PARAPRODUCTS, BLOOM BMO AND SPARSE BMO FUNCTIONS

VALENTIA FRAGKIADAKI AND IRINA HOLMES FAY

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 $\mathcal{R}_B$ and compositions of paraproducts $\Pi\Pi$.

In 1985, Steven Bloom proved [2] 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 $\mu, \lambda$ are two $A_p$ weights ($1 < p < \infty$), if and only if $b$ is in a weighted BMO space determined by the two weights $\mu$ and $\lambda$, namely $b \in BMO(\nu)$, where $\nu := \mu^{1/p}\lambda^{-1/p}$ and

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

In [7] this result was extended to commutators $[b, T]$ in $\mathbb{R}^n$ with Calderón-Zygmund operators $T$. Soon after, [11] gave a different proof which yielded a quantitative result for the upper bound:

$$||[b, T] : L^p(\mu) \to L^p(\lambda)|| \leq ||b||_{BMO(\nu)} \left( [\mu]_{A_p}, [\lambda]_{A_p} \right)^{\max\left(1, \frac{1}{p} \right)}.$$  

The proof in [7] took the route of Hytönen’s representation theorem (the $\mathbb{R}^n$, Calderón-Zygmund operator generalization of Petermichl’s result [14] on the Hilbert transform), and relied heavily on paraproduct decompositions. The proof in [11] 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 suspicion, we suspect that these bounds should be smaller than the ones for commutators: in the one-weight case

$$||[b, H] : L^p(w) \to L^p(w)|| \leq ||b||_{BMO[w]}^{2\max\left(1, \frac{1}{p} \right)},$$

and

$$||\Pi_b : L^p(w) \to L^p(w)|| \leq ||b||_{BMO[w]}^{\max\left(1, \frac{1}{p} \right)},$$

are both known to be sharp – see [8][12] and the references therein – (where throughout this paper $A \leq B$ 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 $\Lambda$ 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 3.3 that

$$||\Pi_b : L^p(\mu) \to L^p(\lambda)|| \leq ||b||_{BMO(\nu)} [\mu]_{A_p}^{\frac{1}{p}} [\lambda]_{A_p} = ||b||_{BMO(\nu)} [\mu]_{A_p}^{\frac{1}{p}} [\lambda]_{A_p}^{\frac{1}{p}},$$

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 $[\mu]_{A_p} [\lambda]_{A_p}$. 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 3.1 via an appeal to the one-weight linear $A_2$ bound for the dyadic square function.

2010 Mathematics Subject Classification. 42B20, 42B35, 47A30.

I. Holmes Fay is supported by Simons Foundation: Mathematics and Physical Sciences-Collaboration Grants for Mathematicians, Award number 853930.
Obviously this bound does not recover the one-weight situation: letting $\mu = \lambda = w$ for some $w \in A_2$, $\nu = 1$ and our bound would give
\[ ||\Pi : L^2(w) \rightarrow L^2(w)|| \leq ||b||_{BMO}[w]^2, \]
when we know that the optimal bound is linear in the $A_2$ characteristic. 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" $\mathcal{A}_S^f := \sum_{Q \in \mathcal{S}(\nu) Q} f Q$, where $\mathcal{S}$ is a sparse collection, and proving that $\mathcal{A}_S^f$ satisfies the bound $[\mu]_{A_p}^{1/(p-1)} [\lambda]_{A_p}$ above. We do this in Theorem 2.6. The domination of the paraproduct is treated in Section 3.

Before all this however, we consider in Section 2 a special type of sparse collections, $\mathcal{T}^D(\mathbb{R}^n)$, which are sparse collections with no "infinitely increasing chains" (a terminology borrowed from [8]). We see that any such collection can be associated with a BMO function
\[ b_S := \sum_{Q \in \mathcal{S}} 1_Q, \]
which satisfies $||b_S||_{BMO} \leq \Lambda$, where $\Lambda$ is the Carleson constant of $\mathcal{S}$ (we show this in Appendix $\Lambda$). Once we have a BMO function, we can immediately talk about paraproducts with symbol $b_S$. In fact, we see in Section 2.3 that these functions allow us to express any sparse operator $\mathcal{A}_S$, $S \in \mathcal{T}^D(\mathbb{R}^n)$, as a sum of paraproducts and a martingale transform:
\[ \mathcal{A}_S f = \Pi_{b_S} f + \Pi_{\tau} f + T_{\tau_S} f, \]
where $T_{\tau_S}$ is a martingale transform:
\[ T_{\tau} = \sum_{J \in \mathcal{D}} (\tau_J)_J h_J, \]
where $(\tau_J)_J := \frac{1}{|J|} \sum_{J \supseteq J} 1_J \leq \Lambda$, $\forall J \in \mathcal{D}$.

As discussed in Section 2.3, 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
\[ \sup_{S \in \mathcal{T}^D} ||\mathcal{A}_S||_{L^p(w) \rightarrow L^p(w)} \approx_{n,p,\Lambda} \sup_{b \in BMO^D} ||1_Q f Q||_{L^p(w) \rightarrow L^p(w)} + \sup_{b \in BMO^D} ||\Pi_{\tau}||_{L^p(w) \rightarrow L^p(w)} + \sup_{\tau \in \mathcal{P}} \frac{||T_{\tau}||_{L^p(w) \rightarrow L^p(w)}}{||\tau||_{\infty}}. \]

The process used to obtain the BMO function $b_S$ associated with $\mathcal{S}$ also works with weights, and obtaining a function in weighted BMO spaces associated with $S \in \mathcal{T}^D(\mathbb{R}^n)$: if $w \in A_p$, the function
\[ b_S^w := \sum_{Q \notin S} (w)_Q 1_Q \]
is in $BMO^D(w)$, with $||b_S^w||_{BMO^D(w)} \leq 2[w]_{A_p}$ $\Lambda^p$. Repeating the process above, we try to express $\mathcal{A}_S$ as a sum of the paraproducts associated with $b_S^w$ and a martingale transform – but we discover instead the operator
\[ \mathcal{A}_S^w f := \sum_{Q \notin S} (w)_Q f Q 1_Q, \]
and its decomposition as
\[ \mathcal{A}_S^w f = \Pi_{b_S}^w f + \Pi_{\tau}^w f + T_{\tau_S}^w f, \]
detailed in Proposition 2.4.

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 $\mathcal{A}_S$, we are able to show that norms of sparse operators are equivalent to certain compositions of paraproducts. In Section 2.4, we see that
\[ \mathcal{A}_S = \Pi_{\tau} \Pi_{\tau_S}, \]
where $\tilde{b}_S$ is another BMO function we can easily associate with $S$:
\[ \tilde{b}_S := \sum_{Q \notin S} \sqrt{Q} h_Q. \]
This provides an upper bound:
\[
\sup_{S \in \mathcal{Y}(\mathbb{R}^n)} \frac{\| \mathcal{A}_S : L^p(w) \to L^p(w) \|}{\Lambda} \leq \sup_{a, \beta \in \text{BMO}^D} \frac{\| \mathcal{A}^\beta : L^p(w) \to L^p(w) \|}{\|a\|_{\text{BMO}} \|b\|_{\text{BMO}}^{\nu}}.
\]

For the other direction, we show in Appendix B – using a bilinear form argument – that for all Bloom weights \(\mu, \lambda, \nu\), BMO functions \(a \in \text{BMO}^D\), \(b \in \text{BMO}^D(\nu)\), and \(\Lambda > 1\),
\[
\| \mathcal{A}_S : L^p(w) \to L^p(w) \| \leq C(n)\|a\|_{\text{BMO}}\|b\|_{\text{BMO}(\nu)} \left(\frac{\Lambda}{\Lambda - 1}\right)^3 \| \mathcal{A}_S : L^p(w) \to L^p(w) \|.
\]

Note that taking \(\mu = \lambda = w\) above, for some \(w \in A_p\), we have the one-weight result
\[
\| \mathcal{A}_S : L^p(w) \to L^p(w) \| \leq \|a\|_{\text{BMO}}\|b\|_{\text{BMO}(\nu)}[w]_{A_p}^{\max(1, \frac{1}{\nu})}.
\]

Moreover, we obtain the equivalence of norms
\[
\sup_{S \in \mathcal{Y}(\mathbb{R}^n)} \| \mathcal{A}_S : L^p(w) \to L^p(w) \| \approx_{A_p, n} \sup_{a, \beta \in \text{BMO}^D} \frac{\| \mathcal{A}^\beta : L^p(w) \to L^p(w) \|}{\|a\|_{\text{BMO}} \|b\|_{\text{BMO}}}.
\]

Section 3 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 \(\text{BMO}_D(w)\) functions with finite Haar expansion, and extending to the general case. So this argument works whenever \(\Pi_n\) 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,
\[
\|b\|_{\text{BMO}(w)} := \sup_{Q \in \mathcal{D}} \frac{1}{w(Q)} \int_Q |b - \langle b \rangle_Q| \, dx,
\]
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 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
\[
\Pi_{b, Q_0} f(x) := \sum_{Q \in \mathcal{Q}_0} \langle b, h_Q \rangle \langle f \rangle_Q h_Q(x), \quad \forall Q_0 \in \mathcal{D},
\]
and construct a sparse collection \(\mathcal{S}(Q_0) \subset \mathcal{D}(Q_0)\) which “ends” at \(Q_0\), and such that \(\mathcal{A}^\alpha f\) pointwise dominates \(\Pi_{b, Q_0} f\) on \(Q_0\). Since the Haar expansion of \(b\) effectively dictates the Haar expansion of \(\Pi_n\) (as well as \(\Pi_b\) and \(\Pi_0\)), this will lead from finite Haar expansion \(b\)’s to collections in \(\text{BMO}_D(\mathbb{R}^n)\).

The second author deeply thanks Cristina Pereyra for several conversations about this work, and for her general support.

1. **Setup and Notations**

1.1. **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 \(k^{th}\) 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 \(Q_0\):
\[
\mathcal{D}(Q_0) := \{Q \in \mathcal{D} : Q \subset Q_0\}.
\]
1.2. **Haar Functions.** Given a dyadic grid $D$ on $\mathbb{R}$, we associate to each $I \in D$ the cancellative Haar function $h_I := h_I^0 = \frac{1}{\sqrt{|I|}}(I_+ - I_-)$, where $I_+$ and $I_-$ are the right and left halves of $I$, respectively. The non-cancellative Haar function is $h_I^1 := \frac{1}{\sqrt{|I|}}I$. The cancellative Haar functions \{h_I\}_{I \in 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

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

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 $Q = I_1 \times I_2 \times \ldots \times I_n$, where every $I_k \in D$ is a dyadic interval with common length $|I_k| = \ell(Q)$, we let

$$h^\epsilon_Q(x) := h^\epsilon_{I_1 \times \ldots \times I_n}(x_1, \ldots, x_n) = \prod_{k=1}^n h^\epsilon_{I_k}(x_k),$$

where $\epsilon_k \in \{0, 1\}$ for all $k$, and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$ is known as the signature of $h^\epsilon_Q$. The function $h^\epsilon_Q$ is cancellative except in one case, when $\epsilon \equiv 1$. As in $\mathbb{R}$, the cancellative Haar functions form an orthonormal basis for $L^2(\mathbb{R}^n, dx)$, and an unconditional basis for $L^p(\mathbb{R}^n, dx)$, $1 < p < \infty$. We often write

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

to mean

$$f = \sum_{Q \in D} (f, h^\epsilon_Q) h^\epsilon_Q,$$

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 $\Gamma_I$ in $\mathbb{R}^n$, $n > 1$.

