# COMPACTNESS OF COMMUTATOR OF RIESZ TRANSFORMS IN THE TWO WEIGHT SETTING

MICHAEL LACEY AND JI LI

Abstract. We characterize the compactness of commutators in the Bloom setting. Namely, for a suitably non-degenerate Calderón–Zygmund operator $T$, and a pair of weights $\mu, \lambda \in A_p$, the commutator $[T, b]$ is compact from $L^p(\mu) \to L^p(\lambda)$ if and only if $b \in VMO_\nu$, where $\nu = (\mu/\lambda)^{1/p}$. This extends the work of the first author, Holmes and Wick. The weighted VMO spaces are different from the classical VMO space. In dimension $d = 1$, compactly supported and smooth functions are dense in $VMO_\nu$, but this need not hold in dimensions $d \geq 2$. Moreover, the commutator in the product setting with respect to little VMO space is also investigated.

## 1. Introduction

Let $\mu$ be a weight on $\mathbb{R}^d$, i.e. a function that is positive almost everywhere and is locally integrable. For $1 < p < \infty$, define $L^p(\mu)$ to be the space of functions $f$ satisfying $\|f\|_{L^p(\mu)} := \left(\int_{\mathbb{R}^d} |f(x)|^p \mu(x)dx\right)^{1/p} < \infty$.

In [1], Bloom considered the behavior of the commutator $[b, H] : L^p(\mu) \to L^p(\lambda)$, where $H$ is the Hilbert transform on $\mathbb{R}$. When $\mu = \lambda = 1$, it is well-known that the boundedness of $[b, H]$ is characterized by $b \in BMO(\mathbb{R})$ [2, 12], and the compactness of $[b, H]$ is characterized by $b \in VMO(\mathbb{R})$ [14]. Bloom worked out the setting for $\mu, \lambda \in A_p(\mathbb{R})$ and $\mu \neq \lambda$. He showed that for $1 < p < \infty$, two weights $\lambda, \mu \in A_p$ and $\nu = (\mu/\lambda)^{1/p}$, $[b, H]$ is bounded from $L^p(\mu)$ to $L^p(\lambda)$ if and only if the symbol $b$ is in the weighted BMO space $BMO_\nu(\mathbb{R})$ (for the definitions of $A_p$ weight and $BMO_\nu(\mathbb{R})$, see section 2). Recently, Holmes, Wick and the first author [4] established this characterization of two-weight boundedness for the commutator of Riesz transforms $[b, R_j]$ in $\mathbb{R}^d$, $d \geq 2$, $j = 1, \ldots, d$, using a new method via representation formula from Hytönen and decomposition via paraproducts. It was further studied by Lerner–Ombrosi–Rivera-Ríos [11] via sparse domination, and by Hytönen [7] via a new weak factorization technique (comparing to [15]).

2000 *Mathematics Subject Classification.* Primary: 42B20 Secondary: 42B25, 42B35.

*Key words and phrases.* Commutator, two weight, compactness, Riesz transform.

The first author is a 2020 Simons Fellow. Research supported in part by grant from the US National Science Foundation, DMS-1949206.

The second author is supported by ARC DP 170101060.

However, characterization of the two-weight compactness of $[b,R_j]$, $j = 1,\dots,d$, ($[b,H]$ in dimension 1) is missing up to now. We fill in this gap as follows.

**Theorem 1.1.** *Suppose* $1 < p < \infty$, *two weights* $\mu,\lambda \in A_p$ *and* $\nu = (\mu/\lambda)^{1/p}$. *Suppose that* $b \in BMO(\mathbb{R}^d)$, $R_j$ *is the* $j$*-th Riesz transform on* $\mathbb{R}^d$, $j = 1,2,\dots,d$. *Then we obtain that* $b \in VMO(\mathbb{R}^d)$ *if and only if the commutator* $[b,R_j] : L^p(\mu) \to L^p(\lambda)$ *is compact.*

Some approaches to this Theorem will not succeed. For instance, Uchiyama [15] used the fact that $VMO(\mathbb{R}^d)$ is the closure of $C_0^\infty(\mathbb{R}^d)$ (smooth functions with compact support) under the $BMO(\mathbb{R}^d)$ norm. However, this is not necessarily true in the two weight setting when $d \geq 2$, see §4. Another recent argument of Hytönen [8] shows that compactness can be extrapolated, recovering compactness of commutators in the one weight setting. It would be interesting to extend that argument to the two weight setting.

The sufficiency of Theorem 1.1 holds more generally. We note that this sufficiency argument holds not just for Riesz transforms but also for general Calderón–Zygmund operators $T$ with the associated kernel $K(x,y)$. We formulate it as follows. Recall that a Calderón-Zygmund operator $T$ (bounded on $L^2(\mathbb{R}^d)$) associated to a $\delta$-standard kernel $K(x,y)$ is an integral operator defined initially on $f \in C_0^\infty(\mathbb{R}^d)$:

$$T(f)(x) := \int_{\mathbb{R}^d} K(x,y)f(y)\,dy, \quad x \notin \operatorname{supp} f,$$

where $K(x,y)$ satisfies the size and smoothness estimates

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

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

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

**Theorem 1.2.** *Suppose* $T$ *is a Calderón–Zygmund operator as above,* $1 < p < \infty$, *two weights* $\mu,\lambda \in A_p$ *and* $\nu = (\mu/\lambda)^{1/p}$, *and* $b \in VMO(\mathbb{R}^d)$. *Then the commutator* $[b,T]$ *is compact from* $L^p(\mu)$ *to* $L^p(\lambda)$.

