

TWO-WEIGHT COMMUTATOR ESTIMATES: GENERAL MULTI-PARAMETER FRAMEWORK

EMIL AIRTA

Abstract: We provide an explicit technical framework for proving very general two-weight commutator estimates in arbitrary parameters. The aim is to both clarify existing literature, which often explicitly focuses on two parameters only, and to extend very recent results to the full generality of arbitrary parameters. More specifically, we study two-weight commutator estimates – Bloom type estimates – in the multi-parameter setting involving weighted product BMO and little BMO spaces, and their combinations.

2010 Mathematics Subject Classification: 42B20.

Key words: iterated commutators, multi-parameter singular integrals, Bloom’s inequality, product BMO, weighted BMO, little BMO.

1. Introduction

Singular integral operators (SIOs) T have the general form

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

Varying the assumptions on the underlying kernel K gives us many fundamental linear transformations arising naturally in pure and applied analysis. One-parameter kernels are singular when $x = y$, while the *multi-parameter* theory deals with kernels with singularities on all hyperplanes of the form $x_i = y_i$, where $x, y \in \mathbb{R}^d$ are written in the form $x = (x_i)_{i=1}^m \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m}$ for a given partition $d = d_1 + \cdots + d_m$. Compare, for example, the one-parameter Cauchy kernel $1/(x - y)^2$ to the bi-parameter kernel

$$\frac{1}{(x_1 - y_1)(x_2 - y_2)},$$

which is the product of Hilbert kernels in both coordinate directions of $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2 = \mathbb{C}$. General multi-parameter kernels do not need to be of the product or convolution form, however. The paper by Fefferman–Stein ([12]) deals with the convolution case, while Journé ([24]) develops more general theory. However, we will be relying on the much more

recent dyadic-probabilistic methodology – see Martikainen [37] for the original bi-parameter theory and Ou [39] for the multi-parameter extensions.

Commutator estimates are a key part of modern harmonic analysis. Coifman–Rochberg–Weiss ([6]) showed that

$$(1.1) \quad \|b\|_{\text{BMO}} \lesssim \|[b, T]\|_{L^p \rightarrow L^p} \lesssim \|b\|_{\text{BMO}}, \text{ where } [b, T]f := bTf - T(bf),$$

for $p \in (1, \infty)$ and for some non-degenerate enough one-parameter SIOs T . In general, commutator estimates, e.g. yield by duality factorizations for Hardy functions, imply various div-curl lemmas relevant for compensated compactness and have connections to recent developments of the Jacobian problem $Ju = f$ in L^p – for the latter see Hytönen [23]. The field of multi-parameter commutator estimates has recently also been very active. For evidence of the activity, see, for example, the paper Duong–Li–Ou–Pipher–Wick [10], which studies the commutators of multi-parameter flag singular integrals. We get to other recent multi-parameter commutator estimates momentarily.

Let μ and λ be two general Radon measures in \mathbb{R}^d . A two-weight problem asks for a characterisation of the boundedness $T: L^p(\mu) \rightarrow L^p(\lambda)$, where T can be an SIO. For the two-weight characterisation for the Hilbert transform $T = H$, where $K(x, y) = 1/(x - y)$, see Lacey [26] and Lacey, Sawyer, Shen, and Uriarte-Tuero [30] (see also Hytönen [22]). The general higher dimensional theory has serious challenges, and there is no characterisation yet in the Riesz transform case. However, recently the corresponding two-weight question in the commutator setting has seen a lot of attention and progress. In these so-called Bloom type variants of the two-weight question we require that μ and λ are Muckenhoupt A_p weights and that the problem involves a function b . The theory then concerns the triple (μ, λ, b) , and the function b will lie in some appropriate weighted BMO space $\text{BMO}(\nu)$ formed using the Bloom weight $\nu := \mu^{1/p} \lambda^{-1/p} \in A_2$. Therefore, this means that for an operator A^b , depending naturally on some function b , the Bloom type questions concern the estimate

$$\|A^b\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|b\|_{\text{BMO}(\nu)}.$$

In the natural commutator setting the corresponding lower bound

$$\|b\|_{\text{BMO}(\nu)} \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|[b, T]\|_{L^p(\mu) \rightarrow L^p(\lambda)}$$

is also of interest. For the Hilbert transform $T = H$ Bloom ([3]) proved such a two-sided estimate – hence the name of the theory.

In the much more recent works of Holmes–Lacey–Wick [16, 17] Bloom’s upper bound was proved for general bounded SIOs in all dimensions \mathbb{R}^d . The lower bound was proved in the Riesz case. Lerner–Ombrosi–Rivera–Ríos ([31]) refined these results – this time the proofs employed sparse domination methods. An iterated commutator of the form $[b, [b, T]]$ is studied by Holmes–Wick [19], when $b \in \text{BMO} \cap \text{BMO}(\nu)$. This iterated case also follows from the so-called Cauchy integral trick of Coifman–Rochberg–Weiss [6]; see Hytönen [21]. This trick only works, though, as it is assumed that $b \in \text{BMO}$. However, this assumption is not valid in the optimal case – a fundamentally improved iterated case is by Lerner–Ombrosi–Rivera–Ríos [32], where $b \in \text{BMO}(\nu^{1/2}) \supsetneq \text{BMO} \cap \text{BMO}(\nu)$. This is optimal: a lower bound is also proved in [32]. In the already mentioned paper [23] by Hytönen lower bounds with very weak non-degeneracy assumptions were shown. Multilinear Bloom type inequalities are studied in the paper Kunwar–Ou [25].

We now get into bi-parameter and multi-parameter theory. Here the recent progress is most often based on the so-called representation theorems as sparse domination methods essentially currently work in one-parameter only (although see Barron–Pipher [2]). In fact, Barron–Conde-Alonso–Ou–Rey ([1]) show that one of the simplest bi-parameter model operators – the dyadic bi-parameter maximal function – cannot satisfy the most natural or useful candidate for bi-parameter sparse domination. A representation theorem represents SIOs by some dyadic model operators (DMOs). To understand the upcoming discussion, we need to discuss some details regarding this. The proofs of representation theorems are based on very careful refinements of various $T1$ theorems (for the original one see David and Journé [9]) and dyadic-probabilistic methods (see Nazarov–Treil–Volberg [38]). Indeed, $T1$ theorems essentially exhibit a decomposition of a standard SIO into its cancellative part and the so-called paraproducts. The one-parameter dyadic representation theorem of Hytönen [20] (extending e.g. Petermichl [41]) then provides a further decomposition of the cancellative part into so-called *dyadic shifts*, which are generalisations of the Haar multipliers

$$f = \sum_{Q \in \mathcal{D}} \langle f, h_Q \rangle h_Q \mapsto \sum_{Q \in \mathcal{D}} \lambda_Q \langle f, h_Q \rangle h_Q, \quad |\lambda_Q| \lesssim 1.$$

On the other hand, a paraproduct refers to an expression obtained by expanding both factors of the usual pointwise product $b \cdot f$ in some resolution of the identity, and dropping some of the terms in the resulting double expansion (so that it is not the full product). The $T1 \in \text{BMO}$ assumptions in $T1$ theorems specifically deal with these paraproducts.

The $T1$ theorems follow from representation theorems, but the real point is that the structural information of representation theorems is key for proofs of many other results.

In the bi-parameter case we have paraproducts and cancellative shifts, but also their hybrid combinations. The latter are new in this setting, and are called partial paraproducts due to their hybrid nature. The pure bi-parameter paraproducts are called full paraproducts. This leads to the following terminology: free of paraproducts (all paraproducts vanish) and free of full paraproducts (the partial paraproducts need not vanish but the full paraproducts do). These can all be phrased with checkable $T1 = 0$ type conditions. Such conditions always hold in the convolution case, and in some works these types of assumptions are made if the technology to handle the various paraproducts is not yet in place. See Martikainen [37] for the bi-parameter representation and Ou [39] for the multi-parameter extension. The following terminology is also convenient: the term SIO refers just to the kernel structure of our operators, while a Calderón–Zygmund operator (CZO) is an SIO satisfying appropriate $T1$ type conditions (and is thus bounded).

We are now ready to start our discussion of bi-parameter and multi-parameter commutators. If T is a bi-parameter CZO in $\mathbb{R}^{d_1+d_2}$ the right thing for $[b, T]$ is that $b \in \text{bmo}(\mathbb{R}^{d_1+d_2})$ – this means that $b(\cdot, x_2)$ and $b(x_1, \cdot)$ are uniformly in the usual BMO (this is one of the many equivalent ways to state this). This so-called little BMO is a certain type of bounded mean oscillation space in bi-parameter, and it arises in commutators of this type, but in many other cases the so-called product BMO (denoted e.g. by BMO_{prod}) of Chang and Fefferman [4, 5] involving general open sets is more fundamental. If T_{d_1} and T_{d_2} are linear one-parameter CZOs in \mathbb{R}^{d_1} and \mathbb{R}^{d_2} , respectively, then for

$$[T_{d_1}^1, [b, T_{d_2}^2]]f = T_{d_1}^1(bT_{d_2}^2 f) - T_{d_1}^1 T_{d_2}^2(bf) - bT_{d_2}^2 T_{d_1}^1 f + T_{d_2}^2(bT_{d_1}^1 f),$$

where $T_{d_1}^1 f(x) = T_{d_1}(f(\cdot, x_2))(x_1)$, the right object is $b \in \text{BMO}_{\text{prod}}(\mathbb{R}^{d_1+d_2})$. In the Hilbert transform $T = H$ case references for these commutators include Ferguson–Sadosky [14] and Ferguson–Lacey [13]. We note that the paper by Ferguson–Lacey ([13]) contains the deep lower bound

$$\|b\|_{\text{BMO}_{\text{prod}}(\mathbb{R}^2)} \lesssim \|[H^1, [b, H^2]]\|_{L^2 \rightarrow L^2}.$$

See also Lacey–Petermichl–Pipher–Wick [27, 28, 29] for the higher dimensional Riesz setting and div-curl lemmas.

By bounding commutators of bi-parameter shifts Ou, Petermichl, and Strouse ([40]) proved that

$$(1.2) \quad \|[b, T]\|_{L^2 \rightarrow L^2} \lesssim \|b\|_{\text{bmo}},$$

when T is a general bi-parameter CZO as in [37] and is free of para-products. This is a very special case of their theorem – we get to the full case later. Holmes–Petermichl–Wick ([18]) removed the paraproduct free assumption of [40] and proved the first bi-parameter Bloom type estimate

$$\|[b, T]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|b\|_{\text{bmo}(\nu)}.$$

Here A_p stands for bi-parameter weights (replace cubes by rectangles in the usual definition) and $\text{bmo}(\nu)$ is the weighted little BMO space defined using the norm

$$\|b\|_{\text{bmo}(\nu)} := \sup_R \frac{1}{\nu(R)} \int_R |b - \langle b \rangle_R|,$$

where the supremum is over all rectangles $R = I_1 \times I_2 \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, $\langle b \rangle_R = \frac{1}{|R|} \int_R b$, and $\nu(R) = \int_R \nu$.

Recently, Li–Martikainen–Vuorinen ([34]) reproved the result of [18] using a short proof based on some improved bi-parameter commutator decompositions from their bilinear bi-parameter theory [35]. Importantly, the new proof also allowed them to handle the iterated little BMO commutator by showing that

$$\|[b_k, \dots [b_2, [b_1, T]] \dots]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \prod_{i=1}^k \|b_i\|_{\text{bmo}(\nu^{1/k})}.$$

They also recently showed the corresponding lower bound in [36] using the median method. In [8] Dalenc and Ou extended [12] by proving that for all one-parameter CZOs

$$\|[T_{d_1}^1, [b, T_{d_2}^2]]\|_{L^p \rightarrow L^p} \lesssim \|b\|_{\text{BMO}_{\text{prod}}}.$$

The two-weight version of this was recently proved in [36]:

$$(1.3) \quad \|[T_{d_1}^1, [b, T_{d_2}^2]]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|b\|_{\text{BMO}_{\text{prod}}(\nu)}.$$

This is the first two-weight Bloom estimate involving the most delicate (and important) bi-parameter BMO space – the product BMO. In the weighted setting it can be defined by using the norm

$$(1.4) \quad \|b\|_{\text{BMO}_{\text{prod}}(\nu, \mathcal{D})} := \sup_{\Omega} \left(\frac{1}{\nu(\Omega)} \sum_{\substack{R \in \mathcal{D} \\ R \subset \Omega}} |\langle b, h_R \rangle|^2 \langle \nu \rangle_R^{-1} \right)^{1/2},$$

where Ω is an open set, \mathcal{D} is a given cartesian product of some dyadic grids in \mathbb{R}^{d_1} and \mathbb{R}^{d_2} , respectively, and the non-dyadic variant is a supremum over all such norms.

Our goal in this paper is to provide a careful proof of the analog of the estimate (1.3) in the case that the appearing singular integrals are multi-parameter, and in the case that we allow more singular integrals in the iteration. We want to provide an explicit proof in the multi-parameter setting, as they are very rare in the literature – often bi-parameter results are proved and the corresponding multi-parameter results are implicitly or explicitly claimed. This practice makes those results available only for a very few experts as often the details of the multi-parameter extensions are actually very challenging – both technically and notationally. The underlying general philosophies can be hard to understand from just the bi-parameter results. So our focus is both on the explicit methodology unveiling the general principles, and also on extending the recent result (1.3) as much as we possibly can.

In [40] the estimate (1.2) is used implicitly as a base case for more complicated multi-parameter commutator estimates. For example, suppose that T_1 and T_2 are paraproduct free linear bi-parameter singular integrals satisfying the assumptions of the representation theorem [37] in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and $\mathbb{R}^{d_3} \times \mathbb{R}^{d_4}$ respectively. Then, according to [40], we have the estimate

$$\begin{aligned} \|[T_1, [b, T_2]]f\|_{L^2(\prod_{i=1}^4 \mathbb{R}^{d_i})} &\lesssim \max \left(\sup_{x_2, x_4} \|b(\cdot, x_2, \cdot, x_4)\|_{\text{BMO}_{\text{prod}}}, \right. \\ &\quad \sup_{x_2, x_3} \|b(\cdot, x_2, x_3, \cdot)\|_{\text{BMO}_{\text{prod}}}, \\ &\quad \sup_{x_1, x_4} \|b(x_1, \cdot, \cdot, x_4)\|_{\text{BMO}_{\text{prod}}}, \\ &\quad \left. \sup_{x_1, x_3} \|b(x_1, \cdot, x_3, \cdot)\|_{\text{BMO}_{\text{prod}}} \right) \|f\|_{L^2(\prod_{i=1}^4 \mathbb{R}^{d_i})} \end{aligned}$$

involving both the product BMO and little BMO philosophies. The paraproduct free assumption can be removed according to [18]. We prove results of this type in the two-weight Bloom case generalising [36] and (1.3). As a byproduct, we get explicit proof of unweighted multi-parameter estimates of [40]. The full methodology is included, which is key.

Statement of the main results. A small restriction in our theorems concerns the fact that the way we handle the hybrid paraproducts (partial paraproducts) requires sparse domination methods in one-parameter. This requires that when our CZOs are not paraproduct free, they are at most bi-parameter – otherwise the partial paraproducts would not be amenable to sparse domination methods. This restriction comes from the methods of [34, 36], which we adapt here. We have not found a way to estimate certain terms without relying on these methods. However, if our CZOs are paraproduct free, they can be CZOs of arbitrarily

many parameters. This recovers multipliers, convolution form CZOs, and others.

Theorem 1.1. *We work in \mathbb{R}^d with m -parameters, i.e., $d = d_1 + \dots + d_m$. For a given $k \leq m$ we want to have k multi-parameter CZOs so that their parameters add up exactly to m . Thus, let $\mathcal{I} = \{\mathcal{I}_i\}_{i=1}^k$ be a partition of $\{1, \dots, m\}$, and for each $i = 1, \dots, k$ let us be given an $\#\mathcal{I}_i$ -parameter CZO T_i in $\prod_{j \in \mathcal{I}_i} \mathbb{R}^{d_j}$. Suppose, in addition, that for all $i = 1, \dots, k$ at least one of the following conditions holds:*

- (1) *the CZO T_i is paraproduct free, or*
- (2) *$\#\mathcal{I}_i \leq 2$.*

Let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_m})$ (m -parameter A_p weights in \mathbb{R}^d) and $1 < p < \infty$. Then for $b: \mathbb{R}^d \rightarrow \mathbb{C}$ we have

$$\|[T_1, [T_2, \dots [b, T_k]]]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{bmo}^\nu(\nu)},$$

where we understand that in this formula T_i acts on the whole space \mathbb{R}^d – i.e., $T_i = T_i^{\mathcal{I}_i}$ (see Subsection 2.1 for this notation). Moreover, $\|b\|_{\text{bmo}^\nu(\nu)}$ is the suitable little product BMO – see (2.2).

The structure of this paper is the following. In the beginning of Section 2 we give the notation which we are going to use in the entire paper. Then we give the definitions and recall some standard estimates.

In Section 3 we introduce expansions of function products and paraproduct operators. The main result of this section is to prove Bloom type upper bounds for these multi-parameter paraproduct operators.

Then we split the study of our main Theorem 1.1. In Section 4 we consider the paraproduct free CZOs by first proving the results for multi-parameter shifts. Then using the representation theorem we get the result for paraproduct free CZOs.

In Section 5 we begin with four parameter product spaces. We prove the case of the main theorem with two bi-parameter CZOs. The strategy of the proof is to use a representation theorem such that it is enough to study commutators with DMOs. We illustrate how to prove Bloom type upper bounds for these commutators by a careful study of a certain special case. Then by iterating previous results and combining with the results of Section 4 we get our main Theorem 1.1.

Acknowledgements. E. A. was supported by the Academy of Finland through the grant 306901, and is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research. The author wishes to thank Kangwei Li for helpful discussions. This work is a part of the PhD thesis of the author supervised by Henri Martikainen.

