Math. Ann. 376 (2020), no. 1-2, 61-102.
DOI 10.1007/s00208-019-01870-z

BOUNDEDNESS RESULTS FOR COMMUTATORS WITH BMO FUNCTIONS VIA WEIGHTED ESTIMATES: A COMPREHENSIVE APPROACH

ÁRPÁD BÉNYI, JOSÉ MARÍA MARTELL, KABE MOEN, ERIC STACHURA, AND RODOLFO H. TORRES

Abstract. We present a unified method to obtain unweighted and weighted estimates of linear and multilinear commutators with *BMO* functions, that is amenable to a plethora of operators and functional settings. Our approach elaborates on a commonly used Cauchy integral trick, recovering many known results but yielding also numerous new ones. In particular, we solve a problem about the boundedness of the commutators of the bilinear Hilbert transform with functions in *BMO*.

Contents

1. Introduction 2
2. Preliminaries 5
2.1. Muckenhoupt weights 5
2.2. The space of functions of bounded mean oscillation 6
2.3. Commutators 7
3. The general theorem 8
3.1. Commutators in the classical setting I 8
3.2. Proof of Theorem 3.1 8
3.3. Commutators in general settings 11
3.4. Commutators in the classical setting II 13
3.5. Commutators vs weighted norm inequalities 15
4. Multilinear operators 16
4.1. Multilinear commutators I 18
4.2. Multilinear commutators II 20
4.3. Multilinear commutators vs weighted estimates 22

Date: June 21, 2019.

2010 *Mathematics Subject Classification.* Primary: 42B20, 47B07; Secondary: 42B25, 47G99.

Key words and phrases. Commutators, spaces of bounded mean oscillations, Muckenhoupt weights, product weights, linearizable operators, linear and multilinear Calderón-Zygmund operators, linear and multilinear fractional integrals, bilinear Hilbert transform, multiparameter Calderón-Zygmund operators, Kato operator.

Bényi and Moen were partially supported by grants from the Simons Foundation No. 246024 and No. 160427, respectively. Martell acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the "Severo Ochoa" Programme for Centres of Excellence in R&D (SEV-2015-0554). He also acknowledges that the research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ ERC agreement no. 615112 HAPDEGMT. Torres was supported in part by NSF grant DMS 1069015.

5. Applications 23
5.1. The Coifman-Rochberg-Weiss result 23
5.2. Fractional integrals 24
5.3. Operators associated with the Kato conjecture 24
5.4. Fractional operators associated with the Kato conjecture 26
5.5. Bilinear Calderón-Zygmund operators 26
5.6. Bilinear fractional integrals 27
5.7. The bilinear Hilbert transform 27
5.8. Multi-parameter operators 28
5.9. Operators on spaces of homogeneous type 31
Appendix A. Multilinear commutators III 31
References 34

1. Introduction

The purpose of this work is to present a panoramic point of view on the topic of commutators of linear, multilinear, and linearizable operators with pointwise multiplication by functions which belong to the John-Nirenberg space of Bounded Mean Oscillation (BMO). We will work in the weighted L^p setting, where the weights will satisfy appropriate Muckenhoupt conditions. We will give a unified method for the weighted estimates of commutators with BMO functions, first in the linear setting, and then in the multilinear setting.

Our approach is based on an idea going back to the pioneer work of Coifman, Rochberg and Weiss [12], who first studied the commutators of Calderón-Zygmund operators with BMO functions. They used what is now sometimes referred to as the *Cauchy integral trick* to obtain boundedness properties of commutators from weighted estimates for the operator itself. The idea is to represent the commutator as the derivative at the origin of an analytic family of operators obtained by conjugating the given operator with complex exponentials. Using the Cauchy integral to represent such derivative, estimates for the commutator follow by Minkowski's inequality and the fact that exponentials of BMO functions are essentially Muckenhoupt weights.

This remarkably simple but elegant approach has been already generalized to other situations. For example, it was substantially extended in the work of Alvarez, Bagby, Kurtz, and Pérez [1], where particular classes of pairs of weights were used, and iterated commutators and some vector valued ones were studied as well. Their work provides a lot of background and motivation for ours. Other applications have been used in the context of spaces of homogeneous type, for example in the work of Hofmann, Mitrea, and Taylor [24]. It is quite remarkable that such a general and *soft* approach to commutator estimates produces, in some cases, sharp weighted estimates in terms of the character of the weights. This was more recently shown by Chung, Pereyra, and Pérez [10]. We will further extend the Cauchy integral trick so it can be applied to many new situations not covered by previous extensions of the method. In particular, we will be able to treat operators which map a Lebesgue space into another one with a different exponent, multilinear operators, and

operators which satisfies weighted estimates but only for limited classes of weights. We will also be able to apply our approach in very general measure theoretic and geometric contexts and multiparameter settings. We will recover many known results and establish several new ones as well. When possible, we have tracked some of the constants involved in our estimates using the existence of sharp bounds to the best of our knowledge. The sharp dependence on the character of the weights in weighted norm inequalities continues to be an area that attracts many investigators and new improvements in some of them have been recently obtained. We want to emphasize, however, that our main objective is to show the generality of our methods and hence we will have preferred clarity in the exposition at the expense of the pursue of best constants, which may be slightly improved in some of our computations. In some cases we have only indicated such possible improvements in a remark leaving the computations to the interested reader.

The main new idea is to replace BMO by its $\exp L$ Orlicz representation denoted by $\mathcal{BMO}$, which allows for a better quantification of some estimates and more flexibility in applications. Of course, in the classical Euclidean setting $BMO = \mathcal{BMO}$, but this may not be the case in more general contexts.

Nowadays, there is a myriad of papers on the topic of weighted norm inequalities for commutators, both linear and multilinear. Although most of the works are based on the same ideas, typically, the arguments have to be reworked every time since the operators or the settings are different. Independently, for these settings and operators it is common to find a well established weighted norm theory. Our results say that, with the exception of some end-point results, the estimates for the commutators follow, almost automatically, from the weighted theory. Thus, in most cases, the ad-hoc arguments for the commutators can be skipped, provided the weighted norm inequalities have been proved. Our main result for linear operators is as follows (we will provide detailed definitions in the next section).

Theorem 1.1. *Let T be a linear operator. Fix $1 < p, q < \infty$, $\theta > 0$, and $1 < s < \infty$. Suppose that there exists an increasing function $\phi : [1, \infty) \longrightarrow [0, \infty)$ such that for all weights w so that $w^\theta \in A_s$, and all $f \in L^p(w^p)$,*

$$\|Tf\|_{L^q(w^q)} \leq \phi\left(\left[w^\theta\right]_{A_s}\right) \|f\|_{L^p(w^p)}. \tag{1.2}$$

Then, for each $b \in \mathcal{BMO}$ and for each weight w and $1 < \eta < \infty$ such that $w^{\theta\eta} \in A_s$,

$$\left\|[T, b]f\right\|_{L^q(w^q)} \leq C(\phi, w, \theta, \eta, s) \|b\|_{\mathcal{BMO}} \|f\|_{L^p(w^p)}; \tag{1.3}$$

and, for every $k \geq 2$,

$$\left\|[T, b]_k f\right\|_{L^q(w^q)} \leq C(\phi, w, \theta, \eta, s, k) \|b\|^k_{\mathcal{BMO}} \|f\|_{L^p(w^p)}. \tag{1.4}$$

Here we follow the approach and notation introduced by Grafakos [22], in particular, $[w]_{A_p}$ denotes the character of the weight (see Section 2.1 below) while ϕ is used to indicate the dependence of the estimate on such quantity. As mentioned earlier, we will sometimes just track such dependence without explicitly specifying the function ϕ. The apparently cumbersome form of the statement of the above theorem allows us to apply it to several families of operators. For example, for classical Calderón-Zygmund operators one simply takes $p = q = \theta = s = 2$. However, other situations require one to work with a power of

a weight in a given class. Consider, for example, the operator $L = -\mathrm{div}(A\nabla)$, where the matrix A is uniformly elliptic, bounded and complex. One can study the Riesz transform, functional calculus, and square function associated with this operator. All these satisfy weighted norm inequalities in $L^p(w^p)$ (or $L^p(u)$ after writing $u = w^p$) for some range of p and for some class of weights which can be written as $w^\theta \in A_s$ with θ, s depending on the index p and the range of the un-weighted estimates. Another familiar situation is the one given by the classical fractional integrals or Riesz potentials. There are natural weighted norm inequalities for some $p < q$ and for weights $w \in A_{p,q}$ or, equivalently, $w^q \in A_{1+q/p'}$ (that is, $w^\theta \in A_s$ for some $\theta > 1$ and $s > 1$). The same kind of behavior occurs for fractional powers of the operators L above. A look at the existing literature (we will give precise reference in later sections) reveals that the commutators of all the operators in the previous families with point-wise multiplication by BMO functions satisfy weighted norm inequalities, but the arguments need to be adapted to each operator in question. Our result says that once the weighted norm theory has been developed, the estimates for commutators follow at once.

Another important feature that we would like to emphasize is that the way the hypotheses are written using the $\mathcal{BMO}$ space is amenable to several immediate generalizations, see Theorem 3.17. For instance, we could replace the basis of cubes in $\mathcal{BMO}$ and the A_p conditions by other families of sets such as dyadic cubes, rectangles, and so on, as well as replace the underlying space $\mathbb{R}^n$ by a space of homogeneous type and more. In general situations, the John-Nirenberg inequality or the self-improvement of the A_p classes – that is, the fact that $w \in A_p$ implies $w^{1+\epsilon} \in A_p$ – may fail (see [15, p. 29]). However, as we will see in the proof of our results, our formulation does not require any of these deep properties. The $\mathcal{BMO}$ space we use here is defined using the exponential norm so that *it carries the John-Nirenberg inequality on its DNA*. Hence, $\mathcal{BMO}$ is in general smaller than the usual BMO space defined using the L^1-averages. Nonetheless, in addition to the classical Euclidean setting, $\mathcal{BMO}$ does coincide with BMO also in some multiparameter setting; see Theorem 5.22 below. Note that the conclusion in Theorem 1.1, is written for weights w such that $w^{\theta\eta} \in A_s$ for some η, which, in the Euclidean setting with the Lebesgue measure and with the Muckenhoupt conditions written in terms of cubes or balls, happens to be true for every $w^\theta \in A_s$ by the fact that A_p-weights satisfy reverse Hölder inequalities. This, however, may not be true for more general measures or geometries. When the reverse Hölder inequality does hold, we can give very precise estimates on the constants appearing in Theorem 1.1; see Theorem 3.22.

We obtain in Theorem 3.31 a partial converse to Theorem 1.1, and multilinear versions in Theorems 4.3 and 4.13. As a corollary we prove the first quantitative results for commutators of multilinear Calderón-Zygmund operators, obtaining estimates in terms of the character of the weights; see Theorem 5.6. A second important corollary is the proof of the boundedness of the commutators of the bilinear Hilbert transform with BMO functions. Although the first results about boundedness of the bilinear Hilbert transform by Lacey and Thiele [31], [32] have been know for some 20 years now, the boundedness of its commutators was a problem that eluded the efforts of several researchers. Our methods provide the result including also some weighted estimates; see Theorems 5.7 and Corollary A.8.

Moreover, combining this result with recent extrapolation ones by Cruz-Uribe and Martell [14], we prove in Corollary 5.8 the boundedness of the commutators in the same range of Lebesgue exponents known for the boundedness of the bilinear Hilbert transform itself.

We collect most of the basic definitions and properties of weights needed in the next section; the expert reader can easily skip this part. The rest of the article is organized as follows: the main results for linear or linearizable operators are presented in Section 3, the general theorems in the multi-linear case are given in Section 4, and a multitude of applications are outlined in Section 5. Finally, Appendix A contains some further results for multilinear operators.

Acknowledgement. Part of this work was carried out while the second and two last named authors were together at the Instituto de Ciencias Matemáticas (ICMAT), Madrid, Spain, as participants of the Research Term on Real Harmonic Analysis and Applications to PDE in 2013. Since then the authors have presented this work at many venues and added several new applications of the methods initially developed. Most recently the last author presented this work at the Mathematical Sciences Research Institute (MSRI), Berkeley, California in February 2017, during the Harmonic Analysis program. The authors would like to thank the ICMAT, the MSRI, and the organizers of the programs mentioned for providing such a great atmosphere to carry out mathematical research. The authors would also like to thank an anonymous referee for some suggestions and corrections.

2. Preliminaries

Let $|E|$ denote the Lebesgue measure of a measurable set $E \subseteq \mathbb{R}^n$. Throughout this paper, we use the notation

$$f_E = \fint_E f\,dx = \frac{1}{|E|}\int_E f(x)\,dx$$

with the understanding that the term is zero if the set E in question has zero or infinite Lebesgue measure. We also use Q to denote cubes with sides parallel to the coordinate axes.

2.1. **Muckenhoupt weights.** By a weight, we mean a non-negative, locally integrable function. We now recall the definition and several properties of Muckenhoupt's A_p classes of weights. For $1 < p < \infty$, a weight w is said to belong to A_p provided

$$[w]_{A_p} := \sup_{Q\subset\mathbb{R}^n}\left(\fint_Q w\,dx\right)\left(\fint_Q w^{1-p'}\,dx\right)^{p-1} < \infty;$$

as usual, p' denotes the conjugate exponent of p. For $p = 1$, w is in A_1 if

$$[w]_{A_1} := \sup_{Q\subset\mathbb{R}^n}\left(\fint_Q w\,dx\right)\|w^{-1}\|_{L^\infty(Q)} < \infty,$$

or, equivalently, $[w]_{A_1} = \|Mw/w\|_{L^\infty(\mathbb{R}^n)}$. Here and elsewhere M denotes the Hardy Littlewood maximal function,

$$Mf(x) = \sup_{Q\ni x}\fint_Q |f(y)|\,dy.$$

We set

$$A_\infty = \bigcup_{p\geq 1} A_p.$$

We say that a weight w belongs to the reverse Hölder class RH_q, $1 < q < \infty$, if

$$[w]_{RH_q} := \sup_{Q\subset\mathbb{R}^n} \left(\fint_Q w^q\,dx\right)^{1/q} \left(\fint_Q w\,dx\right)^{-1} < \infty.$$

When $q = \infty$, we say that $w \in RH_\infty$ if

$$[w]_{RH_\infty} := \sup_{Q\subset\mathbb{R}^n} \|w\|_{L^\infty(Q)} \left(\fint_Q w\,dx\right)^{-1} < \infty.$$

In the case that $q = 1$ we define RH_1 to be all A_∞ weights.

We list a few properties that will be important for us later on. The proofs are standard, and can be found for instance in [21]. For (g), see [28].

Proposition 2.1.

(a) $A_1 \subset A_p \subset A_q$ *for* $1 \leq p \leq q < \infty$ *with* $[w]_{A_1} \geq [w]_{A_p} \geq [w]_{A_q}$.

(b) $RH_\infty \subset RH_q \subset RH_p$ *for* $1 < p \leq q \leq \infty$ *with* $[w]_{RH_\infty} \geq [w]_{RH_q} \geq [w]_{RH_p}$.

(c) *If* $w \in A_p$, $1 < p < \infty$, *then there exists* $1 < q < p$ *such that* $w \in A_q$.

(d) *If* $w \in RH_q$, $1 < q < \infty$, *then there exists* $q < p < \infty$ *such that* $w \in RH_p$.

(e) $A_\infty = \bigcup_{1\leq p<\infty} A_p = \bigcup_{1<q\leq\infty} RH_q$

(f) *If* $1 < p < \infty$, $w \in A_p$ *if and only if* $w^{1-p'} \in A_{p'}$.

(g) *If* $1 \leq q \leq \infty$ *and* $1 \leq s < \infty$, *then* $w \in A_q \cap RH_s$ *if and only if* $w^s \in A_{s(q-1)+1}$.

As we will work in the weighted setting, we need the notion of weighted L^p spaces: $L^p(w) = L^p(\mathbb{R}^n, wdx)$ denotes the collection of measurable functions f on $\mathbb{R}^n$ such that

$$\|f\|_{L^p(w)} := \left(\int_{\mathbb{R}^n} |f(x)|^p w(x)\,dx\right)^{1/p} < \infty.$$

2.2. **The space of functions of bounded mean oscillation.** We recall here the definition of the John-Nirenberg space of functions of bounded mean oscillation, $BMO(\mathbb{R}^n)$, and indicate how it identifies with an appropriate Orlicz space. We say that a locally integrable function $f \in BMO(\mathbb{R}^n)$ if

$$\|f\|_{BMO} := \sup_Q \fint_Q |f - f_Q|\,dx < +\infty;$$

the supremum is taken over the collection of all cubes $Q \subset \mathbb{R}^n$. We will simply write BMO instead of $BMO(\mathbb{R}^n)$ when it is understood that the underlying space on which it is defined is the Euclidean space.