The main idea of proving this argument is to split $[b,T]$ into two parts, $A_\epsilon$ and $B_\epsilon$, where the norm of $A_\epsilon$ is at most $\epsilon$, and $B_\epsilon$ is a compact operator.

Next we provide the argument for the necessity of Theorem 1.1. We note that this necessity argument holds for general operators with the non-degenerate condition on the kernel (formulated in [7]). We state it as follows.

We say that the operator $T$ satisfies the *non-degenerate condition* if: *there exist positive constants* $c_0$ *and* $C_0$ *such that for every* $x \in \mathbb{R}^d$ *and* $r > 0$, *there exists* $y \in B(x,C_0r)\setminus B(x,r)$ *for which the kernel* $K(x,y)$ *satisfies*

$$|K(x,y)| \geq \frac{1}{c_0 r^d}. \tag{1.3}$$

**Theorem 1.4.** *Suppose* $1 < p < \infty$*, two weights* $\mu, \lambda \in A_p$ *and* $\nu = (\mu/\lambda)^{1/p}$*,* $b \in BMO_\nu(\mathbb{R}^d)$*. Suppose that* $T$ *satisfies the non-degenerate condition and that* $[b, T]$ *is compact from* $L^p(\mu)$ *to* $L^p(\lambda)$*. Then* $b \in VMO_\nu(\mathbb{R}^d)$*.*

The main idea of proving this argument is to seek a contradiction, which, in its simplest form, is that there is no bounded operator $T : \ell^p(\mathbb{N}) \to \ell^p(\mathbb{N})$ with $Te_j = Te_k \neq 0$ for all $j, k \in \mathbb{N}$. Here, $e_j$ is the standard basis for $\ell^p(\mathbb{N})$. Thus, the main step is to construct norm one, disjointly supported functions $\{g_j\} \subset L^p(\mu)$ such that $[b, T]g_j \approx \phi \neq 0$.

*Remark* 1.5. We point out that if $T$ is a Calderón–Zygmund operator on $L^2(\mathbb{R}^d)$ and $T$ satisfies the non-degenerate condition, then we obtain that there exist absolute constants $3 \leq A_1 \leq A_2$ such that for any ball $B = B(x_0, r)$, there is another ball $\tilde{B} := B(y_0, r)$ such that $A_1 r \leq |x - y| \leq A_2 r$, and for all $(x, y) \in (B \times \tilde{B})$, $K(x, y)$ does not change sign and (1.3) holds. In fact, this argument is enough for us to deduce the above theorem.

*Remark* 1.6. We also point out that our result and proof hold in a more general setting: spaces of homogeneous type, to cover many examples of Calderón–Zygmund operators beyond the Euclidean setting.

We now study the weighted VMO space $VMO_\nu(\mathbb{R}^d)$, $\nu \in A_2$, which is of independent interest with the unweighted case known around 40 years ago. We show that $VMO_\nu(\mathbb{R}^d)$ has totally different properties between $d = 1$ and $d > 1$.

**Theorem 1.7.** $VMO_\nu(\mathbb{R})$ *is the closure of* $C_0^\infty(\mathbb{R})$ *under the* $BMO_\nu(\mathbb{R})$ *norm. However, this is not necessarily true when* $d \geq 2$*.*

This paper is organized as follows. In Section 2 we provide the definitions for Muckenhoupt weights, weighted BMO and VMO spaces. In Section 3 we prove Theorem 1.1 via showing Theorems 1.2 and 1.4. In Section 4 we prove Theorem 1.7. In the last section, we discuss the two weight compactness in the product setting with respect to little VMO spaces.

Throughout this paper, we use the standard notation "$A \lesssim B$" to denote $A \leq CB$ for some positive constant $C$ that depends only on the dimension $d$.

## 2. Preliminaries

**Definition 2.1.** Let $w(x)$ be a nonnegative locally integrable function on $\mathbb{R}^d$. For $1 < p < \infty$, we say $w$ is an $A_p$-*weight*, written $w \in A_p(\mathbb{R}^d)$, if

$$[w]_{A_p} := \sup_B \left( \fint_B w \right) \left( \fint_B \left( \frac{1}{w} \right)^{1/(p-1)} \right)^{p-1} < \infty.$$

Along the way, we will need different properties of $A_p$ weights, which we will mention as they are needed.

Next we use $A_p$, $1 < p < \infty$, to denote the Muckenhoupt weighted class on $\mathbb{R}^d$ (see the precise definition of $A_p$ in Section 2), and the Muckenhoupt–Wheeden weighted BMO on $\mathbb{R}^d$ is defined as follows.

