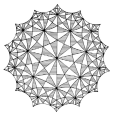
Rev. Mat. Iberoam. 39 (2023), no. 6, 2079–2118
DOI 10.4171/RMI/1400



Paraproducts, Bloom BMO and sparse BMO functions

Irina Holmes Fay and Valentia Fragkiadaki

Abstract. We address $L^p(\mu) \to L^p(\lambda)$ bounds for paraproducts in the Bloom setting. We introduce certain "sparse BMO" functions associated with sparse collections with no infinitely increasing chains, and use these to express sparse operators as sums of paraproducts and martingale transforms – essentially, as Haar multipliers – as well as to obtain an equivalence of norms between sparse operators $\mathcal{A}_S$ and compositions of paraproducts $\Pi_a^* \Pi_b$.

1. Introduction

In 1985, Steven Bloom proved [3] that the commutator $[b, H]f = b \cdot Hf - H(b \cdot f)$, where H is the Hilbert transform, is bounded $L^p(\mu) \to L^p(\lambda)$, where μ and λ are two A_p weights ($1 < p < \infty$), if and only if b is in a weighted BMO space determined by the two weights μ and λ, namely $b \in \mathrm{BMO}(\nu)$, where $\nu := \mu^{1/p}\lambda^{-1/p}$, and

$$\|b\|_{\mathrm{BMO}(\nu)} := \sup_Q \frac{1}{\nu(Q)} \int_Q |b(x) - \langle b \rangle_Q| \, dx.$$

In [8], this result was extended to commutators $[b, T]$ in $\mathbb{R}^n$ with Calderón–Zygmund operators T. Soon after, [12] gave a different proof which yielded a quantitative result for the upper bound:

$$\|[b, T] : L^p(\mu) \to L^p(\lambda)\| \lesssim \|b\|_{\mathrm{BMO}(\nu)} \big([\mu]_{A_p} [\lambda]_{A_p}\big)^{\max(1, 1/(p-1))}. \tag{1.1}$$

The proof in [8] took the route of Hytönen's representation theorem (the $\mathbb{R}^n$, Calderón–Zygmund operator generalization of Petermichl's result [15] on the Hilbert transform), and relied heavily on paraproduct decompositions. The proof in [12] used sparse operators and Lerner's median inequalities to obtain directly a sparse domination result for the commutator $[b, T]$ itself, avoiding paraproducts altogether.

This paper addresses $L^p(\mu) \to L^p(\lambda)$ bounds for the paraproducts. Based on the one-weight situation, we suspect that these bounds should be smaller than the ones for commutators: in the one-weight case,

$$\|[b, H] : L^p(w) \to L^p(w)\| \lesssim \|b\|_{\mathrm{BMO}} \, [w]_{A_p}^{2 \max(1, 1/(p-1))}$$

2020 Mathematics Subject Classification: Primary 42B20; Secondary 42B35, 47A30.
Keywords: BMO, weighted inequalities, paraproducts, sparse operators.

and

$$\|\Pi_b : L^p(w) \to L^p(w)\| \lesssim \|b\|_{\mathrm{BMO}} [w]_{A_p}^{\max(1,1/(p-1))},$$

are both known to be sharp – see [4,14] and the references therein – (throughout this paper, $A \lesssim B$ is used to mean $A \leq C(n)B$, with a constant depending on the dimension and maybe other quantities such as p or Carleson constants Λ of sparse collections, but in any case not depending on any A_p characteristics of the weights involved). In the two-weight Bloom situation, we show in Theorem 4.3 that

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| \lesssim \|b\|_{\mathrm{BMO}(\nu)} [\mu]_{A_p}^{1/(p-1)} [\lambda]_{A_p}.$$

We do not know if this bound is sharp, and this is subject to future investigations, but the bound is smaller than the one in (1.1). In fact, it is strictly smaller with the exception of $p = 2$, when both bounds are $[\mu]_{A_2}[\lambda]_{A_2}$. We can however show that our bound is sharp in one particular instance, namely when $\mu = w$ and $\lambda = w^{-1}$ for some A_2 weight w. We show this in Section 4.1 via an appeal to the one-weight linear A_2 bound for the dyadic square function.

Obviously, this bound does not recover the one-weight situation: letting $\mu = \lambda = w$ for some $w \in A_2$, and $\nu = 1$, our bound would give

$$\|\Pi_b : L^2(w) \to L^2(w)\| \lesssim \|b\|_{\mathrm{BMO}} [w]^2,$$

when we know that the optimal bound is linear in the A_2 characteristic – this was shown in [1] using Bellman function techniques. If the optimal Bloom paraproduct bound is to recover this one-weight situation, we suspect it would need a dependency on $[\nu]_{A_2}$ – as it would need to somehow account for the case $\mu = \lambda$, or $\nu = 1$.

The proof of the Bloom paraproduct bound above relies on dominating the paraproduct by a "Bloom sparse operator" $\mathcal{A}_{\mathcal{S}}^{\nu} f := \sum_{Q \in \mathcal{S}} \langle \nu \rangle_Q \langle f \rangle_Q \mathbb{1}_Q$, where $\mathcal{S}$ is a sparse collection, and proving that $\mathcal{A}_{\mathcal{S}}^{\nu}$ satisfies the bound $[\mu]_{A_p}^{1/(p-1)}[\lambda]_{A_p}$ above. We do this in Theorem 3.9. The domination of the paraproduct is treated in Section 4.

Before all this, however, we consider in Section 3 a special type of sparse collections, $\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$, which are sparse collections with no "infinitely increasing chains" (a terminology borrowed from [9]). We see that any such collection can be associated with a BMO function

$$b_{\mathcal{S}} := \sum_{Q \in \mathcal{S}} \mathbb{1}_Q,$$

which satisfies $\|b_{\mathcal{S}}\|_{\mathrm{BMO}} \leq \Lambda$, where Λ is the Carleson constant of $\mathcal{S}$ (we show this in Appendix A). Once we have a BMO function, we can immediately talk about paraproducts with symbol $b_{\mathcal{S}}$. In fact, we see in Section 3.3 that these functions allow us to express any sparse operator $\mathcal{A}_{\mathcal{S}}$, $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$, as a sum of paraproducts and a martingale transform:

$$\mathcal{A}_{\mathcal{S}} f = \Pi_{b_{\mathcal{S}}} f + \Pi_{b_{\mathcal{S}}}^* f + T_{\tau_{\mathcal{S}}} f,$$

where $T_{\tau_{\mathcal{S}}}$ is a martingale transform:

$$T_{\tau_{\mathcal{S}}} = \sum_{J \in \mathcal{D}} (\tau_{\mathcal{S}})_J \langle f, h_J \rangle h_J, \quad \text{where} \quad (\tau_{\mathcal{S}})_J := \frac{1}{|J|} \sum_{I \in \mathcal{S}, I \subsetneq J} |I| \leq \Lambda, \quad \forall J \in \mathcal{D}.$$

As discussed in Section 3.3, this gives us an upper bound for norms of sparse operators in terms of norms of paraproducts and martingale transforms, and in fact, the equivalence

$$\sup_{\mathcal{S}\in\Upsilon^{\mathcal{D}}} \|\mathcal{A}_{\mathcal{S}}\|_{L^p(w)\to L^p(w)} \simeq_{n,p,\Lambda} \sup_{b\in \mathrm{BMO}^{\mathcal{D}}} \frac{\|\Pi_b\|_{L^p(w)\to L^p(w)}}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}} + \sup_{b\in \mathrm{BMO}^{\mathcal{D}}} \frac{\|\Pi_b^*\|_{L^p(w)\to L^p(w)}}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}} + \sup_{\tau\in\ell^\infty} \frac{\|T_\tau\|_{L^p(w)\to L^p(w)}}{\|\tau\|_\infty}.$$

The process used to obtain the BMO function $b_{\mathcal{S}}$ associated with $\mathcal{S}$ also works with weights, and obtaining a function in weighted BMO spaces associated with $\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$: if $w\in A_p$, the function

$$b_{\mathcal{S}}^w := \sum_{Q\in\mathcal{S}} \langle w\rangle_Q \mathbb{1}_Q$$

is in $\mathrm{BMO}^{\mathcal{D}}(w)$, with

$$\|b_{\mathcal{S}}^w\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \le 2[w]_{A_p}\Lambda^p.$$

Repeating the process above, we try to express $\mathcal{A}_{\mathcal{S}}$ as a sum of the paraproducts associated with $b_{\mathcal{S}}^w$ and a martingale transform, but we discover instead the operator

$$\mathcal{A}_{\mathcal{S}}^w f := \sum_{Q\in\mathcal{S}} \langle w\rangle_Q \langle f\rangle_Q \mathbb{1}_Q,$$

and its decomposition as

$$\mathcal{A}_{\mathcal{S}}^w f = \Pi_{b_{\mathcal{S}}^w} f + \Pi^*_{b_{\mathcal{S}}^w} f + T_{\tau_{\mathcal{S}}^w} f,$$

detailed in Proposition 3.4.

While it would be interesting if the paraproducts and the martingale transform could somehow be "separated" above, giving an independent proof that these operators have the same dependency on $[w]_{A_p}$ by showing each is equivalent to norms of $\mathcal{A}_{\mathcal{S}}$, we are able to show that norms of sparse operators are equivalent to certain compositions of paraproducts. In Section 3.4, we see that

$$\mathcal{A}_{\mathcal{S}} \simeq \Pi^*_{\widetilde{b}_{\mathcal{S}}} \Pi_{\widetilde{b}_{\mathcal{S}}},$$

where $\widetilde{b}_{\mathcal{S}}$ is another BMO function we can easily associate with $\mathcal{S}$ (omitting signatures, see Remark 2.1):

$$\widetilde{b}_{\mathcal{S}} := \sum_{Q\in\mathcal{S}} \sqrt{|Q|}\, h_Q.$$

This provides the following upper bound:

$$\sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)\\ \Lambda_{(\mathcal{S})}=\Lambda}} \frac{\|\mathcal{A}_{\mathcal{S}} : L^p(w)\to L^p(w)\|}{\Lambda} \le \sup_{a,b\in\mathrm{BMO}^{\mathcal{D}}} \frac{\|\Pi_a^*\Pi_b : L^p(w)\to L^p(w)\|}{\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}}}.$$

For the other direction, we show in Appendix B – using a bilinear form argument – that for all Bloom weights μ, λ, ν, BMO functions $a \in \mathrm{BMO}^{\mathcal{D}}$, $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$, and $\Lambda > 1$,

$$\begin{aligned}&\|\Pi_a^* \Pi_b : L^p(\mu) \to L^p(\lambda)\| \\ &\quad \le C(n)\, \|a\|_{\mathrm{BMO}^{\mathcal{D}}}\, \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} \sup_{\substack{\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n) \\ \Lambda_{(\mathcal{S})} = \Lambda}} \Big(\frac{\Lambda}{\Lambda - 1}\Big)^3 \|\mathcal{A}_{\mathcal{S}}^{\nu} : L^p(\mu) \to L^p(\lambda)\|.\end{aligned}$$

Note that taking $\mu = \lambda = w$ above, for some $w \in A_p$, we have the one-weight result

$$\|\Pi_a^* \Pi_b : L^p(w) \to L^p(w)\| \lesssim \|a\|_{\mathrm{BMO}^{\mathcal{D}}}\, \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\, [w]_{A_p}^{\max(1, 1/(p-1))}.$$

Moreover, we obtain the equivalence of norms

$$\sup_{\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)} \|\mathcal{A}_{\mathcal{S}} : L^p(w) \to L^p(w)\| \simeq_{\Lambda, p, n} \sup_{a, b \in \mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)} \frac{\|\Pi_a^* \Pi_b : L^p(w) \to L^p(w)\|}{\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\, \|b\|_{\mathrm{BMO}^{\mathcal{D}}}}.$$

Section 4 gives a proof of a pointwise domination of paraproducts by sparse operators. It relies on first proving certain local pointwise domination results, which are then applied to $\mathrm{BMO}_{\mathcal{D}}(w)$ functions with finite Haar expansion, and extending to the general case. So this argument works whenever Π_b acts between L^p spaces where the Haar system is an unconditional basis – Lebesgue measure or A_p weights. The argument also works with the weighted BMO norm,

$$\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} := \sup_{Q \in \mathcal{D}} \frac{1}{w(Q)} \int_Q |b - \langle b \rangle_Q|\, dx,$$

defined in terms of an $L^1(dx)$ quantity – the Haar system is not unconditional in $L^1(dx)$, but we can choose an ordering of the Haar system that ensures convergence in $L^1(dx)$. The choice to work with b rather than with compactly supported f is motivated by the desire to obtain domination by sparse operators with no infinitely increasing chains. Specifically, we work with restricted paraproducts:

$$\Pi_{b, Q_0} f(x) := \sum_{Q \in Q_0} (b, h_Q)\, \langle f \rangle_Q\, h_Q(x), \quad \forall\, Q_0 \in \mathcal{D},$$

and construct a sparse collection $\mathcal{S}(Q_0) \subset \mathcal{D}(Q_0)$ which "ends" at Q_0, and such that $\mathcal{A}_{\mathcal{S}}^{w} f$ pointwise dominates $\Pi_{b, Q_0} f$ on Q_0. Since the Haar expansion of b effectively dictates the Haar expansion of Π_b (as well as Π_b^* and Γ_b), this will lead from finite Haar expansion b's to collections in $\Upsilon_{\mathcal{D}}(\mathbb{R}^n)$.

2. Setup and notations

2.1. Dyadic grids

By a dyadic grid $\mathcal{D}$ on $\mathbb{R}^n$ we mean a collection of cubes $Q \subset \mathbb{R}^n$ that satisfies

- every $Q \in \mathcal{D}$ has side length 2^k for some $k \in \mathbb{Z}$: $\ell(Q) = 2^k$;
- for a fixed $k_0 \in \mathbb{Z}$, the collection $\{Q \in \mathcal{D} : \ell(Q) = 2^{k_0}\}$ forms a partition on $\mathbb{R}^n$;

- for every $P, Q \in \mathcal{D}$, the intersection $P \cap Q$ is one of $\{P, Q, \emptyset\}$. In other words, two dyadic cubes intersect each other if and only if one contains the other.

  For example, the standard dyadic grid on $\mathbb{R}^n$ is:

$$\mathcal{D}_0 := \{2^{-k}([0,1)^n + m) : k \in \mathbb{Z}, m \in \mathbb{Z}^n\}.$$

We assume such a collection $\mathcal{D}$ is fixed throughout the paper. For every $Q \in \mathcal{D}$ and positive integer $k \geq 1$, we let $Q^{(k)}$ denote the kth dyadic ancestor of Q in $\mathcal{D}$, i.e., the unique $R \in \mathcal{D}$ such that $R \supset Q$ and $\ell(R) = 2^k \ell(Q)$. Given $Q_0 \in \mathcal{D}$, we let $\mathcal{D}(Q_0)$ denote the collection of dyadic subcubes of Q_0:

$$\mathcal{D}(Q_0) := \{Q \in \mathcal{D} : Q \subset Q_0\}.$$

2.2. Haar functions

Given a dyadic grid $\mathcal{D}$ on $\mathbb{R}$, we associate to each $I \in \mathcal{D}$ the cancellative Haar function $h_I := h_I^0 = \frac{1}{\sqrt{|I|}}(\mathbb{1}_{I_+} - \mathbb{1}_{I_-})$, where I_+ and I_- are the right and left halves of I, respectively. The non-cancellative Haar function is $h_I^1 := \frac{1}{\sqrt{|I|}}\mathbb{1}_I$. The cancellative Haar functions $\{h_I\}_{I \in \mathcal{D}}$ form an orthonormal basis for $L^2(\mathbb{R}, dx)$, and an unconditional basis for $L^p(\mathbb{R})$, $1 < p < \infty$. Throughout this paper, we let $(\cdot, \cdot)$ denote inner product in $L^2(dx)$, so we write for example

$$f = \sum_{I \in \mathcal{D}} (f, h_I)\, h_I,$$

where $(f, h_I) = \int f h_I \, dx$ is the Haar coefficient of f corresponding to I.

In $\mathbb{R}^n$, we have $2^n - 1$ cancellative Haar functions and one non-cancellative: for every dyadic cube $Q = I_1 \times I_2 \times \cdots \times I_n$, where every $I_k \in \mathcal{D}$ is a dyadic interval with common length $|I_k| = \ell(Q)$, we let

$$h_Q^\varepsilon(x) := h_{I_1 \times \cdots \times I_n}^{(\varepsilon_1, \ldots, \varepsilon_n)}(x_1, \ldots, x_n) = \prod_{k=1}^n h_{I_k}^{\varepsilon_k}(x_k),$$

where $\varepsilon_k \in \{0, 1\}$ for all k, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ is known as the signature of h_Q^ε. The function h_Q^ε is cancellative except in one case, when $\varepsilon \equiv 1$. As in $\mathbb{R}$, the cancellative Haar functions form an orthonormal basis for $L^2(\mathbb{R}^n, dx)$, and an unconditional basis for $L^p(\mathbb{R}^n, dx)$, $1 < p < \infty$.

Remark 2.1. We often write

$$f = \sum_{Q \in \mathcal{D}} (f, h_Q)\, h_Q$$

to mean

$$f = \sum_{\substack{Q \in \mathcal{D}, \\ \varepsilon \not\equiv 1}} (f, h_Q^\varepsilon)\, h_Q^\varepsilon,$$

omitting the signatures, and understanding that h_Q always refers to a cancellative Haar function. There is really only one instance for us where the signatures matter, and that is in the definition of the paraproduct Γ_b in $\mathbb{R}^n$, $n > 1$.

