STUDIA MATHEMATICA 195 (2) (2009)

Generalized Hörmander conditions and weighted endpoint estimates

by

María Lorente (Málaga), José María Martell (Madrid),
Carlos Pérez (Sevilla) and María Silvina Riveros (Córdoba)

Abstract. We consider two-weight estimates for singular integral operators and their commutators with bounded mean oscillation functions. Hörmander type conditions in the scale of Orlicz spaces are assumed on the kernels. We prove weighted weak-type estimates for pairs of weights (u, Su) where u is an arbitrary nonnegative function and S is a maximal operator depending on the smoothness of the kernel. We also obtain sufficient conditions on a pair of weights (u, v) for the operators to be bounded from $L^p(v)$ to $L^{p,\infty}(u)$. One-sided singular integrals, like the differential transform operator, are considered as well. We also provide applications to Fourier multipliers and homogeneous singular integrals.

1. Introduction. The Calderón–Zygmund decomposition is a powerful tool in harmonic analysis. Since its discovery in [6], many authors have used it to derive boundedness properties of singular integral operators. For instance, using the fact that the Hilbert and Riesz transforms are bounded on L^2, and by means of the Calderón–Zygmund decomposition, one proves that these classical operators are of weak type $(1, 1)$. From this starting point, in the literature one can find many boundedness results for the Hilbert and Riesz transforms: estimates on L^p, one-weight and two-weight norm inequalities, etc.

The Calderón–Zygmund theory generalizes these ideas to provide a general framework allowing one to deal with singular integral operators. A typical Calderón–Zygmund convolution operator T is bounded on $L^2(\mathbb{R}^n)$ and has a kernel K on which various conditions are assumed. In the easiest case, K behaves as the kernel of the Hilbert or Riesz transform. That is, K decays as $|x|^{-n}$ and its gradient as $|x|^{-n-1}$. It was already proved in [15] that these assumptions can be relaxed if we want to show that T is of weak type $(1, 1)$:

2010 *Mathematics Subject Classification*: 42B20, 42B25.
Key words and phrases: Calderón–Zygmund operators, homogeneous singular integrals, multipliers, one-sided operators, commutators, BMO, Hörmander condition of Young type, Muckenhoupt weights, two-weight estimates.

DOI: 10.4064/sm195-2-5

it suffices to require that K satisfies the so-called *Hörmander condition* (we write $K \in H_1$),

$$\int_{|x|>c|y|} |K(x-y) - K(x)|\, dx \le C, \qquad y \in \mathbb{R}^n,\ c > 1.$$

Hence, by interpolation, T is bounded on $L^p(\mathbb{R}^n)$ with $1 < p < \infty$.

The underlying measure dx can be replaced by $w(x)\,dx$ where w is a Muckenhoupt A_p weight; the Hilbert and Riesz transforms are bounded on $L^p(w) = L^p(w(x)\,dx)$ if and only if $w \in A_p$ for $1 < p < \infty$. For $p = 1$, the weak type $(1,1)$ with respect to w holds if and only if $w \in A_1$. The decay assumed before on the kernel and its gradient guarantees the same weighted estimates for the operator T. However, the Hörmander condition does not suffice to derive such estimates, as is proved in [14] (see also [21]). One can relax the decay conditions assumed on the kernel and still prove the previous weighted norm inequalities. Namely, it is enough to require that K satisfies the following *Lipschitz condition* (we write $K \in H_\infty^*$):

$$|K(x-y) - K(x)| \le C\, \frac{|y|^\alpha}{|x|^{\alpha+n}}, \qquad |x| > c|y|.$$

With this condition in hand, one can show *Coifman's estimate* (see [7]): for any $0 < p < \infty$ and any $w \in A_\infty$,

$$\int_{\mathbb{R}^n} |Tf(x)|^p w(x)\, dx \le C \int_{\mathbb{R}^n} Mf(x)^p w(x)\, dx. \tag{1.1}$$

These estimates can be seen as controlling the operator T by the Hardy–Littlewood maximal function M, and this allows one to show that T satisfies most of the weighted estimates that M does (see [13] for more details).

When relaxing the H_∞^* condition, the operators become more singular and less smooth. Thus, the Coifman estimates to be expected will have a worse maximal operator on the right-hand side. For instance, one has a scale of Hörmander conditions based on the Lebesgue spaces L^r for $1 \le r \le \infty$ (see [17], [34] and [37]). A singular integral operator with kernel satisfying the L^r-Hörmander condition, $1 < r \le \infty$, satisfies a Coifman estimate with the maximal operator $M_{r'}$ on the right-hand side (here $M_{r'}f(x) = M(|f|^{r'})(x)^{1/r'}$). These estimates are shown to be sharp in [21]. Let us notice that if r goes to 1 then r' goes to ∞ and the corresponding Coifman estimates get worse. In particular, when $K \in H_1$ the Coifman estimates fail to hold (see [21]).

Sometimes, this scale of Hörmander conditions based on the Lebesgue spaces is not sufficiently fine and gives estimates that are not accurate enough. For instance, let us consider the differential transform operator studied in [16] and [4]:

$$T^{+}f(x)=\sum_{j\in\mathbb{Z}}\nu_j(D_jf(x)-D_{j-1}f(x)), \tag{1.2}$$

where $\|\{\nu_j\}_j\|_\infty<\infty$ and

$$D_jf(x)=\frac{1}{2^j}\int_x^{x+2^j} f(t)\,dt.$$

We see that T^+ is a singular integral operator with kernel K supported in $(-\infty,0)$, and therefore T^+ is a one-sided singular integral operator (that is why we write T^+). In [4] it was shown that $K\in\bigcap_{r\ge1}H_r$ (here H_r is the Hörmander condition associated with L^r, see the precise definition below). Thus one can show that T^+ satisfies a Coifman estimate with M_q on the right-hand side for any $1<q<\infty$. Indeed, exploiting the fact that T^+ is a one-sided operator one can do better: M_qf can be replaced by the pointwise smaller operator M_q^+f (the corresponding one-sided maximal function) and A_∞ by the bigger class A_∞^+ (see the precise definitions and more details below). Notice that one can take any $1<q<\infty$, with the case $q=1$ remaining open (in general, $K\notin H_\infty$). However, there are other maximal operators between M (or M^+) and M_q (or M_q^+): any iteration of the Hardy–Littlewood maximal function, or maximal operators associated with Orlicz spaces lying between L^1 and L^q, such as $L(\log L)^\alpha$, $\alpha>0$.

These ideas motivated [20] where new classes of Hörmander conditions based on Orlicz spaces were introduced. Roughly, given a Young function $\mathcal{A}$, one can associate with the Orlicz space $L^{\mathcal{A}}$ a Hörmander class $H_{\mathcal{A}}$ (see Definition 2.3). Thus, a singular integral operator with kernel in $H_{\mathcal{A}}$ is controlled in the sense of Coifman by the maximal operator $M_{\bar{\mathcal{A}}}$ (which is the maximal function associated with the space $L^{\bar{\mathcal{A}}}$) where $\bar{\mathcal{A}}$ is the conjugate function of $\mathcal{A}$. This was established in [20], together with the one-sided case (see Theorems 2.4 and 3.11 below).

For the differential transform T^+ introduced above one can show that $K\in H_{e^{t^{1/(1+\varepsilon)}}}$ for any $\varepsilon>0$. Thus T^+ satisfies a Coifman type estimate with $M^+_{L(\log L)^{1+\varepsilon}}$ on the right-hand side—in terms of iterations one can write $(M^+)^3$—and this maximal operator is pointwise smaller than M_q^+ for any $1<q<\infty$.

Coifman's estimates are important from the points of view of weighted norm inequalities since they encode a lot of information about the singularity of the operator T (see [9] and [13]). In some sense, T behaves as the maximal operator that controls it. For instance, from (1.1) one shows that T is bounded on $L^p(w)$ for $1<p<\infty$ and $w\in A_p$. Also, T is of weak type $(1,1)$ for weights in A_1, T is bounded on weighted rearrangement invariant function spaces, T satisfies weighted modular inequalities (see [13]), etc. All these one-weight estimates are based on the fact that (1.1) is valid for any weight in A_∞, and the weight always varies within this class.

The situation changes when one works with two-weight inequalities. Let us focus on the endpoint estimates for $p = 1$. In the one-weight case, M is bounded from $L^1(w)$ to $L^{1,\infty}(w)$ for every $w \in A_1$. Also, there is a version of (1.1) in the sense of $L^{1,\infty}(w)$, that is, $\|Tf\|_{L^{1,\infty}(w)} \lesssim \|Mf\|_{L^{1,\infty}(w)}$ for every $w \in A_\infty$ (see [9]). These two facts imply at once that T is of weak type $(1,1)$ for weights in A_1. In the two-weight case, Vitali's covering lemma easily gives that for every weight u (a *weight* is a nonnegative locally integrable function)

$$u\{x \in \mathbb{R}^n : Mf(x) > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^n} |f(x)| Mu(x)\, dx.$$

However, this estimate is not known for singular integral operators with smooth kernel. Even for Hilbert or Riesz transforms the validity of this estimate is an open question. Reasoning as above, one seeks pairs of weights (u, Su) for which these operators are of weak type $(1,1)$, where S will be a maximal operator worse, in principle, than M. For instance, one can put $S = M_q$ for every $1 < q < \infty$: using the fact that $M_q u \in A_1$ and Coifman's estimate (in $L^{1,\infty}$) one easily obtains the estimate proved in [8],

$$\|Tf\|_{L^{1,\infty}(u)} \le \|Tf\|_{L^{1,\infty}(M_q u)} \lesssim \|Mf\|_{L^{1,\infty}(M_q u)} \lesssim \|f\|_{L^1(M_q u)}.$$

As observed before, there are some other maximal operators that lie between M and M_q, such as the iterations of M or $M_{L(\log L)^\alpha}$, $\alpha > 0$. In [27], by means of the Calderón–Zygmund decomposition, it was proved that if T is a singular integral operator with smooth kernel (say $K \in H^*_\infty$), like the Hilbert or Riesz transform, then for any $\varepsilon > 0$ and any weight u,

$$u\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^n} |f(x)| M_{L(\log L)^\varepsilon} u(x)\, dx. \tag{1.3}$$

Note that in terms of iterations one can write M^2.

The goal of this paper is to study estimates like (1.3) for operators T with less smooth kernels. That is, if we require that the kernel of T satisfies a Hörmander condition in the scale of Orlicz spaces, we look for a maximal operator S such that T is of weak type $(1,1)$ with respect to the pair of weights (u, Su). The main technique to be used is the Calderón–Zygmund decomposition. The bad part, where the best possible result is always obtained, is handled by using the smoothness of the kernel. For the good part, one needs a strong two-weight estimate that usually follows from a Coifman estimate (see Theorem 2.6). We also obtain weighted endpoint inequalities for the commutators of such operators with BMO functions. The corresponding Coifman estimates have been studied in [18]. One of our main examples is the differential transform operator presented above, thus we also pay attention to the one-sided operators in which case one can obtain better estimates by replacing a maximal operator by its corresponding one-sided analog.

The paper is organized as follows. The next section contains some preliminaries and definitions. In Section 3 we state our main results on singular integral operators and their commutators with BMO functions, and we consider the one-sided case. Some applications, including the differential transform operator and multipliers, are given in Section 4. Finally, Sections 5 and 6 contain the proofs of our main results.

2. Preliminaries

2.1. *Young functions and Orlicz spaces.* We recall some background on Orlicz spaces, referring the reader to [32] and [3] for a complete account. A function $\mathcal{A} : [0,\infty) \to [0,\infty)$ is a *Young function* if it is continuous, convex, increasing and satisfies $\mathcal{A}(0) = 0$, $\mathcal{A}(\infty) = \infty$. We will assume that the Young functions are normalized so that $\mathcal{A}(1) = 1$. We introduce the following localized and normalized Luxemburg norm associated with the Orlicz space $L^{\mathcal{A}}$: given a cube Q,

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

For instance, when $\mathcal{A}(t) = t^r$ with $r \ge 1$, we have

$$\|f\|_{L^r,Q} = \left(\frac{1}{|Q|}\int_Q |f(x)|^r\, dx\right)^{1/r}.$$

It is well known that if $\mathcal{A}(t) \le C\mathcal{B}(t)$ for $t \ge t_0$ then $\|f\|_{\mathcal{A},Q} \le C\|f\|_{\mathcal{B},Q}$. Thus the behavior of $\mathcal{A}(t)$ for $t \le t_0$ does not matter: if $\mathcal{A}(t) \approx \mathcal{B}(t)$ for $t \ge t_0$ the previous estimate implies that $\|f\|_{\mathcal{A},Q} \approx \|f\|_{\mathcal{B},Q}$. This means that in most cases we will not be concerned about the values of the Young functions for t small.

We can now define the Hardy–Littlewood maximal function associated with $\mathcal{A}$ as

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

If $\mathcal{A}(t) = t$, then $M_{\mathcal{A}} = M$ is the Hardy–Littlewood maximal function. For $\mathcal{A}(t) = t^r$ with $r > 1$ we have $M_{\mathcal{A}}f(x) = M(|f|^r)(x)^{1/r}$.

Given a Young function $\mathcal{A}$, we say that $\mathcal{A}$ is *doubling*, and write $\mathcal{A} \in \Delta_2$, if $\mathcal{A}(2t) \le C\mathcal{A}(t)$ for every $t \ge t_0 > 0$. For $1 < p < \infty$, $\mathcal{A}$ belongs to B_p if there exists $c > 0$ such that

$$\int_c^\infty \frac{\mathcal{A}(t)}{t^p}\frac{dt}{t} < \infty.$$

This condition appears first in [29]; it is known that $\mathcal{A} \in B_p$ if and only if $M_{\mathcal{A}}$ is bounded on $L^p(\mathbb{R}^n)$.

Abusing notation, if $\mathcal{A}(t) = t^r$, $\mathcal{A}(t) = e^{t^\alpha} - 1$ or $\mathcal{A}(t) = t^r(1 + \log^+ t)^\alpha$, the Orlicz norms are respectively written $\|\cdot\|_r = \|\cdot\|_{L^r}$, $\|\cdot\|_{\exp L^\alpha}$, $\|\cdot\|_{L^r(\log L)^\alpha}$, and the corresponding maximal operators $M_r = M_{L^r}$, $M_{\exp L^\alpha}$ and $M_{L^r(\log L)^\alpha}$. For $k \geq 0$, it is known that $M_{L(\log L)^k} f(x) \approx M^{k+1} f(x)$ where M^k is the k-fold iterate of M (see [28], [33] and [13]).

In $\mathbb{R}$, we can also define the one-sided maximal functions associated with a given Young function $\mathcal{A}$:

$$M^+_{\mathcal{A}} f(x) = \sup_{b>x} \|f\|_{\mathcal{A},(x,b)} \quad \text{and} \quad M^-_{\mathcal{A}} f(x) = \sup_{a<x} \|f\|_{\mathcal{A},(a,x)}.$$

The one-sided Hardy–Littlewood maximal functions M^+, M^- correspond to the case $\mathcal{A}(t) = t$.

