Maximal commutators of BMO functions and singular integral operators with non-smooth kernels on spaces of homogeneous type

Guoen Hu\(^a,1\), Dachun Yang\(^b,*,2\)

\(^a\) Department of Applied Mathematics, University of Information Engineering, PO Box 1001-747, Zhengzhou 450002, People’s Republic of China

\(^b\) School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, People’s Republic of China

**Abstract**

Let \(\mathcal{X}\) be a space of homogeneous type in the sense of Coifman and Weiss. In this paper, via a new Cotlar type inequality linking commutators and corresponding maximal operators, a weighted \(L^p(\mathcal{X})\) estimate with general weights and a weak type endpoint estimate with \(A_1(\mathcal{X})\) weights are established for maximal operators corresponding to commutators of BMO(\(\mathcal{X}\)) functions and singular integral operators with non-smooth kernels.

© 2009 Elsevier Inc. All rights reserved.

**1. Introduction**

We work on a space of homogeneous type. Let \(\mathcal{X}\) be a set endowed with a positive Borel regular measure \(\mu\) and a quasi-metric \(d\) satisfying that there exists a constant \(\kappa \geq 1\) such that for all \(x, y, z \in \mathcal{X}\),

\[
d(x, y) \leq \kappa [d(x, z) + d(z, y)].
\]

The triple \((\mathcal{X}, d, \mu)\) is called a space of homogeneous type in the sense of Coifman and Weiss [1], if \(\mu\) satisfies the following doubling condition: there exists a constant \(C \geq 1\) such that for all \(x \in \mathcal{X}\) and \(r > 0\),

\[
\mu(B(x, 2r)) \leq C \mu(B(x, r)) < \infty,
\]

here and in what follows, \(B(x, r) = \{y \in \mathcal{X}: d(y, x) < r\}\). It is easy to see that the above doubling property implies the following strong homogeneity property: there exist positive constants \(c_0\) and \(n\) such that for all \(\lambda \geq 1\), \(r > 0\) and \(x \in \mathcal{X}\),

\[
\mu(B(x, \lambda r)) \leq c_0 \lambda^n \mu(B(x, r));
\]
Moreover, there exist constants $C > 0$ and $N \in [0, n]$ such that for all $x, y \in \mathcal{X}$ and $r > 0$,

$$\mu(B(y, r)) \leq C \left( 1 + \frac{d(x, y)}{r} \right)^N \mu(B(x, r)). \quad (1)$$

We remark that although all balls defined by $d$ satisfy the axioms of complete system of neighborhoods in $\mathcal{X}$, and therefore induce a (separated) topology in $\mathcal{X}$, the balls $B(x, r)$ for $x \in \mathcal{X}$ and $r > 0$ need not to be open with respect to this topology. However, by a well-known result of Macías and Segovia [11], we know that there exists another quasi-metric $\tilde{d}$ such that

(i) there exists a constant $C \geq 1$ such that for all $x, y \in \mathcal{X}$,

$$C^{-1} \tilde{d}(x, y) \leq d(x, y) \leq C \tilde{d}(x, y);$$

(ii) there exist constants $C > 0$ and $\gamma \in (0, 1]$ such that for all $x, x', y \in \mathcal{X}$,

$$|\tilde{d}(x, y) - \tilde{d}(x', y)| \leq C [\tilde{d}(x, x')]^{\gamma} [\tilde{d}(x, y) + \tilde{d}(x', y)]^{1-\gamma}.$$  

The balls corresponding to $\tilde{d}$ are open in the topology induced by $\tilde{d}$. Thus, throughout this paper, we always assume that there exist constants $C > 0$ and $\gamma \in (0, 1]$ such that for all $x, x', y \in \mathcal{X}$,

$$|d(x, y) - d(x', y)| \leq C [d(x, x')]^{\gamma} [d(x, y) + d(x', y)]^{1-\gamma}. \quad (1.2)$$

and that the balls $B(x, r)$ for all $x \in \mathcal{X}$ and $r > 0$ are open.

Also, throughout this paper, we denote by $L^{c}_{\infty}(\mathcal{X})$ the set of bounded functions with bounded support. Let $T$ be an $L^{2}(\mathcal{X})$ bounded linear operator with kernel $K$ in the sense that for all $f \in L^{c}_{\infty}(\mathcal{X})$ and almost all $x \notin \text{supp } f$,

$$Tf(x) = \int_{\mathcal{X}} K(x, y)f(y) d\mu(y), \quad (1.3)$$

where $K$ is a measurable function on $\mathcal{X} \times \mathcal{X} \setminus \{(x, y): x = y\}$. To obtain a weak $(1, 1)$ estimate for certain Riesz transforms, and $L^{p}$-boundedness with $p \in (1, \infty)$ of holomorphic functional calculi of linear elliptic operators on irregular domains, Duong and McIntosh [2] introduced singular integral operators with non-smooth kernels on spaces of homogeneous type via the following generalized approximation to the identity.

**Definition 1.** A family of operators $\{D_{t}\}_{t>0}$ is called an approximation to the identity, if for every $t > 0$, $D_{t}$ is represented by the kernel $a_{t}$ in the following sense: for every function $u \in L^{p}(\mathcal{X})$ with $p \in [1, \infty]$ and almost everywhere $x \in \mathcal{X}$,

$$D_{t}u(x) = \int_{\mathcal{X}} a_{t}(x, y)u(y) d\mu(y),$$

and the kernel $a_{t}$ satisfies that for all $x, y \in \mathcal{X}$ and $t > 0$,

$$|a_{t}(x, y)| \leq h_{t}(x, y) = \frac{1}{\mu(B(x, t^{1/m}))} s(d(x, y)^{m}t^{-1}),$$

where $m > 0$ is a constant and $s$ is a positive, bounded and decreasing function satisfying

$$\lim_{t \to \infty} r^{\delta+m}s(r^{m}) = 0$$

for certain $\delta > N$ with $N$ appearing in (1.1).

