

Sharp estimates for commutators of singular integrals via iterations of the Hardy–Littlewood maximal function

Carlos Pérez

J. Fourier Analysis and Applications, **3** (1997), 108–146.

Departamento de Matemáticas
Universidad Autónoma de Madrid
28049 Madrid, Spain
e-mail: cperezmo@ccuam3.sdi.uam.es

work partially supported by DGICYT grant PB940192, Spain

1 Introduction and description of the main results

The purpose of this paper is to obtain some sharp non standard weighted inequalities for linear and nonlinear commutators of singular integral operators. These estimates provide a further insight into the structure of these operators and in particular they reflect a higher degree of singularity as compared with the standard Calderón–Zygmund singular integral operators.

Let T denote a Calderón–Zygmund singular integral operator and let M be the Hardy–Littlewood maximal function. According to a result of R. Coifman [C], T and M satisfy the following a priori estimate:

Let $0 < p < \infty$ and suppose that $w \in A_\infty(\mathbf{R}^n)$. Then the inequality

$$\int_{\mathbf{R}^n} |Tf(x)|^p w(x) dx \leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} Mf(x)^p w(x) dx, \quad (1)$$

holds for every function f for which the left hand side is finite.

This estimate plays a major role in the modern theory of weighted norm inequalities since as it is well known it follows that T is a bounded operator on $L^p(w)$ whenever $w \in A_p$ and $p > 1$. This extends the previous result of R. A. Hunt, B. Muckenhoupt and R. L. Wheeden in [HMW] whose method works only for the Hilbert Transform. Furthermore, (1) makes explicit the well known **Calderón–Zygmund principle** which establishes that a singular integral operator is controlled by an appropriate maximal function.

There is another aspect of Coifman’s estimate that we shall be exploiting along this paper. It concerns the two weighted inequality problem for singular integrals, say the Hilbert transform, which is completely open. Combining (1) with certain sharp two weighted inequalities for M we can derive two weighted estimates for T with **no** a priori assumption on the weight w . As a sample we quote the following inequality from [Wil] [P3]:

Let T be a Calderón–Zygmund singular integral operator and let $1 < p < \infty$. Then, there exists a constant C such that

$$\int_{\mathbf{R}^n} |Tf(x)|^p w(x) dx \leq C \int_{\mathbf{R}^n} |f(x)|^p M^{[p]+1} w(x) dx, \quad (2)$$

where C is independent of w and f .

1.1 Higher order commutators

In this paper we are going to investigate generalizations of above inequalities (1) and (2) for a large family of singular integral operators. First we shall consider the higher order commutators introduced by R. Coifman, R. Rochberg and G. Weiss in [CRW]. These are linear operators defined for appropriate functions b and f and for $k = 0, 1, 2, \dots$ by

$$T_b^k f(x) = \int (b(x) - b(y))^k K(x, y) f(y) dy$$

which must be understood in the usual sense. When $k = 1$ the operator T_b^1 is usually denoted by $[M_b, T] = M_b \circ T - T \circ M_b$ where M_b is the operator defined by $M_b f = b f$, and b is usually called the ‘‘symbol’’ of the operator. These commutators have proved to be of interest in many contexts and in particular in the theory of P.D.E. We shall only mention the recent results in the theory of non divergence elliptic equations with discontinuous coefficients [CFL1] [CFL2] [DiR].

The main result from [CRW] is the following:

Let $1 < p < \infty$ and let $b \in BMO$, then there exists a constant C such that

$$\|T_b^k f\|_{L^p(\mathbf{R}^n)} \leq C \|b\|_{BMO}^k \|f\|_{L^p(\mathbf{R}^n)}. \quad (3)$$

Throughout the paper $M^k = M \circ \overset{(k)}{!} \circ M$ will denote the Hardy–Littlewood maximal operator M iterated k times.

Following the Calderón–Zygmund principle we shall show that the maximal operator which controls the higher order commutators $T_b^k f$ when b is a BMO function is M^{k+1} , namely in some sense we have that

$$T_b^k \approx M \circ \overset{(k+1)}{!} \circ M$$

when $b \in BMO$. This can be made precise with the following generalization of (1).

THEOREM 1.1 *Let $0 < p < \infty$ and let $w \in A_\infty$ and $b \in BMO$. Then, there exists a constant C such that*

$$\int_{\mathbf{R}^n} |T_b^k f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{kp} [w]_{A_\infty}^{(k+1)p} \int_{\mathbf{R}^n} M^{k+1} f(x)^p w(x) dx. \quad (4)$$

This inequality contains the well known fact that the higher order commutators are bounded on $L^p(w)$, $w \in A_p$, by applying $k + 1$ times Muckenhoupt's Theorem. As we said before (4) can be used as well to get a generalization of inequality (2).

THEOREM 1.2 *Let $1 < p < \infty$ and let $b \in BMO$. Then, there exists a constant C such that for each weight w*

$$\int_{\mathbf{R}^n} |T_b^k f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{kp} \int_{\mathbf{R}^n} |f(x)|^p M^{[(k+1)p]+1} w(x) dx. \quad (5)$$

We remark that the number of iterations of the maximal function needed in both Theorems are optimal (see §5). In fact it follows from the proof of (5) that there is a sharper estimate:

$$\int_{\mathbf{R}^n} |T_b^k f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{kp} \int_{\mathbf{R}^n} |f(x)|^p M_{L(\log L)^{(k+1)p-1+\epsilon}}(w)(x) dx$$

where $\epsilon > 0$, being the result false for $\epsilon = 0$. See §2 for the definition of $M_{L(\log L)^\alpha}$.

Observe that both estimates (4) and (5) show that the operator T_b^k becomes more singular with k since the maximal function on the right hand side of the inequalities needs more “iterations” to balance the inequalities. Also observe that we cannot get the sharp case (5) iterating from the case $k = 1$.