2. Definitions and preliminaries

2.1. Basic notation. We are working with the multi-parameter setting in $\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m}$. We set $d_1 + d_2 + \cdots + d_m = d$. To avoid confusion we need consistent notation. For example, every $x \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m}$ is a tuple (x_1, x_2, \dots, x_m) , where $x_i \in \mathbb{R}^{d_i}$. Similarly, every rectangle $I_1 \times I_2 \times \cdots \times I_m \subset \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m}$ consists of cubes $I_i \subset \mathbb{R}^{d_i}$. Rather than writing each cube separately, we let $I_{\bar{m}}$, where \bar{m} denotes the vector $(1, 2, \dots, m)$, be a rectangle in $\mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m}$ and also for functions we can write $1_{I_1} \otimes 1_{I_2} \otimes \cdots \otimes 1_{I_m} = 1_{I_1 \times \cdots \times I_m} =: 1_{I_{\bar{m}}}$. Since it really does not matter what order tensor form functions are written, we do not distinguish between $1_{I_1} \otimes 1_{I_2}$ and $1_{I_2} \otimes 1_{I_1}$.

We often need operators to be defined only for some of the variables – e.g. for $f: \mathbb{R}^{d_1+\cdots+d_m} \rightarrow \mathbb{C}$ and for some operator U in \mathbb{R}^{d_2} , $U^2 f$ is defined as

$$U^2 f(x) = (Uf(x_1, \cdot, x_3, \dots, x_m))(x_2).$$

Notice that, for example, for $g: \mathbb{R}^{d_1+d_2} \rightarrow \mathbb{C}$ we would also have

$$U^2 g(x_1, x_2) = (Ug(x_1, \cdot))(x_2).$$

Since it is clear from the context we do not make notational differences between these two. Additionally, we always write to the subscript the parameters where the operator is defined. For example, an operator U^1 is defined in \mathbb{R}^{d_1} .

Similarly, for integral pairings:

$$\langle f, g \rangle_1(z) := \langle f(\cdot, z), g \rangle_1,$$

where $f: \mathbb{R}^{d_1+d_2+\cdots+d_m} \rightarrow \mathbb{C}$, $g: \mathbb{R}^{d_1} \rightarrow \mathbb{C}$, and $z \in \mathbb{R}^{d_2+d_3+\cdots+d_m}$. Although for $\langle f, g \rangle$, where $f: \mathbb{R}^d \rightarrow \mathbb{C}$ and $g: \mathbb{R}^d \rightarrow \mathbb{C}$, it makes sense to leave out the parameters since in this case, the output of the pairing is a constant. For example, for $f: \mathbb{R}^{d_1+d_2+\cdots+d_m} \rightarrow \mathbb{C}$ we allow the notation $\langle f, 1_{I_2}/|I_2| \otimes 1_{I_1}/|I_1| \rangle_{2,1}$ and understand it as $\langle f, 1_{I_1}/|I_1| \otimes 1_{I_2}/|I_2| \rangle_{1,2}$.

In addition, if U^{v_1, v_2, \dots, v_n} is an operator in $\mathbb{R}^{d_{v_1}+d_{v_2}+\cdots+d_{v_n}}$ for some subsequence $\bar{v} = (v_i)_{i=1}^n$ of $(1, 2, \dots, m)$, then we simply write $U^{\bar{v}} f = U^{v_1, v_2, \dots, v_n} f$. For example, for brevity and clarity reasons, the operator $U^{2,1}$ is understood as the operator $U^{1,2}$ defined in $\mathbb{R}^{d_1+d_2}$.

Moreover, assume that $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2\}$ is a partition of $\{1, 2, \dots, m\}$, that is, $\bigcup_i \mathcal{I}_i = \{1, 2, \dots, m\}$, the sets \mathcal{I}_i are disjoint, and $\mathcal{I}_1 \neq \emptyset \neq \mathcal{I}_2$. Let $\bar{v} = (i)_{i \in \mathcal{I}_1}$, $x_{\bar{v}} \in \prod_{i \in \mathcal{I}_1} \mathbb{R}^{d_i}$, and let φ be a function on \mathbb{R}^d . Then we write $\varphi_{x_{\bar{v}}}$ as the obvious function defined on $\prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$, where $x_{\bar{v}}$ has been fixed. For example, let $f: \mathbb{R}^{d_1+d_2} \rightarrow \mathbb{C}$ and fix $(x_1, x_2) \in \mathbb{R}^{d_1+d_2}$. Then $f_{x_1}(y_2) = f(x_1, y_2)$ and $f_{x_2}(y_1) = f(y_1, x_2)$ for all $(y_1, y_2) \in \mathbb{R}^{d_1+d_2}$.

We denote dyadic grids in \mathbb{R}^{d_i} by \mathcal{D}^{d_i} and $\mathcal{D}^{\bar{d}} = \mathcal{D}^{d_1, d_2, \dots, d_m} := \mathcal{D}^{d_1} \times \mathcal{D}^{d_2} \times \dots \times \mathcal{D}^{d_m}$. If $I_i \in \mathcal{D}^{d_i}$, then $I_i^{(k_i)}$ denotes the unique dyadic cube $Q_i \in \mathcal{D}^{d_i}$ so that $I_i \subset Q_i$ and $\ell(Q_i) = 2^{k_i} \ell(I_i)$, where $\ell(J_i)$ denotes side length of a cube J_i . Similarly, for rectangles: if $I_{\bar{m}} \in \mathcal{D}^{\bar{d}}$, then $I_{\bar{m}}^{(\bar{k})} = I_1^{(k_1)} \times I_2^{(k_2)} \times \dots \times I_m^{(k_m)}$. In addition, for $I_i \in \mathcal{D}^{d_i}$ we define $\text{ch}(I_i) = \{Q_i \in \mathcal{D}^{d_i} : I_i = Q_i^{(1)}\}$.

For $I_i \in \mathcal{D}^{d_i}$ we denote by h_{I_i} a cancellative L^2 normalised Haar function. Here we suppressed the presence of $\eta \in \{0, 1\}^{d_i} \setminus \{0\}$. In particular, when we write $h_{I_i} h_{I_i}$ it can really stand for $h_{I_i}^{\eta_1} h_{I_i}^{\eta_2}$ for two different η_1, η_2 – however, this causes no problems as we only ever use the following size property $|h_{I_i} h_{I_i}| = 1_{I_i}/|I_i|$. We recall some basic properties: $\int h_{I_i} = 0$, $\langle h_{I_i}^{\eta_1}, h_{I_i}^{\eta_2} \rangle = \delta_{I_i, J_i} \delta_{\eta_1, \eta_2}$, and $h_{I_i}^0 = 1_{I_i}/|I_i|^{\frac{1}{2}} = |h_{I_i}|$.

Martingale representation. Let \mathcal{D}^{d_i} be some dyadic grid in \mathbb{R}^{d_i} and suppose that f is an appropriate function defined on \mathbb{R}^d . Let $x = (x_1, \dots, x_m) \in \mathbb{R}^d$ and $\bar{u}_i = (j)_{j \in \{1, 2, \dots, m\} \setminus \{i\}}$, hence $x_{\bar{u}_i} = (x_j)_{j \in \{1, 2, \dots, m\} \setminus \{i\}} \in \prod_{j \in \{1, 2, \dots, m\} \setminus \{i\}} \mathbb{R}^{d_j}$.

For φ defined on $\mathbb{R}^{d_{\bar{v}}}$ and a rectangle $Q_{\bar{v}} \subset \mathbb{R}^{d_{\bar{v}}}$ we denote the integral average

$$\frac{1}{|Q_{\bar{v}}|} \int_{Q_{\bar{v}}} \varphi$$

by $\langle \varphi \rangle_{Q_{\bar{v}}}$, where \bar{v} is a subsequence of $(1, 2, \dots, m)$. In addition, let $\{\mathcal{I}_1, \mathcal{I}_2\}$ be a partition of $\{1, 2, \dots, m\}$ such that $\mathcal{I}_1 \neq \emptyset \neq \mathcal{I}_2$. We define $\langle f \rangle_{Q_{\bar{v}}, \bar{v}}: \mathbb{R}^{d_{\bar{v}'}} \rightarrow \mathbb{C}$ as

$$\langle f \rangle_{Q_{\bar{v}}, \bar{v}}(y_{\bar{v}'}) = \frac{1}{|Q_{\bar{v}}|} \int_{Q_{\bar{v}}} f_{y_{\bar{v}'}}(y_{\bar{v}}) dy_{\bar{v}},$$

where $y_{\bar{v}'} \in \mathbb{R}^{d_{\bar{v}'}}$, $\bar{v} = (i)_{i \in \mathcal{I}_1}$, and $\bar{v}' = (i)_{i \in \mathcal{I}_2}$.

For all $i = 1, 2, \dots, m$ and $I_i \in \mathcal{D}^{d_i}$ define the one-parameter martingale difference

$$\Delta_{I_i}^i f(x) = (\Delta_{I_i}^i f_{x_{\bar{u}_i}})(x_i) = \sum_{Q_i \in \text{ch}(I_i)} (\langle f_{x_{\bar{u}_i}} \rangle_{Q_i} - \langle f_{x_{\bar{u}_i}} \rangle_{I_i}) 1_{Q_i}(x_i).$$

The multi-parameter martingale difference is defined as iterated one-parameter martingale differences

$$\Delta_{I_{\bar{v}}}^{\bar{v}} f(x) = \Delta_{I_{v_i}}^{v_i} (\Delta_{I_{\bar{v}'}}^{\bar{v}'} f)(x) = \dots = \Delta_{I_{v_i}}^{v_i} (\dots (\Delta_{I_{v_j}}^{v_j} f))(x),$$

where $\bar{v} = (v_i)_{i=1}^n$ is a subsequence of \bar{m} , $I_{\bar{v}} = I_{v_1} \times \dots \times I_{v_n} \in \mathcal{D}^{d_{\bar{v}}}$, \bar{v}'_i is the sequence without the parameter i , and the order of the one-parameter

martingale differences is arbitrary. Notice that we have the following equality

$$\Delta_{I_i}^i f = \sum_{\eta \in \{0,1\}^{d_i} \setminus \{0\}} \langle f, h_{I_i}^\eta \rangle_i \otimes h_{I_i}^\eta = \langle f, h_{I_i} \rangle_i \otimes h_{I_i},$$

where $I_i \in \mathcal{D}^{d_i}$. Naturally, in the multi-parameter situation, we have

$$\begin{aligned} \Delta_{I_{\bar{v}}}^{\bar{v}} f &= \langle \Delta_{I_{\bar{v}'}}^{\bar{v}'} f, h_{I_{v_i}} \rangle_{v_i} \otimes h_{I_{v_i}} = \cdots = \langle f, h_{I_{v_1}} \otimes \cdots \otimes h_{I_{v_n}} \rangle_{I_{v_1}} \otimes \cdots \otimes h_{I_{v_n}} \\ &= \langle f, h_{I_{\bar{v}}} \rangle_{\bar{v}} \otimes h_{I_{\bar{v}}}, \end{aligned}$$

where $\bar{v} = (v_i)_{i=1}^n$ is a subsequence of \bar{m} , and \bar{v}'_i is the sequence without the parameter i . Notice that, if $\bar{v} = \bar{m}$, then

$$\langle f, h_{I_{\bar{v}}} \rangle_{\bar{v}} \otimes h_{I_{\bar{v}}} = \langle f, h_{I_{\bar{m}}} \rangle h_{I_{\bar{m}}}.$$

Define the one-parameter martingale block

$$\Delta_{K_i, k_i}^i f = \sum_{\substack{I_i \in \mathcal{D}^{d_i} \\ I_i^{(k_i)} = K_i}} \Delta_{I_i}^i f$$

and the multi-parameter martingale block

$$\Delta_{K_{\bar{v}}, k_{\bar{v}}}^{\bar{v}} f = \sum_{\substack{I_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}} \\ I_{\bar{v}}^{(k_{\bar{v}})} = K_{\bar{v}}}} \Delta_{I_{\bar{v}}}^{\bar{v}} f,$$

where \bar{v} is a subsequence of \bar{m} .

Square functions. Define the one-parameter square function

$$S_{\mathcal{D}^{d_i}}^i f = \left(\sum_{I_i \in \mathcal{D}^{d_i}} |\Delta_{I_i}^i f|^2 \right)^{\frac{1}{2}} = \left(\sum_{I_i \in \mathcal{D}^{d_i}} |\langle f, h_{I_i} \rangle_i|^2 \otimes \frac{1_{I_i}}{|I_i|} \right)^{\frac{1}{2}}.$$

Since we will work on fixed dyadic grids, we abbreviate $S_{\mathcal{D}^{d_i}}^i$ by S^i . In addition, by the same reasoning, we abbreviate $\sum_{I_i \in \mathcal{D}^{d_i}}$ by \sum_{I_i} if there is no reason to emphasize the dyadic grid where the sum is taken over.

Let \bar{v} be a subsequence of $(1, 2, \dots, m)$. Define the multi-parameter square function

$$S_{\mathcal{D}^{d_{\bar{v}}}}^{\bar{v}} f = \left(\sum_{I_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}}} |\Delta_{I_{\bar{v}}}^{\bar{v}} f|^2 \right)^{\frac{1}{2}}.$$

If \bar{v} is a genuine subsequence, then we have

$$\left(\sum_{I_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}}} |\Delta_{I_{\bar{v}}}^{\bar{v}} f|^2 \right)^{\frac{1}{2}} = \left(\sum_{I_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}}} |\langle f, h_{I_{\bar{v}}} \rangle_{\bar{v}}|^2 \otimes \frac{1_{I_{\bar{v}}}}{|I_{\bar{v}}|} \right)^{\frac{1}{2}}$$

and obviously, for $\bar{v} = \bar{m}$ we have

$$\left(\sum_{I_{\bar{m}} \in \mathcal{D}^{d_{\bar{m}}}} |\Delta_{I_{\bar{m}}}^{\bar{m}} f|^2 \right)^{\frac{1}{2}} = \left(\sum_{I_{\bar{m}} \in \mathcal{D}^{d_{\bar{m}}}} |\langle f, h_{I_{\bar{m}}} \rangle|^2 \frac{1_{I_{\bar{m}}}}{|I_{\bar{m}}|} \right)^{\frac{1}{2}}.$$

Maximal functions. Define the one-parameter dyadic maximal function

$$\begin{aligned} M_{\mathcal{D}^{d_i}}^i f(x) &= M_{\mathcal{D}^{d_i}}^i f_{x_{\bar{u}_i}}(x_i) = \sup_{\substack{Q_i \in \mathcal{D}^{d_i} \\ Q_i \ni x_i}} \frac{1}{|Q_i|} \int_{Q_i} |f_{x_{\bar{u}_i}}(y_i)| dy_i \\ &= \sup_{Q_i \in \mathcal{D}^{d_i}} 1_{Q_i}(x_i) \langle |f_{x_{\bar{u}_i}}| \rangle_{Q_i}. \end{aligned}$$

Similarly, as with the square functions, we suppress the dyadic grid if there is no reason to specify it.

Let $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2\}$ be a partition of $\{1, 2, \dots, m\}$ and we require that $\mathcal{I}_1 \neq \emptyset$. Let $\bar{v} = (i)_{i \in \mathcal{I}_1}$ and $\bar{u} = (i)_{i \in \mathcal{I}_2}$. Define the strong multi-parameter dyadic maximal function

$$\begin{aligned} M^{\bar{v}} f(x) &= M^{\bar{v}} f_{x_{\bar{u}}}(x_{\bar{v}}) = \sup_{Q_{\bar{v}}: Q_{\bar{v}} \ni x_{\bar{v}}} \frac{1}{|Q_{\bar{v}}|} \int_{Q_{\bar{v}}} |f_{x_{\bar{u}}}(y_{\bar{v}})| dy_{\bar{v}} \\ &= \sup_{Q_{\bar{v}}} 1_{Q_{\bar{v}}}(x_{\bar{v}}) \langle |f_{x_{\bar{u}}}| \rangle_{Q_{\bar{v}}}, \end{aligned}$$

where the supremum is taken over the dyadic rectangles in $\mathcal{D}^{d_{\bar{v}}}$. If $\mathcal{I}_2 = \emptyset$, then $f_{x_{\bar{u}}}$ is understood as just f .

Observe that the strong maximal function is dominated by the iterated one-parameter maximal functions. For example, in the bi-parameter case we have

$$\begin{aligned} \frac{1}{|I_1 \times I_2|} \int_{I_1 \times I_2} |f(y_1, y_2)| dy_1 dy_2 &= \frac{1}{|I_2|} \int_{I_2} \langle |f| \rangle_{I_1, 1}(y_2) dy_2 \\ &\leq \frac{1}{|I_2|} \int_{I_2} M^1 f(x_1, y_2) dy_2 \\ &\leq M^2 M^1 f(x_1, x_2) \end{aligned}$$

for all $(x_1, x_2) \in I_1 \times I_2$. Hence, the boundedness of the strong maximal function follows directly from the boundedness of the one-parameter maximal function.

Weights. A weight w ($w \in L^1_{\text{loc}}(\mathbb{R}^d)$ and $0 < w(x) < \infty$ a.e.) belongs to the multi-parameter $A_p(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m})$, $1 < p < \infty$, if

$$[w]_{A_p} := \sup_I \langle w \rangle_I \langle w^{1-p'} \rangle_I^{p-1} < \infty,$$

where the supremum is taken over $I = I_1 \times I_2 \times \cdots \times I_m$, where $I_i \subset \mathbb{R}^{d_i}$ are cubes with sides parallel to the axes. We have $w \in A_p(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m})$ if and only if, for all $i = 1, 2, \dots, m$,

$$\text{ess sup}_{x_{\bar{u}_i}} [w_{x_{\bar{u}_i}}]_{A_p(\mathbb{R}^{d_i})} < \infty,$$

where the supremum is taken over

$$x_{\bar{u}_i} = (x_j)_{j \in \{1, 2, \dots, m\} \setminus \{i\}} \in \prod_{j \in \{1, 2, \dots, m\} \setminus \{i\}} \mathbb{R}^{d_j},$$

and furthermore, we have

$$\max_{i=1, 2, \dots, m} \text{ess sup}_{x_{\bar{u}_i}} [w_{x_{\bar{u}_i}}]_{A_p(\mathbb{R}^{d_i})} \leq [w]_{A_p(\mathbb{R}^d)}.$$

We say that a weight w belongs to one-parameter $A_\infty(\mathbb{R}^{d_i})$ if

$$[w]_{A_\infty(\mathbb{R}^{d_i})} = \sup_{I_i} \left(\frac{1}{|I_i|} \int_{I_i} w \right) \exp \left(\frac{1}{|I_i|} \int_{I_i} \log w^{-1} \right) < \infty,$$

where the supremum is taken over the cubes in \mathbb{R}^{d_i} . Recall that in the one-parameter setting a weight w belongs to $A_\infty(\mathbb{R}^{d_i})$ if $w \in A_p(\mathbb{R}^{d_i})$ for some $p < \infty$.