Duong and McIntosh [2] proved that if $T$ is an $L^{2}(\mathcal{X})$ bounded linear operator with kernel $K$ and satisfies that

(i) there exists an approximation to the identity $\{D_{t}\}_{t>0}$ such that the composite operator $TD_{t}$ with $t > 0$ has an associated kernel $K_{t}$ in the sense (1.3), and there exist positive constants $c_{1}$ and $C$ such that for all $y \in \mathcal{X}$ and $t > 0$,

$$\int_{d(x, y) \geq c_{1}t^{1/m}} |K(x, y) - K_{t}(x, y)| d\mu(x) \leq C,$$
then $T$ is bounded from $L^1(\mathcal{X})$ to $L^{1,\infty}(\mathcal{X})$, that is, there exists a positive constant $C$ such that for any $f \in L^1(\mathcal{X})$ and any $\lambda > 0$,
\[
\mu(\{x \in \mathcal{X}: |Tf(x)| > \lambda\}) \leq C\lambda^{-1}||f||_{L^1(\mathcal{X})}.
\]
An $L^2(\mathcal{X})$ bounded linear operator with kernel $K$ satisfying (i) is called a singular integral operator with non-smooth kernel, since $K$ does not enjoy smoothness in space variables. Martell [12] considered the weighted $L^p(\mathcal{X})$ estimate with $A_p(\mathcal{X})$ weights for $p \in (1, \infty)$ and weighted $L^{1,\infty}(\mathcal{X})$ estimate with $A_1(\mathcal{X})$ weights for $T$. Here and in what follows, $A_p(\mathcal{X})$ with $p \in [1, \infty]$ denotes the weight function class of Muckenhoupt on $\mathcal{X}$; see, for example, [14] (or [6]) for its definition and properties. To be precise, Martell [12] proved that if $T$ is an $L^2(\mathcal{X})$ bounded linear operator, satisfies (i) and

(ii) there exists an approximation to the identity $\{\mathcal{D}_t\}_{t>0}$ such that the composite operator $\mathcal{D}_t T$ with $t > 0$ has an associated kernel $K_t$, and there exist positive constants $c_\alpha$, $C$ and $\alpha$ such that for all $t > 0$ and $x, y \in \mathcal{X}$ with $d(x, y) \geq C t^{1/\alpha}$,
\[
|K_t(x, y) - K_t(x, y)| \leq \frac{1}{\mu(B(x, d(x, y)))} \frac{t^{\alpha/m}}{|d(x, y)|^\beta},
\]
then for any $p \in (1, \infty)$ and $u \in A_p(\mathcal{X})$, $T$ is bounded on $L^p(\mathcal{X}, u)$. Moreover, Martell [12] proved that if $T$ is an $L^2(\mathcal{X})$ bounded linear operator, satisfies (i) and

(iii) there exists an approximation to the identity $\{\mathcal{D}_t\}_{t>0}$ such that the composite operator $TD_t$ with $t > 0$ has an associated kernel $K_t$ in the sense (1.3), and there exist positive constants $C, c_\alpha$ and $\beta$ such that for all $t > 0$ and $x, y \in \mathcal{X}$ with $d(x, y) \geq C t^{1/\alpha}$,
\[
|K_t(x, y) - K_t(x, y)| \leq \frac{1}{\mu(B(y, d(x, y)))} \frac{t^{\beta/m}}{|d(x, y)|^\beta},
\]
then for any $u \in A_1(\mathcal{X})$, $T$ is bounded from $L^1(\mathcal{X}, u)$ to $L^{1,\infty}(\mathcal{X}, u)$. Here and in what follows, $L^p(\mathcal{X}, u)$ means $L^p(\mathcal{X}, u) = L^p(x, d\mu)$, and $L^{1,\infty}(\mathcal{X}, u)$ means $L^{1,\infty}(\mathcal{X}, u) = L^{1,\infty}(x, d\mu)$. Recently, the authors in [10] considered the weighted estimates with general weights for the operator $T$, and proved that if $T$ is an $L^2(\mathcal{X})$ bounded operator which satisfies (ii) and (iii), then for any $p \in (1, \infty)$ and any weight $w$, $T$ is bounded from $L^p(\mathcal{X}, M^{2/p}w)$ to $L^p(\mathcal{X}, w)$ and also from $L^1(\mathcal{X}, M^p w)$ to $L^1(\mathcal{X}, w)$. Here and in what follows, $M$ denotes the Hardy–Littlewood maximal operator and for any $l \in \mathbb{N}$, denote by $M^l$ the $l$-time iterations of $M$. Moreover, for a positive number $\theta$, $[\theta]$ denotes the biggest integer no more than $\theta$.

Now let $b \in BMO(\mathcal{X})$. Define the commutator $T_b$ by
\[
T_b f(x) = b(x)Tf(x) - T(bf)(x),
\]
where $x \in \mathcal{X}$ and $f \in L^1_0(\mathcal{X})$. The maximal operator associated with the commutator $T_b$ is defined by
\[
T^*_b f(x) = \sup_{\epsilon > 0} \left| T_{b, \epsilon} f(x) \right|,
\]
here and in what follows, for any $\epsilon > 0$, $T_{b, \epsilon}$ is the truncated operator defined by
\[
T_{b, \epsilon} f(x) = \int_{d(x, y) > \epsilon} K(x, y) (b(x) - b(y)) f(y) d\mu(y).
\]
The commutator $T_b$ was first considered by Duong and Yan [4]. They showed that if $T$ is an $L^2(\mathcal{X})$ bounded linear operator and satisfies (i) and (ii), then $T_b$ is bounded on $L^p(\mathcal{X})$ for any $p \in (1, \infty)$. Using a general version of the sharp maximal operator introduced by Martell in [12], the authors in [10] proved that if $T$ is an $L^2(\mathcal{X})$ bounded linear operator and satisfies (ii) and (iii), then for any $p \in (1, \infty)$ and weight $w$,
\[
\int_{\mathcal{X}} |T_b f(x)|^p w(x) d\mu(x) \leq C \|b\|_{BMO(\mathcal{X})} \int_{\mathcal{X}} |f(x)|^p M^{2/p} w(x) d\mu(x),
\]
where $C$ is a positive constant depending only on $p$. Moreover, $T_b$ enjoys the weighted weak type endpoint estimate that
\[
\int_{\mathcal{X}} w(x) d\mu(x) \leq C \lambda^{-1} \int_{\mathcal{X}} |f(x)| M^4 w(x) d\mu(x).
\]
Our first purpose of this paper is to prove that the operator $T^*_b$ enjoys a weighted estimate with general weights which is analog with that of the commutator $T_b$. 

Theorem 1. Let $b \in \text{BMO}(\mathcal{X})$, $T$ be an $L^2(\mathcal{X})$ bounded linear operator with kernel $K$ as in (1.3) and $T^*_b$ the maximal operator defined by (1.5). Suppose that $T$ satisfies (ii) and (iii) and that the approximation to the identity $\{\tilde{D}_t\}_{t > 0}$ appeared in (ii) above also satisfies that for all $t > 0$ and $x, y \in \mathcal{X}$ with $d(x, y) \leq ct^{1/m}$,

$$|K^t(x, y)| \leq C \frac{1}{\mu(B(x, t^{1/m}))},$$

where $C$ is a positive constant independent of $t, x$ and $y$. Then for any $p \in (1, \infty)$, there exists a positive constant $C$ depending only on $p$ such that for any weight $w$ and $f \in L^\infty_0(\mathcal{X}),$

$$\int_{\mathcal{X}} |T^*_b f(x)|^p w(x) \, d\mu(x) \leq C \|b\|_{\text{BMO}(\mathcal{X})}^p \int_{\mathcal{X}} |f(x)|^p M_1^{3p/2 + 2} w(x) \, d\mu(x).$$

(1.7)

Although it is still unclear if there exists certain weighted endpoint estimate for $T^*_b$ with general weights, we have the following conclusion, which is new even when $u \equiv 1$.

