J. London Math. Soc. (2) 65 (2002) 672–692
DOI: 10.1112/S0024610702003174

SHARP WEIGHTED ESTIMATES FOR MULTILINEAR COMMUTATORS

C. PÉREZ AND R. TRUJILLO-GONZÁLEZ

Abstract

Multilinear commutators with vector symbol $\vec{b} = (b_1, \ldots, b_m)$ defined by

$$T_{\vec{b}}(f)(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^{m} (b_j(x) - b_j(y)) \right] K(x,y) f(y)\, dy$$

are considered, where K is a Calderón–Zygmund kernel. The following a priori estimates are proved for $w \in A_\infty$. For $0 < p < \infty$, there exists a constant C such that

$$\| \dot{T}_{\vec{b}}(f) \|_{L^p(w)} \leqslant C \| \vec{b} \| \, \| M_{L(\log L)^{1/r}}(f) \|_{L^p(w)}$$

and

$$\sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{ y \in \mathbb{R}^n : |T_{\vec{b}} f(y)| > t \}) \leqslant C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{ y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(\| \vec{b} \| f)(y) > t \}),$$

where

$$\| \vec{b} \| = \prod_{j=1}^{m} \| b_j \|_{osc_{\exp L^{r_j}}},$$

$$\Phi(t) = t \log^{1/r}(e + t), \qquad \frac{1}{r} = \frac{1}{r_1} + \ldots + \frac{1}{r_m},$$

and $M_{L(\log L)^\alpha}$ is an Orlicz type maximal operator. This extends, with a different approach, classical results by Coifman.

As a corollary, it is deduced that the operators $T_{\vec{b}}$ are bounded on $L^p(w)$ when $w \in A_p$, and that they satisfy corresponding weighted $L(\log L)^{1/r}$-type estimates with $w \in A_1$.

1. *Introduction*

The main purpose of this paper is to prove sharp estimates for multilinear commutators involving nonstandard symbols. These will be established by means of appropriate maximal operators that somehow control the commutators. This illustrates the classical Calderón–Zygmund principle, which roughly states that any singular integral operator is controlled by a suitable maximal operator.

1.1. *Background*

Motivated by the work of Calderón on commutators, Coifman, Rochberg and Weiss introduced in **[8]** the operator

$$T_b f(x) = \int_{\mathbb{R}^n} (b(x) - b(y)) K(x,y) f(y)\, dy, \tag{1.1}$$

Received 15 June 2000; revised 26 June 2001.

2000 *Mathematics Subject Classification* 42B20, 42B25.

The first author was partially supported by DGICYT grant PB40192. The second author was partially supported by Gobierno de Canarias PI1999/105.

where K is a kernel satisfying the standard Calderón–Zygmund estimates (see Section 3.1), and b, the 'symbol' of the operator, is any locally integrable function. The operator is called a commutator since $T_b = [b, T] = bT - T(b\cdot)$, where T is the Calderón–Zygmund singular integral operator associated to K. The main result from [**8**] states that $[b, T]$ is a bounded operator on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, when the symbol b is a BMO function.

These commutators have proved to be of interest in many situations. We only mention the recent results in the theory of non-divergent elliptic equations with discontinuous coefficients [**3**, **4**, **9**]. There is also an interesting connection, as pointed out in [**21**], with the nonlinear commutator considered by Rochberg and Weiss in [**23**] and defined by

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

This in turn is related to the Jacobian mapping of vector functions and with nonlinear partial differential equations, as shown in [**13**, **14**].

A natural generalization of the commutator $[b, T]$ is given by

$$T_b^m f(x) = \int_{\mathbb{R}^n} (b(x) - b(y))^m K(x, y) f(y)\, dy, \tag{1.2}$$

where $m \in \mathbb{N}$. The case $m = 0$ recaptures the Calderón–Zygmund singular integral operator.

It was shown in [**21**] that there is an intimate connection between the commutator T_b^m and iterations of the maximal operators. Indeed, the main theorem from [**8**] was sharpened in [**21**] as follows: for any $0 < p < \infty$ and any $w \in A_\infty$, there is a constant C such that

$$\int_{\mathbb{R}^n} |T_b^m f(x)|^p w(x)\, dx \leqslant C \|b\|_{\mathrm{BMO}}^{mp} \int_{R^n} (M^{m+1} f(x))^p w(x)\, dx, \tag{1.3}$$

where M^{m+1} denotes the $m+1$ iterations of the Hardy–Littlewood maximal operator, namely

$$M^m = \overbrace{M \circ \ldots \circ M}^{(m \text{ times})}.$$

Furthermore, this inequality is sharp, since, by the Lebesgue differentiation theorem, M^{m+1} cannot be replaced by the smaller operator M^m.

Estimate (1.3) can also be seen as a generalization of a by now classical result of Coifman for Calderón–Zygmund singular integral operators. See [**6**, **7**].

A weak version of inequality (1.3) is obtained in [**20**]. Indeed, if we let $\Phi_m(t) = t \log^m(e + t)$, for any $w \in A_\infty$ and $b \in$ BMO there is a constant C such that

$$\sup_{t>0} \frac{1}{\Phi_m(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_b^m f(y)| > t\}) \leqslant C \sup_{t>0} \frac{1}{\Phi_m(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M^{m+1} f(y) > t\}). \tag{1.4}$$

A priori inequalities of the form (1.3) (or (1.4)) encode a considerable amount of information about the behavior of the operator. The first observation is that in any case they reflect the higher degree of singularity of T_b^m, as compared with T, since a larger operator than M, namely M^{m+1}, is needed to balance the inequality. As a second instance, if $p > 1$, we can apply Muckenhoupt's theorem $m + 1$ times to conclude from (1.3) the well known fact that high-order commutators are bounded on $L^p(w)$ whenever $w \in A_p$. It should be mentioned that this A_p estimate follows, as is well known, from an estimate due to Strömberg (see [**15**, p. 268; **25**, p. 417]).

This method is powerful, and it can be applied to various situations. As an example, we refer the reader to [11] for a very nice application to commutators with strongly singular integral operators. However, this estimate of Strömberg is not sharp enough to derive (1.3) or (1.4). In fact, the following (($L\log L$)-type) endpoint estimate first deduced in [20] can be shown from (1.4). Let $w \in A_1$ and $b \in \mathrm{BMO}$; then there exists a constant C such that, for all $\lambda > 0$,

$$w(\{y \in \mathbb{R}^n : |T_b^m f(y)| > \lambda\}) \leqslant C \int_{\mathbb{R}^n} \frac{|f(y)|}{\lambda} \left(1 + \log^+ \left(\frac{|f(y)|}{\lambda}\right)\right)^m w(y)\, dy. \tag{1.5}$$

As a third instance, it was shown in [21] that estimate (1.3) implies sharp two weighted inequalities for the commutator of the form

$$\int_{\mathbb{R}^n} |T_b^m f(x)|^p w(x)\, dx \leqslant C \|b\|_{\mathrm{BMO}}^{mp} \int_{R^n} |f(x)|^p M^{[(m+1)p]+1} w(x)\, dx,$$

where no condition on the weight w is assumed. This result is an extension of the case $m = 0$ proved in [18] that generalized some previous partial results by Wilson [26]. The approach considered in [18] is different from that in [26], and it combines (1.3) with certain sharp two weighted estimates for the Hardy–Littlewood maximal function derived in [19].

Finally, there is a relationship between (1.4) and the endpoint behavior of the operator. Indeed, an interesting observation that occurs in the development of the theory of commutators is that the L^p theory, $p > 1$, was developed without an appeal to any endpoint estimate. On the other hand, it is well known that in the classical Calderón–Zygmund L^p theory, $p > 1$, a crucial step is to show that the operators are of weak type $(1, 1)$. However, this is not the case with the commutator $[b, T]$ when $b \in \mathrm{BMO}$ as shown in [20] and estimate (1.5) is the correct replacement. This can be seen as another way of expressing the fact that commutators have a higher degree of singularity than the Calderón–Zygmund singular integral operators.

1.2. *Results of the paper*

In this paper, we obtain similar estimates for a wider class of commutators. Given a Calderón–Zygmund singular integral operator T with kernel K in $\mathbb{R}^n \times \mathbb{R}^n$ and a vector $\vec{b} = (b_1, \ldots, b_m)$ of locally integrable functions, we define the multilinear operator

$$T_{\vec{b}} f(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^m (b_j(x) - b_j(y))\right] K(x, y) f(y)\, dy, \tag{1.6}$$

generalizing the commutator (1.2). We refer the reader to [7] for an extensive study of multilinear operators.