**Definition 2.2.** Suppose $w \in A_\infty$. A function $b \in L^1_{loc}(\mathbb{R}^d)$ belongs to $BMO_w(\mathbb{R}^d)$ if

$$\|b\|_{BMO_w(\mathbb{R}^d)} := \sup_B \frac{1}{w(B)} \int_B |b(x) - b_B|\, dx < \infty,$$

where $b_B := \frac{1}{|B|}\int_B b(x)dx$ and the supremum is taken over all balls $B \subset \mathbb{R}^d$.

The weighted VMO space on $\mathbb{R}^d$ is defined as follows.

**Definition 2.3.** Suppose $w \in A_\infty$. A function $b \in BMO_w(\mathbb{R}^d)$ belongs to $VMO_w(\mathbb{R}^d)$ if

$$\text{(i)}\ \lim_{a\to 0} \sup_{B:\, r_B = a} \frac{1}{w(B)} \int_B |b(x) - b_B|\, dx = 0;$$

$$\text{(ii)}\ \lim_{a\to \infty} \sup_{B:\, r_B = a} \frac{1}{w(B)} \int_B |b(x) - b_B|\, dx = 0;$$

$$\text{(iii)}\ \lim_{a\to \infty} \sup_{B\subset \mathbb{R}^d \setminus B(x_0,a)} \frac{1}{w(B)} \int_B |b(x) - b_B|\, dx = 0;$$

where $x_0$ is any fixed point in $\mathbb{R}^d$.

## 3. Proof of Main result: Theorem 1.1

It is clear that Theorem 1.1 follows from Theorem 1.2 and Theorem 1.4. In what follows, we provide the proofs of these two theorems.

*Proof of Theorem 1.2.* This is seen as follows. Suppose $b \in VMO(\mathbb{R}^d)$ with $\|b\|_{BMO(\mathbb{R}^d)} = 1$. We show that for any fixed $0 < \ < 1$, we have $[b, T] = A + B$, where the norm of $A$ is at most , and $B$ is a compact operator.

Fix $\ > 0$ small enough. Denote the kernel of $T$ by $K(x,y)$. Set $K = \sum_{t=0}^{3} K_t$, where each $K_j$ is a Calderón–Zygmund kernel, and

$$K_0(x,y) = \begin{cases} K(x,y); & 0 < |x-y| < \ ; \\ 0; & |x-y| > 2\ ; \end{cases}$$

$$K_1(x,y) = \begin{cases} K(x,y); & |x-y| > 1/\ ; \\ 0; & |x-y| < 1/(2\ ) \end{cases}$$

$$K_2(x,y),\ 0 \quad = ) \quad |x| > 1/\ \text{ or } |y| > 1/\ ;$$

$$K_3(x,y) \quad \text{is supported on } |x|, |y| < 2/\ :$$

Write $T_j$ for the operator associated to the kernel $K_j$. We claim that

$$\|[b, T_j]\|_{L^p(\,)\to L^p(\,)} \lesssim \ , \qquad j = 0, 1, 2;$$

where  decreases to zero as  does. Consider the case of $j = 1$. The contribution to the  norm estimate above from $BMO$ norm only arises from the weighted oscillation over the cubes of side length greater than 1=2. Indeed, write out the operator as

$$\int (b(x) - b(y))K_1(x; y)f(y)\, dy = \int_{|x-y|\geq 1=2} (b(x) - b(y))K_0(x; y)f(y)\, dy:$$

The oscillation of the symbol on intervals of length smaller than  are irrelevant to the norm properties of the operator. Following any proof of the upper bound for the commutators in the Bloom setting will show that this operator has norm that decreases to zero in .

For $K_1$, the only contribution from b is the oscillation over cubes of side length at least 1=, and for $K_2$, only oscillations for cubes which are either large, or at least a distance 1= from the origin. In each case, the norm is small. We now choose  = () small enough, then we have $\|[b; T_j]\|_{L^p()\,!\,L^p(!)} .$  for $j = 0; 1; 2$.

It remains for us to argue that $[b; T_3]$ is a compact operator. This is not quite trivial, due to the commutator structure, and the delicate nature of the weighted $BMO$ space. For cubes $P; Q$ of the same side length, let $K_{P;Q}(x; y)$ be a smooth kernel supported on $P \times Q$, and with $|K_{P;Q}(x; y)| . |P|^{-2}$. It follows from elementary facts about $BMO$ that the commutator

$$C_{P;Q}f(x) = b(x) \int K_{P;Q}(x; y)f(y)\, dy - \int K_{P;Q}(x; y)b(y)f(y)\, dy$$

is bounded from $L^p() \,!\, L^p()$, with norm that depends only on the relative positions of P and Q. And, clearly, $C_{P;Q}$ has compact range. The operator $T_3$ can be well approximated by a finite sum $\sum_j C_{P_j;Q_j}$. Hence, $[b; T_3]$ is compact. The proof of Theorem 1.2 is complete.

*Proof of Theorem 1.4.* Now assume that $b \in BMO(\mathbb{R}^d)$ such that $[b; T]$ is compact from $L^p()$ to $L^p(!)$. But, for the sake of contradiction, further assume that $b < VMO(\mathbb{R}^d)$.