Given a Young function $\mathcal{A}$, let $\overline{\mathcal{A}}$ denote its associate function: the Young function with the property that $t \leq \mathcal{A}^{-1}(t)\overline{\mathcal{A}}^{-1}(t) \leq 2t$ for $t > 0$. If $\mathcal{A}(t) = t^r$ with $1 < r < \infty$, then $\overline{\mathcal{A}}(t) \approx t^{r'}$; if $\mathcal{A}(t) = t^r \log(e + t)^\alpha$, then $\overline{\mathcal{A}}(t) \approx t^{r'} \log(e + t)^{-\alpha(r'-1)}$.

One has the *generalized Hölder inequality*

$$\frac{1}{|Q|} \int_Q |fg| \leq 2\|f\|_{\mathcal{A},Q} \|g\|_{\overline{\mathcal{A}},Q}.$$

There is a further generalization that turns out to be useful for our purposes (see [26]): If $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are Young functions such that $\mathcal{A}^{-1}(t)\mathcal{B}^{-1}(t)\mathcal{C}^{-1}(t) \leq t$ for all $t \geq t_0 > 0$ (in what follows we assume that $t_0 = 1$ for simplicity of computations)—sometimes, we will equivalently write $\mathcal{A}^{-1}(t)\mathcal{B}^{-1}(t) \leq \overline{\mathcal{C}}^{-1}(t)$—then

$$\|fgh\|_{L^1,Q} \leq C\|f\|_{\mathcal{A},B} \|g\|_{\mathcal{B},Q} \|h\|_{\mathcal{C},Q}, \qquad \|fg\|_{\overline{\mathcal{C}},Q} \leq C\|f\|_{\mathcal{A},Q} \|g\|_{\mathcal{B},Q}. \tag{2.1}$$

Remark 2.1. Let us observe that when $\mathcal{D}(t) = t$, which gives L^1, then $\overline{\mathcal{D}}(t) = 0$ if $s \leq 1$ and $\overline{\mathcal{D}}(t) = \infty$ otherwise. Although $\overline{\mathcal{D}}$ is not a Young function one can see that the space $L^{\overline{\mathcal{D}}}$ coincides with L^∞. On the other hand, as the (generalized) inverse is $\overline{\mathcal{D}}^{-1}(t) \equiv 1$, the previous Hölder inequalities make sense with the appropriate changes if one of the three functions is $\mathcal{D}$ or $\overline{\mathcal{D}}$. We will use this throughout the paper.

Remark 2.2. The convexity of $\mathcal{A}$ implies that $\mathcal{A}(t)/t$ is increasing and thus $t \leq C\mathcal{A}(t)$ for all $t \geq 1$. This yields $\|f\|_{L^1,B} \leq C\|f\|_{\mathcal{A},B}$ for all Young functions $\mathcal{A}$.

2.2. *Muckenhoupt weights.* We recall the definition of the Muckenhoupt classes A_p, $1 \leq p \leq \infty$. Let w be a nonnegative locally integrable function and $1 \leq p < \infty$. We say that $w \in A_p$ if there exists $C_p < \infty$ such that for every ball $B \subset \mathbb{R}^n$,

$$\left(\frac{1}{|B|}\int_B w(x)\,dx\right)\left(\frac{1}{|B|}\int_B w(x)^{1-p'}\,dx\right)^{p-1} \le C_p,$$

when $1 < p < \infty$, and for $p = 1$,

$$\frac{1}{|B|}\int_B w(y)\,dy \le C_1 w(x) \quad \text{for a.e. } x \in B,$$

which can be equivalently written as $Mw(x) \le C_1 w(x)$ for a.e. $x \in \mathbb{R}^n$. Finally, we set $A_\infty = \bigcup_{p\ge 1} A_p$. It is well known that the Muckenhoupt classes characterize the boundedness of the Hardy–Littlewood maximal function on weighted Lebesgue spaces. Namely, $w \in A_p$, $1 < p < \infty$, if and only if M is bounded on $L^p(w)$; and $w \in A_1$ if and only if M maps $L^1(w)$ into $L^{1,\infty}(w)$.

In $\mathbb{R}$, the weighted estimates for the one-sided Hardy–Littlewood maximal function M^+ (and analogously for M^-) are characterized by the classes A_p^+ which are defined as follows. Given $1 < p < \infty$, $w \in A_p^+$ if there exists a constant $C_p < \infty$ such that for all $a < b < c$,

$$\frac{1}{(c-a)^p}\left(\int_a^b w(x)\,dx\right)\left(\int_b^c w(x)^{1-p'}\,dx\right)^{p-1} \le C_p.$$

We say that $w \in A_1^+$ if $M^- w(x) \le C_1 w(x)$ for a.e. $x \in \mathbb{R}$. The class A_∞^+ is defined as the union of all the A_p^+ classes, $A_\infty^+ = \bigcup_{p\ge 1} A_p^+$. The classes A_p^- are defined in a similar way. It is interesting to note that $A_p = A_p^+ \cap A_p^-$, $A_p \subsetneq A_p^+$ and $A_p \subsetneq A_p^-$. See [35], [22], [23], [24] for more definitions and results.

2.3. *Singular integral operators and Hörmander type conditions.* Let T be a singular integral operator of convolution type, that is, T is bounded on $L^2(\mathbb{R}^n)$ and

$$Tf(x) = \text{p.v.} \int_{\mathbb{R}^n} K(x-y) f(y)\,dy$$

where K is a measurable function defined away from 0. Convolution operators are considered for simplicity, but the results presented here can be stated for variable kernels with appropriate changes. The precise statements and the details are left to the reader.

When $n = 1$ and we further assume that the kernel K is supported on $(-\infty, 0)$ we say that T is a *one-sided singular integral* and we write T^+ to emphasize it. The results that we present below for (regular) singular integrals apply to T^+. However, taking advantage of the extra assumption on the kernel, one can be more precise and get better estimates (see Section 3.3).

We introduce various Hörmander type conditions on the kernel K. The weakest one is the so-called Hörmander condition H_1 (we write $K \in H_1$ or say that K satisfies the L^1-*Hörmander condition*): there are constants $c > 1$ and $C > 0$ such that

$$\int_{|x|>c|y|} |K(x-y) - K(x)|\,dx \le C, \qquad y \in \mathbb{R}^n.$$

The strongest one is the classical Lipschitz condition called H_∞^* (this notation is not standard but we keep H_∞ for a weaker L^∞-condition, see the definition below). We say that $K \in H_\infty^*$ if there are $\alpha, C > 0$ and $c > 1$ such that

$$|K(x-y) - K(x)| \le C\,\frac{|y|^\alpha}{|x|^{\alpha+n}}, \qquad |x| > c|y|.$$

Between H_1 and H_∞^* one finds the L^r-*Hörmander conditions* (which are called $H_{L^r} = H_r$ in the definition below). These classes appeared implicitly in [17] where it is shown that the classical L^r-Dini condition for K implies $K \in H_r$ (see also [34] and [37]). However, there are examples of singular integrals like the differential transform operator from ergodic theory defined in (1.2), whose kernel K is in H_r for all $1 \le r < \infty$ but $K \notin H_\infty$. As shown in [18], K satisfies a Hörmander condition in the scale of Orlicz spaces that lies between the intersection of the classes H_r for $1 \le r < \infty$ and H_∞. The same happens with the one-sided discrete square function considered in [36] and [20]. All these things have motivated the definition of the $L^{\mathcal{A}}$-Hörmander conditions in [20]:

Definition 2.3. The kernel K is said to satisfy the $L^{\mathcal{A}}$-*Hörmander condition*, and we write $K \in H_{\mathcal{A}}$, if there exist $c \ge 1$ and $C > 0$ such that for any $y \in \mathbb{R}^n$ and $R > c|y|$,

$$\sum_{m=1}^{\infty} (2^m R)^n \|K(\cdot - y) - K(\cdot)\|_{\mathcal{A},|x|\sim 2^m R} \le C.$$

We say that $K \in H_\infty$ if K satisfies the above condition with $\|\cdot\|_{L^\infty,|x|\sim 2^m R}$ in place of $\|\cdot\|_{\mathcal{A},|x|\sim 2^m R}$.

We have used the notation $|x| \sim s$ for $s < |x| \le 2s$, and

$$\|f\|_{\mathcal{A},|x|\sim s} = \|f\chi_{\{|x|\sim s\}}\|_{\mathcal{A},B(0,2s)}.$$

Note that if $\mathcal{A}(t) = t$ then $H_{\mathcal{A}} = H_1$. On the other hand, since $t \le C\mathcal{A}(t)$ for $t \ge 1$, we have $H_{\mathcal{A}} \subset H_1$, which implies that the classical unweighted Calderón–Zygmund theory can be applied to T. Also, it is easy to see that $H_\infty^* \subset H_\infty \subset H_{\mathcal{A}}$. For convenience, throughout the paper we write $|\cdot| = |\cdot|_\infty$ so that everything is adapted to cubes in place of balls (with appropriate changes everything can be written in terms of balls). For simplicity we also assume that $c = 1$.

Coifman type estimates were proved for kernels in these classes in [20]:

THEOREM 2.4 ([20]). *Let $\mathcal{A}$ be a Young function and let T be a singular integral operator with kernel $K \in H_{\mathcal{A}}$. Then for any $0 < p < \infty$ and $w \in A_\infty$,*

$$\int_{\mathbb{R}^n} |Tf(x)|^p w(x)\,dx \le C \int_{\mathbb{R}^n} M_{\bar{\mathcal{A}}} f(x)^p w(x)\,dx, \qquad f \in L^\infty_c, \tag{2.2}$$

whenever the left-hand side is finite.

Note that this improves the previous results in [17], [34] and [37] (for sharpness issues see also [21]). Similar results are also proved for vector-valued and one-sided operators (see [20]).

REMARK 2.5. Abusing notation, as in Remark 2.1, if $K \in H_\infty$, then (2.2) holds with $M_{\bar{\mathcal{A}}} = M$, where $\bar{\mathcal{A}}(t) = t$. This was obtained in [21] improving the corresponding result for the smaller class H^*_∞.

The previous estimates are useful in applications, as T and $M_{\bar{\mathcal{A}}}$ have a similar behavior (see [13]). For instance, two-weight estimates can be proved in the following way:

THEOREM 2.6 ([18]). *Let $\mathcal{A}$ be a Young function and $1 < p < \infty$. Suppose that there exist Young functions $\mathcal{D}$, $\mathcal{E}$ such that $\mathcal{E} \in B_{p'}$ and $\mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \le \bar{\mathcal{A}}^{-1}(t)$ for $t \ge t_0 > 0$. Set $\mathcal{D}_p(t) = \mathcal{D}(t^{1/p})$. Let T be a linear operator whose adjoint T^* satisfies* (2.2). *Then for any weight u,*

$$\int_{\mathbb{R}^n} |Tf(x)|^p u(x)\,dx \le C \int_{\mathbb{R}^n} |f(x)|^p M_{\mathcal{D}_p} u(x)\,dx. \tag{2.3}$$

REMARK 2.7. Abusing notation, the previous result contains the case $\bar{\mathcal{A}}(t) = t$ for which in (2.2) one has $M_{\bar{\mathcal{A}}} = M$. Then $\mathcal{D}$ and $\mathcal{E}$ are conjugate functions and so (2.3) holds for any $\mathcal{D}_p$ such that $\overline{\mathcal{D}} \in B_{p'}$. In particular, in (2.3) we can take the pair of weights $(u, M_{L(\log L)^{p-1+\delta}} u)$ for any $\delta > 0$: pick $\mathcal{D}(t) = t^p(1 + \log^+ t)^{p-1+\delta}$ whose conjugate function is $\overline{\mathcal{D}}(t) \approx t^{p'}/(1 + \log^+ t)^{1+\delta(p'-1)} \in B_{p'}$.

3. Statements of the main results

3.1. *Singular integral operators.* We are going to obtain endpoint two-weight norm inequalities for singular integral operators where various Hörmander conditions are assumed on the kernel. Namely, we look for the following weak-type $(1,1)$ estimates with pairs of weights (u, Su) where S will be a certain maximal function depending on the smoothness of the kernel:

$$u\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} \le \frac{C}{\lambda} \int_{\mathbb{R}^n} |f(x)|\,Su(x)\,dx. \tag{3.1}$$

THEOREM 3.1. *Let T be a singular integral operator with kernel K.*

(a) *Let $\mathcal{A}$ be a Young function such that $\bar{\mathcal{A}} \in \Delta_2$ and assume that there exists $r > 1$ so that $\liminf_{t\to\infty} \bar{\mathcal{A}}(t)/t^r > 0$. If $K \in H_{\mathcal{A}}$ then* (3.1) *holds for the pairs of weights $(u, M_{\bar{\mathcal{A}}}u)$.*

(b) *Let $\mathcal{A}$ be a Young function and assume that there exist $1 < p < \infty$, and Young functions $\mathcal{D}$ and $\mathcal{E}$ such that $\mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \le \bar{\mathcal{A}}^{-1}(t)$ for $t \ge t_0 > 0$ with $\mathcal{E} \in B_{p'}$. If $K \in H_{\mathcal{A}}$, then* (3.1) *holds for the pairs of weights $(u, M_{\mathcal{D}_p}u)$ with $\mathcal{D}_p(t) = \mathcal{D}(t^{1/p})$.*

(c) *If $K \in H_\infty$, then* (3.1) *holds for the pairs of weights $(u, M_{L(\log L)^\varepsilon}u)$ for any $\varepsilon > 0$.*

REMARK 3.2. In part (c) we improve [32], as we consider a wider class of kernels (recall that $H_\infty^* \subsetneq H_\infty$).

REMARK 3.3. Let us notice that when $\liminf_{t\to\infty} \bar{\mathcal{A}}(t)/t^r > 0$, the pair of weights in (a) is better than the one in (b): one can see that $\bar{\mathcal{A}}(t) \lesssim \mathcal{D}_p(t)$ for $t \ge 1$. Take an arbitrary $t \ge 1$. The fact that $\mathcal{E} \in B_{p'}$ implies $\mathcal{E}(t) \lesssim t^{p'}$. Also, $\bar{\mathcal{A}}(t) \ge t$ as $\bar{\mathcal{A}}$ is a Young function. Then, from the condition assumed on $\bar{\mathcal{A}}$, $\mathcal{D}$ and $\mathcal{E}$ it follows that $\mathcal{D}^{-1}(t) \lesssim t^{1/p}$ and therefore

$$\bar{\mathcal{A}}^{-1}(t) \ge \mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \gtrsim \mathcal{D}^{-1}(t)^p t^{1/p'}/\mathcal{D}^{-1}(t)^{p-1} \gtrsim \mathcal{D}^{-1}(t)^p = \mathcal{D}_p^{-1}(t).$$

REMARK 3.4. We emphasize that in part (a) the associated Coifman estimate in Theorem 2.4 tells us that T is controlled by $M_{\bar{\mathcal{A}}}$. Here we show that the pair of weights of the form $(u, M_{\bar{\mathcal{A}}}u)$ is suitable. In the previous remark, we have observed that in (b) one gets a bigger maximal operator $M_{\mathcal{D}_p}$. In many applications, even if we take p very close to 1, we always obtain a maximal operator pointwise greater than $M_{\bar{\mathcal{A}}}$. This is the case in (c) which covers the classical Hilbert and Riesz transforms. Here, as these operators are controlled by M (in the sense of Coifman), one would wish to show that the pair of weights (u, Mu) is suitable. However, this remains an open question and the best known result is $(u, M_{L(\log L)^\varepsilon}u)$.

