h_I(x-y) - h_I(x))\ dy\|_{L^2_x} \\ &\leq \|\psi_k^{(1)} * \mathbb{D}_j[f]\|_{L^2} + \|C2^{-|j-k|/2}\mathbb{D}_j[f]\|_{L^2} \\ &\leq C2^{-|j-k|/2}\|\mathbb{D}_j[f]\|_{L^2} \end{aligned}$$

and thus (46). □

To demonstrate (45) for $1 < p < 2$ (the exponents $p \neq 2$ are used in Section 2), it suffices by interpolation to prove the weak-type $(1,1)$ inequality

$$|\{x : \mathbb{S}[f] > \alpha\}| \leq \frac{C}{\alpha}\|f\|_{L^1}. \tag{47}$$

To obtain this estimate, we perform a dyadic Calderón-Zygmund decomposition of f at height α, that is we write $f = g + b$ where $\|g\|_{L^\infty} \leq \alpha$, $\|g\|_{L^1} \leq C\|f\|_{L^1}$, and

$$b = \sum_{I\in\mathcal{I}} b_I$$

where $\mathcal{I}$ is a collection of disjoint dyadic intervals with $|\cup_{I\in\mathcal{I}} I| \leq C\|f\|_{L^1}/\alpha$, and where each $b_I(x) = 1_I(x)(f(x) - \frac{1}{|I|}\int_I f)$.

The bound for g follows from the L^2 estimate for $\mathbb{S}$

$$|\{x : \mathbb{S}[g] > \alpha/2\}| \leq C\|\mathbb{S}[g]\|_{L^2}^2/\alpha^2 \leq C\|g\|_{L^2}^2/\alpha^2 \leq C\|g\|_{L^1}/\alpha \leq C\|f\|_{L^1}/\alpha.$$

Thus, (47) will follow from the bound

$$\|\mathbb{S}[b]\|_{L^1(\mathbb{R}\setminus\cup_{I\in\mathcal{I}}2I)} \leq C\|b\|_{L^1}. \tag{48}$$

The left side above is

$$\leq \sum_{I\in\mathcal{I}} \sum_{k=-\infty}^{\infty} \|\psi_k * b_I - \mathbb{E}_k[b_I]\|_{L^1(\mathbb{R}\setminus 2I)}.$$

Any dyadic interval intersecting both I and $\mathbb{R} \setminus 2I$ must contain I. Thus, since each b_I is supported on I and has mean zero, the display above

$$(49) \qquad = \sum_{I \in \mathcal{I}} \sum_{k=-\infty}^{\infty} \|\psi_k * b_I\|_{L^1(\mathbb{R}\setminus 2I)}.$$

For $x \in \mathbb{R} \setminus 2I$,

$$\psi_k * b_I(x) = (1_{\mathbb{R}\setminus[-|I|/2,|I|/2]}\psi_k) * b_I(x)$$

Since ψ is a Schwartz function,

$$\|1_{\mathbb{R}\setminus[-|I|/2,|I|/2]}\psi_k\|_{L^1} \leq C2^k/|I|$$

and so

$$(50) \qquad \|\psi_k * b_I\|_{L^1(\mathbb{R}\setminus 2I)} \leq C(2^k/|I|)\|b_I\|_{L^1}.$$

Since each b_I has mean zero and is supported on I, it follows as in the case $k \geq j$ of the proof of lemma A.1 that, whenever $2^k \geq |I|$, we have

$$|\psi_k * b_I(x)| \leq C(|I|/2^k)2^{-k}(1 + 2^{-k}|\cdot|)^{-2} * |b_I|(x)$$

and so

$$(51) \qquad \|\psi_k * b_I\|_{L^1} \leq C(|I|/2^k)\|b_I\|_{L^1}.$$

Combining (50) and (51), it follows that (49) is

$$\leq C \sum_{I \in \mathcal{I}} \|b_I\|_{L^1}.$$

Thus, since the b_I have disjoint supports, we obtain (48).

After minor modifications, the same argument gives a bound from $L^1(\ell^2)$ to weak L^1 for the dual operator, and so (45) also holds for $2 < p < \infty$.

B. A variational Menshov-Paley-Zygmund theorem

For $\xi, x \in \mathbb{R}$ let

$$\mathcal{C}[f](\xi, x) = \int_{-\infty}^{x} e^{-2\pi i \xi x'} f(x')\, dx'.$$

Menshov, Paley, and Zygmund extended the Hausdorff-Young inequality by proving a version of the bound

$$(52) \qquad \|\mathcal{C}[f]\|_{L_\xi^{p'}(L_x^\infty)} \leq C_p \|f\|_{L^p(\mathbb{R})}$$

for $1 \leq p < 2$. The bound at $p = 2$ is a special case of the much more difficult Theorem 1.1 proved by Carleson and Hunt. Interpolating the variational version, Theorem 1.2, at $p = 2$ with a trivial estimate at $p = 1$, one sees that (52) may be strengthened to the bound

$$(53) \qquad \|\mathcal{C}[f]\|_{L_\xi^{p'}(V_x^r)} \leq C_{p,r} \|f\|_{L^p(\mathbb{R})}$$

for $1 \leq p \leq 2$ and $r > p$. It follows from the same arguments given in Section 2 that this range of r is the best possible. Our interest in this variational bound primarily stems from the fact, which will be proven in Appendix C, that it may be transferred, when $r < 2$, to give a corresponding estimate for certain nonlinear Fourier summation operators. The purpose of the present appendix is to give an easier alternate proof of (53) when $p < 2$.

