Compactness of the Bloom sparse operators and applications

Peng Chen, Michael Lacey, Ji Li and Manasa N. Vempati

Abstract: We establish the characterization of compactness for the sparse operator (associated with symbol in weighted VMO space) in the two weight setting on the spaces of homogeneous type in the sense of Coifman and Weiss. As a direct application we obtain the compactness characterization for the maximal commutators with respect to the weighted VMO functions and the commutator of Calderón–Zygmund operators on the homogeneous spaces. Furthermore, our approach can be applied to compactness characterization for operators in the multilinear Bloom setting.

1 Introduction and statement of main results

In their remarkable result, Coifman–Rochberg–Weiss [7] showed that the commutator of Riesz transforms is bounded on $L^p(\mathbb{R}^n)$ if and only if the symbol $b$ is in the BMO space. See also the subsequent result by Janson [17] and Uchiyama [29]. Later, Bloom [4] obtained the two weight version of the commutator of Hilbert transform $H$ with respect to weighted BMO space. To be more precise, for $1 < p < \infty$, let $\lambda_1, \lambda_2$ be weights in the Muckenhoupt class $A_p$ and consider the weight $\nu = \lambda_1^{1/p} \lambda_2^{-1/p}$. Let $L^p_\nu(\mathbb{R})$ denote the space of functions that are $p$ integrable relative to the measure $w(x)dx$. Then, by [4], there exist constants $0 < c < C < \infty$, depending only on $p, \lambda_1, \lambda_2$, such that

$$c\|b\|_{\text{BMO}_\nu(\mathbb{R})} \leq \|[b, H] : L^p_{\lambda_1}(\mathbb{R}) \to L^p_{\lambda_2}(\mathbb{R})\| \leq C\|b\|_{\text{BMO}_\nu(\mathbb{R})}$$

in which $[b, H](f)(x) = b(x)H(f)(x) - H(bf)(x)$ denotes the commutator of the Hilbert transform $H$ and the function $b \in \text{BMO}_\nu(\mathbb{R})$, i.e., the Muckenhoupt–Wheeden weighted BMO space (introduced in [26], see also the definition in Section 2.4 below). This result provided a characterization of the boundedness of the commutator $[b, H] : L^p_{\lambda_1}(\mathbb{R}) \to L^p_{\lambda_2}(\mathbb{R})$ in terms of a triple of information $b, \lambda_1$ and $\lambda_2$. This result was extended very recently to the commutator of Riesz transform $[b, R_j], j = 1, \ldots, n$, in $\mathbb{R}^n$ by Holmes–Lacey–wick [14] using a different method involving the representation theorem for the Riesz transforms. Recently, Lerner–Ombrosi–Rivera-Ríos [23] also proved this result by using the sparse domination.

The compactness for $[b, H]$ (or $[b, R_j]$) in the Bloom setting was first obtained by the second and third authors [20], which is essentially different from the unweighted setting as studied by Uchiyama [28]. In the weighted case, $C_0^\infty(\mathbb{R}^n)$ need not be contained in the weight BMO space (and hence weighted VMO space) for $n \geq 2$. The proof in [20] relies on the split of Calderón–Zygmund operators into an essential part and the remainder, where the commutator of the remainder has operator norm arbitrarily small and the commutator of the essential part has finite range and hence compact.

1.1 Statement of main results

Inspired by the known result of the second author ([19])—pointwise domination of a Calderón–Zygmund operator via a corresponding sparse operator, and the sparse domination of commutator from Lerner–Ombrosi–Rivera-Ríos [23], it is quite natural to study the compactness
characterization for the Bloom sparse operators associated to the symbol in weighted VMO space.

Precisely, let $0 < \eta < 1$ and let $\mathcal{S}$ be an arbitrary $\eta$-sparse family of dyadic cubes on a space of homogeneous type $(X, d, \mu)$ (details will be provided in Section 2.5), such that $\mathcal{S} \subset \mathcal{D}$, where $\mathcal{D}$ is an arbitrary dyadic system in $X$ (as constructed by Christ [5]). Suppose $b \in \text{BMO}_\nu(X)$ with $\nu \in A_2$. Recall that the Bloom sparse operator associated to $b$ and $\mathcal{S}$, $\mathcal{T}_{\mathcal{S}, b}$ is defined as follows ([23])

$$\mathcal{T}_{\mathcal{S}, b}(f)(x) = \sum_{Q \in \mathcal{S}} |b(x) - b_Q| f_Q \chi_Q(x), \quad \forall f \in L^1_{\text{loc}}(X). \quad (1.1)$$

Along the line of [20], it is natural to consider the following question:

**Q:** Suppose $p \in (1, \infty)$, $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1} \frac{1}{\lambda_2} - \frac{1}{p}$. Suppose $b \in \text{BMO}_\nu(X)$.

(i) If $b \in \text{VMO}_\nu(X)$, then for every $0 < \eta < 1$ and for every $\eta$-sparse family $\mathcal{S}$ in $X$, the Bloom sparse operator $\mathcal{T}_{\mathcal{S}, b}$ as given in (1.1) is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$;

(ii) If for every $0 < \eta < 1$ and for every $\eta$-sparse family $\mathcal{S}$ in $X$, the Bloom sparse operator $\mathcal{T}_{\mathcal{S}, b}$ as given in (1.1) is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$, then we deduce that $b \in \text{VMO}_\nu(X)$.

As an important singular integral of independent interest and an effective tool to prove Part (ii) in Theorem 1.1 above, we will now recall the maximal commutator $C_b$ as studied by Garcia-Cuerva et al in [13, Theorem 2.4], which is bounded from $L^p_{\lambda_1}(\mathbb{R}^n)$ to $L^p_{\lambda_2}(\mathbb{R}^n)$, $1 < p < \infty$, if and only if $b \in \text{BMO}_\nu(\mathbb{R}^n)$ with $\nu = \lambda_1^{1/p} \lambda_2^{-1/p}$ and $\lambda_1, \lambda_2 \in A_p$. This has been revisited by Agcayazi et al [1] in the unweighted $\mathbb{R}^n$ using a different approach and also been generalised to space of homogeneous type by Hu–Yang [15] and by Fu et al [11].

The maximal commutator $C_b$ on $X$ with the symbol $b(x)$ is defined by

$$C_b(f)(x) := \sup_{B \ni x} \frac{1}{\mu(B)} \int_B |b(x) - b(y)||f(y)|d\mu(y),$$

where the supremum is taken over all balls $B \subset X$. We now establish the following compactness characterization for the maximal commutator.

**Theorem 1.2.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1} \frac{1}{\lambda_2} - \frac{1}{p}$. Suppose $b \in L^1_{\text{loc}}(X)$. Then $C_b$ has the following compactness characterization:

(i) if $b \in \text{VMO}_\nu(X)$, then $C_b$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$;

(ii) if $C_b$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$, then $b \in \text{VMO}_\nu(X)$.

Our proof of Theorem 1.1 and Theorem 1.2 goes the following:

$$b \in \text{VMO}_\nu(X) \Rightarrow \forall \mathcal{S}, \mathcal{T}_{\mathcal{S}, b} \text{ compact} \Rightarrow C_b \text{ compact} \Rightarrow b \in \text{VMO}_\nu(X). \quad (1.2)$$

To be more precise, suppose $b \in \text{VMO}_\nu(X)$, we will first proceed to prove the Theorem 1.1 (i) by splitting $\mathcal{T}_{\mathcal{S}, b}$ into an essential part and a remainder term, where we will show that the remainder has norm sufficiently small and the essential part has a finite range and hence is
compact. Next, we prove the Theorem 1.2 (i) by showing that there exists a finite number of sparse families $S_j$ such that $C_b$ can be pointwise dominated by the sum of $T_{S_j,b}$. Hence, the first arrow in (1.2) holds. Then, we will show that if $C_b$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$, we have that $b \in \text{VMO}_\nu(X)$. Finally, if for all sparse families we have, $T_{S,b}$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$, then using the Sparse domination for the maximal commutator $C_b$ (Theorem 1.2 (ii)), we see that $C_b$ is also compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$, and hence Theorem 1.2 (ii) gives us that $b \in \text{VMO}_\nu(X)$.

1.2 Applications

As a direct application of our Theorem 1.1, we obtain the compactness of $[b,T]$ in the Bloom setting on space of homogeneous type where $T$ is a Calderón–Zygmund operator, which gives us an alternate proof of sufficiency part of the result by the second and third author [20].

**Theorem 1.3.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \lambda_1^{\frac{1}{p}}\lambda_2^{-\frac{1}{p}}$. Suppose $b \in L^1_{\text{loc}}(X)$, and that $T$ is a Calderón–Zygmund operator. Then the commutator $[b,T]$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$ if $b \in \text{VMO}_\nu(X)$.

Now consider the non-degeneracy condition on the kernel $K(x,y)$ of the operator $T$ below, which allows us to reverse argument for the compactness. There exist positive constant $c_0$ and $\overline{C}$ such that for every $x \in X$ and $r > 0$, there exists $y \in B(x,\overline{C}r) \setminus B(x,r)$ satisfying

$$|K(x,y)| \geq \frac{1}{c_0 \mu(B(x,r))}, \tag{1.3}$$

To be more precise,

**Theorem 1.4.** Suppose $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \lambda_1^{\frac{1}{p}}\lambda_2^{-\frac{1}{p}}$. Suppose $b \in L^1_{\text{loc}}(X)$, $T$ satisfies the non-degenerate condition (1.3) above, and $[b,T]$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$. Then we deduce that $b \in \text{VMO}_\nu(X)$.

As a corollary to Theorem 1.2, we also have the following two weight compactness argument for the commutator of the Hardy–Littlewood maximal function $Mf(x)$ on $X$ (whose $L^p$ boundedness was given by Bastero–Milman–Ruiz [2, Propositions 4 and 6])

$$Mf(x) := \sup_{B \ni x} \frac{1}{\mu(B)} \int_B |f(y)| \, d\mu(y),$$

where the supremum is taken over all balls $B \subset X$.

**Proposition 1.5.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \lambda_1^{\frac{1}{p}}\lambda_2^{-\frac{1}{p}}$. Suppose $b \in L^1_{\text{loc}}(X)$. Then the commutator $[b,M]$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$ if $b \in \text{VMO}_\nu(X)$.

The main idea used in our main theorem also works for the compactness argument for the sparse operator constructed by [6] in the multilinear weighted setting. The Bloom type bilinear sparse operators $T^B_{S,b}$ and $T^{B,*}_{S,b}$ associated to $b \in \text{VMO}_\nu(X)$, and $S$, are defined as

$$T^B_{S,b}(f,g)(x) = \sum_{Q \in S} b(x) - b_Q |f_Q g_Q| \chi_Q(x), \tag{1.4}$$

$$T^{B,*}_{S,b}(f,g)(x) = \sum_{Q \in S} \frac{1}{\mu(Q)} \int_Q |(b(x) - b_Q)f(x)| \, d\mu(x) g_Q \chi_Q(x), \quad \forall f,g \in L^1_{\text{loc}}(X). \tag{1.5}$$
Then the commutator of the bilinear Calderón–Zygmund operator can be dominated by $T_{S,b}^{B}$ and $T_{S,b}^{B,\ast}$, that is,

$$||b,T[1](f,g)(x)|| := |b(x)T(f,g)(x) - T(bf,g)(x)| \lesssim T_{S,b}^{B}(f,g)(x) + T_{S,b}^{B,\ast}(f,g)(x).$$

Similar domination holds for $[b,T]_{2}(f,g)(x)$.

Our main application to bilinear setting is as follows.

**Theorem 1.6.** Let $p_{1}, p_{2} \in (1, \infty)$, $1/p = 1/p_{1} + 1/p_{2}$ and $\lambda_{1}, \lambda_{2} \in A_{p_{1}}$, $w \in A_{p_{2}}$, $\nu := \lambda_{1}^{1/p} \lambda_{2}^{1/p}$, and $\hat{w} = \lambda_{2}^{1/p} \nu w$. If $b \in \text{VMO}_{\nu}(X)$, then for every $0 < \eta < 1$ and for every $\eta$-sparse family $S$ in $X$, the bilinear Bloom sparse operator $T_{S,b}^{B}$ and $T_{S,b}^{B,\ast}$ as given in (1.4) are compact from $L_{\nu}^{p_{1}}(X) \times L_{\hat{w}}^{p_{2}}(X)$ to $L_{\hat{w}}^{p}(X)$.

Throughout this paper we assume that $\mu(X) = \infty$ and that $\mu(\{x_{0}\}) = 0$ for every $x_{0} \in X$. Also, we denote by $C$ and $\tilde{C}$ positive constants which are independent of the main parameters, but they may vary from line to line. For every $p \in (1, \infty)$, we denote by $p'$ the conjugate of $p$, i.e., $1/p + 1/p' = 1$. If $f \leq Cg$ or $f \geq Cg$, we then write $f \lessgtr g$ or $g \lessgtr f$; and if $f \lessgtr g \lessgtr f$, we write $f \approx g$.

## 2 Preliminaries on Spaces of Homogeneous Type

We say that $(X,d,\mu)$ is a space of homogeneous type in the sense of Coifman and Weiss if $d$ is a quasi-metric on $X$ and $\mu$ is a nonzero measure satisfying the doubling condition. A quasi-metric $d$ on a set $X$ is a function $d : X \times X \to [0,\infty)$ satisfying (i) $d(x,y) = d(y,x) \geq 0$ for all $x, y \in X$; (ii) $d(x,y) = 0$ if and only if $x = y$; and (iii) the quasi-triangle inequality: there is a constant $A_{0} \in [1,\infty)$ such that for all $x, y, z \in X$,

$$d(x,y) \leq A_{0}[d(x,z) + d(z,y)].$$  \hspace{1cm} (2.1)

For any quasi-metric space $(X,d)$ that satisfies the geometric doubling property, there exists a positive integer $\tilde{A}_{0} \in \mathbb{N}$ such that any open ball $B(x,r) := \{y \in X : d(x,y) < r\}$ of radius $r > 0$ can be covered by at most $\tilde{A}_{0}$ balls $B(x_{i},r/2)$ of radius $r/2$. We say that a nonzero measure $\mu$ satisfies the doubling condition if there is a constant $C_{\mu}$ such that for all $x \in X$ and $r > 0$,

$$\mu(B(x,2r)) \leq C_{\mu}\mu(B(x,r)) < \infty,$$  \hspace{1cm} (2.2)

where $B(x,r)$ is the quasi-metric ball by $B(x,r) := \{y \in X : d(x,y) < r\}$ for $x \in X$ and $r > 0$. We point out that the doubling condition (2.2) implies that there exists a positive constant $n$ (the upper dimension of $\mu$) such that for all $x \in X$, $\lambda \geq 1$ and $r > 0$,

$$\mu(B(x,\lambda r)) \leq C_{\mu} \lambda^{n}\mu(B(x,r)).$$  \hspace{1cm} (2.3)

A subset $\Omega \subseteq X$ is open (in the topology induced by $d$) if for every $x \in \Omega$ there exists $\varepsilon > 0$ such that $B(x,\varepsilon) \subseteq \Omega$. A subset $F \subseteq X$ is closed if its complement $X \setminus F$ is open. The usual proof of the fact that $F \subseteq X$ is closed, if and only if it contains its limit points, carries over to the quasi-metric spaces. However, some open balls $B(x,r)$ may fail to be open sets, see [16, Sec 2.1].

Constants that depend only on $A_{0}$ (the quasi-metric constant) and $\tilde{A}_{0}$ (the geometric doubling constant) are referred to as geometric constants.
2.1 A System of Dyadic Cubes

We recall from [16] (see also the previous work by M. Christ [5], as well as Sawyer–Wheeden [27]) the system of dyadic cubes. In a geometrically doubling quasi-metric space \((X,d)\), a countable family

\[
\mathcal{D} = \bigcup_{k \in \mathbb{Z}} \mathcal{D}_k, \quad \mathcal{D}_k = \{Q^k_\alpha : \alpha \in \mathcal{A}_k\},
\]

of Borel sets \(Q^k_\alpha \subseteq X\) is called a system of dyadic cubes with parameters \(\delta \in (0,1)\) and \(0 < c_1 \leq C_1 < \infty\) if it has the following properties:

\[
X = \bigcup_{\alpha \in \mathcal{A}_k} Q^k_\alpha \quad \text{(disjoint union) for all } k \in \mathbb{Z}; \tag{2.4}
\]

if \(\ell \geq k\), then either \(Q^\ell_\beta \subseteq Q^k_\alpha\) or \(Q^k_\alpha \cap Q^\ell_\beta = \emptyset\): \(\tag{2.5}\)

for each \((k,\alpha)\) and each \(\ell \leq k\), there exists a unique \(\beta\) such that \(Q^\ell_\beta \subseteq Q^k_\alpha\); \(\tag{2.6}\)

for each \((k,\alpha)\) there exists at most \(M\) (a fixed geometric constant) \(\beta\) such that

\[
Q^{k+1}_\beta \subseteq Q^k_\alpha, \quad \text{and } Q^k_\alpha = \bigcup_{Q \in \mathcal{D}_{k+1}} Q; \tag{2.7}
\]

\[
B(x^k_\alpha, c_1 \delta^k) \subseteq Q^k_\alpha \subseteq B(x^k_\alpha, C_1 \delta^k) =: B(Q^k_\alpha); \tag{2.8}
\]

\[
\text{if } \ell \geq k \text{ and } Q^k_\beta \subseteq Q^k_\alpha, \text{ then } B(Q^\ell_\beta) \subseteq B(Q^k_\alpha). \tag{2.9}
\]

The set \(Q^k_\alpha\) is called a dyadic cube of generation \(k\) with center point \(x^k_\alpha \in Q^k_\alpha\) and side length \(\delta^k\). The interior and closure of \(Q^k_\alpha\) are denoted by \(\bar{Q}^k_\alpha\) and \(Q^k_\alpha\), respectively.

2.2 Adjacent Systems of Dyadic Cubes

In a geometrically doubling quasi-metric space \((X,d)\), a finite collection \(\{\mathcal{D}^t : t = 1, 2, \ldots, T\}\) of families \(\mathcal{D}^t\) is called a collection of adjacent systems of dyadic cubes with parameters \(\delta \in (0,1), 0 < c_1 \leq C_1 < \infty\) and \(1 \leq C < \infty\) if it has the following properties: individually, each \(\mathcal{D}^t\) is a system of dyadic cubes with parameters \(\delta \in (0,1)\) and \(0 < c_1 \leq C_1 < \infty\); collectively, for each ball \(B(x,r) \subseteq X\) with \(\delta^{k+3} < r \leq \delta^{k+2}, k \in \mathbb{Z}\), there exist \(t \in \{1, 2, \ldots, T\}\) and \(Q \in \mathcal{D}^t\) of generation \(k\) and with center point \(x^k_\alpha\) such that \(\rho(x^t_\alpha, x^k_\alpha) < 2A_0 \delta^k\) and

\[
B(x, r) \subseteq Q \subseteq B(x,Cr). \tag{2.10}
\]

We recall from [16] the following construction.