There is a general extrapolation principle that allows one to pass from pairs of weights (u, Su), with S being a maximal operator, to general pairs of weights (u, v). The main ideas are implicit in [11], [12] and are further exploited in [10]. Below we present a proof in the one-sided case (see Theorem 3.14), which can be easily adapted to the present situation.

THEOREM 3.5. *Let $\mathcal{F}$ be a Young function and assume that a given operator T satisfies*

$$u\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} \le \frac{C}{\lambda} \int_{\mathbb{R}^n} |f(x)| M_{\mathcal{F}}u(x)\, dx \tag{3.2}$$

for every weight u and $\lambda > 0$. Given $1 < p < \infty$, let $\mathcal{G}$, $\mathcal{H}$ be Young functions *such that* $\mathcal{G}^{-1}(t)\mathcal{H}^{-1}(t) \le \mathcal{F}^{-1}(t)$ *for all* $t \ge t_0 > 0$ *and* $\mathcal{H} \in B_{p'}$. *Then for any pair of weights* (u, v) *satisfying*

$$\|u^{1/p}\|_{\mathcal{G},Q}\|v^{-1/p}\|_{L^{p'},Q} \le C \tag{3.3}$$

and for any $\lambda > 0$ *we have*

$$u\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} \le \frac{C}{\lambda^p} \int_{\mathbb{R}^n} |f(x)|^p v(x)\, dx. \tag{3.4}$$

3.2. *Commutators with* BMO *functions.* We are going to consider commutators of singular integral operators with BMO functions. Let us recall that a locally integrable function b is in BMO if

$$\|b\|_{\mathrm{BMO}} = \sup_Q \frac{1}{|Q|} \int_Q |b(x) - b_Q|\, dx < \infty,$$

where the sup is taken over all cubes $Q \subset \mathbb{R}^n$ with sides parallel to the coordinate axes and where b_Q stands for the average of b over Q.

We define the (first-order) commutator by

$$T_b^1 f(x) = [b, T]f(x) = b(x)Tf(x) - T(bf)(x).$$

The higher order commutators T_b^k are defined by induction as $T_b^k = [b, T_b^{k-1}]$ for $k \ge 2$. Note that for every $k \ge 1$,

$$T_b^k f(x) = \text{p.v.} \int_{\mathbb{R}^n} (b(x) - b(y))^k K(x - y) f(y)\, dy.$$

For $k = 0$ we understand that $T_b^0 = T$.

In [18], Coifman's type estimates were proved for commutators of singular integral operators with kernels in the following Hörmander classes that depend on the order of the commutator.

Definition 3.6. Let $\mathcal{A}$ be a Young function and $k \in \mathbb{N}$. We say that the kernel K satisfies the $L^{\mathcal{A},k}$-*Hörmander condition*, and write $K \in H_{\mathcal{A},k}$, if there exist $c \ge 1$ and $C > 0$ (depending on $\mathcal{A}$ and k) such that for all $y \in \mathbb{R}^n$ and $R > c|y|$,

$$\sum_{m=1}^{\infty} (2^m R)^n m^k \|K(\cdot - y) - K(\cdot)\|_{\mathcal{A},|x|\sim 2^m R} \le C.$$

We say that $K \in H_{\infty,k}$ if K satisfies this condition with $\|\cdot\|_{L^\infty,|x|\sim 2^m R}$ in place of $\|\cdot\|_{\mathcal{A},|x|\sim 2^m R}$.

As before, for simplicity we will assume that $c = 1$. For these classes the following Coifman estimates are obtained:

THEOREM 3.7 ([18]). *Let $b \in \mathrm{BMO}$, and $k \geq 0$.*

(a) *Let $\mathcal{A}$, $\mathcal{B}$ be Young functions such that $\bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_k^{-1}(t) \leq t$ for $t \geq t_0 > 0$ with $\bar{\mathcal{C}}_k(t) = e^{t^{1/k}} - 1$. If T is a singular integral operator with kernel $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$ (or, in particular, $K \in H_{\mathcal{B},k}$), then for any $0 < p < \infty$ and $w \in A_\infty$,*

$$\int_{\mathbb{R}^n} |T_b^k f(x)|^p w(x)\,dx \leq C\|b\|_{\mathrm{BMO}}^{pk} \int_{\mathbb{R}^n} M_{\bar{\mathcal{A}}} f(x)^p w(x)\,dx, \qquad f \in L_c^\infty, \tag{3.5}$$

whenever the left-hand side is finite.

(b) *If $K \in H_\infty \cap H_{e^{t^{1/k}},k}$ (or, in particular, $K \in H_{\infty,k}$) then (3.5) holds with M^{k+1}, the $(k+1)$th iteration of M, in place of $M_{\bar{\mathcal{A}}}$.*

This result and Theorem 2.6 can be used to derived endpoint estimates of the form

$$u\{x \in \mathbb{R}^n : |T_b^k f(x)| > \lambda\} \leq C \int_{\mathbb{R}^n} \mathcal{C}_k(\|b\|_{\mathrm{BMO}}^k |f(x)|/\lambda) Su(x)\,dx, \tag{3.6}$$

where $\mathcal{C}_k(t) = t(1 + \log^+ t)^k$.

THEOREM 3.8. *Let T be a singular integral operator with kernel K, $k \in \mathbb{N}$, $b \in \mathrm{BMO}$ and let T_b^k be the kth order commutator of T.*

(a) *Let $\mathcal{A}$, $\mathcal{B}$ be Young functions such that $\bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_k^{-1}(t) \leq t$ for $t \geq t_0 > 0$ with $\bar{\mathcal{C}}_k(t) = e^{t^{1/k}} - 1$. Let $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$ (or, in particular, $K \in H_{\mathcal{B},k}$).*

  (a.1) *If $\bar{\mathcal{A}} \in \Delta_2$ and there exists $r > 1$ with $\liminf_{t\to\infty} \bar{\mathcal{A}}(t)/t^r > 0$, then (3.6) holds for the pairs of weights $(u, M_{\bar{\mathcal{A}}} u)$.*

  (a.2) *Assume that there exist $1 < p < \infty$ and Young functions $\mathcal{D}$ and $\mathcal{E}$ such that $\mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \leq \bar{\mathcal{A}}^{-1}(t)$ for $t \geq t_0 > 0$ with $\mathcal{E} \in B_{p'}$. Then (3.6) holds for the pairs of weights $(u, M_{\mathcal{D}_p} u)$ with $\mathcal{D}_p(t) = \mathcal{D}(t^{1/p})$.*

(b) *If $K \in H_\infty \cap H_{e^{t^{1/k}},k}$ (or, in particular, $K \in H_{\infty,k}$), then (3.6) holds for the pairs of weights $(u, M_{L(\log L)^{k+\varepsilon}} u)$ for any $\varepsilon > 0$.*

REMARK 3.9. In part (b) we obtain the same result as in [30], but considering a weaker condition on the kernel K, since $H_\infty^* \subsetneq H_\infty \cap H_{e^{t^{1/k}},k}$.

REMARK 3.10. Notice that we can view (b) as an extension of (a) when $\mathcal{B}$ corresponds to L^∞ and so $\mathcal{A}(t) = \bar{\mathcal{C}}_k(t) = e^{t^{1/k}} - 1$. Observe that in that case we also have $\bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_k^{-1}(t) \leq t$ (where $\mathcal{B}^{-1}(t) \equiv 1$).

This result can be extended to the multilinear commutators considered in [31]. Given $k \geq 1$, a singular integral operator T with kernel K, and a vector $\vec{b} = (b_1, \ldots, b_k)$ of locally integrable functions, the multilinear commutator is defined as

$$T_{\vec{b}}f(x) = \int_{\mathbb{R}^n} \Big(\prod_{l=1}^{k}(b_l(x) - b_l(y))\Big) K(x,y) f(y)\, dy.$$

When $k = 0$ we understand that $T_{\vec{b}} = T$. Notice that if $k = 1$ and $\vec{b} = b$ then $T_{\vec{b}} = T^1_b$. For $k \geq 1$ if $b_1 = \cdots = b_k = b$ then $T_{\vec{b}} = T^k_b$.

For standard commutators, one assumes that $b \in \mathrm{BMO}$, and by John–Nirenberg's inequality we have $\|b\|_{\mathrm{BMO}} \approx \sup_Q \|b - b_Q\|_{\exp L, Q}$. This can be seen as the supremum of the oscillations of b on the space $\exp L$. As in [31], when dealing with multilinear commutators, the symbols b_j are assumed to be in one of these oscillation spaces. Given $s \geq 1$ we set

$$\|f\|_{\mathrm{Osc}(\exp L^s)} = \sup_Q \|f - f_Q\|_{\exp L^s, Q},$$

and the space $\mathrm{Osc}(\exp L^s)$ is the set of measurable functions $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ such that $\|f\|_{\mathrm{Osc}(\exp L^s)} < \infty$. Let us notice that $\mathrm{Osc}(\exp L^s) \subset \mathrm{Osc}(\exp L^1) = \mathrm{BMO}$. We assume that for each $1 \leq l \leq k$, $b_l \in \mathrm{Osc}(\exp L^{s_l})$ with $s_l \geq 1$, and we set $1/s = 1/s_1 + \cdots + 1/s_k$. For these commutators in [18, Theorem 7.1] it is shown that under the previous conditions, if $K \in H_{\mathcal{B},k}$ and $\overline{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\overline{\mathcal{C}}^{-1}_{1/s}(t) \leq t$ with $\overline{\mathcal{C}}_{1/s}(t) = e^{t^s}$, then $T_{\vec{b}}$ satisfies a Coifman estimate with $M_{\overline{\mathcal{A}}}$ on the right-hand side. In the case $K \in H_{\infty,k}$, the maximal operator is $M_{L(\log L)^{1/s}}$. In this way, we can extend Theorem 3.8 to the multilinear commutators: in (a) we assume $K \in H_{\mathcal{B},k}$ and replace k by $1/s$, and in (b) we assume $K \in H_{\infty,k}$ and replace k by $1/s$. The precise formulation is left to the interested reader. The proof of this result follows the same scheme (see Remark 5.2 below).

3.3. *One-sided operators.* In $\mathbb{R}$ we can consider a smaller class of operators and obtain estimates for the so-called one-sided operators. These are singular integral operators with kernels supported on $(-\infty, 0)$ and we write T^+ to emphasize it. One can also consider operators T^- with kernels supported on $(0, \infty)$; for simplicity we restrict ourselves to the first type.

As one-sided operators are singular integral operators, the previous results can be applied to them. For instance, if $K \in H^*_\infty$ (indeed $K \in H_\infty$ suffices) then T^+ can be controlled by M on $L^p(w)$ for every $0 < p < \infty$ and $w \in A_\infty$, and consequently T^+ is bounded on $L^p(w)$ for every $w \in A_p$, $1 < p < \infty$. These follow from the classical theory of Calderón–Zygmund operators. However, exploiting the fact that the kernel of T^+ is supported on $(-\infty, 0)$ one can do better: in the Coifman estimate we can use the pointwise smaller one-sided maximal operator M^+ and consider a bigger class of weights $w \in A^+_\infty$; thus T^+ is bounded on $L^p(w)$ for every $w \in A^+_p$, $1 < p < \infty$ (note that $A_p \subsetneq A^+_p$).

The same happens with Theorems 2.4, 3.7 and 2.6:

THEOREM 3.11 ([20], [18]). *Let T^+ be a one-sided singular integral operator with kernel K supported in $(-\infty,0)$.*

(i) *Under the assumptions of Theorem* 2.4 *or* 3.7, *one can improve* (2.2) *and* (3.5): *A_∞ can be replaced by the bigger class of weights A_∞^+, and $M_{\bar{\mathcal{A}}}$ by the pointwise smaller operator $M_{\bar{\mathcal{A}}}^+$; in* (b) *of Theorem* 3.7, *M^{k+1} can be replaced by $(M^+)^{k+1}$.*

(ii) *Under the assumptions of Theorem* 2.6, *if the adjoint of T^+ (which is a one-sided operator with kernel supported on $(0,\infty)$) satisfies* (2.2) *for all $0<p<\infty$, $w\in A_\infty^-$ and with $M_{\bar{\mathcal{A}}}^- f$ on the right-hand side, then, for any weight u, T^+ satisfies* (2.3) *with $M_{\mathcal{D}_p}^- u$ in place of $M_{\mathcal{D}_p}u$.*

Here, we can obtain one-sided versions of Theorems 3.1 and 3.8:

THEOREM 3.12. *Let T^+ be a singular integral operator with kernel K supported in $(-\infty,0)$.*

(i) *Under the assumptions of Theorem* 3.1, *T^+ satisfies* (3.1) *for the pairs of weights $(u, M_{\bar{\mathcal{A}}}^- u)$ in* (a), *$(u, M_{\mathcal{D}_p}^- u)$ in* (b), *and $(u, M_{L(\log L)^\varepsilon}^- u)$ in* (c).

(ii) *Under the assumptions of Theorem* 3.8, *$T_b^{+,k}$ (the kth order commutator of T^+) satisfies* (3.6) *for the pairs of weights $(u, M_{\bar{\mathcal{A}}}^- u)$ in* (a.1), *$(u, M_{\mathcal{D}_p}^- u)$ in* (a.2), *and $(u, M_{L(\log L)^{k+\varepsilon}}^- u)$ in* (b).

REMARK 3.13. In (i) when $K\in H_\infty$ we improve the results in [1] where the stronger condition $K\in H_\infty^*$ was assumed. For examples of such kernels see [1].

We can also get an improvement of the estimates in Theorem 3.5 when we start with pairs based on one-sided maximal functions:

THEOREM 3.14. *Let $\mathcal{F}$ be a Young function and assume that an operator T satisfies* (3.2) *with $M_{\mathcal{F}}^-$ in place of $M_{\mathcal{F}}$. Let $1<p<\infty$, and let $\mathcal{G}$, $\mathcal{H}$ be as in Theorem* 3.5. *If (u,v) is a pair of weights such that, for all $a<b<c$ with $b-a<c-b$,*

$$\|u^{1/p}\|_{\mathcal{G},(a,b)}\|v^{-1/p}\|_{L^{p'},(b,c)}\le C,$$

then for all $\lambda>0$,

$$u\{x\in\mathbb{R}:|Tf(x)|>\lambda\}\le\frac{C}{\lambda^p}\int_{\mathbb{R}}|f(x)|^p v(x)\,dx.$$

Let us notice that here one does not need to work with one-sided operators as this abstract result does not use any property of T but the initial two-weight estimate which involves the one-sided maximal function $M_{\mathcal{F}}^-$.

Notice that when applying this result, T will be T^+ or $T_b^{+,k}$ from Theorem 3.12.

To prove Theorem 3.14 we need to find sufficient conditions on (u,v) that guarantee the boundedness of $M_{\mathcal{F}}^-$ from $L^p(v)$ to $L^p(u)$. This result with $M_{\mathcal{F}}$ appears in [11] and here we extend it to the one-sided case. For convenience we state it in terms of $M_{\mathcal{F}}^+$; to pass to $M_{\mathcal{F}}^-$ one just switches the intervals of integration in the corresponding Muckenhoupt type condition.