Note that whenever $P \subsetneq Q$ for some dyadic cubes P and Q, the Haar function h_Q will be constant on P. We denote this constant by

$$h_Q(P) := \text{ the constant value } h_Q \text{ takes on } P \subsetneq Q.$$

It is easy to show that

$$\langle f \rangle_Q = \sum_{R \supsetneq Q} (f, h_R)\, h_R(Q), \quad \forall Q \in \mathcal{D},$$

where throughout the paper

$$\langle f \rangle_Q := \frac{1}{|Q|} \int_Q f \, dx$$

denotes average over Q, and sums such as $\sum_{P \subset Q}$ or $\sum_{R \supset Q}$ are understood to be over dyadic cubes.

2.3. A_p weights

A weight is a locally integrable, a.e. positive function $w(x)$ on $\mathbb{R}^n$. Any such weight immediately gives a measure on $\mathbb{R}^n$ via $dw := w(x)\,dx$ and

$$\int f \, dw := \int f(x)\, w(x)\, dx$$

yields the obvious L^p-spaces associated with the measure w, which we denote by $L^p(w)$:

$$\| f \|_{L^p(w)}^p := \int |f(x)|^p \, w(x)\, dx.$$

Given $1 < p < \infty$, we say $w \in A_p$ if

$$[w]_{A_p} := \sup_Q \langle w \rangle_Q \, \langle w' \rangle_Q^{p-1} < \infty,$$

where the supremum is over cubes $Q \subset \mathbb{R}^n$, p' denotes the Hölder conjugate of p, that is, $1/p + 1/p' = 1$, and the conjugate weight w' is given by

$$w' := w^{1-p'} = w^{-p'/p} = w^{-1/(p-1)}.$$

In fact, $w \in A_p$ if and only if the conjugate weight w' is in $A_{p'}$, with

$$[w']_{A_{p'}} = [w]_{A_p}^{1/(p-1)}.$$

We restrict our attention to dyadic A_p weights, denoted $A_p^{\mathcal{D}}$, and defined in the same way except the supremum is only over dyadic cubes $Q \in \mathcal{D}$. Sometimes we use the standard L^p-duality $(L^p(w))^* = L^{p'}(w)$ with inner product $(\cdot, \cdot)_{dw}$, and other times we think of $(L^p(w))^* \simeq L^{p'}(w')$ with regular Lebesgue inner product $(\cdot, \cdot)$. We refer the reader to Chapter 9 of [7] for a thorough treatment of A_p weights.

2.4. Paraproducts and BMO

We say $b \in \mathrm{BMO}(\mathbb{R}^n)$ if

$$\|b\|_{\mathrm{BMO}(\mathbb{R}^n)} := \sup_Q \frac{1}{|Q|} \int_Q |b(x) - \langle b \rangle_Q| \, dx < \infty,$$

where the supremum is over cubes $Q \subset \mathbb{R}^n$. Given a weight w on $\mathbb{R}^n$, we say that b is in the weighted BMO space $\mathrm{BMO}(w)$, $b \in \mathrm{BMO}(w)$, if

$$\|b\|_{\mathrm{BMO}(w)} := \sup_Q \frac{1}{w(Q)} \int_Q |b(x) - \langle b \rangle_Q| \, dx < \infty.$$

We similarly restrict our attention to dyadic BMO spaces, $\mathrm{BMO}^{\mathcal{D}}$ and $\mathrm{BMO}^{\mathcal{D}}(w)$ for the weighted version, both defined in the same way except the supremum is over dyadic cubes $Q \in \mathcal{D}$.

In $\mathbb{R}$, we have two paraproducts:

$$\Pi_b f(x) := \sum_{I \in \mathcal{D}} (b, h_I) \langle f \rangle_I h_I(x) \quad \text{and} \quad \Pi_b^* f(x) := \sum_{I \in \mathcal{D}} (b, h_I) (f, h_I) \frac{\mathbb{1}_I(x)}{|I|}.$$

They have the property that

$$bf = \Pi_b f + \Pi_b^* f + \Pi_f b,$$

and their boundedness is usually characterized by some BMO-type norm of the symbol b.

In $\mathbb{R}^n$, we have three paraproducts:

$$\Pi_b f(x) := \sum_{Q \in \mathcal{D}} (b, h_Q) \langle f \rangle_Q h_Q(x),$$

$$\Pi_b^* f(x) := \sum_{Q \in \mathcal{D}} (b, h_Q) (f, h_Q) \frac{\mathbb{1}_Q(x)}{|Q|},$$

$$\Gamma_b f(x) := \sum_{Q \in \mathcal{D}} \sum_{\varepsilon, \eta \neq 1; \varepsilon \neq \eta} (b, h_Q^\varepsilon) (f, h_Q^\eta) \frac{1}{\sqrt{|Q|}} h_Q^{\varepsilon + \eta}.$$

The paraproducts Π_b and Π_b^* are adjoints in $L^2(\mathbb{R}^n)$, and Γ_b is self-adjoint. Again, keep in mind that signatures are omitted in the definitions of Π_b and Π_b^* since all Haar functions involved have the same signature. As for Γ_b, signatures are included because they differ – see [8] for details. Generally, in the L^p-situation, we still have

$$(\Pi_b f, g) = (f, \Pi_b^* g),$$

so if we think of $\Pi_b \colon L^p(\mu) \to L^p(\lambda)$ for two A_p weights μ and λ, then its adjoint is $\Pi_b^* \colon L^{p'}(\lambda') \to L^{p'}(\mu')$ – where we are thinking of Banach space duality in terms of $(L^p(\mu))^* \simeq L^{p'}(\mu')$ and $(L^p(\lambda))^* \simeq L^{p'}(\lambda')$, both with regular Lebesgue inner product $(\cdot, \cdot)$.

3. Sparse BMO functions

3.1. Sparse families

Let $0 < \eta < 1$. A collection $\mathcal{S} \subset \mathcal{D}$ is said to be η-*sparse* if for every $Q \in \mathcal{S}$ there is a measurable subset $E_Q \subset Q$ such that the sets $\{E_Q\}_{Q\in\mathcal{S}}$ are pairwise disjoint, and satisfy $|E_Q| \geq \eta|Q|$ for all $Q \in \mathcal{S}$.

Let $\Lambda > 1$. A family $\mathcal{S} \subset \mathcal{D}$ is said to be Λ-*Carleson* if

$$\sum_{P\in\mathcal{S}, P\subset Q} |P| \leq \Lambda|Q|, \quad \forall Q \in \mathcal{D}.$$

It is easy to see that it suffices to impose this condition only on $Q \in \mathcal{S}$. It is also easy to see that any η-sparse collection is $1/\eta$-Carleson. Far less obvious is the remarkable property that any Λ-Carleson collection is $1/\Lambda$-sparse, which is proved in the now classic work [11].

A special type of sparse collection which appears most frequently in practice is defined in terms of so-called "$\mathcal{S}$-children." Suppose a family $\mathcal{S} \subset \mathcal{D}$ has the property that

$$\sum_{P\in \mathrm{ch}_{\mathcal{S}}(Q)} |P| \leq \alpha|Q|, \quad \forall Q \in \mathcal{S},$$

where $\alpha \in (0, 1)$ and $\mathrm{ch}_{\mathcal{S}}(Q)$, the $\mathcal{S}$-children of Q, is the collection of maximal $P \in \mathcal{S}$ such that $P \subsetneq Q$. Then $\mathcal{S}$ is $(1-\alpha)$-sparse: let

$$E_Q := Q \setminus \bigcup_{P\in \mathrm{ch}_{\mathcal{S}}(Q)} P,$$

which are clearly pairwise disjoint, and satisfy $|E_Q| \geq (1-\alpha)|Q|$. A collection that is sparse with respect to the Lebesgue measure is also sparse with respect to any A_p measure w. Recall that (see [7], Proposition 9.1.5) an equivalent definition for $[w]_{A_p}$ is

$$[w]_{A_p} = \sup_{Q\in\mathcal{D}} \sup_{\substack{f\in L^p(Q,w) \\ |Q\cap\{|f|=0\}|=0}} \frac{\langle |f| \rangle_Q^p}{\mathbb{E}_Q^w(|f|^p)},$$

where

$$\mathbb{E}_Q^w f := \frac{1}{w(Q)} \int_Q f \, dw.$$

Taking $f = \mathbb{1}_A$ above, for some measurable subset A of a fixed dyadic cube Q, we get

$$\Big(\frac{|A|}{|Q|}\Big)^p \leq [w]_{A_p} \frac{w(A)}{w(Q)}, \quad \forall A \subset Q, Q \in \mathcal{D}.$$

So, say $\mathcal{S}$ is η-sparse with pairwise disjoint $\{E_Q\}_{Q\in\mathcal{S}}$ subsets $E_Q \subset Q$ and $|E_Q| \geq \eta|Q|$. Then

$$\eta^p \leq \Big(\frac{|E_Q|}{|Q|}\Big)^p \leq [w]_{A_p} \frac{w(E_Q)}{w(Q)},$$

and

$$w(Q) \leq \frac{1}{\eta^p} [w]_{A_p} w(E_Q), \quad \forall Q \in \mathcal{S}. \tag{3.1}$$

3.2. Sparse BMO functions

We borrow the following terminology from [9]: we say a collection $\mathcal{S} \subset \mathcal{D}$ has an *infinitely increasing chain* if there exist $\{Q_K\}_{k\in\mathbb{N}}$, $Q_k \in \mathcal{S}$, such that $Q_k \subsetneq Q_{k+1}$, for all $k \in \mathbb{N}$. The following lemma is also found in [9].

Lemma 3.1. *If a collection $\mathcal{S} \subset \mathcal{D}$ has no infinitely increasing chains, then every $Q \in \mathcal{S}$ is contained in a maximal $Q^* \in \mathcal{S}$, in the sense that there exists no $R \in \mathcal{S}$ such that $R \supsetneq Q$. Any two maximal elements P^* and Q^* of $\mathcal{S}$ are disjoint.*

This type of collections will be important for us, so we let

$$\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$$

denote the set of all sparse collections in $\mathcal{D}$ which have no infinitely increasing chains.

Lemma 3.2. *Let $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ be a sparse collection with no infinitely increasing chains. Then the set of points contained in infinitely many elements of $\mathcal{S}$ has measure* 0.

Proof. Let $\mathcal{S}^*$ denote the collection of maximal elements of $\mathcal{S}$. Since $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$, every $Q \in \mathcal{S}$ is contained in a unique $Q^* \in \mathcal{S}^*$. Any x which belongs to infinitely may elements of $\mathcal{S}$ must then belong to an infinitely decreasing chain

$$x \in \cdots \subsetneq Q_k \subsetneq \cdots \subsetneq Q_2 \subsetneq Q_1 = Q^*$$

terminating at some maximal $Q^* \in \mathcal{S}^*$. Fix any such chain and let A be the set of points contained in all Q_k, that is, $A = \bigcap_{k=1}^{\infty} Q_k$. Then for any $k \in \mathbb{N}$,

$$k|A| \le \sum_{i=1}^{k} |Q_i| \le \Lambda |Q^*|,$$

where Λ is the Carleson constant of $\mathcal{S}$. So $|A| \le \frac{1}{k}\Lambda|Q^*|$ for all $k \in \mathbb{N}$, and then $|A| = 0$.

Alternatively, since $\{Q_k\}$ is a decreasing nest of sets, we have $|A| = \lim_{k\to\infty} |Q_k|$, and $\lim_{k\to\infty} |Q_k| = 0$ because the series

$$\sum_{k=1}^{\infty} |Q_k| \le \sum_{Q\in\mathcal{S}, Q\subset Q^*} |Q| \le \Lambda |Q^*|$$

converges. ∎

The lemma above ensures that the following definition is sound: with every sparse collection $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ with no infinitely increasing chains we associate the function

$$b_{\mathcal{S}} := \sum_{Q\in\mathcal{S}} \mathbb{1}_Q.$$

By Lemma 3.2 we know that $b_{\mathcal{S}}$ is almost everywhere finite: if x is contained in infinitely many elements of $\mathcal{S}$, then $b_{\mathcal{S}}(x) = \infty$, but this can only happen on a set of measure zero.

Note also that $b_{\mathcal{S}}$ is locally integrable: for some $Q_0 \in \mathcal{D}$,

$$\begin{aligned}\langle b_{\mathcal{S}}\rangle_{Q_0} &= \frac{1}{|Q_0|}\Big(\sum_{Q\in\mathcal{S},Q\subset Q_0}|Q| + \sum_{Q\in\mathcal{S},Q\supsetneq Q_0}|Q_0|\Big)\\ &= \underbrace{\frac{1}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subset Q_0}|Q|}_{\leq\Lambda} + \underbrace{\#\{Q\in\mathcal{S}:Q\supsetneq Q_0\}}_{<\infty \text{ because } \mathcal{S}\in\Upsilon^{\mathcal{D}}} < \infty.\end{aligned}$$

Then, for some $Q_0 \in \mathcal{D}$,

$$\begin{aligned}(b_{\mathcal{S}} - \langle b_{\mathcal{S}}\rangle_{Q_0})\,\mathbb{1}_{Q_0} &= \sum_{Q\in\mathcal{S},Q\subset Q_0}\mathbb{1}_Q + \#\{Q\in\mathcal{S}:Q\supsetneq Q_0\}\,\mathbb{1}_{Q_0}\\ &\quad - \frac{\mathbb{1}_{Q_0}}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subset Q_0}|Q| - \#\{Q\in\mathcal{S}:Q\supsetneq Q_0\}\mathbb{1}_{Q_0}\\ &= \sum_{Q\in\mathcal{S},Q\subset Q_0}\mathbb{1}_Q - \frac{\mathbb{1}_{Q_0}}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subset Q_0}|Q|.\end{aligned}$$

In fact, we can reduce this further to

$$(b_{\mathcal{S}} - \langle b_{\mathcal{S}}\rangle_{Q_0})\,\mathbb{1}_{Q_0} = \sum_{Q\in\mathcal{S},Q\subsetneq Q_0}\mathbb{1}_Q - \frac{\mathbb{1}_{Q_0}}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subsetneq Q_0}|Q|, \tag{3.2}$$

which is clear if $Q_0 \notin \mathcal{S}$, and if $Q_0 \in \mathcal{S}$, then $\mathbb{1}_{Q_0} - \frac{\mathbb{1}_{Q_0}}{|Q_0|}|Q_0|$ cancel. A simple estimate then shows that

$$\frac{1}{|Q_0|}\int_{Q_0}|b_{\mathcal{S}} - \langle b_{\mathcal{S}}\rangle_{Q_0}|\,dx \leq \frac{1}{|Q_0|}\,2\sum_{Q\in\mathcal{S},Q\subsetneq Q_0}|Q| \leq 2\Lambda, \quad \forall Q_0\in\mathcal{D},$$

so $b_{\mathcal{S}} \in \mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)$. However, a more careful estimate is possible. We prove[1] the following in Appendix A.

Theorem 3.3. *Let $\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ be a sparse collection with no infinitely increasing chains and Carleson constant Λ. Then the function $b_{\mathcal{S}} = \sum_{Q\in\mathcal{S}}\mathbb{1}_Q$ is in $\mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)$, with*

$$\|b_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)} \leq \Lambda.$$

This process works to yield a weighted BMO function as well: with any $\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ and $w\in A_p^{\mathcal{D}}$ we associate the function

$$b_{\mathcal{S}}^w := \sum_{Q\in\mathcal{S}}\langle w\rangle_Q\,\mathbb{1}_Q.$$

[1] We include this proof simply because we found it interesting, and we hope the process of breaking down the integrals outlined in this proof can be adapted to the weighted situation in a future work.

As before, $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ ensures that $b_{\mathcal{S}}^w$ is a.e. finite, locally integrable, and

$$\mathbb{1}_{Q_0}(b_{\mathcal{S}}^w - \langle b_{\mathcal{S}}^w\rangle_{Q_0}) = \sum_{Q\in\mathcal{S},Q\subsetneq Q_0} \langle w\rangle_Q \mathbb{1}_Q - \frac{\mathbb{1}_{Q_0}}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subsetneq Q_0} w(Q), \quad \forall Q_0 \in \mathcal{D}.$$

By (3.1),

$$\frac{1}{|Q_0|}\sum_{Q\in\mathcal{S},Q\subsetneq Q_0} w(Q) \le [w]_{A_p}\Lambda^p\langle w\rangle_{Q_0},$$

which then easily gives

$$\frac{1}{w(Q_0)}\int_{Q_0} |b_{\mathcal{S}}^w - \langle b_{\mathcal{S}}^w\rangle_{Q_0}|\,dx \le 2[w]_{A_p}\Lambda^p,$$

so

$$b_{\mathcal{S}}^w \in \mathrm{BMO}^{\mathcal{D}}(w), \quad \text{with} \quad \|b_{\mathcal{S}}^w\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \le 2[w]_{A_p}\Lambda^p.$$

In fact, appealing to A_∞ instead, we can eliminate the power of Λ and simply obtain

$$\|b_{\mathcal{S}}^w\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \lesssim [w]_{A_\infty}\Lambda \le [w]_{A_p}\Lambda.$$

3.3. Sparse operators as sums of paraproducts and martingale transforms

For ease of notation, we work in $\mathbb{R}$ below, but the obvious analog for $\mathbb{R}^n$ follows easily in the same way. Consider

$$\mathcal{A}_{\mathcal{S}}^w f := \sum_{I\in\mathcal{S}} \langle w\rangle_I \langle f\rangle_I \mathbb{1}_I,$$

where $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R})$ and w is an $A_p^{\mathcal{D}}$ weight on $\mathbb{R}$, $1 < p < \infty$. A particularly interesting instance of $\mathcal{A}_{\mathcal{S}}^w$ occurs when $w = v \in A_2^{\mathcal{D}}$, where $v := \mu^{1/p}\lambda^{-1/p}$ for two weights $\mu, \lambda \in A_p^{\mathcal{D}}$. We treat this operator in more detail in Section 3.5.

