

# END-POINT ESTIMATES FOR ITERATED COMMUTATORS OF MULTILINEAR SINGULAR INTEGRALS

CARLOS PÉREZ, GLADIS PRADOLINI, RODOLFO H. TORRES,  
AND RODRIGO TRUJILLO-GONZÁLEZ

**ABSTRACT.** Iterated commutators of multilinear Calderón-Zygmund operators and pointwise multiplication with functions in  $BMO$  are studied in products of Lebesgue spaces. Both strong type and weak end-point estimates are obtained, including weighted results involving the vectors weights of the multilinear Calderón-Zygmund theory recently introduced in the literature. Some better than expected estimates for certain multilinear operators are presented too.

## 1. INTRODUCTION AND MAIN RESULTS

The commutator of a linear Calderón-Zygmund operator  $T$  and a  $BMO$  function  $b$ ,

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

was first studied by Coifman, Rochberg, and Weiss [2] who proved that

$$T_b : L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)$$

for all  $1 < p < \infty$ . This can be seen as a bilinear result,

$$BMO(\mathbb{R}^n) \times L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n),$$

since actually

$$(1.1) \quad \|T_b(f)\|_{L^p} \lesssim \|b\|_{BMO} \|f\|_{L^p}.$$

Using duality, the above estimate has as an immediate consequence for  $1 < p < \infty$  the bilinear estimate

$$(1.2) \quad \|gT(f) - fT^*(g)\|_{H^1} \lesssim \|g\|_{L^{p'}} \|f\|_{L^p},$$

where  $p'$  is the dual exponent of  $p$ ,  $H^1$  is the Hardy space, and  $T^*$  is the transpose of  $T$ . Note that (1.2) is a *better than expected estimate*, since trivially by Hölder's inequality and the boundedness of  $T$ ,

$$\|gT(f) - fT^*(g)\|_{L^1} \lesssim \|g\|_{L^{p'}} \|f\|_{L^p}.$$

---

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

*Key words and phrases.* Multilinear singular integrals, Calderón-Zygmund theory, maximal operators, weighted norm inequalities, commutators.

The authors would like to acknowledge the support of the following grants. First author and fourth author: Spanish Ministry of Science and Innovation grant MTM2009-08934 and MTM2008-05891 respectively, Second author: Instituto de Matemática Aplicada del Litoral (IMAL-CONICET)-Universidad Nacional del Litoral, Third author: NSF grant DMS0800492 and a General Research Fund allocation of the University of Kansas.

Both (1.1) and (1.2) put in evidence that some subtle cancellations are taking place. These estimates have found many important applications in other areas of operator theory and partial differential equations.

Another interesting feature of the commutators  $T_b$  is the fact that they fail to satisfy the typical weak end-point  $L^1$  estimate of the Calderón-Zygmund theory. As a remedial feature though, they do satisfy an alternative  $L(\log L)$  end-point estimate, as proved by Pérez in [9] (see [10] for a different proof).

Much of the analysis of linear commutators has been extended to other context such as weighted spaces, spaces of homogeneous type, multiparameter and multilinear settings. Higher order and iterated commutators have been considered too. The literature is by now quite vast. We will only recount here the multilinear situation which is the focus of this article. The purpose of the present work is to prove the optimal results for the iterated commutators and an associated multi(sub)linear operator. In this sense, this article complements and completes the theory developed by Lerner et al in [8], where the reader will find further bibliography in the subject.

Let  $T$  be an  $m$ -linear Calderón-Zygmund operator as defined by Grafakos and Torres in [5] and [7] (see the next section for complete definitions). In particular, such operators satisfy

$$(1.3) \quad T : L^{p_1}(\mathbb{R}^n) \times \cdots \times L^{p_m}(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)$$

whenever  $1 < p_1, \dots, p_m < \infty$  and

$$(1.4) \quad \frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_m},$$

and also the end-point result

$$(1.5) \quad T : L^1(\mathbb{R}^n) \times \cdots \times L^1(\mathbb{R}^n) \rightarrow L^{1/m, \infty}(\mathbb{R}^n).$$

Let  $\mathbf{b} = (b_1, \dots, b_m)$  be in  $BMO^m$ . The commutator of  $\mathbf{b}$  and the  $m$ -linear Calderón-Zygmund operator  $T$ , denoted here by  $T_{\Sigma \mathbf{b}}^1$ , was introduced by Pérez and Torres in [11] and is defined via

$$(1.6) \quad T_{\Sigma \mathbf{b}}(f_1, \dots, f_m) = \sum_{j=1}^m T_{b_j}^j(f_1, \dots, f_m),$$

where each term is the commutator of  $b_j$  and  $T$  in the  $j$ -th entry of  $T$ , that is,

$$\begin{aligned} T_{b_j}^j(\mathbf{f}) &\equiv T_{b_j}^j(f_1, \dots, f_m) \equiv [b_j, T]_j(f_1, \dots, f_m) \\ &\equiv b_j T(f_1, \dots, f_j, \dots, f_m) - T(f_1, \dots, b_j f_j, \dots, f_m). \end{aligned}$$

---

<sup>1</sup>The notation  $T_{\bar{b}}$  was used instead in [11] and [8]. We use the new notation to better differentiate this commutator from the iterated ones we want to study in this article. The notation for both types of commutators is also motivated by the estimates they satisfy.

It was shown in [11] that  $T_{\Sigma \mathbf{b}}$  satisfies the bounds (1.3) for all indices satisfying (1.4) with  $p > 1$ . The result was extended in [8] to all  $p > 1/m$ . The estimates are of the form

$$(1.7) \quad \|T_{\Sigma \mathbf{b}}(\mathbf{f})\|_{L^p} \lesssim \left( \sum_{j=1}^m \|b_j\|_{BMO} \right) \prod_{j=1}^m \|f_j\|_{L^{p_j}}$$

Moreover, weighted- $L^p$  versions of the bounds (1.3) were obtained in [8] for weights in the classes  $A_{\vec{p}}$  (see again the next section for definitions). These classes of weights introduced in [8] are the largest classes of weights for which all  $m$ -linear Calderón-Zygmund operators are bounded.

As it may be expected from the situation in the linear case, the end-point estimate (1.5) does not hold for  $T_{\Sigma \mathbf{b}}$ . Instead the following estimate was also obtained in [8]

$$(1.8) \quad |\{x \in \mathbb{R}^n : |T_{\Sigma \mathbf{b}}(\mathbf{f})(x)| > t^m\}| \leq C(\mathbf{b}) \prod_{j=1}^m \left( \int_{\mathbb{R}^n} \Phi \left( \frac{|f_j(x)|}{t} \right) dx \right)^{1/m},$$

where  $\Phi(t) = t(1 + \log^+ t)$ . The result is still true if the Lebesgue measure is changed by an  $A_{\vec{1}}$  weight (2.9). Note that for  $m = 1$  this is the end-point result in [9]. The estimate (1.8) is sharp in an appropriate sense. It is also the right one from the point of view interpolation as recently shown by Grafakos et al [4].

The results for  $T_{\Sigma \mathbf{b}}$  were obtained in [8] via corresponding ones for the maximal function

$$\mathcal{M}_{\Sigma L(\log L)} = \sum \mathcal{M}_{L(\log L)}^i,$$

where

$$\mathcal{M}_{L(\log L)}^i(\mathbf{f})(x) = \sup_{Q \ni x} \|f_i\|_{L(\log L), Q} \prod_{j \neq i} \frac{1}{|Q|} \int_Q f_j dx.$$

Independently, Tang [13] has also looked at  $T_{\Sigma \mathbf{b}}$ , iterations of it, and vector valued versions, but only for weights in the classical  $A_p$  classes (whose product is still smaller than  $A_{\vec{p}}$ ). He obtained some end-point estimates but with the right-hand side term in (1.8) replaced by a more complicated expression with an extra factor, and without the homogeneity of (1.8), which is crucial to obtain optimality.

We will establish in this article strong bounds for iterated commutators for  $p > 1/m$  allowing the full  $A_{\vec{p}}$  classes and again sharp end-point results when  $p = 1/m$ .

