

2-10-2016

Commutators in the two-weight setting

Irina Holmes

Michael T. Lacey

Brett D. Wick

Washington University in St. Louis, bwick@wustl.edu

Follow this and additional works at: https://openscholarship.wustl.edu/math_facpubs



Part of the [Analysis Commons](#)

Recommended Citation

Holmes, Irina; Lacey, Michael T.; and Wick, Brett D., "Commutators in the two-weight setting" (2016). *Mathematics Faculty Publications*. 41.

https://openscholarship.wustl.edu/math_facpubs/41

This Article is brought to you for free and open access by the Mathematics and Statistics at Washington University Open Scholarship. It has been accepted for inclusion in Mathematics Faculty Publications by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

COMMUTATORS IN THE TWO-WEIGHT SETTING

IRINA HOLMES, MICHAEL T. LACEY[†], AND BRETT D. WICK[‡]

ABSTRACT. Let R be the vector of Riesz transforms on \mathbb{R}^n , and let $\mu, \lambda \in A_p$ be two weights on \mathbb{R}^n , $1 < p < \infty$. The two-weight norm inequality for the commutator $\| [b, R] : L^p(\mu) \mapsto L^p(\lambda) \|$ is shown to be equivalent to the function b being in a BMO space adapted to μ and λ . This is a common extension of a result of Coifman-Rochberg-Weiss in the case of both λ and μ being Lebesgue measure, and Bloom in the case of dimension one.

CONTENTS

1. Introduction and Statement of Main Results	1
2. Notation and Background	3
3. Two-Weight Inequalities for Paraproduct Operators	10
4. Equivalences for Dyadic Bloom BMO	12
5. Two-Weight Inequalities for Commutators with Calderón-Zygmund Operators	14
References	22

1. INTRODUCTION AND STATEMENT OF MAIN RESULTS

The foundational paper of Coifman-Rochberg-Weiss [6] set out a real-variable counterpart to a classical theorem of Nehari [27]. It characterized BMO, the real-variable space of functions with bounded mean oscillation, in terms of commutators with Riesz transforms. Several lines of investigation came out of this work: generalizations to spaces of homogeneous type, [31], multi-parameter extensions, [7, 9, 21, 22], connections to factorization of function spaces, [1, 3, 5, 12], div-curl lemmas, [4, 20] and additional interpretations in operator theory [26, 27, 32].

In 1985, Bloom [2] proved a *two-weight* extension of the Nehari [27] result in one dimension. In particular, for the Hilbert transform,

$$Hf(x) := \text{p.v.} \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(y)}{x-y} dy,$$

a choice of $1 < p < \infty$, and two weights μ and λ in Muckenhoupt's A_p class, (see Section 2.3 for the definitions of these weights), the commutator $[b, H](f) = bHf - H(bf)$ is bounded

2000 *Mathematics Subject Classification*. Primary: 42, 42A, 42B, 42B20, 42B25, 42A50, 42A40,

Key words and phrases. Commutators, Calderón-Zygmund Operators, Bounded Mean Oscillation, Weights.

[†] Research supported in part by a National Science Foundation DMS grant #1265570.

[‡] Research supported in part by a National Science Foundation DMS grant #0955432.

from $L^p(\mu)$ to $L^p(\nu)$ if and only if the function b satisfies

$$\|b\|_{BMO(\nu)} := \sup_Q \left(\frac{\int_Q |b - \langle b \rangle_Q| dx}{\int_Q \nu dx} \right) < \infty,$$

where $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$. If $\mu = \lambda$, the $BMO(\nu)$ space is the classical one, and the result is well-known. But, in full generality, this is a subtle result, as it is a characterization in the triple of μ, λ and b .

The purpose of this paper is to extend Bloom's result to the setting of Coifman-Rochberg-Weiss. Recall that a Calderón-Zygmund operator associated to a kernel $K(x, y)$ is an integral operator:

$$Tf(x) := \int_{\mathbb{R}^n} K(x, y)f(y) dy, \quad x \notin \text{supp } f,$$

and that the kernel satisfies the standard size and smoothness estimates

$$|K(x, y)| \leq \frac{C}{|x - y|^n},$$

$$|K(x + h, y) - K(x, y)| + |K(x, y + h) - K(x, y)| \leq C \frac{|h|^\delta}{|x - y|^{n+\delta}},$$

for all $|x - y| > 2|h| > 0$ and a fixed $\delta \in (0, 1]$.

Our first main result is the following upper bound for the commutator, $[b, T](f) := bTf - T(bf)$, with a Calderón-Zygmund operators.

Theorem 1.1. *Let T be a Calderón-Zygmund operator on \mathbb{R}^n and $\mu, \lambda \in A_p$ with $1 < p < \infty$. Suppose $b \in BMO(\nu)$, where $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$. Then*

$$\|[b, T] : L^p(\mu) \rightarrow L^p(\lambda)\| \leq c \|b\|_{BMO(\nu)},$$

where c is a constant depending on the dimension n , the operator T , and μ, λ , and p .

Recall that the Riesz transforms are defined by:

$$R_j(f)(x) := \text{p.v.} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \int_{\mathbb{R}^n} f(y) \frac{x_j - y_j}{|x - y|^{n+1}} dy, \quad j = 1, \dots, n.$$

Specializing to the Riesz transforms, we are able to characterize $BMO(\nu)$ in terms of the boundedness of the commutators. This gives a joint generalization of Bloom and Coifman-Rochberg-Weiss, which is the main result of the paper.

Theorem 1.2. *For $1 < p < \infty$, and $\mu, \lambda \in A_p$, set $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$. Then there are constants $0 < c < C < \infty$, depending only on n, p, μ and λ , for which*

$$(1.1) \quad c \|b\|_{BMO(\nu)} \leq \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\| \leq C \|b\|_{BMO(\nu)}.$$

Similar to [6], the equivalence above yields a weak-factorization result for weighted Hardy spaces.

Corollary 1.3. *Under the hypotheses and notation of Theorems 1.1 and 1.2, let $\lambda' := \lambda^{1-q}$ and T be a Calderón-Zygmund operator on \mathbb{R}^n . We have the inequality*

$$\|g_1(Tg_2) - (T^*g_1)g_2\|_{H^1(\nu)} \leq c \|g_1\|_{L^q(\lambda')} \|g_2\|_{L^p(\mu)},$$

where c is a constant depending on the dimension, the operator T , and on μ , λ , and p . Conversely, there exists a constant c so that every $f \in H^1(\nu)$ can be written as

$$(1.2) \quad f(x) = \sum_{i=1}^n \sum_{j=1}^{\infty} g_j^i(x) R_i h_j^i(x) + h_j^i(x) R_i g_j^i(x),$$

where R_i is the Riesz transform in the i th variable, and $g_j^i \in L^q(\lambda')$, $h_j^i \in L^p(\mu)$ with

$$\sum_{i=1}^n \sum_{j=1}^{\infty} \|g_j^i\|_{L^q(\lambda')} \|h_j^i\|_{L^p(\mu)} \leq c \|f\|_{H^1(\nu)}.$$

In the special case of the Hilbert transform, and $p = 2$, the paper [13] gives a ‘modern’ proof of Bloom’s result. We follow the outlines of that proof in the current setting. Using a Haar shift representation of T , a commutator can be written out as a sum of several terms. Most of these are paraproducts, with symbol b , but there are error terms as well. For the paraproducts, one needs a two-weight criteria of the boundedness. These criteria come in several different forms, but the additional structure of $\mu, \lambda \in A_p$ forces these ostensibly different criteria to be jointly finite. There are several error terms to handle. A comprehensive treatment of all terms depends upon an H^1 -BMO duality, which fortunately has already been developed.

Here is an outline of the paper. Section 2 collects all the necessary background that will be used throughout the paper. This includes background on A_p weights, and weighted H^1 . Section 3 introduces the paraproduct operators of interest and proves they are bounded in terms of Bloom’s BMO . We are able to mimic certain unweighted proofs by using a duality statement for weighted BMO spaces. In Section 4 we provide a family of equivalent conditions for a function to belong to the dyadic $BMO(\nu)$. Some of these equivalences are more useful when obtaining lower bounds as in Theorem 1.2, while others are more important in the proof of the upper bound in Theorem 1.1. In particular, there is seemingly no canonical form of the definition of Bloom’s BMO space. We have followed Bloom’s presentation in the definition above, and find other forms of the definition more convenient at different points of the proof. We do not track A_p constants, since the sharp bound would depend upon the choice of norm for $BMO(\nu)$. Section 5 contains the proof of Theorems 1.1 and 1.2 and Corollary 1.3. For the upper bound, we will use the Hytönen Representation Theorem, [14], to decompose the Calderón–Zygmund operator into Haar shift operators. Then we carefully analyze the commutator with each Haar shift to prove the desired statement in Theorem 1.1. A similar proof strategy can be found in [7]. For the lower bound, we follow the original proof of Coifman, Rochberg, and Weiss, [6], but with suitable modifications. The proof of Corollary 1.3 is then a standard application of well-known techniques.

2. NOTATION AND BACKGROUND

Throughout this paper, we use the standard notation “ $A \lesssim B$ ” to denote $A \leq cB$ for some constant c that depends only on the dimension n and, in the case of a weighted inequality, on p and the A_p constants of μ , and λ . And, “ $A \approx B$ ” means that $A \lesssim B$ and $B \lesssim A$. We let “ $:=$ ” mean equal by definition.

2.1. Dyadic Grids. Recall the standard dyadic grid on \mathbb{R}^n :

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

For every $\omega = (\omega_j)_{j \in \mathbb{Z}} \in (\{0, 1\}^{\mathbb{Z}})^{\mathbb{Z}}$ we may translate \mathcal{D}^0 by letting

$$\mathcal{D}^\omega := \left\{ Q \dot{+} \omega : Q \in \mathcal{D}^0 \right\},$$

where

$$Q \dot{+} \omega := Q + \sum_{j: 2^{-j} < l(Q)} 2^{-j} \omega_j.$$

Here $l(Q)$ will denote the side length of a cube Q in \mathbb{R}^n . We will only need to pay attention to ω when dealing with \mathbb{E}_ω , which denotes expectation with respect to the standard probability measure on the set of parameters ω . We denote a generic dyadic grid \mathcal{D}^ω on \mathbb{R}^n by \mathcal{D} .

Any such \mathcal{D} has the standard nestedness properties:

- For every $P, Q \in \mathcal{D}$, $P \cap Q$ is one of P , Q , and \emptyset .
- All $Q \in \mathcal{D}$ with $l(Q) = 2^{-k}$ for some fixed $k \in \mathbb{Z}$ partition \mathbb{R}^n .

For every $Q \in \mathcal{D}$ and every non-negative integer k , we denote:

- $Q^{(k)}$: the k^{th} generation ancestor of Q in \mathcal{D} , i.e. the unique element of \mathcal{D} that contains Q and has side length $2^k l(Q)$.
- $Q_{(k)}$: the collection of k^{th} generation descendants of Q in \mathcal{D} , i.e. the 2^{kn} disjoint subcubes of Q in \mathcal{D} with side length $2^{-k} l(Q)$.