THEOREM 3.15. *Let $1 < p < \infty$ and let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be Young functions such that $\mathcal{B}^{-1}(t)\mathcal{C}^{-1}(t) \le \mathcal{A}^{-1}(t)$ for all $t \ge t_0 > 0$, with $\mathcal{C} \in B_p$. If (u,v) is a pair of weights such that, for all $a < b < c$ with $b-a < c-b$,*

$$\|u^{1/p}\|_{L^p,(a,b)}\|v^{-1/p}\|_{\mathcal{B},(b,c)} \le C, \tag{3.7}$$

then

$$\int_{\mathbb{R}} M_{\mathcal{A}}^+ f(x)^p u(x)\,dx \le C \int_{\mathbb{R}} |f(x)|^p v(x)\,dx.$$

4. Applications. In this section we present some applications. As we have already observed, our results include those in [27] and [30] for Calderón–Zygmund singular integrals operators with kernels in H_∞^*. We observed before that weaker conditions on the kernels, say H_∞ for T and $K \in H_\infty \cap H_{e^{t^{1/k}},k}$ (or, in particular, $K \in H_{\infty,k}$) for T_b^k, lead us to the same conclusions.

4.1. *The differential transform operator.* Let us consider the differential transform operator studied in [16] and [4],

$$T^+ f(x) = \sum_{j\in\mathbb{Z}} \nu_j (D_j f(x) - D_{j-1} f(x)), \tag{4.1}$$

where $\|\{\nu_j\}_j\|_\infty < \infty$ and

$$D_j f(x) = \frac{1}{2^j} \int_x^{x+2^j} f(t)\,dt.$$

This operator appears when studying the rate of convergence of the averages $D_j f$. Observe that $D_j f \to f$ a.e. as $j \to -\infty$, and $D_j f \to 0$ as $j \to \infty$. Notice that $T^+ f(x) = K * f(x)$, where

$$K(x) = \sum_{j\in\mathbb{Z}} \nu_j \left(\frac{1}{2^j} \chi_{(-2^j,0)}(x) - \frac{1}{2^{j-1}} \chi_{(-2^{j-1},0)}(x) \right).$$

Since K is supported in $(-\infty,0)$, T^+ is a one-sided singular integral operator (so we write T^+). In [4] it is proved that, for appropriate functions f,

$$T^+ f(x) = \lim_{(N_1,N_2)\to(-\infty,\infty)} \sum_{j=N_1}^{N_2} \nu_j (D_j f(x) - D_{j-1} f(x)) \quad \text{for a.e. } x \in \mathbb{R}.$$

It is known that $K \in \bigcap_{r\ge 1} H_r$ and so T^+ is bounded on $L^p(w)$ for all $w \in A_p^+$, $1 < p < \infty$, and maps $L^1(w)$ into $L^{1,\infty}(w)$ for all $w \in A_1^+$.

When trying to prove Coifman type estimates for T^+, one finds that T^+ is controlled by M_s^+ for every $1 < s < \infty$. In general $K \notin H_\infty$ (see [18] for the case $\{\nu_j\} = \{(-1)^j\}$), thus it is not clear whether one can take $s = 1$, that is, whether T^+ behaves as a one-sided singular integral operator with smooth kernel. This motivated the new Hörmander type conditions in [20], [18]: If one shows that K belongs to some class near L^∞ then one would obtain a maximal operator near M^+. In [18] it was shown that $K \in H_{e^{t^{1/(1+\varepsilon)}}}$ for any $\varepsilon > 0$, and $K \in H_{e^{t^{1/(1+k+\varepsilon)}},k}$ for any $\varepsilon > 0$ and $k \ge 1$. Thus, by Theorems 2.4 and 3.7 for any $k \ge 0$, $\varepsilon > 0$, $0 < p < \infty$, and $w \in A_\infty^+$,

$$\int_{\mathbb{R}} |T_b^{+,k} f(x)|^p w(x)\,dx \le C \int_{\mathbb{R}} M^+_{L(\log L)^{k+1+\varepsilon}} f(x)^p w(x)\,dx.$$

Applying Theorem 3.12 we obtain the following endpoint estimates:

Theorem 4.1. *Let $b \in$ BMO and $k \ge 0$. Let T^+ be the differential transform operator defined above, and let $T_b^{+,k}$ be its kth order commutator. Then, for any $\varepsilon > 0$,*

$$u\{x \in \mathbb{R} : |T_b^{+,k} f(x)| > \lambda\} \le C \int_{\mathbb{R}} \mathcal{C}_k(|f(x)|/\lambda) M^-_{L(\log L)^{k+1+\varepsilon}} u(x)\,dx$$

for all $\lambda > 0$.

Note that this result includes the case $k = 0$ for which $T_b^{+,0} = T^+$.

Remark 4.2. One can write the last estimate in terms of iterations of M^- since $M^-_{L(\log L)^{k+1+\varepsilon}} u(x) \le C(M^-)^{k+3} u(x)$ for $\varepsilon > 0$ small enough. Thus, the previous estimate holds for the pair of weights $(u, (M^-)^{k+3} u)$.

Proof of Theorem 4.1. Given $k \ge 0$ and $\varepsilon > 0$ we fix $\mathcal{D}(t) = t^p(1 + \log^+ t)^{k+1+\varepsilon}$ so that $\mathcal{D}_p(t) \approx t(1 + \log^+ t)^{k+1+\varepsilon}$. We take $1 < p < 1 + \varepsilon/(2(k+2))$, $\mathcal{A}(t) \approx \exp(t^{1/(1+k+\varepsilon/(2p))}) - 1$ and $\mathcal{B}(t) \approx \exp(t^{1/(1+\varepsilon/(2p))}) - 1$. Then, as mentioned before, $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$ (note that for $k = 0$, we just have $K \in H_{\mathcal{A}}$). We also notice that for $k \ge 1$ it follows that $\overline{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\overline{\mathcal{C}}_k^{-1}(t) \lesssim t$ for $t \ge 1$. Next, we pick $\mathcal{E}(t) \approx t^{p'}/(1 + \log^+ t)^{\varepsilon(p'-1)/2-(k+1)}$ and observe that our choice of p guarantees that $\varepsilon(p'-1)/2 - (k+1) > 1$, therefore $\mathcal{E} \in B_{p'}$. Moreover, $\mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \lesssim \overline{\mathcal{A}}^{-1}(t)$ for $t \ge 1$. Then applying Theorem 3.12, that is, the one-sided version of Theorem 3.1(b) when $k = 0$, and the one-sided version of Theorem 3.8(a.2) when $k \ge 1$, we deduce the desired estimate. ∎

As a corollary of Theorem 4.1, applying Theorem 3.14 we get the following weak-type estimates for general pairs of weights (u, v):

COROLLARY 4.3. *Let $b \in \mathrm{BMO}$ and $k \geq 0$. Let T^+ be the differential transform operator defined above, and let $T_b^{+,k}$ be its kth order commutator. For any $\varepsilon > 0$, if (u,v) is a pair of weights such that, for all $a < b < c$ with $b - a < c - b$,*

$$\|u^{1/p}\|_{L^p(\log L)^{(k+2)p-1+\varepsilon},(a,b)}\|v^{-1/p}\|_{L^{p'},(b,c)} \leq C,$$

then for all $\lambda > 0$,

$$u\{x \in \mathbb{R} : |T_b^{+,k} f(x)| > \lambda\} \leq \frac{C}{\lambda^p} \int_{\mathbb{R}} |f(x)|^p v(x)\, dx.$$

This result follows at once from Theorem 3.14. The starting estimate is given by Theorem 4.1, thus $\mathcal{F}(t) = t(1 + \log^+ t)^{k+1+\varepsilon}$ for every $\varepsilon > 0$, and we take $\mathcal{H}(t) = t^{p'}/(1 + \log^+ t)^{1+\delta} \in B_{p'}$ for any $\delta > 0$. This leads to the desired function $\mathcal{G}$. Details are left to the interested reader.

4.2. *An example of a one-sided operator with $K \in H_\infty \cap H_{e^{t^{1/k}},k}$.* We consider the one-sided operator

$$T^+ f(x) = \sum_{j\in\mathbb{Z}} \nu_j (D_j f(x) - D_{j-1} f(x)),$$

where $\|\{\nu_j\}_j\|_\infty < \infty$ and

$$D_j f(x) = \frac{1}{2^j(1+j^2)} \int_x^{x+2^j} f(t)\, dt.$$

Observe that

$$K(x) = \sum_{j\in\mathbb{Z}} \nu_j \bigg(\frac{1}{2^j(1+j^2)}\, \chi_{(-2^j,0)}(x) - \frac{1}{2^{j-1}(1+(j-1)^2)}\, \chi_{(-2^{j-1},0)}(x) \bigg).$$

This operator is similar to the previous one. In [18] it was proved that $K \in H_\infty \cap H_{e^{t^{1/k}},k}$. Thus, by Theorems 2.4 and 3.7 for each $k \geq 0$, $0 < p < \infty$, and $w \in A_\infty^+$,

$$\int_{\mathbb{R}} |T_b^{+,k} f(x)|^p w(x)\, dx \leq C \int_{\mathbb{R}} M^+_{L(\log L)^k} f(x)^p w(x)\, dx. \tag{4.2}$$

Note that on the right-hand side one can alternatively write $(M^+)^{k+1} f$ as $(M^+)^{k+1} f \approx M^+_{L(\log L)^k} f$ a.e. We apply Theorem 3.12, that is, when $k = 0$ we use the one-sided version of Theorem 3.1(c), and when $k \geq 1$ we employ the one-sided version of Theorem 3.8(b). Thus, we deduce the following endpoint estimates: given $b \in \mathrm{BMO}$, for all $k \geq 0$ and $\varepsilon > 0$,

$$u\{x \in \mathbb{R} : |T_b^{+,k} f(x)| > \lambda\} \leq C \int_{\mathbb{R}} \mathcal{C}_k(|f(x)|/\lambda) M^-_{L(\log L)^{k+\varepsilon}} u(x)\, dx. \tag{4.3}$$

Note that taking $\varepsilon > 0$ small enough, $M^-_{L(\log L)^{k+\varepsilon}} u(x) \leq C(M^-)^{k+2} u(x)$.

REMARK 4.4. In terms of iterations of the one-sided Hardy–Littlewood maximal function, notice that in (4.2) we have $k+1$ iterations and in (4.3) we have $k+2$, so we obtain an extra iteration. This is because in Theorems 3.1(c) and 3.8(b) and their corresponding versions for one-sided operators we loose a small power of the logarithm. This also happens with Calderón–Zygmund operators with smooth kernel, like the Hilbert and Riesz transforms: they are controlled, in the sense of Coifman, by M, but the endpoint estimate holds for the pair of weights (u, M^2u)—indeed, one can write $(u, M_{L(\log L)^\varepsilon}u)$ for any $\varepsilon > 0$. It is not known, even for the Hilbert and Riesz transforms, whether the pair of weights (u, Mu) is suitable for the corresponding weak-type estimate.

Notice that in the case of the differential transform operator in both the Coifman inequality and the endpoint estimate the number of iterations for the kth order commutator is $k+3$. This happens as we already have a small power of the logarithm floating around.

From Theorem 3.14 proceeding as in Corollary 4.3 we obtain the following two-weight weak-type estimates: given $b \in \mathrm{BMO}$, for every $k \geq 0$ and for any $\varepsilon > 0$, if (u, v) is a pair of weights such that, for all $a < b < c$ with $b - a < c - b$,

$$\|u^{1/p}\|_{L^p(\log L)^{(k+1)p-1+\varepsilon},(a,b)} \|v^{-1/p}\|_{L^{p'},(b,c)} \leq C,$$

then $T_b^{+,k} f$ maps $L^p(v)$ into $L^{p,\infty}(u)$. This extends the sharp results obtained in [11] for Calderón–Zygmund operators with smooth kernels to the setting of one-sided operators.

4.3. *Multipliers.* Let $m \in L^\infty(\mathbb{R}^n)$ and consider the multiplier operator T defined *a priori* for f in the Schwartz class by $\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi)$. Given $1 < s \leq 2$ and $0 \leq l \in \mathbb{N}$ we say that $m \in M(s, l)$ if

$$\sup_{R>0} R^{|\alpha|} \|D^\alpha m\|_{L^s,|\xi|\sim R} < \infty, \qquad \text{for all } |\alpha| \leq l.$$

In [18] the following was proved. Let $m \in M(s, l)$ with $1 < s \leq 2$, $0 \leq l \leq n$ and $l > n/s$. Then for all $k \geq 0$, $\varepsilon > 0$, $0 < p < \infty$ and $w \in A_\infty$,

$$\int_{\mathbb{R}^n} |T_b^k f(x)|^p w(x)\, dx \leq C \int_{\mathbb{R}^n} M_{n/l+\varepsilon} f(x)^p w(x)\, dx. \tag{4.4}$$

The proof consists in showing that a family of truncations of the kernel $\{K^N\}_N$ are uniformly in $H_{L^r(\log L)^{kr},k}$ with $r' = n/l+\varepsilon$. Thus, taking $\mathcal{A}(t) = t^r$, $\mathcal{B}(t) = t^r(1+\log^+ t)^{kr}$ we have $K^N \in H_\mathcal{B} \cap H_{\mathcal{A},k}$ (this follows easily from $K^N \in H_{L^r(\log L)^{kr},k}$). Notice that $\overline{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\overline{\mathcal{C}}_k^{-1}(t) \lesssim t$ for $t \geq 1$ and therefore (4.4) follows from Theorem 3.7 for the kth order commutators of T^N (which is the operator whose kernel is K^N) with constants that are independent of N. A standard approximation argument leads to the desired estimate for T_b^k. We refer the reader to [18] for more details.

The same argument allows us to apply Theorems 3.1(a) and 3.8(a.1) to T^N. Observe that $\bar{\mathcal{A}}(t) = t^{r'}$; thus choosing $1 < s < r'$ we obtain $\liminf_{t\to\infty} \bar{\mathcal{A}}(t)/t^s = \infty$. Therefore, taking limits we have the following result:

THEOREM 4.5. *Let $m \in M(s,l)$ with $1 < s \le 2$, $0 \le l \le n$ and $l > n/s$. Then for all $k \ge 0$ and $\varepsilon > 0$,*

$$u\{x \in \mathbb{R}^n : |T_b^k f(x)| > \lambda\} \le C \int_{\mathbb{R}^n} \mathcal{C}_k(|f(x)|/\lambda) M_{n/l+\varepsilon} u(x)\, dx.$$

From this estimate one can obtain weak-type estimates for general pairs of weights by using Theorem 3.5. The precise statements are left to the reader.

4.4. *Kernels related to H_r and M_{L^r}.* As was implicit in [34] (see also [17], [37]), and observed in [21], if $K \in H_{L^r}$, that is, if the kernel satisfies the L^r-Hörmander condition, then T is controlled by $M_{L^{r'}}$. In [18] various extensions of that inequality for higher order commutators were considered. In the notation of Theorem 3.7, Table 1 lists the conditions and maximal operators obtained, for $1 < r < \infty$ and $k \ge 0$.