Now, given a cube $Q \subset \mathbb{R}^n$, the localized and normalized $\exp L$-norm is defined as

$$\|f\|_{\exp L,Q} := \inf\left\{\lambda > 0 : \fint_Q \left(e^{\frac{|f(x)|}{\lambda}} - 1\right) dx \le 1\right\}$$

$$= \inf\left\{\lambda > 0 : \fint_Q e^{\frac{|f(x)|}{\lambda}}\, dx \le 2\right\}.$$

Note that,

$$\fint_Q \frac{|f - f_Q|}{\|f - f_Q\|_{\exp L,Q}}\, dx \le \fint_Q \left(e^{\frac{|f-f_Q|}{\|f-f_Q\|_{\exp L,Q}}} - 1\right) dx \le 1$$

or

$$\fint_Q |f - f_Q|\, dx \le \|f - f_Q\|_{\exp L,Q},$$

and hence

$$\|f\|_{BMO} = \sup_Q \fint_Q |f - f_Q|\, dx \le \sup_Q \|f - f_Q\|_{\exp L,Q}. \tag{2.2}$$

On the other hand, the John-Nirenberg inequality says that there exists a dimensional constant C_n such that for every $f \in BMO$ we have

$$\|f\|_{BMO} \le \sup_Q \|f - f_Q\|_{\exp L,Q} \le C_n \|f\|_{BMO}. \tag{2.3}$$

The estimate (2.3) captures the exponential integrability of a BMO function, a fact that is going to be crucial for our results about commutators and weights. Therefore, it is convenient to re-normalize BMO and use the exponential norms as this will make the computations and the constants in our arguments cleaner. Thus, we can define an a priori new Orlicz-type space $\mathcal{BMO}(\mathbb{R}^n)$ (or simply $\mathcal{BMO}$) via the norm

$$\|b\|_{\mathcal{BMO}} := \sup_Q \|b - b_Q\|_{\exp L,Q}.$$

Clearly, (2.3) implies that in the classical Euclidean setting $(BMO, \|\cdot\|_{BMO})$ and $(\mathcal{BMO}, \|\cdot\|_{\mathcal{BMO}})$ are equivalent quasi-normed spaces; nevertheless, the appearance of $\mathcal{BMO}$ in the statements of our results is to emphasize that $\|\cdot\|_{\mathcal{BMO}}$ is used.

2.3. **Commutators.** We recall the notion of commutators with linear operators. Suppose T is a linear operator. Define, whenever it makes sense, the (first order) commutator of T with a measurable function b by

$$T_b^1 f(x) = [T, b]f(x) = T(bf)(x) - b(x)Tf(x) = T\Big(\big(b(\cdot) - b(x)\big)f(\cdot)\Big)(x). \tag{2.4}$$

One can also define the higher order commutators of T with a measurable function b by the recursive formula $T_b^k = [T_b^{k-1}, b]$ for every $k \ge 2$. Instead of T_b^k, the notation $[T, b]_k$ is sometimes used in the literature. It is immediate to see that in such a case,

$$T_b^k f(x) = T\Big(\big(b(\cdot) - b(x)\big)^k f(\cdot)\Big)(x), \qquad k \ge 0,$$

where it is understood that $T_b^0 = T$.

The previous definition can be extended to linearizable operators. We say that a sublinear operator T is linearizable if there exists a Banach space $\mathbb{B}$ and a $\mathbb{B}$-valued linear operator $\mathcal{T}$ such that $Tf(x) = \|\mathcal{T}f(x)\|_{\mathbb{B}}$. In this way we set, for all $k \geq 0$,

$$T_b^k f(x) := \|\mathcal{T}_b^k f(x)\|_{\mathbb{B}} = \left\| \mathcal{T}\Big(\big(b(\cdot) - b(x)\big)^k f(\cdot) \Big)(x), \right\|_{\mathbb{B}} = T\Big(\big(b(\cdot) - b(x)\big)^k f(\cdot) \Big)(x).$$

3. The general theorem

3.1. **Commutators in the classical setting I.** The following result is a generalized version of the classical theorem of Coifman-Rochberg-Weiss [12] which will cover many of the commutators results in the literature. It is a more precise formulation of Theorem 1.1 stated in the introduction. Moreover, the formulation can be further sharpened using stronger properties of the Muckenhoupt weights, but we will do this in Section 3.4. For now, we prefer to state our result in a way that can be generalized to some other situations where the associated Muckenhoupt weights may fail to verify such properties. Recall that we have the identification of spaces $BMO = \mathcal{BMO}$. The operators we consider here act on functions defined on $\mathbb{R}^n$.

Theorem 3.1. *Let T be either a linear or a linearizable operator. Fix $1 \leq p, q < \infty$, $\theta > 0$, and $1 < s < \infty$. Suppose that there exists an increasing function $\phi : [1, \infty) \longrightarrow [0, \infty)$ such that for each $f \in L^p(w^p)$ we have*

$$\|Tf\|_{L^q(w^q)} \leq \phi\left(\left[w^\theta\right]_{A_s} \right) \|f\|_{L^p(w^p)}, \qquad \forall\, w^\theta \in A_s. \tag{3.2}$$

Then, for each $b \in \mathcal{BMO}$, for each weight w and $1 < \eta < \infty$ such that $w^{\theta\eta} \in A_s$ we have

$$\big\|[T, b]f\big\|_{L^q(w^q)} \leq \frac{\eta'\,\theta}{\min\{1, s-1\}}\, \phi\left(4^{\frac{\min\{1,s-1\}}{\eta'}} \left[w^{\theta\eta}\right]_{A_s}^{1/\eta} \right) \|b\|_{\mathcal{BMO}} \|f\|_{L^p(w^p)}; \tag{3.3}$$

and, for every $k \geq 2$,

$$\big\|T_b^k f\big\|_{L^q(w^q)} \leq k! \left(\frac{\eta'\,\theta}{\min\{1, s-1\}} \right)^k \phi\left(4^{\frac{\min\{1,s-1\}}{\eta'}} \left[w^{\theta\eta}\right]_{A_s}^{1/\eta} \right) \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w^p)}. \tag{3.4}$$

3.2. **Proof of Theorem 3.1.** It is well-known that $\mathcal{BMO}$ functions and A_∞ weights are related. Indeed one can show that if $w \in A_\infty$, then $\log w \in \mathcal{BMO}$. The converse is "almost" true: if $b \in \mathcal{BMO}$, then $e^{\lambda b} \in A_\infty$ provided $|\lambda|$ is sufficiently small. A first quantitative version of this result was given by Pereyra in [40, Lemma 3]. We present a similar quantification, computing precisely the parameters involved in terms of $\|b\|_{\mathcal{BMO}}$.

Lemma 3.5. *Let $b \in \mathcal{BMO}$. Then $e^{\lambda b} \in A_{1+|\lambda|\,\|b\|_{\mathcal{BMO}}}$ for every $|\lambda| \leq \|b\|_{\mathcal{BMO}}^{-1}$; moreover,*

$$\left[e^{\lambda b}\right]_{A_{1+|\lambda|\,\|b\|_{\mathcal{BMO}}}} \leq 4^{|\lambda|\,\|b\|_{\mathcal{BMO}}}. \tag{3.6}$$

In particular, for each $1 \leq p < \infty$ and $\lambda \in \mathbb{R}$ verifying

$$|\lambda| \leq \frac{\min\{1, p-1\}}{\|b\|_{\mathcal{BMO}}}, \tag{3.7}$$

we have $e^{\lambda b} \in A_p$ *and*

$$\left[e^{\lambda b}\right]_{A_p} \leq 4^{|\lambda|\,\|b\|_{\mathcal{BMO}}}. \tag{3.8}$$

Proof. By homogeneity considerations we may assume that $\|b\|_{\mathcal{BMO}} = 1$. Then, for every cube Q, we have

$$\fint_Q e^{|b-b_Q|}\,dx \leq 2. \tag{3.9}$$

Now, given $|\lambda| \leq 1$, let $q = 1 + |\lambda|$. We have to prove that $e^{\lambda b} \in A_q$ with $\left[e^{\lambda b}\right]_{A_q} \leq 4^{|\lambda|}$.

The case $\lambda = 0$ is trivial. Assume otherwise that $0 < |\lambda| \leq 1$. By Jensen's inequality and (3.9) we have

$$\fint_Q e^{|\lambda|\,|b-b_Q|}\,dx \leq \left(\fint_Q e^{|b-b_Q|}\,dx\right)^{|\lambda|} \leq 2^{|\lambda|}. \tag{3.10}$$

On the other hand, since $q' - 1 = |\lambda|^{-1}$, (3.9) implies that

$$\fint_Q e^{|\lambda|\,|b-b_Q|\,(q'-1)}\,dx = \fint_Q e^{|b-b_Q|}\,dx \leq 2. \tag{3.11}$$

These estimates and the fact that $q \leq 2$ imply that

$$\begin{aligned}
&\left(\fint_Q e^{\lambda b}\,dx\right)\left(\fint_Q e^{\lambda b\,(1-q')}\,dx\right)^{q-1} \\
&\qquad = \left(\fint_Q e^{\lambda\,(b-b_Q)}\,dx\right)\left(\fint_Q e^{\lambda\,(b-b_Q)(1-q')}\,dx\right)^{q-1} \\
&\qquad \leq \left(\fint_Q e^{|\lambda|\,|b-b_Q|}\,dx\right)\left(\fint_Q e^{|\lambda|\,|b-b_Q|(q'-1)}\,dx\right)^{q-1} \\
&\qquad \leq 2^{|\lambda|}\,2^{q-1} = 4^{|\lambda|}.
\end{aligned}$$

Taking the supremum over all the cubes $Q \subset \mathbb{R}^n$ we conclude that $e^{\lambda b} \in A_q$ with $\left[e^{\lambda b}\right]_{A_q} \leq 4^{|\lambda|}$.

Let us now obtain (3.8). The case $p = 1$ is trivial since $\lambda = 0$. Suppose next that $1 < p \leq 2$; in this case, (3.7) implies that $1 + |\lambda|\,\|b\|_{\mathcal{BMO}} \leq p$ and $|\lambda| \leq \|b\|^{-1}_{\mathcal{BMO}}$. Then, we can apply (3.6) and Proposition 2.1 (*a*) to obtain (3.8). Finally, if $2 < p < \infty$, we observe that (3.7) implies that $|\lambda| \leq \|b\|^{-1}_{\mathcal{BMO}}$ and therefore (3.6) holds. Notice that $1 + |\lambda|\,\|b\|_{\mathcal{BMO}} \leq 2 < p$ and we can get the desired estimate as in the previous case. □

Lemma 3.12. *Fix* $1 < r < \infty$ *and* $1 < \eta < \infty$. *If* $w^\eta \in A_r$ *and* $b \in \mathcal{BMO}$, *then* $w\,e^{\lambda b} \in A_r$ *for every* $\lambda \in \mathbb{R}$ *verifying*

$$|\lambda| \leq \frac{\min\{1, r-1\}}{\eta'\,\|b\|_{\mathcal{BMO}}}. \tag{3.13}$$

Furthermore,

$$\left[w\,e^{\lambda b}\right]_{A_r} \leq [w^\eta]^{\frac{1}{\eta}}_{A_r}\,4^{|\lambda|\,\|b\|_{\mathcal{BMO}}}. \tag{3.14}$$

Proof. Let $U = w\, e^{\lambda b}$. Hölder's inequality yields

$$\begin{aligned}\left(\fint_Q U(x)\,dx\right)\left(\fint_Q U^{1-r'}\,dx\right)^{r-1} &= \left(\fint_Q w\, e^{\lambda b}\,dx\right)\left(\fint_Q w^{1-r'}\, e^{\lambda b\,(1-r')}\,dx\right)^{r-1}\\ &\le \left(\fint_Q w^{\eta}\,dx\right)^{\frac{1}{\eta}}\left(\fint_Q e^{\lambda\,\eta'\,b}\,dx\right)^{\frac{1}{\eta'}}\left(\fint_Q \left(w^{\eta}\right)^{1-r'}\,dx\right)^{\frac{r-1}{\eta}}\left(\fint_Q e^{(1-r')\,\lambda\,\eta'\,b}\,dx\right)^{\frac{r-1}{\eta'}}\\ &\le \left[w^{\eta}\right]_{A_r}^{\frac{1}{\eta}}\left[e^{\lambda\,\eta'\,b}\right]_{A_r}^{\frac{1}{\eta'}} \le \left[w^{\eta}\right]_{A_r}^{\frac{1}{\eta}} 4^{|\lambda|\,\|b\|_{\mathcal{BMO}}},\end{aligned}$$

where we have used that $|\lambda|\,\eta' \le \min\{1, r-1\}/\|b\|_{\mathcal{BMO}}$ which, by Lemma 3.5, implies that $e^{\lambda\,\eta'\,b} \in A_r$ with $\left[e^{\lambda\,\eta'\,b}\right]_{A_r} \le 4^{|\lambda|\,\eta'\,\|b\|_{\mathcal{BMO}}}$. This readily leads to the desired estimate. □

Proof of Theorem 3.1. We start with the case of a linear operator T. Fix $b \in \mathcal{BMO}$ and assume that $\|b\|_{\mathcal{BMO}} = 1$. Fix $1 < \eta < \infty$ and w such that $w^{\eta\,\theta} \in A_s$. Set

$$\delta := \frac{\min\{1, s-1\}}{\theta\,\eta'}.$$

Write, whenever it makes sense, $\Psi(z)f(x) := e^{-z\,b(x)}T(e^{z\,b}\,f)(x)$. Fix $z \in \mathbb{C}$ such that $|z| \le \delta$ and set $W = w\, e^{-\mathrm{Re}\,(z)\,b}$. By Lemma 3.12 we have that

$$\left[W^{\theta}\right]_{A_s} = \left[w^{\theta}\, e^{-\mathrm{Re}\,(z)\,b\,\theta}\right]_{A_s} \le \left[w^{\theta\,\eta}\right]_{A_s}^{\frac{1}{\eta}} 4^{\theta\,|\mathrm{Re}\,(z)|} \le \left[w^{\theta\,\eta}\right]_{A_s}^{\frac{1}{\eta}} 4^{\theta\,\delta}$$

since our choice of δ gives that

$$\theta\,|\mathrm{Re}\,(z)| \le \theta\,\delta = \frac{\min\{1, s-1\}}{\eta'}.$$

Thus, we can apply (3.2) and obtain

$$\|\Psi(z)f\|_{L^q(w^q)} = \|T(e^{z\,b}\,f)\|_{L^q(W^q)} \le \phi\left(\left[W^{\theta}\right]_{A_s}\right)\|e^{z\,b}f\|_{L^p(W^p)},$$

hence

$$\|\Psi(z)f\|_{L^q(w^q)} \le \phi\left(\left[w^{\theta\,\eta}\right]_{A_s}^{\frac{1}{\eta}} 4^{\theta\,\delta}\right)\|f\|_{L^p(w^p)}. \tag{3.15}$$

On the other hand, we notice that

$$\Psi'(z)f(x) = -b(x)\, e^{-z\,b(x)}\, T(e^{z\,b}\,f)(x) + e^{-z\,b(x)}\, T(b\, e^{z\,b}\,f)(x).$$

and by the Cauchy Integral Formula, we have that

$$[T, b]f(x) = \Psi'(0)f(x) = \frac{1}{2\pi i}\int_{|z|=\delta}\frac{\Psi(z)f(x)}{z^2}\,dz.$$

See [1] for a rigorous proof of this commonly used formula.

Next, we use Minkowski's inequality and (3.15) to conclude that

$$\begin{aligned}\|[T, b]f\|_{L^q(w^q)} &\le \frac{1}{2\pi\delta^2}\int_{|z|=\delta}\|\Psi(z)f\|_{L^q(w^q)}|dz|\\ &\le \frac{1}{2\pi\delta^2}\int_{|z|=\delta}\phi\left(\left[w^{\theta\,\eta}\right]_{A_s}^{\frac{1}{\eta}} 4^{\theta\,\delta}\right)\|f\|_{L^p(w^p)}|dz|\end{aligned}$$

$$\leq \frac{1}{\delta}\,\phi\left(\left[w^{\theta\eta}\right]_{A_s}^{\frac{1}{\eta}} 4^{\theta\delta}\right)\|f\|_{L^p(w^p)}.$$

Plugging in the value of the parameter δ in the previous estimate we obtain (3.3).