Theorem 2. Let $b \in \text{BMO}(\mathcal{X}), u \in A_1(\mathcal{X})$ and $T$ be the same as in Theorem 1. Then there exists a positive constant $C$ depending only on $\|b\|_{\text{BMO}(\mathcal{X})}$ and the $A_1(\mathcal{X})$-constant of $u$ such that for any $\lambda > 0$ and $f \in L^\infty_0(\mathcal{X}),$

$$\int_{\{x \in \mathcal{X} : \ T^*_b f(x) > \lambda\}} u(x) \, d\mu(x) \leq C \int_{\mathcal{X}} \frac{|f(x)|}{\lambda} \log^2 \left(2 + \frac{|f(x)|}{\lambda}\right) u(x) \, d\mu(x).$$

Remark 1. It should be pointed out that the operator $T^*_b$ is not a linear operator, and it is not clear if the argument used in [10] to establish the estimate (1.6) also applies to the operator $T^*_b$. We prove Theorem 1 here by establishing a Cotlar type inequality, which shows that $T^*_b$ is controlled by $M(T_b) + M_b T + M_b$ with $M_b$ as in (2.1); see Theorem 3 in Section 3. However, to prove Theorem 2, this Cotlar inequality is not sufficient. We need to employ some inequalities established in [10] to establish certain weighted distribution inequality linking operators $M_b$ and $M^2$ (Lemma 3), and certain weighted distribution inequality linking operators $M(T_b)$, $M^2T$ and $M^3$ (see the estimates (3.11)) via Lemma 4.

Remark 2. As well known, the operator $T^*_b$ is more singular than the operator $T_b$. Thus, it is natural that the iteration time of the Hardy–Littlewood maximal on the right-hand side of (1.7) is one more than that on the right-hand side of (1.6). However, it is still unclear if the iteration time of the Hardy–Littlewood maximal on the right-hand side of (1.7) is optimal.

We now make some conventions. Throughout this paper, $C$ always denotes a positive constant that is independent of the main parameters involved but whose value may differ from line to line. Constants with subscripts, such as $C_1$, do not change in different occurrences. For a fixed $p \in [1, \infty)$, $p'$ denotes the conjugate exponent of $p$, namely, $p' = p/(p - 1)$. For any $f \in L^1_{\text{loc}}(\mathcal{X})$ and $x \in \mathcal{X}$, let $M^2 f$ be the sharp maximal function of Fefferman and Stein defined by

$$M^2 f(x) = \sup_{B \ni x} \frac{1}{\mu(B)} \int_B |f(y) - m_B(f)| \, d\mu(y),$$

(1.8)

where the supremum is taken over all balls containing $x$, and $m_B(f)$ is the mean value of $f$ on $B$, namely, $m_B(f) = \frac{1}{\mu(B)} \int_B f(y) \, d\mu(y)$. For a fixed $q \in (0, 1)$, any suitable function $h$ and $x \in \mathcal{X}$, let $M^q h(x) = [M^{2q}(|h|^q)(x)]^{1/q}$ and $M^q h(x) = [M^{q}(|h|^q)(x)]^{1/q}$.

Some Luxemburg norms are used in our argument. Let $\delta$ be a nonnegative number and $E$ a measurable set with $\mu(E) < \infty$. For any suitable function $f$, define $\|f\|_{L(\log L)^{\delta}, E}$ by

$$\|f\|_{L(\log L)^{\delta}, E} = \inf \left\{ \lambda > 0 : \frac{1}{\mu(E)} \int_E \left| \frac{|f(y)|}{\lambda} \log^\delta \left(2 + \frac{|f(y)|}{\lambda}\right) \right| \, d\mu(y) \leq 1 \right\}.$$

The maximal operator $M_{1(\log L)^{\delta}}$ is defined by

$$M_{1(\log L)^{\delta}} f(x) = \sup_{B \ni x} \|f\|_{L(\log L)^{\delta}, B},$$

where the supremum is taken over all balls $B$ containing $x$. Also we define the norm $\|f\|_{\exp L, E}$ by

$$\|f\|_{\exp L, E} = \inf \left\{ \lambda > 0 : \frac{1}{\mu(E)} \int_E \exp \left( \frac{|f(x)|}{\lambda} \right) \, d\mu(x) \leq 2 \right\}.$$

It is well known that the following generalization of the Hölder inequality
holds for any suitable functions \( f, h \) and measurable set \( E \) with \( \mu(E) < \infty \).

2. Some lemmas

This section is devoted to some lemmas which are used in the proofs of our theorems.

Let \( b \in \text{BMO}(\mathcal{X}) \). For any \( f \in L_0^{\infty}(\mathcal{X}) \) and \( x \in \mathcal{X} \), define the commutator \( M_b \) of the Hardy–Littlewood maximal operator with \( b \) by

\[
M_b f(x) = \sup_{B \ni x} \frac{1}{\mu(B)} \int_B |b(x) - b(y)| |f(y)| \, d\mu(y),
\]

where the supremum is taken over all balls \( B \ni x \). Let \( y \) be the constant as in (1.2). Then there exists an approximation of the identity \( \{S_k\}_{k \in \mathbb{Z}} \) of order \( y \) with bounded support on \( \mathcal{X} \). Namely, \( \{S_k\}_{k \in \mathbb{Z}} \) is a sequence of bounded linear integral operators on \( L^2(\mathcal{X}) \), and there exist positive constants \( C_0 \) and \( C \) such that for all \( k \in \mathbb{Z} \) and all \( x, x', y \) and \( y' \in \mathcal{X} \), the integral kernel of \( S_k \) is a measurable function from \( \mathcal{X} \times \mathcal{X} \) into \( \mathbb{C} \) satisfying

(i) \( S_k(x,y) = 0 \) if \( d(x,y) \geq C2^{-k} \) and \( 0 \leq S_k(x,y) \leq \frac{C_0}{x+y+2^{-k}(y)} \), where for any \( x \in \mathcal{X} \) and \( r > 0 \), \( V_r(x) = \mu(B(x,r)) \);
(ii) \( S_k(x,y) = S_k(y,x) \) for all \( x, y \in \mathcal{X} \);
(iii) \( |S_k(x,y) - S_k(x',y)| \leq \frac{C_0}{x+y+2^{k}(y)} \) for \( d(x,y) \leq \max(\tilde{C}/k,1/k)2^{1-k} \), where \( \kappa \) is the constant appearing in the quasi-triangle inequality satisfied by \( d \);
(iv) \( C_0 V_{2^{-k}}(x) S_k(x,x) > 1 \) for all \( x \in \mathcal{X} \) and \( k \in \mathbb{Z} \);
(v) \( \int_{\mathcal{X}} S_k(x,y) \, d\mu(y) = 1 = \int_{\mathcal{X}} S_k(x,y) \, d\mu(x) \);

see [7] (or [9]) for the details. Define the operator \( \tilde{M}_b \) by setting, for all \( x \in \mathcal{X} \),

\[
\tilde{M}_b f(x) = \sup_{k \in \mathbb{Z}} \int_{\mathcal{X}} S_k(x,y) |b(x) - b(y)| |f(y)| \, d\mu(y).
\]