A now-famous lemma of Christ and Kiselev [3] asserts that if an integral operator

$$Tf(x) = \int_{\mathbb{R}} K(x,y)f(y)\,dy$$

is bounded from $L^p(\mathbb{R})$ to $L^q(X)$ for some measure space X and some $q > p$, thus

$$\|Tf\|_{L^q(X)} \leq A\|f\|_{L^p(\mathbb{R})},$$

then automatically the maximal function

$$T_*f(x) = \sup_{N\in\mathbb{R}} |\int_{y<N} K(x,y)f(y)\,dy|$$

is also bounded from $L^p(\mathbb{R})$ to $L^q(X)$, with a slightly larger constant. Another way to phrase this is as follows. If we define the partial integrals

$$T_{\leq}f(x,N) = \int_{y<N} K(x,y)f(y)\,dy$$

then we have

$$\|T_{\leq}f\|_{L^q_x(L^\infty_N)} \leq C_{p,q}A\|f\|_{L^p(\mathbb{R})}. \tag{54}$$

As was observed by Christ and Kiselev, this may be applied in conjunction with the Hausdorff-Young inequality to obtain (52) for $p < 2$.

The L^∞_N norm can also be interpreted as the V^∞_N norm, and we will now see that V^∞ can be replaced by V^r for $r > p$, thus giving (53) from the Hausdorff-Young inequality.

Lemma B.1. *Under the same assumptions, we have*

$$\|T_{\leq}f\|_{L^q_x(V^r_N)} \leq C_{p,q,r}A\|f\|_{L^p(\mathbb{R})}$$

for any $r > p$.

Proof. This follows by an adaption of the argument by Christ and Kiselev, or by the following argument. Without loss of generality we may take $r < q$, in particular $r < \infty$. We use a bootstrap argument. Let us make the *a priori* assumption that

$$\|T_{\leq}f\|_{L^q_x(V^r_N)} \leq BA\|f\|_{L^p(\mathbb{R})} \tag{55}$$

for *some* constant $0 < B < \infty$; this can be accomplished for instance by truncating the kernel K appropriately. We will show that this a priori bound automatically implies the bound

$$\|T_{\leq}f\|_{L^q_x(V^r_N)} \leq (2^{1/r-1/p}BA + C_{p,q,r}A)\|f\|_{L^p(\mathbb{R})} \tag{56}$$

for some $C_{p,q,r} > 0$. This implies that the best bound B in the above inequality will necessarily obey the inequality

$$B \leq 2^{1/r-1/p}B + C_{p,q,r};$$

since $r > p$, this implies $B \leq C'_{p,q,r}$ for some finite $C'_{p,q,r}$, and the claim follows.

It remains to deduce (56) from (55). Fix f; we may normalize $\|f\|_{L^p(\mathbb{R})} = 1$. We find a partition point N_0 in the real line which halves the L^p norm of f:

$$\int_{-\infty}^{N_0} |f(y)|^p\,dy = \int_{N_0}^{+\infty} |f(y)|^p\,dy = \frac{1}{2}.$$

Write $f_-(y) = f(y)1_{(-\infty,N_0]}(y)$ and $f_+(y) = f(y)1_{[N_0,+\infty)}(y)$, thus $\|f_-\|_{L^p(\mathbb{R})} = \|f_+\|_{L^p(\mathbb{R})} = 2^{-1/p}$. We observe that

$$T_\leq f(x,N) = \begin{cases} T_\leq f_-(x,N) & \text{when } N \leq N_0 \\ Tf_-(x) + T^\leq f_+(x,N) & \text{when } N > N_0 \end{cases}$$

Furthermore, $T_\leq f_-(x,\cdot)$ and $T_\leq f_+(x,\cdot)$ are bounded in L^∞ norm by $O(T_* f(x))$. Thus we have

$$\|T_\leq f(x,\cdot)\|_{V^r_N} \leq (\|T_\leq f_-(x,\cdot)\|^r_{V^r_N} + \|T_\leq f_+(x,\cdot)\|^r_{V^r_N})^{1/r} + O(T_* f(x)).$$