For T linearizable we proceed in the same way but replacing T by the $\mathbb{B}$-valued linear operator $\mathcal{T}$. In such a case we need to work with the Bochner space $L^q_{\mathbb{B}}(w^q)$ whose associated norm is $\|f\|_{L^q_{\mathbb{B}}(w^q)} = \big\|\|f\|_{\mathbb{B}}\big\|_{L^q(w^q)}$. This is still a Banach space and then in the previous argument it is legitimate to use Minkowski's inequality. The rest of the argument extends *mutatis mutandis* and the details are left to the interested reader.

We now consider the kth order commutator. As before, it suffices to assume that T is linear as the linearizable case follows using the same scheme. The proof is very similar after observing that the Cauchy Integral Formula gives

$$T_b^k f(x) = \Psi^{(k)}(0)f(x) = \frac{k!}{2\pi i}\int_{|z|=\delta}\frac{\Psi(z)f(x)}{z^{k+1}}\,dz.$$

This and the previous argument readily leads to the desired estimate. □

3.3. **Commutators in general settings.** As mentioned before, our main result can be easily extended to more general settings. We can replace the Euclidean setting and the Muckenhoupt classes associated with the basis of cubes by any measure space (we do not even need to have a distance) and basis formed by collection of sets. The commutator result is abstract and needs a few ingredients. In several interesting applications, however, one needs to get back to friendlier settings.

Let $(\mathscr{X}, \Sigma, \mu)$ be a measure space where Σ is a σ-algebra and μ is a non-negative measure. A basis $\mathscr{B} \subset \Sigma$ is a collection of sets such that each $B \in \mathscr{B}$ is μ-measurable and $0 < \mu(B) < \infty$. We can define the A_p classes associated with $\mathscr{B}$. A weight w is a non-negative function such that $\int_B w\,d\mu < \infty$ for every $B \in \mathscr{B}$. We say that a weight $w \in A_{p,\mathscr{B}}$, $1 < p < \infty$, if

$$[w]_{A_{p,\mathscr{B}}} := \sup_{B\in\mathscr{B}}\left(\fint_B w\,d\mu\right)\left(\fint_B w^{1-p'}\,d\mu\right)^{p-1} < \infty.$$

Notice that for our results we do not define the "problematic" class $A_{1,\mathscr{B}}$. As before, for a given weight w, we write $L^p(w) = L^p(w\,d\mu)$. The following properties are trivial and we omit the proof.

Proposition 3.16.

(*a*) $A_{p,\mathscr{B}} \subset A_{q,\mathscr{B}}$ *for* $1 < p \leq q < \infty$ *with* $[w]_{A_{p,\mathscr{B}}} \geq [w]_{A_{q,\mathscr{B}}}$.

(*b*) *If* $1 < p < \infty$, $w \in A_{p,\mathscr{B}}$ *if and only if* $w^{1-p'} \in A_{p',\mathscr{B}}$.

We can also define a $\mathcal{BMO}$ space associated with the family $\mathscr{B}$. We say that a measurable function $f \in \mathcal{BMO}_{\mathscr{B}}(\mathscr{X})$ or simply $f \in \mathcal{BMO}_{\mathscr{B}}$ (when the underlying space $\mathscr{X}$ is clear from the context) provided $\int_B |f|\,d\mu < \infty$ for every $B \in \mathscr{B}$ and

$$\|f\|_{\mathcal{BMO}_{\mathscr{B}}} := \sup_{B\in\mathscr{B}}\|f - f_B\|_{\exp L,B} < \infty.$$

Clearly, when $\mathscr{X} = \mathbb{R}^n$, $\mathscr{B}$ denotes the basis of cubes and $d\mu = dx$, we have $\mathcal{BMO}_{\mathscr{B}} = \mathcal{BMO} = BMO$.

The notation f_B is, as before, for the μ-average of f on B, while the norm $\|f\|_{\exp L,B}$ is defined as

$$\inf\left\{\lambda>0:\fint_B\left(e^{\frac{|f|}{\lambda}}-1\right)d\mu\le 1\right\}=\inf\left\{\lambda>0:\fint_B e^{\frac{|f|}{\lambda}}\,d\mu\le 2\right\}.$$

This definition of $\mathcal{BMO}_{\mathscr{B}}$ carries the John-Nirenberg inequality as part of the definition and therefore it is what we need for our purposes. One could also define a $BMO_{\mathscr{B}}$ space using L^1 averages. In general, we have

$$\mathcal{BMO}_{\mathscr{B}}\subset BMO_{\mathscr{B}},$$

and, unless some further properties are assumed, one cannot expect these two spaces to coincide.

We have the following commutator result.

Theorem 3.17. *Let T be either a linear or a linearizable operator. Fix $1\le p,q<\infty$, $\theta>0$, and $1<s<\infty$. Suppose that there exists an increasing function $\phi:[1,\infty)\longrightarrow[0,\infty)$ such that for each $f\in L^p(w^p)$ we have*

$$\|Tf\|_{L^q(w^q)}\le\phi\left(\left[w^\theta\right]_{A_{s,\mathscr{B}}}\right)\|f\|_{L^p(w^p)},\qquad\forall\, w^\theta\in A_{s,\mathscr{B}}. \tag{3.18}$$

Then, for each $b\in\mathcal{BMO}_{\mathscr{B}}$, for each weight w and $1<\eta<\infty$ such that $w^{\theta\eta}\in A_{s,\mathscr{B}}$ we have

$$\left\|[T,b]f\right\|_{L^q(w^q)}\le\frac{\eta'\,\theta}{\min\{1,s-1\}}\,\phi\left(4^{\frac{\min\{1,s-1\}}{\eta'}}\left[w^{\theta\eta}\right]^{1/\eta}_{A_{s,\mathscr{B}}}\right)\|b\|_{\mathcal{BMO}_{\mathscr{B}}}\|f\|_{L^p(w^p)}; \tag{3.19}$$

and, for every $k\ge 2$,

$$\left\|T_b^k f\right\|_{L^q(w^q)}\le k!\left(\frac{\eta'\,\theta}{\min\{1,s-1\}}\right)^k\phi\left(4^{\frac{\min\{1,s-1\}}{\eta'}}\left[w^{\theta\eta}\right]^{1/\eta}_{A_{s,\mathscr{B}}}\right)\|b\|^k_{\mathcal{BMO}_{\mathscr{B}}}\|f\|_{L^p(w^p)}. \tag{3.20}$$

The proof is exactly as before with the only exception that in Lemma 3.5 when $\lambda=0$ one has that $w\in A_{q,\mathscr{B}}$ for every $q>1$ (recall that we have not defined $A_{1,\mathscr{B}}$) and also that, in the second part of that result, the case $p=1$ is excluded. Taking all this into account, one can follow the proof of Theorem 3.1 line by line and the argument goes through with only notational changes.

Remark 3.21. If one forgets about the dependence on the constants, Theorem 3.17 can be re-stated in a qualitative way. Such a result will say that if T is bounded from $L^p(w^p)$ to $L^q(w^q)$ for every w such that $w^\theta\in A_{s,\mathscr{B}}$ then the commutators T_b^k with $b\in\mathcal{BMO}_{\mathscr{B}}$ are bounded from $L^p(w^p)$ to $L^q(w^q)$ for every w such that $w^{\theta\eta}\in A_{s,\mathscr{B}}$ for some $\eta>1$.

Examples of bases $\mathscr{B}$ in $\mathbb{R}^n$ with the Lebesgue measure are the sets of all cubes, or balls, or dyadic cubes or rectangles with sides parallel to the axes. In spaces of homogeneous type, one can also take balls, or Christ's dyadic cubes. All these classes have some additional features that allow one to refine the statement of Theorem 3.17. We will discuss this in the following section.

3.4. **Commutators in the classical setting II.** Our next goal is to sharpen Theorem 3.1 and exploit the fact that the Muckenhoupt weights have self-improving properties. We first notice that in that result the weighted estimates for T are assumed to hold for every weight w in the class $w^\theta \in A_s$. However, in the conclusion, the weighted estimates for the commutators hold provided $w^{\theta\eta} \in A_s$, where $\eta > 1$, which is apparently stronger than $w^\theta \in A_s$. Nevertheless, the self-improvement of the weights gives that $w^\theta \in A_s$ which implies $w^{\theta\eta} \in A_s$ for some $\eta > 1$ which is typically very close to 1. This value of η depends on s and $\left[w^\theta\right]_{A_s}$ and has been precisely quantified in [11]. Let us notice again that this property depends very much on the setting. In the abstract context of the previous section without any further assumptions there is no reason to think that $w^\theta \in A_s$ implies $w^{\theta\eta} \in A_s$.

Theorem 3.22. *Let T be either a linear or a linearizable operator. Fix $1 \le p, q < \infty$, $\theta > 0$, and $1 < s < \infty$. Suppose that there exist an increasing function $\phi : [1, \infty) \longrightarrow [0, \infty)$ such that for each $f \in L^p(w^p)$ we have*

$$\|Tf\|_{L^q(w^q)} \le \phi\left(\left[w^\theta\right]_{A_s}\right) \|f\|_{L^p(w^p)}, \qquad \forall\, w^\theta \in A_s. \tag{3.23}$$

Then, for each $b \in \mathcal{BMO}$ and for every $w^\theta \in A_s$, we have

$$\begin{aligned}\left\|[T,b]f\right\|_{L^q(w^q)} \le{}& \frac{2^{2\max\{s,s'\}+n+2}\,\theta}{\min\{1, s-1\}} \left[w^\theta\right]_{A_s}^{\max\{1,\frac{1}{s-1}\}} \\ & \phi\left(4^{\min\{1,s-1\}}\, 2^s \left[w^\theta\right]_{A_s}\right) \|b\|_{\mathcal{BMO}} \|f\|_{L^p(w^p)};\end{aligned} \tag{3.24}$$

and, for every $k \ge 2$,

$$\begin{aligned}\left\|T_b^k f\right\|_{L^q(w^q)} \le{}& k! \left(\frac{2^{2\max\{s,s'\}+n+2}\,\theta}{\min\{1, s-1\}}\right)^k \left[w^\theta\right]_{A_s}^{k\max\{1,\frac{1}{s-1}\}} \\ & \phi\left(4^{\min\{1,s-1\}}\, 2^s \left[w^\theta\right]_{A_s}\right) \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w^p)}.\end{aligned} \tag{3.25}$$

Before proving this result, we need the following auxiliary lemma due to Pérez [11].

Lemma 3.26. *Let $w \in A_p$, $1 < p < \infty$, and let $\rho_w = 1 + \frac{1}{2^{2p+n+1}\,[w]_{A_p}}$. Then, for any cube $Q \subset \mathbb{R}^n$, we have*

$$\left(\fint_Q w^{\rho_w}\, dx\right)^{\frac{1}{\rho_w}} \le 2 \fint_Q w\, dx. \tag{3.27}$$

Remark 3.28. It was shown in [27] and [26] that one can get a sharper reverse Hölder exponent using the the Fujii-Wilson A_∞-constant. More precisely, (3.27) holds for any $w \in A_\infty$ with $\rho_w = 1 + \frac{1}{2^{n+11}\,[w]_{A_\infty}}$, where $[w]_{A_\infty}$ stands for the Fujii-Wilson A_∞-constant defined as

$$[w]_{A_\infty} := \sup_{Q \subset \mathbb{R}^n} \frac{1}{w(Q)} \int_Q M(w\chi_Q)\, dx.$$

Proof of Theorem 3.22. Fix $u = w^\theta \in A_s$. Write $v = u^{1-s'} = w^{\theta(1-s')} \in A_{s'}$. Apply Lemma 3.26 to u and v to obtain that

$$\left(\fint_Q u^{\rho_u}\,dx\right)^{\frac{1}{\rho_u}} \le 2\fint_Q u\,dx, \qquad \left(\fint_Q v^{\rho_v}\,dx\right)^{\frac{1}{\rho_v}} \le 2\fint_Q v\,dx, \tag{3.29}$$

where

$$\rho_u = 1 + \frac{1}{2^{2\,s+n+1}\,[u]_{A_s}}, \qquad \rho_v = 1 + \frac{1}{2^{2\,s'+n+1}\,[v]_{A_{s'}}} = 1 + \frac{1}{2^{2\,s'+n+1}\,[u]_{A_s}^{s'-1}}.$$

Set $\eta = \min\{\rho_u, \rho_v\}$. We see that $u^\eta \in A_s$ by (3.29) and our choice of η:

$$\begin{aligned}
\left[u^\eta\right]_{A_s}^{\frac{1}{\eta}} &= \sup_Q \left(\fint_Q u^\eta\,dx\right)^{\frac{1}{\eta}} \left(\fint_Q (u^\eta)^{1-s'}\,dx\right)^{\frac{s-1}{\eta}} \\
&\le \sup_Q \left(\fint_Q u^{\rho_u}\,dx\right)^{\frac{1}{\rho_u}} \left(\fint_Q v^{\rho_v}\,dx\right)^{\frac{s-1}{\rho_v}} \\
&\le \sup_Q 2^s \left(\fint_Q u\,dx\right)\left(\fint_Q v\,dx\right)^{(s-1)} = 2^s\,[u]_{A_s}.
\end{aligned}$$

We now estimate η'. We first observe that

$$\eta' = \max\{\rho_u', \rho_v'\} = 1 + \max\left\{2^{2\,s+n+1}\,[u]_{A_s}, 2^{2\,s'+n+1}\,[u]_{A_s}^{s'-1}\right\}.$$

If $1 < s \le 2$ we have that $2^{2\,s+n+1}\,[u]_{A_s} \le 2^{2\,s'+n+1}\,[u]_{A_s}^{s'-1}$ since $1 \le s'-1$ and $[u]_{A_s} \ge 1$. On the other hand, if $2 \le s < \infty$ we analogously observe that $2^{2\,s'+n+1}\,[u]_{A_s}^{s'-1} \le 2^{2\,s+n+1}\,[u]_{A_s}$. Thus,

$$\eta' = 1 + 2^{2\,\max\{s,s'\}+n+1}\,[u]_{A_s}^{\max\{1,s'-1\}} = 1 + 2^{2\,\max\{s,s'\}+n+1}\,[u]_{A_s}^{\max\{1,\frac{1}{s-1}\}}.$$

By (3.3) in Theorem 3.1, the previous estimates, and the fact that $\eta' \ge 1$ we obtain

$$\begin{aligned}
\left\|[T,b]f\right\|_{L^q(w^q)} &\le \frac{\eta'\,\theta}{\min\{1,s-1\}}\,\phi\left(4^{\frac{\min\{1,s-1\}}{\eta'}}\left[w^{\theta\eta}\right]_{A_s}^{1/\eta}\right)\|b\|_{\mathcal{BMO}}\,\|f\|_{L^p(w^p)} \\
&\le \frac{\theta}{\min\{1,s-1\}}\left(1 + 2^{2\,\max\{s,s'\}+n+1}\,[u]_{A_s}^{\max\{1,\frac{1}{s-1}\}}\right) \\
&\qquad \phi\left(4^{\min\{1,s-1\}}\,2^s\left[w^\theta\right]_{A_s}\right)\|b\|_{\mathcal{BMO}}\,\|f\|_{L^p(w^p)} \\
&\le \frac{2^{2\,\max\{s,s'\}+n+2}\,\theta}{\min\{1,s-1\}}\,[u]_{A_s}^{\max\{1,\frac{1}{s-1}\}} \\
&\qquad \phi\left(4^{\min\{1,s-1\}}\,2^s\left[w^\theta\right]_{A_s}\right)\|b\|_{\mathcal{BMO}}\,\|f\|_{L^p(w^p)}.
\end{aligned}$$

The proof of (3.25) is analogous and we leave it to the interested reader. □

Remark 3.30. By using the results from [27] and [26], see Remark 3.28, the estimate (3.24) can be improved to the following:

$$\left\|[T,b]f\right\|_{L^q(w^q)} \le \frac{2^{n+12}\,\theta}{\min\{1,s-1\}}\ \max\left\{[w^\theta]_{A_\infty}, [w^{\theta(1-s')}]_{A_\infty}\right\}$$

$$\phi\left(4^{\min\{1,s-1\}}\,2^s\,\left[w^\theta\right]_{A_s}\right)\,\|b\|_{\mathcal{BMO}}\,\|f\|_{L^p(w^p)}.$$

Further details are left to the interested reader.

3.5. **Commutators vs weighted norm inequalities.** In Theorem 3.1 we proved that if a given linear operators T satisfies weighted norm inequalities, then so do the corresponding commutators with $\mathcal{BMO}$ functions. In many situations, the theories of weighted norm inequalities and the estimates for the commutators run parallel without apparent connection. Theorem 3.1 gives a very powerful one-way bridge between the two theories. It is therefore natural to wonder whether there is in fact a two-way bridge between the two theories, that is, also an abstract principle that allows one to say that estimates for the commutators imply certain norm inequalities.

