Michigan Math. J. 49 (2001)

Sharp Weighted Endpoint Estimates for Commutators of Singular Integrals

Carlos Pérez & Gladis Pradolini

1. Introduction and Statements of the Main Result

The main purpose of this paper is to improve the main result in [P2] by means of a direct proof that avoids the classical good-λ technique considered there. The good-λ method, introduced by Burkholder and Gundy in [BG], is a powerful tool but has the disadvantage that it is essentially adapted to measures satisfying the A_∞ condition, such as the Lebesgue measure. The approach we consider here is more related to the classical argument of Calderón and Zygmund for proving that singular integral operators satisfy the weak-type $(1,1)$-property, an approach whose advantage is that it allows us to consider more general measure. The method, however, must be different because commutators of singular integral operators with BMO functions are not of weak-type $(1,1)$, as shown in [P2].

Let b be a locally integrable function on $\mathbf{R}^n$, usually called the *symbol*, and let T be a Calderón–Zygmund singular integral operator (see [C] or [J]). Consider the commutator operator $[b,T]$ defined for, say, smooth functions f by

$$[b,T]f = bT(f) - T(bf). \tag{1}$$

A now classical result of Coifman, Rochberg, and Weiss [CRW] states that $[b,T]$ is a bounded operator on $L^p(\mathbf{R}^n)$, $1 < p < \infty$, when b is a BMO function. In fact, BMO is also a necessary condition for the commutator $[b,R]$ to be bounded on $L^p(\mathbf{R}^n)$, where $R = (R^1,\dots,R^n)$ is the vector-valued Riesz transform. We will always assume that $b \in \mathrm{BMO}(\mathbf{R}^n)$ unless otherwise noted.

None of the different proofs of this result follows the usual scheme of the classical Calderón–Zygmund theory of singular integral operators T. Indeed, the key result in this theory is that any of these operators satisfies the weak-type $(1,1)$-property, which is derived from the assumption that T is bounded on $L^2(\mathbf{R}^n)$ combined with a mild regularity of the kernel. Once the weak-type $(1,1)$-inequality is obtained, interpolation and duality yield the boundedness of the operator on $L^p(\mathbf{R}^n)$ for all $1 < p < \infty$. However, simple examples show that commutators (with BMO symbols) fail to be of weak-type $(1,1)$, as found in [P2]. To remedy

Received May 3, 2000. Revision received October 11, 2000.

This research was carried out during a stay at the Universidad Autónoma de Madrid. Research of the first author was partially supported by DGESIC grant PB98-0106, Spain. Research of the second author was supported by Universidad Nacional del Litoral (FOMEC) and CONICET, República Argentina.

the situation, it is shown there that commutators satisfy a "$L(\log L)$"-type estimate. To be precise, we have the following result.

THEOREM 1.1 [P2]. *Let T be any Calderón–Zygmund singular integral operator. Then there exists a positive constant C depending upon the BMO norm of b such that, for all functions f and all $\lambda > 0$,*

$$|\{y \in \mathbf{R}^n : |[b,T]f(y)| > \lambda\}| \le C \int_{\mathbf{R}^n} \frac{|f(y)|}{\lambda}\left(1 + \log^+\left(\frac{|f(y)|}{\lambda}\right)\right) dy. \quad (2)$$

The proof of this estimate is based on showing that there is an intimate relationship between commutators and iterations of the Hardy–Littlewood maximal function (in this case, $M^2 = M \circ M$) via the good-λ technique of Burkholder and Gundy [BG]. More precisely, if we let $\Phi(t) = t(1 + \log^+ t)$ and let w be a weight satisfying the A_∞ condition, then there exists a positive constant C depending on the BMO constant of b such that, for any smooth function with compact support f,

$$\begin{aligned} &\sup_{t>0} \frac{1}{\Phi(1/t)} w(\{y \in \mathbf{R}^n : |[b,T]f(y)| > t\}) \\ &\qquad \le C[w]_{A_\infty} \sup_{t>0} \frac{1}{\Phi(1/t)} w(\{y \in \mathbf{R}^n : M^2 f(y) > t\}). \quad (3) \end{aligned}$$

Using this estimate with $w = 1$ and analyzing the behavior of M^2, we obtain the desired estimate (2)—where, in fact, the Lebesgue measure can be replaced by any weight function satisfying the A_1 condition. The L^p versions of these estimates and their consequences are further exploited in [P3].

As mentioned before, we provide a different proof of (2) whose advantage is that it allows us to derive a sharp two-weight inequality that is similar in spirit to the following one for Calderón–Zygmund singular integral operators.

THEOREM 1.2 [P1]. *Let T be any Calderón–Zygmund operator and let $\varepsilon > 0$. Then, for any weight w, function f, and $t > 0$, there is a constant C_ε such that*

$$w(\{x \in \mathbf{R}^n : |Tf(x)| > t\}) \le \frac{C_\varepsilon}{t} \int_{\mathbf{R}^n} |f(x)| M_{L(\log L)^\varepsilon}(w)(x)\, dx. \quad (4)$$

The point here is that no assumption on the weight is assumed. Here $M_A = M_{A(L)}$ denotes a maximal-type function defined by the expression

$$M_{A(L)} f(x) = \sup_{Q \ni x} \|f\|_{A,Q},$$

where A is any Young function and $\|f\|_{A,Q}$ denotes the A-average over Q defined by means of the Luxembourg norm

$$\|f\|_{A,Q} = \inf\left\{\lambda > 0 : \frac{1}{|Q|}\int_Q A\left(\frac{|f|}{\lambda}\right) dx \le 1\right\}. \quad (5)$$

For our applications, the main examples are given by $A(t) = t(1 + \log^+ t)^\alpha$, $\alpha \ge 0$.

