ENDPOINT SPARSE DOMINATION FOR CLASSES OF MULTIPLIER TRANSFORMATIONS

DAVID BELTRAN JORIS ROOS ANDREAS SEEGER

Abstract. We prove endpoint results for sparse domination of translation invariant multiscale operators. The results are formulated in terms of dilation invariant classes of Fourier multipliers based on natural localized $M^{p\to q}$ norms which express appropriate endpoint regularity hypotheses. The applications include new and optimal sparse bounds for classical oscillatory multipliers and multi-scale versions of radial bump multipliers.

1. Introduction

The purpose of this paper is to prove new endpoint bounds in multiscale sparse domination for certain scale invariant classes of translation invariant operators. Interesting partial endpoint sparse bounds are known in some cases (see for example [18, 30]), but they seem to be generally missing in situations where the sharp L^p-bounds rely on Hardy-space or BMO techniques. Model cases for these situations are given by oscillatory Fourier multipliers, for which we obtain optimal endpoint sparse bounds in Theorem 1.6 below, and multi-scale extensions of radial δ-bumps (see Theorem 1.8).

1.1. *Background and definitions.* We begin by reviewing some definitions (see the introduction of [5] for more details). Fix a lattice $\mathfrak{Q}$ of dyadic cubes in the sense of Lerner and Nazarov [38, §2]; this implies, in particular, that the dyadic cubes at a fixed scale are half-open pairwise disjoint cubes, and that every compact set is contained in some $Q \in \mathfrak{Q}$. For $f \in L^1_{\rm loc}$, $Q \in \mathfrak{Q}$ and $1 \le p < \infty$, we set $\langle f\rangle_{Q,p} = (|Q|^{-1}\int_Q |f(y)|^p\,{\rm d}y)^{1/p}$. Given $0<\gamma<1$ a collection $\mathfrak{S} \in \mathfrak{Q}$ is called γ-sparse if for every $Q \in \mathfrak{S}$ there is a measurable subset $E_Q \subset Q$ so that $|E_Q| \ge \gamma|Q|$ and the sets E_Q with $Q \in \mathfrak{S}$ are pairwise disjoint. Given a γ-sparse family $\mathfrak{S}$ of dyadic cubes, and $1 \le p_1, p_2 < \infty$ the corresponding sparse form is defined by

$$\Lambda^{\mathfrak{S}}_{p_1,p_2}(f_1,f_2) = \sum_{Q\in\mathfrak{S}} |Q|\langle f_1\rangle_{Q,p_1}\langle f_2\rangle_{Q,p_2};$$

Date: December 26, 2022.
2020 *Mathematics Subject Classification.* 42B15, 42B20, 42B25.
Key words and phrases. Sparse domination, Fourier multipliers, Endpoint estimates.

this is interesting in the range $p_2 < p_1'$. The maximal (p_1, p_2)-form $\Lambda^*_{p_1,p_2}$ is given by

$$\Lambda^*_{p_1,p_2}(f_1, f_2) = \sup_{\mathfrak{S}:\gamma\text{-sparse}} \Lambda^{\mathfrak{S}}_{p_1,p_2}(f_1, f_2), \tag{1.1}$$

where the supremum is taken over all γ-sparse families of dyadic cubes in $\mathfrak{Q}$. We say that a linear operator $T : C_c^\infty(\mathbb{R}^d) \to \mathcal{D}'(\mathbb{R}^d)$ belongs to the space $\mathrm{Sp}_\gamma(p_1, p_2)$ (or satisfies a (p_1, p_2) sparse bound) if for all f_1, $f_2 \in C_c^\infty$ the inequality

$$|\langle Tf_1, f_2\rangle| \le C\Lambda^*_{p_1,p_2}(f_1, f_2) \tag{1.2}$$

holds with some constant C independent of f_1 and f_2, and we denote by $\|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}$ the best constant in this inequality. The norm $\|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}$ depends on γ, but the space $\mathrm{Sp}_\gamma(p_1, p_2)$ does not. As we keep γ fixed throughout this paper we will drop the subscript γ when using the $\mathrm{Sp}_\gamma(p_1, p_2)$ norm. As mentioned above, the relevant case for applications is $p_2 < p_1'$ (and indeed if T is a convolution operator with compactly supported kernel, a (p_1, p_1') sparse bound follows immediately from the L^{p_1} boundedness of T). We remark that when $p_2 < p_1'$ we can change the a priori assumption of $f_1, f_2 \in C_c^\infty$ to $f_1, f_2 \in \mathbb{V}$ where $\mathbb{V}$ is any subspace dense in L^p for some $p \in (p_1, p_2')$; for example, it is natural to choose $\mathbb{V} = L^{p_1} \cap L^{p_2'}$ (see [5, Lemma A.1]).

The interest in (a pointwise/normed version of) sparse domination started because of its important consequences in weighted inequalities for Calderón–Zygmund operators [35, 36, 20, 38, 32, 37, 6, 22]. For consequences of the bilinear sparse domination (1.2) in weighted theory we refer to the paper by Bernicot, Frey and Petermichl [8]. A detailed exposition of the importance of sparse domination in harmonic analysis can be found in the introduction of [43]; for many further examples beyond Calderón–Zygmund theory see [33, 19, 5] and references therein.

In this paper we shall consider operators that commute with translations. They are defined as a Fourier multiplier transformation $m(D)$ where $\widehat{m(D)f}(\xi) = m(\xi)\widehat{f}(\xi)$. Here we work with $\widehat{f}(\xi) = \int f(y)e^{i\langle y,\xi\rangle}\,\mathrm{d}y$ as the definition of the Fourier transform of $f \in \mathcal{S}(\mathbb{R}^d)$ and denote by $\check{f} \equiv \mathcal{F}^{-1}[f]$ the inverse Fourier transform. For $1 \le p \le q < \infty$ we denote by $M^{p\to q}$ the class of Fourier multipliers for which $m(D)$ is bounded as an operator from L^p to L^q; the norm in $M^{p\to q}$ is just given by the $L^p \to L^q$ operator norm of $m(D)$. A modification is needed for $p = \infty$; then L^∞ is replaced by C_0.

1.2. *Scale invariant classes of multipliers: the main results.* We shall now formulate our three main theorems on sparse domination involving scale invariant classes of multipliers and subsequently discuss new sharp results for oscillatory multipliers and multiscale radial bump multipliers. Our conditions are motivated by p-sensitive endpoint multiplier theorems in [44, 46, 47] (see also an earlier result by Baernstein and Sawyer [1] on H^p multipliers

for $p<1$). The M^p multiplier hypotheses (in particular the one in [47]) can be seen as certain localized Besov-conditions where the Besov spaces are built on suitable Fourier multiplier spaces. Here we will formulate similar conditions which will be relevant for endpoint sparse bounds.

Let $\Phi_0 \in C^\infty(\mathbb{R}^d)$ be supported in $\{x \in \mathbb{R}^d : |x| < 1/2\}$ such that $\Phi_0(x) = 1$ for $|x| \le 1/4$. For $\ell \in \mathbb{Z}$ define

$$\Psi_\ell(x) = \Phi_0(2^{-\ell}x) - \Phi_0(2^{-\ell+1}x) \tag{1.3}$$

which is supported in $\{x \in \mathbb{R}^d : 2^{\ell-3} \le |x| \le 2^{\ell-1}\}$. For a Banach space $\mathfrak{X}$ of distributions (here suitable classes of multipliers) let $B^\alpha_1(\mathfrak{X})$ be the B^α_1-Besov space built on $\mathfrak{X}$, with norm

$$\|h\|_{B^\alpha_1(\mathfrak{X})} = \|h * \widehat{\Phi}_0\|_{\mathfrak{X}} + \sum_{\ell>0} 2^{\ell\alpha}\|h * \widehat{\Psi}_\ell\|_{\mathfrak{X}}.$$

The standard Besov-classes $B^\alpha_{u,1}$ can be recovered by taking $\mathfrak{X} = L^u(\mathbb{R}^d)$; however for our results it is most appropriate to take for $\mathfrak{X}$ a multiplier space such as $M^{r\to q}$ for r between p and q. Let ϕ be a radial C^∞ function supported in $\{\xi \in \widehat{\mathbb{R}}^d : 1/2 < |\xi| < 2\}$ which is not identically zero.[1]

It was proved in [47] that

$$\|m\|_{M^{p\to p}} \le C_{p,r} \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{r})}(M^{r\to r})}, \quad 1<p<r\le 2. \tag{1.4}$$

Inequality (1.4) is related to results in [11, 45] but the latter are not applicable to endpoint estimates in many situations; indeed, they do not give satisfactory results for the oscillatory multipliers in (1.6) below.

We state now versions of these results for (p, q') sparse domination, with three cases, depending on whether $q < 2$, $q = 2$ or $q > 2$. Multipliers for which the right-hand side of (1.4) is finite belong at least to the class $\mathrm{Sp}(p, r')$; this follows from the special case $r = q$ in our first theorem.

Theorem 1.1. *Let $1 < p \le q \le 2$. Then for $p < r \le q$,*

$$\|m(D)\|_{\mathrm{Sp}(p,q')} \le C_{p,r} \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{q})}(M^{r\to q})}.$$

For $q = 2$ we can also let $r = p$ to get a $\mathrm{Sp}(p, 2)$ bound.

Theorem 1.2. *Let $1 < p < 2$. Then*

$$\|m(D)\|_{\mathrm{Sp}(p,2)} \le C_p \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{2})}(M^{p\to 2})}.$$

For $q > 2$ we have the following version.

[1] The assumption that ϕ is radial is convenient but not crucial; one can show that one just needs to assume that for all rays emanating from the origin, the restriction of ϕ to the ray is not identically zero.

Theorem 1.3. *Let $1 < p < 2 < q < p'$. Then for $q' < r \le 2$,*

$$\|m(D)\|_{\mathrm{Sp}(p,q')} \le C_{p,q,r} \times \Big(\sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{q})}(M^{p\to q})} + \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{q'}-\frac{1}{r})}(M^{r\to r})}\Big). \tag{1.5}$$

Remarks. (i) The spaces of multipliers defined by the conditions in the above three theorems are independent of the choice of the radial non-trivial function ϕ and independent of the specific spatial cutoff function. This can be shown by routine but somewhat lengthy calculations. We omit the proof but point out that our choice for ϕ can always be taken as the cutoff function φ used in the Calderón reproducing formula (2.1).

(ii) As noted in [44, p.152] the multipliers in Theorem 1.2 satisfy for $1 < p < 2$ an inequality involving the Lorentz space $L^{p,2}$,

$$\|m(D)f\|_{L^{p,2}} \lesssim_p A\|f\|_p, \qquad \text{with } A = \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{2})}(M^{p\to 2})}.$$

This is shown to be a consequence of the weighted norm inequality

$$\int \mathcal{G}[m(D)f]^2 w\,\mathrm{d}x \lesssim A^2 \int \widetilde{\mathcal{G}}[f]^2 (M[|w|^s])^{1/s}\,\mathrm{d}x, \quad s = (p'/2)',$$

where M is the Hardy-Littlewood maximal operator, and $\mathcal{G}$, $\widetilde{\mathcal{G}}$ are suitable Littlewood–Paley–Stein operators. For earlier closely related variants in non-endpoint cases see [49, Ch.IV], [15].

(iii) The proofs of Theorems 1.1, 1.2, and 1.3 have a similar structure and, in order to avoid repetitions, we shall present them together. They rely on an iteration argument common in sparse domination; one main novelty in this paper is that at every step of the iteration Calderón–Zygmund arguments are combined with atomic decompositions in L^p-spaces of functions on certain dyadic cubes. This use of the iterated atomic decompositions is crucial for the proof of Theorem 1.1, but can be replaced by applications of Littlewood–Paley theory in the proofs of Theorem 1.2 and 1.3.

We can use the embedding $L^u \subset M^{p\to 2}$ for $1 \le p \le 2$, $1/p - 1/2 = 1/u$ to derive a corollary of Theorem 1.2 which uses standard Besov spaces $B_{u,1}^{d(1/p-1/2)} = B_1^{d(1/p-1/2)}(L^u)$.

Corollary 1.4. *Suppose $1 < p < 2$, $1/p - 1/2 \ge 1/u$ and $m \in L^\infty$ satisfies*

$$\sup_{t>0} \|\phi m(t\cdot)\|_{B_{u,1}^{d(\frac{1}{p}-\frac{1}{2})}} < \infty.$$

Then $m(D) \in \mathrm{Sp}(p,2)$.

Another corollary of Theorem 1.2 involves radial multipliers where L^u is replaced by L^2, as a consequence of the Stein–Tomas Fourier restriction theorem.

Corollary 1.5. *Let $d \geq 2$, $1 < p \leq \frac{2(d+1)}{d+3}$. Suppose that*

$$\sup_{t>0} \|\phi h(t\cdot)\|_{B^{d(\frac{1}{p}-\frac{1}{2})}_{2,1}(\mathbb{R})} < \infty.$$

Then $h(|D|) \in \mathrm{Sp}(p,2)$.

1.3. *Oscillatory Fourier multipliers and Miyachi classes.* For $a > 0$, $a \neq 1$ and $b \geq 0$ consider the Fourier multipliers

$$m_{a,b}(\xi) = e^{i|\xi|^a}|\xi|^{-b}\chi_\infty(\xi) \tag{1.6}$$

where $\chi_\infty \in C^\infty(\mathbb{R}^d)$, $\chi_\infty(\xi) = 1$ for $|\xi| \geq 1$ and χ_∞ vanishes in a neighborhood of the origin. It is well-known that the operator $m_{a,b}(D)$ is bounded on $L^p(\mathbb{R}^d)$ for all $p \in (1,\infty)$ if and only if $b \geq ad/2$. Moreover if $0 < b < ad/2$, L^p boundedness holds if and only if $\frac{2ad}{ad+2b} \leq p \leq \frac{2ad}{ad-2b}$, i.e., equivalently, if $b \geq ad|\frac{1}{p} - \frac{1}{2}|$ (see [50], [27], [40]). Miyachi [40] considered classes generalizing the oscillatory multipliers $m_{a,b}$; for $0 < a < 1$ these correspond to translation invariant versions of the pseudo-differential operators with $S^{-b}_{1-a,\delta}$ symbols for which Fefferman [24] had already proved sharp L^p bounds. We say that $m \in \mathrm{FM}(a,b)$ if m is supported in $\{\xi \in \mathbb{R}^d : |\xi| \geq 1\}$ and satisfies the derivative estimates

$$|m^{(\beta)}(\xi)| \leq C_\beta |\xi|^{(a-1)|\beta|-b} \tag{1.7}$$

for all multiindices $\beta \in \mathbb{N}_0^d$. It is proved in [40] that for $1 < p \leq 2$ the multiplier operators $m(D)$ with $m \in \mathrm{FM}(a, ad(\frac{1}{p} - \frac{1}{2}))$ are bounded on L^p; this result is optimal.

In [5] it was shown that for $0 < b < ad/2$, the operator $m_{a,b}(D)$ belongs to $\mathrm{Sp}(p,p)$ in the open range for $\frac{2ad}{ad+2b} < p \leq 2$. For general multipliers in $\mathrm{FM}(a,b)$ it was shown in [4] that the operators belong to $\mathrm{Sp}(p,2)$, in the same p-range. We note that no nontrivial (p_1,p_2) sparse bounds with $p_2 < p_1'$ was obtained at the endpoint $p_1 = \frac{2ad}{ad+2b}$, i.e. $b = ad(\frac{1}{p_1} - \frac{1}{2})$.

We provide a full characterization of the sparse exponent set for the oscillatory multiplier operators $m_{a,b}(D)$ in the relevant parameter case $0 < b < ad/2$, thereby settling the open endpoint problem.

Theorem 1.6. *Let $a \neq 1$ and $0 < b < ad/2$. Let $\triangle(a,b)$ the closed triangle with vertices $Q_1 = (\frac{1}{2}+\frac{b}{da}, \frac{1}{2}-\frac{b}{da})$, $Q_2 = (\frac{1}{2}-\frac{b}{da}, \frac{1}{2}+\frac{b}{da})$, $Q_3 = (\frac{1}{2}+\frac{b}{da}, \frac{1}{2}+\frac{b}{da})$ and $m_{a,b}$ be the oscillatory multiplier in* (1.6)*. Then*

$$m_{a,b}(D) \in \mathrm{Sp}(p_1,p_2) \iff (\tfrac{1}{p_1}, \tfrac{1}{p_2}) \in \triangle(a,b).$$

In particular for the oscillatory multipliers we get the $\mathrm{Sp}(p_1,p_2)$ bound for the endpoint $p_1 = \frac{2ad}{ad+2b}$ in the optimal range $p_1 \leq p_2 \leq p_1'$. We also have a sharp result that applies to the full class $\mathrm{FM}(a,b)$.

Theorem 1.7. *Let $a \neq 1$ and $0 < b < ad/2$. Let $\square(a,b)$ be the closed trapezoid with vertices $Q_1 = (\frac{1}{2} + \frac{b}{da}, \frac{1}{2} - \frac{b}{da})$, $Q_2 = (\frac{1}{2} - \frac{b}{da}, \frac{1}{2} + \frac{b}{da})$, $P_3 =$*

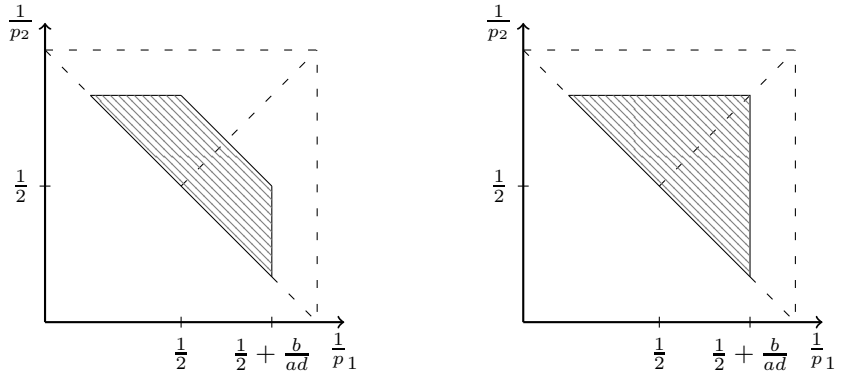

FIGURE 1. Sparse bounds for the general multiplier class $\mathrm{FM}(a,b)$ (left) and for the oscillatory multipliers $m_{a,b}$ (right) for given $a, b > 0$ with $0 < b < ad/2$.

$(\frac{1}{2}, \frac{1}{2}+\frac{b}{da})$, $P_4 = (\frac{1}{2}+\frac{b}{da}, \frac{1}{2})$. *Then*

$$m(D) \in \mathrm{Sp}(p_1,p_2) \textit{ for all } m \in \mathrm{FM}(a,b) \iff (\tfrac{1}{p_1}, \tfrac{1}{p_2}) \in \text{⏢}(a,b).$$

The results of Theorems 1.6 and 1.7 are illustrated in Figure 1. The positive results on the edges $(Q_1Q_3]$ and $[Q_3Q_2)$ of the triangle on the right are new for the oscillatory multipliers. For the $\mathrm{FM}(a,b)$ class the positive results on the edges $(Q_1P_3]$, $[P_3P_4]$, $[P_4Q_2)$ of the trapezoid are new.

Remarks. (i) The general positive result about the multipliers in $\mathrm{FM}(a,b)$ can be derived from Corollary 1.4. Indeed, this corollary implies sharp results for the classes of subdyadic multipliers considered in [3], see also a relevant discussion in [5]. For the extended region of the oscillatory multipliers we need to use Theorem 1.3.

(ii) The methods in this paper can also be used to strengthen results in [4] on sparse bounds for pseudo-differential operators with symbols in the Hörmander classes $S^\nu_{\rho,\delta}$, for $0 < \delta \le \rho < 1$. By [24, 39] these operators are bounded on L^p (here $1 < p < \infty$) provided that $\nu \le -d(1-\rho)|\frac{1}{p} - \frac{1}{2}|$. In the range $-\frac{d}{2}(1-\rho) < \nu < 0$ we now get the full endpoint sparse bounds, extending the results for the multiplier classes $\mathrm{FM}(1-\rho, -\nu)$, that is, the operators belong to $\mathrm{Sp}(p_1,p_2)$ for $(1/p_1, 1/p_2) \in \text{⏢}(1-\rho,-\nu)$.

(iii) The multiplier class in Theorem 1.2 is also relevant in the interesting recent work by Bulj–Kovač [9] and by Stolyarov [51] on lower bounds for other types of oscillatory multipliers; indeed the theorem allows to derive upper $\mathrm{Sp}(p,2)$ bounds in their setting.

1.4. *Multiscale radial bump multipliers.* Let χ be a smooth bump function supported in $(-1/2, 1/2)$ and set for small δ

$$h_\delta(t) = \chi(\delta^{-1}(1-|t|)). \tag{1.8}$$

The multiplier $h_\delta(|\xi|)$ occurs naturally as a building block for the Bochner–Riesz multipliers. It is conjectured that in dimension $d \geq 2$ we have

$$\sup_{0<\delta<1/2} \delta^{d(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}} \|h_\delta(|\cdot|)\|_{M^p\to p} < \infty, \tag{1.9}$$

for $1 \leq p < \frac{2d}{d+1}$. This conjecture is well known in two dimensions (the range is then $1 \leq p < 4/3$, see [12, 25, 21, 48]), and there are partial results in higher dimensions. More specifically, by Tao's arguments in [52] the bound (1.9) for any fixed $p < \frac{2d}{d+1}$ follows from a slightly weaker bound with an additional factor of $c_\epsilon \delta^\epsilon$ for arbitrary $\epsilon > 0$ on the left-hand side, and such estimates with the ϵ-loss have been verified on a partial range of p (see [28] for the latest results and more references).

Here, we consider the multiscale version

$$m_\delta(\xi) = \sum_{k\in\mathbb{Z}} a_k h_\delta(2^k|\xi|). \tag{1.10}$$

From [47] we know that if (1.9) holds for some $p_\circ < \frac{2d}{d+1}$, then we have

$$\sup_{0<\delta<1/2} \delta^{d(\frac{1}{p}-\frac{1}{2})-\frac{1}{2}} \|m_\delta(D)\|_{L^p\to L^p} \lesssim_p \sup_{k\in\mathbb{Z}} |a_k|. \tag{1.11}$$

for $1 < p < p_\circ$.

Our purpose here is to illustrate how Theorems 1.1, 1.2 and 1.3 imply various sharp or essentially sharp sparse domination results in a p-range that will be optimal in two dimensions; in higher dimensions we limit ourselves to the range $1 < p < \frac{2(d+2)}{d+4}$ (i.e. the range dual to Tao's bilinear Fourier extension theorem [53]), as in this range the known sharp $L^p \to L^q$ estimates for $h_\delta(|D|)$ are well documented in the literature [2, 29, 14].

Theorem 1.8. *Let m_δ be as in* (1.10) *and let $d \geq 2$. Then the inequality*

$$\sup_{0<\delta<1/2} \delta^{d(\frac{1}{p_1}-\frac{1}{2})-\frac{1}{2}} \|m_\delta(D)\|_{\mathrm{Sp}(p_1,p_2)} \lesssim_{p_1,p_2} \sup_{k\in\mathbb{Z}} |a_k| \tag{1.12}$$

holds if

(a) $1 < p_1 \leq \frac{2(d+1)}{d+3}$ *and* $p_2 \geq \frac{(d-1)p_1}{d+1-2p_1}$, *or*
(b) $\frac{2(d+1)}{d+3} < p_1 < \frac{2(d+2)}{d+4}$ *and* $p_2 > \frac{(d-1)p_1}{d+1-2p_1}$.

Remarks. (i) Using the building block with $a_0 = 1$ and $a_k = 0$ for $k \neq 0$ one sees that this result is sharp in the sense that inequality (1.12) fails in general if $1 < p_1 < \frac{2d}{d+1}$ and $p_2 < \frac{(d-1)p_1}{d+1-2p_1}$. This can be deduced directly from a corresponding result for Bochner–Riesz operators in [34, §5].

(ii) For the proof of sufficiency in Theorem 1.8 we rely on Theorems 1.2 and 1.3 in the range $p_1 \leq \frac{2(d+1)}{d+3}$ and on Theorem 1.1 in the range $\frac{2(d+1)}{d+3} < p_1 < \frac{2(d+2)}{d+4}$.

1.5. *Necessary conditions for sparse domination of convolution operators.* Necessary conditions for general sparse operators were discussed in Chapter 2 of [5]; here we point out that they can be put in a simple form for scalar operators. This result will be convenient for checking the sharpness of several of the results mentioned above.

Let $T : C_c^\infty(\mathbb{R}^d) \to \mathcal{D}'(\mathbb{R}^d)$ with Schwartz kernel $K \in \mathcal{D}'(\mathbb{R}^d \times \mathbb{R}^d)$. Let $\Psi \in C_c^\infty(\mathbb{R}^d)$ be supported in $\{x \in \mathbb{R}^d : 1 \le |x| \le 2\}$. Define the distribution K_R as the multiplication of K with the C^∞ function $\Psi(R^{-1}(x-y))$ and let T_R denote the linear operator with Schwartz kernel K_R. Define the rescaled kernels

$$K_R^{\mathrm{resc}}(x,y) := R^d K_R(Rx,Ry) = \Psi(x-y) R^d K(Rx,Ry), \tag{1.13}$$

interpreted in the sense of distributions, and let T_R^{resc} be the rescaled version of T_R, with convolution kernel K_R^{resc}.