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

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

It is easy to show that

$$(f)_Q = \sum_{R \subset Q} (f, h_R) h_R(Q), \quad \forall Q \in D,$$

where throughout the paper

$$(f)_Q := \frac{1}{|Q|} \int_Q f \, dx,$$

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

1.3. **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 $w$. We denote these spaces by $L^p(w)$. Given $1 < p < \infty$, we say $w \in A_p$ if

$$[w]_{A_p} := \sup_Q (w)_Q^{p'/p-1} < \infty,$$

where the supremum is over cubes $Q \subset \mathbb{R}^n$, $p'$ denotes the Hölder conjugate of $p$:

$$\frac{1}{p} + \frac{1}{p'} = 1,$$

and

$$w' := w^{1-p'/p} = w^{-p'/p}.$$
In fact, \( w \in A_p \) if and only if the conjugate weight \( w' \) is in \( A_{p'} \), with

\[
[w']_{A_{p'}} = [w]^{1 \over p'}_{A_p}.
\]

We restrict our attention to dyadic \( A_p \) weights, denoted \( A^D_p \), 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 \( (L^p(w))^* = L^q(w) \) with inner product \( \langle \cdot, \cdot \rangle_{dw} \), and other times we think of \( (L^p(w))^* \approx L^{p'}(w') \) with regular Lebesgue inner product \( \langle \cdot, \cdot \rangle \). We refer the reader to Chapter 9 of [6] for a thorough treatment of \( A_p \) weights.

1.4. Paraproducts and BMO. We say \( b \in BMO(\mathbb{R}^n) \) if

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

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

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

We similarly restrict our attention to dyadic BMO spaces, \( BMO^D \) and \( BMO^D(w) \) 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:

\[
\Pi_b(f)(x) := \sum_{l \in \mathcal{D}} (b, h_l) f_l h_l(x),
\]

\[
\Pi^*_b(f)(x) := \sum_{l \in \mathcal{D}} (b, h_l) \frac{f_l(x)}{|l|}.
\]

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:

\[
\Pi_b(f)(x) := \sum_{Q \in \mathcal{D}} (b, h_Q) f_Q h_Q(x),
\]

\[
\Pi^*_b(f)(x) := \sum_{Q \in \mathcal{D}} (b, h_Q) \frac{f_Q(x)}{|Q|},
\]

\[
\Gamma_b(f)(x) := \sum_{Q \in \mathcal{D}} \sum_{\epsilon, \eta \in \mathbb{Z}^n, |\epsilon| \neq |\eta|} (b, h^\epsilon_Q) (f, h^\eta_Q) \frac{1}{|Q|} h^{\epsilon + \eta}.
\]

\( \Pi_b \) and \( \Pi^*_b \) are adjoints in \( L^2(\mathbb{R}^n) \), and \( \Gamma_b \) is self-adjoint. Generally, in the \( L^p \)-situation, we still have

\[
(\Pi_b f, g) = (f, \Pi^*_b g),
\]

so if we think of \( \Pi_b : L^p(\mu) \to L^p(\lambda) \) for two \( A_p \) weights \( \mu, \lambda \), its adjoint is \( \Pi^*_b : L^{p'}(\lambda') \to L^{p'}(\mu') \) – where we are thinking of Banach space duality in terms of \( (L^p(\mu))^* \approx L^{p'}(\mu') \) and \( (L^p(\lambda))^* \approx L^{p'}(\lambda') \), both with regular Lebesgue inner product \( \langle \cdot, \cdot \rangle \).

2. Sparse BMO Functions

2.1. Sparse Families. Let \( 0 < \eta < 1 \). A collection \( \mathcal{S} \subset \mathcal{D} \) is said to be \( \eta \)-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 \( \Lambda \)-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 \( Q \in S \). It is also easy to see that any \( \eta \)-sparse collection is \( 1/p \)-Carleson. Far less obvious is the remarkable property that any \( \Lambda \)-Carleson collection is \( 1/\Lambda \)-sparse, which is proved in the now classic work \cite{10}.

A special type of sparse collection which appears most frequently in practice is defined in terms of so-called “\( S \)-children.” Suppose a family \( S \subseteq D \) has no infinitely increasing chains. Then the \( \Lambda \)-sparse collection is \( 1 \)-sparse with pairwise disjoint \( \eta \)-sparse \( Q \)-children of \( Q \). Any two maximal \( p \)-measure \( \Lambda \)-sparse: let

\[
E_Q := Q \setminus \bigcup_{P \in \text{ch}_\Lambda(Q)} P,
\]

which are clearly pairwise disjoint, and satisfy \(|E_Q| \geq (1-\alpha)|Q|\).

A collection that is sparse with respect to Lebesgue measure is also sparse with respect to any \( A_p \) measure \( w \). Recall that (see \cite{6}, Proposition 9.1.5) an equivalent definition for \([w]_{A_p}\) is

\[
[w]_{A_p} = \sup_{Q \subseteq D} \sup_{f \in L^p(Q, w)} \frac{\langle |f| \rangle_Q}{\mathbb{B}_w^p(|f|)},
\]

where

\[
\mathbb{B}_w^p f := \frac{1}{w(Q)} \int_Q f \, dw.
\]

Taking \( f = 1_A \) above, for some measurable subset \( A \) of a fixed dyadic cube \( Q \), we get

\[
\left( \frac{|A|}{|Q|} \right)^p \leq [w]_{A_p} \frac{w(A)}{w(Q)}, \quad \forall A \subseteq Q, Q \in D.
\]

So, say \( S \) is \( \eta \)-sparse with pairwise disjoint \( \{E_Q\}_{Q \in S} \) subsets \( E_Q \subseteq Q \) and \(|E_Q| \geq \eta|Q|\). Then

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

and

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

\* 

2.2. Sparse BMO Functions. We borrow the following terminology from \cite{8}: we say a collection \( S \subseteq D \) has an infinitely increasing chain if there exist \( \{Q_k\}_{k \in \mathbb{N}} \) \( Q_k \in S \), such that \( Q_k \subseteq Q_{k+1} \), for all \( k \in \mathbb{N} \). The following Lemma is also found in \cite{8}.

**Lemma 2.1.** If a collection \( S \subseteq D \) has no infinitely increasing chains, then every \( Q \in S \) is contained in a maximal \( Q^* \in S \) - in the sense that there exists no \( R \in S \) such that \( R \supseteq Q \). Any two maximal \( p^* \), \( Q^* \) elements of \( S \) are disjoint.

These types of collections will be important for us, so we let

\[
\Upsilon^D(\mathbb{R}^n)
\]

denote the set of all sparse collections in \( D \) which have no infinitely increasing chains.

**Lemma 2.2.** Let \( S \in \Upsilon^D(\mathbb{R}^n) \) be a sparse collection with no infinitely increasing chains. Then the set of points contained in infinitely many elements of \( S \) has measure 0.

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

\[
x \in \ldots \subseteq Q_k \subseteq \ldots \subseteq Q_2 \subseteq Q_1 = Q^*
\]
terminating at some maximal \( Q^* \in S^* \). 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 \( k \in \mathbb{N} \):

\[
k|A| \leq \sum_{i=1}^{k} |Q_i| \leq \Lambda|Q^*|,
\]

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

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

\[
\sum_{k=1}^{\infty} |Q_k| \leq \sum_{Q \in S, Q \subsetneq Q^*} |Q| \leq \Lambda|Q^*|
\]

converges. \( \square \)

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

\[
b_S := \sum_{Q \in S} 1_Q.
\]

By Lemma [2.2] we know that \( b_S \) is almost everywhere finite: if \( x \) is contained in infinitely many elements of \( S \), then \( b_S(x) = \infty \), but this can only happen on a set of measure zero.

Note also that \( b_S \) is locally integrable: for some \( Q_0 \in \mathcal{D} \),

\[
\langle b_S \rangle_{Q_0} = \frac{1}{|Q_0|} \left( \sum_{Q \in S, Q \subsetneq Q_0} |Q| + \sum_{Q \in S, Q \supseteq Q_0} |Q| \right)
\]

\[
= \frac{1}{|Q_0|} \sum_{Q \in S, Q \subsetneq Q_0} |Q| + \sum_{Q \in S, Q \supseteq Q_0} \lambda(Q) \leq \lambda \Lambda < \infty,
\]

\( < \infty \) because \( S \in \mathcal{T}^D \).

Then, for some \( Q_0 \in \mathcal{D} \):

\[
(b_S - \langle b_S \rangle_{Q_0}) 1_{Q_0} = \sum_{Q \in S, Q \subsetneq Q_0} 1_Q + \#(Q \in S : Q \supseteq Q_0) 1_{Q_0} - \frac{1}{|Q_0|} \sum_{Q \in S, Q \supseteq Q_0} |Q| - \#(Q \in S : Q \supseteq Q_0) 1_{Q_0}
\]

\[
= \sum_{Q \in S, Q \subsetneq Q_0} 1_Q - \frac{1}{|Q_0|} \sum_{Q \in S, Q \supseteq Q_0} |Q|.
\]

In fact, we can reduce this further to

\[
(2.2) \quad (b_S - \langle b_S \rangle_{Q_0}) 1_{Q_0} = \sum_{Q \in S, Q \subsetneq Q_0} 1_Q - \frac{1}{|Q_0|} \sum_{Q \in S, Q \supseteq Q_0} |Q|,
\]

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

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

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

**Theorem 2.3.** Let \( S \in \mathcal{T}^D(\mathbb{R}^n) \) be a sparse collection with no infinitely increasing chains and Carleson constant \( \Lambda \). Then the function \( b_S = \sum_{Q \in S} 1_Q \) is in \( BMO^D(\mathbb{R}^n) \), with

\[
\|b_S\|_{BMO^D(\mathbb{R}^n)} \leq \Lambda.
\]
This process works to yield a weighted BMO function as well: with any \( S \in \mathcal{T}^D(\mathbb{R}^n) \) and \( w \in A^D_p \) we associate the function

\[
b^w_S := \sum_{Q \subseteq S} \langle w \rangle_Q 1_Q.
\]

As before, \( S \in \mathcal{T}^D(\mathbb{R}^n) \) ensures that \( b^w_S \) is a.e. finite, locally integrable, and

\[
1_{Q_0}(b^w_S - \langle b^w_S \rangle_{Q_0}) = \sum_{Q \subseteq S, Q \subseteq Q_0} \langle w \rangle_Q 1_Q - \frac{1_{Q_0}}{|Q_0|} \sum_{Q \subseteq S, Q \subseteq Q_0} w(Q), \forall Q_0 \in \mathcal{D}.
\]

By (2.1),

\[
\frac{1}{|Q_0|} \sum_{Q \subseteq S, Q \subseteq Q_0} w(Q) \leq [w]_{A^p} A^p \langle w \rangle_{Q_0},
\]

which then easily gives

\[
\frac{1}{w(Q_0)} \int_{Q_0} [b^w_S - \langle b^w_S \rangle_{Q_0}] dx \leq 2[w]_{A^p} A^p,
\]

so

\[
b^w_S \in BMO^D(w), \text{ with } \|b^w_S\|_{BMO^D(w)} \leq 2[w]_{A^p} A^p.
\]

2.3. Sparse Operators as Sums of Paraproducts and Martingale Transform. 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

\[
\mathcal{A}^w_S f := \sum_{I \subseteq S} \langle w \rangle_I (f) I_I,
\]

where \( S \in \mathcal{T}^D(\mathbb{R}) \) and \( w \) is an \( A^D_p \) weight on \( \mathbb{R} \), \( 1 < p < \infty \). A particularly interesting instance of \( \mathcal{A}^w_S \) occurs when \( w = \nu \in A^D_2 \), where \( \nu := \mu^{1/p} \lambda^{-1/p} \) for two weights \( \mu, \lambda \in A^D_p \). We treat this operator in more detail in Section 2.5.