**Theorem 2.1.** Let \((X,d)\) be a geometrically doubling quasi-metric space. Then there exists a collection \(\{\mathcal{D}^t : t = 1, 2, \ldots, T\}\) of adjacent systems of dyadic cubes with parameters \(\delta \in (0,(96A_0^{-1})^{-1}), c_1 = (12A_0^{-1})^{-1}, C_1 = 4A_0^4\) and \(C = 8A_0^2 \delta^{-3}\). The center points \(x^k_\alpha\) of the cubes \(Q \in \mathcal{D}^t_k\) have, for each \(t \in \{1, 2, \ldots, T\}\), the two properties

\[
\rho(x^t_\alpha, x^k_\beta) \geq (4A_0^{-1})^{-1} \delta^k \quad (\alpha \neq \beta), \quad \min_{\alpha} \rho(x^t_\alpha, x^k_\alpha) < 2A_0 \delta^k \quad \text{for all } x \in X.
\]

We recall from [18, Remark 2.8] that the number \(T\) of the adjacent systems of dyadic cubes as in the theorem above satisfies the estimate

\[
T = T(A_0, \bar{A}_0, \delta) \leq A_0^6(A_0^4/\delta)^{\log_2 \bar{A}_0}. \tag{2.11}
\]

Also, we recall the following result on the smallness of the boundary.
Proposition 2.2. Suppose that $144A^2_\delta \leq 1$. Let $\mu$ be a positive $\sigma$-finite measure on $X$. Then the collection $\{\mathcal{D}^t : t = 1, 2, \ldots, T\}$ may be chosen to have the additional property that $\mu(\partial Q) = 0$ for all $Q \in \bigcup_{t=1}^T \mathcal{D}^t$.

2.3 Muckenhoupt $A_p$ Weights

Definition 2.3. Let $\omega(x)$ be a nonnegative locally integrable function on $X$. For $1 < p < \infty$, we say $\omega$ is an $A_p$ weight, written $\omega \in A_p$, if

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

Here the suprema is taken over all balls $B \subset X$. The quantity $[\omega]_{A_p}$ is called the $A_p$ constant of $\omega$. And $\int_B = \frac{1}{\mu(B)} \int_B$.

Next we note that for $\omega \in A_\infty$ the measure $\omega(x) d\mu(x)$ is a doubling measure on $X$. To be more precise, we have that for all $\lambda > 1$ and all balls $B \subset X$,

$$w(\lambda B) \leq \lambda^{np} [\omega]_{A_p} w(B), \quad (2.12)$$

where $n$ is the upper dimension of the measure $\mu$, as in (2.3).

We also point out that for $\omega \in A_\infty$, there exists $\gamma > 0$ such that for every ball $B$,

$$\mu\left( \left\{ x \in B : \omega(x) \geq \gamma \right\} \right) \geq \frac{1}{\gamma} \mu(B).$$

And this implies that for every ball $B$ and for all $\delta \in (0, 1)$,

$$\int_B w \leq C \left( \int_B w^\delta \right)^{1/\delta}; \quad (2.13)$$

see also [24].

Using the definition of $A_p$ weight and reverse H"older’s inequality, we can easily obtain the following standard properties.

Lemma 2.4. Let $\omega \in A_p(X)$, $p > 1$. Then there exists constants $\hat{C}_1, \hat{C}_2 > 0$ and $\sigma \in (0, 1)$ such that the following holds

$$\hat{C}_1 \left( \frac{\mu(E)}{\mu(B)} \right)^p \leq \frac{\omega(E)}{\omega(B)} \leq \hat{C}_2 \left( \frac{\mu(E)}{\mu(B)} \right)^\sigma$$

for any measurable set $E$ of a quasi metric ball $B$.

2.4 Weighted BMO spaces

Next we recall the definition of the weighted BMO space on space of homogeneous type, while we point out that the Euclidean version was first introduced by Muckenhoupt and Wheeden [26].

Definition 2.5. Suppose $\omega \in A_\infty$. A function $b \in L^1_{\text{loc}}(X)$ belongs to the weighted BMO space $BMO_w(X)$ if

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

where the suprema is taken over all quasi-metric balls $B \subset X$ and $b_B = \frac{1}{\mu(B)} \int_B b(y) d\mu(y)$. 

Also note that the following result, which is a weighted version of the John–Nirenberg theorem, appeared first in Muckenhoupt–Wheeden [26], where the Muckenhoupt $A_p$ characteristic was not tracked. It has been revisited again in [9, Theorem 4.2] with the modern techniques via sparse domination with a sharp quantitative estimate.

**Theorem 2.6 ([26, 9]).** Suppose $1 < p < \infty$ and $w \in A_p(X)$. Let $b \in \text{BMO}_w(X)$. Then for any $1 \leq r \leq p'$, we have

\[
\|b\|_{\text{BMO}_w(X)} \approx \|b\|_{\text{BMO}_{w,r}(X)} := \left( \sup_B \frac{1}{w(B)} \int_B |b(x) - b_B|^r w^{-r}(x) d\mu(x) \right)^{\frac{1}{r}}. \tag{2.14}
\]

In particular, we have $\|b\|_{\text{BMO}_w(X)} \leq \|b\|_{\text{BMO}_{w,r}(X)} \leq C_{\mu,p,r}[w]_{A_p}^{\max\{1, \frac{1}{p-r}\}} \|b\|_{\text{BMO}_w(X)}$, where the constant depends only on $\mu, p$ and $r$.

We recall the median value $\alpha_B(f)$ (see [3]): for any real valued function $f \in L^1_{\text{loc}}(X)$ and any ball $B \subset X$, $\alpha_B(f)$ is the real number such that

\[
\inf_{c \in \mathbb{R}} \frac{1}{\mu(B)} \int_B |f(x) - c| d\mu(x) = \frac{1}{\mu(B)} \int_B |f(x) - \alpha_B(f)| d\mu(x).
\]

Moreover, it is known that $\alpha_B(f)$ satisfies

\[
\mu(\{x \in B : f(x) > \alpha_B(f)\}) \leq \frac{\mu(B)}{2} \tag{2.15}
\]

and

\[
\mu(\{x \in B : f(x) < \alpha_B(f)\}) \leq \frac{\mu(B)}{2}. \tag{2.16}
\]

Denote by $\Omega(h, B)$ the standard mean oscillation

\[
\Omega(h, B) = \frac{1}{\mu(B)} \int_B |h(x) - b_B| d\mu(x).
\]

And it is easy to see that for any ball $B \subset X$,

\[
\Omega(h, B) \approx \frac{1}{\mu(B)} \int_B |h(x) - \alpha_B(b)| d\mu(x), \tag{2.17}
\]

where the implicit constants are independent of the function $b$ and the ball $B$.

We have the following definition of $\text{VMO}_\nu(X)$ as shown in [20].

**Definition 2.7.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, $\nu := \lambda_1^\frac{1}{p} \lambda_2^{-\frac{1}{p}}$ and $b \in \text{BMO}_\nu(X)$. Then $b \in \text{VMO}_\nu(X)$ if and only if $b$ satisfies the following three conditions:

1. \(\lim_{a \to 0} \sup_{B \subseteq X \atop r(B) = a} \frac{1}{\nu(B)} \int_B |b(x) - b_B| d\mu(x) = 0;\)

2. \(\lim_{a \to \infty} \sup_{B \subseteq X \atop r(B) = a} \frac{1}{\nu(B)} \int_B |b(x) - b_B| d\mu(x) = 0;\)

3. \(\lim_{a \to \infty} \sup_{B \subseteq X \atop d(x_0, B) > a} \frac{1}{\nu(B)} \int_B |b(x) - b_B| d\mu(x) = 0,
\]

where $d(x_0, B) = \inf_{x \in B} \{d(x, x_0) : x \in B\}$ for some fixed point $x_0$ in $X$. 


2.5 Sparse Operators on Spaces of Homogeneous Type

Let $\mathcal{D}$ be a system of dyadic cubes on $X$ as in Section 2.1. We recall the sparse family of dyadic cubes on spaces of homogeneous type as studied in [25, 8].

**Definition 2.8 ([8]).** Given $0 < \eta < 1$, a collection $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes is said to be $\eta$-sparse provided that for every $Q \in \mathcal{S}$, there is a measurable subset $E_Q \subset Q$ such that $\mu(E_Q) \geq \eta \mu(Q)$ and the sets $\{E_Q\}_{Q \in \mathcal{S}}$ have only finite overlap.

**Definition 2.9 ([8]).** Given $0 < \eta < 1$, a collection $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes is said to be $\eta$-sparse if for every cube $Q \in \mathcal{D}$,

$$\sum_{P \in \mathcal{S}, P \subset Q} \mu(P) \leq \frac{1}{\eta} \mu(Q).$$

Next, we recall the argument of equivalence of Definition 2.9 and Definition 2.8 on space of homogeneous type. We refer to the original argument on $\mathbb{R}^n$ in [22].

**Theorem 2.10 ([8]).** Given $0 < \eta < 1$ and a collection $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes, the following statements hold:

- If $\mathcal{S}$ is $\eta$-sparse, then $\mathcal{S}$ is $\frac{c}{\eta}$-Carleson, where $c \geq 1$ is an absolute constant;
- If $\mathcal{S}$ is $\frac{1}{\eta}$-Carleson, then $\mathcal{S}$ is $\eta$-sparse.

Note that in general, the doubling measure $\mu$ may not have reverse doubling property, that is, we may not have a uniform constant $c$ such that $\mu(B)^{-1} \mu(\lambda B) \gtrsim \lambda^c$. However, based on the property of sparse family, we have the following argument on reverse doubling within the sparse family.

**Corollary 2.11.** Given $0 < \eta < 1$ and an $\eta$-sparse family $\mathcal{S}$. We split $\mathcal{S}$ into a finite subfamilies $\mathcal{S}_i$ such that the sets $\{E_Q\}_{Q \in \mathcal{S}}$ are disjoint. Let $Q \in \mathcal{S}_i$ and $P$ is a child of $Q$ in $\mathcal{S}_i$. Then we have $\mu(P) \leq (1 - \eta) \mu(Q)$.

We now recall the well-known definition for sparse operator.

**Definition 2.12.** Given $0 < \eta < 1$ and an $\eta$-sparse family $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes. The sparse operators $A_{\mathcal{S}}$ is defined by

$$A_{\mathcal{S}} f(x) := \sum_{Q \in \mathcal{S}} f_Q(x).$$

Following the proof of [25, Theorem 3.1], we obtain that

$$\|A_{\mathcal{S}} f\|_{L^p_w(X)} \leq C_{\eta,n,p}[w]^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p_w(X)}, \quad 1 < p < \infty. \quad (2.18)$$

We recall the following result on space of homogeneous type in [8, Lemma 3.5], where the original version on $\mathbb{R}^n$ was due to [23].

**Lemma 2.13.** Let $\mathcal{D}$ be a dyadic system in $X$ and let $\mathcal{S} \subset \mathcal{D}$ be a $\gamma$-sparse family. Assume that $b \in L^1_{\text{loc}}(X)$. Then there exists a $\frac{\gamma}{2(\gamma+1)}$-sparse family $\hat{\mathcal{S}} \subset \mathcal{D}$ such that $\mathcal{S} \subset \hat{\mathcal{S}}$ and for every cube $Q \in \hat{\mathcal{S}}$,

$$|b(x) - b_Q| \leq C \sum_{R \in \hat{\mathcal{S}}, R \subset Q} \Omega(b, R)\chi_R(x) \quad (2.19)$$

for a.e. $x \in Q$. 

3 Equivalence of VMO Spaces

Here we are going to give equivalent characterization for weighted VMO spaces on $X$. Consider a dyadic system of cubes $\mathcal{D}$ on $X$. Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1^p} \lambda_2^\frac{1}{p}$. Denote $\lambda_1' = (\lambda_1)^{\frac{1}{p-1}}$ and $\lambda_2' = (\lambda_2)^{\frac{1}{p-1}}$. We now provide two new definitions for weighted VMO space on $X$ by $\text{VMO}_{\lambda_1,\lambda_2}(X)$ and $\text{VMO}_{\lambda_1',\lambda_2'}(X)$.

**Definition 3.1.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1^p} \lambda_2^\frac{1}{p}$ and $b \in \text{BMO}_\nu(X)$. Then $b \in \text{VMO}_{\lambda_1,\lambda_2}(X)$ if $b$ satisfies the following three conditions:

(i) $\lim_{a \to 0} \sup_{B \subseteq X \atop r(B) = a} \left( \frac{1}{\lambda_1(B)} \int_B |b(x) - b_B|^{p'} \lambda_2(\nu)(x) d\mu(x) \right)^{\frac{1}{p'}} = 0$;

(ii) $\lim_{a \to 1} \sup_{B \subseteq X \atop r(B) = a} \left( \frac{1}{\lambda_1(B)} \int_B |b(x) - b_B| \lambda_2(\nu)(x) d\mu(x) \right)^{\frac{1}{p}} = 0$;

(iii) $\lim_{a \to \infty} \sup_{B \subseteq X \atop d(x_0, B) > a} \left( \frac{1}{\lambda_1(B)} \int_B |b(x) - b_B| \lambda_2(\nu)(x) d\mu(x) \right)^{\frac{1}{p}} = 0$.

**Definition 3.2.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1^p} \lambda_2^\frac{1}{p}$ and $b \in \text{BMO}_\nu(X)$. Then $b \in \text{VMO}_{\lambda_1',\lambda_2'}(X)$ if $b$ satisfies the following three conditions:

(i) $\lim_{a \to 0} \sup_{B \subseteq X \atop r(B) = a} \left( \frac{1}{\lambda_2'(B)} \int_B |b(x) - b_B|^{p'} \lambda_1'(\nu)(x) d\mu(x) \right)^{\frac{1}{p'}} = 0$;

(ii) $\lim_{a \to 1} \sup_{B \subseteq X \atop r(B) = a} \left( \frac{1}{\lambda_2'(B)} \int_B |b(x) - b_B| \lambda_1'(\nu)(x) d\mu(x) \right)^{\frac{1}{p}} = 0$;

(iii) $\lim_{a \to \infty} \sup_{B \subseteq X \atop d(x_0, B) > a} \left( \frac{1}{\lambda_2'(B)} \int_B |b(x) - b_B| \lambda_1'(\nu)(x) d\mu(x) \right)^{\frac{1}{p}} = 0$.

We are going to show that both characterizations for the weighted VMO space given by Definition 2.7 and Definitions 3.1 and 3.2 on $X$ are equivalent. To be more precise, we have

**Proposition 3.3.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1^p} \lambda_2^\frac{1}{p}$. Then

$$\text{VMO}_\nu(X) = \text{VMO}_{\lambda_1,\lambda_2}(X) = \text{VMO}_{\lambda_1',\lambda_2'}(X).$$

To prove this equivalence we will prove the following Lemmas. The first one is the John–Nirenberg type argument in the Bloom setting.

**Lemma 3.4.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{1}{\lambda_1^p} \lambda_2^\frac{1}{p}$ and $b \in \text{BMO}_\nu(X)$. For any dyadic system in $\mathcal{D}$, and for any $Q \in \mathcal{D}$, we have that

$$\left( \frac{1}{\lambda_1(Q)} \int_Q |b(x) - b_Q|^{p} \lambda_2(\nu)(x) d\mu(x) \right)^{\frac{1}{p}} \lesssim \|b\|_{\text{BMO}_{\nu}^p(X)}$$

(3.1)
Suppose \( \Phi \) is a homogeneous type defined as \( f(x) \) as follows. We let \( \hat{b}(Q) = \sum_{Q \in D} \hat{b}(Q) h_Q \) and \( \Pi^* \) be the adjoint of \( \Pi \). Then

Next we argue that

The proof is due to duality, exploiting the \( H^1 \)-\( BMO \) duality inequality ([8]) to gain the term \( \|b\|_{BMO^2(\nu)} \). 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 \( f \in L^p(X) \) and \( g \in L^q(X) \). Then

\[
| \langle \Pi_b f, g \rangle | = \left| \sum_{Q \in D} \hat{b}(Q) \langle f \rangle_Q \hat{g}(Q) \right| =: | \langle b, \Phi \rangle |
\]

where \( \Phi := \sum_{Q \in D, \epsilon \neq 1} (f)_Q \hat{g}(Q, \epsilon) h^\epsilon_Q \) and \( S_D \Phi \) is the dyadic square function on spaces of homogeneous type defined as

\[
S_D f(x) := \left( \sum_{Q \in D} |\hat{f}(Q)|^2 \frac{1_{Q}(x)}{\mu(Q)} \right)^{\frac{1}{2}},
\]

of which the boundedness was studied in Theorem 6.2 in [8].