2.2. The Haar System. Recall that every dyadic interval $I \subset \mathbb{R}$ is associated with two Haar functions:

$$h_I^0 := \frac{1}{\sqrt{|I|}} (\mathbb{1}_{I_-} - \mathbb{1}_{I_+}), \text{ and } h_I^1 := \frac{1}{\sqrt{|I|}} \mathbb{1}_I.$$

Note that h_I^0 is cancellative, while h_I^1 is non-cancellative. The cancellative Haar functions associated to a dyadic system on \mathbb{R} form an orthonormal basis for $L^2(\mathbb{R})$.

More generally, let $Q = Q_1 \times \cdots \times Q_n$ be a dyadic cube in \mathbb{R}^n – here all Q_i are dyadic intervals in \mathbb{R} with common length $l(Q)$. Then Q is associated with 2^n Haar functions:

$$h_Q^\epsilon(x) := h_{Q_1 \times \cdots \times Q_n}^{(\epsilon_1, \dots, \epsilon_n)}(x_1, \dots, x_n) := \prod_{i=1}^n h_{Q_i}^{\epsilon_i}(x_i),$$

where $\epsilon = (\epsilon_1, \dots, \epsilon_n) \in \{0, 1\}^n$ is called the *signature* of h_Q^ϵ . We write $\epsilon \equiv 1$ when $\epsilon_i = 1$ for all i ; in this case,

$$h_Q^1 := \frac{1}{\sqrt{|Q|}} \mathbb{1}_Q$$

is non-cancellative. All the other $2^n - 1$ Haar functions h_Q^ϵ with $\epsilon \not\equiv 1$ associated with Q are cancellative. Moreover, as in the one-dimensional case, all the cancellative Haar functions associated with a dyadic grid \mathcal{D} on \mathbb{R}^n form an orthonormal basis for $L^2(\mathbb{R}^n)$. In other words, every $f \in L^2(\mathbb{R}^n)$ has the expansion:

$$f = \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{f}(Q, \epsilon) h_Q^\epsilon,$$

where $\widehat{f}(Q, \epsilon) := \langle f, h_Q^\epsilon \rangle$. Throughout this paper we use $\langle \cdot, \cdot \rangle$ to denote the usual inner product on $L^2(\mathbb{R}^n)$.

We make a few simple but useful observations about Haar functions. First, note that h_Q^ϵ is constant on any subcube $P \subsetneq Q$ of Q in \mathcal{D} ; we denote this value by $h_Q^\epsilon(P)$. Then for any integer $k \geq 1$, we can express h_Q^ϵ as:

$$h_Q^\epsilon = \sum_{P \in Q^{(k)}} h_Q^\epsilon(P) \mathbb{1}_P.$$

Second, a simple calculation shows that:

$$h_Q^\epsilon h_Q^\eta = \frac{1}{\sqrt{|Q|}} h_Q^{\epsilon+\eta},$$

where for $\epsilon, \eta \in \{0, 1\}^n$ we define $\epsilon + \eta \in \{0, 1\}^n$ as:

$$(2.1) \quad (\epsilon + \eta)_i := \delta_{(\epsilon_i, \eta_i)} = \begin{cases} 0, & \text{if } \epsilon_i \neq \eta_i \\ 1, & \text{if } \epsilon_i = \eta_i. \end{cases}$$

It is easy to see from this definition that $\epsilon + \eta \equiv 1$ if and only if $\epsilon = \eta$, and $\epsilon + \eta = \epsilon$ if and only if $\eta \equiv 1$.

Third, we note that the average of a function f over a dyadic cube Q :

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

can be expressed as:

$$\langle f \rangle_Q = \sum_{\substack{P \in \mathcal{D}, P \supseteq Q \\ \epsilon \neq 1}} \widehat{f}(P, \epsilon) h_P^\epsilon(Q).$$

In turn, this yields the following useful expression:

$$(2.2) \quad \langle f \rangle_Q - \langle f \rangle_{Q^{(i)}} = \sum_{\substack{P, Q \in \mathcal{D}, \epsilon \neq 1 \\ Q \subsetneq P \subset Q^{(i)}}} \widehat{f}(P, \epsilon) h_P^\epsilon(Q) = \sum_{\substack{1 \leq k \leq i \\ \epsilon \neq 1}} \widehat{f}(Q^{(k)}, \epsilon) h_{Q^{(k)}}^\epsilon(Q),$$

which we shall use later in the proof of our main result.

2.3. A_p Weights. Let w be a weight on \mathbb{R}^n , i.e. w is an almost everywhere positive, locally integrable function. For $1 < p < \infty$, let $L^p(w) := L^p(\mathbb{R}^n; w)$ be the space of functions f that satisfy:

$$\|f\|_{L^p(w)} := \left(\int_{\mathbb{R}^n} |f(x)|^p dw(x) \right)^{\frac{1}{p}} < \infty,$$

where we also use w to denote the measure $w(x) \, dx$. For a cube Q in \mathbb{R}^n , we let

$$w(Q) := \int_Q w(x) \, dx \quad \text{and} \quad \langle w \rangle_Q := \frac{w(Q)}{|Q|}.$$

We say that w belongs to the Muckenhoupt class of A_p weights for some $1 < p < \infty$ provided that:

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

where q denotes the Hölder conjugate of p and the supremum above is over all cubes Q in \mathbb{R}^n with sides parallel to the axes. The quantity $[w]_{A_p}$ is called the A_p (Muckenhoupt) characteristic of w .

If $w \in A_p$, then the ‘conjugate’ weight

$$(2.3) \quad w' := w^{1-q} \in A_q,$$

with A_q characteristic $[w']_{A_q} = [w]_{A_p}^{q-1}$. In other words:

$$1 \leq \langle w \rangle_Q \langle w' \rangle_Q^{p-1} \leq [w]_{A_p} \quad \text{and} \quad 1 \leq \langle w' \rangle_Q \langle w \rangle_Q^{q-1} \leq [w]_{A_p}^{q-1},$$

for all $w \in A_p$ and all cubes Q in \mathbb{R}^n . We shall make much use of the duality relationship:

$$(2.4) \quad (L^p(w))^* \equiv L^q(w'), \quad \text{with pairing } \langle f, g \rangle, \quad \text{for all } f \in L^p(w), g \in L^q(w'),$$

and

$$\|f\|_{L^p(w)} = \sup_{\substack{g \in L^q(w') \\ \|g\|_{L^q(w')} \leq 1}} |\langle f, g \rangle|.$$

The case $p = 2$ is particularly easy to work with, as $w' = w^{-1}$ for $w \in A_2$.

A crucial property of A_p weights that we shall use repeatedly is the $L^p(w)$ -boundedness of the maximal function:

$$Mf := \sup_{Q \text{ cubes in } \mathbb{R}^n} \left(\langle |f| \rangle_Q \mathbb{1}_Q \right).$$

Muckenhoupt [24] showed that

$$(2.5) \quad \|Mf\|_{L^p(w)} \lesssim \|f\|_{L^p(w)},$$

Another pivotal development in A_p weight theory was the Extrapolation Theorem - see [8] - which, in particular, allows one to deduce the $L^p(w)$ -boundedness of an operator for all $w \in A_p$ solely from its $L^2(w)$ -boundedness for all $w \in A_2$. This is an extremely useful tool because, as we shall see, L^2 -estimates for A_2 weights are usually much ‘easier’ than L^p -estimates for A_p weights. We shall use the following form of this theorem:

Theorem 2.1. *Suppose an operator T satisfies:*

$$\|Tf\|_{L^2(w)} \leq AC(w) \|f\|_{L^2(w)}$$

for all $w \in A_2$, for some fixed $A > 0$ and $\alpha > 0$. Then:

$$\|Tf\|_{L^p(w)} \leq AC(w, p) \|f\|_{A_p}$$

for all $1 < p < \infty$ and all $w \in A_p$.

We are keeping track of the constant A here because in the proof of our main result later, we will need to keep track of some constants of the form 2^i appearing in these bounds (see for instance Lemma 2.2).

2.4. Dyadic Square Functions. Given a dyadic grid \mathcal{D} on \mathbb{R}^n , the dyadic square function $S_{\mathcal{D}}$ is defined by:

$$S_{\mathcal{D}}f := \left[\sum_{Q \in \mathcal{D}, \epsilon \neq 1} |\widehat{f}(Q, \epsilon)|^2 \frac{\mathbb{1}_Q}{|Q|} \right]^{\frac{1}{2}}.$$

A crucial property of this operator is the equivalence of norms

$$(2.6) \quad \|f\|_{L^p(w)} \simeq \|S_{\mathcal{D}}f\|_{L^p(w)},$$

for $w \in A_p$, $1 < p < \infty$.

At points below we will also need to have a weighted estimate for a shifted square function. This is the content of the following Lemma.

Lemma 2.2. *For a dyadic grid \mathcal{D} on \mathbb{R}^n and a pair (i, j) of non-negative integers, define:*

$$\widetilde{S}_{\mathcal{D}}^{i,j} f := \left(\sum_{Q \in \mathcal{D}, \epsilon \neq 1} \left(\sum_{P \in (Q^{(j)})_{(i)}} |\widehat{f}(P, \epsilon)| \right)^2 \frac{\mathbb{1}_Q}{|Q|} \right)^{\frac{1}{2}}.$$

Then for every weight $w \in A_p$, with $1 < p < \infty$:

$$(2.7) \quad \left\| \widetilde{S}_{\mathcal{D}}^{i,j} : L^p(w) \rightarrow L^p(w) \right\| \lesssim 2^{\frac{n}{2}(i+j)}.$$

Remark that for $i = j = 0$, this is just the usual dyadic square function $S_{\mathcal{D}}$.

