Bounds on the Walsh model for $M^{q,*}$ Carleson and related operators

Richard Oberlin

Abstract. We prove an extension of the Walsh-analog of the Carleson–Hunt theorem, where the L^∞ norm defining the Carleson maximal operator has been replaced by an L^q maximal-multiplier-norm. Additionally, we consider certain associated variation-norm estimates.

1. Introduction

Given a real-valued function f on $\mathbb{R}^+$ consider the partial Walsh-sum operator, defined for $\xi, x \in \mathbb{R}^+$ by

$$(1.1) \qquad \mathcal{S}[f](\xi, x) = (\hat{f} \, 1_{[0,\xi)})^{\check{}}(x).$$

where $\hat{}$ and $\check{}$ refer to the Walsh–Fourier transform (terminology and notation will be explained in detail in Section 2 and in a paragraph at the end of this section). The operator above can also be written using a wave-packet decomposition:

$$\mathcal{S}[f](\xi, x) = \sum_{P} \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) \, 1_{\omega_{P_u}}(\xi),$$

where we sum over all bitiles P and ϕ_{P_l} is the L^2 normalized wave-packet corresponding to the lower half of P. Additionally, we will need a truncated version of $\mathcal{S}$, defined for each integer k by

$$\mathcal{S}_k[f](\xi, x) = \sum_{P : |I_P| < 2^k} \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) \, 1_{\omega_{P_u}}(\xi).$$

It is well known that the Walsh-analog of the Carleson–Hunt theorem holds, namely that for $1 < p < \infty$,

$$(1.2) \qquad \|\mathcal{S}[f](\xi, x)\|_{L^p_x(L^\infty_\xi)} \leq C_p \, \|f\|_{L^p}.$$

We are interested in versions of the bound above where L^∞ is replaced by various stronger norms.

Given a function m on $\mathbb{R}^+$ and an exponent $1 \le q \le \infty$ consider the Walsh L^q-multiplier norm of m

$$\|m\|_{M^q} = \sup_{g:\|g\|_{L^q}=1} \|(m\hat{g})\check{}\|_{L^q}.$$

Replacing m by a sequence of functions $\{m_k\}_{k\in\mathbb{Z}}$, one can also define a Walsh L^q-maximal-multiplier norm

$$\|m\|_{M^{q,*}} = \sup_{g:\|g\|_{L^q}=1} \|(m_k\hat{g})\check{}(x)\|_{L_x^q(\ell_k^\infty)}.$$

The Walsh-$M^{2,*}$-Carleson theorem was proven by Demeter, Lacey, Tao and Thiele (see [6]). Specifically, they showed that if $1 < p < \infty$ then

(1.3) $$\|\mathcal{S}_k[f](\xi,x)\|_{L_x^p(M_{\xi,k}^{2,*})} \le C_p \|f\|_{L^p}.$$

The main result of this paper is to extend the theorem above to cover[1] exponents $q < 2$. Namely, we will prove:

Theorem 1.1. *Suppose that $1 < p < \infty$ and $1 < q < 2$ satisfy $\frac{1}{p} + \frac{1}{q} < \frac{3}{2}$. Then*

$$\|\mathcal{S}_k[f](\xi,x)\|_{L_x^p(M_{\xi,k}^{q,*})} \le C_{p,q} \|f\|_{L^p}.$$

Since the M^2 norm is equal to the L^∞ norm, one sees that the difference between (1.2) and (1.3) is that the M^2 norm of the former bound is replaced by an $M^{2,*}$ norm in the latter bound, where the $*$ refers to a maximum over truncations. Thus, when approaching Theorem 1.1, one might first ask whether the corresponding bound holds with the M^q norm in place of the $M^{q,*}$ norm. As we will now see, the affirmative answer to this question follows from combining work in [13], which preceded a result for the Fourier-transform [12], with the Walsh-analog of [3].

Given an exponent r and a function m defined on a subset of $\mathbb{R}$ and taking values in some normed linear space (in this paper, the subset of $\mathbb{R}$ will be $\mathbb{R}^+$ or $\mathbb{Z}$, and except for part of Section 9 the normed linear space will be $\mathbb{R}$) consider the r-variation norm

$$\|m\|_{V^r} = \|m\|_{L^\infty} + \sup_{N,\xi_0<\cdots<\xi_N} \Big(\sum_{i=1}^N |m(\xi_i) - m(\xi_{i-1})|^r \Big)^{1/r},$$

where the supremum is over all strictly increasing finite-length sequences in the domain of m. It was proven in [13] (and we will give another proof here, see Section 8) that if $r > 2$ and $p > r'$ then

(1.4) $$\|\mathcal{S}[f](\xi,x)\|_{L_x^p(V_\xi^r)} \le C_{p,r} \|f\|_{L^p}.$$

-----

[1]It seems likely that the range $q > 2$ is also tractable, see comments in [5]. However that case is not of particular interest for applications related to the return times theorem due to the monotonicity of L^q norms in probability spaces.

Applying the method of Coifman, Rubio de Francia, and Semmes [3] to Walsh multipliers, one sees (as in Lemma 3.5 below) that if $r \geq 2$ and $|1/q - 1/2| < 1/r$, then for functions m on $\mathbb{R}^+$,

$$(1.5) \qquad \|m\|_{M^q} \leq C_{q,r} \, \|m\|_{V^r}.$$

Hence, it follows from (1.4) that when $1 < p < \infty$ and $1 < q < \infty$ satisfy $1/p + 1/q < 3/2$,

$$(1.6) \qquad \|\mathcal{S}[f](\xi, x)\|_{L^p_x(M^q_\xi)} \leq C_{p,q} \, \|f\|_{L^p}.$$

It is thus clear that, as in [6], the task at hand is to replace the M^q norm in (1.6) with the $M^{q,*}$ norm. Roughly speaking, in [6] this advance was obtained by incorporating the use of Lemma 3.1 below into a proof of (1.2) (this statement slurs over many technical obstructions; in fact their method required the development of a substantially new proof of (1.2)). We will follow the same approach, but with some necessary refinements which we now detail.

First, we replace (1.5) and the natural L^q extension of Lemma 3.1 with a common extension of the two bounds which is more efficient than their separate applications. We develop a new proof of (1.4) which (as in the proof of (1.2) from [6]) gives pointwise control for the sum over bitiles in a stack of trees in terms of the restrictions of the sum to individual trees. Obtaining this pointwise control for variation, rather than L^∞, norms requires a more careful decomposition into l-overlapping trees, in particular the use of a concept of "l-convexity". This decomposition allows us to obtain an explicit partitioning of $\mathbb{R}^+$ into intervals, on which the sum over a stack of trees agrees with its restriction to an individual tree. Finally, to control the variation-norm for an individual l-overlapping tree we use phase-space projections, as in [5].

1.1. Motivation

A significant part of our interest in Theorem 1.1 is due to its role as a model case for the corresponding Fourier-transform problem. Let Ψ be (say) a Schwartz function on $\mathbb{R}$ and for f defined on $\mathbb{R}$, $\xi, x \in \mathbb{R}$, and $k \in \mathbb{Z}$ consider the truncated partial Fourier sum operator

$$\mathfrak{S}_k[f](\xi, x) = p.v. \int f(x - t) \, e^{2\pi i \xi t} \, \Psi(2^{-k} t) \, \frac{1}{t} \, dt.$$

It was proven in [7] that for $1 < p < \infty$,

$$(1.7) \qquad \|\mathfrak{S}_k[f](\xi, x)\|_{L^p_x(M^{2,*}_{\xi,k})} \leq C_p \, \|f\|_{L^p},$$

and it would be desirable to extend this result, as we have now done for the Walsh model, to cover exponents $q < 2$. One reason for interest in bounds such as (1.7) is their application to the return times problem for the truncated Hilbert transform. It was shown in [7] that bounds similar to (1.7) can be used to deduce that given

a measure preserving system (X, T) and a function $f \in L^p(X)$, $p > 1$, one can obtain a set X' of full measure in X such that for every $x \in X'$, every second measure preserving system (Y, U), and every $g \in L^q(Y)$, $q \geq 2$, the sums

$$(1.8) \qquad \sum_{n=-N}^{N}{}' \frac{1}{n} f(T^n x) \, g(U^n y)$$

converge as $N \to \infty$ for almost every $y \in Y$. An extension of the range of exponents in (1.7) could be used to extend the range of exponents for the pointwise convergence result.

Theorems similar to the convergence result above were originally considered [1] in the context of averages[2], where one is interested in sums

$$(1.9) \qquad \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \, g(U^n y).$$

Here, the relevant analog of $\mathfrak{S}_k$ is

$$\mathfrak{A}_k[f](\xi, x) = \int f(x - t) \, e^{2\pi i \xi t} \, 2^{-k} \, \Psi(2^{-k} t) \, dt.$$

In [7] it was shown that (1.7) holds with $\mathfrak{A}_k$ in place of $\mathfrak{S}_k$ and this bound has been extended ([4], [10]) to cover the range $1/p + 1/q < 3/2$. The Walsh-analog of $\mathfrak{A}_k$ would be

$$\mathcal{A}_k[f](\xi, x) = \sum_{p : |I_p| = 2^k} \langle f, \phi_p \rangle \, \phi_p(x) \, 1_{\omega_p}(\xi),$$

where above we sum over all tiles p. It seems likely that bounds for $\mathcal{A}_k$ could be obtained using a Walsh-analog of the method [4].

1.2. Further results

Using a method from [10], see Section 9, one can obtain a variation-norm version of Lemma 3.1. Substituting this lemma into the proof of Theorem 1.1 gives the stronger

Theorem 1.2. *Suppose that $s > 2$, $1 < p < \infty$. and $1 < q < 2$ satisfy $1/p + 1/q < 3/2$. Then,*

$$(1.10) \qquad \|\mathcal{S}_k[f](\xi, x)\|_{L_x^p(M_{\xi,k}^{q,s})} \leq C_{p,q,s} \, \|f\|_{L^p}$$

where, given an exponent s, we define the s-variation multiplier norm of a sequence of functions $\{m_k\}_{k \in \mathbb{Z}}$

$$\|m\|_{M^{q,s}} = \sup_{g : \|g\|_{L^q} = 1} \left\| (m_k \hat{g})\check{\ }(x) \right\|_{L_x^q(V_k^s)}.$$

One reason for interest in bounds such as (1.10) is that the analogous bounds for $\mathfrak{S}_k$ and $\mathfrak{A}_k$ would yield quantitative information about the convergence in (1.8) and (1.9).

---

[2]Which, at least morally, is an easier setting: one manifestation of this is that the analog of the Carleson–Hunt theorem for $\mathfrak{A} := \mathfrak{A}_0$ is trivial.

Through Corollaries 6.3 and 6.4, we obtain the following variants of (1.4):

Theorem 1.3. *Suppose that $r > 2$ and $p > r'$. Then*

$$\|\mathcal{S}_k[f](\xi,x)\|_{L^p_x(\ell^\infty_k(V^r_\xi))} \leq C_{p,r} \|f\|_{L^p}.$$

Theorem 1.4. *Suppose that $r > 2$ and $1 < p < \infty$. Then*

$$\|\mathcal{S}_k[f](\xi,x)\|_{L^p_x(L^\infty_\xi(V^r_k))} \leq C_{p,r} \|f\|_{L^p}.$$

The Fourier analog

$$(1.11) \qquad \|\mathfrak{S}_k[f](\xi,x)\|_{L^p_x(\ell^\infty_k(V^r_\xi))} \leq C_{p,r} \|f\|_{L^p}.$$

of Theorem 1.3 can be deduced from the Fourier analog of (1.4) by treating $\mathfrak{S}_k$ as a superpositioning of modulated versions of $\mathfrak{S}$. If (1.11) held for exponents $r < 2$ (it does not), then the method of [11] would allow one to deduce bounds of the type (1.7) without using maximal-multiplier estimates such as Lemma 3.1.