Next, we show that

In fact, we write

\[
\langle b, \Phi \rangle = \sum_{Q \in D} \hat{b}(Q) \hat{\phi}(Q)
\]
and define
\[
\Omega_k := \{x \in X : S_D \Phi(x) > 2^k\};
\]
\[
\tilde{\Omega}_k := \{x \in X : M_w(1_{\Omega_k})(x) > \frac{1}{2}\};
\]
\[
B_k := \{Q \in \mathcal{D} : w(Q \cap \Omega_k) > w(Q)/2, \ w(Q \cap \Omega_{k+1}) \leq w(Q)/2\},
\]
where \(M_w\) is the standard weighted Hardy–Littlewood maximal function on \(X\) given by
\[
M_w f(x) := \sup_{B \ni x} \frac{1}{w(B)} \int_B |f(y)| w(y) d\mu(y)
\]
with the supremum is taken over all balls \(B \subset X\). Then using Hölder’s inequality we have
\[
| \langle b, \Phi \rangle | \leq \left| \sum_k \sum_{Q \in B_k, Q \subset \tilde{\Omega}_k} \sum_{Q \in B_k, Q \subset \tilde{\tilde{Q}}} \tilde{b}(Q) \tilde{\phi}(Q) \right|
\]
\[
\leq \sum_k \sum_{Q \in B_k, Q \text{ maximal}} \left( \sum_{Q \in B_k, Q \subset \tilde{\tilde{Q}}} |\tilde{\phi}(Q)|^2 \frac{w(Q)}{\mu(Q)} \right)^{1/2} \left( \sum_{Q \in B_k, Q \subset \tilde{\tilde{Q}}} \left| \tilde{b}(Q) \right|^2 \frac{\mu(Q)}{w(Q)} \right)^{1/2}
\]
\[
\leq \|b\|_{BMO_k^2(\nu)} \sum_k \sum_{Q \in B_k, Q \text{ maximal}} \frac{w(Q)^{1/2}}{\tilde{\phi}(Q)^{1/2}} \left( \sum_{Q \in B_k, Q \subset \tilde{\tilde{Q}}} \left| \tilde{\phi}(Q) \right|^2 \frac{w(Q)}{\mu(Q)} \right)^{1/2}
\]
\[
\leq \|b\|_{BMO_k^2(\nu)} \sum_k \left( \sum_{Q \in B_k, Q \text{ maximal}} \frac{w(Q)^{1/2}}{\tilde{\phi}(Q)^{1/2}} \right) \left( \sum_{Q \in B_k, Q \subset \tilde{\tilde{Q}}} \left| \tilde{\phi}(Q) \right|^2 \frac{w(Q)}{\mu(Q)} \right)^{1/2}
\]
\[
\leq \|b\|_{BMO_k^2(\nu)} \sum_k \frac{w(\tilde{\Omega}_k)^{1/2}}{\left( \sum_{Q \in B_k} \left| \tilde{\phi}(Q) \right|^2 \frac{w(Q)}{\mu(Q)} \right)^{1/2}}.
\]
Now we claim that
\[
\left( \sum_{Q \in B_k} \left| \tilde{b}(Q) \right|^2 \frac{w(Q)}{\mu(Q)} \right)^{1/2} \leq C 2^k w(\tilde{\Omega}_k)^{1/2}.
\]  
(3.8)
In fact, by noting that
\[
\int_{\tilde{\Omega}_k \setminus \tilde{\Omega}_{k+1}} S_D \Phi(x)^2 w(x) dx \leq 2^{2k+2} w(\tilde{\Omega}_k)
\]
and that
\[
\int_{\tilde{\Omega}_k \setminus \tilde{\Omega}_{k+1}} S_D \Phi(x)^2 w(x) d\mu(x) \geq \sum_{Q \in B_k} \left| \tilde{b}(Q) \right|^2 \frac{w(Q \cap (\tilde{\Omega}_k \setminus \tilde{\Omega}_{k+1}))}{\mu(Q)} \geq \frac{1}{2} \sum_{Q \in B_k} \left| \tilde{b}(Q) \right|^2 \frac{w(Q)}{\mu(Q)},
\]
we obtain that the claim (3.8) holds. This yields (3.5).
Now, \(S_D \Phi\) is bilinear in \(f\) and \(g\), and is no more than
\[
(S_D \Phi(x))^2 = \sum_{Q \in \mathcal{D}} \left| \langle f, \Phi \rangle \right|^2 \left| \tilde{g}(Q) \right|^2 \frac{1}{\mu(Q)} \frac{w(x)}{1_Q(x)}
\]
exists some constants $p$ where

$$
\leq (Mf(x))^2 \sum_{Q \in D, \epsilon \neq 1} |g(Q, \epsilon)|^2 \frac{1_Q(x)}{\mu(Q)} = (Mf(x))^2 (S_D g(x))^2.
$$

A straightforward application of Hölder’s inequality, and bounds for the maximal and square functions will complete the proof.

$$
\| S_D \Phi \|_{L^1(\nu)} \leq \int_X (Mf(x))(S_D g(x)) \lambda_1(x)^\frac{1}{p} \lambda_2(x)^{-\frac{1}{p}} d\mu(x)
\leq \| Mf \|_{L^p_\nu(X)} \| S_D g \|_{L^{p'}_\nu(X)} \lesssim \| f \|_{L^p_\nu(X)} \| g \|_{L^{p'}_\nu(X)}.
$$

This gives us the proof of (3.3).

The second set of inequalities (3.4) are similar to the first, by a simple duality argument. Based on (3.3) and (3.4), we see that

$$
\| \Pi_b 1_Q \|_{L^p(\lambda_2)} + \| \Pi_b 1_Q \|_{L^p(\lambda_2)} \lesssim \| b \|_{BMO^p_\nu(\nu)} \lambda_1(Q)^\frac{1}{p},
\| \Pi_b 1_Q \|_{L^q(\lambda'_2)} + \| \Pi_b 1_Q \|_{L^q(\lambda'_2)} \lesssim \| b \|_{BMO^q_\nu(\nu)} \lambda'_2(Q)^\frac{1}{q}.
$$

Then, we have for any $Q \in D$:

$$
\left( \int_Q |b(x) - \langle b \rangle_Q|^p \lambda_2(x) d\mu(x) \right)^\frac{1}{p} = \| \mathbb{1}_Q (\Pi_b 1_Q - \Pi_b 1_Q) \|_{L^p(\lambda_2)}
\leq \| \Pi_b 1_Q \|_{L^p(\lambda_2)} + \| \Pi_b 1_Q \|_{L^p(\lambda_2)}
\lesssim \| b \|_{BMO^p_\nu(\nu)} \lambda_1(Q)^\frac{1}{p}.
$$

This shows that (3.1) holds. Similarly, we get that (3.2) holds.

Based on Lemma 3.4, we have the following

**Lemma 3.5.** Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \frac{\lambda_1^\frac{1}{p}}{\lambda_2^\frac{1}{p}}$ and $b \in BMO_\nu(X)$. Then there exists some constants $c_0, C_0 > 0$ such that for any dyadic system in $D$,

$$
c_0 \left( \frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| d\mu(x) \right) \leq \left( \frac{1}{\lambda'_2(Q)} \int_Q |b(x) - b_Q|^p \lambda'_1(x) d\mu(x) \right)^\frac{1}{p'}, \quad \forall Q \in D, (3.9)
$$

and

$$
\sup_{Q \in D} \left( \frac{1}{\nu(Q)} \int_Q |b(x) - b_Q|^p \lambda'_1(x) d\mu(x) \right)^\frac{1}{p'} \leq C_0 \sup_{Q \in D} \left( \frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| d\mu(x) \right), \quad (3.10)
$$

where $p'$ is the conjugate index of $p$. Similar result holds for the form of the left-hand side of (3.1).

**Proof.** Let us now begin the proof of the lemma. Observe that $\lambda'_1 = (\lambda_1)^{\frac{-1}{p' - 1}} = (\lambda_1)^{\frac{-p'}{p'}}$ and $\lambda'_2 = (\lambda_2)^{\frac{-1}{p' - 1}} = (\lambda_2)^{\frac{-p'}{p'}}$. By using Lemma 3.4, we have

$$
\left( \frac{1}{\lambda'_2(Q)} \int_Q |b(x) - b_Q|^p \lambda'_1(x) d\mu(x) \right)^\frac{1}{p'} \quad (3.11)
$$

Then,

$$
\left( \frac{1}{\lambda'_2(Q)} \int_Q |b(x) - b_Q|^p \lambda'_1(x) d\mu(x) \right)^\frac{1}{p'} \leq c_0 \left( \frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| d\mu(x) \right),
$$

for some $c_0 > 0$.

Thus, we have the desired result.
where the last step follows from the Standard weighted version of John–Nirenberg inequality in Theorem 2.6.

Now let us proceed to prove the other direction. Note that

\[
\frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| \, d\mu(x)
= \frac{1}{\nu(Q)} \int_Q \left( |b(x) - b_Q| \lambda_1^{1/p} f(x) \lambda_2^{1/p} g(x) \right) \, d\mu(x)
\leq \frac{\lambda_2^{1/p}}{\nu(Q)} \left( \frac{1}{\lambda_2^{1/p}} \int_Q |b(x) - b_Q| \lambda_1^{1/p} f(x) \, d\mu(x) \right)^{1/p} \left( \int_Q \lambda_2^{1/p} g(x) \, d\mu(x) \right)^{1/p}
\leq \frac{\lambda_1^{1/p} \lambda_2^{1/p}}{\nu(Q)} \left( \frac{1}{\lambda_2^{1/p}} \int_Q |b(x) - b_Q| \lambda_1^{1/p} f(x) \, d\mu(x) \right)^{1/p}.
\]

Since \( \lambda_1, \lambda_2 \in A_p, \nu = \lambda_1^{1/p} \lambda_2^{1/p} \in A_2 \), we have

\[
\frac{1}{\nu(Q)} \leq \nu^{-1}(Q) \mu(Q)^2 = \frac{\nu^{-1}(Q)}{\mu(Q)^2} \int_Q \lambda_1^{1/p} f(x) \lambda_2^{1/p} g(x) \, d\mu(x)
\leq \frac{1}{\mu(Q)^2} \left( \int_Q \lambda_1^{1/p} f(x) \, d\mu(x) \right)^{1/p} \left( \int_Q \lambda_2^{1/p} g(x) \, d\mu(x) \right)^{1/p}
\leq \frac{1}{\lambda_1^{1/p} \lambda_2^{1/p}}.
\]

This implies that

\[
\frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| \, d\mu(x) \leq \left( \frac{1}{\lambda_2^{1/p}} \int_Q |b(x) - b_Q| \lambda_1^{1/p} f(x) \, d\mu(x) \right)^{1/p}.
\]

This completes the proof of the lemma gives us the desired equivalence. \( \square \)

**Definition 3.6.** Let \( p \in (1, \infty) \) and \( \lambda_1, \lambda_2 \in A_p, \nu := \lambda_1^{1/p} \lambda_2^{1/p} \). We introduce the following 3 versions of weighted atoms:

1. \( \text{supp} \, a(x) \subset B, \int_B a(x) \, d\mu(x) = 0, ||a||_{L_p^2(B)} \leq \nu(B)^{-\frac{1}{p}}; \)
2. \( \text{supp} \, a(x) \subset B, \int_B a(x) \, d\mu(x) = 0, ||a||_{L_p^2(B)} \leq \lambda_1(B)^{-\frac{1}{p}}; \)
3. \( \text{supp} \, a(x) \subset B, \int_B a(x) \, d\mu(x) = 0, ||a||_{L_p^2(B)} \leq \lambda_2(B)^{-\frac{1}{p}}. \)

Then we define \( H^1_{\nu, \text{atom}}(X) = \{ f = \sum_j \beta_j a_j \}, \) where each \( a_j \) is an atom in the form (1) and \( \sum_j |\beta_j| < \infty \). Moreover, \( ||f||_{H^1_{\nu, \text{atom}}(X)} \) is taken to be the infimum of \( \sum_j |\beta_j| \) for all possible representation \( f = \sum_j \beta_j a_j \). Similarly one can define \( H^1_{\lambda_1, \lambda_2, \text{atom}}(X) \) and \( H^1_{\lambda_1, \lambda_2, \text{atom}}(X) \) that link the the atoms in Case (2) and Case (3), respectively.
Moreover, the dyadic version of atoms and atomic Hardy spaces associated with an arbitrary dyadic system $D$ in $X$ is defined via replacing the ball $B$ by a dyadic cube $Q \in D$ as in (1) – (3) above. We denote these dyadic atomic Hardy spaces by $H_{\nu,atom,d}(X)$, $H^{1}_{\lambda_1,\lambda_2,atom,d}(X)$ and $H^{1}_{\lambda'_1,\lambda'_2,atom,d}(X)$.

Lemma 3.7. Let $p \in (1, \infty)$ and $\lambda_1, \lambda_2 \in A_p$, $\nu := \lambda_1^{1/p} \lambda_2^{-1/p}$. Then $H^{1}_\nu(X) = H^{1}_{\lambda_1,\lambda_2,atom}(X) = H^{1}_{\lambda'_1,\lambda'_2,atom}(X)$.

Proof. By noting that the Hardy space is the sum of a finite dyadic Hardy spaces [18], it suffices to show the dyadic version associated with an arbitrary dyadic system $D$ in $X$. That is, it suffices to show $H^{1}_{\nu,d}(X) = H^{1}_{\nu,atom,d}(X) = H^{1}_{\lambda_1,\lambda_2,atom,d}(X) = H^{1}_{\lambda'_1,\lambda'_2,atom,d}(X)$.

For every $f \in H^{1}_{\nu,d}(X)$, we have that $S_{D}f \in L^{p}_{\lambda}(X)$. Hence,

$$f = \sum_{Q} \hat{f}(Q)h_{Q} = \sum_{Q \in B_{k}} \frac{\sum_{Q \subset Q'} \hat{f}(Q)h_{Q}}{\sum_{Q \subset Q'} \hat{f}(Q)h_{Q}} = \sum_{Q \in B_{k}} \beta_{k,Q} a_{k,Q},$$

where

$$\beta_{k,Q} = \lambda_2(Q)^{1/p} \left\| \left( \sum_{Q \subset Q'} \left| \hat{f}(Q) \right|^{2} \frac{1_{Q}(x)}{\mu(Q)} \right)^{1/2} \right\|_{L^{p}_{\lambda_{1}}(X)},$$

and

$$a_{k,Q} = \frac{1}{\beta_{k,Q}} \sum_{Q \subset Q'} \hat{f}(Q)h_{Q}. $$

It is easy to see that each $a_{k,Q}$ satisfies the support condition and cancellation condition. Now we have

$$\|a_{k,Q}\|_{L^{p}_{\lambda_{1}}(X)} = \sup_{\|g\|_{L^{p'}_{\lambda_{1}}(X)} = 1} |(a_{k,Q}, g)| = \sup_{\|g\|_{L^{p'}_{\lambda_{1}}(X)} = 1} \left| \frac{1}{\beta_{k,Q}} \sum_{Q \subset Q'} \hat{f}(Q)g(Q) \right|$$

$$= \sup_{\|g\|_{L^{p'}_{\lambda_{1}}(X)} = 1} \left| \frac{1}{\beta_{k,Q}} \int_{X} \sum_{Q \subset Q'} \hat{f}(Q)\hat{g}(Q) \frac{1_{Q}(x)}{\mu(Q)} \lambda_{1}^{1/p}(x) \lambda_{1}^{-1/p}(x) d\mu(x) \right|$$

$$\leq \sup_{\|g\|_{L^{p'}_{\lambda_{1}}(X)} = 1} \left| \frac{1}{\beta_{k,Q}} \int_{X} \left( \sum_{Q \subset Q'} \left| \hat{f}(Q) \right|^{2} \frac{1_{Q}(x)}{\mu(Q)} \right)^{1/2} \left( \sum_{Q \subset Q'} \left| \hat{g}(Q) \right|^{2} \frac{1_{Q}(x)}{\mu(Q)} \right)^{1/2} \lambda_{1}^{1/p}(x) \lambda_{1}^{-1/p}(x) d\mu(x) \right|$$

$$\leq \sup_{\|g\|_{L^{p'}_{\lambda_{1}}(X)} = 1} \left| \frac{1}{\beta_{k,Q}} \left\| \left( \sum_{Q \subset Q'} \left| \hat{f}(Q) \right|^{2} \frac{1_{Q}(x)}{\mu(Q)} \right)^{1/2} \right\|_{L^{p}_{\lambda_{1}}(X)} \right| \|S_{D}(g)\|_{L^{p'}_{\lambda_{1}}(X)}$$

$$\leq \lambda_2(Q)^{-1/p} \left\| \hat{f}(Q) \right\|_{L^{p}_{\lambda_{1}}(X)}.$$
they have equivalent norms.

This implies that

\[ H_{VMO} \]

Hence, we see that

\[ S_a \subset H_{\nu,d}(X) \]

The other direction is much simpler as we just need to check the uniform boundedness of \( S_D \) on each dyadic atom of the form in Case (2). To be more precise, let \( Q \) be a dyadic cube in \( D \) and \( \text{supp} \, a(x) \subset Q \), \( \int_Q a(x) \, dx = 0 \), \( \|a\|_{L^p_1(X)} \leq \lambda_1(Q)^{-\frac{1}{p}} \). Then by cancellation of \( a \), we see that

\[ S_D(a)(x) = \left[ \sum_{Q' \subset Q} |\hat{a}(Q')|^2 \frac{1_{Q'}(x)}{\mu(Q')} \right]^{\frac{1}{2}}. \]

Hence

\[ \|S_D(a)\|_{L^1_p(X)} = \|S_D(a)\|_{L^1_1(Q)} = \int_Q S_D(a)(x) \lambda_1^{\frac{1}{p}}(x) \lambda_2^{-\frac{1}{p'}}(x) \, d\mu(x) \]

\[ \leq \left( \int_Q S_D(a)(x) \lambda_1^{\frac{1}{p}}(x) \lambda_2^{-\frac{1}{p'}}(x) \, d\mu(x) \right)^{\frac{1}{p}} \left( \int_Q \lambda_1(x) \, d\mu(x) \right)^{\frac{1}{p'}} \]

\[ \lesssim \|a\|_{L^p_2(X)} \lambda_1(Q)^{\frac{1}{p}} \]

\[ \lesssim 1. \]

This implies that \( H_{\nu,d}(X) \subset H_{\nu,d}(X) \). Thus, we see that \( H_{\nu,d}(X) = H_{\nu,d}(X) \) and they have equivalent norms.

By using similar argument, we can obtain the equivalence of the other two Hardy spaces.

We now show Proposition 3.3.

**Proof of Proposition 3.3.** From Lemma 3.5, we see that the norms for the definitions of \( VMO_\nu(X) \), \( VMO_{\lambda_1^1,\lambda_2^1}(X) \) and \( VMO_{\lambda_1^1,\lambda_2^1}(X) \) are equivalent by noting that the Hardy space is the sum of a
finite dyadic Hardy spaces \cite{18}. Moreover, using the standard argument via tent space or discrete sequence spaces, we see that the dual of VMO$_\nu$(X) is $H^1_\nu(X)$, the dual of VMO$_{\lambda_1,\lambda_2}(X)$ is $H^{1}_{\lambda_1,\lambda_2}(X)$ and the dual of VMO$_{\lambda_1,\lambda_2}(X)$ is $H^{1}_{\lambda_1,\lambda_2}(X)$. While in Lemma 3.7 we see that the three Hardy spaces are equivalent. Hence, we obtain that the three VMO spaces are equivalent. \qed

4 Proof of Theorem 1.1

We begin to prove the Theorem 1.1. We denote by $T_{S,b}^*$ the adjoint operator of $T_{S,b}$.

Recall from \cite{12}, we have that $T_{S,b}$ is bounded from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$. So we have $T_{S,b}$ is compact if and only if $T_{S,b}^*$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$. Hence to prove Theorem 1.1, it is enough to show that $T_{S,b}^*$ is compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$.

Our approach can be briefly summarized as the following. We decompose $T_{S,b}^* f(x) = T_{\epsilon,N_i} f(x) + T_{\epsilon} f(x)$, for all $\epsilon > 0$.