(The $O(T_* f(x))$ error comes because the partition used to define $\|T_\leq f(x,\cdot)\|_{V^r_N}$ may have one interval which straddles N_0). We take L^q norms of both sides to obtain

$$\|T_\leq f\|_{L^q_x V^r_N} \leq \|(\|T_\leq f_-(x,\cdot)\|^r_{V^r_N} + \|T_\leq f_-(x,\cdot)\|^r_{V^r_N})^{1/r}\|_{L^q_x} + O(\|T_* f\|_{L^q_x}).$$

The error term is at most $C_{p,q}A$ by the ordinary Christ-Kiselev lemma. For the main term, we take advantage of the fact that $r < q$ to interchange the l^r and L^q norms, thus obtaining

$$\|T_\leq f\|_{L^q_x V^r_N} \leq (\|T_\leq f_-\|^r_{L^q_x V^r_N} + \|T_\leq f_+\|^r_{L^q_x V^r_N})^{1/r} + O(C_{p,q}A).$$

By inductive hypothesis we thus have

$$\|T_\leq f\|_{L^q_x V^r_N} \leq ((2^{-1/p}BA)^r + (2^{-1/p}BA)^r)^{1/r} + O(C_{p,q}A),$$

and the claim follows.

□

C. Variation norms on Lie groups

In this appendix, we will show that certain r-variation norms for curves on Lie groups can be controlled by the corresponding variation norms of their "traces" on the Lie algebra as long as $r < 2$. This follows from work of Terry Lyons [17], we present a self contained proof in this appendix. Combining this fact with the variational Menshov-Paley-Zygmund theorem of Appendix B, we rederive the Christ-Kiselev theorem on the pointwise convergence of the nonlinear Fourier summation operator for $L^p(\mathbb{R})$ functions, $1 \leq p < 2$.

Let G be a connected finite-dimensional Lie group with Lie algebra $\mathfrak{g}$. We give $\mathfrak{g}$ any norm $\|\cdot\|_{\mathfrak{g}}$, and push forward this norm using left multiplication by the Lie group to define a norm $\|x\|_{T_g G} = \|g^{-1}x\|_{\mathfrak{g}}$ on each tangent space $T_g G$ of the group. Observe that this norm structure is preserved under left group multiplication.

We can now define the *length* $|\gamma|$ of a continuously differentiable path $\gamma : [a,b] \to G$ by the usual formula

$$|\gamma| = \int_a^b \|\gamma'(t)\|_{T_{\gamma(t)}G}\, dt.$$

Observe that this notion of length is invariant under left group multiplication, and also under reparameterization of the path γ.

From this notion of length, we can define a metric $d(g,g')$ on G as

$$d(g,g') = \inf_{\gamma:\gamma(a)=g,\gamma(b)=g'} |\gamma|$$

where γ ranges over all differentiable paths from g to g'. It is easy to see that this does indeed give a metric on G.

Given any continuous path $\gamma : [a, b] \to G$ and $1 \leq r < \infty$, we define the *r-variation* $\|\gamma\|_{V^r}$ of γ to be the quantity

$$\|\gamma\|_{V^r} = \sup_{a=t_0<t_1<\ldots<t_n=b} (\sum_{j=0}^{n-1} d(\gamma(t_{j+1}), \gamma(t_j))^r)^{1/r}$$

where the infimum ranges over all partitions of $[a, b]$ by finitely many times $a = t_0, t_1, \ldots, t_n = b$. We can extend this to the $r = \infty$ case in the usual manner as

$$\|\gamma\|_{V^\infty} = \sup_{a=t_0<t_1<\ldots<t_n=b} \sup_{0 \leq j \leq n-1} d(\gamma(t_{j+1}), \gamma(t_j)),$$

and indeed it is clear that the V^∞ norm of γ is simply the diameter of the range of γ. The V^1 norm of γ is finite precisely when γ is rectifiable, and when γ is differentiable it corresponds exactly with the length $|\gamma|$ of γ defined earlier. It is easy to see the monotonicity property

$$\|\gamma\|_{V^p} \leq \|\gamma\|_{V^r} \text{ whenever } 1 \leq r \leq p \leq \infty$$

and the triangle inequalities

$$(\|\gamma_1\|_{V^r}^r + \|\gamma_2\|_{V^r}^r)^{1/r} \leq \|\gamma_1 + \gamma_2\|_{V^r} \leq \|\gamma_1\|_{V^r} + \|\gamma_2\|_{V^r}$$

where $\gamma_1 + \gamma_2$ is the concatenation of γ_1 and γ_2. A key fact about the V^r norms is that they can be subdivided:

Lemma C.1. *Let $\gamma : [a, b] \to G$ be a continuously differentiable curve with finite V^r norm. Then there exists a decomposition $\gamma = \gamma_1 + \gamma_2$ of the curve into two sub-curves such that*

$$\|\gamma_1\|_{V^r}, \|\gamma_2\|_{V^r} \leq 2^{-1/r} \|\gamma\|_{V^r}.$$

Proof. Let $t_* = \sup\{t \in [a, b] : \|\gamma|_{[a,t]}\|_{V^r} \leq 2^{-1/r}\|\gamma\|_{V^r}\}$. Letting $\gamma_1 = \gamma|_{[a,t_*]}$ we have $\|\gamma_1\|_{V^r} = 2^{-1/r}\|\gamma\|_{V^r}$. The bound for $\gamma_2 = \gamma|_{[t_*,b]}$ follows from the left triangle inequality above. □

Given a continuously differentiable curve $\gamma : [a, b] \to G$, we can define its *left trace* $\gamma_l : [a, b] \to \mathfrak{G}$ by the formula

$$\gamma_l(t) = \int_a^t \gamma(s)^{-1}\gamma'(s)\, ds$$

Note that the trace is also a continuously differentiable curve, but taking values now in the Lie algebra $\mathfrak{g}$ instead of G. Clearly γ_l is determined uniquely from γ. The converse is also true after specifying the initial point $\gamma(a)$ of γ, since γ can then be recovered by solving the ordinary differential equation

$$\gamma'(t) = \gamma(t)\gamma_l'(t). \tag{57}$$

This equation is fundamental in the theory of eigenfunctions of a one-dimensional Schrodinger or Dirac operator, or equivalently in the study of the nonlinear Fourier transform; see, for example, [23],[19] for a full discussion. Basically for a fixed potential $f(t)$ and a frequency k, the nonlinear Fourier transform traces out a curve $\gamma(t)$ (depending on k) taking values in a Lie group (e.g. $SU(1, 1)$), and the corresponding left trace is essentially the ordinary linear Fourier transform.

It is easy to see that these curves have the same length (i.e. they have the same V^1 norm):

$$|\gamma| = |\gamma_l|. \tag{58}$$

We now show that something similar is true for the V^r norms provided that $r < 2$.

Lemma C.2. *Let $1 \leq r < 2$, let G be a connected finite-dimensional Lie group, and let $\|\cdot\|_{\mathfrak{g}}$ be a norm on the Lie algebra of G. Then there exist a constant $C > 0$ depending only on these above quantities, such that for all smooth curves $\gamma : [a,b] \to G$, we have*

$$\|\gamma\|_{V^r} \leq \|\gamma_l\|_{V^r} + C\min(\|\gamma_l\|^2_{V^r}, \|\gamma_l\|^r_{V^r}) \tag{59}$$

and

$$\|\gamma_l\|_{V^r} \leq \|\gamma\|_{V^r} + C\min(\|\gamma\|^2_{V^r}, \|\gamma\|^r_{V^r}). \tag{60}$$

An analogous result holds for the *right trace*, $\int_a^t \gamma'(s)\gamma(s)^{-1}\, ds$, once the left-invariant norm on T_gG is replaced by a right-invariant norm.

Proof. We may take $r > 1$ since the claim is already known for $r = 1$ thanks to (58).

It shall suffice to prove the existence of a small $\delta > 0$ such that we have the estimate

$$\|\gamma\|_{V^r} = \|\gamma_l\|_{V^r} + O(\|\gamma_l\|^2_{V^r}) \tag{61}$$

whenever $\|\gamma_l\|_{V^r} \leq \delta$, and similarly

$$\|\gamma_l\|_{V^r} = \|\gamma\|_{V^r} + O(\|\gamma\|^2_{V^r}) \tag{62}$$

whenever $\|\gamma\|_{V^r} \leq \delta$. (We allow the $O()$ constants here to depend on r, the Lie group G, and the norm structure, but not on δ). Let us now see why these estimates will prove the lemma. Let us begin by showing that (61) implies (59). Certainly this will be the case if γ_l has V^r norm less than δ. If instead γ_l has V^r norm larger than δ, we can use Lemma C.1 repeatedly to partition it into $O(\delta^{-r}\|\gamma_l\|^r_{V^r})$ curves, all of whose V^r norms are less than δ. These curves are the left-traces of various components of γ, and thus by (61) these components have a V^r norm bounded by some quantity depending on δ. Concatenating these components together (using the triangle inequality) we obtain the result. A similar argument allows one to deduce (60) from (62).

Next, we observe that to prove the two estimates (61), (62) it suffices to just prove one of the two, for instance (61), as this will also imply (62) for $\|\gamma\|_{V^r}$ sufficiently small by the usual continuity argument (look at the set of times t for which the restriction of γ to $[a,b]$ obeys a suitable version of (62), and use (61) to show that this set is both open and closed if $\|\gamma\|_{V^r}$ is small enough).