Proof. In light of the Extrapolation Theorem 2.1, it suffices to prove an upper bound that is linear in $[w]_{A_2}$ for all A_2 weights w . So let $w \in A_2$ and note that

$$\begin{aligned} \|\widetilde{S}_{\mathcal{D}}^{i,j} f\|_{L^2(w)}^2 &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \left(\sum_{P \in (Q^{(j)})_{(i)}} |\widehat{f}(P, \epsilon)| \right)^2 \langle w \rangle_Q \\ &= \sum_{R \in \mathcal{D}, \epsilon \neq 1} \left(\sum_{P \in R_{(i)}} |\widehat{f}(P, \epsilon)| \right)^2 \sum_{Q \in R_{(j)}} \langle w \rangle_Q. \end{aligned}$$

Now

$$\begin{aligned} \sum_{P \in R_{(i)}} |\widehat{f}(P, \epsilon)| &= \sum_{P \in R_{(i)}} |\widehat{f}(P, \epsilon)| \frac{\langle w^{-1} \rangle_P^{\frac{1}{2}}}{\langle w^{-1} \rangle_P^{\frac{1}{2}}} \\ &\leq \left(\sum_{P \in R_{(i)}} |\widehat{f}(P, \epsilon)|^2 \frac{1}{\langle w^{-1} \rangle_P} \times \sum_{P \in R_{(i)}} \langle w^{-1} \rangle_P \right)^{\frac{1}{2}}, \end{aligned}$$

so, appealing to the square function bound (2.6),

$$\begin{aligned} \|\widetilde{S}_{\mathcal{D}}^{i,j} f\|_{L^2(w)}^2 &\leq 2^{n(i+j)} \sum_{R \in \mathcal{D}, \epsilon \neq 1} \sum_{P \in R_{(i)}} |\widehat{f}(P, \epsilon)|^2 \frac{1}{\langle w^{-1} \rangle_P} \langle w^{-1} \rangle_R \langle w \rangle_R \\ &\leq 2^{n(i+j)} \sum_{P \in \mathcal{D}, \epsilon \neq 1} |\widehat{f}(P, \epsilon)|^2 \frac{1}{\langle w^{-1} \rangle_P} \lesssim 2^{n(i+j)} \|f\|_{L^2(w)}^2. \end{aligned}$$

□

2.5. Hytönen's Representation Theorem. Fix a dyadic grid \mathcal{D}^ω on \mathbb{R}^n . For every pair i, j of non-negative integers, a dyadic shift operator with parameters (i, j) is an operator of the form:

$$S_\omega^{i,j} f := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \in \{0,1\}^n}} \sum_{\substack{P \in R_{(i)} \\ Q \in R_{(j)}}} a_{PQR}^{\epsilon\eta} \widehat{f}(P, \epsilon) h_Q^\eta,$$

where $a_{PQR}^{\epsilon\eta}$ are coefficients with

$$|a_{PQR}^{\epsilon\eta}| \leq \frac{\sqrt{|P||Q|}}{|R|} = 2^{-\frac{n}{2}(i+j)}.$$

The operator \mathbb{S}_ω^{ij} is called cancellative if all Haar functions appearing in its definition are cancellative. Otherwise, \mathbb{S}_ω^{ij} is called non-cancellative. The parameters $\kappa = (i, j)$ are a measure of the *complexity* of the shift. As is well-known, the dependence of norm estimates upon complexity must be tracked, but is only linear in

$$(2.8) \quad \kappa_{ij} := \max(i, j, 1),$$

whereas there is exponential decay in κ , in the celebrated representation theorem of Hytönen [14, 15, 17]:

Theorem 2.3. *Let T be a Calderón-Zygmund operator associated with a δ -standard kernel. Then there exist dyadic shift operators \mathbb{S}_ω^{ij} with parameters (i, j) for all non-negative integers i, j such that*

$$\langle Tf, g \rangle = c \mathbb{E}_\omega \sum_{i,j=0}^{\infty} 2^{-\kappa_{i,j} \frac{\delta}{2}} \langle \mathbb{S}_\omega^{ij} f, g \rangle,$$

for all bounded, compactly supported functions f and g , where c is a constant depending on the dimension n and on T . Here all \mathbb{S}_ω^{ij} with $(i, j) \neq (0, 0)$ are cancellative, but the shifts \mathbb{S}_ω^{00} may be non-cancellative.

The statement of this Theorem involves a *random* choice of grids. However, in all applications of this result, one analyzes the norm behavior of the Haar shift operators, establishing bounds that are uniform with respect to the choice of dyadic grid. The exact manner in which the random dyadic grid are formed is not relevant to us. We will discuss the case $i = j = 0$ in more detail in Section 5.2.

Another useful tool for us will be the weighted estimate below, which can be found in [16, 19, 30].

Theorem 2.4. *Let \mathbb{S}_ω^{ij} be a dyadic shift operator with complexity κ_{ij} . Then for any weight $w \in A_p$ with $p > 1$:*

$$(2.9) \quad \|\mathbb{S}_\omega^{ij} : L^p(w) \rightarrow L^p(w)\| \lesssim \kappa_{ij} [w]_{A_p}^{\max(1, \frac{1}{p-1})}.$$

2.6. Weighted BMO - H^1 duality. For a weight w on \mathbb{R}^n , the weighted BMO space $BMO(w)$ is defined to be the space of all locally integrable functions b that satisfy:

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

where the supremum is over all cubes Q in \mathbb{R}^n with sides parallel to the axes. For a general weight, the definition of the BMO norm is highly dependent on its L^1 average. But, if the weight is A_∞ , one is free to replace the L^1 -norm by larger averages. Namely, defining

$$\|b\|_{BMO^q(w)} := \sup_Q \left(\frac{1}{w(Q)} \int_Q |b - \langle b \rangle_Q|^q dw' \right)^{\frac{1}{q}}.$$

it is shown in [25] that

$$(2.10) \quad \|b\|_{BMO(w)} \leq \|b\|_{BMO^q(w)} \leq C(n, p, [w]_{A_\infty}) \|b\|_{BMO(w)}.$$

The first inequality is a straightforward application of Hölder's inequality, but the second one requires a stopping time argument.

For a dyadic grid \mathcal{D} on \mathbb{R}^n , we define the dyadic versions of the norms above by taking supremum over $Q \in \mathcal{D}$ instead of over all cubes Q in \mathbb{R}^n , and denote these spaces by $BMO_{\mathcal{D}}(w)$ and $BMO_{\mathcal{D}}^q(w)$. Clearly $BMO(w) \subset BMO_{\mathcal{D}}(w)$ for any choice of \mathcal{D} , and the equivalence in (2.10) also holds for the dyadic versions of these spaces.

Now fix a dyadic grid \mathcal{D} on \mathbb{R}^n and a weight $w \in A_{\infty}$. Define the dyadic weighted Hardy space $H_{\mathcal{D}}^1(w)$ (see [11]) to be the space of all Φ that satisfy:

$$\|\Phi\|_{H_{\mathcal{D}}^1(w)} := \|S_{\mathcal{D}}\Phi\|_{L^1(w)} < \infty.$$

The dual space of $H_{\mathcal{D}}^1(w)$ is the weighted Carleson measure space $CM_{\mathcal{D}}^1(w)$, that is, the space of all locally integrable functions g such that:

$$\|g\|_{CM_{\mathcal{D}}^1(w)} := \sup_{Q \in \mathcal{D}} \left(\frac{1}{w(Q)} \sum_{\substack{P \subset Q \in \mathcal{D} \\ \epsilon \neq 1}} \frac{|\widehat{g}(P, \epsilon)|^2}{\langle w \rangle_P} \right)^{\frac{1}{2}} < \infty,$$

with duality pairing $\langle g, \Phi \rangle$ for $g \in CM_{\mathcal{D}}^1(w)$ and $\Phi \in H_{\mathcal{D}}^1(w)$ see [23, 33]. We have then

$$|\langle g, \Phi \rangle| \leq \|g\|_{CM_{\mathcal{D}}^1(w)} \|S_{\mathcal{D}}\Phi\|_{L^1(w)}.$$

Specializing to the case of $w \in A_2$, we have an H^1 - BMO duality.

Lemma 2.5. *If $w \in A_2$, there holds*

$$(2.11) \quad |\langle b, \Phi \rangle| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(w)} \|S_{\mathcal{D}}\Phi\|_{L^1(w)}.$$

Proof. The inequality is equivalent to the assertion that if $b \in BMO_{\mathcal{D}}^2(w)$ for $w \in A_2$, then $b \in CM_{\mathcal{D}}^1(w)$ with $\|b\|_{CM_{\mathcal{D}}^1(w)} \lesssim [w]_{A_2} \|b\|_{BMO_{\mathcal{D}}^2(w)}$. But observe that fixing a cube Q , and expanding b in the Haar basis, we have

$$(b - \langle b \rangle_Q) \mathbb{1}_Q = \sum_{P \subset Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(P, \epsilon) h_P^{\epsilon} =: B_Q.$$

Then, by (2.6),

$$\begin{aligned} \int_Q |b - \langle b \rangle_Q|^2 dw^{-1} &= \|B_Q\|_{L^2(w^{-1})}^2 \\ &\gtrsim \|S_{\mathcal{D}}B_Q\|_{L^2(w^{-1})}^2 \\ &= \sum_{P \subset Q \in \mathcal{D}, \epsilon \neq 1} |\widehat{b}(P, \epsilon)|^2 \langle w^{-1} \rangle_P \\ &\gtrsim \sum_{P \subset Q \in \mathcal{D}, \epsilon \neq 1} |\widehat{b}(P, \epsilon)|^2 \frac{1}{\langle w \rangle_P}. \end{aligned}$$

And so the Lemma follows. □

2.7. Bloom's $BMO(\nu)$. From here on, fix $1 < p < \infty$ and two A_p weights μ and λ on \mathbb{R}^n , and define Bloom's weight

$$\nu := \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}.$$

Lemma 2.6. *The weight ν belongs to the A_2 class. In particular:*

$$[\nu]_{A_2} \leq [\mu]_{A_p}^{\frac{1}{p}} [\lambda]_{A_p}^{\frac{1}{p}}.$$

Moreover

$$(2.12) \quad |\langle b, \Phi \rangle| \lesssim \|b\|_{BMO_D^2(\nu)} \|S_D \Phi\|_{L^1(\nu)}.$$

Proof. The second inequality is immediate from (2.11). For the first, by Hölder's inequality:

$$\langle \nu \rangle_Q \leq \langle \mu \rangle_Q^{\frac{1}{p}} \langle \lambda' \rangle_Q^{\frac{1}{q}} \quad \text{and} \quad \langle \nu^{-1} \rangle_Q \leq \langle \mu' \rangle_Q^{\frac{1}{q}} \langle \lambda \rangle_Q^{\frac{1}{p}},$$

so

$$\langle \nu \rangle_Q \langle \nu^{-1} \rangle_Q \leq \left(\langle \mu \rangle_Q \langle \mu' \rangle_Q^{p-1} \right)^{\frac{1}{p}} \left(\langle \lambda \rangle_Q \langle \lambda' \rangle_Q^{p-1} \right)^{\frac{1}{p}}$$

But the terms in parentheses are at most $[\mu]_{A_p}$ and $[\lambda]_{A_p}$. Hence $\nu \in A_2$. \square

We record here a simple inequality.

$$(2.13) \quad \langle \mu \rangle_Q^{\frac{1}{p}} \langle \lambda' \rangle_Q^{\frac{1}{q}} \lesssim \frac{1}{\langle \mu' \rangle_Q^{\frac{1}{q}} \langle \lambda \rangle_Q^{\frac{1}{p}}} \lesssim \frac{1}{\langle \nu^{-1} \rangle_Q} \lesssim \langle \nu \rangle_Q.$$

3. TWO-WEIGHT INEQUALITIES FOR PARAPRODUCT OPERATORS

3.1. Paraproducts. For a fixed dyadic grid \mathcal{D} on \mathbb{R}^n , the ‘paraproduct’ operators with symbol b are defined by:

$$(3.1) \quad \Pi_b^{\mathcal{D}} f := \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(Q, \epsilon) \langle f \rangle_Q h_Q^\epsilon,$$

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

$$(3.2) \quad \text{and} \quad \Gamma_b^{\mathcal{D}} f := \sum_{Q \in \mathcal{D}} \sum_{\substack{\epsilon, \eta \neq 1 \\ \epsilon \neq \eta}} \widehat{b}(Q, \epsilon) \widehat{f}(Q, \eta) \frac{1}{\sqrt{|Q|}} h_Q^{\epsilon+\eta}.$$

For ease of notation, we fix \mathcal{D} through the rest of this section and suppress the subscript \mathcal{D} from the paraproducts.