Using the \( b^w_S \) function associated with \( S \) and \( w \), we write

\[
(2.3) \quad \mathcal{A}^w_S f = \mathcal{A}^w_S - b^w_S \cdot f + b^w_S \cdot f = \mathcal{A}^w_S f - b^w_S \cdot f + (\Pi_{b^w_S} f + \Pi_{b^w_S} f + \Pi_{b^w_S} f).
\]

Now recall that

\[
\langle b^w_S \rangle_{J_0} = \langle \tau^w_S \rangle_{J_0} + \sum_{J \supseteq J_0} \langle w \rangle_J, \forall J_0 \in \mathcal{D},
\]

where

\[
\langle \tau^w_S \rangle_J := \frac{1}{|J|} \sum_{I \in S, I \subseteq J} w(I), \forall J \in \mathcal{D},
\]

a quantity always bounded if \( w \in A^D_p \):

\[
\langle \tau^w_S \rangle_J \leq [w]_{A^p} A^p \langle w \rangle_J.
\]

So:

\[
\Pi_{b^w_S}(x) = \sum_{J \in \mathcal{D}} (f, h_J) (\tau^w_S)_J h_J(x)
\]

\[
= \sum_{J \in \mathcal{D}} (f, h_J) \langle \tau^w_S \rangle_J + \sum_{K \in S, K \supseteq J} \langle w \rangle_K h_J(x)
\]

\[
= \langle \tau^w_S \rangle_J (f, h_J) h_J(x) + \sum_{J \in \mathcal{D}} (f, h_J) h_J(x) \left( \sum_{K \in S, K \supseteq J} \langle w \rangle_K \right).
\]
The second term can be further explored as

$$
\sum_{J \in \mathcal{D}} (f, h_j) h_J(x) \left( \sum_{K \in \mathcal{S} : K \supseteq J} (w)_K \right) = \sum_{K \in \mathcal{S}} (w)_K \left( \sum_{J \subset K} (f, h_j) h_J(x) \right) = \sum_{K \in \mathcal{S}} (w)_K \left( f(x) - \langle f \rangle_K \right) 1_K(x)
$$

$$
= f(x) \cdot \sum_{K \in \mathcal{S}} (w)_K 1_K(x) - \sum_{K \in \mathcal{S}} (w)_K \langle f \rangle_K 1_K(x)
$$

$$
= f(x) \cdot b^w_S(x) - \mathcal{A}^w_S f(x).
$$

Returning to (2.3):

$$
\mathcal{A}^w_S f = \mathcal{A}^w_S f - b^w_S \cdot f + (\Pi b^w_S f + \Pi^w_{b^w_S} f) + T_{\tau_S} f + f \cdot b^w_S - \mathcal{A}^w_S f,
$$

so we have:

**Proposition 2.4.** Any weighted sparse operator $\mathcal{A}^w_S$, where $w \in A^D_p$ is a weight on $\mathbb{R}$ and $\mathcal{S} \in \mathcal{T}^D(\mathbb{R})$ is a sparse collection with no infinitely increasing chains, may be expressed as

$$
(2.4) \quad \mathcal{A}^w_S f = \Pi b^w_S f + \Pi^w_{b^w_S} f + T_{\tau_S} f,
$$

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

$$
T_{\tau_S} f(x) := \sum_{J \in \mathcal{D}} (\tau^w_S)_J (f, h_J) h_J(x), \quad \text{where} \quad (\tau^w_S)_J := \frac{1}{|J|} \sum_{l \in \mathcal{S} : J \subseteq l} w(l) \leq |w|_{A^p} \Lambda^p w, \forall J \in \mathcal{D}.
$$

**Remark 2.1.** In case $w \equiv 1$, we obtain the unweighted situation

$$
(2.5) \quad \mathcal{A}_S f = \Pi b_S f + \Pi^w_{b_S} f + T_{\tau_S} f,
$$

where $T_{\tau_S}$ is a martingale transform:

$$
T_{\tau_S} = \sum_{J \in \mathcal{D}} (\tau_S)_J (f, h_J) h_J, \quad \text{where} \quad (\tau_S)_J := \frac{1}{|J|} \sum_{l \in \mathcal{S} : J \subseteq l} 1^l \leq \Lambda, \forall J \in \mathcal{D}.
$$

**Remark 2.2.** In fact, (2.4) expresses sparse operators as Haar multipliers: recall that a Haar multiplier is an operator of the form

$$
(\Pi_b + \Pi^w_b) f = \sum_J (b - (b)_J)(f, h_J) h_J.
$$

So, from (2.4):

$$
\mathcal{A}^w_S f(x) = \left[ (b^w_S(x) - (b^w_S)_J) \phi_J(x) + (\tau^w_S)_J \right](f, h_J) h_J(x).
$$

Look more closely now at (2.5): $\mathcal{A}_S = \Pi b_S + \Pi^w_{b_S} + T_{\tau_S}$. This gives an upper bound for $\|\mathcal{A}_S : L^p(w) \to L^p(w)\|$ in terms of the norms of paraproducts and martingale transform – when usually it is the norms of sparse operators that are used as upper bounds:

$$
\|\mathcal{A}_S f\|_{L^p(w)} \leq \|\Pi b_S f\|_{L^p(w)} + \|\Pi^w_{b_S} f\|_{L^p(w)} + \|T_{\tau_S} f\|_{L^p(w)}.
$$
Divide above by \( \Lambda(S) \) := \( \Lambda \), the Carleson constant of \( S \), and recall that \( \| \tau_S \|_\infty \leq \Lambda \):

\[
\frac{\| A_S f \|_{L^p(w)}}{\Lambda} \leq \frac{\| \Pi_{b_L} f \|_{L^p(w)}}{\Lambda} + \frac{\| \Pi_{b_R} f \|_{L^p(w)}}{\Lambda} + \frac{\| T_{\tau_S} f \|_{L^p(w)}}{\Lambda},
\]

from which we can deduce that, for all \( \Lambda > 1 \):

\[
\sup_{S \in \mathcal{Y}_D(\mathbb{R})} \| A_S : L^p(w) \to L^p(w) \| \leq \sup_{b \in \text{BMO}_D} \frac{\| \Pi_b : L^p(w) \to L^p(w) \|}{\| b \|_{\text{BMO}_D}} + \sup_{\tau \in \ell^p_{\infty}} \frac{\| T_{\tau} : L^p(w) \to L^p(w) \|}{\| \tau \|_\infty}.
\]

Given the well-known domination results [9] for the martingale transform and paraproducts:

\[
\sup_{S \in \mathcal{Y}_D(\mathbb{R})} \| A_S : L^p(w) \to L^p(w) \| \approx_{\Lambda,p} \sup_{b \in \text{BMO}_D} \frac{\| \Pi_b : L^p(w) \to L^p(w) \|}{\| b \|_{\text{BMO}_D}} + \sup_{\tau \in \ell^p_{\infty}} \frac{\| T_{\tau} : L^p(w) \to L^p(w) \|}{\| \tau \|_\infty}.
\]

**Remark 2.3.** It would be interesting if the martingale and paraproducts can be “separated” somehow, and to obtain independently that paraproducts and martingale transforms have the same dependency on [\( w \)\]_\( A_p \) by showing they are both equivalent to \( \| A_S \| \). However, we can show that the norms of \( A_S \) are equivalent to norms of certain compositions of paraproducts. We do this next.

### 2.4. Sparse Operators and Compositions of Paraproducts

Consider the composition

\[
\Pi_a \Pi_b f = \sum_{Q \in \mathcal{D}} (a, h_Q)(b, h_Q) (f)_Q 1_Q.\]

We show in Appendix [3] using a bilinear form argument, that:

**Theorem 2.5.** There is a dimensional constant \( C(n) \) such that for all Bloom weights \( \mu, \lambda \in A_p \) (\( 1 < p < \infty \)), \( v := \mu^{1/p} \lambda^{-1/p} \) on \( \mathbb{R}^n \), \( \text{BMO} \) functions \( a \in \text{BMO}_D(\mathbb{R}^n) \), \( b \in \text{BMO}_D(\mathbb{v}) \), and \( \Lambda > 1 \):

\[
\| \Pi_a \Pi_b : L^p(\mu) \to L^p(\lambda) \| \leq C(n) \| a \|_{\text{BMO}_D} \| b \|_{\text{BMO}_D(\mathbb{v})} \sup_{S \in \mathcal{Y}_D(\mathbb{R})} \left( \frac{\Lambda}{\Lambda - 1} \right)^3 \| A_S \| : L^p(\mu) \to L^p(\lambda) \|
\]

Some immediate observations about this result:

- From Theorem 2.6

\[
\| \Pi_a \Pi_b : L^p(\mu) \to L^p(\lambda) \| \leq \| a \|_{\text{BMO}_D} \| b \|_{\text{BMO}_D(\mathbb{v})} \| \mu \|_{A_p} \| \lambda \|_{A_p} \left( \frac{\Lambda}{\Lambda - 1} \right).
\]

- Take \( \mu = \lambda = w \), for some \( w \in A_p \). Then \( v = 1 \) and we obtain in the one-weight situation:

\[
\| \Pi_a \Pi_b : L^p(w) \to L^p(w) \| \leq \| a \|_{\text{BMO}_D} \| b \|_{\text{BMO}_D(\mathbb{v})} \| w \|_{A_p}^{\max(1, \frac{1}{p})}.
\]

- 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 \( b \in \text{BMO}_D(\mathbb{v}), a \in \text{BMO}_D \).
Let $S \in \mathcal{D}^0(\mathbb{R}^n)$. We associated with $S$ the BMO function $b_S = \sum_{Q \in S} 1_Q$. There is another, even more obvious BMO function we can associate with $S$:

$$
\tilde{b}_S := \sum_{Q \in S} \sqrt{|Q|h_Q} = \sum_{e \neq 1} \sqrt{|Q|h'_Q}.
$$

For any $Q_0 \in D$:

$$
\frac{1}{|Q_0|} \int_{Q_0} |\tilde{b}_S - (\tilde{b}_S)_{Q_0}|^2 \, dx = \frac{1}{|Q_0|} \sum_{Q \in S, Q_0 \subseteq S} |Q| = (2^n - 1) \frac{1}{|Q_0|} \sum_{Q < Q_0, Q_0 \subseteq S} |Q| \leq (2^n - 1)\Lambda,
$$

so

$$
\|\tilde{b}_S\|_{BMO} \leq \sqrt{(2^n - 1)\Lambda}.
$$

Moreover,

$$
\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} f = \sum_{Q \in D} (\tilde{b}_S, h'_Q) Q \frac{1_Q}{|Q|} = \sum_{Q \in S} |Q(f) Q \frac{1_Q}{|Q|} = (2^n - 1) \sum_{Q \in S} \langle f \rangle Q \frac{1_Q}{|Q|},
$$

so we may express the sparse operator $A_S$ as

$$
A_S = \frac{1}{2^n - 1} \Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S}.
$$

Then

$$
\frac{\|A_S f\|_{L^p(w)}}{\Lambda} = \frac{\|\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} f\|_{L^p(w)}}{\Lambda} \leq \frac{\|\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} f\|_{L^p(w)}}{\|	ilde{b}_S\|_{BMO}^2} \leq \sup_{a,b \in BMO^p} \frac{\|\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} f\|_{L^p(w)}}{\|a\|_{BMO^p} \|b\|_{BMO^p}},
$$