Proposition 1.9. *Let $T : C_c^\infty(\mathbb{R}^d) \to \mathcal{D}'(\mathbb{R}^d)$ be a continuous linear operator with Schwartz kernel K and let T_R^{resc} be the rescaled version defined in* (1.13)*. Suppose $1 < p_1, p_2 < \infty$ and $T \in \mathrm{Sp}(p_1, p_2)$, with $p_2 < p_1'$. Then T extends to a bounded operator $L^{p_1} \to L^{p_1,\infty}$ and $L^{p_2',1} \to L^{p_2'}$; moreover the operators T_R^{resc} map L^{p_1} to $L^{p_2'}$ with uniform operator norm and*

$$\|T\|_{L^{p_1}\to L^{p_1,\infty}} + \|T\|_{L^{p_2',1}\to L^{p_2'}} + \sup_{R>0} \|T_R^{\mathrm{resc}}\|_{L^{p_1}\to L^{p_2'}} \lesssim \|T\|_{\mathrm{Sp}(p_1,p_2)}.$$

The proof is based on more general results in [5] and will be given in §10.

Structure of the paper. In §2, we present the induction scheme that proves the sparse domination Theorems 1.1, 1.2 and 1.3. In §3 we discuss the atomic decomposition, which is used in §4 to verify the base case for the induction. In §5 we present a Calderón–Zygmund decomposition based on the atomic decomposition. The plan for the proof of the induction step is outlined in §6, with proofs presented in §7 and §8. In §9 we discuss the applications of the main theorems and, in particular, how they imply the positive results in Theorems 1.6 and 1.7, and the positive results on radial multipliers in Corollary 1.5 and Theorem 1.8. Proposition 1.9 and the proof of the necessary conditions for Theorems 1.6 and 1.7 are presented in §10. Finally, §11 contains the proofs of some technical facts which are included for the reader's convenience.

Acknowledgements. The authors would like to thank the Hausdorff Research Institute of Mathematics and the organizers of the trimester program *Interactions between Geometric Measure Theory, Singular Integrals, and PDE* for a pleasant working environment during a visit in January 2022. The work was supported in part by National Science Foundation grants DMS-1954479 (D.B.), DMS-2154835 (J.R.), DMS-2054220 (A.S.), by a grant from the Simons Foundation (ID 855692, J.R.) and by the Agencia Estatal de Investigación through RYC2020-029151-I (D.B.).

2. Structure of the induction argument

In this section we present the proof strategy for Theorems 1.1, 1.2 and 1.3. We will see that sparse bounds for $m(D)$ can be deduced from sparse bounds for finite multi-scale sums of spatially (and frequency) localized pieces of $m(D)$. The proof of the latter is based on an induction on the number of pieces in the multi-scale sums, similarly to our previous work [5]. As in [5], it is useful to work with a modified version of our maximal (p_1, p_2)-form $\Lambda^*_{p_1,p_2}$.

Definition 2.1. *Given a dyadic cube $S_0 \in \mathfrak{Q}$ let*

$$\Lambda^{**}_{S_0,p,q'}(f_1, f_2) := \sup \sum_{S\in\mathfrak{S}} |S|\langle f_1\rangle_{S,p}\langle f_2\rangle_{3S,q'}$$

where the supremum is taken over all γ-sparse collections $\mathfrak{S}$ consisting of cubes in $\mathfrak{Q}(S_0)$, which denotes the subset of $\mathfrak{Q}$ of cubes contained in S_0.

2.1. *Decomposition as a multi-scale sum.* Consider $\eta, \varphi \in \mathcal{S}$ facilitating the Calderón reproducing formula, i.e.,

$$\check{\eta} \text{ has compact support in } \{|x| \le 10^{-d}\} \text{ and } \eta(0) = 0; \tag{2.1a}$$

$$\varphi \text{ has compact support in } \{1/2 < |\xi| < 2\}; \tag{2.1b}$$

$$\sum_{k\in\mathbb{Z}} \eta^2(2^{-k}\xi)\varphi(2^{-k}\xi) = 1, \quad \xi \neq 0. \tag{2.1c}$$

For any $k \in \mathbb{Z}$, let L_k and P_k be defined by

$$\widehat{L_k f}(\xi) = \varphi_k(\xi)\widehat{f}(\xi), \qquad \text{where } \varphi_k(\xi) := \varphi(2^{-k}\xi),$$
$$\widehat{P_k f}(\xi) = \eta_k(\xi)\widehat{f}(\xi), \qquad \text{where } \eta_k(\xi) := \eta(2^{-k}\xi).$$

Let $T_k = m(D)L_k$ and denote by K_k its convolution kernel, that is $K_k = \mathcal{F}^{-1}[\varphi_k m]$. We next perform a spatial decomposition of K_k. Let Φ_0 and Ψ_j be as in (1.3). Let

$$K_k^{(-k)}(x) = \mathcal{F}^{-1}[\varphi_k m](x)\Phi_0(2^k x)$$
$$K_k^{(j)}(x) = \mathcal{F}^{-1}[\varphi_k m](x)\Psi_j(x) \quad \text{ if } j > -k$$

so that we get $K_k = \sum_{j=-k}^{\infty} K_k^{(j)} = K_k^{(-k)} + \sum_{\ell>0} K_k^{(\ell-k)}$. Let $T_k^{(\ell-k)}$ denote the operator with convolution kernel $K_k^{(\ell-k)}$ for $\ell \ge 0$. Note that by (2.1c) we have

$$m(D) = \sum_{k\in\mathbb{Z}} T_k P_k P_k = \sum_{k\in\mathbb{Z}}\sum_{\ell\ge 0} T_k^{(\ell-k)} P_k P_k.$$

It is also convenient to introduce some notation for the operator norm of $T_k^{(\ell-k)}$. We first note that

$$\widehat{K_k^{(\ell-k)}}(2^k\xi) = (\varphi m(2^k\cdot)) * \widehat{\Psi_\ell}(\xi) \tag{2.2}$$

and define the quantities

$$A^{k,\ell}_{p,r,q}[m] \equiv A^{k,\ell}_{p,r,q} = \begin{cases} \|\varphi m(2^k\cdot) * \widehat{\Phi_0}\|_{M^{r\to q}}, & \text{if } \ell = 0, \\ \|\varphi m(2^k\cdot) * \widehat{\Psi_\ell}\|_{M^{r\to q}} 2^{\ell d(\frac{1}{p}-\frac{1}{q})}, & \text{if } \ell > 0, \end{cases} \tag{2.3}$$

and

$$\mathcal{A}_{p,r,q}[m] = \sup_{k\in\mathbb{Z}} \sum_{\ell\geq 0} A^{k,\ell}_{p,r,q}[m], \tag{2.4}$$

assuming that $p < q$ and $p \leq r \leq q$. Then

$$\mathcal{A}_{p,q,r}[m] \leq \sup_{t>0} \|\varphi m(t\cdot)\|_{B_1^{d(\frac{1}{p}-\frac{1}{q})}(M^{r\to q})}. \tag{2.5}$$

Moreover we have for all $k \in \mathbb{Z}$,

$$\sum_{\ell\geq 0} A^{k,\ell}_{p,r_1,q} \lesssim \sum_{\ell\geq 0} A^{k,\ell}_{p,r_2,q}, \quad p \leq r_1 \leq r_2 \leq q. \tag{2.6}$$

This inequality is an immediate consequence of a slightly stronger statement, Corollary 11.3.

The sparse bounds for sums of operators $T_k^{(-k)}$ are reduced to standard sparse bounds for singular integral operators. This only requires the assumption $m \in L^\infty$; note that

$$\|m\|_\infty \lesssim \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^0(M^{p\to p})} \lesssim \sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(1/p-1/q)}(M^{p\to q})}. \tag{2.7}$$

Lemma 2.2. *For any $1 < p \leq q < \infty$, and for any finite subset $F \subset \mathbb{Z}$*

$$\Big|\big\langle \sum_{k\in F} T_k^{(-k)} P_k f_1, f_2\big\rangle\Big| \lesssim \|m\|_\infty \, \Lambda^*_{p,q'}(f_1, f_2),$$

uniformly in $F \subset \mathbb{Z}$, for all $f_1, f_2 \in C_c^\infty$.

The proof is straightforward and will be given in the auxiliary §11.2.

We now introduce operators which are local at a fixed spatial scale. For a fixed finite set $F \subseteq \mathbb{Z}$, let $k_{\min} := \min F$ and $k_{\max} := \max F$. Given $j \in \mathbb{Z}$, it is convenient to define

$$\mathcal{T}_j f \equiv \mathcal{T}_{j,F} f := \sum_{\substack{k\in F \\ k>-j}} T_k^{(j)} P_k P_k f \tag{2.8}$$

and to note that

$$\sum_{k\in F} \sum_{\ell=1}^{N} T_k^{(\ell-k)} P_k P_k = \sum_{-k_{\max}<j\leq -k_{\min}+N} \mathcal{T}_j. \tag{2.9}$$

By construction, the operators $\mathcal{T}_j$ are local at scale 2^j, in the sense that if S is a cube of side length 2^j,

$$\operatorname{supp}(f) \subset S \implies \operatorname{supp}(\mathcal{T}_j f) \subset 3S. \tag{2.10}$$

Indeed, by our definition of the Φ_0 and the Ψ_j in (1.3), $T_k^{(j)} P_k P_k[f\mathbb{1}_S]$ is supported in the set $\{x : \operatorname{dist}(x, S) \leq 2^{j-1} + 10^{-d} 2^{-k+1}\}$. Thus $\mathcal{T}_{j,F}[f\mathbb{1}_S]$ is

supported where $\operatorname{dist}(x,S) \le 2^{j-1}(1+2\cdot 10^{-d})$, and hence in $\{x : \operatorname{dist}(x,S) < 2^j\} \subseteq 3S$.

The key estimate in proving the sparse bounds for $m(D)$ is the following modified sparse bound for sums of $\mathcal{T}_j$, uniformly in the number of terms in the j-sum. Throughout the paper we set

$$L(Q) = \log_2(\text{sidelength}(Q))$$

so that $L(Q) = N$ for a dyadic cube of side length 2^N.

Theorem 2.3. *Let $1 < p < q < \infty$. Given integers $N_1 \le N_2$, a dyadic cube $S_0 \in \mathfrak{Q}$ such that $L(S_0) = N_2$ and a finite subset $F \subseteq \mathbb{Z}$, the inequality*

$$\Big|\big\langle \sum_{j=N_1}^{N_2} \mathcal{T}_{j,F} f_1, f_2\big\rangle\Big| \le c\,\mathcal{C}\,\Lambda^{**}_{S_0,p,q'}(f_1,f_2) \tag{2.11}$$

holds for all $f_1, f_2 \in C_c^\infty$ uniformly in N_1, N_2, F and S_0, where $\mathcal{C}$ is given by

$$\mathcal{C} := \mathcal{A}_{p,r,q}[m] \qquad \text{if } 1 < p \le q < 2, \quad p < r \le q, \tag{2.12a}$$

$$\mathcal{C} := \mathcal{A}_{p,p,2}[m] \qquad \text{if } 1 < p < 2 \quad (\text{and } q = 2), \tag{2.12b}$$

$$\mathcal{C} := \mathcal{A}_{p,p,q}[m] + \mathcal{A}_{q',r,r}[m] \qquad \text{if } 1 < p < 2 < q < p', \quad q' < r \le 2, \tag{2.12c}$$

and $c = c(p,q,r,\gamma,d)$ is a constant depending only on p,q,r,γ,d.

We note that by the definition of $\mathcal{T}_j$ in (2.8), we may assume that the set F featuring in the left-hand side of (2.11) has the property that $k > -N_2 = -L(S_0)$ for $k \in F$.

Also note by (2.10) that in order to prove the theorem we may assume without loss of generality that f_1 is supported in S_0 and f_2 is supported in $3S_0$.

We shall use standard arguments in the theory of sparse domination to make the following

Observation 2.4. *In order to prove Theorems 1.1, 1.2 and 1.3, it suffices to prove Lemma 2.2 and Theorem 2.3.*

The proof of this observation is included in the auxiliary §11.1.

2.2. *Induction scheme for the proof of Theorem 2.3.* We will prove (2.11) by induction on $\mathbf{n}$ where $\mathbf{n}+1$ is the number of terms in the j-sum.

Definition 2.5. *For $\mathbf{n} = 0, 1, 2, \dots$, let $\mathbf{U}(\mathbf{n})$ be the smallest nonnegative constant U so that for all pairs (N_1, N_2) with $0 \le N_2 - N_1 \le \mathbf{n}$, for all finite sets $F \subset \mathbb{Z}$ and for all dyadic cubes $S_0 \in \mathfrak{Q}$ with $L(S_0) = N_2$ we have*

$$\Big|\big\langle \sum_{j=N_1}^{N_2} \mathcal{T}_{j,F} f_1, f_2\big\rangle\Big| \le U\Lambda^{**}_{S_0,p,q'}(f_1,f_2)$$

whenever $\operatorname{supp}(f_1) \subset S_0$.

For the inductive argument we first consider the base case $\mathbf{n} = 0$. We distinguish two situations, $q \geq 2$ and $q < 2$. For fixed j_0 we let

$$\mathcal{C}_{p,r,q}(j_0) = \sup_{k > -j_0} A^{k,j_0+k}_{p,r,q}. \tag{2.13}$$

It is immediate that $\sup_{j_0} \mathcal{C}_{p,r,q}(j_0)$ is bounded by $\mathcal{C}_{(2.12c)}$ if $q \geq 2$, $r = p$, bounded by $\mathcal{C}_{(2.12b)}$ if $r = p$, $q = 2$ and bounded by $\mathcal{C}_{(2.12a)}$ if $q < 2$, $p < r \leq q$.

Lemma 2.6. *Let $1 < p \leq 2 \leq q \leq p'$ and $j_0 \in \mathbb{Z}$. Let S_0 be a dyadic cube with $L(S_0) = j_0$ and let $f_1 \in L^p$ be supported in S_0. Then we have for $f_2 \in L^q_{\mathrm{loc}}$*

$$|\langle \mathcal{T}_{j_0} f_1, f_2 \rangle| \lesssim \mathcal{C}_{p,p,q}(j_0) |S_0| \langle f_1 \rangle_{S_0,p} \langle f_2 \rangle_{3S_0,q'}. \tag{2.14}$$

Lemma 2.7. *Let $1 < p < r \leq q \leq 2$ and $j_0 \in \mathbb{Z}$. Let S_0 be a dyadic cube with $L(S_0) = j_0$ and let $f_1 \in L^p$ be supported in S_0. Then we have for $f_2 \in L^q_{\mathrm{loc}}$*

$$|\langle \mathcal{T}_{j_0} f_1, f_2 \rangle| \lesssim \big(\mathcal{C}_{p,p,q}(j_0) + \mathcal{C}_{p,r,q}(j_0)\big) |S_0| \langle f_1 \rangle_{S_0,p} \langle f_2 \rangle_{3S_0,q'}. \tag{2.15}$$

Both lemmata can be reduced to $L^p \to L^q$ estimates for the operators $\mathcal{T}_{j_0}$; these and the corresponding reduction are stated in §4 (see (4.4) for the reduction argument).

Corollary 2.8. *With $\mathcal{C}$ as in* (2.12),

$$\mathbf{U}(0) \leq c_0(p, r, q, d)\, \mathcal{C}. \tag{2.16}$$

Proof. This is immediate from Lemmata 2.6 and 2.7, in combination with the inequality (2.6). □

Corollary 2.8 is the base case for our induction. We note that the same argument implies $\mathbf{U}(\mathbf{n}) \leq c(\mathbf{n}, p, r, q, d)\mathcal{C}$ for any $\mathbf{n} \geq 0$, but one needs a uniform bound in $\mathbf{n}$. The key is the verification of the induction step, formulated in the following claim.

Proposition 2.9 (Inductive claim)**.** *There is a constant $c = c(p, q, r, \gamma, d)$ such that for all $\mathbf{n} > 0$,*

$$\mathbf{U}(\mathbf{n}) \leq \max\{\mathbf{U}(\mathbf{n} - 1),\, c\mathcal{C}\},$$

with $\mathcal{C}$ as in (2.12).

The proof structure for the inductive claim is presented in §6, with proofs given in §7 and §8. They are based on Calderón–Zygmund decompositions combined with the atomic decomposition outlined in §5. By induction, the conclusion of Theorem 2.3 follows by combining Corollary 2.8 and the inductive claim Proposition 2.9.

3. Atomic and subatomic decompositions

In this section we fix a dyadic reference cube S_0 and outline an atomic decomposition for a function f supported in a dyadic cube S_0 based on estimates for a martingale square function on the cube S_0.

3.1. *Local square functions.* Let $\{\mathbb{E}_n\}_{n\in\mathbb{Z}}$ be the conditional expectation operators associated to the σ-algebra generated by the subfamily $\mathfrak{Q}_n$ of cubes in $\mathfrak{Q}$ with $L(Q) = -n$ (i.e of side length 2^{-n}), that is, $\mathbb{E}_n f(x) = \mathrm{av}_Q f$ for every $x \in Q$ with $Q \in \mathfrak{Q}_n$. Define the martingale difference operator $\mathbb{D}_n := \mathbb{E}_{n+1} - \mathbb{E}_n$ for $n \in \mathbb{Z}$. We shall frequently use the familiar properties

$$\mathbb{D}_k^2 = \mathbb{D}_k, \qquad \text{and} \qquad \mathbb{D}_k\mathbb{D}_{k'} = \mathbb{D}_{k'}\mathbb{D}_k = 0 \text{ for } k \neq k', \tag{3.1}$$

as well as

$$\|P_{k_2}\mathbb{D}_{k_1}\|_{L^p\to L^p} \lesssim 2^{-(k_1-k_2)/p} \quad \text{if } k_1 \geq k_2, \tag{3.2a}$$

$$\|\mathbb{E}_{k_1}P_{k_2}\|_{L^p\to L^p} \lesssim 2^{-(k_1-k_2)/p'} \quad \text{if } k_2 \geq k_1, \tag{3.2b}$$

for $1 \leq p \leq \infty$. From (3.2a), (3.2b) and using duality,

$$\|\mathbb{D}_{k_1}P_{k_2}\|_{L^p\to L^p} + \|P_{k_2}\mathbb{D}_{k_1}\|_{L^p\to L^p} \lesssim 2^{-|k_1-k_2|\min\{1/p,1/p'\}} \tag{3.2c}$$

for $1 \leq p \leq \infty$ and $k_1, k_2 \in \mathbb{Z}$. The bounds (3.2) follow by standard computations exploiting cancellation of P_k and $\mathbb{D}_k$ (see e.g. [5, Ch. 3]). They will allow us for example to interchange $\mathbb{D}_k$ and P_k in L^p bounds for $1 < p < \infty$. We note the reproducing formula

$$f = \mathbb{E}_{1-L(S_0)}f + \sum_{k>-L(S_0)} \mathbb{D}_k f. \tag{3.3}$$

We start from $\mathbb{E}_{1-L(S_0)}$ rather than from $\mathbb{E}_{-L(S_0)}$ because that will be convenient in Section 5. Consider the localized dyadic square function

$$\mathfrak{g}_{S_0}f(x) = |\mathbb{E}_{1-L(S_0)}f(x)| + \Big(\sum_{k>-L(S_0)} |\mathbb{D}_k f(x)|^2\Big)^{1/2}. \tag{3.4}$$

Note that $\mathfrak{g}_{S_0}f$ is supported on S_0, by definition. By a trivial L^2-bound and standard Calderón–Zygmund decomposition (and using Khintchine's inequality), it is well-known that $\mathfrak{g}_{S_0}$ satisfies an L^p bound for all $1 < p < \infty$ with constant only depending on p, d. This is also a special case of Burkholder's square-function estimate for more general martingales [10].

It will also be convenient to work with a slightly larger and more robust square function. Let $Q_k(x)$ be the unique dyadic cube of sidelength 2^{-k} containing x and define a dyadic square function in the spirit of Peetre [42],

$$\mathbb{G}_{S_0}f(x) = |\mathbb{E}_{1-L(S_0)}f(x)| + \Big(\sum_{k>-L(S_0)} \sup_{y\in Q_k(x)} |\mathbb{D}_k f(y)|^2\Big)^{1/2}. \tag{3.5}$$

Since $\mathbb{D}_k f$ is constant on dyadic cubes of sidelength 2^{-k-1} it is easy to see that $\sup_{y\in Q_k(x)} |\mathbb{D}_k f(y)| \lesssim M_{\mathrm{HL}}\mathbb{D}_k f(x)$ where M_{HL} is the Hardy–Littlewood

maximal operator. Using the Fefferman–Stein inequalities [26] for the vector-valued Hardy–Littlewood maximal operator and the bounds for $\mathfrak{g}_{S_0}$ we obtain, for $1 < p < \infty$,

$$\|\mathbb{G}_{S_0} f\|_{L^p(S_0)} \le \mathcal{C}_{\mathrm{sq},p} \|f\|_{L^p(S_0)}, \tag{3.6}$$

where $\mathcal{C}_{\mathrm{sq},p}$ only depends on p and d.

3.2. *Atomic decomposition.* We will now perform an atomic decomposition of f using the local square function $\mathbb{G}_{S_0} f$, following ideas in [13] (see also [47]). Given $\mu \in \mathbb{Z}$, consider the level sets

$$\Omega_\mu \equiv \Omega_\mu[f] := \{x \in S_0 \,:\, \mathbb{G}_{S_0} f(x) > 2^\mu\}, \tag{3.7}$$

and the open sets

$$\widetilde{\Omega}_\mu \equiv \widetilde{\Omega}_\mu[f] := \{x \in \mathbb{R}^d : M_{\mathrm{HL}} \mathbb{1}_{\Omega_\mu[f]}(x) > 2^{-1}(10\sqrt{d})^{-d}\}. \tag{3.8}$$

Note that $\widetilde{\Omega}_\mu$ is not necessarily contained in S_0. Of course, $\Omega_\mu \subseteq \widetilde{\Omega}_\mu$ and $\widetilde{\Omega}_{\mu_2} \subseteq \widetilde{\Omega}_{\mu_1}$ if $\mu_1 \le \mu_2$. By the Hardy–Littlewood theorem, one has

$$|\widetilde{\Omega}_\mu| \le C_d |\Omega_\mu|, \text{ with } C_d = 5^d 2 (10\sqrt{d})^d \tag{3.9}$$

and Chebyshev's inequality and (3.6) imply

$$|\Omega_\mu| \le 2^{-\mu p} \|\mathbb{G}_{S_0} f\|_p^p \le 2^{-\mu p} \mathcal{C}_{\mathrm{sq},p}^p \|f\|_{L^p(S_0)}^p \tag{3.10}$$

for all $1 < p < \infty$. Let $\mathcal{R}_\mu \equiv \mathcal{R}_\mu[f]$ denote the family of all dyadic cubes $R \subsetneq S_0$ satisfying

$$|R \cap \Omega_\mu| > |R|/2, \tag{3.11}$$

$$|R \cap \Omega_{\mu+1}| \le |R|/2. \tag{3.12}$$

Lemma 3.1. *For all $\mu \in \mathbb{Z}$ and all $R \in \mathcal{R}_\mu$ we have $10\sqrt{d}R \subset \widetilde{\Omega}_\mu$.*

Proof. Let $c = 10\sqrt{d}$. For every $x \in cR$, $R \in \mathcal{R}_\mu$, we have by (3.11)

$$\frac{1}{2(10\sqrt{d})^d} < \frac{|R \cap \Omega_\mu|}{c^d |R|} \le \frac{|cR \cap \Omega_\mu|}{|cR|} = \frac{1}{|cR|} \int_{cR} \mathbb{1}_{\Omega_\mu} \le M_{\mathrm{HL}} \mathbb{1}_{\Omega_\mu}(x).$$

By the definition (3.8), this implies $cR \subset \widetilde{\Omega}_\mu$ for all $R \in \mathcal{R}_\mu$. □

The lemma implies in particular that $R \subseteq \widetilde{\Omega}_\mu$ for $R \in \mathcal{R}_\mu$. This and (3.12) further imply

$$|R \cap (\widetilde{\Omega}_\mu \backslash \Omega_{\mu+1})| = |R| - |R \cap \Omega_{\mu+1}| \ge |R|/2. \tag{3.13}$$

We also note that for every dyadic cube $R \subsetneq S_0$ there exists a unique $\mu \in \mathbb{Z}$ such that $R \in \mathcal{R}_\mu$.

Fix $\mu \in \mathbb{Z}$. Let $\widetilde{\mathcal{W}}_\mu = \{W\}$ denote a family of standard dyadic Whitney cubes [49, §VI.1] whose union is the open set $\widetilde{\Omega}_\mu$, which satisfy

$$\operatorname{diam}(W) \le \operatorname{dist}(W, \widetilde{\Omega}_\mu^{\complement}) \le 4 \operatorname{diam}(W). \tag{3.14}$$

If there exists $W \in \widetilde{\mathcal{W}}_\mu$ with $S_0 \subseteq W$, then we set

$$\mathcal{W}_\mu = \{S_0\}.$$

Otherwise we have $W \subsetneq S_0$ for all $W \in \widetilde{\mathcal{W}}_\mu$ intersecting S_0 and we set

$$\mathcal{W}_\mu = \{W \in \widetilde{\mathcal{W}}_\mu \,:\, W \cap S_0 \neq \emptyset\}.$$

The cubes in $\mathcal{R}_\mu$ have a unique ancestor in $\mathcal{W}_\mu$.