It remains to prove (61) for δ sufficiently small. We shall in fact prove the more precise statement

$$\|\log(\gamma(a)^{-1}\gamma(b)) - \gamma_l(b)\|_{\mathfrak{g}} \leq K\|\gamma_l\|^2_{V^r} \tag{63}$$

for some absolute constant $K > 0$ (and for δ sufficiently small), where log is the inverse of the exponential map $\exp : \mathfrak{g} \to G$. Note that it follows from a continuity argument as in the previous paragraph that if δ is sufficiently small then $\gamma(b)^{-1}\gamma(a)$ is sufficiently close to the identity that the logarithm is well-defined. Let us now see why (63) implies (61). Applying the inequality to any segment $[t_j, t_{j+1}]$ in $[a,b]$ we see that

$$\|\log(\gamma(t_j)^{-1}\gamma(t_{j+1})) - (\gamma_l(t_{j+1}) - \gamma_l(t_j))\|_{\mathfrak{g}} \leq K\|\gamma_l|_{[t_j,t_{j+1}]}\|^2_{V^r}$$

and hence (since δ is small)

$$d(\gamma(t_{j+1}),\gamma(t_j)) = \|\gamma_l(t_{j+1}) - \gamma_l(t_j)\|_{\mathfrak{g}} + O(\|\gamma_l|_{[t_j,t_{j+1}]}\|_{V^r}^2).$$

Estimating $O(\|\gamma_l|_{[t_j,t_{j+1}]}\|_{V^r}^2)$ crudely by $\|\gamma_l\|_{V^r} O(\|\gamma_l|_{[t_j,t_{j+1}]}\|_{V^r})$ and taking the l^r sum in the j index, we see that for any partition $a = t_0 < \ldots < t_n = b$ we have

$$(\sum_{j=0}^{n-1} d(\gamma(t_{j+1}),\gamma(t_j))^r)^{1/r} = (\sum_{j=0}^{n-1} \|\gamma_l(t_{j+1}) - \gamma_l(t_j)\|_{\mathfrak{g}}^r)^{1/r} + O(\|\gamma_l\|_{V^r}^2).$$

Taking suprema over all partitions we obtain the result.

It remains to prove (63) for some suitably large K. This we shall do by an induction on scale (or "Bellman function") argument. Let us fix the smooth curve γ. We shall prove the estimate for all subcurves of γ, i.e. for all intervals $[t_1,t_2]$ in $[a,b]$, we shall prove that

$$\|\log(\gamma(t_1)^{-1}\gamma(t_2)) - (\gamma_l(t_2) - \gamma_l(t_1))\|_{\mathfrak{g}} \leq K\|\gamma_l|_{[t_1,t_2]}\|_{V^r}^2. \tag{64}$$

Let us first prove this in the case when the interval $[t_1,t_2]$ is sufficiently short, say of length at most ϵ for some very small ϵ (depending on γ). In that case, we perform a Taylor expansion to obtain

$$\gamma_l(t) = \gamma_l(t_1) + \gamma_l'(t_1)(t-t_1) + \frac{1}{2}\gamma_l''(t_1)(t-t_1)^2 + O_\gamma((t-t_1)^3) \tag{65}$$

and

$$\gamma_l'(t) = \gamma_l'(t_1) + \gamma_l''(t_1)(t-t_1) + O_\gamma((t-t_1)^2) \tag{66}$$

when $t \in [t_1,t_2]$, and where the γ subscript in O_γ means that the constants here are allowed to depend on γ (more specifically, on the C^3 norm of γ), and the $O()$ is with respect to the $\|\|_{\mathfrak{g}}$ norm. Also we remark that as γ is assumed smooth, $\gamma_l'(t_1)$ is bounded away from zero. It is then an easy matter to conclude that

$$\|\gamma_l|_{[t_1,t_2]}\|_{V^r} \geq \frac{1}{2}\|\gamma_l'(t_1)\|_{\mathfrak{g}}|t_2 - t_1| \tag{67}$$

if ϵ is sufficiently small depending on γ. On the other hand, from (57) and (66) we have

$$\gamma'(t) = \gamma(t)(\gamma_l'(t_1) + \gamma_l''(t_1)(t-t_1) + O_\gamma((t-t_1)^2))$$

from which one may conclude that

$$\gamma(t) = \gamma(t_1)\exp(\gamma_l'(t_1)(t-t_1) + \frac{1}{2}\gamma_l''(t_1)(t-t_1)^2 + O(\|\gamma_l'(t_1)^2\|_{\mathfrak{g}}|t-t_1|^2) + O_\gamma((t-t_1)^3))$$

for all $t \in [t_1,t_2]$, if γ is sufficiently small. We rewrite this as

$$\log(\gamma(t_1)^{-1}\gamma(t))$$
$$= \gamma_l'(t_1)(t-t_1) + \frac{1}{2}\gamma_l''(t_1)(t-t_1)^2 + O(\|\gamma_l'(t_1)^2\|_{\mathfrak{g}}|t-t_1|^2) + O_\gamma((t-t_1)^3),$$

and then specialize to the case $t = t_2$. By (65), we have

$$\log(\gamma(t_1)^{-1}\gamma(t_2)) - (\gamma_l(t_2) - \gamma_l(t_1)) = O(\|\gamma_l'(t_1)^2\|_{\mathfrak{g}}|t_2 - t_1|^2) + O_\gamma((t_2-t_1)^3),$$

and hence by (67) we have (64) if $t_2 - t_1$ is small enough (depending on γ) and K is large enough (*independent* of γ).