Before continuing, let us point out that M. Wilson [Wil] was the first author who derived an estimate such as (5) for singular integrals of convolution type $T_b^0 = T$ but only on the range $1 < p \leq 2$. However, Wilson's approach is interesting because is direct and based on sharp weighted estimates for smooth Littlewood–Paley square functions using as a key step a deep result by T. Wolff [CWW] concerning the behavior of the square functions on L^∞ .

Our method is by duality having the advantages that first covers the full range $1 < p < \infty$ and second it is flexible enough to be applied to a wider class of operators such as T_b^k rather than T . Let us give an outline of the proof of Theorem 1.2 which is based on the following steps and which seems to be general enough to be applicable to other (linear) operators:

1. For simplicity denote T_b^k by T and $[(k+1)p] + 1$ by $k(p)$. Now, instead of proving directly (5) we consider the corresponding (equivalent) dual inequality, namely

$$\int_{\mathbf{R}^n} |Tf(x)|^{p'} (M^{k(p)}w(x))^{1-p'} dx \leq C \int_{\mathbf{R}^n} |f(x)|^{p'} w(x)^{1-p'} dx \quad (6)$$

since the adjoint operator to T_b^k is essentially the same.

2. After observing that $(M^{k(p)}w)^{1-p'} \in A_\infty$ (in fact it belongs to \mathbf{RH}_∞) we apply the Calderón–Zygmund principle: we replace the singular integral by a maximal type operator, namely M^{k+1} in our case using Theorem 1.1:

$$\int_{\mathbf{R}^n} |Tf(x)|^{p'} (M^{k(p)}w(x))^{1-p'} dx \leq C \int_{\mathbf{R}^n} M^{k+1}f(x)^{p'} (M^{k(p)}w(x))^{1-p'} dx. \quad (7)$$

3. Therefore everything is reduced to showing a sharp two weighted norm inequalities for the maximal operator M^{k+1}

$$\int_{\mathbf{R}^n} M^{k+1}f(x)^{p'} (M^{k(p)}w(x))^{1-p'} dx \leq C \int_{\mathbf{R}^n} |f(x)|^{p'} w(x)^{1-p'} dx. \quad (8)$$

1.2 The Nonlinear commutator

The second commutator that we are going to consider was introduced by R. Rochberg and G. Weiss in [RW]. This nonlinear operator is defined for appropriate functions by

$$f \rightarrow Nf = T(f \log |f|) - Tf \log |Tf|.$$

N is homogeneous and can be written as a commutator $[\Omega, T] = T \circ \Omega - \Omega \circ T$ where Ω denotes the operation $\Omega f = f \log |f|$. There is a growing interest in studying this operator due to its relationship with the Jacobian mapping and with nonlinear P.D.E. as shown in [IS] [GI] (see also [M]).

The main result from [RW] is the following:

Let $1 < p < \infty$, then there exists a constant C such that

$$\|Nf\|_{L^p(\mathbf{R}^n)} \leq C \|f\|_{L^p(\mathbf{R}^n)}. \quad (9)$$

The theory developed in [RW] is very general. It shows, for instance, that the singular integral T may be replaced by any linear operator bounded on $L^{p_i}(\mathbf{R}^n)$, $i = 1, 2$, with $1 < p_1 < p < p_2 < \infty$. However, to derive A_p type estimates for N such a general framework does not seem to be suitable. We shall be using a different approach based on real variable techniques and in particular on the theory of A_p weights combined with some of the estimates obtained above for the linear commutator $[M_b, T]$. Furthermore and trying to follow the Calderón–Zygmund principle again, we show that the maximal operator which controls N is the Hardy–Littlewood maximal function iterated twice, namely

$$N \approx M \circ M,$$

expression which more precisely means the following:

THEOREM 1.3 *Suppose that $0 < p < \infty$ and that $w \in A_\infty$. Then, there exists a constant C such that*

$$\int_{\mathbf{R}^n} |Nf(x)|^p w(x) dx \leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} M^2 f(x)^p w(x) dx, \quad (10)$$

As an immediate consequence we have the following corollary.

COROLLARY 1.4 *Let $1 < p < \infty$ and let $w \in A_p$. Then, there exists a constant C such that*

$$\int_{\mathbf{R}^n} |Nf(x)|^p w(x) dx \leq C [w]_{A_p}^{3p} \int_{\mathbf{R}^n} |f(x)|^p w(x) dx, \quad (11)$$

Contrary to what we did for the linear commutator T_b^k we cannot apply Theorem 1.3 to derive for N a result in the spirit of Theorem 1.2. The method sketched above breakdowns due to the nonlinearity of N . However and by a direct approach we can still deduce a corresponding estimate.

THEOREM 1.5 *Suppose that $1 < p < \infty$. Then, there exists a constant C such that for each weight w*

$$\int_{\mathbf{R}^n} |Nf(x)|^p w(x) dx \leq C \int_{\mathbf{R}^n} |f(x)|^p M^{[2p]+1} w(x) dx. \quad (12)$$

To get this estimate we show that there is a relationship between N and the linear commutator $[M_b, T]$ and consequently with $M \circ M = M^2$. The observation is that N can be written using the linearity of T as follows (see §4):

$$Nf = T\left(f \log \frac{|f|}{Mf}\right) + [M_{\log Mf}, T](f) - Tf \log \frac{|Tf|}{Mf} = N_1f + N_2f + N_3f.$$

Observe that the symbol of the operator N_2 is the operation $b = b(f) = \log Mf$ which is a *BMO* function with a constant independent of f .