We will consider a more general version of (1) denoted by T_b^m $(m = 0, 1, 2, \dots)$ and usually called higher-order commutators. They are defined by the formula

$$T_b^m = \underbrace{[b, \dots, [b, T]]}_{(m \text{ times})};$$

in the particular case of Calderón–Zygmund operators, they can be expressed by means of its kernel K:

$$T_b^m f(x) = \int_{\mathbf{R}^n} (b(x) - b(y))^m K(x, y) f(y)\, dy,$$

where f is an appropriate test function. As usual, we assume that the kernel K satisfies the so-called standard estimates (cf. [C] or [J]).

Our result is the following.

THEOREM 1.3. *Let T be a Calderón–Zygmund singular integral operator, and let $b \in \mathrm{BMO}$ and $\varepsilon > 0$. Then there exists a positive constant C such that*

$$w(\{x \in \mathbf{R}^n : |T_b^m f(x)| > \lambda\}) \le C \int_{\mathbf{R}^n} \Phi_m\left(\|b\|_{\mathrm{BMO}}^m \frac{|f(x)|}{\lambda} \right) M_{L(\log L)^{m+\varepsilon}}(w)(x)\, dx, \quad (6)$$

where $\Phi_m(t) = t(1 + \log^+ t)^m$. The constant C is independent of the weight w, the function f, and $\lambda > 0$.

Observe that there is no restriction on the class of weights considered. Observe also that, since Φ_m is submultiplicative (i.e., $\Phi_m(ab) \le C\Phi_m(a)\Phi_m(b)$ with $a, b \ge 0$), we have

$$w(\{x \in \mathbf{R}^n : |T_b^m f(x)| > \lambda\}) \le C\Phi_m(\|b\|_{\mathrm{BMO}}^m) \int_{\mathbf{R}^n} \Phi_m\left(\frac{|f(x)|}{\lambda} \right) M_{L(\log L)^{m+\varepsilon}}(w)(x)\, dx.$$

Inequalities similar to (6) have turned out to be very useful in the study of the two-weight problem for singular integral operators (see [CP1; CP2]). On the other hand, it would be interesting to know whether or not this inequality holds when $\varepsilon = 0$.

2. Some Preliminaries and Notation

In this section we summarize a few facts about Orlicz spaces. (For more information, see Bennett and Sharpley [BS] or Rao and Ren [RR].) A function $B: [0, \infty) \to [0, \infty)$ is a *doubling Young function* if:

(a) it is continuous, convex, and increasing;
(b) $B(0) = 0$ and $B(t) \to \infty$ as $t \to \infty$; and
(c) it satisfies $B(2t) \le CB(t)$ for all $t > 0$.

For Orlicz norms we are usually concerned about the behavior of Young functions for t large. Given two functions B and C, we write $B(t) \cong C(t)$ if $B(t)/C(t)$ is bounded and bounded below for $t \ge c > 0$.

Recall that we defined the localized Luxembourg norm by equation (5); an equivalent norm that is often useful in calculations is due to Krasnosel'skiĭ and Rutickiĭ [KR, p. 92] (also see [RR, p. 69]):

$$\|f\|_{A,Q} \le \inf_{\mu>0}\left\{\mu + \frac{\mu}{|Q|}\int_Q A\left(\frac{|f|}{\mu}\right)dx\right\} \le 2\|f\|_{A,Q}. \tag{7}$$

Given a Young function A, we use $\bar{A}$ to denote the complementary Young function associated to A; it has the property that, for all $t > 0$,

$$t \le A^{-1}(t)\bar{A}^{-1}(t) \le 2t.$$

The basic property that we will use is the following generalized Hölder inequality:

$$\frac{1}{|Q|}\int_Q |fg| \le 2\|f\|_{A,Q}\|g\|_{\bar{A},Q}. \tag{8}$$

In particular, we shall work with $A(t) = t(1+\log^+ t)^m$, $m = 1, 2, \dots$, with maximal function denoted by $M_{L(\log L)^m}$. The complementery Young function is given by $\bar{A}(t) \approx \exp(t^{1/m})$, with the corresponding maximal function denoted by $M_{\exp L^{1/m}}$.

The first generalized Young inequality states that $A^{-1}(t)\cdot B^{-1}(t) \le C^{-1}(t)$ for $t > 0$; it follows that

$$C(st) \le A(s) + B(t) \tag{9}$$

holds for all $s, t > 0$.

3. Proof of the Theorem

In this section we prove Theorem 1.3 by induction from the case $m = 1$. We will use the following strong-type version of our estimate derived in [P3].

THEOREM 3.1 [P3]. *Let T be any Calderón–Zygmund singular integral operator, and let $1 < p < \infty$ and $b \in \mathrm{BMO}$. Then for each $\delta > 0$ there exists a positive constant $C = C_\delta$ such that, for all functions g,*

$$\int_{\mathbf{R}^n} |T_b^m g(x)|^p w(x)\,dx \le C_\delta \|b\|_{\mathrm{BMO}}^{mp}\int_{\mathbf{R}^n}|g(y)|^p M_{L(\log L)^{(m+1)p-1+\delta}}(w)(y)\,dy. \tag{10}$$

3.1. The Case $m = 1$

A simple homogeneity shows that we may assume $\|b\|_{\mathrm{BMO}} = 1$. Given that assumption, we need only show that

$$w(\{x \in \mathbf{R}^n : |[b,T]f(x)| > \lambda\}) \le C\int_{\mathbf{R}^n}\Phi\left(\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^{1+\varepsilon}}(w)(x)\,dx,$$

where $\Phi(t) = \Phi_1(t) = t(1+\log^+ t)$.