We will be considering the following class of symbols: for $r \geqslant 1$ and for any locally integrable function f, we define

$$\|f\|_{osc_{\exp L^r}} = \sup_Q \|f - f_Q\|_{\exp L^r, Q},$$

which is the supremum taken over all the cubes Q with sides parallel to the axes. Here $\|g\|_{\exp L^r, Q}$ is the mean value of g on the cube Q with respect to the Young function $\Phi(t) = e^{t^r} - 1$. See Section 2.2 for the precise definition. As usual, f_Q denotes the average of f on Q. For $r \geqslant 1$, we define the space $Osc_{\exp L^r}$ by

$$Osc_{\exp L^r} = \{f \in L^1_{loc}(\mathbb{R}^n) : \|f\|_{Osc_{\exp L^r}} < \infty\}.$$

In the particular case of $r = 1$, $Osc_{\exp L^1}$ coincides with the BMO space by the John–Nirenberg theorem. Also we have $Osc_{\exp L^r} \subsetneq$ BMO for any $r > 1$. Other examples are provided by Trudinger's inequality for Riesz potentials. To be more precise, for any $0 < \alpha < n$ and any $f \in L^{n/\alpha}(\mathbb{R}^n)$, the Riesz potential $I_\alpha f$ of order α belongs to $Osc_{\exp L^{(n/\alpha)'}}$ (cf. [**12**, **27**]).

For the statement of our results, we introduce some notation that will simplify the presentation. Throughout this paper, m will always be the number of symbols of the operator $T_{\vec{b}}$ where $\vec{b} = (b_1, \dots, b_m)$ is a family of m locally integrable functions. If $r_1, \dots, r_m$ are m positive real numbers, we denote

$$\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}$$

and

$$\|\vec{b}\| = \prod_{j=1}^{m} \|b_j\|_{Osc_{\exp L^{r_j}}}.$$

Our main results are the following.

THEOREM 1.1. *Let $0 < p < \infty$, $w \in A_\infty$. Suppose that $T_{\vec{b}}f$ is the commutator* (1.6)*, where b is as above such that $b_i \in Osc_{\exp L^{r_i}}$, $r_i \geqslant 1$, $1 \leqslant i \leqslant m$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x)\,dx \leqslant C\|\vec{b}\|^p \int_{\mathbb{R}^n} (M_{L(\log L)^{1/r}}(f)(x))^p w(x)\,dx \tag{1.7}$$

for all bounded functions f with compact support.

For the precise definition of the operator $M_{L(\log L)^\alpha}$, see Section 2.2.

Inequality (1.7) shows that the maximal operator $M_{L(\log L)^{1/r}}$ is the one that controls the multilinear commutators T_b. Since $r_i \geqslant 1$ for each i, it follows that $M_{L(\log L)^{1/r}}$ is pointwise smaller than $M_{L(\log L)^m}$, and this in turn is known to be equivalent to $m+1$ iterations of the Hardy–Littlewood maximal operator M^{m+1}. Hence we can use Muckenhoupt's theorem again and deduce the boundedness of T_b on $L^p(w)$ for $p > 1$ and any $w \in A_p$.

COROLLARY 1.2. *Let $1 < p < \infty$ and $w \in A_p$. Suppose that $T_{\vec{b}}f$ is the commutator* (1.6)*, where b is as above such that $b_i \in Osc_{\exp L^{r_i}}$, $r_i \geqslant 1$, $1 \leqslant i \leqslant m$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x)\,dx \leqslant C\|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx \tag{1.8}$$

for all bounded functions f with compact support.

As mentioned above, this estimate is a generalized version of Coifman's result in [**6**]. However, our approach is different, and it is based on a pointwise inequality that can roughly be expressed as

$$M^{\#}_\delta(T_{\vec{b}}f)(x) \leqslant C\|\vec{b}\| M_{L(\log L)^{1/r}} f(x) + R(f)(x) \tag{1.9}$$

for an appropriate positive but small enough number δ. R denotes a certain 'remainder' operator that is smoother and less singular than $M_{L(\log L)^{1/r}}$ in some suitable sense. See Lemma 3.1 for a complete and precise statement of this estimate.

Moreover, reasoning as in [**21**, Theorem 2], from this theorem we deduce the following estimate for general weights. As usual, $[\alpha]$ denotes the integer part of α.

THEOREM 1.3. *Let $1 < p < \infty$, and let $T_{\vec{b}}$ be as in Theorem 1.1. Then there exists a constant C such that, for any weight w,*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x)\,dx \leqslant C\|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p M^{[(1/r+1)p]+1}(w)(x)\,dx$$

for all bounded functions f with compact support.

REMARK 1.4. We remark that the number of iterations of the maximal function needed in the theorem is optimal, as can be seen in [**21**, Section 5]. In fact, it follows from the proof of Theorem 1.3 that there is a sharper estimate

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x)\,dx \leqslant C\|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p M_{L(\log L)^{(1/r+1)p-1+\varepsilon}}(w)(x)\,dx,$$

where $\epsilon > 0$, the result being false for $\epsilon = 0$.

The key pointwise estimate (1.9) is also used to derive the next endpoint result. Recall that commutators with BMO functions are not of weak type $(1,1)$.

Also, recall that we denote $\|\vec{b}\| = \prod_{j=1}^m \|b_j\|_{Osc_{\exp L^{r_j}}}$ and $\Phi(t) = t\log^{1/r}(e+t)$, where

$$\frac{1}{r} = \frac{1}{r_1} + \ldots + \frac{1}{r_m}.$$

THEOREM 1.5. *Let $w \in A_1$. There exists a constant $C > 0$ such that, for all $\lambda > 0$,*

$$w(\{y \in \mathbb{R}^n : |T_{\vec{b}}f(y)| > \lambda\}) \leqslant C \int_{\mathbb{R}^n} \Phi\left(\frac{\|\vec{b}\||f(y)|}{\lambda}\right) w(y)\,dy \tag{1.10}$$

for all bounded functions f with compact support and for all $\vec{b}$.

The proof is based on the following result, which generalizes inequality (1.4).

THEOREM 1.6. *Let $w \in A_\infty$. Then there exists a positive constant C such that*

$$\sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_{\vec{b}}f(y)| > t\}) \leqslant C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M_\Phi(\|\vec{b}\|f)(y) > t\}) \tag{1.11}$$

for all bounded functions f with compact support.

In any of the above results, if we choose $b_1 = \ldots = b_m$ and $r_1 = \ldots = r_m = 1$, then we recover completely the corresponding results from [**20**, **21**].

As usual, C denotes a positive constant that can change its value on each statement. If the constant depends on precise parameters, it will be pointed out.

2. *Preliminaries*

In this section we introduce the basic tools needed for the proof of the main results.

2.1. *Fefferman–Stein inequality for A_∞ weights*

A 'weight' will always mean a positive function that is locally integrable. We say that a weight w belongs to the class A_p, $1 < p < \infty$, if there is a constant C such that

$$\left(\frac{1}{|Q|}\int_Q w(y)\,dy\right)\left(\frac{1}{|Q|}\int_Q w(y)^{1-p'}dy\right)^{p-1} \leqslant C$$

for each cube Q, where, as usual, $1/p + 1/p' = 1$. A weight w belongs to the class A_1 if there is a constant C such that

$$\frac{1}{|Q|}\int_Q w(y)\,dy \leqslant C \inf_Q w.$$

We will denote the infimum of the constants C by $[w]_{A_p}$. Observe that $[w]_{A_p} \geqslant 1$ by Jensen's inequality.

Since the A_p classes are increasing with respect to p, the A_∞ class of weights is defined in a natural way by $A_\infty = \bigcup_{p>1} A_p$. However, the following characterization is more interesting. There are positive constants c and ρ such that, for any cube Q and any measurable set E contained in Q,

$$\frac{w(E)}{w(Q)} \leqslant c\left(\frac{|E|}{|Q|}\right)^\rho.$$

A simple fact that will be useful is that, for any $w \in A_p$ and any $m > 0$, $w_m = \min\{w, m\} \in A_p$ with $[w_m]_{A_p} \leqslant C_p[w]_{A_p}$. For more information on A_p weights we refer the reader to [**10**, **25**].

We recall now the definitions of classical maximal operators. If, as usual, M denotes the Hardy–Littlewood maximal operator, we consider, for $\delta > 0$,

$$M_\delta f(x) = M(|f|^\delta)^{1/\delta}(x) = \left(\sup_{Q\ni x}\frac{1}{|Q|}\int_Q |f(y)|^\delta\,dy\right)^{1/\delta},$$

$$M^\#(f)(x) = \sup_{Q\ni x}\inf_c \frac{1}{|Q|}\int_Q |f(y) - c|\,dy \approx \sup_{Q\ni x}\frac{1}{|Q|}\int_Q |f(y) - f_Q|\,dy,$$

and a variant of this sharp maximal operator that will become the main tool in our scheme:

$$M_\delta^\# f(x) = M^\#(|f|^\delta)(x)^{1/\delta}.$$

The main inequality between these operators to be used is a version of the classical one due to Fefferman and Stein (see [**16**, **24**]).