Hence, we say that a weight w belongs to multi-parameter $A_\infty(\mathbb{R}^d)$ if $w_{x_{\bar{u}_i}}$ belongs to $A_\infty(\mathbb{R}^{d_i})$ uniformly for every parameter $i \in \{1, 2, \dots, m\}$.

2.2. Standard estimates. We record some standard estimates. These estimates and some estimates that follow from these are used implicitly in this paper.

First, we record an A_∞ -extrapolation result from Cruz-Uribe–Martell–Pérez [7].

Lemma 2.1. *Let (f, g) be a pair of positive functions defined on \mathbb{R}^d . Suppose that there exists some $0 < p_0 < \infty$ such that for every $w \in A_\infty(\mathbb{R}^d)$ we have*

$$(2.1) \quad \int_{\mathbb{R}^d} f^{p_0} w \leq C([w]_{A_\infty}, p_0) \int_{\mathbb{R}^d} g^{p_0} w.$$

Then, for all $0 < p < \infty$ and $w \in A_\infty(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} f^p w \leq C([w]_{A_\infty}, p) \int_{\mathbb{R}^d} g^p w.$$

In addition, let $\{(f_j, g_j)\}_j$ be a sequence of pairs of positive functions defined on \mathbb{R}^d . Suppose that for some $0 < p_0 < \infty$ the pair (f_j, g_j) satisfies inequality (2.1) for every j . Then, for all $0 < p, q < \infty$ and $w \in A_\infty(\mathbb{R}^d)$ we have

$$\left\| \left(\sum_j (f_j)^q \right)^{\frac{1}{q}} \right\|_{L^p(w)} \lesssim_{[w]_{A_\infty}} \left\| \left(\sum_j (g_j)^q \right)^{\frac{1}{q}} \right\|_{L^p(w)},$$

where $\{(f_j, g_j)\}_j$ is a sequence of pairs of positive functions defined on \mathbb{R}^d .

Lemma 2.2. For $p \in (1, \infty)$ and $w \in A_p(\mathbb{R}^d)$ we have

$$\|f\|_{L^p(w)} \sim_{[w]_{A_p}} \|S^{\bar{v}} f\|_{L^p(w)},$$

where \bar{v} is any subsequence of $(1, 2, \dots, m)$.

Remark 2.3. In what follows we often assume that the appearing functions are nice – a specific choice that works throughout the paper is that this is understood to mean bounded and compactly supported functions.

Lemma 2.4. Let \bar{v} be a subsequence of $(1, 2, \dots, m)$ and \bar{u} be a subsequence of \bar{v} . Let $p \in (0, \infty)$ and $w \in A_\infty(\mathbb{R}^d)$. Then for a nice function f defined on \mathbb{R}^d we have

$$\|f\|_{L^p(w)} \lesssim_{[w]_{A_\infty}} \|S^{\bar{u}} f\|_{L^p(w)} \lesssim_{[w]_{A_\infty}} \|S^{\bar{v}} f\|_{L^p(w)}.$$

For completeness we record the proof.

Proof: Suppose $\bar{v} = (v_1, v_2)$ and $\bar{u} = (v_1)$. Let $(f_j)_j$ be a sequence of nice functions defined on \mathbb{R}^d . Recall the one-parameter result

$$\|f_{x_{\bar{u}}'}\|_{L^2(w_{x_{\bar{u}}'})} \lesssim_{[w]_{A_\infty}} \|S^{v_1} f_{x_{\bar{u}}'}\|_{L^2(w_{x_{\bar{u}}'})},$$

where $x_{\bar{u}}' = (x_i)_{i \in \{1, 2, \dots, m\} \setminus \{v_1\}}$. Using this result we get

$$\begin{aligned} \int_{\mathbb{R}^{d_{v_1}}} \left(\sum_j |f_j|^2 \right)^{\frac{1}{2} \cdot 2} w &= \sum_j \int_{\mathbb{R}^{d_{v_1}}} |f_j|^2 w \\ &\lesssim \sum_j \sum_{I_{v_1}} \int_{\mathbb{R}^{d_{v_1}}} |\Delta_{I_{v_1}}^{v_1} f_j|^2 w \\ &= \int_{\mathbb{R}^{d_{v_1}}} \left(\sum_j \sum_{I_{v_1}} |\Delta_{I_{v_1}}^{v_1} f_j|^2 \right)^{\frac{1}{2} \cdot 2} w \end{aligned}$$

for all $w \in A_\infty(\mathbb{R}^d)$. Here we abbreviated the fixed $x_{\bar{u}}'$.

Now, by the A_∞ -extrapolation Lemma 2.1 we have that

$$\left\| \left(\sum_j |f_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(w)} \lesssim_{[w]A_\infty} \left\| \left(\sum_j \sum_{I_{v_1}} |\Delta_{I_{v_1}}^{v_1} f_j|^2 \right)^{\frac{1}{2}} \right\|_{L^p(w)}$$

for all $w \in A_\infty(\mathbb{R}^d)$ and $0 < p < \infty$. Using this for $f_j = \Delta_{I_{v_2}}^{v_2} f$ we have

$$\begin{aligned} \|f\|_{L^p(w)} &\lesssim_{[w]A_\infty} \left\| \left(\sum_{I_{v_2}} |\Delta_{I_{v_2}}^{v_2} f|^2 \right)^{\frac{1}{2}} \right\|_{L^p(w)} \\ &\lesssim_{[w]A_\infty} \left\| \left(\sum_{I_{v_1}, I_{v_2}} |\Delta_{I_{v_1} \times I_{v_2}}^{v_1, v_2} f|^2 \right)^{\frac{1}{2}} \right\|_{L^p(w)} \end{aligned}$$

for all $w \in A_\infty(\mathbb{R}^d)$ and $0 < p < \infty$. It is clear that we can iterate the previous estimations for any number of parameters. \square

Next, we record the multi-parameter Fefferman–Stein inequality, which follows from the classical one-parameter Fefferman–Stein inequality [11] combined with the fact that strong maximal functions can be bounded with iterated one-parameter maximal functions.

Lemma 2.5 (Fefferman–Stein inequality). *Let \bar{v} be a subsequence of $(1, 2, \dots, m)$. For $p, q \in (1, \infty)$ we have*

$$\left\| \left(\sum_j |M^{\bar{v}} f_j|^q \right)^{\frac{1}{q}} \right\|_{L^p(w)} \lesssim_{[w]A_p} \left\| \left(\sum_j |f_j|^q \right)^{\frac{1}{q}} \right\|_{L^p(w)}.$$

Combining previous results we get the following:

Lemma 2.6. *Let $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2\}$ be a partition of $\{1, 2, \dots, m\}$ such that $\mathcal{I}_1 \neq \emptyset \neq \mathcal{I}_2$. Let $\bar{v} = (i)_{i \in \mathcal{I}_1}$ and $\bar{u} = (i)_{i \in \mathcal{I}_2}$. For $p \in (1, \infty)$ and $w \in A_p(\mathbb{R}^d)$ we have*

$$\begin{aligned} \left\| \sum_{K_{\bar{v}}} M^{\bar{u}'} [\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}] \otimes h_{K_{\bar{v}}} \right\|_{L^p(w)} &\sim \left\| \left(\sum_{K_{\bar{v}}} M^{\bar{u}'} [\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}]^2 \otimes \frac{1_{K_{\bar{v}}}}{|K_{\bar{v}}|} \right)^{\frac{1}{2}} \right\|_{L^p(w)} \\ &\lesssim_{[w]A_p} \|f\|_{L^p(w)}, \end{aligned}$$

where \bar{u}' is a subsequence of \bar{u} .

As explained earlier, we do not specify in the notation the underlying space of the function that $M^{\bar{u}'}$ is operating on.

Proof: Notice that

$$S^{\bar{v}} \left(\sum_{K_{\bar{v}}} M^{\bar{u}'} [\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}] \otimes h_{K_{\bar{v}}} \right) = \left(\sum_{K_{\bar{v}}} M^{\bar{u}'} [\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}]^2 \otimes \frac{1_{K_{\bar{v}}}}{|K_{\bar{v}}|} \right)^{\frac{1}{2}}.$$

Hence by Lemma 2.2 we get the first conclusion

$$\left\| \sum_{K_{\bar{v}}} M^{\bar{u}'}[\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}] \otimes h_{K_{\bar{v}}} \right\|_{L^p(w)} \sim \left\| \left(\sum_{K_{\bar{v}}} M^{\bar{u}'}[\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}]^2 \otimes \frac{1_{K_{\bar{v}}}}{|K_{\bar{v}}|} \right)^{\frac{1}{2}} \right\|_{L^p(w)}.$$

Then using Lemma 2.5 at the right-hand side of the previous estimate we get

$$\begin{aligned} \left\| \left(\sum_{K_{\bar{v}}} M^{\bar{u}'}[\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}]^2 \otimes \frac{1_{K_{\bar{v}}}}{|K_{\bar{v}}|} \right)^{\frac{1}{2}} \right\|_{L^p(w)} &\lesssim \left\| \left(\sum_{K_{\bar{v}}} |\langle f, h_{K_{\bar{v}}} \rangle_{\bar{v}}|^2 \otimes \frac{1_{K_{\bar{v}}}}{|K_{\bar{v}}|} \right)^{\frac{1}{2}} \right\|_{L^p(w)} \\ &= \|S^{\bar{v}} f\|_{L^p(w)} \\ &\lesssim \|f\|_{L^p(w)}, \end{aligned}$$

where at the last step we use again Lemma 2.2. \square

2.3. Product $\text{BMO}^{\bar{v}}$ spaces. Let $\bar{v} = (v_i)_{i=1}^n$ be a subsequence of $(1, 2, \dots, m)$, let w be a multi-parameter A_∞ weight on \mathbb{R}^d , and $b \in L^1_{\text{loc}}(\mathbb{R}^d)$. Also let $\mathcal{D}^{\bar{v}} = \prod_{i=1}^n \mathcal{D}^{v_i}$ be a product of dyadic grids. We say that $b \in \text{BMO}^{\bar{v}}(w; \mathcal{D}^{\bar{v}})$ if for all nice functions φ such that $\|S^{\bar{v}}_{\mathcal{D}^{\bar{v}}} \varphi\|_{L^1(w)} < \infty$ we have

$$|\langle b, \varphi \rangle| \leq C_b \|S^{\bar{v}}_{\mathcal{D}^{\bar{v}}} \varphi\|_{L^1(w)}.$$

Then we denote the best constant C_b by $\|b\|_{\text{BMO}^{\bar{v}}(w; \mathcal{D}^{\bar{v}})}$.

In addition, if $b \in \text{BMO}^{\bar{v}}(w; \mathcal{D}^{\bar{v}})$ for all $\mathcal{D}^{\bar{v}}$, then we say that $b \in \text{BMO}^{\bar{v}}(w)$.

Furthermore, let us define the *little product BMO*. Let $k \leq m$ and let $\mathcal{I} = \{\mathcal{I}_i : i \leq k\}$ be a partition of $\{1, 2, \dots, m\}$ such that $\mathcal{I}_i \neq \emptyset$ for $i = 1, 2, \dots, k$. We say that $b \in \text{bmo}^{\mathcal{I}}(w)$ if for all $\bar{v} = (v_j)_{j=1}^k$ such that $v_j \in \mathcal{I}_j$, we have $b \in \text{BMO}^{\bar{v}}(w)$. Then we set

$$(2.2) \quad \|b\|_{\text{bmo}^{\mathcal{I}}(w)} := \max_{\bar{v}} \|b\|_{\text{BMO}^{\bar{v}}(w)},$$

where the maximum is taken over $\bar{v} = (v_j)_{j=1}^k$ such that $v_j \in \mathcal{I}_j$.

For example, let $m = 3$, $w = 1$, $\mathcal{I}_1 = \{1, 2\}$, and $\mathcal{I}_2 = \{3\}$. Then $b \in \text{bmo}^{\{\{1,2\}, \{3\}\}}$ if $b \in \text{BMO}^{(1,3)}$ and $b \in \text{BMO}^{(2,3)}$.

Remark 2.7. We prefer this more direct square function definition over the typical square sum definition as in the introduction (1.4).

Remark 2.8. If $k = 1$, we have the standard multi-parameter little BMO space. For more details in the bi-parameter framework see e.g. [18, 34].

Proposition 2.9. *Let $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2\}$ be a partition of $\{1, 2, \dots, m\}$ such that $\mathcal{I}_1 \neq \emptyset \neq \mathcal{I}_2$. Let $\bar{v} = (i)_{i \in \mathcal{I}_1}$, $\bar{u} = (i)_{i \in \mathcal{I}_2}$, w be a multi-parameter A_∞ weight on \mathbb{R}^d , and $b \in L^1_{\text{loc}}(\mathbb{R}^d)$. Then $b \in \text{BMO}^{\bar{v}}(w)$ if and only if $b_{x_{\bar{u}}} \in \text{BMO}^{\bar{v}}(w_{x_{\bar{u}}})$ uniformly on $x_{\bar{u}} \in \prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$, that is, we have*

$$(2.3) \quad |\langle b_{x_{\bar{u}}}, f \rangle| \leq C_b \|S^{\bar{v}} f\|_{L^1(w_{x_{\bar{u}}})}$$

for almost every $x_{\bar{u}} \in \prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$ and for every nice f defined on $\prod_{i \in \mathcal{I}_1} \mathbb{R}^{d_i}$.

Proof: First, suppose that $b_{x_{\bar{u}}} \in \text{BMO}^{\bar{v}}(w_{x_{\bar{u}}})$ uniformly on $x_{\bar{u}} \in \prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$. Let φ be a nice function defined on \mathbb{R}^d . Hence, we have

$$\begin{aligned} |\langle b, \varphi \rangle| &\leq \int_{\mathbb{R}^{d_{\bar{u}}}} |\langle b_{x_{\bar{u}}}, \varphi_{x_{\bar{u}}} \rangle| dx_{\bar{u}} \\ &\stackrel{(2.3)}{\leq} C_b \int_{\mathbb{R}^{d_{\bar{u}}}} \|S^{\bar{v}} \varphi_{x_{\bar{u}}}\|_{L^1(w_{x_{\bar{u}}})} dx_{\bar{u}} \\ &= C_b \|S^{\bar{v}} \varphi\|_{L^1(w)}, \end{aligned}$$

as desired.

Suppose, conversely, that $b \in \text{BMO}^{\bar{v}}(w)$. Let $\varphi = \varphi_1 \otimes \varphi_2$ be a function defined on $\prod_{i \in \mathcal{I}_1} \mathbb{R}^{d_i} \times \prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$. Then we have

$$\begin{aligned} (2.4) \quad |\langle \langle b, \varphi_1 \rangle_{\bar{v}}, \varphi_2 \rangle| &= |\langle b, \varphi \rangle| \lesssim \|b\|_{\text{BMO}^{\bar{v}}(w)} \|S^{\bar{v}} \varphi\|_{L^1(w)} \\ &= \int_{\mathbb{R}^{d_{\bar{u}}}} \|S^{\bar{v}} \varphi_1\|_{L^1(w_{y_{\bar{u}}})} \varphi_2(y_{\bar{u}}) dy_{\bar{u}}. \end{aligned}$$

Let $\varphi_1 = f: \prod_{i \in \mathcal{I}_1} \mathbb{R}^{d_i} \rightarrow \mathbb{C}$ and for fixed $x_{\bar{u}} \in \prod_{i \in \mathcal{I}_2} \mathbb{R}^{d_i}$ and $r > 0$ let

$$\varphi_2(y_{\bar{u}}) = \frac{1_{B(x_{\bar{u}}, r)}(y_{\bar{u}})}{|B(x_{\bar{u}}, r)|}.$$

By Lebesgue's differentiation theorem, the left-hand side of (2.4) converges to $|\langle b_{x_{\bar{u}}}, f \rangle|$ for almost every $x_{\bar{u}}$.

By the same argument, the right-hand side of (2.4) converges to $\|S^{\bar{v}} f\|_{L^1(w_{x_{\bar{u}}})}$ for almost every $x_{\bar{u}}$. Hence, we have

$$|\langle b_{x_{\bar{u}}}, f \rangle| \leq \|b\|_{\text{BMO}^{\bar{v}}(w)} \|S^{\bar{v}} f\|_{L^1(w_{x_{\bar{u}}})}$$

for almost every $x_{\bar{u}}$ and for every nice function f defined on $\mathbb{R}^{d_{\bar{v}}}$. \square

We record the following BMO embedding result and we use this fact implicitly later on.

Lemma 2.10. *Let \bar{u} be a subsequence of $(1, 2, \dots, m)$ and \bar{v} be a subsequence of \bar{u} . Suppose $b \in \text{BMO}^{\bar{v}}(w)$. Then we have*

$$|\langle b, f \rangle| \lesssim_{[w]_{A_\infty}} \|b\|_{\text{BMO}^{\bar{v}}(w)} \|S^{\bar{u}} f\|_{L^1(w)},$$

that is, $\text{BMO}^{\bar{v}}(w) \subset \text{BMO}^{\bar{u}}(w)$.

Proof: The claim follows from the definition and Lemma 2.4, namely

$$\begin{aligned} |\langle b, f \rangle| &\leq \|b\|_{\text{BMO}^{\bar{v}}(w)} \|S^{\bar{v}} f\|_{L^1(w)} \\ &\lesssim_{[w]_{A_\infty}} \|b\|_{\text{BMO}^{\bar{v}}(w)} \|S^{\bar{u}} f\|_{L^1(w)}. \end{aligned} \quad \square$$

2.4. Singular integral operators. We define multi-parameter SIOs. For brevity, we give an explicit definition only in the bi-parameter setting. A general m -parameter definition can be found in Journé [24], but in a different operator-valued language. Our definition is as in [37] and is, in fact, equivalent to that given by Journé as proved by Grau de la Herrán [15]. An m -parameter definition using our partial kernel/full kernel language is explicitly given in Ou [39].