Using the $b_{\mathcal{S}}^w$ function associated with $\mathcal{S}$ and w, we write

$$\mathcal{A}_{\mathcal{S}}^w f = \mathcal{A}_{\mathcal{S}}^w f - b_{\mathcal{S}}^w \cdot f + b_{\mathcal{S}}^w \cdot f = \mathcal{A}_{\mathcal{S}}^w f - b_{\mathcal{S}}^w \cdot f + (\Pi_{b_{\mathcal{S}}^w} f + \Pi_{b_{\mathcal{S}}^w}^* f + \Pi_f b_{\mathcal{S}}^w). \tag{3.3}$$

Now recall that

$$\langle b_{\mathcal{S}}^w\rangle_{J_0} = (\tau_{\mathcal{S}}^w)_{J_0} + \sum_{J\in\mathcal{S},J\supset J_0} \langle w\rangle_J, \quad \forall J_0 \in \mathcal{D},$$

where

$$(\tau_{\mathcal{S}}^w)_J := \frac{1}{|J|}\sum_{I\in\mathcal{S},I\subsetneq J} w(I), \quad \forall J \in \mathcal{D},$$

a quantity always bounded if $w \in A_p^{\mathcal{D}}$:

$$(\tau_{\mathcal{S}}^w)_J \le [w]_{A_p}\Lambda^p\langle w\rangle_J.$$

So

$$\begin{aligned}\Pi_f b_{\mathcal{S}}^w(x) &= \sum_{J\in\mathcal{D}} (f,h_J)\,\langle b_{\mathcal{S}}^w\rangle_J\, h_J(x)\\ &= \sum_{J\in\mathcal{D}} (f,h_J)\Big[(\tau_{\mathcal{S}}^w)_J + \sum_{K\in\mathcal{S},K\supset J}\langle w\rangle_K\Big] h_J(x)\\ &= \underbrace{\sum_{J\in\mathcal{D}} (\tau_{\mathcal{S}}^w)_J\,(f,h_J)\,h_J(x)}_{=:\,T_{\tau_{\mathcal{S}}^w} f(x)} + \sum_{J\in\mathcal{D}} (f,h_J)\,h_J(x)\Big(\sum_{K\in\mathcal{S},K\supset J}\langle w\rangle_K\Big).\end{aligned}$$

The second term can be further explored as

$$\begin{aligned}&\sum_{J\in\mathcal{D}} (f,h_J)\,h_J(x)\Big(\sum_{K\in\mathcal{S},K\supset J}\langle w\rangle_K\Big)\\ &\quad= \sum_{K\in\mathcal{S}}\langle w\rangle_K\Big(\sum_{J\subset K}(f,h_J)\,h_J(x)\Big) = \sum_{K\in\mathcal{S}}\langle w\rangle_K\,(f(x)-\langle f\rangle_K)\,\mathbb{1}_K(x)\\ &\quad= f(x)\cdot\sum_{K\in\mathcal{S}}\langle w\rangle_K\,\mathbb{1}_K(x) - \sum_{K\in\mathcal{S}}\langle w\rangle_K\,\langle f\rangle_K\,\mathbb{1}_K(x) = f(x)\cdot b_{\mathcal{S}}^w(x) - \mathcal{A}_{\mathcal{S}}^w f(x).\end{aligned}$$

Returning to (3.3),

$$\mathcal{A}_{\mathcal{S}}^w f = \mathcal{A}_{\mathcal{S}}^w f - b_{\mathcal{S}}^w\cdot f + (\Pi_{b_{\mathcal{S}}^w} f + \Pi^*_{b_{\mathcal{S}}^w} f) + T_{\tau_{\mathcal{S}}^w} f + f\cdot b_{\mathcal{S}}^w - \mathcal{A}_{\mathcal{S}}^w f,$$

so we have the following.

Proposition 3.4. *Any weighted sparse operator $\mathcal{A}_{\mathcal{S}}^w$, where $w\in A_p^{\mathcal{D}}$ is a weight on $\mathbb{R}$ and $\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R})$ is a sparse collection with no infinitely increasing chains, may be expressed as*

$$\mathcal{A}_{\mathcal{S}}^w f = \Pi_{b_{\mathcal{S}}^w} f + \Pi^*_{b_{\mathcal{S}}^w} f + T_{\tau_{\mathcal{S}}^w} f, \tag{3.4}$$

where the first two terms are the paraproducts with symbol $b_{\mathcal{S}}^w$, the sparse $\mathrm{BMO}^{\mathcal{D}}(w)$ function associated with $\mathcal{S}$ and w, and the third term is

$$T_{\tau_{\mathcal{S}}^w} f(x) := \sum_{J\in\mathcal{D}} (\tau_{\mathcal{S}}^w)_J\,(f,h_J)\,h_J(x),$$

$$\text{where}\quad (\tau_{\mathcal{S}}^w)_J := \frac{1}{|J|}\sum_{I\in\mathcal{S},I\subsetneq J} w(I) \le [w]_{A_p}\,\Lambda^p\,\langle w\rangle_J,\quad \forall J\in\mathcal{D}.$$

Remark 3.5. In case $w\equiv 1$, we obtain the unweighted situation

$$\mathcal{A}_{\mathcal{S}} f = \Pi_{b_{\mathcal{S}}} f + \Pi^*_{b_{\mathcal{S}}} f + T_{\tau_{\mathcal{S}}} f, \tag{3.5}$$

where $T_{\tau_{\mathcal{S}}}$ is a martingale transform:

$$T_{\tau_{\mathcal{S}}} = \sum_{J\in\mathcal{D}} (\tau_{\mathcal{S}})_J\,(f,h_J)\,h_J,\quad \text{where}\quad (\tau_{\mathcal{S}})_J := \frac{1}{|J|}\sum_{I\in\mathcal{S},I\subsetneq J}|I|\le\Lambda,\quad \forall J\in\mathcal{D}.$$

Remark 3.6. In fact, (3.4) expresses sparse operators as Haar multipliers: recall that a Haar multiplier is an operator of the form

$$T_\phi f(x) := \sum_{J\in\mathcal{D}} \phi_J(x)\,(f,h_J)\,h_J(x),$$

where $\{\phi_J(x)\}_{J\in\mathcal{D}}$ is a sequence of functions indexed by $\mathcal{D}$. It is known (see [2]) that

$$(\Pi_b + \Pi_b^*) f = \sum_{J\in\mathcal{D}} (b - \langle b\rangle_J)\,(f,h_J)\,h_J.$$

So, from (3.4),

$$\mathcal{A}_{\mathcal{S}}^w f(x) = \sum_{J\in\mathcal{D}} \Big[\underbrace{(b_{\mathcal{S}}^w(x) - \langle b_{\mathcal{S}}^w\rangle_J)\,\mathbb{1}_J(x) + (\tau_{\mathcal{S}}^w)_J}_{\phi_J(x)} \Big]\,(f,h_J)\,h_J(x).$$

Look more closely now at (3.5): $\mathcal{A}_{\mathcal{S}} = \Pi_{b_{\mathcal{S}}} + \Pi_{b_{\mathcal{S}}}^* + T_{\tau_{\mathcal{S}}}$. This gives an *upper bound* for $\|\mathcal{A}_{\mathcal{S}} : L^p(w) \to L^p(w)\|$ in terms of the norms of paraproducts and the martingale transform (when usually it is the norms of sparse operators that are used as upper bounds):

$$\|\mathcal{A}_{\mathcal{S}} f\|_{L^p(w)} \le \|\Pi_{b_{\mathcal{S}}} f\|_{L^p(w)} + \|\Pi_{b_{\mathcal{S}}}^* f\|_{L^p(w)} + \|T_{\tau_{\mathcal{S}}} f\|_{L^p(w)}.$$

Divide above by $\Lambda_{(\mathcal{S})} := \Lambda$, the Carleson constant of $\mathcal{S}$, and recall that $\|b_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}} \le \Lambda$, as well as $\|\tau_{\mathcal{S}}\|_\infty \le \Lambda$:

$$\begin{aligned}\frac{\|\mathcal{A}_{\mathcal{S}} f\|_{L^p(w)}}{\Lambda} &\le \frac{\|\Pi_{b_{\mathcal{S}}} f\|_{L^p(w)}}{\Lambda} + \frac{\|\Pi_{b_{\mathcal{S}}}^* f\|_{L^p(w)}}{\Lambda} + \frac{\|T_{\tau_{\mathcal{S}}} f\|_{L^p(w)}}{\Lambda}\\ &\le \frac{\|\Pi_{b_{\mathcal{S}}} f\|_{L^p(w)}}{\|b_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}}} + \frac{\|\Pi_{b_{\mathcal{S}}}^* f\|_{L^p(w)}}{\|b_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}}} + \frac{\|T_{\tau_{\mathcal{S}}} f\|_{L^p(w)}}{\|\tau_{\mathcal{S}}\|_\infty},\end{aligned}$$

from which we can deduce that, for all $\Lambda > 1$,

$$\begin{aligned}\sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R})\\ \Lambda_{(\mathcal{S})}=\Lambda}} \frac{\|\mathcal{A}_{\mathcal{S}} : L^p(w)\to L^p(w)\|}{\Lambda} &\le \sup_{\substack{b\in\mathrm{BMO}^{\mathcal{D}}\\ \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\neq 0}} \frac{\|\Pi_b : L^p(w)\to L^p(w)\|}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}}\\ &+ \sup_{\substack{b\in\mathrm{BMO}^{\mathcal{D}}\\ \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\neq 0}} \frac{\|\Pi_b^* : L^p(w)\to L^p(w)\|}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}} + \sup_{\substack{\tau\in\ell^\infty\\ \|\tau\|_\infty\neq 0}} \frac{\|T_\tau : L^p(w)\to L^p(w)\|}{\|\tau\|_\infty}.\end{aligned}$$

Given the well-known domination results [10] for the martingale transform and paraproducts, we obtain

$$\begin{aligned}\sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R})\\ \Lambda_{(\mathcal{S})}=\Lambda}} \|\mathcal{A}_{\mathcal{S}} : L^p(w)\to L^p(w)\| &\simeq_{\Lambda,p} \sup_{\substack{b\in\mathrm{BMO}^{\mathcal{D}}\\ \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\neq 0}} \frac{\|\Pi_b : L^p(w)\to L^p(w)\|}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}}\\ &+ \sup_{\substack{b\in\mathrm{BMO}^{\mathcal{D}}\\ \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\neq 0}} \frac{\|\Pi_b^* : L^p(w)\to L^p(w)\|}{\|b\|_{\mathrm{BMO}^{\mathcal{D}}}} + \sup_{\substack{\tau\in\ell^\infty\\ \|\tau\|_\infty\neq 0}} \frac{\|T_\tau : L^p(w)\to L^p(w)\|}{\|\tau\|_\infty}.\end{aligned}$$

Remark 3.7. It would be interesting if the martingale and the paraproducts could be "separated" somehow, and to obtain independently that the paraproducts and the martingale transform have the same dependency on $[w]_{A_p}$ by showing they are both equivalent to $\|\mathcal{A}_{\mathcal{S}}\|$. However, we can show that the norms of $\mathcal{A}_{\mathcal{S}}$ are equivalent to norms of certain compositions of paraproducts. We do this next.

3.4. Sparse operators and compositions of paraproducts

Consider the composition

$$\Pi_a^* \Pi_b f = \sum_{Q \in \mathcal{D}} (a, h_Q)(b, h_Q) \langle f \rangle_Q \frac{\mathbb{1}_Q}{|Q|}.$$

We show in Appendix B, using a bilinear form argument, that:

Theorem 3.8. *There is a dimensional constant $C(n)$ such that, for all Bloom weights on $\mathbb{R}^n$ $\mu, \lambda \in A_p$ $(1 < p < \infty)$, $\nu := \mu^{1/p}\lambda^{-1/p}$, BMO functions $a \in \mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)$ and $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$, and $\Lambda > 1$,*

$$\|\Pi_a^* \Pi_b : L^p(\mu) \to L^p(\lambda)\| \leq C(n) \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} \sup_{\substack{\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n) \\ \Lambda_{(\mathcal{S})} = \Lambda}} \Big(\frac{\Lambda}{\Lambda - 1}\Big)^3 \|\mathcal{A}_{\mathcal{S}}^{\nu} : L^p(\mu) \to L^p(\lambda)\|.$$

Some immediate observations about this result:

- From Theorem 3.9,
$$\|\Pi_a^* \Pi_b : L^p(\mu) \to L^p(\lambda)\| \lesssim \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} [\mu]_{A_p}^{1/(p-1)} [\lambda]_{A_p}.$$
- Take $\mu = \lambda = w$, for some $w \in A_p$. Then $\nu = 1$ and we obtain in the one-weight situation:
$$\|\Pi_a^* \Pi_b : L^p(w) \to L^p(w)\| \lesssim \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}} [w]_{A_p}^{\max(1, 1/(p-1))}.$$
- It is easy to see that $\Pi_a^* \Pi_b = \Pi_b^* \Pi_a$, so the same result holds for $\Pi_b^* \Pi_a$, with $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$ and $a \in \mathrm{BMO}^{\mathcal{D}}$.

Let $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$. We associated with $\mathcal{S}$ the BMO function $b_{\mathcal{S}} = \sum_{Q \in \mathcal{S}} \mathbb{1}_Q$. There is another, even more obvious BMO function we can associate with $\mathcal{S}$:

$$\widetilde{b}_{\mathcal{S}} := \sum_{Q \in \mathcal{S}} \sqrt{|Q|}\, h_Q = \sum_{Q \in \mathcal{S}, \varepsilon \neq 1} \sqrt{|Q|}\, h_Q^{\varepsilon}.$$

For any $Q_0 \in \mathcal{D}$,

$$\frac{1}{|Q_0|} \int_{Q_0} |\widetilde{b}_{\mathcal{S}} - \langle \widetilde{b}_{\mathcal{S}} \rangle_{Q_0}|^2 \, dx = \frac{1}{|Q_0|} \sum_{\substack{Q \subset Q_0, Q \in \mathcal{S} \\ \varepsilon \neq 1}} |Q| = (2^n - 1) \frac{1}{|Q_0|} \sum_{Q \subset Q_0, Q \in \mathcal{S}} |Q| \leq (2^n - 1)\Lambda,$$

so

$$\|\widetilde{b}_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}} \leq \sqrt{(2^n - 1)\Lambda}.$$

Moreover,

$$\Pi_{\widetilde{b}_{\mathcal{S}}}^{*}\,\Pi_{\widetilde{b}_{\mathcal{S}}}\,f=\sum_{\substack{Q\in\mathcal{D}\\ \varepsilon\neq 1}}(\widetilde{b}_{\mathcal{S}},h_Q^{\varepsilon})^2\,\langle f\rangle_Q\,\frac{\mathbb{1}_Q}{|Q|}=\sum_{\substack{Q\in\mathcal{S}\\ \varepsilon\neq 1}}|Q|\,\langle f\rangle_Q\,\frac{\mathbb{1}_Q}{|Q|}=(2^n-1)\sum_{Q\in\mathcal{S}}\langle f\rangle_Q\,\mathbb{1}_Q,$$

so we may express the sparse operator $\mathcal{A}_{\mathcal{S}}$ as

$$\mathcal{A}_{\mathcal{S}}=\frac{1}{2^n-1}\,\Pi_{\widetilde{b}_{\mathcal{S}}}^{*}\,\Pi_{\widetilde{b}_{\mathcal{S}}}.$$

Then

$$\begin{aligned}\frac{\|\mathcal{A}_{\mathcal{S}}f\|_{L^p(w)}}{\Lambda}&=\frac{1}{2^n-1}\,\frac{\|\Pi_{\widetilde{b}_{\mathcal{S}}}^{*}\,\Pi_{\widetilde{b}_{\mathcal{S}}}f\|_{L^p(w)}}{\Lambda}\leq\frac{\|\Pi_{\widetilde{b}_{\mathcal{S}}}^{*}\,\Pi_{\widetilde{b}_{\mathcal{S}}}f\|_{L^p(w)}}{\|\widetilde{b}_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}}^2}\\&\leq\sup_{a,b\in\mathrm{BMO}^{\mathcal{D}}}\frac{\|\Pi_a^{*}\Pi_b f\|_{L^p(w)}}{\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\|b\|_{\mathrm{BMO}^{\mathcal{D}}}},\end{aligned}$$

which means that for all $\Lambda>1$,

$$\sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)\\ \Lambda_{(\mathcal{S})}=\Lambda}}\frac{\|\mathcal{A}_{\mathcal{S}}:L^p(w)\to L^p(w)\|}{\Lambda}\leq\sup_{a,b\in\mathrm{BMO}^{\mathcal{D}}}\frac{\|\Pi_a^{*}\Pi_b:L^p(w)\to L^p(w)\|}{\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}}}.$$