LEMMA 2.1. *Let w be an A_∞ weight. Then there exists a constant c depending upon the A_∞ condition of w such that, for all λ, $\epsilon > 0$,*

$$w(\{y \in \mathbb{R}^n : Mf(y) > \lambda,\ M^\# f(y) \leqslant \lambda\epsilon\}) \leqslant c\epsilon^\rho w\left(\left\{y \in \mathbb{R}^n : Mf(y) > \frac{\lambda}{2}\right\}\right).$$

As a consequence, we have the following estimates for $\delta > 0$.

(a) *Let $\varphi : (0,\infty) \longrightarrow (0,\infty)$ doubling. Then there exists a constant c depending upon the A_∞ condition of w and the doubling condition of φ such that*

$$\sup_{\lambda>0}\varphi(\lambda)w(\{y \in \mathbb{R}^n : M_\delta f(y) > \lambda\}) \leqslant c\sup_{\lambda>0}\varphi(\lambda)w(\{y \in \mathbb{R}^n : M_\delta^\# f(y) > \lambda\}) \quad (2.1)$$

for every function such that the left-hand side is finite.

(b) *Let $0 < p < \infty$. There exists a positive constant C depending upon the A_∞ condition of w and p such that*

$$\int_{\mathbb{R}^n} (M_\delta f(x))^p w(x)\, dx \leqslant C \int_{\mathbb{R}^n} (M_\delta^\# f(x))^p w(x)\, dx \tag{2.2}$$

for every function f such that the left-hand side is finite

2.2. *Orlicz maximal functions*

By a Young function Φ, we will mean a continuous, nonnegative, strictly increasing and convex function on $[0, \infty)$ with

$$\lim_{t\to 0^+} \frac{\Phi(t)}{t} = \lim_{t\to\infty} \frac{t}{\Phi(t)} = 0.$$

We define the Φ-averages of a function f over a cube Q by

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

For Orlicz norms, we are usually only concerned about the behavior of Young functions for large values of t. Given two functions B and C, we write $B(t) \approx C(t)$ if $B(t)/C(t)$ is bounded and bounded below for $t \geqslant c > 0$. We also recall that if $B(t) \leqslant C(t)$ for $t \geqslant c > 0$, then

$$\|f\|_{B,Q} \leqslant C\|f\|_{C,Q},$$

with C an absolute constant. For more information on the subject, see [**22**].

Associated to this average, we can define a maximal operator M_Φ given by

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

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

For example, if $\Phi(x) = e^{x^r} - 1$, then $\|\cdot\|_{\exp L^r,Q}$ and $M_{\exp L^r}$, denote the Φ-average and the maximal operator associated to Φ, respectively. Similarly, we have for $\Phi(x) = x\log^r(e + x)$, $\|\cdot\|_{L(\log L)^r,Q}$ and $M_{L(\log L)^r}$. Observe that by the above remarks, $Mf \leqslant CM_{L(\log L)^r} f$ for any $r > 0$. These examples will be relevant in our work.

Finally, we will be using the following known pointwise inequality. If $m \in \mathbb{N}$, then

$$M_{L(\log L)^m} \sim M^{m+1} = \overbrace{M \circ \ldots \circ M}^{(m+1 \text{ times})},$$

the $m+1$ iterations of the Hardy–Littlewood maximal operator. A generalization of this equivalence can be found in [**2**].

We begin with some technical lemmas on convex functions whose proofs are standard.

LEMMA 2.2 [**17**, Lemma 2.1]. *If $\Phi_0, \Phi_1, \ldots, \Phi_m$ are real-valued, non-negative, non-decreasing, left continuous functions defined on $[0, \infty)$ such that, with the definition*

$$\Phi_i^{-1}(x) = \inf\{y : \Phi_i(y) > x\},$$

it verifies

$$\Phi_1^{-1}(x)\Phi_2^{-1}(x)\ldots\Phi_m^{-1}(x) \leqslant \Phi_0^{-1}(x), \tag{2.3}$$

then, for all $0 \leqslant x_1, x_2, \dots, x_m < \infty$,

$$\Phi_0(x_1 x_2 \dots x_m) \leqslant \Phi_1(x_1) + \Phi_2(x_2) + \dots + \Phi_m(x_m).$$

LEMMA 2.3. *Let* Φ_0 *be a convex function such that* $\Phi_0(0) = 0$, *and let* $\Phi_1, \dots, \Phi_m$ *as in the previous lemma, all verifying* (2.3). *If* $f_1, f_2, \dots, f_m$ *are functions satisfying* $\|f_i\|_{\Phi_i,Q} < \infty$ *for all* $1 \leqslant i \leqslant m$ *and for a given cube* Q, *then*

$$\|f_1 \dots f_m\|_{\Phi_0,Q} \leqslant m \|f_1\|_{\Phi_1,Q} \dots \|f_m\|_{\Phi_m,Q}. \tag{2.4}$$

Proof. Taking $\varepsilon > 0$ small enough, by Lemma 2.2 and convexity,

$$\begin{aligned}
&\frac{1}{|Q|} \int_Q \Phi_0 \left(\frac{f_1(x) f_2(x) \dots f_m(x)}{m(\|f_1\|_{\Phi_1,Q} + \varepsilon)(\|f_2\|_{\Phi_2,Q} + \varepsilon) \dots (\|f_m\|_{\Phi_m,Q} + \varepsilon)} \right) dx \\
&\leqslant \frac{1}{m} \frac{1}{|Q|} \int_Q \Phi_0 \left(\frac{f_1(x) f_2(x) \dots f_m(x)}{(\|f_1\|_{\Phi_1,Q} + \varepsilon)(\|f_2\|_{\Phi_2,Q} + \varepsilon) \dots (\|f_m\|_{\Phi_m,Q} + \varepsilon)} \right) dx \\
&\leqslant \frac{1}{m} \frac{1}{|Q|} \int_Q \left[\Phi_1 \left(\frac{f_1(x)}{\|f_1\|_{\Phi_1,Q} + \varepsilon} \right) + \dots + \Phi_m \left(\frac{f_1(x)}{\|f_m\|_{\Phi_m,Q} + \varepsilon} \right) \right] dx \leqslant 1,
\end{aligned}$$

which implies that

$$\|f_1 f_2 \dots f_m\|_{\Phi_0,Q} \leqslant m(\|f_1\|_{\Phi_1,Q} + \varepsilon) \dots (\|f_m\|_{\Phi_m,Q} + \varepsilon).$$

Now, letting $\varepsilon \to 0$, we get (2.4). □

The main example that we will be using is the following:

$$\frac{1}{|Q|} \int_Q |f_1 \dots f_m g| \leqslant C \|f_1\|_{\exp L^{r_1},Q} \dots \|f_m\|_{\exp L^{r_m},Q} \|g\|_{L(\log L)^{1/r},Q}, \tag{2.5}$$

where $r_1, \dots, r_m \geqslant 1$ and

$$\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}.$$

Indeed, in this case, we can write for any $x \geqslant 0$ that

$$\log^{1/r_1}(1+x) \dots \log^{1/r_m}(1+x) \frac{x}{\log^{1/r_1 + \dots + 1/r_m}(e+x)} \leqslant x.$$

Then (2.3) holds for

$$\Phi_0(x) = x,$$
$$\Phi_i^{-1}(x) = \log^{1/r_i}(1+x),$$

$i = 1, \dots, m$, and

$$\Phi_{m+1}^{-1}(x) = \frac{x}{\log^{1/r_1 + \dots + 1/r_m}(e+x)}.$$

The inverse functions are given by $\Phi_i(x) = e^{x^{r_i}} - 1$, $i = 1, \dots, m$, and $\Phi_{m+1}(x) \approx x \log^{1/r_1 + \dots + 1/r_m}(e+x)$.

2.3. *The oscillation of a function*

We define the oscillation $Osc_\Phi(f, Q)$ of a function f with respect to any Young function Φ as

$$Osc_\Phi(f, Q) = \|f - f_Q\|_{\Phi,Q}.$$

We list the following properties that are very easy to check.

LEMMA 2.4. *For any cube Q and any locally integrable functions f and g, we have the following:*

(i) $Osc_\Phi(f \pm g, Q) \leqslant Osc_\Phi(f, Q) + Osc_\Phi(g, Q)$.

(ii) $Osc_\Phi(\lambda f, Q) = |\lambda| Osc_\Phi(f, Q)$ *for all* $\lambda \in \mathbb{R}$.