The analog of Theorem 1.4 for $\mathfrak{S}_k$ is related to the Wiener–Wintner theorem for the Hilbert transform [8], and was obtained in [12] for the restricted range of exponents $r > 2$ and $p > r'$ using the superposition argument.

The superposition argument does not seem to immediately apply to the Walsh-operator $\mathcal{S}_k$ due to the different method of truncation.

Structure of the paper. We give background information on the Walsh–Fourier transform in Section 2. Machinery is developed in Sections 3 through 7. The machinery is applied to finish the proofs in Section 8. Additional refinements needed for Theorem 1.2 are given in Section 9.

Notational conventions. We use $|\cdot|$ for Lebesgue measure, cardinality, or an understood norm depending on context. Given a rectangle $P = I \times \omega$ we let I_P denote I and ω_P denote ω. The indicator function of a set E is written 1_E. Dyadic intervals are half-open on the right, i.e., of the form $[n2^k, (n+1)2^k)$ for integers n and k.

2. Terminology of Walsh phase-plane analysis

Given a nonnegative real number x and an integer n, let $d_n(x)$ denote the digit which sits in the $(n+1)$st position to the left of the point in the binary expansion of x, i.e.,

$$x = \sum_n d_n(x)\, 2^n$$

(for points x on the dyadic grid, we choose the expansion with $d_n(x) = 0$ for $-n$ sufficiently large). Define the bitwise addition operation $\oplus$:

$$d_n(x \oplus y) = d_n(x) + d_n(y) \mod 2, \quad -\infty < n < \infty,$$

and the corresponding multiplication operation $\otimes$:

$$d_n(x \otimes y) = \sum_m d_{n-m}(x) d_m(y) \mod 2, \quad -\infty < n < \infty.$$

Note that the Lebesgue measure on $\mathbb{R}^+$ is invariant under $\oplus$-translation.

Consider the character

$$e(x) = e^{i\pi d_{-1}(x)}$$

and given a function f on $\mathbb{R}^+$ and $\xi \in \mathbb{R}^+$ define the Walsh–Fourier transform

$$\hat{f}(\xi) = \int_{\mathbb{R}^+} e(\xi \otimes x) f(x) \, dx.$$

It is straightforward to verify that $\hat{1}_{[0,1)} = 1_{[0,1)}$. Thus, after checking the identities $\widehat{f(x \oplus \cdot)} = e(x \otimes \cdot)\hat{f}$, $\widehat{e^{\xi \otimes} \cdot f} = \hat{f}(\xi \oplus \cdot)$, and $\widehat{f(2^k \cdot)} = 2^{-k}\hat{f}(2^{-k}\cdot)$, one sees that (for linear combinations of characteristic functions of dyadic intervals and hence, by density, for general functions $f \in L^2$) $\hat{\ }$ is involutive; however, for metaphorical purposes we will sometimes use the notation $\check{\ }$ in place of $\hat{\ }$.

Given a dyadic "time-interval" $I \subset \mathbb{R}^+$ and a dyadic "frequency interval" $\omega \subset \mathbb{R}^+$, we say that the rectangle $I \times \omega$ is a tile if $|I \times \omega| = 1$ and we say that it is a bitile if $|I \times \omega| = 2$. Each bitile P contains four tiles P_u, P_l, P_s, and P_d which are the upper, lower, left, and right halves respectively. We impose the following partial order on the set of tiles and on the set of bitiles:

$$I_1 \times \omega_1 \leq I_2 \times \omega_2 \iff \omega_2 \subset \omega_1 \text{ and } I_1 \subset I_2.$$

A set of bitiles $\mathbf{P}$ is convex if for all bitiles $P_1 \leq P_2 \leq P_3$ with $P_1, P_3 \in \mathbf{P}$ we also have $P_2 \in \mathbf{P}$. Through the use of standard limiting arguments we can, and will, assume that all bitiles belong to a finite convex set $\mathbf{P}_0$; all constants will be independent of this set.

Associated to each tile is the L^2-normalized Walsh wave-packet

$$(2.1) \qquad\qquad \phi_{I \times \omega}(x) = |I|^{1/2} \, \check{1}_\omega(x \oplus \inf I).$$

Since $\phi_{I \times \omega}$ is supported on I and $\hat{\phi}_{I \times \omega}$ is supported on ω, ϕ_p and $\phi_{p'}$ are orthogonal unless the tiles p and p' have nonempty intersection. Letting $*$ denote the convolution operation,

$$f * g(x) = \int_{\mathbb{R}^+} f(x \oplus y) \, g(y) \, dy,$$

and letting $D_k[f](x)$ denote the average of f over the dyadic interval of length 2^k containing x, we have

$$D_k[f](x) = f * 2^{-k} 1_{[0,2^k)}(x),$$

and, more generally, that for any dyadic interval ω

$$(\hat{f} 1_\omega)^{\check{}}(x) = \langle f, \phi_{I_x \times \omega} \rangle \, \phi_{I_x \times \omega}(x),$$

where I_x is the dyadic interval of length $|\omega|^{-1}$ containing x.

For each bitile P we have the following relations between the wave-packets for the enclosed tiles:

$$(2.2) \qquad \phi_{P_u} = \frac{1}{\sqrt{2}}(\phi_{P_s} - \phi_{P_d})$$

$$(2.3) \qquad \phi_{P_l} = \frac{1}{\sqrt{2}}(\phi_{P_s} + \phi_{P_d}).$$

The relations above can be used to check that our definition (2.1) agrees with that in [6]. If a subset S of $\mathbb{R}^+ \times \mathbb{R}^+$ can be written as the disjoint union of tiles $S = \bigcup_{p \in \mathbf{p}} p$ we define the phase-plane projection

$$\Pi_S f = \sum_{p \in \mathbf{p}} \langle f, \phi_p \rangle \, \phi_p.$$

When S is a bitile, it follows immediately from (2.2) and (2.3) that the projection is independent of the cover used in the definition. This is also true for general sets S, as can be seen by repeatedly appealing to the special case of the bitile. If $S \subset S'$ and $\Pi_S, \Pi_{S'}$ are both defined then

$$(2.4) \qquad \Pi_S \Pi_{S'} = \Pi_{S'} \Pi_S = \Pi_S.$$

3. Some multiplier estimates

In this section we recall an extension of a maximal-multiplier lemma of Bourgain, we recall a multiplier bound of Coifman, Rubio de Francia, and Semmes, and we then prove an estimate which is a hybrid of the two results.

3.1. A maximal-multiplier lemma

Suppose that for every dyadic interval ω we have a coefficient $a_\omega \in \mathbb{R}$. Let $\Xi \subset \mathbb{R}^+$ be a finite collection of frequencies, and for each integer k, consider the Walsh multiplier

$$\mathcal{D}_k(\xi) = \sum_{\substack{|\omega|=2^k \\ \omega \cap \Xi \neq \emptyset}} a_\omega \, 1_\omega(\xi),$$

where, above, we sum over dyadic intervals ω. Building on work of Bourgain [2], the following estimate was proven in [6]:

Lemma 3.1. *Let $r > 2$, and $\Xi \subset \mathbb{R}^+$. Then*

$$\|\mathcal{D}_k\|_{M^{2,*}} \leq C_r \, (1 + \log|\Xi|) \, |\Xi|^{1/2 - 1/r} \, \sup_{\xi \in \Xi} \Big\| \sum_{|\omega|=2^k} a_\omega 1_\omega(\xi) \Big\|_{V_k^r}.$$

In [5] and [10] the Fourier-multiplier version of the estimate above was extended to L^q for $1 < q < 2$. Following the Walsh analog of the argument in [10] (which is part of the proof of Lemma 3.7 below), one would obtain:

Lemma 3.2. *Let $r > 2$, $1 < q < 2$, and $\Xi \subset \mathbb{R}^+$. Then*

$$\|\mathcal{D}_k\|_{M^{q,*}} \le C_{q,r} \left(1 + \log|\Xi|\right) |\Xi|^{1/q - 1/r} \sup_{\xi \in \Xi} \Big\| \sum_{|\omega| = 2^k} a_\omega 1_\omega(\xi) \Big\|_{V_k^r}.$$

We will need to use an estimate which encompasses both Lemma 3.2 and a separate multiplier bound which we will now discuss.

3.2. Multipliers of bounded r-variation

It was shown in [3], see also [15], that if m is a function of bounded r-variation then the associated Fourier multiplier operator is bounded on L^q when $|1/q - 1/2| < 1/r$ and $1 < q < \infty$. The Walsh multiplier analog of this result can be proven by the same method, which we now outline.

The first step is to obtain an estimate for multipliers given by linear combinations of characteristic functions of intervals.

Lemma 3.3. *Let $\epsilon > 0$, let Υ be a finite collection of pairwise disjoint subintervals of $\mathbb{R}^+$, and let $\{b_v\}_{v \in \Upsilon} \subset \mathbb{R}$ be a collection of coefficients. Then for $1 < q < \infty$*

$$\Big\| \sum_{v \in \Upsilon} b_v 1_v \Big\|_{M^q} \le C_{q,\epsilon} |\Upsilon|^{1/q - 1/2| + \epsilon} \sup_{v \in \Upsilon} |b_v|.$$

Proof. In [3], the Fourier multiplier version of this lemma was proven through the use of the Rubio de Francia square function estimate. One could follow the same route here by proving a Walsh analog of the square function estimate, or by instead using (1.4) with r close to 2 as a substitute for the square function estimate. We will instead appeal to an estimate below using a multiple frequency Calderón–Zygmund decomposition (this bears some similarity to the approach in [15]). By duality one may assume $1 < q < 2$. Choosing $r > 2$ and $\epsilon' > 0$ sufficiently small we have $1/q - 1/r + \epsilon' < 1/q - 1/2 + \epsilon$. Taking $\Xi = \{0\}$ and $a_\omega = 1$ for every ω we then have

$$\Big(\sum_{v \in \Upsilon} b_v 1_v \hat{g} \Big)^{\vee}(x) = \lim_{k \to \infty} \Big(\mathcal{D}_k \sum_{v \in \Upsilon} b_v 1_v \hat{g} \Big)^{\vee}(x)$$

almost everywhere and so the lemma follows immediately from Lemma 3.7. □

Next, we see (in a lemma directly from [3]) that functions of bounded r-variation can be written efficiently as sums of functions of the type treated in Lemma 3.3.

Lemma 3.4. *Let m be a compactly supported function on $\mathbb{R}^+$ of bounded r-variation for some $1 \le r < \infty$. Then for each integer $j \ge 0$, one can find a collection Υ_j of pairwise disjoint subintervals of $\mathbb{R}^+$ and coefficients $\{b_v\}_{v \in \Upsilon_j} \subset \mathbb{R}$ so that $|\Upsilon_j| \le 2^j$, $|b_v| \le \|m\|_{V_r} 2^{-j/r}$, and*

$$m = \sum_{j \ge 0} \sum_{v \in \Upsilon_j} b_v 1_v,$$

where the sum over j converges uniformly.