Combined with Theorem 3.8, we have

$$\sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}}(\mathbb{R}^n)\\ \Lambda_{(\mathcal{S})}=\Lambda}}\|\mathcal{A}_{\mathcal{S}}:L^p(w)\to L^p(w)\|\simeq_{\Lambda,p,n}\sup_{a,b\in\mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)}\frac{\|\Pi_a^{*}\Pi_b:L^p(w)\to L^p(w)\|}{\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\|b\|_{\mathrm{BMO}^{\mathcal{D}}}}.$$

3.5. The Bloom sparse operator $\mathcal{A}_{\mathcal{S}}^{\nu}$

Consider

$$\mathcal{A}_{\mathcal{S}}^{\nu}f=\sum_{Q\in\mathcal{S}}\langle\nu\rangle_Q\,\langle f\rangle_Q\,\mathbb{1}_Q$$

for a sparse collection $\mathcal{S}\subset\mathcal{D}(\mathbb{R}^n)$, where $\mu,\lambda\in A_p$ $(1<p<\infty)$ and $\nu:=\mu^{1/p}\lambda^{-1/p}$ are Bloom weights. In looking to bound this operator $L^p(\mu)\to L^p(\lambda)$, the first obvious route is to appeal to the known one-weight bounds for the usual, unweighted sparse operator

$$\mathcal{A}_{\mathcal{S}}f=\sum_{Q\in\mathcal{S}}\langle f\rangle_Q\mathbb{1}_Q.$$

We want something like $\|\mathcal{A}_{\mathcal{S}}^{\nu}f\|_{L^p(\lambda)}\leq C\,\|f\|_{L^p(\mu)}$, and we use duality to express

$$\|\mathcal{A}_{\mathcal{S}}^{\nu}f\|_{L^p(\lambda)}=\sup_{\substack{g\in L^{p'}(\lambda')\\ \|g\|_{L^{p'}(\lambda')}\leq 1}}|(\mathcal{A}_{\mathcal{S}}^{\nu}f,g)|.$$

So we look for a bound of the type $|(\mathcal{A}_{\mathcal{S}}^{\nu} f, g)| \leq C \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')}$. We proceed as follows:

$$\begin{aligned}
|(\mathcal{A}_{\mathcal{S}}^{\nu} f, g)| &= \Big| \sum_{Q \in \mathcal{S}} \langle \nu \rangle_Q \langle f \rangle_Q \langle g \rangle_Q |Q| \Big| \leq \sum_{Q \in \mathcal{S}} \langle |f| \rangle_Q \langle |g| \rangle_Q \nu(Q) \\
&\leq \int \Big(\sum_{Q \in \mathcal{S}} \langle |f| \rangle_Q \langle |g| \rangle_Q \mathbb{1}_Q \Big) d\nu \leq \int (\mathcal{A}_{\mathcal{S}} |f|) (\mathcal{A}_{\mathcal{S}} |g|) \mu^{1/p} \lambda^{-1/p} \, dx \\
&\leq \|\mathcal{A}_{\mathcal{S}} |f|\|_{L^p(\mu)} \|\mathcal{A}_{\mathcal{S}} |g|\|_{L^{p'}(\lambda')} \\
&\leq \|\mathcal{A}_{\mathcal{S}} : L^p(\mu) \to L^p(\mu)\| \cdot \|\mathcal{A}_{\mathcal{S}} : L^{p'}(\lambda') \to L^{p'}(\lambda')\| \cdot \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')}.
\end{aligned}$$

This yields the same dependency on the A_p characteristics of μ and λ as obtained in [12] for commutators:

$$\|\mathcal{A}_{\mathcal{S}}^{\nu} : L^p(\mu) \to L^p(\lambda)\| \lesssim ([\mu]_{A_p} [\lambda]_{A_p})^{\max(1, 1/(p-1))}$$

We give another proof, inspired by the beautiful proof in [5] of the A_2 conjecture for usual unweighted sparse operators, which yields a smaller bound.

Theorem 3.9. *Let $\mathcal{S} \subset \mathcal{D}$ be a sparse collection of dyadic cubes, let $\mu, \lambda \in A_p^{\mathcal{D}}$, for $1 < p < \infty$, be two A_p weights on $\mathbb{R}^n$, and let $\nu := \mu^{1/p} \lambda^{-1/p}$. Then the Bloom sparse operator*

$$\mathcal{A}_{\mathcal{S}}^{\nu} f := \sum_{Q \in \mathcal{S}} \langle \nu \rangle_Q \langle f \rangle_Q \mathbb{1}_Q$$

is bounded $L^p(\mu) \to L^p(\lambda)$ with

$$\begin{aligned}
\|\mathcal{A}_{\mathcal{S}}^{\nu} : L^p(\mu) \to L^p(\lambda)\| &\leq \Lambda^{p'+p-2} (p')^2 [\mu']_{A_{p'}} [\lambda]_{A_p} \\
&= \Lambda^{p'+p-2} (pp') [\mu]_{A_p}^{1/(p-1)} [\lambda]_{A_p},
\end{aligned} \tag{3.6}$$

where Λ is the Carleson constant of $\mathcal{S}$.

Proof. In looking for a bound of the type

$$\|\mathcal{A}_{\mathcal{S}}^{\nu} f\|_{L^p(\lambda)} \leq C \|f\|_{L^p(\mu)},$$

we consider instead $\varphi := f\mu'$: then $\|\varphi\|_{L^p(\mu)} = \|f\|_{L^p(\mu')}$, so we look instead for a bound of the type

$$\|\mathcal{A}_{\mathcal{S}}^{\nu} (f\mu')\|_{L^p(\lambda)} \leq C \|f\|_{L^p(\mu')}.$$

Using the standard $L^p(\lambda)$-$L^{p'}(\lambda)$ duality with the inner product $(\cdot, \cdot)_{d\lambda}$, we write

$$\|\mathcal{A}_{\mathcal{S}}^{\nu} (f\mu')\|_{L^p(\lambda)} = \sup_{g \in L^{p'}(\lambda), \|g\|_{L^{p'}(\lambda)} \leq 1} |(\mathcal{A}_{\mathcal{S}}^{\nu} (f\mu'), g\lambda)|,$$

meaning we finally look for a bound of the type

$$|(\mathcal{A}_{\mathcal{S}}^{\nu} (f\mu'), g\lambda)| \leq C \|f\|_{L^p(\mu')} \|g\|_{L^{p'}(\lambda)}.$$

As in [5], we make use of the weighted dyadic maximal function:

$$M_u^{\mathcal{D}} f(x) := \sup_{Q \in \mathcal{D}} \mathbb{E}_Q^u |f| \mathbb{1}_Q(x),$$

and the following property of being $L^q(u)$-bounded with a constant independent of u.

Theorem 3.10 (See, for example, [9] for a proof.). *For any locally finite Borel measure u on $\mathbb{R}^n$ and any $q \in (1, \infty)$,*

$$\|M_u^{\mathcal{D}} : L^q(u) \to L^q(u)\| \leq q'. \tag{3.7}$$

We bound

$$|(\mathcal{A}_{\mathcal{S}}^{\nu}(f\mu'), g\lambda)| = \Big| \sum_{Q \in \mathcal{S}} \langle \nu \rangle_Q \langle f\mu' \rangle_Q \langle g\lambda \rangle_Q |Q| \Big| \leq \sum_{Q \in \mathcal{S}} \langle |f|\mu' \rangle_Q \langle |g|\lambda \rangle_Q \langle \nu \rangle_Q |Q|.$$

We express the averages involving f and g as weighted averages:

$$\sum_{Q \in \mathcal{S}} \langle |f|\mu' \rangle_Q \langle |g|\lambda \rangle_Q \langle \nu \rangle_Q |Q| = \sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\mu'} |f|) \langle \mu' \rangle_Q (\mathbb{E}_Q^{\lambda} |g|) \langle \lambda \rangle_Q \langle \nu \rangle_Q |Q|.$$

Apply the fact that

$$\langle \nu \rangle_Q \leq \langle \mu \rangle_Q^{1/p} \langle \lambda' \rangle_Q^{1/p'}$$

(an easy consequence of Hölder's inequality), and the fact that for any A_p weight w, we have

$$[w]_{A_p}^{1/p} = \sup_Q \langle w \rangle_Q^{1/p} \langle w' \rangle_Q^{1/p'},$$

to go further:

$$\begin{aligned}
|(\mathcal{A}_{\mathcal{S}}^{\nu}(f\mu'), g\lambda)| &\leq \sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\mu'} |f|) (\mathbb{E}_Q^{\lambda} |g|) \langle \mu' \rangle_Q \langle \lambda \rangle_Q \langle \mu \rangle_Q^{1/p} \langle \lambda' \rangle_Q^{1/p'} |Q| \\
&\leq [\mu]_{A_p}^{1/p} [\lambda]_{A_p}^{1/p} \sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\mu'} |f|) (\mathbb{E}_Q^{\lambda} |g|) \langle \mu' \rangle_Q^{1/p} \langle \lambda \rangle_Q^{1/p'} |Q| \\
&= [\mu]_{A_p}^{1/p} [\lambda]_{A_p}^{1/p} \sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\mu'} |f|) (\mathbb{E}_Q^{\lambda} |g|) \mu'(Q)^{1/p} \lambda(Q)^{1/p'} \\
&\leq [\mu]_{A_p}^{1/p} [\lambda]_{A_p}^{1/p} \Big(\sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\mu'} |f|)^p \mu'(Q) \Big)^{1/p} \Big(\sum_{Q \in \mathcal{S}} (\mathbb{E}_Q^{\lambda} |g|)^{p'} \lambda(Q) \Big)^{1/p'}.
\end{aligned}$$

Applying (3.1), we get that

$$\mu'(Q) \leq [\mu']_{A_{p'}} \Lambda^{p'} \mu'(E_Q) = [\mu]_{A_p}^{p'-1} \Lambda^{p'} \mu'(E_Q)$$

and

$$\lambda(Q) \leq [\lambda]_{A_p} \Lambda^p \lambda(E_Q),$$

so we may later use disjointness of the sets $\{E_Q\}_{Q\in\mathcal{S}}$ in the integrals below:

$$\Big(\sum_{Q\in\mathcal{S}}\big(\mathbb{E}_Q^{\mu'}|f|\big)^p\mu'(Q)\Big)^{1/p}\le[\mu]_{A_p}^{(p'-1)/p}\Lambda^{p'/p}\Big(\sum_{Q\in\mathcal{S}}\big(\mathbb{E}_Q^{\mu'}|f|\big)^p\mu'(E_Q)\Big)^{1/p}$$

$$=[\mu]_{A_p}^{\frac{p'-1}{p}}\Lambda^{p'/p}\Big(\sum_{Q\in\mathcal{S}}\int_{E_Q}(\mathbb{E}_Q^{\mu'}|f|)^p\,d\mu'\Big)^{1/p}\le[\mu]_{A_p}^{\frac{p'-1}{p}}\Lambda^{p'/p}\Big(\sum_{Q\in\mathcal{S}}\int_{E_Q}(M_{\mu'}^{\mathcal{D}}f)^p\,d\mu'\Big)^{1/p}$$

$$\le[\mu]_{A_p}^{(p'-1)/p}\Lambda^{p'/p}\Big(\int_{\mathbb{R}^n}(M_{\mu'}^{\mathcal{D}}f)^p\,d\mu'\Big)^{1/p}=[\mu]_{A_p}^{(p'-1)/p}\Lambda^{p'/p}\|M_{\mu'}^{\mathcal{D}}f\|_{L^p(\mu')}$$

$$\le[\mu]_{A_p}^{(p'-1)/p}\Lambda^{p'/p}p'\|f\|_{L^p(\mu')}\Big(\sum_{Q\in\mathcal{S}}\big(\mathbb{E}_Q^{\lambda}|g|\big)^{p'}\lambda(Q)\Big)^{1/p'}$$

$$\le[\lambda]_{A_p}^{1/p'}\Lambda^{p/p'}\Big(\sum_{Q\in\mathcal{S}}(\mathbb{E}_Q^{\lambda}|g|)^{p'}\lambda(E_Q)\Big)^{1/p'}\le[\lambda]_{A_p}^{1/p'}\Lambda^{p/p'}\Big(\sum_{Q\in\mathcal{D}}\int_{E_Q}(M_{\lambda}^{\mathcal{D}}g)^{p'}\,d\lambda\Big)^{1/p'}$$

$$\le[\lambda]_{A_p}^{1/p'}\Lambda^{p/p'}\|M_{\lambda}^{\mathcal{D}}g\|_{L^{p'}(\lambda)}\le[\lambda]_{A_p}^{1/p'}\Lambda^{p/p'}p\|g\|_{L^{p'}(\lambda)}.$$

Putting these estimates together,

$$\begin{aligned}|(\mathcal{A}_{\mathcal{S}}^{\nu}(f\mu'),g\lambda)|&\le[\mu]_{A_p}^{1/p}[\lambda]_{A_p}^{1/p}[\mu]_{A_p}^{(p'-1)/p}\Lambda^{p'/p}p'\|f\|_{L^p(\mu')}[\lambda]_{A_p}^{1/p'}\Lambda^{p/p'}p\|g\|_{L^{p'}(\lambda)}\\&=[\mu]_{A_p}^{p'/p}[\lambda]_{A_p}\Lambda^{p'/p+p/p'}pp'\|f\|_{L^p(\mu')}\|g\|_{L^{p'}(\lambda)}\\&=[\mu']_{A_{p'}}[\lambda]_{A_p}\Lambda^{p+p'-2}pp'\|f\|_{L^p(\mu')}\|g\|_{L^{p'}(\lambda)},\end{aligned}$$

which proves Theorem 3.9. ∎

4. Paraproducts and Bloom BMO

We show the following pointwise domination result, inspired by ideas in [10] on pointwise domination of the martingale transform. We begin with a paraproduct restricted to a certain dyadic cube:

$$\Pi_{b,Q_0}f:=\sum_{Q\subset Q_0}(b,h_Q)\langle f\rangle_Q h_Q,$$

so everything is the same except we only look inside $\mathcal{D}(Q_0)$.

Theorem 4.1. *There is a dimensional constant $C(n)$ such that the following holds. For every $\Lambda>1$, weight w on $\mathbb{R}^n$, $b\in\mathrm{BMO}^{\mathcal{D}}(w)$, fixed dyadic cube $Q_0\in\mathcal{D}$ and $f\in L^1(Q_0)$, there is a Λ-Carleson sparse collection $\mathcal{S}(Q_0)\subset\mathcal{D}(Q_0)$ (depending on b,w,f) such that*

$$\forall x\in Q_0,\quad|\Pi_{b,Q_0}f(x)|\le C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^2\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\mathcal{A}_{\mathcal{S}(Q_0)}^{w}|f|(x).$$

The same holds for the other paraproducts Π_b^* and Γ_b.

Assuming this, return to the Bloom situation for a moment. We have:

Corollary 4.2. *Given Bloom weights $\mu, \lambda \in A_p^{\mathcal{D}}$, $\nu = \mu^{1/p}\lambda^{-1/p}$, for all $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$,*

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| \leq C(n)\|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} \sup_{\substack{\mathcal{S}\in\Upsilon^{\mathcal{D}} \\ \Lambda_{(\mathcal{S})}=\Lambda}} \Big(\frac{\Lambda}{\Lambda-1}\Big)^2 \|\mathcal{A}_{\mathcal{S}}^{\nu} : L^p(\mu) \to L^p(\lambda)\|.$$

The same holds for the other paraproducts Π_b^* and Γ_b.

The key here is that the weights are A_p weights, and the Haar functions are an unconditional basis for $L^p(w)$ if and only if $w \in A_p$ – see [16] for a first proof of this fact, or the helpful discussions on this in [13]. With this in hand, we note that letting the size of Q_0 in Π_{b,Q_0} go to infinity will sweep all the Haar functions in the "quadrant" of $\mathbb{R}^n$ in which Q_0 lives (since we are in the dyadic situation, there are 2^n "quadrants" of $\mathbb{R}^n$ which are non-overlapping – think of the four quadrants when visualizing $\mathbb{R}^2$ with the standard dyadic grid centered at 0.) So a choice of 2^n cubes, one for each quadrant, and adding up the limits yields the result.

In light of the bound for $\mathcal{A}_{\mathcal{S}}^{\nu}$ in Theorem 3.9, pick some value for Λ, say $\Lambda = 2$. Then we have the following.

Theorem 4.3. *Given Bloom weights $\mu, \lambda \in A_p^{\mathcal{D}}$, $\nu = \mu^{1/p}\lambda^{-1/p}$, for all $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$,*

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| \leq C(n,p)\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)}\,[\mu]_{A_p}^{1/(p-1)}\,[\lambda]_{A_p}.$$

The same holds for the other paraproducts Π_b^* and Γ_b.

Remark 4.4. The result actually follows immediately for Π_b^*, since

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| = \|\Pi_b^* : L^{p'}(\lambda') \to L^{p'}(\mu')\|$$

and

$$\nu' = (\lambda')^{1/p'}(\mu')^{-1/p'} = (\lambda^{-p'/p})^{1/p'}(\mu^{-p'/p})^{-1/p'} = \nu.$$

Remark 4.5. As discussed in the introduction, we do not know if this bound is sharp, but we can show that one particular instance of this inequality is sharp, namely when $\mu = w$ and $\lambda = w^{-1}$ for some $A_2^{\mathcal{D}}$ weight w, in which case the "intermediary" Bloom weight is also $\nu = w$:

$$\|\Pi_b : L^2(w) \to L^2(w^{-1})\| \lesssim \|b\|_{\mathrm{BMO}(w)^{\mathcal{D}}}\,[w]_{A_2}^2. \tag{4.1}$$

Remark 4.6. In terms of necessity, it was shown in [8] that boundedness of Π_b: $L^p(\mu) \to L^p(\lambda)$ does imply that $b \in \mathrm{BMO}^{\mathcal{D}}(\nu)$. An interesting question is: does this also imply that $\mu, \lambda \in A_p$? These questions were partly answered for commutators in [8], but since paraproducts are "less singular" than commutators, it is possible one can go beyond the assumption $\mu, \lambda \in A_p$.