(iii) $Osc_\Phi(|f|, Q) \leqslant 2Osc_\Phi(f, Q)$.

(iv) $Osc_\Phi(\min\{f, g\}, Q) \leqslant \frac{3}{2}(Osc_\Phi(f, Q) + Osc_\Phi(g, Q))$.

(v) $Osc_\Phi(\max\{f, g\}, Q) \leqslant \frac{3}{2}(Osc_\Phi(f, Q) + Osc_\Phi(g, Q))$.

Given a function g, we define, for any $N \in \mathbb{N}$,

$$g^N(x) = \begin{cases} N & g(x) > N \\ g(x) & |g(x)| \leqslant N \\ -N & g(x) < -N. \end{cases} \tag{2.6}$$

Now, from Lemma 2.4, and taking into account that $g^N = \max\{\min\{g, N\}, -N\}$, we infer that, for any locally integrable function, $Osc_\Phi(g^N, Q) \leqslant C Osc_\Phi(g, Q)$, and consequently

$$\|g^N\|_{Osc_\Phi} \leqslant C\|g\|_{Osc_\Phi} \tag{2.7}$$

for all $N \in \mathbb{N}$ and where $C > 0$ is an absolute constant.

3. *Proofs of the main results*

3.1. *Proof of Theorem* 1.1

By a kernel K in $\mathbb{R}^n \times \mathbb{R}^n$, we mean a locally integrable function defined away from the diagonal. We say that K satisfies the standard estimates if there exist positive and finite constants γ and C such that, for all distinct $x, y \in \mathbb{R}^n$ and all z with $2|x - z| < |x - y|$, the following hold:

(i) $|K(x, y)| \leqslant C|x - y|^{-n}$.

(ii) $|K(x, y) - K(z, y)| \leqslant C|(x - z)/(x - y)|^\gamma |x - y|^{-n}$.

(iii) $|K(y, x) - K(y, z)| \leqslant C|(x - z)/(x - y)|^\gamma |x - y|^{-n}$.

We define a linear and continuous operator $T : C_0^\infty(\mathbb{R}^n) \longrightarrow \mathscr{D}'(\mathbb{R}^n)$ associated to the kernel K by

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

where $f \in C_0^\infty(\mathbb{R}^n)$ and x is not in the support of f. T is called a Calderón–Zygmund operator if K satisfies the standard estimates and if it extends to a bounded linear operator on $L^2(\mathbb{R}^n)$. These conditions imply that T is also bounded on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, and is of weak type $(1, 1)$. For more information on this subject, see [**5**, **7**, **16**].

Given any positive integer m, for all $1 \leqslant j \leqslant m$, we denote by C_j^m the family of all finite subsets $\sigma = \{\sigma(1), \ldots, \sigma(j)\}$ of $\{1, \ldots, m\}$ of j different elements. For any $\sigma \in C_j^m$, we associate the complementary sequence σ' given by $\sigma' = \{1, 2, \ldots, m\} \backslash \sigma$.

Let $\vec{b} = (b_1, b_2, \ldots, b_m)$ be a finite family of integrable functions. For all $1 \leqslant j \leqslant m$ and any $\sigma = \{\sigma(1), \ldots, \sigma(j)\} \in C_j^m$, we will denote $\vec{b_\sigma} = (b_{\sigma(1)}, \ldots, b_{\sigma(j)})$ and the product $b_\sigma = b_{\sigma(1)} \ldots b_{\sigma(j)}$. With this notation, we write, for any m-tuple $r = (r_1, \ldots, r_m)$ of positive numbers,

$$\|\vec{b_\sigma}\|_{Osc_{\exp L^{r_\sigma}}} = \|b_{\sigma(1)}\|_{Osc_{\exp L^{r_{\sigma(1)}}}} \ldots \|b_{\sigma(j)}\|_{Osc_{\exp L^{r_{\sigma(j)}}}}.$$

For the product of all the functions, we simply write

$$\|\vec{b}\|_{Osc_{\exp L^r}} = \|b_1\|_{Osc_{\exp L^{r_1}}} \dots \|b_m\|_{Osc_{\exp L^{r_m}}}.$$

For any $\sigma \in C_j^m$, we denote

$$T_{\vec{b_\sigma}} f(x) = \int_{\mathbb{R}^n} (b_{\sigma(1)}(x) - b_{\sigma(1)}(y)) \dots (b_{\sigma(j)}(x) - b_{\sigma(j)}(y)) K(x,y) f(y)\, dy.$$

In the particular case of $\sigma = \{1, \dots, m\}$, we understand $T_{\vec{b_\sigma}}$ as simply $T_{\vec{b}}$.

Finally, observe that $T_{\vec{b}}(f)$ satisfies the following homogeneity on the symbol

$$T_{\vec{b}}\left(\frac{f}{\|\vec{b}\|}\right) = T_{\tilde{\vec{b}}}(f)$$

where

$$\tilde{b} = (b_1/\|b_1\|_{Osc_{\exp L^{r_j}}}, \dots, b_m/\|b_m\|_{Osc_{\exp L^{r_m}}})$$

such that $\|\tilde{b}\| = 1$.

The following lemma gives a pointwise estimate of $M_\delta^\#(T_{\vec{b}} f)$ in terms of other maximal functions of operators of lower order. This result can be understood as the extension of [**20**, Lemma 7.1] to high-order commutators and for the classes of symbol considered in this paper.

LEMMA 3.1. *Let $T_{\vec{b}}$ be as in Theorem 1.1, and let $0 < \delta < \varepsilon < 1$. Then there exists a constant $C > 0$, depending only on δ and ε, such that*

$$M_\delta^\#(T_{\vec{b}} f)(x) \leqslant C \left[\|\vec{b}\| M_{L(\log L)^{1/r}} f(x) + \sum_{j=1}^{m} \sum_{\sigma \in C_j^m} \|b_\sigma\|_{Osc_{\exp L^{r_\sigma}}} M_\varepsilon(T_{\vec{b_{\sigma'}}} f)(x) \right] \quad (3.1)$$

for any bounded function f with compact support.

Recall that

$$\|\vec{b}\| = \prod_{j=1}^{m} \|b_j\|_{Osc_{\exp L^{r_j}}}.$$

Proof of Lemma 3.1. Recall that the operator $T_{\vec{b}} = T_{b_1,\dots,b_m}$ is defined by

$$T_{\vec{b}} f(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^{m} (b_j(x) - b_j(y)) \right] K(x,y) f(y)\, dy.$$

By homogeneity, we can assume that $\|\vec{b}\| = 1$.

We first consider the case $m = 1$. We only have one symbol $b \in Osc_{\exp L^r}$, $r \geqslant 1$, and (3.1) becomes

$$M_\delta^\#(T_b f)(x) \leqslant C\|b\|_{Osc_{\exp L^r}} [M_{L(\log L)^{1/r}} f(x) + M_\varepsilon(Tf)(x)]. \quad (3.2)$$

For any $\lambda \in \mathbb{R}$, we have

$$\begin{aligned} T_b f(x) &= \int_{\mathbb{R}^n} (b(x) - b(y)) K(x,y) f(y)\, dy \\ &= (b(x) - \lambda) Tf(x) - T((b - \lambda) f)(x). \end{aligned}$$

Now, for fixed $x \in \mathbb{R}^n$, for any number c and any ball B centered at x and radius

$R > 0$, since $0 < \delta < 1$ implies that $||\alpha|^\delta - |\beta|^\delta| \leqslant |\alpha - \beta|^\delta$ for any $\alpha, \beta \in \mathbb{R}$, we can estimate that

$$\begin{aligned}\left(\frac{1}{|B|}\int_B \big||T_b f(y)|^\delta - |c|^\delta\big|\,dy\right)^{1/\delta} &\leqslant \left(\frac{1}{|B|}\int_B |T_b f(y) - c|^\delta\,dy\right)^{1/\delta}\\ &\leqslant C\left[\left(\frac{1}{|B|}\int_B |(b(y)-\lambda)Tf(y)|^\delta\,dy\right)^{1/\delta}\right.\\ &\quad\left.+\left(\frac{1}{|B|}\int_B |T((b-\lambda)f)(y) - c|^\delta\,dy\right)^{1/\delta}\right]\\ &= \mathrm{I} + \mathrm{II}.\end{aligned}$$

We analyze each term separately, and we let $\lambda = (b)_{2B}$, the average of b over the ball $2B$. For any $1 < q < \varepsilon/\delta$, we have, by Hölder and Jensen's inequality,

$$\begin{aligned}\mathrm{I} &\leqslant C\left(\frac{1}{|2B|}\int_{2B} |b(y)-\lambda|^{\delta q'}\,dy\right)^{1/\delta q'}\left(\frac{1}{|B|}\int_B |Tf(y)|^{\delta q}\,dy\right)^{1/\delta q}\\ &\leqslant C\|b\|_{Osc_{\exp L^r}} M_{\delta q}(Tf)(x)\\ &\leqslant C M_\varepsilon(Tf)(x).\end{aligned} \tag{3.3}$$

