

END-POINT ESTIMATES, EXTRAPOLATION FOR MULTILINEAR MUCKENHOUT CLASSES, AND APPLICATIONS

KANGWEI LI, JOSÉ MARÍA MARTELL, HENRI MARTIKAINEN, SHELDY OMBROSI,
AND EMIL VUORINEN

ABSTRACT. In this paper we present the results announced in the recent work by the first, second, and fourth authors of the current paper concerning Rubio de Francia extrapolation for the so-called multilinear Muckenhoupt classes. Here we consider the situations where some of the exponents of the Lebesgue spaces appearing in the hypotheses and/or in the conclusion can be possibly infinity. The scheme we follow is similar, but, in doing so, we need to develop a one-variable end-point off-diagonal extrapolation result. This complements the corresponding “finite” case obtained by Duoandikoetxea, which was one of the main tools in the aforementioned paper. The second goal of this paper is to present some applications. For example, we obtain the full range of mixed-norm estimates for tensor products of bilinear Calderón-Zygmund operators with a proof based on extrapolation and on some estimates with weights in some mixed-norm classes. The same occurs with the multilinear Calderón-Zygmund operators, the bilinear Hilbert transform, and the corresponding commutators with BMO functions. Extrapolation along with the already established weighted norm inequalities easily give scalar and vector-valued inequalities with multilinear weights and these include the end-point cases.

1. INTRODUCTION

Recently, the first, second, and fourth authors of the present paper solved in [28] a long standing problem about the extrapolation theorem for multilinear Muckenhoupt classes of weights. A particular and simplified version of the general result established there is the following: suppose that a multivariable operator T satisfies

$$\|T(f_1, \dots, f_m)w\|_{L^p} \lesssim \|f_1w_1\|_{L^{p_1}} \dots \|f_mw_m\|_{L^{p_m}}$$

for some fixed $1 \leq p_1, \dots, p_m < \infty$, and for all $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}}$ (see Section 2), where $\frac{1}{p} = \frac{1}{p_1} + \dots + \frac{1}{p_m}$ and $w = \prod_{i=1}^m w_i$. Then, one gets the same kind of estimates for all $1 < p_1, \dots, p_m < \infty$, and for all $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}}$. The one-variable case (i.e., $m = 1$) is the well-known Rubio de Francia extrapolation

Date: February 12, 2019. *Revised:* March 8, 2020.

2010 Mathematics Subject Classification. 42B20, 42B25, 42B35.

Key words and phrases. Multilinear Muckenhoupt weights, Rubio de Francia extrapolation, multilinear Calderón-Zygmund operators, bilinear Hilbert transform, vector-valued inequalities, mixed-norm estimates.

theorem [33] of which one can find a great deal of extensions and refinements which are adapted to various settings and situations (see [12]).

The first result in the multivariable case [22], due to the second author of this paper and Grafakos, was obtained under the assumption that every individual weight is in the corresponding Muckenhoupt class. This collection of weights is however too small and the paper [27], coauthored by the fourth author of this paper together with Lerner, Pérez, Torres, and Trujillo-González, introduced and studied the classes $A_{\vec{p}}$. In these, rather than looking at the weights in the m -tuple individually, it is assumed that they collectively satisfy a Muckenhoupt-type condition. These are in turn the natural classes for the multilinear Calderón-Zygmund operators in the same way that the classical A_p classes are the natural ones for the linear Calderón-Zygmund operators.

Since the appearance of [27], an associated Rubio de Francia extrapolation theory for the classes $A_{\vec{p}}$ has been sought until recently when the aforementioned paper [28] solved that problem, and even more, obtained extrapolation results for the more general classes $A_{\vec{p}, \vec{r}}$ (see Section 2). That paper, however, considered only the cases in which all the exponents are finite and it was announced that one could consider situations in which some of the exponents p_i 's and/or the q_i 's are infinity. While these end-point scenarios are not natural in the one-variable extrapolation theory, it turns out that for some relevant multilinear applications it is of interest to extend the extrapolation theory so that it includes the end-point cases. The main goal of this paper is to present in a rigorous way these end-point extrapolation results originally announced in [28]. Since the statement of our main extrapolation result for the classes $A_{\vec{p}, \vec{r}}$ (see Theorem 2.1) requires some notation, we postpone that until Section 2. However, we would like to single out the particular case of the original classes $A_{\vec{p}}$, which is of particular interest and also gives an-easy-to-digest illustration of the general case.

Theorem 1.1. *Let $m \geq 2$ and T be an m -variable operator. Given $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$, let $\frac{1}{p} = \frac{1}{p_1} + \dots + \frac{1}{p_m} \in [0, m]$. Assume that given any $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}}$ the inequality*

$$\|T(f_1, \dots, f_m)w\|_{L^p} \lesssim \prod_{i=1}^m \|f_i w_i\|_{L^{p_i}}$$

holds for every $(f_1, \dots, f_m) \in \mathbb{F}$ (a fixed collection of m -tuples of measurable functions), where $w := \prod_{i=1}^m w_i$. Then for all exponents $\vec{q} = (q_1, \dots, q_m)$, with $1 < q_1, \dots, q_m \leq \infty$ and $\frac{1}{q} = \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$, and for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}}$ the inequality

$$\|T(f_1, \dots, f_m)v\|_{L^q} \lesssim \prod_{i=1}^m \|f_i v_i\|_{L^{q_i}}$$

holds for every $(f_1, \dots, f_m) \in \mathbb{F}$, where $v := \prod_{i=1}^m v_i$.

Moreover, for the same family of exponents and weights, and for all exponents $\vec{s} = (s_1, \dots, s_m)$ with $1 < s_1, \dots, s_m \leq \infty$ and $\frac{1}{s} = \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$ the inequality

$$(1.2) \quad \left\| \{T(f_1^j, \dots, f_m^j)v\}_j \right\|_{L_{\ell^s}^q} \lesssim \prod_{i=1}^m \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^{s_i}}^{q_i}}$$

holds for all sequences $\{(f_1^j, \dots, f_m^j)\}_j \subset \mathbb{F}$.

As mentioned, this result is a particular case of Theorem 2.1 and extends [28, Corollary 1.5] where all the p_i 's, q_i 's, s_i 's and p are assumed to be finite. Note that Theorem 1.1 allows us to extrapolate starting with some (or even all) p_i 's being infinity. Not only that, even if we start with all the p_i 's being finite, we can derive estimates with some (but not all) of the q_i 's being infinity.

The statement of Theorem 2.1 and its proof are given in Section 2. For the latter we follow the blueprint established in [28] with very minor changes. However, a key ingredient in the proof is Theorem 2.3, which is a one-variable off-diagonal result of independent interest. This is indeed an extension of [17, Theorem 5.1] (and also of [24]) allowing us to start with or to obtain end-point estimates. Our proof is, however, slightly different and follows the scheme used thoroughly in [12].

Our second goal of this paper is to present some applications of our extrapolation result. In Section 3 we briefly consider some of the examples already treated in [28] and these include multilinear Calderón-Zygmund operators, multilinear sparse forms, bilinear rough singular integral operators, the bilinear Hilbert transform and their commutators with BMO functions. In Section 4, we present a new application which concerns mixed-norm estimates (a topical subject, see e.g. [2, 3] and [15]). Here we work with tensor products $T_n \otimes T_m$, where T_n and T_m are bilinear Calderón-Zygmund operators in \mathbb{R}^n and \mathbb{R}^m . The tensor product is initially defined via

$$(T_n \otimes T_m)(f_1 \otimes f_2, g_1 \otimes g_2)(x) := T_n(f_1, g_1)(x_1)T_m(f_2, g_2)(x_2),$$

where $f_1, g_1: \mathbb{R}^n \rightarrow \mathbb{C}$, $f_2, g_2: \mathbb{R}^m \rightarrow \mathbb{C}$ and $x = (x_1, x_2) \in \mathbb{R}^{n+m}$. Notice that if T_n and T_m are linear, then $T_n \otimes T_m = T_n^1 T_m^2$, where $T_n^1 f(x) = T_n(f(\cdot, x_2))(x_1)$ and $T_m^2 f(x) = T_m(f(x_1, \cdot))(x_2)$. One can then easily obtain the desired estimates by using iterative arguments and Fubini's theorem, and therefore, in the linear case, bi-parameter singular integrals are only interesting if they are not of tensor product type. Unlike the linear case, the theory of tensor products of bilinear operators is already non-trivial. Indeed, Journé in [25] showed that a tensor product of general bilinear operators, both bounded from $L^\infty \times L^2$ to L^2 , needs not be bounded from $L^\infty \times L^2$ to L^2 . On the other hand, he obtained positive results for tensor products of some multilinear singular integral forms from Christ-Journé [10].

As an application of our extrapolation results we are able to obtain (see Corollary 4.3) that

$$\|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} \lesssim \|f_1\|_{L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))} \|f_2\|_{L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}$$

whenever $p_1, p_2, q_1, q_2 \in (1, \infty]$ are such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$. This should be compared with [30] where the first, third, and last authors

of the present paper considered general bilinear bi-parameter Calderón-Zygmund operators (with no tensor form). Some mixed-norm estimates were proved as quick corollaries from the obtained weighted estimates (and operator-valued analysis). However, since the classes of weights considered were of product type $A_{p_1} \times A_{p_2}$ (in place of the corresponding bilinear class $A_{(p_1, p_2)}$), in the mixed-norm estimates one needed to assume that for the case $q < 1$ one has $p_1, p_2 < \infty$. Here we are able to obtain more sophisticated weighted estimates and this restriction can be removed. It should be noticed that here we only work with tensor products of bilinear singular integrals, but we can show somewhat stronger weighted estimates compared to those in [30]: we essentially use weights which are in $A_{(p_1, p_2)}(\mathbb{R}^n)$ and in $A_{p_1}(\mathbb{R}^m) \times A_{p_2}(\mathbb{R}^m)$ and, by extrapolation, this is enough for the mixed-norm inequalities in question.

To conclude with this introduction we would like to mention that some interesting related work has recently appeared while this manuscript was in preparation. Nieraeth in [32] has obtained a result similar to ours with a proof which uses an independent different method. In a nutshell, Theorem 2.1—which was announced in [28] before [32] was posted—is proved by following [28, Proof of Theorem 1.1] with some appropriate changes both in the argument and in the notation. Having said that, the main tool that we need here is Theorem 2.3, which extends [17, Theorem 5.1] (and also [24]) to allow for end-point estimates. We also note that our proof of Theorem 2.3 is somewhat different, and of independent interest, than that in [17] and follows the scheme thoroughly employed in [12, Chapter 3]. On the other hand Duoandikoetxea and Oruetebarria in [19] have obtained mixed-norm estimates of radial-angular type by developing an extrapolation theory for radial weights.

Acknowledgments. K. Li was supported by Juan de la Cierva - Formación 2015 FJCI-2015-24547, by the Basque Government through the BERC 2018-2021 program and by Spanish Ministry of Economy and Competitiveness MINECO through BCAM Severo Ochoa excellence accreditation SEV-2017-0718 and through project MTM2017-82160-C2-1-P funded by (AEI/FEDER, UE) and acronym “HAQMEC”.

J.M. Martell was supported by the Spanish Ministry of Economy and Competitiveness, through the “Severo Ochoa Programme for Centres of Excellence in R&D” (SEV-2015-0554) and the European Research Council through the European Union’s Seventh Framework Programme (FP7/2007-2013)/ ERC agreement no. 615112 HAPDEGMT.

H. Martikainen was supported by the Academy of Finland through the grants 294840 and 306901, and by the three-year research grant 75160010 of the University of Helsinki. He is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research.

S. Ombrosi was supported by CONICET PIP 11220130100329CO, Argentina.

E. Vuorinen was supported by the Jenny and Antti Wihuri Foundation.

The authors would like to thank the anonymous referee for the useful suggestions that helped us to improve the presentation of the paper.

2. END-POINT EXTRAPOLATION FOR MULTILINEAR MUCKENHOUT CLASSSES

Our main result in this section contains the extension to the end-point cases announced in our previous paper [28]. Before we state that result, let us recall some notations and make some conventions. Given a cube Q , its side-length will be denoted by $\ell(Q)$ and for any $\lambda > 0$ we let λQ be the cube concentric with Q whose side-length is $\lambda\ell(Q)$. Let μ be a doubling measure on \mathbb{R}^n , that is, μ is a non-negative Borel regular measure such that $\mu(2Q) \leq C_\mu \mu(Q) < \infty$ for every cube $Q \subset \mathbb{R}^n$. Given a Borel set $E \subset \mathbb{R}^n$ with $0 < \mu(E) < \infty$ we use the notation

$$\int_E f d\mu = \frac{1}{\mu(E)} \int_E f d\mu.$$

This section is devoted to applying off-diagonal extrapolation to get the multi-variable extrapolation in the end-point cases.

Hereafter, $m \geq 2$. Given $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$ and $\vec{r} = (r_1, \dots, r_{m+1})$ with $1 \leq r_1, \dots, r_{m+1} < \infty$, we say that $\vec{r} \preceq_\star \vec{p}$ whenever

$$r_i \leq p_i, \quad i = 1, \dots, m; \quad \text{and} \quad r'_{m+1} \geq p, \quad \text{where} \quad \frac{1}{p} := \frac{1}{p_1} + \dots + \frac{1}{p_m}.$$

This should be compared with the notation $\vec{r} \preceq \vec{p}$ from [28] where one adds the restriction $r'_{m+1} > p$. Analogously, we say that $\vec{r} \prec \vec{p}$ if $\vec{r} \preceq_\star \vec{p}$ and moreover $r_i < p_i$ for every $i = 1, \dots, m$ and $r'_{m+1} > p$. Notice that the fact that $\vec{r} \preceq_\star \vec{p}$ forces that $\sum_{i=1}^{m+1} \frac{1}{r_i} \geq 1$ and also $\frac{1}{p} \leq \sum_{i=1}^m \frac{1}{r_i}$. Hence, if $\sum_{i=1}^m \frac{1}{r_i} > 1$ then we allow p to be smaller than one.

We can now introduce the classes of multilinear Muckenhoupt weights that we consider in the present paper. Given $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$ and $\vec{r} = (r_1, \dots, r_{m+1})$ with $1 \leq r_1, \dots, r_{m+1} < \infty$ so that $\vec{r} \preceq_\star \vec{p}$ we say that $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}, \vec{r}}$, provided $0 < w_i < \infty$ a.e. for every $i = 1, \dots, m$ and

$$[\vec{w}]_{A_{\vec{p}, \vec{r}}} = \sup_Q \left(\int_Q w^{\frac{r'_{m+1} p}{r'_{m+1} - p}} dx \right)^{\frac{1}{p} - \frac{1}{r'_{m+1}}} \prod_{i=1}^m \left(\int_Q w_i^{\frac{r_i p_i}{r_i - p_i}} dx \right)^{\frac{1}{r_i} - \frac{1}{p_i}} < \infty,$$

where $w = \prod_{i=1}^m w_i$. When $p = r'_{m+1}$ the term corresponding to w needs to be replaced by $\text{ess sup}_Q w$ and, analogously, when $p_i = r_i$, the term corresponding to w_i should be $\text{ess sup}_Q w_i^{-1}$. Also, if $p_i = \infty$ the term corresponding to w_i becomes $\left(\int_Q w_i^{-r_i} dx \right)^{\frac{1}{r_i}}$. If $p = \infty$, one necessarily have $r_{m+1} = 1$ and $p_1 = \dots = p_m = \infty$, hence the term corresponding to w must be $\text{ess sup}_Q w$ while the terms corresponding to w_i become $\left(\int_Q w_i^{-r_i} dx \right)^{\frac{1}{r_i}}$. When $r_{m+1} = 1$ and $p < \infty$ the term corresponding to w needs to be replaced by $\left(\int_Q w^p dx \right)^{\frac{1}{p}}$. We remark that similar

classes of weights, to be precise when $r_1 = \dots = r_m = r \geq 1$, were introduced earlier in [7].

Note that with the previous definition $\vec{w} \in A_{\vec{p},(1,\dots,1)}$ means that

$$[\vec{w}]_{A_{\vec{p},(1,\dots,1)}} = \sup_Q \left(\int_Q w^p dx \right)^{\frac{1}{p}} \prod_{i=1}^m \left(\int_Q w_i^{-p'_i} dx \right)^{\frac{1}{p'_i}} < \infty,$$

where $w = \prod_{i=1}^m w_i$ and with the appropriate changes when some $p_i = 1$ or $p = \infty$. In the sequel we will just simply denote $A_{\vec{p},(1,\dots,1)}$ by $A_{\vec{p}}$. We would like to observe our definition of the classes $A_{\vec{p},\vec{r}}$ is slightly different to that in [27] or [28] (notice that e.g. $[w_1^{p_1}, \dots, w_m^{p_m}]_{A_{\vec{p}}}^{\frac{1}{p}}$ with $A_{\vec{p}}$ defined in [27] agrees with our $[\vec{w}]_{A_{\vec{p}}}$ when p_i 's are finite). This change is just cosmetic but it turns out to be useful for understanding the end-point estimates.

It is convenient to write the condition $A_{\vec{p},\vec{r}}$ in a different form. With that goal in mind, given $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$ and $\vec{r} = (r_1, \dots, r_{m+1})$ with $1 \leq r_1, \dots, r_{m+1} < \infty$ so that $\vec{r} \preceq_{\star} \vec{p}$ we set

$$\frac{1}{r} := \sum_{i=1}^{m+1} \frac{1}{r_i}, \quad \frac{1}{p_{m+1}} := 1 - \frac{1}{p}, \quad \text{and} \quad \frac{1}{\delta_i} := \frac{1}{r_i} - \frac{1}{p_i}, \quad i = 1, \dots, m+1.$$

Notice that as observed above we have that $0 < r \leq 1$ and formally $\frac{1}{p_{m+1}} = \frac{1}{p'}$ which could be negative or zero if $p \leq 1$. Note that in this way

$$\sum_{i=1}^{m+1} \frac{1}{p_i} = 1 \quad \text{and} \quad \sum_{i=1}^{m+1} \frac{1}{\delta_i} = \frac{1}{r} - 1 = \frac{1-r}{r}.$$

Also, $\vec{r} \preceq_{\star} \vec{p}$ means that $\delta_i^{-1} \geq 0$ for every $1 \leq i \leq m+1$. On the other hand, $\vec{r} \prec \vec{p}$ means that $\delta_i^{-1} > 0$ for every $1 \leq i \leq m+1$. Notice that with this notation $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p},\vec{r}}$ can be written as

$$[\vec{w}]_{A_{\vec{p},\vec{r}}} = \sup_Q \left(\int_Q w^{\delta_{m+1}} dx \right)^{\frac{1}{\delta_{m+1}}} \prod_{i=1}^m \left(\int_Q w_i^{-\delta_i} dx \right)^{\frac{1}{\delta_i}} < \infty,$$

where $w = \prod_{i=1}^m w_i$.

Again we shall use the abstract formalism of extrapolation families. Hereafter \mathcal{F} will denote a family of $(m+1)$ -tuples (f, f_1, \dots, f_m) of non-negative measurable functions.

Our main result is the following:

Theorem 2.1. *Let \mathcal{F} be a collection of $(m+1)$ -tuples of non-negative functions. Consider a vector $\vec{r} = (r_1, \dots, r_{m+1})$, with $1 \leq r_1, \dots, r_{m+1} < \infty$, and exponents $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$, so that $\vec{r} \preceq_{\star} \vec{p}$. Assume that given any $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p},\vec{r}}$ the inequality*

$$\|fw\|_{L^p} \leq \varphi([\vec{w}]_{A_{\vec{p},\vec{r}}}) \prod_{i=1}^m \|f_i w_i\|_{L^{p_i}}$$

holds for every $(f, f_1, \dots, f_m) \in \mathcal{F}$, where $\frac{1}{p} := \frac{1}{p_1} + \dots + \frac{1}{p_m}$, $w := \prod_{i=1}^m w_i$, and $\varphi \geq 0$ is a non-decreasing function. Then for all exponents $\vec{q} = (q_1, \dots, q_m)$, with $1 < q_1, \dots, q_m \leq \infty$, so that $\vec{r} \prec \vec{q}$, and for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}, \vec{r}}$ the inequality

$$\|fv\|_{L^q} \leq \tilde{\varphi}([\vec{v}]_{A_{\vec{q}, \vec{r}}}) \prod_{i=1}^m \|f_i v_i\|_{L^{q_i}}$$

holds for every $(f, f_1, \dots, f_m) \in \mathcal{F}$, where $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$, $v := \prod_{i=1}^m v_i$, and $\tilde{\varphi} \geq 0$ is a non-decreasing function.

Moreover, for the same family of exponents and weights, and for all exponents $\vec{s} = (s_1, \dots, s_m)$ with $1 < s_1, \dots, s_m \leq \infty$, so that $\vec{r} \prec \vec{s}$

$$(2.2) \quad \left\| \{f^j v\}_j \right\|_{L_{\ell^s}^q} \leq \tilde{\varphi}([\vec{v}]_{A_{\vec{q}, \vec{r}}}) \prod_{i=1}^m \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^{s_i}}^{q_i}}$$

for all $\{(f^j, f_1^j, \dots, f_m^j)\}_j \in \mathcal{F}$ and where $\frac{1}{s} := \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$.

