MULTI-SCALE SPARSE DOMINATION

DAVID BELTRAN JORIS ROOS ANDREAS SEEGER

Abstract. We prove a bilinear form sparse domination theorem that applies to many multi-scale operators beyond Calderón–Zygmund theory, and also establish necessary conditions. Among the applications, we cover large classes of Fourier multipliers, maximal functions, square functions and variation norm operators.

Contents

2020 *Mathematics Subject Classification.* 42B15, 42B20, 42B25.
Key words and phrases. Sparse domination, Fourier multipliers, maximal functions, square functions, variation norm.

1. Introduction

Sparse domination results have received considerable interest in recent years since the fundamental work of Lerner on Calderón–Zygmund operators [75, 76], which provided an alternative proof of the A_2-theorem [57]. The original Banach space domination result was refined and streamlined to a pointwise result [31, 79, 65, 77], but it is the concept of sparse domination in terms of bilinear (or multilinear) forms [18, 33] that has allowed to extend the subject to many operators in harmonic analysis beyond the scope of Calderón–Zygmund theory. Among other examples, one may find the bilinear Hilbert transform [33], singular integrals with limited regularity assumptions [29, 15, 78], Bochner–Riesz operators [16, 68], spherical maximal functions [66], singular Radon transforms [28, 89, 54], pseudo-differential operators [11], maximally modulated singular integrals [38, 8], non-integral square functions [6], and variational operators [36, 35, 17], as well as results in a discrete setting (see for instance [64, 34, 2]).

Many operators in analysis have a multiscale structure, either on the space or frequency side. We consider sums

$$T = \sum_{j\in\mathbb{Z}} T_j,$$

where the Schwartz kernel of T_j is supported in a 2^j neighborhood of the diagonal and where suitable rescalings of the individual operators T_j and their adjoints satisfy uniform $L^p \to L^q$ bounds. Moreover we assume that all partial sums $\sum_{j=N_1}^{N_2} T_j$ satisfy uniform $L^p \to L^{p,\infty}$ and $L^{q,1} \to L^q$ bounds. The goal of this paper is to show bilinear form (p, q')-sparse domination results (with $q' = q/(q-1)$ the dual exponent) and investigate to which extent our assumptions are necessary. We prove such results under a very mild additional regularity assumption on the rescaled pieces; for a precise statement see Theorem 1.1 below. To increase applicability, we cover vector-valued situations, thus consider functions with values in a Banach space B_1 and operators that map simple B_1-valued functions to functions with values in a Banach space B_2. Our results apply to many classes of operators beyond Calderón–Zygmund theory, and cover general classes of convolution operators with weak assumptions on the dyadic frequency localizations, together with associated maximal functions, square functions, variation norm operators, and more. See Theorem 1.4 for a particularly clean result on translation invariant maximal functions. We shall formulate the results with respect to cubes in the standard Euclidean geometry but there are no fundamental obstructions to extend them to other geometries involving nonisotropic dilations (see *e.g.* [28]). Our approach to sparse domination extends ideas in the papers by Lacey [66] on spherical maximal functions and by R. Oberlin [89] on singular Radon transforms to more general situations.

We now describe the framework for our main theorem and first review basic definitions. For a Banach space B let S_B be the space of all B-valued

simple functions on $\mathbb{R}^d$ with compact support, i.e. all functions of the form $f = \sum_{i=1}^N a_i \mathbb{1}_{E_i}$ where $a_i \in B$ and E_i are Lebesgue measurable subsets of $\mathbb{R}^d$ contained in a compact set. For Banach spaces B_1, B_2 we consider the space Op_{B_1,B_2} of linear operators T mapping functions in S_{B_1} to weakly measurable B_2-valued functions (see *e.g.* [56] for an exposition of Banach-space integration theory) with the property that $x \mapsto \langle Tf(x), \lambda\rangle$ is locally integrable for any bounded linear functional $\lambda \in B_2^*$. If $T \in \mathrm{Op}_{B_1,B_2}$, then the integral

$$\big\langle Tf_1, f_2\big\rangle = \int_{\mathbb{R}^d} \langle Tf_1(x), f_2(x)\rangle_{(B_2,B_2^*)} dx$$

is well-defined for all $f_1 \in \mathrm{S}_{B_1}$ and $f_2 \in \mathrm{S}_{B_2^*}$. For a Banach space B and $p, r \in [1,\infty]$ we define the Lorentz space $L^{p,r}_B$ as the space of strongly measurable functions $f : \mathbb{R}^d \to B$ so that the function $x \mapsto |f(x)|_B$ is in the scalar Lorentz space $L^{p,r}$ (and we endow $L^{p,r}_B$ with the topology inherited from $L^{p,r}$). In particular, $L^p_B = L^{p,p}_B$ coincides with the standard Banach space valued L^p space as defined in [56], up to equivalence of norms. If $p \in (1,\infty)$ and $r \in [1,\infty]$, then $L^{p,r}_B$ is normable and we write $\|\cdot\|_{L^{p,r}_B}$ to denote the norm induced by the norm on scalar $L^{p,r}$ defined via the maximal function of the nonincreasing rearrangement [55].

In the definition of sparse forms it is convenient to work with a dyadic lattice $\mathfrak{Q} = \cup_{k\in\mathbb{Z}}\mathfrak{Q}_k$ of cubes, in the sense of Lerner and Nazarov [79, §2]. A prototypical example is when the cubes in the k-th generation $\mathfrak{Q}_k$ are given by

$$\mathfrak{Q}_k = \begin{cases} \{2^{-k}\mathfrak{z} + [-\frac{1}{3}2^{-k}, \frac{1}{3}2^{-k+1})^d : \mathfrak{z} \in \mathbb{Z}^d\} & \text{if } k \text{ is odd,} \\ \{2^{-k}\mathfrak{z} + [-\frac{1}{3}2^{-k+1}, \frac{1}{3}2^{-k})^d : \mathfrak{z} \in \mathbb{Z}^d\} & \text{if } k \text{ is even,} \end{cases}$$

but many other choices are possible. Notice in this example the cubes in $\mathfrak{Q}_k$ have side length 2^{-k}. This family satisfies the three axioms of a *dyadic lattice* in [79]. We briefly review the definition. $\mathfrak{Q}$ is a dyadic lattice if

(i) every child of a cube $Q \in \mathfrak{Q}$ is in $\mathfrak{Q}$,
(ii) every two cubes Q, Q' have a common ancestor in $\mathfrak{Q}$, and
(iii) every compact set in $\mathbb{R}^d$ is contained in a cube in $\mathfrak{Q}$.

For each dyadic lattice there is an $\alpha \in [1,2)$ such that all cubes $Q \in \mathfrak{Q}$ are of side length $\alpha 2^{-k}$ for some $k \in \mathbb{Z}$. Fixing k we then call the cubes of side length $\alpha 2^{-k}$ the k-th generation cubes in $\mathfrak{Q}$. If $Q \in \mathfrak{Q}$ we can, for every $l \geq 0$, tile Q into disjoint subcubes Q of side length equal to 2^{-l} times the side length of Q. We denote this family by $\mathcal{D}_l(Q)$ and let $\mathcal{D}(Q) = \cup_{l\geq 0}\mathcal{D}_l(Q)$, the family of all dyadic subcubes of Q. Then for every $Q \in \mathfrak{Q}$ we have $\mathcal{D}(Q) \subset \mathfrak{Q}$. Note that because of condition (iii) the standard dyadic lattice is not a dyadic lattice in the above sense.

Definition. Let $0 < \gamma < 1$. A collection $\mathfrak{S} \subset \mathfrak{Q}$ is *γ-sparse* if for every $Q \in \mathfrak{S}$ there is a measurable subset $E_Q \subset Q$ such that $|E_Q| \geq \gamma|Q|$ and such that the sets on the family $\{E_Q : Q \in \mathfrak{S}\}$ are pairwise disjoint.

We next review the concept of sparse domination. Given a cube Q, $1 \leq p < \infty$ and a B-valued strongly measurable locally integrable function f we use the notations

$$\mathrm{av}_Q f = |Q|^{-1} \int_Q f(x)dx, \qquad \langle f \rangle_{Q,p,B} = \Big(|Q|^{-1} \int_Q |f(x)|_B^p dx \Big)^{1/p}$$

for the average of f over Q and the L^p norm on Q with normalized measure, thus $\langle f \rangle_{Q,p,B} = (\mathrm{av}_Q |f|_B^p)^{1/p}$. For an operator $T \in \mathrm{Op}_{B_1,B_2}$ we say that pointwise sparse domination [31, 79] by L^p-averages holds if for every $f \in \mathrm{S}_{B_1}$ there are at most 3^d sparse families $\mathfrak{S}_i(f)$ such that

$$|Tf(x)|_{B_2} \leq C \sum_{i=1}^{3^d} \sum_{Q \in \mathfrak{S}_i(f)} \langle f \rangle_{Q,p,B_1} \mathbb{1}_Q(x) \quad \text{for a.e. } x \tag{1.1}$$

and we denote by $\|T\|_{\mathrm{sp}_\gamma(p,B_1,B_2)}$ the infimum over all C such that (1.1) holds for some collection of 3^d γ-sparse families depending on f.

For many operators it is not possible to obtain pointwise sparse domination and the concept of sparse domination of bilinear forms, which goes back to [18] and [33], is an appropriate substitute. Given a γ-sparse collection of cubes $\mathfrak{S}$ and $1 \leq p_1, p_2 < \infty$, one defines an associated sparse (p_1, p_2)-form acting on pairs (f_1, f_2) where f_1 is a simple B_1-valued function and f_2 is a simple B_2^*-valued function. It is given by

$$\Lambda^{\mathfrak{S}}_{p_1,B_1,p_2,B_2^*}(f_1, f_2) = \sum_{Q \in \mathfrak{S}} |Q| \langle f_1 \rangle_{Q,p_1,B_1} \langle f_2 \rangle_{Q,p_2,B_2^*}, \tag{1.2}$$

and will be abbreviated by $\Lambda^{\mathfrak{S}}_{p_1,p_2}(f_1, f_2)$ if the choice of B_1, B_2^* is clear from context. The form (1.2) acts a bi-sublinear form on $(|f_1|_{B_1}, |f_2|_{B_2^*})$. All sparse forms are dominated by a maximal form

$$\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1, f_2) = \sup_{\mathfrak{S}:\gamma\text{-sparse}} \Lambda^{\mathfrak{S}}_{p_1,B_1,p_2,B_2^*}(f_1, f_2), \tag{1.3}$$

again also abbreviated by $\Lambda^*_{p_1,p_2}(f_1, f_2)$ if the choice of B_1, B_2^* is clear from the context. The maximal form may not be a sparse form itself but, obviously, for every f_1, f_2 there exists a sparse family $\mathfrak{S}(f_1, f_2)$ such that $\Lambda^{\mathfrak{S}(f_1,f_2)}_{p_1,B_1,p_2,B_2^*}(f_1, f_2) \geq \frac{1}{2}\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1, f_2)$ (*cf.* [67], [32] for more explicit constructions). Note from (1.2) that for each pair of simple functions (f_1, f_2),

$$\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1, f_2) \leq \gamma^{-1} \|f_1\|_\infty \|f_2\|_\infty \,\mathrm{meas}(\mathrm{supp} f_1 \cup \mathrm{supp} f_2) < \infty.$$

We say that $T \in \mathrm{Op}_{B_1,B_2}$ satisfies a sparse (p_1, p_2) bound if there is a constant C so that for all $f_1 \in \mathrm{S}_{B_1}$ and $f_2 \in \mathrm{S}_{B_2^*}$ the inequality

$$\big|\langle Tf_1, f_2 \rangle\big| \leq C \Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1, f_2) \tag{1.4}$$

is satisfied. The best constant in (1.4) defines a norm $\|\cdot\|_{\mathrm{Sp}_\gamma(p_1,B_1,p_2,B_2^*)}$ on a subspace of Op_{B_1,B_2}. Thus $\|T\|_{\mathrm{Sp}_\gamma(p_1,B_1;p_2,B_2^*)}$ is given by

$$\sup\Big\{\frac{|\langle Tf_1,f_2\rangle|}{\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1,f_2)}: f_1\in \mathrm{S}_{B_1},\ f_2\in \mathrm{S}_{B_2^*},\ f_i\neq 0,\ i=1,2\Big\}, \tag{1.5}$$

where $f_i\neq 0$ means that $f_i(x)\neq 0$ on a set of positive measure. It is then immediate that $\|T\|_{\mathrm{Sp}_\gamma(p_1,B_1,p_2,B_2^*)}\le \|T\|_{\mathrm{sp}_\gamma(p_1,B_1,B_2)}$ for $p_2\ge 1$. It can be shown that the space of operators in Op_{B_1,B_2} for which (1.4) holds for all f_1, f_2 with a finite C does not depend on γ. We denote this space by $\mathrm{Sp}(p_1,B_1,p_2,B_2^*)$ or simply $\mathrm{Sp}(p_1,p_2)$ if the choices of B_1,B_2^* are clear from context. The norms $\|\cdot\|_{\mathrm{Sp}_\gamma(p_1,B_1,p_2,B_2^*)}$, $0<\gamma<1$, are equivalent norms on $\mathrm{Sp}(p_1,B_1,p_2,B_2^*)$. Moreover, if B_1, B_2 are separable Banach spaces and $p_1<p<p_2'$, then all operators in $\mathrm{Sp}(p_1,B_1,p_2,B_2^*)$ extend to bounded operators from $L^p_{B_1}$ to $L^p_{B_2}$.

1.1. **The main result.** For a function f define $\mathrm{Dil}_t f(x)=f(tx)$. For an operator T define the dilated operator $\mathrm{Dil}_t T$ by

$$\mathrm{Dil}_t T=\mathrm{Dil}_t\circ T\circ \mathrm{Dil}_{t^{-1}}.$$

Note that if T is given by a Schwartz kernel $(x,y)\mapsto K(x,y)$, then the Schwartz kernel of $\mathrm{Dil}_t T$ is given by $(x,y)\mapsto t^dK(tx,ty)$.

Basic assumptions. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2}. We shall make the following assumptions.

Support condition. For all $f\in \mathrm{S}_{B_1}$,

$$\mathrm{supp}\ (\mathrm{Dil}_{2^j}T_j)f\subset\{x\in\mathbb{R}^d:\mathrm{dist}(x,\mathrm{supp}\,f)\le 1\}. \tag{1.6}$$

This means that if T_j is given by integration against a Schwartz kernel K_j, then K_j lives on a 2^j-neighborhood of the diagonal.

Weak type (p,p) condition. For all integers $N_1\le N_2$, the sums $\sum_{j=N_1}^{N_2}T_j$ are of weak type (p,p), with uniform bounds,

$$\sup_{N_1\le N_2}\Big\|\sum_{j=N_1}^{N_2}T_j\Big\|_{L^p_{B_1}\to L^{p,\infty}_{B_2}}\le A(p). \tag{1.7a}$$

Restricted strong type (q,q) condition. For all integers $N_1\le N_2$, the sums $\sum_{j=N_1}^{N_2}T_j$ are of restricted strong type (q,q), with uniform bounds,

$$\sup_{N_1\le N_2}\Big\|\sum_{j=N_1}^{N_2}T_j\Big\|_{L^{q,1}_{B_1}\to L^{q}_{B_2}}\le A(q). \tag{1.7b}$$

Single scale (p,q) condition. The operators T_j satisfy the uniform improving bounds

$$\sup_{j\in\mathbb{Z}}\|\mathrm{Dil}_{2^j}T_j\|_{L^p_{B_1}\to L^q_{B_2}}\le A_\circ(p,q). \tag{1.8}$$

Single scale ε-regularity conditions. For some $\varepsilon > 0$ the operators T_j and the adjoints T_j^* satisfy

$$\sup_{|h|\le 1} |h|^{-\varepsilon} \sup_{j\in\mathbb{Z}} \|(\mathrm{Dil}_{2^j} T_j) \circ \Delta_h\|_{L^p_{B_1}\to L^q_{B_2}} \le B, \tag{1.9a}$$

$$\sup_{|h|\le 1} |h|^{-\varepsilon} \sup_{j\in\mathbb{Z}} \|(\mathrm{Dil}_{2^j} T_j^*) \circ \Delta_h\|_{L^{q'}_{B_2^*}\to L^{p'}_{B_1^*}} \le B, \tag{1.9b}$$

where

$$\Delta_h f(x) := f(x+h) - f(x). \tag{1.10}$$

The above hypotheses assume certain boundedness assumptions in Lebesgue or Lorentz spaces of vector-valued functions; it is then implied that all operators T_j map simple B_1-valued functions to B_2-valued functions which are strongly measurable with respect to Lebesgue measure. We formulate our main result for $1 < p \le q < \infty$ and refer to Appendix B for variants with $p = 1$ or $q = \infty$.

Theorem 1.1. *Let $1 < p \le q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} such that*

- *the support condition* (1.6) *holds,*
- *the weak type (p,p) condition* (1.7a) *holds,*
- *the restricted strong type (q,q) condition* (1.7b) *holds,*
- *the single scale (p,q) condition* (1.8) *holds,*
- *the single scale ε-regularity conditions* (1.9a), (1.9b) *hold.*

Define

$$\mathcal{C} = A(p) + A(q) + A_\circ(p,q) \log\big(2 + \tfrac{B}{A_\circ(p,q)}\big). \tag{1.11}$$

Then, for all integers N_1, N_2 with $N_1 \le N_2$,

$$\Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{\mathrm{Sp}_\gamma(p,B_1,q',B_2^*)} \lesssim_{p,q,\varepsilon,\gamma,d} \mathcal{C}. \tag{1.12}$$

The estimate (1.12) implies, via a linearization technique (*cf.* Lemma 4.4) the following variant which leads to a sparse domination result for maximal functions, square functions and variational operators, see Ch.5. Instead of $T_j \in \mathrm{Op}_{B_1,B_2}$ we use the more restrictive assumption that the T_j map functions in S_{B_1} to locally integrable B_2-valued functions. We let $L^1_{B_2,\mathrm{loc}}$ be the space of all strongly measurable B_2-valued functions which are Bochner integrable over compact sets.

Corollary 1.2. *Let $1 < p \le q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators, with $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$, and satisfying the assumptions of Theorem 1.1. Let $\mathcal{C}$ be as in* (1.11)*. Then for all $f \in \mathrm{S}_{B_1}$, all $\mathbb{R}$-valued nonnegative measurable functions ω, and all integers N_1, N_2 with $N_1 \le N_2$,*

$$\int_{\mathbb{R}^d} \Big| \sum_{j=N_1}^{N_2} T_j f(x) \Big|_{B_2} \omega(x)dx \lesssim_{p,q,\varepsilon,\gamma,d} \mathcal{C}\, \Lambda^*_{p,B_1,q',\mathbb{R}}(f,\omega). \tag{1.13}$$

Remarks. (i) We emphasize that the implicit constants in (1.12) and (1.13) are dependent on the input constants in (1.7a), (1.7b), (1.8), (1.9a), (1.9b) but otherwise not dependent on the specific choices of the Banach spaces B_1, B_2. In some applications this enables us to perform certain approximation arguments, where for example the Banach spaces are replaced by finite-dimensional subspaces of large dimension.

(ii) We note that for operators T_j which commute with translations Condition (1.9b) is implied by Condition (1.9a).

(iii) The Hölder-type regularity assumption (1.9) for the operator norm can be further weakened. In applications this will often be used for the situation that an operator T is split into a sum $\sum_{\ell\geq 0} T^\ell$ where each $T^\ell = \sum_j T_j^\ell$ satisfies the assumptions with $A(p), A(q), A_\circ(p,q) = O(2^{-\ell\epsilon'})$ for some $\epsilon' > 0$ and $B = 2^{\ell M}$, for a possibly very large M. The conclusion will then say that $\|T^\ell\|_{\mathrm{S}_{\mathrm{p}_\gamma}(p,q')} = O(\ell 2^{-\ell\epsilon'})$, which can be summed in ℓ, leading to a sparse bound for T.

(iv) In this paper we are mainly interested in applications beyond the Calderón–Zygmund theory and focus on the case $p > 1$ and $q < \infty$. There are some elements in our proof such as the property of $L^{p,\infty}$ being the dual space of $L^{p',1}$ for which there is no analog for $p = 1$ and similarly the failure of a suitable notion of restricted strong type for $q = \infty$; hence Theorem 1.1 does not immediately apply to the situations where $p = 1$ or $q = \infty$. Nevertheless one can formulate variants of the theorem which cover these missing cases. We treat them in Appendix B; indeed they are close to results already covered in other works, in particular [29].

(v) The role of the simple functions is not essential in Theorem 1.1, and the sparse bound can be extended to other classes of functions under appropriate hypotheses; see Lemma A.1.

(vi) We use the Banach space valued formulation only to increase applicability. We emphasize that we make no specific assumptions on the Banach spaces in our formulation of Theorem 1.1 (such as UMD in the theory of Banach space valued singular integrals). In applications to Banach space valued singular integrals, such assumptions are made only because they may be needed to verify L^p-boundedness hypotheses but they are not needed to establish the implication in Theorem 1.1.

1.2. **Necessary conditions.** Under the additional assumption that $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$, together with $p < q$, one has that the weak type (p,p) condition (1.7a) and the restricted strong type condition (1.7b) are necessary for the conclusion of Theorem 1.1 to hold. Moreover, if we strengthen the support condition (1.6) assuming that the Schwartz kernels of T_j are not only supported in $\{|x-y| \lesssim 2^j\}$ but actually in $\{|x-y| \approx 2^j\}$, then we can also show that the single scale (p,q) condition (1.8) is necessary.

We also have an analogous statement for Corollary 1.2. Indeed, as the corollary is proved via the implication

$$(1.12) \implies (1.13),$$

see Lemma 4.4 below, we will simply formulate the necessary conditions for the conclusion in Corollary 1.2, which will also imply those in Theorem 1.1.

To be precise in the general setting, let us formulate the following assumption on a family of operators $\{T_j\}_{j\in\mathbb{Z}}$.

Strengthened support condition. There are $\delta_1 > \delta_2 > 0$ such that for all $j \in \mathbb{Z}$ and all $f \in \mathrm{S}_{B_1}$

$$\mathrm{supp}(\mathrm{Dil}_{2^j} T_j f) \subset \{x : \delta_1 \le \mathrm{dist}(x, \mathrm{supp} f) \le 1\}, \qquad \text{whenever } \mathrm{diam}(\mathrm{supp} f) \le \delta_2. \quad (1.14)$$

If the T_j are given by a Schwartz kernel K_j, then the condition is satisfied provided that

$$\mathrm{supp}(K_j) \subset \{(x,y) : (\delta_1 - \delta_2)2^j \le |x-y| \le 2^j\}.$$

Theorem 1.3. *Suppose that $1 < p < q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators, with $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$, and satisfying the support condition* (1.6). *Assume the conclusion of Corollary 1.2, that is, there exists $\mathcal{C} > 0$ such that for all N_1, N_2 with $N_1 \le N_2$, all $f_1 \in \mathrm{S}_{B_1}$, and all nonnegative simple functions ω*

$$\int_{\mathbb{R}^d} \Big| \sum_{j=N_1}^{N_2} T_j f(x) \Big| \, \omega(x)\, dx \le \mathcal{C}\, \Lambda^*_{p,B_1,q',\mathbb{R}}(f,\omega).$$

Then

(i) Conditions (1.7a) *and* (1.7b) *hold, i.e., there is a constant $c > 0$ only depending on d, p, q, γ such that for all N_1, N_2 with $N_1 \le N_2$,*

$$\Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{L^p_{B_1} \to L^{p,\infty}_{B_2}} \le c\,\mathcal{C}, \qquad \Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{L^{q,1}_{B_1} \to L^q_{B_2}} \le c\,\mathcal{C}\,.$$

(ii) If, in addition, the T_j satisfy the strengthened support condition (1.14) *then condition* (1.8) *holds, i.e., there is a constant $c > 0$ only depending on d, p, q, γ such that*

$$\sup_{j\in\mathbb{Z}} \|\mathrm{Dil}_{2^j} T_j\|_{L^p_{B_1} \to L^q_{B_2}} \le c\,\mathcal{C}.$$

Remarks. (i) Note that in Theorem 1.3 there are no additional assumptions on the Banach spaces. The a priori assumption $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$ enters in the proof of necessary conditions for both Theorem 1.1 and Corollary 1.2.

(ii) There is an alternative version for necessary conditions for Theorem 1.1 where one a priori assumes merely that the T_j belong to Op_{B_1,B_2} (i.e. $T_j f$ is only a priori weakly integrable for $f \in \mathrm{S}_{B_1}$), but where one imposes the assumption that B_2 is reflexive. See Theorem 2.5 below.

(iii) We have no necessity statement regarding the regularity conditions (1.9) in Theorem 1.1, or Corollary 1.2. However, these conditions enter in the conclusion of both Theorem 1.1 and Corollary 1.2 only in a logarithmic way (see (1.11)), hence the gap between necessity and sufficiency appears to be small. Note that the necessary and sufficient conditions are formulated for a uniform statement on a family of operators $\{\sum_{j=N_1}^{N_2} T_j\}_{N_1,N_2}$ but, with the generality of our current formulation, we are unable to prove a necessary condition for sparse domination for a specific operator in this family. Nevertheless, the formulation allows us to show necessary conditions for several specific maximal operators, variation norm operators and other vector-valued variants, in particular those considered in §5.1, §5.2, §5.3 and §7.1.

(iv) The constant c in the conclusion of Theorem 1.3 is independent of the particular pair of Banach spaces B_1, B_2. This is significant for applying the theorem to families of maximal and variational operators where for the necessity conditions one can replace the spaces ℓ^∞, L^∞, V^r by finite-dimensional subspaces of large dimension.

(v) Since $\|T\|_{\mathrm{Sp}_\gamma(p_1,B_1,p_2,B_2^*)} \le \|T\|_{\mathrm{sp}_\gamma(p_1,B_1,B_2)}$ for $p_2 \ge 1$, the necessary conditions in Theorem 1.3 can also be used to prove the impossibility of pointwise sparse domination for many of the operators considered in this paper.

1.3. **An application to maximal functions.** We illustrate Remark (iii) above with a brief discussion about maximal operators associated to a distribution σ compactly supported in $\mathbb{R}^d\backslash\{0\}$ (for example a measure), for which we have necessary conditions for sparse bounds. Denote by $\sigma_t = t^{-d}\sigma(t^{-1}\cdot)$ the t-dilate in the sense of distributions. For a dilation set $E \subset (0,\infty)$ we consider the maximal operator

$$M_E^\sigma f(x) = \sup_{t\in E} |f * \sigma_t(x)|. \tag{1.15}$$

The maximal function is a priori well defined as measurable function if f is in the Schwartz class; alternatively we may just restrict to countable E (see §7.1.1 for comments why this is not a significant restriction).

For the formulation of our theorem we also need the rescaled local operators $M_{E_j}^\sigma$ with

$$E_j = (2^{-j}E) \cap [1,2]. \tag{1.16}$$

A model case is given when E consists of all dyadic dilates of a set in [1.2], *i.e.*

$$E = \bigcup_{j\in\mathbb{Z}} 2^j E^\circ \quad \text{with } E^\circ \subset [1,2].$$

In this case

$$M_{E_j}^\sigma = M_{2^{-j}E\cap[1,2]}^\sigma = M_{E^\circ}^\sigma \quad \text{for all } j \in \mathbb{Z}.$$

Definition. The *Lebesgue exponent set* of the pair (σ, E), denoted by $\mathcal{L}(\sigma, E)$, consists of all $(1/p, 1/q)$ for which

$$\|M_E^\sigma\|_{L^p\to L^{p,\infty}} + \|M_E^\sigma\|_{L^{q,1}\to L^q} + \sup_{j\in\mathbb{Z}} \|M_{E_j}^\sigma\|_{L^p\to L^q} < \infty. \tag{1.17}$$

The *sparse exponent set* of M_E, denoted by $\mathrm{Sp}[M_E^\sigma]$ consists of all pairs $(1/p_1, 1/p_2)$ with $1/p_2 \geq 1/p_1$ for which there is $0 < \gamma < 1$ and a constant C such that

$$\int_{\mathbb{R}^d} M_E^\sigma f(x)\omega(x)dx \leq C\, \Lambda^*_{p_1,p_2}(f, \omega)$$

for all simple f and simple nonnegative ω.

Let $\varepsilon > 0$. We let $\mathcal{E}_{\mathrm{ann}}(\lambda)$ be the space of tempered distributions whose Fourier transform is supported in $\{\xi : \lambda/2 < |\xi| < 2\lambda\}$. We say that the pair (σ, E) satisfies an *ε-regularity condition* if there exists $C \geq 0$, and an exponent $p_0 \geq 1$ such that for all $\lambda > 2$, $j \in \mathbb{Z}$, we have

$$\|M_{E_j}^\sigma f\|_{p_0} \leq C\lambda^{-\varepsilon}\|f\|_{p_0} \quad \text{for all } f \in \mathcal{S} \cap \mathcal{E}_{\mathrm{ann}}(\lambda). \tag{1.18}$$

Remark. The usual lacunary maximal operator correspond to the case where $E = \mathbb{Z}$. Under this assumption, M_E^σ satisfies an ε-regularity condition for some $\varepsilon > 0$ if and only if there is an $\varepsilon' > 0$ such that

$$\widehat{\sigma}(\xi) = O(|\xi|^{-\varepsilon'}).$$

Moreover the condition $\sup_{j\in\mathbb{Z}} \|M_{E_j}^\sigma\|_{L^p\to L^q} < \infty$ is, in this case, equivalent with the L^p improving inequality

$$\|\sigma * f\|_q \lesssim \|f\|_p$$

for all $f \in L^p$.

Denote by $\mathrm{Int}(\Omega)$ the interior of a planar set Ω. Define $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\Phi(x, y) = (x, 1 - y).$$

We will show that, under the assumption of an ε-regularity condition for some $\varepsilon > 0$, the interiors of $\mathcal{L}(\sigma, E)$ and $\mathrm{Sp}[M_E^\sigma]$ are in unique correspondence under Φ (see Figure 1). That is,

$$\mathrm{Int}(\mathrm{Sp}[M_E^\sigma]) = \Phi(\mathrm{Int}(\mathcal{L}(\sigma, E))); \tag{1.19}$$

this can be deduced as a consequence of Corollary 1.2 and Theorem 1.3. The next theorem contains a slightly more precise statement.

Theorem 1.4. *Suppose that σ is a compactly supported distribution supported in $\mathbb{R}^d\setminus\{0\}$, and suppose that (σ, E) satisfies the ε-regularity condition* (1.18) *for some $\varepsilon > 0$. Let $1 < p < q < \infty$. Then the following implications hold:*

$$(\tfrac{1}{p}, \tfrac{1}{q}) \in \mathrm{Int}(\mathcal{L}(\sigma, E)) \Longrightarrow (\tfrac{1}{p}, \tfrac{1}{q'}) \in \mathrm{Sp}[M_E^\sigma]\,, \tag{1.20}$$

$$(\tfrac{1}{p}, \tfrac{1}{q}) \in \mathcal{L}(\sigma, E) \Longleftarrow (\tfrac{1}{p}, \tfrac{1}{q'}) \in \mathrm{Sp}[M_E^\sigma]\,. \tag{1.21}$$

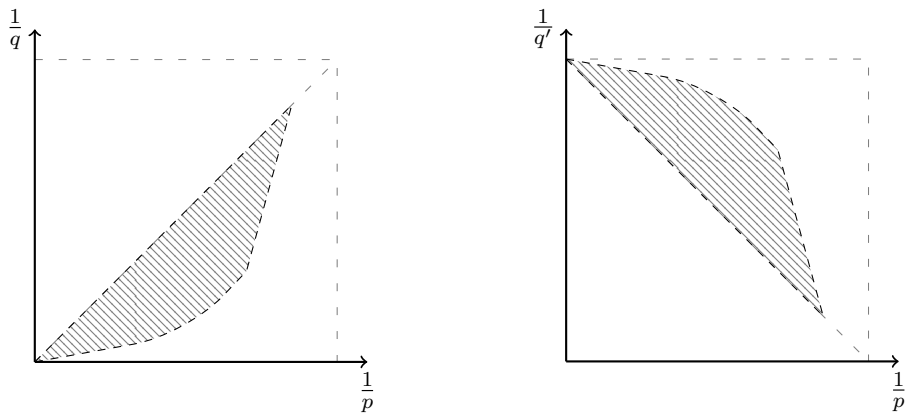

FIGURE 1. Example for $\mathcal{L}(\sigma, E)$ (left) and $\mathrm{Sp}[M_E^\sigma]$ (right). It may occur that the closure of $\mathcal{L}(\sigma, E)$ is not a polygonal region, see for example [93].

Remarks. (i) The correspondence (1.19) is an immediate consequence of Theorem 1.4.

(ii) If σ is as in Theorem 1.4 then similar statements characterizing the sparse exponent set hold for variation norm operators. See the statement of Propositions 7.2.

(iii) In the case of σ being the surface measure on the unit sphere one recovers as a special case the results by Lacey [66] on the lacunary and full spherical maximal functions.

1.4. **Fourier multipliers.** Given a bounded function m we consider the convolution operator $\mathcal{T}$ given on Schwartz functions $f : \mathbb{R}^d \to \mathbb{C}$ by

$$\widehat{\mathcal{T}f}(\xi) = m(\xi)\widehat{f}(\xi), \qquad \xi \in \widehat{\mathbb{R}}^d, \tag{1.22}$$

i.e. $\mathcal{T}f = \mathcal{F}^{-1}[m] * f$ where $\mathcal{F}^{-1}[m]$ is the Fourier inverse of m in the sense of tempered distributions. If $1 \le p < \infty$, we say that $m \in M^p$ if $\mathcal{T}$ extends to a bounded operator on L^p and we define $\|m\|_{M^p}$ to be the $L^p \to L^p$ operator norm of $\mathcal{T}$. A similar definition applies to $p = \infty$; however one replaces L^∞ by the space C_0 of continuous functions that vanish at ∞ (i.e. the closure of the Schwartz functions in the L^∞ norm). By duality we have $M^p = M^{p'}$ for $1/p' = 1 - 1/p$. Moreover, $M^2 = L^\infty$, $M^p \subset L^\infty$ and M^1 is the space of Fourier transforms of finite Borel measures. Similarly, if $1 \le p, q < \infty$, we say that $m \in M^{p,q}$ if $\mathcal{T}$ is bounded from L^p to L^q and we define by $\|m\|_{M^{p,q}}$ to be the $L^p \to L^q$ operator norm of $\mathcal{T}$. For these and other simple facts on Fourier multipliers see [53] or [106].

Let ϕ be a nontrivial radial C_c^∞ function compactly supported in $\widehat{\mathbb{R}}^d \setminus \{0\}$. A natural single scale assumption would be to assume a uniform M^{p_0} bound for the pieces $\phi(t^{-1}\cdot)m$ which is equivalent by dilation-invariance to the

condition

$$\sup_{t>0} \|\phi m(t\cdot)\|_{M^{p_0}} < \infty. \tag{1.23}$$

Inequality (1.23) is a necessary and sufficient condition for $\mathcal{T}$ to be bounded on the homogeneous Besov spaces $\dot{B}^s_{p_0,q}$, for any $s \in \mathbb{R}, 0 < q \leq \infty$; see [107], [110, §2.6]. However, it does *not* imply boundedness on the Lebesgue spaces, except on L^2. Indeed, Littman, McCarthy and Rivière [84] and Stein and Zygmund [107] give examples of m satisfying (1.23) for a $p_0 \neq 2$ for which $m \notin M^p$ for all $p \neq 2$.

The papers by Carbery [24] and by one of the authors [98] provide positive results under an additional dilation invariant regularity condition,

$$\sup_{t>0} \|\phi m(t\cdot)\|_{C^\varepsilon} < \infty, \tag{1.24}$$

where C^ε is the standard Hölder space. Indeed, it is shown in [24, 98] that for $1 < p_0 < 2$, $0 < \varepsilon < 1$,

$$\|m\|_{M^p} \leq C(p,\varepsilon) \sup_{t>0} \big(\|\phi m(t\cdot)\|_{M^{p_0}} + \|\phi m(t\cdot)\|_{C^\varepsilon}\big), \quad p_0 < p < p_0'.$$

If the standard Hölder condition $\|\phi m(t\cdot)\|_{C^\varepsilon} = O(1)$ is replaced by its M^{p_0} variant, $\sup_{t>0} \|\Delta_h[\phi m(t\cdot)]\|_{M^{p_0}} = O(|h|^\varepsilon)$, one obtains a conclusion for $p = p_0$. We will show that for fixed $p \in (p_0, p_0')$, the L^p-boundedness self-improves to a sparse domination inequality.

Theorem 1.5. *Let $1 < p_0 < 2$, $0 < \varepsilon < 1$, and assume that* (1.23) *and* (1.24) *hold. Then for every $p \in (p_0, 2]$ there is a $\delta = \delta(p) > 0$ such that $\mathcal{T} \in \mathrm{Sp}(p-\delta, p'-\delta)$.*

We note that, in view of the compact support, for $p \leq q$ the quantity $\|\phi m(t\cdot)\|_{M^{p,q}}$ can be bounded by $C\|\phi m(t\cdot)\|_{M^p}$ via Young's inequality. In Theorem 1.5 the self-improvement to a sparse bound is due to a tiny bit of regularity as hypothesized in (1.24). This together with (1.23) implies a mild regularity condition for $\phi m(t\cdot)$ measured in the $M^{p,q}$ norm. If one seeks better results on the sparse bound in terms of q a further specification of this regularity is needed. For this we use the iterated difference operators

$$\Delta_h^M = \Delta_h \Delta_h^{M-1} \quad \text{for } M \geq 2,$$

where Δ_h is as in (1.10). With ϕ as above we get the following.

Theorem 1.6. *Let $m \in L^\infty(\mathbb{R}^d)$ and $\mathcal{T}$ as in* (1.22)*. Let $1 < p \leq q < \infty$. Assume that there exists $s > d(1/p - 1/q)$ and an $M \in \mathbb{N}$ such that*

$$\sup_{t>0} \sup_{|h|\leq 1} |h|^{-s} \big\|\Delta_h^M[\phi m(t\cdot)]\big\|_{M^{p,q}} < \infty. \tag{1.25}$$

Then $\mathcal{T} \in \mathrm{Sp}(p, q')$.

One should always take $M > s$. Indeed, note that if $M < s$, then (1.25) implies $m \equiv 0$. We note that the $L^p \to L^q$ conditions (1.8), (1.9) in Theorem 1.1 correspond in the instance of convolution operators to an $M^{p,q}$ condition

of derivatives of order $s > d(1/p - 1/q)$ on the localizations of the Fourier multiplier. Also, for fixed $s > d(1/p-1/q)$, if (1.25) holds with some $M \geq s$, then it holds for all integers $M > s$. For an illustration of this and the broad scope of Theorem 1.6, see the discussion on singular Radon transforms in §7.3.1 and on various classes of Fourier multipliers related to oscillatory multipliers in §7.2 and to radial multipliers in §7.5.

Theorems 1.5 and 1.6 will be deduced in §6.1 from the more precise, but also more technical Theorem 6.1 which expresses the regularity via dyadic decompositions of $\mathcal{F}^{-1}[\phi m(t\cdot)]$. Moreover, there we will cover a version involving Hilbert space valued functions which is useful for sparse domination results for objects such as Stein's square function associated with Bochner–Riesz means.

1.5. **Application to weighted norm inequalities.** It is well known that sparse domination implies a number of weighted inequalities in the context of Muckenhoupt and reverse Hölder classes of weights, and indeed this serves as a first motivation for the subject; see the lecture notes by Pereyra [91] for more information. Here we just cite a general result about this connection which can be directly applied to all of our results on sparse domination and is due to Bernicot, Frey and Petermichl [18]. Recall the definition of the Muckenhoupt class A_t consisting of weights for which

$$[w]_{A_t} = \sup_Q \langle w\rangle_{Q,1}\langle w^{-1}\rangle_{Q,t'-1} < \infty,$$

and the definition of the reverse Hölder class RH_s consisting of weights for which

$$[w]_{\mathrm{RH}_s} = \sup_Q \frac{\langle w\rangle_{Q,s}}{\langle w\rangle_{Q,1}} < \infty.$$

In both cases the supremum is taken over all cubes Q in $\mathbb{R}^d$.

Proposition 1.7 ([18])**.** *If $T \in \mathrm{Sp}(L^p_{B_1}, L^{q'}_{B_2^*})$, then one has the weighted norm inequality*

$$\Big(\int_{\mathbb{R}^d} |Tf(x)|^r_{B_2} w(x)\,dx\Big)^{\frac{1}{r}} \lesssim \|T\|_{\mathrm{Sp}_\gamma(p_1,B_1;p_2,B_2^*)} \times \\ ([w]_{A_{r/p}}[w]_{\mathrm{RH}_{(q/r)'}})^\alpha \Big(\int_{\mathbb{R}^d} |f(x)|^r_{B_1} w(x)\,dx\Big)^{\frac{1}{r}}$$

for all $w \in A_{r/p} \cap \mathrm{RH}_{(q/r)'}$ and $p < r < q$, where $\alpha := \max(\frac{1}{r-p}, \frac{q-1}{q-r})$.

We refer to [18, §6] for more information and a detailed exposition. See also [44] for other weighted norm inequalities.

1.6. **Structure of the paper and notation.** We begin addressing necessary conditions, and prove Theorem 1.3 in §2. In §3 we review useful preliminary facts needed in the proof of Theorem 1.1 regarding the single

scale regularity conditions; in particular, an alternative form for the regularity conditions in (1.9). The proof of Theorem 1.1 is presented in §4. The main part of the argument consists of an induction step, which is contained in §4.4. The implication that yields Corollary 1.2 from Theorem 1.1 is given in §4.3. In §5 we apply Corollary 1.2 to deduce sparse domination results for maximal functions, square functions and variation norm operators, as well as Cotlar-type operators associated to truncations of operators. In the case of maximal functions, the assumptions of Theorem 1.1 can be slightly weakened, and we present this in §5.4. Theorems 1.5 and 1.6 are proved in §6. Finally, in §7 we apply our main theorems to several specific examples, including the proof of Theorem 1.4 in §7.1.1. Moreover, we give several applications of Theorem 1.6 to specific classes of multipliers. For completeness, we include several appendices. Appendix A covers some basic facts on sparse domination. Appendix B covers versions of the main Theorem 1.1 for $p = 1$ and/or $q = \infty$. Some basic facts on Fourier multipliers needed in §6 are covered in Appendix C.

Notation. The notation $A \lesssim B$ will be used to denote that $A \leq C \cdot B$, where the constant C may change from line to line. Dependence of C on various parameters may be denoted by a subscript or will be clear from the context. We use $A \approx B$ to denote that $A \lesssim B$ and $B \lesssim A$.

We shall use the definition $\widehat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^d} e^{-i\langle y,\xi\rangle} f(y)dy$ for the Fourier transform on $\mathbb{R}^d$. We let $\mathcal{F}^{-1}$ be the inverse Fourier transform and use the notation $m(D)f = \mathcal{F}^{-1}[m\widehat{f}]$. We denote by $\mathcal{S} \equiv \mathcal{S}(\mathbb{R}^d)$ the space of Schwartz functions on $\mathbb{R}^d$, by $\mathcal{S}'$ the space of tempered distributions on $\mathbb{R}^d$, and by $\mathcal{E}_{\text{ann}}(\lambda)$ the space of all $f \in \mathcal{S}'$ such that the Fourier transform $\widehat{f}$ is supported in the open annulus $\{\xi \in \hat{\mathbb{R}}^d : \lambda/2 < |\xi| < 2\lambda\}$.

For a d-dimensional rectangle $R = [a_1, b_1] \times \cdots \times [a_d, b_d]$ we denote by x_R the center of R, i.e. the points with coordinates $x_{R,i} = (a_i + b_i)/2$, $i = 1, \ldots, d$. If $\tau > 0$, we denote by τR to be the τ-dilate of R with respect to its center, i.e.

$$\tau R = \Big\{ x \in \mathbb{R}^d \,:\, x_R + \frac{x - x_R}{\tau} \in R \Big\}.$$

We shall use many spatial or frequency decomposition throughout the paper:

- $\{\lambda_k\}_{k\geq 0}$, $\{\widetilde{\lambda}_k\}_{k\geq 0}$ are specific families of Littlewood–Paley type operators that can be used for a reproducing formula (3.1); they are compactly supported and have vanishing moments (*cf.* §3.2);
- $\{\Psi_\ell\}_{\ell\geq 0}$ is an inhomogeneous dyadic decomposition in x-space, compactly supported where $|x| \approx 2^\ell$ if $\ell > 0$ (*cf.* §6.1);
- $\{\eta_\ell\}_{\ell\geq 0}$ is an inhomogeneous dyadic frequency decomposition so that $\widehat{\eta}_\ell$ is supported where $|\xi| \approx 2^\ell$ if $\ell > 0$ (*cf.* §3.3, §7.4).

Similarly, we shall use the following bump functions:

- ϕ is a radial $C^\infty(\widehat{\mathbb{R}}^d)$ function supported in $|\xi| \approx 1$ and not identically zero (*cf.* §6.1);
- θ is a radial $C^\infty(\mathbb{R}^d)$ function supported in $|x| \lesssim 1$ with vanishing moments and such that $\widehat{\theta}(\xi) > 0$ in $|\xi| \approx 1$ (*cf.* §6.1);
- β is any nontrivial $C_c^\infty(\widehat{\mathbb{R}})$ function with compact support in $(0, \infty)$ (*cf.* §7.5).