2 Some preliminaries and notation

We shall introduce in this section some of the necessary tools that we need to prove our results. Recall that a function $B : [0, \infty) \rightarrow [0, \infty)$ is called a Young function if it is continuous, convex and increasing satisfying $B(0) = 0$ and $B(t) \rightarrow \infty$ as $t \rightarrow \infty$. We define the B -average of a function f over a cube Q by means of the Luxemburg norm

$$\|f\|_{B,Q} = \inf\{\lambda > 0 : \frac{1}{|Q|} \int_Q B\left(\frac{|f(y)|}{\lambda}\right) dy \leq 1\}, \quad (13)$$

and recall the following generalization of Hölder's inequality:

$$\frac{1}{|Q|} \int_Q |f(y)g(y)| dy \leq \|f\|_{B,Q} \|g\|_{\bar{B},Q}, \quad (14)$$

where \bar{B} is the complementary Young function associated to B . There is a further generalization which turns out to be useful for our purposes (see [O1]): Let A, B, C be Young functions such that

$$A^{-1}(t) \cdot B^{-1}(t) \leq C^{-1}(t),$$

then

$$\|fg\|_{C,Q} \leq 2 \|f\|_{A,Q} \|f\|_{B,Q} \quad (15)$$

We define a natural maximal operator associated to the Young function associated to B .

DEFINITION 2.1 For each locally integrable function f the maximal operator M_B is defined by

$$M_B f(x) = \sup_{x \in Q} \|f\|_{B,Q},$$

where the supremum is taken over all the cubes containing x .

The main examples that we are going to be using are $B(t) = t(1 + \log^+ t)^\alpha$, $\alpha > 0$, with maximal function denoted by $M_{L(\log L)^\alpha}$. The complementary Young function is given by $\bar{B}(t) \approx e^{t^{1/\alpha}}$ with corresponding maximal function denoted by $M_{\exp(L^{1/\alpha})}$.

The boundedness properties of M_B will play a central role to derive sharp two weighted estimates. We need the following class of Young functions.

DEFINITION 2.2 Let $1 < p < \infty$. We say that a doubling Young function B satisfies the B_p condition if there is a positive constant c such that

$$\int_c^\infty \frac{B(t)}{t^p} \frac{dt}{t} \approx \int_c^\infty \left(\frac{t^{p'}}{\bar{B}(t)} \right)^{p-1} \frac{dt}{t} < \infty.$$

This condition provides with a characterization of those maximal operators M_B which are bounded on $L^p(\mathbf{R}^n)$, $1 < p < \infty$. In fact, we have the following Theorem whose proof can be found in [P1].

THEOREM 2.3 Let $1 < p < \infty$. Suppose that B is a doubling Young function. Then the following are equivalent.

i)

$$B \in B_p; \tag{16}$$

ii) there is a constant c such that

$$\int_{\mathbf{R}^n} M_B f(x)^p dx \leq c \int_{\mathbf{R}^n} |f(x)|^p dx \tag{17}$$

for all functions f ;

iii) there is a constant c such that

$$\int_{\mathbf{R}^n} M_B f(x)^p w(x) dx \leq c \int_{\mathbf{R}^n} |f(x)|^p Mw(x) dx \tag{18}$$

for all functions f and all weights w ;

iv) there is a constant c such that

$$\int_{\mathbf{R}^n} Mf(x)^p \frac{w(x)}{[M_{\bar{B}}(u^{1/p})(x)]^p} dx \leq c \int_{\mathbf{R}^n} |f(x)|^p \frac{Mw(x)}{u(x)} dx, \quad (19)$$

for all functions f and all weights w and u .

In the proof of Theorem 1.2 and for $p > 1$ we shall be working with Young functions of the form $B(t) \approx t^p(\log t)^{-1-\epsilon}$ which satisfy the B_p condition and therefore the associated maximal operators $M_{L^p(\log L)^{-1-\epsilon}}$ are bounded on $L^p(\mathbf{R}^n)$.

We conclude this section with a corollary of this Theorem that will be used later on. The result can be seen as a weighted inequality “dual” to the classical Fefferman–Stein inequality

$$\int_{\mathbf{R}^n} Mf(x)^p w(x) dx \leq c \int_{\mathbf{R}^n} |f(x)|^p Mw(x) dx.$$

If M were a linear operator this inequality would imply

$$\int_{\mathbf{R}^n} Mf(x)^{p'} Mw(x)^{1-p'} dy \leq c \int_{\mathbf{R}^n} |f(x)|^{p'} w(x)^{1-p'} dx,$$

which is false in general, however we have the following sharp replacement.

COROLLARY 2.4 *Let $1 < p < \infty$ and let w, u be weights. Then there exists a constant C independent of the weights such that*

$$\int_{\mathbf{R}^n} Mf(x)^{p'} \frac{u(x)}{M^{[p]+1}w(x)^{p'-1}} dx \leq c \int_{\mathbf{R}^n} |f(x)|^{p'} \frac{Mu(x)}{w(x)^{p'-1}} dx \quad (20)$$

for all f . In particular if $u \in A_1$

$$\int_{\mathbf{R}^n} Mf(x)^{p'} \frac{u(x)}{M^{[p]+1}w(x)^{p'-1}} dx \leq c [u]_{A_1} \int_{\mathbf{R}^n} |f(x)|^{p'} \frac{u(x)}{w(x)^{p'-1}} dx \quad (21)$$

Proof: By part iv) of the Theorem we have that $B \in B_{p'}$ if and only if

$$\int_{\mathbf{R}^n} Mf(x)^{p'} \frac{w(x)}{[M_{\bar{B}}(u^{(p'-1)/p'})(x)]^{p'}} dx \leq c \int_{\mathbf{R}^n} |f(x)|^{p'} \frac{Mw(x)}{u(x)^{p'-1}} dx,$$

with a constant independent of f, u, w . If we choose $\bar{B}(t) \approx t^p \log^{p-1+\delta}(1+t)$ then $B \in B_{p'}$ (see definition 2.2) and

$$[M_{\bar{B}}(u^{(p'-1)/p'})]^{p'} = [M_{L^p(\log L)^{p-1+\delta}}(u^{1/p})]^{p'} = [M_{L(\log L)^{p-1+\delta}}(u)]^{p'-1}.$$

Now, selecting $\delta = [p] - p + 1 > 0$ this equals

$$[M_{L(\log L)^{[p]}}(u)]^{p'-1} \approx (M^{[p]+1}u)^{p'-1}$$

by the fact that

$$M^{k+1}f(x) \approx M_{L(\log L)^k}f(x) \tag{22}$$

whenever $k = 0, 1, 2, \dots$. This follows essentially from Stein's observation in [St1] and the details can be found [P1] p. 151. \square