Lemma 3.2. *For each $R \in \mathcal{R}_\mu$, there exists a unique $W(R) \in \mathcal{W}_\mu$ containing R.*

Proof. If $\mathcal{W}_\mu = \{S_0\}$, then there is nothing to prove. Otherwise, let $c = 10\sqrt{d}$ and $R \in \mathcal{R}_\mu$. By Lemma 3.1, we have $cR \subseteq \widetilde{\Omega}_\mu$. Let x_R denote the center of R and let $W = W(R) \in \mathcal{W}_\mu$ such that $x_R \in W$. With this setup and (3.14) we have

$$\operatorname{dist}(x_R, (cR)^\complement) \leq \operatorname{dist}(x_R, \widetilde{\Omega}_\mu^\complement) \leq \operatorname{diam}(W) + \operatorname{dist}(W, \widetilde{\Omega}_\mu^\complement) \leq 5\operatorname{diam}(W).$$

Noting that $\operatorname{dist}(x_R, (cR)^\complement) = \frac{c\operatorname{diam}(R)}{2\sqrt{d}} = 5\operatorname{diam}(R)$, we obtain that

$$5\operatorname{diam}(R) \leq 5\operatorname{diam}(W).$$

As R and W are dyadic cubes containing x_R, we conclude that $R \subseteq W$. The uniqueness follows from the disjointess of $W \in \mathcal{W}_\mu$. □

Given $\mu \in \mathbb{Z}$ and $W \in \mathcal{W}_\mu$, the sets

$$\mathcal{R}_{W,\mu} := \{R \in \mathcal{R}_\mu \,:\, R \subseteq W\}$$

are disjoint for different W, by disjointness of the W. Also, by Lemma 3.2

$$\mathcal{R}_\mu = \bigcup_{W \in \mathcal{W}_\mu} \mathcal{R}_{W,\mu},$$

where the union is disjoint. We are now ready to define the atoms. First, for each dyadic cube $R \subsetneq S_0$ with $L(R) = -k$ let

$$e_R \equiv e_R[f] := (\mathbb{D}_k f)\mathbb{1}_R = \mathbb{D}_k(f\mathbb{1}_R).$$

We refer to the e_R as *subatoms*; they are pairwise orthogonal and $\int e_R = 0$. The subatoms are building blocks of larger *atoms* which are associated to cubes W. Given $\mu \in \mathbb{Z}$ and $W \in \mathcal{W}_\mu$, these are defined as

$$a_{W,\mu} \equiv a_{W,\mu}[f] := \sum_{R \in \mathcal{R}_{W,\mu}} e_R[f]. \tag{3.15}$$

We refer to the $a_{W,\mu}$ as *atoms*, but note that they have a non-standard normalization with respect to other sources in the literature. Indeed, if we define the coefficients

$$\gamma_{W,\mu} \equiv \gamma_{W,\mu}[f] := \Big(|W|^{-1} \sum_{R \in \mathcal{R}_{W,\mu}} \|e_R[f]\|_2^2\Big)^{1/2}, \tag{3.16}$$

then using orthogonality,

$$\|a_{W,\mu}\|_2 = |W|^{1/2}\gamma_{W,\mu}. \tag{3.17}$$

Note that $|W|^{-1/p}(\gamma_{W,\mu})^{-1}a_{W,\mu}$ corresponds to a $(p,2)$-*atom* in the classical atomic Hardy-space theory developed for $p \le 1$ (see e.g. [17]). Note that for $p \le 2$, Hölder's inequality and (3.17) imply

$$\|a_{W,\mu}\|_p \le |W|^{1/p}\gamma_{W,\mu}. \tag{3.18}$$

In view of (3.3) and the above discussion, we can write the atomic decomposition as

$$f = \mathbb{E}_{1-L(S_0)}f + \sum_{\mu\in\mathbb{Z}}\sum_{W\in\mathcal{W}_\mu} a_{W,\mu}. \tag{3.19}$$

In applications it will be useful to use the fine structure of the $a_{W,\mu}$ and further group subatoms that are at the same scale (see (3.28)).

3.3. *Properties of the atomic decomposition.* The square function $\mathbb{G}_{S_0}$ allows summation of the coefficients $|W|^{1/2}\gamma_{W,\mu}$ in ℓ^2 over the collection $\mathcal{W}_\mu$.

Lemma 3.3. *Let $\mu \in \mathbb{Z}$. Then*

$$\Big(\sum_{W\in\mathcal{W}_\mu} |W|(\gamma_{W,\mu})^2\Big)^{1/2} = \Big(\sum_{R\in\mathcal{R}_\mu} \|e_R\|_2^2\Big)^{1/2} \le 2^{\mu+3/2}|\widetilde{\Omega}_\mu|^{1/2}. \tag{3.20}$$

Proof. The first identity is by definition. Using (3.13),

$$\|e_R\|_2^2 \le |R|\|e_R\|_\infty^2 \le 2|R\cap(\widetilde{\Omega}_\mu\setminus\Omega_{\mu+1})|\,\|e_R\|_\infty^2.$$

Observe that

$$\sum_{\substack{R\in\mathcal{R}_\mu,\\ L(R)=-k}} \|e_R\|_\infty^2 \mathbb{1}_R(x) = \sup_{y\in Q_k(x)} |\mathbb{D}_k f(y)|^2.$$

Thus, the left-hand side of the square of (3.20) is

$$\sum_{k>-L(S_0)}\sum_{\substack{R\in\mathcal{R}_\mu,\\ L(R)=-k}} \|e_R\|_2^2 \le 2\int_{\widetilde{\Omega}_\mu\setminus\Omega_{\mu+1}} \sum_{k>-L(S_0)} \sup_{y\in Q_k(x)} |\mathbb{D}_k f(y)|^2\,\mathrm{d}x.$$

which by definition of $\Omega_{\mu+1}$ is bounded by $2^{2\mu+3}|\widetilde{\Omega}_\mu|$, as claimed. □

Even though the coefficients $\gamma_{W,\mu}$ incorporate ℓ^2 in their definition, there is an ℓ^p-analogue of the above lemma for $1 < p < 2$. For notational convenience, define the auxiliary function

$$F_p(x) := \Big(\sum_{\mu\in\mathbb{Z}}\sum_{W\in\mathcal{W}_\mu} (\gamma_{W,\mu}[f])^p \mathbb{1}_W(x)\Big)^{1/p}. \tag{3.21}$$

The following lemma says that $\|F_p\|_p$ is controlled by $\|f\|_p$.

Lemma 3.4. *Let $1 < p \le 2$ and $\mu \in \mathbb{Z}$. Then*

$$\Big(\sum_{W \in \mathcal{W}_\mu} |W| (\gamma_{W,\mu})^p \Big)^{1/p} \le 2^{\mu+3/2} |\widetilde{\Omega}_\mu|^{1/p} \le C_d^{1/p} 2^{\mu+3/2} |\Omega_\mu|^{1/p} \tag{3.22}$$

with C_d as in (3.9)*. Moreover,*

$$\Big(\sum_{\mu \in \mathbb{Z}} 2^{\mu p} |\Omega_\mu| \Big)^{1/p} \le (2C_d)^{1/p} \mathcal{C}_{\mathrm{sq},p} \|f\|_p \tag{3.23}$$

and

$$\|F_p\|_p \le 2^{3/2} (2C_d)^{1/p} \mathcal{C}_{\mathrm{sq},p} \|f\|_p. \tag{3.24}$$

Proof. Fix $\mu \in \mathbb{Z}$. By Hölder's inequality, the definition of $\mathcal{W}_\mu$, Lemma 3.3 and the estimate (3.9) we have

$$\begin{aligned} \sum_{W \in \mathcal{W}_\mu} |W| (\gamma_{W,\mu})^p &= \sum_{W \in \mathcal{W}_\mu} |W|^{1-p/2} (|W|^{1/2} \gamma_{W,\mu})^p \\ &\le \Big(\sum_{W \in \mathcal{W}_\mu} |W| \Big)^{1-p/2} \Big(\sum_{W \in \mathcal{W}_\mu} |W| (\gamma_{\mu,W})^2 \Big)^{p/2} \\ &\le |\widetilde{\Omega}_\mu|^{1-p/2} (2^{2\mu+3} |\widetilde{\Omega}_\mu|)^{p/2} \le C_d |\Omega_\mu| 2^{\mu p + 3p/2}. \end{aligned}$$

This proves (3.22). To prove (3.24) we use the definition of Ω_μ and (3.6) to estimate

$$\begin{aligned} &\sum_{\mu \in \mathbb{Z}} 2^{\mu p} |\Omega_\mu| = \sum_{\mu \in \mathbb{Z}} (1 - 2^{-p})^{-1} \int_{2^{\mu-1}}^{2^\mu} p \alpha^{p-1} \, \mathrm{d}\alpha \big| \{ x : \mathbb{G}_{S_0} f(x) > 2^\mu \} \big| \\ &\le (1 - 2^{-p})^{-1} \int_0^\infty p \alpha^{p-1} |\{ \mathbb{G}_{S_0} f(x) > \alpha \}| \, \mathrm{d}\alpha \le 2 \|\mathbb{G}_{S_0} f\|_p^p \le 2 C_d \mathcal{C}_{\mathrm{sq},p}^p \|f\|_p^p \end{aligned}$$

which gives (3.23) and further implies

$$\|F_p\|_p^p = \sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_\mu} |W| (\gamma_{W,\mu})^p \le 2^{3p/2} C_d \sum_{\mu \in \mathbb{Z}} 2^{\mu p} |\Omega_\mu| \le 2 C_d 2^{3p/2} \mathcal{C}_{\mathrm{sq},p}^p \|f\|_p^p,$$

as desired. □

3.4. *Fine structure analysis of atomic decompositions.* Note that by (3.17)

$$\Big\| \sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_\mu} a_{W,\mu} \Big\|_2 = \|F_2\|_2 \tag{3.25}$$

via (3.17), so for $p = 2$ Lemma 3.4 recovers the trivial inequality

$$\Big\| \sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_\mu} a_{W,\mu} \Big\|_2 \lesssim \|f\|_2, \tag{3.26}$$

which follows directly from (3.19). There does not seem to be an L^p analogue of this inequality for $1 < p < 2$, because there appears to be no immediate relation between the L^p norms of $\sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_\mu} a_{W,\mu}$ and F_p of the type (3.25). However, we shall rely on other useful analogues where

either the atoms or subatoms are of a fixed scale (see Lemmata 3.7 and 3.9 below); these will then be used in conjunction with a weak orthogonality or Littlewood–Paley type argument based on properties of the operator that is estimated.

We first consider variants where the subatoms correspond to a fixed dyadic scale. For $k \in \mathbb{Z}$, $\mu \in \mathbb{Z}$ and $W \in \mathcal{W}_\mu$, define the families of cubes

$$\mathcal{R}_\mu^k := \{R \in \mathcal{R}_\mu \,:\, L(R) = -k\},$$
$$\mathcal{R}_{W,\mu}^k := \{R \in \mathcal{R}_\mu^k \,:\, R \subseteq W\},$$

the coefficients

$$\gamma_{W,\mu}^k \equiv \gamma_{W,\mu}^k[f] := \Big(|W|^{-1} \sum_{R \in \mathcal{R}_{W,\mu}^k} \|e_R[f]\|_2^2\Big)^{1/2}, \tag{3.27}$$

and the fixed scale atoms

$$a_{W,\mu}^k \equiv a_{W,\mu}^k[f] := \sum_{R \in \mathcal{R}_{W,\mu}^k} e_R[f]. \tag{3.28}$$

Note that if $k \le -L(S_0)$, then $\mathcal{R}_\mu^k = \mathcal{R}_{W,\mu}^k = \emptyset$, $\gamma_{W,\mu}^k = 0$ and $a_{W,\mu}^k = 0$ by definition. We have

$$\gamma_{W,\mu} = \Big(\sum_{k > -L(S_0)} (\gamma_{W,\mu}^k)^2\Big)^{1/2} \qquad \text{and} \qquad a_{W,\mu} = \sum_{k > -L(S_0)} a_{W,\mu}^k.$$

We observe that the inequality (3.18) continues to hold when all subatoms are at a fixed scale.

Lemma 3.5. *Let $k > -L(S_0)$, $\mu \in \mathbb{Z}$ and $W \in \mathcal{W}_\mu$. If $1 < p \le 2$, then*

$$\|a_{W,\mu}^k\|_p \le |W|^{1/p} \gamma_{W,\mu}^k.$$

Proof. By two applications of Hölder's inequality,

$$\begin{aligned}\Big(\sum_{R \in \mathcal{R}_{W,\mu}^k} \|e_R\|_{L^p}^p\Big)^{1/p} &\le \Big(\sum_{R \in \mathcal{R}_{W,\mu}^k} \|e_R\|_{L^2}^2\Big)^{1/2} \Big(\sum_{R \in \mathcal{R}_{W,\mu}^k} |R|\Big)^{\frac{1}{p}-\frac{1}{2}} \\ &\le \Big(\sum_{R \in \mathcal{R}_{W,\mu}^k} \|e_R\|_{L^2}^2\Big)^{1/2} |W|^{\frac{1}{p}-\frac{1}{2}} = \gamma_{W,\mu}^k |W|^{1/p}.\end{aligned}$$

The result now follow by disjointness of the cubes in $R \in \mathcal{R}_{W,\mu}^k$. □

For the remainder of this section we fix parameters

$$Q \in \mathfrak{Q}(S_0) \quad \text{and} \quad \mu_{\min} \in \mathbb{Z} \cup \{-\infty\}. \tag{3.29}$$

For the time being the reader may pretend that $Q = S_0$ and $\mu_{\min} = -\infty$, but we will need the additional localization when combining the atomic

decomposition with an appropriate Calderón–Zygmund decomposition in §5. With this in mind, set $\mathcal{W}_{Q,\mu} = \{W \in \mathcal{W}_\mu : W \subseteq Q\}$ and

$$b_Q = \sum_{\mu > \mu_{\min}} \sum_{W \in \mathcal{W}_{Q,\mu}} a_{W,\mu}, \tag{3.30a}$$

$$b_Q^k = \sum_{\mu > \mu_{\min}} \sum_{W \in \mathcal{W}_{Q,\mu}} a_{W,\mu}^k. \tag{3.30b}$$

Define also

$$\beta_{Q,p} = \Big(\sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_{Q,\mu}} |W| (\gamma_{W,\mu}[f])^p \Big)^{1/p}, \tag{3.31a}$$

$$\beta_{Q,p}^k = \Big(\sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_{Q,\mu}} |W| (\gamma_{W,\mu}^k[f])^p \Big)^{1/p} \tag{3.31b}$$

and observe that $\|F_p \mathbf{1}_Q\|_p = \beta_{Q,p}$. Note that $\mathcal{W}_{Q,\mu}$ and b_Q, b_Q^k, $\beta_{Q,p}$, $\beta_{Q,p}^k$ all depend on the function f. Also observe that the truncation in μ is omitted in the definitions of $\beta_{Q,p}, \beta_{Q,p}^k$. Our first observation is a variant of (3.25) in L^p for a fixed scale k.

Lemma 3.6. *Let $1 < p \le 2$ and $k > -L(S_0)$. Then*

$$\|b_Q^k\|_p \le \beta_{Q,p}^k.$$

Proof. Note that

$$\|b_Q^k\|_{L^p} = \Big(\sum_{\mu > \mu_{\min}} \sum_{W \in \mathcal{W}_{\mu,Q}} \sum_{R \in \mathcal{R}_{W,\mu}^k} \|e_R\|_{L^p}^p \Big)^{1/p}$$

as all the cubes occurring in the definition of b_Q^k are disjoint. By Lemma 3.5,

$$\|b_Q^k\|_{L^p} \le \Big(\sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_{Q,\mu}} |W| (\gamma_{W,\mu}^k)^p \Big)^{1/p} = \beta_{Q,p}^k. \tag{3.32}$$ □

We can sum the coefficients $\{\beta_{Q,p}^k\}_{k \in \mathbb{Z}}$ in ℓ^2.

Lemma 3.7. *If $1 < p \le 2$, then*

$$\Big(\sum_{k > -L(S_0)} (\beta_{Q,p}^k)^2 \Big)^{1/2} \le \|F_p \mathbb{1}_Q\|_p = \beta_{Q,p}.$$

Proof. By the definitions and Minkowski's inequality,

$$\Big(\sum_{k>-L(S_0)}(\beta^k_{Q,p})^2\Big)^{1/2} = \Big(\sum_{k>-L(S_0)}\Big(\sum_{\mu>\mu_{\min}}\sum_{W\in\mathcal{W}_{Q,\mu}}(\gamma^k_{W,\mu})^p|W|\Big)^{2/p}\Big)^{1/2}$$
$$\le \Big(\sum_{\mu>\mu_{\min}}\sum_{W\in\mathcal{W}_{Q,\mu}}\big(\sum_{k\in\mathbb{Z}}(\gamma^k_{W,\mu})^2\big)^{p/2}|W|\Big)^{1/p}$$
$$= \Big(\sum_{\mu>\mu_{\min}}\sum_{W\in\mathcal{W}_{Q,\mu}}(\gamma_{W,\mu})^p|W|\Big)^{1/p} \le \|F_p\mathbb{1}_Q\|_p = \beta_{Q,p},$$

as desired. □

There is a second variant which consists in fixing the scale of the atoms rather in addition to that of the subatoms. Given integers $k\in\mathbb{Z}$ and $n\ge 0$, define

$$b^{k,n}_Q = \sum_{\mu>\mu_{\min}}\sum_{\substack{W\in\mathcal{W}_{Q,\mu},\\ L(W)=-k+n}} a^k_{W,\mu}, \tag{3.33}$$

and

$$\beta^{k,n}_{Q,p} = \Big(\sum_{\mu\in\mathbb{Z}}\sum_{\substack{W\in\mathcal{W}_{Q,\mu},\\ L(W)=-k+n}}|W|(\gamma^k_{W,\mu}[f])^p\Big)^{1/p}. \tag{3.34}$$

Note that by definition, $b^{k,n}_Q = 0$ and $\beta^{k,n}_{Q,p} = 0$ unless $k > -L(S_0)$. Lemma 3.6 continues to hold for these fixed-scale W versions. A crucial observation is that we obtain a gain if we move to a larger Lebesgue exponent $r > p$. This will allow us to think of the case $W = R$ as the dominant contribution. This observation will be crucial in later proofs.

Lemma 3.8. *Let $1 < p \le r \le 2$. Then for $k > -L(S_0)$ and $n \ge 0$,*

$$\|b^{k,n}_Q\|_r \le 2^{-nd(\frac{1}{p}-\frac{1}{r})}2^{kd(\frac{1}{p}-\frac{1}{r})}\beta^{k,n}_{Q,p}.$$

Proof. Arguing as in the proof of Lemma 3.6,

$$\|b^{k,n}_Q\|_{L^r} \le \Big(\sum_{\mu\in\mathbb{Z}}\sum_{\substack{W\in\mathcal{W}_{Q,\mu}\\ L(W)=-k+n}}|W|(\gamma^k_{W,\mu})^r\Big)^{1/r}.$$

Using the embedding $\ell^p\subseteq\ell^r$ for $p<r$,

$$\|b^{k,n}_Q\|_{L^r} \le \Big(\sum_{\mu\in\mathbb{Z}}\sum_{\substack{W\in\mathcal{W}_{Q,\mu}\\ L(W)=-k+n}}(\gamma^k_{W,\mu})^p|W||W|^{\frac{p}{r}-1}\Big)^{1/p}$$
$$\le 2^{(k-n)d(\frac{1}{p}-\frac{1}{r})}\beta^{k,n}_{Q,p}.$$ □

We also have the following variant of Lemma 3.7.

Lemma 3.9. *For every $n \geq 0$,*

$$\Big(\sum_{k > -L(S_0)} (\beta_{Q,p}^{k,n})^p \Big)^{1/p} \leq \|F_p \mathbb{1}_Q\|_p = \beta_{Q,p}.$$

Proof. This follows since k and n are coupled. □

This lemma allows us to sum in ℓ^p rather than ℓ^2 provided that the quantity $L(W) - L(R)$ is constant. In applications, this constitutes a great advantage which permits us to prove endpoint bounds.

4. The base case

Recalling (3.29) we set

$$\mu_{\min} = -\infty$$

throughout this section. We first note the following observations. Let $p \leq r \leq q$. By rescaling (recall (2.2) and (2.3)), we have for $j_0 > -k$,

$$\begin{aligned} \|T_k^{(j_0)}\|_{L^r \to L^q} &= 2^{kd(\frac{1}{r} - \frac{1}{q})} \|\varphi m(2^k \cdot) * \widehat{\Psi_{j_0+k}}\|_{M^{r \to q}} \\ &= 2^{-j_0 d(\frac{1}{p} - \frac{1}{q})} 2^{-kd(\frac{1}{p} - \frac{1}{r})} A_{p,r,q}^{k,j_0+k}. \end{aligned} \tag{4.1}$$

It is well-known that sparse domination for single spatial scale operators follows from certain rescaled $L^p \to L^q$ estimates (see, for instance, [5, §3.1]). In our case, it suffices to verify the $L^p \to L^q$ estimates in the following two lemmata; in both the implicit constants do not depend on j_0 and F. Recall the definition $\mathcal{C}_{p,r,q}(j_0) = \sup_{k > -j_0} A_{p,r,q}^{k,j_0+k}$ (see (2.13)).

Lemma 4.1. *Let $1 < p \leq 2 \leq q \leq p'$ and $j_0 \in \mathbb{Z}$. Let S_0 be a dyadic cube of side length 2^{j_0} and let $f \in L^p$ be supported in S_0. Then*

$$\|\mathcal{T}_{j_0} f\|_q \lesssim 2^{-j_0 d(\frac{1}{p} - \frac{1}{q})} \mathcal{C}_{p,p,q}(j_0) \|f\|_p. \tag{4.2}$$

Lemma 4.2. *Let $1 < p < r \leq q \leq 2$ and $j_0 \in \mathbb{Z}$. Let S_0 be a dyadic cube of side length 2^{j_0} and let $f \in L^p$ be supported in S_0. Then*

$$\|\mathcal{T}_{j_0} f\|_q \lesssim 2^{-j_0 d(\frac{1}{p} - \frac{1}{q})} \big(\mathcal{C}_{p,p,q}(j_0) + \mathcal{C}_{p,r,q}(j_0) \big) \|f\|_p. \tag{4.3}$$

Reduction of Lemmata 2.6, 2.7 to Lemmata 4.1, 4.2. Both reductions use the same argument; we therefore abbreviate by $\mathcal{C}(j_0)$ the respective constants $\mathcal{C}_{p,p,q}(j_0)$, for $q \leq 2$, and $\mathcal{C}_{p,p,q}(j_0) + \mathcal{C}_{p,r,q}(j_0)$, for $q \geq 2$. Keeping in mind that f_1 vanishes in $S_0^{\complement}$ we estimate

$$\begin{aligned} &|\langle \mathcal{T}_{j_0} f_1, f_2 \rangle| = |\langle \mathcal{T}_{j_0} f_1, f_2 \mathbb{1}_{3S_0} \rangle| \leq \|\mathcal{T}_{j_0} f_1\|_q \|f_2 \mathbb{1}_{3S_0}\|_{q'} \\ &\lesssim \mathcal{C}(j_0) 2^{-j_0 d(\frac{1}{p} - \frac{1}{q})} \|f_1\|_p \|f_2 \mathbb{1}_{3S_0}\|_{q'} \lesssim \mathcal{C}(j_0) \langle f_1 \rangle_{S_0,p} \langle f_2 \rangle_{3S_0,q'} |3S_0|, \end{aligned} \tag{4.4}$$

which gives the desired sparse bounds (2.14) and (2.15). □

It remains to prove Lemmata 4.1 and 4.2. The proof of the former is a short standard argument based on Littlewood–Paley inequalities. The proof of the latter is longer and relies on the atomic decomposition discussed in §3.

4.1. *The case $q \geq 2$: Proof of Lemma 4.1.* By the Littlewood-Paley inequality $\|\sum_k L_k P_k F_k\|_q \lesssim \|(\sum_k |F_k|^2)^{1/2}\|_q$ and setting $F_k = P_k f$ and applying (2.1c), we get

$$\|\mathcal{T}_{j_0} f\|_q \lesssim \Big\| \Big(\sum_{k\in\mathbb{F}} |T_k^{(j_0)} P_k f|^2 \Big)^{1/2} \Big\|_q \lesssim \Big(\sum_{k\in\mathbb{F}} \|T_k^{(j_0)} P_k f\|_q^2 \Big)^{1/2} \tag{4.5}$$

where we applied Minkowski's inequality in $L^{q/2}$. By (4.1) with $r = p$ we get

$$\|\mathcal{T}_{j_0} f\|_q \lesssim \mathcal{C}_{p,p,q}(j_0) 2^{-j_0 d(\frac{1}{p}-\frac{1}{q})} \Big(\sum_{k\in\mathbb{Z}} \|P_k f\|_p^2 \Big)^{1/2}.$$

Using $p \leq 2$ and Minkowski's inequality in $L^{2/p}$ we obtain

$$\begin{aligned} \|\mathcal{T}_{j_0} f\|_q &\lesssim \mathcal{C}_{p,p,q}(j_0) 2^{-j_0 d(\frac{1}{p}-\frac{1}{q})} \Big\| \Big(\sum_{k\in\mathbb{Z}} |P_k f|^2 \Big)^{1/2} \Big\|_p \\ &\lesssim \mathcal{C}_{p,p,q}(j_0) 2^{-j_0 d(\frac{1}{p}-\frac{1}{q})} \|f\|_p \end{aligned}$$

and (4.2) is proved. □

4.2. *The case $q < 2$: Proof of Lemma 4.2.* Let $f \in L^p(S_0)$, with $L(S_0) = j_0$, and decompose using (3.3), for fixed $k \in \mathbb{Z}$,

$$f = \mathbb{E}_{1-j_0} f + \sum_{m > -j_0 - k} \mathbb{D}_{k+m} f.$$

Next we split

$$\mathcal{T}_{j_0} f = I + II_1 + II_2,$$

where

$$I = \sum_{k\in\mathbb{F}} T_k^{(j_0)} P_k P_k [\mathbb{E}_{1-j_0} f]$$

and II_1 and II_2 are defined in terms of the additional decomposition $\mathbb{D}_{k+m} f = \sum_{n\geq 0} b_{S_0}^{k+m,n}$ as

$$II_1 = \sum_{m\geq 0} \sum_{0\leq n\leq 2m} \sum_{k\in\mathbb{F}} II_{m,n,k},$$

$$II_2 = \sum_{m<0} \sum_{n\geq 0} \sum_{k\in\mathbb{F}} II_{m,n,k} + \sum_{m\geq 0} \sum_{n>2m} \sum_{k\in\mathbb{F}} II_{m,n,k}, \text{ where}$$

$$II_{m,n,k} = P_k T_k^{(j_0)} P_k \mathbb{D}_{k+m} b_{S_0}^{k+m,n}.$$

It is useful to keep in mind that $II_{m,n,k} = 0$ unless $k > -j_0$ and $k + m > -j_0$. Our goal is to control the L^q norm of the three terms by a constant

times $\mathcal{C}_{p,r,q}(j_0)2^{-j_0 d(1/p-1/q)}\|f\|_{L^p(S_0)}$. For the first two terms we get the better bound with $r=p$.