We would like to remark that this result was announced in [28] and the main difference with [28, Theorem 1.1] is that here we allow the p_i 's and/or the q_i 's to take the value infinity. Also, in the current result we can start with $p = r'_{m+1}$ (including the case $p = \infty$ if $r_{m+1} = 1$) while in the conclusion we obtain $q < \infty$, since $\vec{r} \prec \vec{q}$. We also remark that if we start with $p_{i_0} = r_{i_0}$ for some given i_0 then in the conclusion we can relax $q_{i_0} > r_{i_0}$ to $q_{i_0} \geq r_{i_0}$, see [28, Remark 1.8].

2.1. End-point off-diagonal extrapolation theorem. In this section we present an end-point off-diagonal extrapolation result which is going to play an crucial role in the proof of Theorem 2.1. Next we give the basic properties of weights that we will need below. For proofs and further information, see [18, 20]. By a weight we mean a measurable function v such that $0 < v < \infty$ μ -a.e. For $1 < p < \infty$, we say that $v \in A_p(\mu)$ if

$$[v]_{A_p(\mu)} = \sup_Q \int_Q v \, d\mu \left(\int_Q v^{1-p'} \, d\mu \right)^{p-1} < \infty,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^n$. The quantity $[v]_{A_p(\mu)}$ is called the $A_p(\mu)$ constant of v . Notice that it follows at once from this definition that if $v \in A_p(\mu)$, then $v^{1-p'} \in A_{p'}(\mu)$. When $p = 1$ we say that $v \in A_1(\mu)$ if

$$[v]_{A_1(\mu)} = \sup_Q \left(\int_Q v \, d\mu \right) \operatorname{ess\,sup}_Q v^{-1} < \infty,$$

where the essential supremum is taken with respect to the underlying doubling measure μ . The $A_p(\mu)$ classes are properly nested: for $1 < p < q$, $A_1(\mu) \subsetneq A_p(\mu) \subsetneq A_q(\mu)$. We denote the union of all the $A_p(\mu)$ classes, $1 \leq p < \infty$, by $A_\infty(\mu)$.

Given $1 \leq p \leq \infty$ and $0 < r \leq \infty$ we say that $v \in A_{p,r}(\mu)$ if

$$[v]_{A_{p,r}(\mu)} = \sup_Q \left(\int_Q v^r d\mu \right)^{\frac{1}{r}} \left(\int_Q v^{-p'} d\mu \right)^{\frac{1}{p'}} < \infty,$$

where one has to replace the first term by $\text{ess sup}_Q v$ when $r = \infty$ and the second term by $\text{ess sup}_Q v^{-1}$ when $p = 1$, and these essential suprema are taking with respect to μ . The case $p = 1$ and $r = \infty$ gives a trivial class of weights since $v \in A_{1,\infty}$ amounts to say that $v \approx 1$ μ -a.e.. Assuming that we are not in that case we always have

$$\gamma_{p,r} := \frac{1}{r} + \frac{1}{p'} > 0$$

and one can easily see that $v \in A_{p,r}(\mu)$ if and only if $v^r \in A_{r\gamma_{p,r}}(\mu)$ if and only if $v^{-p'} \in A_{p'\gamma_{p,r}}$ with

$$[v]_{A_{p,r}(\mu)} = [v^r]_{A_{r\gamma_{p,r}}(\mu)}^{\frac{1}{r}} = [v^{-p'}]_{A_{p'\gamma_{p,r}}(\mu)}^{\frac{1}{p'}}$$

when $1 < p \leq \infty$ and $0 < r < \infty$. If $p = 1$ and $0 < r < \infty$ then $\gamma_{p,r} := \frac{1}{r}$ and $v \in A_{p,r}(\mu)$ if and only if $v^r \in A_1(\mu)$ with $[v]_{A_{p,r}(\mu)} = [v^r]_{A_1(\mu)}^{\frac{1}{r}}$. Also, if $1 < p \leq \infty$ and $r = \infty$ then $\gamma_{p,r} := \frac{1}{p'}$ and $v \in A_{p,r}(\mu)$ if and only if $v^{-p'} \in A_1(\mu)$ with $[v]_{A_{p,r}(\mu)} = [v^{-p'}]_{A_1(\mu)}^{\frac{1}{p'}}$. We remark that the current definition of $A_{p,r}$ constant is slightly different with the one that was used in [17] and [28], because we have to take care of the case $r = \infty$.

When μ is the Lebesgue measure we will simply write $A_p, A_{p,r}, \dots$. It is well-known that if $w \in A_\infty$ then $dw = w(x)dx$ is a doubling measure. Besides, since $0 < w < \infty$ a.e., then the Lebesgue measure and w have the same null measure sets hence the essential suprema and infima with respect to the Lebesgue measure and w agree.

We shall use the abstract formalism of extrapolation families. Hereafter \mathcal{F} will denote a family of pairs (f, g) of non-negative measurable functions. This approach to extrapolation has the advantage that, for instance, vector-valued inequalities are an immediate consequence of our extrapolation results. We will discuss applying this formalism to prove norm inequalities for specific operators below. For complete discussion of this approach to extrapolation in the linear setting, see [12].

The main result of this section is the following off-diagonal Rubio de Francia extrapolation result which extends [17, Theorem 5.1] (and also [24]) allowing us to start with or to obtain end-point results:

Theorem 2.3. *Let \mathcal{F} be a collection of pairs of non-negative functions. Let $1 \leq p_0 \leq \infty$, $0 < r_0 \leq \infty$, and $0 < q_0 \leq \infty$. Assume that for all $w \in A_{p_0, r_0}$ we have*

$$(2.4) \quad \|fw\|_{L^{q_0}} \leq \varphi([w]_{A_{p_0, r_0}}) \|gw\|_{L^{p_0}},$$

for all $(f, g) \in \mathcal{F}$ and where $\varphi \geq 0$ is a non-decreasing function. Then for all $1 < p \leq \infty$, $0 < r < \infty$, and $0 < q \leq \infty$ such that

$$(2.5) \quad \frac{1}{q} - \frac{1}{q_0} = \frac{1}{r} - \frac{1}{r_0} = \frac{1}{p} - \frac{1}{p_0},$$

and all $w \in A_{p,r}$ we have

$$(2.6) \quad \|fw\|_{L^q} \leq \tilde{\varphi}([w]_{A_{p,r}}) \|gw\|_{L^p},$$

for all $(f, g) \in \mathcal{F}$ and where $\tilde{\varphi} \geq 0$ is a non-decreasing function.

We observe that this result is on the nature of best possible in the sense that one cannot expect to reach the end-points $p = 1$ and/or $r = \infty$. The fact that one cannot obtain $p = 1$ is well-known. To see that we cannot extrapolate to $r = \infty$ we let $p_0 = q_0 = r_0 = 2$ (again $\gamma_{r_0, p_0} = 1 > 0$) and consider the pairs $(|Hf|, |f|)$ with $f \in L_c^\infty(\mathbb{R})$ and where H is the Hilbert transform. Note that (2.4) follows since the Hilbert transform is bounded on $L^2(w^2)$ for every $w^2 \in A_2$. If we could extrapolate to $r = \infty$, picking $p = q = r = \infty$ we would obtain $\|Hfw\|_{L^\infty} \lesssim \|fw\|_{L^\infty(w)}$ for every $w \in A_{\infty, \infty}$, that is, for any $w^{-1} \in A_1$. Taking in particular $w \equiv 1$ we would obtain that $Hf \in L^\infty(\mathbb{R}^n)$ for every $f \in L_c^\infty(\mathbb{R})$ leading contradiction.

Even more we can see that we cannot extrapolate to $r = \infty$ with $0 < q < \infty$. This requires some extra work. For any $f \in L_c^\infty(\mathbb{R})$, let E_f be the set of Lebesgue points for the function $|Hf|^2 \in L^1(\mathbb{R})$ so that $|\mathbb{R} \setminus E_f| = 0$. Define, for every $x \in \mathbb{R}$ and $0 < \tau < \infty$, the non-negative function $g_{x,\tau} = \tau^{-\frac{1}{2}} \mathbf{1}_{(x-\tau/2, x+\tau/2)}$ so that $\|g_{x,\tau}\|_{L^2(\mathbb{R})} = 1$ and consider the family

$$\mathcal{F} = \{(|Hf| g_{x,\tau}, |f|) : f \in L_c^\infty(\mathbb{R}), x \in E_f, 0 < \tau < \infty\}.$$

Let $q_0 = 1$ and $p_0 = r_0 = 2$, and note that $\gamma_{r_0, p_0} = 1 > 0$. For every $w \in A_{2,2}$ (that is, $w^2 \in A_2$) one has for every pair $(F, G) = (|Hf| g_{x,\tau}, f) \in \mathcal{F}$

$$\begin{aligned} \|Fw\|_{L^1(\mathbb{R})} &= \left\| |Hf| g_{x,\tau} w \right\|_{L^1(\mathbb{R})} \leq \|Hf\|_{L^2(w^2)} \|g_{x,\tau}\|_{L^2(\mathbb{R})} \\ &= \|Hf\|_{L^2(w^2)} \leq C_w \|f\|_{L^2(w^2)} = C_w \|G\|_{L^2(w^2)}. \end{aligned}$$

If we could extrapolate to $r = \infty$ then we would pick $p = \infty$, $q = 2$, and $r = \infty$ so that (2.5) holds to obtain that $\|Fw\|_{L^2(\mathbb{R})} \lesssim \|Gw\|_{L^\infty(\mathbb{R})}$ for all $(F, G) \in \mathcal{F}$ and $w \in A_{\infty, \infty}$, that is, $w^{-1} \in A_1$. We could take again $w \equiv 1$ to see that for every $(F, G) = (|Hf| g_{x,\tau}, f) \in \mathcal{F}$, that is, for every $f \in L_c^\infty$, $x \in E_f$ and $0 < \tau < \infty$

$$\left(\int_{(x_0-\tau/2, x_0+\tau/2)} |Hf(y)|^2 \right)^{\frac{1}{2}} = \|F\|_{L^2(\mathbb{R})} \leq C_0 \|G\|_{L^\infty(\mathbb{R})} = \|f\|_{L^\infty(\mathbb{R})}.$$

Since C_0 does not depend on f , x and τ (note that C is the same for all the pairs in the family \mathcal{F}), and $x \in E_f$ we could let $\tau \rightarrow 0^+$ and conclude that

$$|Hf(x)| \leq C_0 \|f\|_{L^\infty}, \quad \forall x \in E_f.$$

Consequently, we would obtain that $Hf \in L^\infty$ for every $f \in L_c^\infty$ which is again a contradiction.

Proof of Theorem 2.3. The case $1 \leq p_0 < \infty$, $0 < r_0 < \infty$, and $1 < p < \infty$ was proved in [17, Theorem 5.1]. Here we provide a somehow alternative argument which gives that case as well as the desired end-point estimates. Our first observation is that if $\gamma_{p_0, r_0} = 0$ there is nothing to prove. Indeed this only happens when $p_0 = 1$, $r_0 = \infty$ and hence (2.5) gives that $\frac{1}{p} = 1 + \frac{1}{r}$ which contradicts the facts that $p > 1$ and $r < \infty$. Thus, from now on we assume that $\gamma_{p_0, r_0} > 0$ and observe that (2.5) allows us to easily write $\gamma := \gamma_{p, r} = \gamma_{p_0, r_0}$.

Case 1: $\frac{1}{s} := \frac{1}{p} - \frac{1}{p_0} = \frac{1}{q} - \frac{1}{q_0} = \frac{1}{r} - \frac{1}{r_0} > 0$, hence $1 < p < p_0 \leq \infty$, $0 < q < q_0 \leq \infty$ and $0 < r < r_0 \leq \infty$ (note that p_0 and/or r_0 and/or q_0 could be infinity).

Fix $(f, g) \in \mathcal{F}$ and $w \in A_{p, r}$. As observed above, one has $w^{-p'} \in A_{p', \gamma}$. Also, $p'\gamma = 1 + \frac{p'}{r} > 1$ since $r < r_0$, hence if we set $M'h := M(hw^{-p'})w^{p'}$ where M is the Hardy-Littlewood maximal function one has that M' is bounded on $L^{(p'\gamma)'}(w^{-p'})$, write $\|M'\|$ to denote its operator norm and introduce the Rubio de Francia algorithm

$$\mathcal{R}'h = \sum_{k=0}^{\infty} \frac{(M')^{(k)}h}{2^k \|M'\|^k},$$

where for $k \geq 1$, we write $(M')^{(k)} = M' \circ \dots \circ M'$ to denote k iterations of M' and for $k = 0$ is the identity operator. Based on this definition (see [12] for more details) one can readily see that if $0 \leq h \in L^{(p'\gamma)'}(w^{-p'})$ then

$$(2.7) \quad h \leq \mathcal{R}'h, \quad \|\mathcal{R}'h\|_{L^{(p'\gamma)'}(w^{-p'})} \leq 2\|h\|_{L^{(p'\gamma)'}(w^{-p'})}, \quad [\mathcal{R}'h w^{-p'}]_{A_1} \leq 2\|M'\|.$$

Notice that in particular $0 < \mathcal{R}'h < \infty$ a.e. if h is non-trivial.

Next let us observe that without loss of generality we may assume that $0 < \|gw\|_{L^p} < \infty$: if $\|gw\|_{L^p} = \infty$ there would be nothing to prove and if $\|gw\|_{L^p} = 0$ then $g = 0$ a.e. and by (2.4) we would get that $f = 0$ a.e. which would trivially yield the desired estimate. This implies that the auxiliary function $0 \leq h = gw^{p'}/\|gw\|_{L^p}$ clearly satisfies $\|h\|_{L^p(w^{-p'})} = 1$. Set $H = \mathcal{R}'(h^{\frac{p}{(p'\gamma)'}})^{\frac{(p'\gamma)'}{p}} = \mathcal{R}'(h^{\frac{p}{r\gamma}})^{\frac{r\gamma}{p}}$ which satisfies that $0 < H < \infty$ a.e. and by (2.7)

$$(2.8) \quad h \leq H, \quad \|H\|_{L^p(w^{-p'})} \lesssim \|h\|_{L^p(w^{-p'})} = 1, \quad [H^{\frac{p}{r\gamma}} w^{-p'}]_{A_1} \leq 2\|M'\|.$$

Set $W = H^{-\frac{p}{s}} w^{1+\frac{p'}{s}}$ and we claim that $W \in A_{p_0, r_0}$. Assuming this momentarily we note that (2.5) and the definition of s gives $\frac{1}{q} = \frac{1}{q_0} + \frac{1}{s}$. Hence we can use Hölder's inequality, (2.4), and (2.8) to obtain

$$\begin{aligned} \|fw\|_{L^q} &= \|(fW)(H^{\frac{p}{s}} w^{-\frac{p'}{s}})\|_{L^q} \leq \|fW\|_{L^{q_0}} \|H^{\frac{p}{s}} w^{-\frac{p'}{s}}\|_{L^s} \lesssim \|gW\|_{L^{p_0}} \\ &\lesssim \|gw\|_{L^p} \|Hw^{-p'}W\|_{L^{p_0}} = \|gw\|_{L^p} \|H^{1-\frac{p}{s}} w^{-p'(\frac{1}{p}-\frac{1}{s})}\|_{L^{p_0}} \lesssim \|gw\|_{L^p}, \end{aligned}$$

where the last estimate is trivial when $p_0 = \infty$ since $s = p$, and otherwise follows from (2.8) and the fact that $1 - \frac{p}{s} = \frac{p}{p_0}$.

This gives the desired estimate and therefore it remains to show that $W = H^{-\frac{p}{s}} w^{1+\frac{p'}{s}} \in A_{p_0, r_0}$. If $r_0 = \infty$, one can easily see that (2.6) and the definition of s give that $s = r$ and

$$\gamma = 1 + \frac{1}{r} - \frac{1}{p} = 1 - \frac{1}{p_0} = \frac{1}{p'_0}, \quad 1 + \frac{p'}{s} = p' \left(\frac{1}{p'} + \frac{1}{s} \right) = p' \left(1 - \frac{1}{p_0} \right) = \frac{p'}{p'_0}.$$

Hence $W^{-p'_0} = H^{\frac{p}{r\gamma}} w^{-p'}$ and the last condition in (2.8) readily yields $W^{-p'_0} \in A_1$. That is, $W \in A_{p_0, \infty}$ as desired.

Consider next the case $0 < r_0 < \infty$ and note that definition of s implies that

$$1 + \frac{p'}{s} - \frac{p'r\gamma}{s} = 1 + \frac{p'}{s} - \frac{p' + r}{s} = \frac{1}{r} \left(\frac{1}{r} - \frac{1}{s} \right) = \frac{r}{r_0}.$$

On the other hand, (2.6) and the definition of s also yield $\gamma = \frac{1}{r_0} + \frac{1}{p'_0}$ and

$$1 - \frac{p'_0 r \gamma}{s} = r \left(\frac{1}{r} - \frac{p'_0 \gamma}{s} \right) = r \left(\frac{1}{r_0} - \frac{p'_0 \gamma - 1}{s} \right) = \frac{r p'_0}{r_0} \left(\frac{1}{p'_0} - \frac{1}{s} \right) = \frac{r p'_0}{r_0 p'}.$$

Note that the right-hand side is positive and strictly smaller than 1 since in this case $0 < r < r_0 < \infty$ and $1 < p < p_0$. As a consequence, $0 < \frac{p'_0 r \gamma}{s} < 1$ and $(\frac{s}{p'_0 r \gamma})' = \frac{r_0 p'_0}{r p'_0}$. We can then use Hölder's inequality and the previous calculations to conclude that

$$\begin{aligned} \left(\int_Q W^{-p'_0} dx \right)^{\frac{1}{p'_0}} &= \left(\int_Q (H^{\frac{p}{r\gamma}} w^{-p'})^{\frac{p'_0 r \gamma}{s}} w^{-\frac{p'_0 r}{r_0}} dx \right)^{\frac{1}{p'_0}} \\ &\leq \left(\int_Q H^{\frac{p}{r\gamma}} w^{-p'} dx \right)^{\frac{r\gamma}{s}} \left(\int_Q w^{-p'} dx \right)^{\frac{r}{r_0 p'}} \\ &\leq [H^{\frac{p}{r\gamma}} w^{-p'}]_{A_1}^{\frac{r\gamma}{s}} \operatorname{ess\,inf}_Q (H^{\frac{p}{r\gamma}} w^{-p'})^{\frac{r\gamma}{s}} \left(\int_Q w^{-p'} dx \right)^{\frac{r}{r_0 p'}}. \end{aligned}$$

This yields

$$\begin{aligned} \left(\int_Q W^{r_0} dx \right)^{\frac{1}{r_0}} \left(\int_Q W^{-p'_0} dx \right)^{\frac{1}{p'_0}} &\leq [H^{\frac{p}{r\gamma}} w^{-p'}]_{A_1}^{\frac{r\gamma}{s}} \left(\int_Q H^{\frac{r_0 p}{s}} w^{-\frac{p'_0 r_0 r \gamma}{s}} W^{r_0} dx \right)^{\frac{1}{r_0}} \left(\int_Q w^{-p'} dx \right)^{\frac{r}{r_0 p'}} \\ &= [H^{\frac{p}{r\gamma}} w^{-p'}]_{A_1}^{\frac{r\gamma}{s}} \left(\int_Q w^r dx \right)^{\frac{1}{r_0}} \left(\int_Q w^{-p'} dx \right)^{\frac{r}{r_0 p'}} \\ &\leq [H^{\frac{p}{r\gamma}} w^{-p'}]_{A_1}^{\frac{r\gamma}{s}} [w]_{A_{p,r}}^{\frac{r}{r_0}}. \end{aligned}$$

Using (2.8) we conclude as desired that $W \in A_{p_0, r_0}$ and the proof of the present case is complete.

Case 2: $0 < q < \infty$ and $\frac{1}{s} := \frac{1}{p_0} - \frac{1}{p} = \frac{1}{q_0} - \frac{1}{q} = \frac{1}{r_0} - \frac{1}{r} > 0$, hence $1 \leq p_0 < p \leq \infty$, $0 < q_0 < q < \infty$ and $0 < r_0 < r < \infty$ (note that p could be infinity).