Proof. The proof is exactly as in [3] (also see [9]) so we will be short with the details. Choose B so that m is supported on $[0, B]$. Set $V(0) = 0$ and for each $x \in (0, B]$ let

$$V(x) = \sup_{N, 0 = \xi_0 < \ldots < \xi_N = x} \sum_{k=1}^{N} |m(\xi_k) - m(\xi_{k-1})|^r.$$

For $j \geq 0$ and $1 \leq l < 2^j$ let

$$\upsilon_{j,l} = V^{-1}\big([(l-1)\, 2^{-j}\, \|m\|_{V^r}^r, l\, 2^{-j}\, \|m\|_{V^r}^r)\big)$$

and

$$\upsilon_{j,2^j} = V^{-1}\big([(2^j - 1)\, 2^{-j}\, \|m\|_{V^r}^r, \|m\|_{V^r}^r]\big).$$

Set $\tilde{b}_{\upsilon_{j,l}} = 0$ if $\upsilon_{j,l} = \emptyset$, $\tilde{b}_{\upsilon_{j,l}} = m(\upsilon_{j,l})$ if $\upsilon_{j,l}$ is a singleton, and

$$\tilde{b}_{\upsilon_{j,l}} = \frac{1}{|\upsilon_{j,l}|} \int_{\upsilon_{j,l}} m(x)\ dx$$

if $\upsilon_{j,l}$ is an interval containing more than one element. Then letting $\Upsilon_j = \{\upsilon_{j,l}\}_{l=1}^{2^j}$, $b_{\upsilon_{0,l}} = \tilde{b}_{\upsilon_{0,l}}$, and $b_{\upsilon_{j,l}} = \tilde{b}_{\upsilon_{j,l}} - \tilde{b}_{\upsilon_{j-1,l}}$ for $j > 0$, one sees that the requirements of the lemma are satisfied. $\qquad\square$

Finally we combine Lemmas 3.3 and 3.4 to obtain the Walsh analog of a result[3] from [3]:

Lemma 3.5. *Let $1 < q < \infty$, $|1/q - 1/2| < 1/r$, and $\epsilon > 0$. Suppose Υ is a collection of pairwise disjoint subintervals of $\mathbb{R}^+$ and that for each $\upsilon \in \Upsilon$, m_υ is a function supported on υ with $\|m\|_{V^r} < \infty$. Then*

$$\Big\| \sum_{\upsilon \in \Upsilon} m_\upsilon \Big\|_{M^q} \leq C_{q,r,\epsilon}\, |\Upsilon|^{|1/q - 1/2| + \epsilon} \sup_{\upsilon \in \Upsilon} \|m_\upsilon\|_{V^r}.$$

Proof. After a limiting argument, one may assume that all intervals in Υ have finite length. Applying Lemma 3.4 to each m_υ we obtain for $j \geq 0$ a collection $\mathcal{I}_{\upsilon,j}$ of at most 2^j pairwise disjoint subintervals of υ and coefficients $\{b_I\}_{I \in \mathcal{I}_{\upsilon,j}}$ so that

$$m_\upsilon = \sum_{j \geq 0} \sum_{I \in \mathcal{I}_{\upsilon,j}} b_I 1_I.$$

Then

$$\Big\| \sum_{\upsilon \in \Upsilon} m_\upsilon \Big\|_{M^q} \leq \sum_{j \geq 0} \Big\| \sum_{\upsilon \in \Upsilon} \sum_{I \in \mathcal{I}_{\upsilon,j}} b_I 1_I \Big\|_{M^q}.$$

---

[3] Strictly speaking, the authors of [3] considered the case where Υ was a collection of dyadic shells and through an additional Littlewood–Paley argument were able to obtain a norm bound which did not blow up with $|\Upsilon|$. To match our application, we are more flexible with Υ and can accept the resulting loss in the bound.

Applying Lemma 3.3 with the collection of pairwise disjoint intervals $\bigcup_{\upsilon \in \Upsilon} \mathcal{I}_{\upsilon,j}$ we see that each term on the right above is

$$\leq C_{q,\epsilon} \left(2^j |\Upsilon|\right)^{|1/q-1/2|+\epsilon} \sup_{\upsilon \in \Upsilon, I \in \mathcal{I}_{\upsilon,j}} |b_I|$$

$$\leq C_{q,\epsilon} \left(2^j |\Upsilon|\right)^{|1/q-1/2|+\epsilon} 2^{-j/r} \sup_{\upsilon \in \Upsilon} \|m_\upsilon\|_{V^r}.$$

The sum over $j \geq 0$ converges after possibly shrinking ϵ to satisfy $|\frac{1}{q}-\frac{1}{2}|+\epsilon < \frac{1}{r}$.   □

3.3. A hybrid estimate

Our aim here is to prove the following lemma which, except for a restriction on the range of r and a difference in the dependence on $|\Upsilon|$, is a common extension of Lemmas 3.2 and 3.5.

Lemma 3.6. *Let $1 < q \leq 2$, $2 < r < 2q$, $\epsilon > 0$, and $\Xi \subset \mathbb{R}^+$. Suppose that Υ is a collection of pairwise disjoint subintervals of $\mathbb{R}^+$ and that $\{m_\upsilon\}_{\upsilon \in \Upsilon}$ is a collection of functions of bounded r-variation such that each m_υ is supported on υ. Then*

$$\Big\|\mathcal{D}_k \sum_{\upsilon \in \Upsilon} m_\upsilon\Big\|_{M^{q,*}} \leq C_{q,r,\epsilon} \left(|\Xi|+|\Upsilon|\right)^{1/q-1/r+\epsilon} \sup_{\xi \in \Xi} \Big\| \sum_{|\omega|=2^k} a_\omega 1_\omega(\xi)\Big\|_{V_k^r} \sup_{\upsilon \in \Upsilon} \|m_\upsilon\|_{V^r}.$$

A version of the lemma above, but with $\left(|\Xi|+|\Upsilon|\right)^{1/q-1/r+\epsilon}$ replaced by

$$|\Xi|^{1/q-1/r+\epsilon} |\Upsilon|^{1/q-1/2+\epsilon}$$

would follow by combining Lemmas 3.2 and 3.5 and estimating the operator norm of the composition by the product of the operator norms. In our application we will take r arbitrarily close to 2 and $|\Upsilon| = |\Xi|$; thus, the norm bound obtained above improves substantially on the combination of the two prior lemmas; this improvement seems to be necessary to obtain the desired range of exponents in Theorem 1.1.

The new ingredient needed in the proof of Lemma 3.6 is the following hybrid of Lemmas 3.2 and 3.3.

Lemma 3.7. *Let $r > 2$, $1 < q \leq 2$, and $\Xi \subset \mathbb{R}^+$. Suppose that Υ is a collection of pairwise disjoint subintervals of $\mathbb{R}^+$ and that $\{b_\upsilon\}_{\upsilon \in \Upsilon} \subset \mathbb{R}$ is a collection of coefficients. Then*

$$\Big\|\mathcal{D}_k \sum_{\upsilon \in \Upsilon} b_\upsilon 1_\upsilon\Big\|_{M^{q,*}}$$
$$\leq C_{q,r} \left(1 + \log(|\Xi|+|\Upsilon|)\right) \left(|\Xi|+|\Upsilon|\right)^{1/q-1/r} \sup_{\xi \in \Xi} \Big\| \sum_{|\omega|=2^k} a_\omega 1_\omega(\xi)\Big\|_{V_k^r} \sup_{\upsilon \in \Upsilon} |b_\upsilon|.$$

Proof. Through a limiting argument, one may assume that all intervals in Υ have finite length.

The desired bound at $q = 2$ follows immediately from Lemma 3.1, and so by interpolation it suffices to prove a weak-type estimate at $q = 1$. Specifically, given $g \in L^1$ we need to show that for each $\lambda > 0$

$$\left| \left\{ x : \sup_k |(\mathcal{D}_k \sum_{v \in \Upsilon} b_v 1_v \hat{g})\check{\ }(x)| > \lambda \right\} \right| \leq C N^{1/2} B \|g\|_{L^1} / \lambda,$$

where $N = |\Xi| + |\Upsilon|$ and

$$B = (1 + \log(N)) N^{1/2 - 1/r} \sup_{\xi \in \Xi} \big\| \sum_{|\omega| = 2^k} a_\omega 1_\omega(\xi) \big\|_{V_k^r} \sup_{v \in \Upsilon} |b_v|.$$

We start by performing a multiple-frequency Calderón–Zygmund decomposition. Let

$$E = \{ x : M[g](x) > \lambda/(N^{1/2} B) \},$$

where M is the dyadic version of the Hardy–Littlewood maximal operator. Let $\mathcal{I}$ be the collection of maximal dyadic intervals contained in E and

$$\Lambda = \Xi \cup \bigcup_{v \in \Upsilon} \{\inf v, \sup v\}.$$

We now construct the "good function" $\mathfrak{g}$. Let

$$\mathfrak{g}_0 = \sum_{I \in \mathcal{I}} \sum_{\substack{|\omega| = |I|^{-1} \\ \omega \cap \Lambda \neq \emptyset}} \langle g, \phi_{I \times \omega} \rangle \phi_{I \times \omega},$$

where the second sum above is over dyadic intervals ω. Setting

$$\mathfrak{g} = \mathfrak{g}_0 + 1_{\mathbb{R}^+ \setminus E} \, g,$$

we obtain the "bad function"

$$\mathfrak{b} = g - \mathfrak{g},$$

and write $\mathfrak{b}_I$ in place of $1_I \mathfrak{b}$ for each $I \in \mathcal{I}$.

The contribution from the good function is controlled, as usual, by the previously known L^2 bound. Indeed, by the maximality of the intervals I, we have

$$| \langle g, \phi_{I \times \omega} \rangle | \leq 2 |I|^{1/2} \lambda/(N^{1/2} B).$$

Using the orthogonality of the wavepackets and the fact that $|\Lambda| \leq 2N$ we have

$$\|1_I \mathfrak{g}_0\|_{L^2} \leq C |I|^{1/2} \lambda/B.$$

This gives

$$\|\mathfrak{g}_0\|_{L^2}^2 \leq C |E| \lambda^2/B^2 \leq C N^{1/2} \|g\|_{L^1} \lambda/B.$$

Since g is bounded by $\lambda/(N^{1/2} B)$ away from E, we have

$$\|\mathfrak{g} - \mathfrak{g}_0\|_{L^2}^2 \leq N^{-1/2} \|g\|_{L^1} \lambda/B.$$

Thus,

$$\left|\left\{x : \sup_k |(\mathcal{D}_k \sum_{v\in\Upsilon} b_v 1_v \hat{\mathfrak{g}})^{\vee}(x)| > \lambda/2\right\}\right| \leq 4 \left\|\sup_k |(\mathcal{D}_k \sum_{v\in\Upsilon} b_v 1_v \hat{\mathfrak{g}})^{\vee}|\right\|_{L^2}^2 / \lambda^2$$

$$\leq C\,B^2\, \|\mathfrak{g}\|_{L^2}^2 / \lambda^2 \leq C\,N^{1/2}\,B\,\|g\|_{L^1}/\lambda,$$

as desired.

It remains to control the contribution from the bad function. The two important properties of $\mathfrak{b}$ are that it is supported on E and that for each I in $\mathcal{I}$ we have $\langle \mathfrak{b}, \phi_{I\times\omega}\rangle = 0$ for every dyadic interval ω with $|\omega| = |I|^{-1}$ and $\omega\cap\Lambda \neq \emptyset$. We claim that the function

$$h = \Big(\sum_{v\in\Upsilon} b_v 1_v \hat{\mathfrak{b}}\Big)^{\vee}$$

shares these two properties with $\mathfrak{b}$. We first consider the support property. Fix $I \in \mathcal{I}$, suppose $v \in \Upsilon$, and let ω be a maximal dyadic subinterval of v. Then

$$(1_\omega \hat{\mathfrak{b}}_I)^{\vee}(x) = \langle \mathfrak{b}_I, \phi_{J\times\omega}\rangle\, \phi_{J\times\omega}(x),$$

where J is the dyadic interval of length $|\omega|^{-1}$ containing x. We clearly have $\langle \mathfrak{b}_I, \phi_{J\times\omega}\rangle = 0$ if J has empty intersection with I. If $x \notin I$ and J intersects I then we have $I \subsetneq J$ and in particular $|I| < |J|$. By the maximality of ω, the dyadic interval $\tilde\omega$ of length $|I|^{-1}$ containing ω intersects $\{\sup v, \inf v\}$ and hence $\langle \mathfrak{b}_I, \phi_{I\times\tilde\omega}\rangle = 0$. Using the fact that the restriction of $\phi_{J\times\omega}$ to I is a constant multiple of $\phi_{I\times\tilde\omega}$ we see that $\langle \mathfrak{b}_I, \phi_{J\times\omega}\rangle = 0$. Since each v can be written as the union of maximal dyadic subintervals, this implies that $(\sum_{v\in\Upsilon} b_v 1_v \hat{\mathfrak{b}}_I)^{\vee}$ is supported on I and so h is supported on E. To verify the cancellation property of h, we let $I \in \mathcal{I}$ and let ω be a dyadic interval of length $|I|^{-1}$ such that $\omega\cap\Lambda \neq \emptyset$. Then $\langle h, \phi_{I\times\omega}\rangle$ is zero simultaneously with the restriction of $(1_\omega \widehat{1_I h})^{\vee}$ to I. However, $1_I h = (\sum_{v\in\Upsilon} b_v 1_v \hat{\mathfrak{b}}_I)^{\vee}$ and so $1_\omega \widehat{1_I h} = \sum_{v\in\Upsilon} b_v 1_v 1_\omega \hat{\mathfrak{b}}_I$. From the cancellation property of $\mathfrak{b}$, we know that $1_\omega \hat{\mathfrak{b}}_I$ is identically zero.