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

$$
\sup_{S \in \mathcal{D}^0(\mathbb{R}^n) \backslash A_S = \Lambda} \frac{\|A_S : L^p(w) \to L^p(w)\|}{\Lambda} \leq \sup_{a,b \in BMO^p} \frac{\|\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} : L^p(w) \to L^p(w)\|}{\|a\|_{BMO^p} \|b\|_{BMO^p}}.
$$

Combined with (2.6), we have

$$
\sup_{S \in \mathcal{D}^0(\mathbb{R}^n)} \|A_S : L^p(w) \to L^p(w)\| \approx_{\Lambda,p,n} \sup_{a,b \in BMO^p(\mathbb{R}^n)} \frac{\|\Pi_{\tilde{b}_S} \Pi_{\tilde{b}_S} : L^p(w) \to L^p(w)\|}{\|a\|_{BMO^p} \|b\|_{BMO^p}}.
$$

2.5. The Bloom Sparse Operator $A_S^\mu$. Consider

$$
A_S^\mu f = \sum_{Q \in S} \langle \nu \rangle (Q) (\nu) Q \frac{1_Q}{|Q|},
$$

for a sparse collection $S \subset D(\mathbb{R}^n)$, 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 $A_S f = \sum_{Q \in S} (\nu) Q \frac{1_Q}{|Q|}$. We want something like $\|A_S^\mu f\|_{L^p(\lambda)} \leq C\|f\|_{L^p(\mu)}$, and we use duality to express

$$
\|A_S^\mu f\|_{L^p(\lambda)} = \sup_{g \in L^{p'}(\lambda')} \langle (A_S^\mu f, g) \rangle.
$$

So we look for a bound of the type $\|(A_S^\mu f, g)\| \leq C\|f\|_{L^p(\mu)}\|g\|_{L^{p'}(\lambda')}$.

$$
\|(A_S^\mu f, g)\| = \sum_{Q \in S} \langle (\nu) Q (\nu) Q (\nu) \rangle Q \frac{1_Q}{|Q|} \leq \sum_{Q \in S} \langle (\nu) Q (\nu) \rangle Q (\nu) \langle Q \rangle \leq \int \langle (\nu) Q (\nu) \rangle (\nu) Q (\frac{1_Q}{|Q|}) (\nu) d\nu
$$

$$
\leq \int (A_S f) (\mu) (A_S g) \mu^{1/p} \lambda^{-1/p} \, dx \leq \|A_S f\|_{L^p(\mu)} \|A_S g\|_{L^p(\lambda')} \leq \|A_S : L^p(\mu) \to L^p(\lambda')\| \cdot \|A_S : L^p(\lambda') \to L^p(\lambda')\| \cdot \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')}.
$$

This yields the same dependency on the $A_p$ characteristics of $\mu, \lambda$ as obtained in $\Pi$ for commutators:

$$
\|A_S^\mu : L^p(\mu) \to L^p(\lambda)\| \leq (\mu, \lambda)_{A_p}^{\text{max}}(1, \frac{1}{p}).
$$
We give another proof, inspired by the beautiful proof in [4] of the $A_2$ conjecture for usual unweighted sparse operators, which yields a smaller bound.

**Theorem 2.6.** Let $S \subset D$ be a sparse collection of dyadic cubes, $\mu, \lambda \in A_p^D$, $1 < p < \infty$ be two $A_p$ weights on $\mathbb{R}^n$, and $v := \mu^{1/p} \lambda^{-1/p}$. Then the Bloom sparse operator

$$\mathcal{A}_S f := \sum_{Q \in S} (v)_Q (f)_Q 1_Q$$

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

$$\|\mathcal{A}_S f\|_{L^p(\lambda)} \leq \Lambda^{p-p/2}(p')^2 [\mu']_{A_p} [\lambda]_{A_p} = \Lambda^{p-p/2}(p')^2 [\mu]^{1/p}_{A_p} [\lambda]_{A_p},$$

where $\Lambda$ is the Carleson constant of $S$.

**Proof.** In looking for a bound of the type $\|\mathcal{A}_S f\|_{L^p(\lambda)} \leq C \|f\|_{L^p(\mu)}$, consider instead $\varphi := \mu f'$: then $\|\varphi\|_{L^p(\mu)} = \|f\|_{L^p(\mu')}$, so we look instead for a bound of the type $\|\mathcal{A}_S (f \mu')\|_{L^p(\lambda)} \leq C \|f\|_{L^p(\mu')}$. Using the standard $L^p(\lambda) - L^{p'}(\lambda)$ duality with $\langle \cdot, \cdot \rangle_{L^p(\lambda)}$ inner product, we write

$$\|\mathcal{A}_S (f \mu')\|_{L^p(\lambda)} = \sup_{g \in L^{p'}(\lambda)} \|\mathcal{A}_S (f \mu'), g\|_1,$$

meaning we finally look for a bound of the type

$$\|(\mathcal{A}_S f \mu'), g\|_1 \leq C \|f\|_{L^p(\mu')} \|g\|_{L^{p'}(\lambda)}.$$

As in [4], we make use of the weighted dyadic maximal function:

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

and its property of being $L^q(\mu)$-bounded with a constant independent of $\mu$.

**Theorem 2.7.** For any locally finite Borel measure $\mu$ on $\mathbb{R}^n$ and any $q \in (1, \infty)$:

$$\|M^D f\|_{L^q(\mu)} \leq q'.$$

See, for example, [8] for a proof of this fact.

Now:

$$\|(\mathcal{A}_S f \mu'), g\|_1 = \sum_{Q \in S} (v)_Q (f \mu')_Q (g \lambda)_Q |Q| \leq \sum_{Q \in S} (f \mu')_Q (g \lambda)_Q (v)_Q |Q|.$$

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

$$\sum_{Q \in S} (f \mu')_Q (g \lambda)_Q (v)_Q |Q| = \sum_{Q \in S} \left[ \mathbb{E}_Q (f \mu') (g \lambda) (v)_Q \right] |Q|.$$

Apply the fact that $(v)_Q \leq (\mu)^{1/p}(\lambda)^{1/p}$ (an easy consequence of Hölder’s inequality), and the fact that for any $A_p$ weight $w$, we have

$$[w]^{1/p}_A = \sup_Q \mathbb{E}_Q [w]^{1/p}(w')^{1/p}_Q,$$

to go further:

$$\|(\mathcal{A}_S f \mu'), g\|_1 \leq \sum_{Q \in S} \left[ \mathbb{E}_Q (f \mu') (g \lambda) (v)_Q \right] |Q| \leq \sum_{Q \in S} \left[ \mathbb{E}_Q (f \mu') (g \lambda) (v)_Q \right] |Q| \leq \sum_{Q \in S} \left[ \mathbb{E}_Q (f \mu') (g \lambda) (v)_Q \right] |Q| \leq \sum_{Q \in S} \left[ \mathbb{E}_Q (f \mu') (g \lambda) (v)_Q \right] |Q|.$$
Now apply \(2.1\):
\[
\mu'(Q) \leq [\mu']_{A,p} \Lambda^{p'} \mu'(E_Q) = [\mu']_{A,p}^{p-1} \Lambda^{p'} \mu'(E_Q) \quad \text{and} \quad \lambda(Q) \leq [\lambda]_{A,p} \Lambda^{p} \lambda(E_Q),
\]
so we may later use disjointness of the sets \(\{E_Q\}_{Q \in S}\).

\[
\left(\sum_{Q \in S} \left(\mathbb{E}_{Q} f\right)^{p} \mu'(Q)^{1/p} \right)^{1/p} \leq [\mu']_{A,p}^{p-1} \Lambda^{p'} \left(\sum_{Q \in S} \left(\mathbb{E}_{Q} f\right)^{p} \mu'(E_Q)\right)^{1/p} \\
= [\mu']_{A,p}^{p-1} \Lambda^{p'} \left(\sum_{Q \in S} \int_{E_Q} \left(\mathbb{E}_{Q} f\right)^{p} \mu'(Q)^{1/p} \right)^{1/p} \\
\leq [\mu']_{A,p}^{p-1} \Lambda^{p'} \left(\sum_{Q \in S} \int_{E_Q} \left(M_{\mu'}^{D} f\right)^{p} \mu'(Q)^{1/p} \right)^{1/p} \\
\leq [\mu']_{A,p}^{p-1} \Lambda^{p'} p\|f\|_{L^{p}(\mu')}
\]
Putting these estimates together:
\[|\langle \mathcal{A}^w_{S}(f \mu'), g \lambda \rangle| \leq [\mu']_{A,p}^{p-1} \Lambda^{p'} p\|f\|_{L^{p}(\mu')} [\lambda]_{A,p}^{p-1} \Lambda^{p'} \|g\|_{L^{p'}(\lambda)}
\]
\[= [\mu']_{A,p}^{p-1} [\lambda]_{A,p} \Lambda^{p'} p\|f\|_{L^{p}(\mu')} \|g\|_{L^{p'}(\lambda)}
\]
\[= [\mu']_{A,p} [\lambda]_{A,p} \Lambda^{p' - 2} p\|f\|_{L^{p}(\mu')} \|g\|_{L^{p'}(\lambda)},
\]
which proves the theorem. \(\Box\)

3. Paraproducts and Bloom BMO

We show the following pointwise domination result, inspired by ideas in \([9]\) on pointwise domination of the martingale transform.

**Theorem 3.1.** There is a dimensional constant \(C(n)\) such that: for every \(\Lambda > 1\), weight \(w\) on \(\mathbb{R}^n\), \(b \in BMO^D(w)\), fixed dyadic cube \(Q_0 \in D\) and \(f \in L^1(Q_0)\), there is a \(\Lambda\)-Carleson sparse collection \(S(Q_0) \subset D(Q_0)\) (depending on \(b, w, f\)) such that:

\[\forall x \in Q_0 : \|\Pi_{b,Q_0} f(x)\| \leq C(n) \left(\frac{\Lambda}{\Lambda - 1}\right)^2 \|b\|_{BMO^D(w)} \|\mathcal{A}^w_{S(Q_0)} f\| (x).
\]
The same holds for the other paraproducts \(\Pi^*_b\) and \(\Gamma_b\).

Assuming this, return to the Bloom situation for a moment and say \(b \in BMO^D(v)\) has finite Haar expansion. Then there are at most \(2^n\) disjoint dyadic cubes \(\{Q_k\}_{1 \leq k \leq 2^n} \subset D\) such that \(b = \sum_{k} \mathbb{1}_{Q_k}(b, h_Q) h_Q\), and then \(\Pi_b f = \sum_{k} \Pi_{b,Q_k} f\). So, assuming Theorem 3.1 there are \(\Lambda\)-Carleson
Remark 3.1. The same holds for the other paraproducts $\Pi_s$ and $\Gamma_b$.

In light of the bound for $A_S^\nu$ in Theorem 2.6, pick some value for $\Lambda$, say $\Lambda = 2$, and we have:

**Theorem 3.3.** Given Bloom weights $\mu, \lambda \in A_p^D, \nu = \mu^{1/p} \lambda^{-1/p}$, for all $b \in BMO^D(\nu)$:

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

The same holds for the other paraproducts $\Pi_s^\nu$ and $\Gamma_b$.