To deal with II, we split f as usual by $f = f_1 + f_2$, where $f_1 = f\chi_{2B}$ and $f_2 = f - f_1$. This yields

$$\begin{aligned}\mathrm{II} &\leqslant C\left[\left(\frac{1}{|B|}\int_B |T((b-\lambda)f_1)(y)|^\delta\,dy\right)^{1/\delta}\right.\\ &\quad\left.+\left(\frac{1}{|B|}\int_B |T((b-\lambda)f_2)(y) - c|^\delta\,dy\right)^{1/\delta}\right]\\ &= \mathrm{III} + \mathrm{IV}.\end{aligned}$$

For III, since $(b-\lambda)f_1$ is integrable and T is of weak type $(1,1)$, by Kolmogorov's inequality [**25**, p. 104] and (2.5), we get

$$\begin{aligned}\mathrm{III} &\leqslant \frac{C}{|2B|}\int_{2B} |b(y)-\lambda||f(y)|\,dy\\ &\leqslant C\|b-\lambda\|_{\exp L^r, 2B}\|f\|_{L(\log L)^{1/r}, 2B}\\ &\leqslant C\|b\|_{Osc_{\exp L^r}} M_{L(\log L)^{1/r}} f(x)\\ &= C M_{L(\log L)^{1/r}} f(x).\end{aligned} \tag{3.4}$$

For the last term IV, we make the election $c = (T((b-\lambda)f_2))_B$. Hence, by Jensen's inequality, it follows that

$$\begin{aligned}\mathrm{IV} &\leqslant \frac{C}{|B|}\int_B |T((b-\lambda)f_2)(y) - T((b-\lambda)f_2))_B|\,dy\\ &= \frac{C}{|B|}\int_B \left|\left(\int_{\mathbb{R}^n} K(y,w)(b(w)-\lambda)f_2(w)\,dw\right)\right.\\ &\quad\left.-\frac{1}{|B|}\int_B\left(\int_{\mathbb{R}^n} K(z,w)(b(w)-\lambda)f_2(w)\,dw\right)dz\right|\,dy\end{aligned}$$

$$
\begin{aligned}
&= \frac{C}{|B|}\int_B \left| \frac{1}{|B|}\int_B \left[\int_{\mathbb{R}^n\setminus 2B} (K(y,w)-K(z,w))(b(w)-\lambda)f(w)\,dw \right] dz \right| dy \\
&\leqslant \frac{C}{|B|^2}\int_B\int_B \left[\int_{\mathbb{R}^n\setminus 2B} |K(y,w)-K(z,w)||b(w)-\lambda||f(w)|\,dw \right] dzdy \\
&\leqslant \frac{C}{|B|^2}\int_B\int_B \left[\sum_{k=1}^{\infty} \int_{2^kR\leqslant|w-x|<2^{k+1}R} \left| \frac{y-z}{y-w} \right|^{\gamma} \frac{1}{|y-w|^n} |b(w)-\lambda||f(w)|\,dw \right] dzdy \\
&\leqslant C\sum_{k=1}^{\infty} \left(\frac{2R}{2^kR} \right)^{\gamma} \frac{1}{(2^kR)^n} \int_{2^{k+1}B} |b(w)-\lambda||f(w)|\,dw \\
&\leqslant C\sum_{k=1}^{\infty} 2^{-k\gamma} \|b-\lambda\|_{\exp L^r, 2^{k+1}B} \|f\|_{L(\log L)^{1/r}, 2^{k+1}B}. \qquad (3.5)
\end{aligned}
$$

The fifth inequality follows because $y, z \in B$ and $w \in \mathbb{R}^n\setminus B$, and therefore $2|y-w| < |z-w|$. The sixth does since $|y-z| \leqslant 2R$ and $|y-w|^{-1} \leqslant ((2^k-1)R)^{-1} \leqslant C(2^kR)^{-1}$, and the last one is an application of the generalized Hölder inequality (2.5).

Now, we claim that

$$
\|b-\lambda\|_{\exp L^r, 2^{k+1}B} \leqslant Ck\|b\|_{Osc_{\exp L^r}}. \qquad (3.6)
$$

Indeed

$$
\begin{aligned}
\|b-\lambda\|_{\exp L^r, 2^{k+1}B} &\leqslant \|b-(b)_{2^{k+1}B}\|_{\exp L^r, 2^{k+1}B} + \|(b)_{2^{k+1}B}-\lambda\|_{\exp L^r, 2^{k+1}B} \\
&\leqslant \|b-(b)_{2^{k+1}B}\|_{\exp L^r, 2^{k+1}B} + |(b)_{2^{k+1}B}-\lambda| \\
&\leqslant Ck\|b\|_{Osc_{\exp L^r}},
\end{aligned}
$$

where the last estimate follows by the standard inequality $|(b)_{2B} - b_B| \leqslant 2\|b\|_{Osc_{\exp L^r}}$ (cf. [**16**, p. 31]), and then we take the supremum over the balls.

Thus IV is estimated by

$$
\begin{aligned}
\mathrm{IV} &\leqslant C\sum_{k=1}^{\infty} (2^{-k})^{\gamma} k \|b\|_{Osc_{\exp L^r}} M_{L(\log L)^{1/r}} f(x) \\
&\leqslant CM_{L(\log L)^{1/r}} f(x). \qquad (3.7)
\end{aligned}
$$

From (3.4) and (3.7), we conclude that

$$
\mathrm{II} \leqslant CM_{L(\log L)^{1/r}} f(x),
$$

which, together with (3.3), gives (3.2) and proves the lemma for $m = 1$.

Consider now the case $m \geqslant 2$. For any $\vec{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^n$, we have

$$
\begin{aligned}
T_{\vec{b}} f(x) &= \int_{\mathbb{R}^n} (b_1(x)-b_1(y)) \ldots (b_m(x)-b_m(y)) K(x,y) f(y)\,dy \\
&= \int_{\mathbb{R}^n} ((b_1(x)-\lambda_1)-(b_1(y)-\lambda_1)) \ldots ((b_m(x)-\lambda_m) \\
&\quad -(b_m(y)-\lambda_m)) K(x,y) f(y)\,dy \\
&= \sum_{j=0}^{m} \sum_{\sigma\in C_j^m} (-1)^{m-j} (b(x)-\vec{\lambda})_{\sigma} \int_{\mathbb{R}^n} (b(y)-\vec{\lambda})_{\sigma'} K(x,y) f(y)\,dy
\end{aligned}
$$

$$= (b_1(x) - \lambda_1) \dots (b_m(x) - \lambda_m) Tf(x) + (-1)^m T((b_1 - \lambda_1) \dots (b_m - \lambda_m) f)(x) + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} (-1)^{m-j} (b(x) - \vec{\lambda})_\sigma \int_{\mathbb{R}^n} (b(y) - b(x))_{\sigma'} K(x,y) f(y)\, dy$$

Now, expanding $(b(y) - \vec{\lambda})_{\sigma'} = [(b(y) - b(x)) + (b(x) - \vec{\lambda})]_{\sigma'}$ as above, it is easy to see that

$$T_{\vec{b}} f(x) = (b_1(x) - \lambda_1) \dots (b_m(x) - \lambda_m) Tf(x) + (-1)^m T((b_1 - \lambda_1) \dots (b_m - \lambda_m) f)(x) + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} c_{m,j} (b(x) - \vec{\lambda})_\sigma T_{\vec{b_{\sigma'}}} f(x),$$

where $c_{m,j}$ are absolute constants depending only on m and j.

Now, for fixed $x \in \mathbb{R}^n$, for any number c and any ball B centered at x and radius $R > 0$, it follows, since $0 < \delta < 1$,

$$\left(\frac{1}{|B|} \int_B \big| |T_{\vec{b}} f(y)|^\delta - |c|^\delta \big|\, dy \right)^{1/\delta} \leqslant \left(\frac{1}{|B|} \int_B |T_{\vec{b}} f(y) - c|^\delta dy \right)^{1/\delta}$$

$$\leqslant C \left[\left(\frac{1}{|B|} \int_B |(b_1(y) - \lambda_1) \dots (b_m(y) - \lambda_m) Tf(y)|^\delta dy \right)^{1/\delta} + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} \left(\frac{1}{|B|} \int_B \left| (b - \vec{\lambda})_\sigma T_{\vec{b_{\sigma'}}} f(y) \right|^\delta dy \right)^{1/\delta} + \left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \dots (b_m - \lambda_m) f)(y) - c|^\delta dy \right)^{1/\delta} \right]$$

$$= \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

Reasoning as in (3.3) with $\lambda_i = (b_i)_{2B}$, $i = 1, \dots, m$, using this time the standard Hölder inequality for finitely many functions with $1 < q < \varepsilon/\delta$, we have

$$\mathrm{I} \leqslant C M_\varepsilon (Tf)(x) \tag{3.8}$$

and

$$\mathrm{II} \leqslant C \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} M_\varepsilon (T_{\vec{b_{\sigma'}}} f)(x), \tag{3.9}$$

with $C > 0$ depending only on m.