Acknowledgements. We would like to thank Richard Oberlin and Luz Roncal for informative conversations regarding sparse domination. We thank the referee for various valuable suggestions and in particular for recommending to formulate versions of our results for the cases $p = 1$ and $q = \infty$. D.B. was partially supported by NSF grant DMS-1954479. J.R. and A.S. were partially supported by NSF grant DMS-1764295, and A.S. was partially supported by a Simons fellowship. We thank Tuomas Hytönen for pointing out an inaccuracy in the original formula (4.24).

Remark. After we circulated the first version of this paper, José M. Conde-Alonso, Francesco Di Plinio, Ioannis Parissis and Manasa N. Vempati kindly shared their preprint [30], in which they develop a metric theory of sparse domination on spaces of homogeneous type. There is a small overlap with our work, as [30] covers classes of singular Radon transforms, in particular non-isotropic versions of results in §7.3.

2. Necessary conditions

In this section we prove Theorem 1.3 and another partial converse for Theorem 1.1, namely Theorem 2.5 below.

We begin with an immediate and well known, but significant estimate for the maximal sparse forms which will lead to simple necessary conditions. In what follows, let $\mathcal{M}$ denote the Hardy–Littlewood maximal operator.

Lemma 2.1. *The following hold for the maximal forms defined in* (1.3).

(i) For $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$,

$$\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1, f_2) \le \gamma^{-1} \int_{\mathbb{R}^d} (\mathcal{M}[|f_1|^{p_1}_{B_1}](x))^{1/p_1} (\mathcal{M}[|f_2|^{p_2}_{B_2^*}](x))^{1/p_2} dx. \tag{2.1}$$

(ii) If $1 \le p_1 < p$, and $f_1 \in L^{p,1}_{B_1}$, $f_2 \in L^{p'}_{B_2^*}$, then

$$\Lambda^*_{p_1,B_1,p',B_2^*}(f_1, f_2) \lesssim_{p,p_1} \gamma^{-1} \|f_1\|_{L^{p,1}_{B_1}} \|f_2\|_{L^{p'}_{B_2^*}}, \tag{2.2}$$

(iii) If $1 < p < p_2'$, and $f_1 \in L^p_{B_1}$, $f_2 \in L^{p',1}_{B_2^*}$, then

$$\Lambda^*_{p,B_1,p_2,B_2^*}(f_1, f_2) \lesssim_{p,p_2} \gamma^{-1} \|f_1\|_{L^p_{B_1}} \|f_2\|_{L^{p',1}_{B_2^*}}, \tag{2.3}$$

Proof. For a γ-sparse family of cubes we have

$$\Lambda^{\mathfrak{S}}_{p_1,B_1,p_2,B_2^*}(f_1,f_2) \le \sum_{Q\in\mathfrak{S}} \frac{1}{\gamma}\int_{E_Q} (\mathcal{M}[|f_1|^{p_1}_{B_1}](x))^{1/p_1}(\mathcal{M}[|f_2|^{p_2}_{B_2^*}](x))^{1/p_2}dx$$

and (2.1) follows by the disjointness of the sets E_Q and taking supremum over all sparse families.

Now let $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$. For (2.2) we use (2.1), with $p_2 = p'$, together with the fact that for $p_1 < p$ the operator $g \mapsto (\mathcal{M}|g|^{p_1}_{B_1})^{1/p_1}$ maps $L^{p,1}_{B_1}$ to itself; this follows by real interpolation from the fact that it maps L^p to itself, for all $p > p_1$. We can now estimate

$$\begin{aligned}\Lambda^*_{p_1,B_1,p',B_2^*}(f_1,f_2) &\lesssim \gamma^{-1}\|(\mathcal{M}[|f_1|^{p_1}_{B_1}])^{1/p_1}\|_{L^{p,1}}\|(\mathcal{M}[|f_2|^{p'}_{B_2^*}])^{1/p'}\|_{L^{p',\infty}}\\ &\lesssim_{p.p_1} \gamma^{-1}\|f_1\|_{L^{p,1}_{B_1}}\|f_2\|_{L^{p'}_{B_2^*}}.\end{aligned}$$

Since simple B_1-valued functions are dense in $L^{p,1}_{B_1}$ and simple B_2^*-valued functions are dense in $L^{p'}_{B_2^*}$ we get (2.7) for all $f_1 \in L^{p,1}_{B_1}$ and $f_2 \in L^{p'}_{B_2^*}$, by a straightforward limiting argument.

For (2.3) we argue similarly. We use (2.1) with $p_1 = p$, together with the fact that for $p_2 < p'$ the operator $g \mapsto (\mathcal{M}|g|^{p_2}_{B_2^*})^{1/p_2}$ maps $L^{p',1}_{B_2^*}$ to itself, and hence

$$\begin{aligned}\Lambda^*_{p,B_1,p_2,B_2^*}(f_1,f_2) &\lesssim \gamma^{-1}\|(\mathcal{M}[|f_1|^{p}_{B_1}])^{1/p}\|_{L^{p,\infty}}\|(\mathcal{M}[|f_2|^{p_2}_{B_2^*}])^{1/p_2}\|_{L^{p',1}}\\ &\lesssim_{p,p_2} \gamma^{-1}\|f_1\|_{L^{p}_{B_1}}\|f_2\|_{L^{p',1}_{B_2^*}}.\end{aligned}$$

□

The estimates in Lemma 2.1 immediate yield estimates for the forms $\langle Tf_1, f_2\rangle$, since by the definition (1.5)

$$|\langle Tf_1,f_2\rangle| \le \|T\|_{\mathrm{Sp}_\gamma(p_1,B_1;p_2,B_2^*)}\,\Lambda^{\gamma,*}_{p_1,B_1,p_2,B_2^*}(f_1,f_2).$$

We shall now prove Theorem 1.3 in §2.1, and a variant under reflexivity of B_2 in §2.2.

2.1. **The local integrability hypothesis.** If $Tf_1 \in L^1_{B_2,\mathrm{loc}}$, Lemma 2.1 further yields bounds for the $L^p_{B_2}$ or $L^{p,\infty}_{B_2}$ norms of Tf_1 via a duality result for scalar functions.

Lemma 2.2. *Suppose $T : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$. Then the following hold.*

(i) If $1 \le p_1 < p < \infty$ and if for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative simple functions ω

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2}\omega(x)dx \le A_1\Lambda^*_{p_1,B_1,p',\mathbb{R}}(f,\omega),$$

then T extends to a bounded operator from $L^{p,1}_{B_1}$ to $L^p_{B_2}$ so that

$$\|T\|_{L^{p,1}_{B_1}\to L^p_{B_2}} \lesssim_{p_1,p} \gamma^{-1}A_1. \tag{2.4}$$

(ii) If $1 < p < p_2'$ and if for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative simple functions ω

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2} \omega(x) dx \le A_2 \Lambda^*_{p,B_1,p_2,\mathbb{R}}(f,\omega),$$

then T extends to a bounded operator from $L^p_{B_1}$ to $L^{p,\infty}_{B_2}$ so that

$$\|T\|_{L^p_{B_1} \to L^{p,\infty}_{B_2}} \lesssim_{p_2,p} \gamma^{-1} A_2. \tag{2.5}$$

Proof. We rely on Lemma 2.1. For part (i) we use (2.2) to estimate, for $f_2 \in \mathrm{S}_{B_2}$,

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2} \omega(x) dx \lesssim \gamma^{-1} A_1 \|f\|_{L^{p,1}_{B_1}} \|\omega\|_{L^{p'}}.$$

By L^p duality this implies an L^p bound for the locally integrable scalar function $x \mapsto |Tf(x)|_{B_2}$ and consequently $Tf \in L^p_{B_2}$ with

$$\|Tf\|_{L^p_{B_2}} \lesssim \gamma^{-1} A_1 \|f\|_{L^{p,1}_{B_1}}$$

and (2.4) follows.

For part (ii) we argue similarly. We use (2.3) to estimate

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2} \omega(x) dx \lesssim \gamma^{-1} A_2 \|f\|_{L^p_{B_1}} \|\omega\|_{L^{p',1}}.$$

By the duality $(L^{p',1})^* = L^{p,\infty}$ for scalar functions for $1 < p < \infty$ [55] we get

$$\|Tf\|_{L^{p,\infty}_{B_2}} \lesssim \gamma^{-1} A_2 \|f_1\|_{L^p_{B_1}}$$

and (2.5) follows. □

Corollary 2.3. *Assume that $T : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$ and let $1 \le p_1 < p < p_2'$. If for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative simple functions ω*

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2} \omega(x) dx \le A\, \Lambda^*_{p_1,B,p_2,\mathbb{R}}(f,\omega),$$

then T extends to a bounded operator from $L^p_{B_1}$ to $L^p_{B_2}$ so that

$$\|T\|_{L^p_{B_1} \to L^p_{B_2}} \lesssim_{p,p_1,p_2} \gamma^{-1} A. \tag{2.6}$$

Proof. Lemma 2.2 implies T maps boundedly $L^{\tilde p_1}_{B_1} \to L^{\tilde p_1,\infty}_{B_2}$ and $L^{\tilde p_2',1}_{B_1} \to L^{\tilde p_2'}_{B_2}$ for any $p_1 < \tilde p_1 < p$ and $p < \tilde p_2' < p_2'$; the desired $L^p_{B_1} \to L^p_{B_2}$ boundedness for $p_1 < p < p_2'$ then follows by interpolation.

Alternatively, one could deduce this result directly from (2.1). Arguing as in the proofs of (ii) or (iii) in Lemma 2.1, by the Hardy–Littlewood theorem and (2.1) one has

$$\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1,f_2) \lesssim_{p,p_1,p_2} \gamma^{-1} \|f_1\|_{L^p_{B_1}} \|f_2\|_{L^{p'}_{B_2^*}} \tag{2.7}$$

for $1 \le p_1 < p < p_2'$. Then one can argue as in the proof of Lemma 2.2, to deduce (2.6) from (2.7). $\square$

We next turn to the necessity of the condition (1.8) in Corollary 1.2 and Theorem 1.1. In this generality, this type of implication appears to be new in the sparse domination literature. It is inspired by the philosophy of adapting the counterexamples for $L^p \to L^q$ estimates to sparse bounds (see *i.e.* the examples for spherical maximal operators in [66]).

Lemma 2.4. *Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators, with $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$, and satisfying the strengthened support condition* (1.14)*. Let $1 \le p < q \le \infty$ and suppose that for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative simple function ω, the estimate*

$$\int_{\mathbb{R}^d} |T_j f(x)|_{B_2} \omega(x)dx \le \mathcal{C}\, \Lambda^*_{p,B_1,q',\mathbb{R}}(f,\omega)$$

holds uniformly in $j \in \mathbb{Z}$. Then

$$\sup_{j\in\mathbb{Z}} \|\mathrm{Dil}_{2^j} T_j\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim_{\gamma,d,\delta_1,\delta_2,p,q} \mathcal{C}.$$

Proof. Fix $j \in \mathbb{Z}$ and let $S = \mathrm{Dil}_{2^j} T_j$. We first apply a scaling argment. Note that by assumption

$$\begin{aligned}\int |Sf(x)|_{B_2}\omega(x)dx &= 2^{-jd}\int \big|T_j[f(2^{-j}\cdot)](x)\big|_{B_2}\omega(2^{-j}x)dx \\ &\le \mathcal{C}\, 2^{-jd}\Lambda^*_{p,B_1,q',\mathbb{R}}(f(2^{-j}\cdot),\omega(2^{-j}\cdot))\end{aligned}$$

If $\Lambda^{\mathfrak{S}}_{p,B_1,q,\mathbb{R}}$ is a sparse form with a γ-sparse collection of cubes we form the collection $\mathfrak{S}_j$ of dilated cubes $\{2^{-j}y : y \in Q\}$ where $Q \in \mathfrak{S}$. Then

$$2^{-jd}\Lambda^{\mathfrak{S}}_{p,B_1,q',\mathbb{R}}(f(2^{-j}\cdot),\omega(2^{-j}\cdot)) = \Lambda^{\mathfrak{S}_j}_{p,B_1,q',\mathbb{R}}(f,\omega)$$

and therefore we get the estimate

$$\int |Sf(x)|_{B_2}\omega(x)dx \le \mathcal{C}\Lambda^*_{p,B_1,q',\mathbb{R}}(f,\omega). \tag{2.8}$$

Suppose that b is the smallest positive integer such that

$$2^{-b} \le d^{-1/2}\min\{\delta_1/2,\delta_2\}.$$

For $\mathfrak{z} \in \mathbb{Z}^d$ let

$$Q_{\mathfrak{z}} = \{x : 2^{-b}\mathfrak{z}_i \le x_i < 2^{-b}(\mathfrak{z}_i+1),\ i = 1,\dots,d\}$$

and let $f_{\mathfrak{z}} = f\mathbb{1}_{Q_{\mathfrak{z}}}$. Let $R_{\mathfrak{z}}$ the cube of side length 3 centered at $2^{-b}\mathfrak{z}$. Then $Sf_{\mathfrak{z}}$ is supported in $R_{\mathfrak{z}}$. We decompose $R_{\mathfrak{z}}$ into 3^d2^{bd} cubes $R_{\mathfrak{z},\nu}$ of side length 2^{-b}, here $\nu \in \mathcal{I}_{\mathfrak{z}}$ with $\#\mathcal{I}_{\mathfrak{z}} = 3^d2^{bd}$.

Fix $\mathfrak{z},\nu$ and a simple nonnegative function ω with $\|\omega\|_{L^{q'}} \le 1$. We first prove that for $\nu \in \mathcal{I}_{\mathfrak{z}}$

$$\int |Sf_{\mathfrak{z}}(x)|_{B_2}\omega(x)\mathbb{1}_{R_{\mathfrak{z},\nu}}(x)dx \lesssim \mathcal{C}\,\|f_{\mathfrak{z}}\|_{L^p_{B_1}}\|\omega\|_{L^{q'}}. \tag{2.9}$$

In this argument we shall not use strong measurability of $Sf_{\mathfrak{z}}$. By (2.8) we have

$$\int |Sf_{\mathfrak{z}}(x)|_{B_2}\omega(x)\mathbb{1}_{R_{\mathfrak{z},\nu}}(x)dx \le \mathcal{C}\,\Lambda^*_{p,B_1,q',\mathbb{R}}(f_{\mathfrak{z}},\omega\mathbb{1}_{R_{\mathfrak{z},\nu}})$$

and therefore we find a sparse family $\mathfrak{S}_{\mathfrak{z},\nu}$ such that

$$\int |Sf_{\mathfrak{z}}(x)|_{B_2}\omega(x)\mathbb{1}_{R_{\mathfrak{z},\nu}}(x)dx \le 2\mathcal{C}\sum_{Q\in\mathfrak{S}_{\mathfrak{z},\nu}}|Q|\langle f_{\mathfrak{z}}\rangle_{Q,p}\langle\omega\mathbb{1}_{R_{\mathfrak{z},\nu}}\rangle_{Q,q'}. \tag{2.10}$$

By the strengthened support condition, (1.14),

$$Sf_{\mathfrak{z}}\mathbb{1}_{R_{\mathfrak{z},\nu}}\neq 0 \implies \operatorname{dist}(Q_{\mathfrak{z}},R_{\mathfrak{z},\nu})\ge \delta_1-2^{-b}\sqrt{d}. \tag{2.11}$$

Assuming that the left-hand side is not 0 in (2.10), and in view of (2.11), we see that for a cube $Q\in\mathfrak{S}_{\mathfrak{z},\nu}$ we have, recalling that $\delta_1\ge\sqrt{d}\,2^{-b+1}$,

$$\left.\begin{aligned} Q_{\mathfrak{z}}\cap Q\neq\emptyset\\ R_{\mathfrak{z},\nu}\cap Q\neq\emptyset\end{aligned}\right\}\implies \operatorname{diam}(Q)\ge\delta_1-2^{-b}\sqrt{d}\ge 2^{-b}\sqrt{d}.$$

Hence all cubes that contribute to the sum in (2.10) have side length $\ge 2^{-b}$. Denote the cubes in $\mathfrak{S}_{\mathfrak{z},\nu}$ with side length in $[2^\ell,2^{\ell+1})$ by $\mathfrak{S}_{\mathfrak{z},\nu}(\ell)$ and note that for every $\ell\ge -b$ there are at most $C(d)$ many cubes that contribute. Hence we may estimate

$$\begin{aligned}
&\sum_{Q\in\mathfrak{S}_{\mathfrak{z},\nu}}|Q|\langle f_{\mathfrak{z}}\rangle_{Q,p}\langle\omega\mathbb{1}_{R_{\mathfrak{z},\nu}}\rangle_{Q,q'}\\
&\le\sum_{\ell\ge -b}\sum_{Q\in\mathfrak{S}_{\mathfrak{z},\nu}(\ell)}|Q|^{\frac1q-\frac1p}\Big(\int_Q|f_{\mathfrak{z}}(y)|^p_{B_1}dy\Big)^{\frac1p}\Big(\int_Q|\omega(x)\mathbb{1}_{R_{\mathfrak{z},\nu}}(x)|^{q'}dx\Big)^{\frac1{q'}}\\
&\lesssim_d\sum_{\ell\ge -b}2^{\ell d(\frac1q-\frac1p)}\|f_{\mathfrak{z}}\|_{L^p_{B_1}}\|\omega\|_{L^{q'}}\lesssim_{b,d,p,q}\|f_{\mathfrak{z}}\|_{L^p_{B_1}}
\end{aligned}$$

where we used the assumption $q>p$ to sum in ℓ. This establishes (2.9).

By duality combined with (2.9) we have

$$\|Sf_{\mathfrak{z}}\|_{L^q_{B_2}(R_{\mathfrak{z},\nu})}\lesssim\sup_{\substack{\omega\in\mathrm{S}_{\mathbb{R}}\\ \|\omega\|_{L^{q'}}\le 1}}\int|Sf_{\mathfrak{z}}(x)|_{B_2}|\omega(x)|\mathbb{1}_{R_{\mathfrak{z},\nu}}(x)dx\lesssim\mathcal{C}\,\|f_{\mathfrak{z}}\|_{L^p_{B_1}}. \tag{2.12}$$

Considering this for various $\nu\in\mathcal{I}_{\mathfrak{z}}$ we get

$$\|Sf_{\mathfrak{z}}\|_{L^q_{B_2}}\lesssim\sum_{\nu\in\mathcal{I}_{\mathfrak{z}}}\|Sf_{\mathfrak{z}}\|_{L^q_{B_2}(R_{\mathfrak{z},\nu})}\lesssim\sum_{\nu\in\mathcal{I}_{\mathfrak{z}}}\mathcal{C}\,\|f_{\mathfrak{z}}\|_{L^p_{B_1}}\lesssim_{d,\delta_1,\delta_2}\mathcal{C}\,\|f_{\mathfrak{z}}\|_{L^p_{B_1}}.$$

Then

$$\|Sf\|_{L^q_{B_2}} = \Big\| \sum_{\mathfrak{z}\in\mathbb{Z}^d} Sf_{\mathfrak{z}} \Big\|_{L^q_{B_2}} \lesssim C_d 2^{bd} \Big(\sum_{\mathfrak{z}\in\mathbb{Z}^d} \|Sf_{\mathfrak{z}}\|^q_{L^q_{B_2}} \Big)^{1/q} \tag{2.13}$$

$$\lesssim_{d,\delta_1,\delta_2} \mathcal{C} \Big(\sum_{\mathfrak{z}\in\mathbb{Z}^d} \|f_{\mathfrak{z}}\|^q_{L^p_{B_1}} \Big)^{1/q} \lesssim \mathcal{C} \Big(\sum_{\mathfrak{z}\in\mathbb{Z}^d} \|f_{\mathfrak{z}}\|^p_{L^p_{B_1}} \Big)^{1/p} \lesssim \mathcal{C}\, \|f\|_{L^p_{B_1}}. \square$$

Theorem 1.3 now follows from Lemmata 2.2 and 2.4.

2.2. **The reflexivity hypothesis.** In this section we prove a version of Theorem 1.3 where we drop the a priori assumption on T_j sending S_{B_1} to $L^1_{B_2,\mathrm{loc}}$ and thus we can no longer assume the strong B_2 measurability of Tf. We still get a partial converse to Theorem 1.1 if we assume that the Banach space B_2 is reflexive.

Theorem 2.5. *Let B_2 be reflexive and let $1 < p < q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying the support condition* (1.6)*. Assume the conclusion of Theorem 1.1, that is, there exists $\mathcal{C} > 0$ such that for all N_1, N_2 with $N_1 \le N_2$ and every $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$,*

$$\Big| \Big\langle \sum_{j=N_1}^{N_2} T_j f_1, f_2 \Big\rangle \Big| \le \mathcal{C}\, \Lambda^*_{p,B_1,q',B_2^*}(f_1, f_2).$$

Then

(i) Conditions (1.7a) *and* (1.7b) *hold, i.e., there is a constant $c > 0$ only depending on d, p, q, γ such that for all N_1, N_2 with $N_1 \le N_2$,*

$$\Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{L^p_{B_1}\to L^{p,\infty}_{B_2}} \le c\,\mathcal{C}, \qquad \Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{L^{q,1}_{B_1}\to L^q_{B_2}} \le c\,\mathcal{C}.$$

(ii) If, in addition, the T_j satisfy the strengthened support condition (1.14) *then condition* (1.8) *holds, i.e., there is a constant $c > 0$ only depending on d, p, q, γ such that*

$$\sup_{j\in\mathbb{Z}} \|\mathrm{Dil}_{2^j} T_j\|_{L^p_{B_1}\to L^q_{B_2}} \le c\,\mathcal{C}.$$

In the vector valued setting of Theorem 2.5 we need to use a more abstract duality argument which requires some care because of a potential lack of strong local integrability. We briefly discuss the issue of duality.

Let B be a Banach space. Recall that for $1 \le p \le \infty$, $1/p + 1/p' = 1$, the space $L^{p'}_{B^*}$ is embedded in $(L^p_B)^*$ via the canonical isometric homomorphism. In the scalar case this isometry is also surjective when $1 \le p < \infty$, and the proof of this fact relies on the Radon–Nikodym theorem. In the vector-valued case the surjectivity is equivalent with the dual space B^* having the *Radon–Nikodym property* (RNP) with respect to Lebesgue measure (see [56,

Chapter 1.3.b] for the formal definition). Thus under this assumption we have an identification of the dual of L^p_B with $L^{p'}_{B^*}$. To summarize,

$$B^* \in \mathrm{RNP} \iff (L^p_B)^* = L^{p'}_{B^*}, \quad 1 \le p < \infty. \tag{2.14}$$

Similarly, the Radon–Nikodym property for B^* also implies

$$(L^{p,r}_B)^* = L^{p',r'}_{B^*}, \quad 1 < p < \infty, \quad 1 \le r < \infty;$$

this is not stated in [56] but follows by a similar argument as in the scalar case [55, (2.7)], essentially with the exception of the application of the Radon–Nikodym property in place of the scalar Radon–Nikodym theorem. For a detailed discussion of the Radon–Nikodym properties and its applications we refer to [56, Chapter 1.3.c]. The class of spaces which have the Radon–Nikodym property with respect to all σ-finite measure spaces includes all *reflexive* spaces and also all spaces that have a *separable dual* (*cf.* [56, Theorem 1.3.21]). If B is reflexive, so is B^*, and therefore (2.14) holds for reflexive spaces B.

Under the assumption that the double dual B_2^{**} satisfies the Radon–Nikodym property, we can show that the sparse bound implies that Tf can be identified with a B_2^{**} strongly measurable function. This leads to a satisfactory conclusion under the stronger assumption that B_2 is reflexive.

Lemma 2.6. *Assume that $T \in \mathrm{Op}_{B_1,B_2}$ and that B_2^{**} satisfies the Radon–Nikodym property. Then the following hold.*

(i) If $1 \le p_1 < p < \infty$ and $T \in \mathrm{Sp}_\gamma(p_1, B_1, p', B_2^*)$ then T extends to a bounded operator from $L^{p,1}_{B_1}$ to $L^p_{B_2^{**}}$ so that

$$\|T\|_{L^{p,1}_{B_1} \to L^p_{B_2^{**}}} \lesssim_{p_1} \gamma^{-1} \|T\|_{\mathrm{Sp}_\gamma(p_1,B_1,p',B_2^*)}. \tag{2.15}$$

(ii) If $1 < p < p_2'$ and $T \in \mathrm{Sp}_\gamma(p, B_1, p_2, B_2^*)$ then T extends to a bounded operator from $L^p_{B_1}$ to $L^{p,\infty}_{B_2^{**}}$ so that

$$\|T\|_{L^p_{B_1} \to L^{p,\infty}_{B_2^{**}}} \lesssim_{p_2} \gamma^{-1} \|T\|_{\mathrm{Sp}_\gamma(p,B_1,p_2,B_2^*)}. \tag{2.16}$$

Proof. We rely, as in the proof of Lemma 2.2, on Lemma 2.1.

For part (i), we use (2.2) to obtain

$$|\langle Tf_1, f_2\rangle| \lesssim \gamma^{-1} \|T\|_{\mathrm{Sp}_\gamma(p,B_1;p_2,B_2^*)} \|f_1\|_{L^{p,1}_{B_1}} \|f_2\|_{L^{p'}_{B_2^*}}.$$

This inequality establishes the form $f_2 \mapsto \langle Tf_1, f_2\rangle$ as a linear functional on $L^{p'}_{B_2^*}$. Since B_2^{**} has the Radon–Nikodym property and thus $(L^{p'}_{B_2^*})^* = L^p_{B_2^{**}}$, we can identify Tf_1 as a member of $L^p_{B_2^{**}}$. Since

$$\|T\|_{L^{p,1}_{B_1} \to L^p_{B_2^{**}}} = \sup_{\|f_1\|_{L^{p,1}_{B_1}} \le 1} \sup_{\|f_2\|_{L^{p'}_{B_2^*}} \le 1} |\langle Tf_1, f_2\rangle|,$$

we have established (2.15).

Similarly, for part (ii) we use (2.3) to obtain

$$|\langle Tf_1, f_2\rangle| \lesssim \gamma^{-1}\|T\|_{\mathrm{Sp}_\gamma(p,B_1;p_2,B_2^*)}\|f\|_{L^p_{B_1}}\|g\|_{L^{p',1}_{B_2^*}}.$$

Since $L^{p,\infty}_{B_2^{**}}$ can be identified with $(L^{p',1}_{B_2^*})^*$ we then get

$$\|T\|_{L^p_{B_1}\to L^{p,\infty}_{B_2^{**}}} = \sup_{\|f_1\|_{L^p_{B_1}}\le 1}\ \sup_{\|f_2\|_{L^{p',1}_{B_2^*}}\le 1} |\langle Tf_1, f_2\rangle|$$

and obtain (2.16). □

Lemma 2.7. *Assume that B_2^{**} satisfies the Radon–Nikodym property. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in $\mathrm{Op}_{B_1.B_2}$, satisfying the strengthened support condition* (1.14)*. Suppose that $1\le p<q\le\infty$ and*

$$\sup_{j\in\mathbb{Z}}\|T_j\|_{\mathrm{Sp}_\gamma(p,B_1,q',B_2^*)} \le \mathcal{C}.$$

Then

$$\sup_{j\in\mathbb{Z}}\|\mathrm{Dil}_{2^j}T_j\|_{L^p_{B_1}\to L^q_{B_2^{**}}} \lesssim_{\gamma,d,\delta_1,\delta_2,p,q} \mathcal{C}.$$

Proof. We let $S=\mathrm{Dil}_{2^j}T_j$, $R_{\mathfrak{z}}$, $R_{\mathfrak{z},\nu}$, $\nu\in\mathcal{I}_{\mathfrak{z}}$ as in the proof of Lemma 2.4. The proof of (2.9) can be modified just with appropriate notational changes, such as replacing expressions as one the left -hand side of (2.9) with

$$\lambda_{f,\mathfrak{z},\nu}(g) := \langle Sf_{\mathfrak{z}}, g\mathbb{1}_{R_{\mathfrak{z},\nu}}\rangle.$$

This leads to the inequality

$$|\langle Sf_{\mathfrak{z}}, g\mathbb{1}_{R_{\mathfrak{z},\nu}}\rangle| \lesssim \mathcal{C}\|f_{\mathfrak{z}}\|_{L^p_{B_1}}\|g\|_{L^q_{B_2^*}} \tag{2.17}$$

in the place of (2.9). Inequality (2.17) shows that $\lambda_{f,\mathfrak{z},\nu}$ is a continuous linear functional on the space $L^{q'}_{B_2^*}(R_{\mathfrak{z},\nu})$; recall that by assumption $1\le q'<\infty$. By the Radon–Nikodym property of B_2^{**}, the linear functional $\lambda_{f,\mathfrak{z},\nu}$ is identified with a function $Sf_{\mathfrak{z}}$ restricted to $R_{\mathfrak{z},\nu}$, in the space $L^q_{B_2^{**}}(R_{\mathfrak{z},\nu})$. Hence we now get a variant of inequality (2.12), namely

$$\begin{aligned}\|Sf_{\mathfrak{z}}\|_{L^q_{B_2^{**}}} &\lesssim \sum_{\nu\in\mathcal{I}_{\mathfrak{z}}}\|Sf_{\mathfrak{z}}\|_{L^q_{B_2^{**}}(R_{\mathfrak{z},\nu})}\\ &\lesssim \sum_{\nu\in\mathcal{I}_{\mathfrak{z}}}\sup_{\|g\|_{L^{q'}_{B_2^*}}\le 1}|\langle Sf_{\mathfrak{z}}, g\mathbb{1}_{R_{\mathfrak{z},\nu}}\rangle \lesssim_{d,\delta_1,\delta_2} \mathcal{C}\|f_{\mathfrak{z}}\|_{L^p_{B_1}}.\end{aligned}$$

We finish as in (2.13) to bound $\|Sf\|_{L^q_{B_2^{**}}} \lesssim \mathcal{C}\|f_{L^p_{B_1}}$. □

Conclusion of the proof of Theorem 2.5. Since we are assuming that B_2 is reflexive we have that $B_2=B_2^{**}$ satisfies the Radon–Nikodym property. Hence now the necessity of the $L^{q,1}_{B_1}\to L^q_{B_2}$ and $L^p_{B_1}\to L^{p,\infty}_{B_2}$ conditions follow from Lemma 2.6, and the necessity of the single scale $L^p\to L^q$ conditions follows from using the assumption with $N_1=N_2$ and applying Lemma 2.7. □

3. Single scale sparse domination

We collect some preliminary results which are needed in the proof of Theorem 1.1.

3.1. A single scale estimate.

We state an elementary lemma which is used to establish the base case in the induction proof of Theorem 1.1. Recall that for a cube Q, we let $3Q$ denote the cube centered at the center of Q with three times the side length of Q, which is also the union of Q and its neighbors of the same side length.

Lemma 3.1. *Let $T_j \in \mathrm{Op}_{B_1,B_2}$ satisfy (1.6) and (1.8) for some exponents $p, q \in [1,\infty]$. Let Q be a cube of side length 2^j. Then for $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$,*

$$|\langle T_j[f_1 \mathbb{1}_Q], f_2\rangle| \le 3^{d/q'} A_\circ(p,q)|Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'}.$$

Proof. By the support property (1.6), $T_j[f_1\mathbb{1}_Q]$ is supported in $3Q$. By re-scaling we get from (1.8) that

$$\|T_j\|_{L^p\to L^q} \le 2^{-jd(1/p-1/q)} A_\circ(p,q)$$

and thus

$$\begin{aligned}|\langle T_j[f_1\mathbb{1}_Q], f_2\rangle| &= |\langle T_j[f_1\mathbb{1}_Q], f_2\mathbb{1}_{3Q}\rangle| \\ &\le A_\circ(p,q)2^{-jd(1/p-1/q)}\|f_1\mathbb{1}_Q\|_p\|f_2\mathbb{1}_{3Q}\|_{q'} \\ &= A_\circ(p,q)3^{d/q'}|Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'},\end{aligned}$$

as claimed. □

This implies a sparse bound for the single scale operators T_j; indeed the sparse collection is a disjoint collection of cubes.

Corollary 3.2. *For $0<\gamma\le 1$ and $1\le p,q\le\infty$,*

$$\|T_j\|_{\mathrm{S}_{\mathrm{p}_\gamma}(p,q')} \le 3^{d(1/p+1/q')} A_\circ(p,q).$$

Proof. We tile $\mathbb{R}^d$ by a family $\mathfrak{Q}_j$ of dyadic cubes of side length 2^j and estimate

$$\begin{aligned}|\langle T_j f_1, f_2\rangle| &\le \sum_{Q\in\mathfrak{Q}_j} |\langle T_j[f_1\mathbb{1}_Q], f_2\rangle| \le A_\circ(p,q)3^{d/q'}\sum_{Q\in\mathfrak{Q}_j}|Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'} \\ &\le A_\circ(p,q)3^{d(1/q'-1/p')}\sum_{Q\in\mathfrak{Q}_j}|3Q|\langle f_1\rangle_{3Q,p}\langle f_2\rangle_{3Q,q'}.\end{aligned}$$

The family $\{3Q : Q\in\mathfrak{Q}_j\}$ can be split into 3^d subfamilies consisting each of *disjoint* cubes of side length $3\cdot 2^j$. This implies the assertion (for every $0<\gamma\le 1$) since for every $3Q$ involved in each subfamily we can choose $E_{3Q} = 3Q$. □

3.2. **A resolution of the identity.** It will be quite convenient to work with a resolution of the identity using Littlewood–Paley decompositions which are localized in space. We have

$$\mathrm{I} = \sum_{k=0}^{\infty} \Lambda_k \widetilde{\Lambda}_k \tag{3.1}$$

which converges in the strong operator topology on $L^p_{B_1}(\mathbb{R}^d)$, $1 \le p < \infty$. Here Λ_k, $\widetilde{\Lambda}_k$ are convolution operators with convolution kernels λ_k, $\widetilde{\lambda}_k$ such that $\lambda_0 \in C_c^\infty$ has support in $\{x : |x| < 1/2\}$, $\int \lambda_0 = 1$, $\lambda_1 = 2\lambda_0(2\cdot) - \lambda_0$ and $\widetilde{\lambda}_0, \widetilde{\lambda}_1 \in \mathcal{S}$ with $\int \widetilde{\lambda}_0 = 1$, and $\int \widetilde{\lambda}_1 = 0$. Moreover, for $k \ge 1$, $\lambda_k = 2^{(k-1)d}\lambda_1(2^{k-1}\cdot)$, $\widetilde{\lambda}_k = 2^{(k-1)d}\widetilde{\lambda}_1(2^{k-1}\cdot)$. For later applicability we may choose λ_1, so that

$$\int \lambda_1(x) \prod_{i=1}^{d} x_i^{\alpha_i}\, dx = 0, \qquad \text{for } \sum_{i=1}^{d} \alpha_i \le 100d$$

and the same for $\widetilde{\lambda}_1$.

A proof of (3.1) with these specifications can be found in [100, Lemma 2.1] (the calculation there shows that $\widetilde{\lambda}_0$, $\widetilde{\lambda}_1$ can be chosen with compact support as well). For later use we let P_k be the operator given by convolution with $2^{kd}\lambda_0(2^k\cdot)$ for $k \ge 0$, and also set $P_{-1} = 0$, and observe that by our construction

$$\Lambda_k = P_k - P_{k-1}, \quad \text{for } k \ge 1. \tag{3.2}$$

3.3. **Single scale regularity.** In our proof of Theorem 1.1 it will be useful to work with other versions of the regularity conditions (1.9) which are adapted to the dyadic setting. To formulate these, we fix a dyadic lattice of cubes $\mathfrak{Q}$. 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 $\mathfrak{Q}$ of side length in $[2^{-n}, 2^{1-n})$, 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 $\mathbb{D}_n$ by

$$\mathbb{D}_n = \mathbb{E}_n - \mathbb{E}_{n-1} \quad \text{for } n \ge 1.$$

We also use the operators Λ_k, $\widetilde{\Lambda}_k$ in the decomposition (3.1).

Lemma 3.3. *Let $T \in \mathrm{Op}_{B_1,B_2}$.*

(i) Let $1 \le p \le q < \infty$, $0 < \vartheta < 1/p$. Then

$$\|T\mathbb{E}_0\|_{L^p_{B_1}\to L^q_{B_2}} + \sup_{n>0} 2^{n\vartheta}\|T\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim_\vartheta \sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}}. \tag{3.3}$$

(ii) Let $1 \le p \le q \le \infty$, $0 < \vartheta < 1$. Then

$$\sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \approx_\vartheta \|T\|_{L^p_{B_1}\to L^q_{B_2}} + \sup_{0<|h|<1} |h|^{-\vartheta}\|T\Delta_h\|_{L^p_{B_1}\to L^q_{B_2}}. \tag{3.4}$$

(iii) Let $1 < p \le q \le \infty$, $0 < \vartheta < 1 - 1/p$. Then

$$\sup_{k\ge 0} 2^{k\vartheta} \|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim_\vartheta \|T\mathbb{E}_0\|_{L^p_{B_1}\to L^q_{B_2}} + \sup_{n>0} 2^{n\vartheta}\|T\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}}. \quad (3.5)$$

An immediate consequence is the following.

Corollary 3.4. *For $1 < p \le q < \infty$ and $0 < \vartheta < \min\{1/p, 1 - 1/p\}$,*

$$\|T\mathbb{E}_0\|_{L^p_{B_1}\to L^q_{B_2}} + \sup_{n>0} 2^{n\vartheta}\|T\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}} \approx_\vartheta \|T\|_{L^p_{B_1}\to L^q_{B_2}} + \sup_{0<|h|<1} |h|^{-\vartheta}\|T\Delta_h\|_{L^p_{B_1}\to L^q_{B_2}}.$$

Proof of Lemma 3.3. We rely on arguments used before in considerations of variational estimates [58], [59], of basis properties of the Haar system in spaces measuring smoothness [47] and elsewhere. We use

$$\|\Lambda_k\|_{L^p_{B_1}\to L^p_{B_1}} = O(1), \quad \|\widetilde{\Lambda}_k\|_{L^p_{B_1}\to L^p_{B_1}} = O(1), \quad \|\mathbb{E}_n\|_{L^p_{B_1}\to L^p_{B_1}} = O(1)$$

throughout the proof. Since $\vartheta > 0$ we get from (3.1)

$$\|T\mathbb{E}_0\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim \sum_{k\ge 0} \|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim_\vartheta \sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}}.$$

To estimate $T\mathbb{D}_n$ we will need

$$\|\widetilde{\Lambda}_k\mathbb{D}_n\|_{L^p_{B_1}\to L^p_{B_1}} \lesssim \min\{1, 2^{(k-n)/p}\}, \quad (3.6)$$

and only the case $k \le n$ needs a proof. A standard calculation using cancellation of $\mathbb{D}_n$ yields (3.6) for $p = 1$ and the rest follows by interpolation. Consequently we can estimate

$$\begin{aligned}
&2^{n\vartheta}\|T\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}} \le 2^{n\vartheta}\sum_{k\ge 0}\|T\Lambda_k\widetilde{\Lambda}_k\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}}\\
&\le 2^{n\vartheta}\sum_{k\ge 0}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}}\|\widetilde{\Lambda}_k\mathbb{D}_n\|_{L^p_{B_1}\to L^p_{B_1}}\\
&\lesssim \sum_{0\le k\le n} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} 2^{-(n-k)(\frac{1}{p}-\vartheta)} + \sum_{k>n} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} 2^{-(k-n)\vartheta}\\
&\lesssim \sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}},
\end{aligned}$$

where we used $\vartheta < 1/p$ for the first sum. This proves (3.3).

We now turn to (3.4) and estimate the left-hand side. By (3.2) we can write

$$\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \le \|T(\mathrm{I} - P_k)\|_{L^p_{B_1}\to L^q_{B_2}} + \|T(\mathrm{I} - P_{k-1})\|_{L^p_{B_1}\to L^q_{B_2}}.$$

Note that, as $\int \lambda_0 = 1$,

$$(\mathrm{I} - P_{k-1})f(x) = \int 2^{(k-1)d}\lambda_0(2^{k-1}h)\Delta_{-h}f(x)dh,$$

so

$$\|T(\mathrm{I}-P_{k-1})f\|_{L^p_{B_1}\to L^q_{B_2}} = \int 2^{(k-1)d}|\lambda_0(2^{k-1}h)|\|T\Delta_{-h}\|_{L^p_{B_1}\to L^q_{B_2}}dh$$
$$\lesssim \sup_{|h|\le 2^{-k}} \|T\Delta_h\|_{L^p_{B_1}\to L^q_{B_2}}$$

and the same bound for $\|T(\mathrm{I}-P_k)f\|_{L^p_{B_1}\to L^q_{B_2}}$. This establishes that the left-hand side is smaller than the right-hand side in (3.4).

We argue similarly for the converse inequality. We estimate

$$\|T\|_{L^p_{B_1}\to L^q_{B_2}} \le \sum_{k\ge 0} \|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \lesssim \sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}}.$$

For the main terms

$$\|T\Delta_h\|_{L^p_{B_1}\to L^q_{B_2}} \le \sum_{k\ge 0} \|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \|\widetilde{\Lambda}_k\Delta_h\|_{L^p_{B_1}\to L^p_{B_1}}.$$

Now

$$\|\widetilde{\Lambda}_k\Delta_h\|_{L^p_{B_1}\to L^p_{B_1}} \lesssim \|\lambda_k(\cdot+h)-\lambda_k(\cdot)\|_1 \lesssim \min\{1, 2^k|h|\} \tag{3.7}$$

and therefore

$$|h|^{-\vartheta}\|T\Delta_h\|_{L^p_{B_1}\to L^q_{B_2}} \le \sum_{k\ge 0} \|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} 2^{k\vartheta}(2^k|h|)^{-\vartheta}\min\{1, 2^k|h|\}$$
$$\lesssim \sup_{k\ge 0} 2^{k\vartheta}\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}}$$

since $\sum_{k\ge 0}(2^k|h|)^{-\vartheta}\min\{1, 2^k|h|\} \lesssim_\vartheta 1$ if $0<\vartheta<1$. This completes the proof of (3.4).

It remains to prove (3.5). Setting $\mathbb{D}_0 := \mathbb{E}_0$ we observe that $\mathrm{I} = \sum_{n\ge 0}\mathbb{D}_n$ and $\mathbb{D}_n = \mathbb{D}_n\mathbb{D}_n$, and thus

$$\|T\Lambda_k\|_{L^p_{B_1}\to L^q_{B_2}} \le \sum_{n\ge 0} \|T\mathbb{D}_n\|_{L^p_{B_1}\to L^q_{B_2}} \|\mathbb{D}_n\Lambda_k\|_{L^p_{B_1}\to L^p_{B_1}}.$$

We use $\|\mathbb{D}_n\Lambda_k\|_{L^p_{B_1}\to L^p_{B_1}} \lesssim 1$ for $n\ge k$ and

$$\|\mathbb{D}_n\Lambda_k\|_{L^p_{B_1}\to L^p_{B_1}} \lesssim 2^{(n-k)(1-\frac{1}{p})} \text{ for } n<k. \tag{3.8}$$

This is clear for $p=1$ and by interpolation it suffices to show it for $p=\infty$. Let Q be a dyadic cube of side length 2^{-n+1}. Let $\mathrm{Ch}(Q)$ be the set of 2^d dyadic children of Q (i.e. the dyadic sub-cubes of side length 2^{-n}). Let

$$F_{Q,k} = \{x : \mathrm{dist}(x, \partial\tilde{Q}) \le 2^{-k} \text{ for some } \tilde{Q}\in \mathrm{Ch}(Q)\}.$$

Then $|F_{Q,k-1}| \lesssim 2^{-n(d-1)}2^{-k}$. Let $g_{Q,k} = f\mathbb{1}_{Q\setminus F_{Q,k}}$ and observe that by Fubini's theorem and the cancellation and support properties of λ_k

$$\mathbb{E}_n\Lambda_k g_{Q,k}(x) = 0 \quad \text{and} \quad \mathbb{E}_{n-1}\Lambda_k g_{Q,k}(x) = 0 \quad \text{for } x\in Q.$$

Hence for $x \in Q$,

$$\begin{aligned}|\mathbb{D}_n \Lambda_k f(x)| &= |\mathbb{D}_n \Lambda_k (f \mathbb{1}_{F_{Q,k}})(x)| \\ &\lesssim 2^{nd} \int\int |\lambda_k(w-y)||f(y)| dw \mathbb{1}_{F_{Q,k}}(y) dy \\ &\lesssim 2^{nd} |F_{Q,k-1}| \|f\|_\infty \lesssim 2^{n-k} \|f\|_\infty.\end{aligned}$$

This implies (3.8) for $p = \infty$.

To finish we write

$$\begin{aligned}2^{k\vartheta} \|T\Lambda_k\|_{L^p_{B_1} \to L^q_{B_2}} &\leq \sum_{n \geq 0} \|T\mathbb{D}_n\|_{L^p_{B_1} \to L^q_{B_2}} 2^{k\vartheta} \min\{1, 2^{(n-k)(1-1/p)}\} \\ &\lesssim \sup_{n \geq 0} 2^{n\vartheta} \|T\mathbb{D}_n\|_{L^p_{B_1} \to L^q_{B_2}}\end{aligned}$$

where we used $\sum_{n \geq 0} 2^{(k-n)\vartheta} \min\{1, 2^{(n-k)(1-1/p)}\} \lesssim 1$ provided that $0 < \vartheta < 1 - 1/p$. This proves (3.5). □

In the proof of Theorem 1.1 we use the following Corollary.