Arguing as in the previous paragraph, except with h in place of $\mathfrak{b}$, one sees that each $(\mathcal{D}_k \hat{h})^{\vee}$ is supported on E and thus

$$\left|\left\{\sup_k |(\mathcal{D}_k h)^{\vee}(x)| > \lambda/2\right\}\right| \leq |E| \leq N^{1/2}\,B\,\|g\|_{L^1}/\lambda. \qquad \square$$

Proof of Lemma 3.6. Following the argument in Lemma 3.5, but with Lemma 3.7 substituted for Lemma 3.3, we see that

$$\Big\|\mathcal{D}_k \sum_{v\in\Upsilon} m_v\Big\|_{M^{q,*}}$$

$$\leq \sum_{j\geq 0} C_{q,r,\epsilon'}(|\Xi| + 2^j |\Upsilon|)^{1/q - 1/r + \epsilon'} \sup_{\xi\in\Xi} \Big\|\sum_{|\omega|=2^k} a_\omega 1_\omega(\xi)\Big\|_{V_k^r} \sup_{v\in\Upsilon, I\in\mathcal{I}_{v,j}} |b_I|$$

$$\leq \sum_{j\geq 0} C_{q,r,\epsilon'} 2^{j(\frac{1}{q} - \frac{1}{r} + \epsilon')}(|\Xi| + |\Upsilon|)^{\frac{1}{q} - \frac{1}{r} + \epsilon'} \sup_{\xi\in\Xi} \Big\|\sum_{|\omega|=2^k} a_\omega 1_\omega(\xi)\Big\|_{V_k^r} 2^{-j/r} \sup_{v\in\Upsilon} \|m_v\|_{V^r}.$$

The sum over $j \geq 0$ then converges provided that $r < 2q$ and $\epsilon' < \epsilon$ is chosen sufficiently small. $\qquad \square$

4. Tree decompositions

Given a collection of bitiles T, a "top frequency" $\xi_T \in \mathbb{R}^+$, and a dyadic "top interval" $I_T \subset \mathbb{R}^+$, we say that (T, ξ_T, I_T) form a tree if $I_P \subset I_T$ and $\xi_T \in \omega_P$ for every $P \in T$. We say that a tree T is "td-maximal" among trees in a collection $\mathbf{T}$ if it is maximal with respect to inclusion among trees in $\mathbf{T}$ with top data (ξ_T, I_T). Given any convex tree T, we can rewrite $\cup_{P \in T} P$ as a disjoint union of tiles. For such a tree, we abbreviate

$$\Pi_T = \Pi_{\bigcup P : P \in T}.$$

For a convex collection $\mathbf{P}$ of bitiles define

$$\text{size}(\mathbf{P}, f) = \sup_T |I_T|^{-1/2} \|\Pi_T f\|_{L^2},$$

where the sup is over all convex trees $T \subset \mathbf{P}$. Note that (since the L^∞ norm can be controlled by projections to individual subtiles of elements of T) for each convex tree T and $1 \le p \le \infty$ we have

$$\|\Pi_T f\|_{L^p} \le C \, \text{size}(T, f) \, |I_T|^{1/p}.$$

The following lemma was proven in [14]:

Lemma 4.1 (Tree selection). *Assume $\mathbf{P}$ is a finite convex collection of bitiles with* $\text{size}(\mathbf{P}, f) \le 2^{-k}$. *Then we can write $\mathbf{P}$ as the disjoint union of a convex set of bitiles $\mathbf{P}'$ with the union of a collection $\mathbf{T}$ of convex trees such that*

$$(4.1) \qquad \big\| \sum_{T \in \mathbf{T}} 1_{I_T} \big\|_1 \le C \, 2^{2k} \|f\|_2^2,$$

$$(4.2) \qquad \big\| \sum_{T \in \mathbf{T}} 1_{I_T} \big\|_{\text{BMO}} \le C \, 2^{2k} \|f\|_\infty^2,$$

and $\text{size}(\mathbf{P}', f) \le 2^{-k-1}$.

Strictly speaking, the lemma above was proven with a different definition of tree – in [14] a tree is a collection of bitiles with a unique maximal element, we will call this an m-tree. It is easily seen that the lemma for m-trees implies the lemma for trees since every m-tree is a tree and every finite convex tree T contains a convex m-tree T' with $\text{size}(T', f) = \text{size}(T, f)$.

We also note that we may assume that the trees T in the lemma above are td-maximal among trees contained in $\bigcup_{T \in \mathbf{T}} T$; this is accomplished by taking them to be td-maximal among trees contained in $\mathbf{P}$.

Finally, by following the proof of (4.2) one sees that for any $\mathbf{T}' \subset \mathbf{T}$ we also have

$$\big\| \sum_{T \in \mathbf{T}'} 1_{I_T} \big\|_{\text{BMO}} \le C \, 2^{2k} \|f\|_\infty^2.$$

We say that a tree T is *l*-overlapping if $\xi_T \in \omega_{P_l}$ for every $P \in T$ and *u*-overlapping if $\xi_T \in \omega_{P_u}$ for every $P \in T$. We call a set of bitiles $\mathbf{P}$ "*l*-convex"

if $P, P'' \in \mathbf{P}$ and $P_l < P'_l < P''_l$ imply that $P'_l \in \mathbf{P}$. Finally, a collection of l-overlapping trees $\mathbf{T}$ will be called "properly-sorted" if the following conditions hold for each $T \in \mathbf{T}$:

(3) T is l-convex;

(4) for every $T' \in \mathbf{T} \setminus \{T\}$ we have $T' \cap T = \emptyset$;

(5) for every $P \in T$ and $T' \in \mathbf{T}$ with $I_{T'} \cap I_P \neq \emptyset$ we have $\xi_{T'} \notin \omega_{P_u}$.

The importance of condition (3) is that when an l-overlapping tree T is l-convex,

$$(4.6) \qquad \bigcup_{\substack{P \in T \\ x \in I_P}} \omega_{P_u}$$

is an interval for each x. Indeed, suppose that $\xi_1 < \xi_2 < \xi_3$ with ξ_1, ξ_3 in the set (4.6). Then, there are bitiles $P^i \in T$ with $\xi_i \in \omega_{P_u^i}$ and $x \in I_{P^i}$ for $i = 1, 3$; clearly $P_l^3 \leq P_l^1$. Let ω be the smallest dyadic interval with $\xi_T, \xi^2 \in \omega$, let I be the dyadic interval of length $2|\omega|^{-1}$ containing x, and let P^2 be the bitile $I \times \omega$. Then $P_l^3 \leq P_l^2 \leq P_l^1$, so by l-convexity $P^2 \in T$ and hence ξ_2 is in the set (4.6).

The condition (5) is taken from [6]; one of its immediate consequences is that the tiles P_u with $P \in \bigcup_{T \in \mathbf{T}}$ are disjoint. Indeed, $P_u < P'_u$ would imply that the tree containing P'_u has top frequency contained in ω_{P_u}.

We will apply the lemma below to the collection of bitiles $\mathbf{P} \setminus \mathbf{P}'$ from Lemma 4.1.

Lemma 4.2. *Suppose that a finite collection of bitiles $\mathbf{P}$ can be written as the union of convex trees $\mathbf{P} = \bigcup_{T \in \mathbf{T}} T$. Assume that the trees are td-maximal among trees contained in $\mathbf{P}$. Then, we can write*

$$\mathbf{P} = \Big(\bigcup_{T \in \mathbf{T}^u} T \Big) \cup \Big(\bigcup_{T \in \mathbf{T}^l} T \Big),$$

where

(7) *For each $T \in \mathbf{T}^u$, T is a u-overlapping tree which is td-maximal among u-overlapping trees contained in $\bigcup_{T \in \mathbf{T}^u} T$.*

(8) $\mathbf{T}^l$ *is a properly-sorted collection of l-overlapping trees.*

(9) $\bigcup_{T \in \mathbf{T}^u} T \cap \bigcup_{T \in \mathbf{T}^l} T = \emptyset$.

(10) $\{(\xi_T, I_T) : T \in \mathbf{T}^u\} = \{(\xi_T, I_T) : T \in \mathbf{T}^l\} \subset \{(\xi_T, I_T) : T \in \mathbf{T}\}$.

Proof. After throwing out some trees, we may assume that for each $T \in \mathbf{T}$,

$$(4.11) \qquad T \not\subset \bigcup_{T' \in \mathbf{T} \setminus \{T\}} T'.$$

We enumerate $\mathbf{T} = \{T_1, \ldots, T_N\}$ so that for each i, $\xi_{T_i} \leq \xi_{T_{i+1}}$. Set $\mathbf{P}_{N+1} = \mathbf{P}$ and $T_{N+1} = \emptyset$ and for $i = 1, \ldots, N$ let T_i^u be the maximal u-overlapping tree contained in $\mathbf{P}_{i+1}$ with top data (ξ_{T_i}, I_{T_i}), let T_i^l be the maximal l-overlapping tree contained in $\mathbf{P}_{i+1}$ with top data (ξ_{T_i}, I_{T_i}), and let $\mathbf{P}_i = \mathbf{P}_{i+1} \setminus (T_i^u \cup T_i^l)$.

Set $\mathbf{T}^u = \{T_1^u, \ldots, T_N^u\}$ and $\mathbf{T}^l = \{T_1^l, \ldots, T_N^l\}$.

We now verify (3). Suppose that $P, P'' \in T_i^l$ and that $P_l < P_l' < P_l''$. By convexity and td-maximality of the trees, we have $P, P', P'' \in T_i$. It remains to verify that $P' \in \mathbf{P}_{i+1}$, but this follows immediately from the fact that $P'' \in \mathbf{P}_{i+1}$.

To check (5) suppose that $P \in T_i^l$, $\xi_{T_j} \in \omega_{P_u}$, and $I_{T_j} \cap I_P \neq \emptyset$. Since T_i^l is l-overlapping, $\xi_{T_j} > \xi_{T_i}$ and so we have $j > i$. Since $P \in \mathbf{P}_{i+1}$ we know that $P \in \mathbf{P}_{j+1} \setminus T_j^u$; combining this with the fact that $\xi_{T_j} \in \omega_{P_u}$ implies that $I_{T_j} \subsetneq I_P$ contradicting (4.11).

Enlarging the trees in $\mathbf{T}^u$ to obtain td-maximality gives (7). Conditions (9), (10), and (4) are clear by construction. $\qquad\square$

5. Global variation for a single tree

The Lemma below will be used in Section 8 to give pointwise variation-norm estimates which are compatible with Corollary 6.2 and Lemmas 3.2, 3.5, and 3.6.