Fix $(f, g) \in \mathcal{F}$ and $w \in A_{p,r}$. As observed above, one has $w^{-p'} \in A_{p'\gamma}$. Additionally, $p'\gamma = 1 + \frac{p'}{r} > 1$ since $r < \infty$, hence M is bounded on $L^{p'\gamma}(w^{-p'})$ and writing $\|M\|$ for its operator norm one can then introduce the Rubio de Francia algorithm

$$\mathcal{R}h = \sum_{k=0}^{\infty} \frac{M^{(k)}h}{2^k \|M\|^k}.$$

This (see [12] for more details) readily gives that if $0 \leq h \in L^{p'\gamma}(w^{-p'})$ then

$$(2.9) \quad h \leq \mathcal{R}h, \quad \|\mathcal{R}h\|_{L^{p'\gamma}(w^{-p'})} \leq 2\|h\|_{L^{p'\gamma}(w^{-p'})}, \quad [\mathcal{R}h]_{A_1} \leq 2\|M\|.$$

On the other hand, notice that the definition of s and (2.5) give $\frac{1}{q_0} = \frac{1}{q} + \frac{1}{s}$, hence $\frac{q}{q_0} = (\frac{s}{q_0})'$. Then by duality there exists $0 \leq h \in L^{\frac{s}{q_0}}(w^q)$ with $\|h\|_{L^{\frac{s}{q_0}}(w^q)} = 1$ such that

$$\|fw\|_{L^q} = \|f^{q_0}\|_{L^{\frac{q}{q_0}}(w^q)}^{\frac{1}{q_0}} = \left(\int f^{q_0} h w^q dx \right)^{\frac{1}{q_0}}.$$

Set $H = \mathcal{R}\left(h^{\frac{s}{p'q_0\gamma}} w^{\frac{p'+q}{p'\gamma}}\right)^{\frac{p'q_0\gamma}{s}} w^{-\frac{(p'+q)q_0}{s}}$ which satisfies by (2.9)

$$(2.10) \quad h \leq H, \quad \|H\|_{L^{\frac{s}{q_0}}(w^q)} \lesssim \|h\|_{L^{\frac{s}{q_0}}(w^q)} = 1, \quad [H^{\frac{s}{p'q_0\gamma}} w^{\frac{p'+q}{p'\gamma}}]_{A_1} \leq 2\|M\|.$$

Let $W = H^{\frac{1}{q_0}} w^{\frac{q}{q_0}}$ and claim that $W \in A_{p_0, r_0}$. Assuming this momentarily we conclude that

$$\begin{aligned} \|fw\|_{L^q} &= \left(\int f^{q_0} h w^q dx \right)^{\frac{1}{q_0}} = \|fW\|_{L^{q_0}} \lesssim \|gW\|_{L^{p_0}} = \|(gw)(H^{\frac{1}{q_0}} w^{\frac{q}{q_0}-1})\|_{L^{p_0}} \\ &\leq \|gw\|_{L^p} \|H^{\frac{1}{q_0}} w^{\frac{q}{q_0}-1}\|_{L^s} = \|gw\|_{L^p} \|H\|_{L^{\frac{s}{q_0}}(w^q)}^{\frac{1}{q_0}} \lesssim \|gw\|_{L^p}, \end{aligned}$$

where we have used (2.10), that $\frac{1}{p_0} = \frac{1}{p} + \frac{1}{s}$, Hölder's inequality and the fact that $s(\frac{q}{q_0} - 1) = q$ by (2.5) and our choice of s .

To complete this case we just need to see that $W = H^{\frac{1}{q_0}} w^{\frac{q}{q_0}} \in A_{p_0, r_0}$. When $p_0 = 1$ one has $\gamma = r_0^{-1}$, $s = p'$. Hence $\frac{s}{p'q_0\gamma} = \frac{r_0}{q_0}$, and (2.5) yield

$$\frac{p'+q}{p'\gamma} = r_0 q \left(\frac{1}{q} + \frac{1}{p'} \right) = \frac{p'+q}{p'\gamma} = \frac{r_0 q}{q_0},$$

As a result, $W^{r_0} = H^{\frac{s}{p'q_0\gamma}} w^{\frac{p'+q}{p'\gamma}}$ and (2.10) readily gives that $W^{r_0} \in A_1$, that is, $W \in A_{1, r_0}$ as desired.

On the other hand, if $p_0 > 1$ we observe that by (2.5) and the definition of s

$$\frac{p'}{p'_0} = p' \left(1 - \frac{1}{p_0} \right) = p' \left(\frac{1}{p'} - \frac{1}{s} \right) = \frac{q}{q_0} - \frac{q}{s} - \frac{p'}{s} = \frac{q}{q_0} - \frac{p'+q}{s}$$

and also that

$$\begin{aligned} 1 - \frac{r_0 p' \gamma}{s} &= r_0 \left(\frac{1}{r_0} - \frac{p' \gamma}{s} \right) = r_0 \left(\frac{1}{r} + \frac{1}{s} - \frac{p' \gamma}{s} \right) = r_0 \left(\frac{1}{r} - \frac{p' \gamma - 1}{s} \right) \\ &= r_0 \left(\frac{1}{r} - \frac{p'}{rs} \right) = \frac{r_0 p'}{r} \left(\frac{1}{p'} - \frac{1}{s} \right) = \frac{r_0 p'}{r p'_0}. \end{aligned}$$

Note that the right-hand side of the last displayed equation is positive and strictly smaller than 1 since $r_0 < r < \infty$ and $1 < p_0 < p$. As a consequence, $0 < \frac{r_0 p' \gamma}{s} < 1$ and $(\frac{s}{r_0 p' \gamma})' = \frac{r p'_0}{r_0 p'}$. We can then use Hölder's inequality and the previous calculations to conclude that

$$\begin{aligned} \left(\int_Q W^{r_0} dx \right)^{\frac{1}{r_0}} &= \left(\int_Q H^{\frac{r_0}{q_0}} w^{\frac{r_0 q}{q_0}} dx \right)^{\frac{1}{r_0}} = \left(\int_Q \left(H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}} \right)^{\frac{r_0 p' \gamma}{s}} w^{\frac{r_0 p'}{p'_0}} dx \right)^{\frac{1}{r_0}} \\ &\leq \left(\int_Q H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}} dx \right)^{\frac{p' \gamma}{s}} \left(\int_Q w^r dx \right)^{\frac{p'}{r p'_0}} \\ &\leq [H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}}]_{A_1}^{\frac{p' \gamma}{s}} \operatorname{ess\,inf}_Q (H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}})^{\frac{p' \gamma}{s}} \left(\int_Q w^r dx \right)^{\frac{p'}{r p'_0}}. \end{aligned}$$

This and (2.5) imply

$$\begin{aligned} &\left(\int_Q W^{r_0} dx \right)^{\frac{1}{r_0}} \left(\int_Q W^{-p'_0} dx \right)^{\frac{1}{p'_0}} \\ &\leq [H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}}]_{A_1}^{\frac{p' \gamma}{s}} \left(\int_Q w^r dx \right)^{\frac{p'}{r p'_0}} \left(\int_Q H^{\frac{p'_0}{q_0}} w^{\frac{(p'+q)p'_0}{s}} W^{-p'_0} dx \right)^{\frac{1}{p'_0}} \\ &= [H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}}]_{A_1}^{\frac{p' \gamma}{s}} \left(\int_Q w^r dx \right)^{\frac{p'}{r p'_0}} \left(\int_Q w^{-p'} dx \right)^{\frac{1}{p'_0}} \\ &\leq [H^{\frac{s}{p' q_0 \gamma}} w^{\frac{p'+q}{p' \gamma}}]_{A_1}^{\frac{p' \gamma}{s}} [w]_{A_{\infty, r}}^{\frac{p'}{p'_0}}. \end{aligned}$$

Taking the supremum over all cubes and using (2.10) we obtain that $W \in A_{p_0, r_0}$.

Case 3: $q = \infty$ and $\frac{1}{s} := \frac{1}{p_0} - \frac{1}{p} = \frac{1}{q_0} = \frac{1}{r_0} - \frac{1}{r} > 0$, hence $1 \leq p_0 < p \leq \infty$, $0 < q_0 < \infty$, and $0 < r_0 < r < \infty$ (note that p could also be infinity).

Again, fix $(f, g) \in \mathcal{F}$ and $w \in A_{p, r}$. We may assume that $\|gw\|_{L^p} < \infty$, otherwise there is nothing to prove. Note that as above one has $w^{-p'} \in A_{p' \gamma}$. Additionally, $p' \gamma = 1 + \frac{p'}{r} > 1$ since $r < \infty$, hence M is bounded on $L^{p' \gamma}(w^{-p'})$ and writing $\|M\|$ for its operator norm one can then introduce the Rubio de Francia algorithm

$$\mathcal{R}h = \sum_{k=0}^{\infty} \frac{M^{(k)}h}{2^k \|M\|^k}.$$

This (see [12] for more details) readily gives that if $0 \leq h \in L^{p'\gamma}(w^{-p'})$ then

$$(2.11) \quad h \leq \mathcal{R}h, \quad \|\mathcal{R}h\|_{L^{p'\gamma}(w^{-p'})} \leq 2\|h\|_{L^{p'\gamma}(w^{-p'})}, \quad [\mathcal{R}h]_{A_1} \leq 2\|M\|.$$

We fix $x_0 \in \mathbb{R}^n$ and $0 < \tau_0 < \infty$ and let $h := h_{x_0, \tau_0} := |B(x_0, \tau_0)|^{-1} \mathbf{1}_{B(x_0, \tau_0)}$ which is a non-negative function satisfying $\|h\|_{L^1} = 1$. Set $H := \mathcal{R}\left(h^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}\right)^{p'\gamma} w^{-p'}$ which satisfies by (2.11)

$$(2.12) \quad h \leq H, \quad \|H\|_{L^1} \lesssim \|h\|_{L^1} = 1, \quad [H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}]_{A_1} \leq 2\|M\|.$$

Set $W := H^{\frac{1}{q_0}} w$ and claim that $W \in A_{p_0, r_0}$. Assuming this momentarily we conclude that

$$\begin{aligned} \left(\int_{B(x_0, \tau_0)} (f(x)w(x))^{q_0} dx \right)^{\frac{1}{q_0}} &= \left(\int f^{q_0} h w^{q_0} dx \right)^{\frac{1}{q_0}} \leq \|fW\|_{L^{q_0}} \lesssim \|gW\|_{L^{p_0}} \\ &= \|(gw)H^{\frac{1}{q_0}}\|_{L^{p_0}} \leq \|gw\|_{L^p} \|H^{\frac{1}{q_0}}\|_{L^{q_0}} \lesssim \|gw\|_{L^p}, \end{aligned}$$

where we have used (2.12), that $\frac{1}{p_0} = \frac{1}{p} + \frac{1}{q_0}$ and Hölder's inequality. Note that the implicit constants are all independent of x_0 and τ . Hence, since we have that $\|gw\|_{L^p} < \infty$, we conclude that $fw \in L_{\text{loc}}^{q_0}(\mathbb{R}^n)$. Thus, if we write E_{fw} for the Lebesgue points of $(fw)^{q_0}$ we have that $|\mathbb{R}^n \setminus E_{fw}| = 0$. Moreover if we let $x_0 \in E_{fw}$ we conclude that

$$f(x_0)w(x_0) = \lim_{\tau \rightarrow 0^+} \left(\int_{B(x_0, \tau_0)} (f(x)w(x))^{q_0} dx \right)^{\frac{1}{q_0}} \lesssim \|gw\|_{L^p},$$

This eventually shows that $fw \in L^\infty(\mathbb{R}^n)$ with $\|fw\|_{L^\infty} \lesssim \|gw\|_{L^p}$ which is our desired estimate.

We are left with showing that $W := H^{\frac{1}{q_0}} w \in A_{p_0, r_0}$. Consider first the case $p_0 = 1$. In this case we have that (2.5) implies that $p' = q_0$, $\gamma = r_0^{-1}$ and $\frac{r_0}{q_0} = \frac{1}{p'\gamma}$. Thus, $W^{r_0} = H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}$ and (2.12) readily implies that $W^{r_0} \in A_1$, that is, $W \in A_{1, r_0}$ as desired.

Consider next the case $p_0 > 1$ and observe that (2.5) implies $\frac{1}{p'} - \frac{1}{q_0} = \frac{1}{p'_0}$ and

$$\begin{aligned} 1 - \frac{r_0 p' \gamma}{q_0} &= r_0 \left(\frac{1}{r_0} - \frac{p' \gamma}{q_0} \right) = r_0 \left(\frac{1}{r} + \frac{1}{q_0} - \frac{p' \gamma}{q_0} \right) \\ &= r_0 \left(\frac{1}{r} - \frac{p' \gamma - 1}{q_0} \right) = \frac{r_0 p'}{r} \left(\frac{1}{p'} - \frac{1}{q_0} \right) = \frac{r_0 p'}{r p'_0}. \end{aligned}$$

Note that the right-hand side is positive and strictly smaller than 1 since in this case $r_0 < r$ and $1 < p_0 < p$. As a consequence, $0 < \frac{r_0 p' \gamma}{q_0} \leq 1$ and $(\frac{q_0}{r_0 p' \gamma})' = \frac{r p'_0}{r_0 p'}$. Using this and Hölder's inequality we obtain

$$\left(\int_Q W^{r_0} dx \right)^{\frac{1}{r_0}} = \left(\int_Q H^{\frac{r_0}{q_0}} w^{r_0} dx \right)^{\frac{1}{r_0}} = \left(\int_Q (H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}})^{\frac{r_0 p' \gamma}{q_0}} w^{\frac{r_0 p'}{p'_0}} dx \right)^{\frac{1}{r_0}}$$

$$\begin{aligned}
&\leq \left(\int_Q H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}} dx \right)^{\frac{p'\gamma}{q_0}} \left(\int_Q w^r dx \right)^{\frac{p'}{rp_0}} \\
&\leq [H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}]_{A_1}^{\frac{p'\gamma}{q_0}} \operatorname{ess\,inf}_Q (H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}})^{\frac{p'\gamma}{q_0}} \left(\int_Q w^r dx \right)^{\frac{p'}{rp_0}}.
\end{aligned}$$

This and (2.5) imply

$$\begin{aligned}
&\left(\int_Q W^{r_0} dx \right)^{\frac{1}{r_0}} \left(\int_Q W^{-p'_0} dx \right)^{\frac{1}{p'_0}} \\
&\leq [H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}]_{A_1}^{\frac{p'\gamma}{q_0}} \left(\int_Q w^r dx \right)^{\frac{p'}{rp_0}} \left(\int_Q H^{\frac{p'_0}{q_0}} w^{\frac{p'p'_0}{q_0}} W^{-p'_0} dx \right)^{\frac{1}{p'_0}} \\
&= [H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}]_{A_1}^{\frac{p'\gamma}{q_0}} \left(\int_Q w^r dx \right)^{\frac{p'}{rp_0}} \left(\int_Q w^{-p'} dx \right)^{\frac{1}{p'_0}} \\
&\leq [H^{\frac{1}{p'\gamma}} w^{\frac{1}{\gamma}}]_{A_1}^{\frac{p'\gamma}{q_0}} [w]_{A_{p,r}}^{\frac{p'}{p_0}}.
\end{aligned}$$

Taking the supremum over all cubes and using (2.12) we obtain that $W \in A_{p_0, r_0}$. This completes the proof. \square

Remark 2.13. A careful read of the above argument reveals it works in the setting of spaces of homogeneous type, as there we have a theory of Muckenhoupt weights much as in the case of the Euclidean setting. Also, the argument can be easily extended to Muckenhoupt bases (see for instance [12, Chapter 3]). The proof can be easily adapted to that setting. Further details are left to the interested reader.

2.2. Proof of Theorem 2.1. To prove our main result we need some characterization of the $A_{\vec{p}, \vec{r}}$ classes. The following result extends [28, Lemma 3.2] to the end-point cases. The proof can be carried out *mutatis mutandis* (with the appropriate changes of notation) and we leave the details to the interested reader.

Lemma 2.14. *Consider $\vec{p} = (p_1, \dots, p_m)$ with $1 \leq p_1, \dots, p_m \leq \infty$ and $\vec{r} = (r_1, \dots, r_{m+1})$ with $1 \leq r_1, \dots, r_{m+1} < \infty$ such that $\vec{r} \preceq_\star \vec{p}$. Using the previous notation we assume that*

$$\frac{1}{\varrho} := \frac{1}{r_m} - \frac{1}{r'_{m+1}} + \sum_{i=1}^{m-1} \frac{1}{p_i} = \frac{1}{\delta_m} + \frac{1}{\delta_{m+1}} > 0,$$

and for every $1 \leq i \leq m-1$

$$\frac{1}{\theta_i} := \frac{1-r}{r} - \frac{1}{\delta_i} = \left(\sum_{j=1}^{m+1} \frac{1}{\delta_j} \right) - \frac{1}{\delta_i} > 0.$$

Then the following hold:

(i) Given $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}, \vec{r}}$, write $w := \prod_{i=1}^m w_i$ and set

$$\widehat{w} := \left(\prod_{i=1}^{m-1} w_i \right)^\varrho \quad \text{and} \quad W := w^{r_m} \widehat{w}^{-\frac{r_m}{\delta_{m+1}}} = w_m^{r_m} \widehat{w}^{\frac{r_m}{\delta_m}}$$

Then,

$$(i.1) \quad w_i^{\theta_i} \in A_{\frac{1-r}{r}\theta_i} \quad \text{with} \quad \left[w_i^{\theta_i} \right]_{A_{\frac{1-r}{r}\theta_i}} \leq [\vec{w}]_{A_{\vec{p}, \vec{r}}}^{\theta_i}, \quad \text{for every } 1 \leq i \leq m-1.$$

$$(i.2) \quad \widehat{w} \in A_{\frac{1-r}{r}\varrho} \quad \text{with} \quad [\widehat{w}]_{A_{\frac{1-r}{r}\varrho}} \leq [\vec{w}]_{A_{\vec{p}, \vec{r}}}^\varrho.$$

$$(i.3) \quad W \in A_{\frac{p_m}{r_m}, \frac{\delta_{m+1}}{r_m}}(\widehat{w}) \quad \text{with} \quad [W]_{A_{\frac{p_m}{r_m}, \frac{\delta_{m+1}}{r_m}}(\widehat{w})} \leq [\vec{w}]_{A_{\vec{p}, \vec{r}}}^{r_m}.$$

(ii) Given $w_i^{\theta_i} \in A_{\frac{1-r}{r}\theta_i}$, $1 \leq i \leq m-1$, such that

$$\widehat{w} = \left(\prod_{i=1}^{m-1} w_i \right)^\varrho \in A_{\frac{1-r}{r}\varrho}$$

and $W \in A_{\frac{p_m}{r_m}, \frac{\delta_{m+1}}{r_m}}(\widehat{w})$, let us set

$$w_m := W^{\frac{1}{r_m}} \widehat{w}^{-\frac{1}{\delta_m}}.$$

Then $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}, \vec{r}}$ and, moreover,

$$[\vec{w}]_{A_{\vec{p}, \vec{r}}} \leq [W]_{A_{\frac{p_m}{r_m}, \frac{\delta_{m+1}}{r_m}}(\widehat{w})}^{\frac{1}{r_m}} [\widehat{w}]_{A_{\frac{1-r}{r}\varrho}}^{\frac{1}{\varrho}} \prod_{i=1}^{m-1} \left[w_i^{\theta_i} \right]_{A_{\frac{1-r}{r}\theta_i}}^{\frac{1}{\theta_i}}.$$

(iii) For any measurable function $f \geq 0$ and in the context of (i) or (ii) there hold

$$\|fw\|_{L^p} = \left\| \left(f \widehat{w}^{-\frac{1}{r_{m+1}}} \right)^{r_m} W \right\|_{L^{\frac{p}{r_m}}(d\widehat{w})}^{\frac{1}{r_m}}$$

and

$$\|fw_m\|_{L^{p_m}} = \left\| \left(f \widehat{w}^{-\frac{1}{r_m}} \right)^{r_m} W \right\|_{L^{\frac{p_m}{r_m}}(d\widehat{w})}^{\frac{1}{r_m}}.$$

In the previous result we are assuming that δ_{m+1}^{-1} and δ_m^{-1} cannot be simultaneously zero, but they could be zero individually, in which case some of the statements require appropriate interpretations which are left to the interest reader. We also note that the assumption $\varrho^{-1} > 0$, leads to the fact that the class $A_{\frac{p_m}{r_m}, \frac{\delta_{m+1}}{r_m}}(\widehat{w})$ is non-trivial —recall that, as observed above, $W \in A_{1, \infty}(\widehat{w})$ means that $W \approx 1$ almost everywhere.

Proof of Theorem 2.1. The argument is almost identical to that in [28, Proof of Theorem 1.1] with the difference that we allow the p_i 's or q_i 's to be infinity, and this requires us to use Theorem 2.3 in place of [28, Theorem 3.1] (which is really [17, Theorem 5.1]) and Lemma 2.14 in place of [28, Lemma 3.2]. There is however a subtle point that we would like to address here. In the present scenario, to

use Lemma 2.14 we need to assume that $\varrho^{-1} > 0$, while in [28, Lemma 3.2] this condition follows automatically from the assumption that $\delta_{m+1}^{-1} > 0$ since there it was required that $p > r'_{m+1}$ (using the notation there $\vec{r} \preceq \vec{p}$). Thus we consider two possible scenarios $r'_{m+1} > p$ or $r'_{m+1} = p$.

Assume first that $r'_{m+1} > p$ (that is, $\vec{r} \preceq \vec{p}$), hence $\delta_{m+1}^{-1} > 0$. We just follow [28, Proof of Theorem 1.1], an argument which changes one exponent at a time, with the alterations pointed out above after noting that implicit in that scheme it is shown that the intermediate exponents $\vec{t} = (t_1, \dots, t_m)$ satisfy $\vec{r} \preceq \vec{t}$, that is, $r'_{m+1} > t = (\sum_{i=1}^m \frac{1}{t_i})^{-1}$ since in the present case $\vec{r} \preceq \vec{p}$ and $\vec{r} \prec \vec{q}$. This means that in the successive uses of Lemma 2.14 we always have that the corresponding $\varrho^{-1} > 0$ and in the iteration argument we never reach the end-point r'_{m+1} in the target space.