**Remark 3.1.** The result actually follows immediately for $\Pi_b^\nu$, since

$$\|\Pi_b : L^p(\mu) \to L^p(\lambda)\| = \|\Pi_b^\nu : L^p(\mu^\nu) \to L^p(\mu^\nu)\|$$

and

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

**Remark 3.2.** 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 $A_2^D$ weight $w$, in which case the “intermediary” Bloom weight is also $\nu = w$:

$$\|\Pi_b : L^2(w) \to L^2(w^{-1})\| \leq \|b\|_{BMO(w)}\|w\|_{A_2^D}^2$$

(3.1)

**3.1. Proof that the quadratic bound $[w]_{A_2^D}^2$ in (3.1) 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 $BMO(w)$, with

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

To see this, if $Q$ is a cube:

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

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

$$\Pi_w f = \sum_{i \in D} (w, h_i)(f)_i h_i$$

$$\Pi_w^\nu f = \sum_{i \in D} (w, h_i)(f, h_i_\nu) \frac{1}{|I|}$$
If $w \in A_2^D$, these are bounded
\[
||\Pi_w : L^2(w) \to L^2(w^{-1})|| = ||\Pi_w : L^2(w) \to L^2(w^{-1})|| \leq ||w||_{BMO^D(w)}^2 w_{A_2}^2.
\]
Recall the decomposition
\[
f_w = \Pi_w f + \Pi_w f + \Pi_w f
\]
and note that the map $f \mapsto f_w$ is an isometry $L^2(w) \to L^2(w^{-1})$. So
\[
\Pi_f w = \sum_{I \in D} (f, h_I)(w)h_I
\]
is bounded $L^2(w) \to L^2(w^{-1})$:
\[
||\Pi_f w||_{L^2(w^{-1})} \leq (1 + 2||\Pi_w : L^2(w) \to L^2(w^{-1})||) ||f||_{L^2(w)}.
\]
Now look at the $L^2(w)$-norm of the dyadic square function $S_D f := (\sum_j (f, h_j)^2 h_j^{1/2}$:
\[
||S_{D_f}||_{L^2(w)}^2 = \sum_{I \in D} (f, h_I^2) = \left( f, \sum_{I \in D} (f, h_I)(w)h_I \right) = (f, \Pi_f w) \leq ||f||_{L^2(w)} ||\Pi_f w||_{L^2(w)}.
\]
so
\[
||S_{D_f}||_{L^2(w)}^2 \leq (1 + 2||\Pi_w : L^2(w) \to L^2(w^{-1})||) ||f||_{L^2(w)}^2.
\]
Since
\[
\frac{1}{2} \leq \frac{||\Pi_w : L^2(w) \to L^2(w^{-1})||}{||w||_{BMO(w)}},
\]
we have further that
\[
||S_{D_f}||_{L^2(w)}^2 \leq ||f||_{L^2(w)}^2 \left( 2 ||\Pi_w : L^2(w) \to L^2(w^{-1})|| \left( \frac{||\Pi_w : L^2(w) \to L^2(w^{-1})||}{||w||_{BMO(w)}} + \frac{4 ||\Pi_w : L^2(w) \to L^2(w^{-1})||}{||w||_{BMO(w)}} \right) \right),
\]
which yields
\[
\frac{||S_{D_f}||_{L^2(w)}}{||f||_{L^2(w)}} \leq \sqrt{6} \left( \frac{||\Pi_w : L^2(w) \to L^2(w^{-1})||}{||w||_{BMO(w)}} \right)^{1/2} \leq \sqrt{6} \sup_{b \in BMO^D(w)} \left( \frac{||\Pi_b : L^2(w) \to L^2(w^{-1})||}{||b||_{BMO(w)}} \right)^{1/2}.
\]
Finally, the fact that
\[
\sup_{b \in BMO^D(w)} \left( \frac{||\Pi_b : L^2(w) \to L^2(w^{-1})||}{||b||_{BMO(w)}} \right) \geq \frac{1}{6} ||S_D : L^2(w) \to L^2(w)||^2 \approx [w]_{A_2}
\]
shows that any smaller bound in (3.1) would imply a bound for $||S_D : L^2(w) \to L^2(w)||$ smaller than $[w]_{A_2}$, which is well-known to be false.

Going back to (3.2), it is easy to show that
\[
1_Q (b - \langle b \rangle_Q) = 1_Q (\Pi_b 1_Q - \Pi_b 1_Q), \forall Q \in D.
\]
Then
\[
\frac{1}{w(Q)} \int_Q |b - \langle b \rangle_Q| dx = \frac{1}{w(Q)} \int_Q |\Pi_b 1_Q - \Pi_b 1_Q| dx \leq \frac{1}{w(Q)} \left( \left( \int_Q |\Pi_b 1_Q|^2 dw^{-1} \right)^{1/2} w(Q)^{1/2} + \left( \int_Q |\Pi_b 1_Q|^2 dw^{-1} \right)^{1/2} w(Q)^{1/2} \right) \leq \frac{1}{w(Q)^{1/2}} 2 ||\Pi_b : L^2(w) \to L^2(w^{-1})|| ||1_Q||_{L^2(w)},
\]
which gives us
\[
||b||_{BMO^D(w)} \leq 2 ||\Pi_b : L^2(w) \to L^2(w^{-1})||, \forall b \in BMO^D(w).
\]
Now we proceed with the proof of Theorem 3.1 focusing on $\Pi_b$, with the other paraproducts following similarly.

### 3.2. Maximal Truncation of Paraproducts.

Let $b \in BMO_D(\mathbb{R}^n)$. Define the maximal truncation of the paraproduct $\Pi_b$:

$$\Pi_b f(x) := \sup_{P \in D} \left| \sum_{Q \supset P} (b, h_Q)(f) Q h_Q(x) \right|.$$  

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

**Proposition 3.4.** Suppose $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded linear or sublinear operator. If $T$ satisfies

$$\text{supp}(Th_Q) \subset Q, \forall Q \in D,$$

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

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

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

Now we prove some properties of $\Pi_b$.

**Proposition 3.5.** The maximal truncation defined above satisfies the following:

i. $\Pi_b$ dominates $\Pi_b$:

$$|\Pi_b f(x)| \leq \Pi_b f(x), \forall x \in \mathbb{R}^n.$$

ii. $\Pi_b$ is dominated by $M^D \Pi_b$:

$$\Pi_b f(x) \leq M^D(\Pi_b f)(x), \forall x \in \mathbb{R}^n.$$

iii. $\Pi_b$ is strong $(2,2)$:

$$\|\Pi_b f\|_{L^2 \rightarrow L^2} \leq \|b\|_{BMO} \|f\|_{L^2}.$$  

iv. $\Pi_b$ is weak $(1,1)$:

$$|[x \in \mathbb{R}^n : \Pi_b f(x) > \alpha]| \leq \frac{C_n}{\alpha} \|f\|_1.$$

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

$$\Pi_b f(x) = \sum_{Q \in D} (b, h_Q)(f) Q h_Q(x) = \sum_{k \in \mathbb{Z}} (b, h_Q_k)(f) Q_k h_Q_k(x),$$

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

$$\sum_{k > m} (b, h_{Q_k})(f) Q_k h_{Q_k}(x) \leq \Pi_b f(x).$$

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

ii. Let $P \in D$ and define $F_P(x) := \sum_{Q \supset P} (b, h_Q)(f) Q h_Q(x)$. 

If \( x \in P \), then \( |F_P(x)| = |(\Pi_b f)_{P} 1_{P}(x) | \), so
\[
|F_P(x)| \leq |(\Pi_b f)_{P} 1_{P}(x) | \leq M^{D} \Pi_b f(x).
\]
If \( x \notin P \), then there is a unique \( k \geq 0 \) such that \( x \in P^{(k+1)} \setminus P^{(k)} \).
So, there is a unique
\[
P_0 \in \left( P^{(k+1)} \right)_{1}, \quad P_0 \neq P^{(k)}
\]
such that \( x \in P_0 \). Then:
\[
F_P(x) = (b, h_{P^{(k+1)}}) f_{P^{(k+1)}}(x) + \sum_{Q \subseteq P_0} (b, h_{Q}) f_{Q} h_{Q}(P^{(k+1)}) = h_{Q}(P_0)
\]
\[
= (\Pi_b f)_{P_0} 1_{P_0}(x),
\]
so once again \( |F_P(x)| \leq M^{D} \Pi_b f(x) \). 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 \( \Pi_b \) in the unweighted case:
\[
\| \Pi_b \|_{L^2} \leq |\Pi^{D} \Pi_b f|_{L^2} \leq \| \Pi_b f \|_{L^2} \leq \| b \|_{BMO} \| f \|_{L^2}.
\]

\( iv \). Once we verify \( \text{supp}(\Pi_{b}) \subset Q \) for all \( Q \in \mathcal{D} \), we use \( iii \) and Proposition 3.4 to conclude \( iv \).

\[
\Pi_{b} h_{Q}(x) = \sup_{P \in \mathcal{D}} | \sum_{K \subseteq P} (b, h_{R}) (h_{Q})_R h_{R}(x) |
\]
\[
= \sup_{P \in Q} | \sum_{K \subseteq P, P \subseteq \mathcal{Q}} (b, h_{R}) h_{Q} h_{R}(x) |
\]
which is clearly 0 if \( x \notin Q \).

3.3. Proof of Theorem 3.1

Proof. I. The BMO decomposition. We make use of the following modification to the Calderón-Zygmund decomposition used in [5] to essentially reduce a weighted BMO function to a regular BMO function. Given a weight \( w \) on \( \mathbb{R}^n \), a function \( b \in BMO^{D}(w) \), a fixed dyadic cube \( Q_0 \in \mathcal{D} \), and \( \epsilon \in (0, 1) \), let the collection:
\[
\mathcal{E} := \{ \text{maximal subcubes } R \subset Q_0 \text{ s.t. } \langle w \rangle_{R} > \frac{2}{\epsilon} \langle w \rangle_{Q_0} \}
\]
and put
\[
\mathcal{E} := \bigcup_{R \in \mathcal{E}} R.
\]
This is the collection from the usual CZ-decomposition of \( w \), restricted to \( Q_0 \), so we have
\[
\sum_{R \in \mathcal{E}} |R| < \frac{\epsilon}{2} |Q_0|.
\]
But instead of defining the usual “good function” for \( w \), we let
\[
a := 1_{Q_0}(x)b(x) - \sum_{R \in \mathcal{E}} (b(x) - \langle b \rangle_{R}) 1_{R}(x) = \sum_{Q \subset Q_0, Q \in \mathcal{E}} (b, h_{Q}) h_{Q}.
\]
As shown in [5], this function is in unweighted \(BMO\), with:
\[
a \in BMO^D; \quad \|a\|_{BMO^D} \leq \frac{4}{\epsilon} (w)_{Q_0} \|b\|_{BMO^D(w)}.
\]

Moreover,
\[
\forall Q \in D(Q_0), \quad Q \notin E : \langle a \rangle_Q = \langle b \rangle_Q \text{ and } (a, h_Q) = (b, h_Q),
\]
so whenever dealing with a cube \(Q \notin E\), 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 \(BMO\). This has many advantages, since any usage of inequalities involving \(a\) 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):
\[
\|\langle a, h_Q \rangle\| \leq \sqrt{|Q|} \|a\|_{BMO^D}.
\]

It also allows us to use the results on \(\Pi^b_0 f\) 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 \(\epsilon\), such that the set:
\[
F := \{ x \in Q_0 : \Pi^b_0 f(x) > C_0 \|a\|_{BMO^D}(\langle f \rangle)_{Q_0} \} \cup \{ x \in Q_0 : M^D_{\Pi^b_0 f}(x) > C_0 \langle f \rangle_{Q_0} \}
\]
satisfies
\[
|F| < \frac{\epsilon}{2} |Q_0|,
\]
where \(M^D_{\Pi^b_0 f}\) denotes the dyadic maximal function restricted to \(Q_0\), i.e. \(M^D_{\Pi^b_0 f}(x) = \sup_{Q \subset Q_0, Q \notin E} \langle f \rangle_{Q_0} 1_Q(x)\). Let then the collection
\[
\mathcal{T} := \{ \text{maximal subcubes of } Q_0 \text{ contained in } F \}. 
\]