3 Proof of the linear case

The purpose of this section is to prove Theorem 1.2. We recall the definition of a Calderón–Zygmund operator in \mathbf{R}^n . A kernel on $\mathbf{R}^n \times \mathbf{R}^n$ will be a locally integrable function K defined away from the diagonal. K satisfies the standard estimates if there exist positive and finite constants δ, C such that for all distinct $x, y \in \mathbf{R}^n$ and all z such that $2|x-z| < |x-y|$:

- (i) $|K(x, y)| \leq C|x-y|^{-n}$;
- (ii) $|K(x, y) - K(z, y)| \leq C\left(\frac{|x-z|}{|x-y|}\right)^\delta |x-y|^{-n}$;
- (iii) $|K(y, x) - K(y, z)| \leq C\left(\frac{|x-z|}{|x-y|}\right)^\delta |x-y|^{-n}$.

A linear and continuous operator $T : C_0^\infty(\mathbf{R}^n) \rightarrow \mathcal{D}'(\mathbf{R}^n)$ is associated to the kernel K if

$$Tf(x) = \int_{\mathbf{R}^n} K(x, y)f(y) dy,$$

whenever $f \in C_0^\infty(\mathbf{R}^n)$ and x is not in the support of f . There is always such a kernel by the Schwartz kernel Theorem. T is called a Calderón–Zygmund operator

if the associated kernel K satisfies the standard estimates and if it extends to a bounded linear operator on $L^2(\mathbf{R}^n)$. By the standard theory T is also bounded on $L^p(\mathbf{R}^n)$, $1 < p < \infty$, and T is of weak type $(1, 1)$.

The higher order commutators T_b^k $k = 0, 1, 2, \dots$ are generalization of the Calderón–Zygmund operators. Recall that they are defined by $T_b^0 = T$, $T_b^1 = [M_b, T]$ and for $k = 2, \dots$ by $T_b^k = [M_b, T_b^{k-1}]$. Then for $f \in C_0^\infty(\mathbf{R}^n)$ we have

$$T_b^k f(x) = \int (b(x) - b(y))^k K(x - y) f(y) dy$$

when x is not in the support of f .

Our proof of Theorem 1.1 does not follow Coifman’s good- λ inequality. The main tool we use is a variant of the usual sharp maximal operator $M^\#$ of C. Fefferman and E. Stein (cf. [GCRdF]). We denote by M_δ and $M_\delta^\#$ to the operators $M_\delta(f) = M(|f|^\delta)^{1/\delta}$ and $M_\delta^\#(f) = M^\#(|f|^\delta)^{1/\delta}$. Then the Fefferman–Stein’s inequality will be used in the following form.

Let $0 < p, \delta < \infty$ and suppose that $w \in A_\infty(\mathbf{R}^n)$. Then there exists a constant C such that the inequality

$$\int_{\mathbf{R}^n} M_\delta(f)(x)^p w(x) dx \leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} M_\delta^\#(f)(x)^p w(x) dx. \quad (23)$$

holds for every function f for which the left hand side is finite.

The key estimate is the following pointwise inequality which is essentially taken from [P2] where the proof is given for kernels of convolution type. The general case that we are considering here may be treated in the same way with minor modifications and we shall omit the details.

LEMMA 3.1 *For each $b \in BMO$, $0 < \delta < \epsilon < 1$, there exists $C = C_{\delta, \epsilon} > 0$ such that,*

$$M_\delta^\#(T_b^k f)(x) \leq C \|b\|_{BMO} \sum_{j=0}^{k-1} M_\epsilon(T_b^j f)(x) + \|b\|_{BMO}^k M^{k+1} f(x) \quad (24)$$

for all smooth functions f .

It should be mentioned that the idea of relating commutators with the sharp maximal operator is due to J. O. Strömberg (cf. [Ja]). However his original estimate, which is good to get the A_p estimates, is not sharp enough to derive Theorem 1.1.

Proof of Theorem 1.1 Recall that we must prove the inequality

$$\int_{\mathbf{R}^n} |T_b^k f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{kp} [w]_{A_\infty}^{(k+1)p} \int_{\mathbf{R}^n} M^{k+1} f(x)^p w(x) dx. \quad (25)$$

We use an induction argument. The result holds for the case $k = 0$ which is nothing but Coifman's inequality (1), and suppose is true for $k - 1$. Using that L^p is a norm with $p \geq 1$ or the p -triangle inequality for $0 < p < 1$ coupled with the Fefferman–Stein inequality (23) we have

$$\begin{aligned} \|T_b^k f\|_{L^p(w)} &\leq \|M_\delta(T_b^k f)\|_{L^p(w)} \leq [w]_{A_\infty} \|M_\delta^\#(T_b^k f)\|_{L^p(w)} \\ &\leq C [w]_{A_\infty} \|b\|_{BMO} \sum_{j=0}^{k-1} \|M_\epsilon(T_b^j f)\|_{L^p(w)} + C [w]_{A_\infty} \|b\|_{BMO}^k \|M^{k+1} f\|_{L^p(w)}. \end{aligned}$$

Now, since $w \in A_\infty$ there exists $r > 1$ such that $w \in A_r$. Select $\epsilon > 0$ such that $0 < \epsilon < \frac{r}{r-1}$, then by Muckenhoupt's Theorem we continue with

$$\begin{aligned} &C [w]_{A_\infty} \|b\|_{BMO} \sum_{j=0}^{k-1} \|T_b^j f\|_{L^p(w)} + C [w]_{A_\infty} \|b\|_{BMO}^k \|M^{k+1} f\|_{L^p(w)} \\ &\leq C [w]_{A_\infty} \|b\|_{BMO} \sum_{j=0}^{k-1} \|b\|_{BMO}^j [w]_{A_\infty}^{j+1} \|M^{j+1} f\|_{L^p(w)} + C [w]_{A_\infty} \|b\|_{BMO}^k \|M^{k+1} f\|_{L^p(w)} \\ &\leq C [w]_{A_\infty}^{k+1} \|b\|_{BMO}^k \|M^k f\|_{L^p(w)} + C [w]_{A_\infty} \|b\|_{BMO}^k \|M^{k+1} f\|_{L^p(w)} \\ &\leq C [w]_{A_\infty}^{k+1} \|b\|_{BMO}^k \|M^{k+1} f\|_{L^p(w)} \end{aligned}$$

by the Lebesgue differentiation Theorem and since $[w]_{A_\infty} \geq 1$. \square