It was proved in [9] that there exists certain constant \( C \geq 1 \) such that for all \( x \in \mathcal{X} \) and \( f \in L_0^{\infty}(\mathcal{X}) \),

\[
C^{-1} \tilde{M}_b f(x) \leq M_b f(x) \leq C \tilde{M}_b f(x).
\]

Lemma 1. (See [9].) Let \( b \in \text{BMO}(\mathcal{X}) \) and \( M_b \) be as in (2.1). Then,

(i) for any \( q \in (0, 1) \), there exists a positive constant \( C \) depending only on \( q \) such that for any \( f \in L_0^{\infty}(\mathcal{X}) \) and \( x \in \mathcal{X} \),

\[
M_b^q(\tilde{M}_b f(x)) \leq CM^2 f(x);
\]
(ii) for any \( p \in (1, \infty) \) and \( \delta > 0 \), there exists a positive constant \( C \) depending only on \( p \) and \( \delta \) such that for any weight \( w \) and \( f \in L_0^{\infty}(\mathcal{X}) \),

\[
\int_{\mathcal{X}} (M_b f(x))^p w(x) \, d\mu(x) \leq C\|b\|_{\text{BMO}(\mathcal{X})}^p \int_{\mathcal{X}} |f(x)|^p M_{(\log L)^{p+\delta}} w(x) \, d\mu(x);
\]
(iii) there exists a positive constant \( C \) depending on \( \|b\|_{\text{BMO}(\mathcal{X})} \) such that for any weight \( w \) and \( f \in L_0^{\infty}(\mathcal{X}) \),

\[
w(\{x \in \mathcal{X} : M_b f(x) > \lambda\}) \leq C \int_{\mathcal{X}} \frac{|f(x)|}{\lambda} \log \left( 2 + \frac{|f(x)|}{\lambda} \right) M_{4} w(x) \, d\mu(x).
\]

Recall that a nonnegative and locally integrable function \( u \) is said to belong to \( A_\infty(\mathcal{X}) \) if there exist two positive constants \( C_{A_\infty}(u) \) and \( \delta_{A_\infty}(u) \) such that for any ball \( B \) and measurable set \( E \subset B \),

\[
\frac{u(E)}{u(B)} \leq C_{A_\infty}(u) \left( \frac{\mu(E)}{\mu(B)} \right)^{\delta_{A_\infty}(u)},
\]

here and in what follows, \( u(E) = \int_E u(x) \, d\mu(x) \).

Recall also that a function \( \Phi \) on \([0, \infty)\) is said to satisfy the doubling condition, if there exists a positive constant \( C \) such that for any \( t > 0 \), \( \Phi(2t) \leq C \Phi(t) \).
Lemma 2. Let $\Phi$ be an increasing function on $[0, \infty)$, which satisfies the doubling condition. Then for any $u \in A_\infty(\mathcal{X}')$, there exists a positive constant $C$, depending only on $\|b\|_{\text{BMO}(\mathcal{X})}$, $C_{A_\infty}$ and $\delta_{A_\infty}$, such that for any $f \in L^1_{\text{loc}}(\mathcal{X})$,

$$\sup_{\lambda > 0} \Phi(\lambda)u(\{x \in \mathcal{X}': M_\sigma f(x) > \lambda\}) \leq C \sup_{\lambda > 0} \Phi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\}),$$

provided that $\Phi(\lambda)$ depends only on $C$,

$$\sup_{0 < \lambda < R} \Phi(\lambda)u(\{x \in \mathcal{X}': M_\sigma (M_b f)(x) > \lambda\}) < \infty.$$ Proof. Let $M^2$ be as in (1.8). We claim that if $u \in A_\infty(\mathcal{X}')$ and $\Psi$ is an increasing function on $[0, \infty)$ which satisfies the doubling condition, then there exists a positive constant $C$ depending only on $C_{A_\infty}$ such that for any $f \in L^1_{\text{loc}}(\mathcal{X})$,

$$\sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

$$\leq C \left\{ \begin{array}{ll}
sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\}), & \text{if } \mu(\mathcal{X}') = \infty; \\
sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\}) + \Psi(m f)(u)(\mathcal{X}), & \text{if } \mu(\mathcal{X}') < \infty,
\end{array} \right. \leq C \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

provided that for any $R > 0$,

$$\sup_{0 < \lambda < R} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\}) < \infty.$$ (2.3)

In fact, from the proof of Proposition 3.4 in [13], we deduce that there exists a positive constant $C$ such that for any $\gamma \in (0, 1)$, $u \in A_\infty(\mathcal{X}')$ and $\lambda > 0$,

$$u(\{x \in \mathcal{X}': M f(x) > \lambda\}) \leq C \mu(\mathcal{X}') \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\}) + C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

provided that $\lambda > C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$

If $\mu(\mathcal{X}') = \infty$, then for each fixed $u \in A_\infty(\mathcal{X}')$ and $R > 0$, it follows from estimate (2.5) and the doubling condition that

$$\sup_{0 < \lambda < R} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\}) \leq C \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\})$$

$$+ C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

provided that $\mu(\mathcal{X}') < \infty$ or $\mu(\mathcal{X}') = \infty$ and $\lambda > 2m f(x)f$.

On the other hand, if $\mu(\mathcal{X}') < \infty$, the estimate (2.5) tells us that for any $R > 2m f(x)$,

$$\sup_{2m f(x) < \lambda < R} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\}) \leq C \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\})$$

$$+ C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

$$\leq C \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\})$$

$$+ C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

which in turn implies that for any $R > 2m f(x)$,

$$\sup_{0 < \lambda < R} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\}) \leq \sup_{0 < \lambda < 2m f(x)} \Psi(\lambda)u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

$$+ C_{\gamma} \sup_{\lambda > 0} \Psi(\lambda)u(\{x \in \mathcal{X}': M^2 f(x) > \lambda\})$$

$$+ C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} u(\{x \in \mathcal{X}': M f(x) > \lambda\})$$

Choose $\gamma \in (0, 1)$ such that $C_{\gamma} \frac{\delta_{A_\infty}}{\gamma} < 1/2$. Our estimates above imply that if (2.4) is true, then
Thus, when inequality (1.9) that positive constant $C$ such that for any $f$

This in turn gives our desired conclusion (2.3).