The second scenario is that on which $p = r'_{m+1}$, that is, $\delta_{m+1}^{-1} = 0$. Let $\vec{q} = (q_1, \dots, q_m)$ satisfy $\vec{r} \prec \vec{q}$ so that $p = r'_{m+1} > q$. Let $\mathcal{I} = \{i : 1 \leq i \leq m, p_i > q_i\}$ and observe that $\mathcal{I} \neq \emptyset$ since otherwise we get a contradiction:

$$\frac{1}{r'_{m+1}} = \frac{1}{p} = \sum_{i=1}^m \frac{1}{p_i} \geq \sum_{i=1}^m \frac{1}{q_i} = \frac{1}{q} > \frac{1}{r'_{m+1}}.$$

Thus rearranging the f_i 's if needed we may assume that $p_m > q_m > r_m$ —since $\vec{r} \prec \vec{q}$. In particular $\delta_m^{-1} > 0$ and Lemma 2.14 applies. This allows us to follow **Step 1** in [28, Proof of Theorem 1.1] *mutatis mutandis* with the modifications pointed out above to obtain the desired estimate for the exponent $\vec{t} = (t_1, \dots, t_m)$ with $t_i = p_i$ for $1 \leq i \leq m$ and $t_m = q_m$ and the associated class of weights $A_{\vec{t}, \vec{r}}$. Moreover,

$$\frac{1}{t} = \sum_{i=1}^m \frac{1}{t_i} = \sum_{i=1}^{m-1} \frac{1}{p_i} + \frac{1}{q_m} > \sum_{i=1}^m \frac{1}{p_i} = \frac{1}{p} = \frac{1}{r'_{m+1}}.$$

Hence $r'_{m+1} > t$, that is, $\vec{r} \preceq \vec{t}$ and we can extrapolate from this exponent to change the other entries in \vec{t} much as in the first scenario above. Let us observe that, again, at any step in the iteration we will never reach the end-point r'_{m+1} in the target space. Further details are left to the interested reader.

To complete the proof we sketch how to establish (2.2). The proof is almost identical to [28, Proof of Theorem 1.1], the only difference is that here some of the s_i 's could be infinity. If that is the case we just need to observe that by assumption not all the s_i can be infinity (otherwise $s = 0$), hence

$$\begin{aligned} \left(\sum_j \prod_{i=1}^m \|f_i^j v_i\|_{L^{s_i}}^{s_i} \right)^{\frac{1}{s}} &\leq \prod_{i:s_i=\infty} \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^\infty}^{s_i}} \left(\sum_j \prod_{i:s_i \neq \infty} \|f_i^j v_i\|_{L^{s_i}}^{s_i} \right)^{\frac{1}{s}} \\ &\leq \prod_{i:s_i=\infty} \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^\infty}^{s_i}} \prod_{i:s_i \neq \infty} \left(\sum_j \|f_i^j v_i\|_{L^{s_i}}^{s_i} \right)^{\frac{1}{s_i}} = \prod_{i=1}^m \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^{s_i}}^{s_i}}. \end{aligned}$$

With this estimate in hand the proof can be completed in exactly the same manner and we omit the details.

□

3. CALDERÓN-ZYGMUND OPERATORS, THE BILINEAR HILBERT TRANSFORM, AND SPARSE FORMS

In [28] several applications were given to show that extrapolation can be used to provide almost trivial proofs of known results and also to obtain new estimates. With our main result in this paper we can easily complete the picture and obtain the end-point cases. Here we will just indicate the resulting estimates, see [28] and the references therein for more details and the precise definitions.

We start with T being a multilinear Calderón-Zygmund operator. Applying Theorem 1.1 together with the weighted estimates from [27] one immediately obtains (1.2). This should be compared with [28, Section 2.1] where all the q_i 's and s_i 's are assumed to be finite.

Our second example is as follows. Fix $\vec{r} = (r_1, \dots, r_{m+1})$, with $r_i \geq 1$ for $1 \leq i \leq m+1$, and $\sum_{i=1}^{m+1} \frac{1}{r_i} > 1$, and a sparsity constant $\zeta \in (0, 1)$. Let T be an operator so that for every $f_1, \dots, f_m, h \in C_c^\infty(\mathbb{R}^n)$

$$(3.1) \quad \left| \int_{\mathbb{R}^n} h T(f_1, \dots, f_m) dx \right| \lesssim \sup_S \Lambda_{S, \vec{r}}(f_1, \dots, f_m, h),$$

where the sup runs over all sparse families with sparsity constant ζ . In [28, Corollary 2.15] it was shown that for all exponents $\vec{q} = (q_1, \dots, q_m)$, with $1 < q_1, \dots, q_m < \infty$ and $\vec{r} \prec \vec{q}$, for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}, \vec{r}}$, and for all $f_1, \dots, f_m \in C_c^\infty(\mathbb{R}^n)$ one has

$$\|T(f_1, \dots, f_m)v\|_{L^q} \lesssim \prod_{i=1}^m \|f_i v_i\|_{L^{q_i}},$$

where $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m}$ and $v := \prod_{i=1}^m v_i$. From this and Theorem 2.1 we can immediately allow the q_i 's to be infinity (provided that $q < \infty$). Moreover, we obtain (1.2) for all exponents $\vec{q} = (q_1, \dots, q_m)$, with $1 < q_1, \dots, q_m \leq \infty$, $\vec{r} \prec \vec{q}$, and $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$; $\vec{s} = (s_1, \dots, s_m)$, with $1 < s_1, \dots, s_m \leq \infty$, $\vec{r} \prec \vec{q}$, and $\frac{1}{s} := \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$; and for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}, \vec{r}}$. Again, comparing with [28, Corollary 2.15], here we could have that some of the q_i 's (but not all) and/or some of the s_i 's (but not all) are infinity.

A particular case of interest is that of rough bilinear singular integrals introduced by Coifman and Meyer and further studied by [21]. As explained in [28, Section 2.4], either from the weighted estimates obtained in [9] or from the sparse domination from [1] one easily gets an extension of [28, Corollary 2.17] covering the end-point cases.

Our next example is the bilinear Hilbert transform, denoted by BH , which can be framed within the previous class of operators that are controlled by a sparse form as above (see [14, Theorem 2]). From that, repeating the argument in [28, Section

2.5] but invoking Theorem 2.1 we obtain the following estimates. Let $\vec{r} = (r_1, r_2, r_3)$ be such that $1 < r_1, r_2, r_3 < \infty$ and

$$(3.2) \quad \frac{1}{\min\{r_1, 2\}} + \frac{1}{\min\{r_2, 2\}} + \frac{1}{\min\{r_3, 2\}} < 2.$$

Let $\vec{p} = (p_1, p_2)$, $\vec{s} = (s_1, s_2)$ with $1 < p_1, p_2, s_1, s_2 \leq \infty$ be so that $\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2} > 0$, $\frac{1}{s} := \frac{1}{s_1} + \frac{1}{s_2} > 0$. If $\vec{r} \prec \vec{p}$ and $\vec{w} = (w_1, w_2) \in A_{\vec{p}, \vec{r}}$ then

$$\|BH(f, g)w_1w_2\|_{L^p} \lesssim \|fw_1\|_{L^{p_1}} \|gw_1\|_{L^{p_2}}.$$

Moreover, if additionally $\vec{r} \prec \vec{s}$ then

$$\left\| \{BH(f_j, g_j)w_1w_2\}_j \right\|_{L_{\ell^s}^p} \lesssim \left\| \{f_jw_1\}_j \right\|_{L_{\ell^{s_1}}^{p_1}} \left\| \{g_jw_2\}_j \right\|_{L_{\ell^{s_2}}^{p_2}}.$$

Again, here we can allow p_1 or p_2 (but not both) and/or s_1 or s_2 (but not both) to be infinity. The same kind of argument allows us to readily get iterated vector-valued inequalities in spaces of the form $L_{\ell^t}^p$. All these should be compared with the helicoidal method developed in [2, 3, 4], on which they prove all these estimates by some delicate discretization arguments. In a nutshell once we have the estimates from [14, Theorem 2] in hand, our powerful method based on extrapolation easily gives all the desired estimates, including the vector-valued ones, and does not require to use any further fine analysis or decomposition of the operator. This occurs because extrapolation is not something related to operators, is a property about families of functions satisfying weighted norm inequalities.

Finally, our method allows us to deal with commutators of the previous operators with BMO functions. In [28, Section 2.6], the method developed in [5], was further pushed to obtain [28, Theorem 2.22]. The latter in conjunction with Theorem 2.1 easily yields the following extension where both in the hypotheses and the conclusions one can include the end-points:

Corollary 3.3. *Let T be an m -linear operator and let $\vec{r} = (r_1, \dots, r_{m+1})$, with $1 \leq r_1, \dots, r_{m+1} < \infty$. Assume that there exists $\vec{p} = (p_1, \dots, p_m)$, with $1 \leq p_1, \dots, p_m \leq \infty$ and $\vec{r} \preceq_\star \vec{p}$, such that for all $\vec{w} = (w_1, \dots, w_m) \in A_{\vec{p}, \vec{r}}$, we have*

$$(3.4) \quad \|T(f_1, f_2, \dots, f_m)w\|_{L^p} \lesssim \prod_{i=1}^m \|f_iw_i\|_{L^{p_i}},$$

where $\frac{1}{p} := \frac{1}{p_1} + \dots + \frac{1}{p_m}$ and $w := \prod_{i=1}^m w_i$.

Then, for all exponents $\vec{q} = (q_1, \dots, q_m)$, with $1 < q_1, \dots, q_m \leq \infty$ so that $\vec{r} \prec \vec{q}$ and $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$, for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}, \vec{r}}$, for all $\mathbf{b} = (b_1, \dots, b_m) \in \text{BMO}^m$, and for each multi-index α , we have

$$(3.5) \quad \|[T, \mathbf{b}]_\alpha(f_1, f_2, \dots, f_m)v\|_{L^q} \lesssim \prod_{i=1}^m \|b_i\|_{\text{BMO}}^{\alpha_i} \|f_i v_i\|_{L^{q_i}},$$

where $v := \prod_{i=1}^m v_i$. Moreover if $\vec{s} = (s_1, \dots, s_m)$, with $1 < s_1, \dots, s_m \leq \infty$ so that $\vec{r} \prec \vec{s}$ and $\frac{1}{s} := \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$, then

$$(3.6) \quad \left\| \{[T, \mathbf{b}]_\alpha(f_1^j, f_2^j, \dots, f_m^j)v\}_j \right\|_{L_{\ell^s}^p} \lesssim \prod_{i=1}^m \|b_i\|_{\text{BMO}}^{\alpha_i} \left\| \{f_i^j v_i\}_j \right\|_{L_{\ell^{s_i}}^{p_i}}.$$

Proof. We first use (3.4) and Theorem 2.1 to see that for all exponents $\vec{t} = (t_1, \dots, t_m)$, with $1 < t_1, \dots, t_m \leq \infty$ so that $\vec{r} \prec \vec{t}$ and $\frac{1}{t} := \frac{1}{t_1} + \dots + \frac{1}{t_m} > 0$, and for all weights $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{t}, \vec{r}}$ there holds

$$\|T(f_1, f_2, \dots, f_m)v\|_{L^t} \lesssim \prod_{i=1}^m \|f_i v_i\|_{L^{t_i}},$$

where $v := \prod_{i=1}^m v_i$. In particular, this estimate is valid for a particular choice of \vec{t} which additionally satisfies $1 < t_1, \dots, t_m < \infty$. This allows us to invoke [28, Theorem 2.22] to see that (3.6) is valid for the same exponents as in the statement with the additionally assumption that $1 < q_1, \dots, q_m < \infty$. Another use of Theorem 2.1 immediately remove that restriction and also yields the vector-valued inequalities in (3.6). This completes the proof. \square

As a consequence of this result and all the applications considered above we can obtain that (3.5) and (3.6) hold in the following scenarios:

- T is an m -linear Calderón-Zygmund operator, $\vec{q} = (q_1, \dots, q_m)$, $\vec{s} = (s_1, \dots, s_m)$, with $1 < q_i, s_i \leq \infty$, $1 \leq i \leq m$, so that $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$ and $\frac{1}{s} := \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$, and $\vec{v} \in A_{\vec{q}}$. The same applies to a bilinear rough singular integral as in [28, Section 2.4] in the case $m = 2$.
- T is any linear operators satisfying (3.1), $\vec{q} = (q_1, \dots, q_m)$, $\vec{s} = (s_1, \dots, s_m)$, with $\vec{r} \prec \vec{q}$ and $\vec{r} \prec \vec{s}$ so that $\frac{1}{q} := \frac{1}{q_1} + \dots + \frac{1}{q_m} > 0$ and $\frac{1}{s} := \frac{1}{s_1} + \dots + \frac{1}{s_m} > 0$, and $\vec{v} = (v_1, \dots, v_m) \in A_{\vec{q}, \vec{r}}$.

These are respectively extensions of [28, Corollaries 2.25 and 2.26] where we are now able to consider the end-point cases, and these estimates are new to the best of our knowledge. The same occurs in the case of the bilinear Hilbert transform for which we have the following extension [28, Corollary 2.27] which contain new estimates at the end-point cases:

Corollary 3.7. *Assume that $\vec{r} = (r_1, r_2, r_3)$, $1 < r_1, r_2, r_3 < \infty$, verifies (3.2). Let $\vec{p} = (p_1, p_2)$, $\vec{s} = (s_1, s_2)$ with $1 < p_1, p_2, s_1, s_2 \leq \infty$ be so that $\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2} > 0$, $\frac{1}{s} := \frac{1}{s_1} + \frac{1}{s_2} > 0$. If $\vec{r} \prec \vec{p}$, for all weights $\vec{w} = (w_1, w_2) \in A_{\vec{p}, \vec{r}}$, for all $\mathbf{b} = (b_1, b_2) \in \text{BMO}^2$, and for each multi-index $\alpha = (\alpha_1, \alpha_2)$ it follows that*

$$\|[BH, \mathbf{b}]_\alpha(f, g)w_1 w_2\|_{L^p} \lesssim \|b_1\|_{\text{BMO}}^{\alpha_1} \|b_2\|_{\text{BMO}}^{\alpha_2} \|f w_1\|_{L^{p_1}} \|g w_2\|_{L^{p_2}},$$

and, if one further assumes that $\vec{r} \prec \vec{s}$,

$$\left\| \{[BH, \mathbf{b}]_\alpha(f_j, g_j)w_1 w_2\}_j \right\|_{L_{\ell^s}^p} \lesssim \|b_1\|_{\text{BMO}}^{\alpha_1} \|b_2\|_{\text{BMO}}^{\alpha_2} \left\| \{f_j w_1\}_j \right\|_{L_{\ell^{s_1}}^{p_1}} \left\| \{g_j w_2\}_j \right\|_{L_{\ell^{s_2}}^{p_2}}.$$

4. MIXED-NORM ESTIMATES AND TENSOR PRODUCTS OF BILINEAR SINGULAR INTEGRALS

In [30] the first, third and last authors of the present paper showed that if T is a bilinear bi-parameter Calderón-Zygmund operator that is free of full paraproducts (see [30] for the definitions), then

$$\|T(f_1, f_2)w\|_{L^p} \lesssim \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}$$

holds for all $p_1, p_2 \in (1, \infty)$ and weights w_1 and w_2 such that $w_1^{p_1} \in A_{p_1}(\mathbb{R}^n \times \mathbb{R}^m)$ and $w_2^{p_2} \in A_{p_2}(\mathbb{R}^n \times \mathbb{R}^m)$. Here p is defined by $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $w = w_1 w_2$ and $A_p(\mathbb{R}^n \times \mathbb{R}^m)$ is the class of bi-parameter A_p weights (obtained by replacing cubes by rectangles in the usual definition). In addition, the same estimate were obtained for all tensor products $T = T_n \otimes T_m$, where T_n and T_m are bilinear one-parameter Calderón-Zygmund operators in \mathbb{R}^n and \mathbb{R}^m , respectively (we recall the definition below). Tensor products $T_n \otimes T_m$ are examples of bilinear bi-parameter Calderón-Zygmund operators, which are not necessarily free of full paraproducts (if *both* of them are *not* free of paraproducts). Recall that initially $T_n \otimes T_m$ is defined via

$$(T_n \otimes T_m)(f_1 \otimes f_2, g_1 \otimes g_2)(x) := T_n(f_1, g_1)(x_1) T_m(f_2, g_2)(x_2),$$

where $f_1, g_1: \mathbb{R}^n \rightarrow \mathbb{C}$, $f_2, g_2: \mathbb{R}^m \rightarrow \mathbb{C}$ and $x = (x_1, x_2) \in \mathbb{R}^{n+m}$.

The main goal of this section is to prove Theorem 4.2 where we establish weighted estimates for the previous tensor products of bilinear Calderón-Zygmund operators with the new class of weights introduced in Definition 4.1. In turn, we can extrapolate using Theorem 2.1 to prove Corollary 4.3. The latter result contains some mixed-norm end-point estimates that were not proved in [30]. The bottom line is that with our current extrapolation result we can consider the cases in which the outer exponents are infinity even when the inner exponents are in the quasi-Banach range.

Definition 4.1. Given $\vec{p} = (p_1, p_2)$ with $1 < p_1, p_2 < \infty$ let $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. Let $\vec{w} = (w_1, w_2)$ be such that $0 < w_1(x_1, \cdot), w_2(x_1, \cdot) < \infty$ a.e. in \mathbb{R}^m for a.e. $x_1 \in \mathbb{R}^n$ and $0 < w_1(\cdot, x_2), w_2(\cdot, x_2) < \infty$ a.e. in \mathbb{R}^n for a.e. $x_2 \in \mathbb{R}^m$. We say that $\vec{w} \in \mathbb{A}_{\vec{p}, n, m}$ if $w_1(x_1, \cdot)^{p_1} \in A_{p_1}(\mathbb{R}^m)$ and $w_2(x_1, \cdot)^{p_2} \in A_{p_2}(\mathbb{R}^m)$ for a.e. $x_1 \in \mathbb{R}^n$; $(w_1(\cdot, x_2), w_2(\cdot, x_2)) \in A_{\vec{p}}(\mathbb{R}^n)$ for a.e. $x_2 \in \mathbb{R}^m$; and moreover

$$\operatorname{ess\,sup}_{x_2 \in \mathbb{R}^m} [(w_1(\cdot, x_2), w_2(\cdot, x_2))]_{A_{\vec{p}}(\mathbb{R}^n)} < \infty$$

and

$$\operatorname{ess\,sup}_{x_1 \in \mathbb{R}^n} [w_1(x_1, \cdot)^{p_1}]_{A_{p_1}(\mathbb{R}^m)} + \operatorname{ess\,sup}_{x_1 \in \mathbb{R}^n} [w_2(x_1, \cdot)^{p_2}]_{A_{p_2}(\mathbb{R}^m)} < \infty.$$

Here $A_{\vec{p}}(\mathbb{R}^n)$ is the class of bilinear $A_{\vec{p}}$ weights in \mathbb{R}^n as defined in Section 2 and $A_{p_i}(\mathbb{R}^m)$ is the class of A_{p_i} weights in \mathbb{R}^m .

A function $K: \mathbb{R}^{3d} \setminus \Delta \rightarrow \mathbb{C}$, where $\Delta = \{(x, x, x) \in \mathbb{R}^{3d} : x \in \mathbb{R}^d\}$, is called a bilinear singular integral kernel if for some $\alpha \in (0, 1]$ and a constant $C > 0$ it

satisfies the estimates

$$|K(x, y, z)| \leq C \frac{1}{(|x - y| + |x - z|)^{2d}}$$

for $(x, y, z) \in \mathbb{R}^{3d} \setminus \Delta$,

$$\begin{aligned} & |K(x', y, z) - K(x, y, z)| + |K(y, x', z) - K(y, x, z)| \\ & + |K(y, z, x') - K(y, z, x)| \leq C \frac{|x - x'|^\alpha}{(|x - y| + |x - z|)^{2d+\alpha}} \end{aligned}$$

for $(x', y, z), (x, y, z) \in \mathbb{R}^{3d} \setminus \Delta$ such that $|x' - x| \leq \max(|x - y|, |x - z|)/2$. We denote the smallest possible constant C in these inequalities by $\|K\|_{CZ_\alpha}$.

We say that T is a bilinear singular integral in \mathbb{R}^d if there exists a bilinear singular integral kernel K such that

$$\langle T(f_1, f_2), f_3 \rangle = \iiint_{\mathbb{R}^{3d}} K(x, y, z) f_1(y) f_2(z) f_3(x) dx dy dz$$

for all $f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^d)$ whenever two of the functions are disjointly supported, that is, $\text{supp } f_i \cap \text{supp } f_j = \emptyset$ for some $i \neq j$. We say that T is a bilinear Calderón-Zygmund operator in \mathbb{R}^d if T is a bilinear singular integral in \mathbb{R}^d and moreover T is bounded from $L^{p_1}(\mathbb{R}^d) \times L^{p_2}(\mathbb{R}^d)$ into $L^p(\mathbb{R}^d)$ for some $1 < p_1, p_2 \leq \infty$ such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$. For more information about multilinear Calderón-Zygmund operators see [23] where it is shown, among other things, that T is bounded from $L^{q_1}(\mathbb{R}^d) \times L^{q_2}(\mathbb{R}^d)$ into $L^q(\mathbb{R}^d)$ for all $q_1, q_2 \in (1, \infty]$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$.