Commutators are a difference of products, and the product of two functions is decomposed into paraproducts as follows:

$$(3.3) \quad bf = \Pi_b f + \Pi_f b + \Pi_b^* f + \Gamma_b f.$$

To see the decomposition in (3.3), express b and f in terms of the Haar expansions:

$$bf = \sum_{R, Q \in \mathcal{D}} \sum_{\epsilon, \eta \neq 1} \widehat{b}(Q, \epsilon) \widehat{f}(R, \eta) h_Q^\epsilon h_R^\eta$$

and analyze the sum over the three different cases $R \subsetneq Q$, $R \supsetneq Q$, and $Q = R$. The latter case easily yields:

$$\sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(Q, \epsilon) \widehat{f}(Q, \epsilon) \frac{\mathbb{1}_Q}{|Q|} + \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \sum_{\eta \neq 1, \epsilon \neq \eta} \widehat{b}(Q, \epsilon) \widehat{f}(Q, \eta) \frac{1}{\sqrt{|Q|}} h_Q^{\epsilon+\eta} = \Pi_b^* f + \Gamma_b f.$$

To illustrate one of the other two cases:

$$\begin{aligned}
 \sum_{Q \subsetneq R \in \mathcal{D}; \epsilon, \eta \neq 1} \widehat{b}(Q, \epsilon) \widehat{f}(R, \eta) h_Q^\epsilon h_R^\eta &= \sum_{Q \subsetneq R \in \mathcal{D}; \epsilon, \eta \neq 1} \widehat{b}(Q, \epsilon) \widehat{f}(R, \eta) h_Q^\epsilon h_R^\eta(Q) \\
 &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(Q, \epsilon) h_Q^\epsilon \sum_{R \supseteq Q, \eta \neq 1} \widehat{f}(R, \eta) h_R^\eta(Q) \\
 &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(Q, \epsilon) \langle f \rangle_Q h_Q^\epsilon = \Pi_b f.
 \end{aligned}$$

Similarly, the case $R \subsetneq Q$ yields $\Pi_b f$.

We pause here for a moment to remark that the term Γ_b disappears in the one-dimensional case, where the familiar decomposition is $bf = \Pi_b f + \Pi_b^* f + \Gamma_b f$. By definition (2.1), $\epsilon + \eta = 1$ if and only if $\epsilon = \eta$, so while $\Pi_b^* f$ maintains its one-dimensional structure and contains all the non-cancellative Haar functions, the term Γ_b contains only the cancellative Haar functions. Moreover, as in the one-dimensional case, Π_b^* is the adjoint of Π_b in *unweighted* $L^2(\mathbb{R}^n)$, while the third paraproduct Γ_b is self-adjoint in $L^2(\mathbb{R}^n)$:

$$\langle \Pi_b f, g \rangle = \langle f, \Pi_b^* g \rangle \quad \text{and} \quad \langle \Gamma_b f, g \rangle = \langle f, \Gamma_b g \rangle.$$

3.2. Two-weight Inequalities for Paraproducts. Next, we discuss boundedness of the paraproducts as operators from $L^p(\mu) \rightarrow L^p(\lambda)$. Before we proceed, we make the interesting observation that the adjointness statements about the three paraproducts in unweighted $L^2(dx)$ extend to this case, in the sense of Banach space adjoints. Specifically

$$\begin{aligned}
 \text{The adjoint of } \Pi_b : L^p(\mu) \rightarrow L^p(\lambda) &\text{ is } \Pi_b^* : L^q(\lambda') \rightarrow L^q(\mu'); \\
 \text{The adjoint of } \Pi_b^* : L^p(\mu) \rightarrow L^p(\lambda) &\text{ is } \Pi_b : L^q(\lambda') \rightarrow L^q(\mu'); \\
 \text{The adjoint of } \Gamma_b : L^p(\mu) \rightarrow L^p(\lambda) &\text{ is } \Gamma_b : L^q(\lambda') \rightarrow L^q(\mu').
 \end{aligned}$$

Here, λ' is the conjugate weight, as in (2.3).

These follow from (2.4). For instance, the adjoint of $\Pi_b : L^p(\mu) \rightarrow L^p(\lambda)$ is the unique operator $T : L^q(\lambda') \rightarrow L^q(\mu')$ such that

$$\langle \Pi_b f, g \rangle = \langle f, Tg \rangle, \text{ for all } f \in L^p(\mu), g \in L^q(\lambda').$$

But this is just the inner product in unweighted $L^2(dx)$, so $\langle \Pi_b f, g \rangle = \langle f, \Pi_b^* g \rangle$, and $T = \Pi_b^*$. The second statement above follows identically, and the third statement follows from the self-adjointness of Γ_b in $L^2(dx)$.

This is a two-weight result for paraproducts, fundamental for us.

Theorem 3.1. *Let \mathcal{D} be a fixed dyadic grid on \mathbb{R}^n , and suppose $b \in BMO_{\mathcal{D}}^2(\nu)$. Then:*

$$(3.4) \quad \|\Pi_b : L^p(\mu) \rightarrow L^p(\lambda)\| = \|\Pi_b^* : L^q(\lambda') \rightarrow L^q(\mu')\| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)},$$

$$(3.5) \quad \|\Pi_b^* : L^p(\mu) \rightarrow L^p(\lambda)\| = \|\Pi_b : L^q(\lambda') \rightarrow L^q(\mu')\| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)},$$

$$(3.6) \quad \|\Gamma_b : L^p(\mu) \rightarrow L^p(\lambda)\| = \|\Gamma_b : L^q(\lambda') \rightarrow L^q(\mu')\| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)}.$$

Proof. The proof is by duality, exploiting the H^1 -BMO duality inequality (2.12) to gain the term $\|b\|_{BMO_{\mathcal{D}}^2(\nu)}$. This will leave us with a bilinear square function involving f and g , which will be controlled by a product of a maximal function and a linear square function. The

details are as follows. We let $f \in L^p(\mu)$ and $g \in L^q(\lambda')$. Then

$$\begin{aligned} |\langle \Pi_b f, g \rangle| &= \left| \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{b}(Q, \epsilon) \langle f \rangle_Q \widehat{g}(Q, \epsilon) \right| \\ &= |\langle b, \Phi \rangle| \quad \text{where } \Phi := \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \langle f \rangle_Q \widehat{g}(Q, \epsilon) h_Q^\epsilon \\ &\lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \|S_{\mathcal{D}} \Phi\|_{L^1(\nu)} \quad \text{by (2.12)}. \end{aligned}$$

Now, $S_{\mathcal{D}} \Phi$ is bilinear in f and g , and is no more than

$$(S_{\mathcal{D}} \Phi)^2 = \sum_{Q \in \mathcal{D}, \epsilon \neq 1} |\langle f \rangle_Q|^2 |\widehat{g}(Q, \epsilon)|^2 \frac{\mathbb{1}_Q}{|Q|} \leq (Mf)^2 \sum_{Q \in \mathcal{D}, \epsilon \neq 1} |\widehat{g}(Q, \epsilon)|^2 \frac{\mathbb{1}_Q}{|Q|} = (Mf)^2 (S_{\mathcal{D}} g)^2.$$

A straight forward application of Hölder's inequality, and bounds for the maximal and square functions will complete the proof.

$$\begin{aligned} \|S_{\mathcal{D}} \Phi\|_{L^1(\nu)} &\leq \int (Mf)(S_{\mathcal{D}} g) d\mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}} \\ &\leq \|Mf\|_{L^p(\mu)} \|S_{\mathcal{D}} g\|_{L^q(\lambda')} \lesssim \|f\|_{L^p(\mu)} \|g\|_{L^q(\lambda')} \end{aligned}$$

by (2.5), (2.6). This completes the proof of (3.4).

The second set of inequalities (3.5) are equivalent to the first, by a simple duality argument. Concerning the last set of inequalities, (3.6), they are different in that the operator only has cancellative Haar functions. One can bound Haar coefficients by maximal functions, doing so on either f or g . □

4. EQUIVALENCES FOR DYADIC BLOOM BMO

Bloom's BMO space has several equivalent formulations, which is a key fact in proof of the main theorems. Those that we need are summarized here. For a fixed dyadic grid \mathcal{D} on \mathbb{R}^n define the quantities:

$$\begin{aligned} \mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) &:= \sup_{Q \in \mathcal{D}} \left(\frac{1}{\mu(Q)} \int_Q |b - \langle b \rangle_Q|^p d\lambda \right)^{\frac{1}{p}}; \\ \mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda') &:= \sup_{Q \in \mathcal{D}} \left(\frac{1}{\lambda'(Q)} \int_Q |b - \langle b \rangle_Q|^q d\mu' \right)^{\frac{1}{q}}. \end{aligned}$$

We provide several equivalent statements for the dyadic version of Bloom's BMO space $BMO_{\mathcal{D}}(\nu)$.

Theorem 4.1. *Let \mathcal{D} be a fixed dyadic grid on \mathbb{R}^n . The following are equivalent:*

- (1) $b \in BMO_{\mathcal{D}}^2(\nu)$.
- (2) The operator $\Pi_b : L^p(\mu) \rightarrow L^p(\lambda)$ is bounded.
- (3) The operator $\Pi_b^* : L^p(\mu) \rightarrow L^p(\lambda)$ is bounded.
- (4) The operators Π_b and Π_b^* are bounded $L^2(\nu) \rightarrow L^2(\nu^{-1})$.
- (5) $\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) < \infty$.
- (6) $\mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda') < \infty$.
- (7) $b \in BMO_{\mathcal{D}}(\nu)$.

Proof. (1) \Rightarrow (2) and (3). This is the core of Theorem 3.1.

(2) \Rightarrow (5) and (3) \Rightarrow (6): These two assertions are the same by duality, and we consider the first implication. Assuming the paraproduct Π_b is bounded, we test the norm of this operator on indicators of intervals, and get the condition in (5). By the Littlewood-Paley inequalities in weighted L^p spaces, (2.6), we have

$$\begin{aligned} \|\mathbb{1}_Q(b - \langle b \rangle_Q)\|_{L^p(\lambda)} &\lesssim \|S_{\mathcal{D}}[\mathbb{1}_Q(b - \langle b \rangle_Q)]\|_{L^p(\lambda)} \\ &\leq \|S_{\mathcal{D}}(\Pi_b \mathbb{1}_Q)\|_{L^p(\lambda)} \end{aligned}$$

Above, we have the square function of $\Pi_b \mathbb{1}_Q$, which follows from the identity

$$\Pi_b \mathbb{1}_Q = \sum_{P \subset Q, \epsilon \neq 1} \widehat{b}(P, \epsilon) h_P^\epsilon + \sum_{P \supseteq Q, \epsilon \neq 1} \widehat{b}(P, \epsilon) \frac{|Q|}{|P|} h_P^\epsilon.$$

Again by the Littlewood-Paley inequalities,

$$\|S_{\mathcal{D}}(\Pi_b \mathbb{1}_Q)\|_{L^p(\lambda)} \lesssim \|\Pi_b \mathbb{1}_Q\|_{L^p(\lambda)}.$$

But, the assumption of the norm boundedness of the paraproduct implies that we have

$$\lambda(Q)^{-1/p} \|S_{\mathcal{D}}[\mathbb{1}_Q(b - \langle b \rangle_Q)]\|_{L^p(\lambda)} \lesssim \|\Pi_b : L^p(\mu) \rightarrow L^p(\lambda)\|.$$

(5) or (6) \Rightarrow (7): Suppose $\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) < \infty$, the case of $\mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda') < \infty$ being similar. Then

$$\begin{aligned} \int_Q |b - \langle b \rangle_Q| dx &\leq \left(\int_Q |b - \langle b \rangle_Q|^p d\lambda \right)^{\frac{1}{p}} \left(\int_Q d\lambda' \right)^{\frac{1}{q}} \\ &\lesssim \mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) \mu(Q)^{\frac{1}{p}} \lambda'(Q)^{\frac{1}{q}} \\ &\lesssim \mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) \nu(Q) \end{aligned} \quad \text{by (2.13).}$$

That is, we have shown $\|b\|_{BMO_{\mathcal{D}}(\nu)} \leq \mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda)$.