Corollary 3.5. *Let $1 \leq p \leq q < \infty$ and $0 < \vartheta < 1/p$. Then*
(i) For any $n \geq 0$,

$$\|T(\mathrm{I} - \mathbb{E}_n)\|_{L^p_{B_1} \to L^q_{B_2}} \lesssim_\vartheta 2^{-n\vartheta} \sup_{0<|h|<1} |h|^{-\vartheta} \|T\Delta_h\|_{L^p_{B_1} \to L^q_{B_2}}.$$

(ii) If T_j is such that (1.9a) *holds then*

$$\|T_j(\mathrm{I} - \mathbb{E}_{n-j})\|_{L^p_{B_1} \to L^q_{B_2}} \lesssim_\vartheta B 2^{-n\vartheta} 2^{-jd(\frac{1}{p} - \frac{1}{q})}.$$

Proof. We write $\mathrm{I} = \mathbb{E}_n + \sum_{k=1}^\infty \mathbb{D}_{n+k}$, and thus

$$\begin{aligned}\|T(\mathrm{I} - \mathbb{E}_n)\|_{L^p_{B_1} \to L^q_{B_2}} &\leq \sum_{k=1}^\infty \|T\mathbb{D}_{n+k}\|_{L^p_{B_1} \to L^q_{B_2}} \\ &\lesssim_\vartheta \sum_{k=1}^\infty 2^{-(n+k)\vartheta} \sup_{0<|h|<1} |h|^{-\vartheta} \|T\Delta_h\|_{L^p_{B_1} \to L^q_{B_2}}\end{aligned}$$

by combining part (3.3), (3.4) in the statement of Lemma 3.3. We sum and get the assertion. Part (ii) follows by rescaling and the hypotheses. □

We finally discuss a formulation of the regularity condition which involves the Fourier support of the function and is therefore limited to the case where B_1 is a separable Hilbert space, here denoted by $\mathcal{H}$. It is convenient to use a frequency decomposition

$$f = \sum_{\ell \geq 0} \eta_\ell * f, \tag{3.9}$$

with $\widehat{\eta}_0$ is supported in $\{\xi : |\xi| < 1\}$ such that $\widehat{\eta}_0(\xi) = 1$ for $|\xi| \le \frac{3}{4}$, and with η_ℓ defined by $\widehat{\eta}_\ell(\xi) = \widehat{\eta}_0(2^{-\ell}\xi) - \widehat{\eta}_0(2^{1-\ell}\xi)$ for $\ell \ge 1$, *i.e.* we have

$$\operatorname{supp}(\widehat{\eta}_\ell) \subset \{\xi : 2^{\ell-2} < |\xi| < 2^\ell\}, \quad \ell \ge 1. \tag{3.10}$$

Recall that $\mathcal{E}_{\mathrm{ann}}(\lambda)$ denotes the space of tempered distributions whose Fourier transform is supported in $\{\xi : \lambda/2 < |\xi| < 2\lambda\}$.

Lemma 3.6. *Let $\mathcal{H}$ be a separable Hilbert space and $T \in \mathrm{Op}_{\mathcal{H},B_2}$. Suppose that $T : L^p_{\mathcal{H}} \to L^q_{B_2}$ satisfies*

$$\|T\|_{L^p_{\mathcal{H}} \to L^q_{B_2}} \le A, \tag{3.11}$$

and for all $\lambda > 2$ and all $\mathcal{H}$-valued Schwartz functions $f \in \mathcal{E}_{\mathrm{ann}}(\lambda)$,

$$\|Tf\|_{L^q_{B_2}} \le A\lambda^{-\vartheta}\|f\|_{L^p_{\mathcal{H}}}. \tag{3.12}$$

Then

$$\sup_{0<|h|<1} |h|^{-\vartheta}\|T\Delta_h\|_{L^p_{\mathcal{H}} \to L^q_{B_2}} \lesssim_\vartheta A.$$

Proof. By the assumptions (3.11) ($\ell = 0$) and (3.12) ($\ell \ge 1$) we have

$$\|T[\eta_\ell * \Delta_h f]\|_{L^q_{B_2}} \le A 2^{(1-\ell)\vartheta}\|\eta_\ell * \Delta_h f\|_{L^p_{\mathcal{H}}}$$

Arguing as in (3.7) we get

$$\|\eta_\ell * \Delta_h f\|_{L^p_{\mathcal{H}}} = \|\Delta_h \eta_\ell * f\|_{L^p_{\mathcal{H}}} \lesssim \min\{1, 2^\ell|h|\}\|f\|_{L^p_{\mathcal{H}}}.$$

Thus using (3.9) we obtain

$$\|T\Delta_h f\|_{L^q_{B_2}} \le \sum_{\ell=0}^{\infty} \|T[\eta_\ell * \Delta_h f]\|_{L^q_{B_2}} \lesssim A \sum_{\ell=0}^{\infty} 2^{-\ell\vartheta} \min\{1, 2^\ell|h|\}\|f\|_{L^p_{\mathcal{H}}}$$

and after summing in ℓ we arrive at $\|T\Delta_h f\|_{L^q_{B_2}} \lesssim_\vartheta |h|^\vartheta \|f\|_{L^p_{\mathcal{H}}}$. □

4. Proof of the main result

4.1. **A modified version of sparse forms.** We fix a dyadic lattice $\mathfrak{Q}$ in the sense of Lerner and Nazarov, where we assume that the side length of each cube in $\mathfrak{Q}$ is dyadic, *i.e.* of the form 2^k with $k \in \mathbb{Z}$. Also fix $\gamma \in (0,1)$ and $1 < p \le q < \infty$. It will be convenient to use variants $\mathfrak{G}_{Q_0} \equiv \mathfrak{G}_{Q_0,\gamma}$ of the maximal form $\Lambda^*_{p,q'}$ defined in (1.3). The presence of the triple cubes in the new form allows one to exploit more effectively the support condition (1.6).

Definition 4.1. *Given a cube $Q_0 \in \mathfrak{Q}$ let*

$$\mathfrak{G}_{Q_0}(f_1, f_2) = \sup \sum_{Q \in \mathfrak{S}} |Q| \langle f_1 \rangle_{Q,p,B_1} \langle f_2 \rangle_{3Q,q',B_2^*}$$

where the supremum is taken over all γ-sparse collections $\mathfrak{S}$ consisting of cubes in $\mathcal{D}(Q_0)$.

Notational convention. From now on in this proof, the dependence on the Banach spaces B_1, B_2^* will not be explicitly indicated, i.e. $\langle f_1\rangle_{Q,p}$ should be understood as $\langle f_1\rangle_{Q,p,B_1}$ and $\langle f_2\rangle_{Q,q'}$ should be understood as $\langle f_2\rangle_{Q,q',B_2^*}$.

The key step towards proving Theorem 1.1 is to establish a variant in which $\Lambda^*_{p,q'}$ is replaced by $\mathfrak{G}_{Q_0}$, that is,

$$\big|\big\langle \sum_{j=N_1}^{N_2} T_j f_1, f_2\big\rangle\big| \lesssim_{p,q,\varepsilon,d,\gamma} \mathcal{C}\, \mathfrak{G}_{Q_0}(f_1,f_2) \tag{4.1}$$

for $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$ and a sufficiently large cube $Q_0 \in \mathfrak{Q}$. The reader will notice that $\mathfrak{G}_{Q_0}$ does not define a sparse form, and we will show in §4.2 how to finish the proof of Theorem 1.1 given (4.1). The proof of (4.1) will be done by induction, which leads us to the following definition.

Definition 4.2. *For $n = 0, 1, 2, \dots$ let $U(n)$ be the smallest constant U so that for all families of operators $\{T_j\}$ satisfying the assumptions of Theorem 1.1, for all pairs (N_1, N_2) with $0 \le N_2 - N_1 \le n$ and for all dyadic cubes $Q_0 \in \mathfrak{Q}$ of side length 2^{N_2} we have*

$$\big|\big\langle \sum_{j=N_1}^{N_2} T_j f_1, f_2\big\rangle\big| \le U \mathfrak{G}_{Q_0}(f_1,f_2)$$

whenever $f_1 \in \mathrm{S}_{B_1}$ with $\operatorname{supp}(f_1) \subset Q_0$ and $f_2 \in \mathrm{S}_{B_2^*}$.

Thus, in order to show (4.1), it suffices to show that

$$U(n) \lesssim_{p,q,\varepsilon,d,\gamma} \mathcal{C}$$

uniformly in $n \in \mathbb{N}_0$. This will be proven by induction on n. By Lemma 3.1 we have the base case

$$U(0) \le 3^{d/q'} A_\circ(p,q) \tag{4.2}$$

and, more generally, $U(n) \le (n+1)3^{d/q'} A_\circ(p,q)$, which shows the finiteness of the $U(n)$. The proof then reduces to the verification of the following inductive claim.

Claim 4.3. *There is a constant $c = c_{p,q,\varepsilon,d,\gamma}$ such that for all $n > 0$,*

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

with $\mathcal{C}$ defined as in (1.11).

Our proof of the claim is an extension of the proof for sparse bounds of the prototypical singular Radon transforms in [89], which itself builds on ideas in [66]. It is contained in §4.4.

4.2. **Proof of Theorem 1.1 given Claim 4.3.** Fix $N_1 \le N_2$, $f_1 \in \mathrm{S}_{B_1}$, $f_2 \in \mathrm{S}_{B_2^*}$. We choose any dyadic lattice with cubes of dyadic side length as in the previous subsection. By (1.6) we may choose a cube $Q_0 \in \mathfrak{Q}$ of side length $2^{L(Q_0)}$ with $L(Q_0) \ge N_2$ such that f_1 is supported in Q_0. Then $\sum_{j=N_1}^{N_2} T_j f_1$ is supported in $3Q_0$. Define the operators $S_j = T_j$ when $N_1 \le j \le N_2$ and $S_j = 0$ otherwise. Then the assumptions of Theorem 1.1 apply to the family $\{S_j\}$. By (4.2) and Claim 4.3 applied to $S = \sum_{j=N_1}^{L(Q_0)} S_j = \sum_{j=N_1}^{N_2} T_j$ we obtain

$$|\langle Sf_1, f_2\rangle| \le c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C}\, \mathfrak{G}_{Q_0}(f_1, f_2).$$

In order to complete the proof of Theorem 1.1 it remains to replace $\mathfrak{G}_{Q_0}$ by the maximal sparse form $\Lambda^*_{p,q}$. This argument relies on facts in dyadic analysis which we quote from the book by Lerner and Nazarov [79].

We first note that for $\epsilon > 0$ there is a γ-sparse collection $\mathfrak{S}_\epsilon \subset \mathcal{D}(Q_0)$ such that

$$\begin{aligned}\big|\langle Sf_1, f_2\rangle\big| &\le (c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C} + \epsilon) \sum_{Q\in\mathfrak{S}_\epsilon} |Q| \langle f_1\rangle_{Q,p} \langle f_2\rangle_{3Q,q'} \\ &\le 3^{d/p-d}(c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C} + \epsilon) \sum_{Q\in\mathfrak{S}_\epsilon} |3Q| \langle f_1\rangle_{3Q,p} \langle f_2\rangle_{3Q,q'}. \qquad (4.3)\end{aligned}$$

By the Three Lattice Theorem [79, Theorem 3.1] there are dyadic lattices $\mathcal{D}^{(\nu)}$, $\nu = 1, \dots, 3^d$, such that every cube in the collection $3\mathfrak{S}_\epsilon := \{3Q : Q \in \mathfrak{S}_\epsilon\}$ belongs to one of the dyadic lattices $\mathcal{D}^{(\nu)}$. Moreover, each collection

$$\mathfrak{S}_\epsilon^{(\nu)} = 3\mathfrak{S}_\epsilon \cap \mathcal{D}^{(\nu)}$$

is a $3^{-d}\gamma$-sparse collection of cubes in $\mathcal{D}^{(\nu)}$. Each $\mathfrak{S}_\epsilon^{(\nu)}$ is a $3^d\gamma^{-1}$-Carleson family in the sense of [79, Definition 6.2]. By [79, Lemma 6.6] we can write, for each integer $M \ge 2$, the family $\mathfrak{S}_\epsilon^{(\nu)}$ as a union of M sub-families $\mathfrak{S}_{\epsilon,i}^{(\nu)}$, each of which is a $\widetilde{M}$-Carleson family, with $\widetilde{M} = 1 + M^{-1}(3^d\gamma^{-1} - 1)$. By [79, Lemma 6.3] the collections $\mathfrak{S}_{\epsilon,i}^{(\nu)}$ are $\tilde\gamma$-sparse families where $\tilde\gamma = \widetilde{M}^{-1} = (1 + M^{-1}(3^d\gamma^{-1} - 1))^{-1}$. By choosing M large enough we can have $\tilde\gamma > \gamma$ and then, from (4.3), one has

$$\begin{aligned}|\langle Sf_1, f_2\rangle| &\le 3^{d/p-d}(c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C} + \epsilon) \sum_{Q\in\mathfrak{S}_\epsilon} |3Q| \langle f_1\rangle_{3Q,p} \langle f_2\rangle_{3Q,q'} \\ &\le M 3^{d/p}(c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C} + \epsilon) \sup_{\substack{i=1,\dots,M\\ \nu=1,\dots 3^d}} \sum_{R\in\mathfrak{S}_{\epsilon,i}^{(\nu)}} |R| \langle f_1\rangle_{R,p} \langle f_2\rangle_{R,q'} \\ &\le M 3^{d/p}(c_{p,q,\varepsilon,d,\gamma}\mathcal{C} + \epsilon)\, \Lambda^*_{p,q}(f_1, f_2)\end{aligned}$$

which gives the desired γ-sparse bound with $\|S\|_{\mathrm{Sp}_\gamma(p,q')} \le M 3^{d/p} c_{p,q,\varepsilon,d,\gamma}\, \mathcal{C}$.

4.3. **Proof of Corollary 1.2.** Corollary 1.2 is a consequence of Theorem 1.1 and the following lemma, applied to $T = \sum_{j=N_1}^{N_2} T_j$.

Lemma 4.4. *Let $T : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$ and assume that*

$$\|T\|_{\mathrm{Sp}_\gamma(p,B_1;q',B_2^*)} \le \mathcal{C}.$$

Then we have for all $f \in \mathrm{S}_{B_1}$ and all nonnegative simple ω

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2}\omega(x)dx \le \mathcal{C}\,\Lambda^*_{p,B_1,q',\mathbb{R}}(f,\omega). \tag{4.4}$$

Proof. By the monotone convergence theorem we may assume that ω is a compactly supported simple function. Moreover, since $T : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$ we can approximate, in the $L^1_{B_2}(K)$ norm for every compact set K, the function Tf (for $f \in \mathrm{S}_{B_1}$) by simple B_2-valued functions. Thus given $\epsilon > 0$ there is $h \in \mathrm{S}_{B_2}$ such that

$$\int_{\mathbb{R}^d} |Tf(x) - h(x)|_{B_2}\,\omega(x)dx \le \epsilon.$$

Moreover, there is a compactly supported $\lambda \in \mathrm{S}_{B_2^*}$ with $\max_{x\in\mathbb{R}^d} |\lambda(x)|_{B_2^*} \le 1$ (depending on h, ω) such that

$$\int_{\mathbb{R}^d} |h(x)|_{B_2}\omega(x)dx \le \epsilon + \int_{\mathbb{R}^d} \langle h(x), \lambda(x)\rangle\,\omega(x)dx,$$

and we also have

$$\Big|\int_{\mathbb{R}^d} \langle h(x) - Tf(x), \lambda(x)\rangle\,\omega(x)dx\Big| \le \int_{\mathbb{R}^d} |h(x) - Tf(x)|_{B_2}\,\omega(x)dx \le \epsilon.$$

Consequently

$$\int_{\mathbb{R}^d} |Tf(x)|_{B_2}\omega(x)dx \le 3\epsilon + \int_{\mathbb{R}^d} \big\langle Tf(x), \lambda(x)\big\rangle\,\omega(x)dx.$$

Thus in order to show (4.4) it suffices to show

$$\int_{\mathbb{R}^d} \big\langle Tf(x), \lambda(x)\big\rangle\omega(x)dx \le \mathcal{C}\Lambda^*_{p_1,B_1,p_2,\mathbb{R}}(f,\omega) \tag{4.5}$$

for any choice of compactly supported $\lambda \in \mathrm{S}_{B_2^*}$ such that $\|\lambda\|_{L^\infty_{B_2^*}} \le 1$. Let $f_2(x) = \omega(x)\lambda(x)$. Then $f_2 \in \mathrm{S}_{B_2^*}$ with $|f_2(x)|_{B_2^*} \le \omega(x)$ for all $x \in \mathbb{R}^d$. By the hypothesis, applied to f and $f_2 = \omega\lambda$,

$$\int_{\mathbb{R}^d} \big\langle Tf(x), \omega(x)\lambda(x)\big\rangle dx \le \mathcal{C}\,\Lambda^*_{p_1,B_1,p_2,B_2^*}(f,\omega\lambda).$$

Since $\langle\omega\lambda\rangle_{Q,p_2,B_2^*} \le \langle\omega\rangle_{Q,p_2,\mathbb{R}}$, we have established (4.5), and the proof is finished by letting $\epsilon \to 0$. □

4.4. **The inductive step.** In this section we prove Claim 4.3, the key ingredient in the proof of Theorem 1.1. Let Q_0 be a dyadic cube of side length 2^{N_2}. Recall that f_1 is supported in Q_0 and thus

$$\Big\langle \sum_{j=N_1}^{N_2} T_j f_1, f_2 \Big\rangle = \Big\langle \sum_{j=N_1}^{N_2} T_j f_1, f_2 \mathbb{1}_{3Q_0} \Big\rangle.$$

Hence without loss of generality we may assume that f_2 is supported in $3Q_0$.

Let $\mathcal{M}$ denote the Hardy–Littlewood maximal operator and let $\mathcal{M}_p f = (\mathcal{M}|f|^p)^{1/p}$. By the well known weak $(1,1)$ inequality for $\mathcal{M}$,

$$\operatorname{meas}(\{x \in \mathbb{R}^d : \mathcal{M}_p f > \lambda\}) \le 5^d \lambda^{-p} \|f\|_p^p.$$

Define $\Omega = \Omega_1 \cup \Omega_2$ where

$$\begin{aligned} \Omega_1 &= \{x \in 3Q_0 : \mathcal{M}_p f_1(x) > \big(\tfrac{100^d}{1-\gamma}\big)^{1/p} \langle f_1 \rangle_{Q_0,p}\}, \\ \Omega_2 &= \{x \in 3Q_0 : \mathcal{M}_{q'} f_2(x) > \big(\tfrac{100^d}{1-\gamma}\big)^{1/q'} \langle f_2 \rangle_{3Q_0,q'}\}. \end{aligned} \tag{4.6}$$

We then have $|\Omega| \le |\Omega_1| + |\Omega_2| < (1-\gamma)|Q_0|$ and if we set

$$E_{Q_0} = Q_0 \setminus \Omega,$$

then $|E_{Q_0}| > \gamma |Q_0|$.

We perform a Whitney decomposition of Ω. It is shown in [103, VI.1.2] that given any $\beta > \sqrt{d}$, one can write Ω as a union of disjoint dyadic Whitney cubes $W \in \mathcal{W}^\beta \subset \mathfrak{Q}$, with side length $2^{L(W)}$ and $L(W) \in \mathbb{Z}$, so that

$$(\beta - \sqrt{d}) 2^{L(W)} \le \operatorname{dist}(W, \Omega^{\complement}) \le \beta 2^{L(W)+1}$$

for $W \in \mathcal{W}^\beta$. In [103, VI.1.1] the choice of $\beta = 2\sqrt{d}$ is made; here we need to choose β sufficiently large and $\beta = 6\sqrt{d}$ will work for us. We fix this choice and label as $\mathcal{W}$ the corresponding family of Whitney cubes. We then have

$$5 \operatorname{diam}(W) \le \operatorname{dist}(W, \Omega^{\complement}) \le 12 \operatorname{diam}(W) \quad \text{for all } W \in \mathcal{W}. \tag{4.7}$$

We set for $i = 1, 2$,

$$\begin{aligned} f_{i,W} &= f_i \mathbb{1}_W, \\ b_{i,W} &= (f_i - \operatorname{av}_W f_i) \mathbb{1}_W = (\mathrm{I} - \mathbb{E}_{-L(W)}) f_{i,W}, \end{aligned}$$

and

$$\begin{aligned} b_i &= \sum_{W \in \mathcal{W}} b_{i,W}, \\ g_i &= f_i \mathbb{1}_{\Omega^{\complement}} + \sum_{W \in \mathcal{W}} \operatorname{av}_W f_i \, \mathbb{1}_W. \end{aligned}$$

Then we have the Calderón–Zygmund decompositions $f_i = g_i + b_i$ (using the same above family of Whitney cubes for f_1 and f_2). For $i = 1$ we add

an observation, namely that

$$b_1 = \sum_{\substack{W\in\mathcal{W}\\ W\subset Q_0}} b_{1,W}.$$

Since f_1 is supported in Q_0, this follows from the fact that

$$W\cap Q_0\neq\emptyset \quad\Longrightarrow\quad W\subsetneq Q_0. \tag{4.8}$$

Indeed, if (4.8) fails, we must have $Q_0\subseteq W$ as W and Q_0 are dyadic. But then $|Q_0|\le|W|\le|\Omega|<(1-\gamma)|Q_0|$, which is a contradiction.

We note from (4.7) and the definition of Ω that

$$\langle f_1\rangle_{W,p}\lesssim_{d,\gamma}\langle f_1\rangle_{Q_0,p},\qquad \langle f_2\rangle_{W,q'}\lesssim_{d,\gamma}\langle f_2\rangle_{3Q_0,q'} \tag{4.9a}$$

for every $W\in\mathcal{W}$, as a fixed dilate of W intersects $\Omega^{\complement}$. Indeed,

$$\langle f_1\rangle_{Q,p}\lesssim_{d,\gamma}\langle f_1\rangle_{Q_0,p},\qquad \langle f_2\rangle_{Q,q'}\lesssim_{d,\gamma}\langle f_2\rangle_{3Q_0,q'} \tag{4.9b}$$

for every cube Q which contains a $W\in\mathcal{W}$. Moreover, by the definition of g_i and Ω,

$$\|g_1\|_{L^\infty_{B_1}}\lesssim_{d,\gamma}\langle f_1\rangle_{Q_0,p},\qquad \|g_2\|_{L^\infty_{B_2^*}}\lesssim_{d,\gamma}\langle f_2\rangle_{3Q_0,q'}. \tag{4.10}$$

Since $\operatorname{supp}(f_1)\subset Q_0$ and $\operatorname{supp}(f_2)\subset 3Q_0$ we also get $\operatorname{supp}(g_1)\subset Q_0$ and $\operatorname{supp}(g_2)\subset 3Q_0$; here we use (4.8). Since $\|\mathbb{1}_Q\|_{L^{r,1}}\lesssim_r|Q|^{1/r}$ for $r<\infty$, we obtain from (4.10) that for $r_1,r_2<\infty$,

$$\|g_1\|_{L^{r_1,1}_{B_1}}\lesssim_{d,r_1,\gamma}|Q_0|^{1/r_1}\langle f_1\rangle_{Q_0,p},\qquad \|g_2\|_{L^{r_2,1}_{B_2^*}}\lesssim_{d,r_2,\gamma}|Q_0|^{1/r_2}\langle f_2\rangle_{3Q_0,q'}. \tag{4.11}$$

For every dyadic cube $Q\in\mathfrak{Q}$ we have by disjointness of the W

$$\Big\|\sum_{W\subset Q} b_{1,W}\Big\|_{L^r_{B_1}}\lesssim\Big(\sum_{W\subset Q}\|f_{1,W}\|^r_{L^r_{B_1}}\Big)^{1/r}$$

and thus

$$\Big\|\sum_{W\subset Q} b_{1,W}\Big\|_{L^r_{B_1}}\lesssim\Big(\int_Q|f_1(x)|^r_{B_1}dx\Big)^{1/r}. \tag{4.12}$$

Likewise we get for f_2,

$$\Big\|\sum_{W\subset Q} b_{2,W}\Big\|_{L^r_{B_2^*}}\lesssim\Big(\int_Q|f_2(x)|^r_{B_2^*}dx\Big)^{1/r}.$$

We now begin the proof of the induction step in Claim 4.3. Let

$$S\equiv S_{N_1,N_2}=\sum_{j=N_1}^{N_2}T_j.$$

By the Calderón–Zygmund decomposition for f_1 we have

$$|\langle Sf_1,f_2\rangle|\le|\langle Sg_1,f_2\rangle|+|\langle Sb_1,f_2\rangle|. \tag{4.13}$$

Using the $L^{q,1}_{B_1} \to L^q_{B_2}$ boundedness of S (from the restricted strong type (q,q) condition (1.7b)) and (4.11) with $r_1 = q < \infty$ we get

$$\begin{aligned} |\langle Sg_1, f_2\rangle| &\le \|Sg_1\|_{L^q_{B_2}} \|f_2 \mathbb{1}_{3Q_0}\|_{L^{q'}_{B_2^*}} \\ &\le A(q)\|g_1\|_{L^{q,1}_{B_1}} \|f_2 \mathbb{1}_{3Q_0}\|_{L^{q'}_{B_2^*}} \\ &\lesssim_{d,q,\gamma} A(q)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}. \end{aligned} \tag{4.14}$$

Define, for each $W \in \mathcal{W}$ (recalling that the side length of W is $2^{L(W)}$),

$$S_W f = S_{N_1,L(W)}[f\mathbb{1}_W] \equiv \sum_{N_1 \le j \le L(W)} T_j[f\mathbb{1}_W].$$

We decompose the second term in (4.13) as in [89] and write

$$\langle Sb_1, f_2\rangle = I + II + III,$$

where

$$I = \Big\langle \sum_{W\in\mathcal{W}} S_W f_1, f_2 \Big\rangle, \tag{4.15a}$$

$$II = -\Big\langle \sum_{W\in\mathcal{W}} S_W(\mathrm{av}_W[f_1]\mathbb{1}_W), f_2 \Big\rangle, \tag{4.15b}$$

$$III = \Big\langle \sum_{W\in\mathcal{W}} (S - S_W) b_{1,W}, f_2 \Big\rangle. \tag{4.15c}$$

The first term (4.15a) is handled by the induction hypothesis. In view of (4.8), each W that contributes a non-zero summand in (4.15a) is a proper subcube of Q_0. Therefore we have $L(W) - N_1 \le n-1$ and thus by the induction hypothesis,

$$|\langle S_W f_1, f_2\rangle| \le U(n-1)\mathfrak{G}_W(f_1\mathbb{1}_W, f_2).$$

That is, given any $\epsilon > 0$ there is a γ-sparse collection $\mathfrak{S}_{W,\epsilon}$ of subcubes of W such that

$$|\langle S_W f_1, f_2\rangle| \le (U(n-1)+\epsilon) \sum_{Q\in\mathfrak{S}_{W,\epsilon}} |Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'}. \tag{4.16}$$

Because of the γ-sparsity there are measurable subsets E_Q of Q with $|E_Q| \ge \gamma|Q|$ so that the E_Q with $Q \in \mathfrak{S}_{W,\epsilon}$ are disjoint. We combine the various collections $\mathfrak{S}_{W,\epsilon}$ and form the collection $\mathfrak{S}_\epsilon$ of cubes

$$\mathfrak{S}_\epsilon := \{Q_0\} \cup \bigcup_{\substack{W\in\mathcal{W}\\ W\subset Q_0}} \mathfrak{S}_{W,\epsilon}.$$

Observe that the collection $\mathfrak{S}_\epsilon$ is indeed γ-sparse: as defined above, $E_{Q_0} = Q_0 \setminus \Omega$, and therefore $|E_{Q_0}| > \gamma|Q_0|$. By disjointness of the $W \subset Q_0$ the sets E_Q, for $Q \in \mathfrak{S}_\epsilon$ are disjoint; moreover they satisfy $|E_Q| \ge \gamma|Q|$.

We consider the term II in (4.15b). Here we will use that the restricted strong type (q,q) condition (1.7b) implies $\|S_W\|_{L^{q,1}_{B_1}\to L^q_{B_2}} \le A(q)$, and

$$\|\mathrm{av}_W[f_1]\mathbb{1}_W\|_{L^{q,1}_{B_1}} \lesssim |\mathrm{av}_W[f_1]|_{B_1}|W|^{1/q} \lesssim_q \langle f_1\rangle_{W,p}|W|^{1/q}$$

for $q<\infty$. Together with the disjointness of the cubes W and (4.9a), we get

$$\begin{aligned} |II| &\le \sum_{W\in\mathcal{W}} \|S_W(\mathrm{av}_W[f_1]\mathbb{1}_W)\|_{L^q_{B_2}}\|f_2\mathbb{1}_{3W}\|_{L^{q'}_{B_2^*}} \\ &\lesssim_q \sum_{W\in\mathcal{W}} \langle f_1\rangle_{W,p}A(q)|W|^{1/q}(3^d|W|)^{1/q'}\langle f_2\rangle_{3W,q'} \\ &\lesssim_{d,q,\gamma} A(q)\sum_{W\in\mathcal{W}}|W|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'} \\ &\lesssim_{d,q,\gamma} A(q)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\,. \end{aligned} \tag{4.17}$$

Regarding the third term in (4.15c) we claim that

$$|III| \lesssim_{p,q,\varepsilon,d,\gamma} (A(p) + A_\circ(p,q)\log(2+\tfrac{B}{A_\circ(p,q)}))|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\,. \tag{4.18}$$

Taking (4.18) for granted we obtain from (4.14), (4.17), (4.18) and (4.16) that there exist constants $C_1(d,q,\gamma)$ and $C_2(p,q,\varepsilon,d,\gamma)$ such that

$$\begin{aligned} |\langle Sf_1,f_2\rangle| &\le C_1(d,q,\gamma)A(q)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'} \\ &\quad + C_2(p,q,\varepsilon,d,\gamma)(A(p)+A_\circ(p,q)\log(2+\tfrac{B}{A_\circ(p,q)})|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'} \\ &\quad + \sum_{\substack{W\in\mathcal{W}\\ W\subset Q_0}}\sum_{Q\in\mathfrak{S}_{W,\epsilon}} (U(n-1)+\epsilon)|Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'}. \end{aligned}$$

This implies

$$\begin{aligned} |\langle Sf_1,f_2\rangle| &\le \max\{U(n-1)+\epsilon, c_{p,q,\varepsilon,d,\gamma}\,\mathcal{C}\}\sum_{Q\in\mathfrak{S}_\epsilon}|Q|\langle f_1\rangle_{Q,p}\langle f_2\rangle_{3Q,q'} \\ &\le \max\{U(n-1)+\epsilon, c_{p,q,\varepsilon,d,\gamma}\mathcal{C}\}\,\mathfrak{G}_{Q_0}(f_1,f_2) \end{aligned}$$

for all $\epsilon>0$. Letting $\epsilon\to 0$ implies Claim 4.3. We are now coming to the most technical part of the proof, the estimation of the error term III in (4.15c) for which we have to establish the claim (4.18).

Proof of (4.18). We now use the Calderón–Zygmund decomposition for $f_2 = g_2 + \sum_{W\in\mathcal{W}} b_{2,W}$ as described above. We split $III = \sum_{i=1}^4 III_i$ where

$$III_1 = \Big\langle \sum_{W\in\mathcal{W}} Sb_{1,W}, g_2\Big\rangle, \tag{4.19a}$$

$$III_2 = -\sum_{W\in\mathcal{W}} \langle S_W b_{1,W}, g_2\rangle, \tag{4.19b}$$

$$III_3 = \sum_{N_1\le j\le N_2} \sum_{\substack{W\in\mathcal{W}:\\ L(W)<j}} \sum_{\substack{W'\in\mathcal{W}:\\ L(W')\ge j}} \langle T_j b_{1,W}, b_{2,W'}\rangle, \tag{4.19c}$$

$$III_4 = \sum_{N_1\le j\le N_2} \sum_{\substack{W\in\mathcal{W}:\\ L(W)<j}} \sum_{\substack{W'\in\mathcal{W}:\\ L(W')<j}} \langle T_j b_{1,W}, b_{2,W'}\rangle. \tag{4.19d}$$

We use the weak type (p,p) condition (1.7a) that S maps $L^p_{B_1}$ to $L^{p,\infty}_{B_2}$, which is isometrically embedded in $L^{p,\infty}_{B_2^{**}}$. As $p>1$, we obtain using (4.11) for $r_2 = p' < \infty$

$$\begin{aligned}|III_1| &\le \Big\|S[\sum_{W\in\mathcal{W}} b_{1,W}]\Big\|_{L^{p,\infty}_{B_2^{**}}} \|g_2\|_{L^{p',1}_{B_2^*}}\\ &\lesssim_{d,p,\gamma} \|S\|_{L^p_{B_1}\to L^{p,\infty}_{B_2}} \Big\|\sum_{W\in\mathcal{W}} b_{1,W}\Big\|_{L^p_{B_1}} |Q_0|^{1/p'} \langle f_2\rangle_{3Q_0,q'}.\end{aligned}$$

By (4.12) for $r=p$ we obtain

$$|III_1| \lesssim_{d,p,\gamma} A(p)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}.$$

Likewise, the weak type (p,p) condition (1.7a) implies $L^p_{B_1} \to L^{p,\infty}_{B_2}$ boundedness of S_W. Using this and $\operatorname{supp}(S_W b_{1,W}) \subset 3W$, (4.9a), (4.10), and $p>1$ we estimate,

$$\begin{aligned}|III_2| &\le \sum_{W\in\mathcal{W}} \|S_W b_{1,W}\|_{L^{p,\infty}_{B_2^{**}}} \|g_2 \mathbb{1}_{3W}\|_{L^{p',1}_{B_2^*}}\\ &\le A(p) \sum_{W\in\mathcal{W}} \|b_{1,W}\|_{L^p_{B_1}} \|g_2\|_{L^\infty_{B_2^*}} \|\mathbb{1}_{3W}\|_{L^{p',1}_{B_2^*}}\\ &\lesssim_{d,p,\gamma} A(p) \sum_{W\in\mathcal{W}} |W|^{1/p}\langle f_1\rangle_{W,p} \|g_2\|_{L^\infty_{B_2^*}} |W|^{1/p'}\\ &\lesssim_{d,\gamma} A(p) \sum_{W\in\mathcal{W}} |W|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\end{aligned}$$

and hence, by the disjointness of the cubes W,

$$|III_2| \lesssim_{d,p,\gamma} A(p)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}.$$

Next we estimate III_3 and first show that

$$\left.\begin{array}{l}\langle T_j b_{1,W}, b_{2,W'}\rangle \neq 0\\ L(W) < j \leq L(W')\end{array}\right\} \quad \Longrightarrow \quad j \leq L(W') \leq L(W) + 2 \leq j+2. \tag{4.20}$$

To see (4.20) first observe that $T_j b_{1,W}$ is supported on a cube R_W centered at x_W with side length $2^{j+1} + 2^{L(W)}$. Hence, if $\langle T_j b_{1,W}, b_{2,W'}\rangle \neq 0$, then we get from (4.7) and the triangle inequality

$$\begin{aligned} 5\,\mathrm{diam}(W') &\leq \mathrm{dist}(W', \Omega^{\complement}) \leq \mathrm{diam}(W') + \mathrm{diam}(R_W) + \mathrm{dist}(W, \Omega^{\complement})\\ &\leq \mathrm{diam}(W') + (2^{j+1} + 2^{L(W)})\sqrt{d} + 12\sqrt{d}2^{L(W)}. \end{aligned}$$

Hence since $L(W) < j \leq L(W')$ we get $2^{L(W')+1} \leq 13 \cdot 2^{L(W)}$ which gives (4.20). Also, with these specifications $W \subset 3W'$ if $\langle T_j b_{1,W}, b_{2,W'}\rangle \neq 0$. By the single scale (p,q) condition (1.8),

$$\|T_j\|_{L^p \to L^q} \leq 2^{-jd(1/p-1/q)} A_\circ(p,q). \tag{4.21}$$

Hence using Hölder's inequality, (4.21) and (4.9a) we get

$$\begin{aligned} |III_3| &\leq \sum_{N_1 \leq j \leq N_2} \sum_{\substack{W' \in \mathcal{W}:\\ j \leq L(W') \leq j+2}} \sum_{\substack{W \in \mathcal{W}:\\ L(W')-2 \leq L(W) \leq j\\ W \subset 3W'}} |\langle T_j b_{1,W}, b_{2,W'}\rangle|\\ &\leq A_\circ(p,q) \sum_{N_1 \leq j \leq N_2} 2^{-jd(1/p-1/q)} \sum_{\substack{W,W' \in \mathcal{W}:\, W \subset 3W'\\ j \leq L(W') \leq j+2\\ L(W')-2 \leq L(W) \leq j}} \|b_{1,W}\|_{L^p_{B_1}} \|b_{2,W'}\|_{L^{q'}_{B_2^*}}\\ &\lesssim_{d,\gamma} A_\circ(p,q) \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,q'}\\ &\qquad \times \sum_{N_1 \leq j \leq N_2} \sum_{\substack{W,W':\, W \subset 3W'\\ j \leq L(W') \leq j+2\\ L(W')-2 \leq L(W) \leq j}} 2^{-jd(1/p-1/q)} |W|^{1/p} |W'|^{1-1/q}\\ &\lesssim_{d,\gamma} A_\circ(p,q) \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,q'} \sum_{W' \in \mathcal{W}} |W'| \end{aligned}$$

and thus, by disjointness of the W',

$$|III_3| \lesssim_{d,\gamma} A_\circ(p,q) \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,q'} |Q_0|.$$

Finally, consider the term

$$III_4 = \sum_{N_1 \leq j \leq N_2} \sum_{\substack{(W,W') \in \mathcal{W} \times \mathcal{W}\\ L(W) < j\\ L(W') < j}} \langle T_j b_{1,W}, b_{2,W'}\rangle. \tag{4.22}$$

Let $\varepsilon' > 0$ such that

$$\varepsilon' < \min\{1/p, 1/q', \varepsilon\} \tag{4.23}$$

and let ℓ be a positive integer so that

$$\ell < \frac{100}{\varepsilon'} \log\big(2 + \frac{B}{A_\circ(p,q)}\big) \le \ell + 1. \tag{4.24}$$

We split

$$\mathcal{V}_j = (-\infty, j)^2 \cap \mathbb{Z}^2 = \mathcal{V}_{j,1} \cup \mathcal{V}_{j,2} \cup \mathcal{V}_{j,3}$$

into three regions putting

$$\mathcal{V}_{j,1} = \{(L_1, L_2) \in \mathcal{V}_j : j - \ell \le L_1 < j,\ j - \ell \le L_2 < j\},$$

$$\mathcal{V}_{j,2} = \{(L_1, L_2) \in \mathcal{V}_j \setminus \mathcal{V}_{j,1} :\ L_1 \le L_2\},$$
$$\mathcal{V}_{j,3} = \{(L_1, L_2) \in \mathcal{V}_j \setminus \mathcal{V}_{j,1} :\ L_1 > L_2\}.$$

Then $III_4 = \sum_{i=1}^3 IV_i$ where for $i = 1, 2, 3$,

$$IV_i = \sum_{N_1 \le j \le N_2} \sum_{\substack{W, W' \in \mathcal{W} \\ (L(W), L(W')) \in \mathcal{V}_{j,i}}} \langle T_j b_{1,W}, b_{2,W'} \rangle. \tag{4.25}$$

Let $\mathfrak{R}_j$ be the collection of dyadic subcubes of Q_0 of side length 2^j. To estimate IV_1 we tile Q_0 into such cubes and write

$$IV_1 = \sum_{N_1 \le j \le N_2} \sum_{R \in \mathfrak{R}_j} \Big\langle \sum_{\substack{W \subset R \\ j - \ell \le L(W) < j}} T_j b_{1,W}, \sum_{j - \ell \le L(W') < j} b_{2,W'} \mathbb{1}_{3R} \Big\rangle. \tag{4.26}$$

By Hölder's inequality and the single scale (p,q) condition (1.8) (in the form of (4.21)) we get

$$\begin{aligned}
|IV_1| &\le A_\circ(p,q) \sum_{N_1 \le j \le N_2} 2^{-jd(1/p - 1/q)} \\
&\qquad \times \sum_{R \in \mathfrak{R}_j} \Big\| \sum_{\substack{W \subset R \\ j - \ell \le L(W) < j}} b_{1,W} \Big\|_{L^p_{B_1}} \Big\| \sum_{j - \ell \le L(W') < j} b_{2,W'} \mathbb{1}_{3R} \Big\|_{L^{q'}_{B_2^*}} \\
&\le A_\circ(p,q) \sum_{N_1 \le j \le N_2} \sum_{R \in \mathfrak{R}_j} |R|^{-(1/p - 1/q)} \\
&\qquad \times \Big(\sum_{\substack{W \subset R \\ j - \ell \le L(W) < j}} \|b_{1,W}\|^p_{L^p_{B_1}} \Big)^{1/p} \Big(\sum_{\substack{W' \subset 3R \\ j - \ell \le L(W') < j}} \|b_{2,W'}\|^{q'}_{L^{q'}_{B_2^*}} \Big)^{1/q'}
\end{aligned}$$

and using (4.9a) this expression is bounded by $C_{d,\gamma}A_\circ(p,q)$ times

$$\begin{aligned}
&\sum_{N_1\le j\le N_2}\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\sum_{R\in\mathfrak{R}_j}|R|^{-(1/p-1/q)}\\
&\qquad\times\Big(\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}}|W|\Big)^{1/p}\Big(\sum_{\substack{W'\subset 3R\\ j-\ell\le L(W')<j}}|W'|\Big)^{1/q'}\\
&\lesssim_{d,\gamma}\sum_{N_1\le j\le N_2}\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\\
&\qquad\times\sum_{R\in\mathfrak{R}_j}|R|^{-(1/p-1/q)}\Big(\sum_{\substack{W\subset 3R\\ j-\ell\le L(W)<j}}|W|\Big)^{1/p+1-1/q}.
\end{aligned}$$

Using $p\le q$ and the disjointness of the W we see that the last expression is dominated by a constant $\tilde C_{d,\gamma,p,q}$ times

$$\begin{aligned}
&\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\sum_{N_1\le j\le N_2}\sum_{R\in\mathfrak{R}_j}\sum_{\substack{W\subset 3R\\ j-\ell\le L(W)<j}}|W|\\
&\le 3^d\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\sum_{N_1\le j\le N_2}\sum_{R\in\mathfrak{R}_j}\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}}|W|\\
&\lesssim_d\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\sum_{W\in\mathcal{W}}|W|\sum_{\substack{j:N_1\le j\le N_2\\ L(W)<j\le L(W)+\ell}}1\\
&\lesssim_d \ell|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}.
\end{aligned}$$

Thus, using the definition of ℓ in (4.24) we get

$$|IV_1|\lesssim_{d,\gamma,\varepsilon,p,q}A_\circ(p,q)\log\big(2+\frac{B}{A_\circ(p,q)}\big)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}. \tag{4.27}$$

We now turn to the terms IV_2, IV_3 and claim that

$$|IV_2|+|IV_3|\lesssim_{d,\gamma,p,q,\varepsilon}A_\circ(p,q)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}. \tag{4.28}$$

We first note that by the single scale ε-regularity conditions (1.9a), (1.9b), and Corollary 3.5,

$$\|T_j(\mathrm{I}-\mathbb{E}_{s_1-j})\|_{L^p_{B_1}\to L^q_{B_2}}\lesssim_\varepsilon B2^{-jd(\frac1p-\frac1q)}2^{-\varepsilon' s_1}, \tag{4.29a}$$

$$\|T_j^*(\mathrm{I}-\mathbb{E}_{s_2-j})\|_{L^{q'}_{B_2^*}\to L^{p'}_{B_1^*}}\lesssim_\varepsilon B2^{-jd(\frac1p-\frac1q)}2^{-\varepsilon' s_2}. \tag{4.29b}$$

where ε' is as in (4.23).