Lemma 5.1. *Suppose $1 < p < \infty$ and $r > 2$. Then for every tree T which is contained in a convex tree $\overline{T}$ and which is either l-overlapping and l-convex or u-overlapping,*

$$(5.1) \qquad \Big\| \sum_{P \in T} \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{L_x^p(V_\xi^r)} \leq C_{p,r} \, \mathrm{size}(\overline{T}) \, |I_{\overline{T}}|^{1/p}.$$

Proof. Note that, by (2.4), the left side of (5.1) is equal to

$$(5.2) \qquad \Big\| \sum_{P \in T} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{L_x^p(V_\xi^r)},$$

where $g = \Pi_{\overline{T}} f$.

First we consider the case when T is u-overlapping. Then for each ξ we have

$$(5.3) \qquad \sum_{P \in T} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) = \sum_{\substack{P \in T \\ |I_T| \leq 2^k \xi}} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x),$$

where k_ξ is the largest k such that the dyadic interval of length 2^{-k} about ξ_T contains ξ. Then k_ξ is monotonic in ξ on the intervals (ξ_T, ∞) and $[0, \xi_T)$ and so (5.2) is

$$(5.4) \qquad \leq C \Big\| \sum_{\substack{P \in T \\ |I_T| \leq 2^k}} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x) \Big\|_{L_x^p(V_k^r)}.$$

One can check that

$$\sum_{\substack{P \in T \\ |I_T| \leq 2^k}} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x)$$

$$= \sum_{P \in T} \langle g, \phi_{P_l} \rangle \, \phi_{P_l}(x) - \mathrm{sgn}(\phi_{p_T}) D_k \Big[\mathrm{sgn}(\phi_{p_T}) \sum_{P \in T} \langle g, \phi_{P_l} \rangle \, \phi_{P_l} \Big](x),$$

where p_T is a tile satisfying $|I_{p_T}| = |I_T|$, $\xi_T \in \omega_{p_T}$, and D_k is the dyadic averaging operator at scale 2^k as defined in Section 2. Since $r > 2$ and $1 < p < \infty$, we may apply Lépingle's bound for the variation of martingale averages to see that the right side of (5.4) is

$$\leq C_{p,r} \Big\| \sum_{P \in T} \langle g, \phi_{P_l} \rangle \phi_{P_l} \Big\|_{L^p} \leq C_{p,r} \|g\|_{L^p} \leq C_{p,r} \, \text{size}(\overline{T}) \, |I_{\overline{T}}|^{1/p} \,,$$

as desired.

For l-overlapping l-convex T one can check that, for $\xi > \xi_T$, the left side of (5.3)

$$(5.5) \qquad\qquad = \sum_{\substack{P \in T \\ |I_T| = 2^k \xi}} \langle g, \phi_{P_l} \rangle \phi_{P_l}(x) \,,$$

where k_ξ is the largest k such that the dyadic interval of length 2^{1-k} about ξ_T contains ξ. This sum is zero when $\xi \geq \sup \omega_{P_u^{\min}(x)}$ and when $\xi < \inf \omega_{P_u^{\max}(x)}$ where $P^{\min}(x)$ and $P^{\max}(x)$ are, respectively, the smallest and largest bitiles from T which contain x in their time supports. For intermediate ξ it follows from l-convexity that there is a $P \in T$ with $|I_P| = 2^{k_\xi}$ and $x \in I_P$ and hence the right side of (5.5)

$$= \text{sgn}(\phi_{p_T}) D_{k_\xi} \left[\text{sgn}(\phi_{p_T}) g \right] (x).$$

From the monotonicity of the k_ξ it thus follows that (5.2) is

$$\leq C \left\| D_k \left[\text{sgn}(\phi_{p_T}) g \right] (x) \right\|_{L^p_x(V^r_k)} \leq C_{p,r} \|g\|_{L^p} \leq C_{p,r} \, \text{size}(\overline{T}) \, |I_{\overline{T}}|^{1/p}. \qquad \square$$

6. Pointwise variation for stacks of trees

Given a collection of trees $\mathbf{T}$ satisfying certain assumptions, the following lemma allows us to partition $\mathbb{R}^+$ into a collection of intervals $\{v_T\}_{T \in \mathbf{T}}$ such that the restriction of a function of the form

$$(6.1) \qquad\qquad \sum_{P \in \bigcup_{T \in \mathbf{T}} T} c_P \, 1_{\omega_{P_u}}(\cdot)$$

to the interval v_T is $\sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\cdot)$. This partitioning will be used in the current section to obtain variation-norm estimates for functions of the form (6.1) and it will be used in Section 7 to obtain estimates involving the Walsh multiplier operators induced by functions of the form (6.1).

Lemma 6.1. *Suppose that $x \in \mathbb{R}^+$ and that $\mathbf{P}$ is a finite collection of bitiles with $x \in I_P$ for each $P \in \mathbf{P}$ and with $\mathbf{P} = \bigcup_{T \in \mathbf{T}} T$ where $\mathbf{T}$ satisfies one of the following two conditions:*

- *$\mathbf{T}$ is a collection of u-overlapping trees which are td-maximal among u-overlapping trees contained in $\mathbf{P}$.*
- *$\mathbf{T}$ is a collection of properly sorted l-overlapping trees.*

Then there is a collection $\{\upsilon_T\}_{T \in \mathbf{T}}$ of pairwise disjoint intervals covering $\mathbb{R}^+$ such that for each $T \in \mathbf{T}$ and $\xi \in \upsilon_T$

$$(6.2) \qquad \{P \in \mathbf{P} : \xi \in \omega_{P_u}\} \subset T.$$

Proof. We start by proving the lemma under assumption of the first condition. Without loss of generality assume that $|\mathbf{T}| \geq 2$, and that for every $T \in \mathbf{T}$

$$T \not\subset \bigcup_{T' \in \mathbf{T} \setminus \{T\}} T'.$$

By td-maximality we see that $\xi_T \neq \xi_{T'}$ for $T \neq T'$. Enumerate the trees $T_1, \ldots, T_N$ so that $\xi_{T_i} < \xi_{T_{i+1}}$ for $i = 1, \ldots, N-1$.

For each i let P^i be the minimal bitile in $T_i \setminus \bigcup_{i' < i} T_{i'}$, and let $\xi_i^- := \inf(\omega_{P_u^i})$. We claim that $\xi_i^- > \xi_{i'}^-$ whenever $i > i'$. To see this, first note that we may assume that $\omega_{P_u^i} \cap \omega_{P_u^{i'}} \neq \emptyset$ or else the conclusion would follow from the fact that $\xi_{T_i} > \xi_{T_{i'}}$. Since $x \in I_{P^{i'}} \cap I_{P^i}$ we then have $P_u^{i'} \cap P_u^i \neq \emptyset$. Thus, we must have $P_u^i > P_u^{i'}$ since $P_u^i \leq P_u^{i'}$ would imply that $P^i \in T_{i'}$ contradicting the definition of P^i. Then $\omega_{P^i} \subsetneq \omega_{P^{i'}}$ and so $\xi_i^- > \xi_{i'}^-$ as desired.

Set $\upsilon_{T_1} = [0, \xi_2^-), \upsilon_{T_N} = [\xi_N^-, \infty)$, and $\upsilon_{T_i} = [\xi_i^-, \xi_{i+1}^-)$ for $1 < i < N$. From the previous paragraph, we see that $\upsilon_{T_1}, \ldots, \upsilon_{T_N}$ are disjoint and cover $\mathbb{R}^+$.

It remains to check that if $\xi \in \upsilon_{T_i}$ and $P \in \mathbf{P}$ with $\xi \in \omega_{P_u}$ then $P \in T_i$. First assume $1 < i < N$. Choose the minimal i' such that $P \in T_{i'}$. First suppose $i' < i$. Since $\xi \in \omega_{P_u}$, $\xi \geq \inf(\omega_{P^i})$, and $\xi_{T_{i'}} < \xi_{T_i}$, we must have $\omega_{P_u} \cap \omega_{P^i} \neq \emptyset$. The fact that $P^i \notin T_{i'}$ rules out the possibility that $P_u^i \leq P_u$ and so we must have $P_u^i > P_u$. But then $P \in T_i$ as desired. Now suppose $i' > i$. By minimality of i' we have $P \in T_{i'} \setminus \bigcup_{i'' < i'} T_{i''}$. Then, by minimality of $P^{i'}$ we have $P \geq P^{i'}$ and so $\inf(\omega_{P_u}) \geq \xi_{i'}^-$ contradicting the fact that $\xi \in \omega_{P_u}$. The appropriate halves of this argument work when $i = 1$ or N.

Working instead under the second condition, for each $T \in \mathbf{T}$ we let

$$\tilde{J}_T = \bigcup_{\substack{P \in T \\ x \in I_P}} \omega_{P_u}.$$

By (3) each $\tilde{J}_T$ is an interval. From (5) we know that the tiles P_u with $P \in \bigcup_{T \in \mathbf{T}} T$ are disjoint, which gives (6.2) for $\xi \in \tilde{J}_T$. Combining (5) with (4) one sees that the intervals $\tilde{J}_T$ are pairwise disjoint. Finally, the left and right sides of (6.2) are both zero for ξ outside of $\bigcup_{T \in \mathbf{T}} \tilde{J}_T$; thus by choosing $\{J_T\}_{T \in \mathbf{T}}$ to be any collection of pairwise disjoint intervals which cover $\mathbb{R}^+$ and which satisfy $\tilde{J}_T \subset J_T$, we are finished.   □

The following corollary, which can be used to obtain (1.4), follows immediately from the lemma above.

Corollary 6.2. *Suppose that* $\mathbf{P}, \mathbf{T}$ *and* x *satisfy the hypotheses of Lemma* 6.1. *Then, for any collection of coefficients* $\{c_P\}_{P \in \mathbf{P}} \subset \mathbb{R}$,

$$\big\| \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}} \big\|_{V^r} \le C \, |\mathbf{T}|^{1/r} \sup_{T \in \mathbf{T}} \big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \big\|_{V^r}.$$

To prove Theorem 1.3 we will need the estimate below.

Corollary 6.3. *Suppose that* $\mathbf{P}, \mathbf{T}$ *and* x *satisfy the hypotheses of Lemma* 6.1. *Then, for any collection of coefficients* $\{c_P\}_{P \in \mathbf{P}} \subset \mathbb{R}$,

$$\Big\| \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \Big\|_{\ell_k^\infty(V_\xi^r)} \le C \, |\mathbf{T}|^{1/r} \sup_{T \in \mathbf{T}} \big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \big\|_{V^r}.$$

Proof. Thanks to the covering in Lemma 6.1, it suffices to observe that, for each k,

$$(6.3) \qquad \Big\| \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \Big\|_{V_\xi^r} \le C \, \big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi) \big\|_{V_\xi^r}.$$

Let ω_k denote the dyadic interval of length 2^{-k} containing ξ_T. First treating the case where T is u-overlapping, we note that for $\xi \notin \omega_k$ we have

$$(6.4) \qquad \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi),$$

and for $\xi \in \omega_k$ we have

$$\sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi'),$$

where ξ' is any point in $\omega_{k-1} \setminus \omega_k$. Combining these two facts immediately implies (6.3).