So-called $T1$ theorems give conditions under which singular integrals are indeed Calderón-Zygmund operators. For bilinear $T1$ theorems see for example [23] (where multilinear singular integrals are also treated) or [29].

The following theorem contains our weighted estimates for tensor products of bilinear Calderón-Zygmund operators involving the class of weights appearing in Definition 4.1.

Theorem 4.2. *Let T_n and T_m be bilinear Calderón-Zygmund operators in \mathbb{R}^n and \mathbb{R}^m , respectively. Let $\vec{p} = (p_1, p_2)$ be such that $1 < p_1, p_2 < \infty$ and set $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. Then*

$$\|T_n \otimes T_m(f_1, f_2)w_1w_2\|_{L^p} \lesssim \|f_1w_1\|_{L^{p_1}} \|f_2w_2\|_{L^{p_2}}$$

holds for all weights $\vec{w} = (w_1, w_2) \in \mathbb{A}_{\vec{p}, n, m}$ and for all $f_1w_1 \in L^{p_1}(\mathbb{R}^{n+m})$ and $f_2w_2 \in L^{p_2}(\mathbb{R}^{n+m})$.

Given $0 < p, q \leq \infty$, the space $L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ consists of measurable functions $f: \mathbb{R}^{n+m} \rightarrow \mathbb{C}$ such that

$$\|f\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} := \left\| \|f(x_1, x_2)\|_{L_{x_2}^q(\mathbb{R}^m)} \right\|_{L_{x_1}^p(\mathbb{R}^n)} < \infty.$$

Combining Theorem 4.2 with our main extrapolation result, Theorem 2.1, we will easily get the following consequence:

Corollary 4.3. *Let T_n and T_m be bilinear Calderón-Zygmund operators in \mathbb{R}^n and \mathbb{R}^m , respectively. Let $1 < p_1, p_2 \leq \infty$ and $1 < q_1, q_2 < \infty$ be such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$. Then*

$$(4.4) \quad \|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} \lesssim \|f_1\|_{L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))} \|f_2\|_{L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}.$$

for all $f_1, f_2 \in L^\infty(\mathbb{R}^{n+m})$.

Our last goal is to extend the previous result in the following ways. First, we would like to consider the end-points $q_1 = \infty$ or $q_2 = \infty$. Also, we wish to show that (4.4) holds for all $f_1 \in L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$. This is straightforward if $p_1, p_2, q_1, q_2 < \infty$ by a standard density argument. However, such approach fails when some of the exponents is infinity and in that scenario one even needs to make sense of $T_n \otimes T_m(f_1, f_2)$. All this is done in the main result of this section:

Theorem 4.5. *Let T_n and T_m be bilinear Calderón-Zygmund operators in \mathbb{R}^n and \mathbb{R}^m , respectively. Let $1 < p_1, p_2, q_1, q_2 \leq \infty$ be such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$. Then we have*

$$(4.6) \quad \|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} \lesssim \|f_1\|_{L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))} \|f_2\|_{L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}$$

for all $f_1 \in L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$.

To prove the previous results we need some preliminaries. Let $\Omega_d := (\{0, 1\}^d)^\mathbb{Z}$ be the set of $\{0, 1\}^d$ -valued sequences $(\omega_i)_{i \in \mathbb{Z}}$ equipped with the probability measure such that the coordinates are independent and uniformly distributed over $\{0, 1\}^d$. We write \mathbb{E}_ω to denote the associated conditional expectation.

Let

$$\mathcal{D}_0^d := \{2^{-k}([0, 1]^d + m) : k \in \mathbb{Z}, m \in \mathbb{Z}^d\}$$

be the standard dyadic lattice of cubes in \mathbb{R}^d . For $\omega \in \Omega_d$ define the dyadic lattice

$$\mathcal{D}_\omega^d := \{I + \omega : I \in \mathcal{D}_0^d\}, \quad I + \omega := I + \sum_{i: 2^{-i} < \ell(I)} 2^{-i} \omega_i.$$

By a dyadic lattice \mathcal{D}^d we mean that $\mathcal{D}^d = \mathcal{D}_\omega^d$ for some $\omega \in \Omega_d$.

Let \mathcal{D}^d be a dyadic lattice and suppose $I \in \mathcal{D}^d$. The sidelength of I is denoted by $\ell(I)$. If $k = 0, 1, 2, \dots$, then $I^{(k)}$ denotes the k -ancestor, that is, $I^{(k)}$ is the unique cube in \mathcal{D}^d such that $I \subset I^{(k)}$ and $\ell(I^{(k)}) = 2^k \ell(I)$. We define $\text{ch}(I)$ to be the collection of dyadic children of I , that is, those $I' \in \mathcal{D}^d$ such that $(I')^{(1)} = I$.

Suppose $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and $I \in \mathcal{D}^d$. The average $\langle f \rangle_I$ and the martingale difference $\Delta_I f$ are defined by

$$\langle f \rangle_I := \frac{1}{|I|} \int_I f dx, \quad \Delta_I f := \sum_{I' \in \text{ch}(I)} \langle f \rangle_{I'} 1_{I'} - \langle f \rangle_I 1_I.$$

Given $I \in \mathcal{D}^d$ we denote by h_I a cancellative L^2 normalized Haar function. That is, writing $I = I_1 \times \dots \times I_d$ we can define the Haar function h_I^η , $\eta = (\eta_1, \dots, \eta_d) \in$

$\{0, 1\}^d$, by setting $h_I^\eta = h_{I_1}^{\eta_1} \otimes \cdots \otimes h_{I_d}^{\eta_d}$, where $h_{I_i}^0 = |I_i|^{-\frac{1}{2}} 1_{I_i}$ and $h_{I_i}^1 = |I_i|^{-\frac{1}{2}} (1_{I_{i,l}} - 1_{I_{i,r}})$ for every $i = 1, \dots, d$. Here $I_{i,l}$ and $I_{i,r}$ are the left and right halves of the interval I_i respectively. If $\eta \in \{0, 1\}^d \setminus \{0\}^d$ the Haar function is cancellative, that is, $\int h_I^\eta dx = 0$. We usually suppress η and simply write h_I to mean that h_I^η , for some $\eta \in \{0, 1\}^d \setminus \{0\}^d$ and we write h_I^0 to denote h_I^η for $\eta = (0, \dots, 0)$. It is standard to see that for every $I \in \mathcal{D}^d$ one has

$$(4.7) \quad \Delta_I f = \sum_{\eta \in \{0,1\}^d \setminus \{0\}^d} \langle f, h_I^\eta \rangle h_I^\eta.$$

Let $\mathcal{D} = \mathcal{D}^d$ be a dyadic lattice in \mathbb{R}^d and let $k = (k_1, k_2, k_3) \in \mathbb{N}_0^3$ —here and elsewhere $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A bilinear dyadic shift $\mathcal{U}_{\mathcal{D}}^k$ of complexity k is an operator that has three possible different forms. One of them is

$$(4.8) \quad \mathcal{U}_{\mathcal{D}}^k(f_1, f_2) = \sum_{K \in \mathcal{D}} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D} \\ I_i^{(k_i)} = K}} a_{K, (I_i)} \langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2} \rangle h_{I_3},$$

where every $a_{K, (I_i)} = a_{K, I_1, I_2, I_3}$ satisfies

$$|a_{K, (I_i)}| \leq \frac{|I_1|^{\frac{1}{2}} |I_2|^{\frac{1}{2}} |I_3|^{\frac{1}{2}}}{|K|^2}.$$

Another form of the shift is formed by replacing the Haar functions in the above formula by h_{I_1} , $h_{I_2}^0$ and h_{I_3} , and in the third form we have h_{I_1} , h_{I_2} and $h_{I_3}^0$.

A bilinear dyadic paraproduct $\mathcal{U}_{\mathcal{D}}$ has also three possible forms. One of them is

$$(4.9) \quad \mathcal{U}_{\mathcal{D}}(f_1, f_2) = \sum_{K \in \mathcal{D}} a_K \langle f_1 \rangle_K \langle f_2 \rangle_K h_K,$$

where the coefficients satisfy

$$\sup_{K_0 \in \mathcal{D}} \left(\frac{1}{|K_0|} \sum_{\substack{K \in \mathcal{D} \\ K \subset K_0}} |a_K|^2 \right)^{\frac{1}{2}} \leq 1.$$

The second form is $\langle f_1, h_K \rangle \langle f_2 \rangle_K \frac{1}{|K|}$ and third one is $\langle f_1 \rangle_K \langle f_2, h_K \rangle \frac{1}{|K|}$.

We will refer to bilinear dyadic shifts and paraproducts as bilinear dyadic model operators.

Suppose T is a bilinear Calderón-Zygmund operator in \mathbb{R}^d related to a kernel K . We will repeatedly use the following bilinear one-parameter representation theorem from [29]: if $f_1, f_2, f_3 \in L^3(\mathbb{R}^d)$ then

$$(4.10) \quad \langle T(f_1, f_2), f_3 \rangle = C_T \mathbb{E}_\omega \sum_{k \in \mathbb{N}_0^3} \sum_{u=1}^{C_d} 2^{-\max_i k_i \frac{\alpha}{2}} \langle U_{\mathcal{D}_\omega, u}^k(f_1, f_2), f_3 \rangle.$$

Here $|C_T| \lesssim \|T\|_{L^3 \times L^3 \rightarrow L^{3/2}} + \|K\|_{CZ_\alpha}$ and α is the parameter in the Hölder continuity assumptions of the kernel K of T . Also, C_d is a finite constant depending just on d . For each u and $k = (k_1, k_2, k_3) \in \mathbb{N}_0^3$, if $\max k_i > 0$, then $U_{\mathcal{D}_\omega, u}^k$ is a

bilinear dyadic shift of complexity k with respect to the lattice \mathcal{D}_ω , that is, it has the form $\mathcal{U}_{\mathcal{D}_\omega}^k$ (see (4.8)). On the other hand, if $\max k_i = 0$, then $U_{\mathcal{D}_\omega, u}^k$ is either a bilinear dyadic shift of complexity $(0, 0, 0)$ of the form $\mathcal{U}_{\mathcal{D}_\omega}^{(0,0,0)}$ (see (4.8)), or a bilinear dyadic paraproduct of the form $\mathcal{U}_{\mathcal{D}_\omega}$ (see (4.9)).

Let $\vec{p} = (p_1, p_2)$ with $1 < p_1, p_2 \leq \infty$ such that $1/p = 1/p_1 + 1/p_2 > 0$. Given $\vec{k} = (k_1, k_2, k_3)$, let $U_{\mathcal{D}}^k$ be a bilinear dyadic shift of complexity k if $\max k_i > 0$, or if $\max k_i = 0$ assume that $U_{\mathcal{D}}^k$ is either a bilinear dyadic shift of complexity $(0, 0, 0)$ or a bilinear dyadic paraproduct. We claim that for $f_1, f_2 \in L_c^\infty(\mathbb{R}^d)$ we have

$$(4.11) \quad \|U_{\mathcal{D}}^k(f_1, f_2)w\|_{L^p} \lesssim (\max_i k_i + 1) \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}},$$

where $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbb{R}^d)$, $w = w_1 w_2$ and the implicit constant depends on the $A_{(p_1, p_2)}$ characteristic of (w_1, w_2) . To obtain these estimates we first observe that in [29, Section 5] it was shown that bilinear dyadic shifts and paraproducts satisfy the so-called sparse form domination. Second, in [31, Theorem 3.2] it was shown that sparse operators satisfy (4.11) under the further assumption $p_1, p_2 < \infty$. Finally we can apply Theorem 1.1 to conclude as well the case on which either p_1 or p_2 are infinity.

Lemma 4.12. *Let Ω_0 be a probability space and let $\{\mathcal{D}_\omega^d\}_{\omega \in \Omega_0}$ be a collection of dyadic grids in \mathbb{R}^d . Given $0 < p < \infty$ and $\nu \in A_\infty(\mathbb{R}^d)$, for every sequence of functions $\{g_\omega\}_{\omega \in \Omega_0}$ with $g_\omega: \mathbb{R}^d \rightarrow \mathbb{C}$ for every $\omega \in \Omega_0$ one has*

$$\|\mathbb{E}_\omega g_\omega\|_{L^p(\nu)} \lesssim \left\| \mathbb{E}_\omega \left(\sum_{J \in \mathcal{D}_\omega^d} |\Delta_J g_\omega|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\nu)}.$$

Proof. Given \mathcal{D}^d , a dyadic lattice in \mathbb{R}^d , one can first show that

$$(4.13) \quad \|g\|_{L^t(\nu)} \lesssim \left\| \left(\sum_{J \in \mathcal{D}^d} |\Delta_J g|^2 \right)^{\frac{1}{2}} \right\|_{L^t(\nu)}$$

for any $t \in (0, \infty)$ and $\nu \in A_\infty(\mathbb{R}^d)$ where the implicit constant is independent of \mathcal{D}^d . To see this we can first invoke [8] or [16, Theorem 3.4] to obtain the case $t = 2$. In turn, using A_∞ -extrapolation, see [11, Theorem 2.1], we conclude that the same estimate holds for all $0 < t < \infty$.

We next use (4.13) with $t = 1$ to see that for every $\nu \in A_\infty(\mathbb{R}^m)$ we have

$$\begin{aligned} \|\mathbb{E}_\omega g_\omega\|_{L^1(\nu)} &\leq \mathbb{E}_\omega \|g_\omega\|_{L^1(\nu)} \lesssim \mathbb{E}_\omega \left\| \left(\sum_{J \in \mathcal{D}_\omega^d} |\Delta_J g_\omega|^2 \right)^{\frac{1}{2}} \right\|_{L^1(\nu)} \\ &= \left\| \mathbb{E}_\omega \left(\sum_{J \in \mathcal{D}_\omega^d} |\Delta_J g_\omega|^2 \right)^{\frac{1}{2}} \right\|_{L^1(\nu)}. \end{aligned}$$

We use again A_∞ extrapolation, see [11, Theorem 2.1], to obtain the desired estimate. \square

Given a sequence of scalars $\{a_{i,j}\}_{i,j=1}^\infty \subset \mathbb{C}$ and $p, q \in (0, \infty]$, we define

$$\|\{a_{i,j}\}_{i,j}\|_{\ell_j^p(\ell_i^q)} := \left\| \left\{ \|\{a_{i,j}\}_i\|_{\ell_i^q} \right\}_j \right\|_{\ell_j^p}.$$

The following lemma states a variant of the bilinear Marcinkiewicz-Zygmund inequality:

Lemma 4.14. *Let Ω_0 be a probability space. Assume that the sequence of bilinear operators $\{T^\omega\}_{\omega \in \Omega_0}$ satisfies*

$$(4.15) \quad \sup_{\omega \in \Omega_0} \|T^\omega(f_1, f_2)w_1w_2\|_{L^{3/2}} \lesssim \|f_1w_1\|_{L^3} \|f_2w_2\|_{L^3}$$

for all weights $\vec{w} = (w_1, w_2) \in A_{(3,3)}(\mathbb{R}^d)$ and $f_1, f_2 \in L_c^\infty(\mathbb{R}^d)$. Let $1 < s \leq 2$, $\vec{p} = (p_1, p_2)$, and $\vec{q} = (q_1, q_2)$ be such that $1 < p_1, p_2, q_1, q_2 \leq \infty$, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$. Then, the estimate

$$\begin{aligned} & \left\| \mathbb{E}_\omega \left(\sum_{k=1}^\infty \left(\sum_{i,j=1}^\infty |T^\omega(f_{1,i,k}^\omega, f_{2,j,k}^\omega)|^s \right)^{\frac{q}{s}} \right)^{\frac{1}{q}} w_1w_2 \right\|_{L^p} \\ & \lesssim \left\| \left(\mathbb{E}_\omega \|\{f_{1,i,k}^\omega\}_{i,k}\|_{\ell_k^{q_1}(\ell_i^s)}^2 \right)^{\frac{1}{2}} w_1 \right\|_{L^{p_1}} \left\| \left(\mathbb{E}_\omega \|\{f_{2,j,k}^\omega\}_{j,k}\|_{\ell_k^{q_2}(\ell_j^s)}^2 \right)^{\frac{1}{2}} w_2 \right\|_{L^{p_2}} \end{aligned}$$

holds for all weights $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbb{R}^d)$ and for $f_{1,i,k}^\omega, f_{2,j,k}^\omega \in L_c^\infty(\mathbb{R}^d)$.

Proof. Take an arbitrary $\omega \in \Omega_0$. Let $\vec{r} = (r_1, r_2)$ be such that $1 < r_1, r_2 < s$ and set $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2}$. Note that in particular $\frac{1}{r} > \frac{2}{s} \geq 1$. On the other hand, from (4.15) Theorem 1.1 we deduce that T^ω satisfies the estimate

$$\|T^\omega(f_1, f_2)w_1w_2\|_{L^r} \lesssim \|f_1w_1\|_{L^{r_1}} \|f_2w_2\|_{L^{r_2}}.$$

for all $\vec{w} = (w_1, w_2) \in A_{\vec{r}}(\mathbb{R}^d)$, and $f_1, f_2 \in L_c^\infty(\mathbb{R}^d)$. With this in hand, since $0 < r < \max\{r_1, r_2\} < s \leq 2$ we can apply the bilinear Marcinkiewicz-Zygmund inequality in [6, Proposition 5.3] to obtain that

$$\left\| \left(\sum_{i,j=1}^\infty |T^\omega(f_{1,i}, f_{2,j})|^s \right)^{\frac{1}{s}} w_1w_2 \right\|_{L^r} \lesssim \left\| \|\{f_{1,i}\}_i\|_{\ell^s} w_1 \right\|_{L^{r_1}} \left\| \|\{f_{2,j}\}_j\|_{\ell^s} w_2 \right\|_{L^{r_2}},$$

for all $\vec{w} = (w_1, w_2) \in A_{\vec{r}}(\mathbb{R}^d)$ and $f_{1,i}, f_{2,j} \in L_c^\infty(\mathbb{R}^d)$.

By Theorem 1.1, for every $\vec{q} = (q_1, q_2)$, $\vec{t} = (t_1, t_2)$ with $1 < q_1, q_2, t_1, t_2 \leq \infty$ so that $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$ and $\frac{1}{t} = \frac{1}{t_1} + \frac{1}{t_2} > 0$ one can immediately obtain that

$$\begin{aligned} & \left\| \left(\sum_{k=1}^\infty \left(\sum_{i,j=1}^\infty |T^\omega(f_{1,i,k}, f_{2,j,k})|^s \right)^{\frac{q}{s}} \right)^{\frac{1}{q}} w_1w_2 \right\|_{L^t} \\ & \lesssim \left\| \|\{f_{1,i,k}\}_{i,k}\|_{\ell_k^{q_1}(\ell_i^s)} w_1 \right\|_{L^{t_1}} \left\| \|\{f_{2,j,k}\}_{j,k}\|_{\ell_k^{q_2}(\ell_j^s)} w_2 \right\|_{L^{t_2}}. \end{aligned}$$

holds for every $\vec{w} = (w_1, w_2) \in A_{\vec{t}}(\mathbb{R}^d)$ and $f_{1,i,k}, f_{2,j,k} \in L_c^\infty(\mathbb{R}^d)$. Here it is crucial to emphasize all the implicit constants are uniformly bounded on $\omega \in \Omega_0$ because

of (4.15). Applying this for the particular choice $\vec{t} = (2, 2)$ so that $t = 1$ one can easily see that

$$\begin{aligned} & \left\| \mathbb{E}_\omega \left(\sum_{k=1}^{\infty} \left(\sum_{i,j=1}^{\infty} |T^\omega(f_{1,i,k}, f_{2,j,k})|^s \right)^{\frac{q}{s}} \right)^{\frac{1}{q}} w_1 w_2 \right\|_{L^1} \\ & \leq \mathbb{E}_\omega \left\| \left(\sum_{k=1}^{\infty} \left(\sum_{i,j=1}^{\infty} |T^\omega(f_{1,i,k}, f_{2,j,k})|^s \right)^{\frac{q}{s}} \right)^{\frac{1}{q}} w_1 w_2 \right\|_{L^1} \\ & \lesssim \mathbb{E}_\omega \left\| \left\| \{f_{1,i,k}^\omega\}_{i,k}\right\|_{\ell_k^{q_1(\ell_i^s)}} w_1 \right\|_{L^2} \left\| \left\| \{f_{2,j,k}^\omega\}_{j,k}\right\|_{\ell_k^{q_2(\ell_j^s)}} w_2 \right\|_{L^2} \\ & \leq \left\| \left(\mathbb{E}_\omega \left\| \{f_{1,i,k}^\omega\}_{i,k}\right\|_{\ell_k^{q_1(\ell_i^s)}}^2 \right)^{\frac{1}{2}} w_1 \right\|_{L^2} \left\| \left(\mathbb{E}_\omega \left\| \{f_{2,j,k}^\omega\}_{j,k}\right\|_{\ell_k^{q_2(\ell_j^s)}}^2 \right)^{\frac{1}{2}} w_2 \right\|_{L^2} \end{aligned}$$

for every $\vec{w} = (w_1, w_2) \in A_t(\mathbb{R}^d)$ and $f_{1,i,k}, f_{2,j,k} \in L_c^\infty(\mathbb{R}^d)$. We extrapolate from this and Theorem 1.1 readily leads to the desired estimate. \square