Write, with $\mathfrak{R}_j$ as in (4.26),

$$IV_2 = \sum_{N_1\le j\le N_2}\sum_{R\in\mathfrak{R}_j}\sum_{s_2=1}^{\infty}\sum_{s_1=\max\{s_2,\ell+1\}}^{\infty} \Big\langle \sum_{\substack{W\subset R\\ L(W)=j-s_1}} T_j b_{1,W}, \sum_{\substack{W'\subset 3R\\ L(W')=j-s_2}} b_{2,W'}\mathbb{1}_{3R}\Big\rangle. \quad (4.30)$$

Note that for $L(W)=j-s_1$, $b_{1,W}=(I-\mathbb{E}_{s_1-j})f_{1,W}$. By Hölder's inequality and (4.29a) we have for $R\in\mathfrak{R}_j$,

$$\begin{aligned}
&\Big|\Big\langle \sum_{\substack{W\subset R\\ L(W)=j-s_1}} T_j b_{1,W}, \sum_{\substack{W'\subset 3R\\ L(W')=j-s_2}} b_{2,W'}\mathbb{1}_{3R}\Big\rangle\Big| \\
&\le \Big\|T_j(I-\mathbb{E}_{s_1-j})\sum_{\substack{W\subset R\\ L(W)=j-s_1}} f_{1,W}\Big\|_{L^q_{B_2}}\Big\|\sum_{\substack{W'\subset 3R\\ L(W')=j-s_2}} b_{2,W'}\Big\|_{L^{q'}_{B_2^*}} \\
&\lesssim_\varepsilon B2^{-\varepsilon' s_1}|R|^{-(\frac1p-\frac1q)}\Big(\sum_{\substack{W\subset R\\ L(W)=j-s_1}}\|f_{1,W}\|^p_{L^p_{B_1}}\Big)^{1/p}\Big(\sum_{\substack{W'\subset 3R\\ L(W')=j-s_2}}\|b_{2,W'}\|^{q'}_{L^{q'}_{B_2^*}}\Big)^{1/q'}.
\end{aligned}$$

In the above formula for IV_2 we interchange the j-sum and the (s_1,s_2)-sums, write $j=s_1n+i$ with $i=1,\dots,s_1$ and estimate (invoking (4.9a) again)

$$\begin{aligned}
|IV_2| &\lesssim_\varepsilon \sum_{s_2=1}^{\infty}\sum_{s_1=\max\{s_2,\ell+1\}}^{\infty} B2^{-\varepsilon' s_1}\sum_{i=1}^{s_1}\sum_{\substack{n\in\mathbb{Z}\\ s_1n+i\in[N_1,N_2]}}\sum_{R\in\mathfrak{R}_{s_1n+i}}|R|^{-(\frac1p-\frac1q)} \\
&\qquad\times\Big(\sum_{\substack{W\subset R\\ L(W)=s_1n+i-s_1}}\|f_{1,W}\|^p_{L^p_{B_1}}\Big)^{\frac1p}\Big(\sum_{\substack{W'\subset 3R\\ L(W')=s_1n+i-s_2}}\|b_{2,W'}\|^{q'}_{L^{q'}_{B_2^*}}\Big)^{\frac1{q'}} \\
&\lesssim_{d,\gamma} \sum_{s_2=1}^{\infty}\sum_{s_1=\max\{s_2,\ell+1\}}^{\infty} B2^{-\varepsilon' s_1} \\
&\qquad\times\sum_{i=1}^{s_1}\sum_{\substack{n\in\mathbb{Z}\\ s_1n+i\in[N_1,N_2]}}\sum_{R\in\mathfrak{R}_{s_1n+i}}\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}\Gamma(R,n,i),
\end{aligned}$$

where

$$\Gamma(R,n,i)=|R|^{-(\frac1p-\frac1q)}\Big(\sum_{\substack{W\subset R\\ L(W)=s_1n+i-s_1}}|W|\Big)^{\frac1p}\Big(\sum_{\substack{W'\subset 3R\\ L(W')=s_1n+i-s_2}}|W'|\Big)^{1-\frac1q}.$$

We crudely estimate, using $p \le q$,

$$\Gamma(R,n,i) \le |R|^{-(1/p-1/q)}\Big(\sum_{\nu=s_1n+i-s_1}^{s_1n+i-s_2}\sum_{\substack{W\subset 3R\\ L(W)=\nu}}|W|\Big)^{1/p+1-1/q}$$
$$\le 3^{d(1/p-1/q)}\sum_{\nu=s_1n+i-s_1}^{s_1n+i-s_2}\sum_{\substack{W\subset 3R\\ L(W)=\nu}}|W|.$$

For fixed $W\in\mathcal{W}$ consider the set of all triples (R,n,i) such that $s_1n+i-s_1 \le L(W) \le s_1n+i-s_2$, $R\in\mathfrak{R}_{s_1n+i}$ and $W\subset 3R$, and observe that the cardinality of this set is bounded above by $3^d(s_1-s_2+1)$. Combining this with the above estimates and summing over $W\in\mathcal{W}$ we obtain the bound

$$|IV_2| \lesssim_{d,\gamma,\varepsilon} \langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}|Q_0|\sum_{s_2=1}^{\infty}\sum_{s_1=\max\{s_2,\ell+1\}}^{\infty} B2^{-\varepsilon' s_1}(s_1-s_2+1)$$

and the double sum is bounded by

$$C_{\varepsilon,p,q}\Big(\sum_{s_2=1}^{\ell+1}B2^{-\varepsilon'\ell}(\ell+1)+\sum_{s_2=\ell+1}^{\infty}B2^{-\varepsilon' s_2}\Big)$$
$$\lesssim_{\varepsilon,p,q} B2^{-\varepsilon'\ell}(\ell+1)^2 \lesssim_{\varepsilon,p,q} B2^{-\varepsilon'\ell/2} \lesssim_{\varepsilon,p,q} A_\circ(p,q)$$

by the definition of ℓ in (4.24). This establishes (4.28) for the term $|IV_2|$.

The estimation of IV_3 is very similar. We may write

$$IV_3 = \sum_{N_1\le j\le N_2}\sum_{R\in\mathfrak{R}_j}\sum_{s_1=1}^{\infty}\sum_{s_2=\max\{s_1+1,\ell+1\}}^{\infty}$$
$$\Big\langle \sum_{\substack{W\in\mathcal{W}:\\ W\subset R\\ L(W)=j-s_1}} b_{1,W}\mathbb{1}_R, \sum_{\substack{W'\in\mathcal{W}:\\ L(W')=j-s_2}} T_j^*[b_{2,W'}\mathbb{1}_{3R}]\Big\rangle \quad (4.31)$$

By Hölder's inequality and (4.29b) we get for $R\in\mathfrak{R}_j$

$$\Big|\Big\langle \sum_{\substack{W\subset R\\ L(W)=j-s_1}} b_{1,W}, \sum_{L(W')=j-s_2} T_j^*[b_{2,W'}\mathbb{1}_{3R}]\Big\rangle\Big|$$
$$\lesssim_\varepsilon B2^{-\varepsilon' s_2}|R|^{-(\frac1p-\frac1q)}\Big(\sum_{\substack{W\subset R\\ L(W)=j-s_1}}\|b_{1,W}\|^p_{L^p_{B_1}}\Big)^{1/p}\Big(\sum_{\substack{W'\subset 3R\\ L(W')=j-s_2}}\|f_{2,W'}\|^{q'}_{L^{q'}_{B_2^*}}\Big)^{1/q'}$$

and from here on the argument is analogous to the treatment of the term IV_2. □

Remark. It is instructive to observe that the term III_4 can be treated more crudely if one does not aim to obtain the constant $A_\circ(p,q)\log(2+\frac{B}{A_\circ(p,q)})$

in (1.11). More precisely, one simply splits $(-\infty, j)^2 \cap \mathbb{Z}^2$ into two regions $\tilde{\mathcal{V}}_{j,2}$ and $\tilde{\mathcal{V}}_{j,3}$, where

$$\tilde{\mathcal{V}}_{j,2} = \{(L_1, L_2) : L_1 \leq L_2 < j\} \quad \text{and} \quad \tilde{\mathcal{V}}_{j,3} = \{(L_1, L_2) : j > L_1 > L_2\}.$$

Then split $III_4 = \sum_{i=2}^{3} IV_i$, where IV_i are as in (4.25) but with $\mathcal{V}_{j,i}$ replaced by $\tilde{\mathcal{V}}_{j,i}$. One then considers the sum in s_1 in (4.30) to start directly from s_2, and the sum in s_2 in (4.31) to start directly from $s_1 + 1$. Using the same arguments, one obtains

$$|IV_2| + |IV_3| \lesssim_{d,\gamma,p,q,\varepsilon} B|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,q'}$$

instead of (4.28). Note that, as the term IV_1 does not appear in this case (see the bound (4.27)), this yields sparse domination with the constant $\mathcal{C}$ in (1.11) replaced by $A(p) + A(q) + A_\circ(p, q) + B$.

5. Maximal operators, square functions and long variations

In this section we show that Corollary 1.2 yields sparse domination results for maximal functions, ℓ^r-valued variants, r-variation norm operators and maximal and variational truncations of sums of operators. An application of Theorem 1.3 also yields necessary conditions for our sparse domination inequalities. We will formally state necessary conditions only for maximal functions and ℓ^r-valued functions (Theorem 5.1) and leave to the reader the analogous formulations of those conditions for r-variation norm operators (Theorem 5.2), maximal truncations (Theorem 5.3) and variational truncations (Theorem 5.4).

5.1. **Maximal functions and ℓ^r-variants.** Given a family of operators $\{T_j\}_{j\in\mathbb{Z}}$ in Op_{B_1,B_2}, consider the operators

$$S_r Tf(x) = \Big(\sum_{j\in\mathbb{Z}} |T_j f(x)|_{B_2}^r\Big)^{1/r} \tag{5.1}$$

when $1 \leq r < \infty$ and also the maximal operator

$$S_\infty Tf(x) = \sup_{j\in\mathbb{Z}} |T_j f(x)|_{B_2}. \tag{5.2}$$

Theorem 5.1. *Let $1 < p \leq q < \infty$ and $1 \leq r \leq \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying* (1.6).

(i) Suppose that the inequalities

$$\big\|S_r Tf\big\|_{L^{p,\infty}} \leq A(p)\|f\|_{L^p_{B_1}} \qquad \text{and} \qquad \big\|S_r Tf\big\|_{L^q} \leq A(q)\|f\|_{L^{q,1}_{B_1}} \tag{5.3}$$

hold for all $f \in \mathrm{S}_{B_1}$. Moreover, assume that the rescaled operators $\mathrm{Dil}_{2^j} T_j$ satisfy the single scale (p, q) condition (1.8) *and single scale ε-regularity conditions* (1.9a) *and* (1.9b). *Let $\mathcal{C}$ be as in* (1.11). *Then for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative measurable functions ω,*

$$\langle S_r Tf, \omega\rangle \lesssim \mathcal{C}\Lambda^*_{p,q'}(f, \omega). \tag{5.4}$$

(ii) In addition, assume $1 < p < q < \infty$. If the family of operators $\{T_j\}_{j\in\mathbb{Z}}$ satisfies $T_j : \mathrm{S}_{B_1} \to L^1_{B_2,\mathrm{loc}}$ and the strengthened support condition (1.14)*, then the condition*

$$\|S_rT\|_{L^p_{B_1}\to L^{p,\infty}} + \|S_rT\|_{L^{q,1}_{B_1}\to L^q} + \sup_{j\in\mathbb{Z}} \|\mathrm{Dil}_{2^j}T_j\|_{L^p_{B_1}\to L^q_{B_2}} < \infty$$

is necessary for the conclusion (5.4) *to hold.*

Proof of Theorem 5.1. We begin with the proof for $1 \le r < \infty$.

Let $\delta_{kk} = 1$ and $\delta_{jk} = 0$ if $j \neq k$. Let $N_1 \le N_2$ be integers, and for each integer $j \in [N_1, N_2]$, we define the operator H_j sending $L^p_{B_1}$ functions to $\ell^r_{B_2}$-valued functions by

$$H_jf(x,k) = \begin{cases} \delta_{jk}T_jf(x) & \text{if } N_1 \le k \le N_2, \\ 0 & \text{if } k \notin [N_1, N_2]. \end{cases} \tag{5.5}$$

We note that

$$\Big(\sum_{j=N_1}^{N_2} |T_jf(x)|^r_{B_2}\Big)^{1/r} = \Big(\sum_{k=-\infty}^{\infty} \Big|\sum_{j=N_1}^{N_2} H_jf(x,k)\Big|^r_{B_2}\Big)^{1/r}. \tag{5.6}$$

By (5.3) we have

$$\Big\|\sum_{j=N_1}^{N_2} H_j\Big\|_{L^p_{B_1}\to L^{p,\infty}(\ell^r_{B_2})} \le A(p), \qquad \Big\|\sum_{j=N_1}^{N_2} H_j\Big\|_{L^{q,1}_{B_1}\to L^q(\ell^r_{B_2})} \le A(q),$$

where we write $L^{p,\infty}(\ell^r_{B_2})$ to denote $L^{p,\infty}_{\ell^r_{B_2}}$. The adjoint of H_j, acting on $\ell^{r'}_{B_2^*}$-valued functions g, is given by

$$H_j^*g(x) = \sum_{k=N_1}^{N_2} \delta_{jk}T_j^*g_k(x).$$

The assumptions on $\mathrm{Dil}_{2^j}T_j$ can be rewritten as

$$\sup_{j\in\mathbb{Z}} \|\mathrm{Dil}_{2^j}H_j\|_{L^p_{B_1}\to L^q(\ell^r_{B_2})} \le A_\circ(p,q)$$

and

$$\sup_{|h|\le 1} |h|^{-\varepsilon} \sup_{j\in\mathbb{Z}} \|(\mathrm{Dil}_{2^j}H_j)\circ\Delta_h\|_{L^p_{B_1}\to L^q(\ell^r_{B_2})} \le B$$
$$\sup_{|h|\le 1} |h|^{-\varepsilon} \sup_{j\in\mathbb{Z}} \|(\mathrm{Dil}_{2^j}H_j^*)\circ\Delta_h\|_{L^{q'}(\ell^{r'}_{B_2^*})\to L^{p'}_{B_1^*}} \le B.$$

By Corollary 1.2 applied to the sequence $\{H_j\}_{j\in\mathbb{Z}}$ in $\mathrm{Op}_{B_1,\ell^r_{B_2}}$, we get the conclusion

$$\int_{\mathbb{R}^d} \Big(\sum_{k=-\infty}^{\infty} \Big|\sum_{j=N_1}^{N_2} H_jf(x,k)\Big|^r_{B_2}\Big)^{1/r} \omega(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega),$$

which by (5.6) implies

$$\int_{\mathbb{R}^d} \Big(\sum_{j=N_1}^{N_2} |T_j f(x)|_{B_2}^r \Big)^{1/r} \omega(x) dx \lesssim \mathcal{C} \Lambda^*_{p,q'}(f,\omega).$$

We apply the monotone convergence theorem to let $N_1 \to -\infty$ and $N_2 \to \infty$ and obtain the desired conclusion. This is possible since the implicit constant in the conclusion of Corollary 1.2 does not depend on B_1, B_2.

The proof for $r = \infty$ is essentially the same, with notional changes. Since $H_j f(\cdot, k) = 0$ when $k \notin [N_1, N_2]$, we can work with $\ell^\infty_{B_2}$ over the finite set $\mathbb{Z} \cap [N_1, N_2]$. Then there are no complications with the dual space, which is $\ell^1_{B_2^*}$ over $\mathbb{Z} \cap [N_1, N_2]$.

For part (ii) one uses Theorem 1.3 in conjunction with (5.6) and immediately arrives at the desired conclusion, via the monotone convergence theorem. □

5.2. **Variation norms.** We now turn to the variation norms $V^r_{B_2} \equiv V^r_{B_2}(\mathbb{Z})$ defined on B_2-valued functions of the integers $n \mapsto a(n)$. Let $|a|_{V^\infty_{B_2}} = |a|_{L^\infty_{B_2}}$ and, for $1 \le r < \infty$,

$$|a|_{V^r_{B_2}} = \sup_{n_1 < \cdots < n_M} |a(n_1)|_{B_2} + \Big(\sum_{\nu=1}^{M-1} |a(n_{\nu+1}) - a(n_\nu)|_{B_2}^r \Big)^{1/r} \tag{5.7}$$

where the supremum is taken over all positive integers M and all finite increasing sequences of integers $n_1 < ... < n_M$. Similarly, if $I_{N_1,N_2} = [N_1, N_2] \cap \mathbb{Z}$ we define the $V^r_{B_2}(I_{N_1,N_2})$ norm on functions on I_{N_1,N_2} in the same way, restricting $n_1, \dots, n_M$ to I_{N_1,N_2}.

Given a sequence $T = \{T_j\}_{j\in\mathbb{Z}}$ in Op_{B_1,B_2} we define $\mathcal{V}^r T f(x)$ to be the $V^r_{B_2}$ norm of the sequence $j \mapsto T_j f(x)$. The L^p norm of $\mathcal{V}^r T f$ is just the $L^p(V^r_{B_2})$ norm of the sequence $\{T_j f\}_{j\in\mathbb{Z}}$. We define $\mathcal{V}^r_{N_1,N_2} T f(x)$ to be the $V^r_{B_2}$ norm of the sequence $j \mapsto \mathbb{1}_{I_{N_1,N_2}}(j) T_j f(x)$.

The proof of the following theorem is almost identical to that of Theorem 5.1.

Theorem 5.2. *Let $1 < p \le q < \infty$ and $1 \le r \le \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying* (1.6)*. Suppose that the inequalities*

$$\big\|\mathcal{V}^r T f\big\|_{L^{p,\infty}} \le A(p) \|f\|_{L^p_{B_1}} \qquad \text{and} \qquad \big\|\mathcal{V}^r T f\big\|_{L^q} \le A(q) \|f\|_{L^{q,1}_{B_1}} \tag{5.8}$$

hold for all $f \in \mathrm{S}_{B_1}$. Moreover, assume that the rescaled operators $\mathrm{Dil}_{2^j} T_j$ satisfy the single scale (p,q) condition (1.8) *and single scale ε-regularity conditions* (1.9a) *and* (1.9b)*. Let $\mathcal{C}$ be as in* (1.11)*. Then for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative measurable functions ω,*

$$\langle \mathcal{V}^r T f, \omega \rangle \lesssim \mathcal{C} \Lambda^*_{p,q'}(f,\omega). \tag{5.9}$$

Proof. In view of Theorem 5.1, it suffices to consider the case $r < \infty$. Given $N_1 \le N_2$ we define $H_j f(x,k)$ as in (5.5), for $N_1 \le k \le N_2$. Note that for fixed x, $N_1 \le n_1 < \cdots < n_M \le N_2$,

$$\begin{aligned} &|T_{n_1} f(x)|_{B_2} + \Big(\sum_{\nu=1}^{M-1} |T_{n_{\nu+1}} f(x) - T_{n_\nu} f(x)|_{B_2}^r \Big)^{1/r} \\ &= \Big| \sum_{j=N_1}^{N_2} H_j f(x,n_1) \Big|_{B_2} + \Big(\sum_{\nu=1}^{M-1} \Big| \sum_{j=1}^{N} H_j f(x,n_{\nu+1}) - \sum_{j=1}^{N} H_j f(x,n_\nu) \Big|_{B_2}^r \Big)^{1/r}. \end{aligned} \tag{5.10}$$

By (5.8) we have

$$\Big\| \sum_{j=N_1}^{N_2} H_j \Big\|_{L^p_{B_1} \to L^{p,\infty}(V^r_{B_2})} \le A(p), \qquad \Big\| \sum_{j=N_1}^{N_2} H_j \Big\|_{L^{q,1}_{B_1} \to L^q(V^r_{B_2})} \le A(q),$$

where $V^r_{B_2}$ is interpreted to be the space $V^r_{B_2}(I_{N_1,N_2})$ and all the constants in what follows will be independent of N_1 and N_2. The pairing between $V^r_{B_2}(I_{N_1,N_2})$ and its dual is the standard one,

$$\langle a, b \rangle = \sum_{n=N_1}^{N_2} \langle a(n), b(n) \rangle_{(B_2, B_2^*)}$$

and we have,

$$|b|_{(V^r_{B_2}(I_{N_1,N_2}))^*} = \sup_{|a|_{V^r_{B_2}(I_{N_1,N_2})} \le 1} \Big| \sum_{n=N_1}^{N_2} \langle a(n), b(n) \rangle_{(B_2, B_2^*)} \Big|.$$

For $\delta^{(j)} = (\delta_{j,N_1}, \dots, \delta_{j,N_2})$ we have, for $j = N_1, \dots, N_2$,

$$|\delta^{(j)}|_{V^r_{B_2}(I_{N_1,N_2})} = 2^{1/r} \quad \text{and} \quad |\delta^{(j)}|_{(V^r_{B_2}(I_{N_1,N_2}))^*} = 2^{1/r'}.$$

The adjoint of H_j, acting on $(V^r_{B_2}(I_{N_1,N_2}))^*$-valued functions $g = \{g_k\}_{k=N_1}^{N_2}$, is given by

$$H_j^* g(x) = \sum_{k=N_1}^{N_2} \delta_{jk} T_j^* g_k(x).$$

These observations imply

$$\begin{aligned} &\|\mathrm{Dil}_{2^j} H_j\|_{L^p_{B_1} \to L^q(V^r_{B_2})} = 2^{1/r} \|\mathrm{Dil}_{2^j} T_j\|_{L^p_{B_1} \to L^q_{B_2}}, \\ &\|(\mathrm{Dil}_{2^j} H_j) \circ \Delta_h\|_{L^p_{B_1} \to L^q(V^r_{B_2})} = 2^{1/r} \|(\mathrm{Dil}_{2^j} T_j) \circ \Delta_h\|_{L^p_{B_1} \to L^q_{B_2}}, \\ &\|(\mathrm{Dil}_{2^j} H_j^*) \circ \Delta_h\|_{L^{q'}((V^r_{B_2})^*) \to L^{p'}_{B_1^*}} = 2^{1/r'} \|(\mathrm{Dil}_{2^j} T_j^*) \circ \Delta_h\|_{L^{q'}_{B_2^*} \to L^{p'}_{B_1^*}}. \end{aligned}$$

The hypothesis of Theorem 1.1 are then satisfied for the sequence $\{H_j\}_{j\in\mathbb{Z}}$ in $\mathrm{Op}_{B_1,V^r_{B_2}}$. Thus, by Corollary 1.2 we obtain

$$\int_{\mathbb{R}^d} \Big| \sum_{j=N_1}^{N_2} H_j f(x,\cdot)\Big|_{V^r_{B_2}} \omega(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega),$$

which by (5.10) implies

$$\int_{\mathbb{R}^d} \mathcal{V}^r_{N_1,N_2} Tf(x)\omega(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega).$$

As the implicit constant in Corollary 1.2 does not depend on the Banach spaces B_1, B_2 we may apply the monotone convergence theorem and let $N_1 \to -\infty$ and $N_2 \to \infty$ to obtain the desired conclusion (5.9). □

5.3. **Truncations of sums.** We will give a variant of Corollary 1.2 in the spirit of Cotlar's inequality on maximal operators for truncations of singular integrals.

Theorem 5.3. *Let $1 < p \le q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying* (1.6), (1.8), (1.9a), *and* (1.9b). *Moreover, assume that the estimates*

$$\Big\| \sup_{\substack{(n_1,n_2):\\ N_1\le n_1\le n_2\le N_2}} \Big| \sum_{j=n_1}^{n_2} T_j f\big|_{B_2}\Big\|_{L^{p,\infty}} \le A(p)\|f\|_{L^p_{B_1}}, \tag{5.11a}$$

$$\Big\| \sup_{\substack{(n_1,n_2):\\ N_1\le n_1\le n_2\le N_2}} \Big| \sum_{j=n_1}^{n_2} T_j f\big|_{B_2}\Big\|_{L^{q}} \le A(q)\|f\|_{L^{q,1}_{B_1}} \tag{5.11b}$$

hold uniformly for all (N_1,N_2) with $N_1 \le N_2$. Let $\mathcal{C}$ be as in (1.11). *Then for all $f \in \mathrm{S}_{B_1}$, all $\mathbb{R}$-valued nonnegative measurable functions ω, and all integers N_1, N_2 with $N_1 \le N_2$,*

$$\int_{\mathbb{R}^d} \sup_{\substack{(n_1,n_2):\\ N_1\le n_1\le n_2\le N_2}} \Big| \sum_{j=n_1}^{n_2} T_j f(x)\big|_{B_2} \omega(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega).$$

Proof. Define $U(N_1,N_2) = \{(n_1,n_2) : N_1 \le n_1 \le n_2 \le N_2\}$ and $\ell^\infty_{B_2}$ as the space of all bounded B_2-valued functions on $U(N_1,N_2)$. Define operators H_j in $\mathrm{Op}_{B_1,\ell^\infty_{B_2}}$ by

$$H_j f(x,n_1,n_2) = \begin{cases} T_j f(x) & \text{if } N_1 \le n_1 \le j \le n_2 \le N_2,\\ 0 & \text{otherwise.}\end{cases}$$

Then apply Corollary 1.2 to the operators $\sum_{j=N_1}^{N_2} H_j$ as in the proof of Theorems 5.1. □

We also have a variational analogue.

Theorem 5.4. *Let $1 < p \le q < \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying* (1.6), (1.7a), (1.7b), (1.8), (1.9a), *and* (1.9b). *Moreover, assume that the estimates*

$$\Big\| \sup_{M\in\mathbb{N}} \sup_{N_1\le n_1<\dots<n_M\le N_2} \Big(\sum_{\nu=1}^{M-1} \Big| \sum_{j=n_\nu+1}^{n_{\nu+1}} T_j f \Big|_{B_2}^r \Big)^{1/r} \Big\|_{L^{p,\infty}} \le A(p)\|f\|_{L^p_{B_1}} \tag{5.12a}$$

$$\Big\| \sup_{M\in\mathbb{N}} \sup_{N_1\le n_1<\dots<n_M\le N_2} \Big(\sum_{\nu=1}^{M-1} \Big| \sum_{j=n_\nu+1}^{n_{\nu+1}} T_j f \Big|_{B_2}^r \Big)^{1/r} \Big\|_{L^{q}} \le A(p)\|f\|_{L^{q,1}_{B_1}} \tag{5.12b}$$

hold uniformly for all (N_1, N_2) with $N_1 \le N_2$. Let $\mathcal{C}$ be as in (1.11). *Then for all $f \in \mathrm{S}_{B_1}$, all $\mathbb{R}$-valued nonnegative measurable functions ω, and all integers N_1, N_2 with $N_1 \le N_2$,*

$$\int_{\mathbb{R}^d} \sup_{M\in\mathbb{N}} \sup_{N_1\le n_1<\dots<n_M\le N_2} \Big(\sum_{\nu=1}^{M-1} \Big| \sum_{j=n_\nu+1}^{n_{\nu+1}} T_j f(x) \Big|_{B_2}^r \Big)^{1/r} \omega(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega). \tag{5.13}$$

Proof. Let $V^r_{B_2} \equiv V^r_{B_2}(I_{N_1,N_2})$ denote the r-variation space of B_2-valued functions over the integers in $[N_1, N_2]$ and for $N_1 \le j \le N_2$, $N_1 \le n \le N_2$, define the operators $H_j \in \mathrm{Op}_{B_1,V^r_{B_2}}$ by

$$H_j f(x,n) = \begin{cases} T_j f(x) & \text{if } N_1 \le j \le n \le N_2, \\ 0 & \text{if } j > n. \end{cases}$$

Note that, by definition of H_j, $|\sum_{j=N_1}^{N_2} H_j f(x,\cdot)|_{V^r_{B_2}}$ equals to

$$\sup_{\substack{M\in\mathbb{N}\\ N_1\le n_1<\dots<n_M\le N_2}} \Big| \sum_{j=N_1}^{n_1} T_j f(x) \Big|_{B_2} + \Big(\sum_{\nu=1}^{M-1} \Big| \sum_{j=N_1}^{n_{\nu+1}} T_j f(x) - \sum_{j=N_1}^{n_\nu} T_j f(x) \Big|_{B_2}^r \Big)^{1/r} \tag{5.14}$$

and $|H_j f(x,\cdot)|_{V^r_{B_2}} = |T_j f(x)|_{B_2}$. Arguing as in Theorem 5.2, one may apply Corollary 1.2 to the operators $\sum_{j=N_1}^{N_2} H_j$ in $\mathrm{Op}_{B_1,V^r_{B_2}}$. Note, in particular, that in view of (5.14) the conditions (1.7a) and (1.7b) for $\sum_{j=N_1}^{N_2} H_j$ follow from (5.12a) and (5.12b) together with the fact that $\{T_j\}_{j\in\mathbb{Z}}$ in Op_{B_1,B_2} satisfy (1.7a) and (1.7b). This automatically yields (5.13). □

5.4. **Some simplifications for maximal operators.** The goal of this section is to remark that the proof of Theorem 5.1 can be simplified in the case $q \le r \le \infty$. Rather than deducing it from Corollary 1.2, we shall apply the proof method of Theorem 1.1 to the operators S_r and observe that a Calderón–Zygmund decomposition on f_2 is not required for the proof

to work. In particular, this allows us to remove the regularity hypothesis (1.9b) on the adjoints T_j^*. The precise statement reads as follows.

Theorem 5.5. *Let $1 < p \le q < \infty$ and $q \le r \le \infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} satisfying* (1.6)*. Suppose that the inequalities*

$$\big\|S_rTf\big\|_{L^{p,\infty}} \le A(p)\|f\|_{L^p_{B_1}} \qquad \text{and} \qquad \big\|S_rTf\big\|_{L^q} \le A(q)\|f\|_{L^{q,1}_{B_1}} \tag{5.15}$$

hold for all $f \in \mathrm{S}_{B_1}$. Moreover, assume that the rescaled operators $\mathrm{Dil}_{2^j}T_j$ satisfy the single scale (p,q) condition (1.8) *and single scale ε-regularity condition* (1.9a)*. Let $\mathcal{C}$ be as in* (1.11)*. Then for all $f \in \mathrm{S}_{B_1}$ and all $\mathbb{R}$-valued nonnegative measurable functions ω,*

$$\langle S_rTf,\omega\rangle \lesssim \mathcal{C}\Lambda^*_{p,q'}(f,\omega).$$

Proof. We sketch the main changes with respect to the proof of Theorem 1.1. As in Theorem 1.1, it suffices to show

$$\int_{\mathbb{R}^d} S_{r,N_1,N_2}f_1(x)f_2(x)dx \lesssim \mathcal{C}\Lambda^*_{p,q'}(f_1,f_2)$$

uniformly in $N_1 \le N_2$ for all $f_1 \in \mathrm{S}_{B_1}$ and $f_2 \in S_{\mathbb{R}}$, where

$$S_{r,N_1,N_2}f(x) := \Big(\sum_{j=N_1}^{N_2} |T_jf(x)|^r_{B_2}\Big)^{1/r}$$

for $q \le r < \infty$ and

$$S_{\infty,N_1,N_2}f(x) := \sup_{N_1\le j\le N_2} |T_jf(x)|_{B_2}.$$

This will in turn follow from verifying the inductive step in Claim 4.3 for the operators S_{r,N_1,N_2}.

If $r = \infty$, let $\lambda_j(x) \in B_2^*$ with $|\lambda_j(x)|_{B_2^*} \le 1$ such that $|T_jf_1(x)|_{B_2} = \langle T_jf_1,\lambda_j\rangle_{(B_2,B_2^*)}$ and let $x \mapsto j(x)$ be a measurable function such that

$$S_{\infty,N_1,N_2}f_1(x) \le 2\,|T_{j(x)}f_1(x)|_{B_2}.$$

Setting $X_j := \{x : j(x) = j\}$, note that

$$Sf_1(x) \equiv |T_{j(x)}f_1(x)|_{B_2} = \sum_{j=N_1}^{N_2} \langle T_jf_1(x),\lambda_j(x)\mathbb{1}_{X_j}(x)\rangle_{(B_2,B_2^*)}$$

and that the X_j are disjoint measurable sets such that $\sum_j \mathbb{1}_{X_j} \le \mathbb{1}_{3Q_0}$. If $q \le r < \infty$, we linearise the $\ell^r(B_2)$-norm for each x. That is, there exists $\{a_j(x)\}_{j\in\mathbb{Z}} \in \ell^{r'}(B_2^*)$, with $\|a_j(x)\|_{\ell^{r'}(B_2^*)} \le 1$, such that

$$Sf_1(x) \equiv S_{r,N_1,N_2}f_1(x) = \sum_{j=N_1}^{N_2} \langle T_jf_1(x),a_j(x)\rangle_{(B_2,B_2^*)}.$$

Note that we can treat the cases $r = \infty$ and $q \le r < \infty$ together by setting $a_j(x) = \lambda_j(x)1_{X_j}(x)$ for all $x \in 3Q_0$ and all $N_1 \le j \le N_2$, and $a_j(x) = 0$

otherwise; then $\{a_j(x)\}_{j\in\mathbb{Z}} \in \ell^1(B_2^*)$. Clearly, the operator S satisfies the bounds (1.7a), (1.7b) in view of (5.15).

We then perform a Calderón–Zygmund decomposition of f_1 as in (4.13). The first term in (4.13), corresponding to g_1, can be treated analogously. The second term in (4.13), corresponding to $\sum_{W\in\mathcal{W}} b_{1,W}$ can be further split as in (4.15), and I and II can be treated analogously. One is then left with proving (4.18) for III. Rather than performing a Calderón–Zygmund decomposition on f_2, we estimate the term directly.

Indeed, the analysis for III amounts to a simplified version of the analysis of the term III_4 in (4.22). One can define ℓ as in (4.24) and split

$$(-\infty, j)\cap\mathbb{Z} = \mathcal{V}_{j,1}\cup\mathcal{V}_{j,2},$$

where $\mathcal{V}_{j,1} := \{L : j-\ell \le L < j\}$ and $\mathcal{V}_{j,2} = \{L : L < j-\ell\}$. Note that here there is no further need to split $\mathcal{V}_{j,2}$, since we do not make use of a Calderón–Zygmund decomposition of f_2. Write $III = IV_1^\flat + IV_2^\flat$, where for $i = 1, 2$,

$$IV_i^\flat = \Big\langle \sum_{N_1\le j\le N_2} \sum_{\substack{W\in\mathcal{W},\\ L(W)\in\mathcal{V}_{j,i}}} T_j b_{1,W}, f_2\Big\rangle.$$

We first focus on $IV_1^\flat$. By Hölder's inequality with respect to x and j

$$IV_1^\flat \le IV_{1,1}^\flat IV_{1,2}^\flat, \tag{5.16}$$

where

$$IV_{1,1}^\flat = \Big(\sum_{j=N_1}^{N_2}\int \Big|\sum_{j-\ell\le L(W)<j} T_j b_{1,W}(x)\Big|_{B_2}^q dx\Big)^{1/q},$$

$$IV_{1,2}^\flat = \Big(\sum_{j=N_1}^{N_2}\int |a_j(x)|_{B_2}^{q'}|f_2(x)|^{q'}dx\Big)^{1/q'}.$$

Using that $\|a_j(x)\|_{\ell^{q'}(B_2^*)} \le \|a_j(x)\|_{\ell^{r'}(B_2^*)} = 1$ if $1\le r'\le q'$, we get

$$IV_{1,2}^\flat \lesssim \Big(\int_{3Q_0}|f_2(x)|^{q'}dx\Big)^{1/q'} \lesssim |Q_0|^{1-1/q}\langle f_2\rangle_{3Q_0,q'}. \tag{5.17}$$

For the term $IV_{1,1}^\flat$, introduce as in (4.26) the family $\mathfrak{R}_j$ of subcubes of Q_0 of side length 2^j and use the bounded overlap of $3R$ to write

$$IV_{1,1}^\flat \lesssim \Big(\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}\Big\|\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}} T_j b_{1,W}\Big\|_{L^q_{B_2}}^q\Big)^{1/q}.$$

The right-hand side above can then be handled essentially as IV_1 in (4.26) after using the single scale (p,q) condition (1.8) for each T_j (in the form

(4.21)); the only difference is the presence of an ℓ^q-sum. More precisely,

$$\begin{aligned}
IV_{1,1}^\flat &\lesssim A_\circ(p,q)\Big(\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}|R|^{-(\frac{1}{p}-\frac{1}{q})q}\Big(\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}}\|b_{1,W}\|_{L^p_{B_1}}^p\Big)^{q/p}\Big)^{1/q}\\
&\lesssim_d A_\circ(p,q)\langle f_1\rangle_{Q_0,p,B_1}\Big(\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}|R|^{-(\frac{1}{p}-\frac{1}{q})q}\Big(\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}}|W|\Big)^{q/p}\Big)^{1/q}\\
&\lesssim_d A_\circ(p,q)\langle f_1\rangle_{Q_0,p,B_1}\Big(\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}\sum_{\substack{W\subset R\\ j-\ell\le L(W)<j}}|W|\Big)^{1/q}\\
&\lesssim \ell^{1/q}\langle f_1\rangle_{Q_0,p,B_1}|Q_0|^{1/q},
\end{aligned}$$

and combining this with (5.17), the bound for $IV_1^\flat$ immediately follows.

Regarding $IV_2^\flat$, write $IV_2^\flat=\sum_{s=\ell+1}^\infty IV_2^\flat(s)$, where $IV_2^\flat(s)$ has the sum in $L(W)<j-\ell$ further restricted to $L(W)=j-s$. For each fixed s, one can apply Hölder's inequality with respect to x and j as in (5.16),

$$IV_2^\flat(s)\le IV_{2,1}^\flat(s)IV_{2,2}^\flat,$$

where the term $IV_{2,2}^\flat$ (which is independent of s) can be treated as $IV_{1,2}^\flat$ in (5.17). For each $IV_{2,1}^\flat(s)$ we write again

$$IV_{2,1}^\flat(s)\lesssim\Big(\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}\Big\|\sum_{\substack{W\subset R\\ L(W)=j-s}}T_jb_{1,W}\Big\|_{L^q_{B_2}}^q\Big)^{1/q}.$$

This term can now be treated as the term IV_2 in (4.30) using the ε-regularity condition (1.9a) to get a decay of $2^{-s\varepsilon'}$ (as in (4.29a)). The only difference with respect to (4.30) is the presence of the ℓ^q-sum, which introduces no difficulty, as shown above for $IV_{1,1}^\flat$. This completes the proof. □

6. Fourier multipliers

In this section we deduce Theorems 1.5 and 1.6 from a more general result which will lead to more precise sparse domination results and also cover Hilbert space valued versions. We are given two separable Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ and denote by $\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)$ the space of bounded linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ (in our applications one of the Hilbert spaces will be usually $\mathbb{C}$). Consider the translation invariant operator $\mathcal{T}=\mathcal{T}_m$ mapping $\mathcal{H}_1$-valued functions to $\mathcal{H}_2$-valued functions given via a multiplier

$$\widehat{\mathcal{T}f}(\xi)=m(\xi)\widehat{f}(\xi),$$

where $m(\xi) \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ for almost every ξ. For $1 \le p \le q \le \infty$ we write $m \in M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}$ if the inequality

$$\|\mathcal{T} f\|_{L^q(\mathcal{H}_2)} \le C\|f\|_{L^p(\mathcal{H}_1)}$$

holds for all $\mathcal{H}_1$-valued Schwartz functions, and the best constant defines the norm in $M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}$. We may occasionally drop the Hilbert spaces if it is understood from the context and also write M^p for $M^{p,p}$. Note that $m \in M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}$ implies by a duality argument that $m \in M^{q',p'}_{\mathcal{H}_2^*,\mathcal{H}_1^*}$. The $M^{2,2}_{\mathcal{H}_1,\mathcal{H}_2}$ norm is bounded by $\|m\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}}$ where we write $L^\infty_{\mathcal{H}_1,\mathcal{H}_2}$ for $L^\infty_{\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)}$. Also note that by the Marcinkiewicz–Zygmund theorem [56, §2.1b] any scalar multiplier in $M^{p,q}$ extends naturally, for any separable Hilbert space $\mathcal{H}$, to a multiplier

$$m \otimes I_{\mathcal{H}} \in M^{p,q}_{\mathcal{H},\mathcal{H}}, \text{ with } m \otimes I_{\mathcal{H}} : \begin{cases} \mathbb{R}^d \to \mathcal{L}(\mathcal{H},\mathcal{H}), \\ \xi \mapsto (v \mapsto m(\xi)v) \end{cases}$$

and we have $\|m \otimes I_{\mathcal{H}}\|_{M^{p,q}_{\mathcal{H},\mathcal{H}}} \le C\|m\|_{M^{p,q}}$ where C does not depend on the Hilbert space.

6.1. **The main multiplier theorem.** In what follows let ϕ be a radial C^∞ function supported in $\{\xi \in \widehat{\mathbb{R}}^d : 1/2 < |\xi| < 2\}$ (not identically zero). Let $\Psi_0 \in C^\infty(\mathbb{R}^d)$ be supported in $\{x \in \mathbb{R}^d : |x| < 1/2\}$ such that $\Psi_0(x) = 1$ for $|x| \le 1/4$. For $\ell > 0$ define

$$\Psi_\ell(x) = \Psi_0(2^{-\ell}x) - \Psi_0(2^{-\ell+1}x) \tag{6.1}$$

which is supported in $\{x : 2^{\ell-3} \le |x| \le 2^{\ell-1}\}$. Define

$$\mathcal{B}[m] := \sum_{\ell \ge 0} \sup_{t>0} \|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} 2^{\ell d(1/p-1/q)}(1+\ell), \tag{6.2a}$$

$$\mathcal{B}_\circ[m] := \sum_{\ell \ge 0} \sup_{t>0} \|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}}. \tag{6.2b}$$

Theorem 6.1. *Let $1 < p \le q < \infty$, $1/q' = 1 - 1/q$, and assume that $m \in L^\infty_{\mathcal{H}_1,\mathcal{H}_2}$ is such that $\mathcal{B}_\circ[m]$ and $\mathcal{B}[m]$ are finite. Then $\mathcal{T}_m \in \mathrm{Sp}(p, \mathcal{H}_1, q', \mathcal{H}_2^*)$ with*

$$\|\mathcal{T}_m\|_{\mathrm{Sp}_\gamma(p,\mathcal{H}_1,q',\mathcal{H}_2^*)} \lesssim_{d,\gamma,p,q} \mathcal{B}[m] + \mathcal{B}_\circ[m].$$

The implicit constant does not depend on $m, \mathcal{H}_1, \mathcal{H}_2$.

We note that the finiteness of $\mathcal{B}_\circ[m]$ is implied by the finiteness of $\mathcal{B}[m]$ in the case $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}$.

Remark 6.2. The function space of all m with $\mathcal{B}_\circ[m] + \mathcal{B}[m] < \infty$ exhibits familiar properties of similarly defined function spaces in multiplier theory. For example:

(1) The space is invariant under multiplication by a standard smooth symbol of order 0. This fact will be used in the proof of Theorem 6.1 and for the convenience of the reader, the precise statement and proof are contained in §C.1 below.
(2) The finiteness of $\mathcal{B}[m]$ and $\mathcal{B}_\circ[m]$ is independent of the choice of the specific functions ϕ and Ψ. This observation will be convenient in the proof of Theorem 6.1. It can be verified by standard arguments but, for completeness, the proof is provided in C.2 below.

We begin by showing how Theorem 6.1 implies Theorems 1.5 and 1.6. Then we review some known facts and estimates for Fourier multipliers and deduce the proof of Theorem 6.1 from our main Theorem 1.1.

Proof of Theorem 1.5 using Theorem 6.1. We have to check the assumptions of Theorem 6.1. Assumption (1.24) is equivalent with

$$\|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^2} \leq 2^{-\ell\varepsilon}.$$

Thus interpolating (1.24) and (1.23) we get for $p \in (p_0, 2)$,

$$\|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^p} \leq 2^{-\ell\varepsilon(p)} \text{ where } \varepsilon(p) = \varepsilon\big(\tfrac{1}{p_0} - \tfrac{1}{p}\big)\big/\big(\tfrac{1}{p_0} - \tfrac{1}{2}\big).$$

Let $\chi \in C_c^\infty(\widehat{\mathbb{R}}^d \setminus \{0\})$ so that $\chi(\xi) = 1$ in a neighborhood of $\operatorname{supp}(\phi)$. Then by Young's convolution inequality for all $p \in (p_0, 2)$, $q \in [p, \infty]$, $t > 0$,

$$\|\chi([\phi m(t\cdot)] * \widehat{\Psi}_\ell)\|_{M^{p,q}} \lesssim \|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^p} \lesssim 2^{-\ell\varepsilon(p)}.$$

On the other hand we claim that

$$\|(1-\chi)([\phi m(t\cdot)] * \widehat{\Psi}_\ell)\|_{M^{p,q}} \lesssim_N \|\varphi m(t\cdot)\|_1 2^{-\ell N}. \tag{6.3}$$

Indeed, integration by parts in ξ in the integral

$$\int\int e^{ix\cdot\xi}(1-\chi)(\xi)[\phi m(t\cdot)](\zeta) 2^{\ell d}\widehat{\Psi}(2^\ell(\xi-\zeta)) d\xi d\zeta$$

implies the pointwise estimate

$$|\mathcal{F}^{-1}\big((1-\chi)([\phi m(t\cdot)] * \widehat{\Psi}_\ell)\big)(x)| \lesssim_N 2^{-\ell N}(1+|x|)^{-N}\|\phi m(t\cdot)\|_1,$$

whence (6.3) follows from Young's convolution inequality.

Fix $p \in (p_0, 2)$. Combining the two estimates we see that condition (6.2a) holds for a pair of exponents (p_1, q_1) if $p_1 \in (p_0, p)$, $q_1 > p_1$ and

$$d(1/p_1 - 1/q_1) < \varepsilon(p_1).$$

Then Theorem 6.1 gives $\mathcal{T} \in \operatorname{Sp}(p_1, q_1')$. One can then choose $\delta = \delta(p) > 0$ small enough so that $p_1 = p - \delta$ and $q_1' = p' - \delta$ satisfy the above conditions. This concludes the proof. □

Proof of Theorem 1.6 using Theorem 6.1. We need to check that $\mathcal{B}[m] < \infty$, which will follow from showing that

$$\sup_{t>0} \|[\phi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim 2^{-\ell s} \tag{6.4}$$

for some $s > d(1/p - 1/q)$. Here we are in the case $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}$, so this also implies $\mathcal{B}_\circ[m] < \infty$.