4.1. Proof that the quadratic bound $[w]_{A_2}^2$ in (4.1) is sharp (via the one-weight linear A_2 bound for the dyadic square function)

The starting point is a simple observation: given a weight w on $\mathbb{R}^n$, the *weight itself belongs to* $\mathrm{BMO}(w)$, with

$$\|w\|_{\mathrm{BMO}(w)} \leq 2.$$

To see this, if Q is a cube,

$$\frac{1}{w(Q)}\int_Q |w(x)-\langle w\rangle_Q|\,dx \le \frac{1}{w(Q)}(w(Q)+w(Q)) = 2.$$

So we may look at the paraproducts with symbol w: in $\mathbb{R}$, these are

$$\Pi_w f = \sum_{I\in\mathcal{D}} (w,h_I)\,\langle f\rangle_I\, h_I \quad \text{and} \quad \Pi_w^* f = \sum_{I\in\mathcal{D}} (w,h_I)\,(f,h_I)\,\frac{\mathbb{1}_I}{|I|}.$$

If $w\in A_2^{\mathcal{D}}$, these are bounded:

$$\|\Pi_w : L^2(w)\to L^2(w^{-1})\| = \|\Pi_w^* : L^2(w)\to L^2(w^{-1})\| \lesssim \|w\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\,[w]_{A_2}^2 = 2\,[w]_{A_2}^2.$$

Recall the decomposition

$$fw = \Pi_w f + \Pi_w^* f + \Pi_f w,$$

and note that the map $f\mapsto fw$ is an isometry $L^2(w)\to L^2(w^{-1})$. So

$$\Pi_f w = \sum_{I\in\mathcal{D}} (f,h_I)\,\langle w\rangle_I\, h_I$$

is bounded $L^2(w)\to L^2(w^{-1})$:

$$\|\Pi_f w\|_{L^2(w^{-1})} \le \big(1+2\,\|\Pi_w : L^2(w)\to L^2(w^{-1})\|\big)\,\|f\|_{L^2(w)}.$$

Consider the dyadic square function

$$S_{\mathcal{D}} f := \Big(\sum_I (f,h_I)^2\,\frac{\mathbb{1}_I}{|I|}\Big)^{1/2},$$

and look at its $L^2(w)$-norm:

$$\begin{aligned}\|S_{\mathcal{D}} f\|_{L^2(w)}^2 &= \sum_{I\in\mathcal{D}} (f,h_I)^2\,\langle w\rangle_I = \Big(f, \sum_{I\in\mathcal{D}} (f,h_I)\,\langle w\rangle_I\, h_I\Big)\\ &= (f,\Pi_f w) \le \|\Pi_f w\|_{L^2(w^{-1})}\,\|f\|_{L^2(w)},\end{aligned}$$

so

$$\|S_{\mathcal{D}} f\|_{L^2(w)}^2 \le \big(1+2\,\|\Pi_w : L^2(w)\to L^2(w^{-1})\|\big)\,\|f\|_{L^2(w)}^2.$$

Since

$$\frac{1}{2} \le \frac{\|\Pi_w : L^2(w)\to L^2(w^{-1})\|}{\|w\|_{\mathrm{BMO}(w)}} \tag{4.2}$$

(we will show this in a moment) and

$$\frac{1}{2} \le \frac{1}{\|w\|_{\mathrm{BMO}(w)}},$$

we have further that

$$\begin{aligned}\|S_{\mathcal{D}} f\|_{L^2(w)}^2 &\\ &\le \|f\|_{L^2(w)}^2 \Big(2\,\frac{\|\Pi_w : L^2(w) \to L^2(w^{-1})\|}{\|w\|_{\mathrm{BMO}(w)}} + 4\,\frac{\|\Pi_w : L^2(w) \to L^2(w^{-1})\|}{\|w\|_{\mathrm{BMO}(w)}}\Big),\end{aligned}$$

which yields

$$\begin{aligned}\frac{\|S_{\mathcal{D}} f\|_{L^2(w)}}{\|f\|_{L^2(w)}} &\le \sqrt{6}\Big(\frac{\|\Pi_w : L^2(w) \to L^2(w^{-1})\|}{\|w\|_{\mathrm{BMO}(w)}}\Big)^{1/2}\\ &\le \sqrt{6}\sup_{b\in \mathrm{BMO}^{\mathcal{D}}(w)}\Big(\frac{\|\Pi_b : L^2(w) \to L^2(w^{-1})\|}{\|b\|_{\mathrm{BMO}(w)}}\Big)^{1/2}.\end{aligned}$$

Finally, the fact that

$$\sup_{b\in \mathrm{BMO}^{\mathcal{D}}(w)}\Big(\frac{\|\Pi_b : L^2(w) \to L^2(w^{-1})\|}{\|b\|_{\mathrm{BMO}(w)}}\Big) \ge \frac{1}{6}\,\|S_{\mathcal{D}} : L^2(w) \to L^2(w)\|^2 \simeq [w]_{A_2}^2$$

shows that any smaller bound in (4.1) would imply a bound for $\|S_{\mathcal{D}} : L^2(w) \to L^2(w)\|$ smaller than $[w]_{A_2}$, which is well known to be false.

Going back to (4.2), it is easy to show that

$$\mathbb{1}_Q(b - \langle b\rangle_Q) = \mathbb{1}_Q(\Pi_b \mathbb{1}_Q - \Pi_b^*\mathbb{1}_Q), \quad \forall Q \in \mathcal{D}.$$

Then

$$\begin{aligned}\frac{1}{w(Q)}\int_Q |b - \langle b\rangle_Q|\,dx &= \frac{1}{w(Q)}\int_Q |\Pi_b \mathbb{1}_Q - \Pi_b^*\mathbb{1}_Q|\,dx\\ &\le \frac{1}{w(Q)}\Big[\Big(\int_Q |\Pi_b \mathbb{1}_Q|^2\,dw^{-1}\Big)^{1/2} w(Q)^{1/2} + \Big(\int_Q |\Pi_b^*\mathbb{1}_Q|^2\,dw^{-1}\Big)^{1/2} w(Q)^{1/2}\Big]\\ &\le \frac{1}{w(Q)^{1/2}}\,2\,\|\Pi_b : L^2(w) \to L^2(w^{-1})\|\,\|\mathbb{1}_Q\|_{L^2(w)},\end{aligned}$$

which gives us

$$\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \le 2\,\|\Pi_b : L^2(w) \to L^2(w^{-1})\|, \quad \forall b \in \mathrm{BMO}^{\mathcal{D}}(w).$$

Now we proceed with the proof of Theorem 4.1, focusing on Π_b, with the other paraproducts following similarly.

4.2. Maximal truncation of paraproducts

Let $b \in \mathrm{BMO}_{\mathcal{D}}(\mathbb{R}^n)$. Define the maximal truncation of the paraproduct Π_b:

$$\overset{\triangleright}{\Pi}_b f(x) := \sup_{P\in\mathcal{D}}\Big|\sum_{Q\supsetneq P}(b, h_Q)\,\langle f\rangle_Q\,h_Q(x)\Big|.$$

We will need the following result, which may be found in Lemma 2.10 of [13].

Proposition 4.7. *Suppose $T: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded linear or sublinear operator. If T satisfies*

$$\operatorname{supp}(T h_Q) \subset Q, \quad \forall Q \in \mathcal{D},$$

then T is of weak $(1,1)$ type, with

$$|\{x : |Tf(x)| > \alpha\}| \leq C_n B \frac{1}{\alpha} \|f\|_1,$$

where C_n is a dimensional constant and $B := \|T\|_{L^2 \to L^2}$.

Now we prove some properties of $\overset{\triangleright}{\Pi}_b$.

Proposition 4.8. *The maximal truncation defined above satisfies the following.*

(i) $\overset{\triangleright}{\Pi}_b$ *dominates* Π_b:

$$|\Pi_b f(x)| \leq \overset{\triangleright}{\Pi}_b f(x), \quad \forall x \in \mathbb{R}^n.$$

(ii) $\overset{\triangleright}{\Pi}_b$ *is dominated by* $M^{\mathcal{D}} \Pi_b$:

$$\overset{\triangleright}{\Pi}_b f(x) \leq M^{\mathcal{D}}(\Pi_b f)(x), \quad \forall x \in \mathbb{R}^n.$$

(iii) $\overset{\triangleright}{\Pi}_b$ *is strong* $(2,2)$:

$$\|\overset{\triangleright}{\Pi}_b f\|_{2\to 2} \lesssim \|b\|_{\mathrm{BMO}^{\mathcal{D}}} \|f\|_{L^2}.$$

(iv) $\overset{\triangleright}{\Pi}_b$ *is weak* $(1,1)$:

$$|\{x \in \mathbb{R}^n : \overset{\triangleright}{\Pi}_b f(x) > \alpha\}| \leq \frac{C_n}{\alpha} \|b\|_{\mathrm{BMO}^{\mathcal{D}}} \|f\|_1.$$

Proof. (i) Let $x \in \mathbb{R}^n$. Then

$$\Pi_b f(x) = \sum_{Q \in \mathcal{D}} (b, h_Q) \langle f \rangle_Q h_Q(x) = \sum_{k \in \mathbb{Z}} (b, h_{Q_k}) \langle f \rangle_{Q_k} h_{Q_k}(x),$$

where for every $k \in \mathbb{Z}$, Q_k is the unique cube in $\mathcal{D}$ with side length 2^k that contains x. Fix $m \in \mathbb{Z}$:

$$\Big| \sum_{k>m} (b, h_{Q_k}) \langle f \rangle_{Q_k} h_{Q_k}(x) \Big| = \Big| \sum_{Q \supsetneq Q_m} (b, h_Q) \langle f \rangle_Q h_Q(x) \Big| \leq \overset{\triangleright}{\Pi}_b f(x).$$

Taking $m \to -\infty$ finishes the proof.

(ii) Let $P \in \mathcal{D}$ and define

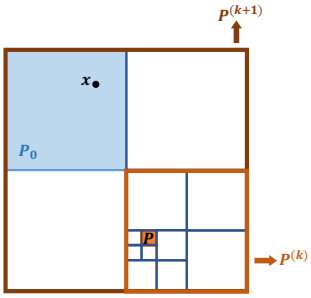
$$F_P(x) := \sum_{Q \supsetneq P} (b, h_Q) \langle f \rangle_Q h_Q(x).$$

If $x \in P$, then $|F_P(x)| = |\langle \Pi_b f\rangle_P|\mathbb{1}_P(x)$, so

$$|F_P(x)| \le \langle|\Pi_b f|\rangle_P \mathbb{1}_P(x) \le M^{\mathcal{D}}\Pi_b f(x).$$

If $x \notin P$, then there is a unique $k \ge 0$ such that

$$x \in P^{(k+1)} \setminus P^{(k)}.$$



So, there is a unique

$$P_0 \in (P^{(k+1)})_{(1)}, \quad P_0 \ne P^{(k)},$$

such that $x \in P_0$. Then

$$F_P(x) = (b, h_{P^{(k+1)}})\langle f\rangle_{P^{(k+1)}} h_{P^{(k+1)}}(x) + \sum_{Q \supsetneq P^{(k+1)}} (b, h_Q)\langle f\rangle_Q \underbrace{h_Q(P^{(k+1)})}_{=h_Q(P_0)}$$

$$= \Big[\sum_{Q \supsetneq P_0} (b, h_Q)\langle f\rangle_Q h_Q(P_0)\Big]\mathbb{1}_{P_0}(x) = \langle \Pi_b f\rangle_{P_0}\mathbb{1}_{P_0}(x),$$

so once again $|F_P(x)| \le M^{\mathcal{D}}\Pi_b f(x)$. This therefore holds for all $x \in \mathbb{R}^n$ and all $P \in \mathcal{D}$, which proves (ii).

(iii) This follows immediately from (ii) and the well-known bound for Π_b in the unweighted case:

$$\|\overset{\triangleright}{\Pi}_b f\|_{L^2} \le \|M^{\mathcal{D}}\Pi_b f\|_{L^2} \lesssim \|\Pi_b f\|_{L^2} \lesssim \|b\|_{\mathrm{BMO}^{\mathcal{D}}}\|f\|_{L^2}.$$

(iv) Once we verify $\mathrm{supp}(\overset{\triangleright}{\Pi}_b h_Q) \subset Q$ for all $Q \in \mathcal{D}$, we use (iii) and Proposition 4.7 to conclude (iv):

$$\overset{\triangleright}{\Pi}_b h_Q(x) = \sup_{P\in\mathcal{D}}\Big|\sum_{R\supsetneq P}(b, h_R)\langle h_Q\rangle_R h_R(x)\Big| = \sup_{P\subsetneq Q}\Big|\sum_{R\supsetneq P, R\subsetneq Q}(b, h_R)h_Q(R)h_R(x)\Big|,$$

which is clearly 0 if $x \notin Q$. ∎

4.3. Proof of Theorem 4.1

I. *The BMO decomposition.*

We make use of the following modification to the Calderón–Zygmund decomposition used in [6] to essentially reduce a weighted BMO function to a regular BMO function. Given a weight w on $\mathbb{R}^n$, a function $b \in \mathrm{BMO}^{\mathcal{D}}(w)$, a fixed dyadic cube $Q_0 \in \mathcal{D}$, and $\varepsilon \in (0, 1)$, let the collection

$$\mathcal{E} := \{\text{maximal subcubes } R \subset Q_0 \text{ such that } \langle w\rangle_R > \tfrac{2}{\varepsilon}\langle w\rangle_{Q_0}\},$$

and put

$$E := \bigcup_{R\in\mathcal{E}} R.$$

This is the collection from the usual CZ-decomposition of w, restricted to Q_0, so we have

$$|E| = \sum_{R\in\mathcal{E}} |R| < \frac{\varepsilon}{2}\,|Q_0|.$$

Now consider the function

$$a(x) := \sum_{Q\subset Q_0, Q\not\subset E} (b, h_Q)\, h_Q(x).$$

We claim that a is in *unweighted BMO*, with

$$\|a\|_{\mathrm{BMO}^{\mathcal{D}}} \le \frac{4}{\varepsilon}\,\langle w\rangle_{Q_0}\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}. \tag{4.3}$$

To see this, we look at the term

$$\mathbb{1}_Q(x)\,(a(x) - \langle a\rangle_Q) = \sum_{P\subset Q} (a, h_P)\, h_P$$

for $Q \in \mathcal{D}$. Since a is supported in Q_0, we only need to look at cubes $Q \in \mathcal{D}$ which intersect Q_0.

Case 1. $Q \subset Q_0$.

If $Q \subset E$, the term is 0. So suppose $Q \not\subset E$, and write

$$\begin{aligned}
\sum_{P\subset Q} (a, h_P)\, h_P &= \sum_{\substack{P\subset Q\\ P\not\subset E}} (b, h_P)\, h_P = \sum_{P\subset Q} (b, h_P)\, h_P - \sum_{\substack{P\subset Q\\ P\subset E}} (b, h_P)\, h_P \\
&= \sum_{P\subset Q} (b, h_P)\, h_P - \sum_{R\in\mathcal{E}, R\subset Q} \sum_{P\subset R} (b, h_P)\, h_P \\
&= \mathbb{1}_Q(b - \langle b\rangle_Q) - \sum_{R\in\mathcal{E}, R\subset Q} \mathbb{1}_R(b - \langle b\rangle_R).
\end{aligned}$$

Then

$$\begin{aligned}
\frac{1}{|Q|}\int_Q |a(x) - \langle a\rangle_Q|\,dx &\le \frac{1}{|Q|}\int_Q |b(x) - \langle b\rangle_Q|\,dx + \frac{1}{|Q|}\sum_{R\in\mathcal{E}, R\subset Q}\int_R |b(x) - \langle b\rangle_R|\,dx \\
&\le \langle w\rangle_Q\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)} + \frac{1}{|Q|}\Big(\sum_{R\in\mathcal{E}, R\subset Q} w(R)\Big)\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)} \\
&\le 2\,\langle w\rangle_Q\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)} \le \frac{4}{\varepsilon}\,\langle w\rangle_{Q_0}\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)},
\end{aligned}$$

where we used that $\mathcal{E}$ is a disjoint collection, and that $\langle w\rangle_Q \le \frac{2}{\varepsilon}\langle w\rangle_{Q_0}$ since $Q \not\subset E$.

Case 2. $Q \supsetneq Q_0$.