Table 1. Examples of H_r-conditions

$H_{\mathcal{B},k}$	$H_{\mathcal{B}} \cap H_{\mathcal{A},k}$	$M_{\bar{\mathcal{A}}} f$
$H_{L^r,k}$	$H_{L^r} \cap H_{L^r(\log L)^{-kr},k}$	$M_{L^{r'}(\log L)^{kr'}} f$
$H_{L^r(\log L)^{kr},k}$	$H_{L^r(\log L)^{kr}} \cap H_{L^r,k}$	$M_{L^{r'}} f$
$H_{L^r(\log L)^k,k}$	$H_{L^r(\log L)^k} \cap H_{L^r(\log L)^{-k(r-1)},k}$	$M_{L^{r'}(\log L)^k} f \approx (M_{L^{r'}})^{k+1}$

Thus, applying Theorems 3.1(a) and 3.8(a.1) we see that T_b^k satisfies (3.6) with the different pairs of weights $(u, M_{\bar{\mathcal{A}}} u)$ in the above table.

4.5. *Homogeneous singular integrals.* Denote by $\Sigma = \Sigma_{n-1}$ the unit sphere in $\mathbb{R}^n$. For $x \neq 0$, we write $x' = x/|x|$. Let $\Omega \in L^1(\Sigma)$. This function can be extended to $\mathbb{R}^n \setminus \{0\}$ as $\Omega(x) = \Omega(x')$ (abusing notation we call both functions Ω). Thus Ω is a function homogeneous of degree 0. We assume that $\int_\Sigma \Omega(x')\, d\sigma(x') = 0$. Set $K(x) = \Omega(x)/|x|^n$ and let T be the operator associated with the kernel K.

Given a Young function $\mathcal{A}$ we define the $L^{\mathcal{A}}$-*modulus of continuity* of Ω as

$$\varpi_{\mathcal{A}}(t) = \sup_{|y|\le t} \|\Omega(\cdot+y) - \Omega(\cdot)\|_{\mathcal{A},\Sigma}.$$

Fix $\Omega \in L^{\mathcal{B}}(\Sigma)$ and T as above. Let $k \ge 0$ and $\mathcal{A}$, $\mathcal{B}$ be Young functions such that $\overline{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\overline{\mathcal{C}}_k^{-1}(t) \le t$ for all $t \ge 1$. If

$$\int_0^1 \varpi_{\mathcal{B}}(t)\,\frac{dt}{t} + \int_0^1 \Big(1+\log\frac{1}{t}\Big)^k \varpi_{\mathcal{A}}(t)\,\frac{dt}{t} < \infty,$$

then it was proved in [18] that $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$ and therefore

$$\int_{\mathbb{R}^n} |T_b^k f(x)|^p w(x)\,dx \le C \int_{\mathbb{R}^n} M_{\overline{\mathcal{A}}} f(x)^p w(x)\,dx$$

for every $0 < p < \infty$ and $w \in A_\infty$.

Once it is known that $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$ one can apply Theorems 3.1 and 3.8 to derive the corresponding two-weight endpoint estimates. The precise statements and further details are left to the interested reader.

5. Proofs of the main results

Proof of Theorem 3.1. Without loss of generality we can assume that u is bounded and has compact support (otherwise we prove the corresponding estimate for $u_N = \min\{u, N\}\chi_{B(0,N)}$ with bounds independent of N and apply the monotone convergence theorem). We assume that $0 \le f \in L_c^\infty(\mathbb{R}^n)$ and consider the standard Calderón–Zygmund decomposition of f at level λ: there exists a collection of maximal (and therefore disjoint) dyadic cubes $\{Q_j\}_j$ (with center x_j and sidelength $2r_j$) such that

$$\lambda < \frac{1}{|Q_j|}\int_{Q_j} f \le 2^n\lambda. \tag{5.1}$$

We write $f = g + h$ where

$$g = f\chi_{\mathbb{R}^n\setminus\bigcup_j Q_j} + \sum_j f_{Q_j}\chi_{Q_j}, \qquad h = \sum_j h_j = \sum_j (f - f_{Q_j})\chi_{Q_j},$$

where f_{Q_j} denotes the average of f over Q_j. Recall that $0 \le g(x) \le 2^n\lambda$ a.e. and also that each h_j has vanishing integral. We set $\tilde{Q}_j = 2Q_j$, $\tilde{\Omega} = \bigcup_j \tilde{Q}_j$, and $\tilde{u} = u\chi_{\mathbb{R}^n\setminus\tilde{\Omega}}$. Then

$$\begin{aligned} u\{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} &\le u(\tilde{\Omega}) + u\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : |Th(x)| > \lambda/2\} \\ &\quad + u\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : |Tg(x)| > \lambda/2\} \\ &= I + II + III. \end{aligned}$$

We estimate each term separately. The estimates for I and II are obtained in the same way in the three cases (a)–(c). We show that

$$I \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^n} f(x)Mu(x)\,dx, \qquad II \lesssim \frac{1}{\lambda} \int_{\mathbb{R}^n} f(x)M_{\bar{\mathcal{A}}}u(x)\,dx, \tag{5.2}$$

where, in case (c), as $K \in H_\infty = H_{L^\infty}$ it is understood that $\bar{\mathcal{A}}(t) = t$, so $M_{\bar{\mathcal{A}}} = M_{L^1} = M$. Observe that both estimates lead us to the desired conclusions in the three cases (a)–(c). Regarding I, Mu is controlled by $M_{\bar{\mathcal{A}}}u$ in (a) (as $\bar{\mathcal{A}}$ is a Young function), by $M_{\mathcal{D}_p}u$ in (b) (since we pointed out in Remark 3.3 that $\mathcal{D}^{-1}(t) \lesssim t^{1/p}$ for $t \geq 1$, which yields $\mathcal{D}_p(t) \geq t$ for $t \geq 1$), and by $M_{L(\log L)^\varepsilon}u$ in (c). For II, $M_{\bar{\mathcal{A}}}u$ is the desired weight in (a); in (b) we observed in Remark 3.3 that $M_{\bar{\mathcal{A}}}u \lesssim M_{\mathcal{D}_p}u$; and in (c) we have $M_{\bar{\mathcal{A}}}u = Mu \leq M_{L(\log L)^\varepsilon}u$.

Let us show the first estimate in (5.2). By (5.1) we have

$$\begin{aligned} I = u\Big(\bigcup_j \tilde{Q}_j\Big) &\leq \sum_j u(\tilde{Q}_j) = 2^n \sum_j \frac{u(\tilde{Q}_j)}{|\tilde{Q}_j|}\,|Q_j| \leq \frac{2^n}{\lambda} \sum_j \frac{u(\tilde{Q}_j)}{|\tilde{Q}_j|} \int_{Q_j} f(x)\,dx \\ &\leq \frac{2^n}{\lambda} \sum_j \int_{Q_j} f(x)Mu(x)\,dx \leq \frac{2^n}{\lambda} \int_{\mathbb{R}^n} f(x)Mu(x)\,dx. \end{aligned}$$

Next, we estimate II: as the functions h_j have vanishing integral,

$$\begin{aligned} II &= u\Big\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : \Big|\sum_j Th_j(x)\Big| > \lambda/2\Big\} \leq \frac{2}{\lambda} \sum_j \int_{\mathbb{R}^n\setminus\tilde{\Omega}} |Th_j(x)|u(x)\,dx \\ &\leq \frac{2}{\lambda} \sum_j \int_{\mathbb{R}^n\setminus\tilde{\Omega}} \Big| \int_{Q_j} (K(x-y) - K(x-x_{Q_j}))h_j(y)\,dy\Big| u(x)\,dx \\ &\leq \frac{2}{\lambda} \sum_j \int_{Q_j} |h_j(y)| \int_{\mathbb{R}^n\setminus\tilde{Q}_j} |K(x-y) - K(x-x_{Q_j})|u(x)\,dx\,dy. \end{aligned}$$

We claim that for every $y \in Q_j$,

$$\int_{\mathbb{R}^n\setminus\tilde{Q}_j} |K(x-y) - K(x-x_{Q_j})|u(x)\,dx \lesssim \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}}u. \tag{5.3}$$

This estimate leads to

$$\begin{aligned} II &\lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}}u \int_{Q_j} |h_j(y)|\,dy \lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}}u \int_{Q_j} f(y)\,dy \\ &\leq \frac{1}{\lambda} \sum_j \int_{Q_j} f(y)M_{\bar{\mathcal{A}}}u(y)\,dy \leq \frac{1}{\lambda} \int_{\mathbb{R}^n} f(y)M_{\bar{\mathcal{A}}}u(y)\,dy. \end{aligned}$$

We obtain (5.3): using the generalized Hölder inequality for $\mathcal{A}$ and $\bar{\mathcal{A}}$ (when $K \in H_\infty$ we understand that $\bar{\mathcal{A}}(t) = t$ and so we have the corresponding

L^1-L^∞ Hölder estimate)

$$\begin{aligned}
&\int_{\mathbb{R}^n\setminus\tilde{Q}_j} |K(x-y)-K(x-x_{Q_j})|u(x)\,dx \\
&\quad\le \sum_{k=1}^{\infty} \int_{|x-x_{Q_j}|\sim 2^k r_j} |K(x-y)-K(x-x_{Q_j})|u(x)\,dx \\
&\quad\lesssim \sum_{k=1}^{\infty} (2^k r_j)^n \|K(\cdot-y)-K(\cdot-x_{Q_j})\|_{\mathcal{A},|x-x_{Q_j}|\sim 2^k r_j} \|u\|_{\bar{\mathcal{A}},|x-x_{Q_j}|\le 2^{k+1} r_j} \\
&\quad\le C \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}} u,
\end{aligned}$$

where in the last estimate we have used the fact that $K \in H_{\mathcal{A}}$.

To complete the proof, it remains to estimate III. Here, the proof is different in each case. We start with (a). As $\liminf_{t\to\infty} \bar{\mathcal{A}}(t)/t^r > 0$, there exists $c = c_r$ such that $\bar{\mathcal{A}}(t) \ge ct^r$ for every $t \ge 1$. On the other hand, since $\bar{\mathcal{A}} \in \Delta_2$ there exist $1 < s < \infty$ (indeed we can take $s > r$) such that $\bar{\mathcal{A}}(t) \le Ct^s$ for every $t \ge 1$ (this follows by iterating the Δ_2-condition). Then taking $p > s$ we have

$$\begin{aligned}
(5.4)\qquad III = u\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : |Tg(x)| > \lambda/2\} &\le \frac{2^p}{\lambda^p} \int_{\mathbb{R}^n} |Tg(x)|^p \tilde{u}(x)\,dx \\
&\le \frac{2^p}{\lambda^p} \int_{\mathbb{R}^n} |Tg(x)|^p M_r\tilde{u}(x)\,dx \lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^n} M_{\bar{\mathcal{A}}} g(x)^p M_r\tilde{u}(x)\,dx,
\end{aligned}$$

where in the last inequality we have used Theorem 2.4 and the fact that $M_r\tilde{u} \in A_1 \subset A_\infty$ as $r > 1$. Notice that one has to check that the left-hand side of (2.2) is finite. Indeed, as we have assumed that $u \in L^\infty$ we have

$$\begin{aligned}
\int_{\mathbb{R}^n} |Tg(x)|^p M_r\tilde{u}(x)\,dx &\le \|u\|_{L^\infty} \int_{\mathbb{R}^n} |Tg(x)|^p\,dx \lesssim \|u\|_{L^\infty} \int_{\mathbb{R}^n} g(x)^p\,dx \\
&\lesssim \|u\|_{L^\infty} \lambda^{p-1} \int_{\mathbb{R}^n} g(x)\,dx = \|u\|_{L^\infty} \lambda^{p-1} \int_{\mathbb{R}^n} f(x)\,dx < \infty,
\end{aligned}$$

where we have used that T is bounded on $L^p(\mathbb{R}^n)$ as $K \in H_{\mathcal{A}} \subset H_1$; and also that f and u are bounded with compact support. We can continue with the estimate of III: as $\bar{\mathcal{A}}(t) \le Ct^s$ for every $t \ge 1$ it follows that

$$\begin{aligned}
(5.5)\qquad III &\lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^n} M_s g(x)^p M_r\tilde{u}(x)\,dx = \frac{1}{\lambda^p} \int_{\mathbb{R}^n} M(g^s)(x)^{p/s} M_r\tilde{u}(x)\,dx \\
&\lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^n} g(x)^p M_r\tilde{u}(x)\,dx \lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^n} g(x)^p M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx,
\end{aligned}$$

where we have used that $M_r\tilde{u} \in A_1$, therefore M is bounded on $L^{p/s}(M_r\tilde{u})$

and also that $\bar{\mathcal{A}}(t) \geq ct^r$ for every $t \geq 1$. We claim that

$$\int_{\bigcup_j Q_j} g(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx \lesssim \int_{\bigcup_j Q_j} f(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx. \tag{5.6}$$

From (5.5), this estimate and the fact that $0 \leq g(x) \leq 2^n\lambda$ a.e. yield

$$\begin{aligned} III &\lesssim \frac{1}{\lambda}\int_{\mathbb{R}^n} g(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx \\ &= \frac{1}{\lambda}\int_{\mathbb{R}^n\setminus\bigcup_j Q_j} f(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx + \frac{1}{\lambda}\int_{\bigcup_j Q_j} g(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx \\ &\lesssim \frac{1}{\lambda}\int_{\mathbb{R}^n} f(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx \leq \frac{1}{\lambda}\int_{\mathbb{R}^n} f(x) M_{\bar{\mathcal{A}}}u(x)\,dx, \end{aligned}$$

which is the desired estimate for III.

To complete the proof of (a) we need to show (5.6). We first prove that for any Young function $\mathcal{C}$, any weight v with $M_{\mathcal{C}}v < \infty$ a.e., and any cube Q,

$$M_{\mathcal{C}}(v\chi_{\mathbb{R}^n\setminus 2Q})(y) \approx \operatorname*{ess\,inf}_{z\in Q} M_{\mathcal{C}}(v\chi_{\mathbb{R}^n\setminus 2Q})(z), \qquad \text{a.e. } y \in Q. \tag{5.7}$$

Let $y \in Q$ and R be any cube such that $y \in R$. If $R \setminus 2Q = \emptyset$ then $\|v\chi_{\mathbb{R}^n\setminus 2Q}\|_{\mathcal{C},R} = 0$. Otherwise, $\ell(R) > \ell(Q)/2$, which implies that $Q \subset 5R$. Then

$$\|v\chi_{\mathbb{R}^n\setminus 2Q}\|_{\mathcal{C},R} \lesssim \|v\chi_{\mathbb{R}^n\setminus 2Q}\|_{\mathcal{C},5R} \leq \operatorname*{ess\,inf}_{z\in Q} M_{\mathcal{C}}(v\chi_{\mathbb{R}^n\setminus 2Q})(z),$$

and taking the supremum over all cubes $R \ni y$ we deduce the desired estimate. Next, we use (5.7) to obtain (5.6):

$$\begin{aligned} \int_{\bigcup_j Q_j} g(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx &= \sum_j \int_{Q_j} g(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx \\ &= \sum_j f_{Q_j} \int_{Q_j} M_{\bar{\mathcal{A}}}(\tilde{u}\chi_{\mathbb{R}^n\setminus 2Q_j})(x)\,dx \\ &\lesssim \sum_j \int_{Q_j} f(x)\,dx \operatorname*{ess\,inf}_{z\in Q_j} M_{\bar{\mathcal{A}}}(\tilde{u}\chi_{\mathbb{R}^n\setminus 2Q_j})(z) \\ &\leq \sum_j \int_{Q_j} f(x) M_{\bar{\mathcal{A}}}(\tilde{u}\chi_{\mathbb{R}^n\setminus 2Q_j})(x)\,dx \\ &= \int_{\bigcup_j Q_j} f(x) M_{\bar{\mathcal{A}}}\tilde{u}(x)\,dx. \end{aligned}$$

This completes the proof of (a).