The main idea of getting contradiction is as follows: on a Hilbert space $\mathcal{H}$, with canonical basis $e_j$, $j \in \mathbb{N}$, an operator T with $Te_j = v$, with non-zero $v \in \mathcal{H}$, is necessarily unbounded. We will see that, for example when $p = 2$, a compact commutator from $L^2() \,!\, L^2(!)$ with symbol $b \in BMO(\mathbb{R}^d)$ n $VMO(\mathbb{R}^d)$ satisfies a variant of this condition.

Suppose that $b < VMO(\mathbb{R}^d)$, then at least one of the three conditions in Definition 2.3 does not hold. The argument is similar in all three cases, and we just present the case that the first condition in Definition 2.3 does not hold. Then there exist $_0 > 0$ and a sequence of balls $fB_jg_{j=1} = fB_j(x_j; r_j)g_{j=1} \subset \mathbb{R}^d$ such that $r_j \,!\, 0$ as $j \,!\, 1$ and that

$$(3.1)\qquad \frac{1}{(B_j)} \int_{B_j} jb(x) - b_{B_j}jdx \geq {}_0:$$

Without loss of generality, we can further assume that

$$(3.2)\qquad 4r_{j_{i+1}} \le r_{j_i}.$$

According to the non-degenerate condition (1.3): there exist constants $3 \le A_1 \le A_2$ and a ball $\tilde{B}_j := B(y_j, r_j)$ such that $A_1 r_j \le |x_j - y_j| \le A_2 r_j$, and for all $(x, y) \in (B_j \times \tilde{B}_j)$, $K(x, y)$ does not change sign and

$$(3.3)\qquad |K(x,y)| \gtrsim \frac{1}{r_j^d}.$$

Let $m_b(\tilde{B}_j)$ be a median value of $b$ on the ball $\tilde{B}_j$. Namely, $m_b(\tilde{B}_j)$ is a real number so that the two sets below have measure at least $\frac{1}{2}|\tilde{B}_j|$.

$$F_{j,1} \subset \{y \in \tilde{B}_j : b(y) \le m_b(\tilde{B}_j)\}; \qquad F_{j,2} \subset \{y \in \tilde{B}_j : b(y) \ge m_b(\tilde{B}_j)\}.$$

Next we define $E_{j,1} = \{x \in B : b(x) \ge m_b(\tilde{B}_j)\}$; $\quad E_{j,2} = \{x \in B : b(x) < m_b(\tilde{B}_j)\}$. Then $B_j = E_{j,1} \cup E_{j,2}$ and $E_{j,1} \cap E_{j,2} = \emptyset$. And it is clear that

$$b(x) - b(y) \ge 0, \quad (x,y) \in E_{j,1} \times F_{j,1}; \qquad b(x) - b(y) < 0, \quad (x,y) \in E_{j,2} \times F_{j,2}.$$

And for $(x, y)$ in $(E_{j,1} \times F_{j,1}) \cup (E_{j,2} \times F_{j,2})$, we have

$$|b(x) - b(y)| = |b(x) - m_b(\tilde{B}_j)| + |m_b(\tilde{B}_j) - b(y)| \ge |b(x) - m_b(\tilde{B}_j)|.$$

We now consider

$$\tilde{F}_{j,1} := F_{j,1} \setminus \bigcup_{\ell=j+1}^{\infty} \tilde{B}_\ell \quad \text{and} \quad \tilde{F}_{j,2} := F_{j,2} \setminus \bigcup_{\ell=j+1}^{\infty} \tilde{B}_\ell, \quad \text{for } j = 1, 2, \ldots.$$

Then, based on the decay condition of the measures of $\{B_j\}$ as in (3.2) we obtain that for each $j$,

$$(3.4)\qquad |\tilde{F}_{j,1}| \ge |F_{j,1}| - \Big|\bigcup_{\ell=j+1}^{\infty} \tilde{B}_\ell\Big| \ge \frac{1}{2}|\tilde{B}_j| - \sum_{\ell=j+1}^{\infty} |\tilde{B}_\ell| \ge \frac{1}{2}|\tilde{B}_j| - \frac{1}{3}|\tilde{B}_j| = \frac{1}{6}|\tilde{B}_j|.$$

Similar estimate holds for $\tilde{F}_{j,2}$.

Now for each $j$, we have that

$$\begin{aligned}
&\frac{1}{\mu(B_j)} \int_{B_j} |b(x) - b_{B_j}|\,dx \\
&\le \frac{2}{\mu(B_j)} \int_{B_j} \big|b(x) - m_b(\tilde{B}_j)\big|\,dx \\
&= \frac{2}{\mu(B_j)} \int_{E_{j,1}} \big|b(x) - m_b(\tilde{B}_j)\big|\,dx + \frac{2}{\mu(B_j)} \int_{E_{j,2}} \big|b(x) - m_b(\tilde{B}_j)\big|\,dx.
\end{aligned}$$

Thus, combining with (3.1) and the above inequalities, we obtain that as least one of the following inequalities holds:

$$\frac{2}{\mu(B_j)}\int_{E_{j,1}}\left|b(x)-m_b(B_j)\right|dx \ge \frac{\varepsilon_0}{2}; \qquad \frac{2}{\mu(B_j)}\int_{E_{j,2}}\left|b(x)-m_b(B_j)\right|dx \ge \frac{\varepsilon_0}{2}.$$