Lemma 4.16. *Let Ω_0 be a probability space and let $\{\mathcal{D}_\omega^d\}_{\omega \in \Omega_0}$ be a collection of dyadic grids in \mathbb{R}^d . For every $\omega \in \Omega_0$, $I \in \mathcal{D}_\omega^d$ and $k \geq 0$ set*

$$\Delta_{I,k} f := \sum_{\substack{J \in \mathcal{D}_\omega^d \\ J^{(k)} = I}} \Delta_J f.$$

If $1 < p < \infty$ and $\nu \in A_p(\mathbb{R}^d)$ then

$$\sup_{k \geq 0} \left\| \left(\mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} (M(\Delta_{I,k} g))^2 \right)^{\frac{1}{2}} \right\|_{L^p(\nu)} \lesssim \|g\|_{L^p(\nu)}.$$

Proof. Fix $\omega \in \Omega_0$ and note that clearly

$$\sum_{I \in \mathcal{D}_\omega^d} |\Delta_{I,k} g|^2 = \sum_{I \in \mathcal{D}_\omega^d} \sum_{\substack{J \in \mathcal{D}_\omega^d \\ J^{(k)} = I}} |\Delta_J g|^2 = \sum_{I \in \mathcal{D}_\omega^d} |\Delta_I g|^2.$$

Hence, for every $\nu \in A_2(\mathbb{R}^d)$, the boundedness of M and the dyadic square function on $L^2(\nu)$ imply

$$\begin{aligned} & \sup_k \left\| \left(\mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} (M(\Delta_{I,k} g))^2 \right)^{\frac{1}{2}} \right\|_{L^2(\nu)}^2 = \sup_k \mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} \int_{\mathbb{R}^d} (M(\Delta_{I,k} g))^2 \nu dx \\ & \lesssim \sup_k \mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} \int_{\mathbb{R}^d} |\Delta_{I,k} g|^2 \nu dx = \mathbb{E}_\omega \left\| \left(\sum_{I \in \mathcal{D}_\omega^d} |\Delta_I g|^2 \right)^{\frac{1}{2}} \right\|_{L^2(\nu)}^2 \lesssim \|g\|_{L^2(\nu)}^2. \end{aligned}$$

We can now invoke Rubio de Francia extrapolation theorem to obtain at once the desired inequality. \square

Lemma 4.17. *Let Ω_0 be a probability space and let $\{\mathcal{D}_\omega^d\}_{\omega \in \Omega_0}$ be a collection of dyadic grids in \mathbb{R}^d . Suppose we have a sequence $\{\gamma_I^\omega\}_{\omega \in \Omega_0, I \in \mathcal{D}_\omega^d}$ of complex numbers*

such that

$$\sup_{\omega \in \Omega_0} \left(\sup_{I_0 \in \mathcal{D}_\omega^d} \frac{1}{|I_0|} \sum_{I \in \mathcal{D}_\omega^d, I \subset I_0} |\gamma_I^\omega|^2 \right)^{\frac{1}{2}} \leq 1.$$

Then, for every $1 < p < \infty$ and $\nu \in A_p(\mathbb{R}^d)$ we have

$$\left\| \left(\mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} |\gamma_I^\omega \langle g \rangle_I|^2 \frac{1_I}{|I|} \right)^{\frac{1}{2}} \right\|_{L^p(\nu)} \lesssim \|g\|_{L^p(\nu)}.$$

Proof. By Rubio de Francia extrapolation we just need to obtain the case $p = 2$. Assuming that $\nu \in A_2(\mathbb{R}^d)$ one can see that if we write $f_\omega = \sum_{I \in \mathcal{D}_\omega^d} \gamma_I^\omega \langle g \rangle_I h_I^\eta$ (for some fixed $\eta \in \{0, 1\}^d \setminus \{0\}^d$) then (4.7) easily yields

$$\begin{aligned} \left\| \left(\mathbb{E}_\omega \sum_{I \in \mathcal{D}_\omega^d} |\gamma_I^\omega \langle g \rangle_I|^2 \frac{1_I}{|I|} \right)^{\frac{1}{2}} \right\|_{L^2(\nu)}^2 &= \mathbb{E}_\omega \left\| \left(\sum_{I \in \mathcal{D}_\omega^d} |\Delta_I f_\omega|^2 \right)^{\frac{1}{2}} \right\|_{L^2(\nu)}^2 \lesssim \mathbb{E}_\omega \|f_\omega\|_{L^2(\nu)}^2 \\ &= \mathbb{E}_\omega \left\| \sum_{I \in \mathcal{D}_\omega^d} \gamma_I^\omega \langle g \rangle_I h_I^\eta \right\|_{L^2(\nu)}^2 \lesssim \|g\|_{L^2(\nu)}^2, \end{aligned}$$

where we have used the $L^2(\nu)$ -boundedness of the dyadic square function and the dyadic paraproduct (with the dyadic-BMO function $b_\omega = \sum_{I \in \mathcal{D}_\omega^d} \gamma_I^\omega h_I^\eta$), see for instance [13]. \square

4.1. Proof of Theorem 4.2. To set the stage we fix $\vec{p} = (p_1, p_2)$ such that $1 < p_1, p_2 < \infty$ and set $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. We also pick $\vec{w} = (w_1, w_2) \in \mathbb{A}_{\vec{p}, n, m}$.

Let f_1, f_2, f_3 be finite linear combinations of tensor products of bounded and compactly supported functions. We apply the representation theorem from [29] (see (4.10)) to both T_n and T_m . For T_n we will use ω_1 to denote the random parameter and U_{ω_1, u_1}^k for the different bilinear model operators. Analogously, for T_m these will be respectively ω_2 and U_{ω_2, u_2}^v . We also let $\mathbb{E}_\omega := \mathbb{E}_{\omega_1} \mathbb{E}_{\omega_2}$. Using this notation (4.10) readily leads to

$$\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle = C_{T_n} C_{T_m} \mathbb{E}_\omega \sum_{k \in \mathbb{N}_0^3} \sum_{v \in \mathbb{N}_0^3} \sum_{u_1=1}^{C_n} \sum_{u_2=1}^{C_m} c_{k,v} \langle U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2), f_3 \rangle,$$

where $c_{k,v} := 2^{-\max_i k_i \frac{\alpha_1}{2} - \max_i v_i \frac{\alpha_2}{2}}$ and α_1 and α_2 are respectively the parameters in the Hölder continuity assumptions of the kernels of T_n and T_m . It is not too hard to show, and this also follows directly from [30], that

$$|\langle U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2), f_3 \rangle| \lesssim \|f_1\|_{L^3} \|f_2\|_{L^3} \|f_3\|_{L^3},$$

uniformly in ω_1, ω_2, k, v . Since finite linear combinations of tensor products of bounded and compactly supported functions are dense in L^3 , this implies via the above representation that $T_n \otimes T_m: L^3 \times L^3 \rightarrow L^{\frac{3}{2}}$ is bounded, and that the representation holds for $f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$.

Take next $f_1, f_2 \in L_c^\infty(\mathbb{R}^{n+m})$. It is easy to see that if $p \geq 1$ we have

$$\|T_n \otimes T_m(f_1, f_2)w\|_{L^p} \lesssim \sum_{k \in \mathbb{N}_0^3} \sum_{v \in \mathbb{N}_0^3} \sum_{u_1, u_2} c_{k,v} \|\mathbb{E}_\omega U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2)w\|_{L^p},$$

and if $p < 1$, then

$$\|T_n \otimes T_m(f_1, f_2)w\|_{L^p}^p \lesssim \sum_{k \in \mathbb{N}_0^3} \sum_{v \in \mathbb{N}_0^3} \sum_{u_1, u_2} c_{k,v}^p \|\mathbb{E}_\omega U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2)w\|_{L^p}^p.$$

We claim that

$$(4.18) \quad \|\mathbb{E}_\omega U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2)w\|_{L^p} \lesssim C_{k,v} \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}$$

holds with $C_{k,v} := 2^{\max_i k_i \frac{\alpha_1}{4} + \max_i v_i \frac{\alpha_2}{4}}$. With this in hand, all the previous observations readily lead us to the desired estimate for $f_1, f_2 \in L_c^\infty(\mathbb{R}^{n+m})$. By density this completes the proof of Theorem 4.2.

We are left with showing our claim (4.18). Let us observe that when $p \geq 1$ we could take the expectation out of the L^p norm and prove a uniform estimate for all ω . However, in the case $p < 1$ we cannot do this, and therefore we need to keep working with the expectation. We consider two cases, whether U_{ω_2, u_2}^v is a bilinear dyadic shift or is a bilinear dyadic paraproduct.

Case 1: Proof of (4.18) when U_{ω_2, u_2}^v is a dyadic bilinear shift.

To simplify the notation we write $U_{\omega_1}^k = U_{\omega_1, u_1}^k$ and $U_{\omega_2}^v = U_{\omega_2, u_2}^v$. In this scenario $U_{\omega_2}^v$ is of the form $\mathcal{U}_{\mathcal{D}_{\omega_2}^m}^v$ (see (4.8)). Recall that dyadic shifts have three different expressions depending where the non-cancellative Haar function is located. For now we consider the case on which the non-cancellative Haar function is in the “second” position, that is,

$$U_{\omega_2}^v(g_1, g_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} \sum_{\substack{J_1, J_2, J_3 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} a_{V, (J_i)}^{\omega_2} \langle g_1, h_{J_1} \rangle, \langle g_2, h_{J_2}^0 \rangle h_{J_3}.$$

The other possibilities of $U_{\omega_2}^v$ will be handled in a similar manner.

If we next write $U_\omega := U_{\omega_1}^k \otimes U_{\omega_2}^v$ one can easily see that

$$(4.19) \quad U_\omega(f_1, f_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} \sum_{\substack{J_1, J_2, J_3 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} a_{V, (J_i)}^{\omega_2} U_{\omega_1}^k(\langle f_1, h_{J_1} \rangle_2, \langle f_2, h_{J_2}^0 \rangle_2) \otimes h_{J_3},$$

where $\langle \cdot, \cdot \rangle_2$ stands for the inner product with respect to the variable x_2 —recall that we always write $x = (x_1, x_2)$ with $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^m$. Here it is useful to remember that the Haar functions live in \mathbb{R}^m , hence they only depend on x_2 .

Let us recall that by assumption $\vec{w} = (w_1, w_2) \in \mathbb{A}_{\vec{p}, n, m}$, see Definition 4.1. For $w = w_1 w_2$ we have by Hölder’s inequality that

$$\operatorname{ess\,sup}_{x_1 \in \mathbb{R}^n} [w(x_1, \cdot)^p]_{A_{2p}(\mathbb{R}^m)} \leq \operatorname{ess\,sup}_{x_1 \in \mathbb{R}^n} [w_1(x_1, \cdot)^{p_1}]_{A_{p_1}(\mathbb{R}^m)}^{\frac{p}{p_1}} [w_2(x_1, \cdot)^{p_2}]_{A_{p_2}(\mathbb{R}^m)}^{\frac{p}{p_2}} < \infty.$$

Hence $w(x_1, \cdot)^p \in A_\infty(\mathbb{R}^m)$ uniformly in $x_1 \in \mathbb{R}^n$. With $x_1 \in \mathbb{R}^n$ fixed, we can invoke Lemma 4.12 with $g_\omega := U_\omega(f_1, f_2)(x_1, \cdot)$ and the weight $w(x_1, \cdot)^p \in A_\infty(\mathbb{R}^m)$. Then, we integrate over $x_1 \in \mathbb{R}^n$ and after some straightforward computations we arrive at

$$(4.20) \quad \iint_{\mathbb{R}^{n+m}} |\mathbb{E}_\omega U_\omega(f_1, f_2)|^p w^p dx \\ \lesssim \iint_{\mathbb{R}^{n+m}} \left(\mathbb{E}_\omega \left[\sum_{\substack{J_3, V \in \mathcal{D}_{\omega_2}^m \\ J_3^{(v_3)} = V}} \left| \sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} a_{V, (J_i)}^{\omega_2} U_{\omega_1}^k(\langle f_1, h_{J_1} \rangle_2, \langle f_2, h_{J_2}^0 \rangle_2) \otimes h_{J_3} \right|^2 \right]^{\frac{1}{2}} \right)^p w^p dx.$$

Let $s = (\frac{8m}{\alpha})'$ so that $\frac{m}{s'} = \frac{\alpha}{8}$ and clearly $1 < s \leq \frac{8}{7} < 2$. Let $J_3, V \in \mathcal{D}_{\omega_2}^m$ be such that $J_3^{(v_3)} = V$. From the size estimate of the coefficients $a_{V, (J_i)}^{\omega_2}$ we get that

$$(4.21) \quad \left| \sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} a_{V, (J_i)}^{\omega_2} U_{\omega_1}^k(\langle f_1, h_{J_1} \rangle_2, \langle f_2, h_{J_2}^0 \rangle_2) \otimes h_{J_3} \right| \\ \leq 2^{\max_i v_i \frac{\alpha}{4}} \left(\sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} |U_{\omega_1}^k(f_{1, J_1, V}, f_{2, J_2, V})|^s \right)^{\frac{1}{s}} \otimes 1_{J_3},$$

where we abbreviated $f_{1, J_1, V} := \frac{|J_1|^{\frac{1}{2}}}{|V|} \langle f_1, h_{J_1} \rangle_2$ and $f_{2, J_2, V} := \frac{|J_2|^{\frac{1}{2}}}{|V|} \langle f_2, h_{J_2}^0 \rangle_2$. Using this in (4.20) we have

$$(4.22) \quad \left(\iint_{\mathbb{R}^{n+m}} |\mathbb{E}_\omega U_\omega(f_1, f_2)|^p w^p dx \right)^{\frac{1}{p}} \\ \lesssim 2^{\max_i v_i \frac{\alpha}{4}} \left(\iint_{\mathbb{R}^{n+m}} \left(\mathbb{E}_\omega \left[\sum_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} |U_{\omega_1}^k(f_{1, J_1, V}, f_{2, J_2, V})|^s \right)^{\frac{2}{s}} \otimes 1_V \right]^{\frac{1}{2}} \right)^p w^p dx \right)^{\frac{1}{p}}.$$

We next introduce some notation. Write $M^2 f(x) = M(f(x_1, \cdot))(x_2)$, where M stands for the Hardy-Littlewood maximal function (over cubes) in \mathbb{R}^m and

$$(4.23) \quad \Delta_{V, v_1}^2 f_1(x) = \sum_{\substack{J_1 \in \mathcal{D}_{\omega_2}^m \\ J_1^{(v_1)} = V}} \Delta_{J_1}(f_1(x_1, \cdot))(x_2), \quad V \in \mathcal{D}_{\omega_2}^m,$$

and when $v_1 = 0$ we will simply write Δ_V^2 . Recalling the definition of $f_{1, J_2, V}$ and (4.7) one can see that

$$|f_{1, J_1, V}| = \frac{|J_1|^{\frac{1}{2}}}{|V|} |\langle f_1, h_{J_1} \rangle_2| = \frac{|J_1|^{\frac{1}{2}}}{|V|} |\langle \Delta_{V, v_1}^2 f_1, h_{J_1} \rangle_2| \leq \frac{1}{|V|} \int_{J_1} |\Delta_{V, v_1}^2 f_1| dx_2,$$

and, analogously,

$$|f_{2,J_2,V}| = \frac{|J_2|^{\frac{1}{2}}}{|V|} |\langle f_2, h_{J_2}^0 \rangle_2| = \frac{1}{|V|} |\langle f_2, 1_{J_2} \rangle_2| \leq \frac{1}{|V|} \int_{J_2} |f_2| dx_2.$$

Thus,

$$(4.24) \quad \sum_{\substack{J_1 \in \mathcal{D}_{\omega_2}^m \\ J_1^{(v_1)} = V}} |f_{1,J_1,V}| \otimes 1_V \leq M^2(\Delta_{V,v_1}^2 f_1) \quad \text{and} \quad \sum_{\substack{J_2 \in \mathcal{D}_{\omega_2}^m \\ J_2^{(v_2)} = V}} |f_{2,J_2,V}| \otimes 1_V \leq M^2 f_2.$$

Recall that $(w_1(\cdot, x_2), w_2(\cdot, x_2)) \in A_{\vec{p}}(\mathbb{R}^n)$ for a.e. $x_2 \in \mathbb{R}^m$ uniformly on x_2 (see Definition 4.1), and that $U_{\omega_1}^k$ satisfies the estimate (4.11). Therefore, we may invoke Lemma 4.14 with the choice $q_1 = 2$, $q_2 = \infty$, $q = 2$ and with s as above to conclude that

$$\begin{aligned} & \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega} \left[\sum_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} |U_{\omega_1}^k(f_{1,J_1,V}, f_{2,J_2,V})|^s \right)^{\frac{2}{s}} \otimes 1_V(x_2) \right]^{\frac{1}{2}} \right)^p w(\cdot, x_2)^p dx_1 \right)^{\frac{1}{p}} \\ & \lesssim (1 + \max_i k_i) \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \sum_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_1 \in \mathcal{D}_{\omega_2}^m \\ J_1^{(v_1)} = V}} |f_{1,J_1,V}|^s \otimes 1_V(x_2) \right)^{\frac{2}{s}} \right)^{\frac{p_1}{2}} w_1(\cdot, x_2)^{p_1} dx_1 \right)^{\frac{1}{p_1}} \\ & \quad \times \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \left(\sup_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_2 \in \mathcal{D}_{\omega_2}^m \\ J_2^{(v_2)} = V}} |f_{2,J_2,V}|^s \otimes 1_V(x_2) \right)^{\frac{2}{s}} \right)^{\frac{p_2}{2}} w_2(\cdot, x_2)^{p_2} dx_1 \right)^{\frac{1}{p_2}} \\ & \leq (1 + \max_i k_i) \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \sum_{V \in \mathcal{D}_{\omega_2}^m} (M^2(\Delta_{V,v_1}^2 f_1))^2 \right)^{\frac{p_1}{2}} w_1(\cdot, x_2)^{p_1} dx_1 \right)^{\frac{1}{p_1}} \\ & \quad \times \left(\int_{\mathbb{R}^n} (M^2 f_2)^{p_2} w_2(\cdot, x_2)^{p_2} dx_1 \right)^{\frac{1}{p_2}}, \end{aligned}$$

where in the second estimate we have used that $s > 1$ and (4.24). This, Hölder's inequality, the fact that by assumption $w_1(x_1, \cdot)^{p_1} \in A_{p_1}(\mathbb{R}^m)$ and $w_2(x_1, \cdot)^{p_2} \in A_{p_2}(\mathbb{R}^m)$ uniformly for a.e. $x_1 \in \mathbb{R}^n$ (see Definition 4.1), and Lemma 4.16 readily give

$$\begin{aligned} & \left(\iint_{\mathbb{R}^{n+m}} \left(\mathbb{E}_{\omega} \left[\sum_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} |U_{\omega_1}^k(f_{1,J_1,V}, f_{2,J_2,V})|^s \right)^{\frac{2}{s}} \otimes 1_V \right]^{\frac{1}{2}} \right)^p w^p dx \right)^{\frac{1}{p}} \\ & \lesssim (1 + \max_i k_i) \left(\iint_{\mathbb{R}^{n+m}} \left(\mathbb{E}_{\omega} \sum_{V \in \mathcal{D}_{\omega_2}^m} (M^2(\Delta_{V,v_1}^2 f_1))^2 \right)^{\frac{p_1}{2}} w_1^{p_1} dx \right)^{\frac{1}{p_1}} \\ & \quad \times \left(\iint_{\mathbb{R}^{n+m}} (M^2 f_2)^{p_2} w_2^{p_2} dx \right)^{\frac{1}{p_2}} \end{aligned}$$

$$\lesssim (1 + \max_i k_i) \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}.$$

Plugging this into (4.22) we arrive at (4.18) in the case when $U_{\omega_1} \otimes U_{\omega_2}$ is of the form (4.19).