Let $\alpha \in (0, 1]$. We say that T is a bi-parameter singular integral operator (SIO) if the kernel representations below are satisfied.

Furthermore, if, in addition to kernel representations, T satisfies also some certain boundedness and cancellation assumptions, $T1$ assumptions, we say that T is a Calderón–Zygmund operator (CZO). These boundedness and cancellation assumptions are equivalent to L^2 -boundedness of T and its partial adjoint defined below.

For Calderón–Zygmund operators, we have the representation theorems [37, 39] using the dyadic model operators, namely paraproducts and shifts. The definitions of these model operators are presented later. We say that CZO T is a paraproduct free Calderón–Zygmund operator if it can be represented using only the dyadic shifts.

2.4.1. Full kernel representation. If $f = f_1 \otimes f_2$ and $g = g_1 \otimes g_2$ with $f_1, g_1: \mathbb{R}^{d_1} \rightarrow \mathbb{C}$, $f_2, g_2: \mathbb{R}^{d_2} \rightarrow \mathbb{C}$, $\text{spt } f_1 \cap \text{spt } g_1 = \emptyset$, and $\text{spt } f_2 \cap \text{spt } g_2 = \emptyset$, then we have the kernel representation

$$\langle Tf, g \rangle = \int_{\mathbb{R}^{d_1+d_2}} \int_{\mathbb{R}^{d_1+d_2}} K(x, y) f(y) g(x) dx dy.$$

The so-called full kernel

$$K: (\mathbb{R}^{d_1+d_2} \times \mathbb{R}^{d_1+d_2}) \setminus \{(x, y) \in \mathbb{R}^{d_1+d_2} \times \mathbb{R}^{d_1+d_2} : x_1 = y_1 \text{ or } x_2 = y_2\} \rightarrow \mathbb{C}$$

is assumed to satisfy the size condition

$$|K(x, y)| \leq C \frac{1}{|x_1 - y_1|^{d_1}} \frac{1}{|x_2 - y_2|^{d_2}},$$

the Hölder condition

$$\begin{aligned} &|K(x, y) - K(x, (y_1, y'_2)) - K(x, (y'_1, y_2)) - K(x, y')| \\ &\leq C \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^{d_1+\alpha}} \frac{|y_2 - y'_2|^\alpha}{|x_2 - y_2|^{d_2+\alpha}} \end{aligned}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$ and $|y_2 - y'_2| \leq |x_2 - y_2|/2$, and the mixed Hölder and size condition

$$|K(x, y) - K(x, (y'_1, y_2))| \leq C \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^{d_1 + \alpha}} \frac{1}{|x_2 - y_2|^{d_2}}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$.

Notice that this implies the kernel representation for T^* , T^{1*} , and $T^{2*} = (T^{1*})^*$, where T^* is the usual adjoint and T^{1*} is the partial adjoint defined by

$$\langle T^{1*}(f_1 \otimes f_2), g_1 \otimes g_2 \rangle = \langle T(g_1 \otimes f_2), f_1 \otimes g_2 \rangle.$$

Say that K^* , K^{1*} , and K^{2*} are the respective kernels of these, then we can write

$$\begin{aligned} K^*((x_1, x_2), (y_1, y_2)) &= K((y_1, y_2), (x_1, x_2)), \\ K^{1*}((x_1, x_2), (y_1, y_2)) &= K((y_1, x_2), (x_1, y_2)), \\ K^{2*}((x_1, x_2), (y_1, y_2)) &= K((x_1, y_2), (y_1, x_2)). \end{aligned}$$

We assume the above size and Hölder conditions also for K^* , K^{1*} , and K^{2*} .

2.4.2. Partial kernel representation. If $f = f_1 \otimes f_2$ and $g = g_1 \otimes g_2$ with $\text{spt } f_1 \cap \text{spt } g_1 = \emptyset$, then we assume the kernel representation

$$\langle Tf, g \rangle = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_1}} K_{f_2, g_2}(x_1, y_1) f_1(y_1) g_1(x_1) dx_1 dy_1.$$

The kernel

$$K_{f_2, g_2} : \{(x_1, y_1) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_1} : x_1 \neq y_1\} \rightarrow \mathbb{C}$$

is assumed to satisfy the size condition

$$|K_{f_2, g_2}(x_1, y_1)| \leq C(f_2, g_2) \frac{1}{|x_1 - y_1|^{d_1}}$$

and the Hölder conditions

$$|K_{f_2, g_2}(x_1, y_1) - K_{f_2, g_2}(x'_1, y_1)| \leq C(f_2, g_2) \frac{|x_1 - x'_1|^\alpha}{|x_1 - y_1|^{d_1 + \alpha}}$$

whenever $|x_1 - x'_1| \leq |x_1 - y_1|/2$ and

$$|K_{f_2, g_2}(x_1, y_1) - K_{f_2, g_2}(x_1, y'_1)| \leq C(f_2, g_2) \frac{|y_1 - y'_1|^\alpha}{|x_1 - y_1|^{d_1 + \alpha}}$$

whenever $|y_1 - y'_1| \leq |x_1 - y_1|/2$. We require the following control on the constant $C(f_2, g_2)$. For every cube $I_2 \subset \mathbb{R}^{d_2}$ we assume that $C(1_{I_2}, 1_{I_2}) + C(\varphi_{I_2}, 1_{I_2}) + C(1_{I_2}, \varphi_{I_2}) \lesssim |I_2|$, where φ_{I_2} is supported on I_2 , $\int \varphi_{I_2} = 0$, and $|\varphi_{I_2}| \leq 1$.

Analogously, we assume similar presentation and properties with K_{f_1, g_1} whenever $\text{spt } f_2 \cap \text{spt } g_2 = \emptyset$.

3. Paraproduct operators and martingale difference expansions of products

We assume that operators in this section are defined in some fixed dyadic grids $\mathcal{D}^{\bar{i}} = \prod_{i=1}^m \mathcal{D}^{d_i}$.

Define the one-parameter paraproduct operators

$$\begin{aligned} A_1^i(b, f) &= \sum_{I_i \in \mathcal{D}^{d_i}} \Delta_{I_i}^i b \Delta_{I_i}^i f, \\ A_2^i(b, f) &= \sum_{I_i \in \mathcal{D}^{d_i}} \Delta_{I_i}^i b E_{I_i}^i f, \\ A_3^i(b, f) &= \sum_{I_i \in \mathcal{D}^{d_i}} E_{I_i}^i b \Delta_{I_i}^i f, \end{aligned}$$

where $E_{I_i} \varphi := \langle \varphi \rangle_{I_i, i} 1_{I_i}$. We call the last term the “illegal” paraproduct.

Then we define the multi-parameter paraproduct operators as iterated one-parameter paraproducts – e.g. for $\bar{v} = (v_i)_{i=1}^n$, $n \leq m$ and $\bar{i} = (i_1, i_2, \dots, i_n)$ with $i_1 = 2$ and $i_j = 3$ for all $j \neq 1$ we have

$$\begin{aligned} A_{\bar{i}}^{\bar{v}}(b, f) &= A_{i_1}^{v_1} A_{i_2}^{v_2} \dots A_{i_n}^{v_n}(b, f) = \sum_{I_{v_1}} A_{i_2}^{v_2} \dots A_{i_n}^{v_n}(\Delta_{I_{v_1}}^{v_1} b, E_{I_{v_1}}^{v_1} f) \\ &= \sum_{I_{v_1}, I_{v_2}} A_{i_3}^{v_3} \dots A_{i_n}^{v_n}(\Delta_{I_{v_1}}^{v_1} E_{I_{v_2}}^{v_2} b, E_{I_{v_1}}^{v_1} \Delta_{I_{v_2}}^{v_2} f) \\ &= \dots = \sum_{I_{v_1}, I_{v_2}, \dots, I_{v_n}} \Delta_{I_{v_1}}^{v_1} E_{I_{v_2}}^{\bar{v}'} b E_{I_{v_1}}^{v_1} \Delta_{I_{v_2}}^{\bar{v}'} f, \end{aligned}$$

where $\bar{v}' = (v_i)_{i \in \{2, 3, \dots, n\}}$.

We write

$$bf = \sum_{i=1}^3 A_i^j(b, f),$$

and we say that this is the one-parameter expansion of the product bf in the parameter j . Then the multi-parameter expansion is obtained by iterating the previous one-parameter expansion. For example, let $\bar{v} = (v_i)_{i=1}^n$ be a subsequence of \bar{m} . Then the expansion in the parameters \bar{v} is

$$bf = \sum_{\bar{i} \in \{1, 2, 3\}^n} A_{\bar{i}}^{\bar{v}}(b, f) = \sum_{i_1, i_2, \dots, i_n=1}^3 A_{i_1}^{v_1} A_{i_2}^{v_2} \dots A_{i_n}^{v_n}(b, f),$$

where the “illegal” paraproduct is the one with $\bar{i} = \{3\}^n$. We want to emphasize that the paraproducts are directly bounded with some BMO assumption if $\bar{i} \neq \{3\}^n$, as we are going to show next, hence the name “illegal”.

Lemma 3.1. *Let \mathcal{J} be a subset of $\{1, 2, \dots, m\}$ with $n \leq m$ elements. Let $\bar{v} = (v_j)_{v_j \in \mathcal{J}}$, $\bar{i} = (i_1, i_2, \dots, i_n) \in \{1, 2, 3\}^n \setminus \{3\}^n$, and $\bar{u} = (u_j)$, $u_j \in \{k \in \mathcal{J} : i_k \neq 3\}$. Also let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_m})$ and $1 < p < \infty$. Then*

$$\|A_{\bar{i}}^{\bar{v}}(b, \cdot)\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim_{[\mu]_{A_p} [\lambda]_{A_p}} \|b\|_{\text{BMO}^{\bar{v}}(\nu)}.$$

Proof: Since we have three different type of one-parameter paraproducts, let $\mathcal{I} = \{\mathcal{I}_i\}_{i=1}^3$ be a partition of \mathcal{J} such that $\mathcal{I}_k = \{j \in \mathcal{J} : i_j = k\}$. Notice we require in the statement that $\mathcal{I}_1 \cup \mathcal{I}_2 \neq \emptyset$. We set $\bar{v}^k = (v_j^k)_{v_j^k \in \mathcal{I}_k}$. If some set $\mathcal{I}_k = \emptyset$ or $\mathcal{J} = \{1, 2, \dots, m\}$, it is fairly obvious what steps are not necessary and we omit the details. By the partition, we are considering the term

$$\begin{aligned} A_{\bar{i}}^{\bar{v}}(b, f) = & \sum_{I_{\bar{v}^1}, I_{\bar{v}^2}, I_{\bar{v}^3}} \left\langle b, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes \frac{1_{I_{\bar{v}^3}}}{|I_{\bar{v}^3}|} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \\ & \otimes h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}}. \end{aligned}$$

Begin the estimation with the dual form

$$\begin{aligned} & \left| \int_{\mathbb{R}^{d_{\bar{v}'}}} \sum_{I_{\bar{v}^1}, I_{\bar{v}^2}, I_{\bar{v}^3}} \left\langle b, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes \frac{1_{I_{\bar{v}^3}}}{|I_{\bar{v}^3}|} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \right. \\ & \quad \left. \times \langle g, h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \right|, \end{aligned}$$

where $\bar{v}' = (j)_{j \in \{1, 2, \dots, m\} \setminus \mathcal{J}}$. Then we fix the variable $x_{\bar{v}'}$ in $\mathbb{R}^{d_{\bar{v}'}}$ and consider the sum inside the integral. For now, in this proof, we do not write $x_{\bar{v}'}$ to the subscript of the functions, i.e. b means $b_{x_{\bar{v}'}}$ and so on. Thus we are estimating

$$\begin{aligned} & \left| \sum_{I_{\bar{v}^1}, I_{\bar{v}^2}, I_{\bar{v}^3}} \left\langle b, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes \frac{1_{I_{\bar{v}^3}}}{|I_{\bar{v}^3}|} \right\rangle \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle \right. \\ & \quad \left. \times \langle g, h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle \right| \\ &= \left| \int_{\mathbb{R}^{d_{\bar{v}^3}}} \sum_{I_{\bar{v}^3}} \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \sum_{I_{\bar{v}^1}, I_{\bar{v}^2}} \langle b_{x_{\bar{v}^3}}, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \rangle \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle \right. \\ & \quad \left. \times \langle g, h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle dx_{\bar{v}^3} \right| \\ &= \left| \int_{\mathbb{R}^{d_{\bar{v}^3}}} \sum_{I_{\bar{v}^3}} \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \langle b_{x_{\bar{v}^3}}, A_{I_{\bar{v}^3}}^{\bar{v}^1, \bar{v}^2}(f, g) \rangle dx_{\bar{v}^3} \right|, \end{aligned}$$

where

$$A_{I_{\bar{v}^3}}^{\bar{v}^1, \bar{v}^2}(f, g) := \sum_{I_{\bar{v}^1}, I_{\bar{v}^2}} \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle \\ \times \langle g, h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}}.$$

By Proposition 2.9 it is enough to show the boundedness of

$$(3.1) \quad \int_{\mathbb{R}^{\bar{v}^3}} \sum_{I_{\bar{v}^3}} \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \|S^{\bar{v}^1, \bar{v}^2}(A_{I_{\bar{v}^3}}^{\bar{v}^1, \bar{v}^2}(f, g))\|_{L^1(\nu_{x_{\bar{v}^3}})} dx_{\bar{v}^3}.$$

Note that $\nu_{x_{\bar{v}^3}}$ actually is $\nu_{x_{\bar{v}'}, x_{\bar{v}^3}}$.

First, observe that

$$S^{\bar{v}^1, \bar{v}^2}(A_{I_{\bar{v}^3}}^{\bar{v}^1, \bar{v}^2}(f, g)) = \left(\sum_{I_{\bar{v}^1}, I_{\bar{v}^2}} \left| \left\langle f, h_{I_{\bar{v}^1}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{I_{\bar{v}^3}} \right\rangle \right|^2 \right. \\ \left. \times \left| \left\langle g, h_{I_{\bar{v}^1}} h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \right\rangle \right|^2 \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \right)^{\frac{1}{2}} \\ \leq \left(\sum_{I_{\bar{v}^1}} [M^{\bar{v}^2} \langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \right)^{\frac{1}{2}} \\ \times \left(\sum_{I_{\bar{v}^2}} [M^{\bar{v}^1} \langle g, h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^2, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \right)^{\frac{1}{2}}.$$

Using the previous inequality we get

$$\sum_{I_{\bar{v}^3}} \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \|S^{\bar{v}^1, \bar{v}^2}(A_{I_{\bar{v}^3}}^{\bar{v}^1, \bar{v}^2}(f, g))\|_{L^1(\nu_{x_{\bar{v}^3}})} \\ \leq \int_{\mathbb{R}^{d_{\bar{v}^1} + d_{\bar{v}^2}}} \sum_{I_{\bar{v}^3}} \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \left(\sum_{I_{\bar{v}^1}} [M^{\bar{v}^2} \langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \right)^{\frac{1}{2}} \\ \times \left(\sum_{I_{\bar{v}^2}} [M^{\bar{v}^1} \langle g, h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^2, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \right)^{\frac{1}{2}} \nu_{x_{\bar{v}^3}} \\ \leq \int_{\mathbb{R}^{d_{\bar{v}^1} + d_{\bar{v}^2}}} \left(\sum_{I_{\bar{v}^1}, I_{\bar{v}^3}} [M^{\bar{v}^2} \langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \otimes \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \\ \times \left(\sum_{I_{\bar{v}^2}, I_{\bar{v}^3}} [M^{\bar{v}^1} \langle g, h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^2, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \nu_{x_{\bar{v}^3}}$$

$$\begin{aligned} &\leq \left\| \left(\sum_{I_{\bar{v}^1}, I_{\bar{v}^3}} [M^{\bar{v}^2} \langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \otimes \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \right\|_{L^p(\mu_{x_{\bar{v}^3}})} \\ &\quad \times \left\| \left(\sum_{I_{\bar{v}^2}, I_{\bar{v}^3}} [M^{\bar{v}^1} \langle g, h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^2, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes \frac{1_{I_{\bar{v}^3}}(x_{\bar{v}^3})}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \right\|_{L^{p'}(\lambda_{x_{\bar{v}^3}}^{1-p'})}. \end{aligned}$$

Putting this estimate back to (3.1) and applying Hölder's inequality we get

$$\begin{aligned} (3.1) &\leq \left\| \left(\sum_{I_{\bar{v}^1}, I_{\bar{v}^3}} [M^{\bar{v}^2} \langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^1, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^1}}}{|I_{\bar{v}^1}|} \otimes \frac{1_{I_{\bar{v}^3}}}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \right\|_{L^p(\mu)} \\ &\quad \times \left\| \left(\sum_{I_{\bar{v}^2}, I_{\bar{v}^3}} [M^{\bar{v}^1} \langle g, h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \rangle_{\bar{v}^2, \bar{v}^3}]^2 \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes \frac{1_{I_{\bar{v}^3}}}{|I_{\bar{v}^3}|} \right)^{\frac{1}{2}} \right\|_{L^{p'}(\lambda^{1-p'})} \\ &\lesssim \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda^{1-p'})}, \end{aligned}$$

where at the last step we apply Lemma 2.6.

Lastly, recall that we fixed $x_{\bar{v}'} \in \mathbb{R}^{d_{\bar{v}'}}$ and the previous bound actually is

$$\|f_{x_{\bar{v}'}}\|_{L^p(\mu_{x_{\bar{v}'}})} \|g_{x_{\bar{v}'}}\|_{L^{p'}(\lambda_{x_{\bar{v}'}}^{1-p'})}.$$

However, by applying Hölder's inequality once more to the integral over $\mathbb{R}^{d_{\bar{v}'}}$ we get

$$|\langle A_i^{\bar{v}}(b, f), g \rangle| \lesssim \|b\|_{\text{BMO}^{\bar{u}}(\nu)} \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda^{1-p'})},$$

where $\bar{u} = (j)_{j \in \mathcal{I}_1 \cup \mathcal{I}_2}$. □

4. Paraproduct free commutators

We assume that operators in this section are defined in some fixed dyadic grids $\mathcal{D}^{\bar{d}} = \prod_{i=1}^m \mathcal{D}^{d_i}$.