For III, we split $f = f_1 + f_2$ with $f_1 = f\chi_{2B}$ and $f_2 = f - f_1$. Thus

$$\mathrm{III} \leqslant C \left[\left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \dots (b_m - \lambda_m) f_1)(y)|^\delta dy \right)^{1/\delta} + \left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \dots (b_m - \lambda_m) f_2)(y) - c|^\delta dy \right)^{1/\delta} \right]$$

$$= \mathrm{IV} + \mathrm{V}.$$

Now, as in (3.4) and again making use of Kolmogorov's inequality and (2.5), we can estimate IV by

$$\begin{aligned} \mathrm{IV} &\leqslant \frac{C}{|2B|}\int_{2B} |b_1(y)-\lambda_1|\ldots|b_m(y)-\lambda_m||f(y)|\,dy \\ &\leqslant C\|b_1-\lambda_1\|_{\exp L^{r_1},2B}\ldots\|b_m-\lambda_m\|_{\exp L^{r_m},2B}\|f\|_{L(\log L)^{1/r},2B} \\ &\leqslant CM_{L(\log L)^{1/r}}f(x), \end{aligned} \tag{3.10}$$

recalling that

$$\frac{1}{r}=\frac{1}{r_1}+\ldots+\frac{1}{r_m}.$$

Finally, for V, choosing $c=(T((b_1-\lambda_1)\ldots(b_m-\lambda_m)f_2))_B$, and repeating the argument used to get (3.5), it follows from (2.5) and (3.6) that

$$\begin{aligned} \mathrm{V} &\leqslant C\sum_{k=1}^{\infty}2^{-k\gamma}\frac{1}{(2^{k+1}R)^n}\int_{2^{k+1}B}\left(\prod_{j=1}^{m}|b_j(w)-\lambda_j|\right)|f(w)|\,dw \\ &\leqslant C\sum_{k=1}^{\infty}2^{-k\gamma}\left(\prod_{j=1}^{m}\|b_j-\lambda_j\|_{\exp L^{r_j},2^{k+1}B}\right)\|f\|_{L(\log L)^{1/r},2^{k+1}B} \\ &\leqslant C\left[\sum_{k=1}^{\infty}2^{-k\gamma}k^m\right]\left(\prod_{j=1}^{m}\|b_j\|_{Osc_{\exp L^{r_j}}}\right)M_{L(\log L)^{1/r}}f(x). \end{aligned} \tag{3.11}$$

Finally, from (3.10) and (3.11), we conclude that

$$\mathrm{III}\leqslant CM_{L(\log L)^{1/r}}f(x),$$

which, together with (3.8) and (3.9), gives (3.1), and the proof of the lemma is finished. □

We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1. We can assume that

$$\int_{\mathbb{R}^n}(M_{L(\log L)^{1/r}}f(x))^p w(x)\,dx<\infty, \tag{3.12}$$

since otherwise there is nothing to be proved.

To apply the Fefferman–Stein inequality (2.2), we first take it for granted that $\|M_\delta(T_{\vec{b}}f)\|_{L^p(w)}$ is finite. We will check this to the end of the proof.

We proceed by induction on m. For $m=1$, by (2.2) and Lemma 3.1, we can estimate that

$$\begin{aligned} \|T_{b_1}f\|_{L^p(w)} &\leqslant \|M_\delta(T_{b_1}f)\|_{L^p(w)} \\ &\leqslant C\|M_\delta^{\#}(T_{b_1}f)\|_{L^p(w)} \\ &\leqslant C\|b_1\|_{Osc_{\exp L^{r_1}}}[\|M_\varepsilon(Tf)\|_{L^p(w)}+\|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)}] \\ &\leqslant C\|b_1\|_{Osc_{\exp L^{r_1}}}[\|Tf\|_{L^p(w)}+\|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)}] \\ &\leqslant C\|b_1\|_{Osc_{\exp L^{r_1}}}[\|Mf\|_{L^p(w)}+\|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)}] \\ &\leqslant C\|b_1\|_{Osc_{\exp L^{r_1}}}\|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)}, \end{aligned}$$

where the fourth inequality follows since $w \in A_\infty$, and therefore there exists $q > 1$ such that $w \in A_q$. Then we can choose ε such that $0 < \varepsilon < p/q$ (we can take $q > p$ if necessary). The fifth holds by the classical estimate (1.3) for $m = 0$ (see [7, Chapter 1]).

Suppose now that for $m - 1$ the theorem is true, and let us prove it for m. The same argument as used above and the induction hypothesis give

$$\begin{aligned}
\|T_{\vec{b}}f\|_{L^p(w)} &\leqslant \|M_\delta(T_{\vec{b}}f)\|_{L^p(w)} \\
&\leqslant C\|M_\delta^{\#}(T_{\vec{b}}f)\|_{L^p(w)} \\
&\leqslant C\Bigg[\|b_1\|_{Osc_{\exp L^{r_1}}}\dots\|b_m\|_{Osc_{\exp L^{r_m}}}\|M_{L(\log L)^{1/r}}f\|_{L^p(w)} \\
&\qquad + \sum_{j=1}^{m-1}\sum_{\sigma\in C_j^m}\|b_\sigma\|_{Osc_{\exp L^{\sigma}}}\|M_\varepsilon(T_{\vec{b}_{\sigma'}})\|_{L^p(w)}\Bigg] \\
&\leqslant C\Bigg[\|b_1\|_{Osc_{\exp L^{r_1}}}\dots\|b_m\|_{Osc_{\exp L^{r_m}}}\|M_{L(\log L)^{1/r}}f\|_{L^p(w)} \\
&\qquad + \sum_{j=1}^{m-1}\sum_{\sigma\in C_j^m}\|b_\sigma\|_{Osc_{\exp L^{\sigma}}}\|b_{\sigma'}\|_{Osc_{\exp L^{\sigma'}}}\|M_{L(\log L)^{1/r_{\sigma'}}}f\|_{L^p(w)}\Bigg] \\
&\leqslant C\|b_1\|_{Osc_{\exp L^{r_1}}}\dots\|b_m\|_{Osc_{\exp L^{r_m}}}\|M_{L(\log L)^{1/r}}f\|_{L^p(w)},
\end{aligned}$$

since $M_{L(\log L)^{1/r_{\sigma'}}} \leqslant CM_{L(\log L)^{1/r}}$.

Let us check now that for appropriate δ we have $\|M_\delta(T_{\vec{b}}f)\|_{L^p(w)} < \infty$. Indeed, as above, since $w \in A_\infty$, there exists $q > 1$ such that $w \in A_q$, and we can choose δ small enough so that $p/\delta > q$. Then, by Muckehoupt's theorem, the proof is reduced to checking that $\|T_b f\|_{L^p(w)} < \infty$.

Suppose that the symbols b_k and the weight w are all bounded functions. Since f has compact support, we can assume that the support of f is contained in the ball $B_R = B(0, R)$. Then we can split the integral as

$$\int_{\mathbb{R}^n}|T_{\vec{b}}f(x)|^p w(x)\,dx = \int_{|x|\leqslant 2R}|T_{\vec{b}}f(x)|^p w(x)\,dx + \int_{|x|>2R}|T_{\vec{b}}f(x)|^p w(x)\,dx.$$

The first integral can easily be estimated by making use of the L^∞-boundedness of the b_k and w and the L^q-boundedness for $q > 1$ of the Calderón–Zygmund operator T.

For the second term, by the properties of the kernel K and the boundedness of the symbols b_k, since $|x| > 2R$, we have the following pointwise estimate:

$$\begin{aligned}
|T_{\vec{b}}f(x)| &\leqslant C\int_{B_R}\frac{|b_1(x)-b_1(y)|\dots|b_m(x)-b_m(y)||f(y)|}{|x-y|^n}dy \\
&\leqslant \frac{C}{|x|^n}\int_{B(0,|x|)}|f(y)|dy \\
&\leqslant CMf(x) \\
&\leqslant CM_{L(\log L)^{1/r}}f(x). \qquad (3.13)
\end{aligned}$$

Thus

$$\int_{|x|>2R} |T_{\vec{b}} f(x)|^p w(x)\,dx \leqslant C \int_{|x|>2R} (M_{L(\log L)^{1/r}} f(x))^p w(x)\,dx,$$

which is finite by assumption (3.12).