(7) \Leftrightarrow (1): As discussed in Section 2.6, this is proved in [25].

(1) \Rightarrow (4): Suppose $b \in BMO_{\mathcal{D}}^2(\nu)$. We make use of the duality $(L^2(\nu))^* \equiv L^2(\nu^{-1})$, with the usual unweighted L^2 inner product as the duality pairing, and let $f, g \in L^2(\nu)$. Then

$$|\langle \Pi_b f, g \rangle| = |\langle b, \Phi \rangle| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \|S_{\mathcal{D}} \Phi\|_{L^1(\nu)},$$

where $\Phi = \sum_{P, \epsilon} \langle f \rangle_P \widehat{g}(P, \epsilon) h_P^\epsilon$. It follows easily that $S_{\mathcal{D}} \Phi \leq (Mf)(S_{\mathcal{D}}g)$, and then from (2.5) and (2.6):

$$\|S_{\mathcal{D}} \Phi\|_{L^1(\nu)} \leq \|Mf\|_{L^2(\nu)} \|S_{\mathcal{D}} f\|_{L^2(\nu)} \lesssim \|f\|_{L^2(\nu)} \|g\|_{L^2(\nu)}.$$

Then

$$\|\Pi_b : L^2(\nu) \rightarrow L^2(\nu^{-1})\| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)}.$$

The same statement for Π_b^* follows by noting that $\Pi_b^* : L^2(\nu) \rightarrow L^2(\nu^{-1})$ is the adjoint of $\Pi_b : L^2(\nu) \rightarrow L^2(\nu^{-1})$.

(4) \Rightarrow (1): Suppose $\Pi_b : L^2(\nu) \rightarrow L^2(\nu^{-1})$ is bounded. Then

$$\|\Pi_b \mathbb{1}_Q\|_{L^2(\nu^{-1})} \leq A\nu(Q)^{\frac{1}{2}} \quad \text{and} \quad \|\Pi_b^* \mathbb{1}_Q\|_{L^2(\nu^{-1})} \leq A\nu(Q)^{\frac{1}{2}}.$$

In this situation we know that both paraproducts are bounded, so we can get to the $BMO_{\mathcal{D}}^2(\nu)$ norm of b faster than in the square function approach, by noting that

$$(4.1) \quad \mathbb{1}_Q(b - \langle b \rangle_Q) = \mathbb{1}_Q(\Pi_b \mathbb{1}_Q - \Pi_b^* \mathbb{1}_Q).$$

Then

$$\begin{aligned} \int_Q |b - \langle b \rangle_Q|^2 d\nu^{-1} &= \int_Q |\Pi_b \mathbb{1}_Q - \Pi_b^* \mathbb{1}_Q|^2 d\nu^{-1} \\ &\leq 2\|\Pi_b \mathbb{1}_Q\|_{L^2(\nu^{-1})}^2 + 2\|\Pi_b^* \mathbb{1}_Q\|_{L^2(\nu^{-1})}^2 \\ &\leq 4A^2\nu(Q), \end{aligned}$$

which gives exactly $\|b\|_{BMO_{\mathcal{D}}^2(\nu)} \leq 2A$. \square

We note here that the equivalence (1) \Leftrightarrow (4) in fact holds for any A_2 weight. Moreover, the strategy used in proving this last equivalence above can be employed to give more precise bounds for the quantities $\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda)$ and $\mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda')$ when $b \in BMO_{\mathcal{D}}^2(\nu)$. For in this case, we know from Theorem 3.1 that

$$\begin{aligned} \|\Pi_b \mathbb{1}_Q\|_{L^p(\lambda)} + \|\Pi_b^* \mathbb{1}_Q\|_{L^p(\lambda)} &\lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \mu(Q)^{\frac{1}{p}}, \\ \|\Pi_b \mathbb{1}_Q\|_{L^q(\mu')} + \|\Pi_b^* \mathbb{1}_Q\|_{L^q(\mu')} &\lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \lambda'(Q)^{\frac{1}{q}}. \end{aligned}$$

Then, using (4.1), we have for any $Q \in \mathcal{D}$:

$$\begin{aligned} \left(\int_Q |b - \langle b \rangle_Q|^p d\lambda \right)^{\frac{1}{p}} &= \|\mathbb{1}_Q(\Pi_b \mathbb{1}_Q - \Pi_b^* \mathbb{1}_Q)\|_{L^p(\lambda)} \\ &\leq \|\Pi_b \mathbb{1}_Q\|_{L^p(\lambda)} + \|\Pi_b^* \mathbb{1}_Q\|_{L^p(\lambda)} \\ &\lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \mu(Q)^{\frac{1}{p}}. \end{aligned}$$

The similar statement for $\mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda')$ follows immediately by considering the paraproducts as operators $L^q(\lambda') \rightarrow L^q(\mu')$. We state this result separately.

Proposition 4.2. *Let $\mu, \lambda \in A_p$ with $1 < p < \infty$ and put $\nu := \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$. Then for any dyadic grid \mathcal{D} on \mathbb{R}^n and any $b \in BMO_{\mathcal{D}}^2(\nu)$:*

$$\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)}, \text{ and } \mathbb{B}_2^{\mathcal{D}}(b, \mu', \lambda') \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)}.$$

5. TWO-WEIGHT INEQUALITIES FOR COMMUTATORS WITH CALDERÓN-ZYGMUND OPERATORS

We prove Theorem 1.1, our upper bound on commutators. By the Hytönen Representation Theorem 2.3,

$$\langle [b, T]f, g \rangle = c(n, T) \mathbb{E}_{\omega} \sum_{i,j=0}^{\infty} 2^{-\kappa_{i,j} \frac{\delta}{2}} \langle [b, \mathbb{S}_{\omega}^{ij}]f, g \rangle,$$

for all bounded, compactly supported f, g , so it suffices to show that the commutators $[b, \mathbb{S}_{\omega}^{ij}]$ are bounded $L^p(\mu) \rightarrow L^p(\lambda)$ uniformly in ω . We claim that for any choice of \mathcal{D}^{ω} :

$$(5.1) \quad \|[b, \mathbb{S}_{\omega}^{ij}] : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \kappa_{ij} \|b\|_{BMO(\nu)},$$

for all non-negative integers i, j , where recall that $\kappa_{i,j}$ is defined in (2.8), and in particular is at most linear in $i + j$. The linear growth in complexity in the second estimate is dominated by the exponential decay in complexity in the first. Hence we conclude an upper bound on the norm of the commutator, completing the proof of Theorem 1.1.

In what follows, consider $\mathcal{D} := \mathcal{D}^\omega$ to be fixed and simply write \mathbb{S}^{ij} . The commutator $[b, \mathbb{S}^{ij}]f = b\mathbb{S}^{ij}f - \mathbb{S}^{ij}(bf)$. Expand the products into paraproducts as in (3.3) and obtain

$$\begin{aligned} [b, \mathbb{S}^{ij}]f &= T_1f + T_2f + \mathcal{R}^{ij}f, \\ \text{where } T_1f &:= (\Pi_b + \Pi_b^* + \Gamma_b)(\mathbb{S}^{ij}f), \quad T_2f := \mathbb{S}^{ij}(\Pi_b + \Pi_b^* + \Gamma_b)f, \\ \text{and } \mathcal{R}^{ij}f &:= \Pi_{\mathbb{S}^{ij}f}b - \mathbb{S}^{ij}\Pi_f b. \end{aligned}$$

In this equality, the principal terms are T_1 and T_2 . Using (2.9) and Theorem 3.1, we easily obtain $\|T_k : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \kappa_{ij}\|b\|_{BMO(\nu)}$ for $k = 1, 2$. So we only need to analyze the remainder term \mathcal{R}^{ij} . In what follows, we will show that

$$(5.2) \quad \|\mathcal{R}^{ij} : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \kappa_{ij}\|b\|_{BMO(\nu)},$$

for all i, j . Then (5.1) follows.

5.1. Remainder Estimate for $(i, j) \neq (0, 0)$. In this case, the dyadic shift \mathbb{S}^{ij} is cancellative:

$$\mathbb{S}^{ij}f := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \sum_{\substack{P \in R^{(i)} \\ Q \in R^{(j)}}} a_{PQR}^{\epsilon\eta} \widehat{f}(P, \epsilon) h_Q^\eta.$$

Then for any $N \in \mathcal{D}$ and $\gamma \neq 1$:

$$\mathbb{S}^{ij}h_N^\gamma = \sum_{\eta \neq 1} \sum_{Q \in (N^{(i)})_{(j)}} a_{NQ}^{\gamma\eta} h_Q^\eta \quad \text{and} \quad \langle \mathbb{S}^{ij}f, h_N^\gamma \rangle = \sum_{\epsilon \neq 1} \sum_{P \in (N^{(j)})_{(i)}} a_{PNP}^{\epsilon\gamma} \widehat{f}(P, \epsilon).$$

These expressions give us the two terms in the remainder as

$$(5.3) \quad \Pi_{\mathbb{S}^{ij}f}b = \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \sum_{\substack{P \in R^{(i)} \\ Q \in R^{(j)}}} a_{PQR}^{\epsilon\eta} \widehat{f}(P, \epsilon) \langle b \rangle_Q h_Q^\eta, \quad \mathbb{S}^{ij}\Pi_f b = \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \sum_{\substack{P \in R^{(i)} \\ Q \in R^{(j)}}} a_{PQR}^{\epsilon\eta} \widehat{f}(P, \epsilon) \langle b \rangle_P h_Q^\eta.$$

From (5.3):

$$\mathcal{R}^{ij}f = \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \sum_{\substack{P \in R^{(i)} \\ Q \in R^{(j)}}} a_{PQR}^{\epsilon\eta} \widehat{f}(P, \epsilon) \left(\langle b \rangle_Q - \langle b \rangle_P \right) h_Q^\eta.$$

The difference in the averages of b is an essential term, but the cubes P and Q above are just descendants of R . They need not intersect.