Let \bar{v} be a subsequence of \bar{m} . Define the multi-parameter shift

$$\mathcal{S}^{\bar{v}} f = \mathcal{S}_{\mathcal{D}^{\bar{d}_{\bar{v}}}}^{\bar{v}, \bar{k}, \bar{l}} f = \sum_{K_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}}} \sum_{\substack{I_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}} \\ I_{\bar{v}}^{(k_{\bar{v}})} = K_{\bar{v}}}} \sum_{\substack{J_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}} \\ J_{\bar{v}}^{(l_{\bar{v}})} = K_{\bar{v}}}} a_{K_{\bar{v}}, I_{\bar{v}}, J_{\bar{v}}} \langle f, h_{I_{\bar{v}}} \rangle_{\bar{v}} \otimes h_{J_{\bar{v}}}.$$

Here $k_{v_i}, l_{v_i} \geq 0$, only finitely many of the coefficients $a_{K_{\bar{v}}, I_{\bar{v}}, J_{\bar{v}}}$ are non-zero, and

$$|a_{K_{\bar{v}}, I_{\bar{v}}, J_{\bar{v}}}| \leq \frac{|I_{\bar{v}}|^{\frac{1}{2}} |J_{\bar{v}}|^{\frac{1}{2}}}{|K_{\bar{v}}|}.$$

First, we record here a standard equality as a lemma, since the notation in the multi-parameter setting needs some explaining.

Lemma 4.1. *Let φ be a locally integrable function defined on $\mathbb{R}^{d_{\bar{v}}}$, where $\bar{v} = (v_i)_{i=1}^n$ is a subsequence of \bar{m} , and let k_{v_i}, l_{v_i} be non-negative integers for $i = 1, 2, \dots, n$. Also let $I_{\bar{v}}, J_{\bar{v}}, K_{\bar{v}} \in \mathcal{D}^{d_{\bar{v}}}$ such that $I_{v_i}^{(k_{v_i})} = K_{v_i} = J_{v_i}^{(l_{v_i})}$ for all $i = 1, 2, \dots, n$. We have*

$$(4.1) \quad \langle \varphi \rangle_{J_{\bar{v}}} - \langle \varphi \rangle_{I_{\bar{v}}} = \sum_{i=1}^n \left(\sum_{t=1}^{l_{v_i}} \langle \Delta_{J_{v_i}^{(t)}}^{v_i} \varphi \rangle_{J_{v_i} \times Q_{\bar{v}'_i}} - \sum_{s=1}^{k_{v_i}} \langle \Delta_{I_{v_i}^{(s)}}^{v_i} \varphi \rangle_{I_{v_i} \times Q_{\bar{v}'_i}} \right),$$

where $\bar{v}'_i = (v_j)_{j \in \{1, 2, \dots, n\} \setminus i}$, and

$$Q_{\bar{v}'_i} = \begin{cases} I_{v_1} \times \cdots \times I_{v_{i-1}} \times J_{v_{i+1}} \times \cdots \times J_{v_n} & \text{if } 1 < i < n, \\ I_{v_1} \times \cdots \times I_{v_{n-1}} & \text{if } i = n, \\ J_{v_2} \times \cdots \times J_{v_n} & \text{if } i = 1. \end{cases}$$

Proof: The case $n = 1$ follows easily from the telescoping nature of the sum. The case $n = 2$ follows from this as follows. For notational simplicity only, let $v_1 = 1$ and $v_2 = 2$. Observe that

$$\langle \varphi \rangle_{J_1 \times J_2} - \langle \varphi \rangle_{I_1 \times I_2} = \langle \varphi \rangle_{J_1 \times J_2} - \langle \varphi \rangle_{I_1 \times J_2} + \langle \varphi \rangle_{I_1 \times J_2} - \langle \varphi \rangle_{I_1 \times I_2}.$$

Since K_i is some parent cube for both I_i and J_i , we can use the one-parameter expansion $\langle \varphi \rangle_{Q_i, i} - \langle \varphi \rangle_{K_i, i} = \sum_{q=1}^{q_i} \langle \Delta_{Q_i^{(q)}}^i \varphi \rangle_{Q_i, i}$, where $Q_i^{(q_i)} = K_i$. Thus, we have

$$\begin{aligned} \langle \varphi \rangle_{J_1 \times J_2} - \langle \varphi \rangle_{I_1 \times I_2} &= \sum_{t=1}^{l_1} \langle \Delta_{J_1^{(t)}}^1 \varphi \rangle_{J_1 \times J_2} - \sum_{s=1}^{k_1} \langle \Delta_{I_1^{(s)}}^1 \varphi \rangle_{I_1 \times J_2} \\ &\quad + \sum_{t=1}^{l_2} \langle \Delta_{J_2^{(t)}}^2 \varphi \rangle_{I_1 \times J_2} - \sum_{s=1}^{k_2} \langle \Delta_{I_2^{(s)}}^2 \varphi \rangle_{I_1 \times I_2}, \end{aligned}$$

as claimed.

We can continue as follows. If the claim holds for a fixed n , we have

$$\langle \varphi \rangle_{J_{\bar{v}}} - \langle \varphi \rangle_{I_{\bar{v}}} = \sum_{i=1}^{n+1} \left(\sum_{t=1}^{l_{v_i}} \langle \Delta_{J_{v_i}^{(t)}}^{v_i} \varphi \rangle_{J_{v_i} \times Q_{\bar{v}'_i}} - \sum_{s=1}^{k_{v_i}} \langle \Delta_{I_{v_i}^{(s)}}^{v_i} \varphi \rangle_{I_{v_i} \times Q_{\bar{v}'_i}} \right)$$

for $\bar{v} = (v_i)_{i=1}^{n+1}$. Indeed, for notational simplicity let again $v_i = i$ for all $i = 1, 2, \dots, n+1$, and notice that we may write

$$\begin{aligned} \langle \varphi \rangle_{J_{\bar{v}}} - \langle \varphi \rangle_{I_{\bar{v}}} &= \langle \varphi \rangle_{J_1 \times J_{\bar{u}}} - \langle \varphi \rangle_{I_1 \times J_{\bar{u}}} + \langle \varphi \rangle_{I_1 \times J_{\bar{u}}} - \langle \varphi \rangle_{I_1 \times I_{\bar{u}}} \\ &=: A + B, \end{aligned}$$

where $\bar{u} = (i)_{i=2}^{n+1}$. For the term A we use the one-parameter expansion and for the term B we use the assumption that the claim holds for

n parameters. Hence, we get

$$\begin{aligned} \langle \varphi \rangle_{J_{\bar{v}}} - \langle \varphi \rangle_{I_{\bar{v}}} &= \sum_{t=1}^{l_1} \langle \Delta_{J_1^{(t)}}^1 \varphi \rangle_{J_1 \times J_{\bar{u}}} - \sum_{s=1}^{k_1} \langle \Delta_{I_1^{(s)}}^1 \varphi \rangle_{I_1 \times J_{\bar{u}}} \\ &\quad + \sum_{i=2}^{n+1} \left(\sum_{t=1}^{l_i} \langle \Delta_{J_i^{(t)}}^i \varphi \rangle_{I_1 \times J_i \times Q_{\bar{u}'_i}} - \sum_{s=1}^{k_i} \langle \Delta_{I_i^{(s)}}^i \varphi \rangle_{I_1 \times I_i \times Q_{\bar{u}'_i}} \right) \\ &= \sum_{i=1}^{n+1} \left(\sum_{t=1}^{l_i} \langle \Delta_{J_i^{(t)}}^i \varphi \rangle_{J_i \times Q_{\bar{v}'_i}} - \sum_{s=1}^{k_i} \langle \Delta_{I_i^{(s)}}^i \varphi \rangle_{I_i \times Q_{\bar{v}'_i}} \right), \end{aligned}$$

where $Q_{\bar{u}'_i}$ and $Q_{\bar{v}'_i}$ are defined as in the statement. \square

The main result of this section is to show the boundedness of the commutators with paraproduct free Calderón–Zygmund operators. By the representation theorem, it is enough to consider the commutators of dyadic shifts (Theorem 4.4). The strategy is to expand the commutator using martingale differences. This leaves us with terms that are compositions of shifts and paraproducts, legal or illegal ones. In the case of illegal paraproducts, we combine some terms together and apply Lemma 4.1. This leads to terms which all fall under the general term (4.2). Before we show in detail how to expand the commutators, we present the general term and show its boundedness.

Assume that $\mathcal{I} = \{\mathcal{I}_i : i = 1, 2, 3, 4, 5\}$ is a partition of $\{1, 2, \dots, m\}$. We set $\bar{v}^i = (v_j^i)_{v_j^i \in \mathcal{I}_i}$. The general term is defined as

$$\begin{aligned} (4.2) \quad & \sum_{\substack{K_{\bar{v}^i} \in \mathcal{D}^{d_{\bar{v}^i}} \\ K_{\bar{v}^5} \in \mathcal{D}^{d_{\bar{v}^5}}}} \sum_{\substack{I_{\bar{v}^i}, J_{\bar{v}^i} \in \mathcal{D}^{d_{\bar{v}^i}} \\ I_{\bar{v}^i}^{(k_{\bar{v}^i})} = K_{\bar{v}^i} \\ J_{\bar{v}^i}^{(l_{\bar{v}^i})} = K_{\bar{v}^i}}} \alpha_{K_{\bar{v}^i}, I_{\bar{v}^i}, J_{\bar{v}^i}} \beta_{I_{\bar{v}^1}, J_{\bar{v}^3}} \\ & \times \left\langle b, h_{I_{\bar{v}^1}^{(\bar{s})}} \otimes \frac{1_{I_{\bar{v}^2}}}{|I_{\bar{v}^2}|} \otimes h_{J_{\bar{v}^3}^{(\bar{t})}} \otimes \frac{1_{J_{\bar{v}^4}}}{|J_{\bar{v}^4}|} \otimes h_{K_{\bar{v}^5}} \right\rangle \\ & \times \left\langle f, h_{I_{\bar{v}^1}} \otimes h_{I_{\bar{v}^2}} \otimes h_{I_{\bar{v}^3}} \otimes h_{I_{\bar{v}^4}} \otimes \frac{1_{K_{\bar{v}^5}}}{|K_{\bar{v}^5}|} \right\rangle \\ & \times \tilde{h}_{J_{\bar{v}^1}} \otimes h_{J_{\bar{v}^2}} \otimes \tilde{h}_{J_{\bar{v}^3}} \otimes h_{J_{\bar{v}^4}} \otimes h_{K_{\bar{v}^5}}, \end{aligned}$$

where $i = 1, 2, 3, 4$, $k_{v_j^i}, l_{v_j^i} \geq 0$ for $j \in \mathcal{I}_i$, $s_j \leq k_{v_j^1}$ for $j \in \mathcal{I}_1$, $t_j \leq l_{v_j^3}$ for $j \in \mathcal{I}_3$, and

$$\begin{aligned} |\alpha_{K_{\bar{v}^i}, I_{\bar{v}^i}, J_{\bar{v}^i}}| &\leq \prod_{j=1}^4 \frac{|I_{\bar{v}^j}|^{\frac{1}{2}} |J_{\bar{v}^j}|^{\frac{1}{2}}}{|K_{\bar{v}^j}|}, \\ |\beta_{I_{\bar{v}^1}, J_{\bar{v}^3}}| &\leq |I_{\bar{v}^1}^{(\bar{s})}|^{-\frac{1}{2}} |J_{\bar{v}^3}^{(\bar{t})}|^{-\frac{1}{2}}, \\ |\tilde{h}_Q| &\leq |Q|^{-\frac{1}{2}} 1_Q. \end{aligned}$$

Moreover, if for example $\mathcal{I}_j = \emptyset$, then the related terms are understood as 1 and we require that $\bigcup_{j=1,3,5} \mathcal{I}_j \neq \emptyset$.

In addition, similar to the proof of Lemma 3.1, we omit the details if some $\mathcal{I}_j = \emptyset$.

Lemma 4.2. *We have*

$$\|\varphi\|_{L^p(\lambda)} \lesssim_{[\nu]_{A_p}} \|b\|_{\text{BMO}^{\bar{u}}(\nu)} \|f\|_{L^p(\mu)},$$

where φ is the term (4.2) defined above and $\bar{u} = (u_j)$, $u_j \in \bigcup_{i=1,3,5} \mathcal{I}_i$.

Proof: We begin by using the size conditions of α and β for the dual form. Hence, we have

$$\begin{aligned} & \left| \sum_{K_{\bar{v}^i}, K_{\bar{v}^5}} \sum_{\substack{I_{\bar{v}^i}^{(k_{\bar{v}^i})} = K_{\bar{v}^i} \\ J_{\bar{v}^i}^{(l_{\bar{v}^i})} = K_{\bar{v}^i}}} \alpha_{K_{\bar{v}^i}, I_{\bar{v}^i}, J_{\bar{v}^i}} \beta_{I_{\bar{v}^1}, J_{\bar{v}^3}} \left\langle b, h_{I_{\bar{v}^1} \times J_{\bar{v}^3}^{(\bar{s})}} \otimes \frac{1_{I_{\bar{v}^2} \times J_{\bar{v}^4}}}{|I_{\bar{v}^2} \times J_{\bar{v}^4}|} \otimes h_{K_{\bar{v}^5}} \right\rangle \right. \\ & \quad \times \left. \left\langle f, h_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3} \times I_{\bar{v}^4}} \otimes \frac{1_{K_{\bar{v}^5}}}{|K_{\bar{v}^5}|} \right\rangle \langle g, \tilde{h}_{J_{\bar{v}^1}} \otimes \tilde{h}_{J_{\bar{v}^3}} \otimes h_{J_{\bar{v}^2} \times J_{\bar{v}^4} \times K_{\bar{v}^5}} \rangle \right| \\ & \leq \sum_{K_{\bar{v}^i}, K_{\bar{v}^5}} \sum_{\substack{I_{\bar{v}^i}^{(k_{\bar{v}^i})} = K_{\bar{v}^i} \\ J_{\bar{v}^i}^{(l_{\bar{v}^i})} = K_{\bar{v}^i}}} \prod_{j=1}^4 \frac{|I_{\bar{v}^j}|^{\frac{1}{2}} |J_{\bar{v}^j}|^{\frac{1}{2}}}{|K_{\bar{v}^j}|} |I_{\bar{v}^1}^{(\bar{s})}|^{-\frac{1}{2}} |J_{\bar{v}^3}^{(\bar{t})}|^{-\frac{1}{2}} |I_{\bar{v}^2} \times J_{\bar{v}^4}|^{-1} \\ & \quad \times \int_{I_{\bar{v}^2} \times J_{\bar{v}^4}} \left[\left| \langle b, h_{I_{\bar{v}^1} \times J_{\bar{v}^3}^{(\bar{s})}} \otimes h_{K_{\bar{v}^5}} \rangle_{\bar{v}^1, \bar{v}^3, \bar{v}^5} \right| \left| \left\langle f, h_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3} \times I_{\bar{v}^4}} \otimes \frac{1_{K_{\bar{v}^5}}}{|K_{\bar{v}^5}|} \right\rangle \right| \right. \\ & \quad \times \left. \left| \langle g, \tilde{h}_{J_{\bar{v}^1}} \otimes \tilde{h}_{J_{\bar{v}^3}} \otimes h_{J_{\bar{v}^2} \times J_{\bar{v}^4} \times K_{\bar{v}^5}} \rangle \right| \right] \\ & = \sum_{K_{\bar{v}^2}, K_{\bar{v}^4}} \sum_{I_{\bar{v}^2}^{(k_{\bar{v}^2})} = K_{\bar{v}^2}} \sum_{J_{\bar{v}^4}^{(l_{\bar{v}^4})} = K_{\bar{v}^4}} \int_{I_{\bar{v}^2} \times J_{\bar{v}^4}} A^b(x_{\bar{v}^2}, \bar{v}^4) dx_{\bar{v}^2}, \end{aligned}$$

where

$$\begin{aligned} A^b(x_{\bar{v}^2}, \bar{v}^4) &:= \sum_{K_{\bar{v}^1}, K_{\bar{v}^3}, K_{\bar{v}^5}} \sum_{\substack{I_{\bar{v}^1}^{(k_{\bar{v}^1}-\bar{s})} = K_{\bar{v}^1} \\ Q_{\bar{v}^3}^{(l_{\bar{v}^3}-\bar{t})} = K_{\bar{v}^3}}} \sum_{I_{\bar{v}^1}^{(\bar{s})} = P_{\bar{v}^1}} \sum_{I_{\bar{v}^3}^{(k_{\bar{v}^3})} = K_{\bar{v}^3}} \sum_{I_{\bar{v}^4}^{(k_{\bar{v}^4})} = K_{\bar{v}^4}} \sum_{J_{\bar{v}^1}^{(l_{\bar{v}^1})} = K_{\bar{v}^1}} \sum_{J_{\bar{v}^2}^{(l_{\bar{v}^2})} = K_{\bar{v}^2}} \\ & \times \prod_{j=1}^4 \frac{|I_{\bar{v}^j}|^{\frac{1}{2}} |J_{\bar{v}^j}|^{\frac{1}{2}}}{|K_{\bar{v}^j}|} |P_{\bar{v}^1}|^{-\frac{1}{2}} |Q_{\bar{v}^3}|^{-\frac{1}{2}} |I_{\bar{v}^2} \times J_{\bar{v}^4}|^{-1} \left| \langle b_{x_{\bar{v}^2}, \bar{v}^4}, h_{P_{\bar{v}^1} \times Q_{\bar{v}^3}} \otimes h_{K_{\bar{v}^5}} \rangle \right| \\ & \times \left| \left\langle f, h_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3} \times I_{\bar{v}^4}} \otimes \frac{1_{K_{\bar{v}^5}}}{|K_{\bar{v}^5}|} \right\rangle \right| \left| \langle g, \tilde{h}_{J_{\bar{v}^1}} \otimes \tilde{h}_{J_{\bar{v}^3}} \otimes h_{J_{\bar{v}^2} \times J_{\bar{v}^4} \times K_{\bar{v}^5}} \rangle \right|. \end{aligned}$$

Then we proceed by replacing f and g with suitable multi-parameter martingale blocks, i.e.