We now conclude the proof of Lemma 2. Let $\sigma \in (0, 1)$ being as in the lemma. Observe that if $\Phi$ is increasing and satisfies the doubling condition, then $\Psi(t) = \Phi(t^{1/\sigma})$ is also increasing and satisfies the doubling condition. Our hypothesis on $\Phi$, via the Lebesgue differential theorem, the estimate (2.2), (i) of Lemma 1 and the claim (2.3), implies that

$$\sup_{\lambda>0} \Phi(\lambda)u(\{x \in \mathcal{X} : Mf(x) > \lambda\})$$

$$\leq C \sup_{\lambda>0} \Phi(\lambda)u(\{x \in \mathcal{X} : Mf(x) > \lambda\}), \quad \text{if } \mu(\mathcal{X}) = \infty;$$

$$\sup_{\lambda>0} \Phi(\lambda)u(\{x \in \mathcal{X} : Mf(x) > \lambda\}) + \sup_{\lambda>0} \Psi(\lambda)u(\mathcal{X}) = \mu(\mathcal{X}) < \infty \text{ and } R > 2m_{\mathcal{X}}(f).$$

(2.6)

Thus, when $\mu(\mathcal{X}) = \infty$, we already obtain the desired estimate. For the case of $\mu(\mathcal{X}) < \infty$, write

$$m_{\mathcal{X}}((M_b f)^{\sigma}) \leq \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} |b(y) - m_{\mathcal{X}}(b)|^{\sigma} (Mf(y))^\sigma d\mu(y)$$

$$+ \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} (M((M_b f)^{\sigma}))^{1/\sigma} d\mu(y) = I + II.$$

An application of the Hölder inequality leads to that for all $x \in \mathcal{X}$,

$$I^{1/\sigma} \leq \left[ \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} |b(y) - m_{\mathcal{X}}(b)|^{\sigma/(1-\sigma)} d\mu(y) \right]^{(1-\sigma)/\sigma} \left[ \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} Mf(y) d\mu(y) \right] \leq CM^2 f(x),$$

where $C$ is a positive constant independent of $f$ and $x$, but depending on $\|b\|_{\text{BMO}(\mathcal{X})}$. This further implies that

$$I^{1/\sigma} \leq C \inf_{x \in \mathcal{X}} M^2 f(x).$$

Recall that $M$ is bounded from $L^1(\mathcal{X})$ to $L^{1,\infty}(\mathcal{X})$. It follows from the Kolmogorov inequality (see [6, p. 485]) and the inequality (1.9) that

$$\|b - M_{\mathcal{X}}(b)\|_{\text{exp},\mathcal{X}} \leq C\|b\|_{\text{BMO}(\mathcal{X})},$$

The John–Nirenberg inequality states that

$$\|b - M_{\mathcal{X}}(b)\|_{\text{exp},\mathcal{X}} \leq C\|b\|_{\text{BMO}(\mathcal{X})},$$

and a well-known estimate (see [13]) tells us that

$$\|f\|_{\text{Log},\mathcal{X}} \leq C \inf_{x \in \mathcal{X}} M^2 f(x).$$

Combining the estimates for terms $I$ and $II$ leads to that

$$\Phi((M_b f)^{\sigma})^{1/\sigma} u(\mathcal{X}) \leq C \Phi \left( \inf_{x \in \mathcal{X}} M^2 f(x) \right) u(\mathcal{X}) \leq C \sup_{\lambda>0} \Phi(\lambda)u(\{x \in \mathcal{X} : M^2 f(x) > \lambda\}).$$

This combined with (2.6) yields the desired estimate when $\mu(\mathcal{X}) < \infty$, and hence, finishes the proof of Lemma 2. \qed

In the following, for any nonnegative integer $k$, we set $\Phi_k(t) = t \log^{-k}(2 + t^{-1})$ for $t \in [0, \infty)$. It is easy to verify that $\Phi_k$ is increasing and satisfies the doubling condition.

**Lemma 3.** Let $b \in \text{BMO}(\mathcal{X})$, $M_b$ be the commutator defined by (2.1) and $u \in A_1(\mathcal{X})$. Then for any positive integer $k$, there exists a positive constant $C$ such that for any $f \in L^{1,\infty}(\mathcal{X}, u) \cap L^{p_0}(\mathcal{X}, u)$ with certain $p_0 \in (1, \infty)$,

$$\sup_{\lambda>0} \Phi_k(\lambda)u(\{x \in \mathcal{X} : M_b f(x) > \lambda\}) \leq C \sup_{\lambda>0} \Phi_k(\lambda)u(\{x \in \mathcal{X} : M^2 f(x) > \lambda\}).$$

(2.7)
Proof. Let \( f \in L^{1,\infty}(\mathcal{X}, u) \cap L^{p_0}(\mathcal{X}, u) \). For all positive integers \( N \) and \( x \in \mathcal{X} \), set

\[
    f_N(x) = f(x)\chi_{\{|f(x)| \leq N\}}(x).
\]

We claim that for any positive integer \( N \),

\[
    \sup_{\lambda > 0} \Phi_k(\lambda)u\left( \{ x \in \mathcal{X} : M_b f_N(x) > \lambda \} \right) \lesssim C \sup_{\lambda > 0} \Phi_k(\lambda)u\left( \{ x \in \mathcal{X} : M^2 f_N(x) > \lambda \} \right)
\]

(2.8)

with \( C \) independent of \( N \). By the Lebesgue dominant convergence theorem, \( M_b f_N \uparrow M_b f \) and \( M^2 f_N \uparrow M^2 f \) pointwise. Here and in what follows, the symbol \( g_N \uparrow g \) means that \( g_N \) increasingly converges to \( g \) pointwise as \( N \to \infty \). So for any \( \lambda > 0 \), we have

\[
    u\left( \{ x \in \mathcal{X} : M_b f_N(x) > \lambda \} \right) \uparrow u\left( \{ x \in \mathcal{X} : M_b f(x) > \lambda \} \right)
\]

and

\[
    u\left( \{ x \in \mathcal{X} : M^2 f_N(x) > \lambda \} \right) \uparrow u\left( \{ x \in \mathcal{X} : M^2 f(x) > \lambda \} \right).
\]

If we can prove (2.8), the estimate (2.7) then follows from (2.8) by taking \( N \to \infty \).