Without lost of generality, we now assume that the first one holds, i.e.,

$$\frac{2}{\mu(B_j)}\int_{E_{j,1}}\left|b(x)-m_b(B_j)\right|dx \ge \frac{\varepsilon_0}{2}.$$

Therefore, for each $j$, from (3.3) and (3.4) we obtain that

$$\begin{aligned}
\frac{\varepsilon_0}{4} &\le \frac{1}{\mu(B_j)}\int_{E_{j,1}}\left|b(x)-m_b(B_j)\right|dx\\
&\lesssim \frac{1}{\mu(B_j)}\frac{|\tilde F_{j,1}|}{|B_j|}\int_{E_{j,1}}\left|b(x)-m_b(B_j)\right|dx\\
&\lesssim \frac{1}{\mu(B_j)}\int_{E_{j,1}}\int_{\tilde F_{j,1}}|K(x,y)||b(x)-b(y)|\,dy\,dx.
\end{aligned}$$

Next, since for $x\in E_{j,1}$ and $y\in \tilde F_{j,1}$, $K(x,y)$ does not change sign and $b(x)-b(y)$ does not change sign either, we obtain that

$$\begin{aligned}
\varepsilon_0 &\lesssim \frac{1}{\mu(B_j)}\left|\int_{E_{j,1}}\int_{\tilde F_{j,1}}K(x,y)\big(b(x)-b(y)\big)\,dy\,dx\right|\\
&\lesssim \frac{1}{\mu(B_j)^{\frac{1}{p}}\lambda'(B_j)^{\frac{1}{p'}}}\int_{E_{j,1}}\left|[b,T](\chi_{\tilde F_{j,1}})(x)\right|dx\\
&= \frac{1}{\lambda'(B_j)^{\frac{1}{p'}}}\int_{E_{j,1}}\left|[b,T]\left(\frac{\chi_{\tilde F_{j,1}}}{\mu(B_j)^{\frac{1}{p}}}\right)(x)\right|dx;
\end{aligned}$$

where $\lambda'(x)=\lambda^{-\frac{1}{p-1}}(x)$, and in the last equality, we use $p'$ to denote the conjugate index of $p$.

Next, by using Hölder's inequality we further have

$$\begin{aligned}
\varepsilon_0 &\lesssim \frac{1}{\lambda'(B_j)^{\frac{1}{p'}}}\int_{E_{j,1}}|[b,T](f_j)(x)|\lambda^{\frac{1}{p}}(x)\lambda^{-\frac{1}{p}}(x)\,dx\\
&\lesssim \frac{1}{\lambda'(B_j)^{\frac{1}{p'}}}\lambda'(E_{j,1})^{\frac{1}{p'}}\left(\int_{\mathbb{R}^d}\left|[b,T](f_j)(x)\right|^p\lambda(x)\,dx\right)^{\frac{1}{p}}\\
&\lesssim \left(\int_{\mathbb{R}^d}\left|[b,T](f_j)(x)\right|^p\lambda(x)\,dx\right)^{\frac{1}{p}};
\end{aligned}$$

where in the above inequalities we denote

$$f_j := \frac{\mathbf{1}_{j\,1}}{(B\ )_j^{\frac{1}{p}}}.$$

This is a sequence of disjointly supported functions, by (3.4), with $\|f_j\|_{L^p()} \eqsim 1$.

Return to the assumption of compactness, and let  be in the closure of $\{[b;T](f_j)\}_j$. We have $\|\|_{L^p()} \gtrsim 1$. And, choose $j_i$ so that

$$\| - [b;T](f_{j_i})\|_{L^p()} \leq 2^{-i}.$$ We

then take non-negative numerical sequence $\{a_i\}$ with

$$\|\{a_i\}\|_{\ell^{p'}} < 1 \quad \text{but} \quad \|\{a_i\}\|_{\ell^1} = 1.$$

Then, $= \sum_i a_i f_{j_i} \in L^p()$, and

$$\Big\| \sum_i a_i - [b;T] \Big\|_{L^p()} \leq \sum_{i=1}^{1} a_i \big\| - [b;T](f_{j_i}) \big\|_{L^p()}$$

$$\leq \|a_i\|_{\ell^{p'}} \Big(\sum_i \| - [b;T](f_{j_i})\|_{L^p()}^p\Big)^{1/p} \lesssim 1.$$

So $\sum_i a_i \in L^p()$. But $\sum_i a_i$ is infinite on a set of positive measure. This is a contradiction that completes the proof.

## 4. Properties for VMO($\mathbb{R}^d$): proof of Theorem 1.7

In this section we prove Theorem 1.7. We split it into two subsections. In the first subsection we show that smooth compactly supported functions are dense in VMO in dimension one. In the second, we construct a counterexample: a nice function f, which is not even in BMO($\mathbb{R}^d$), $d \geq 2$.

4.1. VMO($\mathbb{R}$) **is the closure of** $C_0^1(\mathbb{R})$ **under the** BMO($\mathbb{R}$) **norm.** We provide the following characterization of VMO($\mathbb{R}$), which is parallel to the well-known result in the unweighted setting, however, it is new in this weighted setting.