We will show that for all $\epsilon > 0$, there exists $N_i$ such that $T_{\epsilon,N_i} f(x)$ is a sparse operator with finite range, i.e.,

$$T_{\epsilon,N_i} f(x) = \sum_{k=1}^{N_i} a_k \chi_{Q_k(x)}$$

and we will show that the norm of $T_{\epsilon} f(x)$ is at most $\epsilon$, i.e.,

$$\|T_{\epsilon} f(x)\|_{L^p_{\lambda_2}(X)} \leq \epsilon \|f\|_{L^p_{\lambda_1}(X)}.$$

Recall that $T_{S,b}^*([|f|])(x)$ is given by the following equation

$$T_{S,b}^*([|f|])(x) = \sum_{Q \in S} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x). \quad (4.1)$$

For $\epsilon > 0$, from Definitions 2.7, 3.1, 3.2 and Proposition 3.3, we choose number $N > 0$, $\delta > 0$ and cube $Q_N$ side length $N$ such that

$$\frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| d\mu(x) < \epsilon, \quad \left( \frac{1}{\lambda_1(Q)} \int_Q |b(x) - b_Q|^{p} \lambda_2(x) d\mu(x) \right)^{\frac{1}{p}} < \epsilon,$$

and

$$\left( \frac{1}{\lambda_2(Q)} \int_Q |b(x) - b_Q|^{-p} \lambda_1(x) d\mu(x) \right)^{\frac{1}{p}} < \epsilon$$

when $l(Q) > N$, $l(Q) < \delta$ and $Q \cap Q_N = \emptyset$.

We now write $T_{S,b}^*([|f|])(x)$ as follows.

$$T_{S,b}^*([|f|])(x) = \sum_{Q \supset Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x) \quad (4.2)$$

$$+ \sum_{Q \cap Q_N = \emptyset} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x)$$

$$+ \sum_{Q \subset Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x)$$


\[ + \sum_{Q \subset Q_N, l(Q) > \delta} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q||f(y)|d\mu(y) \right) \chi_Q(x) \]

\[ =: T_1 f(x) + T_2 f(x) + T_3 f(x) + T_4 f(x), \]

where all \( Q \) are in sparse family \( S \) and we omit \( Q \in S \) in each of the summation for brevity.

Given some \( \epsilon > 0 \), to obtain the compactness for our sparse operator \( T_{S, \delta}^* (|f|)(x) \), we will show the norm of \( T_1 f(x), T_2 f(x), T_3 f(x) \) is at most \( \epsilon \) and \( T_4 f(x) \) is a compact operator.

In fact, by noting that there are only finitely many cubes contained in \( Q_N \) such that \( \delta < l(Q) < N \), we obtain that \( T_4 f(x) \) has finite range and hence it is compact.

We will now show that the norm of \( T_1 f(x), T_2 f(x), T_3 f(x) \) is at most \( \epsilon \).

Let us start with the estimate for the norm of \( T_3 f(x) \); i.e.,

\[
\|T_3 f(x)\|_{L^p_{\lambda_2}(X)} \leq \epsilon \|f\|_{L^p_{\lambda_2}(X)}. \tag{4.3}
\]

From Lemma 2.13, we have the following

\[
|b(y) - b_Q| \leq C \sum_{R \subseteq S, R \subseteq Q} \Omega(b, R) \chi_R(y), \quad \text{a.e. } y \in Q. \tag{4.4}
\]

Recall that \( \nu = \lambda_1^{1/p} \lambda_2^{-1/p} \), thus for some \( \epsilon > 0 \), by using (4.4) we have the following

\[
T_3 f(x) \leq \sum_{Q \subset Q_N, l(Q) < \delta} \sum_{R \subseteq S, R \subseteq Q} \left( \frac{1}{\mu(R)} \int_R |b(z) - b_R|d\mu(z) \frac{1}{\mu(Q)} \int_R |f(y)|d\mu(y) \right) \chi_Q(x) \tag{4.5}
\]

\[
\leq \epsilon \sum_{Q \subset Q_N, l(Q) < \delta} \left( \sum_{R \subseteq S, R \subseteq Q} |f| \nu(R) \right) \frac{1}{\mu(Q)} \chi_Q(x)
\]

\[
\leq \epsilon \sum_{Q \subset Q_N, l(Q) < \delta} \frac{1}{\mu(Q)} \left( \int_Q A_\delta(|f|(y)\nu(y))dy \right) \chi_Q(x)
\]

\[
\leq \epsilon A_\delta \left( A_\delta(|f|\nu) \right)(x).
\]

Here we have used from Definition 2.7 that for \( b \in \text{VMO}_\nu(X) \) such that when \( l(Q) < \delta \), we have \( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q|d\mu(y) < \epsilon \). Also we have used Definition 2.12 to obtain the last equation above.

Now observe that from estimate (2.18) for the boundedness of sparse operator, we have that for some constant \( C \)

\[
\|A_\delta f\|_{L^p_{\lambda_2}(X)} \leq C[\lambda_2]_{A_p}^{\max\{1, \frac{p}{p-1}\}} \|f\|_{L^p_{\lambda_2}(X)}, \tag{4.6}
\]

And thus

\[
\|T_3 f\|_{L^p_{\lambda_2}(X)} \leq \epsilon \|A_\delta (A_\delta(|f|\nu))\|_{L^p_{\lambda_2}(X)} \tag{4.7}
\]
\[ \leq \epsilon [\lambda_2]_{A_p}^{\max \{1, \frac{1}{p-1} \}} \| A_{\mathcal{S}}(|f|) \nu \|_{L^p_\nu(X)} = \epsilon [\lambda_2]_{A_p}^{\max \{1, \frac{1}{p-1} \}} \| A_{\mathcal{S}}(|f|) \|_{L^p_\mathcal{S}(X)} \]

\[ \leq \epsilon ([\lambda_1]_{A_p} [\lambda_2]_{A_p})^{\max \{1, \frac{1}{p-1} \}} \| f \|_{L^p_{\mathcal{S}}(X)}. \]

Then this finishes the proof for the control of the norm of \( T_3 f(x) \).

For \( T_2 f \), recall that

\[ T_2 f(x) =: \sum_{Q, Q_N = \emptyset} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x). \]  

(4.8)

Following similar approach in the estimate for \( T_3 f \), we write \(|b(y) - b_Q|\) as in (4.4). Since \( Q \cap Q_N = \emptyset \) and \( R \subset Q \), we have for all \( R \in \mathcal{S} \) in (4.4), \( R \cap Q_N = \emptyset \). According to Definition 2.7, we have that \( \frac{1}{\mu(R)} \int_Q |b(x) - b_R| d\mu(x) \leq \epsilon \). Then following the same arguments in the estimate of the norm \( T_3 f(x) \), we obtain similar control for the norm of \( T_2 f(x) \), i.e.,

\[ \| T_2(f) \|_{L^p_{\mathcal{S}}(X)} \leq \epsilon \| f \|_{L^p_{\mathcal{S}}(X)}. \]  

(4.9)

Now let us show the control for the norm of \( T_1 f(x) \), recall that

\[ T_1 f(x) = \sum_{Q \supset Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| |f(y)| d\mu(y) \right) \chi_Q(x). \]  

(4.10)

We will start with a collection of sparse dyadic cubes \( Q = Q_1 \supset Q_2 \supset Q_3 \supset Q_4 \cdots \supset Q_{\tau Q} \supset Q_{\tau Q+1} = Q_N \), where \( Q_i \) is the “parent” of \( Q_{i+1}, i = 1, 2, \ldots, \tau Q \). For the sake of the sparse property, if the parent of \( Q_{i+1} \) has only one child \( Q_{i+1} \), we should still denote by \( Q_{i+1} \) the parent of \( Q_{i+1} \) since they are the same dyadic cube indeed. Then repeat the process until we find \( Q_i \) such that \( Q_i \) has at least two children and \( Q_{i+1} \) is one of them. For each \( Q_i \), \( i = 1, 2, \ldots, \tau Q \), we denote all its dyadic children except \( Q_{i+1} \) by \( Q_{i,k}, k = 1, 2, \ldots, M_{Q_i} \), where \( M_{Q_i} + 1 \) is the number of the children of \( Q_i \) and less than uniform constant \( M \) in (2.7). Hence for all \( i = 0, 1, 2, \ldots, \tau Q \) and \( k = 1, 2, \ldots, M_{Q_i}, Q_{i,k} \cap Q_N = \emptyset \). Note that \( Q_{i+1} \) and \( Q_{i,k} \) have equivalent measures since it follows from (2.8) that

\[ \mu(Q_{i+1}) \leq \mu(B(Q_i)) \leq C \left( 1 + \frac{d(x_{Q_i}, x_{Q_{i,k}})}{C_1 \delta^{k_i}} \right)^n \left( \frac{C_1}{c_1 \delta} \right)^n \mu(B(x_{Q_{i,k}}, c_1 \delta^{k_i+1}) \leq C 2^n \left( \frac{C_1}{c_1 \delta} \right)^n \mu(Q_{i,k}). \]

And thus there exists uniform constant \( 0 < \tilde{\eta} < 1 \) such that \( \mu(Q_{i+1}) \leq \tilde{\eta} \mu(Q_i) \), which ensure the sparse property of the collection of \( \{Q_i\} \). Then

\[ T_1 f(x) = \sum_{Q \supset Q_N} \sum_{i=1}^{\tau Q} \sum_{k=1}^{M_{Q_i}} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |b(y) - b_{Q_{i,k}}| |f(y)| d\mu(y) \chi_Q(x) \]  

(4.11)

\[ + \sum_{Q \supset Q_N} \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_Q| |f(y)| d\mu(y) \chi_Q(x) \]

\[ + \sum_{Q \supset Q_N} \sum_{i=1}^{\tau Q} \sum_{k=1}^{M_{Q_i}} |b_{Q_{i,k}} - b_Q| \frac{1}{\mu(Q)} \int_{Q_{i,k}} |f(y)| d\mu(y) \chi_Q(x) \]

\[ + \sum_{Q \supset Q_N} |b_Q - b_Q| \frac{1}{\mu(Q)} \int_{Q_N} |f(y)| d\mu(y) \chi_Q(x) \]
If it suffices to prove that
\[
\|I\|_{L^p_{\lambda_x}(X)} + \|II\|_{L^p_{\lambda_x}(X)} + \|III\|_{L^p_{\lambda_x}(X)} + \|IV\|_{L^p_{\lambda_x}(X)} \lesssim \epsilon \|f\|_{L^p_{\lambda_x}(X)}. \tag{4.12}
\]

Let us now begin with the estimate of \(I\). Recall that \(\lambda_1 = \frac{\lambda_2}{\lambda_1}^{\frac{1}{p}}\) and \(\lambda_2 = \lambda_2^{\frac{1}{p}}\).

For an appropriate choice of \(g \in L^p_{\lambda_x}(X)\) of norm one, we have
\[
\|II\|_{L^p_{\lambda_x}(X)} = \left\| \sum_{Q \supseteq Q_N} \left( \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}| |f(y)|d\mu(y) \right) \chi_Q(x) \right\|_{L^p_{\lambda_x}(X)}
= \sup_{\|g\|_{L^p_{\lambda_x}} \leq 1} \left| \sum_{Q \supseteq Q_N} \left( \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}| |f(y)|d\mu(y) \right) \chi_Q(x), g(x) \right|
\leq \sum_{Q \supseteq Q_N} \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}| |f(y)|d\mu(y) \int_{Q} |g(x)|d\mu(x)
\leq \sum_{Q \supseteq Q_N} \frac{1}{\mu(Q)} \left( \int_{Q_N} |b(y) - b_{Q_N}| \lambda_1'(y) \lambda_1(y) \right)^{\frac{1}{p}} \left( \int_{Q_N} |f(y)|^{p} \lambda_1(x) \right) \lambda_2(Q)^{\frac{1}{p}}
\times \left( \int_{Q} |g(x)|^{p} \lambda_2'(x) \right) \lambda_2(Q)^{\frac{1}{p}}
\leq \sum_{Q \supseteq Q_N} \frac{1}{\lambda_2(Q)} \left( \frac{1}{\lambda_2(Q_N)} \right) \int_{Q_N} |b(y) - b_{Q_N}| \lambda_1'(y) \lambda_1(y) \lambda_2'(x) \lambda_1(x) \lambda_2(x)^{\frac{1}{p}}
\|f\|_{L^p_{\lambda_x}(X)} \lambda_2(Q)^{\frac{1}{p}} \lambda_2(Q)^{\frac{1}{p}}.
\]

Observe that since \(Q \supset Q_N\) and thus \(l(Q) > N\), which gives
\[
\left( \frac{1}{\lambda_2(Q)} \int_{Q} |b(x) - b_{Q}| \lambda_1'(x) \lambda_1(x) \lambda_2'(x) \lambda_1(x) \lambda_2(x)^{\frac{1}{p}} \right)^{\frac{1}{p}} < \epsilon.
\]

Also recall that \(\lambda_2\) is doubling and as \(\lambda_2 \in A_p\), there exists some \(\sigma > 0\) such that \(\lambda_2 \in A_{p-\sigma}\) and
\[
\frac{\lambda_2(Q)}{\lambda_2(Q_N)} \leq \left( \frac{\mu(Q)}{\mu(Q_N)} \right)^{p-\sigma} [\lambda_2]_{A_p}. \tag{4.13}
\]

And since all \(Q\) are in sparse family \(S\), it follows from Corollary 2.11 that
\[
\sum_{Q \supseteq Q_N, Q \in S} \left( \frac{\mu(Q_N)}{\mu(Q)} \right) \leq C. \tag{4.14}
\]

So it follows form (4.13) and (4.14) that
\[
\|II\|_{L^p_{\lambda_x}(X)} \leq \epsilon \|f\|_{L^p_{\lambda_x}(X)} \sum_{Q \supseteq Q_N} \frac{\lambda_2(Q)}{\mu(Q_N)} \lambda_2(Q) \lambda_2(Q)^{\frac{1}{p}} \mu(Q) \lambda_2(Q)^{\frac{1}{p}}
\leq \epsilon \|f\|_{L^p_{\lambda_x}(X)} \sum_{Q \supseteq Q_N} [\lambda_2]_{A_p} \mu(Q) \lambda_2(Q)^{\frac{1}{p}} \mu(Q_N) \lambda_2(Q)^{\frac{1}{p}}
\leq \epsilon \|f\|_{L^p_{\lambda_x}(X)} \sum_{Q \supseteq Q_N} [\lambda_2]_{A_p}^{\frac{1}{p}} \mu(Q) \lambda_2(Q)^{\frac{1}{p}} \mu(Q_N) \lambda_2(Q)^{\frac{1}{p}}.
\[ \leq \epsilon \|f\|_{L^p_{\lambda_1}(X)}|\lambda_2|^\frac{2}{p} \sum_{Q \supset Q_N} \left( \frac{\mu(Q)}{\mu(Q)} \right)^{\frac{p}{p}} \]

\[ \leq \epsilon \|f\|_{L^p_{\lambda_1}(X)}|\lambda_2|^\frac{2}{p}. \]

This gives the estimate for the norm of II.

We now estimate the norm of I. We would like to change the order of the summation for \( Q \) and \( k \). Thus we may assume that \( Q_{i,k} = \emptyset \) when \( M \geq k > M_Q \), and the corresponding terms are 0. So we have the following equality

\[ I = \sum_{Q \supset Q_N} \left( \sum_{i=1}^{\tau_Q} \sum_{k=1}^{M_Q} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |b(y) - b_{Q_{i,k}}| |f(y)| d\mu(y) \right) \chi_Q(x) \]

(4.15)

\[ = \sum_{k=1}^{M} \sum_{Q \supset Q_N} \left( \sum_{i=1}^{\tau_Q} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |b(y) - b_{Q_{i,k}}| |f(y)| d\mu(y) \right) \chi_Q(x). \]

Fixing \( k \), then for each \( Q_{i,k} \) where \( i = 1, \ldots, \tau_Q \), following similar approach in the estimate for \( T_{3,f} \), we write \( |b - b_{Q_{i,k}}| \) as in (4.4). Since \( Q_{i,k} \cap Q_N = \emptyset \) and \( R \subset Q_{i,k} \), we have for all \( R \in \mathcal{S} \) in (4.4), \( R \cap Q_N = \emptyset \). Now according to Definition 2.7, we have that \( \frac{1}{\mu(R)} \int_R |b(x) - b_R| d\mu(x) \leq \epsilon. \)

Following similar estimates as showed in equations (4.5) and (4.7), we obtain that

\[ \|I\|_{L^p_{x}(X)} \leq \epsilon([\lambda_1]_{A_p}[\lambda_2]_{A_p})^{\max\{1, \frac{1}{p'}\}} \|f\|_{L^p_{\lambda_1}(X)}. \]

(4.16)

We turn to the estimates for the norm of III and IV. Observe for each fixed \( k \), for each \( Q_{i,k} \) we will obtain the same estimate as follows, which is independent of the cube \( Q_{i,k} \):

\[ \|A_{Q_{i,k}}\|_{L^p_{x}(X)} \leq \epsilon([\lambda_1]_{A_p}[\lambda_2]_{A_p})^{\max\{1, \frac{1}{p'}\}} \|f\|_{L^p_{\lambda_1}(X)}, \]

(4.17)

where

\[ A_{Q_{i,k}}(x) := \sum_{Q \supset Q_N} \left( \sum_{i=1}^{\tau_Q} |b_{Q_{i,k}} - b_Q| \frac{1}{\mu(Q)} \int_{Q_{i,k}} |f(y)| d\mu(y) \right) \chi_Q(x). \]

By using (4.17), we obtain the estimates for the norm of III and IV since the same estimate holds for each \( Q_{i,k} \) where \( k \in \{1, 2, \ldots, M_Q\} \).

Thus, it suffices to show (4.17). Recall the definition of \( Q_i \) and \( Q_{i,k} \): \( Q = Q_1 \supset Q_2 \supset Q_3 \supset Q_4 \ldots \ldots Q_{\tau_Q+1} = Q_N \), where \( Q_i \) is the “parent” of \( Q_{i+1} \), \( i = 1, 2, \ldots, \tau_Q \). The collection \( \{Q_i\}_i \) are sparse. For each \( Q_i \), \( i = 1, 2, \ldots, \tau_Q \), we denote all its dyadic children except \( Q_{i+1} \) by \( Q_{i,k} \), \( k = 1, 2, \ldots, M_Q \).

Observe

\[ |b_{Q_{i,k}} - b_Q| \leq |b_{Q_{i,k}} - b_{Q_{i-1}}| + |b_{Q_{i-1}} - b_{Q_{i-2}}| + \ldots + |b_{Q_2} - b_Q| \]

(4.18)

\[ \leq \frac{1}{\mu(Q_{i,k})} \int_{Q_{i,k}} |b(x) - b_{Q_{i-1}}| d\mu(x) + \ldots + \frac{1}{\mu(Q_2)} \int_{Q_2} |b(x) - b_Q| d\mu(x) \]

\[ \leq C \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \left( \frac{1}{\nu(Q_j)} \int_{Q_j} |b(x) - b_{Q_j}| d\mu(x) \right) \]

\[ \leq C \epsilon \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)}. \]
In the last step above we used the fact that $l(Q_j) > N$ for all $j \in \{0, 1, \ldots, i - 1\}$ because for all these $j$ we have $Q_j \supset Q_N$.