Let us now consider the cases in which $U_{\omega_2}^v$ is a bilinear dyadic shift of some other form. The case in which the non-cancellative Haar function is located in the “first” position (i.e., our operator is written in terms of h_{J_1} , $h_{J_2}^0$ and h_{J_3}) is clearly symmetric to the case we treated already (one just needs to switch the roles of f_1 and f_2). Thus, we only need to consider the case when

$$U_{\omega}(f_1, f_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} \sum_{\substack{J_1, J_2, J_3 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} a_{V, (J_i)}^{\omega_2} U_{\omega_1}^k(\langle f_1, h_{J_1} \rangle_2, \langle f_2, h_{J_2} \rangle_2) \otimes h_{J_3}^0.$$

This time, rather than using the lower bound for the square function estimate, we just proceed as in (4.21), putting absolute values inside and using Hölder’s inequality with s as before. This gives that

$$|U_{\omega}(f_1, f_2)| \leq 2^{\max v_i \alpha / 4} \sum_{V \in \mathcal{D}_{\omega_2}^m} \left(\sum_{\substack{J_1, J_2 \in \mathcal{D}_{\omega_2}^m \\ J_i^{(v_i)} = V}} |U_{\omega_1}^k(f_{1, J_1, V}, f_{2, J_2, V})|^s \right)^{\frac{1}{s}} \otimes 1_V,$$

where $f_{1, J_1, V} := \frac{|J_1|^{\frac{1}{2}}}{|V|} \langle f_1, h_{J_1} \rangle_2$ and $f_{2, J_2, V} := \frac{|J_2|^{\frac{1}{2}}}{|V|} \langle f_2, h_{J_2} \rangle_2$. We can proceed much as before, with the difference that Lemma 4.14 is used with the exponents $q_1 = q_2 = 2$, $q = 1$ and also that since h_{J_2} is cancellative we have the analog of the first estimate in (4.24) for $f_{2, J_2, V}$. This gives that

$$\begin{aligned} & \| \mathbb{E}_{\omega} U_{\omega}(f_1, f_2) w \|_{L^p} \\ & \lesssim 2^{\max v_i \alpha / 4} (1 + \max_i k_i) \left\| \left(\mathbb{E}_{\omega} \sum_{V \in \mathcal{D}_{\omega_2}^m} (M^2(\Delta_{V, v_1}^2 f_1))^2 \right)^{\frac{1}{2}} w_1 \right\|_{p_1} \\ & \quad \times \left\| \left(\mathbb{E}_{\omega} \sum_{V \in \mathcal{D}_{\omega_2}^m} (M^2(\Delta_{V, v_2}^2 f_2))^2 \right)^{\frac{1}{2}} w_2 \right\|_{p_2} \\ & \lesssim 2^{\max v_i \alpha / 4} (1 + \max_i k_i) \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}, \end{aligned}$$

where in the last estimate we have used that $w_1(x_1, \cdot)^{p_1} \in A_{p_1}(\mathbb{R}^m)$ and $w_2(x_1, \cdot)^{p_2} \in A_{p_2}(\mathbb{R}^m)$ uniformly for a.e. $x_1 \in \mathbb{R}^n$ (see Definition 4.1), and Lemma 4.16. This ends our treatment of the case when U_{ω_2} is a dyadic bilinear shift.

Case 2: Proof of (4.18) when U_{ω_2, u_2}^v is a dyadic bilinear paraproduct.

As before to simplify the notation we write $U_{\omega_1}^k = U_{\omega_1, u_1}^k$ and $U_{\omega_2}^v = U_{\omega_2, u_2}^v$. Here we assume that $U_{\omega_2}^v$ is of the form $\mathcal{U}_{\mathcal{D}_{\omega_2}^m}$, see (4.9), and hence $v = (0, 0, 0)$. On the other hand, $U_{\omega_1}^k$ is any model operator. The argument for this case is easier than the previous one: before we used the Marcinkiewicz-Zygmund type estimate in

Lemma 4.14 to deal with the complexity of $U_{\omega_2}^v$ and in this scenario the complexity is zero.

Recall that dyadic bilinear paraproducts have three different expressions depending on where the cancellative function is located. We first consider the case where

$$U_{\omega_2}^v(g_1, g_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} a_V^{\omega_2} \langle g_1 \rangle_V \langle g_2 \rangle_V h_V,$$

and the other cases will be treated below.

We write as before $U_\omega := U_{\omega_1}^k \otimes U_{\omega_2}^v$ so that

$$U_\omega(f_1, f_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} a_V^{\omega_2} U_{\omega_1}^k(\langle f_1 \rangle_{V,2}, \langle f_2 \rangle_{V,2}) \otimes h_V.$$

As shown in the previous case $w_1(x_1, \cdot)^p w_2(x_1, \cdot)^p \in A_\infty(\mathbb{R}^m)$ uniformly in $x_1 \in \mathbb{R}^n$. With $x_1 \in \mathbb{R}^n$ fixed, we can use Lemma 4.12 with $g_\omega := U_\omega(f_1, f_2)(x_1, \cdot)$ and the weight $w_1(x_1, \cdot)^p w_2(x_1, \cdot)^p \in A_\infty(\mathbb{R}^m)$. After that we integrate on x_1 and arrive at

$$\|\mathbb{E}_\omega U_\omega(f_1, f_2)w\|_{L^p} \lesssim \left\| \mathbb{E}_\omega \left(\sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} U_{\omega_1}^k(\langle f_1 \rangle_{V,2}, \langle f_2 \rangle_{V,2}) \otimes h_V|^2 \right)^{\frac{1}{2}} w \right\|_{L^p}.$$

Recall that $(w_1(\cdot, x_2), w_2(\cdot, x_2)) \in A_{\vec{p}}(\mathbb{R}^n)$ for a.e. $x_2 \in \mathbb{R}^m$ uniformly on x_2 (see Definition 4.1) and that U_{ω_1} satisfies (4.11). Thus, we may with fixed $x_2 \in \mathbb{R}^m$ invoke Lemma 4.14 with the choice $q_1 = 2$, $q_2 = \infty$ and $q = 2$ in the very special case where the inner ℓ^s -sums have only one non-zero term. This gives that

$$\begin{aligned} & \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_\omega \left[\sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} U_{\omega_1}^k(\langle f_1 \rangle_{V,2}, \langle f_2 \rangle_{V,2}) \otimes h_V(x_2)|^2 \right]^{\frac{1}{2}} \right)^p w(\cdot, x_2)^p dx_1 \right)^{\frac{1}{p}} \\ & \lesssim (1 + \max_i k_i) \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} \langle f_1 \rangle_{V,2} \otimes h_V(x_2)|^2 \right)^{\frac{p_1}{2}} w_1(\cdot, x_2)^{p_1} dx_1 \right)^{\frac{1}{p_1}} \\ & \quad \times \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \sup_{V \in \mathcal{D}_{\omega_2}^m} |\langle f_2 \rangle_{V,2} \otimes 1_V(x_2)|^2 \right)^{\frac{p_2}{2}} w_2(\cdot, x_2)^{p_2} dx_1 \right)^{\frac{1}{p_2}} \\ & \lesssim (1 + \max_i k_i) \left(\int_{\mathbb{R}^n} \left(\mathbb{E}_{\omega_2} \sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} \langle f_1 \rangle_{V,2} \otimes h_V(x_2)|^2 \right)^{\frac{p_1}{2}} w_1(\cdot, x_2)^{p_1} dx_1 \right)^{\frac{1}{p_1}} \\ & \quad \times \left(\int_{\mathbb{R}^n} (M^2 f_2)^{p_2} w_2(\cdot, x_2)^{p_2} dx_1 \right)^{\frac{1}{p_2}}, \end{aligned}$$

where we have used the pointwise inequality $\sup_{V \in \mathcal{D}_{\omega_2}^m} |\langle f_2 \rangle_{V,2}| \otimes 1_V \leq M^2 f_2$. Using the obtained estimates, Hölder's inequality, the fact that by assumption $w_1(x_1, \cdot)^{p_1} \in A_{p_1}(\mathbb{R}^m)$ and $w_2(x_1, \cdot)^{p_2} \in A_{p_2}(\mathbb{R}^m)$ uniformly for a.e. $x_1 \in \mathbb{R}^n$ (see Definition 4.1), and Lemma 4.17 we conclude that

$$\begin{aligned} \|\mathbb{E}_\omega U_\omega(f_1, f_2)w\|_{L^p} &\lesssim (1 + \max_i k_i) \left\| \left(\mathbb{E}_\omega \sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} \langle f_1 \rangle_{V,2}|^2 \otimes \frac{1_V}{|V|} \right)^{\frac{1}{2}} w_1 \right\|_{L^{p_1}} \\ &\quad \times \left(\iint_{\mathbb{R}^{n+m}} (M^2 f_2)^{p_2} w_2^{p_2} dx \right)^{\frac{1}{p_2}} \lesssim (1 + \max_i k_i) \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}. \end{aligned}$$

This gives the desired estimate in the present case. Let us finally consider the remaining two possibilities for $U_{\omega_2}^v$. By symmetry, we may suppose that

$$U_\omega(f_1, f_2) = \sum_{V \in \mathcal{D}_{\omega_2}^m} a_V^{\omega_2} U_{\omega_1}^k(\langle f_1 \rangle_2, \langle f_2, h_V \rangle_{V,2}) \otimes \frac{1_V}{|V|}.$$

and the other cases is treated similarly switching the roles of f_1 and f_2 . This time we do not use the lower square function estimate and directly apply Lemma 4.14 with exponents $q_1 = q_2 = 2$, $q = 1$ and again the inner ℓ^s -sums have only one non-zero term. Hence, much as before we obtain

$$\begin{aligned} \|\mathbb{E}_\omega U_\omega(f_1, f_2)w\|_{L^p} &\lesssim (1 + \max_i k_i) \left\| \left(\mathbb{E}_\omega \sum_{V \in \mathcal{D}_{\omega_2}^m} |a_V^{\omega_2} \langle f_1 \rangle_{V,2}|^2 \otimes \frac{1_V}{|V|} \right)^{\frac{1}{2}} w_1 \right\|_{L^{p_1}} \\ &\quad \times \left\| \left(\mathbb{E}_\omega \sum_{V \in \mathcal{D}_{\omega_2}^m} |\langle f_2, h_V \rangle_{V,2}|^2 \otimes \frac{1_V}{|V|} \right)^{\frac{1}{2}} w_2 \right\|_{L^{p_2}} \\ &\lesssim (1 + \max_i k_i) \|f_1 w_1\|_{L^{p_1}} \|f_2 w_2\|_{L^{p_2}}, \end{aligned}$$

where the term corresponding to f_1 has been treated as before. For f_2 , using the notation introduced in (4.23), one first sees that (4.7) gives

$$\begin{aligned} |\langle f_2, h_V \rangle_{V,2}| \otimes \frac{1_V}{|V|^{\frac{1}{2}}} &= |\langle \Delta_V^2 f_2, h_V \rangle_{V,2}| \otimes \frac{1_V}{|V|^{\frac{1}{2}}} \\ &\leq \frac{1}{|V|} \int_V |\Delta_V^2 f_2| dx_2 \otimes 1_V \leq M^2(\Delta_V^2 f_2) \end{aligned}$$

and therefore Lemma 4.16 with $k = 0$, along with the fact that $w_2(x_1, \cdot)^{p_2} \in A_{p_2}(\mathbb{R}^m)$ uniformly for a.e. $x_1 \in \mathbb{R}^n$ (see Definition 4.1), yields

$$\begin{aligned} \left\| \left(\mathbb{E}_\omega \sum_{V \in \mathcal{D}_{\omega_2}^m} |\langle f_2, h_V \rangle_{V,2}|^2 \otimes \frac{1_V}{|V|} \right)^{\frac{1}{2}} w_2 \right\|_{L^{p_2}} \\ \leq \left\| \left(\mathbb{E}_\omega \sum_{V \in \mathcal{D}_{\omega_2}^m} (M^2(\Delta_V^2 f_2))^2 \right)^{\frac{1}{2}} w_2 \right\|_{L^{p_2}} \lesssim \|f_2 w_2\|_{L^{p_2}}. \end{aligned}$$

This concludes the proof of (4.18), and hence that of Theorem 4.2. \square

4.2. Proof of Corollary 4.3. We proceed by extrapolation. Fix $1 < q_1, q_2 < \infty$ and let $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} > 0$. Given $f_1, f_2 \in L_c^\infty(\mathbb{R}^{n+m})$ define for every $x_1 \in \mathbb{R}^n$ and for $i = 1, 2$

$$(4.25) \quad F(x_1) := \|T_n \otimes T_m(f_1, f_2)(x_1, \cdot)\|_{L^q(\mathbb{R}^m)} \quad \text{and} \quad F_i(x_1) := \|f_i(x_1, \cdot)\|_{L^{q_i}(\mathbb{R}^m)}$$

Suppose $\vec{w} = (w_1, w_2) \in A_{\vec{q}}(\mathbb{R}^n)$ with $\vec{q} = (q_1, q_2)$ and as usual we define $w = w_1 w_2$. It is easy to see that $(w_1 \otimes 1, w_2 \otimes 1) \in \mathbb{A}_{\vec{q}, n, m}$. From Theorem 4.2 —here it is crucial that $1 < q_1, q_2 < \infty$ — we see that

$$(4.26) \quad \begin{aligned} \|Fw\|_{L^q(\mathbb{R}^n)} &= \|T_n \otimes T_m(f_1, f_2)(w \otimes 1)\|_{L^q(\mathbb{R}^{n+m})} \\ &\lesssim \|f_1(w_1 \otimes 1)\|_{L^{q_1}(\mathbb{R}^{n+m})} \|f_2(w_2 \otimes 1)\|_{L^{q_2}(\mathbb{R}^{n+m})} \\ &= \|F_1 w_1\|_{L^{q_1}(\mathbb{R}^n)} \|F_2 w_2\|_{L^{q_2}(\mathbb{R}^n)}. \end{aligned}$$

We can extrapolate from this estimate with the family \mathcal{F} consisting on the triples (F, F_1, F_2) and Theorem 2.1 (with $\vec{r} = (1, 1, 1)$) gives with $\vec{p} = (p_1, p_2)$ and $1 < p_1, p_2 \leq \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ that

$$\|Fw\|_{L^p(\mathbb{R}^n)} \lesssim \|F_1 w_1\|_{L^{p_1}(\mathbb{R}^n)} \|F_2 w_2\|_{L^{p_2}(\mathbb{R}^n)}$$

for all $(w_1, w_2) \in A_{\vec{p}}(\mathbb{R}^n)$. This in the special case $w_1 \equiv 1$ and $w_2 \equiv 1$ gives us the estimate

$$(4.27) \quad \begin{aligned} \|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} &= \|F\|_{L^p(\mathbb{R}^n)} \\ &\lesssim \|F_1\|_{L^{p_1}(\mathbb{R}^n)} \|F_2\|_{L^{p_2}(\mathbb{R}^n)} \lesssim \|f_1\|_{L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))} \|f_2\|_{L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))} \end{aligned}$$

for all $f_1, f_2 \in L_c^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$. This ends the proof of Corollary 4.3. \square

4.3. Proof of Theorem 4.5. We first observe that a trivial density argument allows us to see that Corollary 4.3 readily gives the desired estimate for all $f_1 \in L^{p_1}(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L^{p_2}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$ provided $1 < p_1, p_2, q_1, q_2 < \infty$. This means that we only need to consider the cases $q_1 = \infty$ or $q_2 = \infty$ and/or $p_1 = \infty$ or $p_2 = \infty$. In these scenarios we need to justify why the operators are well-defined and also get the desired estimates. To accomplish all these we split the argument in two main steps which are almost independent, and each of them is interesting in its own right. In **Case 1** we use sparse domination techniques and the main goal is to treat the cases $q_1 = \infty$ or $q_2 = \infty$ on which our main extrapolation result is not useful. However, we prove more, mostly because in **Case 2** we will need to know that the operators are well-defined for some class of functions. As a result, **Case 1** deals with the exponents $1 < p_1, p_2 \leq \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $1 < q_1, q_2 \leq \infty$ with $0 < \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \leq 1$. We note that the restriction $q \geq 1$ is natural since we use sparse domination and duality, hence we need to be in the Banach range at least for q . In this direction, if $q_1 = \infty$ (resp. $q_2 = \infty$) then $q = q_2 > 1$ (resp. $q = q_1 > 1$), thus in these two cases we do not have a real restriction. In **Case 2**, where $1 < q_1, q_2 < \infty$, we are going to use our extrapolation result again, Theorem 2.1, and in doing so we need to take a detour to show that the operators are well-defined and this is where **Case 1** is invoked in a particular case where the q_i 's are finite. It is important to note that at the end-point cases (i.e., when some of the p_i 's

or q_i 's are infinity) we are able to obtain the desired estimates for functions in the corresponding spaces and not just in $L_c^\infty(\mathbb{R}^{n+m})$. This requires to justify why the operators are well-defined and we do this by duality using the two adjoints of our operators. This makes some of our argument tedious needing to consider several cases quite carefully. If one were interested in obtaining the desired estimates just for compactly supported functions one could skip some of those cases.

Case 1: $1 < p_1, p_2 \leq \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $1 < q_1, q_2 \leq \infty$ with $0 < \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \leq 1$.

We first prove certain vector-valued sparse domination result using dyadic model operators. The starting point is that as observed above we already know that for $f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ we have

$$\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle = C_{T_n} C_{T_m} \mathbb{E}_\omega \sum_{k \in \mathbb{N}_0^3} \sum_{v \in \mathbb{N}_0^3} \sum_{u_1=1}^{C_n} \sum_{u_2=1}^{C_m} c_{k,v} \langle U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v(f_1, f_2), f_3 \rangle,$$

where $c_{k,v} := 2^{-\max_i k_i \frac{\alpha_1}{2} - \max_i v_i \frac{\alpha_2}{2}}$ and α_1 and α_2 are respectively the parameters in the Hölder continuity assumptions of the kernels of T_n and T_m . Denote $U_\omega := U_{\omega_1, u_1}^k \otimes U_{\omega_2, u_2}^v$.

Then, with $1 < q_1, q_2 \leq \infty$ so that $0 < 1/q := 1/q_1 + 1/q_2 \leq 1$ and for any $f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ we write

$$\langle U_\omega(f_1, f_2), f_3 \rangle = \sum_{K \in \mathcal{D}_{\omega_1, u_1}^n} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_{\omega_1, u_1}^n \\ I_i^{(k_i)} = K}} a_{K, (I_i)}^{\omega_1, u_1} \langle U_{\omega_2, u_2}^v(\langle f_1, \tilde{h}_{I_1} \rangle_1, \langle f_2, \tilde{h}_{I_2} \rangle_1), \langle f_3, \tilde{h}_{I_3} \rangle_1 \rangle,$$

where $\tilde{h}_I \in \{h_I, h_I^0\}$. It is not hard to adapt the bilinear sparse domination argument of model operators [29, Section 5] and deduce that for all $f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ we have

$$\begin{aligned} |\langle U_\omega(f_1, f_2), f_3 \rangle| &\lesssim (1 + \max_i k_i) \Lambda_{\mathcal{S}}(\|f_1\|_{L^{q_1}(\mathbb{R}^m)}, \|f_2\|_{L^{q_2}(\mathbb{R}^m)}, \|f_3\|_{L^{q'}(\mathbb{R}^m)}) \\ &:= (1 + \max_i k_i) \sum_{Q \in \mathcal{S}} |Q| \langle \|f_1\|_{L^{q_1}(\mathbb{R}^m)} \rangle_Q \langle \|f_2\|_{L^{q_2}(\mathbb{R}^m)} \rangle_Q \langle \|f_3\|_{L^{q'}(\mathbb{R}^m)} \rangle_Q \end{aligned}$$

for some dyadic grid \mathcal{D}^n and sparse collection of cubes $\mathcal{S} \subset \mathcal{D}^n$ depending on f_1, f_2, f_3, q_1, q_2 . Here it is important to notice that \mathcal{S} and \mathcal{D}^n do not depend on ω (see Lacey-Mena [26] or [29]). Thus,

$$(4.28) \quad |\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle| \leq C \Lambda_{\mathcal{S}}(\|f_1\|_{L^{q_1}(\mathbb{R}^m)}, \|f_2\|_{L^{q_2}(\mathbb{R}^m)}, \|f_3\|_{L^{q'}(\mathbb{R}^m)}).$$

Case 1a: $q_1, q_2 < \infty$ and $p_1, p_2 < \infty$.

This case follows at once from Corollary 4.3 and a standard density argument.

Case 1b: $q_1, q_2 < \infty$ and $p_1 = \infty$ or $p_2 = \infty$.

By symmetry we may assume that $p_1 = \infty$ and hence $p_2 = p$ with $1 < p < \infty$. We proceed by duality and observe first that T_n^{1*} (resp. T_m^{1*}), the adjoint with

respect the first entries of T_n (resp. T_m) is a bilinear Calderón-Zygmund operator in \mathbb{R}^n (resp. \mathbb{R}^m). Also, one can see that $(T_n \otimes T_m)^{1*} = T_n^{1*} \otimes T_m^{1*}$ and therefore the already established estimates give

$$\|(T_n \otimes T_m)^{1*}(f_3, f_2)\|_{L^1(\mathbb{R}^n; L^{q'_1}(\mathbb{R}^m))} \lesssim \|f_3\|_{L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))} \|f_2\|_{L^p(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}$$

for all $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$ and $f_2 \in L^p(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$ since in the current case we have $1 < p', p < \infty$, $1 < q', q_2 < \infty$, and $\frac{1}{q'_1} = \frac{1}{q'} + \frac{1}{q_2} < 1$.

In turn for every $f_1 \in L^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L^p(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$ we can define $T_n \otimes T_m(f_1, f_2) \in L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ by setting for all $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$

$$\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle := \langle (T_n \otimes T_m)^{1*}(f_3, f_2), f_1 \rangle.$$

All these eventually show that (4.6) holds and **Case 1b** is complete.

Case 1c: $q_1 = \infty$ or $q_2 = \infty$ and $p_1, p_2 < \infty$.