We decompose Ψ_ℓ into slighly smaller pieces. Recall that Ψ_1 is supported in $\{x : 1/4 \le |x| \le 1\}$ and $\Psi_\ell(x) = \Psi_1(2^{1-\ell}x)$. We form a partition of unity $\{\varsigma_\nu : \nu \in \mathcal{I}\}$ such that $\sum_{\nu\in\mathcal{I}} \varsigma_\nu(x) = 1$ for $|x| \in [1/8, 2]$, and ς is a C^∞ function supported in a ball $B(x_\nu, r_\nu)$ centered at x_ν, with $|x_\nu| \in [1/4, 1]$ and radius $r_\nu \le 10^{-2}$. Let

$$u_\nu = \frac{\pi}{2}\frac{x_\nu}{|x_\nu|^2}$$

so that $|u_\nu| \in [1, 8]$ and $\langle x_\nu, u_\nu\rangle = \pi/2$. This implies that $|\operatorname{Im}(e^{i\langle x,u_\nu\rangle}-1)| > 1/2$ for $x \in \operatorname{supp}(\varsigma_\nu)$. Define, for M as in (1.25),

$$\Psi_{1,\nu}(x) = \frac{\Psi_1(x)\varsigma_\nu(x)}{(e^{i\langle x,u_\nu\rangle} - 1)^M}, \quad \Psi_{\ell,\nu}(x) = \Psi_{1,\nu}(2^{1-\ell}x)$$

and note that $\Psi_{1,\nu}$ is smooth and $\Psi_\ell(x) = \sum_\nu \Psi_{\ell,\nu}(x)(e^{i\langle x,2^{1-\ell}u_\nu\rangle} - 1)^M$. Hence

$$\phi m(t\cdot) * \widehat{\Psi}_\ell = \sum_\nu \Delta^M_{-2^{1-\ell}u_\nu}[\phi m(t\cdot)] * \widehat{\Psi}_{\ell,\nu}$$

and by assumption we have for some $s > d(1/p - 1/q)$,

$$\|\phi m(t\cdot) * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim \sum_\nu \|\Delta^M_{-2^{1-\ell}u_\nu}[\phi m(t\cdot)]\|_{M^{p,q}} \lesssim 2^{-\ell s}.$$

This implies (6.4) and now Theorem 1.6 follows from Theorem 6.1. □

6.2. **A result involving localizations of Fourier multipliers.** We recall a theorem from [98] (see also [24] for a similar result) which we will formulate in the vector-valued version (see also [51]).

Let ϕ be as before, and fix $1 < p < \infty$. Assume

$$\sup_{t>0} \|\phi m(t\cdot)\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}} \le \mathfrak{a} \tag{6.5a}$$

$$\sup_{t>0} \|\phi m(t\cdot)\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}} \le \mathfrak{a}_\circ, \tag{6.5b}$$

and

$$\sum_{|\alpha|\le d+1} \sup_{t>0} \sup_{\xi\in\widehat{\mathbb{R}}^d} |\partial^\alpha_\xi(\phi m(t\cdot))(\xi)|_{\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)} \le \mathfrak{b}, \tag{6.5c}$$

where $\alpha \in \mathbb{N}_0^d$. Then

$$\|m\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}} \lesssim \mathfrak{a}_\circ + \mathfrak{a}\log(2 + \mathfrak{b}/\mathfrak{a})^{|\frac{1}{p}-\frac{1}{2}|}. \tag{6.6}$$

Of course, in the special case $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}$ the L^2-boundedness condition (6.5b) with $\mathfrak{a}_\circ \le \mathfrak{a}$ is implied by (6.5a) (cf. an analogous remark following Theorem 6.1).

6.3. **Proof of Theorem 6.1.** First assume that m is compactly supported in $\widehat{\mathbb{R}}^d \setminus \{0\}$ without making any quantitative assumption on the support.

Note that by Remark 6.2 we have some freedom to make a convenient choice of the localizing function ϕ, and we will denote this choice by φ. In what follows, let $\theta \in C_c^\infty(\mathbb{R}^d)$ be radial such that θ is supported in $\{x \in \mathbb{R}^d : |x| < 1/2\}$, such that $\int \theta(x)\pi(x)dx = 0$ for all polynomials π of degree at most $10d$, and such that $\widehat{\theta}(\xi) > 0$ for $1/4 \le |\xi| \le 4$. We then choose φ to be a radial C^∞ function supported in $\{\xi \in \widehat{\mathbb{R}}^d : 1/2 < |\xi| < 2\}$ such that

$$\sum_{k\in\mathbb{Z}} \varphi(2^{-k}\xi)\widehat{\theta}(2^{-k}\xi) = 1$$

for all $\xi \neq 0$.

We then decompose $\mathcal{T}$ by writing

$$m(\xi) = \sum_{k=n_1}^{n_2} \widehat{\theta}(2^{-k}\xi)\varphi(2^{-k}\xi)m(\xi)$$

where $n_1, n_2 \in \mathbb{Z}$. We then decompose

$$\mathcal{F}^{-1}[\varphi m(2^k\cdot)](x) = \sum_{\ell\ge 0} \mathcal{F}^{-1}[\varphi m(2^k\cdot)](x)\Psi_\ell(x)$$

which yields

$$\mathcal{F}^{-1}[m\widehat{f}](x) = \sum_{\ell\ge 0} \mathcal{T}^\ell f(x) = \sum_{\ell\ge 0}\sum_{k=n_1}^{n_2} \mathcal{T}^{\ell,k} f(x)$$

where

$$\widehat{\mathcal{T}^{\ell,k} f}(\xi) = \widehat{\theta}(2^{-k}\xi)\ [\varphi m(2^k\cdot)]*\widehat{\Psi}_\ell(2^{-k}\xi).$$

We can write $\mathcal{T}^{\ell,k} f = K_k^\ell * f$ with

$$K_k^\ell(x) = \int \mathcal{F}^{-1}[\varphi(2^{-k}\cdot)m](x-y)\Psi_\ell(2^k(x-y))2^{kd}\theta(2^k y)\,dy.$$

Observe that $K_k^\ell(x)$ is supported in $\{x \in \mathbb{R}^d : |x| \le 2^{\ell+1-k}\}$. We wish to apply Theorem 1.1 to the operators $\mathcal{T}^\ell$ defined by

$$\mathcal{T}^\ell f = \sum_{k=n_1}^{n_2} K_k^\ell * f = \sum_{j=\ell+1-n_2}^{\ell+1-n_1} T_j^\ell f \quad \text{with } T_j^\ell f = K_{\ell+1-j}^\ell * f.$$

The operators T_j^ℓ satisfy the support condition (1.6). To check the conditions (1.7a), (1.7b) we apply the above mentioned theorem from [98] (see (6.6)). We first claim that

$$\|\varphi \sum_{k=n_1}^{n_2} \widehat{K_k^\ell}(s\cdot)\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}} \lesssim \mathfrak{a}_\ell := \sup_{t>0} \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} 2^{\ell d(1/p-1/q)} \quad (6.7)$$

and

$$\|\varphi \sum_{k=n_1}^{n_2} \widehat{K_k^\ell}(s\cdot)\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}} \lesssim \mathfrak{a}_{\circ,\ell} := \sup_{t>0} \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}}, \tag{6.8}$$

uniformly in n_1, n_2. We only give the proof of (6.7) as the the proof of (6.8) is similar but more straightforward. To see this we estimate, using dilation invariance,

$$\|\varphi \sum_{k=n_1}^{n_2} \widehat{K_k^\ell}(s\cdot)\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}} \le \sum_{k=n_1}^{n_2} \|\varphi\widehat{\theta}(2^{-k}s\cdot)\|_{M^p} \|[\varphi m(2^k\cdot)]*\widehat{\Psi}_\ell\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}}.$$

Since $\theta \in \mathcal{S}(\mathbb{R}^d)$ and since all moments of η up to order $10d$ vanish we get

$$\|\varphi\widehat{\theta}(2^{-k}s\cdot)\|_{M^p} \lesssim \min\{(2^{-k}s)^{10d}, (2^{-k}s)^{-10d}\}. \tag{6.9}$$

Moreover,

$$\|[\varphi m(2^k\cdot)]*\widehat{\Psi}_\ell\|_{M^p_{\mathcal{H}_1,\mathcal{H}_2}} \lesssim 2^{\ell d(1/p-1/q)} \|[\varphi m(2^k\cdot)]*\widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} \tag{6.10}$$

and (6.7) follows combining the above. To verify (6.10) we decompose

$$f = \sum_\nu f_\nu,$$

where $f_\nu = f\mathbb{1}_{R_{\ell,\nu}}$ and the $R_{\ell,\nu}$ form a grid of cubes of side length 2^ℓ. Note that the convolution kernel $\mathcal{K}_\ell := \mathcal{F}^{-1}[\varphi m(2^k\cdot)*\widehat{\Psi}_\ell]$ is supported in the ball of radius 2^ℓ centered at the origin. Hence, by Hölder's inequality

$$\begin{aligned}
\|\mathcal{K}_\ell * f\|_{L^p_{\mathcal{H}_2}} &= \Big\| \sum_\nu \mathcal{K}_\ell * f_\nu \Big\|_{L^p_{\mathcal{H}_2}} \lesssim \Big(\sum_\nu \|\mathcal{K}_\ell * f_\nu\|^p_{L^p_{\mathcal{H}_2}} \Big)^{1/p} \\
&\lesssim 2^{\ell d(1/p-1/q)} \Big(\sum_\nu \|\mathcal{K}_\ell * f_\nu\|^p_{L^q_{\mathcal{H}_2}} \Big)^{1/p} \\
&\lesssim 2^{\ell d(1/p-1/q)} \|\widehat{\mathcal{K}_\ell}\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} \Big(\sum_\nu \|f_\nu\|^p_{L^p_{\mathcal{H}_1}} \Big)^{1/p}
\end{aligned}$$

and since $(\sum_\nu \|f_\nu\|^p_{L^p_{\mathcal{H}_1}})^{1/p} = \|f\|_{L^p_{\mathcal{H}_1}}$ we get (6.10).

Straightforward calculation using (6.9) yields

$$\sum_{|\alpha|\le d+1} \sum_{k=n_1}^{n_2} \sup_{t>0} \sup_{\xi\in\hat{\mathbb{R}}^d} |\partial_\xi^\alpha(\varphi\widehat{K_k^\ell}(t\xi))|_{\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)} \le \mathfrak{b}_\ell := \|m\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}} 2^{\ell(d+1)} \tag{6.11}$$

uniformly in n_1, n_2. We combine the two estimates (6.7), (6.8) and (6.11) and using (6.6) we get

$$\|\mathcal{T}^\ell\|_{L^p_{\mathcal{H}_1}\to L^p_{\mathcal{H}_2}} \lesssim (1+\ell)^{|\frac{1}{p}-\frac{1}{2}|}\mathfrak{a}_\ell + \mathfrak{a}_{\circ,\ell}.$$

The L^q estimates are similar. For $m(\xi) \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denote by $m^*(\xi) \in \mathcal{L}(\mathcal{H}_2^*, \mathcal{H}_1^*)$ the adjoint. Note that

$$\|[\varphi m^*(t\cdot)] * \widehat{\Psi}_\ell\|_{L^\infty_{\mathcal{H}_2^*,\mathcal{H}_1^*}} = \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}} \le \mathfrak{a}_{\circ,\ell}.$$

Since $q' \le p'$ the previous calculation gives

$$\begin{aligned}
&\|\mathcal{T}^\ell\|_{L^q_{\mathcal{H}_1}\to L^q_{\mathcal{H}_2}} = \|(\mathcal{T}^\ell)^*\|_{L^{q'}_{\mathcal{H}_2^*}\to L^{q'}_{\mathcal{H}_1^*}} \\
&\lesssim (1+\ell)^{|\frac{1}{q'}-\frac{1}{2}|} 2^{\ell d(1/q'-1/p')} \sup_{t>0} \|[\varphi m^*(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{q',p'}_{\mathcal{H}_2^*,\mathcal{H}_1^*}} + \mathfrak{a}_{\circ,\ell} \\
&\lesssim (1+\ell)^{|\frac{1}{q}-\frac{1}{2}|} 2^{\ell d(1/p-1/q)} \sup_{t>0} \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} + \mathfrak{a}_{\circ,\ell} \\
&= (1+\ell)^{|\frac{1}{q}-\frac{1}{2}|} \mathfrak{a}_\ell + \mathfrak{a}_{\circ,\ell}.
\end{aligned}$$

To summarize,

$$\begin{aligned}
&\Big\|\sum_j T_j^\ell\Big\|_{L^p\to L^p} + \Big\|\sum_j T_j^\ell\Big\|_{L^q\to L^q} \\
&\qquad\qquad \le \mathfrak{a}_{\circ,\ell} + \mathfrak{a}_\ell((1+\ell)^{|\frac{1}{p}-\frac{1}{2}|} + (1+\ell)^{|\frac{1}{q}-\frac{1}{2}|}). \qquad (6.12)
\end{aligned}$$

To verify the single scale (p,q) condition (1.8) we next examine the $L^p_{\mathcal{H}_1} \to L^q_{\mathcal{H}_2}$-norms of the convolution operators $\mathrm{Dil}_{2^j} T_j^\ell$ with convolution kernels $2^{jd} K^\ell_{\ell+1-j}(2^j\cdot)$. We have

$$\|\mathrm{Dil}_{2^j} T_j^\ell\|_{L^p_{\mathcal{H}_1}\to L^q_{\mathcal{H}_2}} = \|\widehat{K^\ell_{\ell+1-j}}(2^{-j}\cdot)\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}}$$

and

$$\widehat{K^\ell_{\ell+1-j}}(2^{-j}\xi) = \widehat{\theta}(2^{-\ell-1}\xi)[\varphi m(2^{\ell+1-j}\cdot)] * \widehat{\Psi}_\ell(2^{-\ell-1}\xi),$$

and we get

$$\|\widehat{K^\ell_{\ell+1-j}}(2^{-j}\cdot)\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} \le \|\theta\|_1 2^{(\ell+1)d(1/p-1/q)} \|[\varphi m(2^{\ell+1-j}\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}}.$$

Hence

$$\sup_j \|\mathrm{Dil}_{2^j} T_j^\ell\|_{L^p\to L^q} \lesssim \mathfrak{a}_\ell \qquad (6.13)$$

Next we turn to the ε-regularity conditions (1.9a) and (1.9b). By translation invariance of the operators T_j^ℓ it suffices to verify (1.9a). Using the above formulas for the Fourier transform of $2^{jd} K^\ell_{\ell+1-j}(2^j\cdot)$ we get

$$\begin{aligned}
&\|(\mathrm{Dil}_{2^j} T_j^\ell) \circ \Delta_h\|_{L^p_{\mathcal{H}_1}\to L^q_{\mathcal{H}_2}} \\
&= \|\widehat{\theta}(2^{-\ell-1}\cdot)[(\varphi m(2^{\ell+1-j}\cdot)) * \widehat{\Psi}_\ell](2^{-\ell-1}\cdot)(e^{i\langle\cdot,h\rangle}-1)\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}} \\
&\le 2^{(\ell+1)d(1/p-1/q)} \big\|\widehat{\theta}[e^{i\langle 2^\ell\cdot,h\rangle}-1]\big\|_{M^p} \big\|[\varphi m(2^{\ell+1-j}\cdot)] * \widehat{\Psi}_\ell\big\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}}.
\end{aligned}$$

Observe that for $0 < \varepsilon < 1$,

$$|h|^{-\varepsilon} \big\|\widehat{\theta}[e^{i\langle 2^\ell\cdot,h\rangle}-1]\big\|_{M^p} \lesssim 2^{\ell\varepsilon}$$

and hence we get

$$\sup_{|h|\le 1} |h|^{-\varepsilon} \sup_j \|(\mathrm{Dil}_{2^j} T_j^\ell)\circ \Delta_h\|_{L^p_{\mathcal{H}_1}\to L^q_{\mathcal{H}_2}} \lesssim 2^{\ell\varepsilon}\mathfrak{a}_\ell \tag{6.14}$$

In view of (6.12), (6.13) and (6.14) we can now apply Theorem 1.1 and obtain

$$\|\mathcal{T}^\ell\|_{\mathrm{Sp}(p,q)} \lesssim \sup_{t>0} \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{L^\infty_{\mathcal{H}_1,\mathcal{H}_2}} +$$
$$\big((1+\ell)^{|\frac{1}{p}-\frac{1}{2}|} + (1+\ell)^{|\frac{1}{q}-\frac{1}{2}|} + (1+\ell)\big) 2^{\ell d(\frac{1}{p}-\frac{1}{q})} \sup_{t>0} \|[\varphi m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}_{\mathcal{H}_1,\mathcal{H}_2}}.$$

The desired conclusionthen follows from summing in $\ell \ge 0$.

Finally, to remove the assumption of m being compactly supported we observe that by Lemma A.1 it suffices to prove the sparse bound

$$\int_{\mathbb{R}^d} \mathcal{F}^{-1}[m\widehat{f_1}](x) f_2(x)dx \le C\big(\mathcal{B}_\circ[m] + \mathcal{B}[m]\big)\Lambda^*_{p_1,p_2}(f_1,f_2) \tag{6.15}$$

for f_i in the dense class $\mathcal{S}_0(\mathbb{R}^d,\mathcal{H}_i)$ of functions whose Fourier transform is compactly supported in $\widehat{\mathbb{R}}^d\backslash\{0\}$. But for those functions we have $\mathcal{F}^{-1}[m\widehat{f_1}] = \mathcal{F}^{-1}[m_{n_1,n_2}\widehat{f_1}]$, where

$$m_{n_1,n_2} = \sum_{k=n_1}^{n_2} \widehat{\theta}(2^{-k}\xi)\varphi(2^{-k}\xi)m(\xi)$$

with suitable $n_1, n_2 \in \mathbb{Z}$ (depending on f_1). By invariance under multiplication by smooth symbols (see Lemma C.3) we have

$$\sup_{n_1,n_2} \mathcal{B}[m_{n_1,n_2}] \lesssim \mathcal{B}[m]$$

and an analogous inequality involving $\mathcal{B}_\circ[m]$. We then get (6.15) for $f_i \in \mathcal{S}_0(\mathbb{R}^d,\mathcal{H}_i)$, $i = 1,2$. A second application of Lemma A.1 yields (6.15) for all $f \in L^p_{\mathcal{H}_1}$ and all $f_2 \in L^{p'}_{\mathcal{H}_2^*}$. □

7. Sample applications

In this section we give a number of specific examples of operators to which Theorem 1.1 and its consequences can be applied. Some of the resulting sparse bounds are well-known and others appear to be new.

7.1. **Operators generated by compactly supported distributions.** In what follows let σ be a distribution which is compactly supported and let $\sigma_t \equiv \mathrm{Dil}_{1/t}\sigma$ denote the t-dilate $t^{-d}\sigma(t^{-1}\cdot)$ given by

$$\langle \sigma_t, f\rangle = \langle \sigma, f(t\cdot)\rangle.$$

Without loss of generality we may assume that the support of σ is contained in $\{x : |x| \le 1\}$, otherwise argue with a rescaling.

Let

$$A_t f(x) = f * \sigma_t(x) \tag{7.1}$$

which is well defined on Schwartz functions as a continous function of (x,t). Many interesting operators in harmonic analysis are generated by dilations of such a single compactly supported distribution (often a measure) and we shall be interested in the corresponding maximal and variational operators. The domain of the dilation parameter t will be either $(0,\infty)$ or $[1,2]$ or a more general subset E of $(0,\infty)$.

7.1.1. *Maximal functions.* We are interested in sparse domination results for the maximal functions, as defined in (1.15),

$$M_E^\sigma f(x) = \sup_{t\in E} |A_t f(x)|$$

where $E \subset (0,\infty)$.

If we assume that f is a Schwartz function then M_E^σ is well defined as a measurable function, but for general L^p functions the measurability of M_E^σ is a priori not clear unless we assume that E is countable. In our statements we will restrict ourselves to a priori estimates, but note that in many applications the proof of L^p bounds shows also a priori estimates for the function $t \mapsto \sigma_t * f(x)$ in suitable subspaces of $C(\mathbb{R})$, for almost all $x \in \mathbb{R}^d$. This observation then ensures the measurability of the maximal functions for f in the relevant L^p classes. In the general case, let $I_{k,n} = [k2^{-n}, (k+1)2^{-n})$ and pick, for each (k,n) such that $E \cap I_{k,n} \neq \emptyset$, a representative $t_{k,n} \in E \cap I_{k,n}$ and let $\widetilde{E}$ consist of these picked $t_{k,n}$. Then $\widetilde{E}$ is countable and we have $M_E^\sigma f(x) = M_{\widetilde{E}}^\sigma f(x)$ for all $x \in \mathbb{R}^d$ and all Schwartz functions f. Thus one can assume that E is countable without loss of generality.

We shall now discuss sparse domination inequalities for the operator M_E^σ. Recall the local variants $M_{E_j}^\sigma$, with the rescaled sets $E_j \subset [1,2]$ as in (1.16). In what follows recall that

$$\Lambda_{p,q'}^{\mathfrak{S}}(f,\omega) = \sum_{Q\in\mathfrak{S}} |Q| \langle f\rangle_{Q,p} \langle \omega\rangle_{Q,q'},$$

with $\Lambda_{p,q'}^*(f,\omega)$ the supremum of all $\Lambda_{p,q'}^{\mathfrak{S}}(f,\omega)$ over all γ-sparse families $\mathfrak{S}$.

Proposition 7.1. *(i) Let $1 < p \le q < \infty$. Let σ be a compactly supported distribution such that*

$$\|M_E^\sigma\|_{L^p\to L^{p,\infty}} + \|M_E^\sigma\|_{L^{q,1}\to L^q} < \infty, \tag{7.2}$$

$$\sup_{j\in\mathbb{Z}} \|M_{E_j}^\sigma\|_{L^p\to L^q} < \infty, \tag{7.3}$$

and assume that there is an $\epsilon > 0$ so that for all $\lambda \ge 2$,

$$\|M_{E_j}^\sigma f\|_q \le C\lambda^{-\varepsilon}\|f\|_p, \quad f \in \mathcal{E}_{\mathrm{ann}}(\lambda). \tag{7.4}$$

Then for all $f \in L^p$ and all simple non-negative functions ω, we have the sparse domination inequality

$$\langle M_E^\sigma f, \omega\rangle \lesssim \Lambda_{p,q'}^*(f,\omega). \tag{7.5}$$

(ii) Conversely, if σ has compact support in $\mathbb{R}^d \setminus \{0\}$ then the sparse bound (7.5) *for $p < q$ implies that conditions* (7.2) *and* (7.3) *hold.*

Proof. We will apply Theorem 5.1 with $r = \infty$, $B_2 = \ell^\infty(E')$, where E' is a finite subset of E, and

$$T_j f(x,t) = \begin{cases} \sigma_t * f(x) & \text{if } t \in E' \cap [2^j, 2^{j+1}) \\ 0 & \text{otherwise.} \end{cases} \tag{7.6}$$

Note that

$$S_\infty T f(x) \equiv \sup_{j\in\mathbb{Z}} |T_j f(x)|_{B_2} = M^\sigma_{E'} f(x), \tag{7.7a}$$

and, with $E'_j = 2^{-j}E' \cap [1,2]$,

$$|\mathrm{Dil}_{2^j} T_j f(x)|_{B_2} = M^\sigma_{E'_j} f(x), \qquad j \in \mathbb{Z}. \tag{7.7b}$$

As σ is supported in $\{x : |x| \le 1\}$, the operators T_j satisfy the support condition (1.6). Moreover, (7.2) and (7.7a) guarantee (5.3) with $r = \infty$, and similarly (7.3) and (7.7b) guarantee the single scale (p,q) condition (1.8). It remains to verify the single scale $\varepsilon-$regularity conditions (1.9) for the operators T_j. But this follows from (7.4) and (7.3) via Lemma 3.6 and the fact that for translation-invariant operators T_j, the conditions (1.9a) and (1.9b) are equivalent (alternatively, one can apply Theorem 5.5 for maximal functions). All hypotheses in the first part of Theorem 5.1 are then satisfied and we thus obtain a sparse bound for the maximal operator $M^\sigma_{E'}$. An application of the monotone convergence theorem then yields the desired sparse bound for M^σ_E and concludes the proof of part (i).

For part (ii) note that the assumption that σ is supported away from the origin corresponds to the strengthened support condition (1.14). Thus we can deduce part (ii) directly from part (ii) of Theorem 5.1. □

Proof of Theorem 1.4. Because of the $L^p \to L^q$ condition on the operators $M^\sigma_{E_j}$ in (1.17) and the ε-regularity assumption (1.18), it follows by interpolation that the condition (7.4) is satisfied for all $(1/p, 1/q)$ in the interior of $\mathcal{L}(\sigma, E)$. Thus Proposition 7.1 establishes the sufficiency of the conditions, that is, (1.20). The converse follows immediately from part (ii) of Proposition 7.1. □

Prototypical examples for Proposition 7.1 are the spherical maximal functions where σ is the surface measure on the sphere (for L^p bounds see the classical results by Stein [105] and Bourgain [19], and for $L^p \to L^q$ bounds see [95, 96]). The proposition covers the results by Lacey [66] for the lacunary and full spherical maximal functions and also the extension to spherical maximal operators with suitable assumptions about various fractal dimensions of E, see [101, 3, 93]. In this context we note that in [5, 45], Lacey's approach was used to establish sparse domination results for two versions

of lacunary spherical maximal functions on the Heisenberg group, defined via the automorphic dilations, and essentially optimal results for the problem considered in [5] can be obtained by combining the sparse technique developed in that paper with recent $L^p \to L^q$ bounds in [94].

One can also cover more singular measures with Fourier conditions (as in [39], [40]) and this leads to questions about the precise range of L^p improving estimates for the local variants of the maximal functions. As an example consider a curve $s \mapsto \gamma(s)$ in $\mathbb{R}^3$ with nonvanishing curvature and torsion, and the measures μ_t given by

$$\langle f, \mu_t \rangle = \int f(t\gamma(s))\chi(s)ds$$

with compactly supported χ. A result in [92], applied in combination with decoupling results in [111, 20] yields that the maximal operators M_E are bounded on $L^p(\mathbb{R}^3)$ for $p > 4$. The optimal result for $p > 3$ was recently obtained in [12] and in [62]. Moreover, the analysis in these papers yield, for the local analogues of these maximal functions (i.e. $E = [1, 2]$), certain $L^p \to L^q$ bounds for some $q > p$. It would be very interesting to find precise ranges of $L^p \to L^q$ boundedness of M_E depending on E, and corresponding sparse bounds for related global maximal functions. Similar questions can be considered in higher dimensions but the optimal bounds are currently unknown (for partial results see [13], [63]).

7.1.2. *Variational operators.* Given $1 \le r \le \infty$ and a set $E \subset (0, \infty)$ we define the r-variation seminorm $|\cdot|_{\mathrm{v}^r(E)}$ and the r-variation norm $|\cdot|_{V^r(E)}$ of a function $a : E \to \mathbb{C}$ by

$$|a|_{\mathrm{v}^r(E)} = \sup_{M\in\mathbb{N}} \sup_{\substack{t_1<\cdots<t_M \\ t_i\in E}} \Big(\sum_{i=1}^{M-1} |a(t_{i+1}) - a(t_i)|^r \Big)^{1/r}$$

$$|a|_{V^r(E)} = \sup_{M\in\mathbb{N}} \sup_{\substack{t_1<\cdots<t_M \\ t_i\in E}} \Big\{ |a(t_1)| + \Big(\sum_{i=1}^{M-1} |a(t_{i+1}) - a(t_i)|^r \Big)^{1/r} \Big\}.$$

Define the r-variation operators $\mathrm{v}^r_E A$, $\mathcal{V}^r_E A$ for the family of operators of convolution with σ_t by taking the r-variation norm in t,

$$\mathrm{v}^r_E Af(x) := |\{\sigma_t * f(x)\}|_{\mathrm{v}^r(E)}, \qquad \mathcal{V}^r_E Af(x) := |\{\sigma_t * f(x)\}|_{V^r(E)}. \tag{7.8}$$

This that the above definition of variation is analogous to the definition in (5.7) where we considered the r-variation for functions integers. The results in §5 mostly apply to the situation where the current E is a subset of $\{2^j : j \in \mathbb{Z}\}$. For general sets $E \subset (0, \infty)$, we will deduce results directly from Corollary 1.2 and Theorem 1.3.

As before we may assume that E is countable (as this does not affect priori estimates). Let $E_j \subset [1, 2]$ be the rescaled sets as in (1.16).

Proposition 7.2. *(i) Let $1 < p \le q < \infty$. Let σ be a compactly supported distribution such that*

$$\|\mathcal{V}^r_E A\|_{L^p\to L^{p,\infty}} + \|\mathcal{V}^r_E A\|_{L^{q,1}\to L^q} < \infty, \tag{7.9}$$

$$\sup_{j\in\mathbb{Z}} \|\mathcal{V}^r_{E_j} A\|_{L^p\to L^q} < \infty \tag{7.10}$$

and assume that there is an $\epsilon > 0$ so that for all $\lambda \ge 2$,

$$\|\mathcal{V}^r_{E_j} f\|_q \le C\lambda^{-\varepsilon}\|f\|_p, \quad f \in \mathcal{E}_{\mathrm{ann}}(\lambda). \tag{7.11}$$

Then for all $f \in L^p$ and all simple nonnegative functions ω, we have the sparse domination inequality

$$\langle \mathcal{V}^r_E f, \omega\rangle \lesssim \Lambda^*_{p,q'}(f,\omega). \tag{7.12}$$

(ii) Conversely, if σ has compact support in $\mathbb{R}^d \setminus \{0\}$ then the sparse bound (7.12) *for $p < q$ implies that conditions* (7.9) *and* (7.10) *hold.*

Proof. We are aiming to apply Corollary 1.2 with $B_2 = V^r(E')$ for any finite $E' \subset E$. With $T_j f(x,t)$ as in (7.6) and $E'(N_1,N_2) = E' \cap [2^{N_1}, 2^{N_2+1}]$, we get

$$\Big|\sum_{j=N_1}^{N_2} T_j f(x)\Big|_{V^r_{E'}} = \mathcal{V}^r_{E'(N_1,N_2)} Af(x) \tag{7.13a}$$

and

$$|\mathrm{Dil}_{2^j} T_j f(x)|_{V^r_{E'}} = \mathcal{V}^r_{2^{-j}E'\cap[1,2]} Af(x). \tag{7.13b}$$

We need to check the assumptions of Corollary 1.2 (i.e. the assumptions of Theorem 1.1). Conditions (1.7a), (1.7b) hold by (7.9) and (7.13a), condition (1.8) holds by (7.10) and (7.13b) and condition (1.9a) follows from (7.10), (7.11), and Lemma 3.6. Condition (1.9a) is equivalent with (1.9b) in the current translation invariant setting.

For the necessity, observe that the assumption that σ is supported away from the origin which corresponds to the strengthened support condition in Theorem 1.3. A sparse bound for $V^r_E A$ implies via (7.13a) a sparse bound for $\sum_{j=N_1}^{N_2} T_j$ for any pair of integers $N_1 \le N_2$. We apply Theorem 1.3 and obtain via (7.13a) and (7.13b) that

$$\|\mathcal{V}^r_{E'(N_1,N_2)} A\|_{L^p\to L^{p,\infty}} + \|\mathcal{V}^r_{E'(N_1,N_2)} A\|_{L^{q,1}\to L^q} \le C,$$
$$\sup_{N_1\le j\le N_2} \|\mathcal{V}^r_{2^{-j}E'\cap[1,2]} A\|_{L^p\to L^q} \le C,$$

with the constant C independent of N_1, N_2 and the particular finite subset E' of E. Applications of the monotone convergence theorem then yield the asserted necessary conditions for $\mathcal{V}^r_E A$, that is, (7.9) and (7.10). □

Proposition 7.2 can be applied to obtain a sparse domination inequality for the r-variation operator associated with the spherical means in $\mathbb{R}^d$. For the necessary global $L^p \to L^p$ bounds see [59] and for $L^p \to L^q$ bounds for

the local variation operators we refer to the recent paper [14]. This addresses a question posed in [66] and [1].

Remark 7.3. In verifying $L^p \to L^{p,\infty}$ and $L^{q,1} \to L^q$ assumptions for the variation operators it is (as shown in [58, 59]) often advantageous to write $V_E^r Af(x) \le V_{\mathrm{dyad}}^r Af(x) + V_{E,\mathrm{sh}}^r Af(x)$ where

$$V_{\mathrm{dyad}}^r Af(x) := \mathcal{V}_{2^{(\mathbb{Z})}}^r Af(x)$$

is the standard variation norm over $2^{(\mathbb{Z})} := \{2^j : j \in \mathbb{Z}\}$, labeled the *dyadic* or *long variation operator* and where

$$V_{E,\mathrm{sh}}^r Af(x) := \Big(\sum_{j\in\mathbb{Z}} |\mathrm{v}_{E\cap[2^j,2^{j+1}]}^r Af(x)|^r\Big)^{1/r},$$

is the so-called *short variation operator* which uses only variation seminorms over E within dyadic intervals. The L^p-boundedness of the long variation operators is usually reduced to Lepingle's theorem [73] (which requires $r > 2$) while the short variation operator is often estimated using a Sobolev embedding inequality (see [58], [59]). We note that it is possible to prove results analogous to Proposition 7.2 for the long variation operator and the short variation operators individually as direct consequences of Theorems 5.2 and 5.1 respectively; the details are left to the reader.

7.1.3. *Lacunary maximal functions for convolutions associated with the wave equation.* In this section we consider a maximal function generated by convolutions with dilates of a tempered distribution, which is not compactly supported (but still concentrated on a compact set). This class is associated with L^p regularity results for solutions of the wave equation. For both simplicity and definiteness of results we shall only consider a lacunary version, but the argument to deduce the sparse bound extends to other sets of dilations and also to variational variants (for which Lemma A.3 would be useful to treat nonlocal error terms).

For $\beta > 0$ define

$$m_\beta(\xi) = \frac{\cos|\xi|}{(1+|\xi|^2)^{\beta/2}}$$

and let

$$\mathcal{M}_{\mathrm{lac}}^\beta f(x) = \sup_{k\in\mathbb{Z}} |m_\beta(2^k D)f(x)|.$$

It was shown by Peral in [90] and Miyachi in [87] that $m_\beta(D)$ is bounded on L^p for $\beta \ge (d-1)|1/p - 1/2|$, $1 < p < \infty$. $L^p \to L^q$ results for m_β go back to [108, 83, 21]; it is known that $m_\beta(D) : L^p \to L^q$ is bounded if either

(W) $1 < p \le 2$, $p \le q \le p'$, $\beta \ge (d-1)\big(\frac{1}{p} - \frac{1}{2}\big) + \frac{1}{p} - \frac{1}{q}$, or
(W′) $1 < p < \infty$, $\max\{p, p'\} \le q < \infty$, $\beta \ge (d-1)\big(\frac{1}{2} - \frac{1}{q}\big) + \frac{1}{p} - \frac{1}{q}$.

Note that (W′) follows from (W) by duality. Moreover it can be shown that $\mathcal{M}_{\mathrm{lac}}^\beta$ is bounded on L^p for $\beta > (d-1)|1/p - 1/2|$ via a single scale

analysis, and either Littlewood–Paley theory for $p \geq 2$ or the result stated in §6.2 for $1 < p < 2$.

We have the following sparse bound for $\mathcal{M}^\beta_{\mathrm{lac}}$ in the non-endpoint case.

Proposition 7.4. *Suppose $1 < p \leq q < \infty$ and that one of the following two conditions holds.*

(W_*) $1 < p \leq 2$, $p \leq q \leq p'$, $\beta > (d-1)\big(\frac{1}{p} - \frac{1}{2}\big) + \frac{1}{p} - \frac{1}{q}$.
(W'_*) $1 < p < \infty$, $\max\{p, p'\} \leq q < \infty$, $\beta > (d-1)\big(\frac{1}{2} - \frac{1}{q}\big) + \frac{1}{p} - \frac{1}{q}$.

Then $\mathcal{M}^\beta_{\mathrm{lac}} \in \mathrm{Sp}(p, q')$.

Proof. Let $K = \mathcal{F}^{-1}[m_\beta(2\cdot)]$ so that the singular support of K is $\{x : |x| = 1/2\}$. Let $K_{\ell,0} = K * \eta_\ell = \mathcal{F}^{-1}[m_\beta \widehat{\eta}_\ell]$, with η_ℓ defined as in (3.9), and split $K_{\ell,0} * f = \mathcal{A}_\ell f + \mathcal{R}_\ell f$ where the convolution kernel R_ℓ of $\mathcal{R}_\ell$ is supported in $\{x : |x| \geq 1\}$. The maximal function associated to R_ℓ is dominated by $2^{-\ell N}$ times the Hardy–Littlewood maximal function of f, similarly the maximal functions associated to A_ℓ are controlled by the Hardy–Littlewood maximal function for small ℓ and therefore satisfy a (p, q') sparse bound by §A.2. We use the notation $A_{\ell,0}$, $R_{\ell,0}$ for the convolution kernels of $\mathcal{A}_\ell$ and $\mathcal{R}_\ell$. Set $K_{\ell,k} = 2^{-kd} K_{\ell,0}(2^{-k}\cdot)$, and similarly define the kernels $A_{\ell,k}$ and $R_{\ell,k}$.

By the $L^p \to L^p$ result for $m_\beta(D)$ together with the multiplier result mentioned in §6.2 one can easily derive for $\ell > 0$ and any $\varepsilon > 0$

$$\Big\| \Big(\sum_{k \in \mathbb{Z}} |K_{\ell,k} * f|^2 \Big)^{1/2} \Big\|_p \lesssim 2^{\ell((d-1)|\frac{1}{p} - \frac{1}{2}| + \varepsilon - \beta)} \|f\|_p$$

for all $1 < p < \infty$ which of course implies

$$\big\| \sup_{k \in \mathbb{Z}} |K_{\ell,k} * f| \big\|_p \lesssim 2^{\ell((d-1)|\frac{1}{p} - \frac{1}{2}| + \varepsilon - \beta)} \|f\|_p.$$

We also have the single scale results

$$2^{\ell\beta} \|K_{\ell,0} * f\|_q \lesssim 2^{\ell((d-1)(\frac{1}{p} - \frac{1}{2}) + \frac{1}{p} - \frac{1}{q})} \|f\|_p \tag{7.14}$$

if $1 < p \leq 2$, $p \leq q \leq p'$, and

$$2^{\ell\beta} \|K_{\ell,0} * f\|_q \lesssim 2^{\ell((d-1)(\frac{1}{2} - \frac{1}{q}) + \frac{1}{p} - \frac{1}{q})} \|f\|_p \tag{7.15}$$

if $1 < p \leq 2$, $p' < q < \infty$ or $2 \leq p < \infty$, $p \leq q < \infty$.

By the above mentioned bounds for the operator $\mathcal{R}_\ell$ and the lacunary maximal operator generated by it we can replace $K_{\ell,k}$ and $K_{\ell,0}$ by $A_{\ell,k}$ and $A_{\ell,0}$, respectively,

Note that the exponents for L^p-boundedness and for L^q boundedness, i.e. $(d-1)|1/p - 1/2|$, $(d-1)|1/q - 1/2|$ are not larger than the exponents in the displayed inequalities (7.14) and (7.15) in their respective ranges. An application of Proposition 7.1 gives the desired sparse results for the maximal function generated by the $A_{\ell,k}$ and then also for the maximal function

generated by convolution with $K_{\ell,k}$. Summing in ℓ we can complete the proof of the proposition. □

Remark. The multiplier m_β can be replaced by other variants such as

$$m_{\beta,1}(\xi) = \frac{\sin|\xi|}{|\xi|}\frac{1}{(1+|\xi|^2)^{(\beta-1)/2}}, \qquad m_{\beta,2}(\xi) = \frac{J_{\beta-1/2}(|\xi|)}{|\xi|^{\beta-1/2}}\,.$$

7.2. **General classes of multipliers.** It is well known that the classical Mikhlin–Hörmander multiplier theorem [53, 103] can be interpolated with the L^2-estimate for multiplier transformations $m(D)$ with bounded multipliers [81, 82]. In particular one gets for $1 < p \le 2$,

$$\|m\|_{M^p} \lesssim \sup_{t>0} \|\phi m(t\cdot)\|_{L^r_\alpha}, \quad 1/r = 1/p - 1/2, \quad \alpha > d/r\,, \tag{7.16}$$

where $\|g\|_{L^r_\alpha} = \|(1+|D|^2)^{\alpha/2} g\|_r$ and ϕ is a nontrivial radial function supported with compact support away from the origin.

We give a sparse bound for this class of multipliers.

Proposition 7.5. *Let $1 < p \le 2$, $1/r = 1/p - 1/2$ and let m satisfy*

$$\sup_{t>0} \|\phi m(t\cdot)\|_{L^r_\alpha} \le A. \tag{7.17}$$

Suppose one of the following holds:

(i) $1 < p \le q \le 2$, *and* $\alpha > d(1/p - 1/2)$.
(ii) $2 \le q < \infty$ *and* $\alpha > d(1/p - 1/q)$.

Then

$$\|m(D)\|_{\mathrm{Sp}_\gamma(p,q')} \lesssim_{p,r,q,\alpha,\gamma} A.$$

Proof. We deduce this result from Theorem 1.6. Observe the inequality

$$\|g\|_{M^{p,2}} \le \|g\|_{L^r}, \qquad 1/r = 1/p - 1/2,$$

valid for $1 \le p \le 2$ which follows by interpolation from the standard cases $p = 1$ and $p = 2$. In view of the embedding $B^0_{r,1} \hookrightarrow L^r$ (see [110] for the definition and properties of Besov spaces) we get, for $1 \le p \le 2$,

$$\|g\|_{M^{p,2}} \lesssim \|g\|_{B^0_{r,1}}, \qquad 1/r = 1/p - 1/2. \tag{7.18}$$

Interpolating Bernstein's theorem $B^{d/2}_{2,1} \hookrightarrow \widehat{L^1}$ (which follows from the Cauchy–Schwarz inequality and Plancherel's theorem) with the embedding $B^0_{\infty,1} \hookrightarrow L^\infty$, we also have for $1 \le p \le 2$,

$$\|g\|_{M^{p,p}} \lesssim \|g\|_{B^{d/r}_{r,1}}, \qquad 1/r = 1/p - 1/2. \tag{7.19}$$

A further interpolation of (7.18) and (7.19) yields for $1 \le p \le q \le 2$

$$\|g\|_{M^{p,q}} \lesssim \|g\|_{B^{d(\frac{1}{q}-\frac{1}{2})}_{r,1}}\,, \qquad 1/r = 1/p - 1/2.$$

Finally, we have for $M \ge s$, the well-known inequality

$$\|\Delta_h^M g\|_{B^\beta_{r,1}} \lesssim |h|^s \|g\|_{B^{s+\beta}_{r,1}}$$

which we shall use for $\beta = d(1/q - 1/2)$ and which can be deduced from standard L^1-convolution inequalities.

Now let $r = 2p/(2-p)$, *i.e.* $1/r = 1/p - 1/2$. Applying the above inequalities to $g = \phi m(t\cdot)$ we get for $M \geq s$,

$$\sup_{|h|\leq 1} |h|^{-s} \|\Delta_h^M [\phi m(t\cdot)]\|_{M^{p,q}} \lesssim \|\phi m(t\cdot)\|_{B_{r,1}^{d(\frac{1}{q}-\frac{1}{2})+s}}$$

Now since $\alpha > d(1/p - 1/2)$ we can find $s > d(1/p - 1/q)$ such that

$$\alpha > d(1/q - 1/2) + s > d(1/p - 1/2).$$

Thus if α is as in the display, then $L^r_\alpha \hookrightarrow B_{r,1}^{d(1/q-1/2)+s}$ and an application of Theorem 1.6 yields the sparse bound stated in part (i).

For part (ii) let $2 \leq q < \infty$ and observe that

$$\|g\|_{M^{p,q}} \lesssim_K \|g\|_{M^{p,2}} \quad \text{if } \operatorname{supp}(g) \subset K, \quad K \text{ compact.} \tag{7.20}$$

To see this take a Schwartz function υ whose Fourier transform equals 1 on K and observe that by Young's inequality convolution with υ maps L^2 into L^q. We see from (7.20) and (7.18) that for such compactly supported g and $M > s$,

$$|h|^{-s} \|\Delta_h^M g\|_{M^{p,q}} \lesssim |h|^{-s} \|\Delta_h^M g\|_{B^0_{r,1}} \lesssim \|g\|_{B^s_{r,1}}.$$

This we use for $g = \phi m(t\cdot)$ and $\alpha > s > d(1/p - 1/q)$. Then part (ii) follows by the embedding $L^r_\alpha \hookrightarrow B^s_{r,1}$ and an application of Theorem 1.6. □

Remark 7.6. The assumption $p \leq 2$ is not a significant restriction. Indeed observe that by definition of the sparse operator classes we have $T \in \mathrm{Sp}(p_1, p_2)$ if and only if $T^* \in \mathrm{Sp}(p_2, p_1)$. For multiplier transformations we have $m(D)^* = m(-D)$ and $m(-D)f(-x) = m(D)[f(-\cdot)](x)$ which implies that $m(D) \in \mathrm{Sp}(p_1, p_2)$ if and only if $m(D) \in \mathrm{Sp}(p_2, p_1)$.