Suppose that a given operator T has the property that every commutator T_b^k is bounded on $L^2(\mathbb{R}^n)$ with an appropriate control on the constants. We wonder whether this implies that T is bounded on $L^2(w)$ for some collection of weights. The best that one could expect the class of weights to be is A_2 (or some A_p), since we know that for the Riesz or Hilbert transforms the condition A_2 is necessary and sufficient. However, this class is too big. In fact, there are examples of operators whose commutators behave as the ones of the Riesz or Hilbert transforms for which the class of weights cannot be A_2. For example, consider the Riesz transform $\nabla L^{-1/2}$ where $Lf = -\mathrm{div}(A\nabla f)$ is a second order divergence form elliptic operator with bounded and complex matrix A. All commutators of L are bounded on $L^2(\mathbb{R}^n)$ with the appropriate constants; see [4] and Section 5 below. However, if $\nabla L^{-1/2}$ were bounded on $L^2(w)$ for every $w \in A_2$, then the Rubio de Francia extrapolation theorem would imply that $\nabla L^{-1/2}$ is bounded on $L^p(\mathbb{R}^n)$ for every $1 < p < \infty$ and this is not true for all such operators L; see [2]. This shows that the class of weights considered for such a converse should be small enough so that extrapolation does not lead to invalid estimates.

Theorem 3.31. *Fix $1 < p < \infty$ and suppose that T is a linear operator.*

(*a*) *Suppose there is a $\lambda_0 > 0$ and $\phi : [0, \lambda_0] \longrightarrow [0, \infty)$ an increasing function such that for each $b \in \mathcal{BMO}$ we have*

$$\|Tf\|_{L^p(e^{\lambda b})} \le \phi\big(|\lambda|\,\|b\|_{\mathcal{BMO}}\big)\,\|f\|_{L^p(e^{\lambda b})}, \qquad \forall\,\lambda,\ |\lambda| \le \frac{\lambda_0}{\|b\|_{\mathcal{BMO}}}. \tag{3.32}$$

Then, for each $k \ge 0$ and $b \in \mathcal{BMO}$, one has

$$\|T_b^k f\|_{L^p(\mathbb{R}^n)} \le \phi(\lambda_0)\,k!\,\left(\frac{p}{\lambda_0}\right)^k\,\|b\|_{\mathcal{BMO}}^k\,\|f\|_{L^p(\mathbb{R}^n)}.$$

(*b*) *Conversely, suppose that there exists $\lambda_0, C_0 > 0$ such that, for every $b \in \mathcal{BMO}$ and for every $k \ge 0$, one has*

$$\|T_b^k f\|_{L^p(\mathbb{R}^n)} \le C_0\,k!\,\left(\frac{p}{\lambda_0}\right)^k\,\|b\|_{\mathcal{BMO}}^k\,\|f\|_{L^p(\mathbb{R}^n)}. \tag{3.33}$$

Then, there exists an increasing function $\phi : [0, \lambda_0) \longrightarrow [0, \infty)$ *such that for every* $b \in \mathcal{BMO}$ *we have*

$$\|Tf\|_{L^p(e^{\lambda b})} \leq \phi\big(|\lambda|\,\|b\|_{\mathcal{BMO}}\big)\,\|f\|_{L^p(e^{\lambda b})}, \qquad \forall\,\lambda,\ |\lambda| < \frac{\lambda_0}{\|b\|_{\mathcal{BMO}}}. \tag{3.34}$$

Proof. We shall first prove (a). The case $k = 0$ follows from (3.32) with $\lambda = 0$. Let $k \geq 1$. By homogeneity we may assume that $\|b\|_{\mathcal{BMO}} = 1$. We set $\delta = \lambda_0/p$. Write, as before $\Psi(z)f(x) := e^{-z\,b(x)}T(e^{z\,b}\,f)(x)$. The Cauchy Integral Formula gives

$$T_b^k f(x) = \Psi^{(k)}(0)f(x) = \frac{k!}{2\pi i}\int_{|z|=\delta}\frac{\Psi(z)f(x)}{z^{k+1}}\,dz.$$

For every $|z| = \delta$ we note that $p\,|\mathrm{Re}\,(z)| \leq p\,\delta \leq \lambda_0$. Thus we can apply (3.32) with $e^{-p\,\mathrm{Re}\,(z)\,b}$ and write

$$\begin{aligned}\|T_b^k f\|_{L^p(\mathbb{R}^n)} \leq \frac{k!}{2\,\pi\,\delta^{k+1}}\int_{|z|=\delta}\|\Psi(z)f\|_{L^p(\mathbb{R}^n)}|dz| &\leq \frac{k!}{\delta^k}\sup_{|z|=\delta}\|T(e^{z\,b}\,f)\|_{L^p(e^{-p\,\mathrm{Re}\,(z)\,b})}\\ &\leq \frac{k!}{\delta^k}\,\|f\|_{L^p(\mathbb{R}^n)}\sup_{|z|=\delta}\phi\big(p\,|\mathrm{Re}\,(z)|\big) \leq \frac{k!}{\delta^k}\,\phi(\lambda_0),\end{aligned}$$

which is the desired estimate.

Let us now turn to proving (b). By homogeneity we may assume again that $\|b\|_{\mathcal{BMO}} = 1$. Define $\Psi(z)$ as before and and fix x. Since $\Psi(z)f(x)$ is holomorphic, we can write

$$|\Psi(z)f(x)| = \left|\sum_{k=0}^{\infty}\frac{\Psi^{(k)}(0)f(x)}{k!}\,z^k\right| \leq \sum_{k=0}^{\infty}\frac{|T_b^k f(x)|}{k!}\,|z|^k.$$

We now use (3.33) to obtain that if $|z| < \lambda_0/p$ then

$$\|\Psi(z)f\|_{L^p(\mathbb{R}^n)} \leq \sum_{k=0}^{\infty}\frac{1}{k!}\,|z|^k\,\|T_b^k f\|_{L^p(\mathbb{R}^n)} \leq C_0\,\|f\|_{L^p(\mathbb{R}^n)}\sum_{k=0}^{\infty}|z|^k\,\left(\frac{p}{\lambda_0}\right)^k = \phi(|z|\,p)\,\|f\|_{L^p(\mathbb{R}^n)}$$

where $\phi : [0, \lambda_0) \longrightarrow [0, \infty)$ is the increasing function defined by the formula $\phi(s) = C_0\,(1 - s/\lambda_0)^{-1}$, $0 \leq s < \lambda_0$. Let us now take $z = -\lambda/p$ with $|\lambda| < \lambda_0$ (which implies that $|z| < \lambda_0/p$) and $f(x) = g(x)\,e^{\lambda\,b(x)/p}$. The previous estimate yields

$$\|Tg\|_{L^p(e^{\lambda\,b})} = \|\Psi(-\lambda/p)f\|_{L^p(\mathbb{R}^n)} \leq \phi(|\lambda|)\,\|f\|_{L^p(\mathbb{R}^n)} = \phi(|\lambda|)\,\|g\|_{L^p(e^{\lambda\,b})}.$$

This completes the proof. □

4. Multilinear operators

We turn now our attention to multilinear operators with the main goal of obtaining the natural counterparts to the general results from the linear setting, Theorem 3.1 and Theorem 3.22. The situation here changes slightly since the good classes of weights come in two flavors: product weights and vector weights.

Let us thus begin by defining the main objects that we will be dealing with in this setting. T will always denote an m-linear operator from $X_1 \times \cdots \times X_m$ into Y, where X_j, $1 \leq j \leq m$, and Y are some normed spaces. In our statements the X_j and Y will be again appropriate

weighted Lebesgue spaces. For $\mathbf{f} = (f_1, f_2, \dots, f_m) \in X_1 \times X_2 \times \cdots \times X_m$ and for a measurable vector $\mathbf{b} = (b_1, b_2, \dots, b_m)$, and $1 \le j \le m$, we define, whenever it makes sense, the (first order) commutators

$$[T, \mathbf{b}]_{e_j}(\mathbf{f}) = T(f_1, \dots, b_j f_j, \dots f_m) - b_j T(f_1, \dots, f_j, \dots f_m);$$

we denoted by e_j the basis element taking the value 1 at component j and 0 in every other component, therefore expressing the fact that the commutator acts as a linear one in the jth variable and leaving the rest of the entries of $\mathbf{f}$ untouched. Then, if $k \in \mathbb{N}$, we define

$$[T, \mathbf{b}]_{ke_j} = [\cdots [[T, \mathbf{b}]_{e_j}, \mathbf{b}]_{e_j} \cdots, \mathbf{b}]_{e_j},$$

where the commutator is performed k times. Finally, if $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m) \in (\mathbb{N} \cup \{0\})^m$ is a multi-index, we define

$$[T, \mathbf{b}]_\alpha = [\cdots [[T, \mathbf{b}]_{\alpha_1 e_1}, \mathbf{b}]_{\alpha_2 e_2} \cdots, \mathbf{b}]_{\alpha_m e_m}. \tag{4.1}$$

Informally, if the multilinear operator T has a kernel representation of the form

$$T(\mathbf{f})(x) = \int_{\mathbb{R}^{nm}} K(x, y_1, \dots, y_m) f_1(y_1) \cdots f_m(y_m) dy_1 \dots dy_m,$$

then $[T, \mathbf{b}]_\alpha(\mathbf{f})(x)$ can be expressed in a similar way, with kernel

$$\prod_{j=1}^m (b_j(y_j) - b_j(x))^{\alpha_j} K(x, y_1, \dots, y_m).$$

Next, we define the appropriate vector-weights classes of Lerner, Ombrosi, Pérez, Torres and Trujillo-González [33].

Definition 4.2. Let $\mathbf{P} = (p_1, \dots, p_m)$ with $1 < p_1, \dots, p_m < \infty$ satisfying

$$\frac{1}{p_1} + \cdots + \frac{1}{p_m} = \frac{1}{p}.$$

Let $\mathbf{w} = (w_1, \dots, w_m)$ and set

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

We say that the vector weight $\mathbf{w} \in A_{\mathbf{P}}$ if

$$[\mathbf{w}]_{A_{\mathbf{P}}} := \sup_Q \left(\fint_Q \nu_w \, dx \right)^{1/p} \prod_{j=1}^m \left(\fint_Q w_j^{1-p_j'} \, dx \right)^{1/p_j'} < \infty.$$

We observe that, by Hölder's inequality, we have

$$\left(\fint_Q \nu_{\mathbf{w}} \, dx \right)^{1/p} \prod_{j=1}^m \left(\fint_Q w_j^{1-p_j'} \, dx \right)^{1/p_j'} \le \prod_{j=1}^m \left(\fint_Q w_j \, dx \right)^{1/p_j} \left(\fint_Q w_j^{1-p_j'} \, dx \right)^{1/p_j'}.$$

and thus

$$\prod_{j=1}^m A_{p_j} \subset A_{\mathbf{P}}.$$

It was shown in [33] that the above inclusion is strict.

4.1. **Multilinear commutators I.** We are now able to give a multilinear version of Theorem 3.1 where each weight belongs, individually, to a Muckenhoupt class.

Theorem 4.3. *Let T be an m-linear operator. Fix $1 < p_j < \infty$, $1 < s_j < \infty$, and $\theta_j > 0$, $1 \le j \le m$. Let $1 < p < \infty$ be so that*

$$\frac{1}{p} = \sum_{j=1}^{m} \frac{1}{p_j},$$

and suppose that there exist increasing functions $\phi_j : [1,\infty) \to [0,\infty)$ such that for all $\boldsymbol{w} = (w_1, \dots, w_m)$ satisfying $w_j^{\theta_j} \in A_{s_j}$, $1 \le j \le m$, we have

$$\|T\boldsymbol{f}\|_{L^p(\prod_{j=1}^m w_j^p)} \le \prod_{j=1}^{m} \phi_j\left([w_j^{\theta_j}]_{A_{s_j}}\right) \|f_j\|_{L^{p_j}\left(w_j^{p_j}\right)}. \tag{4.4}$$

Then, for all $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$, for each multi-index α and vector weight $\boldsymbol{w} = (w_1, \dots, w_m)$ such that for all $1 \le j \le m$, $w_j^{\eta_j\theta_j} \in A_{s_j}$ with some $1 < \eta_j < \infty$, we have

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\prod_{j=1}^m w_j^p)} \le \alpha! \prod_{j=1}^{m} \frac{1}{\delta_j^{\alpha_j}} \phi_j\left(4^{\theta_j\delta_j}[w_j^{\theta_j\eta}]_{A_{s_j}}^{1/\eta}\right) \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}\left(w_j^{p_j}\right)}, \tag{4.5}$$

where $\delta_j = \frac{\min\{1, s_j - 1\}}{\eta_j' \theta_j}$, $1 \le j \le m$.

Proof. We adapt the proof of Theorem 3.1 to the multilinear setting. Let then $\mathbf{b} \in \mathcal{BMO}^m$, and without loss of generality assume that they are real valued and $\|b_j\|_{\mathcal{BMO}} = 1$, $1 \le j \le m$.

Define, as before, for $z = (z_1, \dots, z_n)$, the following holomorphic map

$$\Psi(z)(\mathbf{f}) = e^{-\sum_{j=1}^m z_j b_j} T\left(e^{z_1 b_1} f_1, e^{z_2 b_2} f_2, \dots, e^{z_m b_m} f_m\right).$$

The Cauchy integral formula adapted to several complex variables now yields

$$[T, \mathbf{b}]_\alpha \mathbf{f} = D^\alpha \Psi_z(\mathbf{f})\big|_{z=0} = \frac{\alpha!}{(2\pi i)^m} \int_{b_0 P(0,\vec{\delta})} \frac{\Psi_z(\mathbf{f})}{z_1^{\alpha_1+1} z_2^{\alpha_2+1} \cdots z_m^{\alpha_m+1}} \, dz_1 dz_2 \cdots dz_m, \tag{4.6}$$

where we denoted by $b_0 P(0, \vec{\delta}) = \{(z_1, z_2, \dots, z_m) \in \mathbb{C}^m : |z_j| = \delta_j, 1 \le j \le m\}$ the so-called distinguished boundary of the polydisc $P(0, \vec{\delta}) = \{(z_1, z_2, \dots, z_m) \in \mathbb{C}^m : |z_j| < \delta_j, 1 \le j \le m\}$.

Fix now the weights w_j so that $w_j^{\theta_j\eta_j} \in A_{s_j}$ for $1 \le j \le m$ and let us define $v_j = w_j e^{-\mathrm{Re}(z_j) b_j}$. Note that

$$v_j^{\theta_j} = w_j^{\theta_j} e^{-\mathrm{Re}(z_j)\theta_j b_j}.$$

Since $w_j^{\theta_j\eta_j} \in A_{s_j}$ and

$$|\mathrm{Re}(z_j)\theta_j| \le \delta_j \theta_j \le \frac{\min\{1, s_j - 1\}}{\eta_j'},$$

by Lemma 3.12 we obtain that $v_j^{\theta_j} \in A_{s_j}$ and

$$[v_j^{\theta_j}]_{A_{s_j}} \le [w_j^{\theta_j\eta_j}]_{A_{s_j}}^{1/\eta_j} 4^{\delta_j\theta_j}.$$

By our hypothesis condition (4.4), we can write

$$\begin{aligned}\|\Psi_z \mathbf{f}\|_{L^p(\prod_{j=1}^m w_j^p)} &= \|T(e^{z_1 b_1} f_1, \dots, e^{z_m b_m} f_m)\|_{L^p(\prod_{j=1}^m v_j^p)} \\ &\le \prod_{j=1}^m \phi_j \left([v_j^{\theta_j}]_{A_{s_j}}\right) \|e^{z_j b_j} f_j\|_{L^{p_j}\left(v_j^{p_j}\right)},\end{aligned}$$

which, by the increasing property of the functions ϕ_j and the fact that

$$\|e^{z_j b_j} f_j\|_{L^{p_j}\left(v_j^{p_j}\right)} = \|f_j\|_{L^{p_j}\left(w_j^{p_j}\right)}$$

further yields the uniform estimate over $b_0 P(0, \vec{\delta})$:

$$\|\Psi_z \mathbf{f}\|_{L^p(\prod_{j=1}^m w_j^p)} \le \prod_{j=1}^m \phi_j \left(4^{\theta_j \delta_j} [w_j^{\theta_j \eta_j}]_{A_{s_j}}^{1/\eta_j}\right) \|f_j\|_{L^{p_j}\left(w_j^{p_j}\right)}.$$

Now, by Minkowski's inequality, we obtain that

$$\begin{aligned}\|[T, \mathbf{b}]_\alpha \mathbf{f}\|_{L^p(\prod_{j=1}^m w_j^p)} &\le \frac{\alpha!}{(2\pi)^m} \int_{b_0 P(0,\vec{\delta})} \frac{\|\Psi_z \mathbf{f}\|_{L^p(\prod_{j=1}^m w_j^p)}}{\prod_{j=1}^m |z_j|^{\alpha_j+1}} |dz_1| \cdots |dz_m| \\ &\le \frac{\alpha!}{(2\pi)^m} \prod_{j=1}^m \frac{2\pi\delta_j}{\delta_j^{\alpha_j+1}} \phi_j \left(4^{\theta_j \delta_j} [w_j^{\theta_j \eta_j}]_{A_{s_j}}^{1/\eta}\right) \|f_j\|_{L^{p_j}\left(w_j^{p_j}\right)},\end{aligned}$$

which is (4.5). □

The statement of Theorem 4.3 and its argument can now be adapted in a straightforward way to obtain a multilinear version of Theorem 3.22. We leave this task to the interested reader. However, we would like to make several remarks on Theorem 4.3.