We consider the standard Calderón–Zygmund decomposition of f at level λ and obtain a collection of dyadic non-overlapping cubes $Q_j = Q_j(x_{Q_j}, r_j)$ that satisfy

$$\lambda < \frac{1}{|Q_j|}\int_{Q_j}|f| \le 2^n\lambda. \tag{11}$$

We set $\Omega = \Omega_\lambda = \bigcup_j Q_j$; then $|f(x)| \le \lambda$ a.e. $x \in \mathbf{R}^n \setminus \Omega$. We write $f = g + h$, where g is defined by

$$g(x) = \begin{cases} f(x) & \text{if } x \in \mathbf{R}^n \setminus \Omega, \\ f_{Q_j} & \text{if } x \in Q_j. \end{cases}$$

As usual, we use the notation $f_Q = \frac{1}{|Q|}\int_Q f$ for a locally integrable function f and a cube Q. Observe that $|g(x)| \le 2^n\lambda$ a.e.

We split the "bad part" h as $h = \sum_j h_j$, where $h_j(x) = (f(x) - f_{Q_j})\chi_{Q_j}(x)$. We will use the notation $w^*(x) = w(x)\chi_{\mathbf{R}^n\setminus\tilde{\Omega}}(x)$ and $w_j(x) = w(x)\chi_{\mathbf{R}^n\setminus 3Q_j}$, where $\tilde{Q}_j = 3Q_j$ and $\tilde{\Omega} = \bigcup_j \tilde{Q}_j$. Then

$$\begin{aligned} &w(\{x \in \mathbf{R}^n : |[b,T]f(x)| > \lambda\}) \\ &\quad\le w(\{x \in \mathbf{R}^n \setminus \tilde{\Omega} : |[b,T]g(x)| > \lambda/2\}) + w(\tilde{\Omega}) \\ &\quad\quad + w(\{x \in \mathbf{R}^n \setminus \tilde{\Omega} : |[b,T]h(x)| > \lambda/2\}) \\ &\quad = \mathrm{I} + \mathrm{II} + \mathrm{III}. \end{aligned}$$

As we will see from the proof, part I (precisely the piece associated to the "good part") is the one that carries a higher degree of singularity. Now we use Theorem 3.1, with $m = 1$ and with p, δ such that $1 < p < 1 + \varepsilon/2$ and $\delta = \varepsilon - 2(p-1) > 0$. Then

$$\begin{aligned} \mathrm{I} &\le \frac{C}{\lambda^p}\int_{\mathbf{R}^n} |[b,T]g(x)|^p w^*(x)\,dx \\ &\le \frac{C}{\lambda^p}\int_{\mathbf{R}^n} |g(x)|^p M_{L(\log L)^{1+\varepsilon}}(w^*)(x)\,dx \\ &\le \frac{C}{\lambda}\int_{\mathbf{R}^n} |g(x)| M_{L(\log L)^{1+\varepsilon}}(w^*)(x)\,dx \\ &= \frac{C}{\lambda}\left(\int_{\mathbf{R}^n\setminus\Omega} |f(x)| M_{L(\log L)^{1+\varepsilon}}(w)(x)\,dx + \int_\Omega |g(y)| M_{L(\log L)^{1+\varepsilon}}(w^*)(y)\,dy\right). \end{aligned}$$

It is clear that we need only estimate the second term in the last expression; to do so, we use the following fact:

> For arbitrary Young function A, nonnegative measure w with $M_A w(x) < \infty$ a.e., cube Q, and $R > 1$, we have
>
> $$M_A(\chi_{\mathbf{R}^n\setminus RQ}w)(y) \approx M_A(\chi_{\mathbf{R}^n\setminus RQ}w)(z)$$
>
> for each $y, z \in Q$; hence
>
> $$M_A(\chi_{\mathbf{R}^n\setminus RQ}w)(y) \approx \inf_{y\in Q} M_A(\chi_{\mathbf{R}^n\setminus RQ}w)(y) \tag{12}$$
>
> for each $y \in Q$.

This is an observation whose proof follows exactly as for the case of the Hardy–Littlewood maximal operator M, which corresponds to the case $A(t) = t$ (see e.g. [GR, p. 159]).

Hence we can continue estimating the second term with

$$\begin{aligned}\int_\Omega |g(x)| M_{L(\log L)^{1+\varepsilon}}(w^*)(x)\,dx &\le \sum_j \int_{Q_j} |f_{Q_j}| M_{L(\log L)^{1+\varepsilon}}(w_j)(x)\,dx \\ &= \sum_j \left(\int_{Q_j} |f(x)|\,dx\right) \frac{1}{|Q_j|}\int_{Q_j} M_{L(\log L)^{1+\varepsilon}}(w_j)(x)\,dx \\ &\le C \sum_j \left(\int_{Q_j} |f(x)|\,dx\right) \inf_{Q_j} M_{L(\log L)^{1+\varepsilon}}(w_j) \\ &= C \sum_j \int_{Q_j} |f(x)| M_{L(\log L)^{1+\varepsilon}}(w)(x)\,dx \\ &\le C \int_{\mathbf{R}^n} |f(x)| M_{L(\log L)^{1+\varepsilon}}(w)(x)\,dx.\end{aligned}$$