We now prove (2.8). By Lemma 2, it suffices to prove that for any \( \sigma \in (0, 1) \) and positive integer \( N \),

\[
    \sup_{\lambda > 0} \Phi_k(\lambda)u\left( \{ x \in \mathcal{X} : M_\sigma (M_b f_N(x)) > \lambda \} \right) < \infty.
\]

(2.9)

It follows from (iii) of Lemma 1 that for all fixed \( \tau > 0 \),

\[
    u\left( \{ x \in \mathcal{X} : M_b f_N(x) > \tau \} \right) \leq C \int_{\mathcal{X}} \left[ \frac{|f_N(x)|}{\tau} \log \left( 2 + \frac{|f_N(x)|}{\tau} \right) \right] u(x) d\mu(x)
\]

\[
= C \int_0^{N/\tau} u\left( \{ x \in \mathcal{X} : |f_N(x)| > s \tau \} \right) d(s \log(2 + s))
\]

\[
+ C \int_{N/\tau}^1 u\left( \{ x \in \mathcal{X} : |f_N(x)| > s \tau \} \right) d(s \log(2 + s))
\]

\[
\leq CN^{-1} u\left( \{ x \in \mathcal{X} : |f(x)| > N^{-1} \} \right) \tau^{-1} \log(2 + \tau^{-1})
\]

\[
+ C \tau^{-1} \int_{1/(N\tau)}^1 s^{-1} d(s \log(2 + s))
\]

\[
\leq CN^3 \tau^{-1} \log(2 + \tau^{-1}) \| f \|_{L^{1,\infty}(\mathcal{X}, u)}.
\]

Notice that there exist two positive constants \( C \) and \( C_\sigma \) such that for any \( \lambda > 0 \),

\[
    u\left( \{ x \in \mathcal{X} : M_\sigma h(x) > \lambda \} \right) \leq C_\lambda^{-1} \sup_{\tau \geq C_\sigma \lambda} \tau u\left( \{ x \in \mathcal{X} : |h(x)| > \tau \} \right).
\]

In fact, if \( (\mathcal{X}, d, \mu) \) is the Euclidean space, this inequality was proved in [8], and the same idea also works for the space of homogeneous type. We thus have that

\[
    u\left( \{ x \in \mathcal{X} : M_\sigma (M_b f_N(x)) > \lambda \} \right) \leq CN^3 \lambda^{-1} \sup_{\tau \geq C_\sigma \lambda} \tau u\left( \{ x \in \mathcal{X} : |h(x)| > \tau \} \right)
\]

\[
\leq CN^3 \lambda^{-1} \| f \|_{L^{1,\infty}(\mathcal{X}, u)}.
\]

This establishes (2.9) and hence, finishes the proof of Lemma 3. \( \square \)

Let \( k \) be a positive integer and \( \bar{D}_{\lambda} \) the approximation to the identity as in Definition 1. Define the sharp maximal operator \( M_{\bar{D}_{\lambda}}^k \) by

\[
    M_{\bar{D}_{\lambda}}^k f(x) = \sup_{B \ni x} \| f - \bar{D}_{\lambda} f \|_{L^k(\log L)^k B},
\]

where the supremum is taken over all balls \( B \ni x \), \( r_B \) is the radius of \( B \) and \( r_B = \frac{m}{2} \). For the case that \( k = 0 \), this operator was introduced by Martell [12]; for \( k \in \mathbb{N} \), this operator was introduced in [10]. It was proved by Duong and Yan [5] that
such sharp maximal operators when \( k = 0 \) play an important role in the theory of some new BMO-type spaces; see also [3]. In what follows, let \( \mathbb{Z}_+ = \mathbb{N} \cup \{0\} \).

**Lemma 4.** Let \( k, l \in \mathbb{Z}_+ \), \( \Phi_k(t) = t \log^k (2 + t^{-1}) \) and \( u \in A_1(\mathcal{X}) \). Then, there exists a positive constant \( C \) such that

\[
\sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} f(x) > \lambda \} \right) \leq C \left\{ \sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} D_{L}^{k} f(x) > \lambda \} \right) \right. \]

\[
\quad + \left. \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : \| f \|_{L}^{k} g(x) > \lambda \} \right) \right\} u(\mathcal{X}),
\]

provided that \( f \in L^{p_0}(\mathcal{X}), u \cap L^{p_1}(\mathcal{X}) \) with \( p_0, p_1 \in (1, \infty) \) and

\[
\sup_{\lambda > 0} \Phi_{l}(\lambda) u \left( \{ x \in \mathcal{X} : |f(x)| > \lambda \} \right) < \infty.
\]

**Proof.** By [10, Theorem 2.2], we know that there exists a positive constant \( C \) such that for any \( h \in L^{p_0}(\mathcal{X}) \) with \( p_0 \in (1, \infty) \),

\[
\sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} h(x) > \lambda \} \right) \leq C \left\{ \sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} D_{L}^{k} h(x) > \lambda \} \right) \right. \]

\[
\quad + \left. \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : \| h \|_{L}^{k} g(x) > \lambda \} \right) \right\} u(\mathcal{X}),
\]

provided that

\[
\sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} h(x) > \lambda \} \right) < \infty.
\]

Now let \( f \in L^{p_0}(\mathcal{X}), u \cap L^{p_1}(\mathcal{X}) \). For all positive integers \( N \) and \( x \in \mathcal{X} \), set

\[
\tilde{f}_N(x) = f(x) \chi_{\{ x \in \mathcal{X} : |f(x)| \leq N \}}.
\]

We claim that

\[
\sup_{\lambda > 0} \Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} \tilde{f}_N(x) > \lambda \} \right) < \infty.
\]

To prove this, for all positive integers \( N, \lambda > 0 \) and \( x \in \mathcal{X} \), set

\[
\tilde{f}_N^1(x) = \tilde{f}_N(x) \chi_{\{ x \in \mathcal{X} : |\tilde{f}_N(x)| > \lambda/2 \}}(x)
\]

and \( \tilde{f}_N^2(x) = \tilde{f}_N(x) \chi_{\{ x \in \mathcal{X} : |\tilde{f}_N(x)| \leq \lambda/2 \}}(x) \). Recall that the operator \( M \) satisfies the weighted weak type endpoint estimate that for any \( \lambda > 0 \),

\[
u \left( \{ x \in \mathcal{X} : M^{k+1} f(x) > \lambda \} \right) \leq C \int_{\mathcal{X}} \left| \frac{f(x)}{\lambda} \right| \log^k \left( 2 + \frac{|f(x)|}{\lambda} \right) u(x) d\mu(x),
\]

see [10, Lemma 2.2]. An argument similar to that used in the proof of Lemma 3 gives us that

\[
\Phi_{k+l+1}(\lambda) u \left( \{ x \in \mathcal{X} : M^{k+1} \tilde{f}_N^1(x) > \lambda/2 \} \right)
\]

\[
\leq C \Phi_{k+l+1}(\lambda) \int_{\mathcal{X}} \left| \frac{\tilde{f}_N^1(x)}{\lambda} \right| \log^k \left( 2 + \frac{|\tilde{f}_N^1(x)|}{\lambda} \right) u(x) d\mu(x)
\]

\[
\leq C \Phi_{k+l+1}(\lambda) \int_0^{1/2} \nu \left( \{ x \in \mathcal{X} : |\tilde{f}_N^1(x)| > \lambda t \} \right) d(t \log^k (2 + t))
\]

\[
+ C \Phi_{k+l+1}(\lambda) \int_{1/2}^{N/\lambda} \nu \left( \{ x \in \mathcal{X} : |\tilde{f}_N^1(x)| > \lambda t \} \right) d(t \log^k (2 + t))
\]
On the other hand, by (2.12) again, we have that for any

\[ C \Phi_{k+1}(\lambda)u\left( \{ x \in \mathcal{X} : |\tilde{f}_N(x)| > \lambda/2 \} \right) \]

\[ \leq C \left( \sup_{x \in \mathcal{X}} |\tilde{f}_N(x)| \right) d\left( t \log^k (2 + t) \right) \]

\[ + C \sup_{x \in \mathcal{X}} |\tilde{f}_N(x)| \left( 2 + \log^k (2 + t) \right) \]

\[ \leq C \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + t) \right) \]

\[ \leq C \log^{k+1} N \sup_{x \in \mathcal{X}} |\tilde{f}_N(x)| \]