Proof of Theorem 1.2

As we showed in §1.1 everything is reduced to prove the following two weighted norm inequality for M^{k+1}

$$\int_{\mathbf{R}^n} M^{k+1}(f)^{p'} (M^{[(k+1)p]+1} w)^{1-p'} \leq C \int_{\mathbf{R}^n} |f|^{p'} w^{1-p'}, \quad (26)$$

which is equivalent to

$$\int_{\mathbf{R}^n} M^{k+1}(f w^{1/p})^{p'} (M^{[(k+1)p]+1} w)^{1-p'} \leq C \int_{\mathbf{R}^n} |f|^{p'}.$$

By (22) everything is reduced to prove

$$\int_{\mathbf{R}^n} M_{L(\log L)^k} (fw^{1/p})^{p'} (M^{[(k+1)p]+1}w)^{1-p'} \leq C \int_{\mathbf{R}^n} |f|^{p'}.$$

Now, using the notation $\Phi_k(t) = t(1 + \log^+ t)^k$ we have for large t that $\Phi_k^{-1}(t) \approx \frac{t}{(\log t)^k}$ and for $\epsilon > 0$ we put

$$\begin{aligned} \Phi_k^{-1}(t) &\approx \frac{t}{(\log t)^k} = \frac{t^{1/p}}{(\log t)^{k + \frac{p-1+\epsilon}{p}}} \times t^{1/p'} (\log t)^{\frac{p-1+\epsilon}{p}} \\ &= A^{-1}(t) \times B^{-1}(t), \end{aligned}$$

where $A(t) \approx t^p (\log t)^{(k+1)p-1+\epsilon}$ and $B(t) \approx \frac{t^{p'}}{(\log t)^{1+(p'-1)\epsilon}}$ (see [O2]). Therefore by the generalized Hölder's inequality (15) we have

$$\begin{aligned} M_{L(\log L)^k} (fw^{1/p}) &\leq M_{L^{p'}(\log L)^{-(1+(p'-1)\epsilon)}}(f) M_{L^p(\log L)^{(k+1)p-1+\epsilon}}(w^{1/p}) \\ &\approx M_{L^{p'}(\log L)^{-(1+(p'-1)\epsilon)}}(f) M_{L(\log L)^{(k+1)p-1+\epsilon}}(w)^{1/p}. \end{aligned}$$

Using this and selecting $\epsilon > 0$ such that $(k+1)p - 1 + \epsilon = [(k+1)p]$ we have

$$\begin{aligned} &\int_{\mathbf{R}^n} M_{L(\log L)^k} (fw^{1/p})^{p'} (M^{[(k+1)p]+1}w)^{1-p'} \\ &\leq \int_{\mathbf{R}^n} M_{L^{p'}(\log L)^{-(1+(p'-1)\epsilon)}}(f)^{p'} M_{L(\log L)^{(k+1)p-1+\epsilon}}(w)^{p'/p} (M^{[(k+1)p]+1}w)^{1-p'} \\ &\approx \int_{\mathbf{R}^n} M_{L^{p'}(\log L)^{-(1+(p'-1)\epsilon)}}(f)^{p'} M_{L(\log L)^{[(k+1)p]}}(w)^{p'-1} (M^{[(k+1)p]+1}w)^{1-p'} \\ &\approx \int_{\mathbf{R}^n} M_{L^{p'}(\log L)^{-(1+(p'-1)\epsilon)}}(f)^{p'} \end{aligned}$$

using (22) and then Theorem 2.3. □

4 Proof of the nonlinear case

For the Proof of Theorem 1.3 we need the following two lemmas about classes of weights.

Recall that a weight v belongs to \mathbf{RH}_∞ if there exists a positive constant c such that for all cubes Q

$$\sup_Q v \leq \frac{c}{|Q|} \int_Q v(y) dy.$$

LEMMA 4.1 *Let $u \in A_1$, and $v \in \mathbf{RH}_\infty$. Then*

$$uv \in A_\infty.$$

Indeed, since $u \in A_1 \subset A_\infty$, there exists a $q > 1$ such that

$$\left(\frac{1}{|Q|} \int_Q u(y)^q dy \right)^{1/q} \leq \frac{C}{|Q|} \int_Q u(y) dy \leq C \inf_Q u$$

Then for the same q we have

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q (u(y)v(y))^q dy \right)^{1/q} &\leq \sup_Q v \left(\frac{1}{|Q|} \int_Q u(y)^q dy \right)^{1/q} \\ &\leq \frac{C}{|Q|} \int_Q v(y) dy \inf_Q u \leq \frac{C}{|Q|} \int_Q v(y)u(y) dy, \end{aligned}$$

and this means that $vu \in A_\infty$.

LEMMA 4.2 *Let $\lambda > 0$, then if $M\mu$ is finite almost everywhere*

$$(M\mu)^{-\lambda} \in \mathbf{RH}_\infty.$$

For the proof of this let $r > \lambda + 1$; then $[M\mu]^{\frac{\lambda}{r-1}} \in A_1$ and therefore

$$\begin{aligned} \sup_Q (M\mu)^{-\lambda} &= \sup_Q [(M\mu)^{\frac{\lambda}{r-1}}]^{1-r} = [\inf_Q (M\mu)^{\frac{\lambda}{r-1}}]^{1-r} \\ &\leq C \left[\frac{1}{|Q|} \int_Q (M\mu)^{\frac{\lambda}{r-1}} \right]^{1-r} \leq C \frac{1}{|Q|} \int_Q (M\mu)^{-\lambda} \end{aligned}$$

by Hölder's inequality.