For II we have

$$\begin{aligned}\mathrm{II} = w(\tilde\Omega) \le C \sum_j \frac{w(\tilde Q_j)}{|\tilde Q_j|}|Q_j| &\le \frac{C}{\lambda}\sum_j \frac{w(\tilde Q_j)}{|\tilde Q_j|}\int_{Q_j}|f(x)|\,dx \\ &\le \frac{C}{\lambda}\sum_j \int_{Q_j} |f(x)| Mw(x)\,dx \\ &\le \frac{C}{\lambda}\int_{\mathbf{R}^n} |f(x)| Mw(x)\,dx.\end{aligned}$$

Observe that this part is smoother than I since we obtain a smaller operator M on the right-hand side. Similarly, part III is smoother than I but rougher than II, as we now show. Indeed, first note that

$$\begin{aligned}[b,T]h(x) &= \sum_j [b,T]h_j(x) \\ &= \sum_j (b(x)-b_{Q_j})Th_j(x) - \sum_j T((b-b_{Q_j})h_j)(x),\end{aligned}$$

where (as before) $b_Q = \frac{1}{|Q|}\int_Q b$. Then

$$\begin{aligned}\mathrm{III} \le\ & w\left(\left\{x \in \mathbf{R}^n \setminus \tilde\Omega : \left|\sum_j (b(x)-b_{Q_j})Th_j(x)\right| > \frac{\lambda}{4}\right\}\right) \\ &+ w\left(\left\{x \in \mathbf{R}^n \setminus \tilde\Omega : \left|\sum_j T((b-b_{Q_j})h_j)(x)\right| > \frac{\lambda}{4}\right\}\right) \\ =\ & A + B.\end{aligned}$$

Using the standard estimates of the kernel K, we have

$$
\begin{aligned}
A &\leq \frac{C}{\lambda}\int_{\mathbf{R}^n\setminus\tilde{\Omega}}\sum_j |b(x)-b_{Q_j}||Th_j(x)|w(x)\,dx\\
&\leq \frac{C}{\lambda}\sum_j\int_{\mathbf{R}^n\setminus 3Q_j}|b(x)-b_{Q_j}|w(x)\int_{Q_j}|h_j(y)||K(x-y)-K(x-x_{Q_j})|\,dy\,dx\\
&\leq \frac{C}{\lambda}\sum_j\int_{Q_j}|h_j(y)|\int_{\mathbf{R}^n\setminus 3Q_j}|K(x-y)-K(x-x_{Q_j})||b(x)-b_{Q_j}|w_j(x)\,dx\,dy\\
&\leq \frac{C}{\lambda}\sum_j\int_{Q_j}|h_j(y)|\sum_{k=1}^{\infty}\\
&\qquad\int_{2^k r_j\leq|x-x_{Q_j}|<2^{k+1}r_j}\frac{|y-x_{Q_j}|}{|x-x_{Q_j}|^{n+1}}|b(x)-b_{Q_j}|w_j(x)\,dx\,dy\\
&\leq \frac{C}{\lambda}\sum_j\left(\int_{Q_j}|h_j(y)|\,dy\right)\sum_{k=1}^{\infty}\frac{2^{-k}}{(2^{k+1}r_j)^n}\int_{|x-x_{Q_j}|<2^{k+1}r_j}|b(x)-b_{Q_j}|w_j(x)\,dx.
\end{aligned}
$$

To control the sum on k, we use standard estimates together with the generalized Hölder inequality and the John–Nirenberg theorem. Indeed, if $y\in Q_j$ then we have

$$
\begin{aligned}
&\sum_{k=1}^{\infty}\frac{2^{-k}}{(2^{k+1}r_j)^n}\int_{|x-x_{Q_j}|<2^{k+1}r_j}|b(x)-b_{Q_j}|w_j(x)\,dx\\
&\qquad\leq C\sum_{k=1}^{\infty}\frac{2^{-k}}{(2^{k+1}r_j)^n}\int_{2^{k+1}Q_j}|b(x)-b_{2^{k+1}Q_j}|w_j(x)\,dx\\
&\qquad\quad+\sum_{k=1}^{\infty}\frac{2^{-k}}{(2^{k+1}r_j)^n}\int_{2^{k+1}Q_j}|b_{2^{k+1}Q_j}-b_{Q_j}|w_j(x)\,dx\\
&\qquad\leq C\sum_{k=1}^{\infty}2^{-k}\|b-b_{2^{k+1}Q_j}\|_{\exp L,2^{k+1}Q_j}\|w_j\|_{L\log L,2^{k+1}Q_j}\\
&\qquad\quad+\sum_{k=1}^{\infty}2^{-k}(k+1)M(w_j)(y)\\
&\qquad\leq C\left(M_{L(\log L)}(w_j)(y)\sum_{k=1}^{\infty}2^{-k}+M(w_j)(y)\sum_{k=1}^{\infty}2^{-k}(k+1)\right)\\
&\qquad\leq CM_{L\log L}(w_j)(y).
\end{aligned}
$$