$$\begin{aligned}
A^b(x_{\bar{v}^2, \bar{v}^4}) &= \sum_{K_{\bar{v}^1}, K_{\bar{v}^3}, K_{\bar{v}^5}} \sum_{\substack{P_{\bar{v}^1}^{(k_{\bar{v}^1} - \bar{s})} = K_{\bar{v}^1} \\ Q_{\bar{v}^3}^{(l_{\bar{v}^3} - \bar{t})} = K_{\bar{v}^3}}} \sum_{\substack{I_{\bar{v}^1}^{(\bar{s})} = P_{\bar{v}^1} \\ J_{\bar{v}^3}^{(\bar{t})} = Q_{\bar{v}^3}}} \sum_{\substack{I_{\bar{v}^3}^{(k_{\bar{v}^3})} = K_{\bar{v}^3} \\ J_{\bar{v}^1}^{(l_{\bar{v}^1})} = K_{\bar{v}^1}}} \sum_{\substack{I_{\bar{v}^4}^{(k_{\bar{v}^4})} = K_{\bar{v}^4} \\ J_{\bar{v}^2}^{(l_{\bar{v}^2})} = K_{\bar{v}^2}}} \\
&\times \prod_{j=1}^4 \frac{|I_{\bar{v}j}|^{\frac{1}{2}} |J_{\bar{v}j}|^{\frac{1}{2}}}{|K_{\bar{v}j}|} |P_{\bar{v}^1}|^{-\frac{1}{2}} |Q_{\bar{v}^3}|^{-\frac{1}{2}} |I_{\bar{v}^2} \times J_{\bar{v}^4}|^{-1} |\langle b_{x_{\bar{v}^2, \bar{v}^4}}, h_{P_{\bar{v}^1} \times Q_{\bar{v}^3}} \otimes h_{K_{\bar{v}^5}} \rangle| \\
&\times \left| \left\langle \Delta_{K_{\bar{v}^1} \times K_{\bar{v}^2} \times K_{\bar{v}^3} \times K_{\bar{v}^4}, (k_{\bar{v}j})_{j=1,2,3,4}} f, h_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3} \times I_{\bar{v}^4}} \otimes \frac{1_{K_{\bar{v}^5}}}{|K_{\bar{v}^5}|} \right\rangle \right| \\
&\times |\langle \Delta_{K_{\bar{v}^2} \times K_{\bar{v}^4} \times K_{\bar{v}^5}, (l_{\bar{v}^2}, l_{\bar{v}^4}, \bar{0})} g, \tilde{h}_{J_{\bar{v}^1}} \otimes \tilde{h}_{J_{\bar{v}^3}} \otimes h_{J_{\bar{v}^2} \times J_{\bar{v}^4} \times K_{\bar{v}^5}} \rangle| \\
&\leq \sum_{K_{\bar{v}^1}, K_{\bar{v}^3}, K_{\bar{v}^5}} \sum_{\substack{P_{\bar{v}^1}^{(k_{\bar{v}^1} - \bar{s})} = K_{\bar{v}^1} \\ Q_{\bar{v}^3}^{(l_{\bar{v}^3} - \bar{t})} = K_{\bar{v}^3}}} \sum_{\bar{v}^1} |P_{\bar{v}^1}|^{\frac{1}{2}} |Q_{\bar{v}^3}|^{\frac{1}{2}} |K_{\bar{v}^5}|^{\frac{1}{2}} \\
&\times |\langle b_{x_{\bar{v}^2, \bar{v}^4}}, h_{P_{\bar{v}^1} \times Q_{\bar{v}^3}} \otimes h_{K_{\bar{v}^5}} \rangle| \\
&\times \langle |\Delta_{K_{\bar{v}^1} \times K_{\bar{v}^2} \times K_{\bar{v}^3} \times K_{\bar{v}^4}, (k_{\bar{v}j})_{j=1,2,3,4}} f| \rangle_{P_{\bar{v}^1} \times I_{\bar{v}^2} \times K_{\bar{v}^3} \times K_{\bar{v}^4} \times K_{\bar{v}^5}} \\
&\times \langle |\Delta_{K_{\bar{v}^2} \times K_{\bar{v}^4} \times K_{\bar{v}^5}, (l_{\bar{v}^2}, l_{\bar{v}^4}, \bar{0})} g| \rangle_{K_{\bar{v}^1} \times K_{\bar{v}^2} \times Q_{\bar{v}^3} \times J_{\bar{v}^4} \times K_{\bar{v}^5}},
\end{aligned}$$

where we summed up rectangles of levels \bar{s} , \bar{t} , $k_{\bar{v}^3}$, $l_{\bar{v}^1}$, $k_{\bar{v}^4}$, $l_{\bar{v}^2}$ after the modulus is taken inside of the pairings of martingale blocks of f and g .

Now, using Proposition 2.9 for A^b with fixed $x_{\bar{v}^2, \bar{v}^4}$ we get

$$\begin{aligned}
A^b(x_{\bar{v}^2, \bar{v}^4}) &\lesssim \|b\|_{\text{BMO}^{\bar{u}}(\nu)} \int_{\mathbb{R}^{d_{\bar{v}^1} + d_{\bar{v}^3} + d_{\bar{v}^5}}} \left(\sum_{K_{\bar{v}^1}, K_{\bar{v}^3}} \right. \\
&\times [M^{\bar{v}^1, \bar{v}^3, \bar{v}^5} \langle |\Delta_{K_{\bar{v}^1} \times K_{\bar{v}^2} \times K_{\bar{v}^3} \times K_{\bar{v}^4}, (k_{\bar{v}j})_{j=1,2,3,4}} f| \rangle_{I_{\bar{v}^2} \times K_{\bar{v}^4}, \bar{v}^2, \bar{v}^4}]^2 1_{K_{\bar{v}^1} K_{\bar{v}^3}} \Big)^{\frac{1}{2}} \\
&\times \left(\sum_{K_{\bar{v}^5}} [M^{\bar{v}^1, \bar{v}^3, \bar{v}^5} \langle |\Delta_{K_{\bar{v}^2} \times K_{\bar{v}^4} \times K_{\bar{v}^5}, (l_{\bar{v}^2}, l_{\bar{v}^4}, \bar{0})} g| \rangle_{K_{\bar{v}^2} \times J_{\bar{v}^4}, \bar{v}^2, \bar{v}^4}]^2 1_{K_{\bar{v}^5}} \right)^{\frac{1}{2}} \nu_{x_{\bar{v}^2, \bar{v}^4}},
\end{aligned}$$

where again we summed over the rectangles $P_{\bar{v}^1}$, $Q_{\bar{v}^3}$, and $\bar{u} = (u_j)$, $u_j \in \bigcup_{i=1,3,5} \mathcal{I}_i$. Notice that here we needed the requirement that $\bigcup_{i=1,3,5} \mathcal{I}_i \neq \emptyset$.

Hence, we can conclude that

$$\begin{aligned}
|\langle \varphi, g \rangle| &\lesssim \|b\|_{\text{BMO}^{\bar{u}}(\nu)} \int \sum_{K_{\bar{v}2}, K_{\bar{v}4}} \sum_{I_{\bar{v}2}^{(k_{\bar{v}2})} = K_{\bar{v}2}} \sum_{J_{\bar{v}4}^{(l_{\bar{v}4})} = K_{\bar{v}4}} \left(\sum_{K_{\bar{v}1}, K_{\bar{v}3}} \right. \\
&\times [M^{\bar{v}1, \bar{v}3, \bar{v}5} \langle |\Delta_{K_{\bar{v}1} \times K_{\bar{v}2} \times K_{\bar{v}3} \times K_{\bar{v}4}, (k_{\bar{v}j})_{j=1,2,3,4}} f \rangle|_{I_{\bar{v}2} \times K_{\bar{v}4}, \bar{v}2, \bar{v}4}]^2 1_{K_{\bar{v}1} \times K_{\bar{v}3}} \Big)^{\frac{1}{2}} \\
&\times \left(\sum_{K_{\bar{v}5}} [M^{\bar{v}1, \bar{v}3, \bar{v}5} \langle |\Delta_{K_{\bar{v}2} \times K_{\bar{v}4} \times K_{\bar{v}5}, (l_{\bar{v}2}, l_{\bar{v}4}, \bar{0})} g \rangle|_{K_{\bar{v}2} \times J_{\bar{v}4}, \bar{v}2, \bar{v}4}]^2 1_{K_{\bar{v}5}} \right)^{\frac{1}{2}} 1_{I_{\bar{v}2} \times J_{\bar{v}4}} \nu \\
&=: \|b\|_{\text{BMO}^{\bar{u}}(\nu)} \times \text{I}.
\end{aligned}$$

Using standard estimates we get

$$\begin{aligned}
\text{I} &\lesssim \int \left(\sum_{\substack{K_{\bar{v}1} K_{\bar{v}2} \\ K_{\bar{v}3}, K_{\bar{v}4}}} [M^{\bar{m}} \Delta_{K_{\bar{v}1} \times K_{\bar{v}2} \times K_{\bar{v}3} \times K_{\bar{v}4}, (k_{\bar{v}j})_{j=1,2,3,4}} f]^2 1_{K_{\bar{v}1} \times K_{\bar{v}2} \times K_{\bar{v}3} \times K_{\bar{v}4}} \right)^{\frac{1}{2}} \\
&\times \left(\sum_{K_{\bar{v}2}, K_{\bar{v}4}, K_{\bar{v}5}} [M^{\bar{m}} \Delta_{K_{\bar{v}2} \times K_{\bar{v}4} \times K_{\bar{v}5}, (l_{\bar{v}2}, l_{\bar{v}4}, \bar{0})} g]^2 1_{K_{\bar{v}2} \times K_{\bar{v}4} \times K_{\bar{v}5}} \right)^{\frac{1}{2}} \nu \\
&\lesssim \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda^{1-p'})}. \quad \square
\end{aligned}$$

For simplicity, we begin with the case of two iterations.

Theorem 4.3. *Let $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2\}$ be some partition of $\{1, 2, \dots, m\}$ such that $\mathcal{I}_1 \neq \emptyset \neq \mathcal{I}_2$. Let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_m})$, $\bar{u} = (u_i)_{i \in \mathcal{I}_1}$, and $\bar{v} = (v_i)_{i \in \mathcal{I}_2}$. We have*

$$\|[\mathcal{S}^{\bar{u}}, [b, \mathcal{S}^{\bar{v}}]]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{BMO}^{\mathcal{I}}(\nu)}$$

for $1 < p < \infty$.

Before the proof, we make a small remark. We can begin with the commutator $[b, \mathcal{S}^{\bar{m}}]$. However, in this case, b is in the little BMO space. In [34] this is proved for bi-parameter operators, i.e. the case $m = 2$. The method used there can be applied to the multi-parameter case. Hence, the result of [34] regarding the first order shift case extends to the multi-parameter framework. We omit the details.

Proof: We say that the number of parameters in \mathcal{I}_1 is n . We begin by expanding appearing products in all of the parameters. Hence, we have

$$\begin{aligned}
[\mathcal{S}^{\bar{u}}[b, \mathcal{S}^{\bar{v}}]]f &= \mathcal{S}^{\bar{u}}(b\mathcal{S}^{\bar{v}}f) - \mathcal{S}^{\bar{u}}\mathcal{S}^{\bar{v}}(bf) - b\mathcal{S}^{\bar{u}}\mathcal{S}^{\bar{v}}f + \mathcal{S}^{\bar{v}}(b\mathcal{S}^{\bar{u}}f) \\
&= \sum_{\substack{i_{\bar{u}} \in \{1,2,3\}^n \\ i_{\bar{v}} \in \{1,2,3\}^{m-n}}} \mathcal{S}^{\bar{u}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{v}}f)) - \mathcal{S}^{\bar{u}}\mathcal{S}^{\bar{v}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, f)) \\
&\quad - A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}}\mathcal{S}^{\bar{v}}f) + \mathcal{S}^{\bar{v}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}}f)).
\end{aligned}$$

Now, if $i_{\bar{u}} \neq \{3\}^n$ and $i_{\bar{v}} \neq \{3\}^{m-n}$, then each individual term is bounded by combining boundedness of the multi-parameter shifts with Lemma 3.1. Hence, it is enough to consider terms in the following sums

$$(4.3) \quad \sum_{\substack{i_{\bar{u}}=\{3\}^n \\ i_{\bar{v}} \in \{1,2,3\}^{m-n} \setminus \{3\}^{m-n}}} + \sum_{\substack{i_{\bar{u}} \in \{1,2,3\}^n \setminus \{3\}^n \\ i_{\bar{v}}=\{3\}^{m-n}}} + \sum_{\substack{i_{\bar{u}}=\{3\}^n \\ i_{\bar{v}}=\{3\}^{m-n}}}.$$

The terms in the first two sums are similar. Hence, considering the first sum, we are essentially handling the second one simultaneously and we choose to deal with the first one.

Fix $i_{\bar{u}} = \{3\}^n$, $i_{\bar{v}} \in \{1,2,3\}^{m-n} \setminus \{3\}^{m-n}$. We pair $\mathcal{S}^{\bar{u}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{v}} f))$ with $A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}} \mathcal{S}^{\bar{v}} f)$ and $\mathcal{S}^{\bar{v}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}} f))$ with $\mathcal{S}^{\bar{u}} \mathcal{S}^{\bar{v}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, f))$. It is enough to study

$$A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}} f) - \mathcal{S}^{\bar{u}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, f)),$$

since $\mathcal{S}^{\bar{u}} \mathcal{S}^{\bar{v}} = \mathcal{S}^{\bar{v}} \mathcal{S}^{\bar{u}}$ and $\mathcal{S}^{\bar{v}}$ is bounded.

We remark that when considering the second sum in (4.3), the terms need to be paired in the other order and then $\mathcal{S}^{\bar{u}}$ can be left out by a similar argument. Generally, we pair the terms so that we get rid of the shifts on the parameters where the paraproduct operator is a legal one.

Let us recall the definition of the paraproduct operator $A_{i_{\bar{v}}}^{\bar{v}}$. Let $\mathcal{J}_s = \{j \leq m-n : i_{v_j} = s\}$, for $s = 1, 2, 3$, $\bar{v}^1 = (v_j)_{j \in \mathcal{J}_1}$, $\bar{v}^2 = (v_j)_{j \in \mathcal{J}_2}$, and $\bar{v}^3 = (v_j)_{j \in \mathcal{J}_3}$. Thus

$$\begin{aligned} A_{i_{\bar{v}}}^{\bar{v}}(b, f) &= \sum_{K_{\bar{v}^1}, K_{\bar{v}^2}, K_{\bar{v}^3}} \left\langle b, h_{K_{\bar{v}^1,2}} \otimes \frac{1_{K_{\bar{v}^3}}}{|K_{\bar{v}^3}|} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \\ &\quad \times \left\langle f, h_{K_{\bar{v}^1}} \otimes \frac{1_{K_{\bar{v}^2}}}{|K_{\bar{v}^2}|} \otimes h_{K_{\bar{v}^3}} \right\rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3} \otimes h_{K_{\bar{v}^1}} h_{K_{\bar{v}^2}} \otimes h_{K_{\bar{v}^3}}, \end{aligned}$$

where $K_{\bar{v}^1,2}$ denotes the rectangle $K_{\bar{v}^1} \times K_{\bar{v}^2}$. Also recall the definition of multi-parameter shift

$$\mathcal{S}^{\bar{u}} f = \sum_{K_{\bar{u}}} \sum_{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}}} \sum_{J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \langle f, h_{I_{\bar{u}}} \rangle_{\bar{u}} \otimes h_{J_{\bar{u}}}.$$

Hence, we have

$$\begin{aligned} (4.4) \quad & A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, \mathcal{S}^{\bar{u}} f) - \mathcal{S}^{\bar{u}}(A_{i_{\bar{u}}}^{\bar{u}} A_{i_{\bar{v}}}^{\bar{v}}(b, f)) = \sum_{\substack{K_{\bar{v}^1}, K_{\bar{v}^2}, K_{\bar{v}^3} \\ K_{\bar{u}}}} \sum_{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}}} \sum_{J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \\ & \times \left[\left\langle b, h_{K_{\bar{v}^1,2}} \otimes \frac{1_{K_{\bar{v}^3}}}{|K_{\bar{v}^3}|} \otimes \frac{1_{J_{\bar{u}}}}{|J_{\bar{u}}|} \right\rangle - \left\langle b, h_{K_{\bar{v}^1,2}} \otimes \frac{1_{K_{\bar{v}^3}}}{|K_{\bar{v}^3}|} \otimes \frac{1_{I_{\bar{u}}}}{|I_{\bar{u}}|} \right\rangle \right] \\ & \times \left\langle f, h_{K_{\bar{v}^1}} \otimes \frac{1_{K_{\bar{v}^2}}}{|K_{\bar{v}^2}|} \otimes h_{K_{\bar{v}^3}} \otimes h_{I_{\bar{u}}} \right\rangle h_{K_{\bar{v}^1}} h_{K_{\bar{v}^2}} \otimes h_{K_{\bar{v}^3}} \otimes h_{J_{\bar{u}}}, \end{aligned}$$

where \bar{v}^1 , \bar{v}^2 , \bar{v}^3 , and $K_{\bar{v}^1,2}$ are defined as above.

By Lemma 4.1, we can write (4.4) as

$$\begin{aligned}
& \sum_{\substack{K_{\bar{v}1}, K_{\bar{v}2}, K_{\bar{v}3} \\ K_{\bar{u}}}} \sum_{\substack{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}} \\ J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}}} \sum_{J_{\bar{u}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \\
& \times \left[\sum_{j=1}^n \left(\sum_{t=1}^{l_{u_j}} \langle \Delta_{J_{u_j}^{(t)}}^{u_j} b_{K_{\bar{v}}} \rangle_{J_{u_j} \times Q_{\bar{u}'_j}} - \sum_{s=1}^{k_{u_j}} \langle \Delta_{I_{u_j}^{(s)}}^{u_j} b_{K_{\bar{v}}} \rangle_{I_{u_j} \times Q_{\bar{u}'_j}} \right) \right] \\
& \times \left\langle f, h_{K_{\bar{v}1}} \otimes \frac{1_{K_{\bar{v}2}}}{|K_{\bar{v}2}|} \otimes h_{K_{\bar{v}3}} \otimes h_{I_{\bar{u}}} \right\rangle h_{K_{\bar{v}1}} h_{K_{\bar{v}1}} \otimes h_{K_{\bar{v}2}} \otimes h_{K_{\bar{v}3}} \otimes h_{J_{\bar{u}}},
\end{aligned}$$

where $b_{K_{\bar{v}}} = \langle b, h_{K_{\bar{v}1,2}} \otimes 1_{K_{\bar{v}3}} | K_{\bar{v}3} \rangle_{\bar{v}^1, \bar{v}^2, \bar{v}^3}$ and $Q_{\bar{u}'_j} = I_{(u_i)_{1 \leq i < j}} \times J_{(u_i)_{j < i \leq n}}$.