Estimation of $\|I\|_q$. For the term I we take $r=p$ and estimate, using (3.2b) and (4.1) with $r=p$,

$$\begin{aligned}\|I\|_q &\lesssim \sum_{k\in F} \|T_k^{(j_0)}\|_{L^p\to L^q}\|P_k\mathbb{E}_{1-j_0}\|_{L^p\to L^p}\|f\|_p\\ &\lesssim 2^{-j_0 d(\frac1p-\frac1q)}\sum_{k\in F} A_{p,p,q}^{k,j_0+k}2^{-(k+j_0)/p'}\|f\|_p\\ &\lesssim 2^{-j_0 d(\frac1p-\frac1q)}\mathcal{C}_{p,p,q}(j_0)\sum_{\ell>0}2^{-\ell/p'}\|f\|_p\end{aligned}$$

and we get the desired bound.

Estimation of $\|II_1\|_q$. By the almost-orthogonality of the P_k and the resulting inequality $\|\sum_k P_kF_k\|_q\lesssim(\sum_k\|F_k\|_q^q)^{1/q}$ we get

$$\|II_1\|_q\lesssim\sum_{m\ge0}\sum_{0\le n\le 2m}\Big(\sum_{k\in F}\|II_{m,n,k}\|_q^q\Big)^{1/q}. \tag{4.6}$$

Now, by (4.1), (3.2) and Lemma 3.8, we have for all $p\le r\le q$ that

$$\begin{aligned}\|II_{m,n,k}\|_q &\lesssim \|T_k^{(j_0)}\|_{L^r\to L^q}\|P_k\mathbb{D}_{k+m}\|_{L^r\to L^r}\|b_{S_0}^{k+m,n}\|_r\\ &\lesssim \mathcal{C}_{p,r,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}2^{-kd(\frac1p-\frac1r)}2^{-|m|/r'}2^{-nd(\frac1p-\frac1r)}2^{(k+m)d(\frac1p-\frac1r)}\beta_{S_0,p}^{k+m,n}.\end{aligned} \tag{4.7}$$

We use this with $r=p$ and estimate the right-hand side of (4.6) by a constant times

$$\begin{aligned}&2^{-j_0 d(\frac1p-\frac1q)}\mathcal{C}_{p,p,q}(j_0)\sum_{m\ge0}\sum_{0\le n\le 2m}2^{-m/p'}\Big(\sum_{k>-j_0}(\beta_{S_0,p}^{k+m,n})^q\Big)^{1/q}\\ &\qquad\lesssim 2^{-j_0 d(\frac1p-\frac1q)}\mathcal{C}_{p,p,q}(j_0)\sum_{m\ge0}\sum_{0\le n\le 2m}2^{-m/p'}\|f\|_p\end{aligned}$$

using $\ell^q\subseteq\ell^p$, Lemma 3.9 and (3.24). Altogether

$$\begin{aligned}\|II_1\|_q &\lesssim \sum_{m\ge0}\sum_{0\le n\le 2m}\Big(\sum_{k>-j_0}\|II_{m,n,k}\|_q^q\Big)^{1/q}\\ &\lesssim \mathcal{C}_{p,p,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}\sum_{m\ge0}(1+2m)2^{-m/p'}\|f\|_p\\ &\lesssim \mathcal{C}_{p,p,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}\|f\|_p,\end{aligned}$$

which finishes the estimation of $\|II_1\|_q$.

Estimation of $\|II_2\|_q$. By the almost orthogonality of the P_k we have

$$\Big\|\sum_{k\in F} P_k II_{m,n,k}\Big\|_q \lesssim \Big(\sum_{k\in F}\|II_{m,n,k}\|_q^q\Big)^{1/q} \lesssim \Big(\sum_{k\in F}\|II_{m,n,k}\|_q^r\Big)^{1/r}$$

We use (4.7) with $r>p$, followed by the embedding $\ell^r\subseteq\ell^p$, Lemma 3.9 and (3.24) to deduce

$$\begin{aligned}&\Big(\sum_{k\in F}\|II_{m,n,k}\Big\|_q^r\Big)^{1/r}\\ &\lesssim \mathcal{C}_{p,r,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}2^{-|m|/r'}2^{(m-n)d(\frac1p-\frac1r)}\Big(\sum_{k\in F}(\beta^{k+m,n}_{S_0,p})^r\Big)^{1/r}\\ &\lesssim \mathcal{C}_{p,r,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}2^{-|m|/r'}2^{(m-n)d(\frac1p-\frac1r)}\|f\|_p\end{aligned}$$

and therefore, using $r>p$,

$$\|II_2\|_q\lesssim\sum_{m\in\mathbb{Z}}\sum_{n\geq\max\{0,2m\}}\Big(\sum_{k>-j_0}\|II_{m,n,k}\|_q^q\Big)^{1/q}\lesssim\mathcal{C}_{p,r,q}(j_0)2^{-j_0 d(\frac1p-\frac1q)}\|f\|_p.$$

This finishes the estimation of $\|II_2\|_q$ and the proof of the lemma. □

5. Combination of atomic and Calderón–Zygmund decompositions

Let S_0 be a dyadic cube, let f_1 and f_2 be given functions. Assume that f_1 is supported in S_0 and that f_2 is supported in $3S_0$. In analogy to the Calderón–Zygmund decomposition we decompose the functions f_1, f_2 given some threshold parameters $\alpha_1,\alpha_2>0$; these will be defined as

$$\alpha_1=\langle f_1\rangle_{S_0,p},\qquad \alpha_2=\langle f_2\rangle_{3S_0,q'}. \tag{5.1}$$

The decomposition of f_2 will be essentially based on a Calderón–Zygmund decomposition at level α_2, see (5.12), (5.13) below. We describe the decomposition of f_1 which is more involved and essentially based on the atomic decomposition introduced in §3. The idea of combining atomic and Calderón–Zygmund decompositions was previously used in [31], and can be traced back to [16], although here we need a different variant.

In the proof we will use two large constants U_1, U_2 which need to significantly exceed various constants in standard maximal or square function inequalities, or combinations thereof; we shall see that any choice of U_1, U_2 with

$$U_1\geq(1-\gamma)^{-1/p}(2^{100}d)^{d/p}\mathcal{C}_{\mathrm{sq},p}\,, \tag{5.2a}$$

$$U_2\geq(1-\gamma)^{-1/q'}(2^{100}d)^{d/q'}\,, \tag{5.2b}$$

and $\mathcal{C}_{\mathrm{sq},p}$ as in (3.6), will work.

We start writing

$$f_1 = g_1 + b_1, \tag{5.3}$$

with the "good" function defined as

$$g_1 = \mathbb{E}_{-L(S_0)+1} f_1 + \sum_{k>-L(S_0)} g_1^k, \tag{5.4a}$$

where

$$g_1^k = \sum_{\mu\in\mathbb{Z}\,:\,2^\mu\le U_1\alpha_1} \sum_{R\in\mathcal{R}_\mu^k} e_R \tag{5.4b}$$

and the "bad" function defined as $b_1 = f_1 - g_1$, i.e. $b_1 = \sum_{k>-L(S_0)} b_1^k$ with

$$b_1^k = \sum_{\mu\in\mathbb{Z}\,:\,2^\mu> U_1\alpha_1} \sum_{R\in\mathcal{R}_\mu^k} e_R. \tag{5.5}$$

Note that $b_1^k = g_1^k = 0$ for $k \le -L(S_0)$. Clearly,

$$|\mathbb{E}_{-L(S_0)+1} f_1(x)| \le 2^{d/p}\alpha_1 \le U_1\alpha_1 \quad \text{ for all } x \in S_0.$$

Furthermore, the square function associated with the $\{g_1^k\}_{k>-L(S_0)}$ is pointwise bounded by $2U_1\alpha_1$. This is analogous to the L^∞ estimate for the "good" function in a standard Calderón–Zygmund decomposition.

Lemma 5.1. *For almost every $x \in S_0$ we have*

$$\Big(\sum_{k>-L(S_0)} |g_1^k(x)|^2\Big)^{1/2} \le 2U_1\alpha_1.$$

Proof. Fix $x \in S_0$. Let $\digamma$ be a finite family of indices with $k > -L(S_0)$. It suffices to show that

$$\Big(\sum_{k\in\digamma} |g_1^k(x)|^2\Big)^{1/2} \le 2U_1\alpha_1. \tag{5.6}$$

Let

$$\mathfrak{R}_x = \Big\{R : R \in \bigcup_{2^\mu\le U_1\alpha_1} \mathcal{R}_\mu,\ x \in R,\ L(R) = -k \text{ for some } k \in \digamma\Big\}.$$

Note that these are the only cubes contributing to $\sum_{k\in\digamma} |g_1^k(x)|^2$. We can assume that $\mathfrak{R}_x \neq \emptyset$, as otherwise $\sum_{k\in\digamma} |g_1^k(x)|^2 = 0$ and the inequality is trivial. Next, let $\mathfrak{R}_{x,k} = \{R \in \mathfrak{R}_x, L(R) = -k\}$ and let $k_\circ(x)$ be the maximal integer $k \in \digamma$ for which $\mathfrak{R}_{x,k}$ is non-empty. Note $k_\circ(x)$ exists as $\mathfrak{R}_x$ is non-empty and $\digamma$ is finite. Moreover, observe that $\mathfrak{R}_{x,k}$ is either empty or consists only of one (half-open) cube. Let $R_{x,k_\circ(x)} \in \mathfrak{R}_{x,k_\circ(x)}$. By definition, there exists a unique μ_x with $2^{\mu_x} \le U_1\alpha_1$ and $R_{x,k_\circ(x)} \in \mathcal{R}_{\mu_x}$. Moreover, in view of (3.12), there exists $w_x \in R_{x,k_\circ(x)}\backslash\Omega_{\mu_x+1}$. Thus, $\mathbb{G}_{S_0} f_1(w_x) \le 2^{\mu_x+1} \le 2U_1\alpha_1$. Note that by the maximality of $k_\circ(x)$, we have $w_x \in R$ for all $R \in \mathfrak{R}_x$.

Consequently,

$$\Big(\sum_{k\in F}|g_1^k(x)|^2\Big)^{\frac12}=\Big(\sum_{k\in F}\Big|\sum_{R\in\mathfrak{R}_{x,k}}e_R(x)\Big|^2\Big)^{\frac12}\le \mathbb{G}_{S_0}f_1(w_x)\le 2\,U_1\alpha_1.\quad\square$$

The above lemma will be used in the proof of the sparse bound for dealing with the term that involves g_1: see §7.1 for the case $q<2$ and §7.2 for the case $q\ge 2$.

We need a further refined decomposition of the bad parts b_1^k. Recall that by Lemma 3.4 the function $F_{1,p}=F_p$ satisfies

$$\|F_{1,p}\|_{L^p(S_0)}\le 2^{3/2}(2C_d)^{1/p}\mathcal{C}_{\mathrm{sq},p}|S_0|^{1/p}\alpha_1, \tag{5.7}$$

where $C_d=5^d2(10\sqrt{d})^d$ (defined in (3.9)).

Our next goal is to perform a Calderón–Zygmund decomposition so that this inequality continues to hold for smaller cubes. We now bring in the second function f_2. Let $\mu(\alpha_1)$ be the smallest integer μ such that $2^\mu>U_1\alpha_1$. Define

$$\mathcal{O}=\mathcal{O}_1\cup\mathcal{O}_2, \tag{5.8}$$

where

$$\mathcal{O}_1:=\widetilde{\Omega}_{\mu(\alpha_1)}[f_1]\cup\{x: M_{\mathrm{HL}}(F_{1,p}^p)(x)>U_1^p\alpha_1^p\}, \tag{5.9a}$$

$$\mathcal{O}_2:=\{x: M_{\mathrm{HL}}(|f_2|^{q'})(x)>U_2^{q'}\alpha_2^{q'}\}. \tag{5.9b}$$

Then set

$$\widetilde{\mathcal{O}}=\{x: M_{\mathrm{HL}}\mathbb{1}_{\mathcal{O}}(x)>2^{-10d}(\sqrt{d})^{-d}\}\,.$$

The following relation between the sizes of $\widetilde{\mathcal{O}}$ and S_0 is key in order to prove sparse bounds.

Lemma 5.2. *If in the definitions* (5.8), (5.9) *we make the choices of* U_1,U_2 *as in* (5.2a), (5.2b), *then*

$$|\widetilde{\mathcal{O}}|<(1-\gamma)|S_0|.$$

Proof. By the weak type inequality for the Hardy–Littlewood maximal function

$$|\widetilde{\mathcal{O}}|\le 5^d2^{10d}(\sqrt{d})^d|\mathcal{O}|.$$

Moreover, by the definition of α_2,

$$|\mathcal{O}_2|\le\frac{5^d\|f_2\|_{q'}^{q'}}{U_2^{q'}\alpha_2^{q'}}=\frac{5^d3^d|S_0|}{U_2^{q'}}.$$

Furthermore, by (3.9) and (3.10),

$$|\widetilde{\Omega}_{\mu(\alpha_1)}[f_1]|\le C_d2^{-\mu(\alpha_1)p}\mathcal{C}_{\mathrm{sq},p}^p\|f_1\|_{L^p(S_0)}^p<C_dU_1^{-p}\mathcal{C}_{\mathrm{sq},p}^p|S_0|\,;$$

here we used that $2^{-\mu(\alpha_1)p} < U_1^{-p}\alpha_1^{-p}$. Finally, using (5.7), we obtain

$$\big|\{x : M_{\mathrm{HL}}(F_{1,p}^p)(x) > U_1^p\alpha_1^p\}\big| \le \frac{5^d\|F_{1,p}\|_p^p}{U_1^p\alpha_1^p} \le \frac{5^d 2^{3p/2} 2C_d\mathfrak{C}_{\mathrm{sq},p}^p|S_0|}{U_1^p}\,.$$

Altogether,

(5.10)

$$|\widetilde{\mathcal{O}}| \le |\mathcal{O}_1| + |\mathcal{O}_2| \le 5^d 2^{10d} d^{d/2}\Big(\frac{C_d\mathfrak{C}_{\mathrm{sq},p}^p}{U_1^p} + \frac{5^d 2^{1+3p/2} C_d\mathfrak{C}_{\mathrm{sq},p}^p}{U_1^p} + \frac{15^d}{U_2^{q'}}\Big)|S_0|.$$

For large choices of U_1, U_2 we get the conclusion of the lemma, and one checks that the choices of U_1, U_2 made in (5.2a), (5.2b) achieve this. □

Let $\mathcal{Q} := \{Q\}$ denote the family of dyadic Whitney cubes whose union is the open set $\widetilde{\mathcal{O}}$, which satisfy

$$5\,\mathrm{diam}(Q) \le \mathrm{dist}(Q, \widetilde{\mathcal{O}}^{\complement}) \le 12\,\mathrm{diam}(Q). \tag{5.11}$$

We note that here we adapt the standard Whitney decomposition with different constants - it will be important that the constant on the left-hand side is greater than 3 which ensures the family of triple dilates of Whitney cubes has bounded overlap (see [41] and [5, §4.4] for more details). Note that by Lemma 5.2 we have $|Q| < |S_0|$ and thus either $Q \cap S_0 = \emptyset$ or $Q \subseteq S_0$, since Q and S_0 are dyadic cubes.

We describe a decomposition of f_2 into a good and a bad part which is analogous to the usual Calderón–Zygmund decomposition at level α_2. Define

$$g_2(x) = f_2(x)\mathbb{1}_{\mathcal{O}^{\complement}}(x) + \sum_{Q\in\mathcal{Q}}\Big(\frac{1}{|Q|}\int_Q f_2(w)\,\mathrm{d}w\Big)\,\mathbb{1}_Q(x) \tag{5.12}$$

and let $b_2 = f_2 - g_2$ which gives $b_2 = \sum_{Q\in\mathcal{Q}} b_{2,Q}$ with

$$b_{2,Q}(x) = \Big(f_2(x) - \frac{1}{|Q|}\int_Q f_2(w)\,\mathrm{d}w\Big)\mathbb{1}_Q(x). \tag{5.13}$$

We have the standard Calderón–Zygmund properties.

Lemma 5.3. *(i) For all $Q \in \mathcal{Q}$, $\big(\frac{1}{Q}\int_Q |f_2(x)|^{q'}\,\mathrm{d}x\big)^{1/q'} \lesssim \alpha_2$.*
(ii) For almost every $x \in 3Q$, $|g_2(x)| \lesssim \alpha_2$.

The proof is immediate from the definition of $\mathcal{O}_2$, by the standard reasoning from Calderón–Zygmund theory (see for example [49]). We omit the details.

Next we record the following relation between cubes in $\mathcal{W}_\mu$ for $2^\mu > U_1\alpha_1$ and cubes in $\mathcal{Q}$; note that the family $\mathcal{Q}$ does not depend on μ.

Lemma 5.4. *Let $\mu \in \mathbb{Z}$ such that $2^\mu > U_1\alpha_1$. For every $W \in \mathcal{W}_\mu[f_1]$ there exists a unique $Q \in \mathcal{Q}$ such that $W \subset Q$.*

Proof. We first note that if $W \in \mathcal{W}_\mu$, then $W \subseteq \mathcal{O}_1$. This follows from the definition of $\mathcal{W}_\mu$, as $W \subseteq \widetilde{\Omega}_\mu \subseteq \widetilde{\Omega}_{\mu(\alpha_1)}$ for any $\mu \in \mathbb{Z}$ with $2^\mu > U_1\alpha_1$. Furthermore, $cW \subseteq \widetilde{\mathcal{O}}$ for sufficiently small $c \geq 1$. This follows because if $y \in cW$, then

$$M_{\mathrm{HL}} \mathbb{1}_{\mathcal{O}}(y) \geq \frac{1}{|cW|} \int_{cW} \mathbb{1}_{\mathcal{O}}(w)\,\mathrm{d}w \geq \frac{1}{|cW|} \int_{W} \mathbb{1}_{\mathcal{O}}(w)\,\mathrm{d}w = \frac{|W|}{|cW|} = c^{-d}$$

where we used that $W \subseteq \mathcal{O}$. Thus, the claim holds provided $1 \leq c < 2^{10}\sqrt{d}$. This claim implies that if x_W denotes the center of W, then

$$\operatorname{dist}(x_W, \widetilde{\mathcal{O}}^{\complement}) \geq \operatorname{dist}(x_W, (cW)^{\complement}) = \frac{\operatorname{diam}(cW)}{2\sqrt{d}} = \frac{c}{2\sqrt{d}} \operatorname{diam}(W).$$

Furthermore, as $x_W \in W \subseteq \widetilde{\mathcal{O}}$, there exists $Q \in \mathcal{Q}$ such that $x_W \in Q$. As $Q \subseteq \widetilde{\mathcal{O}}$,

$$\operatorname{dist}(x_W, \widetilde{\mathcal{O}}^{\complement}) \leq \operatorname{diam}(Q) + \operatorname{dist}(Q, \widetilde{\mathcal{O}}^{\complement}) \leq 13 \operatorname{diam}(Q)$$

where in the last inequality we have used (5.11). Consequently,

$$\frac{c}{2\sqrt{d}} \operatorname{diam}(W) \leq 13 \operatorname{diam}(Q).$$

As long as $13 \leq \frac{c}{2\sqrt{d}}$, we have that $\operatorname{diam}(W) \subseteq \operatorname{diam}(Q)$. Thus, we require a choice of c such that $26\sqrt{d} \leq c < 2^{10}\sqrt{d}$. Since W and Q are dyadic, this implies that $W \subseteq Q$, and as the cubes in $\mathcal{Q}$ have disjoint interior, the cube Q is unique. □

At this point we set once and for all (throughout the proof of Proposition 2.9 in §6-§8),

$$\mu_{\min} = \log_2(U_1\alpha_1)$$

and recalling the definitions of $b^k_{1,Q} = b^k_Q$, $b^{k,n}_{1,Q} = b^{k,n}_Q$ from §3.4 (with $f = f_1$) we then have

$$b^k_1 = \sum_{Q \in \mathcal{Q}} b^k_{1,Q} = \sum_{n \geq 0} \sum_{Q \in \mathcal{Q}} b^{k,n}_{1,Q}. \tag{5.14}$$

Note that the families

$$\mathcal{W}_{Q,\mu} = \{W \in \mathcal{W}_\mu \,:\, W \subset Q\}$$

are disjoint for different Q. For the cubes $Q \in \mathcal{Q}$, we have a standard stopping time condition for the function $F_{1,p}$.

Lemma 5.5. *For every $Q \in \mathcal{Q}$, we have*

$$\beta_{1,Q,p} = \|F_{1,p}\mathbb{1}_Q\|_p = \Big(\sum_{\mu \in \mathbb{Z}} \sum_{W \in \mathcal{W}_{Q,\mu}} (\gamma_{W,\mu}[f_1])^p |W|\Big)^{1/p} \lesssim |Q|^{1/p}\alpha_1. \tag{5.15}$$

Proof. Let $Q \in \mathcal{Q}$. By (5.11) we have that $cQ \cap \widetilde{\mathcal{O}}^{\complement} \neq \emptyset$ provided that c is sufficiently large, say $c = 100\sqrt{d}$. Let $x^* \in cQ \cap \widetilde{O}^{\complement} \subseteq cQ \cap O^{\complement}$. Then we have

$$\frac{1}{c^d|Q|}\int_Q |F_{1,p}|^p \leq \frac{1}{|cQ|}\int_{cQ} |F_{1,p}|^p \leq M_{\mathrm{HL}}(F_{1,p}^p)(x^*) \leq U_1^p \alpha_1^p,$$

as desired. □

Combining this with Lemma 3.7 and Lemma 3.9 we obtain the key estimates

$$\Big(\sum_{k>-L(S_0)} (\beta_{1,Q,p}^k)^2\Big)^{1/2} \leq \beta_{1,Q,p} \lesssim |Q|^{1/p}\alpha_1. \tag{5.16}$$

and

$$\Big(\sum_{k>-L(S_0)} (\beta_{1,Q,p}^{k,n})^p\Big)^{1/p} \leq \beta_{1,Q,p} \lesssim |Q|^{1/p}\alpha_1. \tag{5.17}$$

In the proof of the sparse bounds, the case $q \geq 2$ will only require the decomposition in k but not in n. Correspondingly, Lemma 3.6 and (5.16) will be essential in the proof of Proposition 6.4 in §8.2. The case $q < 2$ is more subtle and requires decomposition in the n-parameter. It will be essential in our argument that for $r > p$ the L^r norms of $b_{1,Q}^{k,n}$ exhibit exponential decay in n. Correspondingly, Lemma 3.8 and (5.17) will be of central importance in the proof of Proposition 6.3 in §8.1.

6. The induction step

Let $\mathbf{n} \geq 1$. This section is devoted to reducing the proof of the inductive claim (Proposition 2.9) to a couple of main estimates.

Recall from (2.8) that $\mathcal{T}_j f = \sum_{\substack{k\in\Gamma \\ k>-j}} T_k^{(j)} P_k^2 f$ and define

$$\mathcal{T} = \sum_{j=N_1}^{N_2} \mathcal{T}_j \qquad \text{and} \qquad \mathcal{T}^Q f = \sum_{j=N_1}^{L(Q)} \mathcal{T}_j[f\mathbb{1}_Q] \tag{6.1}$$

for $Q \in \mathcal{Q}\cup\{S_0\}$, and note that $\mathcal{T}^{S_0} = \mathcal{T}$. Note that by Lemma 5.2, if $Q \in \mathcal{Q}$ is such that $Q \cap S_0 \neq \emptyset$ then $Q \subsetneq S_0$. In particular $L(Q) < L(S_0) = N_2$, so $L(Q) - N_1 < \mathbf{n}$ which puts us in the position to apply the induction hypothesis to the operators $\mathcal{T}^Q$.

Next, the decomposition $f_1 = g_1 + b_1$ as in (5.3) and (3.30a) give

$$\begin{aligned}|\langle \mathcal{T} f_1, f_2\rangle| &\le |\langle \mathcal{T} g_1, f_2\rangle| + |\langle \mathcal{T} b_1, f_2\rangle| \\ &\le |\langle \mathcal{T}^{S_0} g_1, f_2\rangle| + \Big|\big\langle \sum_{Q\in\mathcal{Q}} \mathcal{T}^Q b_1, f_2\big\rangle\Big| + \Big|\big\langle \sum_{N_1\le j\le N_2} \sum_{\substack{Q\in\mathcal{Q},\\ L(Q)<j}} \mathcal{T}_j b_{1,Q}, f_2\big\rangle\Big| \\ &\le \sum_{Q\in\mathcal{Q}\cup\{S_0\}} |\langle \mathcal{T}^Q g_1, f_2\rangle| + \sum_{Q\in\mathcal{Q}} |\langle \mathcal{T}^Q f_1, f_2\rangle| + \Big|\big\langle \sum_{N_1\le j\le N_2} \sum_{\substack{Q\in\mathcal{Q},\\ L(Q)<j}} \mathcal{T}_j b_{1,Q}, f_2\big\rangle\Big|, \end{aligned} \tag{6.2}$$

where in the last step we used again $b_1 = f_1 - g_1$. We state four propositions that will be proved in the four subsequent sections. For the good part, i.e. the first term in (6.2) we have the following propositions. Here we use the notation $j_0 = L(S_0)$ as in §4.