Then we can continue the estimate of A using (12) as follows:

$$\begin{aligned}
A &\le \frac{C}{\lambda}\sum_j \int_{Q_j} |h_j(y)| M_{L\log L}(w_j)(y)\,dy \\
&\le \frac{C}{\lambda}\Big(\sum_j \int_{Q_j} |f(y)| M_{L\log L}(w)(y)\,dy + \sum_j \int_{Q_j} |f_{Q_j}| M_{L\log L}(w_j)(y)\,dy\Big) \\
&\le \frac{C}{\lambda}\Big(\int_{\mathbf{R}^n} |f(y)| M_{L\log L}(w)(y)\,dy \\
&\qquad + \sum_j \int_{Q_j} |f(x)|\,dx \frac{1}{|Q_j|}\int_{Q_j} M_{L\log L}(w_j)(y)\,dy\Big) \\
&\le \frac{C}{\lambda}\Big(\int_{\mathbf{R}^n} |f(y)| M_{L\log L}(w)(y)\,dy + \sum_j \int_{Q_j} |f(x)| M_{L\log L}(w)(x)\,dx\Big) \\
&\le \frac{C}{\lambda}\int_{\mathbf{R}^n} |f(y)| M_{L\log L}(w)(y)\,dy.
\end{aligned}$$

To estimate B, we combine inequality (4) for singular integrals together with (again) observation (12):

$$\begin{aligned}
B &= w^*\Big(\Big\{x \in \mathbf{R}^n : \Big|T\Big(\sum_j (b-b_{Q_j})h_j\Big)(x)\Big| > \frac{\lambda}{4}\Big\}\Big) \\
&\le \frac{C}{\lambda}\int_{\mathbf{R}^n} \Big|\sum_j (b(x)-b_{Q_j})h_j)(x)\Big| M_{L(\log L)^\varepsilon}(w^*)(x)\,dx \\
&\le \frac{C}{\lambda}\sum_j \int_{Q_j} |b(x)-b_{Q_j}||f(x)-f_{Q_j}| M_{L(\log L)^\varepsilon}(w_j)(x)\,dx \\
&\le \frac{C}{\lambda}\sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x)\Big(\int_{Q_j} |b(x)-b_{Q_j}||f(x)|\,dx \\
&\qquad\qquad + \int_{Q_j} |b(x)-b_{Q_j}||f_{Q_j}|\,dx\Big) \\
&= B_1 + B_2.
\end{aligned}$$

The estimate for B_2 is simple since, by (12),

$$\begin{aligned}
B_2 &= \frac{C}{\lambda}\sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x)\int_{Q_j} |b(x)-b_{Q_j}||f_{Q_j}|\,dx \\
&\le \frac{C}{\lambda}\sum_j \frac{1}{|Q_j|}\int_{Q_j} |b(x)-b_{Q_j}|\int_{Q_j} |f(x)| M_{L(\log L)^\varepsilon}(w_j)(x)\,dx \\
&\le C\int_{\mathbf{R}^n} |f(x)| M_{L(\log L)^\varepsilon}(w)(x)\,dx.
\end{aligned}$$

For B_1 we have, by the generalized Hölder inequality (8),

$$B_1 = \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x) \int_{Q_j} |b(x) - b_{Q_j}||f(x)|\, dx$$
$$\leq \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x)|Q_j|\|f\|_{L\log L, Q_j}.$$

Now, combining formula (7) with (11) and recalling that $\Phi(t) = t(1 + \log^+ t)$, we have

$$\begin{aligned}\frac{1}{\lambda}|Q_j|\|f\|_{L\log L, Q_j} &\leq \frac{1}{\lambda}|Q_j| \inf_{\mu>0}\left\{\mu + \frac{\mu}{|Q_j|}\int_{Q_j} \Phi\left(\frac{|f(x)|}{\mu}\right) dx\right\} \\ &\leq |Q_j| + \int_{Q_j} \Phi\left(\frac{|f(x)|}{\lambda}\right) dx \\ &\leq \frac{1}{\lambda}\int_{Q_j} |f(x)|\, dx + \int_{Q_j} \Phi\left(\frac{|f(x)|}{\lambda}\right) dx \\ &\leq 2\int_{Q_j} \Phi\left(\frac{|f(x)|}{\lambda}\right) dx.\end{aligned}$$

Then

$$\begin{aligned}B_1 &\leq C\int_{Q_j} \Phi\left(\frac{|f(x)|}{\lambda}\right) M_{L(\log L)^\varepsilon}(w_j)(x)\, dx \\ &\leq C\int_{\mathbf{R}^n} \Phi\left(\frac{|f(x)|}{\lambda}\right) M_{L(\log L)^\varepsilon}(w)(x)\, dx.\end{aligned}$$

This concludes the proof of the case $m = 1$. □

3.2. The General Case

We will use an induction argument and will omit some technical arguments that are similar to the case $m = 1$. Again, a simply homogeneity argument using that $T_b^m(f/\|b\|_{\mathrm{BMO}}^m) = T_{b/\|b\|_{\mathrm{BMO}}}^m(f)$ shows that we may assume $\|b\|_{\mathrm{BMO}} = 1$. We consider again the Calderón–Zygmund decomposition of f at level λ. Then, with the same notation as in the proof of the case $m = 1$, we have

$$\begin{aligned}w(\{y \in \mathbf{R}^n : |T_b^m f(y)| > \lambda\}) &\leq w(\{y \in \mathbf{R}^n \setminus \tilde{\Omega} : |T_b^m g(y)| > \lambda/2\}) + w(\tilde{\Omega}) \\ &\quad + w(\{y \in \mathbf{R}^n \setminus \tilde{\Omega} : |T_b^m h(y)| > \lambda/2\}) \\ &= \mathrm{I} + \mathrm{II} + \mathrm{III}.\end{aligned}$$