**Theorem 4.1.** *For* $\in A_2(\mathbb{R})$*, we have* $\overline{C_0^1}^{BMO(\mathbb{R})} = VMO(\mathbb{R})$.

This elementary Lemma is needed.

**Lemma 4.2.** *Let* $\in A_2(\mathbb{R}^d)$*. Then,*

(1) *we have* $(\mathbb{R}^d) = 1$*;*

(2) *there is a* $0 < d^0 = d_{[]_{A_2}} < d$ *so that for any* $T > 1$*, we have*

$$\inf_{\substack{Q\,:\,Q\subset[-T;T]^d \\ \ell(Q)\leq 1}} \frac{(Q)}{\ell(Q)^{d+d^0}} > 0. \tag{4.3}$$

Standard examples show that the second result is optimal. Let $Q_T = [-T;T]^d$. We will systematically suppress the dependence of various constants on the $A_2$ constants of the weights.

*Proof.* For both, we argue by contradiction.

Assume $(R^d) < 1$. The $A_2$ product is always at least one. So, for all $k \in N$ we have $^{-1}(Q_{2^k}) \gtrsim 2^{2kd}$. And, so by equidistribution of $^{-1}$,

$$^{-1}(Q_{2^{(k+1)}} \cap Q_{2^k}) \gtrsim 2^{2kd}:$$

Now, $v^{-1}$ is an $A_2$ weight, so by Muckenhoupt's theorem, the maximal function $M(f)$ is bounded from $L^2(^{-1})$ to $L^2(^{-1})$, where $M$ is the standard Hardy–Littlewood maximal function on $R^d$. Apply it to $\mathbf{1}_{Q_1}$, so see that

$$\|M(\mathbf{1}_{Q_1})\|^2_{L(^{-1})} \gtrsim \sum_{k\in N} 2^{-2kd-1}(Q_{2^{k+1}} \cap Q_{2^k}) = 1:$$

This is a contradiction. So $(R^d) = 1$.

Given $\in A_2$, we have $^{-1} \in A_2$, so there is a $p > 1$ so that $^{-1}$ is in the Reverse Holder class $RH_p$. Choose $d^0$ so that $d^0 p = d$. Again, argue by contraction. Fix $T$ so that the infimum in (4.3) is zero. Then, we can find a sequence of cubes $Q_j \subset Q_T$ so that each $Q_j$ contains a set $E_j$ with $\{E_j\}$ being pairwise disjoint, and $2|E_j| > |Q_j|$. Finally, by equidistribution of $A_2$ weights, $(x) \leq \ell(Q_j)^{d^0}$ for $x \in E_j$. Then, it follows that

$$^{-p}(Q_T) \geq \sum_j \int_{E_j} {}^{-p}(x)\, dx \gtrsim \sum_j 1 = 1:$$

But, $^{-p}$ must be locally in $L^1$, so we have a contradiction, which yields (2) holds.

**Lemma 4.4.** *For* $\in A_2(R)$ *we have* $C_0^2(R) \subset VMO(R)$.

*Proof.* From Definition 2.3, for every $b \in C_0^2(R)$, it suffices to check the three conditions. The first of these conditions is that the contribution from oscillations on large scales tends to zero. This follows from compactness and $(R) = 1$.

The second concerns medium scales. Oscillations should be bounded, but that follows from $b$ being bounded. Third, oscillations at small scales should vanish. For $b \in C_0^2(R)$, we can assume that $\|Db\|_1 \leq 1$. We have

$$b(x) - b(y) = Db(y)\,(x - y) + O(|x - y|^2):$$

That is, in the direction of $\nabla b(y)$, the difference grows like $|x - y|$, but is otherwise of small order. Then, it follows that for a interval $I$ with side length at most one,

$$\frac{1}{|I|}\int_I\int_I |b(x) - b(y)|\, dy\, dx \lesssim |I|^2:$$

This integral has to be divided by $(I) \gtrsim \ell(I)^{1+d^0}$, where $0 < d^0 < 1$. But then from (4.3), the conclusion follows.

**Lemma 4.5.** *For $v \in A_2(\mathbb{R})$ we have* $\overline{C_0^2}^{BMO(\mathbb{R})} = \overline{C_0^1}^{BMO(\mathbb{R})}$.

*Proof.* For the proof, we need only show that $\overline{C_0^2}^{BMO(\mathbb{R})} \subset \overline{C_0^1}^{BMO(\mathbb{R})}$: Fix $b \in C_0^2(\mathbb{R})$, and a smooth non-negative compactly supported kernel $\varphi$, with integral equal to one. Set $\varphi_j(x) = 2^j \varphi(x2^j)$, for $j \in \mathbb{N}$. Then, $\varphi_j * b$ converges to $b$ in the $BMO(\mathbb{R})$ norm, as follows from the proof of Lemma 4.4. It is also in $C_0^1(\mathbb{R})$. Hence, $b \in \overline{C_0^1}^{BMO(\mathbb{R})}$.

*Proof of Theorem 4.1.* By Lemma 4.5, it suffices to verify that $\overline{C_0^2}^{BMO(\mathbb{R})} = VMO(\mathbb{R})$. But, the inclusion '$\subset$' is the content of Lemma 4.4. So we should show that every function $f \in VMO(\mathbb{R})$ can be approximated by a function in $C_0^2(\mathbb{R})$.