First use the well-known weak (1, 1) inequality for the dyadic maximal function:
\[
|\{ x \in \mathbb{R}^n : M^D \phi(x) > \alpha \}| \leq \frac{C_1(n)}{\alpha} \|\phi\|_1,
\]
applied to \(\phi = f 1_{Q_0}\). For all \(x \in Q_0\), \(M^D(f 1_{Q_0})(x) = M^D_{\Pi^b_0 f}(x)\), so
\[
|\{ x \in Q_0 : M^D_{\Pi^b_0 f}(x) > C_0 \langle f \rangle_{Q_0} \}| \leq \frac{C_1}{C_0} |Q_0|.
\]

Since \(a \in BMO^D\) we can apply the weak (1, 1) inequality for \(\Pi^b_0\), according to Proposition 3.5:
\[
|\{ x \in \mathbb{R}^n : \Pi^b_0 \phi(x) > \alpha \}| \leq \frac{C_2(n)}{\alpha} \|a\|_{BMO^D} \|\phi\|_1,
\]
and let again \(\phi = f 1_{Q_0}\). By the definition of \(a\), in this case, \(\Pi^b_0 f\) sums only over \(Q \subset Q_0\), so regardless of \(x\) we have \(\Pi^b_0 \phi = \Pi^b_0 (f 1_{Q_0})\). Same holds for \(\Pi^b_0\):
\[
\Pi^b_0 \phi(x) = \sup_{P \in D} \sum_{Q^0 \subset P} (a, h_Q)(\phi) h_Q(x) = \sup_{P \in Q_0} \sum_{Q^0 \subset P, Q^0 \subset Q} (b, h_Q)(\phi) h_Q(x),
\]
so
\[
|\{ x \in Q_0 : \Pi^b_0 f(x) > C_0 \|a\|_{BMO^D}(\langle f \rangle)_{Q_0} \}| = |\{ x \in Q_0 : \Pi^b_0 (f 1_{Q_0})(x) > C_0 \|a\|_{BMO^D}(\langle f \rangle)_{Q_0} \}| \leq \frac{C_2}{C_0} \|a\|_{BMO^D} \| f 1_{Q_0} \|_1 = \frac{C_2}{C_0} |Q_0|.
\]

Then, as we wished,
\[
|F| \leq \frac{C_1 + C_2}{C_0} |Q_0| < \frac{\epsilon}{2} |Q_0|,
\]
if we choose $C_0$ large enough:

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

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

(3.3) $$\left| \bigcup_{R \in \mathcal{G}} R \right| < \epsilon |Q_0|$$

We show that:

(3.4) $$1_{Q_0}(x)\Pi_{b,Q_0} f(x) \leq 2C_0\|b\|_{BMO^p(\mathcal{E})} 1_{Q_0}(x) + \sum_{R \in \mathcal{G}} 1_R(x)\Pi_{b,R} f(x).$$

Since $\|b\|_{BMO^p} \leq \frac{4}{\epsilon} (w)_{Q_0}\|b\|_{BMO^p(\mathcal{E})}$, this yields

$$1_{Q_0}(x)\Pi_{b,Q_0} f(x) \leq \frac{C_0}{\epsilon} (w)_{Q_0}\|b\|_{BMO^p(\mathcal{E})} 1_{Q_0}(x) + \sum_{R \in \mathcal{G}} 1_R(x)\Pi_{b,R} f(x).$$

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 $\{R' \subset R \text{ satisfying } |R'| < \epsilon |R| \}$ and

$$1_R\Pi_{b,R} f(x) \leq \frac{C_0}{\epsilon} (w)_{Q_0}\|b\|_{BMO^p(\mathcal{E})} 1_{Q_0}(x) + \sum_{R'} 1_{R'}\Pi_{b,R'} f(x).$$

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

$$\Pi_{b,Q_0} f(x) \leq \frac{C_0}{\epsilon} \|b\|_{BMO^p(\mathcal{E})} \sum_{Q \subset S(Q_0)} \langle w \rangle_{Q} |f| Q_0 1_Q(x).$$

Recall that $C_0 \sim \frac{C(n)}{\epsilon}$:

$$\Pi_{b,Q_0} f(x) \leq \frac{C(n)}{\epsilon^2} \|b\|_{BMO^p(\mathcal{E})} \mathcal{A}_S (Q_0) |f| (x).$$

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

$$\sum_{P \in ch_b(Q_0)} |P| < \epsilon |Q|, \forall Q \in \mathcal{S}(Q_0),$$

so $\mathcal{S}(Q_0)$ is $\frac{1}{\epsilon}$-Carleson. So we choose $\epsilon = \frac{1}{\Lambda - 1}$ and we have the desired sparse collection with Carleson constant $\Lambda$ such that

$$\Pi_{b,Q_0} f(x) \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^2 \|b\|_{BMO^p(\mathcal{E})} \mathcal{A}_S (Q_0) |f| (x).$$

**III. Proof of (3.4).** We start by noting that

$$\Pi_{b,Q_0} f(x) = \sum_{P \subset Q_0} (b, h_P)(f) p h_P(x)$$

$$= \sum_{P \subset Q_0} (b, h_P)(f) p h_P(x) + \sum_{R \in \mathcal{G}} \sum_{P \subset R} (b, h_P)(f) p h_P(x),$$

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

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

Now, we have to account for the relationship to the collection $\mathcal{F}$ and its union $F$.

**Case 1:** $x \notin F$. 

In this case, \( \Pi_a f(x) \leq C_0 \|a\|_{BMO^\infty} \langle \|f\| \rangle_{Q_0} \), and since \( \Pi^\circ \) dominates \( \Pi \):

\[
|\Pi_a f(x)| \leq \Pi^\circ f(x) \leq C_0 \|a\|_{BMO^\infty} \langle \|f\| \rangle_{Q_0},
\]

so we have

\[
|\Pi_b Q_0, f(x)| \leq C_0 \|a\|_{BMO^\infty} \langle \|f\| \rangle_{Q_0} + \left| \sum_{R \in \mathcal{E}} \Pi_{b,R} f(x) \right|.
\]

- **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 \( R \supseteq R_0, R \in \mathcal{F} \). 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

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

  which gives (3.4) in this case.

- **Case 1b**: If \( x \notin E \), then the second part of the sum is 0 and we are done, having simply

  \[
  |\Pi_{b,Q_0} f(x)| \leq C_0 \|a\|_{BMO^\infty} \langle \|f\| \rangle_{Q_0}.
  \]

**Case 2**: \( x \in F \).

Then there is a *unique* \( P \in \mathcal{F} \) such that \( x \in P \). Look first at the term \( \Pi_a f(x) = \sum_{Q \subset F} (a, h_Q)(f) Q h_Q(x) \).

Since \( x \in P \), this can be expressed as

\[
\Pi_a f(x) = \sum_{Q \supset \hat{P}} (a, h_Q)(f) Q h_Q(x) + \sum_{Q \subset P} (a, h_Q)(f) Q h_Q(x),
\]

where \( \hat{P} \) denotes the dyadic parent of \( P \). The first term we split into two:

\[
|\sum_{Q \supset \hat{P}} (a, h_Q)(f) Q h_Q(x)| \leq |\sum_{Q \supset \hat{P}} (a, h_Q)(f) Q h_Q(x)| + |(a, h_{\hat{P}})(f) h_{\hat{P}}(x)|.
\]

\[
= A(x) + B
\]

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

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

  \[
  A \leq C_0 \|a\|_{BMO^\infty} \langle \|f\| \rangle_{Q_0}.
  \]
Let us now look at the term $B$. If $\hat{P} \subset E$, then $B = 0$. Otherwise, since $|(a, h_Q)| \lesssim \sqrt{|Q|}\|a\|_{BMO^D}$,

$$B \leq \sqrt{|\hat{P}|}\|a\|_{BMO^D}|\langle f \rangle|_{\hat{P}} \frac{1}{\sqrt{|\hat{P}|}} = \|a\|_{BMO^D}|\langle f \rangle|_{\hat{P}}.$$

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

So

$$\sum_{Q \supset P} (a, h_Q)(f)Qh_Q(x) \lesssim C_0\|a\|_{BMO^D}|\langle f \rangle|_{Q_0},$$

giving us that

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

where the term $C$ is defined as

$$C := \sum_{Q \supset P, Q \in E} (b, h_Q)(f)Qh_Q(x) + \sum_{R \in \mathcal{E}} \Pi_{b, R}f(x).$$

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

$$C = \left[ \sum_{Q \supset P, Q \in E} (b, h_Q)(f)Qh_Q(x) + \sum_{R \in \mathcal{E}, R \subset P} \Pi_{b, R}f(x) \right] + \sum_{R \in \mathcal{E}, R \supset P} \Pi_{b, R}f(x) = \Pi_{b, P}f(x) = \Pi_{b, R_0}f(x).$$

- **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 3.3.** One can also use Theorem 3.1 to obtain a full $\mathbb{R}^n$ domination, losing the requirement for no infinitely increasing chains. Say $f$ is such that $\text{supp}(f) \subset Q_0$ for some $Q_0 \in \mathcal{D}$ (or, for general compactly supported functions, $\text{supp}(f)$ is contained in at most $2^n$ disjoint $Q_k \in \mathcal{D}$). Then

$$\Pi_{b}f(x) = \Pi_{b, Q_0}f(x) + \left( \sum_{Q \supset Q_0} (b, h_Q) \frac{1}{|Q|} h_Q(x) \right) \int_{Q_0} f.$$

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

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

To see this, let $Q \in \mathcal{D}$ and apply the decomposition to $b$ over $Q$:

$$E := \{\text{maximal subcubes } R \subset Q_0 \text{ s.t. } \langle w \rangle_R > 2\langle w \rangle_Q\}; \quad E := \bigcup_{R \in \mathcal{E}} R;$$

$$a := \sum_{R \subset Q_0, R \in \mathcal{E}} (b, h_Q) \in BMO^D \text{ with } \|a\|_{BMO^D} \leq 4\langle w \rangle_Q \|b\|_{BMO^D(w)}.$$

Since $Q$ itself is not selected for $E$, $Q \notin E$, so $(a, h_Q) = (b, h_Q)$. Finally, then:

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

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

$$\Pi_{b}f(x) = \left( \sum_{Q \supset Q_0^{(k)}} (b, h_Q) \frac{1}{|Q|} h_Q(x) \right) \int_{Q_0^{(k)}} f.$$
By Theorem 3.1, there is a $\Lambda$ so

$$\sum_{Q \supset Q_0} |(b, h_Q)| \frac{1}{|Q|} \frac{1}{\sqrt{|Q|}} (\int_{Q_0} |f|)$$

$$\leq \sum_{Q \supset Q_0} \langle w \rangle_Q \| b \|_{BMO^w} \frac{1}{|Q|} \int_{Q_0} |f|$$

$$= \| b \|_{BMO^w} \sum_{Q \supset Q_0} \langle w \rangle_Q (|f|)_Q.$$ 

Then

$$|\Pi_b f(x)| \leq \sum_{Q \supset Q_0} |(b, h_Q)| \frac{1}{|Q|} \frac{1}{\sqrt{|Q|}} (\int_{Q_0} |f|)$$

$$\leq \sum_{Q \supset Q_0} \langle w \rangle_Q \| b \|_{BMO^w} \frac{1}{|Q|} \int_{Q_0} |f|$$

$$= \| b \|_{BMO^w} \sum_{Q \supset Q_0} \langle w \rangle_Q (|f|)_Q.$$ 

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

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

so

$$|\Pi_b f(x)| \leq |\Pi_{b, Q_0} f(x)| + \| b \|_{BMO^w} \sum_{Q \supset Q_0} \langle w \rangle_Q (|f|)_Q.$$ 

By Theorem 3.1 there is a $\Lambda$-Carleson sparse collection $S(Q_0)$ such that

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

So form a sparse collection $S$ as follows:

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

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

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

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

$$x \notin Q_0 : |\Pi_b f(x)| \leq \| b \|_{BMO^w} \mathcal{A}^w_S f(x);$$

$$x \in Q_0 : |\Pi_b f(x)| \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^2 \| b \|_{BMO^w} \mathcal{A}^w_S f(x).$$

So indeed

$$|\Pi_b f(x)| \leq \| b \|_{BMO^w} \mathcal{A}^w_S f(x), \forall x \in \mathbb{R}^n,$$

for all compactly supported $f$.