Now these terms are expanded to a desired form, i.e. there is a cancellative Haar function on some parameter in \mathcal{I}_1 and \mathcal{I}_2 paired with the function b . For example, the first term in the above pair with $j = 2$ equals to

$$\begin{aligned}
& \sum_{t=1}^{l_{u_2}} \sum_{\substack{K_{\bar{v}1}, K_{\bar{v}2}, K_{\bar{v}3} \\ K_{\bar{u}}}} \sum_{\substack{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}} \\ J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}}} \sum_{J_{\bar{u}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \langle h_{J_{u_2}^{(t)}} \rangle_{J_{u_2}} \\
& \times \left\langle b, h_{K_{\bar{v}1,2}} \otimes \frac{1_{K_{\bar{v}3}}}{|K_{\bar{v}3}|} \otimes \frac{1_{I_{u_1}}}{|I_{u_1}|} \otimes h_{J_{u_2}^{(t)}} \otimes \frac{1_{J_{\bar{u}'}}}{|J_{\bar{u}'}|} \right\rangle \\
& \times \left\langle f, h_{K_{\bar{v}1}} \otimes \frac{1_{K_{\bar{v}2}}}{|K_{\bar{v}2}|} \otimes h_{K_{\bar{v}3}} \otimes h_{I_{\bar{u}}} \right\rangle h_{K_{\bar{v}1}} h_{K_{\bar{v}1}} \otimes h_{K_{\bar{v}2}} \otimes h_{K_{\bar{v}3}} \otimes h_{J_{\bar{u}}},
\end{aligned}$$

where $\bar{u}' = (u_j)_{j=3}^n$ and $|\langle h_{J_{u_2}^{(t)}} \rangle_{J_{u_2}}| = |J_{u_2}^{(t)}|^{-\frac{1}{2}}$. Hence, these terms are bounded by Lemma 4.2.

Let us also expand the last term of (4.3). Here we can not do any reductions and we sum everything together. Hence, the term equals to

$$\begin{aligned}
(4.5) \quad & \sum_{K_{\bar{u}}, K_{\bar{v}}} \sum_{\substack{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}} \\ J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}}} \sum_{\substack{I_{\bar{v}}^{(k_{\bar{v}})} = K_{\bar{v}} \\ J_{\bar{v}}^{(l_{\bar{v}})} = K_{\bar{v}}}} a_{K_{\bar{v}}, I_{\bar{v}}, J_{\bar{v}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \\
& \times [\langle b \rangle_{I_{\bar{u}} \times J_{\bar{v}}} - \langle b \rangle_{I_{\bar{u}} \times I_{\bar{v}}} - \langle b \rangle_{J_{\bar{u}} \times J_{\bar{v}}} + \langle b \rangle_{J_{\bar{u}} \times I_{\bar{v}}}] \\
& \times \langle f, h_{I_{\bar{u}}} \otimes h_{I_{\bar{v}}} \rangle h_{J_{\bar{u}}} \otimes h_{J_{\bar{v}}}.
\end{aligned}$$

Here we proceed similarly, as with the previous terms, but now expanding on both parameter sets \mathcal{I}_1 and \mathcal{I}_2 .

First, we apply Lemma 4.1 in the parameters \mathcal{I}_1 . Hence, we have

$$\begin{aligned}
 & \langle b \rangle_{I_{\bar{u}} \times J_{\bar{v}}} - \langle b \rangle_{I_{\bar{u}} \times I_{\bar{v}}} - \langle b \rangle_{J_{\bar{u}} \times J_{\bar{v}}} + \langle b \rangle_{J_{\bar{u}} \times I_{\bar{v}}} \\
 &= (\langle b \rangle_{J_{\bar{u}} \times I_{\bar{v}}} - \langle b \rangle_{I_{\bar{u}} \times I_{\bar{v}}}) - (\langle b \rangle_{J_{\bar{u}} \times J_{\bar{v}}} - \langle b \rangle_{I_{\bar{u}} \times J_{\bar{v}}}) \\
 (4.6) \quad &= \sum_{j_1=1}^n \left(\sum_{t_1=1}^{l_{u_{j_1}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)}}^{u_{j_1}} b_{I_{\bar{v}}} \rangle_{J_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} - \sum_{s_1=1}^{k_{u_{j_1}}} \langle \Delta_{I_{u_{j_1}}^{(s_1)}}^{u_{j_1}} b_{I_{\bar{v}}} \rangle_{I_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} \right) \\
 &\quad - \sum_{j_1=1}^n \left(\sum_{t_1=1}^{l_{u_{j_1}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)}}^{u_{j_1}} b_{J_{\bar{v}}} \rangle_{J_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} - \sum_{s_1=1}^{k_{u_{j_1}}} \langle \Delta_{I_{u_{j_1}}^{(s_1)}}^{u_{j_1}} b_{J_{\bar{v}}} \rangle_{I_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} \right),
 \end{aligned}$$

where $b_{I_{\bar{v}}} = \langle b \rangle_{I_{\bar{v}}, \bar{v}}$, $b_{J_{\bar{v}}} = \langle b \rangle_{J_{\bar{v}}, \bar{v}}$, and $Q_{\bar{u}'_{j_1}} = I_{(u_j)_{1 \leq j < j_1}} \times J_{(u_j)_{j_1 < j \leq n}}$. Then we pair these terms in the other order and apply Lemma 4.1 in the parameters \mathcal{I}_2 . Then, for example, the pair of the first and third term on the right-hand side of (4.6) equals to

$$\begin{aligned}
 & \sum_{j_1=1}^n \left(\sum_{t_1=1}^{l_{u_{j_1}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)}}^{u_{j_1}} b_{I_{\bar{v}}} \rangle_{J_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} - \langle \Delta_{J_{u_{j_1}}^{(t_1)}}^{u_{j_1}} b_{J_{\bar{v}}} \rangle_{J_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} \right) \\
 (4.7) \quad &= \sum_{j_1=1}^n \sum_{t_1=1}^{l_{u_{j_1}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)}}^{u_{j_1}} (\langle b \rangle_{I_{\bar{v}}, \bar{v}} - \langle b \rangle_{J_{\bar{v}}, \bar{v}}) \rangle_{J_{u_{j_1}} \times Q_{\bar{u}'_{j_1}}} \\
 &= \sum_{j_1=1}^n \sum_{t_1=1}^{l_{u_{j_1}}} \sum_{j_2=1}^{m-n} \left(\sum_{s_2=1}^{k_{v_{j_2}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)} \times I_{v_{j_2}}^{(s_2)}}^{u_{j_1}, v_{j_2}} b \rangle_{J_{u_{j_1}} \times I_{v_{j_2}} \times Q_{\bar{u}'_{j_1}} \times Q_{\bar{v}'_{j_2}}} \right. \\
 &\quad \left. - \sum_{t_2=1}^{l_{v_{j_2}}} \langle \Delta_{J_{u_{j_1}}^{(t_1)} \times J_{v_{j_2}}^{(t_2)}}^{u_{j_1}, v_{j_2}} b \rangle_{J_{u_{j_1}} \times J_{v_{j_2}} \times Q_{\bar{u}'_{j_1}} \times Q_{\bar{v}'_{j_2}}} \right),
 \end{aligned}$$

where $Q_{\bar{v}'_{j_2}} = I_{(v_j)_{1 \leq j < j_2}} \times J_{(v_j)_{j_2 < j \leq m-n}}$.

Thus, for example, the first term on the right-hand side of (4.7) with $j_1 = 2$ and $j_2 = 2$ related to (4.5) equals to

$$\begin{aligned}
 & \sum_{t_1=1}^{l_{u_2}} \sum_{s_2=1}^{k_{v_2}} \sum_{K_{\bar{u}}, K_{\bar{v}}} \sum_{\substack{I_{\bar{u}}^{(k_{\bar{u}})} = K_{\bar{u}} \\ J_{\bar{u}}^{(l_{\bar{u}})} = K_{\bar{u}}}} \sum_{\substack{I_{\bar{v}}^{(k_{\bar{v}})} = K_{\bar{v}} \\ J_{\bar{v}}^{(l_{\bar{v}})} = K_{\bar{v}}}} a_{K_{\bar{v}}, I_{\bar{v}}, J_{\bar{v}}} a_{K_{\bar{u}}, I_{\bar{u}}, J_{\bar{u}}} \langle h_{J_{u_2}^{(t_1)}} \rangle_{J_{u_2}} \langle h_{I_{v_2}^{(s_2)}} \rangle_{I_{v_2}} \\
 (4.8) \quad & \times \left\langle b, \frac{1_{I_{u_1}}}{|I_{u_1}|} \otimes h_{J_{u_2}^{(t_1)}} \otimes \frac{1_{J_{\bar{u}'}}}{|J_{\bar{u}'}|} \otimes \frac{1_{I_{v_1}}}{|I_{v_1}|} \otimes h_{I_{v_2}^{(s_2)}} \otimes \frac{1_{J_{\bar{v}'}}}{|J_{\bar{v}'}|} \right\rangle \\
 & \times \langle f, h_{I_{\bar{u}}} \otimes h_{I_{\bar{v}}} \rangle h_{J_{\bar{u}}} \otimes h_{J_{\bar{v}}},
 \end{aligned}$$

where $\bar{u}' = (u_j)_{j=3}^n$ and $\bar{v}' = (v_j)_{j=3}^{m-n}$. Hence, the boundedness of these terms follows from Lemma 4.2. \square

Theorem 4.4. *Let $k \leq m$ be an integer and let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m})$ and $1 < p < \infty$. For all partitions $\mathcal{I} = \{\mathcal{I}_i : i \leq k\}$ of $\{1, 2, \dots, m\}$ we have*

$$\|[\mathcal{S}^{\bar{v}^1}, [\mathcal{S}^{\bar{v}^2}, \dots [b, \mathcal{S}^{\bar{v}^k}]]]\|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{bmo}^\mathcal{I}(\nu)},$$

where $\bar{v}^i = (v_j^i)_{v_j^i \in \mathcal{I}_i}$.

Proof: We consider here the case $k = 3$. The aim is to show that the strategy and techniques used in the case $k = 2$ work here also. The general case follows similarly.

Let $n_1 + n_2 + n_3 = m$ and say the number of parameters in \bar{v}^1 is n_1 and so on. By definition

$$\begin{aligned} [\mathcal{S}^{\bar{v}^1}, [\mathcal{S}^{\bar{v}^2}, [b, \mathcal{S}^{\bar{v}^3}]]]f &= \mathcal{S}^{\bar{v}^1} \mathcal{S}^{\bar{v}^2} (b \mathcal{S}^{\bar{v}^3} f) - \mathcal{S}^{\bar{v}^1} \mathcal{S}^{\bar{v}^2} \mathcal{S}^{\bar{v}^3} (bf) \\ &\quad - \mathcal{S}^{\bar{v}^1} (b \mathcal{S}^{\bar{v}^3} \mathcal{S}^{\bar{v}^2} f) + \mathcal{S}^{\bar{v}^1} \mathcal{S}^{\bar{v}^3} (b \mathcal{S}^{\bar{v}^2} f) \\ &\quad - \mathcal{S}^{\bar{v}^2} (b \mathcal{S}^{\bar{v}^3} \mathcal{S}^{\bar{v}^1} f) + \mathcal{S}^{\bar{v}^2} \mathcal{S}^{\bar{v}^3} (b \mathcal{S}^{\bar{v}^1} f) \\ &\quad + b \mathcal{S}^{\bar{v}^3} \mathcal{S}^{\bar{v}^2} \mathcal{S}^{\bar{v}^1} f - \mathcal{S}^{\bar{v}^3} (b \mathcal{S}^{\bar{v}^2} \mathcal{S}^{\bar{v}^1} f). \end{aligned}$$

As in the case $k = 2$, we expand in all of the parameters

$$[\mathcal{S}^{\bar{v}^1}, [\mathcal{S}^{\bar{v}^2}, [b, \mathcal{S}^{\bar{v}^3}]]]f = \sum_{\substack{\bar{i}_1 \in \{1,2,3\}^{n_1} \\ \bar{i}_2 \in \{1,2,3\}^{n_2} \\ \bar{i}_3 \in \{1,2,3\}^{n_3}}} \sum_{j=1}^8 \sigma_{j, \bar{i}_1, \bar{i}_2, \bar{i}_3},$$

where, for example,

$$\sigma_{1, \bar{i}_1, \bar{i}_2, \bar{i}_3} = \mathcal{S}^{\bar{v}^1} \mathcal{S}^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, \mathcal{S}^{\bar{v}^3} f)).$$

Again, Lemma 3.1 combined with boundedness of the shifts yields that each individual term of

$$\sum_{\substack{\bar{i}_1 \in \{1,2,3\}^{n_1} \setminus \{3\}^{n_1} \\ \bar{i}_2 \in \{1,2,3\}^{n_2} \setminus \{3\}^{n_2} \\ \bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}}} \sum_{j=1}^8 \sigma_{j, \bar{i}_1, \bar{i}_2, \bar{i}_3}$$

is directly bounded. Hence, we need to consider terms in the sum

$$\begin{aligned}
 (4.9) \quad & \sum_{\substack{\bar{i}_1=\{3\}^{n_1} \\ \bar{i}_2 \in \{1,2,3\}^{n_2} \setminus \{3\}^{n_2} \\ \bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}}} + \sum_{\substack{\bar{i}_1 \in \{1,2,3\}^{n_1} \setminus \{3\}^{n_1} \\ \bar{i}_2=\{3\}^{n_2} \\ \bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}}} + \sum_{\substack{\bar{i}_1 \in \{1,2,3\}^{n_1} \setminus \{3\}^{n_1} \\ \bar{i}_2 \in \{1,2,3\}^{n_2} \setminus \{3\}^{n_2} \\ \bar{i}_3=\{3\}^{n_3}}} \\
 & + \sum_{\substack{\bar{i}_1=\{3\}^{n_1} \\ \bar{i}_2=\{3\}^{n_2} \\ \bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}}} + \sum_{\substack{\bar{i}_1=\{3\}^{n_1} \\ \bar{i}_2 \in \{1,2,3\}^{n_2} \setminus \{3\}^{n_2} \\ \bar{i}_3=\{3\}^{n_3}}} \\
 & + \sum_{\substack{\bar{i}_1 \in \{1,2,3\}^{n_1} \setminus \{3\}^{n_1} \\ \bar{i}_2=\{3\}^{n_2} \\ \bar{i}_3=\{3\}^{n_3}}} + \sum_{\substack{\bar{i}_1=\{3\}^{n_1} \\ \bar{i}_2=\{3\}^{n_2} \\ \bar{i}_3=\{3\}^{n_3}}}.
 \end{aligned}$$

In the first three sums we can reduce to cases of one shift by pairing two terms. For example, let $\bar{i}_1 = \{3\}^{n_1}$, $\bar{i}_2 \in \{1,2,3\}^{n_2} \setminus \{3\}^{n_2}$, and $\bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}$. Then the first pair of terms is

$$S^{\bar{v}^1} S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} f)) - S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} S^{\bar{v}^1} f)).$$

Here it is enough to study

$$S^{\bar{v}^1} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, f)) - (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^1} f))$$

since $S^{\bar{v}^2}$ and $S^{\bar{v}^3}$ are bounded and order of the shifts is interchangeable.

Then the following three sums reduces to cases of two shifts by summing four terms. For example, let $\bar{i}_1 = \{3\}^{n_1}$, $\bar{i}_2 = \{3\}^{n_2}$, and $\bar{i}_3 \in \{1,2,3\}^{n_3} \setminus \{3\}^{n_3}$. Then the first sum of four terms is

$$\begin{aligned}
 & S^{\bar{v}^1} S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} f)) - S^{\bar{v}^1} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} S^{\bar{v}^2} f)) \\
 & - S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} S^{\bar{v}^1} f)) + A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^3} S^{\bar{v}^2} S^{\bar{v}^1} f).
 \end{aligned}$$

By similar arguments as before, it is enough to consider

$$\begin{aligned}
 & S^{\bar{v}^1} S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, f)) - S^{\bar{v}^1} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^2} f)) \\
 & - S^{\bar{v}^2} (A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^1} f)) + A_{\bar{i}_1}^{\bar{v}^1} A_{\bar{i}_2}^{\bar{v}^2} A_{\bar{i}_3}^{\bar{v}^3} (b, S^{\bar{v}^2} S^{\bar{v}^1} f).
 \end{aligned}$$

Notice that this type of terms is similar to terms in the case $k = 2$. In the latter example there is an additional legal paraprodut in the parameters \mathcal{I}_3 compared to the last term in the previous proof. However, we have already taken this into account in the general term (4.2) and the boundedness follows by a similar expansion as in the case $k = 2$.