For a Calderón-Zygmund operator  $T$  and  $\mathbf{b} = (b_1, \dots, b_m)$  in  $BMO^m$ , we define the iterated commutators  $T_{\Pi \mathbf{b}}$  to be

$$(1.9) \quad T_{\Pi \mathbf{b}}(f_1, \dots, f_m) \equiv [b_1, [b_2, \dots [b_{m-1}, [b_m, T]_m]_{m-1} \dots]_2]_1(\mathbf{f}).$$

To clarify the notation, if  $T$  is associated in the usual way with a Calderón-Zygmund kernel  $K$ , then at a formal level

$$(1.10) \quad T_{\Pi \mathbf{b}}(\mathbf{f})(x) = \int_{(\mathbb{R}^n)^m} \prod_{j=1}^m (b_j(x) - b_j(y_j)) K(x, y_1, \dots, y_m) f_1(y_1) \dots f_m(y_m) dy_1 \dots dy_m.$$

(See also (1.14) below for another explicit formula in the bilinear case.)

We will prove the following strong type bound for  $T_{\Pi\mathbf{b}}$ .

**Theorem 1.1.** *Let  $T$  be an  $m$ -linear Calderón-Zygmund operator;  $\vec{w} \in A_{\vec{p}}$  with*

$$\frac{1}{p} = \frac{1}{p_1} + \dots + \frac{1}{p_m}$$

and  $1 < p_j < \infty, j = 1, \dots, m$ ; and  $\mathbf{b} \in BMO^m$ . Then, there exists a constant  $C$  such that

$$(1.11) \quad \|T_{\Pi\mathbf{b}}(\mathbf{f})\|_{L^p(\nu_{\vec{w}})} \leq C \prod_{j=1}^m \|b_j\|_{BMO} \prod_{j=1}^m \|f_j\|_{L^{p_j}(w_j)}.$$

At the end-point we obtain the following estimate.

**Theorem 1.2.** *Let  $T$  be an  $m$ -linear Calderón-Zygmund operator;  $\vec{w} \in A_{\vec{1}}$ , and  $\mathbf{b} \in BMO^m$ . Then, there exists a constant  $C$  depending on  $\mathbf{b}$  such that*

$$(1.12) \quad \nu_{\vec{w}}(\{x \in \mathbb{R}^n : |T_{\Pi\mathbf{b}}(\mathbf{f})(x)| > t^m\}) \leq C \prod_{j=1}^m \left( \int_{\mathbb{R}^n} \Phi^{(m)} \left( \frac{|f_j(x)|}{t} \right) w_j(x) dx \right)^{1/m},$$

where  $\Phi(t) = t(1 + \log^+ t)$  and  $\Phi^{(m)} = \overbrace{\Phi \circ \dots \circ \Phi}^m$ .

Moreover, the estimate is sharp in the sense that  $\Phi^{(m)}$  can not be replaced by  $\Phi^{(k)}$  for any  $k < m$ .

To prove the sharpness of theorem above we adapt some ideas from [8]. For simplicity, we consider  $m = 2, n = 1$ ,  $T$  one of the bilinear operators for  $n = 1$  obtained from the (linear) Riesz transforms in  $n = 2$ , as it is done for example in [8], and the functions  $b_1(x) = b_2(x) = \log |1 + x|$  and  $f_1 = f_2 = \chi_{(0,1)}$ . We can prove, for example, that the estimate

$$|\{x \in \mathbb{R}^n : |T_{\Pi\mathbf{b}}(\mathbf{f})(x)| > t^2\}| \leq C \left( \int_{\mathbb{R}^n} \Phi \left( \frac{|f_1|}{t} \right) \right)^{1/2} \left( \int_{\mathbb{R}^n} \Phi^{(2)} \left( \frac{|f_2|}{t} \right) \right)^{1/2}$$

is false. In fact, if the inequality above were to hold, by the homogeneity we would have that

$$|\{x \in \mathbb{R}^n : |T_{\Pi\mathbf{b}}(\mathbf{f})(x)| > t^2\}|^2 \leq C \int_{\mathbb{R}^n} \Phi \left( \frac{|f_1|}{t^2} \right) \int_{\mathbb{R}^n} \Phi^{(2)}(f_2),$$

and hence, since  $\Phi$  is a Young function

$$\sup_{\lambda > 0} \frac{1}{\Phi(1/\lambda)} |\{x \in \mathbb{R}^n : |T_{\Pi\mathbf{b}}(\mathbf{f})(x)| > \lambda\}|^2 < \infty.$$

However, using the fact that  $\Phi^{-1}(x) \cong x/(\log x)$  for  $x > e$ , it is easy to check that

$$\begin{aligned}
 & \sup_{\lambda > 0} \frac{1}{\Phi(1/\lambda)} |\{x \in \mathbb{R}^n : |T_{\Pi \mathbf{b}}(\mathbf{f})(x)| > \lambda\}|^2 \\
 & \leq \sup_{\lambda > 0} \frac{1}{\Phi(1/\lambda)} |\{x > e : \frac{\log^2(1+x)}{x^2} > \lambda\}|^2 \\
 & = \sup_{\lambda > 0} \frac{1}{\Phi(1/\lambda^2)} |\{x > e : \frac{\log(x)}{x} > \lambda\}|^2 \\
 & \geq \sup_{\lambda > 0} \frac{1}{\Phi(1/\lambda^2)} |\{x > e : \frac{C}{\Phi^{-1}(x)} > \lambda\}|^2 \\
 & \geq C \sup_{0 < \lambda < C/e} \frac{(\Phi(C/\lambda) - e)^2}{\Phi((C/\lambda)^2)} \\
 & \geq \frac{C}{4} \sup_{0 < \lambda < C/2e} \frac{(\Phi(C/\lambda))^2}{\Phi((C/\lambda)^2)} \\
 & \geq \frac{C}{4} \sup_{0 < \lambda < C/2e} \log(C/\lambda) = \infty.
 \end{aligned}$$

As in the linear case and the particular multilinear case studied in [8], the proofs of the two main theorems will be based on corresponding estimates on a maximal function that controls the commutator, the operator  $\mathcal{M}_{L(\log L)}$  given by

$$(1.13) \quad \mathcal{M}_{L(\log L)}(\mathbf{f})(x) = \sup_{Q \ni x} \prod_{j=1}^m \|f_j\|_{L(\log L), Q},$$

where the supremum is taken over all cubes  $Q$  containing  $x$ . Strong bounds for this operator were already obtained in [8] but not weak-type ones. We present in this article the right end-point distributional estimate it satisfies (see Theorem 4.1). This operator and the estimates it satisfies are crucial in this paper.

Our analysis will show that in fact one can also study commutators where only  $k < m$  factors appear in (1.10), and which are controlled by an appropriate modification of the maximal function  $\mathcal{M}_{L(\log L)}$ . We will concentrate only in the case where there are  $m$  functions in  $BMO$ , which is the most difficult one, and leave other generalizations to the interested reader. See, however Section 3 below.

The next section contains some basic definitions and further background related to the classes  $A_{\vec{p}}$  of vector weights and several multilinear maximal functions from [8]. Nevertheless, the reader already familiar with the subject can skip Section 2 and move directly to Section 3, where a key pointwise estimate involving the maximal function  $\mathcal{M}_{L(\log L)}$ , Theorem 3.1, is combined with the classical Fefferman-Stein inequality to prove the strong bounds in Theorem 1.1. Likewise, the proof of Theorem 1.2 is obtained

using a new weak type estimate for the maximal function  $\mathcal{M}_{L(\log L)}$ , Theorem 4.1, which is presented in Section 4.