Using Equation (4.18) we get the following

$$A_{Q_{i,k}} \leq C \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |f(y)|d\mu(y)x_Q(x).$$  (4.19)

Hence we have

$$\|A_{Q_{i,k}}\|_{L^p_{\lambda_2}(X)}$$  (4.20)

$$\leq \left| \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |f(y)|d\mu(y)x_Q(x) \right|_{L^p_{\lambda_2}(X)}$$

$$\leq \left( \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \left( \frac{\nu(Q_j)}{\mu(Q_j)} \right)^p \left( \int_{Q_{i,k}} |f(y)|d\mu(y) \right)^p \frac{1}{\mu(Q)} \left( \frac{\mu(Q)}{\mu(Q_j)} \right)^{\sigma'} \lambda_2(Q) \right)^{\frac{1}{p}}$$

$$\times \left( \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \left( \frac{\mu(Q_{i,k})}{\mu(Q)} \right)^{p'} \left( \int_{Q} |g(y)|d\mu(y) \right)^{p'} \left( \lambda_2(Q) \right)^{-\sigma'} \lambda_2(Q) \right)^{\frac{1}{p'}}$$

$$\leq \epsilon A^p B^{\frac{1}{p}},$$

where $\sigma' = \frac{\sigma}{p}$. Now observe that

$$B \leq \sum_{Q \supset Q_N} \left[ \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \left( \frac{\mu(Q_i)}{\mu(Q)} \right)^{\sigma'} \left( \frac{1}{\lambda_2(Q)} \int_{Q} |g(y)|d\mu(y) \right)^{p'} \lambda_2(Q) \right]$$  (4.21)

$$\leq C_1 C_2 \sum_{Q \supset Q_N} \inf_{x \in Q} M_{\lambda_2}^{p'}(\lambda_2^{-1}(x)\lambda_2(E(Q)))$$

$$\leq C_1 C_2 \sum_{Q \supset Q_N} \int_{E(Q)} M_{\lambda_2}^{p'}(\lambda_2^{-1}(x)\lambda_2(x)d\mu(x)$$

$$\leq C_1 C_2 \int_{\mathbb{R}^n} M_{\lambda_2}^{p'}(\lambda_2^{-1}(x)\lambda_2(x)d\mu(x)$$

$$\leq C_1 C_2 \|g\|_{L_{\lambda_2}^{p'}(X)}^{p'}$$

$$= C_1 C_2 \|g\|_{L_{\lambda_2}^{p'}(X)}^{p'},$$

where we use the facts that $\{Q_i\}_i$ are sparse family and thus there is a constant $C_1$ such that

$$\sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \left( \frac{\mu(Q_i)}{\mu(Q)} \right)^{\sigma'} \leq C_1;$$

and that there is a constant $C_2$ such that (by Lemma 2.4)

$$\frac{\lambda_2(Q)}{\lambda_2(E(Q))} \leq C_2,$$
where $E(Q)$ is the set in the cube $Q$ in some $\eta$-sparse collection $\mathcal{S}$ such that $\mu(E_Q) \geq \eta \mu(Q)$.

We also have the following estimate for $A$

$$A \leq \sum_{Q \supset Q_N} \tau_0 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \frac{\lambda_1(Q_j)X_2(Q_j)}{\mu(Q_j)} \left\| f \right\|_{L^p_{\lambda_1}(Q_{i,k})} \frac{\lambda_1(Q_j)}{\mu(Q_j)} \left( \frac{\mu(Q_i)}{\mu(Q_{i,k})} \right)^p \left( \frac{\mu(Q_j)}{\mu(Q_{i,k})} \right)^{p-1} \frac{1}{\mu(Q)^p} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p-1} \lambda_2(Q) \quad (4.22)$$

$$\leq \sum_{Q \supset Q_N} \tau_0 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \frac{\lambda_1(Q_j)X_2(Q_j)}{\mu(Q_j)} \left\| f \right\|_{L^p_{\lambda_1}(Q_{i,k})} \frac{\lambda_1(Q_j)}{\mu(Q_j)} \left( \frac{\mu(Q_i)}{\mu(Q_{i,k})} \right)^p \left( \frac{\mu(Q_j)}{\mu(Q_{i,k})} \right)^{p-1} \frac{1}{\mu(Q)^p} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p-1} \lambda_2(Q)$$

$$\times \frac{\lambda_2(Q)}{\lambda_2(Q_j)} \left( \frac{\mu(Q_i)}{\mu(Q_{i,k})} \right)^p \left( \frac{\mu(Q_j)}{\mu(Q_{i,k})} \right)^{p-1} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p-1} \| f \|_{L^p_{\lambda_1}(Q_{i,k})}$$

$$\leq \sum_{Q \supset Q_N} \tau_0 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \frac{\lambda_1(Q_j)X_2(Q_j)}{\mu(Q_j)} \left\| f \right\|_{L^p_{\lambda_1}(Q_{i,k})} \frac{\lambda_1(Q_j)}{\mu(Q_j)} \left( \frac{\mu(Q_i)}{\mu(Q_{i,k})} \right)^p \left( \frac{\mu(Q_j)}{\mu(Q_{i,k})} \right)^{p-1} \frac{1}{\mu(Q)^p} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p-1} \lambda_2(Q)$$

$$\leq \lambda_1^2 \lambda_2^2 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \log \left( \frac{\mu(Q_i)}{\mu(Q_j)} \right) \left( \frac{\mu(Q_j)}{\mu(Q_{i,k})} \right)^{p-1} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p-1} \| f \|_{L^p_{\lambda_1}(Q_{i,k})}$$

$$\leq \lambda_1^2 \lambda_2^2 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \left\| f \right\|_{L^p_{\lambda_1}(Q_{i,k})}$$

$$\leq C \lambda_1^2 \lambda_2^2 \sum_{i=1}^{\tau_0} \sum_{j=1}^{\tau_0} \left\| f \right\|_{L^p_{\lambda_1}(X)}$$

Then (4.17) follows from (4.21), (4.22) and (4.20). This completes the proof of Theorem 1.1.

## 5 Proof of Theorem 1.2

In this section we give the proof for the Theorem 1.2. We will first show the sparse domination for the maximal commutator $C_b$.

### 5.1 Sparse domination of the maximal commutator $C_b$, and proof of part (i) in Theorem 1.2

To begin with, we show that the maximal commutator $C_b$ is bounded from above by sparse operators. This in turns gives the proof for part (i) of Theorem 1.2.

To be more precise, we will show that part (i) of the Theorem 1.2 follows from the proof of Theorem 1.1 and from the result of Theorem 5.2 below, as which shows that

$$|C_b(f)(x)| \leq C \sum_{i=1}^{T} \left( T_{S_{i,b}}(\left| f \right|)(x) + T_{S_{i,b}}^*(\left| f \right|)(x) \right),$$

where $T_{S_{i,b}}$ and $T_{S_{i,b}}^*$ are sparse operators and hence compact from $L^p_{\lambda_1}(X)$ to $L^p_{\lambda_2}(X)$ if $b \in VMO_{\nu}(X)$.

We now begin with the sparse domination.

Given a ball $B_0 \subset X$, for $x \in B_0$ we define a local grand maximal truncated operator $\mathcal{M}_{B_0}$ as follows:

$$\mathcal{M}_{B_0} f(x) := \sup_{B \ni x, B \subset B_0} \text{ess sup}_{\xi \in B} \mathcal{M}(f \chi_{A_0 B_0 \setminus A_0 B})(\xi).$$
Using the idea of [21, Lemma 3.2], we can obtain the following lemma.

**Lemma 5.1.** For a.e. \( x \in B_0 \),
\[
\mathcal{M}(f\chi_{A_0B_0})(x) \leq C\|\mathcal{M}\|_{L^1 \to L^{\infty}}|f(x)| + \mathcal{M}_{B_0}f(x).
\]

**Proof.** Suppose that \( x \in B_0 \), and let \( x \) be a point of approximate continuity of \( \mathcal{M}(f\chi_{A_0B_0}) \) (see, e.g., [10], p. 46). Then for every \( \varepsilon > 0 \), the sets
\[
E_s(x) = \{y \in B(x,s) : |\mathcal{M}(f\chi_{A_0B_0})(y) - \mathcal{M}(f\chi_{A_0B_0})(x)| < \varepsilon \}
\]
satisfy
\[
\lim_{s \to 0} \frac{\mu(E_s(x))}{\mu(B(x,s))} = 1.
\]
Then for a.e. \( y \in E_s(x) \),
\[
\mathcal{M}(f\chi_{A_0B_0})(x) \leq \mathcal{M}(f\chi_{A_0B_0})(y) + \varepsilon \leq \mathcal{M}(f\chi_{A_0B(x,s)})(y) + \mathcal{M}_{B_0}f(x) + \varepsilon.
\]
Therefore, applying the weak type \((1,1)\) of \( \mathcal{M} \) yields
\[
\mathcal{M}(f\chi_{A_0B_0})(x) \leq \text{ess inf}_{y \in E_s(x)} \mathcal{M}(f\chi_{A_0B(x,s)})(y) + \mathcal{M}_{B_0}f(x) + \varepsilon
\]
\[
\leq C\|\mathcal{M}\|_{L^1 \to L^{\infty}} \frac{1}{\mu(E_s(x))} \int_{A_0B(x,s)} |f(x)|d\mu(x) + \mathcal{M}_{B_0}f(x) + \varepsilon.
\]
Assuming additionally that \( x \) is a Lebesgue point of \( f \) and letting subsequently \( s \to 0 \) and \( \varepsilon \to 0 \), we completes the proof of this lemma. \( \square \)

Then we have the following result.

**Theorem 5.2.** For every compactly supported \( f \in L^\infty(X) \), there exists \( T \) dyadic systems \( \mathcal{D}^t, t = 1,2,\ldots,T \) and \( \eta \)-sparse families \( \mathcal{S}_t \subset \mathcal{D}^t \) such that for a.e. \( x \in X \),
\[
|C_b(f)(x)| \leq C \sum_{t=1}^T \left( T_{\mathcal{S}_{t,b}}(|f|)(x) + T_{\mathcal{S}_{t,b}}(|f|)(x) \right).
\]

**Proof.** We recall from Section 2.2, for each ball \( B(x,r) \subseteq X \) with \( \delta^{k+3} < r \leq \delta^{k+2}, k \in \mathbb{Z} \), there exist \( t \in \{1,2,\ldots,T\} \) and \( Q \in \mathcal{D}^t \) of generation \( k \) and with center point \( t^x_n \) such that \( \rho(t^x_n,t^x_{n+1}) < 2A_0\delta_k \) and \( B(x,r) \subseteq Q \subset B(x,Cr) \). Here and in what follows, \( A_0 \) denotes the constant in \((2.1)\).

Fix a ball \( B_0 \subset X \), then it is clear that there exist a positive constant \( C_0 \), \( t_0 \in \{1,2,\ldots,T\} \) and \( Q_0 \in \mathcal{D}^{t_0} \) such that \( 3A_0B_0 \subseteq Q_0 \subseteq C(3A_0B_0) \). We now show that there exists a \( \frac{1}{8C_{A_0,B_0}} \)-sparse family \( \mathcal{F}^{t_0} \subset \mathcal{D}^{t_0}(B_0) \) such that for a.e. \( x \in B_0 \),
\[
|C_b(f\chi_{3A_0B_0})(x)| \leq C \sum_{Q \in \mathcal{F}^{t_0}} \left( |b(x) - b_{3Q_0}| ||f||_{3A_0Q} + |(b - b_{3Q_0})f||_{3A_0Q} \right)\chi_Q(x).
\]

Here, \( R_Q \) is the dyadic cube in \( \mathcal{D}^t \) for some \( t \in \{1,2,\ldots,T\} \) such that \( 3A_0Q \subset R_Q \subset C(3A_0Q) \).

It suffices to prove the following recursive claim: there exist pairwise disjoint cubes \( P_j \in \mathcal{D}^{t_0}(B_0) \) such that \( \sum_j \mu(P_j) \leq \frac{1}{4} \mu(B_0) \) and
\[
|C_b(f\chi_{3A_0B_0})(x)|\chi_{B_0} \leq C \left( |b(x) - b_{3Q_0}| ||f||_{3A_0B_0} + |(b - b_{3Q_0})f||_{3A_0B_0} \right)
\]

(5.4)
\[ a.e. \text{ on } B_0. \text{ Here we have a } \frac{1}{16} \text{-sparse family since the sets } E_Q = Q \setminus \bigcup_j P_j, \text{ and then we can appeal to the discussion after Definition 2.9.} \]

Now observe that for arbitrary pairwise disjoint cubes \( P_j \in \mathcal{D}^{io}(B_0) \),

\[
\begin{align*}
|C_b(f\chi_{3A_0B_0})(x)|_{\mathcal{B}_0} & = |C_b(f\chi_{3A_0B_0})(x)|_{\mathcal{B}_0 \setminus \bigcup_j P_j} + \sum_j |C_b(f\chi_{3A_0B_0})(x)|_{P_j} \\
& \leq |C_b(f\chi_{3A_0B_0})(x)|_{\mathcal{B}_0 \setminus \bigcup_j P_j} + \sum_j |C_b(f\chi_{3A_0B_0\setminus 3A_0P_j})(x)|_{P_j} + \sum_j |C_b(f\chi_{3A_0P_j})(x)|_{P_j}.
\end{align*}
\]

Hence, in order to prove the recursive claim (5.4), it suffices to show that one can select pairwise disjoint cubes \( P_j \in \mathcal{D}^{io}(B_0) \) with \( \sum_j \mu(P_j) \leq \frac{1}{16} \mu(B_0) \) and such that for a.e. \( x \in B_0 \),

\[
|C_b(f\chi_{3A_0B_0})(x)|_{\mathcal{B}_0 \setminus \bigcup_j P_j} + \sum_j |C_b(f\chi_{3A_0B_0\setminus 3A_0P_j})(x)|_{P_j} \leq C \left( |b(x) - b_{Q_0}| f|_{3A_0B_0} + |(b - b_{Q_0}) f|_{3A_0B_0} \right). \tag{5.5}
\]

To see this, by definition, we obtain that

\[
|C_b(f\chi_{3A_0B_0})(x)|_{\mathcal{B}_0 \setminus \bigcup_j P_j} + \sum_j |C_b(f\chi_{3A_0B_0\setminus 3A_0P_j})(x)|_{P_j} \leq \sup_{B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |b(x) - b(y)||f(y)|\chi_{3A_0B_0}(y)d\mu(y)\chi_{\mathcal{B}_0 \setminus \bigcup_j P_j}(x) + \sum_j \sup_{B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |b(x) - b(y)||f(y)|\chi_{3A_0B_0\setminus 3A_0P_j}(y)d\mu(y)\chi_{P_j}(x) \leq |b(x) - b_{Q_0}| \sup_{B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |f(y)|\chi_{3A_0B_0}(y)d\mu(y)\chi_{\mathcal{B}_0 \setminus \bigcup_j P_j}(x) + \sum_j \sup_{B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |b(y) - b_{Q_0}| |f(y)|\chi_{3A_0B_0}(y)d\mu(y)\chi_{\mathcal{B}_0 \setminus \bigcup_j P_j}(x) + \sum_j \sup_{B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |b(y) - b_{Q_0}| |f(y)|\chi_{3A_0B_0\setminus 3A_0P_j}(y)d\mu(y)\chi_{P_j}(x) \leq |b(x) - b_{Q_0}| f|_{3A_0B_0} + |(b - b_{Q_0}) f|_{3A_0B_0}. \tag{5.6}
\]

We now choose \( \alpha \) such that the set \( E := E_1 \cup E_2 \), with

\[
E_1 = \{ x \in B_0 : |f(x)| > \alpha |f|_{3A_0B_0} \} \cup \{ x \in B_0 : \mathcal{M}_{B_0} f(x) > \alpha C |f|_{3A_0B_0} \},
\]

and

\[
E_2 = \{ x \in B_0 : |(b(x) - b_{Q_0}) f(x)| > \alpha |(b - b_{Q_0}) f|_{3A_0B_0} \} \cup \{ x \in B_0 : \mathcal{M}_{B_0} ((b - b_{Q_0}) f)(x) > \alpha C |(b - b_{Q_0}) f|_{3A_0B_0} \},
\]
will satisfy

$$\mu(E) \leq \frac{1}{2^{n+1}} \mu(B_0).$$

We now apply the Calderón–Zygmund decomposition to the function $\chi_E$ on $B_0$ at the height $\lambda = \frac{1}{2^{n+1}}$, where $n$ is the upper dimension of the measure $\mu$ as in (2.3), to obtain the pairwise disjoint cubes $P_j \in \mathcal{D}^{\mu_0}(B_0)$ such that

$$\chi_E(x) \leq \frac{1}{2^{n+1}} \quad \text{a.e. } x \notin \cup_j P_j$$

and hence we have that $\mu(E \setminus \cup_j P_j) = 0$. Moreover, we have that

$$\sum_j \mu(P_j) = \mu\left(\bigcup_j P_j\right) \leq 2^{n+1} \mu(E) \leq \frac{1}{16} \mu(B_0),$$

and that

$$\frac{1}{2^{n+1}} \leq \frac{1}{\mu(P_j)} \int_{P_j} \chi_E(x) d\mu(x) = \frac{\mu(P_j \cap E)}{\mu(P_j)} \leq \frac{1}{2},$$

which implies that

$$P_j \cap E^c \neq \emptyset.$$

Therefore, we observe that for each $P_j$, since $P_j \cap E^c \neq \emptyset$, we have that

$$\mathcal{M}_{B_0}\left((b - b_{Q_0})f\right)(x) \leq \alpha C |(b - b_{Q_0})f|_{3A_{B_0}}$$

for some $x \in P_j$, which implies that

$$\text{ess sup}_{\xi \in P_j} \text{ess sup}_{B \ni \xi} \frac{1}{\mu(B)} \int_B |b(y) - b_{Q_0}| |f(y)| \chi_{3A_{B_0} \setminus 3A_0} P_j(y) d\mu(y) \leq \alpha C |(b - b_{Q_0})f|_{3A_{B_0}}.$$

Similarly, we have

$$\text{ess sup}_{\xi \in P_j} \text{ess sup}_{B \ni \xi} \frac{1}{\mu(B)} \int_B |f(y)| \chi_{3A_{B_0} \setminus 3A_0} P_j(y) d\mu(y) \leq \alpha C |f|_{3A_{B_0}}.$$

Also, by Lemma 5.1, we have that

$$\sup_{B \ni x} \frac{1}{\mu(B)} \int_B |f(y)| \chi_{3A_{B_0}}(y) d\mu(y) \leq C |f(x)| + \mathcal{M}_{B_0} f(x).$$

and

$$\sup_{B \ni x} \frac{1}{\mu(B)} \int_B |b(y) - b_{Q_0}| |f(y)| \chi_{3A_{B_0}}(y) d\mu(y) \leq C |b(x) - b_{Q_0}| |f(x)| + \mathcal{M}_{B_0}\left((b - b_{Q_0})f\right)(x).$$

Since $\mu(E \setminus \cup_j P_j) = 0$, we have that from the definition of the set $E$, the following estimates

$$|f(x)| \leq \alpha |f|_{3A_{B_0}}, \quad |(b(x) - b_{Q_0})f(x)| \leq \alpha |(b - b_{Q_0})f|_{3A_{B_0}}$$

hold for $\mu$-almost every $x \in B_0 \setminus \cup_j P_j$, and also

$$\mathcal{M}_{B_0} f(x) \leq \alpha C |f|_{3A_{B_0}}, \quad \mathcal{M}_{B_0}\left((b - b_{Q_0})f\right)(x) \leq \alpha C |(b - b_{Q_0})f|_{3A_{B_0}}.$$
hold for \( \mu \)-almost every \( x \in B_0 \setminus \bigcup_j P_j \).