Similarly,

$$\sum_{P\subset Q} (a, h_P)\, h_P = \sum_{\substack{P\subset Q_0\\ P\not\subset E}} (b, h_P)\, h_P = \mathbb{1}_{Q_0}(b - \langle b\rangle_{Q_0}) - \sum_{R\in\mathcal{E}} \mathbb{1}_R(b - \langle b\rangle_R),$$

so

$$\begin{aligned}\frac{1}{|Q|}\int_Q |a(x)-\langle a\rangle_Q|\,dx &\le \frac{1}{|Q|}\int_{Q_0}|b(x)-\langle b\rangle_{Q_0}|\,dx+\frac{1}{|Q|}\sum_{R\in\mathcal{E}}\int_R |b(x)-\langle b\rangle_R|\,dx\\ &\le \frac{|Q_0|}{|Q|}\,\langle w\rangle_{Q_0}\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)}+\frac{1}{|Q|}\sum_{R\in\mathcal{E}} w(R)\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)}\\ &\le 2\,\langle w\rangle_{Q_0}\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)} \le \frac{4}{\varepsilon}\,\langle w\rangle_{Q_0}\,\|b\|_{\mathrm{BMO}_{\mathcal{D}}(w)}.\end{aligned}$$

Remark 4.9. This BMO decomposition is similar to that of [6], where the auxiliary BMO function was defined as

$$a'(x) := \mathbb{1}_{Q_0}(x)\,b(x)-\sum_{R\in\mathcal{E}}(b(x)-\langle b\rangle_R)\,\mathbb{1}_R(x),$$

the classical format of the "good function" in the Calderón–Zygmund decomposition. This function is also in unweighted BMO, and has the structure

$$a'(x)=\begin{cases} b(x), & \text{if } x\in Q_0\setminus E,\\ \langle b\rangle_R, & \text{if } x\in R,\ R\in\mathcal{E}.\end{cases}$$

In relation to our function a, note that

$$a'(x)=\mathbb{1}_{Q_0}(x)\,\langle b\rangle_{Q_0}+a(x).$$

The main advantage in working with a here instead is that $\langle a\rangle_{Q_0}=0$, and the paraproduct $\Pi_a f$ has a simple structure:

$$\Pi_a f=\sum_{\substack{Q\subset Q_0\\ Q\not\subset E}}(b,h_Q)\,\langle f\rangle_Q\,h_Q.$$

The main advantage of a' was that

$$\forall\, Q\in\mathcal{D}(Q_0),\ Q\not\subset E,\quad \langle a'\rangle_Q=\langle b\rangle_Q\quad\text{and}\quad (a',h_Q)=(b,h_Q),$$

so whenever dealing with a cube $Q\not\subset E$, we can replace any average or Haar coefficient of b – the function in weighted BMO – with the average or Haar coefficient of a' – the function in *unweighted* BMO. With our function a, the Haar coefficient property is preserved and it is all we need here:

$$\forall\, Q\in\mathcal{D}(Q_0),\ Q\not\subset E,\quad \langle a\rangle_Q=\langle b\rangle_Q-\langle b\rangle_{Q_0}\quad\text{and}\quad (a,h_Q)=(b,h_Q).$$

This has many advantages, since any usage of inequalities involving a or a' will not add any extra A_p characteristics. For instance, we can use the well-known bound for Haar coefficients of BMO functions (resulting from applying the John–Nirenberg theorem to replace the L^1 norm in the BMO definition with the L^2 norm):

$$|(a,h_Q)|\lesssim\sqrt{|Q|}\,\|a\|_{\mathrm{BMO}^{\mathcal{D}}}.$$

It also allows us to use the results on $\overset{\triangleright}{\Pi}_a f$ from the previous section.

II. *Use the properties of the maximal truncation of* unweighted *BMO paraproducts.*

We claim that there exists a constant C_0, depending on the dimension n and on ε, such that the set:

$$F := \left\{x \in Q_0 : \overset{\triangleright}{\Pi}_a f(x) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}\right\} \bigcup \left\{x \in Q_0 : M_{Q_0}^{\mathcal{D}} f(x) > C_0 \langle |f| \rangle_{Q_0}\right\}$$

satisfies

$$|F| < \frac{\varepsilon}{2} |Q_0|,$$

where $M_{Q_0}^{\mathcal{D}}$ denotes the dyadic maximal function restricted to Q_0, i.e.,

$$M_{Q_0}^{\mathcal{D}} f(x) = \sup_{Q \subset Q_0} \langle |f| \rangle_Q \mathbb{1}_Q(x).$$

Let then the collection

$$\mathcal{F} := \{\text{maximal subcubes of } Q_0 \text{ contained in F}\},$$

and note that $F = \bigcup_{Q \in \mathcal{F}} Q$, a disjoint union.

First use the well-known weak $(1,1)$ inequality for the dyadic maximal function:

$$|\{x \in \mathbb{R}^n : M^{\mathcal{D}} \varphi(x) > \alpha\}| \leq \frac{C_1(n)}{\alpha} \|\varphi\|_1,$$

applied to $\varphi = f \mathbb{1}_{Q_0}$. For all $x \in Q_0$, $M^{\mathcal{D}}(f \mathbb{1}_{Q_0})(x) = M_{Q_0}^{\mathcal{D}} f(x)$, so

$$|\{x \in Q_0 : M_{Q_0}^{\mathcal{D}} f(x) > C_0 \langle |f| \rangle_{Q_0}\}| \leq \frac{C_1}{C_0} |Q_0|.$$

Since $a \in \mathrm{BMO}^{\mathcal{D}}$, we can apply the weak $(1,1)$ inequality for $\overset{\triangleright}{\Pi}_a$ according to Proposition 4.8 (iv):

$$|\{x \in \mathbb{R}^n : \overset{\triangleright}{\Pi}_a \varphi(x) > \alpha\}| \leq \frac{C_2(n)}{\alpha} \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|\varphi\|_1,$$

and let again $\varphi = f \mathbb{1}_{Q_0}$. By the definition of a, in this case, $\Pi_a f$ sums only over $Q \subset Q_0$, so regardless of x we have $\Pi_a \varphi = \Pi_a(f)$. The same holds for $\overset{\triangleright}{\Pi}_a$:

$$\overset{\triangleright}{\Pi}_a \varphi(x) = \sup_{P \in \mathcal{D}} \left| \sum_{Q \supsetneq P} (a, h_Q) \langle \varphi \rangle_Q h_Q(x) \right| = \sup_{P \subsetneq Q_0} \left| \sum_{Q \supsetneq P, Q \subset Q_0, Q \not\subset E} (b, h_Q) \langle \varphi \rangle_Q h_Q(x) \right|,$$

so

$$\begin{aligned}
&|\{x \in Q_0 : \overset{\triangleright}{\Pi}_a f(x) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}\}| \\
&\qquad = |\{x \in Q_0 : \overset{\triangleright}{\Pi}_a (f \mathbb{1}_{Q_0})(x) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}\}| \\
&\qquad \leq |\{x \in \mathbb{R}^n : \overset{\triangleright}{\Pi}_a (f \mathbb{1}_{Q_0})(x) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}\}| \\
&\qquad \leq \frac{C_2}{C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}} \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|f \mathbb{1}_{Q_0}\|_1 = \frac{C_2}{C_0} |Q_0|.
\end{aligned}$$

Then, as we wished,

$$|F| \le \frac{C_1 + C_2}{C_0} |Q_0| < \frac{\varepsilon}{2} |Q_0|$$

if we choose C_0 large enough:

$$C_0 = \frac{C(n)}{\varepsilon}.$$

Join the collections $\mathcal{E}$ and $\mathcal{F}$ into

$$\mathcal{G} := \{\text{maximal subcubes of } Q_0 \text{ contained in } E \cup F\},$$

which then satisfies

$$\Big| \bigcup_{R \in \mathcal{G}} R \Big| < \varepsilon |Q_0|. \tag{4.4}$$

Note that while it is not true that $\mathcal{G}$ is the union on $\mathcal{E}$ and $\mathcal{F}$, it is true that $G = E \cup F$, where $G = \bigcup_{Q \in \mathcal{G}} Q$.

We show that:

$$\mathbb{1}_{Q_0}(x) |\Pi_{b,Q_0} f(x)| \le 2C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0} \mathbb{1}_{Q_0}(x) + \sum_{R \in \mathcal{G}} \mathbb{1}_R(x) |\Pi_{b,R} f(x)|. \tag{4.5}$$

Since $\|a\|_{\mathrm{BMO}^{\mathcal{D}}} \le \frac{4}{\varepsilon} \langle w \rangle_{Q_0} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}$, this yields

$$\mathbb{1}_{Q_0}(x) |\Pi_{b,Q_0} f(x)| \lesssim \frac{C_0}{\varepsilon} \langle w \rangle_{Q_0} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \langle |f| \rangle_{Q_0} \mathbb{1}_{Q_0}(x) + \sum_{R \in \mathcal{G}} \mathbb{1}_R(x) |\Pi_{b,R} f(x)|.$$

Once we have this, we recurse on the terms of the second sum, and repeat the argument: for each $R \in \mathcal{G}$, construct a disjoint collection $\{R'\} \subset R$ satisfying $|\cup R'| < \varepsilon |R|$ and

$$\mathbb{1}_R |\Pi_{b,R} f(x)| \lesssim \frac{C_0}{\varepsilon} \langle w \rangle_R \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \langle |f| \rangle_R \mathbb{1}_R(x) + \sum_{R'} |\Pi_{b,R'} f(x)|.$$

So we construct the collection $\mathcal{S}(Q_0)$ recursively, starting with Q_0 as its first element, its $\mathcal{S}$-children are $\mathcal{G}$, and so on. We have

$$\mathbb{1}_{Q_0}(x) |\Pi_{b,Q_0} f(x)| \lesssim \frac{C_0}{\varepsilon} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \underbrace{\sum_{Q \in \mathcal{S}(Q_0)} \langle w \rangle_Q \langle |f| \rangle_Q \mathbb{1}_Q(x)}_{= \mathcal{A}^w_{\mathcal{S}(Q_0)} |f|(x)}.$$

Recall that $C_0 \sim C(n)/\varepsilon$, thus

$$|\Pi_{b,Q_0} f(x)| \lesssim \frac{C(n)}{\varepsilon^2} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \mathcal{A}^w_{\mathcal{S}(Q_0)} |f|(x).$$

The collection $\mathcal{S}(Q_0)$ satisfies the $\mathcal{S}$-children definition of sparse collections:

$$\sum_{P \in \mathrm{ch}_{\mathcal{S}}(Q)} |P| < \varepsilon |Q|, \quad \forall Q \in \mathcal{S}(Q_0),$$

so $\mathcal{S}(Q_0)$ is $\frac{1}{1-\varepsilon}$-Carleson. Then, choosing $\varepsilon = (\Lambda - 1)/\Lambda$, we have the desired sparse collection with Carleson constant Λ such that

$$|\Pi_{b,Q_0} f(x)| \leq C(n)\Big(\frac{\Lambda}{\Lambda - 1}\Big)^2 \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \mathcal{A}^w_{\mathcal{S}(Q_0)} |f|(x).$$

III. *Proof of* (4.5).

We start by noting that

$$\begin{aligned}\Pi_{b,Q_0} f(x) &= \sum_{P \subset Q_0} (b, h_P) \langle f \rangle_P h_P(x) \\ &= \underbrace{\sum_{P \subset Q_0, P \not\subset E} (b, h_P) \langle f \rangle_P h_P(x)}_{\Pi_a f(x)} + \sum_{R \in \mathcal{E}} \underbrace{\sum_{P \subset R} (b, h_P) \langle f \rangle_P h_P(x)}_{\Pi_{b,R} f(x)},\end{aligned}$$

so we may decompose $\Pi_{b,Q_0} f$ as

$$\mathbb{1}_{Q_0}(x)\, \Pi_{b,Q_0} f(x) = \Pi_a f(x) + \sum_{R \in \mathcal{E}} \Pi_{b,R} f(x).$$

Now, we have to account for the relationship to the collection $\mathcal{F}$ and its union F. For the different cases below, it may be helpful to consult Figure 1 to visualize the position of x.

Case 1: $x \notin F$.

In this case, $\overset{\triangleright}{\Pi}_a f(x) \leq C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}$, and since $\overset{\triangleright}{\Pi}_a$ dominates Π_a,

$$|\Pi_a f(x)| \leq \overset{\triangleright}{\Pi}_a f(x) \leq C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0},$$

so we have, for $x \notin F$,

$$|\Pi_{b,Q_0} f(x)| \leq C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0} + \Big| \sum_{R \in \mathcal{E}} \Pi_{b,R} f(x) \Big|.$$

– *Case* 1a. If $x \in E$, there is a *unique* $R_0 \in \mathcal{E}$ such that $x \in R_0$. But then $R_0 \in \mathcal{G}$: say $R_0 \notin \mathcal{G}$; since $R_0 \subset E$, it must have been absorbed by a larger $R \supsetneq R_0$, $R \in \mathcal{F}$. Then $R_0 \subset R \subset F$, which contradicts $x \notin F$. So then

$$\sum_{R \in \mathcal{E}} \Pi_{b,R} f(x) = \Pi_{b,R_0} f(x),$$

and for $x \in E \setminus F$,

$$|\Pi_{b,Q_0} f(x)| \leq C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0} + |\Pi_{b,R_0} f(x)|, \quad R_0 \in \mathcal{G},$$

which gives (4.5) in this case.

– *Case* 1b. If $x \notin E$, then the second part of the sum is 0 and we are done, having simply $|\Pi_{b,Q_0} f(x)| \leq C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}$.

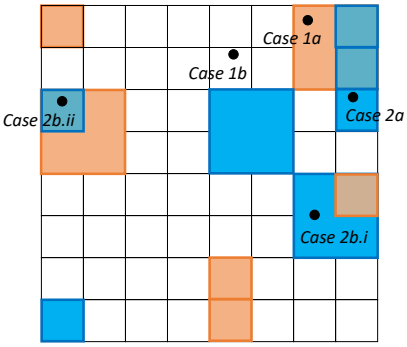


Figure 1. Orange cubes: $\mathcal{E}$; blue cubes: $\mathcal{F}$.

Case 2: $x \in F$.

Then there is a *unique* $P \in \mathcal{F}$ such that $x \in P$. Look first at the term $\Pi_a f(x) = \sum_{Q \subset Q_0} (a, h_Q) \langle f \rangle_Q h_Q(x)$. Since $x \in P$, this can be expressed as

$$\Pi_a f(x) = \sum_{Q \supset \hat{P}} (a, h_Q) \langle f \rangle_Q h_Q(x) + \sum_{Q \subset P} (a, h_Q) \langle f \rangle_Q h_Q(x),$$

where $\hat{P}$ denotes the dyadic parent of P. All sums here are understood to be over dyadic cubes contained in Q_0, since all Haar coefficients of a are 0 elsewhere. We split the first term into two:

$$\Big| \sum_{Q \supset \hat{P}} (a, h_Q) \langle f \rangle_Q h_Q(x) \Big| \le \underbrace{\Big| \sum_{Q \supsetneq \hat{P}} (a, h_Q) \langle f \rangle_Q h_Q(x) \Big|}_{=:A(x)} + \underbrace{|(a, h_{\hat{P}}) \langle f \rangle_{\hat{P}} h_{\hat{P}}(x)|}_{=:B}.$$

The term A is constant on $\hat{P}$, so if $A(x) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}$, then

$$A(y) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0} \quad \text{for all } y \in \hat{P}.$$

This would force $\overset{\triangleright}{\Pi}_a f(y) > C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}$ for all $y \in \hat{P}$, so $\hat{P} \subset F$ – but this contradicts maximality of P in $\mathcal{F}$. Therefore,

$$A \le C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0}.$$

Let us now look at the term B. If $\hat{P} \subset E$, then $B = 0$. Otherwise, since $|(a, h_Q)| \lesssim \sqrt{|Q|}\, \|a\|_{\mathrm{BMO}^{\mathcal{D}}}$,

$$B \le \sqrt{|\hat{P}|}\, \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{\hat{P}} \frac{1}{\sqrt{|\hat{P}|}} = \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{\hat{P}},$$

but $\langle |f| \rangle_{\hat{P}} \leq C_0 \langle |f| \rangle_{Q_0}$ – otherwise, $M^{\mathcal{D}}_{Q_0} f(y) > C_0 \langle |f| \rangle_{Q_0}$ for all $y \in \hat{P}$, which would force $\hat{P} \subset F$, again contradicting maximality of P in $\mathcal{F}$.

So

$$\Big| \sum_{Q \supset \hat{P}} (a, h_Q) \langle f \rangle_Q h_Q(x) \Big| \lesssim C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0},$$

giving us that

$$|\Pi_{b,Q_0} f(x)| \lesssim C_0 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \langle |f| \rangle_{Q_0} + |C|,$$

where the term C is defined as

$$C := \sum_{Q \subset P, Q \not\subset E} (b, h_Q) \langle f \rangle_Q h_Q(x) + \sum_{R \in \mathcal{E}} \Pi_{b,R} f(x).$$

We claim that

$$C = \Pi_{b,R_0} f(x),$$

where R_0 is the unique element of $\mathcal{G}$ such that $x \in R_0$:

Case 2a. If $P \cap E = \emptyset$, then $R_0 = P$ and $C = \Pi_{b,P} f(x) = \Pi_{b,R_0} f(x)$ (the first term is $\Pi_{b,P}$ and the second term is 0).