If T is instead l-overlapping then for $\xi \notin \omega_{k-1}$ we have (6.4) and for $\xi \in \omega_{k-1}$

$$(6.5) \qquad \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = 0,$$

and hence (6.3). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Finally, Theorem 1.4 is obtained from the following bound:

Corollary 6.4. *Suppose that* $\mathbf{P}, \mathbf{T}$ *and* x *satisfy the hypotheses of Lemma* 6.1. *Then for any collection of coefficients* $\{c_P\}_{P \in \mathbf{P}} \subset \mathbb{R}$,

$$\Big\| \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \Big\|_{L_\xi^\infty(V_k^r)} \le C \sup_{T \in \mathbf{T}} \big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \big\|_{V^r}.$$

Proof. Again using the covering in Lemma 6.1, it suffices to observe that, for each ξ and T,

$$(6.6) \qquad \Big\| \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \Big\|_{V_k^r} \leq C \Big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi') \Big\|_{V_{\xi'}^r}.$$

First suppose that T is u-overlapping and let $k_\xi = \sup\{k' : \xi \in \omega_{k'}\}$ (here ω_k is as defined in Corollary 6.3). Then for $k \leq k_\xi + 1$,

$$\sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P,$$

and for $k > k_\xi + 1$,

$$\sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{\substack{P \in T \\ |I_P| \leq 2^{k_\xi}}} c_P.$$

Thus

$$\Big\| \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \Big\|_{V_k^r} \leq \Big\| \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \Big\|_{V_k^r}.$$

For each integer k, let ξ_k be the left endpoint of the largest dyadic interval ω such that ξ_T is in the right half of ω and $|\omega| \leq 2^{-k+1}$ (if no such dyadic interval exists, let $\xi_k = \xi_T$). Then the points ξ_k are monotonic in k and

$$\sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P = \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi_k).$$

Thus

$$(6.7) \qquad \Big\| \sum_{\substack{P \in T \\ |I_P| < 2^k}} c_P \Big\|_{V_k^r} \leq \Big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi') \Big\|_{V_{\xi'}^r},$$

as desired.

For l-overlapping T we have (6.4) if $k > \max\{k' : \xi \in \omega_{k'-1}\}$ and (6.5) otherwise; hence we obtain (6.6). $\square$

7. Pointwise maximal multiplier estimates for stacks of trees

Suppose that $\mathbf{P}, \mathbf{T}, x$, and $\{c_P\}_{P \in \mathbf{P}}$ are as in the hypotheses of Lemma 6.1. It follows from Lemmas 6.1 and 3.5 that if $1 < q < \infty$ and $|1/q - 1/2| < 1/r$, then

$$\Big\| \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}} \Big\|_{M^q} \leq C_{q,r,\epsilon} |\mathbf{T}|^{|1/q - 1/2| + \epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \Big\|_{V^r}.$$

The aim of the present section is to extend this M^q bound to an $M^{q,*}$ bound through the use of Lemma 3.6.

Lemma 7.1. *Suppose that* $\mathbf{P}, \mathbf{T}, x$, *and* $\{c_P\}_{P \in \mathbf{P}}$ *are as above and that* $1 < q \leq 2$, $2 < r < 2q$, *and* $\epsilon > 0$. *Then*

$$\Big\| \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}} \Big\|_{M^{q,*}} \leq C_{q,r,\epsilon} \, |\mathbf{T}|^{1/q - 1/r + \epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \Big\|_{V^r}.$$

Proof. We start by assuming the first condition in Lemma 6.1, i.e., that $\mathbf{P} = \bigcup_{T \in \mathbf{T}} T$ where $\mathbf{T}$ is a collection of u-overlapping trees which are td-maximal among all trees contained in $\mathbf{P}$. Without loss of generality we may also assume that $x \in I_T$ for each $T \in \mathbf{T}$. Let $\Xi = \{\xi_T : T \in \mathbf{T}\}$.

If $P \in T$ and $|I_P| \geq 2^k$, then we have ω_{P_u} contained in the dyadic interval of length 2^{-k} about ξ_T. This implies that for

$$\xi \in \mathbb{R}^+ \setminus \bigcup_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} \omega$$

we have

$$(7.1) \qquad \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}}(\xi).$$

If ω is any dyadic interval of length 2^{-k} and $\xi \in \omega$ then

$$(7.2) \qquad \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P.$$

Combining (7.1) and (7.2) we see that

$$(7.3) \qquad \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P \, 1_{\omega_{P_u}}(\xi)$$

$$= \Big(1 - \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi) \Big) \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}}(\xi) \ + \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi) \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P.$$

The right side of (7.3) is the sum of three terms each of which we will bound separately. For the first term, we argue as indicated in the discussion at the beginning of this section. Specifically, applying Lemma 6.1, we obtain a collection of intervals $\{v_T\}_{T \in \mathbf{T}}$ so that

$$(7.4) \qquad \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{P \in T} c_P \, 1_{\omega_{P_u}}(\xi)$$

for $\xi \in v_T$. We then apply Lemma 3.5 with the collection of intervals $\Upsilon = \{v_T\}_{T \in \mathbf{T}}$ to obtain

$$\Big\| \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}} \Big\|_{M^q} \leq C_{q,r,\epsilon} \, |\mathbf{T}|^{1/q - 1/r + \epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \Big\|_{V^r}.$$

For the second term, we note that when Ξ is as above and $a_\omega = 1$ for each dyadic interval ω,

$$(7.5) \qquad \sum_{\substack{|\omega|=2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega \sum_{P \in \mathbf{P}} c_P 1_{\omega_{P_u}} = \mathcal{D}_k \sum_{P \in \mathbf{P}} c_P 1_{\omega_{P_u}}.$$

Combining (7.5) and (7.4), we see that Lemma 3.6 gives

$$\Big\| \sum_{\substack{|\omega|=2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega \sum_{P \in \mathbf{P}} c_P 1_{\omega_{P_u}} \Big\|_{M^{q,*}}$$

$$\leq C_{q,r,\epsilon} \left(|\Xi| + |\{v_T\}_{T \in \mathbf{T}}| \right)^{1/q-1/r+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P 1_{\omega_{P_u}} \Big\|_{V^r}$$

$$\leq C_{q,r,\epsilon} |\mathbf{T}|^{1/q-1/r+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P 1_{\omega_{P_u}} \Big\|_{V^r}.$$

Finally, for the last term we note that with Ξ as above and

$$(7.6) \qquad a_\omega = \sum_{P \in \mathbf{P},\, \omega \subsetneq \omega_{P_u}} c_P$$

we have

$$\sum_{\substack{|\omega|=2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P = \mathcal{D}_k$$

and so, applying Lemma 3.2,

$$\Big\| \sum_{\substack{|\omega|=2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P \Big\|_{M^{q,*}}$$

$$(7.7) \qquad \leq C_{q,r,\epsilon} |\mathbf{T}|^{1/q-1/r+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{|\omega|=2^k} 1_\omega(\xi_T) \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P \Big\|_{V_k^r}.$$

For each k,

$$\sum_{|\omega|=2^k} 1_\omega(\xi_T) \sum_{\substack{P \in \mathbf{P} \\ \omega \subsetneq \omega_{P_u}}} c_P = \sum_{\substack{P \in \mathbf{P} \\ |I_P|<2^k}} c_P 1_{\omega_{P_u}}(\xi_T) = \sum_{\substack{P \in T' \\ |I_P|<2^k}} c_P$$

where $T' \in \mathbf{T}$ is the tree containing the maximal element of $\mathbf{P}$ satisfying $\xi_T \in \omega_{P_u}$, $x \in I_P$, and the last identity follows from td-maximality of the tree T'. Therefore, the right side of (7.7) is

$$\leq C_{q,r,\epsilon} |\mathbf{T}|^{1/q-1/r+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{\substack{P \in T \\ |I_P|<2^k}} c_P \Big\|_{V_k^r}$$

$$\leq C_{q,r,\epsilon} |\mathbf{T}|^{1/q-1/r+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P 1_{\omega_{P_u}}(\xi) \Big\|_{V_\xi^r},$$

where the second inequality follows from (6.7). Thus, we obtain the proof of the lemma under the assumption of the first condition.

We now assume the second condition, i.e., that $\mathbf{T}$ is a properly sorted collection of l-overlapping trees. Again, assume that $x \in I_T$ for each $T \in \mathbf{T}$. Consider

$$\sum_{\substack{P \in \mathbf{P} \\ |I_P| > 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) \, .$$

If $P \in T$ and $I_P > 2^k$ then ω_P is contained in the dyadic interval of length 2^{-k} containing ξ_T and so, letting $\Xi = \{\xi_T : T \in \mathbf{T}\}$ we have

$$1_{\omega_{P_u}}(\xi) = 1_{\omega_{P_u}}(\xi) \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi) \, ,$$

where we sum over dyadic intervals ω. If $P \in T$ with $|I_P| \leq 2^k$ then ω_{P_u} does not intersect the dyadic interval of length 2^{-k} about ξ_T and, furthermore, by (5) does not intersect the dyadic interval of length 2^{-k} about any $\xi_{T'} > \xi_T$ for $T' \in \mathbf{T}$. This gives

$$1_{\omega_{P_u}}(\xi) \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi) = 0 \, ,$$

and hence

$$\sum_{\substack{P \in \mathbf{P} \\ |I_P| > 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi) \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}}(\xi).$$

Thus,

$$\sum_{\substack{P \in \mathbf{P} \\ |I_P| \leq 2^k}} c_P \, 1_{\omega_{P_u}}(\xi) = \left(1 - \sum_{\substack{|\omega| = 2^{-k} \\ \omega \cap \Xi \neq \emptyset}} 1_\omega(\xi)\right) \sum_{P \in \mathbf{P}} c_P \, 1_{\omega_{P_u}}(\xi) \, ,$$

and the remaining argument follows exactly that for the first two terms in (7.3).   □

8. Proof of theorems

Theorem 1.1 is established by using Lemma 7.1 to apply the following proposition with $\| \cdot \|_{\mathcal{N}} = \| \cdot \|_{M^{q,*}}$, $\eta_{P,k} = 1_{(-\infty, 2^k)}(|I_P|)$, and r sufficiently close to 2. Using Corollary 6.2 to apply the proposition with $\| \cdot \|_{\mathcal{N}_{\xi,k}} = \| \cdot \|_{\ell_k^\infty(V_\xi^r)}$ and $\eta_{P,k} = 1$ establishes (1.4). Using Corollary 6.3 to apply the proposition with $\| \cdot \|_{\mathcal{N}_{\xi,k}} = \| \cdot \|_{\ell_k^\infty(V_\xi^r)}$ and $\eta_{P,k} = 1_{(-\infty, 2^k)}(|I_P|)$ establishes Theorem 1.3. Using Corollary 6.4 to apply the proposition with $\| \cdot \|_{\mathcal{N}_{\xi,k}} = \| \cdot \|_{L_\xi^\infty(V_k^r)}$ and $\eta_{P,k} = 1_{(-\infty, 2^k)}(|I_P|)$ establishes Theorem 1.4.

Proposition 8.1. *Let $r > 2$, $1 < p < \infty$, and let $\{\eta_{P,k}\}_{P \in \mathbf{P}_0, k \in \mathbb{Z}} \subset \mathbb{R}$ be a collection of coefficients. Suppose that for all $\mathbf{P}, \mathbf{T}, x$, and $\{c_P\}_{P \in \mathbf{P}}$ as in the hypotheses of Corollary 6.2, a norm $\| \cdot \|_{\mathcal{N}}$ acting on functions defined on $\mathbb{R}^+ \times \mathbb{Z}$ satisfies*

$$(8.1) \qquad \big\| \sum_{P \in \mathbf{P}} \eta_{P,k} c_P \, 1_{\omega_{P_u}} \big\|_{\mathcal{N}} \leq C \, |\mathbf{T}|^\alpha \sup_{T \in \mathbf{T}} \big\| \sum_{P \in T} c_P \, 1_{\omega_{P_u}} \big\|_{V^r}$$

for some $\alpha < \min(1 - 1/p, 1/2)$. Then,

$$\big\| \sum_{P \in \mathbf{P}_0} \eta_{P,k} \, \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{L^p_x(\mathcal{N}_{\xi,k})} \leq C_{p,r} \, \|f\|_{L^p}$$

Proof. We will prove a restricted weak-type estimate; the full result follows by interpolation. Specifically, we suppose that $|f| \leq 1_F$ and $\lambda > 0$ and want to show

$$\Big| \big\{ x : \big\| \sum_{P \in \mathbf{P}_0} \eta_{P,k} \, \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{\mathcal{N}_{\xi,k}} > \lambda \big\} \Big| \leq C_{p,r} \, |F|/\lambda^p.$$

The inequality above will be demonstrated by covering the set on the left side by "exceptional sets" E_1, E_2^u, E_2^l, E_3 of acceptably small measure.