By re-normalizing the weights and writing $w_j := w_j^{p_j}$, the statement of our theorem can be reformulated in terms of the weight $\nu_{\mathbf{w}}$ introduced in Definition 4.2. Ignoring the precise constants in terms of the increasing functions ϕ_j in Theorem 4.3, we obtain the following.

Corollary 4.7. *Let T be an m-linear operator. Fix $1 < p_j < \infty$, $1 < s_j < \infty$, and $\theta_j > 0$, $1 \le j \le m$. Let $1 < p < \infty$ be so that*

$$\frac{1}{p} = \sum_{j=1}^m \frac{1}{p_j},$$

and suppose that for all $\boldsymbol{w} = (w_1, \dots, w_m)$ satisfying $w_j^{\theta_j/p_j} \in A_{s_j}$, $1 \le j \le m$, we have

$$\|T\boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \prod_{j=1}^m \|f_j\|_{L^{p_j}(w_j)}. \tag{4.8}$$

Then, for all $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$, *for each multi-index* α *and vector weight* $\boldsymbol{w} = (w_1, \dots, w_m)$ *such that for all* $1 \le j \le m$, $w_j^{\eta_j \theta_j / p_j} \in A_{s_j}$ *with some* $1 < \eta_j < \infty$, *we have*

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \prod_{j=1}^m \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}(w_j)}. \tag{4.9}$$

The following is a variation on the corollary just stated, in which we express the conditions on the weights in a simpler manner.

Corollary 4.10. *Let* T *be an* m*-linear operator. Fix* $1 < p_j < \infty$, $1 \le j \le m$, *and* $\theta > 0$. *Let* $1 < p < \infty$ *be so that*

$$\frac{1}{p} = \sum_{j=1}^m \frac{1}{p_j},$$

and suppose that for all $\boldsymbol{w} = (w_1, \dots, w_m)$ *satisfying* $w_j^\theta \in A_{p_j}$, $1 \le j \le m$, *we have*

$$\|T\boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \prod_{j=1}^m \|f_j\|_{L^{p_j}(w_j)}. \tag{4.11}$$

Then, for all $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$, *for each multi-index* α *and vector weight* $\boldsymbol{w} = (w_1, \dots, w_m)$ *as above, we have*

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \prod_{j=1}^m \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}(w_j)}. \tag{4.12}$$

4.2. **Multilinear commutators II.** The unsatisfying aspect of the statement of Theorem 4.3 or of its subsequent corollaries is that the dependence of the constants in the estimates is on the Muckenhoupt constant of each one of the weights w_j that comprise the vector weight $\mathbf{w}$. We present here a version with constants depending only on the $A_{\mathbf{P}}$ constant of $\mathbf{w}$.

Theorem 4.13. *Let* T *be an* m*-linear operator. Fix* $1 < p_j < \infty$ *and let* $1 < p < \infty$ *be so that*

$$\frac{1}{p} = \sum_{j=1}^m \frac{1}{p_j}.$$

Suppose that there exists an increasing functions $\phi : [1, \infty) \to [0, \infty)$ *such that for all* $\boldsymbol{w} = (w_1, \dots, w_m) \in A_{\boldsymbol{P}}$, *we have*

$$\|T\boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \phi\left([\boldsymbol{w}]_{A_{\boldsymbol{P}}}\right) \prod_{j=1}^m \|f_j\|_{L^{p_j}(w_j)}. \tag{4.14}$$

Then, for all $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$ *and for each multi-index* α, *we have*

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \alpha! \, \phi\left(c_{\boldsymbol{P}} [\boldsymbol{w}]_{A_{\boldsymbol{P}}}\right) [\boldsymbol{w}]_{A_{\boldsymbol{P}}}^{|\alpha| \max\{p, p_1', \dots, p_m'\}} \prod_{j=1}^m \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}(w_j)}, \tag{4.15}$$

where $c_{\boldsymbol{P}} = 4^{1 + \sum_{j=1}^m \min\{1/p_j, 1/p_j'\}}$.

For the convenience of notation, we will only prove this result for bilinear operators ($m = 2$). Before we begin our argument, let us recall that if $\mathbf{w} = (w_1, w_2)$ and $\mathbf{P} = (p_1, p_2)$, then we have [33] the following equivalence:

$$\mathbf{w} \in A_{\mathbf{P}} \Leftrightarrow \nu_{\mathbf{w}} = w_1^{p/p_1} w_2^{p/p_2} \in A_{2p}, \sigma_1 := w_1^{1-p_1'} \in A_{2p_1'}, \text{ and } \sigma_2 := w_2^{1-p_2'} \in A_{2p_2'}.$$

Moreover, by [18, Lemma 3.1], we have $[\sigma_j]_{A_{2p_j'}} \leq [\mathbf{w}]_{A_{\mathbf{P}}}^{p_j'}$, $j = 1, 2$, and since $p/p_1' + p/p_2' = 2p - 1$, Hölder's inequality gives

$$\left(\fint_Q \nu_{\mathbf{w}}\, dx\right)\left(\fint_Q \nu_{\mathbf{w}}^{1-(2p)'}\, dx\right)^{2p-1} \leq \left(\fint_Q \nu_{\mathbf{w}}\, dx\right)\left(\fint_Q \sigma_1\, dx\right)^{p/p_1'}\left(\fint_Q \sigma_2\, dx\right)^{p/p_2'} \leq [\mathbf{w}]_{A_{\mathbf{P}}}^{p},$$

that is, $[\nu_{\mathbf{w}}]_{A_{2p}} \leq [\mathbf{w}]_{A_{\mathbf{P}}}^{p}$.

Proof. As before, and without loss of generality, we assume that b_j, $j = 1, 2$ are real valued and normalized so that their $\mathcal{BMO}$ norms are equal to 1. By repeating the argument with the Cauchy integral trick in Theorem 4.3, given $\mathbf{w} \in A_{\mathbf{P}}$ we see that everything reduces to showing that for some appropriate $\delta_1, \delta_2 > 0$ (to be chosen later) and for $|z_1| = \delta_1$, $|z_2| = \delta_2$, we have

$$\widetilde{\mathbf{w}} := (\widetilde{w}_1, \widetilde{w}_2) := (w_1 e_{b_1}, w_2 e_{b_2}) := (w_1 e^{-\mathrm{Re}(z_1) p_1 b_1}, w_2 e^{-\mathrm{Re}(z_2) p_2 b_2}) \in A_{\mathbf{P}}.$$

Note that as observed above $\nu_{\mathbf{w}} \in A_{2p}$ and $\sigma_j := w_j^{1-p_j'} \in A_{2p_j'}$, for $j = 1, 2$. Using now Lemma 3.26 and writing, for a given weight w, $\rho(w)$ instead of ρ_w, we can find $r = r(\mathbf{w}) = \min\{\rho(\nu_{\mathbf{w}}), \rho(\sigma_1), \rho(\sigma_2)\} > 1$ so that

$$r' \sim \max\{[\nu_{\mathbf{w}}]_{A_{2p}}, [\sigma_1]_{A_{2p_1}}, [\sigma_2]_{A_{2p_2}}\} \leq [\mathbf{w}]_{A_{\mathbf{P}}}^{\max\{p, p_1', p_2'\}} \tag{4.16}$$

and the following reverse Hölder inequalities hold:

$$\left(\fint_Q \nu_{\mathbf{w}}^{r}\, dx\right)^{1/r} \leq 2 \fint_Q \nu_{\mathbf{w}}\, dx \tag{4.17}$$

and for $j = 1, 2$

$$\left(\fint_Q \sigma_j^{r}\, dx\right)^{1/r} \leq 2 \fint_Q \sigma_j\, dx. \tag{4.18}$$

Using these, Hölder's inequality and regrouping terms, we get

$$\left(\fint_Q \nu_{\widetilde{\mathbf{w}}}\, dx\right)^{1/p} \prod_{j=1}^{2}\left(\fint_Q \widetilde{w}_j^{1-p_j'}\, dx\right)^{1/p_j'} = \left(\fint_Q \nu_{\mathbf{w}}\, e_{b_1}^{p/p_1}\, e_{b_2}^{p/p_1}\, dx\right)^{1/p} \prod_{j=1}^{2}\left(\fint_Q \sigma_j\, e_{b_j}^{1-p_j'}\, dx\right)^{1/p_j'}$$

$$
\begin{aligned}
&\leq \left(\fint_Q v_{\mathbf{w}}^r\,dx\right)^{1/(pr)} \left(\fint_Q e_{b_1}^{r'p/p_1}\, e_{b_2}^{r'p/p_1}\,dx\right)^{1/(pr')} \\
&\qquad \prod_{j=1}^2 \left(\fint_Q \sigma_j^r\,dx\right)^{1/(p_j'r)} \left(\fint_Q e_{b_j}^{r'(1-p_j')}\,dx\right)^{1/p_j'r'} \\
&\leq 4[\mathbf{w}]_{A_{\mathbf{P}}} \prod_{j=1}^2 \left(\fint_Q e_{b_j}^{r'}\,dx\right)^{1/(p_jr')} \left(\fint_Q e_{b_j}^{r'(1-p_j')}\,dx\right)^{1/p_j'r'} \\
&\leq 4[\mathbf{w}]_{A_{\mathbf{P}}} \prod_{j=1}^2 \left[e^{-\mathrm{Re}(z_j)p_jr'b_j}\right]_{A_{p_j}}^{1/(p_jr')} \\
&\leq 4^{1+\delta_1+\delta_2}[\mathbf{w}]_{A_{\mathbf{P}}},
\end{aligned}
$$

where the last estimates uses Lemma 3.5 as long as we assume that $\delta_j \leq \frac{\min\{1,p_j-1\}}{p_jr'}$, $j=1,2$. Notice that this choice implies that $\delta_j \leq \min\{1/p_j, 1/p_j'\}$. On the other hand, recalling (4.16) and assuming further than $\delta_j \sim [\mathbf{w}]_{A_{\mathbf{P}}}^{-\max\{p,p_1',p_2'\}}$ we eventually obtain

$$
\|[T,\mathbf{b}]_\alpha \mathbf{f}\|_{L^p(v_{\mathbf{w}})} \lesssim \alpha!\delta_1^{-\alpha_1}\delta_2^{-\alpha_2}\phi\left(c_{\mathbf{P}}[\mathbf{w}]_{A_{\mathbf{P}}}\right)\|f_1\|_{L^{p_1}(w_1)}\|f_2\|_{L^{p_2}(w_2)}.
$$

with $c_{\mathbf{P}} = 4^{1+\min\{1/p_1,1/p_1'\}+\min\{1/p_2,1/p_2'\}}$. This easily gives the desired estimate. □

4.3. **Multilinear commutators vs weighted estimates.** In this subsection we give a multilinear version of Theorem 3.31, thus showing that there is a two way bridge between weighted estimates for commutators of the operator and weighted estimates for the operator itself in the multilinear setting as well. We do not include the proof as it follows the linear situation very closely using the Cauchy integral formula for several complex variables (4.6) and a multivariable Taylor series.

Theorem 4.19. *Let $\boldsymbol{P} = (p_1,\dots,p_m) \in (1,\infty)^m$ and $p > 1$ be such that $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_m}$, and let T be a multilinear operator.*

(*a*) *Suppose there are $\lambda_0 > 0$ and a function $\Phi : [0,\lambda_0]^m \to [0,\infty)$ that is increasing in each variable such that for each $\boldsymbol{b} = (b_1,\dots,b_m) \in \mathcal{BMO}^m$ we have*

$$
\|T\boldsymbol{f}\|_{L^p\left(\exp(\sum_{j=1}^\infty \frac{p\lambda_j}{p_j}b_j)\right)} \leq \Phi(|\lambda_1|\,\|b_1\|_{\mathcal{BMO}},\dots,|\lambda_m|\,\|b_m\|_{\mathcal{BMO}}) \prod_{j=1}^m \|f_j\|_{L^{p_j}\left(\exp(\lambda_j b_j)\right)}
$$

for $|\lambda_j| \leq \lambda_0/\|b_j\|_{\mathcal{BMO}}$. Then, for each multi-index α and $\boldsymbol{b} \in \mathcal{BMO}^m$

$$
\|[T,\boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\mathbb{R}^n)} \leq \Phi(\lambda_0,\dots,\lambda_0)\alpha! \prod_{j=1}^m \left(\frac{p_j\|b_j\|_{\mathcal{BMO}}}{\lambda_0}\right)^{\alpha_j} \|f_j\|_{L^{p_j}(\mathbb{R}^n)}.
$$

(*b*) *Conversely, suppose there exist* $C_0, \lambda_0 > 0$ *such that for* $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$ *and for every multi-index* α *we have*

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\mathbb{R}^n)} \leq C_0 \alpha! \prod_{j=1}^m \left(\frac{p_j \|b_j\|_{\mathcal{BMO}}}{\lambda_0} \right)^{\alpha_j} \|f_j\|_{L^{p_j}(\mathbb{R}^n)}.$$

Then, there exists a function $\Phi : [0, \lambda_0)^m \to [0, \infty)$*, increasing in each variable, such that for every* $\boldsymbol{b} \in \mathcal{BMO}^m$ *we have*

$$\|T\boldsymbol{f}\|_{L^p\left(\exp\left(\sum_{j=1}^\infty \frac{p\lambda_j}{p_j} b_j\right)\right)} \leq \Phi(|\lambda_1| \, \|b_1\|_{\mathcal{BMO}}, \dots, |\lambda_m| \, \|b_m\|_{\mathcal{BMO}}) \prod_{j=1}^m \|f_j\|_{L^{p_j}\left(\exp(\lambda_j b_j)\right)},$$

for all $\lambda_1, \dots, \lambda_m$ *satisfying* $|\lambda_j| < \lambda_0 / \|b_j\|_{\mathcal{BMO}}$.

After this work was completed and communicated at many venues, we have learned that Kunwar and Ou [29] have extended some of our methods to the two-weight setting. Actually, there is also a growing literature in the subject of two-weighted estimates and numerous preprints using different methods have recently appeared, but those topics are beyond the interest of our current work.

5. Applications

The goal of this section is to apply our general results connecting weighted estimates for an operator and those of its commutators in various situations, some explored and some unexplored before. We begin with the case of linear or linearizable operators.

5.1. **The Coifman-Rochberg-Weiss result.** As a first application, from Theorem 3.22 we recover the classical Coifman-Rochberg-Weiss (see [12]) whose quantitative control of the constants appeared in [10].