Proof of Theorem 1.3

As we mentioned earlier we split the nonlinear commutator N in following way

$$\begin{aligned} Nf &= T(f \log |f|) - Tf \log |Tf| = \\ &= T(f \log |f|) - T(f \log Mf) + T(f \log Mf) - Tf \log Mf + Tf \log Mf - Tf \log |Tf| \\ &= T\left(f \log \frac{|f|}{Mf}\right) + [\log Mf, T]f - Tf \log \frac{|Tf|}{Mf} = N_1f + N_2f + N_3f. \end{aligned}$$

Then for any weight w

$$\|Nf\|_{L^p(w)} \leq \|N_1f\|_{L^p(w)} + \|N_2f\|_{L^p(w)} + \|N_3f\|_{L^p(w)}.$$

To estimate N_1 we write

$$N_1f = T\left(Mf \frac{f}{Mf} \log\left(\frac{|f|}{Mf}\right)\right)$$

and we apply Theorem 1.1 with $k = 0$ since by hypothesis $w \in A_\infty$:

$$\begin{aligned} \int_{\mathbf{R}^n} |N_1f|^p w &\leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} \left(M\left(Mf \frac{f}{Mf} \log\left(\frac{|f|}{Mf}\right)\right)\right)^p w \\ &\leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} (M^2f)^p w \end{aligned}$$

since $|t \log t| \leq \frac{1}{e}$, $0 < t \leq 1$.

For N_2 we can apply an small variation of Theorem 1.1 with $k = 1$. Indeed we need to consider a commutator but where the symbol depends upon the function itself namely,

$$[b, T]f = b(f)T(f) - T(b(f)f).$$

In our context $b(f) = \log Mf \in BMO$ with a constant independent of f and hence by Theorem 1.1 (observe that there is no change in the proof since Lemma 4.2 still works)

$$\int_{\mathbf{R}^n} |N_2f|^p w = \int_{\mathbf{R}^n} |[\log Mf, T]f|^p w \leq C \|\log Mf\|_{BMO}^{2p} [w]_{A_\infty}^p \int_{\mathbf{R}^n} (M^2f)^p w$$

$$\leq C [w]_{A_\infty}^p \int_{\mathbf{R}^n} (M^2 f)^p w.$$

For the last term we split \mathbf{R}^n in two disjoint sets A and B where $A = \{y \in \mathbf{R}^n : |Tf(y)| \leq Mf(y)\}$ and $B = \{y \in \mathbf{R}^n : |Tf(y)| > Mf(y)\}$. Writing

$$N_3 f = Mf \frac{|Tf|}{Mf} \log \frac{|Tf|}{Mf} =$$

Using in B that $\log t \leq C_\epsilon \frac{t^\epsilon}{\epsilon}$, $t > 1$, $\epsilon > 0$ we have the following

$$\int_{\mathbf{R}^n} |N_3 f|^p w \leq C_\epsilon \int_A (Mf)^p w + C_\epsilon \int_B |Tf|^{p(\epsilon+1)} (Mf)^{-\epsilon p} w$$

We would like to apply again Theorem 1.1 with $k = 0$. To do this we must show that $w (Mf)^{-\epsilon p} w \in A_\infty$ for ϵ small enough and with a constant independent of f . (recall that ϵ is still available). Indeed, since $w \in A_\infty$ $w \in A_q$ for some $q > 1$ and by the factorization (cf. [GCRdF] p. 436) theorem $w = w_1 w_2^{1-q}$ where w_1 and w_2 are A_1 weights. Then

$$w (Mf)^{-\epsilon p} = w_1 w_2^{1-q} (Mf)^{-\epsilon p} = w_1 (w_2 (Mf)^{\frac{\epsilon p}{q-1}})^{1-q}.$$

By the factorization theorem it is enough to show that $w_2 (Mf)^{\frac{\epsilon p}{q-1}} \in A_1$ for ϵ small enough. To do this we fix a cube Q , and arbitrary a.e. $x \in Q$. Then

$$\frac{1}{|Q|} \int_Q w_2 (Mf)^{\frac{\epsilon p}{q-1}} \leq \left(\frac{1}{|Q|} \int_Q w_2^r \right)^{1/r} \left(\frac{1}{|Q|} \int_Q (Mf)^{\frac{\epsilon r' p}{q-1}} \right)^{1/r'}. \quad (27)$$

Now, since $w_2 \in A_\infty$ we can pick $r > 1$ such that we can continue with

$$\frac{C}{|Q|} \int_Q w_2 \left(\frac{1}{|Q|} \int_Q (Mf)^{\frac{\epsilon r' p}{q-1}} \right)^{1/r'},$$

and if we pick ϵ with $0 < \epsilon < \frac{q-1}{p r'}$ then $(Mf)^{\frac{\epsilon r' p}{q-1}} \in A_1$ and then this is less or equal than

$$\frac{C}{|Q|} \int_Q w_2 (Mf(x))^{\frac{\epsilon p}{q-1}} \leq C w_2(x) (Mf(x))^{\frac{\epsilon p}{q-1}}.$$

An important observation is that the A_∞ norm does not depend on f . □

Proof of Theorem 1.5 Proceeding as before we have

$$\|Nf\|_{L^p(w)} \leq \|N_1f\|_{L^p(w)} + \|N_2f\|_{L^p(w)} + \|N_3f\|_{L^p(w)},$$

and the proof of the two first pieces are similar to the previous case. Recall that there is no assumption on w . For N_1 we combine Theorem 1.2 for $k = 0$, the fact that $|t \log t| \leq \frac{1}{e}$, $0 < t \leq 1$, together with the classical Fefferman–Stein inequality

$$\int_{\mathbf{R}^n} (Mf)^p w \leq c \int_{\mathbf{R}^n} |f|^p Mw$$

to obtain

$$\begin{aligned} \int_{\mathbf{R}^n} |N_1f|^p w &= \int_{\mathbf{R}^n} |T(Mf \frac{f}{Mf} \log(\frac{|f|}{Mf}))|^p w \\ &\leq C \int_{\mathbf{R}^n} (Mf)^p M^{[p]+1} w \leq C \int_{\mathbf{R}^n} |f(y)|^p M^{[p]+2} w \\ &\leq C \int_{\mathbf{R}^n} |f|^p M^{[2p]+1} w. \end{aligned}$$