We continue our analysis of the remainder in terms of the relative sizes of i and j , but the cases of $i \leq j$ and $j \leq i$ are dual, and so we only consider the former. Each $Q \in R_{(j)}$ is contained in a unique $N \in R_{(i)}$, and then $Q \in N_{(j-i)}$. (Note that $N = Q$ if $i = j$.) Rewrite $\mathcal{R}^{ij}f$ by grouping the Q 's this way:

$$\mathcal{R}^{ij}f = \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \sum_{P, N \in R^{(i)}} \widehat{f}(P, \epsilon) \sum_{Q \in N_{(j-i)}} a_{PQR}^{\epsilon\eta} \left(\langle b \rangle_Q - \langle b \rangle_P \right) h_Q^\eta,$$

and write

$$\langle b \rangle_Q - \langle b \rangle_P = (\langle b \rangle_Q - \langle b \rangle_N) + (\langle b \rangle_N - \langle b \rangle_R) + (\langle b \rangle_R - \langle b \rangle_P).$$

Note that the first term disappears if $i = j$, and also the expansion (2.2) applies to each of the terms in the parentheses above. The remainder is the sum of three terms.

$$\begin{aligned} \mathcal{R}^{ij} f &= \sum_{k=1}^{j-i} \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{P \in R^{(i)} \\ N \in R^{(j-k)}}} \widehat{f}(P, \epsilon) \widehat{b}(N, \gamma) \sum_{Q \in N^{(k)}} a_{PQR}^{\epsilon \eta} h_N^\gamma(Q) h_Q^\eta \\ &+ \sum_{k=1}^i \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{P \in R^{(i)} \\ N \in R^{(i-k)}}} \widehat{f}(P, \epsilon) \widehat{b}(N, \gamma) \sum_{Q \in N^{(j-i+k)}} a_{PQR}^{\epsilon \eta} h_N^\gamma(Q) h_Q^\eta \\ &- \sum_{k=1}^i \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{N \in R^{(i-k)} \\ Q \in R^{(j)}}} \left(\sum_{P \in N^{(k)}} \widehat{f}(P, \epsilon) h_N^\gamma(P) \right) \widehat{b}(N, \gamma) a_{PQR}^{\epsilon \eta} h_Q^\eta. \end{aligned}$$

We relabel the second term by replacing k with $k - j + i$, and then combine it with the first term. Finally, we may write

$$(5.4) \quad \mathcal{R}^{ij} f = \sum_{k=1}^j A_k f - \sum_{k=1}^i B_k f,$$

where:

$$A_k f := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{P \in R^{(i)} \\ N \in R^{(j-k)}}} \widehat{f}(P, \epsilon) \widehat{b}(N, \gamma) \sum_{Q \in N^{(k)}} a_{PQR}^{\epsilon \eta} h_N^\gamma(Q) h_Q^\eta,$$

and

$$B_k f := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{N \in R^{(i-k)} \\ Q \in R^{(j)}}} \left(\sum_{P \in N^{(k)}} a_{PQR}^{\epsilon \eta} \widehat{f}(P, \epsilon) h_N^\gamma(P) \right) \widehat{b}(N, \gamma) h_Q^\eta.$$

It suffices to prove that $\|A_k : L^2(\mu) \rightarrow L^2(\lambda)\| + \|B_k : L^2(\mu) \rightarrow L^2(\lambda)\| \lesssim \|b\|_{BMO(\nu)}$, because then from (5.4) we obtain:

$$\|\mathcal{R}^{ij} : L^2(\mu) \rightarrow L^2(\lambda)\| \leq j \|A_k : L^2(\mu) \rightarrow L^2(\lambda)\| + i \|B_k : L^2(\mu) \rightarrow L^2(\lambda)\| \lesssim \kappa_{ij} \|b\|_{BMO(\nu)},$$

which is nothing other than (5.2). We now turn to computing the norms of A_k and B_k .

We begin with A_k and again proceed by duality. We let $f \in L^p(\mu)$ and $g \in L^q(\lambda')$ and appeal to H^1 -BMO duality, as expressed in (2.12), to get the BMO_ν norm.

$$|\langle A_k f, g \rangle| = |\langle b, \Phi \rangle| \lesssim [\nu]_{A_2} \|b\|_{BMO_\nu^2(\nu)} \|S_{\mathcal{D}} \Phi\|_{L^1(\nu)},$$

where, as before, Φ is a bilinear expression involving f and g .

$$\Phi := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{P \in R^{(i)} \\ N \in R^{(j-k)}}} \widehat{f}(P, \epsilon) \left(\sum_{Q \in N^{(k)}} a_{PQR}^{\epsilon \eta} h_N^\gamma(Q) \widehat{g}(Q, \eta) \right) h_N^\gamma.$$

This is a Haar series, and we pass to its square function, summing over the cubes N .

$$(S_{\mathcal{D}}\Phi)^2 \lesssim 2^{-n(i+j)} \sum_{\substack{N \in \mathcal{D} \\ \epsilon, \eta \neq 1}} \left(\sum_{P \in (N^{(j-k)})_{(i)}} |\widehat{f}(P, \epsilon)| \sum_{Q \in N_{(k)}} \frac{1}{\sqrt{|N|}} |\widehat{g}(Q, \eta)| \right)^2 \frac{\mathbb{1}_N}{|N|}.$$

The term $2^{-n(i+j)}$ comes from the decay of the Haar shift coefficients. The sum involving g is bounded by

$$\sum_{Q \in N_{(k)}} \frac{1}{\sqrt{|N|}} |\widehat{g}(Q, \eta)| \leq \frac{2^{\frac{kn}{2}}}{|N|} \sum_{Q \in N_{(k)}} \int_Q |g| dx = 2^{\frac{kn}{2}} \langle |g| \rangle_N.$$

So the square function is bounded by

$$\begin{aligned} (S_{\mathcal{D}}\Phi)^2 &\lesssim 2^{-n(i+j-k)} (Mg)^2 \sum_{N \in \mathcal{D}, \epsilon \neq 1} \left(\sum_{P \in (N^{(j-k)})_{(i)}} |\widehat{f}(P, \epsilon)| \right)^2 \frac{\mathbb{1}_N}{|N|} \\ &= 2^{-n(i+j-k)} (Mg)^2 \left(\widetilde{S}_{\mathcal{D}}^{i, j-k} f \right)^2. \end{aligned}$$

The maximal function is controlled by Muckenhoupt's bound, and the square function by the estimate (2.7). We have

$$\begin{aligned} \|S_{\mathcal{D}}f\|_{L^1(\nu)} &\lesssim 2^{-\frac{n}{2}(i+j-k)} \int_{\mathbb{R}^n} (Mg) (\widetilde{S}_{\mathcal{D}}^{i, j-k} f) d\nu \\ &\leq 2^{-\frac{n}{2}(i+j-k)} \|Mg\|_{L^q(\lambda')} \left\| \widetilde{S}_{\mathcal{D}}^{i, j-k} f \right\|_{L^p(\mu)} \\ &\lesssim 2^{-\frac{n}{2}(i+j-k)} \|g\|_{L^q(\lambda')} 2^{\frac{n}{2}(i+j-k)} \|f\|_{L^p(\mu)} = \|f\|_{L^p(\mu)} \|g\|_{L^q(\lambda')}. \end{aligned}$$

This completes the proof of $\|A_k : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \|b\|_{BMO(\nu)}$.

Similarly for B_k :

$$|\langle B_k f, g \rangle| = |\langle b, \Phi \rangle| \lesssim \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \|S_{\mathcal{D}}\Phi\|_{L^1(\nu)},$$

where

$$\Phi := \sum_{\substack{R \in \mathcal{D} \\ \epsilon, \eta, \gamma \neq 1}} \sum_{\substack{N \in R_{(i-k)} \\ Q \in R_{(j)}}} \left(\sum_{P \in N_{(k)}} \widehat{f}(P, \epsilon) h_N^\gamma(P) a_{PQR}^{\epsilon\eta} \right) \widehat{g}(Q, \eta) h_N^\gamma.$$

The analysis of the square function $S_{\mathcal{D}}\Phi$ is symmetric with respect to the roles of f and g . The proof is analogous, and so omitted.

5.2. Remainder Estimate for $i = j = 0$. A precise analysis of the case $i = j = 0$ in Theorem 2.3 is given in [14], where it is shown that \mathbb{S}^{00} is of the form

$$\mathbb{S}^{00} = \mathbb{S}_c^{00} + \Pi_a + \Pi_d^*,$$

where \mathbb{S}_c^{00} is a *cancellative* dyadic shift with parameters $(0, 0)$, and Π_a, Π_d^* are paraproducts with symbols $a, d \in BMO_{\mathcal{D}}$ with $\|a\|_{BMO_{\mathcal{D}}} \leq 1$; $\|d\|_{BMO_{\mathcal{D}}} \leq 1$. (The definition of the paraproduct is in (3.1).) Here $BMO_{\mathcal{D}}$ denotes the *unweighted* dyadic BMO space. The functions a and d come from the T1 theorem of David-Journé. The remainder \mathcal{R}^{00} then has the form

$$\mathcal{R}^{00} = \mathcal{R}_c^{00} + \mathcal{R}_a + \mathcal{R}_d^*,$$

where

$$\mathcal{R}_c^{00} f := \Pi_{\mathbb{S}_c^{00}} f b - \mathbb{S}_c^{00} \Pi_f b; \quad \mathcal{R}_a f := \Pi_{\Pi_a f} b - \Pi_a \Pi_f b; \quad \mathcal{R}_d^* f := \Pi_{\Pi_d^* f} b - \Pi_d^* \Pi_f b.$$

Now, \mathbb{S}_c^{00} is cancellative, so

$$\mathbb{S}_c^{00} f = \sum_{R \in \mathcal{D}} \sum_{\epsilon, \eta \neq 1} a_R^{\epsilon \eta} \widehat{f}(R, \epsilon) h_R^\eta,$$

for some $|a_R^{\epsilon \eta}| \leq 1$. It is easy to check that

$$\Pi_{\mathbb{S}_c^{00}} f b = \mathbb{S}_c^{00} \Pi_f b = \sum_{R \in \mathcal{D}} \sum_{\epsilon, \eta \neq 1} a_R^{\epsilon \eta} \widehat{f}(R, \epsilon) \langle b \rangle_R h_R^\eta.$$

So the term $\mathcal{R}_c^{00} = 0$, and we only need to look at \mathcal{R}_a and \mathcal{R}_d^* .