Corollary 5.1. *Let* T *be either a linear or a linearizable operator. Fix* $1 < p < \infty$*, and suppose that there exist an increasing function* $\phi : [1, \infty) \longrightarrow [0, \infty)$ *such that for each* $f \in L^p(w)$ *there holds*

$$\|Tf\|_{L^p(w)} \leq \phi\left([w]_{A_p}\right) \|f\|_{L^p(w)}, \qquad \forall\, w \in A_p.$$

Then, for each $b \in \mathcal{BMO}$ *and for every* $w \in A_p$ *and, for every* $k \geq 1$*, there holds*

$$\left\|T_b^k f\right\|_{L^p(w)} \leq k! \left(\frac{2^{2\max\{p,p'\}+n+2} \, p}{\min\{1, p-1\}} \right)^k [w]_{A_p}^{k \max\{1, \frac{1}{p-1}\}} \phi\left(4^{\min\{1,p-1\}} 2^p [w]_{A_p}\right) \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w)}.$$

If we now take T a Calderón-Zygmund operator, the A_2-conjecture proved by Hytönen in [25] gives the result in [10]:

$$\|Tf\|_{L^p(w)} \leq C \, [w]_{A_p}^{\max\{1, \frac{1}{p-1}\}} \|f\|_{L^p(w)}, \qquad \forall\, w \in A_p,$$

and therefore for the commutators one obtains

$$\left\|T_b^k f\right\|_{L^p(w)} \leq C_k [w]_{A_p}^{(k+1)\max\{1, \frac{1}{p-1}\}} \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w)}.$$

5.2. **Fractional integrals.** We begin by recalling the notion of the Riesz potentials or fractional integrals. For a more thorough introduction see [42, Chapter 5]. Given $0 < \alpha < n$, we write

$$I_\alpha f(x) = (-\Delta)^{-\alpha/2} f(x) = c_\alpha \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy$$

where c_α is a constant depending on n and α. The relevant class of weights is $A_{p,q}$ defined as follows: $w \in A_{p,q}$ if

$$[w]_{A_{p,q}} := \sup_{Q \subset \mathbb{R}^n} \left(\fint_Q w^q\, dx \right) \left(\fint_Q w^{-p'}\, dx \right)^{q/p'} < \infty.$$

The weighted norm inequalities for I_α were obtained by B. Muckenhoupt and R. Wheeden [38] and the sharp behavior in terms of the weight constants by [30]. The precise estimate is as follows: for every $0 < \alpha < n$, $1 < p < n/\alpha$, $1/p - 1/q = \alpha/n$, and $w \in A_{p,q}$ one has

$$\|I_\alpha f\|_{L^q(w^q)} \le C_{p,q,\alpha} [w]_{A_{p,q}}^{(1-\frac{\alpha}{n})\max\{1, \frac{p'}{q}\}} \|f\|_{L^p(w^p)}.$$

It is easy to see that $w \in A_{p,q}$ iff $w^q \in A_{q\,\frac{n-\alpha}{n}}$ and moreover

$$\left[w^q\right]_{A_{q\,\frac{n-\alpha}{n}}} = [w]_{A_{p,q}}.$$

The commutators of fractional integrals and BMO functions were first studied in [8]. We can use Theorem 3.22 and obtain the following quantitative weighted result which was shown in [16] for $k = 1$, but seems to be new when $k > 1$.

Corollary 5.2. *Fix $0 < \alpha < n$, $1 < p < n/\alpha$ and $1/p - 1/q = \alpha/n$. For every $k \ge 1$ and $b \in \mathcal{BMO}$ one has*

$$\left\| (I_\alpha)_b^k f \right\|_{L^q(w^q)} \le C_{p,q,\alpha} k! \left(\frac{2^{2\max\{s,s'\}+n+2} q}{\min\{1, s-1\}} \right)^k [w]_{A_{p,q}}^{(k+1-\frac{\alpha}{n})\max\{1, \frac{p'}{q}\}} \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w^p)},$$

for every $w \in A_{p,q}$ and where $s = q\,\frac{n-\alpha}{n}$.

5.3. **Operators associated with the Kato conjecture.** The following estimate models the behavior of operators associated with the Kato conjecture (see below).

Corollary 5.3. *Let $1 \le r_- < p < r_+ \le \infty$ and assume that*

$$\|Tf\|_{L^p(w)} \le C_w \|f\|_{L^p(w)}, \qquad \forall\, w \in A_{\frac{p}{r_-}} \cap RH_{(\frac{r_+}{p})'}.$$

Then, for every $b \in \mathcal{BMO}$ and for every $k \ge 1$

$$\|T_b^k f\|_{L^p(w)} \le C_k \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w)}, \qquad \forall\, w \in A_{\frac{p}{r_-}} \cap RH_{(\frac{r_+}{p})'}.$$

Notice that if $r_+ = \infty$ it is understood that the condition RH_1 is vacuous. This result follows easily from Theorem 3.22 (this time without paying attention to constants) after observing that Proposition 2.1 (g) yields

$$w \in A_{\frac{p}{r_-}} \cap RH_{\left(\frac{r_+}{p}\right)'} \qquad \Longleftrightarrow \qquad w^{(\frac{r_+}{p})'} \in A_{\left(\frac{r_+}{p}\right)'\left(\frac{p}{r_-}-1\right)+1}.$$

Moreover, $A_{\frac{p}{r_-}} \cap RH_{(\frac{r_+}{p})'}$ coincides with the class $A_{p,(r_-,r'_+)}$ introduced in Appendix A where one takes $m = 1$, with constants

$$[w^{(\frac{r_+}{p})'}]_{A_{\left(\frac{r_+}{p}\right)'\left(\frac{p}{r_-}-1\right)+1}} = [w]^{\frac{pr_+}{r_+-p}}_{A_{p,(r_-,r'_+)}}.$$

Let A be an $n \times n$ matrix of complex and L^∞-valued coefficients defined on $\mathbb{R}^n$. We assume that this matrix satisfies the following ellipticity (or "accretivity") condition: there exist $0 < \lambda \leq \Lambda < \infty$ such that

$$\lambda\,|\xi|^2 \leq \operatorname{Re} A(x)\,\xi \cdot \bar{\xi} \qquad \text{and} \qquad |A(x)\,\xi \cdot \bar{\zeta}| \leq \Lambda\,|\xi|\,|\zeta|,$$

for all $\xi, \zeta \in \mathbb{C}^n$ and almost every $x \in \mathbb{R}^n$. We have used the notation $\xi \cdot \zeta = \xi_1\,\zeta_1 + \cdots + \xi_n\,\zeta_n$ and therefore $\xi \cdot \bar{\zeta}$ is the usual inner product in $\mathbb{C}^n$. Note that then $A(x)\,\xi \cdot \bar{\zeta} = \sum_{j,k} a_{j,k}(x)\,\xi_k\,\bar{\zeta}_j$. Associated with this matrix we define the second order divergence form operator

$$Lf = -\operatorname{div}(A\,\nabla f),$$

which is understood in the standard weak sense as a maximal-accretive operator on the space $L^2(\mathbb{R}^n, dx)$ with domain $D(L)$ by means of a sesquilinear form. Associated to this operator we can consider a functional calculus $\varphi(L)$ where φ is holomorphic and bounded in some sector, the Riesz transform $\nabla L^{-1/2}$, and some square functions. The L^p theory for these operators was developed in the monograph [2]. The weighted norm inequalities were obtained in [4] using a generalized Calderón-Zygmund theory from [3]. As part of these results it was obtained that the commutators of these operators with $\mathcal{BMO}$ functions also satisfy weighted norm inequalities. It should be pointed out that the proof of the weighted norm inequalities for the associated commutators followed from the developed Calderón-Zygmund theory with a somehow technical adaptation of the proof of the weighted norm inequalities for the corresponding operators. As we are going to see next, a simple application of Corollary 5.3 gives the same estimates without any extra effort. In order to apply Corollary 5.3 we just need to recall that the following weighted norm inequalities hold for $\varphi(L)$ and $\nabla L^{-1/2}$ (see [4]):

$$\|\varphi(L)f\|_{L^p(w)} \leq C_w\,\|f\|_{L^p(w)}, \qquad \forall\, p_- < p < p_+,\ w \in A_{\frac{p}{p_-}} \cap RH_{(\frac{p_+}{p})'};$$

and

$$\|\nabla L^{-1/2} f\|_{L^p(w)} \leq C_w\,\|f\|_{L^p(w)}, \qquad \forall\, q_- < p < q_+,\ w \in A_{\frac{p}{q_-}} \cap RH_{(\frac{q_+}{p})'};$$

where (p_-, p_+) and (q_-, q_+) are respectively the maximal intervals where the semigroup $\{e^{-t\,L}\}_{t>0}$ and its gradient $\{\sqrt{t}\,\nabla e^{-t\,L}\}_{t>0}$ are uniformly bounded on $L^p(\mathbb{R}^n)$. This clearly allows us to apply Corollary 5.3 and obtain the corresponding commutators results as desired. Similar results can be obtained for the square functions associated with L (see [4] for full details and references).

Remark 5.4. A careful examination of [6, Theorem 1.1, Proposition 6.4] gives that for the family of operators in [6, Theorem 1.1] (or the sparse operators in [6, Proposition 6.4]) the

following quantitative estimates hold:

$$\|Tf\|_{L^p(w)} \le c_p [w]_{A_{p,(r_-,r_+')}}^{\max\left\{\left(\frac{p}{r_-}\right)', \left(\frac{p'}{r_+'}\right)'\right\}} \|f\|_{L^p(w)}, \tag{5.5}$$

for all $r_- < p < r_+$ and $w \in A_{\frac{p}{r_-}} \cap RH_{(\frac{r_+}{p})'}$, and were we have employed the notation from Appendix A with $m = 1$. Thus Corollary 5.3, via the method of Theorem 3.22, combined with inequality (5.5) implies there exists some constant C such that

$$\|T_b^k f\|_{L^p(w)} \lesssim k! \, [w]_{A_{p,(r_-,r_+')}}^{\max\left\{\left(\frac{p}{r_-}\right)', \left(\frac{p'}{r_+'}\right)'\right\} + k \max\left\{r_-\left(\frac{p}{r_-}\right)', p\left(\frac{r_+}{p}\right)'\right\}} \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w)},$$

for all $r_- < p < r_+$ and $w \in A_{\frac{p}{r_-}} \cap RH_{(\frac{r_+}{p})'}$.

5.4. **Fractional operators associated with the Kato conjecture.** Very much as before we can consider the fractional operators $L^{-\alpha/2}$. The weighted norm inequalities for these were proved in [5]. Using the same notation as before, if $p_- < p < q < p_+$ and $1/p - 1/q = \alpha/n$, then

$$\|L^{-\alpha/2} f\|_{L^q(w^q)} \le C_w \|f\|_{L^p(w^p)}, \qquad \forall\, w \in A_{1+\frac{1}{p_-}-\frac{1}{p}} \cap RH_{q\left(\frac{q_+}{q}\right)'}.$$

Note that Proposition 2.1 (g) gives that the previous estimates hold for a class of weights that can be written as $w^\theta \in A_s$ for some $\theta > 1$ and $s > 1$. Then, we can apply Theorem 3.22 and conclude that, under the same hypotheses, for every $b \in \mathcal{BMO}$ and $k \ge 1$

$$\|(L^{-\alpha/2})_b^k f\|_{L^q(w^q)} \le C_w \, \|b\|_{\mathcal{BMO}}^k \|f\|_{L^p(w^p)}, \qquad \forall\, w \in A_{1+\frac{1}{p_-}-\frac{1}{p}} \cap RH_{q\left(\frac{q_+}{q}\right)'}.$$

These estimates were proved in [5] using a Calderón-Zygmund type result that allows one to extend the unweighted estimates for the fractional operators to the commutators. The present method immediately produces the same results once the weighted estimates have been obtained.

We continue our applications by considering the case of bilinear operators. Several of the statements can be generalized to m-linear operators, $m \ge 2$, an easy task that is left to the interested reader.

5.5. **Bilinear Calderón-Zygmund operators.** We say that a bilinear operator T a priori defined from $\mathcal{S} \times \mathcal{S}$ into $\mathcal{S}'$ of the form

$$T(f,g)(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x,y,z) f(y) g(z) \, dydz$$

is a bilinear Calderón-Zygmund operator if it can be extended as a bounded operator from $L^{p_1} \times L^{p_2}$ to L^p for all $1 < p_1, p_2 < \infty$ with $1/p_1 + 1/p_2 = 1/p$, and its distributional kernel K coincides, away from the diagonal $\{(x,y,z) \in \mathbb{R}^{3n} : x = y = z\}$, with a function $K(x,y,z)$ locally integrable which satisfies estimates of the form

$$|\partial^\alpha K(x,y,z)| \lesssim \big(|x-y| + |x-z| + |y-z|\big)^{-2n-|\alpha|}, |\alpha| \le 1.$$

The estimates on K above are not the most general that one can impose in such theory, see [23]. An immediate consequence of Theorem 4.13 and the known weighted boundedness

of the bilinear Calderón-Zygmund operators [33] leads to the following result. The bounds are sharp for the operator T ($|\alpha| = 0$), see [35].

Theorem 5.6. *Let T be a bilinear Calderón-Zygmund operator, $1 < p_1, p_2, p < \infty$ be such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $\boldsymbol{b} = (b_1, b_2) \in \mathcal{BMO}^2$, $\boldsymbol{w} = (w_1, w_2) \in A_{\boldsymbol{P}}$ and α a multi-index. Then, we have*

$$\|[T, \boldsymbol{b}]_{\alpha} \boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \alpha! [\boldsymbol{w}]_{A_{\boldsymbol{P}}}^{(|\alpha|+1)\max\{p, p_1', p_2'\}} \prod_{j=1}^{2} \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}(w_j)}.$$

The quantitative estimates in the above theorem are new. The limitations in the value of p in our methods come from the use of Minkowski's inequality. However, for (non-quantitative) boundedness results the range of exponents can be extended so that $1/2 < p < \infty$. See for example [33] for $|\alpha| = 1$ and [41] for $|\alpha| = 2$. Alternatively, from Theorem 5.6 the full range of exponents could be obtained by extrapolation (although not with optimal constants); see the recent work [34, Corollary 1.5].

5.6. **Bilinear fractional integrals.** Consider now bilinear operators with positive kernels of the form

$$BI_s(f, g)(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{1}{(|x - y| + |x - z|)^{2n-s}} f(y) g(z)\, dy dz,$$

with $0 < s < 2n$. The appropriate class of vector weights to study these operators is defined as follows. For $1 < p_1, p_2 < \infty$ with $1/p_1 + 1/p_2 > s/n$, let again $\mathbf{P} = (p_1, p_2)$ and $q > 0$ be such that

$$\frac{1}{q} = \frac{1}{p_1} + \frac{1}{p_2} - \frac{s}{n}.$$

The vector weight $\mathbf{w} = (w_1, w_2)$ is said to belong to the class $A_{\mathbf{P},q}$ if

$$[\mathbf{w}]_{A_{\mathbf{P},q}} := \sup_{Q} \left(\fint_Q w_1^q w_2^q dx \right)^{1/q} \left(\fint_Q w_1^{-p_1'} dx \right)^{1/p_1'} \left(\fint_Q w_2^{-p_2'} dx \right)^{1/p_2'} < \infty.$$

It was shown by Moen [37] and Chen and Xue [9] that $BI_\alpha : L^{p_1}(w_1^{p_1}) \times L^{p_2}(w_2^{p_2}) \to L^q(w_1^q w_2^q)$ as long as $\mathbf{w} \in A_{\mathbf{P},q}$. The dependence on the weighted constant was shown in [36]. Using this fact and the $A_{\mathbf{P},q}$ version of our multilinear result, Theorem 4.13, we immediately recover the boundedness of the commutators of BI_s with $\mathbf{b} \in \mathcal{BMO}^2$ from the work in [9]. Specifically, for α a multi-index, we have

$$[BI_s, \mathbf{b}]_\alpha : L^{p_1}(w_1^{p_1}) \times L^{p_2}(w_2^{p_2}) \to L^q(w_1^q w_2^q),$$

with appropriate quantitative estimates on the commutator operator norms similar to the ones stated in Corollary 5.2.

5.7. **The bilinear Hilbert transform.** The bilinear Hilbert transform, defined via

$$BHT(f, g)(x) = \text{p.v} \int_{\mathbb{R}} f(x - t) g(x + t) \frac{dt}{t},$$

is a celebrity in harmonic analysis. It is a bilinear operator whose multiplier, unlike the ones for bilinear Calderón-Zygmund operators which are singular only at the origin, is singular

along a line when viewed in the frequency plane. Let $1 < p_1, p_2, p < \infty$ be such that $1/p_1 + 1/p_2 = 1/p$ and $\mathbf{w} = (w_1, w_2)$ such that $w_1^{2p_1} \in A_{p_1}$ and $w_2^{2p_2} \in A_{p_2}$. Under these conditions, Culiuc, Di Plinio and Ou [17] proved that

$$BHT : L^{p_1}(w_1^{p_1}) \times L^{p_2}(w_2^{p_2}) \to L^p(w_1^p w_2^p).$$

In view of Corollary 4.10, we immediately get the following new boundedness result for the commutator of BHT.

Theorem 5.7. *Let $\boldsymbol{b} \in \mathcal{BMO}^2$, $\boldsymbol{w} = (w_1, w_2)$ be such that $w_j^2 \in A_{p_j}$, where $1 < p_j < \infty$, $j = 1, 2$ and $1 < p < \infty$ be so that $1/p = 1/p_1 + 1/p_2$. Then, for any multi-index α we have*

$$[BHT, \boldsymbol{b}]_\alpha : L^{p_1}(w_1) \times L^{p_2}(w_2) \to L^p(v_{\boldsymbol{w}}).$$

Combining this result with the extrapolation method from [14] we can obtain weighted estimates for the commutators for $\frac{2}{3} < p \leq 1$.