This, together with the trivial estimate

\[ u\left( \{ x \in \mathcal{X} : M^{k+1} \tilde{f}_N(x) > \lambda \} \right) \leq u\left( \{ x \in \mathcal{X} : M^{k+1} \tilde{f}_N(x) > \lambda/2 \} \right). \]

leads to (2.11).

Now applying (2.10) and (2.11), we have that for any positive integer N,

\[ \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + t) \right) \]

\[ \leq C \left( \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + t) \right) \right) \]

(2.13)

On the other hand, by (2.12) again, we have that for any \( \lambda > 0 \),

\[ \Phi_{k+1}(\lambda)u\left( \{ x \in \mathcal{X} : M^{k+1} (\tilde{f}_N - f) > \lambda \} \right) \]

\[ \leq C \left( \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + |\tilde{f}_N(x) - f(x)|) \right) \right) \]

\[ \leq C \int_{\mathcal{X}} \left| f(x) \right| \left( 2 + \log^k (2 + |\tilde{f}_N(x) - f(x)|) \right) \]

\[ \leq C \int_{\mathcal{X}} \left| f(x) \right| \log^k (2 + |\tilde{f}_N(x) - f(x)|) dx \]

which in turn implies that

\[ \lim_{N \to \infty} \sup_{\lambda > 0} \Phi_{k+1}(\lambda)u\left( \{ x \in \mathcal{X} : M^{k+1} (\tilde{f}_N - f) > \lambda \} \right) = 0. \]

Recall that for any ball \( B \) and locally integrable function \( f \in L^p(\mathcal{X}) \) with \( p \in [1, \infty) \),

\[ \frac{1}{\mu(B)} \int_B |\tilde{D}_x (h)(y)| \mu(y) \leq C \inf_{x \in B} Mh(x). \]

see Lemma 3.5 in [12]. Therefore, as \( N \to \infty \), we have

\[ \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + t) \right) \]

\[ \leq C \left( \sup_{x \in \mathcal{X}} \left( |\tilde{f}_N(x)| + \log^k (2 + t) \right) \right) \]

(2.13)

This along with (2.13) then gives the desired conclusion. \( \square \)
3. Proofs of Theorems 1 and 2

We begin with the following Cotlar type inequality, which is new even for the Euclidean space and has independent interest.

**Theorem 3.** Let $b \in \text{BMO}(\mathcal{X})$, $T_b$ and $T_b^+$ be the operators defined by (1.4) and (1.5), respectively. Then there exists a positive constant $C$ such that for any $f \in L_0^\infty(\mathcal{X})$ and almost every $x \in \mathcal{X}$,

$$T_b^+ f(x) \leq C(M(T_b f)(x) + M_b(T f)(x) + M_b f(x)).$$

**Proof.** For any fixed $f \in L_0^\infty(\mathcal{X})$, by Theorem 1.5 in [10], we know that $T_b f$ is finite almost everywhere. Let $x$ be a point in $\mathcal{X}$ such that $|T_b f(x)| < \infty$. For any $\epsilon > 0$ and $x \in \mathcal{X}$, write

$$T_{b, \epsilon} f(x) = \tilde{D}_{e^n}(T_b f)(x) - (\tilde{D}_{e^n} T_b f)(x) + (\tilde{D}_{e^n} T_b)(f) - \tilde{D}_{e^n}(T_b f)(x),$$

where $(\tilde{D}_{e^n} T_b)$ denotes the commutator generated by $b$ and the composite operator $\tilde{D}_{e^n}$. It is obvious that for all $x \in \mathcal{X}$,

$$|T_{b, \epsilon} f(x)| \leq CM(T_b f)(x).$$

Let $K_{\epsilon}(x, y) = K(x, y)\chi_{\mathcal{X} \times \mathcal{X} : d(x, y) > \epsilon}(x, y)$. As in [2, p. 249], a straightforward computation leads to that

$$|(\tilde{D}_{e^n} T_b - T_{b, \epsilon}) f(x)| \leq \int_{d(x, y) < \epsilon} |b(x) - b(y)| |K_{e^n}(x, y) - K_{\epsilon}(x, y)| |f(y)| d\mu(y)$$

$$+ \int_{d(x, y) \leq \epsilon} |b(x) - b(y)| |K_{e^n}(x, y) - K_{\epsilon}(x, y)| |f(y)| d\mu(y)$$

$$\leq C \sum_{k=0}^{\infty} 2^{-k\mu} \int_{d(x, y) \leq 2^{k+1}\epsilon} |b(x) - b(y)| |f(y)| d\mu(y)$$

$$+ C \frac{1}{\mu(B(x, \epsilon))} \int_{d(x, y) \leq \epsilon} |b(x) - b(y)| |f(y)| d\mu(y) \leq CM_b f(x).$$

Let

$$\tilde{D}_{t, b} f(x) = \int_{\mathcal{X}} a_t(x, y) (b(x) - b(y)) f(y) d\mu(y).$$

Notice that for any $t > 0$,

$$|\tilde{D}_{t, b} f(x)| \leq CM_b f(x)$$

with a positive constant $C$ independent of $t$, and that

$$(\tilde{D}_{e^n} T_b)(f)(x) - \tilde{D}_{e^n}(T_b f)(x) = \tilde{D}_{e^n}(T_b f)(x).$$

We then obtain that for any $\epsilon > 0$,

$$|T_{b, \epsilon} f(x)| \leq C (M(T_b f)(x) + M_b(T f)(x) + M_b f(x))$$

with a positive constant $C$ independent of $\epsilon$, $f$ and $x$. This finishes the proof of Theorem 3. \qed