Proposition 6.1. *Let $1 < p \le q \le 2$. For all $Q \in \mathcal{Q}\cup\{S_0\}$,*

$$|\langle \mathcal{T}^Q g_1, f_2\rangle| \lesssim \|m\|_\infty\, |Q| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}\,.$$

Proposition 6.2. *Let $2 < q \le p' < \infty$, $q' < r \le 2$. For all $Q \in \mathcal{Q}\cup\{S_0\}$,*

$$|\langle \mathcal{T}^Q g_1, f_2\rangle| \lesssim \mathcal{A}_{q',r,r}[m]\, |Q| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}\,.$$

Propositions 6.1 and 6.2 will be proved in §7.1 and §7.2, respectively. Note that disjointness of the cubes in $Q \in \mathcal{Q}$ implies $\sum_{Q\in\mathcal{Q}} |Q| \le |S_0|$ and thus Propositions 6.1 and 6.2 yield

$$\sum_{Q\in\mathcal{Q}\cup\{S_0\}} |\langle \mathcal{T}^Q g_1, f_2\rangle| \lesssim |S_0| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}.$$

The terms involving $\mathcal{T}^Q f_1$ for $Q \in \mathcal{Q}$ are estimated using the inductive hypothesis exactly as described in [5, §4.4], as $L(Q) - N_1 < n$. More precisely, given any $\epsilon > 0$, for each $Q \in \mathcal{Q}$, there exists a γ-sparse family of cubes $\mathcal{S}_Q^\epsilon \subseteq \mathcal{D}(Q)$ such that

$$|\langle \mathcal{T}^Q f_1, f_2\rangle| \le (\mathbf{U}(\mathbf{n}-1) + \epsilon) \sum_{\tilde{Q}\in\mathcal{S}_Q^\epsilon} |\tilde{Q}| \langle f_1\rangle_{\tilde{Q},p} \langle f_2\rangle_{3\tilde{Q},q'}$$

holds. By disjointness of the $Q \in \mathcal{Q}$ and Lemma 5.2, the resulting family

$$\mathcal{S}^\epsilon = \{S_0\} \cup \bigcup_{Q\in\mathcal{Q}: Q\subseteq S_0} \mathcal{S}_Q^\epsilon$$

is γ-sparse. We then get the desired result from the following propositions which take care of the third term in (6.2).

Proposition 6.3. *Let $1 < p < r \le q \le 2$. Then*

$$\Big|\big\langle \sum_{N_1\le j\le N_2} \sum_{\substack{Q\in\mathcal{Q},\\ L(Q)<j}} \mathcal{T}_j[b_{1,Q}], f_2\big\rangle\Big| \lesssim \mathcal{A}_{p,r,q}[m] |S_0| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}.$$

Proposition 6.4. *Let $2 \le q \le p' < \infty$. Then*

$$\Big|\Big\langle \sum_{N_1 \le j \le N_2} \sum_{\substack{Q \in \mathcal{Q}, \\ L(Q) < j}} \mathcal{T}_j[b_{1,Q}], f_2 \Big\rangle\Big| \lesssim \mathcal{A}_{p,p,q}[m]\, |S_0| \langle f_1 \rangle_{S_0,p} \langle f_2 \rangle_{3S_0,q'}.$$

Propositions 6.3 and 6.4 will be proved in §8.1 and §8.2, respectively. Notice that the main induction step for Theorem 1.1 follows from Propositions 6.1 and 6.3. For Theorem 1.2 it follows from Propositions 6.1 and 6.4 and for Theorem 1.3 it follows from Propositions 6.2 and 6.4.

7. The good part

Here we prove Proposition 6.1 and Proposition 6.2. By the definition of $\mathcal{T}_j$ in (2.8), using (2.10) and substituting j by $\ell - k$,

$$\langle \mathcal{T}^Q g_1, f_2 \rangle = \sum_{k \in \mathcal{F}} \sum_{\substack{\ell > 0 \\ N_1 \le \ell - k \le L(Q)}} \langle T_k^{(\ell-k)} P_k P_k[g_1 \mathbb{1}_Q], f_2 \mathbb{1}_{3Q} \rangle.$$

Next, using the decomposition $g_1 \mathbb{1}_Q = (\mathbb{E}_{1-j_0} f_1)\mathbb{1}_Q + \sum_{k' > -L(Q)} g_1^{k'} \mathbb{1}_Q$ as in (5.4a),

$$|\langle \mathcal{T}^Q g_1, f_2 \rangle| \le I + II,$$

where

$$I = \sum_{k \in \mathcal{F}} \Big| \sum_{\substack{\ell > 0 \\ N_1 \le \ell - k \le L(Q)}} \langle P_k T_k^{(\ell-k)} P_k[\mathbb{E}_{1-j_0}(f_1 \mathbb{1}_Q)], f_2 \mathbb{1}_{3Q} \rangle \Big|, \tag{7.1}$$

$$II = \Big| \sum_{k \in \mathcal{F}} \sum_{k' > -L(Q)} \sum_{\substack{\ell > 0 \\ N_1 \le \ell - k \le L(Q)}} \langle P_k T_k^{(\ell-k)} P_k[g_1^{k'} \mathbb{1}_Q], f_2 \mathbb{1}_{3Q} \rangle \Big| \tag{7.2}$$

and $g_1^{k'}$ is as in (5.4b). The main contribution to II is given by the terms with $|k - k'| \lesssim 1$. Therefore we substitute $k' = k + \nu$ with $\nu \in \mathbb{Z}$ and estimate

$$II \le \sum_{\nu \in \mathbb{Z}} \Big| \sum_{k \in \mathcal{F}} \sum_{\substack{\ell > 0 \\ N_1 \le \ell - k \le L(Q)}} \langle P_k T_k^{(\ell-k)} P_k[g_1^{k+\nu} \mathbb{1}_Q], f_2 \mathbb{1}_{3Q} \rangle \Big|.$$

From here on the terms I and II will each be estimated differently depending on whether $q \le 2$ or $q > 2$.

7.1. *The case $q \le 2$: Proof of Proposition 6.1.* By (2.2) and (1.3) we have for each fixed $k \in \mathcal{F}$ and integers $0 < L_1 \le L_2$ that

$$\Big\| \sum_{L_1 < \ell \le L_2} T_k^{(\ell-k)} \Big\|_{2 \to 2} \le 2 \sup_{\ell > 0} \|\varphi m(2^k \cdot) * 2^\ell \widehat{\Phi_0}(2^\ell \cdot)\|_\infty \lesssim \|m\|_\infty, \tag{7.3}$$

uniformly in k, L_1, L_2. Using this and the Cauchy–Schwarz inequality we bound

$$II \lesssim \|m\|_\infty \sum_{\nu \in \mathbb{Z}} \sum_{k \in \mathcal{F}} \|P_k \mathbb{D}_{k+\nu}[g_1^{k+\nu} \mathbb{1}_Q]\|_2 \|P_k[f_2 \mathbb{1}_{3Q}]\|_2,$$

which by another application of the Cauchy–Schwarz inequality and (3.2) is

$$\lesssim \|m\|_\infty \sum_{\nu\in\mathbb{Z}} 2^{-|\nu|/2} \Big\| \Big(\sum_{k>-j_0} |g_1^k|^2 \Big)^{1/2} \Big\|_{L^2(Q)} \Big\| \Big(\sum_{k\in\mathbb{Z}} |P_k[f_2 \mathbf{1}_{3Q}]|^2 \Big)^{1/2} \Big\|_2 .$$

By Lemma 5.1 we have

$$\Big\| \Big(\sum_{k>-j_0} |g_1^k|^2 \Big)^{1/2} \Big\|_{L^2(Q)} \lesssim \alpha_1 |Q|^{\frac12}.$$

Moreover using $q' \geq 2$,

$$\Big\| \Big(\sum_{k\in\mathbb{Z}} |P_k[f_2 \mathbf{1}_{3Q}]|^2 \Big)^{1/2} \Big\|_2 \lesssim \|f_2\|_{L^2(3Q)} \lesssim |Q|^{\frac12} \langle f_2\rangle_{3Q,q'} \lesssim |Q|^{\frac12}\alpha_2, \tag{7.4}$$

where the last step follows from Lemma 5.3 in the case $Q \in \mathcal{Q}$ (and is void if $Q = S_0$). Thus we obtain $II \lesssim \|m\|_\infty |Q| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}$, as desired.

To bound I we again use (7.3), the Cauchy–Schwarz inequality and (3.2) to arrive at

$$\begin{aligned} I &\lesssim \|m\|_\infty \sum_{k\in F} \|P_k[(\mathbb{E}_{1-j_0} f_1)\mathbb{1}_Q]\|_2 \|P_k[f_2 \mathbb{1}_{3Q}]\|_2 \\ &\leq \|m\|_\infty \Big\| \Big(\sum_{k\in\mathbb{Z}} |P_k(\mathbb{E}_{1-L(Q)} f_1)\mathbb{1}_Q|^2 \Big)^{1/2} \Big\|_{L^2(Q)} \Big\| \Big(\sum_{k\in\mathbb{Z}} |P_k[f_2 \mathbf{1}_{3Q}]|^2 \Big)^{1/2} \Big\|_2 \\ &\lesssim \|m\|_\infty \|\mathbb{E}_{1-j_0} f_1\|_{L^2(Q)} |Q|^{\frac12}\alpha_2 \lesssim \|m\|_\infty |Q| \langle f_1\rangle_{S_0,p} \langle f_2\rangle_{3S_0,q'}, \end{aligned}$$

where in the last two steps we have used (7.4) and that

$$\|\mathbb{E}_{1-j_0} f_1\|_{L^2(Q)} \lesssim |Q|^{\frac12} \langle f_1\rangle_{S_0,1} \leq |Q|^{\frac12} \langle f_1\rangle_{S_0,p}.$$

This concludes the proof of Proposition 6.1. □

7.2. *The case $q > 2$: Proof of Proposition 6.2.* Here we assume $2 < q \leq p' < \infty$ and let $r \in (q', 2]$. We begin with estimating the term II. Note that by Fubini's theorem,

$$II \leq \sum_{\nu\in\mathbb{Z}} \Big| \int \sum_{k\in F} P_k[g_1^{k+\nu}\mathbb{1}_Q](x) \sum_{\substack{\ell>0 \\ N_1\leq \ell-k\leq L(Q)}} (T^{(\ell-k)})^* P_k[f_2\mathbb{1}_{3Q}](x)\, \mathrm{d}x \Big|.$$

By the Cauchy–Schwarz inequality applied to the summation in k and Hölder's inequality applied to the integration in x, we obtain that the previous display is no greater than

$$\sum_{\nu\in\mathbb{Z}} \mathcal{E}^1_{q,\nu} \cdot \mathcal{E}^2_{q'},$$

where

$$\mathcal{E}^1_{q,\nu} = \Big\| \Big(\sum_{k\in F} \big| P_k[g_1^{k+\nu} \mathbb{1}_Q] \big|^2 \Big)^{1/2} \Big\|_q, \tag{7.5a}$$

$$\mathcal{E}^2_{q'} = \Big\| \Big(\sum_{k\in F} \Big| \sum_{\substack{\ell>0 \\ N_1\le \ell-k\le L(Q)}} (T_k^{(\ell-k)})^* P_k[f_2 \mathbb{1}_{3Q}] \Big|^2 \Big)^{1/2} \Big\|_{q'}. \tag{7.5b}$$

By Lemma 5.1 we have $\mathcal{E}^1_{q,\nu} \le C_q \alpha_1 |Q|^{1/q}$, uniformly in m. Moreover for $q = 2$ we can use Fubini's theorem and (3.2) to see that $\mathcal{E}^1_{2,\nu} \lesssim 2^{-|\nu|/2}\alpha_1|Q|^{1/2}$. By log-convexity of the L^q-norm, we deduce $\mathcal{E}^1_{q,\nu} \lesssim_q 2^{-|\nu|\varepsilon(q)}\alpha_1|Q|^{1/q}$ with $\varepsilon(q) < 1/q$ for $2 < q < \infty$ and thus

$$\sum_{\nu\in\mathbb{Z}} \mathcal{E}^1_{q,\nu} \lesssim_q \alpha_1 |Q|^{1/q}, \quad 2 \le q < \infty. \tag{7.6}$$

We shall now prove that

$$\mathcal{E}^2_{q'} \lesssim \mathcal{A}_{q',r,r} \alpha_2 |Q|^{1/q'}. \tag{7.7}$$

By averaging with Rademacher functions the desired bound will follow if we show that for any sequence $\{a_k\}_{k\in\mathbb{Z}}$ with $\sup_k |a_k| \le 1$ and subsets $\Lambda(k)$ of nonnegative integers we can show that

$$\Big\| \sum_{k\in F} a_k \sum_{\ell\in\Lambda(k)} (T_k^{(\ell-k)})^* P_k \Big\|_{q'\to q'} \lesssim \mathcal{A}_{q',r,r}.$$

This can be established by showing that the associated multipliers

$$h(\xi) = \sum_{k\in F} a_k \sum_{\ell\in\Lambda(k)} \big([\phi m(2^k\cdot)] * \widehat{\Psi_\ell} \big)(2^{-k}\xi) \eta(2^{-k}\xi) \tag{7.8}$$

have $M^{q'\to q'}$ norm bounded by a constant times $\mathcal{A}_{q',r,r}$, which by (1.4) (with p replaced by q') follows from the following lemma.

Lemma 7.1. *Suppose $q' < r \le 2$. Then the multipliers in* (7.8) *satisfy*

$$\sup_{t>0} \|\phi h(t\cdot)\|_{B_1^{d(\frac{1}{q'}-\frac{1}{r})}(M^{r\to r})} \lesssim \mathcal{A}_{q',r,r}, \tag{7.9}$$

with the implicit constant independent of F, of the sets $\Lambda(k) \subset \mathbb{N}_0$ and of $\{a_k\}$ in the unit ball of c_0.

The proof is straightforward but somewhat technical, and therefore postponed to §11.4. This finishes the proof of (7.7) and therefore we obtain the desired estimate for the term II. The bound for the term I is slightly simpler. We again use Fubini's theorem, the Cauchy–Schwarz inequality and Hölder's inequality to obtain that

$$I \le \mathcal{E}^1_q \cdot \mathcal{E}^2_{q'},$$

where

$$\mathcal{E}_q^1 = \Big\| \Big(\sum_{k\in F} |P_k[(\mathbb{E}_{1-j_0} f_1)\mathbb{1}_Q]|^2 \Big)^{1/2} \Big\|_q .$$

By Littlewood–Paley theory and Hölder's inequality,

$$\mathcal{E}_q^1 \lesssim \|\mathbb{E}_{1-j_0} f_1\|_{L^q(Q)} \lesssim \langle f_1\rangle_{S_0,1} |Q|^{1/q} \le \langle f_1\rangle_{S_0,p} |Q|^{1/q}.$$

Combined with (7.7) we obtain the desired bound for I which concludes the proof of Proposition 6.2. □

8. The bad part

Here we prove Proposition 6.3 and Proposition 6.4. By the definition of $\mathcal{T}_j$ in (2.8), using $b_1 = \sum_{k'>-j_0} \sum_{Q\in\mathcal{Q}} b_{1,Q}^{k'}$ and substituting j by $\ell - k$ and k' by $k+\nu$,

$$\Big| \Big\langle \sum_{N_1\le j\le N_2} \sum_{\substack{Q\in\mathcal{Q},\\ L(Q)<j}} \mathcal{T}_j[b_{1,Q}], f_2 \Big\rangle \Big| \le$$

$$\sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \Big| \sum_{\substack{Q\in\mathcal{Q}\\ L(Q)<\ell-k\le N_2}} \langle T_k^{(\ell-k)} P_k b_{1,Q}^{k+\nu}, P_k f_2\rangle \Big|.$$

For fixed $j \le j_0$ we tile S_0 with a family $\mathcal{B}_j = \{B\}$ of dyadic cubes B such that $L(B) = j$. For convenience we also set $\mathcal{B}_j = \emptyset$ if $j > j_0$. Then the previous display is

$$\le \sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \sum_{B\in\mathcal{B}_{\ell-k}} \Big| \Big\langle T_k^{(\ell-k)} P_k \Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu} \Big], (P_k f_2)\mathbb{1}_{3B} \Big\rangle \Big|. \tag{8.1}$$

Let us also recall the scaling relation

$$\begin{aligned} \|T_k^{(\ell-k)}\|_{r\to q} &= 2^{kd(\frac1r-\frac1q)} \|\varphi m(2^k\cdot) * \widehat{\Psi_\ell}\|_{M^{r\to q}} \\ &= 2^{-\ell d(\frac1p-\frac1q)} 2^{kd(\frac1r-\frac1q)} A_{p,r,q}^{k,\ell}. \end{aligned} \tag{8.2}$$

8.1. *The case $q \le 2$: Proof of Proposition 6.3.* In view of (2.6) it is no loss of generality to assume that $r > p > 1$ is chosen very close to p; indeed, it will be convenient to assume

$$d(1/p - 1/r) < 1 - 1/r, \tag{8.3}$$

which is admissible since $p > 1$. By Hölder's inequality and (8.2) we can bound (8.1) by

$$\sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \sum_{B\in\mathcal{B}_{\ell-k}} A_{p,r,q}^{k,\ell} 2^{-\ell d(\frac1p-\frac1q)} 2^{kd(\frac1r-\frac1q)} \times$$

$$\Big\| P_k \mathbb{D}_{k+\nu} \Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu} \Big] \Big\|_r \|(P_k f_2)\mathbb{1}_{3B}\|_{q'}.$$

We next use (3.2) and write $b_{1,Q}^{k+\nu}=\sum_{n\ge 0} b_{1,Q}^{k+\nu,n}$ so that the above is

$$\lesssim \sum_{n\ge 0}\sum_{\nu\in\mathbb{Z}}\sum_{\ell>0}\sum_{k\in F}\sum_{B\in\mathcal{B}_{\ell-k}} A_{p,r,q}^{k,\ell}\, 2^{-\ell d(\frac1p-\frac1q)}2^{kd(\frac1r-\frac1q)}2^{-|\nu|/r'} \times \Big\| \sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu,n}\Big\|_r \|(P_k f_2)\mathbb{1}_{3B}\|_{q'}. \tag{8.4}$$

By Lemma 11.4 and disjointness of the $Q\in\mathcal{Q}$, we have for $B\in\mathcal{B}_{\ell-k}$

$$\Big\| \sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu,n}\Big\|_r \lesssim |B|^{\frac1r-\frac1q}\Big(\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{r}}\|b_{1,Q}^{k+\nu,n}\|_r^q\Big)^{1/q}.$$

By Lemma 3.6, $\|b_{1,Q}^{k+\nu,n}\|_r\le 2^{(k+\nu-n)d(\frac1p-\frac1r)}\beta_{1,Q,p}^{k+m,n}$. Noting that

$$2^{-\ell d(\frac1p-\frac1q)}2^{kd(\frac1r-\frac1q)}|B|^{\frac1r-\frac1q}\times 2^{kd(\frac1p-\frac1r)}|Q|^{\frac1p}|Q|^{-\frac1r}=(|Q|/|B|)^{\frac1p-\frac1r}\le 1$$

for $Q\subseteq B$ and $r>p$, we obtain that (8.4) is bounded by a constant times

$$\sum_{n\ge 0}\sum_{\nu\in\mathbb{Z}} 2^{-(n-\nu)d(\frac1p-\frac1r)}2^{-|\nu|/r'}\times$$

$$\sum_{k\in F}\sum_{\ell>0}\sum_{B\in\mathcal{B}_{\ell-k}} A_{p,r,q}^{k,\ell}\Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} |Q|^{1-\frac{q}{p}}(\beta_{1,Q,p}^{k+\nu,n})^q\Big)^{1/q}\|(P_k f_2)\mathbb{1}_{3B}\|_{q'}. \tag{8.5}$$

In view of (8.3) we now fix $n\ge 0$ and $\nu\in\mathbb{Z}$. Using Hölder's inequality with exponents $(\frac1q,\frac{1}{q'})$ on the summation over B, (8.5) is estimated by

$$\mathcal{A}_{p,r,q}\sum_{k\in F}\Big(\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{p}}(\beta_{1,Q,p}^{k+\nu,n})^q\Big)^{1/q}\|P_k f_2\|_{q'}.$$

Using Hölder's inequality again with exponents $(\frac1q,\frac{1}{q'})$ on the summation over k, this is then estimated by

$$\mathcal{A}_{p,r,q}\Big(\sum_{k\in F}\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{p}}(\beta_{1,Q,p}^{k+\nu,n})^q\Big)^{1/q}\Big(\sum_{k\in\mathbb{Z}}\|P_k f_2\|_{q'}^{q'}\Big)^{1/q'}.$$

Since $q'\ge 2$ we have

$$\Big(\sum_{k\in\mathbb{Z}}\|P_k f_2\|_{q'}^{q'}\Big)^{1/q'}\le \Big\|\Big(\sum_{k\in\mathbb{Z}}|P_k f_2|^2\Big)^{1/2}\Big\|_{q'}\lesssim |S_0|^{\frac{1}{q'}}\langle f_2\rangle_{3S_0,q'}.$$

By the embedding $\ell^p\subseteq\ell^q$ for $p\le q$ in the k-sum and (5.17), we estimate

$$\Big(\sum_{k\in F}\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{p}}(\beta_{1,Q,p}^{k+\nu,n})^q\Big)^{1/q}\le \Big(\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{p}}\big(\sum_{k\in F}(\beta_{1,Q,p}^{k+\nu,n})^p\big)^{q/p}\Big)^{1/q}\lesssim |S_0|^{\frac1q}\alpha_1.$$

Summing over $n\ge 0$ and $\nu\in\mathbb{Z}$ using (8.3) concludes the argument. □

8.2. *The case $q \geq 2$: Proof of Proposition 6.4.* We decompose $f_2 = g_2 + \sum_{Q' \in \mathcal{Q}} b_{2,Q'}$, as in (5.12), (5.13), so that (8.1) is bounded by $I + II$, where

$$I = \sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \sum_{B\in\mathcal{B}_{\ell-k}} \Big|\Big\langle T_k^{(\ell-k)} P_k\Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu}\Big], (P_k g_2)\mathbb{1}_{3B}\Big\rangle\Big|,$$

$$II = \sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \sum_{B\in\mathcal{B}_{\ell-k}} \Big|\Big\langle T_k^{(\ell-k)} P_k\Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu}\Big], \mathbb{1}_{3B} \sum_{Q'\in\mathcal{Q}} P_k b_{2,Q'}\Big\rangle\Big|.$$

8.2.1. The term I. By the Cauchy–Schwarz inequality,

$$|I| \leq \sum_{\nu\in\mathbb{Z}} \sum_{\ell>0} \sum_{k\in F} \sum_{B\in\mathcal{B}_{\ell-k}} \Big\| T_k^{(\ell-k)}\big[P_k\mathbb{D}_{k+\nu} \sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu}\big]\Big\|_2 \|(P_k g_2)\mathbb{1}_{3B}\|_2.$$

By a standard localization argument, for $p \leq 2 \leq q$,

$$\|\varphi m(2^k\cdot) * \widehat{\Psi_\ell}\|_{M^{p\to 2}} \lesssim 2^{\ell d(\frac12-\frac1q)} \|\varphi m(2^k\cdot) * \widehat{\Psi_\ell}\|_{M^{p\to q}} = 2^{-\ell d(\frac1p-\frac12)} A_{p,p,q}^{k,\ell}$$

and thus by the scaling relation (8.2),

$$\|T_k^{(\ell-k)}\|_{p\to 2} \leq 2^{-(\ell-k)d(\frac1p-\frac12)} A_{p,p,q}^{k,\ell}. \tag{8.6}$$

By (3.2) and Lemma 11.4 with the exponent pair $(p,2)$,

$$\begin{aligned}
\Big\|P_k\mathbb{D}_{k+\nu}\Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu}\Big]\Big\|_p &\lesssim 2^{-|\nu|/p'} \Big\| \sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} b_{1,Q}^{k+\nu}\Big\|_p \\
&\lesssim 2^{-|\nu|/p'} |B|^{\frac1p-\frac12} \Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} |Q|^{1-\frac2p} \|b_{1,Q}^{k+\nu}\|_p^2\Big)^{1/2} \\
&\lesssim 2^{-|\nu|/p'} 2^{(\ell-k)d(\frac1p-\frac12)} \Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} |Q|^{1-\frac2p} (\beta_{1,Q,p}^{k+\nu})^2\Big)^{1/2}, \qquad (8.7)
\end{aligned}$$

where for the last line we used that $|B| = 2^{(\ell-k)d}$ and Lemma 3.6. Combining (8.6) and (8.7) we obtain

$$\begin{aligned}
|I| \leq \sum_{\nu\in\mathbb{Z}} 2^{-|\nu|/p'} \sum_{\ell>0} \sum_{k\in F} A_{p,p,q}^{k,\ell} \quad &\times \\
&\sum_{B\in\mathcal{B}_{\ell-k}} \Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}} |Q|^{1-\frac2p} (\beta_{1,Q,p}^{k+\nu})^2\Big)^{1/2} \|(P_k g_2)\mathbb{1}_{3B}\|_2
\end{aligned}$$

and after applying the Cauchy–Schwarz inequality to the sum over B we get

$$|I| \leq \mathcal{A}_{p,p,q} \sum_{\nu\in\mathbb{Z}} 2^{-|\nu|/p'} \sum_{k\in F} \Big(\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac2p} (\beta_{1,Q,p}^{k+\nu})^2\Big)^{1/2} \|P_k g_2\|_2.$$

Applying the Cauchy–Schwarz inequality in k and using $\sum_{\nu\in\mathbb{Z}} 2^{-|\nu|/p'} \lesssim 1$,

$$\le \mathcal{A}_{p,p,q}\Big(\sum_{Q\in\mathcal{Q}}\sum_{k>-j_0}|Q|^{1-\frac{2}{p}}(\beta^k_{1,Q,p})^2\Big)^{1/2}\Big\|\Big(\sum_{k\in\mathbb{Z}}|P_kg_2|^2\Big)^{1/2}\Big\|_2.$$

We apply (5.16) to get

(8.8)
$$\Big(\sum_{Q\in\mathcal{Q}}|Q|^{1-\frac{2}{p}}\sum_{k>-j_0}(\beta^k_{1,Q,p})^2\Big)^{1/2}\le\Big(\sum_{Q\in\mathcal{Q}}|Q|^{1-\frac{2}{p}}|Q|^{\frac{2}{p}}(\alpha_1)^2\Big)^{1/2}\lesssim\alpha_1|S_0|^{\frac12}.$$

By the almost orthogonality of the P_k and Lemma 5.3 (ii),

$$\Big(\sum_{k\in\mathbb{Z}}\|P_kg_2\|_2^2\Big)^{1/2}\lesssim\|g_2\|_{L^2(3S_0)}\lesssim|S_0|^{\frac12}\alpha_2.$$

In summary we obtain $|I|\lesssim\mathcal{A}_{p,p,q}|S_0|\alpha_1\alpha_2$ as claimed.