For the general case, we will truncate the symbols b_k and the weight w as follows (cf. [**7**, p. 40]). We denote by $\vec{b}^N$ the vector of truncated elements by N, that is, $\vec{b}^N = (b_1^N, \dots, b_m^N)$, where each b_k^N is the truncation of b_k as defined in (2.6). Observe that in our case (2.7) becomes

$$\|b_k^N\|_{Osc_{\exp L^{r_k}}} \leqslant C\|b_k\|_{Osc_{\exp L^{r_k}}}, \tag{3.14}$$

where $C > 0$ is a constant independent of N. Analogously, we consider the truncations of the weight w by $w_N = \inf\{w, N\}$ that satisfy

$$[w_N]_{A_\infty} \leqslant C[w]_{A_\infty}. \tag{3.15}$$

Then (1.7) holds for the operator $T_{\vec{b}^N}$ and the weight w_N. Combining (3.14) and (3.15), this estimate gives

$$\int_{\mathbb{R}^n} |T_{\vec{b}^N} f(x)|^p w_N(x)\,dx \leqslant C \left(\prod_{j=1}^m \|b_j\|^p_{Osc_{\exp L^{r_j}}} \right) \int_{\mathbb{R}^n} (M_{L(\log L)^{1/r}} f(x))^p w(x)\,dx.$$

Next, taking into account the fact that f has compact support, we deduce that any product $b_{i_1}^N \dots b_{i_k}^N f$ converges in any L^q for $q > 1$ to $b_{i_1} \dots b_{i_k} f$ as $N \to \infty$. Hence the classical L^q-boundedness of the operator T gives, at least for a subsequence, that $|T_{\vec{b}^N} f(x)|^p w_N(x)$ converges pointwise almost everywhere to $|T_{\vec{b}} f(x)|^p w(x)$, and by Fatou's lemma we conclude that the theorem holds for this general case. The theorem is proved. □

3.2. *Proof of Theorem* 1.5

We adapt here some of the arguments from [**20**]. Since the proof of Theorem 1.5 is based on Theorem 1.6, we prove this first. Namely we must show that

$$\sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} w(\{y \in \mathbb{R}^n \colon |T_{\vec{b}} f(y)| > t\}) \leqslant C \sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} w(\{y \in \mathbb{R}^n \colon M_\Phi(\|\vec{b}\| f)(y) > t\}) \tag{3.16}$$

for all bounded functions f with compact support. Recall that $\Phi(t) = \Phi_{\vec{b}}(t) = t \log^{1/r}(e + t)$.

In fact, we are going to prove something stronger than (3.16), namely the following: For every $\vec{b}$, $\varphi : (0, \infty) \longrightarrow (0, \infty)$ doubling with $\varphi(t) \leqslant Ct$, $t > 0$, and for every $0 < \delta < 1$ there exists a constant C such that

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}} f)(y) > t\}) \leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(\|\vec{b}\| f)(y) > t\}) \tag{3.17}$$

for all bounded functions f with compact support.

By the Lebesgue differentiation theorem, and taking

$$\varphi(t) = \Phi\left(\frac{1}{t}\right)^{-1} = \frac{t}{\log^{1/r}\left(e + \frac{1}{t}\right)}$$

it is clear that (3.17) implies (3.16).

By making use of the weighted version of the Fefferman–Stein Lemma 2.1, and more precisely estimate (2.1), we have

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}} f)(y) > t\}) \leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_{\vec{b}} f)(y) > t\}) \tag{3.18}$$

whenever the left-hand side is finite. Therefore (3.17) will follow from

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_{\vec{b}} f)(y) > t\}) \leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(\|\vec{b}\| f)(y) > t\}). \tag{3.19}$$

We first check that the left-hand side of (3.18) is finite for all bounded functions f with compact support. By proceeding as in the proof of Theorem 1.1, we can assume that b and w are bounded. For the general case of unbounded symbols and unbounded weight, we reproduce the argument used in the proof of Theorem 1.1, this time taking into account the weak (1,1) boundedness of the operator T that gives the convergence in measure.

Suppose that $\operatorname{supp} f \subset B_R = B(0, R)$. Since $0 < \delta < 1$, it follows that

$$\begin{aligned} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}} f)(y) > t\}) &\leqslant C\varphi(t) |\{y \in \mathbb{R}^n : M_\delta(\chi_{B_{2R}} T_{\vec{b}} f)(y) > t/2\}| \\ &\quad + C\varphi(t) |\{y \in \mathbb{R}^n : M_\delta(\chi_{\mathbb{R}^n \setminus B_{2R}} T_{\vec{b}} f)(y) > t/2\}| \\ &= \mathrm{I} + \mathrm{II}. \end{aligned}$$

For I, we use the fact that M is of weak type (1,1) and that $\varphi(t) \leqslant Ct$. Then

$$\begin{aligned} \mathrm{I} &\leqslant Ct |\{y \in \mathbb{R}^n : M(\chi_{B_{2R}} T_{\vec{b}} f)(y) > t/2\}| \\ &\leqslant C \int_{B_{2R}} |T_{\vec{b}} f(y)| \, dy \\ &\leqslant C R^{n/2} \left(\int_{\mathbb{R}^n} |Tf(y)|^2 dy \right)^{1/2}, \end{aligned}$$

which is finite since T is a Calderón–Zygmund operator and the fact is used that the symbols b_k are bounded.

For II, we take into account the pointwise estimate (3.13) and the well known fact that $(Mf)^\delta \in A_1$. Then we have

$$\begin{aligned} \mathrm{II} &\leqslant Ct |\{y \in \mathbb{R}^n : M_\delta(Mf)(y) > Ct\}| \\ &\leqslant Ct |\{y \in \mathbb{R}^n : Mf(y) > Ct\}| \\ &\leqslant C \int_{\mathbb{R}^n} |f(y)| \, dy < \infty. \end{aligned}$$

Combining the homogeneity and the linearity of $T_{\vec{b}}$, it is easy to see that we can assume that $\|\vec{b}\| = 1$ in both (1.10) and (3.16).

To prove (3.19), we proceed by induction on m.

3.3. *The case* $m = 1$

This case is essentially taken from [**20**], and we repeat it, with minor modifications, for the sake of completeness. In this case, the operator T_b is simply defined by one single function b:

$$T_b f = [b, T]f = bT(f) - T(bf),$$

where T is any Calderón–Zygmund operator. Recall that, by homogeneity, we can assume that $\|b\| = \|b\|_{Osc_{\exp L^r}} = 1$, and therefore what we must prove is

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#([b, T]f)(y) > t\}) \leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) \quad (3.20)$$

for all bounded functions f with compact support. Now, applying Lemma 3.1 with any α such that $\delta < \alpha < 1$, we find that the left-hand side of (3.20) is estimated by

$$\begin{aligned} C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : C[M_{L(\log L)^{1/r}}(f)(y) + M_\alpha(Tf)(y)] > t\}) &\leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) \\ &\quad + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\alpha(Tf)(y) > t\}), \end{aligned}$$

where we have also used the doubling condition of φ.

Next, considering the estimate

$$M_\alpha^\#(Tf)(y) \leqslant C_\alpha Mf(y), \quad (3.21)$$

which holds for all $0 < \alpha < 1$ (see [**1**, Theorem 2.1]), if we further select α such that $0 < \delta < \alpha < 1$, then the Fefferman–Stein lemma yields

$$\begin{aligned} \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#([b, T]f)(y) > t\}) &\leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) \\ &\quad + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\alpha^\#(Tf)(y) > t\}) \\ &\leqslant C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) \\ &\quad + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M(f)(y) > t\}) \\ &\leqslant C \sup_{t>0} C\varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}), \end{aligned}$$

since trivially $M(f) = M_L(f) \leqslant M_{L(\log L)^{1/r}}(f)$. This finishes the proof of (3.20).