**Proof of Theorem 1.** By the homogeneity, we may assume that $\|b\|_{\text{BMO}(\mathcal{X})} = 1$. Repeating the proof of Theorem 1.3 for the case $k = 1$ in [10], with the estimate

$$\int_{\mathcal{X}} (Mf(x))^{p'} (M|f|^{p+1} \omega(x))^{1-p'} d\mu(x) \leq C \int_{\mathcal{X}} |f(x)|^{p'} \omega(x)^{1-p'} d\mu(x)$$

replaced by the refined inequality

$$\int_{\mathcal{X}} (M_k f(x))^{p'} (M_{1+\log L})^{p-1} \omega(x)^{1-p'} d\mu(x) \leq C_{k, p} \int_{\mathcal{X}} |f(x)|^{p'} w^{1-p'}(x) d\mu(x)$$

for $p \in (1, \infty)$ and $\delta > 0$, we obtain that for any $p \in (1, \infty)$, weight $w$ and $f \in L_0^\infty(\mathcal{X})$, we have

$$\int_{\mathcal{X}} (M f(x))^{p'} (M|f|^{p+1} \omega(x))^{1-p'} d\mu(x) \leq C \int_{\mathcal{X}} |f(x)|^{p'} \omega(x)^{1-p'} d\mu(x).$$
\[
\int_X |Tf(x)|^p w(x) d\mu(x) \leq C \int_X |f(x)|^p M_{L(\log L)^{1+\delta}}(x) w(x) d\mu(x).
\] (3.1)

Choose \( \delta \in (0, 1/2) \). It then follows from (3.1) and (ii) of Lemma 1 that for any \( p \in (1, \infty) \), weight \( w \) and \( f \in L_0^\infty(X) \),

\[
\int_X (M_b(Tf)(x))^p w(x) d\mu(x) \leq C \int_X |f(x)|^p M_{L(\log L)^{1+\delta}} w(x) d\mu(x)
\]

\[
\leq C \int_X |f(x)|^p M_{[3p]^{ \frac{1}{2}}} w(x) d\mu(x).
\] (3.2)

On the other hand, by (1.6), we have that

\[
\int_X (M(T_b f)(x))^p w(x) d\mu(x) \leq C \int_X |f(x)|^p M_{[3p]^{ \frac{1}{2}}} w(x) d\mu(x).
\] (3.3)

The estimates (3.2) and (3.3), and (ii) of Lemma 1, via Theorem 3, lead to the desired conclusion. \( \square \)

**Proof of Theorem 2.** Obviously, it suffices to prove that for any \( f \in L_0^\infty(X) \),

\[
u\{x \in X : T^p f(x) > 1\} \leq C \int_X |f(x)| \log^2 (2 + |f(x)|) u(x) d\mu(x)
\]

with \( C \) depending only on \( u \) and \( \|b\|_{\text{BMO}(X)} \). Notice that by (iii) of Lemma 1,

\[
u\{x \in X : M_b f(x) > 1\} \leq C \int_X |f(x)| \log (2 + |f(x)|) u(x) d\mu(x).
\]

Thus by Theorem 3, we see that the proof of Theorem 2 can be reduced to proving that

\[
u\{x \in X : M(T_b f)(x) > 1\} \leq C \int_X |f(x)| \log^2 (2 + |f(x)|) u(x) d\mu(x).
\] (3.4)

and that

\[
u\{x \in X : M(T_b f)(x) > 1\} \leq C \int_X |f(x)| \log (2 + |f(x)|) u(x) d\mu(x).
\] (3.5)

Now we claim that for any \( u \in A_1(X) \) and \( f \in L_0^\infty(X) \),

\[
\sup_{\lambda > 0} \Phi_2(\lambda) u\{x \in X : M^2(Tf)(x) > \lambda\} \leq C \int_X |f(x)| \log^2 (2 + |f(x)|) u(x) d\mu(x).
\] (3.6)

To prove this, we recall that for any \( f \in L_0^\infty(X) \) and \( u \in A_1(X) \),

\[
Tf \in L^{1,\infty}(X, u) \cap \bigcap_{1 < p < \infty} L^p(X, u).
\]

On the other hand, we know that \( T \) enjoys the sharp function estimate that for any \( x \in X \),

\[
M_{D_{L},L(\log L)}^2(Tf)(x) \leq CM^3 f(x);
\]

see [10, (3.1)]. So by Lemma 4 with \( k = 1 \) and \( l = 0 \), we obtain

\[
\sup_{\lambda > 0} \Phi_2(\lambda) u\{x \in X : M^2(Tf)(x) > \lambda\}
\]

\[
\leq C \begin{cases} 
\sup_{\lambda > 0} \Phi_2(\lambda) u\{x \in X : M^2 f(x) > \lambda\}, & \text{if } \mu(X) = \infty; \\
\sup_{\lambda > 0} \Phi_2(\lambda) u\{x \in X : M^2 f(x) > \lambda\} + \Phi_2(\|Tf\|_{L(\log L)(X)}) u(X), & \text{if } \mu(X) < \infty.
\end{cases}
\] (3.7)

For the case of \( \mu(X) < \infty \), we have by Lemma 3.1 in [10] that

\[
\|Tf\|_{L(\log L)(X)} \leq C \inf_{x \in X} M^3 f(x).
\] (3.8)
Combining the estimates (3.7) and (3.8) then yields
\[
\sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M^2(Tf)(x) > \lambda\}\right) \leq C \sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M^3f(x) > \lambda\}\right)
\]
\[
\leq C \sup_{\lambda > 0} \Phi_2(\lambda) \int_{\mathcal{X}} \frac{|f(x)|}{\lambda} \log^2(2 + \frac{|f(x)|}{\lambda}) u(x) \, d\mu(x)
\]
\[
\leq C \int_{\mathcal{X}} |f(x)| \log^2(2 + |f(x)|) u(x) \, d\mu(x),
\]
where in the penultimate inequality, we again invoked (2.12). Thus, (3.6) holds.

The estimate (3.4) is an easy consequence of (3.6) and Lemma 3 with \(k = 2\).

To prove (3.5), we observe that for any \(\lambda > 0\),
\[
(\mu(\mathcal{X}))^{-1} \|T_b f\|_{L^1(\mathcal{X})} \leq C \inf_{\lambda < \infty} \left( M^2(Tf)(x) + M^3f(x) \right),
\]
and so
\[
\Phi_2(\|T_b f\|_{L^1(\mathcal{X})} (\mu(\mathcal{X}))^{-1}) u(\mathcal{X}) \leq C \sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M^2(Tf)(x) + M^3f(x) > \lambda\}\right).
\]

The estimates (3.9) and (3.10), along with the sharp function estimate that
\[
M^2(T_b f)(x) \leq C \|b\|_{\text{BMO}(\mathcal{X})} (M^2(Tf)(x) + M^3f(x))
\]
(see [10, Lemma 5.2]), then give us that
\[
\sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M(T_b f)(x) > \lambda\}\right) \leq C \sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M^2(Tf)(x) > \lambda\}\right)
\]
\[
+ \sup_{\lambda > 0} \Phi_2(\lambda)u\left(\{x \in \mathcal{X} : M^3f(x) > \lambda\}\right).
\]

This via (3.6) and (2.12) leads to the estimate (3.5) and then finishes the proof of Theorem 2. \(\Box\)

Acknowledgment

The authors wish to express their deep thanks to the referee for his/her several valuable suggestions which improve the presentation of this article.

References