But using the notation $\varphi_j$ from the previous proof, we have $\varphi_j * b$ converges to $b$ in $BMO$ norm. And $\varphi * b \in \overline{C_0^2}^{BMO(\mathbb{R})}$ so the proof is complete.

4.2. **Nice functions are not necessarily in $VMO(\mathbb{R}^d)$ when $d > 1$.** We construct an example of a smooth function $g$ and weight $\in A_2$ such that $g$ is not even in $BMO(\mathbb{R}^d)$ when $d > 1$.

The point is that $A_2$ weights can vanish at a point, with the vanishing order allowed to be as large as $d$ on $\mathbb{R}^d$. There is no such requirement for smooth functions, of course. Let $d = 2$. Take $w(x) := |x|^{2-\alpha}$ for $\alpha \in (0, 1)$. Then $w \in A_2$ and moreover, we get

$$\frac{1}{|B_r|}\int_{B_r} w(x)\,dx = \frac{1}{|B_r|}\int_{B_r} |x|^{2-\alpha}\,dx \approx r^{2-\alpha}.$$

Take

$$g(x) := \begin{cases} x_1\, e^{-\frac{1}{(1-|x|)^2}}, & |x| < 1; \\ 0, & |x| \geq 1. \end{cases}$$

Then it is easy to see that $g \in C_0^1(\mathbb{R}^2)$, $\int_{B_r} g(x)\,dx = 0$ for any ball $B_r$ centered at the origin, and $\frac{1}{|B_r|}\int_{B_r} |g(x)|\,dx \approx r$ for $0 < r < 1$. (That is, $g$ has a zero of order 1 at the origin.)

The function $g(x)$ is not in $BMO_w(\mathbb{R}^2)$, since

$$\frac{1}{w(B_r)}\int_{B_r} |g(x) - g_{B_r}|\,dx = \frac{1}{w(B_r)}\int_{B_r} |g(x)|\,dx \approx \frac{1}{r^{1-\alpha}},$$

and this goes to $\infty$ as $r \to 0^+$.

The example prompts the referee to raise the question of identifying a class of smooth functions $\Phi$ on $\mathbb{R}^d$ for which $\overline{\Phi}^{BMO_w} = VMO_w$. A satisfactory answer here would make the discussion of the compactness of the operators in the two weight setting parallel to that of the unweighted theory. In dimensions $d \geq 2$, the description of $\Phi$ should linked to the reverse Hölder index of the weight $w$, and sensitive to the zeros of the weight. Doing so seems to require some insights which we do not currently have.

## 5. Product Setting: weighted little vmo space and compactness

We show that our main results and methods above can be applied to the product setting to characterize the two weight compactness for commutators with respect to little vmo spaces. Note that the characterization of the two weight boundedness of commutators with respect to little bmo spaces was obtained in [5]. For the sake of simplicity, we only consider the commutator of double Riesz transforms and a symbol b in little bmo, that is $[b; \mathrm{P}^{(1)}_j \mathrm{P}^{(2)}_k]$, where $\mathrm{P}^{(1)}_j$ is the jth Riesz transform on $R^{n_1}$ and $\mathrm{P}^{(2)}_k$ is the kth Riesz transform on $R^{n_2}$.

To begin with, we now recall the product $A_p(R^{n_1} \times R^{n_2})$ weights.

**Definition 5.1.** Let $w(x_1; x_2)$ be a nonnegative locally integrable function on $R^{n_1} \times R^{n_2}$. For $1 < p < 1$, we say w is a product $A_p$ *weight*, written as $w \in A_p(R^{n_1} \times R^{n_2})$, if

$$[w]_{A_p(R^{n_1} \times R^{n_2})} := \sup_R \left( - \!\!\!\!\! \int_R w \right) \left( - \!\!\!\!\! \int_R \left( \frac{1}{w} \right)^{\frac{1}{p-1}} \right)^{p-1} < 1 :$$

Here the supremum is taken over all rectangles $R := I_1 \times I_2 \subset R^{n_1} \times R^{n_2}$, where $I_i$ is a cube in $R^{n_i}$ for $i = 1; 2$. The quantity $[w]_{A_p(R^{n_1} \times R^{n_2})}$ is called the $A_p$ *constant of* w.

Next we recall the weighted little bmo and vmo spaces on $R^{n_1} \times R^{n_2}$.

**Definition 5.2.** For $1 < p < 1$ and $w \in A_p(R^{n_1} \times R^{n_2})$, the weighted little bmo space $bmo_w(R^{n_1} \times R^{n_2})$ is the space of all locally integrable functions b on $R^{n_1} \times R^{n_2}$ such that

$$\kappa b \kappa_{bmo_w(R^{n_1} \times R^{n_2})} = \sup_R \frac{1}{w(R)} \int_R j b(x_1; x_2) - b_R j dx_1 dx_2 < 1 ;$$

where the supremum is taken over all rectangles $R := I_1 \times I_2 \subset R^{n_1} \times R^{n_2}$, where $I_i$ is a cube in $R^{n_i}$ for $i = 1; 2$.