To show (b), we only have to estimate III. The argument is very similar, the main change consists in proving (5.5) with $\mathcal{D}_p$ in place of $\bar{\mathcal{A}}$. Once we have that, the above argument adapts trivially and the desired estimate for III follows. Note that our hypotheses guarantee that we can apply Theorem 2.4 to the adjoint of T—observe that $T^* = \tilde{T}$ where $\tilde{T}$ is the singular operator with kernel $\tilde{K}(x) = K(-x) \in H_{\mathcal{A}}$—and so Theorem 2.6 yields

$$(5.8)\qquad III = u\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : |Tg(x)| > \lambda/2\} \le \frac{2^p}{\lambda^p} \int_{\mathbb{R}^n} |Tg(x)|^p \tilde{u}(x)\, dx$$
$$\lesssim \frac{1}{\lambda^p} \int_{\mathbb{R}^n} g(x)^p M_{\mathcal{D}_p} \tilde{u}(x)\, dx.$$

As just mentioned, the ideas used before apply directly and the desired estimate follows at once.

Finally, we show (c). Given $\varepsilon > 0$ we pick $p > 1$ and $\delta > 0$ so that $p - 1 + \delta = \varepsilon$ (note that p is taken very close to 1 and δ very small). Then, by Remark 2.7, (2.3) holds for the pair of weights $(u, M_{L(\log L)^\varepsilon} u)$. Thus, the previous case *mutatis mutandis* leads us to the desired estimate. ∎

Remark 5.1. There is another argument to derive (c): Given $\varepsilon > 0$ we pick $p > 1$ and $\delta > 0$ so that $p - 1 + 2\delta = \varepsilon$. Let $\bar{\mathcal{A}}(t) = t(1 + \log^+)^{\delta/p}$. Note that $H_\infty \subset H_{\mathcal{A}}$ and so $K \in H_{\mathcal{A}}$. We take $\mathcal{D}(t) = t^p(1 + \log^+ t)^{p-1+2\delta}$ and $\mathcal{E}(t) \approx t^{p'}/(1 + \log^+ t)^{1+\delta(p'-1)} \in B_{p'}$. Then we can apply (b) to obtain the desired estimate for the pair of weights $(u, M_{\mathcal{D}_p} u)$. To conclude we observe that $\mathcal{D}_p(t) = \mathcal{D}(t^{1/p}) = t(1 + \log^+ t)^\varepsilon$.

Proof of Theorem 3.8. The argument follows the scheme of the proof of Theorem 3.1, which corresponds to the case $k = 0$, and we only give the main changes. We proceed by induction to obtain (a). The proof of (b) follows as in Theorem 3.1 from (a.2) by a suitable choice of $\mathcal{A}$ and $\mathcal{B}$ (see Remark 5.1).

We assume that the cases $m = 0, 1, \ldots, k-1$ are proved and we show the desired estimate for T_b^k. Thus, we fix a weight $u \in L_c^\infty$ and $0 \le f \in L_c^\infty$. By homogeneity we can also assume that $\|b\|_{\mathrm{BMO}} = 1$.

We recall some properties of BMO to be used later. Given $b \in \mathrm{BMO}$, a cube Q, $j \ge 0$ and $q > 0$, by John–Nirenberg's theorem we have

$$(5.9)\quad \|(b - b_Q)^j\|_{L^q,Q} \le \|(b - b_Q)^j\|_{\bar{\mathcal{C}}_j,Q} = \|b - b_Q\|^j_{\exp L,Q} \le C\|b\|^j_{\mathrm{BMO}}.$$

On the other hand, for every $l \ge 1$ and $b \in \mathrm{BMO}$, we have

$$(5.10)\qquad |b_Q - b_{2^l Q}| \le \sum_{m=1}^{l} |b_{2^{m-1}Q} - b_{2^m Q}| \le 2^n \sum_{m=1}^{l} \|b - b_{2^m Q}\|_{L^1,2^m Q}$$
$$\le 2^n l \|b\|_{\mathrm{BMO}}.$$

We perform the Calderón–Zygmund decomposition of f at level λ. Let g, $h=\sum_j h_j$, Q_j, $\tilde{Q}_j$, $\tilde{\Omega}$ and $\tilde{u}$ be as in the proof of Theorem 3.1. Then

$$\begin{aligned} u\{x\in\mathbb{R}^n : |T_b^k f(x)|>\lambda\} &\le u(\tilde{\Omega}) + u\{x\in\mathbb{R}^n\setminus\tilde{\Omega} : |T_b^k h(x)|>\lambda/2\} \\ &\quad + u\{x\in\mathbb{R}^n\setminus\tilde{\Omega} : |T_b^k g(x)|>\lambda/2\} \\ &= I + II + III, \end{aligned}$$

and we estimate each term separately. For I we obtain the first estimate in (5.2) exactly as before. Then

$$I \lesssim \frac{1}{\lambda}\int_{\mathbb{R}^n} f(x)Mu(x)\,dx \le \int_{\mathbb{R}^n} \mathcal{C}_k(|f(x)|/\lambda)Mu(x)\,dx$$

and we observe that Mu is pointwise controlled by either $M_{\bar{\mathcal{A}}}u$, $M_{\mathcal{D}_p}u$ or $M_{L(\log L)^{k+\varepsilon}}u$. So the desired estimate follows in all cases.

Next, we estimate II by using the induction hypothesis and the conditions assumed on the kernel. As in [28] we can write

$$\text{(5.11)}\qquad \begin{aligned} T_b^k h(x) = \sum_j T_b^k h_j(x) &= \sum_{m=0}^{k-1} C_{k,m} T_b^m\Big(\sum_j (b-b_{Q_j})^{k-m}h_j\Big)(x) \\ &\quad + \sum_j (b(x)-b_{Q_j})^k T h_j(x) = F_1(x)+F_2(x), \end{aligned}$$

and we estimate each function in turn.

For F_1 we would like to use the induction hypothesis. We start with (a.1). If $0\le m\le k-1$ then $H_{\mathcal{A},k}\subset H_{\mathcal{A},m}$ and so $K\in H_{\mathcal{B}}\cap H_{\mathcal{A},m}$. Also, as $\bar{\mathcal{C}}_k(t)\le\bar{\mathcal{C}}_m(t)$ we have

$$\bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_m^{-1}(t) \le \bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_k^{-1}(t) \le t.$$

Thus the hypotheses on (a) are satisfied for every $0\le m\le k-1$ and therefore

$$\begin{aligned} &u\{x\in\mathbb{R}^n\setminus\tilde{\Omega} : |F_1(x)|>\lambda/4\} \\ &\qquad \le \sum_{m=0}^{k-1} \tilde{u}\Big\{x : \Big|T_b^m\Big(\sum_j (b-b_{Q_j})^{k-m}h_j\Big)(x)\Big| > \lambda/C\Big\} \\ &\qquad \lesssim \sum_{m=0}^{k-1} \int_{\mathbb{R}^n} \mathcal{C}_m\Big(\Big|\sum_j (b-b_{Q_j})^{k-m}h_j\Big|/\lambda\Big) M_{\bar{\mathcal{A}}}\tilde{u}\,dx \\ &\qquad \lesssim \sum_{m=0}^{k-1}\sum_j \int_{Q_j} \mathcal{C}_m(|b-b_{Q_j}|^{k-m}|h_j|/\lambda) M_{\bar{\mathcal{A}}}\tilde{u}\,dx \\ &\qquad \lesssim \sum_{m=0}^{k-1}\sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}}\tilde{u} \int_{Q_j} \mathcal{C}_m(|b-b_{Q_j}|^{k-m}|h_j|/\lambda)\,dx, \end{aligned}$$

where in the last estimate we have used (5.7). As $\mathcal{C}_k^{-1}(t)\bar{\mathcal{C}}_{k-m}^{-1}(t) \lesssim \mathcal{C}_m^{-1}(t)$, Young's inequality implies

$$\begin{aligned}(5.12)\qquad &\int_{Q_j} \mathcal{C}_m(|b-b_{Q_j}|^{k-m}|h_j|/\lambda)\,dx \\ &\lesssim \int_{Q_j} \mathcal{C}_k(|h_j|/(c\lambda))\,dx + \int_{Q_j} \bar{\mathcal{C}}_{k-m}(c|b-b_{Q_j}|^{k-m})\,dx \\ &\lesssim \int_{Q_j} \mathcal{C}_k(|h_j|/\lambda)\,dx + \int_{Q_j} e^{c|b-b_{Q_j}|}\,dx \lesssim \int_{Q_j} \mathcal{C}_k(|h_j|/\lambda)\,dx + |Q_j|,\end{aligned}$$

as $\|b\|_{\mathrm{BMO}} = 1$ implies, by John–Nirenberg's theorem, that $\|b-b_{Q_j}\|_{\exp L, Q_j} \le c^{-1}$. Moreover, since $\mathcal{C}_k(t)^\delta$, $0<\delta<1$, is concave and so subadditive it follows that $\mathcal{C}_k$ is quasi-subadditive—that is, $\mathcal{C}_k(t_1+t_2) \lesssim \mathcal{C}_k(t_1)+\mathcal{C}_k(t_2)$. Therefore, by Jensen's inequality for $\mathcal{C}_k$,

$$\int_{Q_j} \mathcal{C}_k(|h_j|/\lambda)\,dx \le \int_{Q_j} \mathcal{C}_k(f/\lambda)\,dx + |Q_j|\mathcal{C}_k(f_{Q_j}/\lambda) \le 2\int_{Q_j} \mathcal{C}_k(f/\lambda)\,dx.$$

Also, (5.1) implies

$$|Q_j| \le \frac{1}{\lambda}\int_{Q_j} f\,dx \le \int_{Q_j} \mathcal{C}_k(f/\lambda)\,dx.$$

Plugging these estimates into (5.12) we obtain

$$\begin{aligned}u\{x\in\mathbb{R}^n\setminus\tilde{\Omega} : |F_1(x)|>\lambda/4\} &\lesssim \sum_{m=0}^{k-1}\sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}}\tilde{u}\int_{Q_j}\mathcal{C}_k(f/\lambda)\,dx \\ &\lesssim \sum_j \int_{Q_j} \mathcal{C}_k(f/\lambda) M_{\bar{\mathcal{A}}}\tilde{u}\,dx \le \int_{\mathbb{R}^n} \mathcal{C}_k(f/\lambda) M_{\bar{\mathcal{A}}} u\,dx.\end{aligned}$$

This gives the desired estimate for F_1 in case (a.1). Notice that the same computations hold in case (a.2) upon replacing everywhere $M_{\bar{\mathcal{A}}}$ by $M_{\mathcal{D}_p}$.

Next, we estimate F_2:

$$\begin{aligned}&u\{x\in\mathbb{R}^n\setminus\tilde{\Omega} : |F_2(x)|>\lambda/4\} \le \frac{4}{\lambda}\sum_j \int_{\mathbb{R}^n\setminus\tilde{\Omega}} |b(x)-b_{Q_j}|^k|Th_j(x)|u(x)\,dx \\ &\quad\le \frac{4}{\lambda}\sum_j \int_{\mathbb{R}^n\setminus\tilde{\Omega}} |b(x)-b_{Q_j}|^k \int_{Q_j} |K(x-y)-K(x-x_{Q_j})|\,|h_j(y)|\,dy\,u(x)\,dx \\ &\quad\le \frac{4}{\lambda}\sum_j \int_{Q_j} |h_j(y)| \int_{\mathbb{R}^n\setminus\tilde{Q}_j} |K(x-y)-K(x-x_{Q_j})|\,|b(x)-b_{Q_j}|^k u(x)\,dx\,dy.\end{aligned}$$

We claim that for every cube Q (whose center is x_Q) and every $y \in Q$,

$$\int_{\mathbb{R}^n \setminus 2Q} |K(x-y) - K(x-x_Q)|\, |b(x) - b_Q|^k u(x)\, dx \lesssim \operatorname*{ess\,inf}_Q M_{\bar{\mathcal{A}}} u. \tag{5.13}$$

This estimate applied to each Q_j implies

$$\begin{aligned} u\{x \in \mathbb{R}^n \setminus \tilde{\Omega} : |F_2(x)| > \lambda/4\} &\lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}} u \int_{Q_j} |h_j(y)|\, dy \\ &\lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{Q_j} M_{\bar{\mathcal{A}}} u \int_{Q_j} f(y)\, dy \le \frac{1}{\lambda} \sum_j \int_{Q_j} f(y) M_{\bar{\mathcal{A}}} u(y)\, dy \\ &\le \frac{1}{\lambda} \int_{\mathbb{R}^n} f(y) M_{\bar{\mathcal{A}}} u(y)\, dy \le \int_{\mathbb{R}^n} \mathcal{C}_k(|f(x)|/\lambda) M_{\bar{\mathcal{A}}} u(y)\, dy. \end{aligned}$$

Note that this leads to the desired estimate in (a.1) and also in (a.2) (we observed in Remark 3.3 that $M_{\bar{\mathcal{A}}} u \lesssim M_{\mathcal{D}_p} u$). Collecting the inequalities obtained for F_1 and F_2 we complete the estimate of II.

We show (5.13). Let Q be a cube with center x_Q and sidelength $2r$. Using (5.10), the generalized Hölder inequality for $\mathcal{A}$ and $\bar{\mathcal{A}}$, and also for $\bar{\mathcal{A}}$, $\mathcal{B}$ and $\bar{\mathcal{C}}_k$, and (5.9), we obtain

$$\begin{aligned} &\int_{\mathbb{R}^n \setminus 2Q} |K(x-y) - K(x-x_Q)|\, |b(x) - b_Q|^k u(x)\, dx \\ &\quad\lesssim \sum_{l=1}^{\infty} \int_{|x-x_Q| \sim 2^l r} |K(x-y) - K(x-x_Q)|\, |b(x) - b_{2^{l+1}Q}|^k u(x)\, dx \\ &\qquad + \sum_{l=1}^{\infty} l^k \int_{|x-x_Q| \sim 2^l r} |K(x-y) - K(x-x_Q)| u(x)\, dx \\ &\quad\lesssim \sum_{l=1}^{\infty} (2^l r)^n \|K(\cdot - y) - K(\cdot - x_Q)\|_{\mathcal{B}, |x-x_{Q_j}| \sim 2^l r} \\ &\qquad\qquad \times \|(b - b_{2^{l+1}Q})^k\|_{\bar{\mathcal{C}}_k, 2^{l+1}Q} \|u\|_{\bar{\mathcal{A}}, 2^{l+1}Q} \\ &\qquad + \sum_{l=1}^{\infty} (2^l r)^n l^k \|K(\cdot - y) - K(\cdot - x_Q)\|_{\mathcal{A}, |x-x_Q| \sim 2^l r} \|u\|_{\bar{\mathcal{A}}, 2^{l+1}Q} \\ &\quad\lesssim \operatorname*{ess\,inf}_Q M_{\bar{\mathcal{A}}} u, \end{aligned}$$

where we have used that $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$.