For N_2 we use Theorem 1.2 since the BMO norm of $\log Mf$ is independent of f

$$\begin{aligned} \int_{\mathbf{R}^n} |N_2f|^p w &= \int_{\mathbf{R}^n} |[\log Mf, T]f|^p w \\ &\leq \|\log Mf\|_{BMO}^{2p} \int_{\mathbf{R}^n} |f|^p M^{[2p]+1} w \leq C \int_{\mathbf{R}^n} |f|^p M^{[2p]+1} w. \end{aligned}$$

For the last term N_3 we start as above with

$$\begin{aligned} \int_{\mathbf{R}^n} |N_3f|^p w &\leq C_\epsilon \int_{\mathbf{R}^n} (Mf)^p w + C_\epsilon \int_{\mathbf{R}^n} |Tf|^{p(\epsilon+1)} (Mf)^{-\epsilon p} w, \\ &\leq C_\epsilon \int_{\mathbf{R}^n} |f|^p Mw + C_\epsilon \int_{\mathbf{R}^n} |Tf|^{p(\epsilon+1)} (Mf)^{-\epsilon p} w. \end{aligned}$$

The key point of the proof is to understand the last term. We are tempted in applying Theorem 1.1 with $k = 0$ replacing T by M which would finish the proof of the Theorem; however, there is no assumption on w in such a way that we cannot say that the weight on the right hand side, namely $(Mf)^{-\epsilon p} w$, is an A_∞ weight. We may argue as follows.

Let $p_\epsilon = p(\epsilon + 1)$, then there exist a function $g \in L^{(p_\epsilon)'}$ with unit norm such that

$$\left(\int_{\mathbf{R}^n} |Tf|^{p_\epsilon} (Mf)^{-\epsilon p} w \right)^{\frac{1}{p_\epsilon}} = \left\| Tf (Mf)^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}} \right\|_{L^{p_\epsilon}(\mathbf{R}^n)} = \int_{\mathbf{R}^n} Tf (Mf)^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}} g.$$

Since the adjoint operator T^* is also a Calderón–Zygmund operator bounded on all the L^p spaces as well we can equal last expression to

$$\begin{aligned} \int_{\mathbf{R}^n} f T^*(g (Mf)^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}}) &= \int_{\mathbf{R}^n} f \frac{(M^k w)^{\frac{1}{p_\epsilon}}}{(Mf)^{\frac{\epsilon}{\epsilon+1}}} T^*(g (Mf)^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}}) \frac{(Mf)^{\frac{\epsilon}{\epsilon+1}}}{(M^k w)^{\frac{1}{p_\epsilon}}} \\ &\leq \left(\int_{\mathbf{R}^n} |f|^{p_\epsilon} \frac{M^k w}{(Mf)^{\frac{\epsilon}{\epsilon+1} p_\epsilon}} \right)^{\frac{1}{p_\epsilon}} \left(\int_{\mathbf{R}^n} |T^*(g (Mf)^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}})|^{(p_\epsilon)'} \frac{(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}}{(M^k w)^{\frac{(p_\epsilon)'}}}{p_\epsilon}} \right)^{\frac{1}{(p_\epsilon)'}} \\ &= I \times II, \end{aligned}$$

where k is an integer to be chosen in a moment. To estimate I we simply use the Lebesgue differentiation Theorem

$$I = \left(\int_{\mathbf{R}^n} |f|^{p_\epsilon} |f|^p \frac{M^k w}{(Mf)^{p_\epsilon}} \right)^{\frac{1}{p_\epsilon}} \leq \left(\int_{\mathbf{R}^n} |f|^p M^k w \right)^{\frac{1}{p_\epsilon}},$$

where k is still available. For the last term II we are going to replace T^* by the Hardy–Littlewood maximal function using again Theorem 1.1 with $k = 0$ and since T^* is also a Calderón–Zygmund operator. All we have to do is to show that the weight

$$u = \frac{(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}}{(M^k w)^{\frac{(p_\epsilon)'}}{p_\epsilon}} = (Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'} (M^k w)^{1-(p_\epsilon)'}$$

is an A_∞ weight with constant independent of f . To do this observe first that $(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}$ is in A_1 since $\frac{\epsilon}{\epsilon+1} (p_\epsilon)' = \frac{\epsilon}{\epsilon+1/p'} < 1$ and $(M^k w)^{1-(p_\epsilon)'}$ is in RH_∞ by Lemma 4.2 where the constants are independent of both f and w . Therefore $u \in A_\infty$ by Lemma 4.1 above and we have applying Theorem 1.1 that

$$II \leq C \left(\int_{\mathbf{R}^n} M(g Mf^{-\frac{\epsilon}{\epsilon+1}} w^{\frac{1}{p_\epsilon}})^{(p_\epsilon)'} \frac{(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}}{(M^k w)^{(p_\epsilon)'-1}} \right)^{\frac{1}{(p_\epsilon)'}}.$$

Finally we can apply Corollary 2.4 with $k = [p_\epsilon] + 1$ using as we pointed out above that $(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}$ is in A_1 . Then

$$II \leq C \left(\int_{\mathbf{R}^n} |g|^{(p_\epsilon)'} (Mf)^{-\frac{\epsilon}{\epsilon+1} (p_\epsilon)'} w^{\frac{(p_\epsilon)'}{p_\epsilon}} \frac{(Mf)^{\frac{\epsilon}{\epsilon+1} (p_\epsilon)'}}{w^{(p_\epsilon)'-1}} \right)^{\frac{1}{(p_\epsilon)'}}$$

$$= C \left(\int_{\mathbf{R}^n} |g|^{(p\epsilon)'} \right)^{\frac{1}{(p\epsilon)'}} = C.$$