Before we conclude this introduction, we would like to consider analogs of (1.2) in the multilinear setting in view of (1.11) and put in evidence again some *better than expected estimates*, which are implied by the commutator results and which also motivate in part our study of commutators. For simplicity we consider the following particular case. For a bilinear Calderón-Zygmund operator  $T$ , we can write

$$(1.14) \quad T_{\Pi\mathbf{b}}(f_1, f_2) = b_1 b_2 T(f_1, f_2) - b_2 T(b_1 f_1, f_2) - b_1 T(f_1, b_2 f_2) + T(b_1 f_1, b_2 f_2).$$

We can use duality to obtain the surprising quad-linear estimate

$$(1.15) \quad \begin{aligned} & \|hb_2 T(f_1, f_2) - f_1 T^{*1}(hb_2, f_2) - hT(f_1, b_2 f_2) + f_1 T^{*1}(h, b_2 f_2)\|_{H^1} \\ & \lesssim \|b_2\|_{BMO} \|h\|_{L^{p'}} \|f_1\|_{L^q} \|f_2\|_{L^r}, \end{aligned}$$

for  $1/q + 1/r = 1/p$ ,  $1 < p, q, r < \infty$ , and where  $T^{*1}$  is the transpose of  $T$  in the first variable. Notice that this is again an improvement (now both in the target and the range) over the trivial estimate

$$S : L^\infty(\mathbb{R}^n) \times L^{p'}(\mathbb{R}^n) \times L^q(\mathbb{R}^n) \times L^r(\mathbb{R}^n) \rightarrow L^1(\mathbb{R}^n),$$

where

$$S(b, h, f_1, f_2) = hbT(f_1, f_2) - f_1 T^{*1}(hb, f_2) - hT(f_1, bf_2) + f_1 T^{*1}(h, bf_2),$$

and which follows by Hölder's inequality and the boundedness of  $T$ . The better estimate obtained reflects again the presence of certain hidden cancellations. Though we will not carry their study here any further, it would be interested to see if estimates like (1.15) are amenable to some analysis similar to the one generated in the linear case as consequence of (1.2).

**Acknowledgement.** We would like to thank Hua Wang for pointing out some typos and a gap in a previous version of this manuscript which we have now corrected and filled in. Also we would to thank the referee for the valuable comments.

## 2. SOME BACKGROUND DEFINITIONS AND ESTIMATES

**2.1. Calderón-Zygmund operators.** Following [5] we will assume here that  $T$  is a bounded  $m$ -linear Calderón-Zygmund operator. That is,  $T$  satisfies the bounds (1.3) and (1.5) and its Schwartz kernel  $K$  satisfies away from the diagonal  $x = y_1 = \dots = y_m$  in  $(\mathbb{R}^n)^{m+1}$ ,

$$(2.1) \quad |K(y_0, y_1, \dots, y_m)| \leq \frac{A}{\left(\sum_{k,l=0}^m |y_k - y_l|\right)^{mn}}$$

and also

$$(2.2) \quad |K(y_0, \dots, y_j, \dots, y_m) - K(y_0, \dots, y'_j, \dots, y_m)| \leq \frac{A|y_j - y'_j|^\varepsilon}{\left(\sum_{k,l=0}^m |y_k - y_l|\right)^{mn+\varepsilon}},$$

for some  $\varepsilon > 0$  and all  $0 \leq j \leq m$ , whenever  $|y_j - y'_j| \leq \frac{1}{2} \max_{0 \leq k \leq m} |y_j - y_k|$ . In particular for  $x \notin \cap \text{supp } f_j$ ,

$$T(f_1, \dots, f_m)(x) = \int K(x, y_1, \dots, y_m) f_1(y_1) \dots f_m(y_m) dy_1 \dots dy_m.$$

**2.2. Orlicz norms.** For  $\Phi(t) = t(1 + \log^+ t)$  and a cube  $Q$  in  $\mathbb{R}^n$  we will consider the average  $\|f\|_{\Phi, Q}$  of a function  $f$  given by the Luxemburg norm

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

We will need the several basic estimates from the theory of Orlicz spaces. We first recall that

$$(2.3) \quad \|f\|_{\Phi, Q} > 1 \quad \text{if and only if} \quad \frac{1}{|Q|} \int_Q \Phi(|f(x)|) dx > 1.$$

Next, we note that the generalized Hölder inequality in Orlicz spaces together with the John-Nirenberg inequality implies that

$$(2.4) \quad \frac{1}{|Q|} \int_Q |b(y) - b_Q| f(y) dy \leq C \|b\|_{BMO} \|f\|_{L(\log L), Q},$$

an estimate that we shall use in several occasions without further comment.

We will also use the maximal function

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

where the supremum is taken over all the cubes containing  $x$ . This operator satisfies the pointwise equivalence

$$(2.5) \quad M_{L(\log L)} f(x) \approx M^2 f(x),$$

where  $M$  is the Hardy-Littlewood maximal function, and we will also employ several times the Kolmogorov inequality

$$(2.6) \quad \|f\|_{L^p(Q, \frac{dx}{|Q|})} \leq C \|f\|_{L^q(Q, \frac{dx}{|Q|})},$$

for  $0 < p < q < \infty$ . See, e.g. [14] and the reference in [8].

**2.3. Sharp maximal functions.** For  $\delta > 0$ ,  $M_\delta$  is the maximal function

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

In addition,  $M^\#$  is the sharp maximal function of Fefferman and Stein [3],

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

and

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

We will also use from [3], the inequality

$$(2.7) \quad \int_{\mathbb{R}^n} (M_\delta f(x))^p w(x) dx \leq C \int_{\mathbb{R}^n} (M_\delta^\# f(x))^p w(x) dx,$$

for all function  $f$  for which the left-hand side is finite, and where  $0 < p, \delta < \infty$  and  $w$  is a weight in  $A_\infty$ . Moreover, if  $\varphi : (0, \infty) \rightarrow (0, \infty)$  is doubling, then there exists a constant  $c$  (depending on the  $A_\infty$  constant of  $w$  and the doubling constant of  $\varphi$ ) such that

$$(2.8) \quad \sup_{\lambda > 0} \varphi(\lambda) w(\{y \in \mathbb{R}^n : M_\delta f(y) < \lambda\}) \leq c \sup_{\lambda > 0} \varphi(\lambda) w(\{y \in \mathbb{R}^n : M_\delta^\# f(y) < \lambda\}),$$

again for every function  $f$  such that the left hand side is finite.

**2.4. Multiple weights.** Following the notation in [8], for  $m$  exponents  $p_1, \dots, p_m$ , we will often write  $p$  for the number given by  $\frac{1}{p} = \frac{1}{p_1} + \dots + \frac{1}{p_m}$ , and  $\vec{P}$  for the vector  $\vec{P} = (p_1, \dots, p_m)$ .

Let  $1 \leq p_1, \dots, p_m < \infty$ , a multiple weight  $\vec{w} = (w_1, \dots, w_m)$ , is said to satisfy the *multilinear  $A_{\vec{P}}$  condition* if for

$$\nu_{\vec{w}} = \prod_{j=1}^m w_j^{p/p_j}.$$

it holds that

$$(2.9) \quad \sup_Q \left( \frac{1}{|Q|} \int_Q \nu_{\vec{w}} \right)^{1/p} \prod_{j=1}^m \left( \frac{1}{|Q|} \int_Q w_j^{1-p'_j} \right)^{1/p'_j} < \infty.$$

When  $p_j = 1$ ,  $\left( \frac{1}{|Q|} \int_Q w_j^{1-p'_j} \right)^{1/p'_j}$  is understood as  $(\inf_Q w_j)^{-1}$ .

One can check that  $A_{(1, \dots, 1)}$  is contained in  $A_{\vec{P}}$  for each  $\vec{P}$ , however the classes  $A_{\vec{P}}$  are not increasing with the natural partial order. As mentioned in the introduction, these are the largest classes of weights for which the multilinear Calderón-Zygmund



operators are bounded on Lebesgue spaces, as proved in [8]) improving on the results in [6] and [11]. In fact, one has

$$\prod_{j=1}^m A_{p_j} \subset A_{\vec{p}},$$

with strict containment. Moreover, in general  $\vec{w} \in A_{\vec{p}}$  does not imply  $w_j \in L_{\text{loc}}^1$  for any  $j$ , but instead

$$(2.10) \quad \vec{w} \in A_{\vec{p}} \iff \begin{cases} w_j^{1-p'_j} \in A_{mp'_j}, & j = 1, \dots, m \\ \nu_{\vec{w}} \in A_{mp}, \end{cases}$$

where the condition  $w_j^{1-p'_j} \in A_{mp'_j}$  in the case  $p_j = 1$  is understood as  $w_j^{1/m} \in A_1$ .