This proves (64) when the interval $[t_1,t_2]$ is small enough. By (67), it also proves (64) when $\|\gamma_l|_{[t_1,t_2]}\|_{V^r}$ is sufficiently small. To conclude the proof of (64) in general, we now assert the following inductive claim: if (64) holds whenever $\|\gamma_l|_{[t_1,t_2]}\|_{V^r} < \epsilon$ and some given $0 < \epsilon \leq \delta$, then it also holds whenever $\|\gamma_l|_{[t_1,t_2]}\|_{V^r} < 2^{1/r}\epsilon$, providing that K is sufficiently large (*independent* of ϵ) and δ is sufficiently small (depending on K, but *independent* of ϵ). Iterating this we will obtain the claim (64) for all intervals $[t_1, t_2]$ in $[a, b]$.

It remains to prove the inductive claim. Let $[t_1, t_2]$ be any subinterval of $[a, b]$ such that the quantity $A = \|\gamma_l|_{[t_1,t_2]}\|_{V^r}$ is less than $2^{1/r}\epsilon$. Applying Lemma C.1, we may subdivide $[t_1, t_2] = [t_1, t_*] \cup [t_*, t_2]$ such that

$$\|\gamma_l|_{[t_1,t_*]}\|_{V^r}, \|\gamma_l|_{[t_*,t_2]}\|_{V^r} \le 2^{-1/r} A < \epsilon \le r.$$

By the inductive hypothesis, we thus have

$$\|\log(\gamma(t_1)^{-1}\gamma(t_*)) - (\gamma_l(t_*) - \gamma_l(t_1))\|_{\mathfrak{g}} \le K 2^{-2/r} A^2$$

and

$$\|\log(\gamma(t_*)^{-1}\gamma(t_2)) - (\gamma_l(t_2) - \gamma_l(t_*))\|_{\mathfrak{g}} \le K 2^{-2/r} A^2.$$

In particular, we have

$$\begin{aligned}
\|\log(\gamma(t_1)^{-1}\gamma(t_*)\|_{\mathfrak{g}} &\le \|\gamma_l(t_*) - \gamma_l(t_1)\|_{\mathfrak{g}} + K 2^{-2/r} A^2 \\
&\le \|\gamma_l|_{[t_1,t_*]}\|_{V^r} + O(KA^2) \\
&= O(A(1 + KA)) \\
&= O(A(1 + K\delta)) \\
&= O(A)
\end{aligned}$$

if δ is sufficiently small depending on K. Similarly we have

$$\|\log(\gamma(t_*)^{-1}\gamma(t_2)\|_{\mathfrak{g}} = O(A)$$

and hence by the Baker-Campbell-Hausdorff formula (if δ is sufficiently small)

$$\|\log(\gamma(t_1)^{-1}\gamma(t_2) - \log(\gamma(t_1)^{-1}\gamma(t_*) - \log(\gamma(t_*)^{-1}\gamma(t_2)\|_{\mathfrak{g}} = O(A^2).$$

By the triangle inequality, we thus have

$$\|\log(\gamma(t_1)^{-1}\gamma(t_2)) - (\gamma_l(t_2) - \gamma_l(t_1))\|_{\mathfrak{g}} \le 2K 2^{-2/r} A^2 + O(A^2).$$

We now use the hypothesis $r < 2$, which forces $2 \times 2^{-2/r} < 1$. If K is large enough (depending on r, but independently of δ, A, or ϵ) we thus have (64). This closes the inductive argument. □

Letting w, v be any elements of the Lie algebra $\mathfrak{g}$, one can define a nonlinear Fourier summation operator associated to G, w, v by means of the left trace

$$\begin{aligned}
&\mathcal{NC}[f](k, 0) = I \\
&\frac{\partial}{\partial x}\mathcal{NC}[f](k, x) = \mathcal{NC}[f](k, x)\left(\mathrm{Re}(e^{-2\pi ikx} f(x))w + \mathrm{Im}(e^{-2\pi ikx} f(x))v\right)
\end{aligned}$$

or (giving a different operator) by the right trace

$$\begin{aligned}
&\mathcal{NC}[f](k, 0) = I \\
&\frac{\partial}{\partial x}\mathcal{NC}[f](k, x) = \left(\mathrm{Re}(e^{-2\pi ikx} f(x))w + \mathrm{Im}(e^{-2\pi ikx} f(x))v\right)\mathcal{NC}[f](k, x).
\end{aligned}$$