Combining all these inequalities we get that

$$\int_{\mathbf{R}^n} |Tf|^{p(\epsilon+1)} Mf^{-\epsilon p} w \leq C \int_{\mathbf{R}^n} |f|^p M^{[p\epsilon]+1} w = C \int_{\mathbf{R}^n} |f|^p M^{[p]+1} w$$

with ϵ small enough. Therefore we have

$$\int_{\mathbf{R}^n} |N_3 f(y)|^p w(y) dy \leq \int_{\mathbf{R}^n} |f(y)|^p M^{[p]+1} w(y) dy$$

which combined with the estimates for N_1 and N_2 yield the final result. Observe that the piece corresponding to N_3 behaves more as the usual singular integral and that the “worst” piece corresponds to N_2 . \square

5 A counterexample

We end the paper by showing that Theorem 1.2, and consequently the others, is optimal.

Consider the classical Hilbert transform

$$Hf(x) = pv \int_{\mathbf{R}} \frac{f(y)}{x-y} dy,$$

and let $m = 1, 2, \dots$ be the largest exponent for which the following inequality does not hold

$$\int_{\mathbf{R}^n} |H_b^k f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{kp} \int_{\mathbf{R}^n} |f(x)|^p M^m w(x) dx. \quad (28)$$

By duality this is equivalent to showing

$$\int_{\mathbf{R}^n} |H_b^k f(x)|^{p'} M^m w(x)^{1-p'} dx \leq C \|b\|_{BMO}^{kp'} \int_{\mathbf{R}^n} |f(x)|^{p'} w(x)^{1-p'} dx$$

Consider the *BMO* function $b(x) = \log|x|$ and let $f = w = \chi_{(0,1)}$ so that the right hand is equal to a finite constant. For the left hand side we use that for $x > e$

$$|H_b^k f(x)| \approx \frac{(\log x)^k}{x} \approx M^{k+1} f(x).$$

Then

$$\begin{aligned} & \int_{\mathbf{R}} |H_b^k f(x)|^{p'} M^m w(x)^{1-p'} dx \geq \\ & \geq \int_{x>e} \left(\frac{(\log x)^k}{x} \right)^{p'} \left(\frac{(\log x)^{m-1}}{x} \right)^{1-p'} dx = \int_{x>1} x^{kp'-(m-1)(p'-1)+1} \frac{dx}{x}, \end{aligned}$$

which becomes unbounded when $m \leq [(k+1)p]$. Therefore (28) is false when $m = [(k+1)p]$ and then inequality (5) in Theorem 1.2 is optimal.

References

- [CWW] S. Y. A. Chang, J. M. Wilson, and T. H. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helvetici **60** (1985), 217–286.
- [CFL1] F. Chiarenza, M. Frasca y P. Longo, *Interior $W^{2,p}$ estimates for non divergence elliptic equations with discontinuous coefficients*, Recherche Mat. **40** (1991), 149–168.
- [CFL2] F. Chiarenza, M. Frasca y P. Longo, *$W^{2,p}$ -solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients*, Trans. Amer. Math. Soc. **334** (1993), 841–853.
- [C] R. Coifman, *Distribution function inequalities for singular integrals*, Proc. Acad. Sci. U.S.A. **69** (1972), 2838–2839.
- [CRW] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. **103** (1976), 611–635.
- [DiR] G. Di Fazio y M. A. Ragusa, *Interior estimates in Morrey spaces for strong solutions to nondivergence form equations with discontinuous coefficients*, J. of Functional Analysis **112** (1993), 241–256.
- [GCRdF] J. Garcia-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North Holland Math. Studies **116**, North Holland, Amsterdam, (1985).

- [GI] L. Greco y T. Iwaniec, *New inequalities for the Jacobian*, Ann. Inst. Henri Poincaré, **11** (1994), 17–35.
- [HMW] R. A. Hunt, B. Muckenhoupt and R. L. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. **176** (1973), 227–252.
- [IS] T. Iwaniec y C. Sbordone, *Weak minima of variational integrals*, J. Reine Angew Math. **454**, 143–161.
- [Ja] S. Janson, *Mean oscillation and commutators of singular integral operators*, Ark. Mat. **16**, (1978), 263–270.
- [J] J. L. Journé, *Calderón–Zygmund operators, pseudo–differential operators and the Cauchy integral of Calderón*, Lect. Notes Math. **994**, Springer Verlag, (1983).
- [M] M. Milman, *Extrapolation and Optimal Decompositions*, Lect. Notes Math. **1580**, Springer Verlag, (1995).
- [O1] R. O’Neil, *Fractional integration in Orlicz spaces*. Trans. Amer. Math. Soc. **115**, 300–328 (1963).
- [O2] R. O’Neil, *Integral transforms and tensor products on Orlicz spaces and $L_{p,q}$ spaces*. J. D’Anal. Math. **21**, 1–276 (1968).
- [P1] C. Pérez, *On sufficient conditions for the boundedness of the Hardy–Littlewood maximal operator between weighted L^p –spaces with different weights*, Proc. of the London Math. Soc. (3) **71** (1995), 135–157.
- [P2] C. Pérez, *Endpoint estimates for commutators of singular integral operators*, J. of Functional Analysis **128** (1995), 163–185.
- [P3] C. Pérez, *Weighted norm inequalities for singular integral operators*, J. London Math. Soc. **49** (1994), 296–308.
- [RW] R. Rochberg and G. Weiss, *Derivatives of analytic families of Banach spaces*, Ann. of Math. **118** (1983).
- [St1] E. M. Stein, *Note on the class $L \log L$* , Studia Math. **32** (1969), 305–310.

[Wil] J. M. Wilson, *Weighted norm inequalities for the continuous square functions*,
Trans. Amer. Math. Soc. **314** (1989), 661–692.