Observe that in the linear case ( $m = 1$ ) both conditions included in (2.10) represent the same  $A_p$  condition. However, when  $m \geq 2$  neither of the conditions in (2.10) implies the other. We refer the reader to [8] for more details on this multilinear weights.

### 3. PROOF OF THEOREM 1.1

The proof of the Theorem 1.1 will rely on a pointwise estimate using sharp maximal functions. The technique of comparing commutators with sharp maximal operators has by now a long history of successful applications (see the comments in [8] p.15 and the references therein).

To state the pointwise result in great generality we need to introduce some additional for  $m$ -linear iterated commutators involving  $j$   $BMO$  functions with  $j < m$ . Following [12], for positive integers  $m$  and  $j$  with  $1 \leq j \leq m$ , we denote by  $C_j^m$  the family of all finite subsets  $\sigma = \{\sigma(1), \dots, \sigma(j)\}$  of  $\{1, \dots, m\}$  of  $j$  different elements, where we always take  $\sigma(k) < \sigma(l)$  if  $k < l$ . For any  $\sigma \in C_j^m$ , we associate the complementary sequence  $\sigma' \in C_{m-j}^m$  given by  $\sigma' = \{1, \dots, m\} \setminus \sigma$  with the convention  $C_0^m = \emptyset$ . Given an  $m$ -tuple of functions  $\mathbf{b}$  and  $\sigma \in C_j^m$ , we also use the notation  $\mathbf{b}_\sigma$  for the  $j$ -tuple obtained from  $\mathbf{b}$  given by  $(b_{\sigma(1)}, \dots, b_{\sigma(j)})$ .

Similarly to (1.9), we define for a Calderón-Zygmund operator  $T$ ,  $\sigma \in C_j^m$ , and  $\mathbf{b}_\sigma = (b_{\sigma(1)}, \dots, b_{\sigma(j)})$  in  $BMO^j$ , the iterated commutator

$$(3.1) \quad T_{\Pi \mathbf{b}_\sigma}(f_1, \dots, f_m) = [b_{\sigma(1)}, [b_{\sigma(2)}, \dots [b_{\sigma(j-1)}, [b_{\sigma(j)}, T]_{\sigma(j)}]_{\sigma(j-1)} \dots]_{\sigma(2)}]_{\sigma(1)}(\mathbf{f}).$$

That is, formally

$$T_{\Pi \mathbf{b}_\sigma}(\mathbf{f})(x) = \int_{(\mathbb{R}^n)^m} \left( \prod_{i=1}^j (b_{\sigma(i)}(x) - b_{\sigma(i)}(y_{\sigma(i)})) \right) K(x, y_1, \dots, y_m) \prod_{i=1}^m f_i(y_i) d\mathbf{y},$$

where  $d\mathbf{y} = dy_1 \dots dy_m$ . Clearly  $T_{\Pi \mathbf{b}_\sigma} = T_{\Pi \mathbf{b}}$  as defined before when  $\sigma = \{1, 2, \dots, m\}$ , while  $T_{\Pi \mathbf{b}_\sigma} = T_{b_j}^j$  when  $\sigma = \{j\}$ .

The pointwise estimate that will serve our purposes is given by the next result. The statement and its proof are similar to those of [12, Lemma 3.1] and [13, Lemma 3.4],

which the reader may consult for further details. The difference here is the use of the operator  $\mathcal{M}_{L(\log L)}$ , which eventually allows to use vector weights.

**Theorem 3.1.** *Let  $T_{\Pi \mathbf{b}}$  be a multilinear commutator with  $\mathbf{b} \in BMO^m$  and let  $0 < \delta < \varepsilon$ , with  $0 < \delta < 1/m$ . Then, there exists a constant  $C > 0$ , depending on  $\delta$  and  $\varepsilon$ , such that*

$$(3.2) \quad M_{\delta}^{\#}(T_{\Pi \mathbf{b}}(\mathbf{f}))(x) \leq C \prod_{j=1}^m \|b_j\|_{BMO} \left( \mathcal{M}_{L(\log L)}(\mathbf{f})(x) + M_{\varepsilon}(T(\mathbf{f}))(x) \right) \\ + C \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} \prod_{i=1}^j \|b_{\sigma(i)}\|_{BMO} M_{\varepsilon}(T_{\Pi \mathbf{b}_{\sigma}}(\mathbf{f}))(x)$$

for all  $m$ -tuples  $\mathbf{f} = (f_1, \dots, f_m)$  of bounded measurable functions with compact support.

*Proof.* The way to interpret (3.2) is

$$M_{\delta}^{\#}(T_{\Pi \mathbf{b}}(\mathbf{f}))(x) \lesssim \prod_{j=1}^m \|b_j\|_{BMO} \mathcal{M}_{L(\log L)}(\mathbf{f})(x) + \text{“lower order terms”},$$

as it will become apparent in its application. Given the heavy technical notation and for simplicity in the exposition, we only present the case  $m = 2$ . As the reader may soon see, the general case is only notationally more complicated and can be obtained with a similar procedure. Hence, we will limit our selves to establish the following version of (3.2).

For  $b_1, b_2 \in BMO$  we will show that

$$M_{\delta}^{\#}(T_{\Pi \mathbf{b}}(f_1, f_2))(x) \leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \left( \mathcal{M}_{L(\log L)}(f_1, f_2)(x) + M_{\varepsilon}(T(f_1, f_2))(x) \right) \\ + C \left( \|b_2\|_{BMO} M_{\varepsilon}(T_{b_1}^1(f_1, f_2))(x) + \|b_1\|_{BMO} M_{\varepsilon}(T_{b_2}^2(f_1, f_2))(x) \right).$$

For any constants  $\lambda_1$  and  $\lambda_2$ , write

$$T_{\Pi \mathbf{b}}(\mathbf{f})(x) = (b_1(x) - \lambda_1)(b_2(x) - \lambda_2)T(f_1, f_2)(x) - (b_1(x) - \lambda_1)T(f_1, (b_2 - \lambda_2)f_2)(x) \\ - (b_2(x) - \lambda_2)T((b_1 - \lambda_1)f_1, f_2)(x) + T((b_1 - \lambda_1)f_1, (b_2 - \lambda_2)f_2)(x). \\ = -(b_1(x) - \lambda_1)(b_2(x) - \lambda_2)T(f_1, f_2)(x) + (b_1(x) - \lambda_1)T_{b_2 - \lambda_2}^2(f_1, f_2)(x) \\ + (b_2(x) - \lambda_2)T_{b_1 - \lambda_1}^1(f_1, f_2)(x) + T((b_1 - \lambda_1)f_1, (b_2 - \lambda_2)f_2)(x).$$

Also, if we fix  $x \in \mathbb{R}^n$ , a cube  $Q$  centered at  $x$  and a constant  $c$ , then since  $0 < \delta < 1/2$ , we can estimate

$$\left( \frac{1}{|Q|} \int_Q |T_{\Pi \mathbf{b}}(\mathbf{f})(z)|^{\delta} - |c|^{\delta} dz \right)^{1/\delta} \leq \left( \frac{1}{|Q|} \int_Q |T_{\Pi \mathbf{b}}(\mathbf{f})(z) - c|^{\delta} dz \right)^{1/\delta} \\ \leq \left( \frac{C}{|Q|} \int_Q |(b_1(x) - \lambda_1)(b_2(x) - \lambda_2)T(f_1, f_2)(z)|^{\delta} dz \right)^{1/\delta}$$

$$\begin{aligned}
 & + \left( \frac{C}{|Q|} \int_Q |(b_1(x) - \lambda_1) T_{b_2 - \lambda_2}^2(f_1, f_2)(z)|^\delta dz \right)^{1/\delta} \\
 & + \left( \frac{C}{|Q|} \int_Q |(b_2(x) - \lambda_2) T_{b_1 - \lambda_1}^1(f_1, f_2)(z)|^\delta dz \right)^{1/\delta} \\
 & + \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1, (b_2 - \lambda_2)f_2)(z) - c|^\delta dz \right)^{1/\delta} \\
 & = I + II + III + IV.
 \end{aligned}$$