We can draw two conclusions from this duality argument. First, the range $1 \leq p \leq 2$, $q \geq p'$ in Proposition 7.5 could be deduced from the result in the range $1 \leq p \leq 2$, $2 \leq q \leq p'$. Second, the result in Proposition 7.5 also implies a result in the range $2 \leq p \leq q < \infty$. Namely, in this case, if $1/r = 1/2 - 1/q$ and $\alpha > d(1/2 - 1/q)$ then one gets $m(D) \in \mathrm{Sp}(p, q')$ under the assumption (7.17).

7.2.1. *Miyachi classes and subdyadic Hörmander conditions.* We now discuss some consequences for multiplier classes considered by Miyachi [88] and their corresponding versions under a subdyadic Hörmander-type formulation [10]. Given $a > 0, b \in \mathbb{R}$, let $\mathrm{Miy}(a, b)$ denote the class of smooth functions $m : \mathbb{R}^d \to \mathbb{C}$ supported on $\{\xi : |\xi| \geq 1\}$ and satisfying the differential inequalities

$$|\partial^\iota m(\xi)| \lesssim_\iota |\xi|^{-b+|\iota|(a-1)} \tag{7.21}$$

for all $|\xi| \geq 1$ and all multiindices $\iota \in \mathbb{N}_0^d$ satisfying $|\iota| \leq \lfloor d/2 \rfloor + 1$. The oscillatory multipliers $m_{a,b}$ defined below in (7.26) are considered model cases, at least in regards to the $L^p \to L^p$ boundedness properties. It is

known that multipliers in $\mathrm{Miy}(a,b)$ belong to M^p whenever $b \ge ad|1/p-1/2|$ and $1<p<\infty$, see [43, 88]. It has also been observed that these endpoint results are special cases of Hörmander-type multiplier theorems involving certain endpoint Besov spaces, see [4, 99]. Sparse bounds for multipliers in $\mathrm{Miy}(a,b)$ in the non-endpoint range $b > ad|1/p-1/2|$ were obtained by Cladek and the first author in [11] via a single scale analysis, under the additional assumption that (7.21) hold for all multiindices $\iota \in \mathbb{N}_0^d$. We note that in the range $0<a<1$ they also extended these results to larger closely related classes of pseudo-differential operators, *cf.* [41, 11].

The subdyadic Hörmander-type classes, also extending the class $\mathrm{Miy}(a,b)$ are obtained by replacing the pointwise condition (7.21) by

$$\sup_B \ \mathrm{dist}(B,0)^{b+(1-a)|\iota|}\Big(\frac{1}{|B|}\int_B |D^\iota m(\xi)|^2 d\xi\Big)^{1/2} < \infty \tag{7.22}$$

for all $\iota \in \mathbb{N}_0^d$ with $|\iota| \le \lfloor d/2\rfloor + 1$. Here the supremum is taken over all euclidean balls B in $\mathbb{R}^d$ with $\mathrm{dist}(B,0)\ge 1$ such that $r(B)\sim \mathrm{dist}(B,0)^{1-a}$, where $r(B)$ denotes the radius of B. This class was considered in [10] which contains sharp weighted inequalities of Fefferman–Stein type that can be used to recover the sharp L^p estimates. In [11, §3] the question was raised whether the results on sparse bounds for multiplier transformations in the Miyachi class can be extended to multipliers satisfying a subdyadic condition above, in the sense that it is sufficient to assume that (7.21) or (7.22) hold for all $|\iota| \le \lfloor d/2\rfloor + 1$ rather than for all $\iota \in \mathbb{N}_0^d$. We shall see that this is the case, and that such and more general multi-scale results can be obtained from Proposition 7.5. The following simple observation will be helpful; note that condition (7.22) (and therefore (7.21)) implies (7.23).

Lemma 7.7. *Let $a>0$, $2\le r<\infty$, $b>ad/r$. Suppose m_k are supported in $\{\xi : |\xi|\ge 1\}$ and suppose that there is a constant C such that*

$$\sup_{t>1} t^{b-a|\imath|}\Big(t^{-d}\int_{t/2\le|\xi|\le 2t} \big[|\xi|^{|\imath|}\partial^\imath m_k(\xi)|\big]^r d\xi\Big)^{1/r} \le C \tag{7.23}$$

for all multiindices $\imath$ with $|\imath|\le \lfloor d/r\rfloor+1$ and for all $k\in\mathbb{Z}$. Then the family $\{m_k\}$ satisfies condition

$$\sup_{k\in\mathbb{Z}}\sup_{t>0} t^{b-a\alpha}\|\phi m_k(t\cdot)\|_{L^r_\alpha} < \infty. \tag{7.24}$$

Proof. A change of variable shows that the condition (7.23) is equivalent to

$$\big\|\partial^\imath[\phi m_k(s\cdot)]\big\|_r \lesssim s^{a|\imath|-b}$$

for all multiindices $\imath$ with $|\imath|\le\lfloor d/r\rfloor+1$. Pick $\alpha\in(d/r,\lfloor d/r\rfloor+1)$ such that $\alpha a<b$. Then the condition implies

$$\sup_s \|\phi m_k(s\cdot)\|_{L^r_\alpha} s^{b-a\alpha} < \infty,$$

which implies (7.24) in view of the assumption on the supports since $\alpha > d/r$. □

We shall now formulate a result for families of multipliers satisfying condition (7.24). For simplicity of our statements, we consider only the case $p \le 2$ and argue by duality for $p > 2$ (see Remark 7.6).

Proposition 7.8. *Let $1 < p \le 2$, $r = \frac{2p}{2-p}$ (i.e. $\frac{1}{r} = \frac{1}{p} - \frac{1}{2}$), and let, for $k \in \mathbb{Z}$, m_k be supported in $\{\xi : |\xi| \ge 1\}$. Let $a, b \ge 0$ such that $b > ad(\frac{1}{p} - \frac{1}{2})$ and suppose that either*

(i) $1 < p \le q \le 2$ *and* $b > ad(\frac{1}{p} - \frac{1}{2})$, *or*
(ii) $2 \le q < \infty$ *and* $b > ad(\frac{1}{p} - \frac{1}{q})$.

Let $\alpha > d(\frac{1}{p} - \frac{1}{2})$ and assume $\sup_{k\in\mathbb{Z}} \sup_{t>0} t^{b-a\alpha} \|\phi m_k(t\cdot)\|_{L^r_\alpha} < \infty$.

Then $m := \sum_{k\in\mathbb{Z}} m_k(2^k\cdot) \in M_p$ and $m(D) \in \mathrm{Sp}(p, q')$

Proof. We split, by a dyadic decomposition $m_k(\xi) = \sum_{n=1}^\infty m_{k,n}(\xi)$ where $m_{k,n}$ is supported in an annulus $\{\xi : |\xi| \approx 2^n\}$, for all $k \in \mathbb{Z}$; in fact we can set $m_{k,n}(\xi) = m_k(\xi)\widehat{\eta}_n(\xi)$ with $\widehat{\eta}_n$ as in (3.10). Observe that $m_k\widehat{\eta}_0 = 0$ by the support properties of m_k and $\widehat{\eta}_0$. Now form $m^n(\xi) = \sum_{k\in\mathbb{Z}} m_{k,n}(2^k\xi)$. We wish to apply Proposition 7.5 to m^n, for every $n \ge 0$.

Fix any $t > 0$ and use the assumption to compute

$$\|\phi m^n(t\cdot)\|_{L^r_\alpha} \lesssim \sum_{\substack{k\in\mathbb{Z}:\\ c_1 2^n \le 2^k t \le C_1 2^n}} \|\phi m_{k,n}(2^k t\cdot)\|_{L^r_\alpha} \lesssim 2^{-n(b-a\alpha)}.$$

By (7.16) we obtain $\|m^n\|_{M^p} \lesssim 2^{-n(b-a\alpha)}$ and similarly, by Proposition 7.5 we obtain $\|m^n(D)\|_{\mathrm{Sp}_\gamma(p,q')} \lesssim 2^{-n(b-a\alpha)}$. The desired bounds follow by summing in n as $\alpha < b/a$. □

As a consequence we can obtain a sparse bound for the lacunary maximal function $\sup_k |m(2^k D)f|$ and indeed a square function that dominates it.

Corollary 7.9. *Let p, r, q, a, b as in Proposition 7.8. Let m be supported in $\{\xi : |\xi| \ge 1\}$ satisfying $\sup_{t>0} t^{b-a\alpha}\|\phi m(t\cdot)\|_{L^r_\alpha} < \infty$. Then we have the (p, q')-sparse bound*

$$\int_{\mathbb{R}^d} \Big(\sum_{k\in\mathbb{Z}} |m(2^k D)f(x)|^2\Big)^{1/2} \omega(x)dx \lesssim \Lambda^*_{p,q'}(f, \omega).$$

Proof. Consider the multiplier $m_v(\xi) = \sum_{k\in\mathbb{Z}} r_k(v)m(2^k\xi)$ where $(r_k)_{k\in\mathbb{N}}$ denotes the sequence of Rademacher functions defined on the unit interval. Then by Proposition 7.8 applied to $m_k(\xi) = r_k(v)m(\xi)$ we obtain

$$\Big|\int_{\mathbb{R}^d} m_v(D)f_1(x)f_2(x)dx\Big| \lesssim \Lambda^*_{p,q'}(f_1, f_2), \tag{7.25}$$

with the implicit constant independent of v. Let $u_v(x) = \frac{m_v(D)f(x)}{|m_v(D)f(x)|}$ so that u_v is unimodular, and we also get by (7.25) with $f_2 = \omega u_v$

$$\int_{\mathbb{R}^d} |m_v(D)f|\omega(x)dx = \int_{\mathbb{R}^d} m_v(D)f(x)\,\omega(x)u_v(x)dx \\ \lesssim \Lambda^*_{p,q'}(f, u_v\omega) = \Lambda^*_{p,q'}(f,\omega).$$

Integrating in v and using Fubini's theorem and Khinchine's inequality, one obtains

$$\int_{\mathbb{R}^d} \Big(\sum_{k\in\mathbb{Z}} |m_k(2^kD)f|^2\Big)^{1/2}\omega(x)dx \lesssim \int_{\mathbb{R}^d}\int_0^1 |m_v(D)f(x)|\,dv\,\omega(x)\,dx \\ = \int_0^1\int_{\mathbb{R}^d} |m_v(D)f(x)|\,\omega(x)\,dx\,dv \lesssim \Lambda^*_{p,q'}(f,\omega)$$

and the proof is complete. □

Remark. Similar results can be obtained for versions of the previous multiplier classes if $a < 0$ and m is supported in $\{\xi : |\xi| \le 1\}$. We omit the statements.

7.2.2. *Multiscale variants of oscillatory multipliers.* Given $a > 0$, $a \neq 1$, $b \in \mathbb{R}$, consider the oscillatory Fourier multipliers

$$m_{a,b}(\xi) = \chi_\infty(\xi)|\xi|^{-b}e^{i|\xi|^a}, \tag{7.26}$$

where $\chi_\infty \in C^\infty(\mathbb{R}^d)$ is such that $\chi(\xi) = 0$ for $|\xi| \le 1$ and $\chi_\infty(\xi) = 1$ for $|\xi| \ge 2$. As already mentioned the operators $m_{a,b}(D)$ are sometimes considered model cases of the class $\mathrm{Miy}(a,b)$, known to be bounded on L^p if and only if $b \ge ad|1/p-1/2|$ and $1 < p < \infty$; see [104], [43], [88]. This result is sharp when $a \neq 1$; the case $a = 1$ forms an exceptional case corresponding to the wave multipliers considered in §7.1.3; we exclude it in this section.

Given a sequence $(c_k)_{k\in\mathbb{Z}}$ with $|c_k| \le 1$ we form the multiscale variant

$$m(\xi) = \sum_{k\in\mathbb{Z}} c_k m_{a,b}(2^k\xi) \tag{7.27}$$

which is bounded on L^p for $b > ad|1/p - 1/2|$. Proposition 7.8 shows that for $1 < p \le 2$ we have $m(D) \in \mathrm{Sp}(p,2)$ for $q \le 2$, but in order to get a $\mathrm{Sp}(p,q')$ bound for $q > 2$ we had to impose the more restrictive condition $b > ad(1/p - 1/q)$. We show that this estimate can be improved, in particular an additional restriction is not necessary for $q \le p'$ and in this range we can upgrade the $\mathrm{Sp}(p,2)$ bound to an $\mathrm{Sp}(p,p)$ bound for the multipliers in (7.27) (see Figure 2).

This improvement relies on special features of the multipliers $m_{a,b}$ which are not shared by a general multiplier in the class $\mathrm{Miy}(a,b)$. Unlike in the proof of Proposition 7.8 we can no longer rely on analyzing the problem on the multiplier side. Instead we have to analyze Schwartz kernels and employ stationary phase estimates, taking advantage of the fact that the

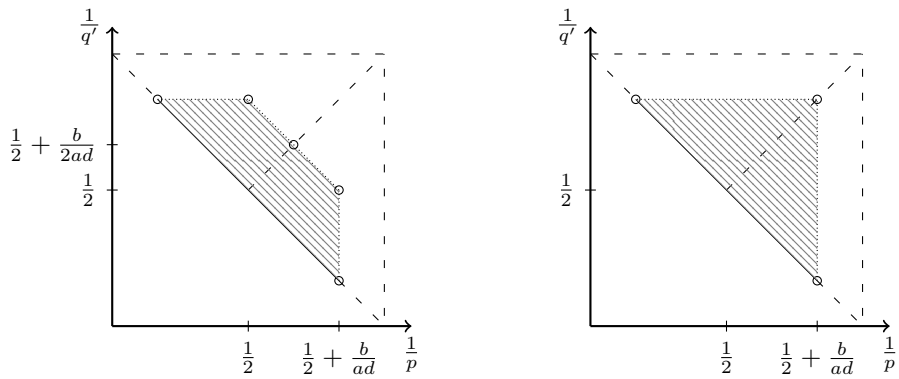

FIGURE 2. Sparse bounds for a general multiplier in $\mathrm{Miy}(a,b)$ (left) and for the oscillatory multipliers $m_{a,b}$ (right) for given $a,b>0$. The condition (ii) in Proposition 7.8 can be relaxed for the specific $m_{a,b}$ (Proposition 7.10).

Hessian of the phase function $\xi \mapsto |\xi|^a$ is nondegenerate when $a \neq 1$, $a>0$. Incidentally, this also reveals that the $m_{a,b}$ satisfy better $L^p \to L^q$ mapping properties than a general multiplier in $\mathrm{Miy}(a,b)$ when $1<p\le 2$, $2<q\le p'$. It is therefore more natural to base the proof directly on Theorem 6.1 rather than on the formulation in Theorem 1.6.

Proposition 7.10. *Let $1<p\le 2$, $a\in(0,\infty)\setminus\{1\}$, and m as in* (7.27), *with* $\sup_k |c_k|\le 1$. *Let $b>ad(1/p-1/2)$. Then $m(D)\in \mathrm{Sp}(p,p)$.*

Proof. We decompose as in the proof of Proposition 7.8. Recall that $\widehat{\eta}_0$ is supported in $\{|\xi|<1\}$ and $m_{a,b}$ in $\{|\xi|>1\}$, hence $\widehat{\eta}_0 m_{a,b}=0$, and we can write $m=\sum_{n=1}^\infty m^n$ where

$$m^n(\xi)=\sum_{k\in\mathbb{Z}} c_k m_{a,b}(2^k\xi)\widehat{\eta}_n(2^k\xi).$$

We shall show that

$$\|m^n(D)\|_{\mathrm{Sp}_\gamma(p,p)} \lesssim 2^{-n\varepsilon(p)} \tag{7.28}$$

with $\varepsilon(p)>0$ and then sum in n.

To verify the claim (7.28) we use Theorem 6.1. For this we have to analyze, for radial $\phi\in C_c^\infty$ supported in $\{\xi : 1/2<|\xi|<2\}$, the expression

$$\|[\phi m^n(t\cdot)]*\widehat{\Psi}_\ell\|_{M^{p,p'}} \le \sum_{\substack{k\in\mathbb{Z}\\ 2^{n-3}\le 2^k t\le 2^{n+1}}} \|[\phi m_{a,b}(2^k t\cdot)\widehat{\eta}_n(2^k t\cdot)]*\widehat{\Psi}_\ell\|_{M^{p,p'}}$$

and show that for some $\varepsilon>0$

$$\sum_{\ell\ge 0}\sup_{t>0}\|[\phi m^n(t\cdot)]*\widehat{\Psi}_\ell\|_{M^{p,p'}}\, 2^{\ell(d(1/p-1/p')+\varepsilon)} \lesssim 2^{-n\varepsilon(p)}. \tag{7.29}$$

To this end, fix k, t with $2^{n-3} \le 2^k t \le 2^{n+1}$ and analyze the Fourier inverse of $\phi m_{a,b}(2^k t\cdot)\widehat{\eta}_n(2^k t\cdot)$, i.e.

$$K_n(x) = (2\pi)^{-d} \int \phi(\xi)\widehat{\eta}_n(2^k t\xi)\chi_\infty(2^k t\xi)(2^k t|\xi|)^{-b} e^{i\langle x,\xi\rangle + i(2^k t)^a|\xi|^a} d\xi.$$

The phase function $\langle x,\xi\rangle + (2^k t)^a|\xi|^a$ becomes stationary on the support of ϕ only when $|x| \approx (2^k t)^a \approx 2^{na}$ and the Hessian of $|\xi|^a$ is nondegenerate there. Thus by integration by parts we see that there are constants $c_1 < 1$, $C_1 > 1$ such that

$$|K_n(x)| \le \begin{cases} C_N 2^{-n(b+N)} & \text{for } |x| \le c_1 2^{na} \\ C_N 2^{-nb}|x|^{-N} & \text{for } |x| \ge C_1 2^{na} \end{cases}$$

and by the method of stationary phase

$$|K_n(x)| \lesssim 2^{-n(b+ad/2)}, \quad \text{for } \ c_1 2^{na} \le |x| \le C_1 2^{na}.$$

This implies for $2^{n-3} \le 2^k t \le 2^{n+1}$ and suitable C, independently of k, t,

$$\begin{aligned} &\|[\phi m^n(2^k t\cdot)] * \widehat{\Psi}_\ell\|_{M^{1,\infty}} \\ &\qquad \lesssim \begin{cases} 2^{-n(b+ad/2)} & \text{for } |\ell - na| \le C \\ \min\{C_N 2^{-n(b+N)},\, C_N 2^{-nb} 2^{-\ell N}\} & \text{for } |\ell - na| > C. \end{cases} \end{aligned} \tag{7.30}$$

We also have the $M^{2,2}$ bound

$$\begin{aligned} &\|[\phi m^n(2^k t\cdot)] * \widehat{\Psi}_\ell\|_{\infty} \\ &\quad \lesssim \begin{cases} 2^{-nb} & \text{for } |\ell - na| \le C \\ \min\{C_N 2^{-n(b+N-d)}, C_N\, 2^{-nb} 2^{-\ell(N-d)}\} & \text{for } |\ell - na| > C. \end{cases} \end{aligned} \tag{7.31}$$

Interpolating (7.30) and (7.31) we get

$$\|[\phi m^n(2^k t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,p'}} \lesssim \begin{cases} 2^{-n(b+ad(1/p-1/2))} & \text{for } |\ell - na| \le C \\ 2^{-nb} C_N \min\{2^{-nN}, 2^{-\ell N}\} & \text{for } |\ell - na| > C. \end{cases}$$

Only the five terms with $2^{n-3} \le 2^k t \le 2^{n+1}$ make a contribution. We sum those terms, then take a supremum in t (observing that the displayed bound above is independent of t) and then sum in $\ell \ge 0$. We obtain

$$\begin{aligned} &\sum_{\ell \ge 0} \sup_{t>0} \|[\phi m^n(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,p'}} 2^{\ell(d(\frac{1}{p}-\frac{1}{p'})+\epsilon)} \\ &\qquad\qquad \lesssim_\epsilon 2^{-n(b+ad(\frac{1}{p}-\frac{1}{2}))} 2^{na(d(\frac{1}{p}-\frac{1}{p'})+\varepsilon)} \lesssim 2^{-n(b-ad(\frac{1}{p}-\frac{1}{2}-\epsilon))}. \end{aligned}$$

Since we assume $b > ad(1/p - 1/2)$ this leads to (7.29) and then to the claim (7.28) via Theorem 6.1. □

Given Proposition 7.10, we can now derive an improved sparse bound for a lacunary maximal function and a corresponding square function associated

with the multipliers $m_{a,b}$; thus for these examples we improve on Corollary 7.9.

Corollary 7.11. *Let $1 < p \le 2$, $a > 0$, $b > ad(1/p - 1/2)$. Then*

$$\int_{\mathbb{R}^d} \Big(\sum_{k\in\mathbb{Z}} |m_{a,b}(2^k D) f(x)|^2 \Big)^{1/2} \omega(x) dx \lesssim \Lambda^*_{p,p}(f,\omega).$$

Proof. Choose $c_k = \pm 1$ in (7.26). Then Proposition 7.10 together with a randomization argument exactly as in the proof of Corollary 7.9 yields the assertion. □

7.3. **Prototypical versions of singular Radon transforms.** Let σ be a bounded Borel measure supported in $\{x : |x| \le 1\}$ and satisfying

$$\int d\sigma = 0 \quad \text{and} \quad \sup_{\xi\in\widehat{\mathbb{R}}^d} (1+|\xi|)^b |\widehat{\sigma}(\xi)| < \infty \quad \text{for some } b > 0. \tag{7.32}$$

Let $\{a_j\}_{j\in\mathbb{Z}}$ satisfy

$$|a_j| \le 1 \tag{7.33}$$

and define

$$\mathcal{S}^{N_1,N_2} f(x) = \sum_{j=N_1}^{N_2} a_j 2^{-jd} \sigma(2^{-j}\cdot) * f(x)$$

and

$$\mathcal{S} f(x) = \lim_{\substack{N_2\to\infty \\ N_1\to-\infty}} \mathcal{S}^{N_1,N_2} f(x). \tag{7.34}$$

This is the "prototypical" singular Radon transform considered by R. Oberlin [89], see also Duoandikoetxea and Rubio de Francia [39]. It is easy to see using the cancellation of the kernel that the limit exists pointwise for C_c^∞ functions.

In addition, we assume that σ is L^{p_0} improving, i.e.

$$\|\sigma * f\|_q \le A \|f\|_{p_0} \tag{7.35}$$

for some q with $p_0 < q < \infty$. The following result is due to R. Oberlin.

Proposition 7.12 ([89])**.** *Let σ be as in (7.32), and $\{a_j\}_{j\in\mathbb{Z}}$, $\mathcal{S}$ be as in (7.33), (7.34). Let $1 < p_0 < p < q < \infty$ and assume that (7.35) holds. Then $\mathcal{S}$ satisfies the (p,q')-sparse bound*

$$|\langle \mathcal{S} f, \omega\rangle| \lesssim \Lambda^*_{p,q'}(f,\omega).$$

The same sparse bound holds for the operators $\mathcal{S}_{N_1,N_2}$, uniformly in N_1, N_2.

We emphasize that Oberlin also proved certain endpoint estimates for $p = p_0$, working with local Orlicz norms in the definition of sparse forms.

One can extend Proposition 7.12 to cover associated maximal truncation and variational truncation operators defined by

$$\mathcal{S}_* f(x) = \sup_{N_1<N_2} \Big| \sum_{j=N_1}^{N_2} a_j 2^{-jd} \sigma(2^{-j}\cdot) * f(x) \Big|,$$

$$\mathcal{V}_*^r \mathcal{S} f(x) = \sup_{M\in\mathbb{N}} \sup_{n_1<\cdots<n_M} \Big(\sum_{i=1}^{M-1} \Big| \sum_{j=n_i+1}^{n_{i+1}} a_j 2^{-jd} \sigma(2^{-j}\cdot) * f(x) \Big|^r \Big)^{1/r}.$$

Proposition 7.13. *Let $1 < p_0 < p < q < \infty$, $r > 2$ and σ, $\{a_j\}$ be as in* (7.32), (7.33), (7.35). *Then $\mathcal{S}_*$ and $\mathcal{V}_*^r\mathcal{S}$ satisfy the sparse bounds*

$$|\langle \mathcal{S}_* f, \omega\rangle| + |\langle \mathcal{V}_*^r \mathcal{S} f, \omega\rangle| \lesssim \Lambda^*_{p,q'}(f,\omega).$$

Proof. We apply Theorems 5.3 and 5.4. To verify the assumptions (5.11a), (5.11b) see [39, Theorem E]. To verify assumptions (5.12a), (5.12b) see [59, Theorem 1.2]. Interpolation arguments using the Fourier decay assumption in (7.32), and Lemma 3.6 can be used to establish the additional Hölder condition in (1.9). □

The setup above is also similar in spirit to the theorems on truncations of rough singular integrals with bounded kernels [37]. We have been deliberately short in our presentation as the results in this section are essentially known. For a more detailed exposition the reader may consult §7.4 below, in which a singular Radon transform built on spherical integrals, and other versions of maximal functions associated to singular Radon transforms, are considered.

7.3.1. *An approach via Fourier multipliers.* In order to understand the scope of our multiplier theorems, it is instructive to deduce the sparse bounds for the prototypical singular Radon transform $\mathcal{S}$ in Proposition 7.12 from Theorem 6.1 (or Theorem 1.6). Since σ is a finite Borel measure we have $\|\phi\widehat{\sigma}(2^j t\cdot)\|_{M^{q,q}} = O(1)$ for $1 \le q \le \infty$. By (7.32) and interpolation with $L^2 \to L^2$ bounds we have for some $\varepsilon_0(q) > 0$

$$\|\phi\widehat{\sigma}(2^j t\cdot)\|_{M^{q,q}} \lesssim C_q \min\{(2^j t)^{\varepsilon_0(q)}, (2^j t)^{-\varepsilon_0(q)}\}, \quad 1 < q < \infty, \tag{7.36}$$

using either cancellation or decay, and by Young's inequality we get the same bound for $\|\phi\widehat{\sigma}(2^j t\cdot)\|_{M^{p,q}}$ when $1 \le p \le q$, $1 < q < \infty$. This takes care of the term $\ell = 0$ in the condition (6.2a). To verify the remaining hypothesis of Theorem 6.1 it suffices to check that for $\ell > 0$ the condition

$$\sup_{t>0} \sum_{j\in\mathbb{Z}} \|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim 2^{-\ell(d(1/p-1/q)+\varepsilon)} \tag{7.37}$$

is satisfied, as the condition (6.2b) trivially follows by the assumption (7.32).

Since $\widehat{\sigma}$ is smooth we have for $2^j t \le 1$

$$\|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{r,s}} \lesssim C_N 2^{-\ell N}, \quad 2^j t \le 1,$$

for $1 \le r \le s \le \infty$ and therefore by interpolation with (7.36) and taking geometric means we see that there is an $\epsilon_1(r,s)$ such that $\epsilon_1(r,s) > 0$ if $1 < r \le s < \infty$ and

$$\|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{r,s}} \lesssim C_N 2^{-\ell d}(2^j t)^{\epsilon_1(r,s)}, \quad 2^j t \le 1. \tag{7.38}$$

The contributions for $2^j t \ge 1$ are more interesting. Since σ is supported in $\{x : |x| \le 1\}$ we have the kernel estimate

$$|\mathcal{F}^{-1}[\phi\widehat{\sigma}(2^j t\cdot)](x)| \lesssim_N |x|^{-N} \quad \text{for } |x| \ge 2^{j+1}t$$

and hence

$$\|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim 2^{-\ell N} \quad \text{for } 2^\ell \ge 2^{j+4}t \ge 1. \tag{7.39}$$

For $2^\ell \le 2^{j+4}t$ we do a rescaling argument to estimate

$$\|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p_0,q}} \lesssim \|\widehat{\sigma}(2^j t\cdot)\|_{M^{p_0,q}} = (2^j t)^{-d(1/p_0-1/q)}\|\widehat{\sigma}\|_{M^{p_0,q}}$$

and by assumption $\widehat{\sigma} \in M^{p_0,q}$. Interpolating with the $M^{q,q}$ estimate in (7.36) we get for $p_0 < p \le q$

$$\|[\phi\widehat{\sigma}(2^j t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim (2^j t)^{-d(1/p-1/q)-\epsilon(p,q)} \quad \text{for } 2^{j+4}t \ge 2^\ell. \tag{7.40}$$

Combining (7.38), (7.39) and (7.40) and summing in j we get (7.37) for a suitable $\varepsilon = \varepsilon(p,q) > 0$.

7.4. **Densities on spheres: Maximal singular integrals.** As discussed in §7.3 the Corollary 1.2 covers classes of singular Radon transforms and also associated maximal operators for truncations. Here we will consider a natural singular integral variant of the spherical maximal function, and obtain a sparse domination inequality analogous to the one for spherical maximal functions with specific dilation sets in [3, 93]. Let σ be the surface measure on the unit sphere $\{x : |x| = 1\}$ in $\mathbb{R}^d$ for $d \ge 2$ and $\mu = \chi\sigma$ with a choice of smooth χ such that

$$\int d\mu = 0.$$

For every $t \in [1,2]$ we consider, for fixed $t \in [1,2]$, the prototypical singular Radon transform as in the previous section

$$\mathcal{S}_t^{N_1,N_2} f = \sum_{j=N_1}^{N_2} \mu_{2^j t} * f, \qquad \mathcal{S}_t f = \lim_{\substack{N_2\to\infty,\\ N_1\to-\infty}} \mathcal{S}_t^{N_1,N_2} f \tag{7.41}$$

and then form, for $E \subset [1,2]$, the maximal function

$$\mathcal{S}_E f(x) = \sup_{t\in E} |\mathcal{S}_t f(x)|. \tag{7.42}$$

For $0 \le \beta \le \alpha \le 1$ define $\mathcal{R}(\beta,\alpha) \subset [0,1]^2$ as the union of the interior of the convex hull of the points

$$Q_1 = (0,0), \qquad Q_{2,\beta} = (\tfrac{d-1}{d-1+\beta}, \tfrac{d-1}{d-1+\beta}),$$
$$Q_{3,\beta} = (\tfrac{d-\beta}{d-\beta+1}, \tfrac{1}{d-\beta+1}), \qquad Q_{4,\alpha} = (\tfrac{d(d-1)}{d^2+2\alpha-1}, \tfrac{d-1}{d^2+2\alpha-1})$$

with the open segment connecting Q_1 and $Q_{2,\beta}$.

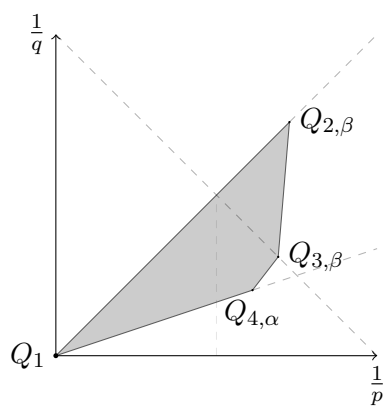

FIGURE 3. The region $\mathcal{R}(\beta,\alpha)$ with $\beta = 0.75$, $\alpha = 0.9$, $d = 3$.

For $E \subset [1,2]$ denote by $\dim_{\mathrm{M}} E$ the upper Minkowski dimension of E and by $\dim_{\mathrm{qA}} E$ the quasi-Assouad dimension of E (see [93] for definitions and background, and for a discussion of classes of sets E for which the single-scaled $L^p \to L^q$ results described above are sharp).

Proposition 7.14. *Let $d \ge 2$, $E \subset [1,2]$ and $(1/p, 1/q) \in \mathcal{R}(\beta,\alpha)$ with $\beta = \dim_{\mathrm{M}} E$, $\alpha = \dim_{\mathrm{qA}} E$. Then there is the (p,q')-sparse domination inequality*

$$|\langle \mathcal{S}_E f, \omega\rangle| \le C\Lambda^*_{p,q'}(f,\omega).$$

The two-dimensional version of our operator models a maximal operator associated to a family of Hilbert transforms on curves considered in [49, 50] where nonisotropic dilations are used (see also the previous papers [86, 48] for related problems). In this nonisotropic case one could also consider more general situations, i.e. when E is not a subset of $[1,2]$ (see also the prior work [86] on maximal functions) but this involves multi-parameter structures for which sparse domination result are difficult and in some cases are proved to not hold [7].

Proof of Proposition 7.14. Using the density Lemma A.1 we may assume that $f \in C_c^\infty$. It is then easy to see that for any bounded set $U \in \mathbb{R}^n$ we have $\mu_{2^j t} * f(x) = 0$ for all $x \in U$, $t \in [1,2]$ and sufficiently large j. Moreover using the cancellation of μ and the smoothness of f we see that

$\mu_{2^j t} * f(x) = O(2^j)$ as $j \to -\infty$. Thus we see that for $f \in C_c^\infty$ the function $\mathcal{S}_t f$ is well defined and

$$\lim_{\substack{N_2\to\infty,\\ N_1\to-\infty}} \sup_{t\in[1,2]} |\mathcal{S}_t f - \mathcal{S}_t^{N_1,N_2} f| = 0,$$

where the limit is uniform on compact sets. It is therefore sufficient to prove a sparse bound for the maximal function $\sup_{t\in E} |S_t^{N_1,N_2} f|$ which is uniform in N_1, N_2. In what follows we will drop the superscript in $\mathcal{S}_t^{N_1,N_2}$ but assume that we still working with a truncated sum depending on N_1, N_2.

To verify conditions (1.8), (1.9) in Corollary 1.2 we first note that for $(\frac{1}{p}, \frac{1}{q}) \in \mathcal{R}(\beta, \alpha)$ there is $\varepsilon(p,q) > 0$ such that for $\lambda > 2$

$$\| \sup_{t\in E} |\mu_t * f|\|_q \lesssim_{p,q} \lambda^{-\varepsilon(p,q)} \|f\|_p, \quad f \in \mathcal{E}_{\mathrm{ann}}(\lambda). \tag{7.43}$$

This is coupled with an elementary $L^p \to L^q$ estimate with constant $O(1)$ estimate for functions with frequency support near the origin to yield (1.9a) via Lemma 3.6; this also settles (1.9a) by translation invariance. For inequality (7.43) we may refer to [93, Cor. 2.2].

It remains to verify (1.7a) and (1.7b) which follow by verifying the L^p boundedness of $\mathcal{S}_E$ for $(\frac{1}{p}, \frac{1}{p})$ on the open interval $(Q_1 Q_{2,\beta})$. To accomplish this we make a further decomposition on the frequency side. Let η_ℓ be as in (3.9), (3.10), and set $\eta_{\ell,j} = 2^{-jd}\eta_\ell(2^{-j}\cdot)$, so that $\widehat{\eta_{0,j}}$ is supported where $|\xi| \lesssim 2^{-j}$ and $\widehat{\eta_{\ell,j}}$ is supported where $|\xi| \approx 2^{\ell-j}$. Setting

$$\mathcal{S}_E^\ell f(x) = \sup_{t\in E} \Big| \sum_{j=N_1}^{N_2} \mu_{2^j t} * \eta_{\ell,j} * f(x) \Big|$$

it then suffices to show

$$\|\mathcal{S}_E^\ell f\|_p \lesssim_p 2^{-\ell\delta(p)} \|f\|_p \tag{7.44}$$

with $\delta(p) > 0$ for $(\frac{1}{p}, \frac{1}{p}) \in (Q_1 Q_{2,\beta})$. The estimate for $\ell = 0$ reduces to standard singular integral theory; this uses the cancellation of μ. Thus from now on we assume $\ell > 0$.

We shall first discuss the case when either $d \geq 3$ or $d = 2$, $\beta < 1$ where we use arguments as in [101]. Because of $|\widehat{\mu}(\xi)| \lesssim \min\{|\xi|, |\xi|^{-(d-1)/2}\}$ we get

$$\sup_{\xi\in\widehat{\mathbb{R}}^d} \Big| \sum_{j=N_1}^{N_2} \widehat{\mu_{2^j t}}(\xi) \widehat{\eta_{\ell,j}}(\xi) \Big| \lesssim 2^{-\ell(d-1)/2},$$

which implies an L^2 boundedness result for the operators $\mathcal{S}_t^\ell$ with constant $O(2^{-\ell(d-1)/2})$, uniformly in $t \in [1,2]$.

We also have the L^p boundedness result

$$\Big\| \sum_{j=N_1}^{N_2} \mu_{2^j t} * \eta_{\ell,j} * f \Big\|_p \lesssim C_p \|f\|_p, \quad 1 < p < \infty$$

which is a consequence of results on isotropic singular Radon transforms as, say, in [39]. By interpolation we get for all $\varepsilon > 0$

$$\Big\| \sum_{j=N_1}^{N_2} \mu_{2^j t} * \eta_{\ell,j} * f \Big\|_p \lesssim C_{\varepsilon,p} 2^{\ell\varepsilon(1-1/p)} \min(2^{-\ell\frac{d-1}{p}}, 2^{-\ell\frac{d-1}{p'}}) \|f\|_p, \quad 1 < p < \infty.$$

The same estimate with $\mu_{2^j t} * \eta_{\ell,j}$ replaced by $2^{-\ell} 2^{-jd} [\frac{d}{dt}\mu_t * \eta_\ell](2^{-j}\cdot)$ also holds. We cover the set E with $O(2^{\ell(\beta+\varepsilon)})$ intervals of length $2^{-\ell}$ and argue as in [101, p.119] to obtain

$$\|\mathcal{S}_E^\ell f\|_p \lesssim_\varepsilon 2^{\ell(\frac{\beta}{p}+\varepsilon)} \min(2^{-\ell\frac{d-1}{p}}, 2^{-\ell\frac{d-1}{p'}}) \|f\|_p.$$

This gives (7.44), provided that $d \geq 3$ or $d = 2$, $\beta < 1$.

For the case $d = 2$, $\beta = 1$, we need to show L^p boundedness for $p > 2$. By a Sobolev embedding argument this follows from the inequality

$$\Big(\int_1^2 \Big\| \sum_{j=N_1}^{N_2} \mu_{2^j t} * \eta_{\ell,j} * f \Big\|_p^p dt \Big)^{1/p} + 2^{-\ell} \Big(\int_1^2 \Big\| \frac{\partial}{\partial t} \sum_{j=N_1}^{N_2} \mu_{2^j t} * \eta_{\ell,j} * f \Big\|_p^p dt \Big)^{1/p} \lesssim 2^{-\ell/p - \ell a(p)} \|f\|_p \quad (7.45)$$

where $a(p) > 0$ for $2 < p < \infty$. By Littlewood–Paley theory we see that the bound for the first term in (7.45) reduces to

$$\Big(\int_1^2 \Big\| \Big(\sum_{j=N_1}^{N_2} |\mu_{2^j t} * \eta_{\ell,j} * f_j|^2 \Big)^{1/2} \Big\|_p^p dt \Big)^{1/p} \lesssim 2^{-\ell/p - \ell a(p)} \Big\| \Big(\sum_j |f_j|^2 \Big)^{1/2} \Big\|_p \quad (7.46)$$

for $2 < p < \infty$. (7.46) is established by a local smoothing argument as in [49] (see in particular an isotropic version of Corollary 3.6 of that paper). We thus have established the bound for the first term in (7.45), and the argument for the second term is analogous. Finally from (7.45) we obtain (7.44) by another application of Littlewood–Paley theory (applying the inequality to functions f_j with $\widehat{f_j}$ supported where $|\xi| \approx 2^{\ell-j}$). □

7.5. **On radial Fourier multipliers.** We consider radial Fourier multipliers on $\widehat{\mathbb{R}}^d$ with $d \geq 2$, of the form $m(\xi) = h(|\xi|)$ where h satisfies the condition $\sup_{t>0} \|\beta h(t\cdot)\|_{L^2_\alpha} < \infty$ for suitable α; here L^2_α is the usual Sobolev space on the real line and β is any nontrivial C_c^∞ function with compact support in $(0, \infty)$. By duality we only need to consider the range $p \leq 2$.

The inequality

$$\|h(|\cdot|)\|_{M^{p,q}} \lesssim \sup_{t>0} t^{d(\frac{1}{p}-\frac{1}{q})} \|\beta h(t\cdot)\|_{L^2_\alpha}, \quad \alpha > d(1/q - 1/2) \quad (7.47)$$

is known to hold for $1 < p < \frac{2(d+1)}{d+3}$, $p \leq q \leq 2$ and one may conjecture that it holds for $\frac{2(d+1)}{d+3} < p \leq \frac{2d}{d+1}$ and $p \leq q < \frac{d-1}{d+1}p'$. Indeed, as a straightforward consequence of the Stein–Tomas restriction theorem and Littlewood–Paley

theory one gets for the endpoint $p = \frac{2(d+1)}{d+3}$, $q = 2$, $\alpha = 0$, a complete characterization of radial Fourier multipliers in $M^{p,2}$; namely

$$\|h(|\cdot|)\|_{M^{p,2}} \approx \sup_{t>0} t^{d(\frac{1}{p}-\frac{1}{2})} \Big(\int_t^{2t} |h(s)|^2 \frac{ds}{s} \Big)^{1/2},$$

see *e.g.* [46]. The case $p = q = \frac{2d}{d-1}$ has been settled only in two dimensions in [22, 25], but remains open in three and higher dimensions. Note that as a special case one has the Bochner–Riesz conjecture when $h(s) = (1-s^2)_+^\lambda$. For partial $L^p \to L^p$ results in higher dimensions (via the connection [25] with Stein's square function) we refer to [27, 97, 70, 69], *cf.* §7.6.2 below.

We formulate sparse bounds for the multipliers satisfying (7.47); in fact our hypotheses will involve the single scale variant

$$\|g(|\cdot|)\|_{M^{p,q}} \le C(\alpha)\|g\|_{L^2_\alpha},\ \alpha > d(1/q-1/2),\ \operatorname{supp}(g) \subset [1/2,2]. \tag{7.48}$$

Typically, the assumption (7.48) will be applied to g of the form $\beta h(t\cdot)$.

Theorem 6.1 leads to the following result.

Proposition 7.15. *Let $1 < p \le q \le 2$ and let T_h be the convolution operator with multiplier $h(|\cdot|)$. Then*

(i) Assume (7.48) *holds for a specific exponent pair (p,q) with $1 < p \le \frac{2d}{d+1}$, $p \le q \le \min\{\frac{d-1}{d+1}p', 2\}$, and all $\alpha > d(1/q-1/2)$. Then*

$$\|T_h\|_{\mathrm{Sp}_\gamma(p,q')} \le C_b \sup_{t>0} \|\beta h(t\cdot)\|_{L^2_b}, \quad b > d(1/p-1/2). \tag{7.49}$$

(ii) In particular, (7.49) *holds true for $1 < p \le \frac{2(d+1)}{d+3}$, $p \le q \le 2$.*

Proof. We need to verify the assumptions of Theorem 6.1. This amounts to veryfing the finiteness of the condition (6.2a). Setting $g = \beta h(t\cdot)$ and fixing $b > d(1/p-1/2)$ this follows from proving that for $\ell \ge 1$ we have

$$\|g(|\cdot|) * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim 2^{-\ell(\frac{d}{p}-\frac{d}{q}+\varepsilon(b))}\|g\|_{L^2_b} \tag{7.50}$$

for some $\varepsilon(b) > 0$.

Let υ_0 be supported in $\{s \in \mathbb{R} : |s| < 1/2\}$ such that $\int \upsilon_0(s)ds = 1$. For $n \ge 1$ let $\upsilon_n(s) = \upsilon_0(2^n s) - \upsilon_0(2^{n-1}s)$, and define $g_n(s) = g * \upsilon_n(s)$. By assumption (7.48), we have $\|g_n(|\cdot|)\|_{M^{p,q}} \lesssim \|g_n\|_{L^2_\alpha}$ for any $\alpha > d(1/q-1/2)$ and hence also

$$\|g_n(|\cdot|) * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim \|g_n(|\cdot|)\|_{M^{p,q}} \lesssim \|g_n\|_{L^2_\alpha} \lesssim 2^{-n(b-\alpha)}\|g\|_{L^2_b}. \tag{7.51}$$

For our fixed choice of $b > d(1/p-1/2)$, we choose $b_1 < b$ such that $d(1/p-1/2) < b_1 < b$, so that (7.51) holds for the choice $\alpha = b_1 - d(1/p-1/q) > d(1/q-1/2)$.