Corollary 5.8. *Let $\boldsymbol{b} \in \mathcal{BMO}^2$. Given $1 < p_1,\ p_2 < \infty$, let $p > 2/3$ be so that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. Then, for any multi-index α we have*

$$[BHT, \boldsymbol{b}]_\alpha : L^{p_1}(w_1) \times L^{p_2}(w_2) \to L^p(v_{\boldsymbol{w}}), \tag{5.9}$$

for all $\boldsymbol{w} = (w_1, w_2)$ such that $w_j \in A_{\max\{1, p_j/2\}} \cap RH_{\max\{1, 2/p_j\}}$.
In particular,

$$[BHT, \boldsymbol{b}]_\alpha : L^{p_1}(|x|^{-a}) \times L^{p_2}(|x|^{-a}) \longrightarrow L^p(|x|^{-a}), \tag{5.10}$$

if $a = 0$ or if

$$1 - \min\left\{\max\left\{1, \frac{p_1}{2}\right\}, \max\left\{1, \frac{p_2}{2}\right\}\right\} < a < \min\left\{1, \frac{p_1}{2}, \frac{p_2}{2}\right\}. \tag{5.11}$$

As a result, (5.10) holds for all $a \in [0, 1/2)$.

We would like to emphasize that (5.10) with $a = 0$ gives that $[BHT, \mathbf{b}]_\alpha$ satisfies the same unweighted estimates as BHT. The proof of Corollary 5.8 follows that of [14, Corollary 1.23] by simply replacing BHT by $[BHT, \mathbf{b}]_\alpha$. The bottom line is that Theorem 5.7 provides the initial weighted norm inequalities needed to apply the extrapolation method and, in turn, these are the same for BHT and $[BHT, \mathbf{b}]_\alpha$. The same idea allows us to get a version of [14, Theorem 1.18] valid for $[BHT, \mathbf{b}]_\alpha$, and also some vector-valued inequalities. Further details and the precise statements are left to the interested reader.

In Appendix A we also state some additional weighted estimates for the commutators of the BHT with weights in some classes generalizing $A_{\mathbf{P}}$, see Corollary A.8.

5.8. **Multi-parameter operators.** An interesting application of our general result in a general setting is the boundedness of multi-parameter commutators of Calderón-Zygmund operators of product type. Moreover, we shall discover that if we consider the basis of rectangles $\mathcal{R}$, then the two flavors of spaces $BMO_{\mathcal{R}}$ and $\mathcal{BMO}_{\mathcal{R}}$ coincide.

For simplicity, we suppose that our underlying space is $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, and we let $\mathcal{R}$ denote the family of rectangles with sides parallel to the axes. As explained in Subsection 3.3, we

can consider the following two versions of spaces: $\mathcal{BMO}_{\mathcal{R}}$, defined via the norm

$$\|f\|_{\mathcal{BMO}_{\mathcal{R}}} = \sup_{R\in\mathcal{R}} \|f - f_R\|_{\exp L,R},$$

and $BMO_{\mathcal{R}}$, defined via the norm

$$\|f\|_{BMO_{\mathcal{R}}} = \sup_{R\in\mathcal{R}} \fint_R |f - f_R|\, dx.$$

The space $BMO_{\mathcal{R}}$ is sometimes called "little BMO"–since it is smaller than the $BMO(\mathbb{R}^2)$ space defined on cubes–and denoted *bmo*. Now, since

$$\|f\|_{BMO_{\mathcal{R}}} \leq \|f\|_{\mathcal{BMO}_{\mathcal{R}}},$$

we immediately obtain that $\mathcal{BMO}_{\mathcal{R}} \subset bmo$. It turns out that the converse inclusion

$$bmo \subset \mathcal{BMO}_{\mathcal{R}}$$

is also true, a "little" fact previously unobserved to the best of our knowledge. In order to prove it, we need a few preliminary lemmas. In what follows, we will write $A_{2,\mathscr{B}} = A_{2,\mathscr{B}}(\mathbb{R}^2)$ to denote the class of A_2 weights defined with respect to the family $\mathscr{B}$, see again Proposition 3.16; and we will write $A_2 = A_2(\mathbb{R})$ for the usual A_2 class of weights, see Subsection 2.1.

The lemma we first state below is proved in [21, pp. 406-408] for cubes, but the exact same proof actually holds for any family $\mathscr{B}$. Our interest will eventually be in applying these lemmas to the case where $\mathscr{B} = \mathcal{R}$.

Lemma 5.12. *Let f be a real valued locally integrable function. Then, e^f is in $A_{2,\mathscr{B}}$ if and only if*

$$\sup_{B\in\mathscr{B}} \fint_B e^{|f-f_B|}\, dx = C_0 < \infty.$$

Moreover, if $e^f \in A_2$, then $C_0 \leq 2[e^f]_{A_{2,\mathscr{B}}}$.

Lemma 5.13. *Let f be a real valued locally integrable function. If*

$$\sup_{B\in\mathscr{B}} \fint_B e^{|f-f_B|}\, dx = C_0 < \infty,$$

then $f \in \mathcal{BMO}_{\mathscr{B}}$ and $\|f\|_{\mathcal{BMO}_{\mathscr{B}}} \leq \max\{1, \log_2 C_0\}$.

Proof. If $C_0 \leq 2$, then clearly, by definition, $\|f\|_{\mathcal{BMO}_{\mathscr{B}}} \leq 1$. Assume thus that $C_0 > 2$ and let $\lambda \geq 1$. By Jensen's inequality, we have

$$\fint_B e^{\frac{|f-f_B|}{\lambda}}\, dx \leq \left(\fint_B e^{|f-f_B|}\, dx\right)^{\frac{1}{\lambda}} \leq C_0^{\frac{1}{\lambda}}.$$

Taking $\lambda = \log_2 C_0$ we obtain $\|f\|_{\mathcal{BMO}_{\mathscr{B}}} \leq \log_2 C_0$. □

We point out that, if $f \in \mathcal{BMO}_{\mathscr{B}}$, then $\fint_B e^{|f-f_B|}\, dx$ may not be finite, but obviously

$$\fint_B e^{\frac{|f-f_B|}{\|f\|_{\mathcal{BMO}_{\mathscr{B}}}}}\, dx \leq 2.$$

We have the following immediate corollary of the previous two lemmas.

Corollary 5.14. *Let f be a real valued locally integrable function. If $e^f \in A_{2,\mathscr{B}}$, then $f \in \mathcal{BMO}_{\mathscr{B}}$ and*

$$\|f\|_{\mathcal{BMO}_{\mathscr{B}}} \leq 1 + \log_2 [e^f]_{A_{2,\mathscr{B}}}. \tag{5.15}$$

Conversely, we recall that as observed above (3.8) in Lemma 3.5 is valid for arbitrary bases provided $p > 1$. Thus, given $f \in \mathcal{BMO}_{\mathscr{B}}$, for all $\lambda \geq \|f\|_{\mathcal{BMO}_{\mathscr{B}}}$, we then have

$$\left[e^{\frac{f}{\lambda}}\right]_{A_{2,\mathscr{B}}} \leq 4^{\frac{\|f\|_{\mathcal{BMO}_{\mathscr{B}}}}{\lambda}}. \tag{5.16}$$

Lemma 5.17. *A weight $w(x,y)$ is in $A_{2,\mathcal{R}}(\mathbb{R}^2)$ if and only if $w(x,\cdot)$ and $w(\cdot,y)$ are uniformly in $A_2(\mathbb{R})$ for almost every $x, y \in \mathbb{R}$.*

The proof of this result can be found in [21, pp. 453-459]. The arguments there show that

$$\operatorname{ess\,sup}_x [w(x,\cdot)]_{A_2} \leq [w]_{A_{2,\mathcal{R}}}, \qquad \operatorname{ess\,sup}_y [w(\cdot,y)]_{A_2} \leq [w]_{A_{2,\mathcal{R}}}, \tag{5.18}$$

and

$$[w]_{A_{2,\mathcal{R}}} \leq C_n \operatorname{ess\,sup}_x [w(x,\cdot)]_{A_2} \operatorname{ess\,sup}_y [w(\cdot,y)]_{A_2}, \tag{5.19}$$

where C_n is a dimensional constant.

A similar statement holds for *bmo* functions. The following result was proved in [13, pp. 279-281].

Lemma 5.20. *Let f be a locally integrable function on $\mathbb{R}^2$. Then, f is in bmo if and only if $f(x,\cdot)$ and $f(\cdot,y)$ are uniformly in $BMO(\mathbb{R})$ for almost every $x, y \in \mathbb{R}$. Moreover,*

$$ess\,sup_x \|f(x,\cdot)\|_{BMO} + ess\,sup_y \|f(\cdot,y)\|_{BMO} \approx \|f\|_{bmo}. \tag{5.21}$$

Finally, we are able to prove the following

Theorem 5.22. *Let f be a locally integrable function on $\mathbb{R}^2$. Then, $f \in$ bmo if and only if $f \in \mathcal{BMO}_{\mathcal{R}}$.*

Proof. It suffices to show that $\|f\|_{\mathcal{BMO}_{\mathcal{R}}} \lesssim \|f\|_{bmo}$. Let then $f \in bmo$. By (5.21)

$$\|f(x,\cdot)\|_{BMO(\mathbb{R})} + \|f(\cdot,y)\|_{BMO(\mathbb{R})} \lesssim \|f\|_{bmo(\mathbb{R}^2)}.$$

By the John-Nirenberg inequality in $\mathbb{R}$ we have that, for some *fixed* $c > 0$,

$$\|f(x,\cdot)\|_{\mathcal{BMO}(\mathbb{R})} + \|f(\cdot,y)\|_{\mathcal{BMO}(\mathbb{R})} \leq c\|f\|_{bmo(\mathbb{R}^2)}.$$

By (5.16) applied in $\mathbb{R}$,

$$\left[e^{\frac{f(x,\cdot)}{c\|f\|_{bmo}}}\right]_{A_2(\mathbb{R})}, \quad \left[e^{\frac{f(\cdot,y)}{c\|f\|_{bmo}}}\right]_{A_2(\mathbb{R})} \leq 4,$$

which combined with (5.19) gives

$$\left[e^{\frac{f}{c\|f\|_{bmo}}}\right]_{A_{2,\mathcal{R}}} \leq 16C_n.$$

Now, by (5.15) , $\frac{f}{c\|f\|_{bmo}} \in \mathcal{BMO}_{\mathcal{R}}$ and

$$\Big|\Big|\frac{f}{c\|f\|_{bmo}}\Big|\Big|_{\mathcal{BMO}_{\mathcal{R}}} \leq 5 + \log_2 C_n,$$

and finally

$$\|f\|_{\mathcal{BMO}_{\mathcal{R}}} \leq (5 + \log_2 C_n)\, c\, \|f\|_{bmo}.$$

□

The following result was obtained by Ferguson and Sadosky [20, Theorem 2.1] on the 2-dimensional torus. The version we state follows from Theorems 5.22 and 3.17. The needed weighted estimates to use our approach in the 2-dimensional torus were obtained by Cotlar and Sadosky [13] and in the Euclidean setting are due to Fefferman and Stein [19].

Theorem 5.23. *Let H_j denote the 1-dimensional Hilbert transforms in the j-th variable, $j = 1, 2$ and let $b \in bmo$. Then, for all $f \in L^p(\mathbb{R}^2)$, $1 < p < \infty$, we have*

$$\|[b, H_1H_2]\|_{L^p(\mathbb{R}^2)} \lesssim \|b\|_{bmo}\|f\|_{L^p(\mathbb{R}^2)}.$$

The results above about *bmo* could be easily generalized to other product settings $\mathbb{R}^n \times \mathbb{R}^m$ and if T is Calderón-Zygmund operator of product type in the sense of [19] then we can recover with similar arguments the following known result.

Theorem 5.24. *Let T be a Calderón-Zygmund operator of product type and let b be a function in bmo$(\mathbb{R}^n \times \mathbb{R}^m)$. Then, for all $f \in L^p(\mathbb{R}^n \times \mathbb{R}^m)$, $1 < p < \infty$, we have*

$$\|[b, T]\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim \|b\|_{bmo}\|f\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}.$$

In fact, an even more general version of this theorem holds and was obtained through different methods for certain operators of Journé type by Ou, Petermichl, and Strouse [39, Theorem 7].

5.9. **Operators on spaces of homogeneous type.** All the previous applications, as long as the operators can be defined, can be extended to the case where the underlying Euclidean space is replaced by a space of homogeneous type. The details are left to the interested reader.

Appendix A. Multilinear commutators III

Shortly after this paper was posted on arXiv, a new extrapolation result was obtained in [34] associated to the classes of weights $A_{\mathbf{P},\mathbf{R}}$ which generalize the classes $A_{\mathbf{P}}$ introduced above. As a matter of fact, [34, Section 2.5] borrowed some of the key ideas from the present paper and sketched an argument that yields a commutator result along the lines of Theorem 4.13. We present here the complete argument with the quantitative bounds.

We begin by introducing the class of weights $A_{\mathbf{P},\mathbf{R}}$. Let $\mathbf{R} = (r_1, \dots, r_{m+1}) \in [1, \infty)^{m+1}$, and $\mathbf{P} = (p_1, \dots, p_m)$ with

$$\text{(A.1)}\quad r_j < p_j < \infty, \quad j = 1, \dots, m; \qquad \text{and} \qquad r'_{m+1} > p, \quad \text{where} \quad \frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_m}.$$

Let $\mathbf{w} = (w_1, \dots, w_m)$ and set

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

We say that the vector weight $\mathbf{w} \in A_{\mathbf{P},\mathbf{R}}$ if

$$[\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}} := \sup_{Q} \left(\fint_Q \nu_w^{\Delta_{m+1}/p}\, dx \right)^{1/\Delta_{m+1}} \prod_{j=1}^{m} \left(\fint_Q w_j^{-\Delta_j/p_j}\, dx \right)^{1/\Delta_j} < \infty,$$

where

$$\frac{1}{\Delta_j} = \frac{1}{r_j} - \frac{1}{p_j}, \quad j = 1, \dots, m; \qquad \text{and} \qquad \frac{1}{\Delta_{m+1}} = \frac{1}{p} - \frac{1}{r'_{m+1}}.$$

It is straightforward to see that $A_{\mathbf{P},\mathbf{R}} = A_{\mathbf{P}}$ whenever $\mathbf{R} = (1, \dots, 1)$. See also [7] for other similar classes of weights.

We next state the aforementioned result (the present formulation differs from [34, Theorem 2.21] where the emphasis was put on extrapolation):

Theorem A.2. *Let T be an m-linear operator. Let $\boldsymbol{R} = (r_1, \dots, r_{m+1}) \in [1, \infty)^{m+1}$, and $\boldsymbol{P} = (p_1, \dots, p_m)$ so that (A.1) holds and assume further that $p > 1$. Suppose that there exists an increasing functions $\phi : [1, \infty) \to [0, \infty)$ such that for all $\boldsymbol{w} = (w_1, \dots, w_m) \in A_{\boldsymbol{P},\boldsymbol{R}}$, we have*

$$\|Tf\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \phi\left([\boldsymbol{w}]_{A_{\boldsymbol{P},\boldsymbol{R}}}\right) \prod_{j=1}^{m} \|f_j\|_{L^{p_j}(w_j)}. \tag{A.3}$$

Then, for all $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathcal{BMO}^m$ and for each multi-index α, we have

$$\|[T, \boldsymbol{b}]_\alpha \boldsymbol{f}\|_{L^p(\nu_{\boldsymbol{w}})} \lesssim \alpha!\, \phi\left(c_{\boldsymbol{P},\boldsymbol{R}} [\boldsymbol{w}]_{A_{\boldsymbol{P},\boldsymbol{R}}}\right) [\boldsymbol{w}]_{A_{\boldsymbol{P},\boldsymbol{R}}}^{|\alpha| \max\{\Delta_1, \dots, \Delta_{m+1}\}} \prod_{j=1}^{m} \|b_j\|_{\mathcal{BMO}}^{\alpha_j} \|f_j\|_{L^{p_j}(w_j)}, \tag{A.4}$$

where $c_{\boldsymbol{P},\boldsymbol{R}} = 2^{\frac{r}{1-r} + 2\sum_{j=1}^{m} \min\{p_j/\Delta_j, p/\Delta_{m+1}\}/p_j}$.