We analyze each term separately selecting appropriate constants. Let  $Q^* = 3Q$  and let  $\lambda_j = (b_j)_{Q^*}$  be the average of  $b_j$  on  $Q^*$ ,  $j = 1, 2$ . For any  $1 < q_1, q_2, q_3 < \infty$  with  $1 = 1/q_1 + 1/q_2 + 1/q_3$  and  $q_3 < \varepsilon/\delta$  we have by Hölder's and Jensen's inequalities,

$$\begin{aligned}
 I & \leq C \left( \frac{1}{|Q|} \int_Q |b_1(z) - \lambda_1|^{\delta q_1} dz \right)^{1/\delta q_1} \left( \frac{1}{|Q|} \int_Q |b_2(z) - \lambda_2|^{\delta q_2} dz \right)^{1/\delta q_2} \\
 & \quad \times \left( \frac{1}{|Q|} \int_Q |T(f_1, f_2)(z)|^{\delta q_3} dz \right)^{1/\delta q_3}.
 \end{aligned}$$

From the fact that, for every  $q > 0$

$$\begin{aligned}
 \frac{1}{|Q|} \int_Q |b_j(z) - (b_j)_{Q^*}|^q dz & \leq C \frac{1}{|Q|} \int_Q |b_j(z) - (b_j)_Q|^q dz + C |(b_j)_Q - (b_j)_{Q^*}|^q \\
 & \leq C \|b_j\|_{BMO}^q
 \end{aligned}$$

we obtain that

$$\begin{aligned}
 I & \leq C \|b_1\|_{BMO} \|b_2\|_{BMO} M_{\delta q_3}(T(f_1, f_2))(x) \\
 & \leq C \|b_1\|_{BMO} \|b_2\|_{BMO} M_\varepsilon(T(f_1, f_2))(x),
 \end{aligned}$$

which is an appropriate estimate for what we want to obtain.

Since  $II$  and  $III$  are symmetric we only study  $II$ . Let  $1 < t_1, t_2 < \infty$  with  $1 = 1/t_1 + 1/t_2$  and  $t_2 < \varepsilon/\delta$  then, by Hölder's and Jensen's inequalities,

$$\begin{aligned}
 II & \leq C \left( \frac{1}{|Q|} \int_Q |b_1(z) - \lambda_1|^{\delta t_1} dz \right)^{1/\delta t_1} \left( \frac{1}{|Q|} \int_Q |T_{b_2 - \lambda_2}^2(f_1, f_2)(z)|^{\delta t_2} dz \right)^{1/\delta t_2} \\
 & \leq C \|b_1\|_{BMO} M_{\delta t_2}(T_{b_2 - \lambda_2}^2(f_1, f_2))(x) \\
 & \leq C \|b_1\|_{BMO} M_\varepsilon(T_{b_2 - \lambda_2}^2(f_1, f_2))(x) \\
 & = C \|b_1\|_{BMO} M_\varepsilon(T_{b_2}^2(f_1, f_2))(x).
 \end{aligned}$$

Similarly,

$$III \leq C \|b_2\|_{BMO} M_\varepsilon(T_{b_1 - \lambda_1}^1(f_1, f_2))(x) = C \|b_2\|_{BMO} M_\varepsilon(T_{b_1}^1(f_1, f_2))(x).$$

It only remain to study the last term  $IV$ . We split each  $f_i$  as  $f_i = f_i^0 + f_i^\infty$  where  $f_i^0 = f\chi_{Q^*}$  and  $f_i^\infty = f_i - f_i^0$ . Let

$$c = \sum_{j=1}^3 c_j,$$

where

$$\begin{aligned} c_1 &= T(f_1^0, (b_2 - \lambda_2)f_2^\infty)(x), \\ c_2 &= T(f_1^\infty, (b_2 - \lambda_2)f_2^0)(x), \\ c_3 &= T(f_1^\infty, (b_2 - \lambda_2)f_2^\infty)(x). \end{aligned}$$

Then,

$$\begin{aligned} IV &= \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1, (b_2 - \lambda_2)f_2)(z) - c|^\delta dz \right)^{1/\delta} \\ &\leq \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^0)(z)|^\delta dz \right)^{1/\delta} \\ &\quad + \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^\infty)(z) - c_1|^\delta dz \right)^{1/\delta} \\ &\quad + \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1^\infty, (b_2 - \lambda_2)f_2^0)(z) - c_2|^\delta dz \right)^{1/\delta} \\ &\quad + \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1^\infty, (b_2 - \lambda_2)f_2^\infty)(z) - c_3|^\delta dz \right)^{1/\delta} \\ &= IV_1 + IV_2 + IV_3 + IV_4 \end{aligned}$$

We choose  $1 < p < 1/(2\delta)$ . Since  $p\delta < 1/2$ , we can estimate  $IV_1$  using Hölder's inequality and the fact that  $T$  is a Calderón-Zygmund operator

$$\begin{aligned} IV_1 &\leq \left( \frac{C}{|Q|} \int_Q |T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^0)(z)|^{p\delta} dz \right)^{1/(p\delta)} \\ &\leq C \|T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^0)\|_{L^{1/2, \infty}(Q, \frac{dx}{|Q|})} \\ &\leq C \frac{1}{|Q|} \int_Q |(b_1(z) - \lambda_1)f_1^0(z)| dz \frac{1}{|Q|} \int_Q |(b_2(z) - \lambda_2)f_2^0(z)| dz \\ &\leq C \|b_1\|_{BMO} \|f_1\|_{L(\log L), Q} \|b_2\|_{BMO} \|f_2\|_{L(\log L), Q} \\ &\leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \mathcal{M}_{L(\log L)}(f_1, f_2)(x). \end{aligned}$$

Since  $IV_2$  and  $IV_3$  are symmetric, we consider for example  $IV_2$ , and estimate

$$\begin{aligned}
 & |T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^\infty)(z) - T((b_1 - \lambda_1)f_1^0, (b_2 - \lambda_2)f_2^\infty)(x)| \\
 & \leq \int_{3Q} |(b_1(y_1) - \lambda_1)f_1(y_1)| \int_{\mathbb{R}^n \setminus 3Q} \frac{|x - z|^\varepsilon |(b_2(y_2) - \lambda_2)f_2(y_2)| dy_2}{(|z - y_1| + |z - y_2|)^{2n+\varepsilon}} dy_1 \\
 & \leq \int_{3Q} |b_1(y_1) - \lambda_1| f_1(y_1) |dy_1 \sum_{k=1}^{\infty} \frac{|Q|^{\varepsilon/n}}{((3^k|Q|)^{1/n})^{2n+\varepsilon}} \int_{(3^{k+1}Q)} |(b_2(y_2) - \lambda_2)f_2(y_2)| dy_2 \\
 & \leq C \sum_{k=1}^{\infty} \frac{|Q|^{\varepsilon/n}}{((3^k|Q|)^{1/n})^{2n+\varepsilon}} \left( \int_{(3^{k+1}Q)} |b_1(y_1) - \lambda_1| f_1(y_1) |dy_1 \right) \times \\
 & \quad \times \left( \int_{(3^{k+1}Q)} |b_2(y_2) - \lambda_2| f_2(y_2) |dy_2 \right) \\
 & \leq C \sum_{k=1}^{\infty} \frac{k^2}{3^{\varepsilon k}} \|b_1\|_{BMO} \|b_2\|_{BMO} \|f_1\|_{L(\log L), 3^{k+1}Q} \|f_2\|_{L(\log L), 3^{k+1}Q} \\
 & \leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \mathcal{M}_{L(\log L)}(f_1, f_2)(x).
 \end{aligned}$$

Fiinally, the term  $IV_4$  is estimated in similar way and we deduce

$$\begin{aligned}
 & |T((b_1 - \lambda_1)f_1^\infty, (b_2 - \lambda_2)f_2^\infty)(z) - T((b_1 - \lambda_1)f_1^\infty, (b_2 - \lambda_2)f_2^\infty)(x)| \leq \\
 & \leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \mathcal{M}_{L(\log L)}(f_1, f_2)(x).
 \end{aligned}$$