To complete the proof we need to estimate III. The proof is almost identical to that of Theorem 3.1. For the case (a.1), in (5.4) we apply Theorem 3.7 in place of Theorem 2.4. Once we have that estimate, the proof follows the same computations once we check that $|T_b^k g|^p M_r \tilde{u} \in L^1(\mathbb{R}^n)$

(we show this below). For the case (a.2) we need to show that T_b^k satisfies the corresponding estimate in (5.8). But this follows from Theorem 2.6 as we can apply Theorem 3.7 to the adjoint of T_b^k —note that $(T_b^k)^* = (T^*)^k_{-b}$ and T^* is a singular integral operator with kernel $\tilde{K}(x) = K(-x) \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$.

As just mentioned, we only need to check that $|T_b^k g|^p M_r \tilde{u} \in L^1(\mathbb{R}^n)$. Since $u \in L_c^\infty$, it suffices to see that $T_b^k g \in L^p(\mathbb{R}^n)$ for p large enough. This is trivial if one assumes that $b \in L^\infty$ as our assumption on K implies that $K \in H_1$ and thus T is bounded on $L^p(\mathbb{R}^n)$ for every $1 < p < \infty$:

$$\|T_b^k g\|_{L^p(\mathbb{R}^n)} = \Big\| \sum_{m=0}^k C_{m,k} b^{k-m} T(b^m g) \Big\|_{L^p(\mathbb{R}^n)} \lesssim \|b\|_{L^\infty}^k \|g\|_{L^p(\mathbb{R}^n)}$$
$$\le \|b\|_{L^\infty}^k \lambda^{(p-1)/p} \|f\|_{L^1(\mathbb{R}^n)}^{1/p} < \infty.$$

Thus, we obtain (3.6) with $Su = M_{\bar{\mathcal{A}}} u$ under the additional assumption that $b \in L^\infty$. We pass to an arbitrary $b \in \mathrm{BMO}$: for any $N > 0$ we define $b_N(x) = b(x)$ if $-N \le b(x) \le N$, $b_N(x) = N$ if $b(x) > N$, and $b_N(x) = -N$ if $b(x) < -N$. It is not hard to prove that $|b_N(x) - b_N(y)| \le |b(x) - b(y)|$ and hence $\|b_N\|_{\mathrm{BMO}} \le 2\|b\|_{\mathrm{BMO}}$. Therefore, as $b_N \in L^\infty$ we can use (3.6) with b_N in place of b and so

$$u\{x \in \mathbb{R}^n : |T_{b_N}^k f(x)| > \lambda\} \le C \int_{\mathbb{R}^n} \mathcal{C}_k(\|b_N\|_{\mathrm{BMO}}^k |f(x)|/\lambda) M_{\bar{\mathcal{A}}} u(x)\, dx \tag{5.14}$$
$$\le C \int_{\mathbb{R}^n} \mathcal{C}_k(\|b\|_{\mathrm{BMO}}^k |f(x)|/\lambda) M_{\bar{\mathcal{A}}} u(x)\, dx$$

where C does not depend on N. Since $f \in L_c^\infty$ it follows that for $0 \le m \le k$, $(b_N)^m f \to b^m f$ as $N \to \infty$ in L^q for $q > 1$. The fact that T is bounded on L^q implies $T((b_N)^m f) \to T(b^m f)$ as $N \to \infty$ in L^q. Passing to a subsequence the convergence is almost everywhere and so using the equality

$$T_{b_N}^k f(x) = \sum_{m=0}^k C_{m,k} b_N(x)^{k-m} T(b_N^m f)(x)$$

it follows that $T_{b_{N_j}}^k f(x) \to T_b^k f(x)$ for a.e. $x \in \mathbb{R}^n$ as $j \to \infty$. Thus, we clearly have $\chi_{\{T_b^k f > \lambda\}}(x) \le \liminf_{j\to\infty} \chi_{\{T_{b_{N_j}}^k f > \lambda\}}(x)$ a.e. Consequently, Fatou's lemma and (5.14) lead to the desired estimate for T_b^k. This completes the proof of (a).

To obtain (b), we proceed as in Remark 5.1. Given $\varepsilon > 0$ we pick $p > 1$ and $\delta > 0$ so that $(k+1)p - 1 + 2\delta = k + \varepsilon$. Let $\mathcal{A}(t) = \exp(t^{1/(k+\delta/p)}) - 1$ and $\mathcal{B}(t) = \exp(t^{p/\delta}) - 1$. Note that $\bar{\mathcal{A}}^{-1}(t)\mathcal{B}^{-1}(t)\bar{\mathcal{C}}_k^{-1}(t) \lesssim t$. Also, $H_\infty \subset H_{\mathcal{B}}$ and $H_{e^{t^{1/k}},k} \subset H_{\mathcal{A},k}$ (as $\mathcal{A}(t) \lesssim e^{t^{1/k}} - 1$ for $t \ge 1$). Then $K \in H_{\mathcal{B}} \cap H_{\mathcal{A},k}$. We apply (a.2) with $\mathcal{D}(t) = t^p(1+\log^+ t)^{(k+1)p-1+2\delta}$ and $\mathcal{E}(t) \approx$

$t^{p'}/(1+\log^+ t)^{1+\delta(p'-1)} \in B_{p'}$ (note that $\mathcal{D}^{-1}(t)\mathcal{E}^{-1}(t) \le \bar{\mathcal{A}}^{-1}(t)$). Then we obtain the desired estimate for the pair of weights $(u, M_{\mathcal{D}_p}u)$. To conclude we observe that $\mathcal{D}_p(t) = \mathcal{D}(t^{1/p}) = t(1+\log^+ t)^{k+\varepsilon}$. ∎

REMARK 5.2. The proof for the multilinear commutators follows the same scheme; we just indicate some of the changes, leaving the details to the reader. To estimate II we use ideas from [31] and replace (5.11) by

$$\begin{aligned} |T_{\vec{b}}h(x)| &\lesssim \sum_{\sigma_1,\sigma_2} \Big|T_{\vec{b}_{\sigma_2}}\Big(\sum_j \pi_{\sigma_1}(\vec{b}-\vec{\lambda}^j)h_j\Big)(x)\Big| \\ &\quad + \sum_j |\pi_{\{1,\dots,k\}}(\vec{b}-\vec{\lambda}^j)|\,|Th_j(x)| \\ &= F_1(x)+F_2(x), \end{aligned}$$

where the first sum runs over all partitions σ_1, σ_2 of $\{1,\dots,k\}$ with $\sigma_1 \ne \emptyset$; $T_{\vec{b}_{\sigma_2}}$ is the multilinear commutator associated with the vector $\vec{b}_{\sigma_2} = (b_{\sigma_2(l)})_l$; $\pi_{\sigma_1}(\vec{v}) = \prod_l v_{\sigma_1(l)}$ and $\pi_{\{1,\dots,k\}}(\vec{v}) = \prod_{l=1}^k v_l$; and $\vec{\lambda}^j = ((b_1)_{Q_j},\dots (b_k)_{Q_j})$. With this in hand we estimate F_1 using the induction hypothesis since $\#\sigma_2 \le k-1$, and we estimate F_2 using that $K \in H_{\mathcal{B},k}$ (see [18]).

The estimate for III is obtained by using [18, Theorem 7.1] and observing that $(T_{\vec{b}})^* = (T^*)_{-\vec{b}}$ and that T^* is a singular integral operator with kernel $\tilde{K}(x) = K(-x) \in H_{\mathcal{B},k}$.

6. Proofs in the one-sided case

Proof of Theorem 3.12(i). The proof follows the same scheme of the proof of Theorem 3.1. We will only highlight some of the details. Again, we can assume that u is bounded and has compact support, and also $0 \le f \in L_c^\infty(\mathbb{R})$. Let

$$\Omega = \{x \in \mathbb{R} : M^+ f(x) > \lambda\} = \bigcup_j I_j = \bigcup_j (a_j, b_j),$$

where $I_j = (a_j, b_j)$ are the connected components of Ω that satisfy (see [25])

$$\frac{1}{|I_j|}\int_{I_j} f(y)\,dy = \lambda.$$

Note that if $x \notin \Omega$, then for all $h \ge 0$,

$$\frac{1}{h}\int_x^{x+h} f(y)\,dy \le \lambda.$$

Therefore $f(x) \le \lambda$ for a.e $x \in \mathbb{R} \setminus \Omega$. Let $\widehat{I_j^-} = (c_j, a_j)$ with c_j chosen so that $|\widehat{I_j^-}| = 2|I_j|$ and set

$$\tilde{\Omega} = \bigcup_j (\widehat{I_j^-} \cup I_j) = \bigcup_j \tilde{I}_j.$$

We write $\tilde{u} = u\chi_{\mathbb{R}\setminus\tilde{\Omega}}$ and $f = g + h$ where

$$g = f\chi_{\mathbb{R}\setminus\Omega} + \sum_{j=1}^{\infty} f_{I_j}\chi_{I_j}, \qquad h = \sum_{j=1}^{\infty} h_j = \sum_{j=1}^{\infty} (f - f_{I_j})\chi_{I_j}.$$

Observe that $0 \le g(x) \le \lambda$ for a.e. x, and h_j has vanishing integral. Thus

$$\begin{aligned} u\{x \in \mathbb{R} : |T^+ f(x)| > \lambda\} &\le u(\tilde{\Omega}) + u\{x \in \mathbb{R} \setminus \tilde{\Omega} : |T^+ h(x)| > \lambda/2\} \\ &\quad + u\{x \in \mathbb{R} \setminus \tilde{\Omega} : |T^+ g(x)| > \lambda/2\} \\ &= I + II + III. \end{aligned}$$

Now we proceed in the same way as in the proof of Theorem 3.1. We estimate I:

$$I = u(\tilde{\Omega}) = u\Big(\bigcup_j \tilde{I}_j\Big) \le \sum_j (u(\widehat{I_j^-}) + u(I_j)).$$

For each j we have

$$u(\widehat{I_j^-}) = \frac{u(\widehat{I_j^-})}{|I_j|}\,|I_j| = \frac{u(\widehat{I_j^-})}{|I_j|}\,\frac{1}{\lambda}\int_{I_j} |f(x)|\,dx \le \frac{3}{\lambda}\int_{I_j} |f(x)| M^- u(x)\,dx,$$

and then

$$\sum_j u(\widehat{I_j^-}) \lesssim \frac{1}{\lambda}\int_{\Omega} |f(x)| M^- u(x)\,dx \le \frac{1}{\lambda}\int_{\mathbb{R}} |f(x)| M^- u(x)\,dx.$$

On the other hand, since M^+ is of weak type $(1,1)$ with respect to the pair of weights $(u, M^- u) \in A_1^+$ (see [22]),

$$\sum_j u(I_j) = u(\Omega) \le \frac{C}{\lambda}\int_{\mathbb{R}} f(x) M^- u(x)\,dx,$$

and therefore

$$I \le \frac{C}{\lambda}\int_{\mathbb{R}} f(x) M^- u(x)\,dx.$$

Observe that, as before, $M^- u$ is controlled by $M_{\mathcal{A}}^- u$ in (a), by $M_{\mathcal{D}_p}^- u$ in (b), and by $M_{L(\log L)^\varepsilon}^- u$ in (c).

We estimate II. Let $r_j = |I_j| = |\widehat{I_j^-}|/2$. We use the fact that h_j is supported in I_j and has vanishing integral, and that K is supported in $(-\infty, 0)$:

$$II \le \frac{2}{\lambda} \sum_j \int_{\mathbb{R}\setminus\tilde{\Omega}} |Th_j(x)| u(x)\, dx$$

$$\le \frac{2}{\lambda} \sum_j \int_{I_j} |h_j(y)| \int_{\mathbb{R}\setminus\tilde{I}_j} |K(x-y) - K(x-a_j)| u(x)\, dx\, dy$$

$$= \frac{2}{\lambda} \sum_j \int_{I_j} |h_j(y)| \int_{-\infty}^{c_j} |K(x-y) - K(x-a_j)| u(x)\, dx\, dy.$$

Then it suffices to prove that for every $y \in I_j$,

$$\int_{-\infty}^{c_j} |K(x-y) - K(x-a_j)| u(x)\, dx \lesssim \operatorname*{ess\,inf}_{I_j} M_{\bar{\mathcal{A}}}^- u, \tag{6.1}$$

which readily leads to the desired estimate

$$II \lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{I_j} M_{\bar{\mathcal{A}}}^- u \int_{I_j} |h_j(y)|\, dy \lesssim \frac{1}{\lambda} \sum_j \operatorname*{ess\,inf}_{I_j} M_{\bar{\mathcal{A}}}^- u \int_{I_j} f(y)\, dy$$

$$\le \frac{1}{\lambda} \sum_j \int_{I_j} f(y) M_{\bar{\mathcal{A}}}^- u(y)\, dy \le \frac{1}{\lambda} \int_{\mathbb{R}} f(y) M_{\bar{\mathcal{A}}}^- u(y)\, dy.$$

We show (6.1). Let $y, z \in I_j$. Using the generalized Hölder inequality for $\mathcal{A}$ and $\bar{\mathcal{A}}$, and the fact that $K \in H_{\mathcal{A}}$, we obtain

$$\int_{-\infty}^{c_j} |K(x-y) - K(x-a_j)| u(x)\, dx$$

$$= \sum_{k=1}^{\infty} \int_{a_j - 2^{k+1} r_j}^{a_j - 2^k r_j} |K(x-y) - K(x-a_j)| u(x)\, dx$$

$$\lesssim \sum_{k=1}^{\infty} 2^k r_j \|K(\cdot - y) - K(\cdot - a_j)\|_{\mathcal{A}, |x-a_j| \sim 2^k r_j} \times \|u \chi_{(a_j - 2^{k+1} r_j, a_j - 2^k r_j)}\|_{\bar{\mathcal{A}}, |x-a_j| \sim 2^k r_j}$$

$$\le \sum_{k=1}^{\infty} 2^k r_j \|K(\cdot - y) - K(\cdot - a_j)\|_{\mathcal{A}, |x-a_j| \sim 2^k r_j} \|u\|_{\bar{\mathcal{A}}, (a_j - 2^{k+1} r_j, z)}$$

$$\lesssim M_{\bar{\mathcal{A}}}^- u(z).$$

To estimate III we first claim that for any Young function $\mathcal{C}$, any weight v with $M_{\mathcal{C}}^{-}v<\infty$ a.e., and any interval $I=(a,b)$ we have

$$M_{\mathcal{C}}^{-}(v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)})(y)\approx \operatorname*{ess\,inf}_{z\in I} M_{\mathcal{C}}^{-}(v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)})(z), \qquad \text{a.e. } y\in I, \tag{6.2}$$

where $\widehat{I}^{-}=(c,a)$ with c so that $|\widehat{I}^{-}|=2|I|$. Assuming this the proofs of the three cases (a)–(c) adapt readily to the one-sided setting. For (a) one uses the fact that $M_r^{-}\tilde{u}\in A_1^{+}\subset A_\infty^{+}$ and therefore M^{+} is bounded on $L^q(M_r^{-}\tilde{u})$ for every $1<q<\infty$. For (b) we apply Theorem 3.11(i) to the one-sided operator $T^{-}=(T^{+})^{*}$ (whose kernel $\tilde{K}(x)=K(-x)$ is supported on $(0,\infty)$) and then Theorem 3.11(ii) to deduce (5.8) with $M_{\mathcal{D}_p}^{-}\tilde{u}$ on the right-hand side. In case (c) we only need to adapt Remark 5.1 to this setting.