From (10) with p and δ such that

$$1 < p < 1 + \varepsilon/(m+1) \quad \text{and} \quad \delta = \varepsilon - (m+1)(p-1) > 0,$$

we have

$$\begin{aligned} \mathrm{I} &\le \frac{C}{\lambda}\int_{\mathbf{R}^n} |g(y)| M_{L(\log L)^{m+\varepsilon}}(w^*)(y)\,dy \\ &\le \frac{C}{\lambda}\int_{\mathbf{R}^n} |f(y)| M_{L(\log L)^{m+\varepsilon}}(w)(y)\,dy, \end{aligned}$$

as in the case $m = 1$. Similarly, for II we have

$$\mathrm{II} \le \frac{C}{\lambda}\int_{\mathbf{R}^n} |f(x)| Mw(x)\,dx. \tag{13}$$

To estimate III, we split the operator as in [GHST]:

$$\begin{aligned} T_b^m h_j(x) &= \int_{\mathbf{R}^n} (b(x)-b(y))^m K(x-y) h_j(y)\,dy \\ &= \sum_{l=0}^{m} C_{l,m}(b(x)-\alpha)^{m-l}\int_{\mathbf{R}^n}(b(y)-\alpha)^l K(x-y)h_j(y)\,dy \\ &= C(b(x)-\alpha)^m Th_j(x) + T((b-\alpha)^m h_j)(x) \\ &\quad + \sum_{l=1}^{m-1} C_{l,m}(b(x)-\alpha)^{m-l}\int_{\mathbf{R}^n}(b(y)-\alpha)^l K(x-y)h_j(y)\,dy, \end{aligned}$$

where α is a number to be chosen soon. Then the last term is further broken as follows:

$$\begin{aligned} &\sum_{l=1}^{m-1} C_{l,m}(b(x)-\alpha)^{m-l}\int_{\mathbf{R}^n}(b(y)-\alpha)^l K(x-y)h_j(y)\,dy \\ &\quad = \sum_{l=1}^{m-1} C_{l,m}\sum_{h=0}^{m-l} C_{h,m,l}\int_{\mathbf{R}^n}(b(x)-b(y))^h(b(y)-\alpha)^{m-h}K(x-y)h_j(y)\,dy \\ &\quad = \sum_{h=0}^{m-1} C_{m,h}\int_{\mathbf{R}^n}(b(x)-b(y))^h(b(y)-\alpha)^{m-h}K(x-y)h_j(y)\,dy \\ &\quad = CT((b-\alpha)^m h_j)(x) + \sum_{h=1}^{m-1} C_{m,h}T_b^h((b-\alpha)^{m-h}h_j)(x). \end{aligned}$$

If we now take $\alpha = b_{Q_j}$ then we obtain

$$\begin{aligned} \sum_j T_b^m h_j(x) &= C\sum_j (b(x)-b_{Q_j})^m Th_j(x) + \sum_j T((b-b_{Q_j})^m h_j)(x) \\ &\quad + \sum_{h=1}^{m-1} C_{m,h}T_b^h\Big(\sum_j (b-b_{Q_j})^{m-h}h_j\Big)(x). \end{aligned}$$

Thus,

$$\begin{aligned}
\mathrm{III} &\le w\left(\left\{y \in \mathbf{R}^n \setminus \tilde{\Omega} : \left|\sum_j (b(y) - b_{Q_j})^m T h_j(y)\right| > \frac{\lambda}{6}\right\}\right) \\
&\quad + w\left(\left\{y \in \mathbf{R}^n \setminus \tilde{\Omega} : \left|\sum_j T((b - b_{Q_j})^m h_j)(y)\right| > \frac{\lambda}{6}\right\}\right) \\
&\quad + w\left(\left\{y \in \mathbf{R}^n \setminus \tilde{\Omega} : \left|\sum_{h=1}^{m-1} C_{m,h} T_b^h\left(\sum_j (b - b_{Q_j})^{m-h} h_j\right)(y)\right| > \frac{\lambda}{6}\right\}\right) \\
&= A + B + C.
\end{aligned}$$

To estimate A, we proceed as in the case $m = 1$ to obtain

$$\begin{aligned}
A &\le \frac{C}{\lambda} \int_{\mathbf{R}^n \setminus \tilde{\Omega}} \sum_j |b(x) - b_{Q_j}|^m |T h_j(x)| w(x)\, dx \\
&\le \frac{C}{\lambda} \sum_j \left(\int_{Q_j} |h_j(y)|\, dy\right) \sum_{k=1}^{\infty} \frac{2^{-k}}{(2^{k+1} r_j)^n} \\
&\quad \times \int_{|x - x_{Q_j}| < 2^{k+1} r_j} |b(x) - b_{Q_j}|^m w_j(x)\, dx.
\end{aligned}$$