The proof is complete.  $\square$

We note that we can also obtain analogous estimates to (3.2) for  $m$ -linear commutators involving  $j < m$  functions in  $BMO$ . That is estimates of the form

$$(3.3) \quad M_\delta^\#(T_{\Pi \mathbf{b}_\sigma}(\mathbf{f}))(x) \lesssim \prod_{k=1}^j \|b_{\sigma(k)}\|_{BMO} \mathcal{M}_{L(\log L)_\sigma}(\mathbf{f})(x) + \text{“lower order terms”},$$

where  $\mathcal{M}_{L(\log L)_\sigma}$  denotes the analog of  $\mathcal{M}_{L(\log L)}$  but with only log factors in the  $\mathbf{f}_\sigma$  functions. (Note that  $\mathcal{M}_{L(\log L)_\sigma} = \mathcal{M}_{L(\log L)}^j$  when  $\sigma = \{j\}$ .) The lower order terms are now of the form

$$\prod_{k=1}^l \|b_{\eta'(k)}\|_{BMO} M_\varepsilon(T_{\Pi \mathbf{b}_\eta}(\mathbf{f}))(x)$$

for  $l < j$ , where  $\eta$  is subset of  $\sigma$  of cardinality  $l$ , and  $\eta \cup \eta' = \sigma$ . Note also that trivially

$$(3.4) \quad \mathcal{M}_{L(\log L)_\sigma}(\mathbf{f})(x) \leq \mathcal{M}_{L(\log L)}(\mathbf{f})(x).$$

These pointwise estimates are the key for the strong and weak-type estimates with multiple weights. In particular, they yield an appropriate version of the following Coifman-Fefferman type inequalities ([1]).

**Theorem 3.2.** *Let  $0 < p < \infty$ , let  $w$  be a weight in  $A_\infty$ , and suppose that  $\mathbf{b} \in BMO^m$ . Then, there exists a constant  $C_w$  (independent of  $\mathbf{b}$ ) and a constant  $c_w(\mathbf{b})$  such that*

$$(3.5) \quad \int_{\mathbb{R}^n} |T_{\Pi \mathbf{b}}(\mathbf{f})(x)|^p w(x) dx \leq C_w \prod_{j=1}^m \|b_j\|_{BMO} \int_{\mathbb{R}^n} \mathcal{M}_{L(\log L)}(\mathbf{f})(x)^p w(x) dx,$$

and

$$(3.6) \quad \sup_{t>0} \frac{1}{\Phi^{(m)}(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_{\Pi \mathbf{b}}(\mathbf{f})(y)| > t^m\}) \leq c_w(\mathbf{b}) \sup_{t>0} \frac{1}{\Phi^{(m)}(\frac{1}{t})} w(\{y \in \mathbb{R}^n : \mathcal{M}_{L(\log L)}(\mathbf{f})(y) > t^m\}),$$

for all  $\mathbf{f} = (f_1, \dots, f_m)$  bounded with compact support.

*Proof.* The proof of these types of estimates is by now standard. We refer the reader to [12, Theorem 1.6] and [8, Corollary 3.8 and Theorem 3.19]. The arguments there can be followed step by step in this new case. We briefly indicate such arguments in the case  $m = 2$ , but, as the reader will immediately notice, an iterative procedure using (3.3) and (3.4) can be followed to obtain the general case.

As already mentioned the philosophy in the approach is that terms involving  $M_\epsilon$  can actually be treated as “lower order terms”. In fact, as in [8] and using the Fefferman-Stein inequality (2.7),

$$(3.7) \quad \begin{aligned} \|T_{\Pi \mathbf{b}}(\vec{f})\|_{L^p(w)} &\leq \|M_\delta(T_{\Pi \mathbf{b}}(\vec{f}))\|_{L^p(w)} \\ &\leq C \|M_\delta^\#(T_{\Pi \mathbf{b}}(\vec{f}))\|_{L^p(w)}. \end{aligned}$$

Using the pointwise estimate in the previous theorem and again the Fefferman-Stein inequality we can continue from (3.7) with

$$\begin{aligned} &\leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \left( \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)} + \|M_\epsilon(T(\vec{f}))\|_{L^p(w)} \right) \\ &\quad + C (\|b_2\|_{BMO} \|M_\epsilon(T_{b_1}^1(f_1, f_2))\|_{L^p(w)} + \|b_1\|_{BMO} \|M_\epsilon(T_{b_2}^2(f_1, f_2))\|_{L^p(w)}) \\ &\leq C \|b_1\|_{BMO} \|b_2\|_{BMO} \left( \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)} + \|M_\epsilon^\#(T(\vec{f}))\|_{L^p(w)} \right) \\ &\quad + C (\|b_2\|_{BMO} \|M_\epsilon^\#(T_{b_1}^1(f_1, f_2))\|_{L^p(w)} + \|b_1\|_{BMO} \|M_\epsilon^\#(T_{b_2}^2(f_1, f_2))\|_{L^p(w)}). \end{aligned}$$

If we take  $\epsilon$  small, we can now repeat the procedure using the results in [8] and estimate

$$\|M_\epsilon^\#(T(\vec{f}))\|_{L^p(w)} \leq C \|\mathcal{M}(\vec{f})\|_{L^p(w)} \leq C \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)};$$

and for  $\epsilon < \epsilon'$ ,

$$\begin{aligned} \|M_\epsilon^\#(T_{b_1}^1(f_1, f_2))\|_{L^p(w)} &\leq C \|b_1\|_{BMO} \left( \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)} + \|M_{\epsilon'}(T(f_1, f_2))\|_{L^p(w)} \right) \\ &\leq C \|b_1\|_{BMO} \left( \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)} + \|M_{\epsilon'}^\#(T(f_1, f_2))\|_{L^p(w)} \right) \\ &\leq C \|b_1\|_{BMO} \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)}. \end{aligned}$$

Similarly,

$$\|M_\epsilon^\#(T_{b_2}^2(f_1, f_2))\|_{L^p(w)} \leq C \|b_2\|_{BMO} \|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(w)}.$$

The desired inequality now follows.

We observe that to use the Fefferman-Stein inequality as argued in [8, pp.32-33], one needs to verify that certain terms in the left-hand side of the inequalities are finite when the right-hand side ones are (when the right-hand are infinite there is nothing to prove). However, if one assumes  $\mathbf{b}$  in  $(L^\infty)^m$ , then everything is clear because of the boundedness properties of  $T$ . The passage to  $\mathbf{b}$  in  $BMO^m$  is standard, and combining it with Fatou's lemma, one gets the desired result.

The proof of (3.6) also follows the pattern for the corresponding estimate relating  $T_{\Sigma \mathbf{b}}$  and  $\mathcal{M}_{\Sigma L(\log L)}$  in [8, pp. 33-35]. We also briefly indicate some of the details needed when  $m = 2$ . To further simplify the presentation, and since we do not intend to keep track of the exact dependence on  $\mathbf{b}$ , we assume that the  $BMO$  norms of the functions  $b_j$ s are equal to one. Note that the doubling properties of  $\Phi$  will produce a constant  $c(\mathbf{b})$  in the general case. It should be noted though that, unlike the strong case, such constant is not multilinear in  $\mathbf{b}$ .

Using the pointwise estimate for  $M_\delta^\#(T_{\Pi \mathbf{b}}(\vec{f}))$  we get

$$\begin{aligned} \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{y \in \mathbb{R}^n | T_{\Pi \mathbf{b}}(\mathbf{f})(y)| > t^2\}) &\leq \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{y \in \mathbb{R}^n | M_\delta(T_{\Pi \mathbf{b}}(\mathbf{f}))(y) > t^2\}) \\ &\leq \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_\delta^\#(T_{\Pi \mathbf{b}}(\vec{f}))(x) > t^2\}) \\ &\leq C \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{\mathcal{M}_{L(\log L)}(f_1, f_2)(x) > t^2\}) \\ &\quad + C \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_\epsilon(T(f_1, f_2))(x) > t^2\}) \\ &\quad + C \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_\epsilon(T_{b_1}^1(f_1, f_2))(x) > t^2\}) \\ &\quad + C \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_\epsilon(T_{b_2}^2(f_1, f_2))(x) > t^2\}) \\ &= I + II + III + IV. \end{aligned}$$

We claim that the main term is  $I$ , which will give the desired result. In fact, from the estimates

$$(3.8) \quad M_\epsilon^\#(T(f_1, f_2))(x) \leq \mathcal{M}(\vec{f})(x) \leq \mathcal{M}_{L \log L}(\vec{f})(x)$$

and the weak-type version of the Fefferman-Stein inequality (2.8) we easily get that  $II \lesssim I$ .