**Definition 5.3.** For $1 < p < 1$ and $w \in A_p(R^{n_1} \times R^{n_2})$, the weighted little vmo space $bmo_w(R^{n_1} \times R^{n_2})$ is the space of all b in $bmo_w(R^{n_1} \times R^{n_2})$ such that

(i) $\displaystyle \lim_{a ! 0} \sup_{R:\ diam(R)=a} \frac{1}{w(R)} \int_R j b(x_1; x_2) - b_R j dx_1 dx_2 = 0;$

(ii) $\displaystyle \lim_{a ! 1} \sup_{R:\ diam(R)=a} \frac{1}{w(R)} \int_R j b(x_1; x_2) - b_R j dx_1 dx_2 = 0;$

(ii) $\displaystyle \lim_{a ! 1} \sup_{R \subset R^d n B((x_0^{(1)}; x_0^{(2)}); a)} \frac{1}{w(R)} \int_R j b(x_1; x_2) - b_R j dx_1 dx_2 = 0;$

where $(x_0^{(1)}; x_0^{(2)})$ is any fixed point in $R^{n_1} \times R^{n_2}$. where the supremum is taken over all rectangles $R := I_1 \times I_2 \subset R^{n_1} \times R^{n_2}$, where $I_i$ is a cube in $R^{n_i}$ for $i = 1; 2$.

**Theorem 5.4.** *Suppose* $1 < p < 1$, *two weights* $;\ \in A_p(R^{n_1} \times R^{n_2})$ *and* $= (=)^{1=p}$. *Suppose that* $b \in bmo(R^{n_1} \times R^{n_2})$, $\mathrm{P}^{(1)}_j$ *is the* j*th Riesz transform on* $R^{n_1}$ *and*

$P_k^{(2)}$ *is the* $k$*th Riesz transform on* $\mathbb{R}^{n_2}$*,* $j = 1; 2; \ldots; n_1$; $k = 1; 2; \ldots; n_2$. *Then we obtain that* $b \in \mathrm{vmo}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ *if and only if the commutator* $[b; P_j^{(1)} P_k^{(2)}] : L^p() \to L^p()$ *is compact.*

We point out that when $b \in \mathrm{vmo}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, the compactness argument follows from the same idea and technique as in the proof of Theorem 1.2 with the splitting of the Riesz transforms $P_j^{(1)}$ and $P_k^{(2)}$ into four parts respectively. The reverse argument follows directly from the approach in the proof of Theorem 1.4. For the details, we omit here.

**Acknowledgment:** The authors would like to thank the referees for their careful reading, patient reviewing, valuable corrections and constructive comments, which improved the exposition of the manuscript.

## References


[1] S. Bloom, A commutator theorem and weighted BMO, Trans. Amer. Math. Soc., **292** (1985), 103–122.

[2] R. R. Coifman, R. Rochberg and G. Weiss, Factorization theorems for Hardy spaces in several variables, Ann. of Math., (2) **103** (1976), 611–635.

[3] A.G. de la Herrán and T. Hytönen, Dyadic representation and boundedness of nonhomogeneous Calderón-Zygmund operators with mild kernel regularity, Michigan Math. J., **67** (2018), no. 4, 757–786.

[4] I. Holmes, M. Lacey and B. D. Wick, Commutators in the two-weight setting, Math. Ann., **367** (2017), 51–80.

[5] I. Holmes, S. Petermichl and B.D. Wick, Weighted little bmo and two-weight inequalities for Journé commutators, Anal. PDE, **11** (2018), no. 7, 1693–1740.

[6] T. Hytönen, The sharp weighted bound for general Calderón–Zygmund operators, Ann. of Math., (2) **175** (2012), no. 3, 1473–1506.

[7] T. Hytönen, The $L^p \to L^q$ boundedness of commutators with applications to the Jacobian operator, arXiv:1804.11167.

[8] T. Hytönen, Extrapolation of compactness on weighted spaces, arXiv:2003.01606v2.

[9] T. Hytönen, M. Lacey, H. Martikainen, T. Orponen, M. Reguera, E. Sawyer, I. Uriarte-Tuero, Weak and strong type estimates for maximal truncations of Calderón–Zygmund operators on weighted spaces, J. Anal. Math., **118** (2012), no. 1, 177–220.

[10] A.K. Lerner, A simple proof of the $A_2$ conjecture, Int. Math. Res. Not. IMRN, 2013, no. 14, 3159–3170.

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

[12] Z. Nehari, On bounded bilinear forms, Ann. of Math., **65** (1957), 153–162.

[13] S. Treil, Sharp $A_2$ estimates of Haar shifts via Bellman function. (English summary) Recent trends in analysis, 187–208, Theta Ser. Adv. Math., 16, Theta, Bucharest, 2013.

[14] A. Uchiyama, On the compactness of operators of Hankel type, Tôhoku Math. J., **30** (1978), 163–171.

[15] A. Uchiyama, The factorization of $H^p$ on the space of homogeneous type, Pacific J. Math., **92** (1981), 453–468.

School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA
*Email address*: lacey@math.gatech.edu

Department of Mathematics, Macquarie University, NSW 2109, Australia
*Email address*: ji.li@mq.edu.au