8.2.2. The term II. By the scaling relation (8.2) with $r=p$, (3.2) and using $|B|=2^{d(\ell-k)}$,

$$\Big|\Big\langle T_k^{(\ell-k)}P_k\Big[\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}}b^{k+\nu}_{1,Q}\Big],\mathbb{1}_{3B}\sum_{Q'\in\mathcal{Q}}P_kb_{2,Q'}\Big\rangle\Big|$$
$$\lesssim 2^{-|\nu|/p'}|B|^{-(\frac1p-\frac1q)}A^{k,\ell}_{p,p,q}\Big\|\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}}b^{k+\nu}_{1,Q}\Big\|_p\Big\|\sum_{Q'\in\mathcal{Q}}P_kb_{2,Q'}\Big\|_{L^{q'}(3B)}.$$

By disjointness of the cubes in $\mathcal{Q}$, Lemma 11.4 and Lemma 3.6,

$$\Big\|\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}}b^{k+\nu}_{1,Q}\Big\|_p\lesssim|B|^{\frac1p-\frac12}\Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}}|Q|^{1-\frac2p}(\beta^{k+\nu}_{1,Q,p})^2\Big)^{1/2}\tag{8.9}$$

and similarly, applying Hölder's inequality as in the proof of Lemma 11.4 we also get

$$\Big\|\sum_{Q'\in\mathcal{Q}}P_kb_{2,Q'}\Big\|_{L^{q'}(3B)}\lesssim|B|^{\frac{1}{q'}-\frac12}\Big(\sum_{\substack{Q'\in\mathcal{Q},\\ Q'\subseteq 3B}}|Q'|^{1-\frac{2}{q'}}\|P_kb_{2,Q'}\|_{q'}^2\Big)^{1/2}.$$

Combining these estimates we get

$$|II|\lesssim\sum_{\nu\in\mathbb{Z}}2^{-|\nu|/p'}\sum_{k\in F}\sum_{\ell>0}A^{k,\ell}_{p,p,q}\times\tag{8.10}$$
$$\sum_{B\in\mathcal{B}_{\ell-k}}\Big(\sum_{\substack{Q\in\mathcal{Q},\\ Q\subsetneq B}}|Q|^{1-\frac2p}(\beta^{k+\nu}_{1,Q,p})^2\Big)^{1/2}\Big(\sum_{\substack{Q'\in\mathcal{Q}:\\ Q'\subseteq 3B}}|Q'|^{1-\frac{2}{q'}}\|P_kb_{2,Q'}\|_{q'}^2\Big)^{1/2}$$

and by the Cauchy–Schwarz inequality applied to the sums over B and k,

$$|II|\lesssim\mathcal{A}_{p,p,q}\Big(\sum_{Q\in\mathcal{Q}}|Q|^{1-\frac2p}\sum_{k>-j_0}(\beta^k_{1,Q,p})^2\Big)^{1/2}\Big(\sum_{Q'\in\mathcal{Q}}\sum_{k\in\mathbb{Z}}|Q'|^{1-\frac{2}{q'}}\|P_kb_{2,Q'}\|_{q'}^2\Big)^{1/2}$$

Using $q' \le 2$ and Lemma 5.3 (i),

$$\Big(\sum_{k\in\mathbb{Z}} \|P_k b_{2,Q'}\|_{q'}^2\Big)^{1/2} \lesssim \Big\|\Big(\sum_{k\in\mathbb{Z}} |P_k b_{2,Q'}|^2\Big)^{1/2}\Big\|_{q'} \lesssim \|b_{2,Q'}\|_{q'} \lesssim |Q'|^{\frac{1}{q'}}\alpha_2$$

and thus

$$\Big(\sum_{Q'\in\mathcal{Q}}\sum_{k\in\mathbb{Z}} |Q'|^{1-\frac{2}{q'}}\|P_k b_{2,Q'}\|_{q'}^2\Big)^{1/2} \lesssim \Big(\sum_{Q'\in\mathcal{Q}} |Q'|\alpha_2^2\Big)^{1/2} \lesssim \alpha_2|S_0|^{\frac12}.$$

Together with (8.8) this gives $|II| \lesssim \mathcal{A}_{p,p,q}|S_0|\alpha_1\alpha_2$ as desired. □

9. Applications

9.1. *The classes* $\mathrm{FM}(a,b)$. We prove the positive result in Theorem 1.7. Condition (1.7) can be reformulated as

$$\sum_{|\alpha|\le n} \|\partial^\alpha[\phi m(t\cdot)]\|_\infty \lesssim_n t^{-b+na}$$

for all $n\in\mathbb{N}$ and for all $t > 1/8$ (and we have $\phi m(t\cdot) = 0$ for small t). Using a standard interpolation result for Sobolev spaces ([7, Chapter 5.4]) we also get for any $s \ge 0$

$$\|\phi m(t\cdot)\|_{B^s_{\infty,1}} \lesssim_s t^{-b+sa}, \tag{9.1}$$

for all $t > 1/8$, with the implicit constant independent of t.

We now verify $m(D) \in \mathrm{Sp}(\rho_1,\rho_2)$ for $(1/\rho_1, 1/\rho_2)$ on the edge $[P_3, P_4]$, thus satisfying $1/\rho_1 - 1/\rho_2' = b/da$. Here $P_4 = (1/p_4, 1/2)$ with $1/p_4 = 1/2 + b/ad$. The results in the remaining parts of the trapezoid then follow by Hölder's inequality.

Observe that $\frac12 \le \frac{1}{\rho_1} \le \frac{1}{p_4}$. First, we claim that the cases $\rho_1 = 2$, $\rho_1 = p_4$ follow from Theorem 1.2. By duality we only need to discuss the case $\rho_1 = p_4$. Then, taking $b = ad(1/p_4 - 1/2)$ and $s = d(1/p_4 - 1/2)$ in (9.1) we get

$$\sup_{t>0} \|\phi m(t\cdot)\|_{B^{d(1/p_4-1/2)}_{\infty,1}} < \infty. \tag{9.2}$$

Next, using the compact support of ϕ and some calculations about the contributions away from the support of ϕ we have

$$\sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(1/p_4-1/2)}(M^{p_4\to 2})} \lesssim \|m\|_\infty + \sup_{t>0} \|\phi m(t\cdot)\|_{B^{d(1/p_4-1/2)}_{\infty,1}}, \tag{9.3}$$

which establishes that $m \in \mathrm{Sp}(p_4, 2)$, as desired. In the remaining case $\frac12 < \frac{1}{\rho_1} < \frac{1}{p_4}$ we use Theorem 1.3 (with the parameters $p = \rho_1$, $q' = \rho_2$; note that $\rho_2 < 2$). To this end it suffices to verify that

$$\sup_{t>0} \|\phi m(t\cdot)\|_{B_1^{d(1/\rho_1-1/\rho_2')}(M^{\rho_1\to\rho_2'})} < \infty, \tag{9.4}$$

$$\sup_{t>0} \|\phi m(t\cdot)\|_{B^{d(1/\rho_2-1/2)}_{\infty,1}} < \infty. \tag{9.5}$$

Note that (9.5) follows from (9.2) since $B^{d(1/p_4-1/2)}_{\infty,1} \hookrightarrow B^{d(1/\rho_2-1/2)}_{\infty,1}$. Finally, (9.4) follows from (9.3) and (9.2). This is because $1/\rho_1 - 1/\rho_2' = 1/p_4 - 1/2$ and since $M^{p_4\to 2} = M^{2\to p_4'} \hookrightarrow M^{\rho_1\to\rho_2'}$ by interpolation (observe that $(1/\rho_1, 1/\rho_2')$ lies on the line between $(1/p_4, 1/2)$ and $(1/2, 1/p_4')$).

9.2. *Oscillatory multipliers.* We next turn to the proof of the positive result in Theorem 1.6. Consider the oscillatory multipliers $m_{a,b}$. It is our goal to establish the endpoint $\mathrm{Sp}(p_1,p_1)$ bound, for $1/p_1 = 1/2 + b/(ad)$, as the remaining bounds then follow by Hölder's inequality. Since $m_{a,b}$ belongs to $\mathrm{FM}(a,b)$ we have

$$\sup_{t>0} \|\phi m_{a,b}(t\cdot)\|_{B_1^{d(1/p_1-1/2)}(M^{2\to 2})} < \infty$$

and in order to apply Theorem 1.3 it remains to verify that

(9.6)
$$\sup_{t>0} \Big[\|\phi m_{a,b}(t\cdot) * \widehat{\Phi_0}\|_{M^{p_1\to p_1'}} + \sum_{\ell>0} 2^{\ell d(1/p_1-1/p_1')} \|\phi m_{a,b}(t\cdot) * \widehat{\Psi_\ell}\|_{M^{p_1\to p_1'}} \Big] < \infty.$$

We sketch the argument; a similar calculation appears in [5, Chapter 7.2.2]. Note that $\phi m_{a,b}(t\cdot) = 0$ for $t \ll 1$ and that the inequality is trivial for $t \approx 1$. For $t \gg 1$ we have

$$K_t(x) := \mathcal{F}^{-1}[\phi m_{a,b}(t\cdot)](x) = (2\pi)^{-d} \int \frac{\phi(\xi)}{t^b|\xi|^b} e^{it^a|\xi|^a - i\langle x,\xi\rangle}\, \mathrm{d}\xi\,.$$

For $a \neq 1$ the Hessian of $\xi \mapsto |\xi|^a$ has full rank. Thus we get by stationary phase the bound $|K_t(x)| \lesssim t^{-b-ad/2}$ if $|x| \approx t^a$, moreover $|K_t(x)| \lesssim_N t^{-N-b}$ if $|x| \ll t^a$ and $|K_t(x)| \lesssim_N t^{-b}|x|^{-N}$ for $|x| \gg t^a$ for all $N \geq 0$. These estimates give $M^{1\to\infty}$ bounds for $\phi m_{a,b}(t\cdot) * \widehat{\Psi_\ell}$ while we also have the trivial bound $O(t^{-b})$ for the M^2 norm. Interpolation shows that (for suitable constants $c(a) < C(a)$)

$$\|(\phi m_{a,b}(t\cdot)) * \widehat{\Psi_\ell}\|_{M^{p\to p'}} \lesssim_N \begin{cases} t^{-b-da(\frac{1}{p}-\frac{1}{2})} & \text{if } c(a)t^a \leq 2^\ell \leq C(a)t^a \\ t^{-b-Na(\frac{1}{p}-\frac{1}{2})} & \text{if } 2^\ell \leq c(a)t^a \\ t^{-b}2^{-\ell N(\frac{1}{p}-\frac{1}{2})} & \text{if } 2^\ell \geq C(a)t^a \end{cases}$$

and since $1/p_1 = 1/2 + b/(ad)$ the inequality (9.6) follows.

9.3. *The results for radial multipliers.* We use the Fefferman–Stein argument [23] based on the L^2 Stein–Tomas restriction theorem. For radial $m(\xi) = a(|\xi|)$, $1 \leq p \leq \frac{2(d+1)}{d+3}$, by Plancherel's theorem we may write $\|m(D)f\|_2^2$ as a constant times

$$\int |m(\xi)\widehat{f}(\xi)|^2\, \mathrm{d}\xi = \int_0^\infty |a(\rho)|^2 \rho^d \int_{\mathbb{S}^{d-1}} |\widehat{f}(\rho\theta)|^2\, \mathrm{d}\sigma(\theta) \frac{\mathrm{d}\rho}{\rho}.$$

Therefore, by the Stein–Tomas theorem applied to the integral over $\mathbb{S}^{d-1}$,

(9.7)
$$\|m\|_{M^p\to 2} \lesssim \Big(\int_0^\infty |a(\rho)|^2 \rho^{2d(\frac{1}{p}-\frac{1}{2})}\frac{\mathrm{d}\rho}{\rho}\Big)^{1/2} = c_d\Big(\int |m(\xi)|^2|\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi\Big)^{1/2}$$

where c_d is the surface measure of $\mathbb{S}^{d-1}$ raised to the power $-\frac{1}{2}$.

9.3.1. Proof of Corollary 1.5. We use Theorem 1.2 in conjunction with (9.7). We set $m_t(\xi) = \phi(|\xi|)h(t|\xi|)$. We let $\mathrm{d}\mu_p(\xi) = |\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi$. It suffices to prove the estimate

$$\sup_{t>0}\|\phi(\cdot)h(t|\cdot|)\|_{B^\alpha_1(L^2(\mathrm{d}\mu_p))} \lesssim \sup_{t>0}\|\phi h(t\cdot)\|_{B^\alpha_{2,1}(\mathbb{R})} \tag{9.8}$$

for $\alpha>0$, $1\le p<2$ and apply it for $\alpha = d(1/p-1/2)$ and $1<p\le\frac{2(d+1)}{d+3}$.

We are assuming that ϕ is supported in $(1/2,2)$ and it is convenient to choose χ to be a radial function supported in $\{\xi : 1/4<|\xi|<4\}$ such that $0\le\chi\le 1$ and $\chi(\xi)=1$ for $1/3\le|\xi|\le 3$. For $\ell>0$, let

$$I_{\ell,t} = \Big(\int \big|(\phi(|\cdot|)h(t|\cdot|))*\widehat{\Psi_\ell}(\xi)\big|^2\chi(\xi)|\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi\Big)^{1/2},$$
$$II_{\ell,t} = \Big(\int \big|(\phi(|\cdot|)h(t|\cdot|))*\widehat{\Psi_\ell}(\xi)\big|^2(1-\chi(\xi))|\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi\Big)^{1/2}$$

and let $I_{0,t}$, $II_{0,t}$ be the analogous expressions with Φ_0 in place of Ψ_ℓ.

First note that $|\xi|^{2d(1/p-1)}\approx 1$ on the support of χ, and therefore

$$\sum_{\ell\ge 0}2^{\ell\alpha}I_{\ell,t}\le C_\alpha\|\phi(|\cdot|)h(t|\cdot|)\|_{B^\alpha_{2,1}(\mathbb{R}^d)}.$$

We observe the inequality

$$\|\chi g(|\cdot|)\|_{W^m_2(\mathbb{R}^d)}\le C_m\|g\|_{W^m_2(\mathbb{R})} \tag{9.9}$$

which follows by application of the product and chain rules; by real interpolation we get for all $\alpha>0$

$$\|\chi g(|\cdot|)\|_{B^\alpha_{2,1}(\mathbb{R}^d)}\lesssim_\alpha C_\alpha\|g\|_{B^\alpha_{2,1}(\mathbb{R})}$$

and hence,

$$\sum_{\ell\ge 0}2^{\ell\alpha}I_{\ell,t}\lesssim_\alpha \|\phi h(t\cdot)\|_{B^\alpha_{2,1}(\mathbb{R})}.$$

It remains to estimate the term $II_{\ell,t}$. Note that for $|\xi|\ge 3$ and $N>d$,

$$\begin{aligned}\big|\phi(|\cdot|)h(t|\cdot|))*\widehat{\Psi_\ell}(\xi)\big| &\le \int_{|y|\ge|\xi|-2}|\phi(|\xi-y|)h(t|\xi-y|)|2^{\ell d}(2^\ell|y|)^{-N}\,\mathrm{d}y\\ &\lesssim \Big(\int_{|y|\ge|\xi|-2}|\phi(|\xi-y|)h(t|\xi-y|)|^2\,\mathrm{d}y\Big)^{1/2}2^{\ell(d-N)}|\xi|^{\frac{d-2N}{2}}\end{aligned}$$

by the Cauchy–Schwarz inequality. Since $2d(\frac{1}{p}-1)$ is negative, we get, for $N>d$,

$$\Big(\int_{|\xi|\geq 3}\big|(\phi(|\cdot|)h(t|\cdot|))*\widehat{\Psi_\ell}(\xi)\big|^2|\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi\Big)^{1/2}$$
$$\lesssim_N 2^{\ell(d-N)}\|\phi(|\cdot|)h(t\cdot)\|_{L^2(\mathbb{R}^d)}\lesssim_N 2^{\ell(d-N)}\|\phi h(t\cdot)\|_{L^2(\mathbb{R})}.$$

Similarly, we have

$$\Big(\int_{|\xi|\leq\frac{1}{3}}\big|(\phi(|\cdot|)h(t|\cdot|))*\widehat{\Psi_\ell}(\xi)\big|^2|\xi|^{2d(\frac{1}{p}-1)}\,\mathrm{d}\xi\Big)^{\frac12}$$
$$\lesssim_N \|\phi h(t\cdot)\|_{L^2(\mathbb{R})}\Big(\int_{|\xi|\leq\frac13}\int_{|y|\geq\frac16}2^{2\ell d}|y2^\ell|^{-2N}\,\mathrm{d}y\,|\xi|^{-\frac{2d}{p'}}\,\mathrm{d}\xi\Big)^{\frac12}$$
$$\lesssim_N 2^{\ell(d-N)}\|\phi h(t\cdot)\|_{L^2(\mathbb{R})}.$$

Altogether we get for $p<2$ (i.e. $-2d/p'>-d$),

$$2^{\ell\alpha}II_{\ell,t}\lesssim 2^{\ell(\alpha+d-N)}\|\phi h(t\cdot)\|_{L^2(\mathbb{R})}.$$

The same applies for the term $II_{0,t}$. Combining the estimates we obtain (9.8) which concludes the proof of the corollary. □

9.3.2. Proof of Theorem 1.8. Let $u_\delta(\xi)=h_\delta(|\xi|)=\chi(\delta^{-1}(1-|\xi|))$ in what follows. To use the notation in our main theorems we are setting $p_1=p$ and $p_2=q'$.

The case $p\leq\frac{2(d+1)}{d+3}$. We are seeking to prove a $\mathrm{Sp}(p,q')$ bound, under the assumption $1<p\leq\frac{2(d+1)}{d+3}$ and $\frac{1}{q'}\leq\frac{d+1}{d-1}\frac1p-\frac{2}{d-1}$ (which is equivalent with the condition $\frac1q\geq\frac{d+1}{d-1}\frac{1}{p'}$). Since $\mathrm{Sp}(p_1,p_2)\subset\mathrm{Sp}(p_1,p_3)$ for $p_3\geq p_2$ we just need to consider the endpoint line with $\frac{1}{q'}=\frac{d+1}{d-1}\frac1p-\frac{2}{d-1}$; note that under this assumption we have $2\leq q<\infty$ for $1<p\leq\frac{2(d+1)}{d+3}$ and, moreover, $q=2$ if and only $p=\frac{2(d+1)}{d+3}$. In the latter case we use Theorem 1.2 while for $1<p<\frac{2(d+1)}{d+3}$ we have $q>2$ and use Theorem 1.3.

In order to establish the assertion we have to prove, for $1<p\leq\frac{2(d+1)}{d+3}$ and $\frac{1}{q'}=\frac{d+1}{d-1}\frac1p-\frac{2}{d-1}$ the inequality

$$\sum_{\ell>0}2^{\ell d(\frac1p-\frac1q)}\|u_\delta*\widehat{\Psi_\ell}\|_{M^{p\to q}}\lesssim\delta^{-d(\frac1p-\frac12)+\frac12}. \tag{9.10}$$

Furthermore for p in the open range $1<p<\frac{2(d+1)}{d+3}$ (where $q'<2$) we use Theorem 1.3 and also have to prove, for suitable $q'<r\leq2$,

$$\sum_{\ell>0}2^{\ell d(\frac{1}{q'}-\frac1r)}\|u_\delta*\widehat{\Psi_\ell}\|_{M^{r\to r}}\lesssim\delta^{-d(\frac1p-\frac12)+\frac12}. \tag{9.11}$$

From the L^2 restriction theorem and (9.7) we get

$$\|u_\delta * \widehat{\Psi_\ell}\|_{M^{p_\circ}\to 2} \lesssim_N \delta^{1/2}\min\{1,(2^\ell\delta)^{-N}\},\quad p_\circ=\tfrac{2(d+1)}{d+3}. \tag{9.12}$$

By stationary phase and integration by parts arguments we get

$$\|u_\delta * \widehat{\Psi_\ell}\|_{M^{1\to\infty}} \lesssim 2^{-\ell\frac{d-1}{2}}\delta\min\{1,(2^\ell\delta)^{-N}\}. \tag{9.13}$$

This implies

$$2^{\ell d(\frac1p-\frac1q)}\delta^{d(\frac1p-\frac12)-\frac12}\|u_\delta * \widehat{\Psi_\ell}(D)\|_{L^p\to L^q} \lesssim (2^\ell\delta)^{\alpha(q)}\min\{1,(2^\ell\delta)^{-N}\}, \tag{9.14}$$
$$\text{for } 2\le q\le\infty,\ \tfrac1q=\tfrac{d+1}{d-1}\tfrac{1}{p'},\ \alpha(q)>0.$$

The relation between p and q can be rewritten as $\frac{1}{q'}=\frac{d+1}{d-1}\frac1p-\frac{2}{d-1}$. Thus, given (9.14) we obtain (9.10) for $1\le p<\frac{2(d+1)}{d+3}$ after summing in ℓ.

To verify (9.14) observe that by (9.12) we have (9.14) for $(\frac1p,\frac1q)$ equal to $(\frac{d+2}{2(d+1)},\frac12)$, with $\alpha(2)=d(\frac{1}{p_\circ}-\frac12)=\frac{d}{d+1}$. By (9.13) we have (9.14) for $(\frac1p,\frac1q)$ equal to $(1,0)$, with $\alpha(\infty)=\frac{d+1}{2}$. Thus by interpolation we get (9.14) for all $2\le q\le\infty$, with $\frac1q=\frac{d+1}{d-1}\frac{1}{p'}$, and $\alpha(q)>0$; more precisely $\alpha(q)=\frac{2d}{q(d+1)}+(d+1)(\frac12-\frac1q)$.

We still have to verify (9.11), but only when $p<\frac{2(d+1)}{d+3}$. Observe that $q'<2$ for $p<\frac{2(d+1)}{d+3}$. Choose r with $q'<r\le 2$, and r very close to q'. Here it suffices to use classical non-endpoint estimates which give

$$2^{\ell d(\frac{1}{q'}-\frac1r)}\|u_\delta * \widehat{\Psi_\ell}(D)\|_{L^r\to L^r}\le \tag{9.15}$$
$$C_{N,\varepsilon}2^{\ell d(\frac{1}{q'}-\frac1r)}\min\{1,(2^\ell\delta)^{-N}\}\begin{cases}\delta^{-d(\frac1r-\frac12)+\frac12-\varepsilon} & \text{if } 1\le r\le\frac{2(d+1)}{d+3}\\ \delta^{-(d+1)(\frac1r-\frac12)-\varepsilon} & \text{if } \frac{2(d+1)}{d+3}\le r\le 2\end{cases},$$

with a better result in two dimensions:

$$2^{\ell 2(\frac{1}{q'}-\frac1r)}\|u_\delta * \widehat{\Psi_\ell}(D)\|_{L^r(\mathbb{R}^2)\to L^r(\mathbb{R}^2)}\le \tag{9.16}$$
$$C_N2^{\ell 2(\frac{1}{q'}-\frac1r)}\min\{1,(2^\ell\delta)^{-N}\}\begin{cases}\delta^{-2(\frac1r-\frac12)+\frac12-\varepsilon} & \text{if } 1\le r<4/3\\ \delta^{-\varepsilon} & \text{if } 4/3\le r\le 2\end{cases}.$$

In dimension $d\ge 3$ we have to show that for r sufficiently close to q' the right-hand side of (9.15) is dominated by

$$\min\{(2^\ell\delta)^{-\epsilon_1},(2^\ell\delta)^{\epsilon_1}\}\delta^{-d(\frac1p-\frac12)+\frac12} \tag{9.17}$$

under the assumption that $\frac{1}{q'}=\frac{d+1}{d-1}\frac1p-\frac{2}{d-1}$, for some $\epsilon_1>0$ depending on p,q. Since $p<q'$ this is immediate for $q'<\frac{2(d+1)}{d+3}$. When $2>q'\ge\frac{2(d+1)}{d+3}$ the goal is accomplished once

$$d(\tfrac1p-\tfrac12)-\tfrac12-(d+1)(\tfrac{1}{q'}-\tfrac12)>0,\qquad \text{for } \tfrac1p=\tfrac{d-1}{(d+1)q'}+\tfrac{2}{d+1}$$

and the given range of q'. The displayed inequality holds for all $q' \in [\frac{2(d+1)}{d+3}, 2)$ if and only if $q' > \frac{6d+2}{5d+1}$. Since $\frac{2(d+1)}{d+3} > \frac{6d+2}{5d+1}$ for $d \geq 3$ we get the bound (9.17) for r close to q' in dimension $d \geq 3$.