We begin by treating the case $\lambda < 1$. There we set

$$E_1 := \{ M[1_F] \geq c\lambda^p \},$$

where M is the dyadic Hardy–Littlewood maximal operator. By the weak-type $(1,1)$ estimate for M we have $|E_1| \leq C|F|/\lambda^p$. Since we only need to bound the $\mathcal{N}$-norm for $x \notin E_1$, we can assume that for every $P \in \mathbf{P}_0$ we have $I \not\subset E_1$ and hence, since the L^∞ norm of the phase-space projection onto any convex T can be controlled by the L^∞ norm of the projection onto a subtile of an element of T,[4]

$$\operatorname{size}(\mathbf{P}_0, f) \leq C \, \lambda^p.$$

For each $n \geq 0$ we apply Lemma 4.1 with $\mathbf{P} = \mathbf{P}_n$ and set $\mathbf{P}_{n+1} := \mathbf{P}'$ so that

$$\operatorname{size}(\mathbf{P}_{n+1}, f) \leq C \, 2^{-(n+1)} \lambda^p.$$

For each n, we apply Lemma 4.2 to the collection of bitiles $\mathbf{P}_n \setminus \mathbf{P}_{n+1}$ and obtain collections of u-overlapping trees $\mathbf{T}_n^u$ and l-overlapping trees $\mathbf{T}_n^l$.

Fix $s > 0$ large and $\epsilon > 0$ small with magnitudes to be determined later and let

$$\gamma_n = c \, 2^{-n} \lambda^p \, (2^{2n} \lambda^{-p})^{1/s} \, 2^{\epsilon n}.$$

---

[4]If T was allowed to be, say, an arbitrary u-overlapping tree, then the phase-plane projection onto T could exhibit singular-integral type behavior. In that situation, in order to obtain the analogous size estimate, it would be necessary to consider a Calderón–Zygmund decomposition of f, or perhaps even use a John–Nirenberg type estimate, depending on context. Here, since we are assuming that T is convex, there is cancellation between scales and so the phase-plane projection is better behaved, and is in particular pointwise dominated by a maximal average over appropriate scales.

Each $T \in \mathbf{T}_n^u$ is contained, by construction, in a convex tree $\overline{T}$ with size$(\overline{T}, f) \leq C 2^{-n} \lambda^p$ and $I_{\overline{T}} = I_T$. Also, we have

$$(8.2) \qquad \big\| \sum_{T \in \mathbf{T}_n^u} 1_{I_T} \big\|_{L^1} \leq C \, 2^{2n} \, \lambda^{-2p} \, |F|.$$

Thus, letting

$$E_2^u := \bigcup_{n \geq 0} \bigcup_{T \in \mathbf{T}_n^u} \big\{ x : \big\| \sum_{P \in T} \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{V_\xi^r} > \gamma_n \big\},$$

we apply Lemma 5.1 to obtain

$$|E_2^u| \leq \sum_{n \geq 0} \sum_{T \in \mathbf{T}_n^u} \gamma_n^{-s} \big\| \sum_{P \in T} \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{L_x^s(V_\xi^r)}^s$$

$$\leq C \sum_{n \geq 0} \sum_{T \in \mathbf{T}_n^u} \gamma_n^{-s} \, (2^{-n} \lambda^p)^s \, |I_T| \leq C \sum_{n \geq 0} \gamma_n^{-s} \, (2^{-n} \lambda^p)^s \, 2^{2n} \, \lambda^{-2p} \, |F|$$

$$\leq C \sum_{n \geq 0} 2^{-s\epsilon n} \, |F|/\lambda^p \leq C |F|/\lambda^p.$$

Defining E_2^l analogously, we obtain the same bound.

Let

$$\beta_n = c \, 2^{2n} \, \lambda^{-p} \, 2^{\epsilon n}$$

and

$$E_3 := \bigcup_{n \geq 0} \big\{ x : \sum_{T \in \mathbf{T}_n^u} 1_{I_T}(x) > \beta_n \big\}.$$

Applying (8.2), we have

$$|E_3| \leq C \sum_{n \geq 0} \beta_n^{-1} \, 2^{2n} \, \lambda^{-2p} \, |F| \leq C \sum_{n \geq 0} 2^{-\epsilon n} \, |F|/\lambda^p \leq C \, |F|/\lambda^p.$$

Recall that the trees in $\mathbf{T}_n^u$ and $\mathbf{T}_n^l$ have shared top data and so E_3 also gives control over $T \in \mathbf{T}_n^l$.

Fix $x \notin E_1 \cup E_2^u \cup E_2^l \cup E_3$; we need to show that

$$\big\| \sum_{P \in \mathbf{P}_0} \eta_{P,k} \, \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{\mathcal{N}_{\xi,k}} \leq \lambda.$$

Since every $P \in \mathbf{P}_0$ with $\langle f, \phi_{P_l} \rangle \neq 0$ is in $\mathbf{P}_n \setminus \mathbf{P}_{n+1}$ for some n, the left side above is

$$(8.3) \qquad \leq \sum_{n \geq 0} \big\| \sum_{P \in \mathbf{P}_n \setminus \mathbf{P}_{n+1}} \eta_{P,k} \, \langle f, \phi_{P_l} \rangle \, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \big\|_{\mathcal{N}_{\xi,k}}.$$

For each n, we have

$$\mathbf{P}_n \setminus \mathbf{P}_{n+1} = \Big(\bigcup_{T \in \mathbf{T}_n^u} T \Big) \cup \Big(\bigcup_{T \in \mathbf{T}_n^l} T \Big),$$

and so by (9) we see that the nth term in (8.3) is

$$(8.4) \quad \leq \Big\| \sum_{P \in \bigcup_{T \in \mathbf{T}_n^u} T} \eta_{P,k} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{\mathcal{N}_{\xi,k}}$$

$$+ \Big\| \sum_{P \in \bigcup_{T \in \mathbf{T}_n^l} T} \eta_{P,k} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{\mathcal{N}_{\xi,k}}.$$

Letting $\tilde{\mathbf{T}}_n^u = \{(T \cap \{P \in \mathbf{P}_0 : x \in I_P\}, \xi_T, I_T) : T \in T_n^u, x \in I_T\}$ and similarly for $\tilde{\mathbf{T}}_n^l$, the expression above is clearly

$$(8.5) \quad = \Big\| \sum_{P \in \bigcup_{T \in \tilde{\mathbf{T}}_n^u} T} \eta_{P,k} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{\mathcal{N}_{\xi,k}}$$

$$+ \Big\| \sum_{P \in \bigcup_{T \in \tilde{\mathbf{T}}_n^l} T} \eta_{P,k} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{\mathcal{N}_{\xi,k}}.$$

Noting that $\mathbf{T}_n^l$ is still properly sorted, and that the $\mathbf{T}_n^u$ are still td-maximal among u-overlapping trees contained in $\bigcup_{T \in \tilde{\mathbf{T}}_n^u} T$, we may apply (8.1) with $c_P = \langle f, \phi_{P_l} \rangle \phi_{P_l}(x)$ to see that the display above is

$$\leq C \left(|\tilde{\mathbf{T}}_n^l| + |\tilde{\mathbf{T}}_n^u| \right)^\alpha \sup_{T \in \tilde{\mathbf{T}}_n^u \cup \tilde{\mathbf{T}}_n^l} \Big\| \sum_{P \in T} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{V_\xi^r}$$

$$\leq C \beta_n^\alpha \gamma_n \leq c \, 2^{-n(1-2\alpha)} \, 2^{n(2/s + (\alpha+1)\epsilon)} \, \lambda^{p(1-\alpha)} \, \lambda^{-p/s}.$$

Since $\alpha < \min(1/2, 1 - 1/p)$, we may choose ϵ sufficiently small and s sufficiently large so that the right side above is $\leq c \, 2^{-\tilde{\epsilon} n} \lambda$ for some $\tilde{\epsilon} > 0$, and hence, summing over n and choosing c sufficiently small, we obtain the desired bound for $\lambda < 1$.

In the case that $\lambda \geq 1$ we set $E_1 = \emptyset$ and use the bound size$(\mathbf{P}_0, f) \leq C$. We decompose $\mathbf{P}_0$ as in the case $\lambda < 1$ so that size$(\mathbf{P}_n, f) \leq C 2^{-n}$. Letting

$$\gamma_n = c \, 2^{-n} \, (2^{2n} \lambda^p)^{1/s} \, 2^{\epsilon n},$$

we define E_2^u and E_2^l as above and obtain $|E_2^u|, |E_2^l| \leq C \, |F|/\lambda^p$.

Interpolating the bounds

$$\Big\| \sum_{T \in \mathbf{T}_n^u} 1_{I_T} \Big\|_{L^1} \leq C \, 2^{2n} \, |F| \quad \text{and} \quad \Big\| \sum_{T \in \mathbf{T}_n^u} 1_{I_T} \Big\|_{\text{BMO}} \leq C \, 2^{2n},$$

we see that

$$\Big\| \sum_{T \in \mathbf{T}_n^u} 1_{I_T} \Big\|_{L^t} \leq C \, 2^{2n} \, |F|^{1/t},$$

where $t < \infty$ is some fixed exponent which will be chosen sufficiently large in a manner to be determined. Then, for each β,

$$\Big| \Big\{ \sum_{T \in \mathbf{T}_n^u} 1_{I_T} > \beta \Big\} \Big| \leq C \, \beta^{-t} \, 2^{2tn} \, |F|,$$

so, letting
$$\beta_n = c\, 2^{2n}\, \lambda^{p/t}\, 2^{\epsilon n},$$
we define E_3 as above and have $|E_3| \le C|F|/\lambda^p$.

For $x \notin E_2^u \cup E_2^l \cup E_3$ we thus have
$$\Big\| \sum_{P \in \mathbf{P}_n \setminus \mathbf{P}_{n+1}} \eta_{P,k}\, \langle f, \phi_{P_l} \rangle\, \phi_{P_l}(x) 1_{\omega_{P_u}}(\xi) \Big\|_{\mathcal{N}_{\xi,k}}$$
$$\le C\, \beta_n^\alpha\, \gamma_n \le c\, 2^{-n(1-2\alpha)}\, 2^{n(2/s+(\alpha+1)\epsilon)}\, \lambda^{p(\alpha/t+1/s)}.$$

Summing over n, this is less than or equal to λ provided that s and t are chosen sufficiently large, and ϵ and c are chosen sufficiently small. $\qquad\square$

9. Variation-norm estimates for multipliers

The following is an s-variation-norm analog of Lemma 7.1. By taking r sufficiently close (depending on p q and s) to 2, it implies Theorem 1.2 through the use of Proposition 8.1.

Lemma 9.1. *Suppose that $\mathbf{P}, \mathbf{T}, x$, and $\{c_P\}_{P \in \mathbf{P}}$ are as in Lemma 7.1 and that $1 < q \le 2$, $2 < r < 2q$, $\epsilon > 0$ and $r < s$. Then*
$$\Big\| \sum_{\substack{P \in \mathbf{P} \\ |I_P| < 2^k}} c_P 1_{\omega_{P_u}} \Big\|_{M^{q,s}} \le C_{q,r,s,\epsilon} |\mathbf{T}|^{(1/2-1/r)\frac{s}{s-2}+1/q-1/2+\epsilon} \sup_{T \in \mathbf{T}} \Big\| \sum_{P \in T} c_P 1_{\omega_{P_u}} \Big\|_{V^r}.$$

Except for Lemma 3.1 (which is a key element in the proof of Lemma 3.7), each step in the proof of Lemma 7.1 is insensitive to the difference between the $M^{q,s}$ and the $M^{q,*}$ norms. Thus, to establish Lemma 9.1 it suffices to prove the following variation-norm extension of Lemma 3.1.