By symmetry we may assume that $q_1 = \infty$, hence $1 < q = q_2 < \infty$. Our first goal is to see that the above sparse estimate (4.28) holds for any $f_1 \in L^\infty(\mathbb{R}^{n+m})$ and $f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$. As observed before $(T_n \otimes T_m)^{1*}(f_3, f_2) = T_n^{1*} \otimes T_m^{1*}(f_3, f_2)$ and by Case 1a with exponents $\tilde{p}_1 = \tilde{p}_2 = \tilde{q}_1 = \tilde{q}_2 = 2$ we have

$$\begin{aligned} \|(T_n \otimes T_m)^{1*}(f_3, f_2)\|_{L^1(\mathbb{R}^{n+m})} &= \|(T_n \otimes T_m)^{1*}(f_3, f_2)\|_{L^1(\mathbb{R}^n; L^1(\mathbb{R}^m))} \\ &\lesssim \|f_3\|_{L^2(\mathbb{R}^n; L^2(\mathbb{R}^m))} \|f_2\|_{L^2(\mathbb{R}^n; L^2(\mathbb{R}^m))} < \infty. \end{aligned}$$

Let $R_1 \subset R_2 \subset \dots$ be an increasing sequence of rectangles such that $\bigcup_k R_k = \mathbb{R}^{n+m}$ and note that by duality

$$\begin{aligned} \langle T_n \otimes T_m(f_1, f_2), f_3 \rangle &= \langle (T_n \otimes T_m)^{1*}(f_3, f_2), f_1 \rangle \\ &= \lim_{k \rightarrow \infty} \langle (T_n \otimes T_m)^{1*}(f_3, f_2), 1_{R_k} f_1 \rangle = \lim_{k \rightarrow \infty} \langle T_n \otimes T_m(1_{R_k} f_1, f_2), f_3 \rangle. \end{aligned}$$

Hence since all the quantities involved are finite we can choose k (depending on f_1, f_2, f_3) so that

$$|\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle| \leq 2 |\langle T_n \otimes T_m(1_{R_k} f_1, f_2), f_3 \rangle|.$$

At this point we can invoke (4.28) —which is valid with $q_1 = \infty$, $q_2 = q$ as observed above— for $1_{R_k} f_1, f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ and find a sparse family \mathcal{S} and a dyadic grid \mathcal{D}^n with $\mathcal{S} \subset \mathcal{D}^n$ (both \mathcal{S} and \mathcal{D}^n depending on f_1, f_2, f_3 , and on k which ultimately depends on these three functions) such that

$$\begin{aligned} |\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle| &\lesssim \Lambda_{\mathcal{S}}(\|1_{R_k} f_1\|_{L^\infty(\mathbb{R}^m)}, \|f_2\|_{L^q(\mathbb{R}^m)}, \|f_3\|_{L^{q'}(\mathbb{R}^m)}) \\ &\leq \Lambda_{\mathcal{S}}(\|f_1\|_{L^\infty(\mathbb{R}^m)}, \|f_2\|_{L^q(\mathbb{R}^m)}, \|f_3\|_{L^{q'}(\mathbb{R}^m)}). \end{aligned}$$

All in one, we have been able to show that for any $f_1 \in L^\infty(\mathbb{R}^{n+m})$, $f_2, f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ there is a sparse domination formula as in (4.28) with a possible larger constant.

To proceed, let $w \in A_\infty(\mathbb{R}^n)$ and note that for any $f_1 \in L^\infty(\mathbb{R}^{n+m})$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$ we have

$$\begin{aligned} \|T_n \otimes T_m(f_1, f_2)\|_{L^q(w(x_1)dx_1, \mathbb{R}^n; L^q(\mathbb{R}^m))} &= \|T_n \otimes T_m(f_1, f_2)\|_{L^q(w \otimes 1, \mathbb{R}^{n+m})} \\ &= \sup |\langle T_n \otimes T_m(f_1, f_2), f_3 (w \otimes 1)^{\frac{1}{q}} \rangle|, \end{aligned}$$

where the sup runs over all $f_3 \in L_c^\infty(\mathbb{R}^{n+m})$ with $\|f_3\|_{L^{q'}} = 1$. Fix such a function f_3 and write $\tilde{f}_1 = \|f_1\|_{L^\infty(\mathbb{R}^m)}$, $\tilde{f}_2 = \|f_2\|_{L^q(\mathbb{R}^m)}$, and $\tilde{f}_3 = \|f_3\|_{L^{q_3}(\mathbb{R}^m)}$. By the previous argument we know that there exists a sparse family \mathcal{S} and a dyadic grid \mathcal{D}^n with $\mathcal{S} \subset \mathcal{D}^n$ (both \mathcal{S} and \mathcal{D}^n depending on f_1, f_2, f_3) such that

$$\begin{aligned} |\langle T_n \otimes T_m(f_1, f_2), f_3 (w \otimes 1)^{\frac{1}{q}} \rangle| &\lesssim \Lambda_{\mathcal{S}}(\tilde{f}_1, \tilde{f}_2, \tilde{f}_3 w^{\frac{1}{q}}) \\ &= \sum_{Q \in \mathcal{S}} w(Q) \langle \tilde{f}_1 \rangle_Q \langle \tilde{f}_2 \rangle_Q \frac{1}{w(Q)} \int_Q (\tilde{f}_3 w^{-\frac{1}{q'}}) w dx_1 \\ &\lesssim \sum_{Q \in \mathcal{S}} w(E_Q) \langle \tilde{f}_1 \rangle_Q \langle \tilde{f}_2 \rangle_Q \frac{1}{w(Q)} \int_Q (\tilde{f}_3 w^{-\frac{1}{q'}}) w dx_1 \\ &\lesssim \int_{\mathbb{R}^n} M(\tilde{f}_1, \tilde{f}_2) M_w^{\mathcal{D}^n}(\tilde{f}_3 w^{-\frac{1}{q'}}) w dx_1 \\ &\lesssim \|M(\tilde{f}_1, \tilde{f}_2)\|_{L^q(w, \mathbb{R}^n)} \|M_w^{\mathcal{D}^n}(\tilde{f}_3 w^{-\frac{1}{q'}})\|_{L^{q'}(w, \mathbb{R}^n)} \\ &\lesssim \|M(\tilde{f}_1, \tilde{f}_2)\|_{L^q(w, \mathbb{R}^n)}. \end{aligned}$$

In the previous computations we have used that $\{E_Q\}_{Q \in \mathcal{S}}$ is the pairwise disjoint family associated with the sparse family \mathcal{S} for which we have $|E_Q| \approx |Q|$. Also, since $w \in A_\infty$ it follows that $w(Q) \approx w(E_Q)$ since $|E_Q| \approx |Q|$. Finally, $M_w^{\mathcal{D}^n}$ is the \mathcal{D}^n -dyadic Hardy-Littlewood maximal function with underlying measure $w dx_1$ which is bounded in $L^{q'}(w, \mathbb{R}^n)$ for every $1 \leq q < \infty$. Gathering all the obtained estimates we have concluded that for all $w \in A_\infty(\mathbb{R}^n)$

$$\begin{aligned} \|\|T_n \otimes T_m(f_1, f_2)\|_{L^q(\mathbb{R}^m)}\|_{L^q(w, \mathbb{R}^n)} &= \|T_n \otimes T_m(f_1, f_2)\|_{L^q(w(x_1)dx_1, \mathbb{R}^n; L^q(\mathbb{R}^m))} \\ &\lesssim \|M(\tilde{f}_1, \tilde{f}_2)\|_{L^q(w, \mathbb{R}^n)}. \end{aligned}$$

Using then A_∞ -extrapolation, see [11, Theorem 2.1], we obtain for every $0 < r < \infty$

$$\begin{aligned} \|T_n \otimes T_m(f_1, f_2)\|_{L^r(\mathbb{R}^n; L^q(\mathbb{R}^m))} &= \|\|T_n \otimes T_m(f_1, f_2)\|_{L^q(\mathbb{R}^m)}\|_{L^r(\mathbb{R}^n)} \\ &\lesssim \|M(\tilde{f}_1, \tilde{f}_2)\|_{L^r(\mathbb{R}^n)}. \end{aligned}$$

At this point, given $1 < p_1, p_2 < \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ we can apply the previous estimate with $r = p$ and the well-known estimates for the bilinear maximal operator (see [27]) to conclude that

$$\begin{aligned} (4.29) \quad \|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} &\lesssim \|M(\tilde{f}_1, \tilde{f}_2)\|_{L^p(\mathbb{R}^n)} \lesssim \|\tilde{f}_1\|_{L^{p_1}(\mathbb{R}^n)} \|\tilde{f}_2\|_{L^{p_2}(\mathbb{R}^n)} \\ &= \|f_1\|_{L^{p_1}(\mathbb{R}^n; L^\infty(\mathbb{R}^m))} \|f_2\|_{L^{p_2}(\mathbb{R}^n; L^q(\mathbb{R}^m))}. \end{aligned}$$

for all $f_1 \in L^\infty(\mathbb{R}^{n+m})$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$.

Next, given $f_1 \in L^{p_1}(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$ consider $f_1^N(x_1, x_2) = f_1(x_1, x_2)1_{\{x_1: \|f(x_1, \cdot)\|_{L^\infty(\mathbb{R}^m)} < N\}}$ for every $(x_1, x_2) \in \mathbb{R}^{n+m}$ and $N \geq 1$. It is straightforward to see that $\{f_1^N\}_{N \geq 1}$ is Cauchy sequence in $L^{p_1}(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and hence this allows us to define

$$T_n \otimes T_m(f_1, f_2) := \lim_{N \rightarrow \infty} T_n \otimes T_m(f_1^N, f_2)$$

where the convergence is in $L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$. Since $f_1^N \in L^\infty(\mathbb{R}^{n+m})$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$ we can invoke (4.29) to see that (4.6) holds for all $f_1 \in L^{p_1}(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$. Next, since $p_2, q < \infty$ we can readily extend this estimate by density to all $f_1 \in L^{p_1}(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_2 \in L^{p_2}(\mathbb{R}^n; L^q(\mathbb{R}^m))$ and this completes the proof of (4.6) in the present case.

Case 1d: $q_1 = \infty$ or $q_2 = \infty$, and $p_1 = \infty$ or $p_2 = \infty$.

Again by symmetry we may assume that $q_1 = \infty$, hence $1 < q = q_2 < \infty$. Consider first the case $p_1 = \infty$, thus $1 < p_2 = p < \infty$. As before, from **Case 1a** we readily get

$$\|(T_n \otimes T_m)^{1*}(f_3, f_2)\|_{L^1(\mathbb{R}^n; L^1(\mathbb{R}^m))} \lesssim \|f_3\|_{L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))} \|f_2\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))},$$

for every $f_2 \in L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ and $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$. Then for every $f_1 \in L^\infty(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_2 \in L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ and we can define $T_n \otimes T_m(f_1, f_2) \in L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ by setting for all $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$

$$\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle := \langle (T_n \otimes T_m)^{1*}(f_3, f_2), f_1 \rangle.$$

All these readily imply the desired estimate.

In the second case, that is, when $p_2 = \infty$ and hence $1 < p_1 = p < \infty$ we can show much as before that **Case 1c** yields

$$\|(T_n \otimes T_m)^{2*}(f_1, f_3)\|_{L^1(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))} \lesssim \|f_1\|_{L^p(\mathbb{R}^n; L^\infty(\mathbb{R}^m))} \|f_3\|_{L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))},$$

for every $f_1 \in L^p(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$. As a consequence, for any $f_1 \in L^p(\mathbb{R}^n; L^\infty(\mathbb{R}^m))$ and $f_2 \in L^\infty(\mathbb{R}^n; L^q(\mathbb{R}^m))$ and we can define $T_n \otimes T_m(f_1, f_2) \in L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))$ by setting for all $f_3 \in L^{p'}(\mathbb{R}^n; L^{q'}(\mathbb{R}^m))$

$$\langle T_n \otimes T_m(f_1, f_2), f_3 \rangle := \langle (T_n \otimes T_m)^{2*}(f_1, f_3), f_2 \rangle.$$

This gives the desired estimate in the present scenario completing **Case 1d** and hence **Case 1**.

Case 2: $1 < p_1, p_2 \leq \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} > 0$ and $1 < q_1, q_2 < \infty$.

By Corollary 4.3 and a standard density argument we only need to treat the cases where either $p_1 = \infty$ or $p_2 = \infty$. By symmetry we just explain the case $p_1 = \infty$, $1 < p_2 = p < \infty$. Let $f_1 \in L^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$. Pick $1 < \tilde{q}_2 < \infty$ large enough so that $\frac{1}{\tilde{q}} = \frac{1}{q_1} + \frac{1}{\tilde{q}_2} < 1$. Since $f_2 \in L_c^\infty(\mathbb{R}^{n+m}) \subset L^p(\mathbb{R}^n; L^{\tilde{q}_2}(\mathbb{R}^m))$, our choices of exponents allows us to invoke **Case 1** we know that $T_n \otimes T_m(f_1, f_2)$ is a well-defined function in $L^p(\mathbb{R}^n; L^{\tilde{q}}(\mathbb{R}^m))$ —we would like to emphasize that this is the only place in this proof on which we use **Case 1** and it is done just qualitatively and with the exponents q_i 's being finite. At this point we proceed as in the proof of

Corollary 4.3 and define F , F_1 and F_2 as in (4.25). Our goal is to show the validity of (4.26). For \vec{w} as above, we may assume that $\|f_1(w_1 \otimes 1)\|_{L^{q_1}(\mathbb{R}^{n+m})} = \|F_1 w_1\|_{L^{q_1}(\mathbb{R}^n)}$ and $\|f_2(w_2 \otimes 1)\|_{L^{q_2}(\mathbb{R}^{n+m})} = \|F_2 w_2\|_{L^{q_2}(\mathbb{R}^n)}$ are finite, otherwise there is nothing to prove. That means that we can invoke again Theorem 4.2, and (4.26) follows in the same manner. We can then extrapolate and in the special case $w_1 \equiv 1$ and $w_2 \equiv 1$ obtain (4.27) in the present scenario:

$$\|T_n \otimes T_m(f_1, f_2)\|_{L^p(\mathbb{R}^n; L^q(\mathbb{R}^m))} \lesssim \|f_1\|_{L^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))} \|f_2\|_{L^p(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}.$$

The important fact is that this estimate holds for any $f_1 \in L^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L_c^\infty(\mathbb{R}^{n+m})$. In turn, since $1 < p_2, q_2 < \infty$ we can run a standard density argument to conclude that it also holds for all $f_1 \in L^\infty(\mathbb{R}^n; L^{q_1}(\mathbb{R}^m))$ and $f_2 \in L^p(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))$. This completes the proof. \square

REFERENCES

- [1] A. Barron. Weighted estimates for rough bilinear singular integrals via sparse domination. *New York J. Math.*, 23:779–811, 2017. [18](#)
- [2] C. Benea, C. Muscalu, Multiple vector valued inequalities via the helicoidal method, *Anal. PDE* 9 (2016) 1931–1988. [3](#), [19](#)
- [3] C. Benea, C. Muscalu, Quasi-Banach valued inequalities via the helicoidal method, *J. Funct. Anal.* 273(2017) 1295–1353. [3](#), [19](#)
- [4] C. Benea and C. Muscalu. Sparse domination via the helicoidal method. *arXiv:1707.05484*. [19](#)
- [5] Á. Bényi, J.M. Martell, K. Moen, E. Stachura, and R.H. Torres. Boundedness results for commutators with BMO functions via weighted estimates: a comprehensive approach. *Math. Ann.* 376 (2020) 61–102. [19](#)
- [6] D. Carando, M. Mazzitelli, S. Ombrosi, Multilinear Marcinkiewicz-Zygmund inequalities, *J. Fourier Anal. Appl.* 25 (2019) 51–85. [26](#)
- [7] L. Chaffee, R. Torres, 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. [6](#)
- [8] S.-Y. A. Chang, J. M. Wilson, T. H. Wolff, Some weighted norm inequalities concerning the Schrödinger operators, *Comment. Math. Helv.* 60 (1985) 217–246. [25](#)
- [9] P. Chen, D. He, and L. Song. Weighted inequalities for bilinear rough singular integrals from $L^2 \times L^2 \rightarrow L^1$. *J. Geom. Anal.* 29 (2019) 402–412. [18](#)
- [10] M. Christ, J.-L. Journé, Polynomial growth estimates for multilinear singular integral operators, *Acta Math.* 159 (1987) 51–80. [3](#)
- [11] D. Cruz-Uribe, J.M. Martell, C. Pérez, Extrapolation results for A_∞ weights and applications, *J. Funct. Anal.* 213 (2004) 412–439. [25](#), [38](#)
- [12] D. Cruz-Uribe, J.M. Martell, and C. Pérez. *Weights, extrapolation and the theory of Rubio de Francia*, volume 215 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer Basel AG, Basel, 2011. [2](#), [3](#), [4](#), [8](#), [10](#), [12](#), [14](#), [15](#)
- [13] D. Cruz-Uribe, J.M. Martell, C. Pérez, Sharp weighted estimates for classical operators. *Adv. Math.* 229 (2012), no. 1, 408–441. [28](#)
- [14] A. Culiuc, F. Di Plinio, and Y. Ou. Domination of multilinear singular integrals by positive sparse forms. *J. Lond. Math. Soc.* (2) 98 (2018), 369–392. [18](#), [19](#)
- [15] F. Di Plinio, Y. Ou, Banach-valued multilinear singular integrals, *Indiana Univ. Math. J.*, 67 (2018) 1711–1763. [3](#)

- [16] K. Domelevo, P. Ivanisvili, S. Petermichl, S. Treil, A. Volberg, On the failure of lower square function estimates in the non-homogeneous weighted setting, *Math. Ann.* 374 (2019) 1923–1952. [25](#)
- [17] J. Duoandikoetxea, Extrapolation of weights revisited: New proofs and sharp bounds, *J. Funct. Anal.* 260 (2011) 1886–1901. [3](#), [4](#), [8](#), [10](#), [16](#)
- [18] J. Duoandikoetxea, *Fourier analysis*, volume 29 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2001. [7](#)
- [19] J. Duoandikoetxea, O. Oruetebarria, Weighted mixed-norm inequalities through extrapolation, *Math. Nachr.* 292 (2019) 1482–1489. [4](#)
- [20] L. Grafakos, *Modern Fourier Analysis*, volume 250 of Graduate Texts in Mathematics. Springer, New York, 2nd edition, 2008. [7](#)
- [21] L. Grafakos, D. He, and P. Honzík, Rough bilinear singular integrals. *Adv. Math.*, 326:54–78, 2018. [18](#)
- [22] L. Grafakos, J.M. Martell, Extrapolation of weighted norm inequalities for multivariable operators and applications, *J. Geom. Anal.* 14 (2004) 19–46. [2](#)
- [23] L. Grafakos, R. Torres, Multilinear Calderón-Zygmund theory, *Adv. Math.* 165 (2002) 124–164. [22](#)
- [24] E. Harboure, R. Macías, C. Segovia, Extrapolation results for classes of weights, *Amer. J. Math.* 110 (1988) 383–397. [3](#), [4](#), [8](#)
- [25] J.-L. Journé, Two problems of Calderón-Zygmund theory on product-spaces, *Ann. Inst. Fourier* 38 (1988) 111–132. [3](#)
- [26] M. T. Lacey, D. Mena, The sparse $T1$ theorem, *Houston J. Math.* 43 (2017) 111–127. [36](#)
- [27] A.K. Lerner, S. Ombrosi, C. Pérez, R.H. Torres, R. Trujillo-González. New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory, *Adv. Math.* 220 (2009) 1222–1264. [2](#), [6](#), [18](#), [38](#)
- [28] K. Li, J. M. Martell, S. Ombrosi, Extrapolation for multilinear Muckenhoupt classes and applications to the bilinear Hilbert transform, preprint, arXiv:1802.03338, 2018. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#), [15](#), [16](#), [17](#), [18](#), [19](#), [20](#)
- [29] K. Li, H. Martikainen, Y. Ou, E. Vuorinen, Bilinear representation theorem, *Trans. Amer. Math. Soc.*, 371 (2019) 4193–4214. [22](#), [24](#), [25](#), [28](#), [36](#)
- [30] K. Li, H. Martikainen, E. Vuorinen, Bilinear Calderón-Zygmund theory on product spaces, *J. Math. Pures Appl.* (2019). <https://doi.org/10.1016/j.matpur.2019.10.007>. [3](#), [4](#), [21](#), [28](#)
- [31] K. Li, K. Moen, W. Sun, The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators, *J. Fourier Anal. Appl.* 20 (2014) 751–765. [25](#)
- [32] B. Nieraeth, Quantitative estimates and extrapolation for multilinear weight classes, *Math. Ann.* 375 (2019) 453–507. [4](#)
- [33] J.L. Rubio de Francia. Factorization theory and A_p weights. *Amer. J. Math.*, 106(3):533–547, 1984. [2](#)

KANGWEI LI, *Current address:* CENTER FOR APPLIED MATHEMATICS, TIANJIN UNIVERSITY, WEIJIN ROAD 92, 300072 TIANJIN, CHINA

AND

BCAM, BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO 14, E-48009 BILBAO, SPAIN

Email address: `kli@tju.edu.cn`

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`

HENRI MARTIKAINEN, DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND

Email address: `henri.martikainen@helsinki.fi`

SHELDY OMBROSI, DEPARTMENT OF MATHEMATICS, UNIVERSIDAD NACIONAL DEL SUR, BAHÍA BLANCA, ARGENTINA

Email address: `sombrosi@uns.edu.ar`

EMIL VUORINEN, CENTRE FOR MATHEMATICAL SCIENCES, UNIVERSITY OF LUND, P.O.B. 118, 22100 LUND, SWEDEN

Email address: `j.e.vuorin@gmail.com`