If $y \in Q_j$ then, by the generalized Hölder inequality and the John–Nirenberg theorem (recall that $\|b\|_{\mathrm{BMO}} = 1$), we have

$$\begin{aligned}
&\sum_{k=1}^{\infty} \frac{2^{-k}}{(2^{k+1} r_j)^n} \int_{|x - x_{Q_j}| < 2^{k+1} r_j} |b(x) - b_{Q_j}|^m w_j(x)\, dx \\
&\quad \le \sum_{k=1}^{\infty} \frac{2^{-k}}{(2^{k+1} r_j)^n} \int_{2^{k+1} Q_j} |b(x) - b_{2^{k+1} Q_j}|^m w_j(x)\, dx \\
&\qquad + \sum_{k=1}^{\infty} \frac{2^{-k}}{(2^{k+1} r_j)^n} \int_{2^{k+1} Q_j} |b_{2^{k+1} Q_j} - b_{Q_j}|^m w_j(x)\, dx \\
&\quad \le C \sum_{k=1}^{\infty} 2^{-k} \|(b - b_{2^{k+1} Q_j})^m\|_{\exp L^{1/m}, 2^{k+1} Q_j} \|w_j\|_{L(\log L)^m, 2^{k+1} Q_j} \\
&\qquad + C \sum_{k=1}^{\infty} 2^{-k} (k+1) M(w_j)(y) \\
&\quad \le C M_{L(\log L)^m}(w_j)(y).
\end{aligned}$$

Then

$$
\begin{aligned}
A &\leq \frac{C}{\lambda} \sum_j \int_{Q_j} |h_j(y)| M_{L(\log L)^m}(w_j)(y)\, dy \\
&\leq \frac{C}{\lambda} \Big(\sum_j \int_{Q_j} |f(y)| M_{(L\log L)^m}(w)(y)\, dy \\
&\qquad + \sum_j \int_{Q_j} |f_{Q_j}| M_{L(\log L)^m}(w_j)(y)\, dy \Big) \\
&\leq \frac{C}{\lambda} \int_{\mathbf{R}^n} |f(y)| M_{L(\log L)^m}(w)(y)\, dy,
\end{aligned}
$$

where we have used observation (12). Again, this observation combined with inequality (4) yields

$$
\begin{aligned}
B &= w^*\Big(\Big\{x \in \mathbf{R}^n : \Big|T\Big(\sum_j (b - b_{Q_j})^m h_j\Big)(x)\Big| > \frac{\lambda}{6}\Big\}\Big) \\
&\leq \frac{C}{\lambda} \sum_j \int_{Q_j} |b(x) - b_{Q_j}|^m |f(x) - f_{Q_j}| M_{L(\log L)^\varepsilon}(w_j)(x)\, dx \\
&\leq \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x) \Big(\int_{Q_j} |b(x) - b_{Q_j}|^m |f(x)|\, dx \\
&\qquad + \int_{Q_j} |b(x) - b_{Q_j}|^m |f_{Q_j}|\, dx \Big) \\
&= B_1 + B_2.
\end{aligned}
$$

But it is easy to see that

$$
B_2 \leq C \int_{\mathbf{R}^n} |f(x)| M_{L(\log L)^\varepsilon} w(x)\, dx;
$$

on the other hand, by the generalized Hölder inequality (8) we have

$$
\begin{aligned}
B_1 &= \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x) \int_{Q_j} |b(x) - b_{Q_j}|^m |f(x)|\, dx \\
&\leq \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x) |Q_j| \|(b - b_{Q_j})^m\|_{\exp L^{1/m}, Q_j} \|f\|_{L(\log L)^m, Q_j} \\
&\leq \frac{C}{\lambda} \sum_j \inf_{Q_j} M_{L(\log L)^\varepsilon}(w_j)(x) |Q_j| \|f\|_{L(\log L)^m, Q_j}.
\end{aligned}
$$

Recalling that $\Phi_m(t) = t(1 + \log^+ t)^m$, by (7) we have

$$\frac{1}{\lambda}|Q_j|\|f\|_{L(\log L)^m,Q_j} \le \frac{1}{\lambda}|Q_j| \inf_{\mu>0}\left\{\mu + \frac{\mu}{|Q_j|}\int_{Q_j}\Phi^m\left(\frac{|f(x)|}{\mu}\right)dx\right\}$$
$$\le 2\int_{Q_j}\Phi^m\left(\frac{|f(x)|}{\lambda}\right)dx$$

and consequently

$$B_1 \le C\sum_j \int_{Q_j}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^\varepsilon}(w_j)(x)\,dx$$
$$\le C\int_{\mathbf{R}^n}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^\varepsilon}(w)(x)\,dx.$$

To conclude the proof of the theorem, we are left with the estimate for C, where we will use the induction argument:

$$w\left(\left\{y\in\mathbf{R}^n\setminus\tilde{\Omega} : \left|\sum_{h=1}^{m-1}C_{m,h}T_b^h\left(\sum_j (b-b_{Q_j})^{m-h}h_j\right)(y)\right| > \frac{\lambda}{6}\right\}\right)$$
$$\le w\left(\left\{y\in\mathbf{R}^n\setminus\tilde{\Omega} : \left|\sum_{h=1}^{m-1}C_{m,h}T_b^h\left(f\sum_j (b-b_{Q_j})^{m-h}\chi_{Q_j}\right)(y)\right| > \frac{\lambda}{12}\right\}\right)$$
$$+ w\left(\left\{y\in\mathbf{R}^n\setminus\tilde{\Omega} : \left|\sum_{h=1}^{m-1}C_{m,h}T_b^h\left(\sum_j (b-b_{Q_j})^{m-h}f_{Q_j}\chi_{Q_j}\right)(y)\right| > \frac{\lambda}{12}\right\}\right)$$
$$= C_1 + C_2.$$