We recall the A_p bounds for paraproduct operators, which is classical. Namely, for $a \in BMO_{\mathcal{D}}$ and a weight $w \in A_p$ with $1 < p < \infty$ we have:

$$(5.5) \quad \|\Pi_a f\|_{L^p(w)} \lesssim \|a\|_{BMO_{\mathcal{D}}} \|f\|_{L^p(w)}.$$

Let us look at the term $\mathcal{R}_a f$:

$$\mathcal{R}_a f = \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \sum_{R \supseteq Q, \eta \neq 1} \widehat{f}(R, \eta) \left[\langle b \rangle_Q - \langle b \rangle_R \right] h_R^\eta(Q).$$

We write

$$\langle b \rangle_Q - \langle b \rangle_R = \sum_{\substack{N \in \mathcal{D} \\ Q \subsetneq N \subseteq R}} \sum_{\gamma \neq 1} \widehat{b}(N, \gamma) h_N^\gamma(Q),$$

and express $\mathcal{R}_a f$ as a sum of three terms, which we analyze separately. Specifically, we look at the cases $N \subsetneq R$, $N = R$ with $\gamma \neq \eta$ and $N = R$ with $\gamma = \eta$. For the first case:

$$\begin{aligned} A &:= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \sum_{R \supseteq Q, \eta \neq 1} \sum_{Q \subsetneq N \subseteq R} \sum_{\gamma \neq 1} \widehat{b}(N, \gamma) \widehat{f}(R, \eta) h_N^\gamma(Q) h_R^\eta(Q) \\ &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \sum_{N \supseteq Q, \gamma \neq 1} \widehat{b}(N, \gamma) \langle f \rangle_N h_N^\gamma(Q) \\ &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) \langle \Pi_b f \rangle_Q h_Q^\epsilon = \Pi_a \Pi_b f. \end{aligned}$$

The second case similarly gives:

$$\begin{aligned} B &:= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \sum_{R \supseteq Q, \eta \neq 1} \sum_{\gamma \neq 1, \gamma \neq \eta} \widehat{b}(R, \gamma) \widehat{f}(R, \eta) h_R^\gamma(Q) h_R^\eta(Q) \\ &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) \langle \Gamma_b f \rangle_Q h_Q^\epsilon = \Pi_a \Gamma_b f. \end{aligned}$$

Here, Γ_b is defined in (3.2). Finally, the case $N = R, \gamma = \eta$ yields:

$$\begin{aligned} C &:= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \sum_{R \supseteq Q, \eta \neq 1} \widehat{b}(R, \eta) \widehat{f}(R, \eta) \frac{1}{|R|} \\ &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) h_Q^\epsilon \left(\langle \Pi_b^* f \rangle_Q - \frac{1}{|Q|} \sum_{P \subseteq Q, \eta \neq 1} \widehat{b}(P, \eta) \widehat{f}(P, \eta) \right) \\ &= \Pi_a \Pi_b^* f - \Lambda_{a,b} f, \end{aligned}$$

$$\text{where } \Lambda_{a,b} f := \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) \frac{1}{|Q|} \left(\sum_{P \subseteq Q, \eta \neq 1} \widehat{b}(P, \eta) \widehat{f}(P, \eta) \right) h_Q^\epsilon.$$

In summary

$$(5.6) \quad \mathcal{R}_a = \Pi_a \Pi_b + \Pi_a \Gamma_b + \Pi_a \Pi_b^* - \Lambda_{a,b}.$$

A similar analysis of \mathcal{R}_d^* shows that

$$(5.7) \quad \mathcal{R}_d^* = \Lambda_{d,b}^* - \Pi_b \Pi_d^* - \Gamma_b \Pi_d^* - \Pi_b^* \Pi_d^*,$$

$$\text{where } \Lambda_{d,b}^* f := \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{d}(Q, \epsilon) \widehat{f}(Q, \epsilon) \frac{1}{|Q|} \left(\sum_{P \subseteq Q, \eta \neq 1} \widehat{b}(P, \eta) h_P^\eta \right).$$

We need to compute the $L^p(\nu)$ -norm for the terms in (5.6) and (5.7). For the terms that involve the composition of paraproducts and Γ , first use the two-weight inequalities of Theorem 3.1, and then (5.5). It remains to show that $\Lambda_{a,b}$ is bounded.

Lemma 5.1. *Let $a \in BMO_{\mathcal{D}}$ and $b \in BMO_{\mathcal{D}}^2(\nu)$. These inequalities hold.*

$$(5.8) \quad \|\Lambda_{a,b} : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \|a\|_{BMO_{\mathcal{D}}} \|b\|_{BMO_{\mathcal{D}}^2(\nu)},$$

$$(5.9) \quad \|\Lambda_{a,b}^* : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \|a\|_{BMO_{\mathcal{D}}} \|b\|_{BMO_{\mathcal{D}}^2(\nu)}.$$

Proof. The proof of the first set of inequalities is given, with the other set following by similar reasoning. We argue by duality, so that the appeal to weighted H^1 -BMO duality is easy. For $f \in L^p(\mu)$, and $g \in L^{p'}(\lambda')$, we have

$$\begin{aligned} \langle \Lambda_{a,b} f, g \rangle &= \sum_{Q \in \mathcal{D}, \epsilon \neq 1} \widehat{a}(Q, \epsilon) \frac{1}{|Q|} \sum_{P \subseteq Q, \eta \neq 1} \widehat{b}(P, \eta) \widehat{f}(P, \eta) \widehat{g}(Q, \epsilon) \\ &= \sum_{P \in \mathcal{D}, \eta \neq 1} \widehat{b}(P, \eta) \widehat{f}(P, \eta) \Psi_P \end{aligned}$$

$$\text{where } \Psi_P := \sum_{Q \supseteq P, \epsilon \neq 1} \widehat{a}(Q, \epsilon) \frac{1}{|Q|} \widehat{g}(Q, \epsilon).$$

Recall that if we multiply the coefficients $\widehat{a}(Q, \epsilon)$ by choices of signs, we do not increase the BMO norm of a . The same remark applies to $b \in BMO(\nu)$. Therefore, since f and g are fixed, we are free to assume that *each individual summand above is non-negative*. This only requires that we modify the Haar wavelet expansions of b and a by choices of signs, but this fact is suppressed in the notation.

The key fact that this gives us is a control of the terms Ψ_P , namely

$$\Psi_P \leq \langle \Pi_a^* g \rangle_P \leq \inf_{x \in P} M(\Pi_a^* g)(x).$$

Here, we are using the adjoint paraproduct of applied to g , as defined in (3.1). And, then, using the weighted H^1 -BMO duality, as expressed in (2.11), we have

$$\langle \Lambda_{a,b} f, \cdot \rangle \leq \|b\|_{BMO_{\mathcal{D}}^2(\nu)} \|S\Phi\|_{L^1(\nu)},$$

where $\Phi := \sum_{P \in \mathcal{D}, \eta \neq 1} \widehat{f}(P, \eta) \Psi_P h_P^\epsilon.$

And, last of all, using the definition of μ and Hölder's inequality, and weighted inequalities for the maximal function and paraproduct operators,

$$\begin{aligned} \|S\Phi\|_{L^1(\nu)} &\lesssim \int S f \cdot M(\Pi_a^* g) \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}} dx \\ &\leq \|S f\|_{L^p(\mu)} \|M(\Pi_a^* g)\|_{L^{p'}(\lambda')} \\ &\lesssim \|a\|_{BMO_{\mathcal{D}}} \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')}. \end{aligned}$$

□

Now we may combine the results in (5.8) and (5.9) with the rest of the terms in (5.6) and (5.7), which are controlled by (5.5) and Theorem 3.1, and obtain: $\|\mathcal{R}_a : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \|b\|_{BMO(\nu)}$, and $\|\mathcal{R}_d^* : L^p(\mu) \rightarrow L^p(\lambda)\| \lesssim \|b\|_{BMO(\nu)}$. Then (5.2) for the non-cancellative case follows.

5.3. Characterization of Bloom BMO by Commutators with the Riesz Transforms. In this section we prove Theorem 1.2. Note the the first part follow directly from Theorem 1.1, and so it only remains to prove the lower bound. Suppose

$$\|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\| < \infty, \quad i = 1, \dots, n,$$

where R_i are the Riesz transforms. Then, since $[b, R_i R_j] = [b, R_i] R_j + R_i [b, R_j]$ and $R_i : L^r(w) \rightarrow L^r(w)$ is bounded for all $i = 1, \dots, n$ and all $w \in A_r$ with $1 < r < \infty$ [28], we have that $[b, K]$ is bounded $L^p(\mu) \rightarrow L^p(\lambda)$ for all K that are polynomials in the Riesz transforms.

We employ the standard computation in [6]. Let $\{Y_k\}$ be an orthonormal basis for the space of spherical harmonics of degree n . Then

$$\sum_k |Y_k(x)|^2 = c_n |x|^{2n},$$

and, by homogeneity,

$$Y_k(x - y) = \sum_{|\alpha|+|\beta|=n} a_{\alpha\beta}^k x^\alpha y^\beta,$$

where we are using standard multi-index notation. As shown in [6]:

$$(5.10) \quad |Q| \left| (b - \langle b \rangle_Q) \mathbb{1}_Q \right| (x) = \frac{1}{c_n} \sum_{k, \alpha, \beta} a_{\alpha\beta}^k x^\alpha \Gamma_Q(x) ([b, R^{(k)}] y^\beta \mathbb{1}_Q(y)) (x),$$

for all cubes Q centered at the origin, where $\Gamma_Q := \mathbb{1}_Q \operatorname{sgn}(b - \langle b \rangle_Q)$ and $R^{(k)}$ is the polynomial in the Riesz transforms associated with $Y_k(x) |x|^{-n}$. Note that, since Q is centered at the origin

$$|x^\alpha| \lesssim l(Q)^{|\alpha|},$$

for all $x \in Q$. Then from (5.10):

$$\begin{aligned}
 |Q| \left(\int_Q |b - \langle b \rangle_Q|^p d\lambda \right)^{\frac{1}{p}} &\lesssim \sum_{k,\alpha,\beta} l(Q)^{|\alpha|} \|[b, R^{(k)}]y^\beta \mathbb{1}_Q(y)\|_{L^p(\lambda)} \\
 &\lesssim \sum_{k,\alpha,\beta} l(Q)^{|\alpha|} \|[b, R^{(k)}] : L^p(\mu) \rightarrow L^p(\lambda)\|_{L^p(\lambda)} \|y^\beta \mathbb{1}_Q(y)\|_{L^p(\mu)} \\
 &\lesssim \sum_{\alpha,\beta} l(Q)^{|\alpha|} \left(\sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\| \right) l(Q)^{|\beta|} \mu(Q)^{\frac{1}{p}} \\
 &\lesssim |Q| \left(\sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\| \right) \mu(Q)^{\frac{1}{p}}.
 \end{aligned}$$

Since the argument is translation-invariant, we may conclude that

$$\left(\frac{1}{\mu(Q)} \int_Q |b - \langle b \rangle_Q|^p d\lambda \right)^{\frac{1}{p}} \lesssim \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|,$$

for all cubes $Q \subset \mathbb{R}^n$. Then

$$\mathbb{B}_1^{\mathcal{D}}(b, \mu, \lambda) \lesssim \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|$$

for all dyadic grids \mathcal{D} on \mathbb{R}^n . By Theorem 4.1,

$$(5.11) \quad b \in BMO_{\mathcal{D}}(\nu) \text{ with } \|b\|_{BMO_{\mathcal{D}}(\nu)} \lesssim \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|,$$

for all \mathcal{D} .