3.4. *The general case*

Suppose now that (3.19) holds for $m-1$, and let us prove it for m. Recall that $\Phi(t) = \Phi_b(t) = t \log^{1/r}(e + t)$ with

$$\frac{1}{r} = \frac{1}{r_1} + \ldots + \frac{1}{r_m}.$$

Then, by Lemma 3.1,

$$\begin{aligned}
&\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_\delta^\#(T_{\vec b}f)(y)>t\})\\
&\quad\leqslant \sup_{t>0}\varphi(t)w\left(\left\{y\in\mathbb{R}^n : C\left[M_\Phi(f)(y)+\sum_{j=1}^m\sum_{\sigma\in C_j^m}\|b_\sigma\|_{Osc_{\exp L^{r_\sigma}}}M_{\varepsilon'}(T_{\vec b_{\sigma'}}f)(y)\right]>t\right\}\right)\\
&\quad\leqslant C_m\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_\Phi(f)(y)>t\})\\
&\qquad+C_m\sum_{j=1}^m\sum_{\sigma\in C_j^m}\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_{\varepsilon'}(T_{\vec b_{\sigma'}}(\|b_\sigma\|_{Osc_{\exp L^{r_\sigma}}}f))(y)>t\}).
\end{aligned}$$

$\varepsilon'<1$, and we have already checked that the distribution set on the left-hand side is finite, so combining the Fefferman–Stein lemma with the induction hypothesis on (3.19) with $T_{\vec b_{\sigma'}}$, we can estimate the last expression by

$$\begin{aligned}
&C_m\sum_{j=1}^m\sum_{\sigma\in C_j^m}\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_{\varepsilon'}^\#(T_{\vec b_{\sigma'}}(\|b_\sigma\|_{Osc_{\exp L^{r_\sigma}}}f))(y)>t\})\\
&\quad\leqslant C_m\sum_{j=1}^m\sum_{\sigma\in C_j^m}\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_{\Phi_{\vec b_{\sigma'}}}(\|b_{\sigma'}\|_{Osc_{\exp L^{r_{\sigma'}}}}\|b_\sigma\|_{Osc_{\exp L^{r_\sigma}}}f)(y)>t\})\\
&\quad\leqslant C_m\sum_{j=1}^m\sum_{\sigma\in C_j^m}\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_{\Phi_{\vec b_{\sigma'}}}(f)(y)>t\}),
\end{aligned}$$

since

$$\|\vec b_{\sigma'}\|_{Osc_{\exp L^{r_{\sigma'}}}}\|\vec b_\sigma\|_{Osc_{\exp L^{r_\sigma}}}=\|\vec b\|=1.$$

Finally, using the trivial observation that

$$M_{\Phi_{\vec b_{\sigma'}}}(f)\leqslant M_{\Phi_{\vec b_\sigma}}(f)=M_\Phi(f),$$

we have

$$\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_\delta^\#(T_{\vec b}f)(y)>t\})\leqslant C_m\sup_{t>0}\varphi(t)w(\{y\in\mathbb{R}^n : M_\Phi(f)(y)>t\}),$$

and claim (3.19) is proved.

We need the following lemma concerning estimates of the maximal operator M_Φ, which is a more general version than the one given in [**20**, Lemma 8.3]. The proof is standard, and we shall omit it.

LEMMA 3.2. *Let $w\in A_1$. Then there exists a positive constant C such that, for any $t>0$ and any locally integrable function f,*

$$w(\{y\in\mathbb{R}^n : M_\Phi f(y)>t\})\leqslant C\int_{\mathbb{R}^n}\Phi\left(\frac{|f(y)|}{t}\right)w(y)\,dy.$$

We are now in position to prove Theorem 1.5.

Proof of Theorem 1.5. By homogeneity, it is enough to assume that $t=\|b\|=1$, and hence we must prove that

$$w(\{y\in\mathbb{R}^n : |T_bf(y)|>1\})\leqslant C\int_{\mathbb{R}^n}\Phi(|f(y)|)w(y)\,dy.$$

Now, since Φ is submultiplicative, namely $\Phi(ab) \leqslant 2\Phi(a)\Phi(b)$, $a, b \geqslant 0$, we have by Theorem 1.6 and Lemma 3.2,

$$\begin{aligned} w(\{y \in \mathbb{R}^n : |T_b f(y)| > 1\}) &\leqslant C \sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} w(\{y \in \mathbb{R}^n : |T_b f(y)| > t\}) \\ &\leqslant C \sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} w(\{y \in \mathbb{R}^n : M_\Phi f(y) > t\}) \\ &\leqslant C \sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} \int_{\mathbb{R}^n} \Phi\left(\frac{|f(y)|}{t}\right) w(y)\, dy \\ &\leqslant C \sup_{t>0} \frac{1}{\Phi\left(\frac{1}{t}\right)} \int_{\mathbb{R}^n} \Phi(|f(y)|)\Phi\left(\frac{1}{t}\right) w(y)\, dy \\ &\leqslant C \int_{\mathbb{R}^n} \Phi(|f(y)|) w(y)\, dy, \end{aligned}$$

and the proof is concluded. □

References

1. J. Alvarez and C. Pérez, 'Estimates with A_∞ weights for various singular integral operators', *Boll. Un. Mat. Ital.* A 8 (1994) 123–133.
2. M. Carozza and A. Passarelli, 'Composition of maximal operators', *Publ. Mat.* 40 (1996) 397–409.
3. F. Chiarenza, M. Frasca and P. Longo, 'Interior $W^{2,p}$ estimates for non divergence elliptic equations with discontinuous coefficients', *Ricerche Mat.* 40 (1991) 149–168.
4. F. Chiarenza, M. Frasca and 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.
5. M. Christ, *Lectures on singular integral operators*, Regional Conferences Series in Mathematics 77 (American Mathematical Society, Providence, RI, 1990).
6. R. Coifman, 'Distribution function inequalities for singular integrals', *Proc. Nat. Acad. Sci. U.S.A.* 69 2838–2839.
7. R. Coifman and Y. Meyer, *Wavelets. Calderón–Zygmund and multilinear operators* Cambridge Studies in Advanced Mathematics 48 (Cambridge University Press, Cambridge, 1997).
8. R. Coifman, R. Rochberg and G. Weiss, 'Factorization theorems for Hardy spaces in several variables', *Ann. of Math.* 103 (1976) 611–635.
9. G. Di Fazio and M. A. Ragusa, 'Interior estimates in Morrey spaces for strong solutions to nondivergence form equations with discontinuous coefficients', *J. Funct. Anal.* 112 (1993) 241–256.
10. J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North Holland Mathematical Studies 116 (North-Holland, Amsterdam, 1985).
11. J. García-Cuerva, E. Harboure, C. Segovia and J. L. Torrea, 'Weighted norm inequalities for commutators of strongly singular integrals', *Indiana Univ. Math. J.* 40 (1991) 1397–1420.
12. A. E. Gatto and S. Vági, 'On the exponential integrability of fractional integrals on spaces of homogeneous type', *Colloq. Math.* 64 (1993) 121–127.
13. L. Greco and T. Iwaniec, 'New inequalities for the Jacobian', *Ann. Inst. H. Poincaré* 11 (1994) 17–35.
14. T. Iwaniec and C. Sbordone, 'Weak minima of variational integrals', *J. Reine Angew. Math.* 454 143–161.
15. S. Janson, 'Mean oscillation and commutators of singular integral operators', *Ark. Mat.* 16 (1978) 263–270.
16. J. L. Journé, *Calderón–Zygmund operators, pseudo-differential operator, and the Cauchy integral of Calderón*, Lecture Notes in Mathematics 994 (Springer, New York, 1983).
17. R. O'neil, 'Fractional integration in Orlicz spaces', *Trans. Amer. Math. Soc.* 115 (1963) 300–328.
18. C. Pérez, 'Weighted norm inequalities for singular integral operators', *J. London Math. Soc.* 49 (1994) 296–308.

19. C. PÉREZ, '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.

20. C. PÉREZ, 'Endpoint estimates for commutators of singular integral operators', *J. Funct. Anal.* 128 (1995) 163–185.

21. C. PÉREZ, 'Sharp estimates for commutators of singular integrals via iterations of the Hardy–Littlewood maximal function', *J. Fourier Anal. Appl.* 3 (1997) 743–756.

22. M. M. RAO and Z. D. REN, *Theory of Orlicz spaces* (Marcel Dekker, New York, 1991).

23. R. ROCHBERG and G. WEISS, 'Derivatives of analytic families of Banach spaces', *Ann. of Math.* 118 (1983) 315–347.

24. E. M. STEIN, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series 43 (Princeton University Press, Princeton, NJ, 1993).

25. A. TORCHINSKY, *Real variable methods in harmonic analysis* (Academic Press, New York, 1988).

26. J. M. WILSON, 'Weighted norm inequalities for the continuous square functions', *Trans. Amer. Math. Soc.* 314 (1989) 661–692.

27. W. P. ZIEMER, *Weakly differentiable functions*, Graduate Texts in Mathematics 120 (Springer, 1989).

Departmento de Análisis Matematico
Facultad de Matemáticas
Universidad de Sevilla
41080 Sevilla
Spain

carlos.perez@uam.es

Departmento de Análisis Matemático
Universidad de La Laguna
38271 La Laguna
S/C de Tenerife
Spain

rotrujil@ull.es