Lemma 9.2. *Let $r > 2, \epsilon > 0$, and $\Xi \subset \mathbb{R}^+$. Then*
$$\|\mathcal{D}_k\|_{M^{2,s}} \le C_{r,s,\epsilon}\, |\Xi|^{(1/2-1/r)\frac{s}{s-2}+\epsilon} \sup_{\xi \in \Xi} \Big\| \sum_{|\omega|=2^k} a_\omega 1_\omega(\xi) \Big\|_{V_k^r}.$$

To prove Lemma 9.2, one follows the method used to prove Lemma 3.1 in [6] with some refinements which we will now elaborate. The main advance needed is the following variation-norm version of Proposition 4.2 from [6].

Proposition 9.3. *Let H be a Hilbert space, let A be a finite measure space, and let $2 < r < s$ and $\delta > 0$. Suppose that we are given a function g from A to H such that for each $a \in A$, $|g(a)| \le \delta$, and such that for each $h \in H$,*

(9.1)
$$\| \langle g(a), h \rangle \|_{L^2(A)} \le |h|.$$

Then, for each sequence $\{c_k\}_{k \in \mathbb{Z}}$ of points in H,

(9.2)
$$\| \langle g(a), c_k \rangle \|_{L_a^2(V_k^s)} \le C_{r,s}\, (\delta^2 |A|)^{(1/2-1/r)\frac{s}{s-2}}\, \|c_k\|_{V^r}.$$

The proof of Proposition 9.3 uses the same method as that of Lemma 3.2 in [10]. However, since the statement is more general here, we will repeat the argument.

Proof. Let

$$\| \cdot \|_{\tilde{V}_s} = \| \cdot \|_{V_s} - \| \cdot \|_{L^\infty}.$$

By Proposition 4.2 of [6], it suffices to prove (9.2) with the $\tilde{V}^s$ norm in place of the V^s norm. By a limiting argument, we may also assume that our sequence $\{c_k\}_{k=1}^M$ has finite length, provided that $C_{r,s}$ is independent of M.

For each $\lambda > 0$ we cover $\{c_k\}_{k=1}^M$ with respect to λ-jumps as follows. Set $l(\lambda, 1) = 1$. Suppose that $l(\lambda, 1) < \cdots < l(\lambda, L)$ have been chosen, and let $B(c_{l(\lambda,L)}, \lambda)$ denote the ball of radius λ centered at $c_{l(\lambda,L)}$. If $\{c_k : k > l(\lambda, L)\} \subset B(c_{l(\lambda,L)}, \lambda)$ then stop and set $L_\lambda = L$ and $l(\lambda, L+1) = \infty$. Otherwise, let $l(\lambda, L+1)$ be chosen minimally with $l(\lambda, L + 1) > l(\lambda, L)$ and $c_{l(\lambda,L+1)} \notin B(c_{l(\lambda,L)}, \lambda)$. This process will stop, yielding some $L_\lambda \leq M$. It is clear that

(9.3) $$\lambda(L_\lambda - 1)^{1/r} \leq \|c_k\|_{V^r}.$$

We now define a recursive "parent" function based on the covering above. Fix some $\lambda_0 < \min\{|c - c'|: c, c' \in \{c_k\}_{k=1}^M$ and $c \neq c'\}$. For $k = 1, \ldots, M$ define $\rho(-1, k) = k$. Once $\rho(n, k)$ has been defined for $n = -1, \ldots, L$ set $\rho(L + 1, k) = l(2^{L+1}\lambda_0, m)$, where m is the unique integer satisfying

$$l(2^{L+1}\lambda_0, m) \leq \rho(L, k) < l(2^{L+1}\lambda_0, m + 1).$$

Notice that we have

$$|c_{\rho(n,k)} - c_{\rho(n+1,k)}| < 2^{n+1}\lambda_0,$$

and in particular $c_{\rho(0,k)} = c_k$. Also note that $\rho(n, k) = 1$ whenever $2^n \lambda_0 \geq \text{diameter}(\{c_k\}_{k=1}^M)$. Thus

$$c_k = c_1 + \sum_{n=0}^{\infty} c_{\rho(n,k)} - c_{\rho(n+1,k)}.$$

Finally, by induction, one sees that $\rho(n, k)$ is nondecreasing in k for each fixed n.

We have

$$\| \langle g(a), c_k \rangle \|_{L_a^2(\tilde{V}_k^s)} \leq \sum_{n=0}^{\infty} \| \langle g(a), (c_{\rho(n,k)} - c_{\rho(n+1,k)}) \rangle \|_{L_a^2(\tilde{V}_k^s)}.$$

Observe that the right-hand side above

$$= \sum_{n:L_{2^n \lambda_0} > 1} \| \langle g(a), c_{\rho(n,k)} - c_{\rho(n+1,k)}) \rangle \|_{L_a^2(\tilde{V}_k^s)}.$$

Using the monotonicity of the $\rho(n, \cdot)$ and the fact that the range of $\rho(n, \cdot)$ is contained in $\{l(2^n\lambda_0, m) : m = 1, \ldots, L_{2^n\lambda_0}\}$ we see that the expression above is

$$\leq 2 \sum_{n:L_{2^n \lambda_0} > 1} \left\| \left(\sum_{m=1}^{L_{2^n\lambda_0}} |\langle g(a), (c_{l(2^n\lambda_0, m)} - c_{\tilde{\rho}(n+1,l(2^n\lambda_0, m)))} \rangle |^s \right)^{1/s} \right\|_{L_a^2},$$

where we let $\tilde{\rho}(n+1, l(2^n\lambda_0, m))$ denote $l(2^{n+1}\lambda_0, i)$, where i is the unique integer satisfying

$$l(2^{n+1}\lambda_0, i) \leq l(2^n\lambda_0, m) < l(2^{n+1}\lambda_0, i+1).$$

Estimating ℓ^s by ℓ^2, switching the order of integration, and using (9.1), we see that the nth term in the outer sum above is

$$\leq C\, 2^n\, \lambda_0\, L_{2^n\lambda_0}^{1/2} \leq C\, (2^n\lambda_0)^{1-r/2}\, \|c_k\|_{V^r}^{r/2}.$$

We can also estimate the nth term by

$$\Big\| \Big(\sum_{m=1}^{L_{2^n\lambda_0}} (\delta|c_{l(2^n\lambda_0, m)} - c_{\tilde{\rho}(n+1, (2^n\lambda_0, m))}|)^s \Big)^{1/s} \Big\|_{L_a^2}$$

$$\leq \delta\, |A|^{1/2} \Big(\sum_{m=1}^{L_{2^n\lambda_0}} \big|c_{l(2^n\lambda_0, m)} - c_{\tilde{\rho}(n+1, l(2^n\lambda_0, m))}\big|^s \Big)^{1/s}$$

$$\leq \delta\, |A|^{1/2}\, (2^n\lambda_0)^{1-r/s}\, \|c_k\|_{V^r}^{r/s}.$$

Choosing whichever of the two bounds is favorable for each n and summing gives the desired result. $\qquad\square$

Through the averaging argument in the proof of Corollary 4.3 from [6], one sees that Proposition 9.3 implies the following two corollaries.

Corollary 9.4. *Let* $r > 2, 1 < q \leq 2$, *and* $\Xi \subset \mathbb{R}^+$. *If no two elements of* Ξ *are contained in the same dyadic interval of length* 1 *then*

$$\|1_{(-\infty,0]}(k)\mathcal{D}_k\|_{M^{2,s}} \leq C_{r,s}\, |\Xi|^{(1/2-1/r)\frac{s}{s-2}} \sup_{\xi\in\Xi} \big\| \sum_{|\omega|=2^k} a_\omega 1_\omega(\xi) \big\|_{V_k^r}.$$

Corollary 9.5. *Let* $r > 2, 1 < q \leq 2$, *and* $\Xi \subset \mathbb{R}^+$. *Suppose that for each dyadic interval* ω *of length* 1 *and each* $k \in \mathbb{Z}$ *we have a coefficient* $a_{\omega,k} \in \mathbb{R}$. *Then*

$$\big\| \sum_{\substack{|\omega|=1 \\ \omega\cap\Xi\neq\emptyset}} a_{\omega,k} 1_\omega \big\|_{M^{2,s}} \leq C_{r,s}\, |\Xi|^{(1/2-1/r)\frac{s}{s-2}} \sup_{\xi\in\Xi} \big\| \sum_{|\omega|=1} a_{\omega,k} 1_\omega(\xi) \big\|_{V_k^r}.$$

Finally, to see that Corollaries 9.4 and 9.5 imply Lemma 9.2, one argues almost exactly as in the proof of Theorem 4.3 of [10] (the substitution of the Walsh–Paley transform for the Fourier transform allows minor technical simplifications).

References

[1] Bourgain, J.: Return times of dynamical systems. Unpublished manuscript.

[2] Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. (With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.) *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 5–45.

[3] COIFMAN, R., RUBIO DE FRANCIA, J. AND SEMMES, S.: Multiplicateurs de Fourier de $L^p(\mathbb{R})$ et estimations quadratiques. *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988), no. 8, 351–354.

[4] DEMETER, C.: Improved range in the return times theorem. *Canad. Math. Bull.* **55** (2012), no. 4, 708–722.

[5] DEMETER, C.: On some maximal multipliers in L^p. *Rev. Mat. Iberoam.* **26** (2010), no. 3, 947–964.

[6] DEMETER, C., LACEY, M., TAO, T. AND THIELE, C.: The Walsh model for M_2^* Carleson. *Rev. Mat. Iberoam.* **24** (2008), no. 3, 721–744.

[7] DEMETER, C., LACEY, M., TAO, T. AND THIELE, C.: Breaking the duality in the return times theorem. *Duke Math. J.* **143** (2008), no. 2, 281–355.

[8] LACEY, M. AND TERWILLEGER, E.: A Wiener–Wintner theorem for the Hilbert transform. *Ark. Mat.* **46** (2008), no. 2, 315–336.

[9] LACEY, M.: *Issues related to Rubio de Francia's Littlewood–Paley inequality.* NYJM Monographs 2, State University of New York University at Albany, Albany, NY, 2007.

[10] NAZAROV, F., OBERLIN, R. AND THIELE, C.: A Calderón–Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain. *Math. Res. Lett.* **17** (2010), no. 2-3, 529–545.

[11] OBERLIN, R.: A maximal Marcinkiewicz multiplier theorem. *Proc. Amer. Math. Soc.* **141** (2013), no. 6, 2081–2083.

[12] OBERLIN, R., SEEGER, A., TAO, T., THIELE, C. AND WRIGHT, J.: A variation norm Carleson theorem. *J. Eur. Math. Soc.* **14** (2012), no. 2, 421–464.

[13] OBERLIN, R., SEEGER, A., TAO, T., THIELE, C. AND WRIGHT, J.: A variation norm Carleson theorem – Walsh case. Unpublished manuscript.

[14] OBERLIN, R. AND THIELE, C.: New uniform bounds for a Walsh model of the bilinear Hilbert transform. *Indiana Univ. Math. J.* **60** (2011), no. 5, 1693–1712.

[15] TAO, T. AND WRIGHT, J.: Endpoint multiplier theorems of Marcinkiewicz type. *Rev. Mat. Iberoamericana* **17** (2001), no. 3, 521–558.

RICHARD OBERLIN: Mathematics Department, Louisiana State University, Baton Rouge, LA, USA.
E-mail: oberlin@math.lsu.edu