In dimension $d = 2$, the previous argument is not strong enough to cover the range $\frac{6}{5} \leq q' \leq \frac{14}{11}$. We need to use the better estimate (9.16). Now the bound (9.17) is immediate for $q' < 4/3$ and if $4/3 \leq q' < 2$ and $\frac{1}{p} = \frac{1}{3q'} + \frac{2}{3}$ we have $2(\frac{1}{p} - \frac{1}{2}) - \frac{1}{2} = \frac{2}{3q'} - \frac{1}{6} > 0$ so that the desired bound follows in this case as well.

We have thus checked the assumptions of Theorem 1.2 (when $p = \frac{2(d+1)}{d+3}$) and Theorem 1.3 (when $p < \frac{2(d+1)}{d+3}$) and the theorem is proved for the case $p \leq \frac{2(d+1)}{d+3}$.

The case $\frac{2(d+1)}{d+3} < p < \frac{2(d+2)}{d+4}$. The proof relies on an inequality in [14, Proposition 2.4], which in its dual formulation says

$$\|u_\delta\|_{M^{\rho\to q}} \lesssim \delta^{-d(\frac{1}{q}-\frac{1}{2})+\frac{1}{2}},\ \tfrac{1}{q'} = \tfrac{d+1}{d-1}\tfrac{1}{\rho} - \tfrac{2}{d-1},\ \tfrac{2(d+1)}{d+3} \leq \rho < \tfrac{2(d+2)}{d+4}. \tag{9.18}$$

By averaging we can replace u_δ with $u_\delta * \widehat{\Psi_\ell}$ or $u_\delta * \widehat{\Phi_0}$ in (9.18), and after an additional interpolation with (9.12) (the case with $r = \frac{2(d+1)}{d+3}$ and $q = 2$) we also get

$$\|u_\delta * \widehat{\Psi_\ell}\|_{M^{r\to q}} \lesssim_{N,q} \min\{1, (2^\ell\delta)^{-N}\}\delta^{-d(\frac{1}{q}-\frac{1}{2})+\frac{1}{2}}, \tag{9.19}$$

first for $\frac{1}{q'} = \frac{d+1}{d-1}\frac{1}{r} - \frac{2}{d-1}$, $\frac{2(d+1)}{d+3} \leq r < \frac{2(d+2)}{d+4}$. Then, using the compact support of u_δ and the auxiliary Lemma 11.2 we also get (9.19) in the range $\frac{1}{q'} \leq \frac{d+1}{d-1}\frac{1}{r} - \frac{2}{d-1}$ and again $\frac{2(d+1)}{d+3} \leq r < \frac{2(d+2)}{d+4}$.

Now assuming $q' > \frac{(d-1)p}{d+1-2p}$ we can find $r > p$ such that $r < \frac{2(d+2)}{d+4}$ and $q' \geq \frac{(d-1)r}{d+1-2r}$ (which is equivalent with $\frac{1}{q'} \leq \frac{d+1}{d-1}\frac{1}{r} - \frac{2}{d-1}$) and then (choosing N large enough)

$$\sum_{\ell>0} 2^{\ell d(\frac{1}{p}-\frac{1}{q})}\|u_\delta * \widehat{\Psi_\ell}\|_{M^{\rho\to q}} \lesssim \delta^{-d(\frac{1}{p}-\frac{1}{2})+\frac{1}{2}}.$$

This leads to

$$\sup_{t>0} \|\phi m_\delta(t\cdot)\|_{B_1^{d/p-d/q}(M^{r\to q})} \lesssim \delta^{-d(\frac{1}{p}-\frac{1}{2})+\frac{1}{2}} \sup_{k\in\mathbb{Z}} |a_k|$$

and thus Theorem 1.1 can be applied to complete the proof of part (i) in Theorem 1.8. □

10. Necessary conditions

In this section we first give a proof of Proposition 1.9 and then discuss the sharpness of the results on oscillatory multipliers and the classes $\mathrm{FM}(a, b)$.

10.1. *Proof of Proposition 1.9.* The statements about the $L^{p_1} \to L^{p_1,\infty}$ and $L^{p_2',1} \to L^{p_2'}$ operator norms are already proved in [5, Theorem 2.5]. Moreover by [5, Lemma 2.4] we have the bound $\|T_R\|_{L^{p_1}\to L^{p_2'}} \lesssim \|T_R\|_{\mathrm{Sp}(p_1,p_2)}$, with the implicit constant independent of R. It thus suffices to prove

$$\|T_R\|_{\mathrm{Sp}(p_1,p_2)} \le (2\pi)^{-d}\|\widehat{\Psi}\|_1\|T\|_{\mathrm{Sp}(p_1,p_2)}. \tag{10.1}$$

We have $|\langle Tf_1, f_2\rangle| \le \|T\|_{\mathrm{Sp}(p_1,p_2)}\Lambda^*_{p_1,p_2}(f_1,f_2)$ for all $f_1, f_2 \in C_c^\infty$, and the same inequality holds with T replaced by T_R. To see (10.1) we write

$$\Psi(R^{-1}(x-y)) = (2\pi)^{-d}\int \widehat{\Psi}(\omega)e^{iR^{-1}\langle\omega,x\rangle}e^{-iR^{-1}\langle\omega,y\rangle}\,\mathrm{d}\omega$$

and thus we have for $f_1, f_2 \in C_c^\infty$

$$\langle T_R f_1, f_2\rangle = (2\pi)^{-d}\int \widehat{\Psi}(\omega)\langle T[f_{1,\omega}], f_{2,\omega}\rangle\,\mathrm{d}\omega$$

where $f_{1,\omega}(y) = f_1(y)e^{-iR^{-1}\langle\omega,y\rangle}$ and $f_{2,\omega}(x) = f_2(x)e^{iR^{-1}\langle\omega,x\rangle}$. Since $\langle f_{i,\omega}\rangle_{p_i} = \langle f_i\rangle_{p_i}$ we get

$$|\langle T_R f_1, f_2\rangle| \le (2\pi)^{-d}\|\widehat{\Psi}\|_1\|T\|_{\mathrm{Sp}(p_1,p_2)}\Lambda^*_{p_1,p_2}(f_1,f_2)$$

which shows (10.1). □

10.2. *Sharpness of results on Miyachi and oscillatory multipliers.* Proposition 1.9 and stationary phase calculations such as in §9.2 can be used to show that the condition $(1/p_1, 1/p_2) \in \Delta(a,b)$ in Theorem 1.6 is necessary for $m_{a,b}(D) \in \mathrm{Sp}(p_1,p_2)$. See also calculations in the proof of [5, Prop. 7.10] and related arguments for the multiplier in (10.3) below. We now construct an example completing the proof of Theorem 1.7.

Proposition 10.1. *Let $0 < b < ad/2, a \ne 1$. There is $m \in \mathrm{FM}(a,b)$ such that $m(D) \in \mathrm{Sp}(p_1,p_2)$ if and only if $(1/p_1, 1/p_2) \in \square(a,b)$.*

Proof. Let $\varphi_\circ \in C_c^\infty$ be supported in $\{\xi : \frac{3}{4} < |\xi| < \frac{5}{4}\}$ such that $\varphi_\circ(\xi) = 1$ for $\frac{7}{8} < |\xi| < \frac{9}{8}$. Let $\eta_\circ = \mathcal{F}^{-1}[\varphi_\circ]$. Let $\mathcal{N}_1$, $\mathcal{N}_2$ be two infinite disjoint subsets of $\mathbb{N}$ such that $\mathcal{N} := \mathcal{N}_1 \cup \mathcal{N}_2$ is well separated in the sense that $n \ge 1 + \frac{10}{|1-a|}$ and $|n-\tilde n| > 1 + \frac{10}{|1-a|}$ if $n, \tilde n \in \mathcal{N}$ and $n \ne \tilde n$. Define

$$m = m_1 + m_2,$$

where

$$m_1(\xi) = \sum_{k\in\mathcal{N}_1} 2^{-kb}\varphi_\circ(2^{-k}\xi)e^{-i2^{-k(1-a)}\xi_1}, \tag{10.2}$$

$$m_2(\xi) = \sum_{k\in\mathcal{N}_2} 2^{-kb}\varphi_\circ(2^{-k}\xi)e^{i2^{-k(2-a)}|\xi|^2/2} \tag{10.3}$$

and note that $m \in \mathrm{FM}(a,b)$. We remark that for the purpose of $L^p \to L^p$ inequalities m_1 behaves better than the oscillatory multipliers $m_{a,b}$, indeed if $b > 0$ then $m_1(D)$ maps $L^p \to L^p$ for all $1 < p < \infty$; yet $m_1(D)$ provides an example for the sharpness of the line through P_3 and P_4 in Theorem 1.7.

Let $K = \mathcal{F}^{-1}[m]$. Assuming that m belongs to $\mathrm{Sp}(p, q')$ we must show that $(1/p, 1/q')$ are on or below the line connecting $P_3 = (\frac{1}{2}, \frac{1}{2} + \frac{b}{da})$, and $P_4 = (\frac{1}{2} + \frac{b}{da}, \frac{1}{2})$, that is $1/p + 1/q' \le 1 + b/(da)$ or equivalently, $b \ge da(1/p - 1/q)$. Moreover, we must show that $(1/p, 1/q')$ lies on or to the left of the segment $\overline{Q_1P_4}$ i.e. satisfies $b \ge da(\frac{1}{p} - \frac{1}{2})$.

Let T_R^{resc} be the convolution operator with kernel $\Psi(x)R^dK(Rx)$ then by Proposition 1.9 we get that T_R^{resc} is bounded from $L^p \to L^q$ with operator norm uniformly bounded in R. Here we may use a suitable $\Psi \in C_c^\infty$ supported in $\{x : 1/2 < |x| < 2\}$ such that $\Psi(x) = 1$ for $2^{-1/2} \le |x| \le 2^{1/2}$. We shall use this for the parameters

$$R_n = 2^{-n(1-a)}. \tag{10.4}$$

We also let κ_n be the convolution kernel of $T_{R_n}^{\mathrm{resc}}$. We shall show the following lower bounds:

$$\text{For } n \in \mathcal{N}_1\text{:} \quad \|T_{R_n}^{\mathrm{resc}}\|_{L^p\to L^q} \gtrsim 2^{-n(b-ad(\frac{1}{p}-\frac{1}{q}))}. \tag{10.5}$$

$$\text{For } n \in \mathcal{N}_2\text{:} \quad \|T_{R_n}^{\mathrm{resc}}\|_{L^p\to L^q} \gtrsim 2^{-n(b-ad(\frac{1}{p}-\frac{1}{2}))}. \tag{10.6}$$

These imply after letting $n \to \infty$ within $\mathcal{N}_1$, $\mathcal{N}_2$, that the conditions $b \ge ad(1/p - 1/q)$, $b \ge ad(1/p - 1/2)$, are indeed necessary for $m(D) \in \mathrm{Sp}(p, q')$. Since for convolution operators the $\mathrm{Sp}(p_1, p_2)$ and $\mathrm{Sp}(p_2, p_1)$ norms coincide we get that $(1/p_1, 1/p_2) \in \square(a, b)$ is necessary for $m(D)$ to belong to the class $\mathrm{Sp}(p_1, p_2)$. A calculation yields

$$\kappa_n(x) = \Psi(x) \sum_{k\in\mathcal{N}} K_{n,k}(x),$$

where $K_{n,k}$ is defined by the following:

$$\text{For } k \in \mathcal{N}_1 : \ K_{n,k}(x) = 2^{-kb}2^{(k-n(1-a))d}\eta_\circ(2^{-n(1-a)}2^k x - 2^{ka}e_1). \tag{10.7}$$

$$\text{For } k \in \mathcal{N}_2 : \ \widehat{K_{n,k}}(\xi) = 2^{-kb}\varphi_\circ(2^{n(1-a)-k}\xi)e^{i(2^{2n(1-a)-k(2-a)}|\xi|^2/2}. \tag{10.8}$$

We let $n \in \mathcal{N}$ and decompose

$$\kappa_n = \kappa_n^{\mathrm{main}} + u_n + \sum_{k\ne n} \kappa_{n,k},$$

where

$$\kappa_n^{\mathrm{main}}(x) = K_{n,n}(x), \tag{10.9a}$$

$$u_n(x) = (\Psi(x) - 1)K_{n,n}(x), \tag{10.9b}$$

$$\kappa_{n,k}(x) = \Psi(x)K_{n,k}(x). \tag{10.9c}$$

We first consider the case $n \in \mathcal{N}_1$ and show a lower bound for the $L^p \to L^q$ norm of the operator $T_{R_n,\mathrm{main}}^{\mathrm{resc}}$ with convolution kernel $\kappa_n^{\mathrm{main}} := K_{n,n}$. By scaling and translation we have

$$\|T_{R_n,\mathrm{main}}^{\mathrm{resc}}\|_{L^p\to L^q} = \|\varphi_\circ\|_{M^{p\to q}} 2^{-n(b-ad(\frac{1}{p}-\frac{1}{q}))}, \quad n \in \mathcal{N}_1. \tag{10.10}$$

Moreover, for $n \in \mathcal{N}_2$,

$$\begin{aligned}\|T^{\mathrm{resc}}_{R_n,\mathrm{main}}\|_{L^p\to L^q} &= \|2^{-nb}\varphi_\circ(2^{-na}\cdot)e^{-i2^{-na}|\cdot|^2/2}\|_{M^{p\to q}}\\ &= 2^{-nb}2^{nad(\frac{1}{p}-\frac{1}{q})}\|\varphi_\circ e^{i2^{na}|\cdot|^2/2}\|_{M^{p\to q}}.\end{aligned}$$

Applying the method of stationary phase we get

$$|\mathcal{F}^{-1}[\varphi_\circ e^{i2^{na}|\cdot|^2/2}](x)| \approx 2^{-nad/2} \quad \text{for } ||x|-2^{na}| \le 2^{na/4}.$$

Hence

$$\begin{aligned}\|\varphi_\circ e^{i2^{na}|\cdot|^2/2}\|_{M^{p\to q}} &\gtrsim \Big(\int_{||x|-2^{na}|\le 2^{-na/4}} \big|\mathcal{F}^{-1}[\varphi_\circ e^{i2^{na}|\cdot|^2/2}]\big|^q\,\mathrm{d}x\Big)^{1/q}\\ &\gtrsim 2^{nad(\frac{1}{q}-\frac{1}{2})}\end{aligned}$$

and combining the above we get for $q \ge p$,

$$\|T^{\mathrm{resc}}_{R_n,\mathrm{main}}\|_{L^p\to L^q} \gtrsim 2^{-n(b-ad(\frac{1}{p}-\frac{1}{2}))}, \quad n \in \mathcal{N}_2. \tag{10.11}$$

In order to deduce (10.5), (10.6) from (10.10), (10.11) we show error bounds for the convolution operators with kernels u_n and $\Psi K_{n,k}$.

The contributions for $K_{n,k}$ are negligible for $k, n \in \mathcal{N}$ with $k \ne n$. Indeed from (10.7) it is immediate that for $k \in \mathcal{N}_1$, $n \in \mathcal{N}$, $k \ne n$,

$$|K_{n,k}(x)| \lesssim_N 2^{-kb}\frac{2^{(k-n(1-a))d}}{(1+2^{k-n(1-a)}|x-2^{(n-k)(1-a)}e_1|)^N}.$$

Now consider $k \in \mathcal{N}_2$ and if $n \ne k$ for $n \in \mathcal{N}$, then $(n-k)(1-a) \notin [-10,10]$. We have then

$$K_{n,k}(x) = (2\pi)^{-d}2^{-kb}\int \varphi_\circ(2^{n(1-a)-k}\xi)e^{i\phi_{n,k}(x,\xi)}\,\mathrm{d}\xi,$$

where $\phi_{n,k}(x,\xi) = 2^{2n(1-a)-k(2-a)}|\xi|^2/2 - \langle x,\xi\rangle$. Compute that for $x \in \operatorname{supp}\Psi$, $|\xi| \approx 2^{k-n(1-a)}$

$$|\nabla_\xi\phi_{n,k}(x,\xi)| \approx \begin{cases}2^{(n-k)(1-a)} & \text{if } (n-k)(1-a) \ge 10\\ 1 & \text{if } (n-k)(1-a) \le -10.\end{cases}$$

This implies after an N-fold integration by parts for $|x| \approx 1$, $k \in \mathcal{N}_2$, $n \in \mathcal{N}$, $|(n-k)(1-a)| \ge 10$

$$|K_{n,k}(x)| \lesssim_N \begin{cases}2^{-kb}2^{-(n-k)(1-a)d}2^{-ka(N-d)} & \text{for } (n-k)(1-a) \ge 10,\\ 2^{-kb}2^{-ka(N-d)}2^{(n-k)(1-a)(N-d)} & \text{for } (n-k)(1-a) \le -10.\end{cases}$$

Finally, by the support properties of $(1-\Psi)$ and $\varphi_\circ$, an integration by parts also yields

$$|u_n(x)| \lesssim_N 2^{-nb+na(N-d)}(1+|x|)^{-N}$$

for all $n \in \mathcal{N}$.

The above estimates and the resulting consequences for upper bounds for the corresponding $L^p \to L^q$ operator norms (obtained via Young's inequality) show that those terms are small compared to the lower bounds in (10.10), (10.11) and as a consequence we obtain (10.5), (10.6). □

11. Proofs of some auxiliary facts

11.1. *Proof of Observation 2.4.* We first note that the estimate (2.11) immediately implies the analogous estimate with $L(S_0) \geq N_2$, by writing S_0 as a disjoint union of cubes $Q \in \mathfrak{Q}(S_0)$ with $L(Q) = N_2$, applying the estimate on each such Q, and noting that by the disjointness of such Q,

$$\mathfrak{S}_{S_0} := \bigcup_{\substack{Q \in \mathfrak{Q}(S_0) \\ L(Q)=N_2}} \bigcup_{\substack{\mathfrak{S}_Q \subseteq \mathfrak{Q}(Q) \\ \mathfrak{S}_Q : \gamma-\text{sparse}}} \mathfrak{S}_Q \tag{11.1}$$

is a γ-sparse collection of cubes in $\mathfrak{Q}(S_0)$.

Secondly, following the argument in [5, §4.2] (based on results from [38]), one can replace $\Lambda^{**}_{S_0,p,q'}$ in (2.11) by an actual maximal sparse form $\Lambda^{*}_{S_0,p,q'}$ as in (1.1); we omit the details. Lastly, in view of (2.9), the sums $\sum_{j=N_1}^{N_2} \mathcal{T}_j$ in (2.11) can be replaced by $\sum_{k\in F} \sum_{\ell=1}^{N} P_k T_k^{(\ell-k)} P_k$.

To summarize the above reductions we see that Theorem 2.3 implies that the inequality

$$\Big| \big\langle \sum_{k\in F} \sum_{\ell=1}^{N} P_k T_k^{(\ell-k)} P_k f_1, f_2 \big\rangle \Big| \lesssim \mathcal{C}\, \Lambda^{*}_{p,q'}(f_1, f_2) \tag{11.2}$$

holds uniformly in N and F, for all C_c^∞ functions $f_1,\ f_2$.

We now use a limiting argument from [5] together with Lemma 2.2 to show that (11.2) can be upgraded to

$$\big|\langle m(D) f_1, f_2\rangle\big| \lesssim \mathcal{C}\, \Lambda^{*}_{p,q'}(f_1, f_2), \tag{11.3}$$

which in conjunction with (2.5) leads to the statements of Theorems 1.1, 1.2 and 1.3.

To this end we use (2.1c) to decompose, for $f \in \mathcal{S}$,

$$m(D) f = \sum_{k\in\mathbb{Z}} P_k m(D) L_k P_k f \tag{11.4}$$

with convergence in the sense of tempered distributions. We now apply [5, Lemma A.1] for the subspace $\mathcal{V} \equiv \mathcal{V}_1 = \mathcal{V}_2$ consisting of all $f \in \mathcal{S}$ for which $\widehat{f}$ is compactly supported in $\mathbb{R}^d \setminus \{0\}$ (and use that these are dense in L^ρ for $1 < \rho < \infty$). This lemma tells us that it suffices to prove the inequality

$$|\langle m(D) f_1, f_2\rangle| \lesssim \mathcal{C}\, \Lambda^{*}_{p,q'}(f_1, f_2) \tag{11.5}$$

for all $f_1, f_2 \in \mathcal{V}$. For fixed $f_1 \in \mathcal{V}$ the k-sum in (11.4) reduces to a sum over indices in a finite set $F(f_1)$. It therefore suffices to prove the inequality

$$\Big|\big\langle \sum_{k\in F} P_k m(D) L_k P_k f_1, f_2\big\rangle\Big| \lesssim \mathcal{C}\, \Lambda^*_{p,q'}(f_1,f_2) \tag{11.6}$$

for all finite families $F \subset \mathbb{Z}$ and for all $f_1, f_2 \in \mathcal{V}$. Again, by [5, Lemma A.1] it follows that (11.6) for all $f_1, f_2 \in \mathcal{V}$ is equivalent to (11.6) for all $f_1 \in \widetilde{\mathcal{V}}_1$, $f_2 \in \widetilde{\mathcal{V}}_2$ where, for given ρ with $p < \rho < q'$, $\widetilde{\mathcal{V}}_1$ is any dense subspace of L^ρ, and $\widetilde{\mathcal{V}}_2$ is any dense subspace of $L^{\rho'}$. It thus suffices to prove (11.6) for all $f_1, f_2 \in C_c^\infty$ and all finite families $F \subset \mathbb{Z}$.

Let $f_1, f_2 \in C_c^\infty$ so that the union of the supports of f_1 and f_2 is contained in a set of diameter R. Let $k_{\max} = \max F$. Observe that for all $k \in F$,

$$\langle T_k^{(\ell-k)} f_1, f_2\rangle = 0 \quad \text{if } \ 2^{\ell-k_{max}-3} > R \ \text{ and } \ \ell > 0\ .$$

Then, we have $\sum_{k\in F} P_k m(D) L_k P_k f_1 = \lim_{N\to\infty} \sum_{k\in F}\sum_{\ell=0}^N P_k T_k^{(\ell-k)} P_k f_1$. The terms for $\ell = 0$ are taken care of by Lemma 2.2; note that by (2.7) $\mathcal{C}$ can be used both in Lemma 2.2 and Theorem 2.3. We have thus shown that (11.6) follows from (11.2) for $f_1, f_2 \in C_c^\infty$ and the proof is complete. □

11.2. *Proof of Lemma 2.2.* We are proving that for any $1 < p \le q < \infty$ the inequality

$$|\langle \sum_{k\in F} T_k^{(-k)} P_k f_1, f_2\rangle| \le C\|m\|_\infty\, \Lambda^*_{p,q'}(f_1,f_2) \tag{11.7}$$

holds uniformly in F.

The assertion can be derived for example from [5, Theorem 1.1]. The verification of the hypotheses on that theorem is similar to the computations in [5, §6.3] and it is included for completeness. Let $m_F = \sum_{k\in F} \widehat{K}_k^{(-k)}\eta(2^{-k}\cdot)$, so that $T_{m_F} = \sum_{k\in F} T_k^{(-k)} P_k$. The support condition [5, (1.6)] clearly holds. The boundedness conditions [5, (1.7)] follow from the standard Hörmander multiplier theorem after verifying that

$$\sum_{|\alpha|\le d+1} \sup_{t>0}\sup_{\xi\in\mathbb{R}^d} |\partial_\xi^\alpha(\varphi m_F(t\cdot))(\xi)| \lesssim \|m\|_\infty \tag{11.8}$$

uniformly in F. We need to analyze the derivatives of

$$\varphi m_F(t\xi) = \sum_{\substack{k\in F\\ 2^{-k}t\sim 1}} \varphi(\xi)\int \varphi(\omega) m(2^k\omega)\widehat{\Phi_0}(2^{-k}t\xi-\omega)\,\mathrm{d}\omega$$

and (11.8) follows after straightforward computation.

For the hypothesis [5, (1.8)], we have

$$\begin{aligned}\|\mathrm{Dil}_{2^{-k}} T_k^{(-k)} P_k\|_{L^p\to L^q} &= \|\eta\big([\varphi m(2^k\cdot)] * \widehat{\Phi_0}\big)\|_{M^{p\to q}}\\ &\lesssim \|\eta\|_1\|\Phi_0\|_\infty \|\mathcal{F}^{-1}[\varphi m(2^k\cdot)]\|_\infty \lesssim \|m\|_\infty,\end{aligned}$$

for all $1 < p \le q < \infty$, using the compact support of Φ_0 and φ. Finally, the hypothesis [5, (1.9)] also follows from noting that

$$\|\mathrm{Dil}_{2^{-k}} T_k^{(-k)} P_k\|_{L^p \to L^q} \le \|\eta[e^{i\langle\cdot,h\rangle} - 1]\|_{M^p} \|[\varphi m(2^k\cdot)] * \widehat{\Phi_0}\|_{M^{p\to q}},$$

the previous bound $\|[\varphi m(2^k\cdot)] * \widehat{\Phi_0}\|_{M^{p\to q}} \lesssim \|m\|_\infty$ and that

$$\|\eta[e^{i\langle\cdot,h\rangle} - 1]\|_{M^p} \le \|\check{\eta}(\cdot + h) - \check{\eta}\|_1 \lesssim |h|$$

for any $1 < p < \infty$. □

11.3. *Some embeddings for multiplier classes.* We begin with a simple observation for compactly supported multipliers.