To complete the proof of (i) we show (6.2). Fix $y,z\in I=(a,b)$ and write $\widehat{I}^{-}=(c,a)$, where we recall that $|\widehat{I}^{-}|=2|I|$. Observe that if $c\le t<y$ then $(t,y)\subset(c,b)=\widehat{I}^{-}\cup I$. Thus,

$$M_{\mathcal{C}}^{-}(v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)})(y)=\sup_{t<y}\|v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}\|_{\mathcal{C},(t,y)}=\sup_{t<c}\|v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}\|_{\mathcal{C},(t,y)}.$$

Given $t<c$ and $\lambda>0$, it follows that

$$\begin{aligned}\frac{1}{y-t}\int_t^y \mathcal{C}\Big(\frac{v(x)\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}(x)}{\lambda}\Big)\,dx &= \frac{1}{y-t}\int_t^c \mathcal{C}\Big(\frac{v(x)\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}(x)}{\lambda}\Big)\,dx\\ &= \frac{z-t}{y-t}\,\frac{1}{z-t}\int_t^c \mathcal{C}\Big(\frac{v(x)\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}(x)}{\lambda}\Big)\,dx\\ &\le \frac{3}{2}\,\frac{1}{z-t}\int_t^z \mathcal{C}\Big(\frac{v(x)\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}(x)}{\lambda}\Big)\,dx\end{aligned}$$

and therefore $\|v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}\|_{\mathcal{C},(t,y)}\le \frac{3}{2}\|v\chi_{\mathbb{R}\setminus(\widehat{I}^{-}\cup I)}\|_{\mathcal{C},(t,z)}$, which in turn gives the desired estimate. ∎

Using the previous ideas the proof of Theorem 3.12(ii) can be obtained by adapting the proof of Theorem 3.8. Further details are left to the reader.

Proof of Theorem 3.14. By homogeneity it suffices to consider the case $\lambda=1$. Let $1<p<\infty$ and $\Omega=\{x:|Tf(x)|>1\}$. Then, by duality, there exists $G\in L^{p'}(u)$ with $\|G\|_{L^{p'}(u)}=1$ such that

$$u(\Omega)^{1/p}=\|\chi_\Omega\|_{L^p(u)}=\int_\Omega G(x)u(x)\,dx.$$

Using the hypotheses and Hölder's inequality we obtain

$$u(\Omega)^{1/p}\lesssim\int_{\mathbb{R}^n}|f(x)|M_{\mathcal{F}}^{-}(Gu)(x)\,dx\le\|f\|_{L^p(v)}\|M_{\mathcal{F}}^{-}(Gu)\|_{L^{p'}(v^{1-p'})}.$$

Our hypotheses guarantee that we can apply Theorem 3.15 (indeed the corresponding version for $M_{\mathcal{F}}^-$) to infer that $M_{\mathcal{F}}^-$ maps $L^{p'}(u^{1-p'})$ into $L^{p'}(v^{1-p'})$. Therefore,

$$u(\Omega)^{1/p} \lesssim \|f\|_{L^p(v)} \|Gu\|_{L^{p'}(u^{1-p'})} = \|f\|_{L^p(v)} \|G\|_{L^{p'}(u)} = \|f\|_{L^p(v)}. \ \blacksquare$$

REMARK 6.1. As observed before, this proof can be easily adapted to yield Theorem 3.5: one only needs to see that (3.3) guarantees that $M_{\mathcal{F}}$ is bounded from $L^{p'}(u^{1-p'})$ into $L^{p'}(v^{1-p'})$ (see [11]). For further results and a deep treatment of extrapolation results of this kind the reader is referred to [10].

Proof of Theorem 3.15. We use ideas from [33]. First we observe that it suffices to assume that u is bounded and compactly supported (otherwise we work with $u_R = u\chi_{\{|x|\le R,\, u(x)\le R\}}$ and let $R \to \infty$).

Fix f continuous with compact support, and for $k \in \mathbb{Z}$ define $\Omega_k = \{x \in \mathbb{R} : 2^k < M_{\mathcal{A}}^+ f(x) < 2^{k+2}\}$. For any $x \in \Omega_k$ there exists $c_x > x$ such that $2^k < \|f\|_{\mathcal{A},(x,c_x)} < 2^{k+2}$. Using the continuity of the integral it is easy to show that there exists $\delta_x \in (x, c_x)$ (which can be taken sufficiently close to x satisfying $\delta_x < (c_x + x)/2$) such that $[x, \delta_x) \subset \Omega_k$ and $2^k < \|f\|_{\mathcal{A},(\delta_x,c_x)} < 2^{k+2}$. We write $I_{x,k}^- = [x, \delta_x)$ and $I_{x,k}^+ = (\delta_x, c_x)$ and therefore

$$\Omega_k = \bigcup_{x\in\Omega_k} I_{x,k}^- \quad \text{and} \quad 2^k < \|f\|_{\mathcal{A},I_{x,k}^+} < 2^{k+2}. \tag{6.3}$$

As in [5] (see also [33, Lemma 2]) there exists a finite subcollection of pairwise disjoint intervals $\{I_{j,k}^-\}_{j\in J}$ such that

$$u(\Omega_k) \le 3 \sum_{j\in J} u(I_{j,k}^-).$$

This and (6.3) yield

$$\begin{aligned}
\int_{\mathbb{R}} M_{\mathcal{A}}^+ f(x)^p u(x)\,dx &\le \sum_{k\in\mathbb{Z}} \int_{\Omega_k} (M_{\mathcal{A}}^+ f)^p u\,dx \lesssim \sum_{k\in\mathbb{Z}} 2^{kp} u(\Omega_k) \\
&\lesssim \sum_{k\in\mathbb{Z}} 2^{kp} \sum_{j\in J} u(I_{j,k}^-) \le \sum_{k\in\mathbb{Z}} \sum_{j\in J} \|f\|_{\mathcal{A},I_{j,k}^+}^p u(I_{j,k}^-) \\
&= \sum_{k\in\mathbb{Z}} \sum_{j\in J} \|fv^{1/p}v^{-1/p}\|_{\mathcal{A},I_{j,k}^+}^p u(I_{j,k}^-) \\
&\lesssim \sum_{k\in\mathbb{Z}} \sum_{j\in J} \|fv^{1/p}\|_{\mathcal{C},I_{j,k}^+}^p \|v^{-1/p}\|_{\mathcal{B},I_{j,k}^+}^p \|u^{1/p}\|_{L^p,I_{j,k}^-}^p |I_{j,k}^-|.
\end{aligned}$$

Note that

$$\|fv^{1/p}\|_{\mathcal{C},I_{j,k}^+} \le 2\|fv^{1/p}\|_{\mathcal{C},I_{j,k}^-\cup I_{j,k}^+} \le 2M_{\mathcal{C}}(fv^{1/p})(x), \qquad x \in I_{j,k}^-.$$

This, (3.7), and the fact that the intervals $I^-_{j,k}$ are pairwise disjoint and contained in Ω_k imply

$$\int_{\mathbb{R}} M^+_{\mathcal{A}} f(x)^p u(x)\,dx \lesssim \sum_{k\in\mathbb{Z}}\sum_{j\in J} \|fv^{1/p}\|^p_{\mathcal{C},I^+_{j,k}} |I^-_{j,k}|$$
$$\lesssim \sum_{k\in\mathbb{Z}}\sum_{j\in J} \int_{I^-_{j,k}} M_{\mathcal{C}}(fv^{1/p})(x)^p\,dx \le \sum_{k\in\mathbb{Z}} \int_{\Omega_k} M_{\mathcal{C}}(fv^{1/p})(x)^p\,dx$$
$$\le 2\int_{\mathbb{R}} M_{\mathcal{C}}(fv^{1/p})(x)^p\,dx \lesssim \int_{\mathbb{R}} |f(x)|^p v(x)\,dx,$$

where in the last estimate we have used the fact that $\mathcal{C} \in B_p$ and consequently $M_{\mathcal{C}}$ is bounded on $L^p(\mathbb{R})$ (see [29]). This completes the proof. ∎

Acknowledgements. The first and last authors are partially supported by MCYT Grant MTM2005-08350-C03-02. The first author is also supported by Junta de Andalucía Grant FQM354. The second author is partially supported by CSIC PIE 200850I015 , by MEC Grant MTM2007-60952, and by UAM-CM Grant CCG07-UAM/ESP-1664. The third author is partially supported by MTM2006-05622. The last author is partially supported by CONICET, Agencia Córdoba Ciencias and SECYT-UNC.

References

[1] H. Aimar, L. Forzani and F. J. Martín-Reyes, *On weighted inequalities for one-sided singular integrals*, Proc. Amer. Math. Soc. 125 (1997), 2057–2064.

[2] J. Alvarez, R. Bagby, D. Kurtz and C. Pérez, *Weighted estimates for commutators of linear operators*, Studia Math. 104 (1993), 195–209.

[3] C. Bennett and R. Sharpley, *Interpolation of Operators*, Academic Press, New York, 1998.

[4] A. L. Bernardis, M. Lorente, F. J. Martín-Reyes, M. T. Martínez, A. de la Torre and J. L. Torrea, *Differential transforms in weighted spaces*, J. Fourier Anal. Appl. 12 (2006), 83–103.

[5] J. Bliendtner and P. Loeb, *A reduction technique for limit theorems in analysis and probability theory*, Ark. Mat. 30 (1992), 25–43.

[6] A. P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Math. 88 (1952), 85–139.

[7] R. Coifman, *Distribution function inequalities for singular integrals*, Proc. Nat. Acad. Sci. U.S.A. 69 (1972), 2838–2839.

[8] A. Cordoba and C. Fefferman, *A weighted norm inequality for singular integrals*, Studia Math. 57 (1976), 97–101.

[9] D. Cruz-Uribe, J. M. Martell and C. Pérez, *Extensions of Rubio de Francia's extrapolation theorem*, in: Proc. 7th Int. Conf. on Harmonic Analysis and Partial Differential Equations (El Escorial, 2004), Collect. Math. 2006, 195–231.

[10] —, —, —, *Weights, Extrapolation and the Theory of Rubio de Francia*, preprint, 2008.

[11] D. Cruz-Uribe and C. Pérez, *Sharp two-weight, weak-type norm inequalities for singular integral operators*, Math. Res. Lett. 6 (1999), 417–427.

[12] —, —, *Two weight extrapolation via the maximal operator*, J. Funct. Anal. 174 (2000), 1–17.

[13] G. P. Curbera, J. García-Cuerva, J. M. Martell and C. Pérez, *Extrapolation with weights, rearrangement invariant function spaces, modular inequalities and applications to singular integrals*, Adv. Math. 203 (2006), 256–318.

[14] S. Hofmann, *Singular integrals with power weights*, Proc. Amer. Math. Soc. 110 (1990), 343–353.

[15] L. Hörmander, *Estimates for translation invariant operators in L^p spaces*, Acta Math. 104 (1960), 93–140.

[16] R. L. Jones and J. Rosenblatt, *Differential and ergodic transform*, Math. Ann. 323 (2002), 525–546.

[17] D. S. Kurtz and R. L. Wheeden, *Results on weighted norm inequalities for multipliers*, Trans. Amer. Math. Soc. 255 (1979), 343–362.

[18] M. Lorente, J. M. Martell, M. S. Riveros and A. de la Torre, *Generalized Hörmander's condition, commutators and weights*, J. Math. Anal. Appl. 342 (2008), 1399–1425.

[19] M. Lorente and M. S. Riveros, *Weighted inequalities for commutators of one-sided singular integrals*, Comment. Math. Univ. Carolin. 43 (2002), 83–101.

[20] M. Lorente, M. S. Riveros and A. de la Torre, *Weighted estimates for singular integral operators satisfying Hörmander's conditions of Young type*, J. Fourier Anal. Apl. 11 (2005), 497–509.

[21] J. M. Martell, C. Pérez, and R. Trujillo-González, *Lack of natural weighted estimates for some singular integral operators*, Trans. Amer. Math. Soc. 357 (2005), 385–396

[22] F. J. Martín-Reyes, P. Ortega and A. de la Torre, *Weighted inequalities for one-sided maximal functions*, Trans. Amer. Math. Soc. 319 (1990), 517–534.

[23] F. J. Martín-Reyes, L. Pick and A. de la Torre, *A_∞^+ condition*, Canad. J. Math. 45 (1993), 1231–1244.

[24] F. J. Martín-Reyes and A. de la Torre, *One sided BMO spaces*, J. London Math. Soc. (2) 49 (1994), 529–542.

[25] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207–226.

[26] R. O'Neil, *Fractional integration in Orlicz spaces*, ibid. 115 (1963), 300–328.

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

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

[29] —, *On sufficient conditions for the boundedness of the Hardy–Littlewood maximal operator between weighted L^p-spaces with different weights*, Proc. London Math. Soc. (3) 71 (1995), 135–157.

[30] C. Pérez and G. Pradolini, *Sharp weighted endpoint estimates for commutators of singular integrals*, Michigan Math. J. 49 (2001), 23–37.

[31] C. Pérez and R. Trujillo-González, *Sharp weighted estimates for multilinear commutators*, J. London Math. Soc. (2) 65 (2002), 672–692.

[32] M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, Monographs Textbooks Pure Appl. Math. 146, Dekker, New York, 1991.

[33] M. S. Riveros, L. de Rosa and A. de la Torre, *Sufficient conditions for one-sided operators*, J. Fourier Anal. Appl. 6 (2000), 607–621.

[34] J. L. Rubio de Francia, F. J. Ruiz and J. L. Torrea, *Calderón–Zygmund theory for vector-valued functions*, Adv. Math. 62 (1986), 7–48.
[35] E. Sawyer, *Weighted inequalities for the one-sided Hardy–Littlewood maximal functions*, Trans. Amer. Math. Soc. 297 (1986), 53–61.
[36] A. de la Torre and J. L. Torrea, *One-sided discrete square function*, Studia Math. 156 (2003), 243–260.
[37] D. K. Watson, *Weighted estimates for singular integrals via Fourier transform estimates*, Duke Math. J. 60 (1990), 389–399.

Departamento de Análisis Matemático
Facultad de Ciencias
Universidad de Málaga
Campus de Teatinos, s/n
29071 Málaga, Spain
E-mail: lorente@anamat.cie.uma.es

Instituto de Ciencias Matemáticas
CSIC-UAM-UC3M-UCM
Consejo Superior de Investigaciones Científicas
C/ Serrano 121
28006 Madrid, Spain
E-mail: chema.martell@uam.es

Departamento de Análisis Matemático
Facultad de Matemáticas
Universidad de Sevilla
41080, Sevilla, Spain
E-mail: carlosperez@us.es

FaMAF
Universidad Nacional de Córdoba
CIEM (CONICET)
5000 Córdoba, Argentina
E-mail: sriveros@mate.uncor.edu

Received November 21, 2008
Revised version March 5, 2009 (6466)