By the induction hypothesis, the theorem holds for $k < m$; then

$$C_1 \le C\sum_{h=1}^{m-1}\int_{\mathbf{R}^n}\Phi_h\left(\frac{|f(x)|}{\lambda}\left|\sum_j (b(x)-b_{Q_j})^{m-h}\chi_{Q_j}(x)\right|\right) M_{L(\log L)^{h+\varepsilon}}(w^*)(x)\,dx$$
$$\le C\sum_{h=1}^{m-1}\sum_j\int_{Q_j}\Phi_h\left(\frac{|f(x)|}{\lambda}|b(x)-b_{Q_j}|^{m-h}\right)M_{L(\log L)^{h+\varepsilon}}(w_j)(x)\,dx$$
$$\le C\sum_{h=1}^{m-1}\sum_j\inf_{Q_j}M_{L(\log L)^{h+\varepsilon}}(w_j)\int_{Q_j}\Phi_h\left(\frac{|f(x)|}{\lambda}|b(x)-b_{Q_j}|^{m-h}\right)dx.$$

Let $\psi_k(t) = \exp t^{1/k} - 1$. Then $\Phi_m^{-1}(t)\cdot\psi_{m-h}^{-1}(t) \le C\Phi_h^{-1}(t)$, because $\Phi_k^{-1}(t) \approx t/(\log t)^k$ and $\psi_k^{-1}(t) \approx (\log t)^k$. Then, combining (9) with the John–Nirenberg theorem and (11), we obtain

$$\begin{aligned}
&\int_{Q_j} \Phi_h\left(\frac{|f(x)|}{\lambda}|b(x)-b_{Q_j}|^{m-h}\right)dx \\
&\quad\le \int_{Q_j} \Phi_m\left(\frac{|f(x)|}{\lambda}\right)dx + \int_{Q_j} \psi_{m-h}(|b(x)-b_{Q_j}|^{m-h})\,dx \\
&\quad\le \int_{Q_j} \Phi_m\left(\frac{|f(x)|}{\lambda}\right)dx + C|Q_j| \\
&\quad\le C\int_{Q_j} \Phi_m\left(\frac{|f(x)|}{\lambda}\right)dx.
\end{aligned}$$

Hence C_1 can finally be estimated by

$$\begin{aligned}
C_1 &\le C\sum_{h=1}^{m-1}\sum_j \inf_{Q_j} M_{L(\log L)^{h+\varepsilon}}(w_j)\int_{Q_j}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)dx \\
&\le C\sum_{h=1}^{m-1}\sum_j \int_{Q_j}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^{h+\varepsilon}}(w)(x)\,dx \\
&\le C\int_{\mathbf{R}^n}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^{m-1+\varepsilon}}(w)(x)\,dx.
\end{aligned}$$

We may control C_2 in similar way by observing that (9) and Jensen's inequality yield

$$\begin{aligned}
\int_{Q_j} \Phi_h\left(\frac{|f_{Q_j}|}{\lambda}|b(x)-b_{Q_j}|^{m-h}\right)dx &\le \int_{Q_j}\Phi_m\left(\frac{|f|_{Q_j}}{\lambda}\right)dx + C|Q_j| \\
&\approx |Q_j| \approx \frac{1}{\lambda}\int_{Q_j}|f| \\
&\le C\int_{Q_j}\Phi_m\left(\frac{|f(x)|}{\lambda}\right)dx.
\end{aligned}$$

The proof is complete. □

References

[BS] C. Bennett and R. Sharpley, *Interpolation of operators*, Academic Press, New York, 1988.

[BG] D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasi-linear operators on martingales*, Acta Math. 124 (1970), 249–304.

[C] M. Christ, *Lectures on singular integral operators*, CBMS Regional Conf. Ser. in Math., 77, Amer. Math. Soc., Providence, RI, 1990.

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

[CP1] D. Cruz-Uribe and C. Pérez, *Sharp two-weight, weak-type norm inequalities for singular integral operators*, Math. Res. Lett. 6 (1999), 417–428.

[CP2] ———, *Two-weight, weak-type norm inequalities for fractional integrals, Calderón–Zygmund operators and commutators,* Indiana Univ. Math. J. 49 (2000), 697–721.

[GHST] J. Garcia-Cuerva, E. Harboure, C. Segovia, and J. L. Torrea, *Weighted norm inequalities for commutators of strongly singular integrals,* Indiana Univ. Math. J. 40 (1991), 1397–1420.

[GR] J. Garcia-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics,* North-Holland Math. Stud., 116, North-Holland, Amsterdam, 1985.

[J] J. L. Journé, *Calderón–Zygmund operators, pseudodifferential operators and the Cauchy integral of Calderón,* Lecture Notes in Math., 994, Springer-Verlag, Berlin, 1983.

[KR] M. A. Krasnosel'skiĭ and Ya. B. Rutickiĭ, *Convex functions and Orlicz spaces,* Noordhoff, Groningen, 1961.

[P1] C. Pérez, *Weighted norm inequalities for singular integral operators,* J. London Math. Soc. (2) 49 (1994), 296–308.

[P2] ———, *Endpoint estimates for commutators of singular integral operators,* J. Funct. Anal. 128 (1995), 163–185.

[P3] ———, *Sharp estimates for commutators of singular integrals via iterations of the Hardy–Littlewood maximal functions,* J. Fourier Anal. Appl. 3 (1997), 743–756.

[RR] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces,* Monogr. Textbooks Pure Appl. Math., 146, Dekker, New York, 1991.

C. Pérez
Departamento de Matemáticas
Universidad Autónoma de Madrid
28049 Madrid
Spain

carlos.perez@uam.es

G. Pradolini
Universidad Nacional del Litoral – CONICET
Güemes 3450
3000 Santa Fe
Argentina

gpradoli@intec.unl.edu.ar