**Remark 3.4.** If we let $f \equiv 1$ in $\Pi_{b, Q_0} f$, we have

$$\Pi_{b, Q_0} 1(x) = \sum_{Q \supset Q_0} (b, h_Q) h_Q(x) = (b(x) - (b)_{Q_0}) 1_{Q_0}.$$ 

So, applying Theorem 3.1 to the function $f \equiv 1$ essentially gives us that local mean oscillations of functions in $BMO^D(w)$ can be dominated by one of the sparse BMO functions in Section 2.2

**Corollary 3.6.** There is a dimensional constant $C(n)$ such that for all $\Lambda > 1$, weights $w$ on $\mathbb{R}^n$, $b \in BMO^D(w)$ and $Q_0 \in \mathcal{D}$, there is a $\Lambda$-Carleson sparse collection $S(Q_0) \subset \mathcal{D}(Q_0)$ such that

$$|(b(x) - (b)_{Q_0}) 1_{Q_0}(x)| \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^2 \| b \|_{BMO^w} \sum_{Q \supset S(Q_0)} \langle w \rangle_Q 1_{Q}(x)$$

$$= C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^2 \| b \|_{BMO^w} b^w_{S(Q_0)}(x).$$
Appendix A. Proof of Theorem 2.3

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

$$b_S := \sum_{Q \in S} 1_Q,$$

and we wish to show that

$$\|b_S\|_{BMO} \leq \Lambda,$$

where $\Lambda$ is the Carleson constant of $S$.

Proof. Let $Q_0 \in D$ be fixed. We wish to estimate $\frac{1}{|Q_0|} \int_{Q_0} |b - \langle b \rangle_{Q_0}| \, dx$, and recall that

$$(b_S - \langle b_S \rangle_{Q_0})_{Q_0} = \sum_{Q \in S, Q \subset Q_0} 1_Q - (\tau_S)_{Q_0} 1_{Q_0},$$

where $(\tau_S)_P := \frac{1}{|P|} \sum_{Q \in S, Q \subset P} |Q| \leq \Lambda \forall P \in D$.

In fact,

If $P \in S$, then $(\tau_S)_P \leq \Lambda - 1$.

With $Q_0 \in D$ fixed, here we are only looking at $S(Q_0) := \{Q \in S : Q \subset Q_0\}$. We define the collections as sets:

$S_1 := \text{ch}_S(Q_0)$ (the $S$-children of $Q_0$) and $S_1 := \bigcup_{Q \in S_1} Q$;

$S_2 := \{Q_2 \in \text{ch}_S(Q_1) : Q_1 \in S_1\}$ and $S_2 := \bigcup_{Q \in S_2} Q$, so $S_2$ are the "$S$-grandchildren" of $Q_0$, the second generation of $S$-cubes in $Q_0$. Generally,

$S_k := \{Q_k \in \text{ch}_S(Q_{k-1}) : Q_{k-1} \in S_{k-1}\}$ and $S_k := \bigcup_{Q \in S_k} Q$.

Note that:

- Each $S_k$ is a disjoint union of $Q_k \in S_k$, as each $S_k$ is a pairwise disjoint collection.
- The sets $S_k$ satisfy $Q_0 \supset S_1 \supset S_2 \supset \ldots$
- Moreover

$$\bigcap_{k=1}^\infty S_k = \emptyset,$$

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

For ease of notation, denote for now

$$\theta := (\tau_S)_{Q_0} = \frac{1}{|Q_0|} \sum_{Q \in S, Q \subset Q_0} |Q| \leq \Lambda.$$

We have:

$$\frac{1}{|Q_0|} \int_{Q_0} |b_S - \langle b_S \rangle_{Q_0}| \, dx = \frac{1}{|Q_0|} \int_{Q_0} \left\| \sum_{Q \in S, Q \subset Q_0} 1_Q - \theta \right\| \, dx$$

$$= \frac{1}{|Q_0|} \int_{Q_0 \setminus S_1} |\theta| \, dx + \frac{1}{|Q_0|} \int_{S_1} \left\| \sum_{Q \in S, Q \subset Q_0} 1_Q - \theta \right\| \, dx$$
Since $S_1$ is a disjoint union of $Q_k \in S_1$:

\[
\frac{1}{|Q_0|} \int_{S_1} \sum_{Q \in S, Q \subseteq Q_0} 1_Q(x) - \theta \, dx = \frac{1}{|Q_0|} \sum_{Q \in S} \int_{Q} \sum_{Q \in S, Q \subseteq Q_0} 1_Q(x) - \theta \, dx \\
= \frac{1}{|Q_0|} \sum_{Q \in S} \int_{Q} \sum_{Q \in S, Q \subseteq Q_0} 1_Q(x) + 1 - \theta \, dx \\
= \frac{1}{|Q_0|} \left[ \sum_{Q \in S} \left( \int_{Q} (1 - \theta) \, dx + \sum_{Q \in S, Q \subseteq Q_0} \int_{Q} \sum_{Q \in S} 1_Q(x) + 1 - \theta \, dx \right) \right] \\
= \frac{1}{|Q_0|} [1 - \theta] \sum_{Q \in S, Q \subseteq Q_0} \left[ |Q_1 \setminus S_2| + \frac{1}{|Q_0|} \sum_{Q \in S} \int_{Q} \sum_{Q \in S} 1_Q(x) + 2 - \theta \, dx \right].
\]

So

\[
\frac{1}{|Q_0|} \int_{Q_0} |b_S - b_{S(Q)}| \, dx = \theta \frac{|Q_0 \setminus S_1|}{|Q_0|} + [1 - \theta] \frac{|S_1 \setminus S_2|}{|Q_0|} + \frac{1}{|Q_0|} \sum_{Q \in S} \int_{Q} \sum_{Q \in S} 1_Q(x) + 2 - \theta \, dx.
\]

We can apply the same reasoning to each $Q_2 \in S_2$:

\[
\int_{Q_2} \sum_{Q \in S} 1_Q(x) + 2 - \theta \, dx = \int_{Q_2 \setminus S_3} |2 - \theta| \, dx + \sum_{Q \in S} \sum_{Q \subseteq Q_2} \int_{Q} \sum_{Q \subseteq Q_0} 1_Q(x) + 3 - \theta \, dx,
\]

and we can conclude inductively

\[
(A.1) \quad \frac{1}{|Q_0|} \int_{Q_0} |b_S - b_{S(Q)}| \, dx = \theta \frac{|Q_0 \setminus S_1|}{|Q_0|} + [1 - \theta] \frac{|S_1 \setminus S_2|}{|Q_0|} + \frac{1}{|Q_0|} \sum_{Q \in S} \left[ |S_2 \setminus S_3| + \ldots \right].
\]

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

\[
\theta \frac{|Q_0 \setminus S_1|}{|Q_0|} + (1 - \theta) \frac{|S_1 \setminus S_2|}{|Q_0|} + (2 - \theta) \frac{|S_2 \setminus S_3|}{|Q_0|} + \ldots \\
= \frac{1}{|Q_0|} \left( \theta |Q_0 \setminus S_1| + (1 - \theta) |S_1 \setminus S_2| + (2 - \theta) |S_2 \setminus S_3| + (3 - \theta) |S_3 \setminus S_4| + \ldots \right) \\
= \frac{1}{|Q_0|} \left( \theta |Q_0 \setminus S_1| + (1 - \theta) |S_1| + |S_2| + |S_3| + \ldots \right).
\]

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

\[
a_k = c |S_1 \setminus S_2| + (c + 1)|S_2 \setminus S_3| + \ldots + (c + k - 1)|S_{k+1} \setminus S_{k+2}| \\
= c |S_1| - c |S_2| + (c + 1)|S_2| - (c + 1)|S_3| + \ldots + (c + k - 1)|S_{k+1}| - (c + k - 1)|S_{k+2}| \\
= c |S_1| + |S_2| + |S_3| + \ldots + |S_k| - (c + k - 1)|S_{k+1}|,
\]

where $c = (1 - \theta) \geq 0$. 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:

\[
\sum_{k=1}^{\infty} k|S_k| = |S_1| + 2|S_2| + 3|S_3| + \ldots \\
= |S_1| + |S_2| + |S_3| + \ldots \leq \Lambda |S_1| \\
+ |S_2| + |S_3| + \ldots \leq \Lambda |S_2| \\
+ |S_3| + |S_4| + \ldots \leq \Lambda |S_3| \\
+ \ldots \leq \Lambda (|S_1| + |S_2| + \ldots) \leq \Lambda^2 |S_1|.
\]
So
\[ \lim_{k \to \infty} a_k = c|S_1| + \sum_{k=2}^{\infty} |S_k| - \lim_{k \to \infty} (c + k - 1)|S_{k+1}| = 0 \]
and
\[ \frac{1}{|Q_0|} \int_{Q_0} |b_S - \langle b_S \rangle_{Q_0}| dx = \frac{1}{|Q_0|} \left( \theta|Q_0 \setminus S_1| + (1 - \theta)|S_1| + |S_2| + |S_3| + \ldots \right) \]
holds.

Now,
\[ |S_2| + |S_3| + \ldots = \sum_{Q \in S_1} \left( \sum_{Q_0 \subseteq S_0} \frac{|Q|}{|Q_0|} \right) \leq (\Lambda - 1) \sum_{Q \in S_1} |Q_0| = (\Lambda - 1)|S_1|. \]

So
\[ \frac{1}{|Q_0|} \int_{Q_0} |b_S - \langle b_S \rangle_{Q_0}| dx \leq \frac{1}{|Q_0|} \left( \theta|Q_0 \setminus S_1| + (1 - \theta)|S_1| + (\Lambda - 1)|S_1| \right) \]
\[ = \frac{1}{|Q_0|} \left( \theta|Q_0 \setminus S_1| + (\Lambda - \theta)|S_1| \right) \]
\[ \leq \frac{1}{|Q_0|} \left( \Lambda |Q_0 \setminus S_1| + \Lambda |S_1| \right) \]
\[ = \frac{\Lambda}{|Q_0|} (|Q_0 \setminus S_1| + |S_1|) \]
\[ = \Lambda. \]

Generally, if \( n < \theta \leq (n + 1) \) for some \( n \in \mathbb{N} \): the right hand side of (A.1) becomes
\[ \frac{1}{|Q_0|} \left[ \theta|Q_0 \setminus S_1| + (\theta - 1)|S_1 \setminus S_2| + \ldots + (\theta - n)|S_n \setminus S_{n+1}| + (n + 1 - \theta)|S_{n+1} \setminus S_1| + (n + 2 - \theta)|S_{n+2} \setminus S_2| + \ldots \right] \]
\[ \leq \frac{1}{|Q_0|} \left[ \theta|Q_0 \setminus S_1| + (\theta - 1)|S_1 \setminus S_2| + \ldots + (\theta - n)|S_n \setminus S_{n+1}| + (\Lambda + n - \theta)|S_{n+1}| \right] \]
\[ \leq \frac{1}{|Q_0|} \left[ \Lambda |Q_0 \setminus S_1| + \Lambda |S_1 \setminus S_2| + \ldots + \Lambda |S_n \setminus S_{n+1}| + \Lambda |S_{n+1}| \right] \]
\[ \leq \frac{\Lambda}{|Q_0|} \left( |Q_0 \setminus S_1| + |S_1 \setminus S_2| + \ldots + |S_n \setminus S_{n+1}| + |S_{n+1}| \right) \]
\[ = \Lambda. \]

\[ \square \]