Combining these fact with (5.6), we see that (5.5) holds, which further implies that (5.3) holds.

We now consider the partition of the space as follows. Suppose \( f \) is supported in a ball \( B_0 \subset X \). We have

\[
X = \bigcup_{j=0}^{\infty} 2^j B_0.
\]

First, we note that the ball \( B_0 \) is covered by \( 3A_0 B_0 \). Consider the annuli \( U_j := 2^j B_0 \setminus 2^{j-1} B_0 \) for \( j \geq 1 \). It is clear that we can choose the balls \( \{ \bar{B}_j, \ell \}_{\ell=1}^{L_j} \) with radius \( 2^{j-1} r_B \) to cover \( U_j \), satisfying that the center of each the ball \( \bar{B}_j, \ell \) is in \( U_j \neq \emptyset \) and that \( \sup_j L_j \leq C_{A_0, \mu} \), where \( C_{A_0, \mu} \) is an absolute constant depending on \( A_0 \) and \( C_{\mu} \) only, here \( C_{\mu} \) is the constant as in (2.2).

Moreover, we also have that for each such \( \bar{B}_j, \ell \), the enlargement \( 3A_0 \bar{B}_j, \ell \) covers \( B_0 \). Also, we note that for each \( \bar{B}_j, \ell \), there exist a positive constant \( C, t_{j, \ell} \in \{1, 2, \ldots, T\} \) and \( \bar{Q}_{j, \ell} \in \mathcal{D}_{t_{j, \ell}} \) such that \( 3A_0 \bar{B}_j, \ell \subseteq \bar{Q}_{j, \ell} \subseteq C(3A_0 \bar{B}_j, \ell) \).

We now apply (5.3) to each \( \bar{B}_j, \ell \), then we obtain a \( \frac{1}{16} \)-sparse family \( \bar{F}_{j, \ell} \subset \mathcal{D}_{t_{j, \ell}} (\bar{B}_j, \ell) \) such that (5.3) holds for a.e. \( x \in \bar{B}_j, \ell \).

Now we set \( \mathcal{F} = \bigcup_{j, \ell} \bar{F}_{j, \ell} \). Note that the balls \( \bar{B}_{j, \ell} \) are overlapping at most \( 4C_{A_0, \mu} \) times. Then we obtain that \( \mathcal{F} \) is a \( \frac{1}{64C_{A_0, \mu}} \)-sparse family and for a.e. \( x \in X \),

\[
|C_b(f)(x)| \leq C \sum_{Q \in \mathcal{F}} \left( |b(x) - b_{R_Q}| |f|_{3A_0Q} + |(b - b_{R_Q})f|_{3A_0Q} \right) \chi_Q(x) \quad \text{(5.7)}
\]

Since \( 3A_0Q \subset R_Q \), and it is clear that \( \mu(R_Q) \leq \overline{C} \mu(3A_0Q) \), we obtain that \( |f|_{3A_0Q} \leq C |f|_{R_Q} \). Next, we further set

\[
S_t = \{ R_Q \in \mathcal{D}^t : Q \in \mathcal{F} \}, \quad t \in \{1, 2, \ldots, T\},
\]

and from the fact that \( \mathcal{F} \) is \( \frac{1}{64C_{A_0, \mu}} \)-sparse, we can obtain that each family \( S_t \) is \( \frac{1}{64C_{A_0, \mu}} \)-sparse. Now we let

\[
\eta = \frac{1}{64C_{A_0, \mu} C}.
\]

Then it follows that

\[
|C_b(f)(x)| \leq C \sum_{t=1}^{T} \sum_{R \in S_t} \left( |b(x) - b_R| |f|_R + |(b - b_R)f|_R \right) \chi_R(x), \quad \text{(5.8)}
\]

finishing the proof. \( \square \)

### 5.2 Proof of part (ii) in Theorem 1.2

Now we are left with proving part (ii) of Theorem 1.2. So let us begin the proof of part (ii).

**Proof.** (Theorem 1.2(ii)) So we will start by assuming \( b \in BMO_v(X) \) such that \( C_b \) is compact from \( L^2_{\lambda_0}(X) \) to \( L^\infty_{\lambda_0}(X) \). We will use the method of proof by contradiction and hence let us suppose that \( b \notin VMO_v(X) \). Here the main idea for the contradiction is that, on any Hilbert space \( H \), with canonical basis \( e_j, j \in \mathbb{N} \), an operator \( T \) with \( Te_j = v \), with non-zero \( v \in H \) is necessarily unbounded. For \( b \in BMO_v(X) \setminus VMO_v(X) \), we show a variant of this condition for the \( C_b \) giving us the contradiction.
Compactness Characterization for Bloom Sparse Operators

As we assume that \( b \notin VMO_{\nu}(X) \), at least one of the three conditions presented in Definition 2.7 does not hold. Since we have a similar argument for all three conditions, let us suppose that the first condition in Definition 2.7 does not hold.

That is, there exists some \( \delta_0 > 0 \) and a sequence of balls \( \{Q_j\}_{j \in I} \subset X \) such that \( r(Q_j) \to 0 \) as \( j \to \infty \) and we have that

\[
\frac{1}{\nu(Q_j)} \int_{Q_j} |b(x) - b_{Q_j}| \, d\mu(x) \geq \delta_0.
\]  

(5.9)

We will also further assume without loss of generality that

\[
4r(Q_{j+1}) \leq r(Q_j).
\]  

(5.10)

Next, for each \( Q_j \) we choose another ball \( \tilde{Q}_j \) such that \( r(\tilde{Q}_j) = r(Q_j) \), \( Q_j \cap \tilde{Q}_j = \emptyset \) and that \( d(Q_j, \tilde{Q}_j) := \inf_{x \in Q_j, y \in \tilde{Q}_j} d(x, y) \leq 5r(Q_j) \).

Let us denote by \( m_b(\tilde{Q}_j) \) be a median value of \( b \) on the ball \( \tilde{Q}_j \). That is \( m_b(\tilde{Q}_j) \) is a real number such that the two sets we give below have a measure at least \( \frac{1}{2} \mu(\tilde{Q}_j) \).

\[
F_{j,1} \subset \{ y \in \tilde{Q}_j : b(y) \leq m_b(\tilde{Q}_j) \}, \quad F_{j,2} \subset \{ y \in \tilde{Q}_j : b(y) \geq m_b(\tilde{Q}_j) \}.
\]  

(5.11)

Also define the sets

\[
E_{j,1} \subset \{ x \in Q_j : b(x) \geq m_b(Q_j) \}, \quad E_{j,2} \subset \{ x \in Q_j : b(x) < m_b(Q_j) \}.
\]  

(5.12)

So we have that \( Q_j = E_{j,1} \cup E_{j,2} \) and we also have that \( E_{j,1} \cap E_{j,2} = \emptyset \). Also we have the following

\[
b(x) - b(y) \geq 0, \quad (x, y) \in E_{j,1} \times F_{j,1}.
\]

and we also have

\[
b(x) - b(y) < 0, \quad (x, y) \in E_{j,2} \times F_{j,2}.
\]

Also for all \( (x, y) \in E_{j,1} \times F_{j,1} \cup E_{j,1} \times F_{j,1} \), we have that

\[
|b(x) - b(y)| = |b(x) - m_b(\tilde{Q}_j)| + |m_b(\tilde{Q}_j) - b(y)| \geq |b(x) - m_b(\tilde{Q}_j)|.
\]

Let us also define the following sets

\[
\tilde{F}_{j,1} := F_{j,1} \setminus \cup_{l=1}^{j+1} \tilde{Q}_l \quad \tilde{F}_{j,2} := F_{j,2} \setminus \cup_{l=1}^{j+1} \tilde{Q}_l \quad \forall j = 1, 2, \ldots.
\]  

(5.13)

Now using the decay condition for lengths of \( \{Q_j\} \) as given by equation (5.10), we have for each \( j \) the following;

\[
\mu(\tilde{F}_{j,1}) \geq \mu(F_{j,1}) - \mu(\cup_{l=1}^{j+1} \tilde{Q}_l) \geq \frac{1}{2} \mu(\tilde{Q}_j) - \sum_{l=j+1}^{\infty} \mu(\tilde{Q}_l) \geq \frac{1}{2} \mu(\tilde{Q}_j) - \frac{1}{3} \mu(\tilde{Q}_j) = \frac{1}{6} \mu(\tilde{Q}_j).
\]  

(5.14)

We can obtain a similar estimate for the set \( \tilde{F}_{j,2} \). Observe now for every \( j \), we have the following

\[
\frac{1}{\nu(Q_j)} \int_{Q_j} |b(x) - b_Q| \, d\mu(x) \leq \frac{2}{\nu(Q_j)} \int_{Q_j} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x)
\]  

(5.15)

\[
= \frac{2}{\nu(Q_j)} \int_{E_{j,1}} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x) + \frac{2}{\nu(Q_j)} \int_{E_{j,2}} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x)
\]
From the equation (5.9) we have that at least one of these inequalities holds
\[
\frac{2}{\nu(Q_j)} \int_{E_{j,1}} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x) \geq \frac{\delta_0}{2}, \quad \frac{2}{\nu(Q_j)} \int_{E_{j,2}} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x) \geq \frac{\delta_0}{2}.
\]

Let us suppose that the first of these inequalities holds, i.e.,
\[
\frac{2}{\nu(Q_j)} \int_{E_{j,1}} |b(x) - m_b(\tilde{Q}_j)| \, d\mu(x) \geq \frac{\delta_0}{2}.
\]

Hence for every \( j \), using (5.14) we have that
\[
\frac{\delta_0}{4} \leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \left| \int_{\tilde{F}_{j,1}} \frac{1}{\mu(Q_j)} (b(x) - m_b(\tilde{Q}_j)) \, d\mu(y) \right| \, d\mu(x)
\]
\[
\leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \left| \int_{\tilde{F}_{j,1}} \frac{1}{\mu(Q_j)} (b(x) - m_b(\tilde{Q}_j)) \, d\mu(y) \right| \, d\mu(x),
\]
\[
\leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \sum_{i=1}^\infty |C_b(x|^\chi_{\tilde{F}_{j,1}}(x))| \, d\mu(x).
\]

And, then,
\[
\delta_0 \leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \left| \int_{\tilde{F}_{j,1}} \frac{1}{\mu(Q_j)} (b(x) - m_b(\tilde{Q}_j)) \, d\mu(y) \right| \, d\mu(x)
\]
\[
\leq \frac{1}{\lambda_1(Q_j)^p \lambda_2'(Q_j)^p} \int_{E_{j,1}} |C_b(x|^\chi_{\tilde{F}_{j,1}}(x))| \, d\mu(x)
\]
\[
= \frac{1}{\lambda_2'(Q_j)^p} \int_{E_{j,1}} \left| C_b \left( \frac{\chi_{\tilde{F}_{j,1}}(x)}{\lambda_1(Q_j)^{\frac{1}{p}}} \right) \right| \, d\mu(x).
\]

Consider \( f_j =: \frac{\chi_{\tilde{F}_{j,1}}}{\lambda_1(Q_j)^{\frac{1}{p}}} \), observe that this is a sequence of disjointly supported functions using the equation (5.14) with \( \|f_j\|_{L^p_{\lambda_1}(X)} \simeq 1 \). Now using the Hölder’s inequality we get
\[
\delta_0 \leq \frac{1}{\lambda_2'(Q_j)^p} \int_{E_{j,1}} |C_b(f_j(x))| \lambda_2^{-\frac{1}{p}} \lambda_2^{-\frac{1}{p}} \, d\mu(x)
\]
\[
\leq \frac{1}{\lambda_2'(Q_j)^p} \lambda_2^{-\frac{1}{p}} \left( \int_X |C_b(f_j(x))|^p \lambda_2(x) \, d\mu(x) \right)^{\frac{1}{p}}
\]
\[
\leq \left( \int_X |C_b(f_j(x))|^p \lambda_2(x) \, d\mu(x) \right)^{\frac{1}{p}}.
\]

Let us consider \( \psi \) in the closure of \( \{C_b(f_j)\}_j \), then we have \( \|\psi\|_{L^p_{\lambda_2}(X)} \simeq 1 \). Now choose some \( j_i \) such that
\[
\|\psi - C_b(f_{j_i})\|_{L^p_{\lambda_2}(X)} \leq 2^{-i}.
\]

To complete the proof consider a non-negative numerical sequence \( \{c_i\} \) with \( \|\{c_i\}\|_{\ell^p} < \infty \) but \( \|\{c_i\}\|_{\ell^p} = \infty \). Then consider \( \phi = \sum c_i f_{j_i} \in L^p_{\lambda_1}(X) \) and
\[
\| \sum c_i \psi - C_b \phi \|_{L^p_{\lambda_1}(X)} \leq \| \sum c_i (\psi - C_b(f_{j_i})) \|_{L^p_{\lambda_1}(X)}.
\]
\[
\leq \|c_i\|_{p^\prime} \left[ \sum_i \|\psi - C_b(f_{j_i})\|_{L^p_{\lambda_i}(X)}^p \right]^{\frac{1}{p}} \leq 1.
\]

Hence we conclude that \( \sum_i c_i \psi \in L^p_{\lambda_i}(X) \), but \( \sum_i c_i \psi \) is an infinite on set of positive measure which is the contradiction that completes our proof. \( \square \)

### 6 Proof of Theorems 1.3 and 1.4

In this section we prove the proof of Theorems 1.3 and 1.4. In fact, Theorem 1.3 follows from the proof of Theorem 1.1 and from the result of sparse domination of \([b,T]\) on spaces of homogeneous type \([8]\). We now show Theorem 1.4.

**Proof of Theorem 1.4.** So we will start by assuming \( b \in BMO_\nu(X) \) such that \([B,T]\) is compact from \( L^p_{\lambda}(X) \) to \( L^p_{\lambda}(X) \). We will use the method of proof by contradiction and hence let us suppose that \( b \notin VMO_\nu(X) \). Here the main idea for the contradiction is that, on any Hilbert space \( H \), with canonical basis \( e_j, j \in \mathbb{N} \), an operator \( T \) with \( Te_j = v \), with non-zero \( v \in H \) is necessarily unbounded. For \( b \in BMO_\nu(X) \setminus VMO_\nu(X) \), we show a variant of this condition for the \([B,T]\) giving us the contradiction.

As we assume that \( b \notin VMO_\nu(X) \), at least one of the three conditions presented in Definition 2.7 does not hold. Since we have a similar argument for all three conditions, let us suppose that the first condition in Definition 2.7 does not hold. Then we can choose a sequence of balls \( \{Q_j\} \) such that (5.9) and \( (5.10) \) hold.

Note that \( T \) satisfies the non-degenerate condition (1.3). This implies that there exist positive constants \( 3 \leq A_1 \leq A_2 \) such that for any ball \( B := B(x_0, r) \subset X \), there exist balls \( \tilde{B} := B(y_0, r) \) such that \( A_1r \leq d(x_0, y_0) \leq A_2r \), and for all \( (x, y) \in (\tilde{B} \times \tilde{B}) \), \( K(x,y) \) does not change sign and

\[
|K(x,y)| \geq \frac{1}{\mu(B)}. \quad (6.1)
\]

If the kernel \( K(x,y) := K_1(x,y) + iK_2(x,y) \) is complex-valued, where \( i^2 = -1 \), then at least one of \( K_i \) satisfies (6.1)’

Thus, we see that for any ball \( Q_j \) in the selected sequence that satisfies (5.9), there is another ball \( Q_j \) such that the kernel \( K \) satisfies (6.1) above.