Case 2b. If $P \cap E \neq \emptyset$:

– *Case* 2b.i. If P contains some elements of $\mathcal{E}$, then again $R_0 = P$ and we can "fill in the blanks" in the first term with the $\Pi_{b,R}$'s from the second term:

$$C = \Big[\sum_{Q \subset P, Q \not\subset E} (b, h_Q) \langle f \rangle_Q h_Q(x) + \sum_{R \in \mathcal{E}, R \subset P} \Pi_{b,R} f(x) \Big] + \underbrace{\sum_{R \in \mathcal{E}, R \not\subset P} \Pi_{b,R} f(x)}_{=0}$$

$$= \Pi_{b,P} f(x) = \Pi_{b,R_0} f(x).$$

– *Case* 2b.ii. If $P \subset S_0$ for some $S_0 \in \mathcal{E}$, then $R_0 = S_0$ and the first term in C is 0 (because $P \subset E$), and the second term is $\sum_{R \in \mathcal{E}} \Pi_{b,R} f(x) = \Pi_{b,S_0} f(x) = \Pi_{b,R_0} f(x)$. This concludes the proof. ∎

Remark 4.10. One can also use Theorem 4.1 to obtain a full $\mathbb{R}^n$ domination, losing the requirement for no infinitely increasing chains. Say f is such that $\mathrm{supp}(f) \subset Q_0$ for some $Q_0 \in \mathcal{D}$ (or, for general compactly supported functions, $\mathrm{supp}(f)$ is contained in at most 2^n disjoint $Q_k \in \mathcal{D}$). Then

$$\Pi_b f(x) = \Pi_{b,Q_0} f(x) + \Big(\sum_{Q \supsetneq Q_0} (b, h_Q) \frac{1}{|Q|} h_Q(x) \Big) \int_{Q_0} f.$$

Note that, as an application of the modified CZ-decomposition used in Part I of the proof above, one can obtain

$$|(b, h_Q)| \lesssim \sqrt{|Q|} \langle w \rangle_Q \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}, \quad \forall Q \in \mathcal{D}, b \in \mathrm{BMO}^{\mathcal{D}}(w).$$

To see this, let $Q \in \mathcal{D}$ and apply the decomposition to b over Q (so here Q plays the role of Q_0):

$$\mathcal{E} := \{\text{maximal subcubes } R \subset Q_0 \text{ such that } \langle w\rangle_R > 2\langle w\rangle_Q\}; \quad E := \bigcup\nolimits_{R\in\mathcal{E}} R,$$
$$a := \sum_{P\subset Q; P\not\subset E} (b, h_P)h_P \in \mathrm{BMO}^{\mathcal{D}} \quad \text{with} \quad \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \le 4\langle w\rangle_Q \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}.$$

Since Q itself is not selected for $\mathcal{E}$, $Q \not\subset E$, so $(a, h_Q) = (b, h_Q)$. Finally,

$$|(b, h_Q)| = |(a, h_Q)| \lesssim \sqrt{|Q|}\,\|a\|_{\mathrm{BMO}^{\mathcal{D}}} \le \sqrt{|Q|}\,4\langle w\rangle_Q \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}.$$

Returning to $\Pi_b f$, suppose first that $x \notin Q_0$. Then there is a unique $k \ge 1$ such that $x \in Q_0^{(k)} \setminus Q_0^{(k-1)}$, and

$$\Pi_b f(x) = \Big(\sum_{Q\supset Q_0^{(k)}} (b, h_Q)\,\frac{1}{|Q|}\,h_Q(x)\Big)\int_{Q_0} f.$$

Then, for $x \notin Q_0$,

$$\begin{aligned}|\Pi_b f(x)| &\le \sum_{Q\supset Q_0^{(k)}} |(b, h_Q)|\,\frac{1}{|Q|}\,\frac{1}{\sqrt{|Q|}}\Big(\int_{Q_0}|f|\Big)\\ &\lesssim \sum_{Q\supset Q_0^{(k)}} \langle w\rangle_Q \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\,\frac{1}{|Q|}\int_{Q_0}|f| = \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \sum_{Q\supset Q_0^{(k)}} \langle w\rangle_Q \langle |f|\rangle_Q.\end{aligned}$$

If, on the other hand, $x \in Q_0$,

$$\Pi_b f(x) = \Pi_{b,Q_0} f(x) + \sum_{Q\supsetneq Q_0} (b, h_Q)\,\frac{1}{|Q|}\,h_Q(Q_0)\int_{Q_0} f,$$

so

$$|\Pi_b f(x)| \lesssim |\Pi_{b,Q_0} f(x)| + \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \sum_{Q\supsetneq Q_0} \langle w\rangle_Q \langle |f|\rangle_Q.$$

By Theorem 4.1, there is a Λ-Carleson sparse collection $\mathcal{S}(Q_0)$ such that

$$|\Pi_{b,Q_0} f(x)| \le C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^2 \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\,\mathcal{A}^w_{\mathcal{S}(Q_0)}|f|(x).$$

So form a sparse collection $\mathcal{S}$ as follows:

$$\mathcal{S} := \mathcal{S}(Q_0) \bigcup \bigcup_{k=1}^{\infty} Q_0^{(k)},$$

with $Q_0^{(k-1)}$ being the only $\mathcal{S}$-child of $Q_0^{(k)}$ for all $k \ge 1$. It is easy to see that $\mathcal{S}$ is $(\Lambda+1)$-Carleson. Moreover, the associated sparse operator

$$\mathcal{A}^w_{\mathcal{S}} f = \mathcal{A}^w_{\mathcal{S}(Q_0)} f + \sum_{Q\supsetneq Q_0} \langle w\rangle_Q \langle f\rangle_Q \mathbb{1}_Q$$

appears exactly in the previous inequalities, which can be expressed as:

$$x \notin Q_0: \quad |\Pi_b f(x)| \lesssim \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \mathcal{A}_{\mathcal{S}}^{w} |f|(x);$$
$$x \in Q_0: \quad |\Pi_b f(x)| \lesssim C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^2 \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \mathcal{A}_{\mathcal{S}}^{w} |f|(x).$$

So indeed

$$|\Pi_b f(x)| \lesssim \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \mathcal{A}_{\mathcal{S}}^{w} |f|(x), \quad \forall x \in \mathbb{R}^n,$$

for all compactly supported f.

Remark 4.11. If we let $f \equiv 1$ in $\Pi_{b,Q_0} f$, we have

$$\Pi_{b,Q_0} 1(x) = \sum_{Q \subset Q_0} (b, h_Q)\, h_Q(x) = (b(x) - \langle b \rangle_{Q_0}) \mathbb{1}_{Q_0}(x).$$

So, applying Theorem 4.1 to the function $f \equiv 1$ essentially gives us that local mean oscillations of functions in $\mathrm{BMO}^{\mathcal{D}}(w)$ can be dominated by one of the sparse BMO functions in Section 3.2:

Corollary 4.12. *There is a dimensional constant $C(n)$ such that, for all $\Lambda > 1$, weights w on $\mathbb{R}^n$, $b \in \mathrm{BMO}^{\mathcal{D}}(w)$ and $Q_0 \in \mathcal{D}$, there is a Λ-Carleson sparse collection $\mathcal{S}(Q_0) \subset \mathcal{D}(Q_0)$ such that*

$$\begin{aligned} |(b(x) - \langle b \rangle_{Q_0}) \mathbb{1}_{Q_0}(x)| &\le C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^2 \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)} \sum_{Q \in \mathcal{S}(Q_0)} \langle w \rangle_Q \mathbb{1}_Q(x) \\ &= C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^2 \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\, b^{w}_{\mathcal{S}(Q_0)}(x). \end{aligned}$$

A. Proof of Theorem 3.3

Recall that we are given $\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n)$ and the associated function

$$b_{\mathcal{S}} := \sum_{Q \in \mathcal{S}} \mathbb{1}_Q,$$

and we wish to show that

$$\|b_{\mathcal{S}}\|_{\mathrm{BMO}^{\mathcal{D}}} \le \Lambda,$$

where Λ is the Carleson constant of $\mathcal{S}$.

Proof. Let $Q_0 \in \mathcal{D}$ be fixed. We wish to estimate

$$\frac{1}{|Q_0|} \int_{Q_0} |b - \langle b \rangle_{Q_0}|\, dx,$$

and recall that

$$(b_{\mathcal{S}} - \langle b_{\mathcal{S}} \rangle_{Q_0}) \mathbb{1}_{Q_0} = \sum_{Q \in \mathcal{S}, Q \subsetneq Q_0} \mathbb{1}_Q - (\tau_{\mathcal{S}})_{Q_0} \mathbb{1}_{Q_0},$$

where

$$(\tau_{\mathcal{S}})_P := \frac{1}{|P|} \sum_{Q\in\mathcal{S}, Q\subsetneq P} |Q| \le \Lambda, \quad \forall P \in \mathcal{D}.$$

In fact,

$$\text{if } P \in \mathcal{S}, \text{ then } (\tau_{\mathcal{S}})_P \le \Lambda - 1.$$

With $Q_0 \in \mathcal{D}$ fixed, here we are only looking at

$$\mathcal{S}(Q_0) := \{Q \in \mathcal{S} : Q \subset Q_0\}.$$

We define the collections as sets:

$$\mathcal{S}_1 := \mathrm{ch}_{\mathcal{S}}(Q_0) \text{ (the } \mathcal{S}\text{-children of } Q_0), \quad \text{and} \quad S_1 := \bigcup_{Q_1\in\mathcal{S}_1} Q_1;$$

$$\mathcal{S}_2 := \{Q_2 \in \mathrm{ch}_{\mathcal{S}}(Q_1) : Q_1 \in \mathcal{S}_1\}, \qquad \text{and} \quad S_2 := \bigcup_{Q_2\in\mathcal{S}_2} Q_2,$$

so $\mathcal{S}_2$ are the "$\mathcal{S}$-grandchildren" of Q_0, the second generation of $\mathcal{S}$-cubes in Q_0. Generally,

$$\mathcal{S}_k := \{Q_k \in \mathrm{ch}_{\mathcal{S}}(Q_{k-1}) : Q_{k-1} \in \mathcal{S}_{k-1}\} \quad \text{and} \quad S_k := \bigcup_{Q_k\in\mathcal{S}_k} Q_k.$$

Note that:

- Each S_k is a disjoint union of $Q_k \in \mathcal{S}_k$, as each $\mathcal{S}_k$ is a pairwise disjoint collection.
- The sets S_k satisfy $Q_0 \supset S_1 \supset S_2 \supset \cdots$
- Moreover,

$$\Big|\bigcap_{k=1}^{\infty} S_k\Big| = 0,$$

since $\bigcap_{k=1}^{\infty} S_k$ is exactly the set of all x contained in infinitely many elements of $\mathcal{S}(Q_0)$. We can also see this directly, as

$$\sum_{k=1}^{\infty} |S_k| \le \Lambda |Q_0|,$$

and the series converges.

For ease of notation, denote for now

$$\theta := (\tau_{\mathcal{S}})_{Q_0} = \frac{1}{|Q_0|} \sum_{Q\in\mathcal{S}, Q\subsetneq Q_0} |Q| \le \Lambda.$$

We have:

$$\begin{aligned}\frac{1}{|Q_0|}\int_{Q_0} |b_{\mathcal{S}} - \langle b_{\mathcal{S}}\rangle_{Q_0}|\,dx &= \frac{1}{|Q_0|}\int_{Q_0} \Big|\sum_{Q\in\mathcal{S}, Q\subsetneq Q_0} \mathbb{1}_Q(x) - \theta\Big|\,dx \\ &= \frac{1}{|Q_0|}\int_{Q_0\setminus S_1} |\theta|\,dx + \frac{1}{|Q_0|}\int_{S_1} \Big|\sum_{Q\in\mathcal{S}, Q\subsetneq Q_0} \mathbb{1}_Q(x) - \theta\Big|\,dx.\end{aligned}$$

Since S_1 is a disjoint union of $Q_1 \in \mathcal{S}_1$,

$$\begin{aligned}
&\frac{1}{|Q_0|}\int_{S_1}\Big|\sum_{Q\in\mathcal{S},Q\subsetneq Q_0}\mathbb{1}_Q(x)-\theta\Big|\,dx\\
&\quad=\frac{1}{|Q_0|}\sum_{Q_1\in\mathcal{S}_1}\int_{Q_1}\Big|\sum_{Q\in\mathcal{S},Q\subset Q_1}\mathbb{1}_Q(x)-\theta\Big|\,dx\\
&\quad=\frac{1}{|Q_0|}\sum_{Q_1\in\mathcal{S}_1}\int_{Q_1}\Big|\sum_{Q\in\mathcal{S},Q\subsetneq Q_1}\mathbb{1}_Q(x)+1-\theta\Big|\,dx\\
&\quad=\frac{1}{|Q_0|}\Big[\sum_{Q_1\in\mathcal{S}_1}\Big(\int_{Q_1\setminus S_2}|1-\theta|\,dx+\sum_{\substack{Q_2\in\mathcal{S}_2\\ Q_2\subsetneq Q_1}}\int_{Q_2}\Big|\sum_{\substack{Q\in\mathcal{S}\\ Q\subset Q_2}}\mathbb{1}_Q(x)+1-\theta\Big|\,dx\Big)\Big]\\
&\quad=\frac{1}{|Q_0|}|1-\theta|\underbrace{\sum_{Q_1\in\mathcal{S}_1}|Q_1\setminus S_2|}_{|S_1\setminus S_2|}+\frac{1}{|Q_0|}\sum_{Q_2\in\mathcal{S}_2}\int_{Q_2}\Big|\sum_{\substack{Q\in\mathcal{S}\\ Q\subsetneq Q_2}}\mathbb{1}_Q(x)+2-\theta\Big|\,dx.
\end{aligned}$$

So

$$\begin{aligned}
\frac{1}{|Q_0|}\int_{Q_0}|b_{\mathcal{S}}-\langle b_{\mathcal{S}}\rangle_{Q_0}|\,dx&=\theta\,\frac{|Q_0\setminus S_1|}{|Q_0|}+|1-\theta|\,\frac{|S_1\setminus S_2|}{|Q_0|}\\
&\quad+\frac{1}{|Q_0|}\sum_{Q_2\in\mathcal{S}_2}\int_{Q_2}\Big|\sum_{\substack{Q\in\mathcal{S}\\ Q\subsetneq Q_2}}\mathbb{1}_Q(x)+2-\theta\Big|\,dx.
\end{aligned}$$

We can apply the same reasoning to each $Q_2 \in \mathcal{S}_2$:

$$\begin{aligned}
&\int_{Q_2}\Big|\sum_{\substack{Q\in\mathcal{S}\\ Q\subsetneq Q_2}}\mathbb{1}_Q(x)+2-\theta\Big|\,dx\\
&\qquad=\int_{Q_2\setminus S_3}|2-\theta|\,dx+\sum_{\substack{Q_3\in\mathcal{S}_3\\ Q_3\subsetneq Q_2}}\int_{Q_3}\Big|\sum_{\substack{Q\in\mathcal{S}\\ Q\subsetneq Q_3}}\mathbb{1}_Q(x)+3-\theta\Big|\,dx,
\end{aligned}$$

and we can conclude inductively

$$\begin{aligned}
\frac{1}{|Q_0|}\int_{Q_0}|b_{\mathcal{S}}-\langle b_{\mathcal{S}}\rangle_{Q_0}|\,dx&\\
=\theta\,\frac{|Q_0\setminus S_1|}{|Q_0|}+|1-\theta|\,\frac{|S_1\setminus S_2|}{|Q_0|}+|2-\theta|\,\frac{|S_2\setminus S_3|}{|Q_0|}+\cdots&
\end{aligned}\tag{A.1}$$

Suppose for a moment that $\theta \le 1$. Then (A.1) becomes

$$\begin{aligned}\theta\,\frac{|Q_0\setminus S_1|}{|Q_0|} + (1-\theta)\,\frac{|S_1\setminus S_2|}{|Q_0|} + (2-\theta)\,\frac{|S_2\setminus S_3|}{|Q_0|} + \cdots \\ &= \frac{1}{|Q_0|}\big(\theta\,|Q_0\setminus S_1| + (1-\theta)\,|S_1| - (1-\theta)\,|S_2| + (2-\theta)\,|S_2| \\ &\quad - (2-\theta)\,|S_3| + (3-\theta)\,|S_3| - \cdots\big) \\ &= \frac{1}{|Q_0|}\big(\theta\,|Q_0\setminus S_1| + (1-\theta)\,|S_1| + |S_2| + |S_3| + \cdots\big).\end{aligned}$$

Remark A.1. Thoroughly, we have above a sequence of partial sums

$$\begin{aligned}a_k &= c\,|S_1\setminus S_2| + (c+1)\,|S_2\setminus S_3| + \cdots + (c+k-1)\,|S_k\setminus S_{k+1}| \\ &= c\,|S_1| - c\,|S_2| + (c+1)\,|S_2| - (c+1)\,|S_3| \\ &\quad + \cdots + (c+k-1)\,|S_k| - (c+k-1)\,|S_{k+1}| \\ &= c\,|S_1| + |S_2| + |S_3| + \cdots + |S_k| - (c+k-1)\,|S_{k+1}|,\end{aligned}$$

where $c = (1-\theta) \ge 0$. We know that

- the series $\sum_{k=1}^{\infty} |S_k|$ converges, by the Carleson property;
- the "remainder" $(c+k-1)|S_{k+1}| \to 0$ as $k \to \infty$, because the series $\sum_{k=1}^{\infty} k|S_k|$ also converges:

$$\begin{aligned}\sum_{k=1}^{\infty} k\,|S_k| &= |S_1| + 2|S_2| + 3|S_3| + \cdots \\ &= |S_1| + |S_2| + |S_3| + \ldots && (\le \Lambda|S_1|) \\ &\quad + |S_2| + |S_3| + \cdots && (\le \Lambda|S_2|) \\ &\quad + |S_3| + |S_4| + \cdots && (\le \Lambda|S_3|) \\ &\quad + \cdots \\ &\le \Lambda(|S_1| + |S_2| + \cdots) \le \Lambda^2|S_1|.\end{aligned}$$

So

$$\lim_{k\to\infty} a_k = c\,|S_1| + \sum_{k=2}^{\infty} |S_k| - \underbrace{\lim_{k\to\infty} (c+k-1)\,|S_{k+1}|}_{=0}$$

and

$$\frac{1}{|Q_0|}\int_{Q_0} |b_{\mathcal{S}} - \langle b_{\mathcal{S}}\rangle_{Q_0}|\,dx = \frac{1}{|Q_0|}\big(\theta\,|Q_0\setminus S_1| + (1-\theta)\,|S_1| + |S_2| + |S_3| + \cdots\big)$$

holds.