Proof. The proof follows the ideas of Theorem 4.13 and we only consider the bilinear case. Without loss of generality, we assume that b_1, b_2 are real valued and normalized so that their $\mathcal{BMO}$ norms are equal to 1. As before, we are going to use the Cauchy integral trick, and given $\mathbf{w} \in A_{\mathbf{P},\mathbf{R}}$ everything reduces to showing that for some appropriate $\delta_1, \delta_2 > 0$ (to be chosen later) and for $|z_1| = \delta_1$, $|z_2| = \delta_2$, we have

$$\widetilde{\mathbf{w}} := (\widetilde{w}_1, \widetilde{w}_2) := (w_1 e_{b_1}, w_2 e_{b_2}) := (w_1 e^{-\mathrm{Re}(z_1) p_1 b_1}, w_2 e^{-\mathrm{Re}(z_2) p_2 b_2}) \in A_{\mathbf{P},\mathbf{R}}.$$

By [34, Lemma 5.3(i)] it follows that $\nu_{\mathbf{w}}^{\Delta_3/p} \in A_{\frac{1-r}{r}\Delta_3}$ and $w_j^{\theta_j/p_j} \in A_{\frac{1-r}{r}\theta_j}$ or, equivalently, $w_j^{-\Delta_j/p_j} \in A_{\frac{1-r}{r}\Delta_j}$ for $j = 1, 2$, where

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} > \frac{1}{p_1} + \frac{1}{p_2} + 1 - \frac{1}{p} = 1; \qquad \frac{1}{\theta_j} = \frac{1-r}{r} - \frac{1}{\Delta_j}, \quad j = 1, 2.$$

Moreover,

$$\left[\nu_{\mathbf{w}}^{\Delta_3/p}\right]_{A_{\frac{1-r}{r}\Delta_3}} \le [\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}^{\Delta_3}; \qquad \left[w_j^{-\Delta_j/p_j}\right]_{A_{\frac{1-r}{r}\Delta_j}} = \left[w_j^{\theta_j/p_j}\right]_{A_{\frac{1-r}{r}\theta_j}}^{\frac{1-r}{r}\Delta_j-1} \le [\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}^{\Delta_j}, \quad j=1,2.$$

Using now Lemma 3.26 and writing, for a given weight w, $\rho(w)$ instead of ρ_w, we can find $\varrho = \varrho(\mathbf{w}) = \min\{\rho(\nu_{\mathbf{w}}^{\Delta_3/p}), \rho(w_1^{-\Delta_1/p_1}), \rho(w_2^{-\Delta_2/p_2})\} > 1$ so that

$$\varrho' \sim \max\left\{\left[\nu_{\mathbf{w}}^{\Delta_3/p}\right]_{A_{\frac{1-r}{r}\Delta_3}}, \left[w_1^{-\Delta_1/p_1}\right]_{A_{\frac{1-r}{r}\Delta_1}}, \left[w_2^{-\Delta_2/p_2}\right]_{A_{\frac{1-r}{r}\Delta_2}}\right\} \le [\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}^{\max\{\Delta_1,\Delta_2,\Delta_3\}} \tag{A.5}$$

and the following reverse Hölder inequalities hold:

$$\left(\fint_Q \nu_{\mathbf{w}}^{\Delta_3\varrho/p}dx\right)^{1/\varrho} \le 2\fint_Q \nu_{\mathbf{w}}^{\Delta_3/p}dx \tag{A.6}$$

and, for $j=1,2$,

$$\left(\fint_Q w_j^{-\Delta_j\varrho/p_j}\,dx\right)^{1/\varrho} \le 2\fint_Q w_j^{-\Delta_j/p_j}\,dx. \tag{A.7}$$

Using all these and regrouping terms, we get

$$\begin{aligned}
&\Big(\fint_Q \nu_{\widetilde{\mathbf{w}}}^{\Delta_3/p}dx\Big)^{1/\Delta_3}\prod_{j=1}^2\Big(\fint_Q \widetilde{w}_j^{-\Delta_i/p_i}dx\Big)^{1/\delta_j}\\
&\quad= \left(\fint_Q \nu_{\mathbf{w}}^{\Delta_3/p}e_{b_1}^{\Delta_3/p_1}e_{b_1}^{\Delta_3/p_2}dx\right)^{1/\Delta_3}\prod_{j=1}^2\Big(\fint_Q w_j^{-\Delta_j/p_j}e_{b_j}^{-\Delta_j/s_j}dx\Big)^{1/\Delta_j}\\
&\quad\le \left(\fint_Q \nu_{\mathbf{w}}^{\Delta_3\varrho/p}dx\right)^{1/(\Delta_3\varrho)}\prod_{j=1}^2\Big(\fint_Q w_j^{-\Delta_j\varrho/p_j}dx\Big)^{1/(\Delta_j\varrho)}\\
&\qquad\left(\fint_Q \prod_{j=1}^2 e_{b_j}^{\Delta_3\varrho'/p_j}dx\right)^{1/(\Delta_3\varrho')}\prod_{j=1}^2\Big(\fint_Q e_{b_j}^{-\Delta_j\varrho'/p_j}dx\Big)^{1/(\Delta_j\varrho')}\\
&\quad\le 2^{\frac{1-r}{r}}[\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}\prod_{j=1}^2\left(\fint_Q e_{b_j}^{\Delta_3\varrho'/p}dx\right)^{p/(\Delta_3\varrho' p_j)}\Big(\fint_Q e_{b_j}^{-\Delta_j\varrho'/p_j}dx\Big)^{1/(\Delta_j\varrho')}\\
&\quad\le 2^{\frac{1-r}{r}}[\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}\prod_{j=1}^2\left[e_{b_j}^{\Delta_3\varrho'/p}\right]_{A_{1+\Delta_3 p_j/(\Delta_j p)}}^{p/(\Delta_3\varrho' p_j)}\\
&\quad\le 2^{\frac{1-r}{r}+2(\delta_1+\delta_2)}[\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}},
\end{aligned}$$

where the last estimate holds by Lemma 3.5 provided

$$\delta_j \le \frac{1}{\varrho' p_j}\min\left\{\frac{p_j}{\Delta_j}, \frac{p}{\Delta_3}\right\}.$$

Notice that this choice implies that $\delta_j \le \min\{p_j/\Delta_j, p/\Delta_3\}/p_j$. On the other hand, recalling (A.5) and assuming further that $\delta_j \sim [\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}^{\max\{\Delta_1,\Delta_2,\Delta_3\}}$ we eventually obtain

$$\|[T,\mathbf{b}]_\alpha \mathbf{f}\|_{L^p(\nu_{\mathbf{w}})} \lesssim \alpha! \delta_1^{-\alpha_1}\delta_2^{-\alpha_2}\phi\left(c_{\mathbf{P},\mathbf{R}}[\mathbf{w}]_{A_{\mathbf{P},\mathbf{R}}}\right)\|f_1\|_{L^{p_1}(w_1)}\|f_2\|_{L^{p_2}(w_2)}.$$

with $c_{\mathbf{P},\mathbf{R}} = 2^{\frac{r}{1-r}+2\min\{p_1/\Delta_1,p/\Delta_3\}/p_1+2\min\{p_2/\Delta_2,p/\Delta_3\}/p_2}$. This easily gives the desired estimate. □

Using Theorem A.2 and [17, Theorem 3] we can immediately obtain the following weighted estimates for the commutators of the BHT (note that, as before, one can remove the restriction $p > 1$ by using extrapolation, see [34, Corollary 2.26]):

Corollary A.8. *Let $\boldsymbol{R} = (r_1, r_2, r_3)$ be such that $1 < r_1, r_2, r_3 < \infty$ and*

$$\frac{1}{\min\{r_1,2\}} + \frac{1}{\min\{r_2,2\}} + \frac{1}{\min\{r_3,2\}} < 2. \tag{A.9}$$

Let $\boldsymbol{P} = (p_1, p_2)$ be such that $r_1 < p_1 < \infty$, $r_2 < p_2 < \infty$, and $1 < p < r_3'$ where $1/p = 1/p_1 + 1/p_2$. For every $\boldsymbol{b} \in \mathcal{BMO}^2$, $\boldsymbol{w} = (w_1, w_2) \in A_{\boldsymbol{P},\boldsymbol{R}}$ and any multi-index α we have

$$[BHT, \boldsymbol{b}]_\alpha : L^{p_1}(w_1) \times L^{p_2}(w_2) \to L^p(\nu_{\boldsymbol{w}}).$$

References

[1] J. Alvarez, R. Bagby, D. S. Kurtz, and C. Pérez. *Weighted estimates for commutators of linear operators.* Studia Math. 104(21): 195–209, 1993. 2, 10

[2] P. Auscher. *On necessary and sufficient conditions for L^p-estimates of Riesz transforms associated to elliptic operators on $\mathbb{R}^n$ and related estimates.* Mem. Amer. Math. Soc. 186(871): xviii+75, 2007. 15, 25

[3] P. Auscher and J. M. Martell. *Weighted norm Inequalities, off-diagonal estimates, and elliptic operators. Part I: General Operator Theory and weights.* Adv. Math. 212: 225–276, 2007. 25

[4] P. Auscher and J. M. Martell. *Weighted norm inequalities, off-diagonal estimates and elliptic operators. III. Harmonic analysis of elliptic operators.* J. Funct. Anal. 241(2):703–746, 2006. 15, 25

[5] P. Auscher and J. M. Martell. *Weighted norm inequalities for fractional operators*. Indiana Univ. Math. J. 57: 1845–1869, 2008. 26

[6] F. Bernicot, D. Frey, and S. Petermichl. *Sharp weighted norm estimates beyond Calderón-Zygmund theory*. Anal. PDE, 9(5), 1079–1113, 2016. 25

[7] L. Chaffee, R.H. Torres, and X. Wu. *Multilinear weighted norm inequalities under integral type regularity conditions.* Harmonic analysis, partial differential equations and applications, 193–216, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, 2017. 32

[8] S. Chanillo. *A note on commutators.* Indiana Univ. Math. J. 31: 7–16, 1982. 24

[9] X. Chen and Q. Xue. *Weighted estimates for a class of multilinear fractional type operators.* J. Math. Anal. Appl. 362(2): 355–373, 2010. 27

[10] D. Chung, M. C. Pereyra, and C. Pérez. *Sharp bounds for general commutators on weighted Lebesgue spaces.* Trans. Amer. Math. Soc. 364(3): 1163–1177, 2012. 2, 23

[11] G. Citti, L. Grafakos, C. Pérez, A. Sarti, and X. Zhong. *Harmonic and Geometric Analysis*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser, Basel, 2014. 13

[12] R. R. Coifman, R. Rochberg, and G. Weiss. *Factorization theorems for Hardy spaces in several variables.* Ann. of Math. 103: 611–635, 1976. 2, 8, 23

[13] M. Cotlar and C. Sadosky. *Two distinguished subspaces of product BMO and Nehari-AAK theory for Hankel operators on the torus.* Integral Equations Operator Theory 26(3): 273–304, 1996. 30, 31

[14] D. V. Cruz-Uribe and J. M. Martell. *Limited range multilinear extrapolation with applications to the bilinear Hilbert transform.* To appear in Math. Ann., http://doi.org/10.1007/s00208-018-1640-9. 5, 28

[15] D. V. Cruz-Uribe, J. M. Martell, and C. Pérez. *Weights, Extrapolation, and the Theory of Rubio de Francia*, volume 215. Birkhäuser, 2010. 4

[16] D. V. Cruz-Uribe and K. Moen. *Sharp norm inequalities for commutators of classical operators*. Publ. Mat. 56(1): 147–190, 2012. 24

[17] A. Culiuc, F. Di Plinio, and Y. Ou. *Domination of multilinear singular integrals by positive sparse forms.* arXiv:1603.05317. 28, 34

[18] W. Damián, A. K. Lerner, and C. Pérez. *Sharp weighted bounds for multilinear maximal functions and Calderón-Zygmund operators.* J. Fourier Anal. Appl. 21(1): 161-181, 2015. 21

[19] R. Fefferman and E.M. Stein. *Singular integrals on product spaces.* Adv. Math. 45: 117-143, 1982. 31

[20] S. H. Ferguson and C. Sadosky. *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures.* J. Anal. Math. 81: 239–267, 2000. 31

[21] J. Garcia-Cuerva and J. L. Rubio de Francia. *Weighted Norm Inequalities and Related Topics*, volume 116. North Holland Mathematics Studies, 1985. 6, 29, 30

[22] L. Grafakos. *Classical and Modern Fourier Analysis* Pearson/Prentice Hall, 2004. 3

[23] L. Grafakos and R.H. Torres. *Multilinear Calderón-Zygmund theory.* Adv. Math. 165: 124–164, 2002. 26

[24] S. Hofmann, M. Mitrea, and M. Taylor. *Singular integrals and elliptic boundary problems on regular Semmes-Kenig-Toro domains.* Int. Math. Res. Not. 2010(14): 2567–2865, 2010. 2

[25] T. Hytönen. *The sharp weighted bound for general Calderón-Zygmund operators*. Ann. of Math., 175: 1473–1506, 2012. 23

[26] T. Hytönen and C. Pérez. *Sharp weighted bounds involving A_∞*. Anal. PDE, 6(4): 777-818, 2013. 13, 14

[27] T. Hytönen, C. Pérez, and E. Rela. *Sharp reverse Hölder property for A_∞ weights on spaces of homogeneous type.* J. Funct. Anal., 263(12): 3883-3899, 2012. 13, 14

[28] R. Johnson and C. J. Neugebauer. *Change of variable results for A_p and reverse Hölder RH_r - classes*. Trans. Amer. Math. Soc. 328: 639–666, 1991. 6

[29] I. Kunwar and Y. Ou. *Two-weight inequalities for multilinear commutators*. arXiv:1710.07392. 23

[30] M. Lacey, K. Moen, C. Pérez, and R.H. Torres. *Sharp weighted bounds for fractional integral operators.* J. Func. Anal. 259: 1073–1097, 2010. 24

[31] M. Lacey and C. Thiele. *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$.* Ann. of Math. 146(2): 693-724, 1997. 4

[32] M. Lacey and C. Thiele. *On Calderón's conjecture.* Ann. of Math. 149(2): 475-496, 1999. 4

[33] A. K. Lerner, S. Ombrosi, C. Pérez, R.H. Torres, and R. Trujillo-González. *New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory.* Adv. Math. 220: 1222–1264, 2009. 17, 21, 27

[34] K. Li, J.M. Martell, and S. Ombrosi. *Extrapolation for multilinear Muckenhoupt classes and applications to the bilinear Hilbert transform.* arXiv:1802.03338. 27, 31, 32, 34

[35] K. Li, K. Moen, and W. Sun. *Sharp weighted bounds for multilinear maximal functions and Calderón-Zygmund operators.* J. Fourier Anal. Appl. 20(4): 751–765, 2014. 27

[36] K. Li, K. Moen, and W. Sun. *Sharp weighted inequalities for multilinear fractional maximal operators and fractional integrals.* Math. Nachr. 288(5-6): 619–632, 2015. 27

[37] K. Moen. *Weighted inequalities for multilinear fractional integral operators.* Collect. Math. 60(2): 213–238, 2009. 27

[38] B. Muckenhoupt and R. Wheeden. *Weighted norm inequalities for fractional integrals.* Trans. Amer. Math. Soc. 192: 261–274, 1974. 24

[39] Y. Ou, S. Petermichl, and E. Strouse, *Higher order Journé commutators and characterizations of multi-parameter BMO*, Adv. Math. 291 (2016), 24-58. 31

[40] M. C. Pereyra. *Weighted inequalities and dyadic harmonic analysis.* Excursions in Harmonic Analysis. Volume 2, 281–306, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, 2013. 8

[41] C. Pérez, G. Pradolini, R.H. Torres, and R. Trujillo-González. *End-point estimates for iterated commutators of multilinear singular integrals.* Bull. Lond. Math. Soc. 46(1): 26–42, 2014. 27

[42] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, 1970. 24

Árpád Bényi, Department of Mathematics, 516 High St, Western Washington University, Bellingham, WA 98225, USA

Email address: arpad.benyi@wwu.edu

José María Martell, Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM, Consejo Superior de Investigaciones Científicas, C/ Nicolás Cabrera, 13-15, E-28049 Madrid, Spain

Email address: chema.martell@icmat.es

Kabe Moen, Department of Mathematics, University of Alabama, P.O. Box 870350, Tuscaloosa, AL 35487, USA

Email address: kabe.moen@ua.edu

Eric Stachura, Department of Mathematics, Kennesaw State University, 850 Polytechnic Lane, Marietta, GA 30060, USA

Email address: estachur@kennesaw.edu

Rodolfo H. Torres, Department of Mathematics, University of Kansas, 1460 Jayhawk Blvd, Lawrence, Kansas 66045-7523, USA

Email address: torres@ku.edu