We let $\epsilon = (b-b_1)(2d)^{-1}$ and use (7.51) for $n \ge \ell(1+\epsilon)^{-1}$. Since

$$\frac{d(\frac{1}{p}-\frac{1}{q})+b-b_1}{1+\epsilon} - d(\frac{1}{p}-\frac{1}{q}) = \frac{b-b_1-\epsilon d(\frac{1}{p}-\frac{1}{q})}{1+\epsilon} \le \frac{b-b_1}{1+\epsilon} := \varepsilon(b) > 0$$

we get

$$\sum_{n\geq \frac{\ell}{1+\epsilon}} \|g_n(|\cdot|)*\widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim 2^{-\frac{\ell}{1+\epsilon}(\frac{d}{p}-\frac{d}{q}+b-b_1)}\|g\|_{L^2_b} \lesssim 2^{-\ell(\frac{d}{p}-\frac{d}{q}+\varepsilon(b))}\|g\|_{L^2_b}$$

For $n \leq \ell/(1+\epsilon)$ we observe that any derivative of order k of $g_n(|\cdot|)$ is $O(2^{kn}\|g\|_1)$ and an N-fold integration by parts gives $|\mathcal{F}^{-1}[g_n(|\cdot|)](x)| \leq C_N 2^{(n-\ell)N}$ for $|x| \approx 2^\ell$, for all $N \in \mathbb{N}$. We use this with $N := 10d\frac{1+\epsilon}{\epsilon}$. By Young's inequality

$$\sum_{n\leq \frac{\ell}{1+\varepsilon_1}} \|\mathcal{F}^{-1}[g_n(|\cdot|)*\widehat{\Psi}_\ell]\|_{M^{p,q}} \lesssim C_N \ell 2^{\ell(d-d(\frac{1}{p}-\frac{1}{q}))}2^{-\ell(\frac{\epsilon}{1+\epsilon}N)} \lesssim C(\epsilon)2^{-8d\ell}$$

and (7.50) is verified. □

As an example in the above class of multiplier transformations we consider a multi-scale version of Bochner–Riesz operators. The Bochner–Riesz means of the Fourier integral are defined by

$$\widehat{S_t^\lambda f}(\xi) = (1-|\xi|^2/t^2)_+^\lambda \widehat{f}(\xi) \tag{7.52}$$

and are conjectured to be bounded from $L^p \to L^q$ if $\lambda > d(1/q-1/2)-1/2$ and $1 \leq p \leq q \leq \min\{\frac{d-1}{d+1}p', 2\}$, with operator norm $O(t^{d(1/p-1/q)})$. One may reduce to $t=1$ by scaling, and if $h_\lambda(s) = (1-s^2)_+^\lambda$ then $h_\lambda \in L^2_\nu$ for $\lambda > \nu + 1/2$. Therefore, Proposition 7.15 immediate leads to sparse bounds for operators such as $\sum_{k=-\infty}^{\infty} \pm\big(S^\lambda_{2^k} - S^\lambda_{2^{k+1}}\big)$. with uniform bounds in the choice of the sequence of signs. After a standard averaging argument using Rademacher functions this implies sparse bounds for lacunary square functions The vector-valued version of Theorem 6.1 leads to sparse domination for the lacunary square-function

$$\Big(\sum_{k\in\mathbb{Z}} |S^\lambda_{2^k}f - S^\lambda_{2^{k+1}}f|^2\Big)^{1/2}$$

and consequently to sparse bounds for lacunary Bochner–Riesz maximal functions $M_\lambda f = \sup_{k\in\mathbb{Z}} |S^\lambda_{2^k}f|$. These results can be viewed as a natural multi-scale generalization of the sparse domination results for Bochner–Riesz means in [16, 68]. In this context, we remark that there are sharper endpoint sparse domination result for Bochner–Riesz means [61] which yield back some of the known weak type (p,p) endpoint bounds for

$$\lambda = d(1/p-1/2)-1/2.$$

However, currently there is no sparse bound for analogous endpoints which involve multiple frequency scales. We intend to return to this question in the future.

7.6. **Stein's square function.** In [102] Stein introduced the square function defined via Bochner–Riesz means by,

$$\begin{aligned} G^\alpha f(x) &= \Big(\int_0^\infty \Big| \frac{\partial S_t^\alpha f(x)}{\partial t} \Big|^2 t\, dt \Big)^{1/2} \\ &= c_\alpha \Big(\int_0^\infty |S_t^{\alpha-1} f(x) - S_t^\alpha f(x)|^2 \frac{dt}{t} \Big)^{1/2} \end{aligned}$$

in order to establish pointwise convergence and strong summability results. Another important connection was established in [25], namely that an L^p-boundedness result for G^α implies that the condition $\sup_{t>0} \|\beta h(t\cdot)\|_{L^2_\alpha} < \infty$ is sufficient for $h(|\cdot|) \in M_p$. Moreover, G^α also controls maximal operators associated to radial Fourier multipliers [23].

The expression $G^\alpha f(x)$ is almost everywhere equivalent to many alternative square functions, which can be obtained via versions of Plancherel's theorem with respect to the t-variable; see the paper by Kaneko and Sunouchi [60]. We distinguish the cases $1 < p \le 2$, in which by a result of Sunouchi [109] we have L^p boundedness for $\alpha > d(1/p - 1/2) + 1/2$, and the more subtle case $2 \le p < \infty$, where $L^p(\mathbb{R}^d)$ boundedness for $d \ge 2$ is conjectured for $p > \frac{2d}{d-1}$ and $\alpha > d(1/2 - 1/p)$, and known if $d = 2$ [22]. L^p boundedness in the latter problem is closely related to the multiplier problem discussed in §7.5; see [27, 97, 70, 69] for partial results and [71, 72] for certain endpoint and weighted bounds.

We recall some basic decompositions of the Bochner–Riesz means. One splits

$$(1 - |\xi|^2)_+^{\alpha-1} - (1 - |\xi|^2)_+^\alpha = \sum_{n \ge 0} 2^{-n(\alpha-1)} u_n(|\xi|),$$

where $u_0(0) = 0$, the u_n are smooth, and for $n \ge 1$ we have

$$\operatorname{supp}(u_n) \subset (1 - 2^{-n+1}, 1 - 2^{-n-1})$$

and $|\frac{d^j}{ds^j} u_n(s)| \le C_j 2^{nj}$ for $j \in \mathbb{N}_0$. Let $K_n = \mathcal{F}^{-1}[u_n(|\cdot|)]$, $K_{n,s} = s^d K_n(s\cdot)$ and

$$G_n f(x) = \Big(\int_0^\infty |K_{n,s} * f(x)|^2 \frac{ds}{s} \Big)^{1/2},$$

so that

$$G^\alpha f(x) \lesssim \sum_{n=0}^\infty 2^{-n(\alpha-1)} G_n f(x).$$

We shall rely on the standard pointwise estimates obtained by stationary phase calculations,

$$|K_n(x)| \lesssim_N (1 + |x|)^{-\frac{d+1}{2}} (1 + 2^{-n}|x|)^{-N}. \tag{7.53}$$

7.6.1. *The case* $1 < p \le 2$. A pointwise sparse domination result for $\alpha > (d+1)/2$ was proved by Carro and Domingo-Salazar [26]. For $1/2 < \alpha \le (d+1)/2$ we have L^p boundedness ($p \le 2$) only in the restricted range $\frac{2d}{2\alpha+2d-1} < p \le 2$ by Sunouchi's result which is sharp. Thus in this range we are seeking sparse domination results for the forms $\langle G^\alpha f_1, f_2\rangle$. Theorem 6.1 yields the following.

Proposition 7.16. *Let* $d \ge 2$, $\frac12 < \alpha \le \frac{d+1}{2}$. *Then for* $\frac{2d}{2\alpha+2d-1} < p \le 2$ *we have the* (p,p)*-sparse domination inequality*

$$|\langle G^\alpha f, \omega\rangle| \le C\Lambda^*_{p,p}(f,\omega).$$

Proof. The operators G_n are defined through smooth kernels and therefore the result in [26] yields pointwise sparse bounds, with norms depending on n. This settles the case of small values of n. For large values of n, given $\varepsilon > 0$ we have to show

$$|\langle G_n f, \omega\rangle| \lesssim_\varepsilon 2^{n(\frac{d}{p}-\frac{d}{2}-\frac12+\varepsilon)}\, \Lambda^{\gamma,*}_{p,p}(f,\omega) \tag{7.54}$$

since in the assumed range of p we have $\alpha - 1 > d/p - d/2 - 1/2$ and therefore we can sum in n to obtain the result for G^α. Let $\mathcal{H}$ be the Hilbert space $L^2(\mathbb{R}^+, \frac{ds}{s})$. By the linearization argument in §4.3, the inequality (7.54) follows, for a scalar function f_1 and an $\mathcal{H}$-valued function $f_2 = \{f_{2,s}\}$, from

$$\Big|\iint K_{n,s} * f_1(x) f_{2,s}(x) \frac{ds}{s} dx\Big| \lesssim_\varepsilon 2^{n(\frac{d}{p}-\frac{d}{2}-\frac12+\varepsilon)}\, \Lambda^{\gamma,*}_{p,\mathbb{C},p,\mathcal{H}^*}(f_1,f_2).$$

By Theorem 6.1 this follows from

$$\sup_{t>0} \big\|[\beta(|\cdot|)u_n(t|\cdot|)] * \widehat{\Psi}_\ell\big\|_{M^{p,p'}_{\mathbb{C},\mathcal{H}^*}} 2^{\ell(\frac{d}{p}-\frac{d}{p'}+\epsilon_1)} \lesssim_\varepsilon 2^{n(\frac{d}{p}-\frac{d}{2}-\frac12+\varepsilon)} \tag{7.55}$$

for some $\epsilon_1 > 0$. To verify (7.55) we argue by interpolation and reduce to the cases $p = 2$ and $p = 1$. It will be helpful to observe that for $1/8 \le t/s \le 8$ we can replace the kernel K_n on the left hand side of the inequality in (7.53) by $\mathcal{F}^{-1}[\beta(|\cdot|)u_n(\frac{t}{s}|\cdot|)]$. Thus

$$\Big(\int |\mathcal{F}^{-1}[\beta(|\cdot|)u_n(\tfrac{t}{s}|\cdot|)](x)|^2 \tfrac{ds}{s}\Big)^{1/2} \lesssim_N (1+|x|)^{-\frac{d+1}{2}}(1+2^{-n}|x|)^{-N}. \tag{7.56}$$

Here we used that given t the integrand is zero unless $s \in [t/2, 2t]$. Hence for any $\epsilon_2 > 0$ (which we choose $\ll \min\{\varepsilon, \epsilon_1\}$) we get

$$\sup_{t>0} \big\|[\beta(|\cdot|)u_n(t|\cdot|)] * \widehat{\Psi}_\ell\big\|_{M^{p,q}_{\mathbb{C},\mathcal{H}^*}} \lesssim_{\epsilon_2,N} 2^{-\ell N} \quad \text{if } \ell > n(1+\epsilon_2). \tag{7.57}$$

If $p = 2$ we have also have, for $\ell \le n(1+\epsilon_2)$,

$$\begin{aligned}\big\|[\beta(|\cdot|)u_n(t|\cdot|)] * \widehat{\Psi}_\ell\big\|_{M^{2,2}_{\mathbb{C},\mathcal{H}^*}} 2^{\ell\epsilon_1} &\lesssim 2^{\ell\epsilon_1} \sup_\xi \Big(\int |\beta(|\xi|)u_n(\tfrac{t}{s}|\xi|)|^2 \tfrac{ds}{s}\Big)^{1/2} \\ &\lesssim 2^{\ell\epsilon_1} 2^{-n/2} \lesssim 2^{n\epsilon_1(1+\epsilon_2)} 2^{-n/2}.\end{aligned} \tag{7.58}$$

Furthermore, for $p=1$ and $\ell \le n(1+\epsilon_2)$ we use (7.56) to see that

$$\begin{aligned}&\big\|[\beta(|\cdot|)u_n(t|\cdot|)] * \widehat{\Psi}_\ell\big\|_{M^{1,\infty}_{\mathbb{C},\mathcal{H}^*}} 2^{\ell(d+\epsilon_1)} \\ &\lesssim 2^{\ell(d+\epsilon_1)} 2^{-n(d+1)/2} \lesssim 2^{n(1+\epsilon_2)(d+\epsilon_1)-(d+1)/2} \lesssim 2^{n(d-1)/2+\varepsilon}.\end{aligned} \tag{7.59}$$

Combining (7.58), (7.59) with (7.57) we obtain the cases of (7.55) for $p=1$ and $p=2$ and (7.55) follows by interpolation for $1 \le p \le 2$. □

7.6.2. *The case* $2<p<\infty$. The reduction to sparse bounds will be similar as in the case $p \le 2$, but the input information is more subtle. Instead of the pointwise bounds (7.53) we now use that

$$\Big\|\Big(\int_{1/8}^{8} |K_{n,s} * f|^2 \tfrac{ds}{s}\Big)^{1/2}\Big\|_p \lesssim 2^{n(\frac{d}{2}-\frac{d}{r}-1)}\|f\|_r, \tag{7.60}$$

for $2 \le r \le p$, $p \ge \frac{2(d+1)}{d-1}$, which was proved using the Stein–Tomas restriction theorem [27, 97, 71]. We then obtain a satisfactory result for $\alpha \ge \frac{d}{d+1}$.

Proposition 7.17. *Let $d \ge 2$, $\frac{d}{d+1} \le \alpha \le \frac{d}{2}$. Then for $\frac{2d}{d-2\alpha} < p < \infty$ we have the $(2,p)$-sparse domination inequality*

$$|\langle G^\alpha f, \omega\rangle| \le C\Lambda^*_{2,p}(f,\omega).$$

Proof. Note that in the given p-range, $p > \frac{2(d+1)}{d-1}$ when $\alpha \ge \frac{d}{d+1}$. We use the notation in the proof of the preceding proposition. By linearization (see the argument in §4.3) it suffices to prove

$$\Big|\iint K_{n,s} * f_1(x) f_{2,s}(x) \frac{ds}{s} dx\Big| \lesssim_\varepsilon 2^{n(\frac{d}{2}-\frac{d}{p}-1+\varepsilon)}\, \Lambda^{\gamma,*}_{2,\mathbb{C},p,\mathcal{H}^*}(f_1,f_2)$$

and by Theorem 6.1 this follows, given $\varepsilon>0$ from

$$\big\|[\beta(|\cdot|)u_n(t|\cdot|)] * \widehat{\Psi}_\ell\big\|_{M^{2,p}_{\mathbb{C},\mathcal{H}^*}} 2^{\ell(\frac{d}{2}-\frac{d}{p}+\epsilon_1)} \lesssim_\varepsilon 2^{n(\frac{d}{2}-\frac{d}{p}-1+\varepsilon)} \tag{7.61}$$

for some $\epsilon_1>0$, uniformly in $t>0$. For $\ell < n(1+\epsilon_2)$ the left hand side is bounded by

$$2^{n(1+\epsilon_2)(\frac{d}{2}-\frac{d}{p}+\epsilon_1)}\big\|\beta(|\cdot|)u_n(t|\cdot|)\big\|_{M^{2,p}_{\mathbb{C},\mathcal{H}^*}}. \tag{7.62}$$

Using (7.60) for $r=2$ we get

$$\Big\|\Big(\int_{t/8}^{8t} |\mathcal{F}^{-1}[\beta(|\cdot|)u_n(\tfrac{t}{s}|\cdot|)\widehat{f}]|^2 \tfrac{ds}{s}\Big)^{1/2}\Big\|_p \lesssim 2^{-n}\|f\|_2$$

and thus the expression in (7.62) is $O(2^{n(1+\epsilon_2)(\frac{d}{2}-\frac{d}{p}+\epsilon_1)-n})$. Finally, we choose $\epsilon_1,\epsilon_2 \ll \varepsilon$ and combine this with the error estimates (7.57). This completes the verification of (7.61). □

Appendix A. Facts about sparse domination

For completeness, we collect a number of auxiliary results, some of them well-known, about sparse domination.

A.1. **Replacing simple functions.** It is often convenient to replace the spaces S_{B_1} and $\mathrm{S}_{B_2^*}$ in the definition of the $\mathrm{Sp}_\gamma(p_1, B_1, p_2, B_2^*)$ norms by other suitable test function classes such as the spaces of compactly supported C^∞ functions or Schwartz functions. This is justified by the following Lemma.

Lemma A.1. *Suppose $1 \le p_1 < p_2'$ and $p_1 < p < p_2'$ and let $T \in \mathrm{Sp}_\gamma(p_1, p_2)$. Let $\mathcal{V}_1$ be a dense subspace of $L^p_{B_1}$ and $\mathcal{V}_2$ be a dense subspace of $L^{p'}_{B_2^*}$. Then*

$$\|T\|_{\mathrm{Sp}_\gamma(p_1,B_1,p_2,B_2^*)} = \sup\Big\{\frac{|\langle Tf_1, f_2\rangle|}{\Lambda^*_{p_1,B_1,p_2,B_2^*}(f_1,f_2)} : f_i \in \mathcal{V}_i,\ f_i \neq 0,\ i = 1,2\Big\}. \tag{A.1}$$

Proof. We first assume that $\mathcal{V}_1 = L^p_{B_1}$ and $\mathcal{V}_2 = L^{p'}_{B_2^*}$. In what follows we omit the reference to B_1, B_2^*. The right-hand side of (A.1) dominates $\|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}$, defined in (1.5). In order to verify the reverse inequality we have to show that given $\epsilon > 0$ and given $f_1 \in L^p$, $f_2 \in L^{p'}$ we have the inequality

$$|\langle Tf_1, f_2\rangle| \le (\|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)} + \epsilon)\, \Lambda^*_{p_1,p_2}(f_1,f_2). \tag{A.2}$$

This is clear if one of the f_i is zero almost everywhere. We may thus assume that $\|f_1\|_p > 0$, $\|f_2\|_{p'} > 0$. For any $\varepsilon_1 > 0$ choose $g_1 \in \mathrm{S}_{B_1}$, $g_2 \in \mathrm{S}_{B_2^*}$ so that $\|f_1 - g_1\|_p \le \epsilon_1$, $\|f_2 - g_2\|_{p'} \le \epsilon_1$, and also $\|g_1\|_p \le 2\|f_1\|_p$, $\|g_2\|_{p'} \le 2\|f_2\|_{p'}$ and estimate, using the definition of $\|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}$,

$$\begin{aligned}
&|\langle Tf_1, f_2\rangle| \le |\langle T[f_1 - g_1], f_2\rangle| + |\langle Tf_1, f_2 - g_2\rangle| + |\langle Tg_1, g_2\rangle| \\
&\le \|T\|_{p-p}\big(\|f_1 - g_1\|_p\|f_2\|_{p'} + \|f_1\|_p\|f_2 - g_2\|_{p'}\big) + \|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}\Lambda^*_{p_1,p_2}(g_1,g_2) \\
&\le \|T\|_{p-p}\big(\epsilon_1\|f_2\|_{p'} + \|f_1\|_p\epsilon_1\big) + \|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}\Lambda^*_{p_1,p_2}(g_1,g_2).
\end{aligned}$$

Moreover, for $p_1 < p < p_2'$, one has using (ii) in Lemma A.2. that

$$\begin{aligned}
&\Lambda^*_{p_1,p_2}(g_1,g_2) \le \Lambda^*_{p_1,p_2}(g_1 - f_1, g_2) + \Lambda^*_{p_1,p_2}(f_1, g_2 - f_2) + \Lambda^*_{p_1,p_2}(f_1,f_2) \\
&\le C_1(p,p_1,p_2)(\|g_1 - f_1\|_p\|g_2\|_{p'} + \|f_1\|_p\|g_2 - f_2\|_{p'}) + \Lambda^*_{p_1,p_2}(f_1,f_2) \\
&\le C_1(p,p_1,p_2)(2\|f_2\|_{p'}\epsilon_1 + \|f_1\|_p\epsilon_1) + \Lambda^*_{p_1,p_2}(f_1,f_2)
\end{aligned}$$

and thus

$$|\langle Tf_1, f_2\rangle| \le \|T\|_{\mathrm{Sp}_\gamma(p_1,p_2)}\, \Lambda^*_{p_1,p_2}(f_1,f_2) + \mathcal{E},$$

with $\mathcal{E} \le C(f_1, f_2, p, p_1, p_2, T)\epsilon_1$. Choosing a suitable ϵ_1 depending on ϵ we obtain the assertion (A.2), for the case $\mathcal{V}_1 = L^p$, $\mathcal{V}_2 = L^p$.

In the general case we replace the couple of pairs $(\mathrm{S}_{B_1}, \mathrm{S}_{B_2^*})$ and $(L^p, L^{p'})$ by the couple of pairs $(\mathcal{V}_1, \mathcal{V}_2)$ and $(L^p, L^{p'})$ and see that a repetition of the above arguments settles this case as well. $\square$

A.2. **The Hardy–Littlewood maximal function.** It is a well-known fact that the Hardy–Littlewood maximal operator, denoted by $\mathcal{M}$, satisfies a sparse domination inequality. We have not been able to identify the original reference for this fact and refer to Lerner's expository lecture [74] instead. This constitutes a first nontrivial example for sparse domination and we include a standard proof for completeness.

Lemma A.2. *Let $f \in L^1_{loc}(\mathbb{R}^d)$. Then there exist γ-sparse families $\mathfrak{S}_i(f)$, $i = 1, \dots, 3^d$, such that*

$$\mathcal{M}f(x) \le 2^d(1-\gamma)^{-1} \sum_{i=1}^{3^d} \sum_{Q \in \mathfrak{S}_i(f)} \langle f \rangle_{Q,1} \mathbb{1}_Q(x).$$

Proof. Let $\mathfrak{D}$ be a dyadic lattice and let $\mathcal{M}^{\mathfrak{D}}$ denote the dyadic maximal function associated to $\mathfrak{D}$, that is, $\mathcal{M}^{\mathfrak{D}} f(x) := \sup_{\substack{Q \ni x \\ Q \in \mathfrak{D}}} \langle f \rangle_Q$. Fix $a \in \mathbb{R}$ a constant to be determined later, and for each $k \in \mathbb{Z}$, define the sets

$$\Omega_k := \{x \in \mathbb{R}^d : \mathcal{M}^{\mathfrak{D}} f(x) > a^k\} = \bigcup_{Q_k^j \in \mathfrak{D}} Q_k^j$$

where $\{Q_k^j\}_j \subseteq \mathfrak{D}$ are the maximal disjoint dyadic cubes for Ω_k, that is,

$$a^k \le \langle f \rangle_{Q_k^j} \le a^k 2^d. \tag{A.3}$$

Define the sets $E_k^j := Q_k^j \backslash \Omega_{k+1}$, and note that the family of sets $\{E_k^j\}_{k,j}$ is pairwise disjoint and $|E_k^j| > (1 - \frac{2^d}{a})|Q_k^j|$; for the last claim, note that

$$\begin{aligned} |Q_k^j \cap \Omega_{k+1}| &= \sum_i |Q_k^j \cap Q_{k+1,i}| = \sum_{i\,:\,Q_{k+1}^i \subset Q_k^j} |Q_{k+1}^i| \\ &< \sum_{i\,:\,Q_{k+1}^i \subset Q_k^j} \frac{1}{a^{k+1}} \int_{Q_{k+1}^i} |f| \le \frac{1}{a^{k+1}} \int_{Q_k^j} |f| \le \frac{2^d}{a} |Q_k^j|, \end{aligned}$$

using the disjointness of the cubes Q_{k+1}^i and (A.3) for Q_{k+1}^i and Q_k^j. Thus, choosing $a = 2^d(1-\gamma)^{-1}$, $\mathfrak{S}(f) := \{Q_k^j\}_{k,j}$ is a γ-sparse family, and moreover

$$\begin{aligned} \mathcal{M}^{\mathfrak{D}} f(x) &= \sum_{k \in \mathbb{Z}} \mathcal{M}^{\mathfrak{D}} f(x) \mathbb{1}_{\Omega_k \backslash \Omega_{k+1}}(x) \le \sum_{k \in \mathbb{Z}} \sum_{Q_j^k} a^{k+1} \mathbb{1}_{E_k^j}(x) \\ &\le a \sum_{k,j} \langle f \rangle_{Q_k^j} \mathbb{1}_{Q_k^j}(x) = a \sum_{Q \in \mathfrak{S}(f)} \langle f \rangle_Q \mathbb{1}_Q(x) \end{aligned}$$

by (A.3). Finally, the result for the maximal function $\mathcal{M}$ follows from the 3^d-trick (see [79, Theorem 3.1]), which ensures that there exist 3^d dyadic

lattices $\mathfrak{D}_i$, $i = 1, \dots, 3^d$ such that

$$\mathcal{M}f(x) \le \sum_{i=1}^{3^d} \mathcal{M}^{\mathfrak{D}_i} f(x). \qquad \square$$

Remark. If f has values in a Banach space B, the same argument applies to $f \mapsto \mathcal{M}(|f|_B)$. However, there are more interesting vector-valued extensions such as in the Fefferman–Stein theorem [42], and corresponding general sparse domination results with additional hypotheses on the Banach space are discussed in a paper by Hänninen and Lorist [52].

A.3. **Operators associated with dilates of Schwartz functions.** It is convenient in many applications to observe that maximal functions and variation operators generated by convolution operators with Schwarz functions satisfy sparse bounds. We choose to deduce the variational statements as a consequence of our Theorems in §5, but it could also be based *e.g.* on [37]. For the definition of the dyadic and short variation operators we refer to Remark 7.3; here V^r_{sh} is understood with $E = (0, \infty)$.

Lemma A.3. *Let $K \in C^2(\mathbb{R}^d)$ be a convolution kernel satisfying, for all multiindices $\alpha \in \mathbb{N}_0^d$ with $\sum_i |\alpha_i| \le 2$,*

$$|\partial^\alpha K(x)| \le (1 + |x|)^{-d-2}.$$

Let $K_t(x) = t^{-d} K(t^{-1}x)$ and let $\mathcal{K}_t f = K_t * f(x)$. Then for $1 < p \le q < \infty$

$$\begin{aligned}
&|\langle \sup_{t>0} |\mathcal{K}_t f|, \omega\rangle| \lesssim \Lambda^*_{p,q'}(f, \omega)\\
&|\langle V^r_{\mathrm{dyad}} \mathcal{K} f, \omega\rangle| \lesssim \Lambda^*_{p,q'}(f, \omega),\ 2 < r < \infty,\\
&|\langle V^r_{\mathrm{sh}} \mathcal{K} f, \omega\rangle| \lesssim \Lambda^*_{p,q'}(f, \omega),\ 2 \le r < \infty.
\end{aligned}$$

Proof. Since $\sup_{t>0} |\mathcal{K}_t * f|$ is pointwise dominated by the Hardy-Littlewood maximal function the sparse bound for $\langle \sup_{t>0} |\mathcal{K}_t f|, \omega\rangle$ can be directly deduced from the sparse bound for the Hardy–Littlewood maximal function $\mathcal{M}$ in §A.2.

For the variation norm inequalities we decompose $\mathcal{K} = \sum_{n=0}^\infty \mathcal{K}^n$ where $\mathcal{K}^n$ denotes convolution with $K^n := K\Psi_n$ (here Ψ_n is supported where $|x| \approx 2^n$ when $n > 0$, see (6.1)). We can form the long and short variation operators with respect to the family of operators $\{\mathcal{K}^n_t\}_{t>0}$ where $\mathcal{K}^n_t$ denotes convolution with $K^n_t := t^{-d} K^n(t^{-1}\cdot)$. Using the pointwise bound on ∇K and results in [58] or [59] we have

$$\begin{aligned}
\|V^r_{\mathrm{dyad}} \mathcal{K}^n f\|_p &\lesssim_p 2^{-n} \|f\|_p, \quad 1 < p \le \infty,\ r > 2,\\
\|V^r_{\mathrm{sh}} \mathcal{K}^n f\|_p &\lesssim_p 2^{-n} \|f\|_p, \quad 1 < p \le \infty,\ r \ge 2.
\end{aligned}$$

The kernel K^n_t is supported in $\{x : |x| \le 2^{n-1} t\}$ and from our assumptions it is easy to see that the rescaled estimate

$$\big\| V^r_{[1,2]} \{2^{nd} K^n_t (2^n \cdot) * f\} \big\|_q \lesssim 2^{-n} \|f\|_p, \quad 1 \le p \le q \le \infty,\ 1 \le r \le \infty,$$

holds. Moreover, using the bound for ∇K and $\nabla^2 K$ we also get

$$\big\|V^r_{[1,2]}\{2^{nd}K^n_t(2^n\cdot)*\Delta_h f\}\big\|_q \lesssim 2^{n(1/r-1)}|h|^{1/r}\|f\|_p$$

for $1\le p\le q\le\infty$, $1<r\le\infty$. Applications of Theorem 5.2 and Theorem 5.1 (for ℓ^r-sums and with the choice of B_2 a subspace of $V^r_{[1,2]}$ of large finite dimension) together with the monotone convergence theorem yield

$$|\langle V^r_{\mathrm{dyad}}\mathcal{K}^n f,\omega\rangle| \lesssim 2^{-n}\Lambda^*_{p,q'}(f,\omega),\ r>2,$$
$$|\langle V^r_{\mathrm{sh}}\mathcal{K}^n f,\omega\rangle| \lesssim 2^{-n(1-1/r)}\Lambda^*_{p,q'}(f,\omega),\ r\ge 2.$$

The proof is completed by summation in n. □

Appendix B. Sparse domination: Cases where $p=1$ or $q=\infty$

Here we describe analogues of our main result Theorem 1.1 which cover cases where $p=1$ or $q=\infty$; we refer to Remark (iv) following the statement of Theorem 1.1 for an explanation of why these cases need to be treated separately. We formulate three different results, one for $p=1$, $q<\infty$, one for $q=\infty$, $p>1$, and one for $p=1$, $q=\infty$. This allows us to recover the classical case of Calderón–Zygmund operators, although we do not claim universality of sparse-domination results here: for example, we do not recover the sparse domination for Carleson-type operators from [38, 9, 8], neither the works for $p=1$ by Conde-Alonso, Culiuc, Di Plinio and Ou [29] and by Lerner [78] which also treat results on rough singular integrals, nor the works for $q=\infty$ which can often be upgraded to stronger pointwise sparse domination results of the type (1.1) (see in particular [77], [80], [85]).

We will sketch the proofs of our results, indicating only what modifications need to be made compared to the proof of Theorem 1.1. Theorems B.1, B.2 and B.3 below have applications to maximal operators, square functions and long variation operators (as formulated in §5) similar to those of Theorem 1.1. We leave the details to the interested reader.

B.1. **The case $p=1, q<\infty$.** If $p=1$, one can drop the condition of weak type $(1,1)$. Our variant of Theorem 1.1 is then as follows.

Theorem B.1. *Let $1\le q<\infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} such that*

- *the support condition* (1.6) *holds,*
- *the restricted strong type (q,q) condition* (1.7b) *holds,*
- *the single scale $(1,q)$ condition* (1.8) *holds,*
- *the single scale ε-regularity conditions* (1.9a), (1.9b) *hold with $p=1$.*

Define

$$\mathcal{C}=A(q)+A_\circ(1,q)\log\big(2+\tfrac{B}{A_\circ(1,q)}\big).$$

Then, for all integers N_1, N_2 with $N_1 \le N_2$,

$$\Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{\mathrm{Sp}_\gamma(1,B_1,q',B_2^*)} \lesssim_{q,\varepsilon,\gamma,d} \mathcal{C}.$$

Proof. We argue as in the proof of Theorem 1.1, with the decomposition (4.13) and the bound (4.14). The terms (4.15a), (4.15b) are handled exactly as in the proof of Theorem 1.1. Consider the splitting of III as in (4.19). The terms (4.19c) and (4.19d) are estimated as in the proof of Theorem 1.1. We are thus only left with estimating the term

$$III_1 + III_2 = \sum_{N_1 \le j \le N_2} \sum_{\substack{W \in \mathcal{W} \\ L(W) < j}} \langle T_j b_{1,W}, g_2 \rangle.$$

The argument in the proof of Theorem 1.1 does no longer work; recall that for $p > 1$ these terms were bounded immediately via the weak type (p,p) condition (1.7a) and the duality of $L^{p,\infty}_{B_2^{**}}$ and $L^{p',1}_{B_2^*}$. Instead, here we will bound $III_1 + III_2$ using (1.8) and the regularity condition (1.9a), close in spirit to the bounds of the terms IV_1 and IV_2 (defined in (4.25)) in the proof of Theorem 1.1.

We let $0 < \varepsilon' < \min\{1/q', \varepsilon\}$ and $\ell > 0$ be as in (4.24), that is,

$$2^\ell < \frac{100}{\epsilon'}\big(2 + \frac{B}{A_\circ(1,q)}\big) \le 2^{\ell+1}.$$

Let $\mathfrak{R}_j$ be the collection of dyadic subcubes of Q_0 of side length 2^j. We tile Q_0 into such cubes and write

$$III_1 + III_2 = \sum_s \widetilde{III}_s,$$

where

$$\widetilde{III}_s = \sum_{N_1 \le j \le N_2} \sum_{R \in \mathfrak{R}_j} \Big\langle \sum_{\substack{W \subset R \\ L(W) = j-s}} T_j b_{1,W}, g_2 \mathbb{1}_{3R} \Big\rangle. \tag{B.1}$$

We first note

$$\|T_j\|_{L^1_{B_1} \to L^q_{B_2}} \lesssim A_\circ(1,q) 2^{-jd(1-\frac{1}{q})}, \tag{B.2a}$$

$$\|T_j(\mathrm{I} - \mathbb{E}_{s-j})\|_{L^1_{B_1} \to L^q_{B_2}} \lesssim_\varepsilon B 2^{-jd(1-\frac{1}{q})} 2^{-\varepsilon' s}. \tag{B.2b}$$

where (B.2a) follows from the single scale $(1,q)$ condition (1.8) and (B.2b) follows from the single scale ε-regularity condition (1.9a) and Corollary 3.5.

For $L(W) = j - s$, we have $b_{1,W} = (I - \mathbb{E}_{s-j}) f_{1,W}$. Let $R \in \mathfrak{R}_j$. By (B.2a) we get

$$\begin{aligned}
&\Big|\Big\langle \sum_{\substack{W\subset R\\ L(W)=j-s}} T_j b_{1,W}, g_2 \mathbb{1}_{3R}\Big\rangle\Big| \\
&\le \Big\| T_j \sum_{\substack{W\subset R\\ L(W)=j-s}} b_{1,W}\Big\|_{L^q_{B_2}} \big\| g_2 \mathbb{1}_{3R}\big\|_{L^{q'}_{B_2^*}} \\
&\lesssim A_\circ(1,q) |R|^{-(1-\frac{1}{q})} \sum_{\substack{W\subset R\\ L(W)=j-s}} \|b_{1,W}\|_{L^1_{B_1}} |R|^{1/q'} \langle f_2\rangle_{3Q_0,q'} \\
&\lesssim A_\circ(1,q) \sum_{\substack{W\subset R\\ L(W)=j-s}} |W| \langle f_1\rangle_{Q_0,1} \langle f_2\rangle_{3Q_0,q'}, \qquad \text{(B.3a)}
\end{aligned}$$

and by (B.2b),

$$\begin{aligned}
&\Big|\Big\langle \sum_{\substack{W\subset R\\ L(W)=j-s}} T_j b_{1,W}, g_2 \mathbb{1}_{3R}\Big\rangle\Big| \\
&\le \Big\| T_j (I - \mathbb{E}_{s-j}) \sum_{\substack{W\subset R\\ L(W)=j-s}} f_{1,W}\Big\|_{L^q_{B_2}} \big\| g_2 \mathbb{1}_{3R}\big\|_{L^{q'}_{B_2^*}} \\
&\lesssim_\varepsilon B 2^{-\varepsilon' s} |R|^{-(1-\frac{1}{q})} \sum_{\substack{W\subset R\\ L(W)=j-s}} \|f_{1,W}\|_{L^1_{B_1}} |R|^{1/q'} \langle f_2\rangle_{3Q_0,q'} \\
&\lesssim B 2^{-\varepsilon' s} \sum_{\substack{W\subset R\\ L(W)=j-s}} |W| \langle f_1\rangle_{Q_0,1} \langle f_2\rangle_{3Q_0,q'}. \qquad \text{(B.3b)}
\end{aligned}$$

Note that in obtaining the above bounds we have used (4.9a) and (4.10).

In the above definition (B.1) for $\widetilde{III}_s$ we write $j = sn + i$ with $i = 1, \dots, s$ so that

$$\begin{aligned}
|\widetilde{III}_s| &\lesssim_\varepsilon \min\{A_\circ(1,q), B 2^{-\varepsilon' s}\} \\
&\qquad \sum_{i=1}^{s} \sum_{\substack{n\in\mathbb{Z}\\ sn+i\in[N_1,N_2]}} \sum_{R\in\mathfrak{R}_{sn+i}} \sum_{\substack{W\subset R\\ L(W)=sn+i-s}} |W| \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,q'}.
\end{aligned}$$

Now interchange the order of summation; here consider for fixed $W \in \mathcal{W}$ the set of all triples (R, n, i) such that $L(W) = s(n-1) + i$, $R \in \mathfrak{R}_{sn+i}$ and $W \subset 3R$, and observe that the cardinality of this set is 1. Combining this with the above estimates and summing over the disjoint cubes $W \in \mathcal{W}$ we

obtain the bound

$$|III_1 + III_2| \lesssim_{d,\gamma,\varepsilon} \langle f_1\rangle_{Q_0,1}\langle f_2\rangle_{3Q_0,q'} \sum_{W\in\mathcal{W}} |W| \sum_{s=1}^{\infty} \min\{A_\circ(1,q), B2^{-\varepsilon' s}\}$$
$$\lesssim A_\circ(1,q)\log\big(2+\frac{B}{A_\circ(1,q)}\big)|Q_0|\langle f_1\rangle_{Q_0,1}\langle f_2\rangle_{3Q_0,q'},$$

as desired. □

In the spirit of Section 2, it is possible to deduce that the sparse bound in Theorem B.1 implies that the multi-scale sums $\sum_{j=N_1}^{N_2} T_j$ are of weak-type $(1,1)$. The proof of this fact is slightly different than the one given in Section 2 for $p>1$, as it cannot rely on the duality between $L^{p,\infty}$ and $L^{p,1}$. We refer to the reader to [29, Appendix B] for details.

B.2. **The case $p>1$, $q=\infty$.** If $q=\infty$ one can drop the restricted strong type (q,q) condition (1.7b). Our variant of Theorem 1.1 is then as follows.

Theorem B.2. *Let $1<p\le\infty$. Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} such that*

- *the support condition* (1.6) *holds,*
- *the weak type (p,p) condition* (1.7a) *holds,*
- *the single scale (p,∞) condition* (1.8) *holds,*
- *the single scale ε-regularity conditions* (1.9a), (1.9b) *hold with $q=\infty$.*

Define

$$\mathcal{C} = A(p) + A_\circ(p,\infty)\log\big(2+\tfrac{B}{A_\circ(p,\infty)}\big).$$

Then, for all integers N_1, N_2 with $N_1\le N_2$,

$$\Big\|\sum_{j=N_1}^{N_2} T_j\Big\|_{\mathrm{Sp}_\gamma(p,B_1,1,B_2^*)} \lesssim_{p,\varepsilon,\gamma,d} \mathcal{C}.$$

Proof. Again we argue as in the proof of Theorem 1.1 and describe the main induction step. Using the previous notations we now decompose

$$\langle Sf_1, f_2\rangle = \mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV} \tag{B.4}$$

where

$$\mathrm{I} = \sum_{W\in\mathcal{W}} \langle S_W f_1, f_2\rangle$$
$$\mathrm{II} = \langle (S-\sum_{W\in\mathcal{W}} S_W) f_1, g_2\rangle$$
$$\mathrm{III} = \langle (S-\sum_{W\in\mathcal{W}} S_W) g_1, b_2\rangle$$
$$\mathrm{IV} = \langle (S-\sum_{W\in\mathcal{W}} S_W) b_1, b_2\rangle.$$

Note that the numbering here is slightly different from the one in the proof of Theorem 1.1. We deal with the term I using the induction hypothesis as in the proof of Theorem 1.1 and, using the argument therein, it suffices to show that the terms II, III and IV are bounded by $c\mathcal{C}|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}$.

We first consider $\mathrm{II} = \mathrm{II}_1 - \mathrm{II}_2$ where

$$\mathrm{II}_1 = \langle Sf_1, g_2\rangle, \qquad \mathrm{II}_2 = \sum_W \langle S_W f_1, g_2\rangle.$$

Here we use the weak type (p,p) assumption (1.7a) for $p>1$ and (4.11) for $r_2 = p' < \infty$ to get

$$\begin{aligned}\langle Sf_1, g_2\rangle &\le \|Sf_1\|_{L^{p,\infty}_{B_2^{**}}} \|g_2 \mathbb{1}_{3Q_0}\|_{L^{p',1}_{B_2^*}}\\ &\lesssim A(p)\|f_1\|_p \langle f_2\rangle_{3Q_0,1}|Q_0|^{1-1/p}\\ &\lesssim A(p)|Q_0|\langle f_1\rangle_{p,Q_0}\langle f_2\rangle_{3Q_0,1}\end{aligned}$$

and

$$\begin{aligned}\sum_{W\in\mathcal{W}} \langle S_W f_1, g_2\rangle &\le \sum_W \|S_W[f_1\mathbb{1}_W]\|_{L^{p,\infty}_{B_2^{**}}} \|g_2\|_{L^\infty_{B_2^*}} \|\mathbb{1}_{3W}\|_{L^{p',1}_{B_2^*}}\\ &\lesssim A(p) \sum_{W\in\mathcal{W}} \|f_1\mathbb{1}_W\|_p \langle f_2\rangle_{3Q_0,1}|W|^{1-1/p}\\ &\lesssim A(p) \sum_{W\in\mathcal{W}} |W|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}\end{aligned}$$

and hence, by the disjointness of the cubes $W\in\mathcal{W}$,

$$|\mathrm{II}| \le |\mathrm{II}_1| + |\mathrm{II}_2| \lesssim A(p)|Q_0|\langle f_1\rangle_{p,Q_0}\langle f_2\rangle_{3Q_0,1}. \tag{B.5}$$

The last term IV corresponds exactly to the sum $III_3 + III_4$ in the proof of Theorem 1.1, defined in (4.19c), (4.19d), and it is therefore treated in the same way; here the weak type and restricted strong type assumptions are not used. In particular, we obtain

$$|\mathrm{IV}| \lesssim A_\circ(p,\infty)\log\big(2 + \frac{B}{A_\circ(p,\infty)}\big)|Q_0|\langle f_1\rangle_{p,Q_0}\langle f_2\rangle_{3Q_0,1}. \tag{B.6}$$

It remains to bound the term III. By the definition of g_1 and S_W we have

$$\sum_{W\in\mathcal{W}} S_W g_1 = \sum_{W\in\mathcal{W}}\sum_{j=N_1}^{L(W)} T_j[\mathrm{av}_W[f_1]\mathbb{1}_W] = \sum_{j=N_1}^{N_2}\sum_{\substack{W\in\mathcal{W}\\ j\le L(W)}} T_j[\mathrm{av}_W[f_1]\mathbb{1}_W]$$

and thus we may split $\mathrm{III} = \mathrm{III}_1 + \mathrm{III}_2$ where

$$\begin{aligned}\mathrm{III}_1 &= \langle S[g_1\mathbb{1}_{\Omega^\complement}], b_2\rangle\\ \mathrm{III}_2 &= \sum_{j=N_1}^{N_2}\sum_{\substack{W\in\mathcal{W}\\ L(W)<j}} \langle T_j[\mathrm{av}_W[f_1]\mathbb{1}_W], b_2\rangle.\end{aligned}$$

Let $\mathfrak{R}_j$ be the collection of dyadic subcubes of Q_0 of side length 2^j. We tile Q_0 into such cubes and write

$$\mathrm{III}_1 = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \langle T_j[g_1 \mathbb{1}_{\Omega^{\complement}\cap R}], b_2\rangle = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \langle T_j[g_1 \mathbb{1}_{\Omega^{\complement}\cap R}], \sum_{W'\in\mathcal{W}} b_{2,W'}\mathbb{1}_{3R}\rangle.$$

Next, note that in order to have $\langle T_j[g_1 \mathbb{1}_{\Omega^{\complement}\cap R}], \sum_{W'\in\mathcal{W}} b_{2,W'}\mathbb{1}_{3R}\rangle \neq 0$ we must have that $\Omega^{\complement}\cap R \neq \emptyset$ and $W'\cap 3R \neq \emptyset$. As $W'\in\mathcal{W}$, the above implies

$$5\,\mathrm{diam}(W') \leq \mathrm{dist}(W', \Omega^{\complement}) \leq 3\,\mathrm{diam}(R)$$

and therefore $L(W') < j$. Thus,

$$\mathrm{III}_1 = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \langle T_j[g_1 \mathbb{1}_{\Omega^{\complement}\cap R}], \sum_{\substack{W'\in\mathcal{W}\\ L(W')<j}} b_{2,W'}\mathbb{1}_{3R}\rangle. \tag{B.7}$$

We next decompose $\mathrm{III}_2 = \mathrm{III}_{2,1} + \mathrm{III}_{2,2}$, where

$$\mathrm{III}_{2,1} := \sum_{j=N_1}^{N_2} \big\langle T_j\big[\sum_{\substack{W\in\mathcal{W}\\ L(W)<j}} \mathrm{av}_W[f_1]\mathbb{1}_W\big], \sum_{\substack{W'\in\mathcal{W}\\ L(W')\geq j}} b_{2,W'}\big\rangle, \tag{B.8a}$$

$$\mathrm{III}_{2,2} := \sum_{j=N_1}^{N_2} \sum_{\substack{W\in\mathcal{W}\\ L(W)<j}} \big\langle T_j[\mathrm{av}_W[f_1]\mathbb{1}_W], \sum_{\substack{W'\in\mathcal{W}\\ L(W')<j}} b_{2,W'}\big\rangle. \tag{B.8b}$$

The term $\mathrm{III}_{2,1}$ can be treated as in the estimation of the term III_3 in the proof of Theorem 1.1, defined in (4.19c), as cancellation does not play a role in this argument. The geometry expressed in (4.20) is crucial, i.e. we have likewise

$$\left.\begin{array}{l} \langle T_j[\mathrm{av}_W f_1 \mathbb{1}_W], b_{2,W'}\rangle \neq 0 \\ L(W) < j \leq L(W') \end{array}\right\} \quad \Longrightarrow \quad j \leq L(W') \leq L(W) + 2 \leq j+2. \tag{B.9}$$

This implies

$$
\begin{aligned}
|\mathrm{III}_{2,1}| &\le \sum_{N_1\le j\le N_2} \sum_{\substack{W'\in\mathcal{W}:\\ j\le L(W')\le j+2}} \sum_{\substack{W\in\mathcal{W}:\\ L(W')-2\le L(W)\le j\\ W\subset 3W'}} |\langle T_j[\mathrm{av}_W[f_1]\mathbb{1}_W], b_{2,W'}\rangle| \\
&\le A_\circ(p,\infty) \sum_{N_1\le j\le N_2} 2^{-jd/p} \sum_{\substack{W,W'\in\mathcal{W}:\, W\subset 3W'\\ j\le L(W')\le j+2\\ L(W')-2\le L(W)\le j}} \|\mathrm{av}_W[f_1]\mathbb{1}_W\|_{L^p_{B_1}} \|b_{2,W'}\|_{L^1_{B_2^*}} \\
&\lesssim A_\circ(p,\infty)\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1} \\
&\qquad \times \sum_{N_1\le j\le N_2} \sum_{\substack{W,W':\, W\subset 3W'\\ j\le L(W')\le j+2\\ L(W')-2\le L(W)\le j}} 2^{-jd/p}|W|^{1/p}|W'| \\
&\lesssim A_\circ(p,\infty)\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1} \sum_{W'\in\mathcal{W}} |W'| \lesssim A_\circ(p,\infty)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.
\end{aligned}
$$

The terms III_1 and $\mathrm{III}_{2,2}$ can be treated in a similar way as in the estimation of the terms IV_1, IV_3 (defined in (4.25)) in the proof of Theorem 1.1. Let $0<\varepsilon'<\min\{1/p,\varepsilon\}$ and $\ell>0$ be as in (4.24). Then we split

$$
\mathrm{III}_1 = \mathrm{III}_1^{\mathrm{lg}} + \mathrm{III}_1^{\mathrm{sm}}, \qquad \mathrm{III}_{2,2} = \mathrm{III}_{2,2}^{\mathrm{lg}} + \mathrm{III}_{2,2}^{\mathrm{sm}}
$$

where

$$
\mathrm{III}_1^{\mathrm{lg}} = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=1}^{\ell} \Big\langle T_j[g_1\mathbb{1}_{\Omega^{\mathfrak{C}}\cap R}], \sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'}\mathbb{1}_{3R}\Big\rangle
$$

$$
\mathrm{III}_1^{\mathrm{sm}} = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=\ell+1}^{\infty} \Big\langle g_1\mathbb{1}_{\Omega^{\mathfrak{C}}\cap R}, T_j^*\big[\sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'}\mathbb{1}_{3R}\big]\Big\rangle
$$

and

$$
\mathrm{III}_{2,2}^{\mathrm{lg}} = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=1}^{\ell} \Big\langle \sum_{\substack{W\subset R\\ L(W)<j}} T_j[\mathrm{av}_W[f_1]\mathbb{1}_W], \sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'}\mathbb{1}_{3R}\Big\rangle,
$$

$$
\mathrm{III}_{2,2}^{\mathrm{sm}} = \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=\ell+1}^{\infty} \Big\langle \sum_{\substack{W\subset R\\ L(W)<j}} \mathrm{av}_W[f_1]\mathbb{1}_W, T_j^*\big[\sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'}\mathbb{1}_{3R}\big]\Big\rangle.
$$

Observe that the terms $\mathrm{III}_1^{\mathrm{sm}}$, $\mathrm{III}_{2,2}^{\mathrm{sm}}$ involve very small cubes W' for which the cancellation of $b_{2,W'}$ can be most effectively used. The terms $\mathrm{III}_1^{\mathrm{lg}}$, $\mathrm{III}_{2,2}^{\mathrm{lg}}$ involve larger cubes; for these terms it is more effective to use the single scale (p,q) conditions (1.8).