Lemma 11.1. *Let $1 \le p \le q \le \infty$ and $m \in M^{p\to q}$ be supported in a compact set E. Then $m \in M^{1\to\infty}$ and $\|m\|_{M^{1\to\infty}} \lesssim_E \|m\|_{M^{p\to q}}$.*

Proof. Let $\chi \in C_c^\infty$ be such that χ is supported on a compact subset of diameter less than twice the diameter of E, such that $\chi(\xi) = 1$ on a neighborhood of E. Since $\chi \in M^{1\to p} \cap M^{q\to\infty}$ we get

$$\begin{aligned}
\|\mathcal{F}^{-1}[m\widehat{f}]\|_\infty &\le \|\chi\|_{M^{q\to\infty}} \|\mathcal{F}^{-1}[m\widehat{f}]\|_q \\
\|\mathcal{F}^{-1}[m\widehat{f}]\|_q &\le \|m\|_{M^{p\to q}} \|\|\mathcal{F}^{-1}[\chi\widehat{f}]\|_p \\
\|\mathcal{F}^{-1}[\chi\widehat{f}]\|_p &\le \|\chi\|_{M^{1\to p}} \|f\|_1
\end{aligned}$$

and putting the three inequality together we deduce the assertion. □

Lemma 11.2. *For $r_1 \le r_2 \le q$ let $g \in M^{1\to\infty}$ be supported in a compact set E. Let Φ be a Schwartz function and $\Phi_\ell(x) = \Phi(2^{-\ell}x)$. Then*

$$\|g * \widehat{\Phi_\ell}\|_{M^{r_1\to q}} \lesssim \|g * \widehat{\Phi_\ell}\|_{M^{r_2\to q}} + C_N 2^{-\ell N} \|g\|_{M^{1\to\infty}}.$$

Proof. Let $E_\circ$ be a compact set which contains a neighborhood of E and let $\chi \in C_c^\infty$ such that $\chi(\xi) = 1$ for ξ in a neighborhood of $E_\circ$. Clearly $\chi \in M^{r_1\to r_2}$ and therefore

$$\big\|\chi(g * \widehat{\Phi_\ell})\big\|_{M^{r_1\to q}} \lesssim \|g * \widehat{\Phi_\ell}\|_{M^{r_2\to q}}. \tag{11.9}$$

Next we will examine the multiplier

$$(1 - \chi(\xi))\big(g * \widehat{\Phi_\ell}(\xi)\big) = (1 - \chi(\xi))\big\langle g, \widehat{\Phi_\ell}(\xi - \cdot)\big\rangle$$

where $\langle g, \cdot\rangle$ refers to the standard pairing of a tempered distribution g with a Schwartz function. We have $M^{1\to\infty} = \mathcal{F}L^\infty \hookrightarrow L^2_{-N}$ with any $N > d/2$, where $\|f\|^2_{L^2_{-N}} = \int (1 + |\xi|^2)^{-N} |\widehat{f}(\xi)|^2 \,\mathrm{d}\xi$. Since g is supported in an open

subset of $E_\circ$ we have for $\xi \notin \operatorname{supp}(\chi)$

$$\begin{aligned}
&\Big|\partial_\xi^\gamma\big[(1-\chi(\xi))\big\langle g, \widehat{\Phi_\ell}(\xi-\cdot)\big\rangle\big]\Big| \\
&\lesssim_N \|g\|_{L^2_{-N}} \sum_{\beta+\beta'=\gamma} |\partial_\xi^{\beta'}(1-\chi(\xi))| \big\|\partial_\xi^\beta \widehat{\Phi_\ell}(\xi-\cdot)\big\|_{L^2_N(E_\circ)} \\
&\lesssim_{N,\gamma,N_1} \|g\|_{L^2_{-N}} 2^{\ell(d+|\gamma|+N)} \sup_{\eta\in K_\circ} |2^\ell \operatorname{dist}(\xi,\eta)|^{-N_1} \\
&\lesssim_{N,|\gamma|,N_1} \|g\|_{M^{1\to\infty}} 2^{\ell(d+|\gamma|+N)-N_1} \sup_{\eta\in K_\circ} |\operatorname{dist}(\xi,\eta)|^{-N_1}.
\end{aligned}$$

We use that $\operatorname{dist}(\operatorname{supp}(1-\chi), E_\circ) > 0$. We apply the displayed inequality for any $|\gamma| \le 2d$ and then choose $N_1 > 3d + N + N_2$ to get (for all $1 \le r_1 \le q \le \infty$)

$$\big\|(1-\chi)\big(g * \widehat{\Phi_\ell}\big)\big\|_{M^{r_1\to q}} \lesssim_N 2^{-\ell N_2} \|g\|_{M^{1\to\infty}}. \tag{11.10}$$

The desired estimate follows from (11.9) and (11.10). □

Corollary 11.3. *Let $A^{k,\ell}_{p,r,q}$ be as in (2.3). For $p \le r_1 \le r_2 \le q$ we have*

$$A^{k,\ell}_{p,r_1,q} \lesssim A^{k,\ell}_{p,r_2,q} + C_N 2^{-\ell N} \sum_{\tilde\ell \ge 0} 2^{-\tilde\ell d(\frac{1}{p}-\frac{1}{q})} A^{k,\tilde\ell}_{p,r_2,q}. \tag{11.11}$$

Proof. Use Lemma 11.2 with $g = \varphi m(2^k\cdot)$, expand for the error term $g = g * \widehat{\Phi_0} + \sum_{\tilde\ell>0} g * \widehat{\Psi_{\tilde\ell}}$ and invoke Lemma 11.1. □

11.4. *Proof of Lemma 7.1.* Define

$$\begin{aligned}
u_\lambda(\xi) &= \phi(\xi)\eta(\lambda\xi) \\
\rho_{\ell,k}(\xi) &= [\phi m(2^k\cdot)] * \widehat{\Psi_\ell}(\xi) \\
\rho_{\ell,k,t}(\xi) &= \rho_{\ell,k}(t2^{-k}\xi)
\end{aligned}$$

and verify that

$$\phi(\xi)h(t\xi) = \sum_{k\in F} a_k u_{t2^{-k}}(\xi) \sum_{\ell\in\Lambda(k)} \rho_{\ell,k,t}(\xi).$$

We shall use that η is a Schwartz function which vanishes at the origin and thus get the estimate

$$|\mathcal{F}^{-1}[u_\lambda](x)| \le C_N \begin{cases} \lambda(1+|x|)^{-N} & \text{if } \lambda \le 1 \\ \lambda^{-N}(1+|x|)^{-N} & \text{if } \lambda \ge 1 \end{cases} \tag{11.12a}$$

for all N. Consequently,

$$\|\mathcal{F}^{-1}[u_\lambda]\|_1 \lesssim_N \min\{\lambda, \lambda^{-N}\}. \tag{11.12b}$$

For any $t>0$ we have

$$\|\phi h(t\cdot)\|_{B_1^{d(\frac{1}{q'}-\frac{1}{r})}(M^{r\to r})} \le \sum_{n\ge 0} 2^{nd(\frac{1}{q'}-\frac{1}{r})} \sum_{k\in F}\sum_{\ell\in\Lambda(k)} \|[u_{t2^{-k}}\rho_{\ell,k,t}]*\widehat{\Psi_n}\|_{M^{r\to r}}. \tag{11.13}$$

We split the sets $\Lambda(k)=\Lambda^*(k,t,n)\cup\Lambda_*(k,t,n)$ where

$$\Lambda^*(k,t,n)=\big\{\ell\in\Lambda(k): 2^\ell\ge 2^{n-5}\min\{1,2^kt^{-1}\}\big\}$$
$$\Lambda_*(k,t,n)=\Lambda(k)\setminus\Lambda^*(k,t,n).$$

We first argue that in (11.13) the terms with $\ell\in\Lambda_*(k,t,n)$ are negligible. For $\ell\in\Lambda_*(k,t,n)$ we have $2^n\ge 2^{\ell+5}\max\{1,2^{-k}t\}$. We use crude estimates for $\mathcal{F}^{-1}[\rho_{\ell,k,t}]$ and take advantage of support properties. Write $\rho_{\ell,k}(\xi)=\int\widehat{\Psi}_\ell(\xi-\omega)\phi(\omega)m(2^k\omega)\,\mathrm{d}\omega$ and by a $(d+1)$-fold integration by parts we get

$$|\mathcal{F}^{-1}[\rho_{\ell,k}](w)|\le 2^{\ell(d+1)}(1+|w|)^{-d-1}\|\phi m(2^k\cdot)\|_\infty,$$

and we have $\mathcal{F}^{-1}[\rho_{\ell,k}](w)=0$ for $|w|>2^{\ell-1}$. Hence we obtain

$$\begin{aligned}&|\mathcal{F}^{-1}[(u_{t2^{-k}}\rho_{\ell,k,t})*\widehat{\Psi_n}](x)|\\ &\le|\Psi_n(x)|\int|\mathcal{F}^{-1}[u_{t2^{-k}}](y)|(2^kt^{-1})^d|\mathcal{F}^{-1}[\rho_{\ell,k}](2^kt^{-1}(x-y))|\,\mathrm{d}y\\ &\le|\Psi_n(x)|\int_{|x-y|\le 2^{-k}t2^{\ell-1}}\frac{\min\{t2^{-k},(t2^{-k})^{-N}\}}{(1+|y|)^N}\frac{2^{\ell(d+1)}(2^kt^{-1})^d}{(1+2^kt^{-1}|x-y|)^{d+1}}\,\mathrm{d}y.\end{aligned}$$

We invoke the condition $2^\ell\le 2^{n-5}\min\{1,2^kt^{-1}\}$ for $\ell\in\Lambda_*(k,t,n)$. For $x\in\operatorname{supp}\Psi_n$ we have $|x|\ge 2^{n-3}$. Thus in the above integral we can use $|y|\ge|x|-|x-y|\ge 2^{n-3}-2^{-k}t2^{\ell-1}\ge 2^{n-3}-2^{n-6}\,2^{-k}t\min\{1,2^kt^{-1}\}\ge 2^{n-4}$ and hence

$$\begin{aligned}&\|\mathcal{F}^{-1}[(u_{t2^{-k}}\rho_{\ell,k,t})*\widehat{\Psi_n}]\|_1\\ &\lesssim 2^{\ell(d+1)}\int_{|y|\ge 2^{n-4}}\frac{\min\{t2^{-k},(t2^{-k})^{-N}\}}{(1+|y|)^N}\int\frac{(2^kt^{-1})^d}{(1+2^kt^{-1}|x-y|)^{d+1}}\,\mathrm{d}x\,\mathrm{d}y\\ &\lesssim 2^{\ell(d+1)}2^{-n(N-d)}\min\{t2^{-k},(t2^{-k})^{-N}\}.\end{aligned}$$

Consequently, using $\|[u_{t2^{-k}}\rho_{\ell,k,t}]*\widehat{\Psi_n}\|_{M^{r\to r}}\le\|\mathcal{F}^{-1}[(u_{t2^{-k}}\rho_{\ell,k,t})*\widehat{\Psi_n}]\|_1$ in (11.13) we get (assuming $N>3d$)

$$\begin{aligned}&\sum_{n\ge0}2^{nd(\frac{1}{q'}-\frac{1}{r})}\sum_{k\in F}\sum_{\ell\in\Lambda_*(k,t,n)}\|[u_{t2^{-k}}\rho_{\ell,k,t}]*\widehat{\Psi_n}\|_{M^{r\to r}}\\ &\lesssim\sum_{n\ge0}2^{nd(\frac{1}{q'}-\frac{1}{r})}2^{-n(N-d)}\sum_k\min\{t2^{-k},(t2^{-k})^{-N}\}\sum_{\ell\le n-5}2^{\ell(d+1)}\\ &\lesssim\sum_{n\ge0}2^{n(3d-N)}\sum_k\min\{t2^{-k},(t2^{-k})^{-N}\}\lesssim 1.\end{aligned}\tag{11.14}$$

We now turn to the main terms with $\ell \in \Lambda^*(k,t,n)$ in (11.13), i.e. the terms with $2^n < 2^{\ell+5}\max\{1, 2^{-k}t\}$. Notice that

$$\|[u_{t2^{-k}}\rho_{\ell,k,t}] * \widehat{\Psi_n}\|_{M^{r\to r}} \lesssim \|u_{t2^{-k}}\|_1 \|\rho_{\ell,k,t}\|_{M^{r\to r}} = \|u_{t2^{-k}}\|_1 \|\rho_{\ell,k}\|_{M^{r\to r}}$$
$$\lesssim \min\{t2^{-k}, (t2^{-k})^{-N}\} A^{k,\ell}_{q',r,r} 2^{-\ell d(\frac{1}{q'}-\frac{1}{r})}.$$

Hence in (11.13) we can estimate the terms with $\ell \in \Lambda^*(k,t,n)$ as

$$\begin{aligned}
&\sum_{n\geq 0} 2^{nd(\frac{1}{q'}-\frac{1}{r})} \sum_{k\in F} \sum_{\ell\in\Lambda^*(k,t)} \|[u_{t2^{-k}}\rho_{\ell,k,t}] * \widehat{\Psi_n}\|_{M^{r\to r}} \\
&\lesssim \sum_{\substack{k\in F\\ 2^{-k}t\leq 1}} 2^{-k}t \sum_{\ell>0} A^{k,\ell}_{q',r,r} 2^{-\ell d(\frac{1}{q'}-\frac{1}{r})} \sum_{1\leq 2^n\leq 2^{\ell+5}} 2^{nd(\frac{1}{q'}-\frac{1}{r})} \\
&\quad + \sum_{\substack{k\in F\\ 2^{-k}t>1}} (2^{-k}t)^{-N} \sum_{\ell\geq 0} A^{k,\ell}_{q',r,r} 2^{-\ell d(\frac{1}{q'}-\frac{1}{r})} \sum_{\substack{1\leq 2^n\\ \leq 2^{\ell+2}2^{-k}t}} 2^{nd(\frac{1}{q'}-\frac{1}{r})} \\
&\lesssim \sup_k \sum_{\ell\geq 0} A^{k,\ell}_{q',r,r}.
\end{aligned}$$

This finishes the proof of (7.9) and thus the proof of the lemma. □

11.5. *An elementary lemma.* The following elementary lemma is repeatedly used in the induction step for constructing sparse families of cubes.

Lemma 11.4. *Let $\mathcal{Q}$ be a family of cubes with bounded overlap and let $\{f_Q\}_{Q\in\mathcal{Q}}$ be a family of functions such that* $\operatorname{supp} f_Q \subseteq Q$. *Then, for all $1 \leq p \leq q < \infty$,*

$$\Big\| \sum_{Q\in\mathcal{Q}} f_Q \Big\|_p \lesssim \big(\sum_{Q\in\mathcal{Q}} |Q| \big)^{\frac{1}{p}-\frac{1}{q}} \Big(\sum_{Q\in\mathcal{Q}} |Q|^{1-\frac{q}{p}} \|f_Q\|_p^q \Big)^{1/q}.$$

Proof. By assumption, there is a constant C such that every x is contained in at most C of the cubes in $\mathcal{Q}$. We may split $\mathcal{Q}$ into $O(C)$ disjoint families $\mathcal{Q}_\nu$ and it suffices to prove the inequality for each $\mathcal{Q}_\nu$. From Hölder's inequality,

$$\begin{aligned}
\Big\| \sum_{Q\in\mathcal{Q}_\nu} f_Q \Big\|_p &= \Big(\sum_{Q\in\mathcal{Q}_\nu} \|f_Q\|_p^p \Big)^{1/p} = \Big(\sum_{Q\in\mathcal{Q}_\nu} |Q|^{1-p/q} |Q|^{p/q-1} \|f_Q\|_p^p \Big)^{1/p} \\
&\leq \Big(\sum_{Q\in\mathcal{Q}_\nu} |Q| \Big)^{1/p-1/q} \Big(\sum_{Q\in\mathcal{Q}_\nu} |Q|^{1-q/p} \|f_Q\|_p^q \Big)^{1/q}.
\end{aligned}$$

□

References

[1] Albert Baernstein, II and Eric T. Sawyer, *Embedding and multiplier theorems for $H^P(\mathbf{R}^n)$*, Mem. Amer. Math. Soc. **53** (1985), no. 318, iv+82.

[2] Jong-Guk Bak, *Sharp estimates for the Bochner-Riesz operator of negative order in $\mathbf{R}^2$*, Proc. Amer. Math. Soc. **125** (1997), no. 7, 1977–1986.

[3] David Beltran and Jonathan Bennett, *Subdyadic square functions and applications to weighted harmonic analysis*, Adv. Math. **307** (2017), 72–99.

[4] David Beltran and Laura Cladek, *Sparse bounds for pseudodifferential operators*, J. Anal. Math. **140** (2020), no. 1, 89–116.
[5] David Beltran, Joris Roos, and Andreas Seeger, *Multi-scale sparse domination*, Mem. Amer. Math. Soc., to appear, arXiv:2009.000227.
[6] Cristina Benea and Frédéric Bernicot, *Conservation de certaines propriétés à travers un contrôle épars d'un opérateur et applications au projecteur de Leray-Hopf*, Ann. Inst. Fourier (Grenoble) **68** (2018), no. 6, 2329–2379.
[7] Colin Bennett and Robert Sharpley, *Interpolation of operators*, Pure and Applied Mathematics, vol. 129, Academic Press, Inc., Boston, MA, 1988.
[8] Frédéric Bernicot, Dorothee Frey, and Stefanie Petermichl, *Sharp weighted norm estimates beyond Calderón-Zygmund theory*, Anal. PDE **9** (2016), no. 5, 1079–1113.
[9] Aleksandar Bulj and Vjekoslav Kovač, *Asymptotic behavior of L^p estimates for a class of multipliers with homogeneous unimodular symbols*, Trans. Amer. Math. Soc., to appear. arXiv:2022.04035.
[10] Donald L. Burkholder, *Martingale transforms*, Ann. Math. Statist. **37** (1966), 1494–1504.
[11] Anthony Carbery, *Variants of the Calderón-Zygmund theory for L^p-spaces*, Rev. Mat. Iberoamericana **2** (1986), no. 4, 381–396.
[12] Lennart Carleson and Per Sjölin, *Oscillatory integrals and a multiplier problem for the disc*, Studia Math. **44** (1972), 287–299.
[13] Sun-Yung A. Chang and Robert Fefferman, *The Calderón-Zygmund decomposition on product domains*, Amer. J. Math. **104** (1982), no. 3, 455–468.
[14] Yonggeun Cho, Youngcheol Kim, Sanghyuk Lee, and Yongsun Shim, *Sharp L^p-L^q estimates for Bochner-Riesz operators of negative index in $\mathbb{R}^n$, $n \geq 3$*, J. Funct. Anal. **218** (2005), no. 1, 150–167.
[15] Michael Christ, *On almost everywhere convergence of Bochner-Riesz means in higher dimensions*, Proc. Amer. Math. Soc. **95** (1985), no. 1, 16–20.
[16] ———, *Weak type $(1,1)$ bounds for rough operators*, Ann. of Math. (2) **128** (1988), no. 1, 19–42.
[17] Ronald R. Coifman and Guido Weiss, *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 4, 569–645.
[18] José M. Conde-Alonso, Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *A sparse domination principle for rough singular integrals*, Anal. PDE **10** (2017), no. 5, 1255–1284.
[19] José M. Conde-Alonso, Francesco Di Plinio, Ioannis Parissis, and Manasa N. Vempati, *A metric approach to sparse domination*, Ann. Mat. Pura Appl. (4) **201** (2022), no. 4, 1639–1675.
[20] José M. Conde-Alonso and Guillermo Rey, *A pointwise estimate for positive dyadic shifts and some applications*, Math. Ann. **365** (2016), no. 3-4, 1111–1135.
[21] A. Córdoba, *A note on Bochner-Riesz operators*, Duke Math. J. **46** (1979), no. 3, 505–511.
[22] Francesco Di Plinio, Tuomas P. Hytönen, and Kangwei Li, *Sparse bounds for maximal rough singular integrals via the Fourier transform*, Ann. Inst. Fourier (Grenoble) **70** (2020), no. 5, 1871–1902.
[23] Charles Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36.
[24] ———, *L^p bounds for pseudo-differential operators*, Israel J. Math. **14** (1973), 413–417.
[25] ———, *A note on spherical summation multipliers*, Israel J. Math. **15** (1973), 44–52.
[26] Charles Fefferman and Elias M. Stein, *Some maximal inequalities*, Amer. J. Math. **93** (1971), 107–115.
[27] ———, *H^p spaces of several variables*, Acta Math. **129** (1972), no. 3-4, 137–193.

[28] Shaoming Guo, Changkeun Oh, Hong Wang, Shukun Wu, and Ruixiang Zhang, *The Bochner-Riesz problem: an old approach revisited*, arxiv.org/abs/2104.11188.

[29] Susana Gutiérrez, *A note on restricted weak-type estimates for Bochner-Riesz operators with negative index in* $\mathbf{R}^n$, $n \geq 2$, Proc. Amer. Math. Soc. **128** (2000), no. 2, 495–501.

[30] Robert Kesler and Michael T. Lacey, *Sparse endpoint estimates for Bochner-Riesz multipliers on the plane*, Collect. Math. **69** (2018), no. 3, 427–435.

[31] Jongchon Kim and Andreas Seeger, *Riesz means of Fourier series and integrals: strong summability at the critical index*, Trans. Amer. Math. Soc. **372** (2019), no. 4, 2959–2999.

[32] Michael T. Lacey, *An elementary proof of the* A_2 *bound*, Israel J. Math. **217** (2017), no. 1, 181–195.

[33] ———, *Sparse bounds for spherical maximal functions*, J. Anal. Math. **139** (2019), no. 2, 613–635.

[34] Michael T. Lacey, Dario Mena, and Maria Carmen Reguera, *Sparse bounds for Bochner-Riesz multipliers*, J. Fourier Anal. Appl. **25** (2019), no. 2, 523–537.

[35] Andrei K. Lerner, *On an estimate of Calderón-Zygmund operators by dyadic positive operators*, J. Anal. Math. **121** (2013), 141–161.

[36] ———, *A simple proof of the* A_2 *conjecture*, Int. Math. Res. Not. IMRN (2013), no. 14, 3159–3170.

[37] ———, *On pointwise estimates involving sparse operators*, New York J. Math. **22** (2016), 341–349.

[38] Andrei K. Lerner and Fedor Nazarov, *Intuitive dyadic calculus: the basics*, Expo. Math. **37** (2019), no. 3, 225–265.

[39] Nicholas Michalowski, David J. Rule, and Wolfgang Staubach, *Weighted norm inequalities for pseudo-pseudodifferential operators defined by amplitudes*, J. Funct. Anal. **258** (2010), no. 12, 4183–4209.

[40] Akihiko Miyachi, *On some singular Fourier multipliers*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **28** (1981), no. 2, 267–315.

[41] Richard Oberlin, *Sparse bounds for a prototypical singular Radon transform*, Canad. Math. Bull. **62** (2019), no. 2, 405–415.

[42] Jaak Peetre, *On spaces of Triebel-Lizorkin type*, Ark. Mat. **13** (1975), 123–130.

[43] María Cristina Pereyra, *Dyadic harmonic analysis and weighted inequalities: the sparse revolution*, New trends in applied harmonic analysis. Vol. 2—harmonic analysis, geometric measure theory, and applications, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, 2019, pp. 159–239.

[44] Andreas Seeger, *A limit case of the Hörmander multiplier theorem*, Monatsh. Math. **105** (1988), no. 2, 151–160.

[45] ———, *Some inequalities for singular convolution operators in* L^p*-spaces*, Trans. Amer. Math. Soc. **308** (1988), no. 1, 259–272.

[46] ———, *Estimates near* L^1 *for Fourier multipliers and maximal functions*, Arch. Math. (Basel) **53** (1989), no. 2, 188–193.

[47] ———, *Remarks on singular convolution operators*, Studia Math. **97** (1990), no. 2, 91–114.

[48] ———, *Endpoint inequalities for Bochner-Riesz multipliers in the plane*, Pacific J. Math. **174** (1996), no. 2, 543–553.

[49] Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.

[50] ———, L^p *boundedness of certain convolution operators*, Bull. Amer. Math. Soc. **77** (1971), 404–405.

[51] Dmitriy Stolyarov, *On Fourier multipliers with rapidly oscillating symbols*, arXiv:2022.04881.

[52] Terence Tao, *The weak-type endpoint Bochner-Riesz conjecture and related topics*, Indiana Univ. Math. J. **47** (1998), no. 3, 1097–1124.
[53] ———, *A sharp bilinear restrictions estimate for paraboloids*, Geom. Funct. Anal. **13** (2003), no. 6, 1359–1384.

David Beltran: Departament d'Anàlisi Matemàtica, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain
Email address: david.beltran@uv.es

Joris Roos: Department of Mathematics and Statistics, University of Massachusetts Lowell, Lowell, MA 01854, USA
Email address: joris_roos@uml.edu

Andreas Seeger: Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Dr, Madison, WI-53706, USA
Email address: seeger@math.wisc.edu