Now,

$$|S_2| + |S_3| + \cdots = \sum_{Q_1\in\mathcal{S}_1} \underbrace{\Big(\sum_{Q\in\mathcal{S}, Q\subsetneq Q_1} |Q|\Big)}_{\le(\Lambda-1)|Q_1| \text{ because } Q_1\in\mathcal{S}} \le (\Lambda-1)\sum_{Q_1\in\mathcal{S}_1} |Q_1| = (\Lambda-1)|S_1|.$$

So

$$\begin{aligned}\frac{1}{|Q_0|}\int_{Q_0}|b_S-\langle b_S\rangle_{Q_0}|\,dx &\le \frac{1}{|Q_0|}\left(\theta\,|Q_0\setminus S_1|+(1-\theta)\,|S_1|+(\Lambda-1)\,|S_1|\right)\\ &= \frac{1}{|Q_0|}\left(\theta\,|Q_0\setminus S_1|+(\Lambda-\theta)\,|S_1|\right)\le \frac{1}{|Q_0|}\left(\Lambda\,|Q_0\setminus S_1|+\Lambda\,|S_1|\right)\\ &= \frac{\Lambda}{|Q_0|}\left(|Q_0\setminus S_1|+|S_1|\right)=\Lambda.\end{aligned}$$

Generally, if $n<\theta\le(n+1)$ for some $n\in\mathbb{N}$, the right-hand side of (A.1) becomes

$$\begin{aligned}&\frac{1}{|Q_0|}\Big[\theta\,|Q_0\setminus S_1|+(\theta-1)\,|S_1\setminus S_2|+\cdots+(\theta-n)\,|S_n\setminus S_{n+1}|\\ &\qquad+\underbrace{(n+1-\theta)\,|S_{n+1}\setminus S_{n+2}|+(n+2-\theta)\,|S_{n+2}\setminus S_{n+3}|+\cdots}_{(n+1-\theta)\,|S_{n+1}|+\underbrace{|S_{n+2}|+|S_{n+3}|+\cdots}_{\le(\Lambda-1)\,|S_{n+1}|}}\Big]\\ &\le\frac{1}{|Q_0|}\big[\theta|Q_0\setminus S_1|+(\theta-1)|S_1\setminus S_2|+\cdots+(\theta-n)|S_n\setminus S_{n+1}|+(\Lambda+n-\theta)|S_{n+1}|\big]\\ &\le\frac{1}{|Q_0|}\big[\Lambda\,|Q_0\setminus S_1|+\Lambda\,|S_1\setminus S_2|+\cdots+\Lambda\,|S_n\setminus S_{n+1}|+\Lambda\,|S_{n+1}|\big]\\ &\le\frac{\Lambda}{|Q_0|}\left(|Q_0\setminus S_1|+|S_1\setminus S_2|+\cdots+|S_n\setminus S_{n+1}|+|S_{n+1}|\right)=\Lambda.\end{aligned}$$

■

B. Proof of Theorem 3.8

Say we have $a\in \mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)$, $b\in \mathrm{BMO}^{\mathcal{D}}(w)$, where w is a weight on $\mathbb{R}^n$, and a fixed $Q_0\in\mathcal{D}$. We look at

$$\Pi_a^*\Pi_{b,Q_0}f:=\sum_{Q\subset Q_0}(a,h_Q)\,(b,h_Q)\,\langle f\rangle_Q\,\frac{\mathbb{1}_Q}{|Q|}$$

and the inner product

$$(\Pi_a^*\Pi_{b,Q_0}f,g)=\sum_{Q\subset Q_0}(a,h_Q)\,(b,h_Q)\,\langle f\rangle_Q\,\langle g\rangle_Q.$$

Within Q_0, we form the local CZ-decompositions of f and g, and the BMO decomposition of b:

$$\begin{aligned}&\mathcal{E}_1:=\{\text{maximal subcubes } R\in\mathcal{D}(Q_0) \text{ such that } \langle|f|\rangle_R>\tfrac{3}{\varepsilon}\langle|f|\rangle_{Q_0}\},\quad E_1:=\textstyle\bigcup_{R\in\mathcal{E}_1}R;\\ &\mathcal{E}_2:=\{\text{maximal subcubes } R\in\mathcal{D}(Q_0) \text{ such that } \langle|g|\rangle_R>\tfrac{3}{\varepsilon}\langle|g|\rangle_{Q_0}\},\quad E_2:=\textstyle\bigcup_{R\in\mathcal{E}_2}R;\\ &\mathcal{E}_3:=\{\text{maximal subcubes } R\in\mathcal{D}(Q_0) \text{ such that } \langle w\rangle_R>\tfrac{3}{\varepsilon}\langle w\rangle_{Q_0}\},\quad E_3:=\textstyle\bigcup_{R\in\mathcal{E}_3}R.\end{aligned}$$

Based on $\mathcal{E}_3$, we define

$$\widetilde{b} := \mathbb{1}_{Q_0} b - \sum_{R\in\mathcal{E}_3} (b - \langle b\rangle_R)\mathbb{1}_R = \sum_{\substack{Q\subset Q_0\\ Q\not\subset E_3}} (b, h_Q)\, h_Q,$$

which satisfies $\widetilde{b} \in \mathrm{BMO}^{\mathcal{D}}(\mathbb{R}^n)$, with

$$\|\widetilde{b}\|_{\mathrm{BMO}^{\mathcal{D}}} \le \frac{6}{\varepsilon}\, \langle w\rangle_{Q_0}\, \|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}.$$

Moreover, $(b, h_Q) = (\widetilde{b}, h_Q)$ for all $Q \subset Q_0$, $Q \not\subset E_3$. Each collection $\mathcal{E}_i$ satisfies

$$\sum_{R\in\mathcal{E}_i} |R| \le \frac{\varepsilon}{3}\, |Q_0|.$$

Finally, let

$$E := E_1 \cup E_2 \cup E_3 \quad \text{and} \quad \mathcal{E} := \{\text{maximal subcubes } R \in \mathcal{D}(Q_0) \text{ such that } R \subset E\}.$$

Then

$$\sum_{R\in\mathcal{E}} |R| \le \varepsilon\, |Q_0|.$$

Now look at $(\Pi_a^* \Pi_{b,Q_0} f, g)$ and split the sum as

$$\begin{aligned}(\mathrm{B.1}) \qquad &|(\Pi_a^* \Pi_{b,Q_0} f, g)|\\ &\le \sum_{\substack{Q\subset Q_0\\ Q\not\subset E}} |(a, h_Q)|\, |(b, h_Q)|\, \langle |f|\rangle_Q \langle |g|\rangle_Q + \sum_{R\in\mathcal{E}} |(\Pi_a^* \Pi_{b,R} f, g)|.\end{aligned}$$

For every $Q \subset Q_0$, $Q \not\subset E$, we have

$$\langle |f|\rangle_Q \le \frac{3}{\varepsilon}\langle |f|\rangle_{Q_0}, \quad \langle |g|\rangle_Q \le \frac{3}{\varepsilon}\langle |g|\rangle_{Q_0}, \quad \text{and} \quad (b, h_Q) = (\widetilde{b}, h_Q),$$

so:

$$\begin{aligned}&\sum_{\substack{Q\subset Q_0\\ Q\not\subset E}} |(a, h_Q)|\, |(b, h_Q)|\, \langle |f|\rangle_Q\, \langle |g|\rangle_Q\\ &\qquad \le \frac{9}{\varepsilon^2}\, \langle |f|\rangle_{Q_0} \langle |g|\rangle_{Q_0} \sum_{Q\subset Q_0, Q\not\subset E} |(a, h_Q)|\, |(\widetilde{b}, h_Q)|\\ &\qquad \le \frac{9}{\varepsilon^2}\, \langle |f|\rangle_{Q_0} \langle |g|\rangle_{Q_0} \underbrace{\Big(\sum_{Q\subset Q_0} |(a, h_Q)|^2\Big)^{1/2}}_{\le C(n)\sqrt{|Q_0|}\,\|a\|_{\mathrm{BMO}^{\mathcal{D}}}} \underbrace{\Big(\sum_{Q\subset Q_0} |(\widetilde{b}, h_Q)|^2\Big)^{1/2}}_{\substack{\le C(n)\sqrt{|Q_0|}\,\|\widetilde{b}\|_{\mathrm{BMO}^{\mathcal{D}}}\\ \le C(n)\frac{6}{\varepsilon}\langle w\rangle_{Q_0}\sqrt{|Q_0|}\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}}},\end{aligned}$$

where $C(n)$ is the dimensional constant arising from using the John–Nirenberg theorem. Finally, we have

$$\sum_{\substack{Q\subset Q_0\\ Q\notin E}} |(a,h_Q)|\,|(b,h_Q)|\,\langle|f|\rangle_Q\,\langle|g|\rangle_Q \leq \frac{C(n)}{\varepsilon^3}\,\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\,\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\,\langle|f|\rangle_{Q_0}\,\langle|g|\rangle_{Q_0}\,\langle w\rangle_{Q_0}|Q_0|.$$

Now we recurse on the $\sum_{R\in\mathcal{E}}$ terms in (B.1) and form $\mathcal{S}(Q_0)$ by adding Q_0 first, $\mathcal{E}$ are the $\mathcal{S}$-children of Q_0, and so on. The collection $\mathcal{S}(Q_0)$ satisfies the $\mathcal{S}$-children definition of sparseness, with

$$\sum_{R\in \mathrm{ch}_{\mathcal{S}}(Q)} |R| \leq \varepsilon |Q| \quad \text{for all } Q\in\mathcal{S}(Q_0),$$

so it is $\frac{1}{1-\varepsilon}$-Carleson. So, if we choose $\varepsilon = (\Lambda-1)/\Lambda$, we have

$$\Big|\sum_{Q\subset Q_0}(a,h_Q)(b,h_Q)\langle f\rangle_Q\langle g\rangle_Q\Big| \leq C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^3\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}\underbrace{\sum_{Q\in\mathcal{S}(Q_0)}\langle w\rangle_Q\langle|f|\rangle_Q\langle|g|\rangle_Q|Q|}_{=(\mathcal{A}^w_{\mathcal{S}(Q_0)}|f|,|g|)}.$$

We summarize this below:

Proposition B.1. *There is a dimensional constant $C(n)$ such that, for all $a\in\mathrm{BMO}^{\mathcal{D}}$, $b\in\mathrm{BMO}^{\mathcal{D}}(w)$, where w is a weight on $\mathbb{R}^n$, fixed $Q_0\in\mathcal{D}$ and $\Lambda>1$, there is a Λ-Carleson sparse collection $\mathcal{S}(Q_0)\subset\mathcal{D}(Q_0)$ such that*

$$\Big|\sum_{Q\subset Q_0}(a,h_Q)(b,h_Q)\langle f\rangle_Q\langle g\rangle_Q\Big| \leq C(n)\Big(\frac{\Lambda}{\Lambda-1}\Big)^3\|a\|_{\mathrm{BMO}^{\mathcal{D}}}\|b\|_{\mathrm{BMO}^{\mathcal{D}}(w)}(\mathcal{A}^w_{\mathcal{S}(Q_0)}|f|,|g|).$$

Say now we have Bloom weights $\mu,\lambda\in A_p$ $(1<p<\infty)$, $\nu:=\mu^{1/p}\lambda^{-1/p}$ on $\mathbb{R}^n$ and $a\in\mathrm{BMO}^{\mathcal{D}}$, $b\in\mathrm{BMO}^{\mathcal{D}}(\nu)$. Suppose further that a has finite Haar expansion. Then there are at most 2^n disjoint dyadic cubes $Q_k\in\mathcal{D}$, $1\leq k\leq 2^n$, such that

$$a=\sum_k\sum_{Q\subset Q_k}(a,h_Q)\,h_Q,$$

and then

$$(\Pi_a^*\Pi_b f,g)=\sum_k(\Pi_a^*\Pi_{b,Q_k}f,g).$$

Given $\Lambda > 1$, by Proposition B.1, there is for each k a Λ-Carleson sparse collection $\mathcal{S}(Q_k) \subset \mathcal{D}(Q_k)$ such that

$$\left|(\Pi_a^* \Pi_{b,Q_k} f, g)\right| \leq C(n) \left(\frac{\Lambda}{\Lambda - 1}\right)^3 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} (\mathcal{A}^{\nu}_{\mathcal{S}(Q_k)} |f|, |g|).$$

Then

$$\left|(\Pi_a^* \Pi_b f, g)\right| \leq C(n) \left(\frac{\Lambda}{\Lambda - 1}\right)^3 \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} (\mathcal{A}^{\nu}_{\mathcal{S}} |f|, |g|),$$

where

$$\mathcal{S} := \bigcup_k \mathcal{S}(Q_k)$$

is a Λ-Carleson sparse collection in $\Upsilon^{\mathcal{D}}(\mathbb{R}^n)$.

Take now $f \in L^p(\mu)$ and $g \in L^{p'}(\lambda')$. By a simple application of Hölder's inequality,

$$|(\mathcal{A}^{\nu}_{\mathcal{S}} |f|, |g|)| \leq \|\mathcal{A}^{\nu}_{\mathcal{S}} : L^p(\mu) \to L^p(\lambda)\| \, \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')}.$$

Then

$$\begin{aligned} &\|\Pi_a^* \Pi_b : L^p(\mu) \to L^p(\lambda)\| \\ &\quad \leq C(n) \|a\|_{\mathrm{BMO}^{\mathcal{D}}} \|b\|_{\mathrm{BMO}^{\mathcal{D}}(\nu)} \sup_{\substack{\mathcal{S} \in \Upsilon^{\mathcal{D}}(\mathbb{R}^n) \\ \Lambda_{(\mathcal{S})} = \Lambda}} \left(\frac{\Lambda}{\Lambda - 1}\right)^3 \|\mathcal{A}^{\nu}_{\mathcal{S}} : L^p(\mu) \to L^p(\lambda)\| \end{aligned}$$

holds for all a with finite Haar expansion, and therefore for all a. This ends the proof of Theorem 3.8.

Acknowledgments. I. Fay Holmes deeply thanks Cristina Pereyra for several conversations about this work, and for her general support. Both authors are grateful to the two anonymous referees for an extremely careful read of our work which led to a much improved write-up.

Funding. I. Fay Holmes is supported by Simons Foundation Collaboration Grants for Mathematicians, Award no. 853930.

References

[1] Beznosova, O. V.: Linear bound for the dyadic paraproduct on weighted Lebesgue space $L_2(w)$. *J. Funct. Anal.* **255** (2008), no. 4, 994–1007.

[2] Blasco, O.: Dyadic BMO, paraproducts and Haar multipliers. In *Interpolation theory and applications*, pp. 11–18. Contemp. Math. 445, Amer. Math. Soc., Providence, RI, 2007.

[3] Bloom, S.: A commutator theorem and weighted BMO. *Trans. Amer. Math. Soc.* **292** (1985), no. 1, 103–122.

[4] Chung, D.: Sharp estimates for the commutators of the Hilbert, Riesz transforms and the Beurling–Ahlfors operator on weighted Lebesgue spaces. *Indiana Univ. Math. J.* **60** (2011), no. 5, 1543–1588.

[5] Cruz-Uribe, D., Martell, J. M. and Pérez, C.: Sharp weighted estimates for classical operators. *Adv. Math.* **229** (2012), no. 1, 408–441.

[6] Duong, X. T., Holmes, I., Li, J., Wick, B. D. and Yang, D.: Two weight commutators in the Dirichlet and Neumann Laplacian settings. *J. Funct. Anal.* **276** (2019), no. 4, 1007–1060.

[7] Grafakos, L.: *Classical and modern Fourier analysis*. Pearson Education, Upper Saddle River, NJ, 2004.

[8] Holmes, I., Lacey, M. T. and Wick, B. D.: Commutators in the two-weight setting. *Math. Ann.* **367** (2017), no. 1-2, 51–80.

[9] Hytönen, T.: *Dyadic analysis and weights*. Lecture notes from a course at University of Helsinki, Spring 2014. http://wiki.helsinki.fi/download/attachments/130069291/dyadic.pdf.

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

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

[12] Lerner, A. K., Ombrosi, S. and Rivera-Ríos, I. P.: On pointwise and weighted estimates for commutators of Calderón–Zygmund operators. *Adv. Math.* **319** (2017), 153–181.

[13] Pereyra, M. C.: Lecture notes on dyadic harmonic analysis. In *Second summer school in analysis and mathematical physics (Cuernavaca, 2000)*, pp. 1–60. Contemp. Math. 289, Amer. Math. Soc., Providence, RI, 2001.

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

[15] Petermichl, S.: Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol. *C. R. Acad. Sci. Paris Sér. I Math.* **330** (2000), no. 6, 455–460.

[16] Treil, S. and Volberg, A.: Wavelets and the angle between past and future. *J. Funct. Anal.* **143** (1997), no. 2, 269–308.

Received March 14, 2022; revised September 2, 2022. Published online November 25, 2022.

Irina Holmes Fay
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA;
irinaholmes@tamu.edu

Valentia Fragkiadaki
Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA;
valeria96@tamu.edu