We note that the terms $\mathrm{III}_1^{\mathrm{lg}}$ and $\mathrm{III}_{2,2}^{\mathrm{lg}}$ behave very similarly, and also the terms $\mathrm{III}_1^{\mathrm{sm}}$ and $\mathrm{III}_{2,2}^{\mathrm{sm}}$.

Indeed, if h_R, $\widetilde{h}_R$ denote either of the first functions on the bilinear form,

$$h_R(x) = g_1(x)\mathbb{1}_{\Omega^{\complement}\cap R}, \quad \widetilde{h}_R(x) = \sum_{\substack{W\subset R\\ L(W)<j}} \mathrm{av}_W[f_1]\mathbb{1}_W(x),$$

then it follows from the definition of $\Omega^{\complement}$ and the disjointness of the $W \in \mathcal{W}$ that h_R, $\widetilde{h}_R$ share the relevant property

$$\|h_R\|_{L^r_{B_1}}, \|\widetilde{h}_R\|_{L^r_{B_1}} \le |R|^{1/r}\langle f_1\rangle_{Q_0,p}, \quad 1 \le r \le \infty,$$

which we will use with $r = p$.

By the above considerations, the hypothesis (1.8) and (4.9a) we have

$$\begin{aligned}
\mathrm{III}_1^{\mathrm{lg}} &= \sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}\sum_{s=1}^{\ell}\langle T_j h_R, \sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s_2}} b_{2,W'}\mathbb{1}_{3R}\rangle \\
&\lesssim \sum_{s=1}^{\ell}\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j} A_\circ(p,\infty)2^{-jd/p}\|h_R\|_{L^p_{B_1}}\Big\|\sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'}\mathbb{1}_{3R}\Big\|_{L^1_{B_2^*}} \\
&\lesssim \sum_{s=1}^{\ell} A_\circ(p,\infty)\langle f_1\rangle_{Q_0,p}\sum_{j=N_1}^{N_2}\sum_{R\in\mathfrak{R}_j}\sum_{\substack{W'\subset 3R\\ L(W')=j-s}}\|b_{2,W'}\|_{L^1_{B_2^*}} \\
&\lesssim A_\circ(p,\infty)\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}\sum_{s=1}^{\ell}\sum_{j=N_1}^{N_2}\sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}}|W'| \\
&\lesssim \ell A_\circ(p,\infty)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}
\end{aligned}$$

and hence, by the definition of ℓ,

$$\mathrm{III}_1^{\mathrm{lg}} \lesssim A_\circ(p,\infty)\log(2+\frac{B}{A_\circ(p,\infty)})|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.$$

Similarly we show (after replacing h_R with $\widetilde{h}_R$ in the above calculation)

$$\mathrm{III}_{2,2}^{\mathrm{lg}} \lesssim A_\circ(p,\infty)\log(2+\frac{B}{A_\circ(p,\infty)})|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.$$

For the estimation of $\mathrm{III}_1^{\mathrm{sm}}$, $\mathrm{III}_{2,2}^{\mathrm{sm}}$ we use the ε-regularity property (1.9b) and Corollary 3.5 to get

$$\|T_j^*(\mathrm{I}-\mathbb{E}_{s-j})\|_{L^1_{B_1}\to L^{p'}_{B_2}} \lesssim_\varepsilon B2^{-jd/p}2^{-\varepsilon s}. \tag{B.10}$$

Moreover we use the formula $b_{2,W'} = (I - \mathbb{E}_{s-j}) f_{2,W'}$, valid for $L(W') = j-s$. Thus, via Hölder's inequality

$$\sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=\ell+1}^{\infty} \langle h_R, T_j^*\big[\sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} b_{2,W'} \mathbb{1}_{3R}\big]\rangle$$

$$\lesssim \sum_{j=N_1}^{N_2} \sum_{s=\ell+1} \|h_R\|_{L^p_{B_1}} \|T_j^*(I-\mathbb{E}_{s-j})\|_{L^1_{B_2^*}\to L^{p'}_{B_1}} \Big\| \sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s}} f_{2,W'} \mathbb{1}_{3R}\Big\|_{L^1_{B_2^*}}$$

$$\lesssim \sum_{j=N_1}^{N_2} \sum_{R\in\mathfrak{R}_j} \sum_{s=\ell+1}^{\infty} |R|^{1/p} \langle f_1\rangle_{Q_0,p} B 2^{-\varepsilon' s} 2^{-jd/p} \Big\| \sum_{\substack{W'\in\mathcal{W}\\ L(W')=j-s_2}} f_{2,W'} \mathbb{1}_{3R}\Big\|_{L^1_{B_2^*}}$$

$$\lesssim \sum_{j=N_1}^{N_2} \sum_{s=\ell+1}^{\infty} B 2^{-\varepsilon' s} \langle f_1\rangle_{Q_0,p} \sum_{R\in\mathfrak{R}_j} \sum_{\substack{W'\subseteq 3R\\ L(W')=j-s}} \|f_{2,W'}\|_{L^1_{B_2^*}}$$

$$\lesssim \sum_{j=N_1}^{N_2} \sum_{s=\ell+1}^{\infty} B 2^{-\varepsilon' s} \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,1} \sum_{R\in\mathfrak{R}_j} \sum_{\substack{W'\subseteq 3R\\ L(W')=j-s}} |W'|.$$

We sum in W' and then use $\sum_{s=\ell+1}^{\infty} B 2^{-\varepsilon' s} \lesssim A_\circ(p,\infty)$ to obtain

$$|\mathrm{III}_1^{\mathrm{sm}}| \lesssim A_\circ(p,\infty) |Q_0| \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,1}.$$

In exactly the same way (replacing h_R by $\widetilde{h}_R$) we obtain

$$|\mathrm{III}_{2,2}^{\mathrm{sm}}| \lesssim A_\circ(p,\infty) |Q_0| \langle f_1\rangle_{Q_0,p} \langle f_2\rangle_{3Q_0,1}.$$

This concludes the proof. □

B.3. **The case $p = 1$ and $q = \infty$.** In this case we can get rid of both the weak (p,p) and restricted strong type (q,q) hypotheses, but we shall still assume either a weak-type estimate (r,r) or restricted strong type (r,r) for some $1 < r < \infty$.

Theorem B.3. *Let $\{T_j\}_{j\in\mathbb{Z}}$ be a family of operators in Op_{B_1,B_2} such that*

- *the support condition* (1.6) *holds,*
- *there exists $r \in (1,\infty)$ so that either the weak type (r,r) condition* (1.7a) *holds or the restricted strong type (r,r) condition* (1.7b) *holds,*
- *the single scale $(1,\infty)$ condition* (1.8) *holds,*
- *the single scale ε-regularity conditions* (1.9a), (1.9b) *hold with $p = 1$ and $q = \infty$.*

Define

$$\mathcal{C} = A(r) + A_\circ(1,\infty) \log\big(2 + \tfrac{B}{A_\circ(1,\infty)}\big).$$

Then, for all integers N_1, N_2 with $N_1 \le N_2$,

$$\Big\| \sum_{j=N_1}^{N_2} T_j \Big\|_{\mathrm{Sp}_\gamma(1,B_1,1,B_2^*)} \lesssim_{r,\varepsilon,\gamma,d} \mathcal{C}.$$

Sketch of proof. We use the terminology in the proofs of Theorems B.1 and B.2. An examination of the proofs reveals that it only remains to establish the inequality

$$|\langle (S - \sum_{W\in\mathcal{W}} S_W) g_1, g_2\rangle| \lesssim A(r)|Q_0|\langle f_1\rangle_{Q_0,1}\langle f_2\rangle_{3Q_0,1}, \tag{B.11}$$

either under the restricted strong type (r,r) assumption (1.7b), or under the weak type (r,r) assumption (1.7a). Here we will strongly use (4.10) for both g_1 and g_2.

We first verify (B.11) assuming (1.7b). By Hölder's inequality,

$$\begin{aligned}|\langle Sg_1, g_2\rangle| &\lesssim \|Sg_1\|_{L^r_{B_1}}\|g_2\|_{L^{r'}_{B_2^*}} \lesssim A(r)\|g_1\|_{L^{r,1}_{B_1}}|Q_0|^{1/r'}\|g_2\|_{L^\infty_{B_2^*}}\\ &\lesssim A(r)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.\end{aligned}$$

Moreover for each $W \in \mathcal{W}$

$$\begin{aligned}|\langle S_W g_1, g_2\rangle| &= |\langle S_W[g_1\mathbb{1}_W], g_2\mathbb{1}_{3W}\rangle| \le \|S_W[g_1\mathbb{1}_W]\|_{L^r_{B_1}}\|g_2\mathbb{1}_{3W}\|_{L^{r'}_{B_2^*}}\\ &\lesssim A(r)\|g_1\mathbb{1}_W\|_{L^{r,1}_{B_1}}|W|^{1/r'}\|g_2\mathbb{1}_{3W}\|_{L^\infty_{B_2^*}} \lesssim A(r)|W|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1},\end{aligned}$$

and by summing over the disjoint cubes $W \in \mathcal{W}$ we obtain

$$\sum_{W\in\mathcal{W}} \big|\langle S_W g_1, g_2\rangle\big| \lesssim A(r)|Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.$$

Combining the two bounds yields (B.11) (under the assumption (1.7b)).

We now verify (B.11) assuming (1.7a). First, by Hölder's inequality for Lorentz spaces,

$$\begin{aligned}|\langle Sg_1, g_2\rangle| &\lesssim \|Sg_1\|_{L^{r,\infty}_{B_1}}\|g_2\|_{L^{r',1}_{B_2^*}} \lesssim A(r)\|g_1\|_{L^r_{B_1}}|Q_0|^{1/r'}\|g_2\|_{L^\infty_{B_2^*}}\\ &\lesssim |Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.\end{aligned}$$

Similarly, for all $W \in \mathcal{W}$,

$$|\langle S_W g_1, g_2\rangle| \lesssim |W|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}$$

and then after summation

$$\sum_{W\in\mathcal{W}} |\langle S_W g_1, g_2\rangle| \lesssim |Q_0|\langle f_1\rangle_{Q_0,p}\langle f_2\rangle_{3Q_0,1}.$$

This yields (B.11) (under the assumption (1.7a)). □

Appendix C. Facts about Fourier multipliers

For completeness, we provide proofs of the facts stated in the remark after the definition of the $\mathcal{B}[m]$. The proofs will be given for scalar multipliers but they carry over to the setting with $\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)$-valued multipliers. We start with the following simple observations.

Lemma C.1. *Let $\Psi \in C_c^\infty(\mathbb{R}^d)$ be supported in $\{x \in \mathbb{R}^d : 1/2 < |x| < 2\}$. Let $\Phi \in C_c^\infty(\mathbb{R}^d)$ be supported in $\{x \in \mathbb{R}^d : |x| < 2\}$. Let $N > d$ and κ be such that*

$$\sup_{x\in\mathbb{R}^d} (1+|x|)^N |\kappa(x)| \le 1.$$

Then the following holds.

(i) Let $1 \le \rho \le R/8$. Then

$$\big\| \big[\widehat{\kappa}(m * R^d\widehat{\Psi}(R\cdot))\big] * \rho^d\widehat{\Phi}(\rho\cdot)\big\|_{M^{p,q}} \lesssim R^{d-N}\big\| m * R^d\widehat{\Psi}(R\cdot)\big\|_{M^{p,q}}$$

(ii) Let $1 \le \rho \le R/8$. Then

$$\big\| \big[\widehat{\kappa}(m * \rho^d\widehat{\Phi}(\rho\cdot))\big] * R^d\widehat{\Psi}(R\cdot)\big\|_{M^{p,q}} \lesssim R^{d-N}\big\| m * \rho^d\widehat{\Phi}(\rho\cdot)\big\|_{M^{p,q}}$$

Proof. Let $K = \mathcal{F}^{-1}[m]$ and set $\|K\|_{cv(p,q)} := \|\widehat{K}\|_{M^{p,q}}$. The expression in (i) is equal to

$$\Big\| \Phi(\rho^{-1}\cdot) \int \kappa(y) K(\cdot - y)\Psi(R^{-1}(\cdot - y))\, dy\Big\|_{cv(p,q)}.$$

Observe that by the support properties of Φ, Ψ the integral in y is extended over $R/2 - 2\rho \le |y| \le 2R + 2\rho$, hence $|y| \in (R/4, 4R)$. Thus the displayed quantity is bounded by

$$\begin{aligned}&\int_{R/4\le|y|\le 4R} |\kappa(y)| \|K(\cdot - y)\Psi(R^{-1}(\cdot - y))\|_{cv(p,q)}\, dy\\ &\le \int_{R/4\le|y|\le 4R} |\kappa(y)|\, dy\, \|m * R^d\widehat{\Psi}(R\cdot)\|_{M^{p,q}}\end{aligned}$$

and the desired bound follows from the hypothesis on κ. Part (ii) is proved in the same way. □

Lemma C.2. *Let Ψ_n, $n \ge 0$, be as in §6.1. Let $N > d$ and let χ be such that $\|\partial_\xi^\alpha \chi\|_1 \le A$ for all $\alpha \in \mathbb{N}_0$ such that $|\alpha| \le N$. Let $h \in L^1(\widehat{\mathbb{R}}^d)$ be supported in $\{\xi \in \widehat{\mathbb{R}}^d : 1/2 \le |\xi| \le 2\}$. Then*

$$\|(h\chi) * \widehat{\Psi}_\ell\|_{M^{p,q}} \lesssim A \sum_{n=0}^\infty C_{N-d}(n,\ell)\|h * \widehat{\Psi}_n\|_{M^{p,q}}$$

for any $\ell \ge 0$, where

$$C_{N_1}(n,\ell) := \begin{cases} 1 & \text{if } \ell - 5 \le n \le \ell + 5,\\ 2^{-\ell N_1} & \text{if } 0 \le n < \ell - 5,\\ 2^{-nN_1} & \text{if } \ell + 5 < n.\end{cases} \tag{C.1}$$

Proof. We write $(h\chi)*\widehat{\Psi}_\ell = \sum_{n=0}^\infty [(h*\widehat{\Psi}_n)\chi]*\widehat{\Psi}_\ell$. The result then follows by noting that $|\mathcal{F}^{-1}[\chi](x)| \lesssim (1+|x|)^{-N}$ and an application of Lemma C.1. □

C.1. **Multiplication by smooth symbols.** The above observations can be applied to show that the space defined by the finiteness of $\mathcal{B}[m]$ in (6.2a) is invariant under multiplication with multipliers satisfying a standard symbol of order 0 assumption. There is of course also a corresponding similar and immediate statement for $\mathcal{B}_\circ[m]$.

Lemma C.3. *Let $a \in C^\infty(\widehat{\mathbb{R}}^d)$. Then*

$$\mathcal{B}[am] \lesssim \mathcal{B}[m] \sum_{|\alpha|\le 2d+1} \sup_{\xi\in\widehat{\mathbb{R}}^d} |\xi|^{|\alpha|}|\partial^\alpha a(\xi)|,$$

where $\alpha \in \mathbb{N}_0$. Consequently, if $|\partial^\alpha a(\xi)| \lesssim_\alpha (1+|\xi|)^{-|\alpha|}$ for all $\xi \in \widehat{\mathbb{R}}^d$ and all $\alpha \in \mathbb{N}_0^d$, we have $\mathcal{B}[am] \lesssim \mathcal{B}[m]$.

Proof. Let $\widetilde{\phi} \in C_c^\infty(\widehat{\mathbb{R}}^d)$ be supported in $\{\xi \in \widehat{\mathbb{R}}^d : 1/4 \le |\xi| \le 4\}$ and such that $\widetilde{\phi}(\xi) = 1$ for $1/2 \le |\xi| \le 2$. Let $a^t(\xi) = \widetilde{\phi}(\xi)a(t\xi)$.

$$[\varphi a(t\cdot)m(t\cdot)] * \widehat{\Psi}_\ell = \sum_{n=0}^\infty \left[\left([\varphi m(t\cdot)] * \widehat{\Psi}_n\right)a^t\right] * \widehat{\Psi}_\ell.$$

We have $\sum_{|\alpha|\le 2d+1} |\partial^\alpha a^t(\xi)| \lesssim 1$, uniformly in t. By Lemma C.2 with $N_1 = 2d+1$,

$$\begin{aligned}
&\sum_{\ell=0}^\infty (1+\ell)2^{\ell d(1/p-1/q)} \|[\phi a(t\cdot)m(t\cdot)] * \widehat{\Psi}_\ell\|_{M^{p,q}} \\
&\lesssim \sum_{\ell=0}^\infty \sum_{n=0}^\infty C_{d+1}(n,\ell)(1+\ell)2^{\ell d(1/p-1/q)} \|[\phi m(t\cdot)] * \widehat{\Psi}_n\|_{M^{p,q}} \\
&\lesssim \sum_{n=0}^\infty (1+n)2^{nd(1/p-1/q)} \|[\varphi m(t\cdot)] * \widehat{\Psi}_n\|_{M^{p,q}}
\end{aligned}$$

where in the last line we used that

$$\sup_{n\ge 0} \sum_{\ell=0}^\infty \frac{1+\ell}{1+n} 2^{(\ell-n)d(\frac{1}{p}-\frac{1}{q})} C_{d+1}(n,\ell) < \infty. \qquad \square$$

C.2. **Independence of ϕ, Ψ in the finiteness of $\mathcal{B}[m]$.** The previous argument in Lemma C.3 can also be used to show that the space defined by the finiteness of $\mathcal{B}[m]$, $\mathcal{B}_\circ[m]$ is independent of the specific choices of ϕ, Ψ in §6.1. We only give the argument for $\mathcal{B}[m]$ and a similar reasoning applies to $\mathcal{B}_\circ[m]$.

Lemma C.4. *Denote the left hand side of* (6.2a) *by* $\mathcal{B}[m,\phi,\Psi]$. *Given two choices of* (ϕ,Ψ) *and* $(\widetilde{\phi},\widetilde{\Psi})$ *with the specifications in the first paragraph of §6.1, there is a constant* $C=C(\phi,\Psi)>1$ *such that*

$$C^{-1}\mathcal{B}[m,\phi,\Psi]\le \mathcal{B}[m,\widetilde{\phi},\widetilde{\Psi}]\le C\mathcal{B}[m,\phi,\Psi].$$

Proof. We show the second inequality. Note that $\int_0^\infty |\phi(s\xi)|^2\frac{ds}{s}\ge c>0$ for $\xi\neq 0$. Let β^s be defined by

$$\widehat{\beta^s}(\xi)=\frac{\widetilde{\phi}(s^{-1}\xi)\phi(\xi)}{\int_0^\infty |\phi(\sigma s^{-1}\xi)|^2\frac{d\sigma}{\sigma}}.$$

We then have, in view of the support conditions on ϕ and $\widetilde{\phi}$,

$$\tilde{\phi}(\xi)=\int_{1/4}^4 \widehat{\beta^s}(s\xi)\phi(s\xi)\frac{ds}{s}$$

and hence

$$\begin{aligned}\big\|\widetilde{\phi}m(t\cdot)*\widehat{\widetilde{\Psi}}_\ell\big\|_{M^{p,q}}&\le\int_{1/4}^4\big\|[\widehat{\beta^s}(s\cdot)\phi(s\cdot)m(t\cdot)]*\widehat{\widetilde{\Psi}}_\ell\big\|_{M^{p,q}}\frac{ds}{s}\\&=\int_{1/4}^4 s^{d(\frac1p-\frac1q)}\big\|[\widehat{\beta^s}\phi(\cdot)m(ts^{-1}\cdot)]*s^{-d}\widehat{\widetilde{\Psi}}_\ell(s^{-1}\cdot)\big\|_{M^{p,q}}\frac{ds}{s}\\&\lesssim\sum_{n=0}^\infty\int_{1/4}^4\big\|\big(\widehat{\beta^s}\big([\phi(\cdot)m(ts^{-1}\cdot)]*\widehat{\Psi}_n\big)\big)*s^{-d}\widehat{\widetilde{\Psi}}_\ell(s^{-1}\cdot)\big\|_{M^{p,q}}\frac{ds}{s}.\end{aligned}$$

By Lemma C.2 this is dominated by

$$C_{N-d}(n,\ell)\int_{1/4}^4\big\|[\phi(\cdot)m(ts^{-1}\cdot)]*\widehat{\Psi}_n\big\|_{M^{p,q}}\frac{ds}{s},$$

where $C_{N-d}(n,\ell)$ is as in (C.1). It is now easy to see that for $N\ge 2d+1$

$$\begin{aligned}&\sum_{\ell\ge0}2^{\ell d(\frac1p-\frac1q)}(1+\ell)\big\|\widetilde{\phi}m(t\cdot)*\widehat{\widetilde{\Psi}}_\ell\big\|_{M^{p,q}}\\&\lesssim\int_{1/4}^4\sum_{\ell\ge0}\sum_{n\ge0}C_{N-d}(n,\ell)2^{\ell d(\frac1p-\frac1q)}(1+\ell)\big\|[\phi(\cdot)m(ts^{-1}\cdot)]*\widehat{\Psi}_n\big\|_{M^{p,q}}\frac{ds}{s}\\&\lesssim\sup_\tau\sum_{n=0}^\infty 2^{nd(\frac1p-\frac1q)}\big\|[\phi(\cdot)m(\tau\cdot)]*\widehat{\Psi}_n\big\|_{M^{p,q}}.\end{aligned}$$

This establishes the inequality $\mathcal{B}[m,\widetilde{\phi},\widetilde{\Psi}]\le C\mathcal{B}[m,\varphi,\Psi]$ and the converse follows by interchanging the roles of (ϕ,Ψ) and $(\widetilde{\phi},\widetilde{\Psi})$. □

References

[1] AimPL, *Sparse domination of singular integral operators*, American Institute of Mathematics Problem List, edited by Dario Mena, available at `http://aimpl.org/sparsedomop`.

[2] Theresa C. Anderson, Bingyang Hu, and Joris Roos, *Sparse bounds for discrete singular Radon transforms*, Colloq. Math. **165** (2021), no. 2, 199–217.

[3] Theresa C. Anderson, Kevin Hughes, Joris Roos, and Andreas Seeger, *$L^p \to L^q$ bounds for spherical maximal operators*, Math. Z. **297** (2021), no. 3-4, 1057–1074.

[4] Albert Baernstein, II and Eric T. Sawyer, *Embedding and multiplier theorems for $H^P(\mathbb{R}^n)$*, Mem. Amer. Math. Soc. **53** (1985), no. 318, iv+82. MR 776176

[5] Sayan Bagchi, Sourav Hait, Luz Roncal, and Sundaram Thangavelu, *On the maximal function associated to the spherical means on the Heisenberg group*, New York J. Math. **27** (2021), 631–675.

[6] Julian Bailey, Gianmarco Brocchi, and Maria Carmen Reguera, *Quadratic sparse domination and weighted estimates for non-integral square functions*, 2020, arXiv:2007.15928.

[7] Alex Barron, José M. Conde-Alonso, Yumeng Ou, and Guillermo Rey, *Sparse domination and the strong maximal function*, Adv. Math. **345** (2019), 1–26. MR 3897437

[8] David Beltran, *Geometric control of oscillatory integrals*, Ph.D. thesis, University of Birmingham, 2017.

[9] ———, *A Fefferman-Stein inequality for the Carleson operator.* Rev. Mat. Iberoam. **34** (2018), no. 1, 221–244. MR 3763345

[10] David Beltran and Jonathan Bennett, *Subdyadic square functions and applications to weighted harmonic analysis*, Adv. Math. **307** (2017), 72–99. MR 3590514

[11] David Beltran and Laura Cladek, *Sparse bounds for pseudodifferential operators*, J. Anal. Math. **140** (2020), no. 1, 89–116. MR 4094458

[12] David Beltran, Shaoming Guo, Jonathan Hickman and Andreas Seeger, *Sharp L^p bounds for the helical maximal function.* Preprint February 2021, arXiv 2102.08272. To appear in Amer. J. Math.

[13] ———, *Sobolev improving for averages over curves in $\mathbb{R}^4$.* Adv. of Math., **393** (2021), Paper No. 108089, 85. MR 4340226

[14] David Beltran, Richard Oberlin, Luz Roncal, Andreas Seeger, and Betsy Stovall, *Variation bounds for spherical averages*, Math. Ann. **382** (2022), 459–512. MR 4377310

[15] 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. MR 3897969

[16] Cristina Benea, Frédéric Bernicot, and Teresa Luque, *Sparse bilinear forms for Bochner Riesz multipliers and applications*, Trans. London Math. Soc. **4** (2017), no. 1, 110–128. MR 3653057

[17] Cristina Benea and Camil Muscalu, *Sparse domination via the helicoidal method*, Rev. Mat. Iberoam. **37** (2021), 2027–2118. MR 4310287

[18] Frédéric Bernicot, Dorothee Frey, and Stefanie Petermichl, *Sharp weighted norm estimates beyond Calderón-Zygmund theory*, Anal. PDE **9** (2016), no. 5, 1079–1113. MR 3531367

[19] Jean Bourgain, *Averages in the plane over convex curves and maximal operators*, J. Analyse Math. **47** (1986), 69–85. MR 874045

[20] Jean Bourgain and Ciprian Demeter, *The proof of the l^2 decoupling conjecture*, Ann. of Math. (2) **182** (2015), no. 1, 351–389. MR 3374964

[21] Philip Brenner, *On $L_p - L_{p'}$ estimates for the wave-equation*, Math. Z. **145** (1975), no. 3, 251–254. MR 387819

[22] Anthony Carbery, *The boundedness of the maximal Bochner-Riesz operator on $L^4(\mathbb{R}^2)$*, Duke Math. J. **50** (1983), no. 2, 409–416. MR 705033

[23] ———, *Radial Fourier multipliers and associated maximal functions*, Recent progress in Fourier analysis (El Escorial, 1983), North-Holland Math. Stud., vol. 111, North-Holland, Amsterdam, 1985, pp. 49–56. MR 848141

[24] ———, *Variants of the Calderón-Zygmund theory for L^p-spaces*, Rev. Mat. Iberoamericana **2** (1986), no. 4, 381–396. MR 913694

[25] Anthony Carbery, George Gasper, and Walter Trebels, *Radial Fourier multipliers of $L^p(\mathbb{R}^2)$*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 10, , Phys. Sci., 3254–3255. MR 747595

[26] María J. Carro and Carlos Domingo-Salazar, *Stein's square function G_α and sparse operators*, J. Geom. Anal. **27** (2017), no. 2, 1624–1635. MR 3625166

[27] Michael Christ, *On almost everywhere convergence of Bochner-Riesz means in higher dimensions*, Proc. Amer. Math. Soc. **95** (1985), no. 1, 16–20. MR 796439

[28] Laura Cladek and Yumeng Ou, *Sparse domination of Hilbert transforms along curves*, Math. Res. Lett. **25** (2018), no. 2, 415–436. MR 3826828

[29] 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. MR 3668591

[30] José M. Conde-Alonso, Francesco Di Plinio, Ioannis Parissis and Manasa M. Vempati, *A metric approach to sparse domination.* Preprint, arXiv 2009.00336

[31] 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. MR 3521084

[32] Amalia Culiuc, Francesco Di Plinio, Yumeng Ou, *A sparse estimate for multisublinear forms involving vector-valued maximal functions.* Bruno Pini Mathematical Analysis Seminar 2017, 168–184, Bruno Pini Math. Anal. Semin., **8**, Univ. Bologna, Alma Mater Stud., Bologna, 2017. MR3893587

[33] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *Domination of multilinear singular integrals by positive sparse forms*, J. Lond. Math. Soc. (2) **98** (2018), no. 2, 369–392. MR 3873113

[34] Amalia Culiuc, Robert Kesler, and Michael T. Lacey, *Sparse bounds for the discrete cubic Hilbert transform*, Anal. PDE **12** (2019), no. 5, 1259–1272. MR 3892403

[35] Fernanda Clara de França Silva and Pavel Zorin-Kranich, *Sparse domination of sharp variational truncations*, 2016, arXiv:1604.05506.

[36] Francesco Di Plinio, Yen Q. Do, and Gennady N. Uraltsev, *Positive sparse domination of variational Carleson operators*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **18** (2018), no. 4, 1443–1458. MR 3829751

[37] Francesco Di Plinio, Tuomas 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.

[38] Francesco Di Plinio and Andrei K. Lerner, *On weighted norm inequalities for the Carleson and Walsh-Carleson operator*, J. Lond. Math. Soc. (2) **90** (2014), no. 3, 654–674. MR 3291794

[39] Javier Duoandikoetxea and José L. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates*, Invent. Math. **84** (1986), no. 3, 541–561. MR 837527

[40] Javier Duoandikoetxea and Ana Vargas, *Maximal operators associated to Fourier multipliers with an arbitrary set of parameters*, Proc. Roy. Soc. Edinburgh Sect. A **128** (1998), no. 4, 683–696. MR 1635404

[41] Charles Fefferman, *L^p bounds for pseudo-differential operators*, Israel J. Math. **14** (1973), 413–417. MR 336453

[42] Charles Fefferman and Elias M. Stein, *Some maximal inequalities*, Amer. J. Math. **93** (1971), 107–115. MR 0284802 (44 #2026)

[43] ———, *H^p spaces of several variables*, Acta Math. **129** (1972), no. 3-4, 137–193. MR 447953

[44] Dorothee Frey, Bas Nieraeth, *Weak and strong type $A_1 - A_\infty$ estimates for sparsely dominated operators*, J. Geom. Anal. **29** (2019), no. 1, 247–282.

[45] Pritam Ganguly and Sundaram Thangavelu, *On the lacunary spherical maximal function Heisenberg group*, arXiv:1912.11302. J. Funct. Anal. **280** (2021), no. 3, Paper No. 108832, 32 pp.

[46] Gustavo Garrigós and Andreas Seeger, *Characterizations of Hankel multipliers*, Math. Ann. **342** (2008), no. 1, 31–68. MR 2415314

[47] Gustavo Garrigós, Andreas Seeger, and Tino Ullrich, *The Haar system as a Schauder basis in spaces of Hardy-Sobolev type*, J. Fourier Anal. Appl. **24** (2018), no. 5, 1319–1339. MR 3856231

[48] Shaoming Guo, Jonathan Hickman, Victor Lie, and Joris Roos, *Maximal operators and Hilbert transforms along variable non-flat homogeneous curves*, Proc. Lond. Math. Soc. (3) **115** (2017), no. 1, 177–219. MR 3669936

[49] Shaoming Guo, Joris Roos, Andreas Seeger, and Po-Lam Yung, *A maximal function for families of Hilbert transforms along homogeneous curves*, Math. Ann. **377** (2020), no. 1-2, 69–114. MR 4099627

[50] ———, *Maximal functions associated with families of homogeneous curves: L^p bounds for $p \leq 2$*, Proc. Edinb. Math. Soc. (2) **63** (2020), no. 2, 398–412. MR 4089381

[51] Shaoming Guo, Joris Roos, and Po-Lam Yung, *Sharp variation-norm estimates for oscillatory integrals related to Carleson's theorem*, Analysis & PDE **13** (2020), no. 5, 1457–1500.

[52] Timo S. Hänninen and Emiel Lorist, *Sparse domination for the lattice Hardy-Littlewood maximal operator*, Proc. Amer. Math. Soc. **147** (2019), no. 1, 271–284. MR 3876748

[53] Lars Hörmander, *Estimates for translation invariant operators in L^p spaces*, Acta Math. **104** (1960), 93–140. MR 121655

[54] Bingyang Hu, *Sparse domination of singular Radon transform*, J. Math. Pures Appl. (9) **139** (2020), 235–316.

[55] Richard A. Hunt, *On $L(p, q)$ spaces*, Enseign. Math. (2) **12** (1966), 249–276. MR 223874

[56] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 63, Springer, Cham, 2016. MR 3617205

[57] Tuomas P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. MR 2912709

[58] Roger L. Jones, Robert Kaufman, Joseph M. Rosenblatt, and Máté Wierdl, *Oscillation in ergodic theory*, Ergodic Theory Dynam. Systems **18** (1998), no. 4, 889–935. MR 1645330

[59] Roger L. Jones, Andreas Seeger, and James Wright, *Strong variational and jump inequalities in harmonic analysis*, Trans. Amer. Math. Soc. **360** (2008), no. 12, 6711–6742. MR 2434308

[60] Makoto Kaneko and Gen-ichirô Sunouchi, *On the Littlewood-Paley and Marcinkiewicz functions in higher dimensions*, Tohoku Math. J. (2) **37** (1985), no. 3, 343–365. MR 799527

[61] Robert Kesler and Michael T. Lacey, *Sparse endpoint estimates for Bochner-Riesz multipliers on the plane*, Collect. Math. **69** (2018), no. 3, 427–435. MR 3842215

[62] Hyerim Ko, Sanghyuk Lee and Sewook Oh, *Maximal estimate for average over space curve*. Preprint February 2021, arXiv 2102.07175.

[63] ———, *Sharp smoothing properties of averages over curves*, arXiv 2105.01628.

[64] Ben Krause and Michael T. Lacey, *Sparse bounds for random discrete Carleson theorems*, 50 years with Hardy spaces, Oper. Theory Adv. Appl., vol. 261, Birkhäuser/Springer, Cham, 2018, pp. 317–332. MR 3792103

[65] Michael T. Lacey, *An elementary proof of the A_2 bound*, Israel J. Math. **217** (2017), no. 1, 181–195. MR 3625108

[66] ———, *Sparse bounds for spherical maximal functions*, J. Anal. Math. **139** (2019), no. 2, 613–635. MR 4041115

[67] Michael T. Lacey, Dario Mena Arias. *The sparse T1 theorem.* Houston J. Math. **43** (2017), no. 1, 111–127. MR3647935.

[68] 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. MR 3917956

[69] Sanghyuk Lee, *Square function estimates for the Bochner-Riesz means*, Anal. PDE **11** (2018), no. 6, 1535–1586. MR 3803718

[70] Sanghyuk Lee, Keith M. Rogers, and Andreas Seeger, *Improved bounds for Stein's square functions*, Proc. Lond. Math. Soc. (3) **104** (2012), no. 6, 1198–1234. MR 2946085

[71] ———, *Square functions and maximal operators associated with radial Fourier multipliers*, Chapter 12 in Advances in analysis: the legacy of Elias M. Stein, Princeton Math. Ser., vol. 50, Princeton Univ. Press, Princeton, NJ, 2014, pp. 273–302. MR 3329855

[72] Sanghyuk Lee and Andreas Seeger, *On radial Fourier multipliers and almost everywhere convergence*, J. Lond. Math. Soc. (2) **91** (2015), no. 1, 105–126. MR 3338611

[73] Dominique Lépingle, *La variation d'ordre p des semi-martingales*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 4, 295–316. MR 420837

[74] Andrei K. Lerner, *On pointwise estimates involving sparse operators*, 10th International Conference on Harmonic Analysis and Partial Differential Equations, El Escorial, Madrid (Spain) June 12-17, 2016. `matematicas.uam.es/~AFA/Escorial/2016/SlidesTalks/Lerner-Escorial-2016.pdf`.

[75] ———, *On an estimate of Calderón-Zygmund operators by dyadic positive operators*, J. Anal. Math. **121** (2013), 141–161. MR 3127380

[76] ———, *A simple proof of the A_2 conjecture*, Int. Math. Res. Not. IMRN (2013), no. 14, 3159–3170. MR 3085756

[77] ———, *On pointwise estimates involving sparse operators*, New York J. Math. **22** (2016), 341–349. MR 3484688

[78] ———, *A weak type estimate for rough singular integrals*, Rev. Mat. Iberoam. **35** (2019), no. 5, 1583–1602. MR 4018107

[79] Andrei K. Lerner and Fedor Nazarov, *Intuitive dyadic calculus: the basics*, Expo. Math. **37** (2019), no. 3, 225–265. MR 4007575

[80] Andrei K. Lerner; Sheldy Ombrosi. *Some remarks on the pointwise sparse domination.* J. Geom. Anal. 30 (2020), no. 1, 1011–1027. MR 4058547

[81] Jacques-Louis Lions and Jaak Peetre, *Sur une classe d'espaces d'interpolation*, Inst. Hautes Études Sci. Publ. Math. (1964), no. 19, 5–68. MR 165343

[82] Walter Littman, *Multipliers in L^p and interpolation*, Bull. Amer. Math. Soc. **71** (1965), 764–766. MR 179544

[83] ———, *$L^p - L^q$-estimates for singular integral operators arising from hyperbolic equations*, Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), 1973, pp. 479–481. MR 0358443

[84] Walter Littman, Charles McCarthy, and Nestor Rivière, *L^p-multiplier theorems*, Studia Math. **30** (1968), 193–217. MR 231126

[85] Emiel Lorist, *On pointwise ℓ^2 sparse domination in a space of homogeneous type*, J. Geom. Anal. **31** (2021), 9366-9405. MR 4302224

[86] Gianfranco Marletta and Fulvio Ricci, *Two-parameter maximal functions associated with homogeneous surfaces in $\mathbb{R}^n$*, Studia Math. **130** (1998), no. 1, 53–65. MR 1623004

[87] Akihiko Miyachi, *On some estimates for the wave equation in L^p and H^p*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **27** (1980), no. 2, 331–354. MR 586454

[88] ———, *On some Fourier multipliers for $H^p(\mathbb{R}^n)$*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **27** (1980), no. 1, 157–179. MR 573335

[89] Richard Oberlin, *Sparse bounds for a prototypical singular Radon transform*, Canad. Math. Bull. **62** (2019), no. 2, 405–415. MR 3952528

[90] Juan C. Peral, *L^p estimates for the wave equation*, J. Functional Analysis **36** (1980), no. 1, 114–145. MR 568979

[91] Maria Cristina Pereyra, *Dyadic harmonic analysis and weighted inequalities: the sparse revolution*, New Trends in Applied Harmonic Analysis, Volume 2, Applied and Numerical Harmonic Analysis, vol. 50, Birkhäuser, 2019, pp. 273–302.

[92] Malabika Pramanik and Andreas Seeger, *L^p regularity of averages over curves and bounds for associated maximal operators*, Amer. J. Math. **129** (2007), no. 1, 61–103. MR 2288738

[93] Joris Roos and Andreas Seeger, *Spherical maximal functions and fractal dimensions of dilation sets*, Amer. J. Math. **145** (2023), 1077-1110. MR 4621382

[94] Joris Roos, Andreas Seeger, and Rajula Srivastava, *Lebesgue space estimates for spherical maximal functions on Heisenberg groups*, Int. Math. Res. Notices IMRN **24** (2022), 19222–19257. MR 4523247 arXiv 2103.09734 (2021). To appear in IMRN.

[95] Wilhelm Schlag, *A generalization of Bourgain's circular maximal theorem*, J. Amer. Math. Soc. **10** (1997), no. 1, 103–122. MR 1388870

[96] Wilhelm Schlag and Christopher D. Sogge, *Local smoothing estimates related to the circular maximal theorem*, Math. Res. Lett. **4** (1997), no. 1, 1–15. MR 1432805

[97] Andreas Seeger, *On quasiradial Fourier multipliers and their maximal functions*, J. Reine Angew. Math. **370** (1986), 61–73. MR 852510

[98] ———, *Some inequalities for singular convolution operators in L^p-spaces*, Trans. Amer. Math. Soc. **308** (1988), no. 1, 259–272. MR 955772

[99] ———, *Remarks on singular convolution operators*, Studia Math. **97** (1990), no. 2, 91–114. MR 1083340

[100] Andreas Seeger and Terence Tao, *Sharp Lorentz space estimates for rough operators*, Math. Ann. **320** (2001), no. 2, 381–415. MR 1839769

[101] Andreas Seeger, Stephen Wainger, and James Wright, *Pointwise convergence of spherical means*, Math. Proc. Cambridge Philos. Soc. **118** (1995), no. 1, 115–124. MR 1329463

[102] Elias M. Stein, *Localization and summability of multiple Fourier series*, Acta Math. **100** (1958), 93–147. MR 105592

[103] ———, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 0290095 (44 #7280)

[104] ———, *L^p boundedness of certain convolution operators*, Bull. Amer. Math. Soc. **77** (1971), 404–405. MR 276757

[105] ———, *Maximal functions. I. Spherical means*, Proc. Nat. Acad. Sci. U.S.A. **73** (1976), no. 7, 2174–2175. MR 420116

[106] Elias M. Stein and Guido Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, N.J., 1971, Princeton Mathematical Series, No. 32. MR 0304972

[107] Elias M. Stein and Antoni Zygmund, *Boundedness of translation invariant operators on Hölder spaces and L^p-spaces*, Ann. of Math. (2) **85** (1967), 337–349. MR 215121

[108] Robert S. Strichartz, *Convolutions with kernels having singularities on a sphere*, Trans. Amer. Math. Soc. **148** (1970), 461–471. MR 256219

[109] Gen-ichirô Sunouchi, *On the Littlewood-Paley function g^* of multiple Fourier integrals and Hankel multiplier transformations*, Tohoku Math. J. (2) **19** (1967), 496–511. MR 225097

[110] Hans Triebel, *Theory of function spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983. MR 781540

[111] Thomas Wolff, *Local smoothing type estimates on L^p for large p*, Geom. Funct. Anal. **10** (2000), no. 5, 1237–1288. MR 1800068

David Beltran: Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Dr, Madison, WI-53706, USA

Current address: Department of d'Análisi Mathemática, Universitat de Valencia, Dr. Moliner 50, 46100 Burjassot, Spain

Email address: david.beltran@uv.es

Joris Roos: Department of Mathematics and Statistics, University of Massachusetts Lowell, 1 University Ave. Lowell, MA 01854, USA, and School of Mathematics, The University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Rd, Edinburgh EH9 3FD, UK

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