By using the same argument as in the proof of part (ii) in Theorem 1.2, we can construct the sets \( \tilde{F}_{j,1} \) and \( \tilde{F}_{j,2} \) that satisfies (5.14). Again, let us suppose that the first of these inequalities holds, i.e.,

\[
\frac{2}{\nu(Q_j)} \int_{E_{j,1}} \left| b(x) - m_b(\tilde{Q}_j) \right| \, d\mu(x) \geq \frac{\delta_0}{2}.
\]

Hence for every \( j \), using the non degenerate condition (6.1) and using (5.14) we have that

\[
\delta_0 \leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \left| b(x) - m_b(\tilde{Q}_j) \right| \, d\mu(x) \leq \frac{1}{\nu(Q_j)} \frac{\mu(\tilde{F}_{j,1})}{\mu(Q_j)} \int_{E_{j,1}} \left| b(x) - m_b(\tilde{Q}_j) \right| \, d\mu(x) \quad (6.2)
\]

Using this and noting that \( K(x,y) \geq \frac{1}{\mu(Q_j)} \), \( b(x) - b(y) \geq b(x) - m_b(\tilde{Q}_j) > 0 \) when \( x \in E_{j,1} \) and \( y \in \tilde{F}_{j,1} \), we obtain that

\[
\delta_0 \leq \frac{1}{\nu(Q_j)} \int_{E_{j,1}} \int_{\tilde{F}_{j,1}} \frac{1}{\mu(Q_j)} \left( b(x) - m_b(\tilde{Q}_j) \right) x_{\tilde{F}_{j,1}}(y) \, d\mu(y) \, d\mu(x) \quad (6.3)
\]
Consider \( f_j =: \frac{\chi_{F_j,1}}{\lambda_1(Q_j)^{\frac{1}{p}}} \) observe that this is a sequence of disjointly supported functions using the equation \((5.14)\) with \( ||f_j||_{L_{\lambda_1(X)}^p} \simeq 1 \). Now using Hölder’s inequality we get

\[
\begin{align*}
\delta_0 \lesssim & \frac{1}{\lambda_2(Q_j)^{\frac{1}{p}}} \int_{E_{j,1}} ||b, T|| f_j(x) \lambda_2^\frac{1}{p}(x) \lambda_2^\frac{1}{p}(x) d\mu(x) \\
\lesssim & \frac{1}{\lambda_2(Q_j)^{\frac{1}{p}}} \lambda_2^\frac{1}{p}(E_{j,1}) \left( \int_X ||b, T||(f_j(x)) \lambda_2(x) d\mu(x) \right)^\frac{1}{p} \\
\lesssim & \left( \int_X ||b, T||(f_j(x)) \lambda_2(x) d\mu(x) \right)^\frac{1}{p}
\end{align*}
\]

Let us consider \( \psi \) in the closure of \( \{ [b, T](f_j) \} \), then we have \( ||\psi||_{L_{\lambda_2(X)}^p} \gtrsim 1 \). Now choose some \( j \) such that

\[
||\psi - [b, T](f_j)||_{L_{\lambda_2}^p(X)} \leq 2^{-i}.
\]

To complete the proof consider a non-negative numerical sequence \( \{c_i\} \) with \( ||c_i||_{L_{p'}} < \infty \) but \( ||c_i||_{L_1} = \infty \). Then consider \( \phi = \sum_i c_i f_j \in L_{\lambda_1}^p(X) \) and

\[
||\sum_i c_i \psi - [b, T](\phi)||_{L_{\lambda_1}^p(X)} \leq ||\sum_i c_i (\psi - [b, T](f_j))||_{L_{\lambda_2}^p(X)} \leq ||c_i||_{L_{p'}} \left( \sum_i ||\psi - [b, T](f_j)||_{L_{\lambda_2}^p(X)} \right)^\frac{1}{p} \lesssim 1.
\]

Hence we conclude that \( \sum_i c_i \psi \in L_{\lambda_1}^p(X) \), but \( \sum_i c_i \psi \) is an infinite on set of positive measure which is contradiction that completes our proof. \( \square \)

7 Compact Bilinear Commutator

7.1 Proof of Theorem 1.6

We begin to prove the Theorem 1.6. We first prove \( T_{S,b}^{B,*} \) is compact. The compactness of \( T_{S,b}^B \) can be archived by the dual argument for \( p > 1 \) and a simple derivation for \( 1/2 < p \leq 1 \), which we put at the end of this section.

For \( 1/2 < p \leq 1 \), note that

\[
 ||T_{S,b}^{B,*}(f, g)||_{L_p(\hat{w})}^p = \int_X \left( \sum_{Q \in S} \frac{1}{\mu(Q)} |(b(y) - b_Q) f(y) g(y) \chi_Q(x) \hat{w}(x) d\mu(x) \right)^p \leq \sum_{Q \in S} \left( \frac{1}{\mu(Q)} |(b(y) - b_Q) f(y) g(y) \chi_Q(x) \hat{w}(x) d\mu(x) \right)^p g_Q^p \int_Q \hat{w}(x) d\mu(x)
\]

which completes our proof.
Recall that $T$ when $(Q > N)$, which gives that $4 < N$, we choose number $\delta > 0$, or $l(Q) < \delta$, or $Q \cap Q_N = \emptyset$.

We now write $T_{S_b}^{B,*}(f,g)(x)$ as follows.

$$T_{S_b}^{B,*}(f,g)(x) = \sum_{Q > Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| f(y) d\mu(y) \right) g_Q \chi_Q(x)$$

For $\epsilon > 0$, from Definitions 2.7, 3.1, 3.2 and Proposition 3.3, we choose number $N > 0$, $\delta > 0$ and cube $Q_N$ side length $N$ such that these conditions hold

$$\frac{1}{\nu(Q)} \int_Q |b(x) - b_Q| d\mu(x) < \epsilon,$$

$$\left( \frac{1}{\lambda_1(Q)} \int_Q |b(x) - b_Q|^p \lambda_2(x) d\mu(x) \right)^{\frac{1}{p}} < \epsilon,$$

$$\left( \frac{1}{\lambda_2(Q)} \int_Q |b(x) - b_Q|^p \lambda_1(x) d\mu(x) \right)^{\frac{1}{p}} < \epsilon$$

when $l(Q) > N$, or $l(Q) < \delta$, or $Q \cap Q_N = \emptyset$.

We now write $T_{S_b}^{B,*}(f,g)(x)$ as follows.

$$T_{S_b}^{B,*}(f,g)(x) = \sum_{Q > Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q| f(y) d\mu(y) \right) g_Q \chi_Q(x)$$

For $T_4(f,g)(x)$, note that there are only finitely many cubes contained in $Q_N$ such that $\delta < l(Q) < N$, which gives that $T_4(f,g)(x)$ has finite range, and thus it is a compact operator.
We will now show that the norm of $T_1(f, g)(x), T_2(f, g)(x), T_3(f, g)(x)$ is at most $\epsilon$. Let us start with the estimate for the norm of $T_3(f, g)(x)$; i.e.,

$$
\|T_3(f, g)(x)\|_{L_p^p(X)} \leq \epsilon \|f\|_{L_{\lambda_1}^1(X)} \|g\|_{L_{\mu}^q(X)}. \quad (7.3)
$$

Recall that $\nu = \lambda_1^{1/p_1} \lambda_2^{-1/p_2}$, thus for some $\epsilon > 0$, by Lemma 2.13, we have the following

$$
T_3(f, g)(x) \leq \sum_{Q \subset Q_N \atop l(Q) < \delta} \sum_{R \in S \atop R \subset Q} \left( \frac{1}{\mu(R)} \int_R |b(z) - b_R|d\mu(z) \frac{1}{\mu(Q)} \int_R |f(y)|d\mu(y) \right) g_Q \chi_Q(x) \quad (7.4)
$$

$$
= \sum_{Q \subset Q_N \atop l(Q) < \delta} \sum_{R \in S \atop R \subset Q} \frac{1}{\nu(R)} \int_R |b(z) - b_R|d\mu(z) \left( \frac{1}{\mu(R)} \int_R |f(y)|\nu(R)d\mu(y) \right) \frac{1}{\mu(Q)} g_Q \chi_Q(x)
$$

$$
\leq \epsilon \sum_{Q \subset Q_N \atop l(Q) < \delta} \left( \sum_{R \in S \atop R \subset Q} \frac{1}{\mu(Q)} \int_Q A_S(|f|)(y)\nu(y)dy \right) g_Q \chi_Q(x)
$$

$$
\leq \epsilon A_S^B \left( A_S(|f|)\nu, g \right)(x).
$$

Here we have used from Definition 2.7 that for $b \in VMO_\nu(X)$ such that when $l(Q) < \delta$, we have $\frac{1}{\nu(Q)} \int_Q |b(y) - b_Q|d\mu(y) < \epsilon$. Here $A_S^B$ is the classical bilinear sparse operator.

Now observe that from classical weighted boundedness of sparse operator and bilinear sparse operator, we have that

$$
\|T_3(f, g)(x)\|_{L_p^p(X)} \leq \epsilon \|A_S^B (A_S(|f|)\nu, g)(x)\|_{L_p^p(X)} \quad (7.5)
$$

$$
\leq \epsilon \|A_S^B (A_S(|f|)\nu)\|_{L_{\lambda_1}^1(X)} \|g\|_{L_{\mu}^q(X)}
$$

$$
= \epsilon \|A_S^B (|f|)\|_{L_{\lambda_1}^1(X)} \|g\|_{L_{\mu}^q(X)}
$$

$$
\leq \epsilon \|A_S^B (|f|)\|_{L_{\lambda_1}^1(X)} \|g\|_{L_{\mu}^q(X)}.
$$

Then this finishes the proof for the control of the norm of $T_3(f, g)(x)$.

For $T_2(f, g)$, recall that

$$
T_2(f, g)(x) = \sum_{Q \cap Q_N = \emptyset} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q|f(y)d\mu(y) \right) g_Q \chi_Q(x) \quad (7.6)
$$

Following similar approach in the estimate for $T_3(f, g)$, we write $|b - b_Q|$ as in (4.4). Since $Q \cap Q_N = \emptyset$ and $R \subset Q$, we have for all $R \in S$ in (4.4), $R \cap Q_N = \emptyset$. Now according to Definition 2.7, we have that $\frac{1}{\nu(R)} \int_{R} |b(x) - b_R|d\mu(x) \leq \epsilon$. Then following the same arguments as we did for the control of the norm $T_3(f, g)(x)$, we obtain similar estimate for the norm of $T_2(f, g)(x)$, i.e.,

$$
\|T_2(f, g)\|_{L_p^p(X)} \leq \epsilon \|f\|_{L_{\lambda_1}^1(X)} \|g\|_{L_{\mu}^q(X)}. \quad (7.7)
$$
Now let us show the control for the norm of $T_1(f,g)(x)$, recall that

$$
T_1(f,g)(x) = \sum_{Q \ni Q_N} \left( \frac{1}{\mu(Q)} \int_Q |b(y) - b_Q||f(y)|d\mu(y) \right) g_Q \chi_Q(x).
$$

(7.8)

We will start with a collection of sparse dyadic cubes $Q = Q_1 \supset Q_2 \supset Q_3 \supset Q_4 \ldots \supset Q_{\tau_Q} \supset Q_{\tau_Q+1} = Q_N$, where $Q_i$ is the “parent” of $Q_{i+1}$, $i = 1, 2, \ldots, \tau_Q$. For the sake of the sparse property, if the parent of $Q_{i+1}$ has only one child $Q_{i+1}$, we should still denote by $Q_{i+1}$ the parent of $Q_{i+1}$ since they are the same dyadic cube indeed. Then repeat the process until we find $Q_i$ such that $Q_i$ has at least two children and $Q_{i+1}$ is one of them. For each $Q_i$, $i = 1, 2, \ldots, \tau_Q$, we denote all its dyadic children except $Q_{i+1}$ by $Q_{i,k}$, $k = 1, 2, \ldots, M_{Q_i}$ where $M_{Q_i} + 1$ is the number of the children of $Q_i$ and less than uniform constant $M$ in (2.7). Hence for all $i = 0, 1, 2, \ldots, \tau_Q$ and $k = 1, 2, \ldots, M_{Q_i}$, $Q_{i,k} \cap Q_N = \emptyset$. Note that $Q_{i+1}$ and $Q_{i,k}$ have equivalent measures since it follows from (2.8) that

$$
\mu(Q_{i+1}) \leq \mu(B(Q_i)) \leq C(1 + \frac{d(x_{Q_i}, x_{Q_{i+1}})}{C_1 \delta_{Q_i}})^n \mu(B(x_{Q_{i,k}}), c_1 \delta_{Q_{i,k}}^{k_i+1}) \leq C2^n (\frac{C_1}{c_1 \delta})^n \mu(Q_{i,k}).
$$

And thus there exists uniform constant $0 < \tilde{\eta} < 1$ such that $\mu(Q_{i+1}) \leq \tilde{\eta}\mu(Q_i)$, which ensure the sparse property of the collection of $\{Q_i\}$. Then

$$
T_1(f,g)(x) \leq \sum_{Q \ni Q_N} \left( \sum_{i=1}^{\tau_Q} \sum_{k=1}^{M_{Q_i}} \frac{1}{\mu(Q)} \int_{Q_{i,k}} |b(y) - b_{Q_{i,k}}||f(y)|d\mu(y) \right) g_Q \chi_Q(x) \tag{7.9}
$$

$$
+ \sum_{Q \ni Q_N} \left( \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}|f(y)|d\mu(y) \right) g_Q \chi_Q(x)
$$

$$
+ \sum_{Q \ni Q_N} \left( \sum_{i=1}^{\tau_Q} \sum_{k=1}^{M_{Q_i}} \int_{Q_{i,k}} |f(y)|d\mu(y) \right) g_Q \chi_Q(x)
$$

$$
+ \sum_{Q \ni Q_N} \left( \int_{Q_N} |f(y)|d\mu(y) \right) g_Q \chi_Q(x)
$$

$$= : I + II + III + IV.
$$

It suffices to prove that

$$
\|I\|_{L^p_w(X)} + \|II\|_{L^p_w(X)} + \|III\|_{L^p_w(X)} + \|IV\|_{L^p_w(X)} \lesssim C\|f\|_{L^p_{\lambda_1}(X)}\|g\|_{L^p_{\lambda_2}(X)},
$$

(7.10)

where the implicit constants depends only on $\lambda_1$, $\lambda_2$, $w$.

Let us now begin with the estimate of the norm of $II$. Recall that $\lambda_1' = \lambda_1^{\frac{1}{p_1-1}}$, $\lambda_2' = \lambda_2^{\frac{1}{p_1-1}}$, $w' = \frac{1}{w^{p_2-1}}$ and $\tilde{w}' = \frac{1}{\tilde{w}^{p_2-1}}$.

$$
\|II\|_{L^p_w(X)} \tag{7.11}
$$

$$
= \left\| \sum_{Q \ni Q_N} \left( \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}|f(y)|d\mu(y) \right) g_Q \chi_Q(x) \right\|_{L^p_w(X)}
$$

$$
= \sup_{\|h\|_{L^p_{w'}} \leq 1} \left( \sum_{Q \ni Q_N} \left( \frac{1}{\mu(Q)} \int_{Q_N} |b(y) - b_{Q_N}|f(y)|d\mu(y) \right) g_Q \chi_Q(x), h(x) \right) \right.
$$
\[ \leq \sup_{\|h\|_{L_{p'}^w} \leq 1} \sum_{Q \supseteq Q_N} \frac{1}{\mu(Q)} \int_{Q} |b(y) - b_{Q_N}| \|f(y)\|_{L^p} g_{Q_N} \int_{Q} |h(x)| \, d\mu(x) \]

\[ \leq \sup_{\|h\|_{L_{p'}^w} \leq 1} \sum_{Q \supseteq Q_N} \frac{1}{\mu(Q)} \left( \int_{Q} |b(y) - b_{Q_N}| |f(y)| \, d\mu(y) \right)^{\frac{1}{p_1}} \left( \int_{Q} |f(x)|^{p_1} \lambda_1(x) \, d\mu(x) \right)^{\frac{1}{p_1}} \]

\[ \times \frac{1}{\mu(Q)} \left( \int_{Q} |g(x)|^{p_2} w(x) \, d\mu(x) \right)^{\frac{1}{p_2}} w'(Q)^{\frac{1}{p_2}} \left( \int_{Q} |h(x)|^{p_2} \hat{w}'(x) \, d\mu(x) \right)^{\frac{1}{p_2}} \hat{w}(Q)^{\frac{1}{p_2}} \]

\[ \leq \sup_{\|h\|_{L_{p'}^w} \leq 1} \sum_{Q \supseteq Q_N} \frac{1}{\mu(Q)} \left( \frac{1}{\lambda_2'(Q_N)} \int_{Q} |b(y) - b_{Q_N}| |f(y)| \, d\mu(y) \right)^{\frac{1}{p_1}} \lambda_2'(Q_N)^{\frac{1}{p_1}} \lambda_2'(Q_N)^{\frac{1}{p_1}} \hat{w}(Q)^{\frac{1}{p_2}}. \]

Observe that since \( Q \supseteq Q_N \) and thus \( l(Q) > N \),

\[ \left( \frac{1}{\lambda_2'(Q_N)} \int_{Q} |b(x) - b_{Q_N}| \lambda_1'(x) \, d\mu(x) \right)^{\frac{1}{p_1}} < \epsilon. \]

Recall that there exists some \( \sigma > 0 \) such that \( \lambda_2 \in A_{p_1 - \sigma} \) as \( \lambda_2 \in A_{p_1} \), and that

\[ \frac{\lambda_2(Q)}{\lambda_2(Q_N)} \leq \left( \frac{\mu(Q)}{\mu(Q_N)} \right)^{p_1 - \sigma} \left[ \lambda_2 \right]_{A_{p_1}}. \] (7.12)

And noting that \( 1 = \frac{p_1}{p_1} + \frac{p_2}{p_2} \), by Hölder’s inequality we have

\[ \hat{w}(Q)^{\frac{1}{p_2}} = \left( \int_Q \lambda_2^{\frac{p_1}{p_2}} w^{\frac{p_2}{p_2}} \, d\mu \right)^{\frac{1}{p_2}} \leq \left( \int_Q \lambda_2^2 \, d\mu \right)^{\frac{1}{p_1}} \left( \int_Q w \, d\mu \right)^{\frac{1}{p_2}} = \lambda_2(Q)^{\frac{1}{p_1}} w(Q)^{\frac{1}{p_2}}. \]

So we have

\[ \|II\|_{L_{p'}^w(X)} \leq \epsilon \|g\|_{L_{p_2}^w(X)} \sum_{Q \supseteq Q_N} \frac{\lambda_2(Q)^{\frac{p_1}{p_2}} \lambda_2(Q_N)^{\frac{p_1}{p_2}}}{\mu(Q_N)} \frac{\mu(Q)}{\lambda_2(Q)^{\frac{p_1}{p_2}}} \frac{\lambda_2(Q_N)^{\frac{p_1}{p_2}}}{\mu(Q)} \]

\[ \times \frac{\hat{w}(Q)^{\frac{1}{p_2}} w'(Q)^{\frac{1}{p_2}} w(Q)^{\frac{1}{p_2}}}{\mu(Q)} \]

\[ \leq \epsilon \|g\|_{L_{p_2}^w(X)} \sum_{Q \supseteq Q_N} \left[ \lambda_2 \right]_{A_{p_1}}^{\frac{p_1}{p_2}} \frac{\mu(Q)}{\mu(Q_N)} \left( \frac{\mu(Q)}{\mu(Q_N)} \right)^{p_1 - \sigma} \left[ \lambda_2 \right]_{A_{p_1}}^{\frac{p_1}{p_2}} |w|_{A_{p_2}}^{\frac{1}{p_2}} \sum_{Q \supseteq Q_N} \left( \frac{\mu(Q)}{\mu(Q_N)} \right)^{\frac{p_1}{p_2}} \]

\[ \leq \epsilon \|g\|_{L_{p_2}^w(X)} \left[ \lambda_2 \right]_{A_{p_1}}^{\frac{p_1}{p_2}} |w|_{A_{p_2}}^{\frac{1}{p_2}}. \]

This gives the control for the norm of \( II \).