To estimate *III* we invoke again (2.8) and the results for the  $M^\#$  function of commutators of lower order to get for  $\epsilon < \epsilon'$ ,

$$\begin{aligned} III &\leq C \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_\epsilon^\#(T_{b_1}^1(f_1, f_2))(x) > t^2\}) \\ &\leq \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{\mathcal{M}_{L(\log L)}(\vec{f})(x) > t^2\}) \\ &\quad + \sup_{t>0} \frac{1}{\Phi^{(2)}(\frac{1}{t})} w(\{M_{\epsilon'}(T(f_1, f_2))(x) > t^2\}). \end{aligned}$$

Iterating the procedure and using (3.8) we arrive to  $III \lesssim I$ . The term *IV* is completely analogous. Again, to be able to apply (2.8) some justification is needed. But one can always assume the weight to be bounded and use a limiting process. We refer to [8] and omit the rest of the details.  $\square$

*Proof of Theorem 1.1* We can now easily finish the proof of Theorem 1.1. Since for  $\vec{w}$  in  $A_{\vec{p}}$ , the weight  $\nu_{\vec{w}}$  is in  $A_\infty$ , we can use one more result from [8] on strong bounds for  $\mathcal{M}_{L(\log L)}$  and conclude from (3.5) that

$$\begin{aligned} \int_{\mathbb{R}^n} |T_{\Pi \mathbf{b}}(\mathbf{f})(x)|^p \nu_{\vec{w}}(x) dx &\leq C_{\nu_{\vec{w}}} \prod_{j=1}^m \|b_j\|_{BMO} \int_{\mathbb{R}^n} (\mathcal{M}_{L(\log L)}(\mathbf{f})(x))^p \nu_{\vec{w}}(x) dx \\ &\leq C_{\nu_{\vec{w}}} \prod_{j=1}^m \|b_j\|_{BMO} \prod_{j=1}^m \|f_j\|_{L^{p_j}(w_j)}. \end{aligned}$$

#### 4. PROOF OF THEOREM 1.2

We start with a new weak type end-point estimate for  $\mathcal{M}_{L(\log L)}$ .

**Theorem 4.1.** *Let  $\vec{w} \in A_{\vec{1}}$ . Then there exists a constant  $C$  such that*

$$(4.1) \quad \nu_{\vec{w}}(\{x \in \mathbb{R}^n : \mathcal{M}_{L(\log L)}(\mathbf{f})(x) > t^m\}) \leq C \prod_{j=1}^m \left( \int_{\mathbb{R}^n} \Phi^{(m)} \left( \frac{|f_j(x)|}{t} \right) w_j(x) dx \right)^{1/m}.$$

Moreover, this estimate is sharp in the sense that  $\Phi^{(m)}$  can not be replaced by  $\Phi^{(k)}$  for any  $k < m$ .

*Proof.* Our goal is to estimate  $|\Omega| = |\{\mathcal{M}_{L(\log L)}(f_1, f_2, \dots, f_m) > 1\}|$ . The set  $\Omega$  is open and we may assume it to be not empty. It is enough then to control the size of every compact set  $F$  contained in  $\Omega$ .

For  $x \in F$  there exists a cube  $Q$  with  $x \in Q$  such that

$$(4.2) \quad \prod_{j=1}^m \|f_j\|_{\Phi, Q} > 1.$$



Thus, by a covering argument, we can extract a finite family of disjoint cubes  $\{Q_i\}$  whose dilations cover  $F$  for which

$$(4.3) \quad |F| \leq C \sum_i |Q_i|$$

and  $\{Q_i\}$  satisfies

$$\prod_{j=1}^m \|f_j\|_{\Phi, Q_i} > 1.$$

We use again the notation  $C_h^m$  for the family of all subset  $\sigma = (\sigma(1), \dots, \sigma(h))$  extracted from the set of indices  $\{1, \dots, m\}$  using  $1 \leq h \leq m$  different elements. Given  $\sigma \in C_h^m$  and a cube  $Q_i$ , we say that  $i \in B_\sigma$  if  $\|f_{\sigma(k)}\|_{\Phi, Q_i} > 1$  for  $k = 1, \dots, h$  and  $\|f_{\sigma(k)}\|_{\Phi, Q_i} \leq 1$  for  $k = h + 1, \dots, m$ .

Let us consider  $\sigma \in C_h^m$  and  $i \in B_\sigma$ . Denote

$$\Pi_k = \prod_{j=1}^k \|f_{\sigma(j)}\|_{\Phi, Q_i}$$

and  $\Pi_0 = 1$ . Then it is easy to check that  $\Pi_k > 1$  for every  $1 \leq k \leq m$ . It follows that

$$1 < \Pi_k = \|f_{\sigma(k)}\|_{\Phi, Q_i} \Pi_{k-1} = \|f_{\sigma(k)} \Pi_{k-1}\|_{\Phi, Q_i}$$

or, equivalently (by (2.3))

$$(4.4) \quad \frac{1}{|Q_i|} \int_{Q_i} \Phi(f_{\sigma(k)} \Pi_{k-1}) > 1.$$

In particular,

$$(4.5) \quad 1 < \frac{1}{|Q_i|} \int_{Q_i} \Phi(f_{\sigma(m)} \Pi_{m-1}) \leq \frac{1}{|Q_i|} \int_{Q_i} \Phi(f_{\sigma(m)}) \Phi(\Pi_{m-1}).$$

Now, by taking into account the following equivalence

$$\|f\|_{\Phi, Q} \simeq \inf_{\mu > 0} \left\{ \mu + \frac{\mu}{|Q|} \int_Q \Phi(|f|/\mu) \right\},$$

if  $1 \leq j \leq m - h - 1$ , by (4.4) we get

$$\begin{aligned} \Phi^{(j)}(\Pi_{m-j}) &= \Phi^{(j)}(\|f_{\sigma(m-j)} \Pi_{m-j-1}\|_{\Phi, Q_i}) \\ &\leq C \Phi^{(j)} \left( 1 + \frac{1}{|Q_i|} \int_{Q_i} \Phi(f_{\sigma(m-j)} \Pi_{m-j-1}) \right) \\ &\leq C \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(j+1)}(f_{\sigma(m-j)}) \Phi^{(j+1)}(\Pi_{m-j-1}). \end{aligned}$$

From (4.5), by iterating the inequality above, we obtain

$$1 < C \frac{1}{|Q_i|} \int_{Q_i} \Phi(f_{\sigma(m)}) \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(2)}(f_{\sigma(m-1)}) \Phi^{(2)}(\Pi_{m-2})$$

$$\begin{aligned}
&\leq C \left( \prod_{j=0}^{m-h-1} \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(j+1)}(f_{\sigma(m-j)}) \right) \Phi^{(m-h)}(\Pi_h) \\
&\leq C \left( \prod_{j=0}^{m-h-1} \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(j+1)}(f_{\sigma(m-j)}) \right) \left( \prod_{j=1}^h \Phi^{(m-h)}(\|f_{\sigma(j)}\|_{\Phi, Q_i}) \right).
\end{aligned}$$

since  $\Phi$  is submultiplicative.