Above, $k, x \in \mathbb{R}$, $\mathcal{NC}[f]$ takes values in G, I is the identity element of G, and Re, Im are the real and imaginary parts of a complex number. An example of interest is given by $G = SU(1, 1)$,

$$w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and

$$v = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Combining Lemma C.2 with the variational Menshov-Paley-Zygmund theorem of the previous section, we obtain a variational version of the Christ-Kiselev theorem [4]. Namely, we see that for $1 \le p < 2$ and $r > p$

$$\|1_{|\mathcal{NC}[f]| \le 1} \mathcal{NC}[f]\|_{L_k^{p'}(V_x^r)} \le C_{p,r,G,w,v} \|f\|_{L^p(\mathbb{R})}$$

and

$$\|1_{|\mathcal{NC}[f]| \ge 1} \mathcal{NC}[f]\|^{1/r}_{L_k^{p'/r}(V_x^r)} \le C_{p,r,G,w,v} \|f\|_{L^p(\mathbb{R})}.$$

Note that the usual logarithms are hidden in the d metric we have placed on the Lie group G.

Extending these estimates to the case $p = 2$ is an interesting and challenging problem, even when $r = \infty$, which would corresponds to a nonlinear Carleson theorem. Lemma C.2 cannot be extended to any exponent $r \ge 2$. Sandy Davie and the fifth author of this paper have an unpublished example of a curve in the Lie group $SU(1,1)$ with trace in the subspace of $\mathfrak{su}(1,1)$ of matrices vanishing on the diagonal so that the 2-variation of the curve is not controlled by the 2-variation of the trace.

Terry Lyons' machinery [18] via iterated integrals faces an obstruction in a potential application to a nonlinear Carleson theorem becasue of the unboundedness results for the iterated integrals shown in [20].

D. An application to ergodic theory

Wiener-Wintner type theorems is an area in ergodic theory that is most closely related to the study of Carleson's operator. In [14], Lacey and Terwilleger prove the following singular integral variant of the Wiener-Wintner theorem:

Theorem D.1. *For $1 < p$, all measure preserving flows $\{T_t : t \in \mathbb{R}\}$ on a probability space (X, μ) and functions $f \in L^p(\mu)$, there is a set $X_f \subset X$ of probability one, so that for all $x \in X_f$ we have that the limit*

$$\lim_{s \to 0} \int_{s < |t| < 1/s} e^{i\theta t} f(T_t x) \frac{dt}{t} \quad .$$

exists for all $\theta \in \mathbb{R}$.

One idea to approach such convergence results is to study quantitative estimates in the parameter s that imply convergence, as pioneered by Bourgain's paper [1] in similar context. We first need to pass to a mollified variant of the above theorem:

Theorem D.2. *Let ϕ be a function on $\mathbb{R}$ in the Wiener space, i.e. the Fourier transform $\widehat{\phi}$ is in $L^1(\mathbb{R})$. For $1 < p$, all measure preserving flows $\{T_t : t \in \mathbb{R}\}$ on a probability space (X, μ) and functions $f \in L^p(\mu)$, there is a set $X_f \subset X$ of probability one, so that for all $x \in X_f$ we have that the limits*

$$\lim_{s \to \infty} \int e^{i\theta t} f(T_t x) \phi(st) \frac{dt}{t} \quad ,$$

$$\lim_{s \to 0} \int e^{i\theta t} f(T_t x) \phi(st) \frac{dt}{t} \quad .$$

exist for all $\theta \in \mathbb{R}$.

This theorem clearly follows from an a priori estimate

$$\left\| \sup_\theta \left\| \int e^{i\theta t} f(T_t x)\phi(st)\frac{dt}{t} \right\|_{V^r(s)} \right\|_{L^p(x)} \le C\|f\|_p$$

for $r > \max(2, p')$. Here we have written $V^r(s)$ for the variation norm taken in the parameter s of the expression inside, and likewise for $L^p(x)$. The variation norm is the strongest norm widely used in this context, while Lacey and Terwilliger use a weaker oscillation norm in the proof of their Theorem.

By a standard transfer method, involving replacing f by translates $T_y f$ and an averaging procedure in y, the a priori estimate can be deduced from an analogous estimate on the real line

$$\| \sup_\xi \| \int e^{\xi i t} f(x+t)\phi(st)\frac{dt}{t} \|_{V^r(s)} \|_{L^p(x)} \le C\|f\|_p \quad . \tag{68}$$

The main purpose of this appendix is to show how this estimate (68) can be decuded from the main theorem of this paper by an averaging argument. We write the $V^r(s)$ norm explicitly and expand ϕ into a Fourier integral to obtain for the left hand side of (68) the expression

$$\| \sup_\xi \sup_{s_0 < s_1 < \dots < s_K} \left(\sum_{k=1}^K \left| \int \int e^{\xi i t} f(x+t) e^{i\eta(s_k - s_{k-1})t} \frac{dt}{t} \widehat{\phi}(\eta)\, d\eta \right|^r \right)^{1/r} \|_{L^p(x)} \quad .$$

Now pulling the integral in η out of the various norms and considering only positive η (with the case of negative η being similar) and defining $\xi_k = \xi + \eta s_k$ we obtain the upper bound

$$\int_{\eta > 0} \| \sup_{\xi_0 < \xi_1 < \dots < \xi_K} \left(\sum_{k=1}^K \left| \int \int e^{i(\xi_k - \xi_{k-1})t} f(x+t) \frac{dt}{t} \right|^r \right)^{1/r} \|_{L^p(x)} |\widehat{\phi}(\eta)|\, d\eta \quad .$$

Now applying the variational Carleson estimate and doing the trivial integral in η bounds this term by a constant times $\|f\|_p$.