Let us now prove the control for the norm of \( I \). Again, we would like to change the order of the summation for \( Q \) and \( k \). Thus we may assume that \( Q_{i,k} = \emptyset \) when \( M \geq k > M_Q \), and the
corresponding terms are 0.

\[ I \sum_{k=1}^{M} \sum_{Q \supseteq Q_N} \left( \sum_{i=1}^{\tau_Q} \frac{1}{\mu(Q)} \int_{Q_{i,k}} \left| b(y) - b_{Q_{i,k}} \right| f(y) |d\mu(y)| \right) g_Q \chi_Q(x). \]  

(7.13)

Fixing \( k \), then for each \( Q_{i,k} \) where \( i = 1, \ldots, \tau_Q \), following similar approach in the estimate for \( T_3(f, g) \), we write \( |b - b_{Q_{i,k}}| \) as in (4.4). Since \( Q_{i,k} \cap Q_N = \emptyset \) and \( R \subset Q_{i,k} \), we have for all \( R \in \hat{S} \) in (4.4), \( R \cap Q_N = \emptyset \). Now according to Definition 2.7, we have that \( \frac{1}{\nu(R)} \int_R |b(x) - b_R| |d\mu(x)| \leq \epsilon \).

Following similar estimates in equations (7.4) and (7.5), we obtain that

\[ \| I \|_{L^p_v(X)} \leq \epsilon \| \tilde{w} \|_{A_p} \left( \| \lambda_1 \|_{A_p} \right)^{\max\{1, \frac{\nu(Q)}{\nu_1-1} \}} \| f \|_{L^p_{\lambda_1}(X)} \| g \|_{L^p_{\nu}(X)}. \]  

(7.14)

We now turn to the estimates for the norm of III and IV. Observe that for each fixed \( k \) and for each \( Q_{i,k} \) we will obtain the same estimate as follows

\[ \| A_{Q_{i,k}} \|_{L^p_v(X)} \leq \epsilon \| \tilde{w} \|_{A_p} \left( \| \lambda_1 \|_{A_p} \right)^{\max\{1, \frac{\nu(Q)}{\nu_1-1} \}} \| f \|_{L^p_{\lambda_1}(X)} \| g \|_{L^p_{\nu}(X)}, \]  

(7.15)

where

\[ A_{Q_{i,k}}(x) = \sum_{Q \supseteq Q_N} \left( \sum_{i=1}^{\tau_Q} \left| b_{Q_{i,k}} - b_Q \right| \frac{1}{\mu(Q)} \int_{Q_{i,k}} \left| f(y) \right| |d\mu(y)| \right) g_Q \chi_Q(x). \]

Hence, the estimates of the norms of III and IV follow from (7.15).

Thus, it suffices to prove (7.15). Recall the definition of \( Q_i \) and \( Q_{i,k} \): \( Q = Q_1 \supseteq Q_2 \supseteq Q_3 \supseteq Q_4 \ldots \supseteq Q_{\tau_Q} = Q_N \), where \( Q_i \) is the “parent” of \( Q_{i+1}, i = 1, 2, \ldots, \tau_Q \). The collection \{\( Q_i \)\} is sparse. For each \( Q_i, i = 1, 2, \ldots, \tau_Q \), we denote all its dyadic children except \( Q_{i+1} \) by \( Q_{i,k}, k = 1, 2, \ldots, M_{Q_i} \). By using (4.18) we have that

\[ |b_{Q_{i,k}} - b_Q| \leq C \epsilon \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)}, \]

which implies

\[ A_{Q_{i,k}} \leq C \sum_{Q \supseteq Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \frac{1}{\mu(Q)} \int_{Q_{i,k}} \left| f(y) \right| |d\mu(y)| g_Q \chi_Q(x). \]  

(7.16)

Hence we have

\[ \| A_{Q_{i,k}} \|_{L^p_v(X)} \]  

(7.17)

\[ \leq \left\| \sum_{Q \supseteq Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \frac{1}{\mu(Q)} \int_{Q_{i,k}} \left| f(y) \right| |d\mu(y)| g_Q \chi_Q(x) \right\|_{L^p_v(X)} \]

\[ \leq \sup_{h \in L^p_v(X)} \left( \left\| \sum_{Q \supseteq Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \frac{\nu(Q_j)}{\mu(Q_j)} \frac{1}{\mu(Q)} \int_{Q_{i,k}} \left| f(y) \right| |d\mu(y)| g_Q \chi_Q(x), h(x) \right\| \right) \]

\[ \leq \epsilon \left( \sum_{Q \supseteq Q_N} \sum_{i=1}^{\tau_Q} \sum_{j=1}^{i-1} \left( \frac{\nu(Q_j)}{\mu(Q_j)} \right)^{p_1} \left\| \int_{Q_{i,k}} \left| f(y) \right| |d\mu(y)| \right\|_{L^{p_1}_{\lambda_1}} \frac{1}{\mu(Q)} \left( \frac{\mu(Q)}{\mu(Q_{i,k})} \right)^{p_1} w(Q)^{-\frac{p_1}{p_2}} \hat{w}(Q)^{(1-\frac{1}{p_1})p_1} \right)^{\frac{1}{p_1}} \]
where we use the facts that

\[
\begin{align*}
&\times \left( \sum_{Q > Q_N} \sum_{i=1}^{\tau_Q} \frac{\mu(Q_{i,k})}{\mu(Q)} \right)^{p_2\sigma'/2} \left( \frac{1}{\mu(Q)} \int_Q |g(y)| d\mu(y) \right)^{p_2} w(Q) \quad \frac{1}{p_2} \\
&\times \left( \sum_{Q > Q_N} \sum_{i=1}^{\tau_Q} \frac{\mu(Q_{i,k})}{\mu(Q)} \right)^{p'\sigma'/2} \left( \int_Q |h(y)| d\mu(y) \right)^{p'} (\hat{w}(Q))^{-p'} \hat{w}(Q) \quad \frac{1}{p'} \\
= : \epsilon A_B A_B^B D^{\frac{1}{p'}},
\end{align*}
\]

where \( \sigma' = \frac{\sigma}{p} \). Now observe that

\[
D \leq \sum_{Q > Q_N} \left[ \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \left( \frac{\mu(Q_i)}{\mu(Q)} \right)^{p'\sigma'} \right] \left( \frac{1}{\hat{w}(Q)} \int_Q |h(y)| d\mu(y) \right)^{p'} \hat{w}(Q) \tag{7.18}
\]

\[
\leq C_1 C_2 \sum_{Q > Q_N} \inf_{x \in Q} \mathcal{M}_{\hat{w}}^p(|h| \hat{w}^{-1})(x) \hat{w}(E(Q))
\]

\[
\leq C_1 C_2 \sum_{Q > Q_N} \int_{E(Q)} \mathcal{M}_{\hat{w}}^p(|h| \hat{w}^{-1})(x) \hat{w}(x) d\mu(x)
\]

\[
\leq C_1 C_2 \int_{E(Q)} \mathcal{M}_{\hat{w}}^p(|h| \hat{w}^{-1})(x) \hat{w}(x) d\mu(x)
\]

\[
\leq C_1 C_2 \|h \hat{w}^{-1}\|_{L_{\hat{w}}^p(X)}^{p'}
\]

where we use the facts that

\[
\sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \left( \frac{\mu(Q_i)}{\mu(Q)} \right)^{p'\sigma'} \leq C_1
\]

with \( C_1 \) an absolute positive constant, that

\[
\mu(E(Q)) = \int_{E(Q)} \frac{1}{\hat{w}(E(Q))} \lambda_2(E(Q))^{\frac{1}{2p_1}} \lambda_2(E(Q))^{\frac{1}{2p_2}} w'(E(Q))^{\frac{1}{2p_2}} d\mu \leq \frac{1}{\hat{w}(E(Q))} \lambda_2(E(Q))^{\frac{1}{2p_1}} w'(E(Q))^{\frac{1}{2p_2}}
\]

and that

\[
\frac{\hat{w}(Q)}{\hat{w}(E(Q))} \leq \frac{\hat{w}(Q) \lambda_2(E(Q))^{\frac{1}{2p_1}} w'(E(Q))^{\frac{1}{2p_2}}}{\mu(E(Q))^{\frac{1}{2p}}} \leq C [\lambda_2]_{A_{p_1}} [w]_{A_{p_2}}^{\frac{1}{2p_2}}.
\]

Here \( E(Q) \) is the measurable subset of \( Q \) in some \( \eta \) Sparse collection of cubes \( S \) such that \( \mu(E_Q) \geq \eta \mu(Q) \).

Similarly we can estimate \( B \) term as

\[
B \leq \sum_{Q > Q_N} \left[ \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \left( \frac{\mu(Q_i)}{\mu(Q)} \right)^{p_2\sigma'} \right] \left( \frac{1}{\mu(Q)} \int_Q |g(y)| d\mu(y) \right)^{p_2} w(Q) \tag{7.19}
\]

\[
\leq C_1 C_2 \sum_{Q > Q_N} \inf_{x \in Q} \mathcal{M}^{p_2}(|g|)(x) w(E(Q))
\]

\[
\leq C_1 C_2 \sum_{Q > Q_N} \int_{E(Q)} \mathcal{M}^{p_2}(|g|)(x) w(x) d\mu(x)
\]
Compactness Characterization for Bloom Sparse Operators

\[ \leq C_1 C_2 \int_{\mathbb{R}^n} M^{p_2}(|g|)(x) w(x) d\mu(x) \]
\[ \leq C_1 C_2 [w]_{A_{p_2}} \|g\|^{p_2}_{L^{p_2}_{w}(X)} \]

We also have the following estimate for \( A \)

\[ A \leq \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \frac{\lambda_1(Q_i) \lambda_2(Q_j)^{p_1-1}}{\mu(Q_j)^{p_1}} \|f\|^{p_1}_{L_{\lambda_1}^{p_1}(Q_i,k)} \lambda_1^{p_1} \frac{1}{\mu(Q)} \left( \frac{\mu(Q)}{\mu(Q_i,k)} \right)^{p_1 \sigma} \lambda_2(Q) \]
\[ \leq \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \frac{\lambda_1(Q_i) \lambda_2(Q_j)^{p_1-1}}{\mu(Q_j)^{p_1}} \lambda_1(Q_i) \frac{\lambda_1(Q_j)^{p_1}}{\mu(Q_j)^{p_1}} \lambda_2(Q_j)^{p_1-1} \lambda_1(Q_i) \frac{\lambda_1(Q_j)^{p_1}}{\mu(Q_j)^{p_1}} \lambda_2(Q_j) \frac{\lambda_1(Q_j)^{p_1}}{\mu(Q_j)^{p_1}} \lambda_2(Q) \]
\[ \leq \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \frac{\lambda_1(Q_i) \lambda_2(Q_j)^{p_1-1}}{\mu(Q_j)^{p_1}} \lambda_1(Q_i) \frac{\lambda_1(Q_j)^{p_1}}{\mu(Q_j)^{p_1}} \lambda_2(Q_j) \frac{\lambda_1(Q_j)^{p_1}}{\mu(Q_j)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \supset Q_N} \sum_{i=1}^{\tau_Q} \log \left( \frac{\mu(Q)}{\mu(Q_i)} \right) \frac{\lambda_1^{p_1}}{\mu(Q_i)^{p_1}} \lambda_2(Q) \]

Then (7.15) follows from (7.19), (7.20) and (7.17). This completes the proof of the compactness of \( \mathcal{T}^{B^*}_{S,b} \).

Now let us consider \( \mathcal{T}^{B*}_{S,b} \). For \( p > 1 \), the compactness of \( \mathcal{T}^{B*}_{S,b} \) can be archived by the dual argument and the result of \( \mathcal{T}^{B^*}_{S,b} \). For \( 1/2 < p \leq 1 \), note that

\[ \|\mathcal{T}^{B^*}_{S,b}(f,g)\|_{L^p(\tilde{w})}^p = \int_X \left( \sum_{Q \in S} |b(x) - b_Q| f_Q g_Q \chi_Q(x) \right)^p \tilde{w}(x) d\mu(x) \]
\[ \leq \sum_{Q \in S} \int_X (|b(x) - b_Q| f_Q g_Q \chi_Q(x))^p \tilde{w}(x) d\mu(x) \]
\[ = \sum_{Q \in S} f_Q^p g_Q^p \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq \sum_{Q \in S} f_Q^p g_Q^p \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]

and

\[ \sum_{Q \in S} f_Q^p g_Q^p \lambda_1(Q)^{\frac{p}{p_2}} w(Q)^{\frac{p}{p_2}} \leq \left( \sum_{Q \in S} f_Q^p \lambda_1(Q)^{\frac{p}{p_2}} \right)^{\frac{p}{p_2}} \left( \sum_{Q \in S} g_Q^p w(Q)^{\frac{p}{p_2}} \right)^{\frac{p}{p_2}} \]
\[ \leq \|Mf\|^{p}_{L_{\lambda_1}^{p_1}(X)} \|Mg\|^{p}_{L_{\lambda_2}^{p_2}(X)} \]
\[ \leq C \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq C \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
\[ \leq C \lambda_1 \lambda_2 \sum_{Q \in S} \int_Q (|b(x) - b_Q|)^p \chi_Q^p(x) w(x)^{\frac{p}{p_2}} d\mu(x) \]
Thus we can follow the similar argument for linear case or $T_{S,b}^{B,*}$, just noting that when the side-length of $Q$ is small enough or large enough or $Q$ is far away, \[
\left( \frac{1}{\lambda_1(Q)} \int_Q \left( |b(x) - b_Q| \right)^{p_1} \lambda_2(x) d\mu(x) \right)^{1/p_1}
\]
is less than any given $\epsilon > 0$.
This completes the proof of the whole theorem.

Acknowledgments: P. Chen is supported by NNSF of China, Grant No. 12171489 and Guangdong Natural Science Foundation 2022A1515011157. Lacey is a 2020 Simons Fellow, and is supported in part by grant from the US National Science Foundation, DMS-1949206. J. Li is supported by ARC DP 220100285.

References

[1] M. Agcayazi, A. Gogatishvili, K. Koca and R. Mustafayev, A note on maximal commutators and commutators of maximal functions, *J. Math. Soc. Japan*, 67 (2) (2015), 581–593.
[2] J. Bastero, M. Milman and F. J. Ruiz, Commutators for the maximal and sharp functions, *Proc. Amer. Math. Soc.*, 128 (2000), 3329–3334.
[3] P. Chen, X. Duong, J. Li and Q.Y. Wu, Compactness of Riesz transform commutator on stratified Lie groups, *J. Funct. Anal.*, 277 (6) (2019), 1639–1676.
[4] S. Bloom, A commutator theorem and weighted BMO, *Trans. Amer. Math. Soc.*, 292 (1985), 103–122.
[5] M. Christ, A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral, *Colloq. Math.*, 60/61 (1990), no. 2, 601–628.
[6] A. Culiuc, F. Di Plinio and Y. Ou, Domination of multilinear singular integrals by positive sparse forms, *J. London Math. Soc.*, 98 (2018), no. 2, 369–392.
[7] R. R. Coifman and G. Weiss, Factorization theorems for Hardy spaces in several variables, *Ann. of Math.* (2) 103 (1976), 611-635.
[8] X. Duong, R. Gong, M.-J. S. Kuffner, J. Li, B.D. Wick and D. Yang, Two weight commutators on spaces of homogeneous type and applications, *J. Geom. Anal.*, 31 (2021), 980–1038.
[9] X. Duong, I. Holmes, J. Li, B.D. Wick and D. Yang, Two weight commutators in the Dirichlet and Neumann Laplacian settings, *J. Funct. Anal.*, 276 (2019), 1007–1060.
[10] L.C. Evans and R.F. Gariepy, Measure theory and fine properties of functions. *Studies in Advanced Mathematics. CRC Press, Boca Raton, FL*, (1992).
[11] Z. Fu, E. Pozzi and Q. Wu, Commutators of maximal functions on spaces of homogeneous type and their weighted, local versions, *arXiv:1907.09953*.
[12] R. Gong, M. N. Vempati and Q. Wu, A note on two weight commutators of maximal functions on spaces of homogeneous type, *arXiv:2012.00575*.
[13] J. Garcia-Cuerva, E. Harboure, C. Segovia, J.L. Torrea, Weighted norm inequalities for commutators of strongly singular integrals, *Indiana Univ. Math. J.*, 40 (1991), 1397–1420.
[14] I. Holmes, M. Lacey and B. D. Wick, Commutators in the two-weight setting, *Math. Ann.*, 367 (2017), no. 1–2, 51–80.
[15] G. Hu and D. Yang, Maximal commutators of BMO functions and singular integral operators with non-smooth kernels on spaces of homogeneous type, *J. Math. Anal. Appl.*, 354 (2009), 249–262.
[16] T. Hytönen and A. Kairema, Systems of dyadic cubes in a doubling metric space, *Colloq. Math.* **126** (2012), no. 1, 1–33.
[17] S. Janson, Mean oscillation and commutators of singular integral operators, *Ark. Mat.*, **16** (1978), no. 2, 263–270.
[18] A. Kairema, J. Li, C. Pereyra and L. A. Ward, Haar bases on quasi-metric measure spaces, and dyadic structure theorems for function spaces on product spaces of homogeneous type, *J. Funct. Anal.*, **271** (2016), 1793–1843.
[19] M. Lacey, An elementary proof of the $A^2$ Bound, *Israel Journal of Mathematics*, **217** (2017), 181–195.
[20] M. Lacey and J. Li, Compactness of commutator of Riesz transforms in the two weight setting, *J. Math. Anal. Appl.*, **508** (2022), 125869.
[21] A.K. Lerner, On pointwise estimates involving sparse operators, *New York J. Math.*, **22** (2016), 341–349.
[22] A.K. Lerner and F. Nazarov, Intuitive dyadic calculus: the basics, *arXiv*: 1508.05639.
[23] 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.
[24] A.K. Lerner, S. Ombrosi and I.P. Rivera-Ríos, Commutators of singular integrals revisited, *Bull. London Math. Soc.*, **51** (2019), 107–119.
[25] K. Moen, Sharp weighted bounds without testing or extrapolation, *Archiv der Mathematik*, **99** (2012), 457–466.
[26] B. Muckenhoupt and R. L. Wheeden, Weighted bounded mean oscillation and the Hilbert transform, *Studia Math.*, **54** (1975/76), 221–237.
[27] E. Sawyer and R. L. Wheeden, Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces, *Amer. J. Math.*, **114** (1992), 813–874.
[28] A. Uchiyama, On the compactness of operators of Hankel type, *Tohoku Math. J.*, **30** (1978), 163–171.
[29] A. Uchiyama, The factorization of $H^p$ on the space of homogeneous type, *Pacific J. Math.*, **92** (1981), 453–468.

Peng Chen, Department of Mathematics, Sun Yat-sen University, Guangzhou, 510275, China.
*E-mail*: chenpeng30@mail.sysu.edu.cn

Michael Lacey, School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA
*E-mail*: lacey@math.gatech.edu

Ji Li, School of Mathematical and Physical Sciences, Macquarie University, NSW, 2109, Australia.
*E-mail*: ji.li@mq.edu.au

Manasa N. Vempati, School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA
*E-mail*: nvempati6@gatech.edu