Thus, since  $i \in B_\sigma$ , we have  $\|f_{\sigma(j)}\|_{\Phi, Q_i} > 1$  for  $j = 1, \dots, h$ , and it follows

$$(4.6) \quad 1 < C \left( \prod_{j=0}^{m-h-1} \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(j+1)}(f_{\sigma(m-j)}) \right) \left( \prod_{j=1}^h \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(m-h+1)}(f_{\sigma(j)}) \right).$$

Now, since for  $1 \leq h \leq m$  and  $0 \leq j \leq m-h-1$  we have that  $\Phi^{(j+1)}(t) \leq \Phi^{(m-h)}(t) \leq \Phi^{(m)}(t)$  and  $\Phi^{(m-h+1)}(t) \leq \Phi^{(m)}(t)$ , we deduce

$$1 < C \prod_{j=1}^m \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(m)}(f_j)$$

or equivalently

$$|Q_i| < C \prod_{j=1}^m \left( \int_{Q_i} \Phi^{(m)}(f_j) \right)^{1/m}.$$

Thus, going back to (4.3) it follows that

$$\begin{aligned}
\nu_{\bar{w}}(F)^m &\approx \left( \sum_i \nu_{\bar{w}}(Q_i) \right)^m \\
&\leq \left( \sum_{h=1}^m \sum_{\sigma \in C_h^m} \sum_{i \in B_\sigma} \nu_{\bar{w}}(Q_i) \right)^m \\
&\leq C \left( \sum_{h=1}^m \sum_{\sigma \in C_h^m} \sum_{i \in B_\sigma} \prod_{j=1}^m \inf_{Q_i} w_j^{1/m} |Q_i|^{1/m} \left( \frac{1}{|Q_i|} \int_{Q_i} \Phi^{(m)}(f_j) \right)^{1/m} \right)^m \\
&\leq C \prod_{j=1}^m \left( \int_{\mathbb{R}^n} \Phi^{(m)}(f_j(y)) w_j(y) dy \right)
\end{aligned}$$

which concludes the proof of (4.1).

We now prove that the estimate (4.1) is sharp in the sense stated in theorem.

We claim that the following estimate is false

$$(4.7) \quad |\{x : \mathcal{M}_{L(\log L)}(\mathbf{f}) > \lambda^m\}| \leq C \left( \prod_{j=1}^m \|\Phi^{(m-1)}\left(\frac{|f_j|}{\lambda}\right)\|_{L^1} \right)^{1/m}$$

We let  $\lambda = 1$  and then the estimate to be studied is

$$(4.8) \quad |\{x : \mathcal{M}_{L(\log L)}(\mathbf{f}) > 1\}|^m \leq C \prod_{j=1}^m \|\Phi^{(m-1)}(|f_j|)\|_{L^1}$$

for any  $\mathbf{f}$  with all the components positives. Hence by the same homogeneity we replace  $f_1$  by  $\frac{f_1}{\lambda^m}$

$$(4.9) \quad |\{x : \mathcal{M}_{L(\log L)}(\mathbf{f}) > \lambda^m\}|^m \leq C \int_{\mathbb{R}} \Phi^{(m-1)}\left(\frac{f_1}{\lambda^m}\right) \prod_{j=2}^m \int_{\mathbb{R}} \Phi^{(m-1)}(f_j)$$

Now, let  $f_j = \chi_{(0,1)}$ . If (4.9) holds, since  $\Phi$  is a Young function, we conclude

$$(4.10) \quad \sup_{\lambda > 0} \frac{1}{\Phi^{m-1}(\lambda^{-m})} |\{x \in \mathbb{R} : \mathcal{M}_{L(\log L)}(\mathbf{f})(x) > \lambda^m\}|^m \leq C.$$

However, observe that, by definition of  $\|\cdot\|_{\Phi, Q}$ , it follows for any subset  $A$  that  $\|\chi_A\|_{\Phi, Q} = \frac{1}{\Phi^{-1}(\frac{|Q|}{|A \cap Q|})}$ . Hence, if  $x > e$  we have

$$\mathcal{M}_{L(\log L)}(\mathbf{f})(x) = \sup_{Q \ni x} \prod_{j=1}^m \|f_j\|_{L(\log L), Q} \geq \|\chi_{(0,1)}\|_{L(\log L), (0,x)}^m = \frac{1}{\Phi^{-1}(x)^m}.$$

Thus, taking into account that  $\Phi^{(k)}(t) \cong t(1 + \log^+ t)^k$ , the left-hand side of (4.10) is bigger than

$$\begin{aligned} \sup_{\lambda > 0} \frac{1}{\Phi^{m-1}(\lambda^{-m})} |\{x > e : \frac{1}{\Phi^{-1}(x)} > \lambda\}|^m &\geq \sup_{0 < \lambda < 1/e} \frac{(\Phi(\frac{1}{\lambda}) - e)^m}{\Phi^{m-1}(\frac{1}{\lambda^m})} \\ &\geq \frac{1}{2^m} \sup_{0 < \lambda < \frac{1}{2e}} \frac{(\Phi(\frac{1}{\lambda}))^m}{\Phi^{m-1}(\frac{1}{\lambda^m})} \\ &\geq \frac{C}{m^{m-1} 2^m} \sup_{0 < \lambda < \frac{1}{2e}} \log \frac{1}{\lambda} \\ &= \infty \end{aligned}$$

□

Given (3.6) and (4.1) the proof of Theorem 1.2 is almost routine. The reader can see [8, pp.38-39] and easily adapt the arguments.

## REFERENCES

- [1] R.R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. **51** (1974), 241–250.
- [2] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. **103** (1976), 611–635.
- [3] C. Fefferman and E.M. Stein,  *$H^p$  spaces of several variables*, Acta Math., **129** (1972), 137–193.
- [4] L. Grafakos, L. Liu, C. Pérez and R.H. Torres, *The multilinear strong maximal function*, submitted.

- [5] L. Grafakos and R.H. Torres, *Multilinear Calderón-Zygmund theory*, Adv. Math. **165** (2002), no. 1, 124–164.
- [6] L. Grafakos and R.H. Torres, *Maximal operator and weighted norm inequalities for multilinear singular integrals*, Indiana Univ. Math. J., **51** (2002), no. 5, 1261–1276.
- [7] L. Grafakos and R.H. Torres, *On multilinear singular integrals of Calderón-Zygmund type*, Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000). Publ. Mat. 2002, Vol. Extra, 57–91.
- [8] A. Lerner, S. Ombrosi, C. Pérez, R. Torres and R. Trujillo-González, *New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory* Adv. Math. **220** (4) (2009), 1222-1264
- [9] C. Pérez, *Endpoint estimates for commutators of singular integral operators*, J. Funct. Anal. **128** (1995), 163–185.
- [10] C. Pérez and G. Pradolini, *Sharp weighted endpoint estimates for commutators of singular integral operators*, Michigan Math. J., **49** (2001), 23–37.
- [11] C. Pérez and R.H. Torres, *Sharp maximal function estimates for multilinear singular integrals*, Contemp. Math., **320** (2003), 323–331.
- [12] C. Pérez and R. Trujillo-González, *Sharp weighted estimates for multilinear commutators*, J. London Math. Soc., **65** (2002), 672–692.
- [13] L. Tang, *Weighted estimates for vector-valued commutators of multilinear operators*, Proc. Roy. Soc. Edinburgh Sect. A **138** (2008), 897–922.
- [14] M. Wilson, *Littlewood-Paley Theory and Exponential-Square Integrability*, Lectures Notes in Math., Springer Verlag, (To appear October 2007).

CARLOS PÉREZ, DEPARTAMENTO DE ANÁLISIS MATEMÁTICO, FACULTAD DE MATEMÁTICAS,  
UNIVERSIDAD DE SEVILLA, 41080 SEVILLA, SPAIN  
*E-mail address:* carlosperez@us.es

G. PRADOLINI, UNIVERSIDAD NACIONAL DEL LITORAL - CONICET, GÜEMES 3450, 3000  
SANTA FE, ARGENTINA  
*E-mail address:* gladis.pradolini@gmail.com

RODOLFO H. TORRES, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF KANSAS, 405 SNOW  
HALL 1460 JAYHAWK BLVD, LAWRENCE, KANSAS 66045-7523, USA  
*E-mail address:* torres@math.ku.edu

RODRIGO TRUJILLO-GONZÁLEZ, DEPARTAMENTO DE ANÁLISIS MATEMÁTICO, UNIVERSIDAD DE  
LA LAGUNA, 38271 LA LAGUNA, S.C. DE TENERIFE, SPAIN  
*E-mail address:* rotrujil@ull.es