**Appendix B. Proof of Theorem 2.5**

Say we have \( a \in BMO^D(\mathbb{R}^n) \), \( b \in BMO^D(w) \) where \( w \) is a weight on \( \mathbb{R}^n \), and a fixed \( Q_0 \in \mathcal{D} \). We look at
\[ \Pi^*_a \Pi_{b,Q_0} f := \sum_{Q \subset Q_0} (a, h_Q)(b, h_Q)(f)_Q \frac{1_Q}{|Q|} \]
and the inner product
\[ (\Pi^*_a \Pi_{b,Q_0} f, g) = \sum_{Q \subset Q_0} (a, h_Q)(b, h_Q)(f)_Q (g)_Q. \]
Within $Q_0$ we form the local CZ-decompositions of $f$ and $g$, and the BMO decomposition of $b$: 

$$\begin{align*}
\mathcal{E}_1 &:= \{\text{maximal subcubes } R \in \mathcal{D}(Q_0) \text{ s.t. } \langle f \rangle_R > \frac{3}{\epsilon} \langle f \rangle_{Q_0} \}; \quad E_1 := \cup_{R \in \mathcal{E}_1} R; \\
\mathcal{E}_2 &:= \{\text{maximal subcubes } R \in \mathcal{D}(Q_0) \text{ s.t. } \langle g \rangle_R > \frac{3}{\epsilon} \langle g \rangle_{Q_0} \}; \quad E_2 := \cup_{R \in \mathcal{E}_2} R; \\
\mathcal{E}_3 &:= \{\text{maximal subcubes } R \in \mathcal{D}(Q_0) \text{ s.t. } \langle w \rangle_R > \frac{3}{\epsilon} \langle w \rangle_{Q_0} \}; \quad E_3 := \cup_{R \in \mathcal{E}_3} R.
\end{align*}$$

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

$$\tilde{b} := 1_{Q_0} b - \sum_{R \in \mathcal{E}_3} (b - \langle b \rangle_R) 1_R = \sum_{Q \subset Q_0, Q \neq E_3} (b, h_Q) 1_Q,$$

which satisfies $\tilde{b} \in \text{BMO}^D(\mathbb{R}^n)$ with 

$$||\tilde{b}||_{\text{BMO}^D} \leq \frac{6}{\epsilon} \langle w \rangle_{Q_0} ||b||_{\text{BMO}^D(w)}.$$ 

Moreover, $(b, h_Q) = (\tilde{b}, h_Q)$ for all $Q \subset Q_0, Q \notin E_3$. Each collection $\mathcal{E}_i$ satisfies 

$$\sum_{R \in \mathcal{E}_i} |R| \leq \epsilon |Q_0|.$$ 

Finally, let 

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

Then 

$$\sum_{R \in \mathcal{E}} |R| \leq \epsilon |Q_0|.$$ 

Now look at $(\Pi_n^1 \Pi_{b}^f, g)$ and split the sum as 

\begin{equation}
(B.1) \quad |(\Pi_n^1 \Pi_{b}^f, g)| \leq \sum_{Q \subset Q_0 \cap E} |(a, h_Q)| |(b, h_Q)| \langle f \rangle_{Q} \langle g \rangle_{Q} + \sum_{R \in \mathcal{E}} |(\Pi_n^1 \Pi_{b}^f, g)|.
\end{equation}

For every $Q \subset Q_0, Q \notin E$, we have: 

$$\langle f \rangle_{Q} \leq \frac{3}{\epsilon} \langle f \rangle_{Q_0}, \quad \langle g \rangle_{Q} \leq \frac{3}{\epsilon} \langle g \rangle_{Q_0}, \text{ and } (b, h_Q) = (\tilde{b}, h_Q),$$ 

so: 

$$\sum_{Q \subset Q_0 \cap E} |(a, h_Q)| |(b, h_Q)| \langle f \rangle_{Q} \langle g \rangle_{Q} \leq \frac{9}{\epsilon^2} \langle f \rangle_{Q_0} \langle g \rangle_{Q_0} \sum_{Q \subset Q_0 \cap E} |(a, h_Q)| |(\tilde{b}, h_Q)|$$

$$\leq \frac{9}{\epsilon^2} \langle f \rangle_{Q_0} \langle g \rangle_{Q_0} \left( \sum_{Q \subset Q_0} |(a, h_Q)|^2 \right)^{1/2} \left( \sum_{Q \subset Q_0} |(\tilde{b}, h_Q)|^2 \right)^{1/2},$$

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

$$\sum_{Q \subset Q_0 \cap E} |(a, h_Q)| |(b, h_Q)| \langle f \rangle_{Q} \langle g \rangle_{Q} \leq \frac{C(n)}{\epsilon^3} |a||\text{BMO}^D||b||\text{BMO}^D(w) \langle f \rangle_{Q_0} \langle g \rangle_{Q_0} \langle w \rangle_{Q_0} |Q_0|.$$ 

Now we recurse on the $\sum_{R \in \mathcal{E}}$ terms in (B.1) and form $S(Q_0)$ by adding $Q_0$ first, $E$ are the $S$-children of $Q_0$, and so on. The collection $S(Q_0)$ satisfies the $S$-children definition of sparseness, with $\sum_{R \in S(Q_0) \cap E} |R| \leq \epsilon |Q|$ for all $Q \in S(Q_0)$, so it is $\frac{1}{2\epsilon}$-Carleson. So, if we choose $\epsilon = \frac{A}{\Lambda - 1}$, we have 

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

$$= (\mathcal{A}_{S(Q_0)}^{f, g}(w)).$$
We summarize this below:

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

\[
\left| \sum_{Q \subset Q_0} (a, h_Q(b, h_Q(f), g)|_Q \right| \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^3 \|a\|_{BMO^D} \|b\|_{BMO^D(w)}(\mathcal{A}^w_{S(Q_0)}[f], |g|).
\]

\[
\sum_{Q \subset Q_0} \eta_Q(b_Q, f_Q)|_Q \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^3 \|a\|_{BMO^D} \|b\|_{BMO^D(w)}(\mathcal{A}^w_{S(Q_0)}[f], |g|).
\]

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 \( a \in BMO^D \), \( b \in BMO^D(\nu) \). 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 \( a = \sum_k \sum_{Q \subset Q_k} (a, h_Q) \), and then

\[
(P_n^* \Pi_b f, g) = \sum_k (P_n^* \Pi_b Q_k f, g)
\]

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

\[
\left| (P_n^* \Pi_b Q_k f, g) \right| \leq C(n) \left( \frac{\Lambda}{\Lambda - 1} \right)^3 \|a\|_{BMO^D} \|b\|_{BMO^D(\nu)}(\mathcal{A}^\nu_{S(Q_k)}[f], |g|).
\]

Then

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

where \( S := \cup_k S(Q_k) \) is a \( \Lambda \)-Carleson sparse collection in \( \mathcal{T}^D(\mathbb{R}^n) \).

Take now \( f \in L^p(\mu) \) and \( g \in L^r(\nu') \). By a simple application of Hölder’s inequality:

\[
|(\mathcal{A}^\nu_{S}[f], |g|)| \leq \|\mathcal{A}^\nu_{S} : L^p(\mu) \rightarrow L^p(\lambda)\| \|f\|_{L^p(\mu)} \|g\|_{L^r(\nu')}.
\]

Then

\[
\|P_n^* \Pi_b : L^p(\mu) \rightarrow L^p(\lambda)\| \leq C(n) \|a\|_{BMO^D} \|b\|_{BMO^D(\nu)} \sup_{S \subset \mathcal{T}^D(\mathbb{R}^n)} \left( \frac{\Lambda}{\Lambda - 1} \right)^3 \|\mathcal{A}^\nu_{S} : L^p(\mu) \rightarrow L^p(\lambda)\|
\]

holds for all \( a \) with finite Haar expansion, and therefore for all \( a \). This proves Theorem 2.3.
References

[1] Oscar Blasco, *Dyadic BMO, paraproducts and Haar multipliers*, Interpolation theory and applications, Contemp. Math., vol. 445, Amer. Math. Soc., Providence, RI, 2007, pp. 11–18, DOI 10.1090/conm/445/08590. MR2381883

[2] Steven Bloom, *A commutator theorem and weighted BMO*, Trans. Amer. Math. Soc. 292 (1985), no. 1, 103–122, DOI 10.2307/2000172. MR805955

[3] Daewon Chung, *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, DOI 10.1512/iumj.2011.60.4453. MR2997000

[4] David Cruz-Uribe, José María Martell, and Carlos Pérez, *Sharp weighted estimates for classical operators*, Adv. Math. 229 (2012), no. 1, 408–441, DOI 10.1016/j.aim.2011.08.013. MR2854179

[5] Xuan Thinh Duong, Irina Holmes, Ji Li, Brett D. Wick, and Dongyong Yang, *Two weight commutators in the Dirichlet and Neumann Laplacian settings*, J. Funct. Anal. 276 (2019), no. 4, 1007–1060, DOI 10.1016/j.jfa.2018.12.003. MR3906299

[6] Loukas Grafakos, *Classical and modern Fourier analysis*, Pearson Education, Inc., Upper Saddle River, NJ, 2004. MR2449250

[7] Irina Holmes, Michael T. Lacey, and Brett D. Wick, *Commutators in the two-weight setting*, Math. Ann. 367 (2017), no. 1-2, 51–80, DOI 10.1007/s00208-016-1378-1. MR3606434

[8] Tuomas Hytönen, *Dyadic Analysis and Weights; Lecture Notes from a course at University of Helsinki* (2014), available at http://wiki.helsinki.fi/download/attachments/130069291/dyadic.pdf.

[9] Michael T. Lacey, *An elementary proof of the $A_2$ bound*, Israel J. Math. 217 (2017), no. 1, 181–195, DOI 10.1007/s11856-016-1442-x. MR3625108

[10] Andrei K. Lerner and Fedor Nazarov, *Intuitive dyadic calculus: the basics*, Expo. Math. 37 (2019), no. 3, 225–265, DOI 10.1016/j.exmath.2018.01.001. MR4007575

[11] Andrei K. Lerner, Sheldy Ombrosi, and Israel P. Rivera-Ríos, *On pointwise and weighted estimates for commutators of Calderón-Zygmund operators*, Adv. Math. 319 (2017), 153–181, DOI 10.1016/j.aim.2017.08.022. MR3695871

[12] María Cristina Pereyra, *Dyadic harmonic analysis and weighted inequalities: the sparse revolution*, New trends in applied harmonic analysis. Vol. 2, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, [2019] ©2019, pp. 159–239, DOI 10.1007/978-3-030-32353-0. MR431199

[13] ______, *Lecture notes on dyadic harmonic analysis*, Second Summer School in Analysis and Mathematical Physics (Cuernavaca, 2000), Contemp. Math., vol. 289, Amer. Math. Soc., Providence, RI, 2001, pp. 1–60, DOI 10.1090/conm/289/04874. MR1864538

[14] Stefanie Petermichl, *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, DOI 10.1016/S0764-4442(00)00162-2 (English, with English and French summaries). MR1756958

Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
Email address: valeria96@tamu.edu

Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
Email address: irinaholmes@tamu.edu