To see how this implies that $b \in BMO(\nu)$, recall that there exist 2^n dyadic grids \mathcal{D}^ω such that for any cube $Q \subset \mathbb{R}^n$ there is $Q^\omega \in \mathcal{D}^\omega$ such that

$$Q \subset Q^\omega \text{ and } l(Q^\omega) \leq 6l(Q).$$

See the proof in [18, Proof of Theorem 1.10]. Now from (5.11)

$$|\langle b \rangle_{Q^\omega} - \langle b \rangle_Q| \leq \frac{1}{|Q|} \int_Q |b - \langle b \rangle_{Q^\omega}| dx \lesssim \frac{\nu(Q^\omega)}{|Q|} \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|,$$

so

$$\int_Q |b - \langle b \rangle_Q| dx \lesssim \nu(Q^\omega) \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|.$$

But, using the doubling property of A_p weights

$$\nu(Q^\omega) \leq [\nu]_{A_2} \left(\frac{|Q^\omega|}{|Q|} \right)^2 \nu(Q) \lesssim \nu(Q),$$

hence

$$\sup_Q \left(\frac{1}{\nu(Q)} \int_Q |b - \langle b \rangle_Q| dx \right) \lesssim \sum_{i=1}^n \|[b, R_i] : L^p(\mu) \rightarrow L^p(\lambda)\|,$$

which is exactly the lower bound in (1.1).

5.4. **Proof of Corollary 1.3.** The first part of the proof will use the following duality statement, which can be found in [10]: $(H^1(\nu))^* \equiv BMO^2(\nu)$, in the sense that every element of $(H^1(\nu))^*$ is of the form

$$\Lambda_b : H^1(\nu) \ni h \mapsto \Lambda_b h := \int_{\mathbb{R}^n} b(x)h(x) dx,$$

for a unique $b \in BMO^2(\nu)$, with $\|b\|_{BMO^2(\nu)} \lesssim \|\Lambda_b\|$. In terms of maximal functions, the weighted Hardy space $H^1(\nu)$ is defined as follows: let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x) dx = 1$ and set $\varphi_r(x) := r^{-n}\varphi(x/r)$ for $r > 0$ and $x \in \mathbb{R}^n$. Then

$$H^1(\nu) := \{f \in \mathcal{S}'(\mathbb{R}^n) : f^* \in L^1(\nu)\}, \text{ with } \|f\|_{H^1(\nu)} := \|f^*\|_{L^1(\nu)},$$

where $f^*(x) := \sup_{r>0} |f * \varphi_r(x)|$ is the maximal function. There are many equivalent ways to define the weighted Hardy spaces – in terms of the square function, or in terms of an atomic decomposition – see [11, 29].

Now let $f := g_1(Tg_2) - (T^*g_1)g_2$. Then for $g_1 \in L^q(\lambda')$, $g_2 \in L^p(\mu)$, and any $b \in BMO^2(\nu)$:

$$(5.12) \quad \left| \int_{\mathbb{R}^n} b(x)f(x) dx \right| = \left| \int_{\mathbb{R}^n} g_1(x)[b, T]g_2(x) dx \right| \\ \leq \|g_1\|_{L^q(\lambda')} \|[b, T]g_2\|_{L^p(\lambda)} \\ \lesssim \|g_1\|_{L^q(\lambda')} \|g_2\|_{L^p(\mu)} \|b\|_{BMO^2(\nu)} \quad \text{by Theorem 1.1.}$$

Then clearly $f \in H^1(\nu)$, with $\|f\|_{H^1(\nu)} \lesssim \|g_1\|_{L^q(\lambda')} \|g_2\|_{L^p(\mu)}$.

The second statement follows identically as in [6], with the appropriate modifications. We consider the Banach space of functions $f \in L^1(\nu)$ which admit a decomposition as in (1.2), normed by

$$\|f\|_{H^1(\nu)} := \inf \left\{ \sum_{i=1}^n \sum_{j=1}^{\infty} \|g_j^i\|_{L^q(\lambda')} \|h_j^i\|_{L^p(\mu)} \right\},$$

where the infimum is over all possible decompositions of f . Part one of this corollary shows that this is a subspace of $H^1(\nu)$. Now (5.12) and Theorem 1.2 show that

$$\sup \left\{ \left| \int_{\mathbb{R}^n} b(x)f(x) dx \right| : \|f\|_{H^1(\nu)} = 1 \right\} \approx \|b\|_{BMO^2(\nu)},$$

which implies the norms $\|f\|_{H^1(\nu)}$ and $\|f\|_{H^1(\nu)}$ are equivalent (see [4] for the simple functional analysis argument that yields this). This completes the proof.

REFERENCES

- [1] N. Arcozzi, R. Rochberg, E. Sawyer, and B. D. Wick, *Bilinear forms on the Dirichlet space*, Anal. PDE **3** (2010), no. 1, 21–47. [↑1](#)
- [2] S. Bloom, *A commutator theorem and weighted BMO*, Trans. Amer. Math. Soc. **292** (1985), no. 1, 103–122. [↑1](#)
- [3] C. Cascante and J. M. Ortega, *On a characterization of bilinear forms on the Dirichlet space*, Proc. Amer. Math. Soc. **140** (2012), no. 7, 2429–2440. [↑1](#)
- [4] R. Coifman, P.-L. Lions, Y. Meyer, and S. Semmes, *Compensated compactness and Hardy spaces*, J. Math. Pures Appl. (9) **72** (1993), no. 3, 247–286. [↑1](#), [22](#)
- [5] R. R. Coifman and T. Murai, *Commutators on the potential-theoretic energy spaces*, Tohoku Math. J. (2) **40** (1988), no. 3, 397–407. [↑1](#)

- [6] R. R. Coifman, R. Rochberg, and Guido Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635. ↑[1](#), [2](#), [3](#), [20](#), [22](#)
- [7] L. Dalenc and Y. Ou, *Upper Bound for Multi-parameter Iterated Commutators* (2014), 1–25 pp., available at <http://arxiv.org/abs/1401.5994>. ↑[1](#), [3](#)
- [8] O. Dragičević, L. Grafakos, M. C. Pereyra, and S. Petermichl, *Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces*, Publ. Math **49** (2005), no. 1, 73–91. ↑[6](#)
- [9] S. H. Ferguson and M. T. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143–160. ↑[1](#)
- [10] J. García-Cuerva, *Weighted Hardy spaces*, Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979, pp. 253–261. ↑[22](#)
- [11] J. García-Cuerva and J. M. Martell, *Wavelet characterization of weighted spaces*, J. Geom. Anal. **11** (2001), no. 2, 241–264. ↑[9](#), [22](#)
- [12] J. B. Garnett and R. H. Latter, *The atomic decomposition for Hardy spaces in several complex variables*, Duke Math. J. **45** (1978), no. 4, 815–845. ↑[1](#)
- [13] I. Holmes, M. T. Lacey, and B. D. Wick, *Bloom’s Inequality: Commutators in a Two-Weight Setting* (2015), 1–9 pp., available at <http://arxiv.org/abs/1505.07947>. ↑[3](#)
- [14] T. Hytönen, *Representation of singular integrals by dyadic operators, and the A_2 theorem* (2011), available at <http://arxiv.org/abs/1108.5119>. ↑[3](#), [8](#), [17](#)
- [15] ———, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. ↑[8](#)
- [16] T. P. Hytönen, M. T. Lacey, H. Martikainen, T. Orponen, M. Reguera, E. T. Sawyer, and I. Uriarte-Tuero, *Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on A_p weighted spaces*, J. Anal. Math. **118** (2012), no. 1, 177–220. ↑[8](#)
- [17] T. Hytönen, C. Pérez, S. Treil, and A. Volberg, *Sharp weighted estimates for dyadic shifts and the A_2 conjecture*, Journal für die reine und angewandte Mathematik **2014** (2012), no. 687, 43–86. ↑[8](#)
- [18] T. Hytönen and C. Pérez, *Sharp weighted bounds involving A_∞* , Anal. PDE **6** (2013), no. 4, 777–818. ↑[21](#)
- [19] M. T. Lacey, *On the A_2 inequality for Calderón-Zygmund operators*, Recent advances in harmonic analysis and applications, Springer Proc. Math. Stat., vol. 25, Springer, New York, 2013, pp. 235–242. ↑[8](#)
- [20] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, *Multi-parameter Div-Curl lemmas*, Bull. Lond. Math. Soc. **44** (2012), no. 6, 1123–1131. ↑[1](#)
- [21] ———, *Multiparameter Riesz commutators*, Amer. J. Math. **131** (2009), no. 3, 731–769. ↑[1](#)
- [22] M. T. Lacey and E. Terwilleger, *Hankel operators in several complex variables and product BMO*, Houston J. Math. **35** (2009), no. 1, 159–183. ↑[1](#)
- [23] M.-Y. Lee, C.-C. Lin, and Y.-C. Lin, *A wavelet characterization for the dual of weighted Hardy spaces*, Proc. Amer. Math. Soc. **137** (2009), no. 12, 4219–4225. ↑[9](#)
- [24] Benjamin Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226. ↑[6](#)
- [25] B. Muckenhoupt and R. L. Wheeden, *Weighted bounded mean oscillation and the Hilbert transform*, Studia Math. **54** (1975/76), no. 3, 221–237. ↑[8](#), [13](#)

- [26] F. Nazarov, G. Pisier, S. Treil, and A. Volberg, *Sharp estimates in vector Carleson imbedding theorem and for vector paraproducts*, J. Reine Angew. Math. **542** (2002), 147–171. [↑1](#)
- [27] Z. Nehari, *On bounded bilinear forms*, Ann. of Math. (2) **65** (1957), 153–162. [↑1](#)
- [28] S. Petermichl, *The sharp weighted bound for the Riesz transforms*, Proc. Amer. Math. Soc. **136** (2008), no. 4, 1237–1249. [↑20](#)
- [29] J.-O. Strömberg and A. Torchinsky, *Weighted Hardy spaces*, Lecture Notes in Mathematics, vol. 1381, Springer-Verlag, Berlin, 1989. [↑22](#)
- [30] S. Treil, *Sharp A_2 estimates of Haar shifts via Bellman function* (2011), 1-23 pp., available at <http://arxiv.org/abs/1105.2252>. [↑8](#)
- [31] A. Uchiyama, *The factorization of H^p on the space of homogeneous type*, Pacific J. Math. **92** (1981), no. 2, 453–468. [↑1](#)
- [32] ———, *On the compactness of operators of Hankel type*, Tôhoku Math. J. (2) **30** (1978), no. 1, 163–171. [↑1](#)
- [33] S. Wu, *A wavelet characterization for weighted Hardy Spaces*, Rev. Mat. Iberoamericana **8** (1992), no. 3, 329–349. [↑9](#)

IRINA HOLMES, SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY STREET, ATLANTA, GA USA 30332-0160

E-mail address: irina.holmes@math.gatech.edu

MICHAEL T. LACEY, SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY STREET, ATLANTA, GA USA 30332-0160

E-mail address: lacey@math.gatech.edu

BRETT D. WICK, SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, 686 CHERRY STREET, ATLANTA, GA USA 30332-0160

E-mail address: wick@math.gatech.edu