We conclude this appendix with two remarks.

1) To prove the Lacey Terwilleger theorem D.1 from the mollified version, one may approximate the characteristic functions used as cutoff functions by Wiener space functions so that the difference is small in L^1 norm. Then at least for f in L^∞ one can show convergence of the limits by an approximation argument, even though one will not recover the full strength of the quantitative estimate in the Wiener space setting. The result for f in L^∞ can then be used as a dense subclass result in other L^p spaces, which can be handled by easier maximal function estimates and further approximation arguments.

2) The classical version of the Wiener-Wintner theorem does not invoke singular integrals but more classical averages of the type

$$\frac{1}{2s} \int_{|t|<s} e^{i\theta t} f(T_t x)\, dt \quad .$$

We note that the same technique as above may be applied to these easier averages.

References

[1] Jean Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. (1989), no. 69, 5–45, With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.

[2] Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.

[3] Michael Christ and Alexander Kiselev, *Maximal functions associated to filtrations*, J. Funct. Anal. **179** (2001), no. 2, 409–425.

[4] ———, *WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials*, J. Funct. Anal. **179** (2001), no. 2, 426–447.

[5] Ciprian Demeter, Michael T. Lacey, Terence Tao, and Christoph Thiele, *Breaking the duality in the return times theorem*, Duke Math. J. **143** (2008), no. 2, 281–355.

[6] ———, *The Walsh model for M_2^* Carleson*, Rev. Mat. Iberoamericana **24** (2008), no. 3, 721–744.

[7] Charles Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. (2) **98** (1973), 551–571.

[8] Loukas Grafakos, Terence Tao, and Erin Terwilleger, *L^p bounds for a maximal dyadic sum operator*, Math. Z. **246** (2004), no. 1-2, 321–337.

[9] Richard A. Hunt, *On the convergence of Fourier series*, Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255.

[10] Roger L. Jones, Andreas Seeger, and James Wright, *Strong variational and jump inequalities in harmonic analysis*, Trans. Amer. Math. Soc. **360** (2008), no. 12, 6711–6742.

[11] Roger L. Jones and Gang Wang, *Variation inequalities for the Fejér and Poisson kernels*, Trans. Amer. Math. Soc. **356** (2004), no. 11, 4493–4518 (electronic).

[12] Carlos E. Kenig and Peter A. Tomas, *Maximal operators defined by Fourier multipliers*, Studia Math. **68** (1980), no. 1, 79–83.

[13] Michael Lacey, *Issues related to Rubio de Francia's Littlewood-Paley inequality*, NYJM Monographs, **2**. State University of New York, University at Albany, Albany, NY, (2007) 36 pp. (electronic).

[14] Michael Lacey and Erin Terwilleger, *A Wiener-Wintner theorem for the Hilbert transform*, Ark. Mat. **46** (2008), no. 2, 315–336.

[15] Michael Lacey and Christoph Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370.

[16] Dominique Lépingle, *La variation d'ordre p des semi-martingales*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 4, 295–316.

[17] Terry Lyons, *Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young*, Math. Res. Lett. **1** (1994), no. 4, 451–464.

[18] Terry Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), 215–310.

[19] Camil Muscalu, Terence Tao, and Christoph Thiele, *A Carleson theorem for a Cantor group model of the scattering transform*, Nonlinearity **16** (2003), no. 1, 219–246.

[20] Camil Muscalu, Terence Tao, and Christoph Thiele, *A counterexample to a multilinear endpoint question of Christ and Kiselev*, Math. Res. Lett. **10** (2003) no. 2-3, 237–246.

[21] Jinghua Qian, *The p-variation of partial sum processes and the empirical process*, Ann. Probab. **26** (1998), no. 3, 1370–1383.

[22] J.L. Rubio de Francia, *A Littlewood-Paley inequality for arbitrary intervals*, Revista Mat. Iberoamericana **1** (1985), no. 2, 1-14.

[23] Terence Tao and Christoph Thiele, *Nonlinear Fourier analysis*, IAS/Park City Math. Ser., to appear.

[24] Christoph Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006.

[25] Norbert Wiener, *The quadratic variation of a function and its Fourier coefficients.* MIT Journal of Math. and Physics **3**(1924) pp. 72–94.