In the last term in (4.9) we need to expand

$$\begin{aligned}
 (4.10) \quad & \langle b \rangle_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times J_{\bar{v}^3}} - \langle b \rangle_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3}} - \langle b \rangle_{I_{\bar{v}^1} \times J_{\bar{v}^2} \times J_{\bar{v}^3}} + \langle b \rangle_{I_{\bar{v}^1} \times J_{\bar{v}^2} \times I_{\bar{v}^3}} \\
 & - \langle b \rangle_{J_{\bar{v}^1} \times I_{\bar{v}^2} \times J_{\bar{v}^3}} + \langle b \rangle_{J_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3}} + \langle b \rangle_{J_{\bar{v}^1} \times J_{\bar{v}^2} \times J_{\bar{v}^3}} - \langle b \rangle_{J_{\bar{v}^1} \times J_{\bar{v}^2} \times I_{\bar{v}^3}}.
 \end{aligned}$$

Now, we apply Lemma 4.1 three times. First, we apply the lemma in parameters \mathcal{I}_1 – e.g. the sum of the first and the fifth terms equals to

$$\sum_{j_1=1}^{n_1} \left(\sum_{s_1=1}^{k_{v_{j_1}^1}} \langle \Delta_{I_{v_{j_1}^1}} b \rangle_{I_{v_{j_1}^1} \times Q_{\bar{v}_{j_1}^1} \times I_{\bar{v}^2} \times J_{\bar{v}^3}} - \sum_{t_1=1}^{l_{v_{j_1}^1}} \langle \Delta_{J_{v_{j_1}^1}} b \rangle_{J_{v_{j_1}^1} \times Q_{\bar{v}_{j_1}^1} \times I_{\bar{v}^2} \times J_{\bar{v}^3}} \right),$$

where $Q_{\bar{v}_{j_1}^1} = I_{(v_u^1)_{1 \leq u < j_1}} \times J_{(v_u^1)_{j_1 < u \leq n_1}}$. Then we switch pairs such that we can apply the lemma in parameters \mathcal{I}_2 . For example, the sum of related terms of the first and the third terms in (4.10) equals to

$$\begin{aligned} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \sum_{s_1=1}^{k_{v_{j_1}^1}} \sum_{s_2=1}^{k_{v_{j_2}^2}} & \left(\sum_{v_{j_1}^1} \langle \Delta_{I_{v_{j_1}^1}} \Delta_{I_{v_{j_2}^2}} b \rangle_{I_{v_{j_1}^1} \times I_{v_{j_2}^2} \times Q_{\bar{v}_{j_1}^1} \times Q_{\bar{v}_{j_2}^2} \times J_{\bar{v}^3}} \right. \\ & \left. - \sum_{t_2=1}^{l_{v_{j_2}^2}} \langle \Delta_{I_{v_{j_1}^1}} \Delta_{J_{v_{j_2}^2}} b \rangle_{I_{v_{j_1}^1} \times J_{v_{j_2}^2} \times Q_{\bar{v}_{j_1}^1} \times Q_{\bar{v}_{j_2}^2} \times J_{\bar{v}^3}} \right). \end{aligned}$$

Finally, we pair terms such that we apply Lemma 4.1 in parameters \mathcal{I}_3 . Hence, for example, the sum of the related terms of the first and the second terms in (4.10) equals to

$$\begin{aligned} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \sum_{j_3=1}^{n_3} \sum_{s_1=1}^{k_{v_{j_1}^1}} \sum_{s_2=1}^{k_{v_{j_2}^2}} & \sum_{t_3=1}^{l_{v_{j_3}^3}} \left(\sum_{v_{j_1}^1} \langle \Delta_{I_{v_{j_1}^1}} \Delta_{I_{v_{j_2}^2}} \Delta_{J_{v_{j_3}^3}} b \rangle_{I_{v_{j_1}^1} \times I_{v_{j_2}^2} \times J_{v_{j_3}^3} \times Q_{\bar{v}_{j_1}^1} \times Q_{\bar{v}_{j_2}^2} \times Q_{\bar{v}_{j_3}^3}} \right. \\ & \left. - \sum_{s_3=1}^{k_{v_{j_3}^3}} \langle \Delta_{I_{v_{j_1}^1}} \Delta_{I_{v_{j_2}^2}} \Delta_{I_{v_{j_3}^3}} b \rangle_{I_{v_{j_1}^1} \times I_{v_{j_2}^2} \times I_{v_{j_3}^3} \times Q_{\bar{v}_{j_1}^1} \times Q_{\bar{v}_{j_2}^2} \times Q_{\bar{v}_{j_3}^3}} \right). \end{aligned} \quad (4.11)$$

Now, each appearing term is fully expanded. For example, for fixed s_1, s_2, t_3 and $j_1 = 1, j_2 = 1, j_3 = 1$, the term to be estimated related to the first term in (4.11) equals

$$\begin{aligned} \sum_{K_{\bar{v}^1}, K_{\bar{v}^2}, K_{\bar{v}^3}} \sum_{I_{\bar{v}^1}^{(k_{\bar{v}^1})} = K_{\bar{v}^1}} \sum_{I_{\bar{v}^2}^{(k_{\bar{v}^2})} = K_{\bar{v}^2}} \sum_{I_{\bar{v}^3}^{(k_{\bar{v}^3})} = K_{\bar{v}^3}} & a_{K_{\bar{v}^1}, I_{\bar{v}^1}, J_{\bar{v}^1}} a_{K_{\bar{v}^2}, I_{\bar{v}^2}, J_{\bar{v}^2}} a_{K_{\bar{v}^3}, I_{\bar{v}^3}, J_{\bar{v}^3}} \\ & \times \langle h_{I_{v_1^1}} \rangle_{I_{v_1^1}} \langle h_{I_{v_2^2}} \rangle_{I_{v_2^2}} \langle h_{J_{v_3^3}} \rangle_{J_{v_3^3}} \\ & \times \left\langle b, h_{I_{v_1^1} \times I_{v_2^2} \times J_{v_3^3}} \otimes \frac{1_{J_{\bar{v}_1^1} \times J_{\bar{v}_2^2} \times J_{\bar{v}_3^3}}}{|J_{\bar{v}_1^1} \times J_{\bar{v}_2^2} \times J_{\bar{v}_3^3}|} \right\rangle \\ & \times \langle f, h_{I_{\bar{v}^1} \times I_{\bar{v}^2} \times I_{\bar{v}^3}} \rangle h_{J_{\bar{v}^1} \times J_{\bar{v}^2} \times J_{\bar{v}^3}}, \end{aligned}$$

where $\bar{v}_1^{j'} = (v_u^j)_{u=2}^{n_j}$ for $j = 1, 2, 3$. It is easy to see that these terms have the form of the general term. Hence, by Lemma 4.2 these terms are bounded assuming $b \in \text{bmo}^{\mathcal{I}}$. \square

By the representation theorem of the multi-parameter singular integrals [39] we get the following result:

Corollary 4.5. *Let $k \leq m$ be an integer and let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_m})$ and $1 < p < \infty$. For all partitions $\mathcal{I} = \{\mathcal{I}_i : i \leq k\}$ of $\{1, 2, \dots, m\}$ we have*

$$\| [T^{\bar{v}^1}, [T^{\bar{v}^2}, \dots [b, T^{\bar{v}^k}]]] \|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{bmo}^{\mathcal{I}}(\nu)},$$

where $\bar{v}^i = (v_j^i)_{v_j^i \in \mathcal{I}_i}$ and $T^{\bar{v}^i}$ are paraproduct free multi-parameter Calderón-Zygmund operators.

5. Commutators involving paraproducts

In this section we consider the space $\mathbb{R}^d = \mathbb{R}^{d_1+d_2+d_3+d_4}$ and operators which are defined in some fixed grids \mathcal{D}^{d_i} .

Next, we define the other two bi-parameter dyadic model operators: partial and full paraproducts.

Partial paraproduct. Let $k_1, l_1 \geq 0$. We define

$$P^{1,2}g = P_{\mathcal{D}^{d_1}, \mathcal{D}^{d_2}}^{1,2,(k_1,l_1)}g = \sum_{\substack{K_1 \in \mathcal{D}^{d_1} \\ K_2 \in \mathcal{D}^{d_2}}} \sum_{\substack{I_1, J_1 \in \mathcal{D}^{d_1} \\ I_1^{(k_1)} = K_1 \\ J_1^{(l_1)} = K_1}} \langle a_{K_1, I_1, J_1}, h_{K_2} \rangle h_{J_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes \langle g, h_{I_1} \otimes h_{K_2} \rangle_{1,2}.$$

Here only finitely many of the coefficients $\langle a_{K_1, I_1, J_1}, h_{K_2} \rangle$ are non-zero, and

$$\|a_{K_1, I_1, J_1}\|_{\text{BMO}(\mathcal{D}^{d_2})} \leq \frac{|I_1|^{\frac{1}{2}} |J_1|^{\frac{1}{2}}}{|K_1|}.$$

Also we have partial paraproducts of the form

$$P^{1,2}g = P_{\mathcal{D}^{d_1}, \mathcal{D}^{d_2}}^{1,2,(k_2,l_2)}g = \sum_{\substack{K_1 \in \mathcal{D}^{d_1} \\ K_2 \in \mathcal{D}^{d_2}}} \sum_{\substack{I_2, J_2 \in \mathcal{D}^{d_2} \\ I_2^{(k_2)} = K_2 \\ J_2^{(l_2)} = K_2}} \langle a_{K_2, I_2, J_2}, h_{K_1} \rangle \frac{1_{K_1}}{|K_1|} \otimes h_{J_2} \otimes \langle g, h_{K_1} \otimes h_{I_2} \rangle_{1,2},$$

where $k_2, l_2 \geq 0$ and

$$\|a_{K_2, I_2, J_2}\|_{\text{BMO}(\mathcal{D}^{d_1})} \leq \frac{|I_2|^{\frac{1}{2}} |J_2|^{\frac{1}{2}}}{|K_2|}.$$

Full paraproduct. We define

$$\Pi^{1,2}f = \Pi_{\mathcal{D}^{d_1}, \mathcal{D}^{d_2}}^{1,2}f = \sum_{\substack{K_1 \in \mathcal{D}^{d_1} \\ K_2 \in \mathcal{D}^{d_2}}} \langle a, h_{K_1 \times K_2} \rangle \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{K_2}}{|K_2|} \otimes \langle f, h_{K_1} \otimes h_{K_2} \rangle_{1,2}.$$

Here only finitely many of the coefficients $\langle a, h_{K_1 \times K_2} \rangle$ are non-zero, and

$$\|a\|_{\text{BMO}(\mathcal{D}^{d_1} \times \mathcal{D}^{d_2})} \leq 1.$$

Also Π^* , Π^{1*} , Π^{2*} are full paraproducts, where e.g.

$$\Pi^{1*}f = \sum_{\substack{K_1 \in \mathcal{D}^{d_1} \\ K_2 \in \mathcal{D}^{d_2}}} \langle a, h_{K_1 \times K_2} \rangle h_{K_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes \left\langle f, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \right\rangle_{1,2}$$

is the partial adjoint in the first parameter of the above $\Pi^{1,2}$. Later on, we abbreviate $\langle a, h_{K_1 \times K_2} \rangle$ by a_{K_1, K_2} .

Theorem 5.1. *Let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_4})$ and $1 < p < \infty$. We have*

$$\| [T^{1,2}, [b, T^{3,4}]] \|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{bmo}^{\{\{1,2\}\{3,4\}\}}(\nu)},$$

where $T^{1,2}$ and $T^{3,4}$ are bi-parameter Calderón–Zygmund operators in $\mathbb{R}^{d_1+d_2}$ and $\mathbb{R}^{d_3+d_4}$, respectively.

Proof: By the representation theorem [37], we are considering the following collection of commutators:

$[\mathcal{S}^{1,2}, [b, \mathcal{S}^{3,4}]]$	$[\mathcal{S}^{1,2}, [b, P^{3,4}]]$	$[\mathcal{S}^{1,2}, [b, \Pi^{3,4}]]$
$[P^{1,2}, [b, \mathcal{S}^{3,4}]]$	$[P^{1,2}, [b, P^{3,4}]]$	$[P^{1,2}, [b, \Pi^{3,4}]]$
$[\Pi^{1,2}, [b, \mathcal{S}^{3,4}]]$	$[\Pi^{1,2}, [b, P^{3,4}]]$	$[\Pi^{1,2}, [b, \Pi^{3,4}]]$

By definition, for all model operators we have

$$\begin{aligned} [U^{1,2}, [b, V^{3,4}]]f &= U^{1,2}bV^{3,4}f - U^{1,2}V^{3,4}bf - bV^{3,4}U^{1,2}f + V^{3,4}bU^{1,2}f \\ &=: I - II - III + IV. \end{aligned}$$

Now, the forms of U and V determine how we expand the terms. We expand the products in the parameters where a cancellative Haar function is paired with b . In the parameters where b is paired with a non-cancellative Haar function we do not expand at all.

As explained earlier, by Lemma 3.1 the terms where b is paired with the cancellative Haar functions on parameters 1 or 2, and 3 or 4, are directly bounded with the correct BMO condition. For the other terms, we need to pair them depending on the expansion.

We only demonstrate the general strategy with a case involving both full and partial paraproducts.

We are considering model operators of the following form

$$\begin{aligned}\Pi^{1,2}f &= \sum_{K_1, K_2} \langle a, h_{K_1 \times K_2} \rangle \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes \left\langle f, h_{K_1} \otimes \frac{1_{K_2}}{|K_2|} \right\rangle_{1,2}, \\ P^{3,4}g &= \sum_{K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle \langle g, h_{I_3} \otimes h_{K_4} \rangle_{3,4} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|}.\end{aligned}$$

Here only finitely many of the coefficients $\langle a, h_{K_1 \times K_2} \rangle$ and $\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle$ are non-zero, and these coefficients have the following bounds

$$\begin{aligned}\|a\|_{\text{BMO}(\mathcal{D}^{d_1} \times \mathcal{D}^{d_2})} &\leq 1, \\ \|a_{K_3, I_3, J_3}\|_{\text{BMO}(\mathcal{D}^{d_4})} &\leq \frac{|I_3|^{\frac{1}{2}} |J_3|^{\frac{1}{2}}}{|K_3|}.\end{aligned}$$

As explained earlier, we expand the appearing terms in the following way:

$$\begin{aligned}I &= \sum_{i_1, i_3=1}^3 \Pi^{1,2} A_{i_1}^1 A_{i_3}^3 (b, P^{3,4}f), \quad II = \sum_{i_1, i_3, i_4=1}^3 \Pi^{1,2} P^{3,4} A_{i_1}^1 A_{i_3, i_4}^{3,4} (b, f), \\ III &= \sum_{i_2, i_3=1}^3 A_{i_2}^2 A_{i_3}^3 (b, \Pi^{1,2} P^{3,4}f), \quad IV = \sum_{i_2, i_3, i_4=1}^3 P^{3,4} A_{i_2}^2 A_{i_3, i_4}^{3,4} (b, \Pi^{1,2}f).\end{aligned}$$

Boundedness of the model operators combined with Lemma 3.1 implies that each term is directly bounded whenever we do not have the “illegal” paraproducts in the parameters 1 or 2 and 3 or 4. For the rest of the terms, we group them as follows

$$\begin{aligned}(5.1) \quad & \sum_{i_3=1}^2 [\Pi^{1,2} A_3^1 A_{i_3}^3 (b, P^{3,4}f) - A_3^2 A_{i_3}^3 (b, \Pi^{1,2} P^{3,4}f)] \\ & + \sum_{i_1=1}^2 [\Pi^{1,2} A_{i_1}^1 A_3^3 (b, P^{3,4}f) - \Pi^{1,2} P^{3,4} A_{i_1}^1 A_{3,3}^{3,4} (b, f)] \\ & + \sum_{(i_3, i_4) \in \{1, 2, 3\}^2 \setminus \{3\}^2} [P^{3,4} A_3^2 A_{i_3, i_4}^{3,4} (b, \Pi^{1,2}f) - \Pi^{1,2} P^{3,4} A_3^1 A_{i_3, i_4}^{3,4} (b, f)]\end{aligned}$$

$$\begin{aligned}
& + \sum_{i_2=1}^2 [P^{3,4} A_{i_2}^2 A_{3,3}^{3,4}(b, \Pi^{1,2} f) - A_{i_2}^2 A_3^3(b, \Pi^{1,2} P^{3,4} f)] \\
& + [\Pi^{1,2} A_3^1 A_3^3(b, P^{3,4} f) - \Pi^{1,2} P^{3,4} A_3^1 A_{3,3}^{3,4}(b, f) \\
& \quad - A_3^2 A_3^3(b, \Pi^{1,2} P^{3,4} f) + P^{3,4} A_3^2 A_{3,3}^{3,4}(b, \Pi^{1,2} f)].
\end{aligned}$$

We begin with the first pair. Since $P^{3,4}$ is bounded, it is enough to consider the boundedness of $\Pi^{1,2} A_3^1 A_{i_3}^3(b, f) - A_3^2 A_{i_3}^3(b, \Pi^{1,2} f)$. We show the case $i_3 = 1$, and the other case can be handled similarly.

First, notice that by the one-parameter expansion $\sum_{I_1 \subset K_1} \Delta_{I_1}^1 \varphi = (\varphi - \langle \varphi \rangle_{K_1,1}) 1_{K_1}$, we have

$$\begin{aligned}
& (\langle \varphi \rangle_{K_1,1} - \langle \varphi \rangle_{K_2,2}) 1_{K_1 \times K_2} \\
& = (\langle \varphi \rangle_{K_1,1} - \langle \varphi \rangle_{K_1 \times K_2,1,2} + \langle \varphi \rangle_{K_1 \times K_2,1,2} - \langle \varphi \rangle_{K_2,2}) 1_{K_1 \times K_2} \\
& = \sum_{I_2 \subset K_2} \left\langle \varphi, \frac{1_{K_1}}{|K_1|} \otimes h_{I_2} \right\rangle_{1,2} \otimes h_{I_2} - \sum_{I_1 \subset K_1} \left\langle \varphi, h_{I_1} \otimes \frac{1_{K_2}}{|K_2|} \right\rangle_{1,2} \otimes h_{I_1}.
\end{aligned}$$

Using the previous observation, we have

$$\begin{aligned}
& \Pi^{1,2} A_3^1 A_1^3(b, f) - A_3^2 A_1^3(b, \Pi^{1,2} f) \\
& = \sum_{K_1, K_2, K_3} a_{K_1, K_2} \left[\langle \langle b_{K_3} \rangle_{K_1,1} \langle f, h_{K_1} \otimes h_{K_3} \rangle_{1,3} \rangle_{K_2,2} \right. \\
& \quad \left. - \langle b_{K_3} \rangle_{K_2,2} \left\langle f, h_{K_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes h_{K_3} \right\rangle_{1,2,3} \right] \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{K_3} h_{K_3} \\
& = \sum_{K_1, K_2, K_3} a_{K_1, K_2} \left\langle (\langle b_{K_3} \rangle_{K_1,1} - \langle b_{K_3} \rangle_{K_2,2}) 1_{K_1 \times K_2} \langle f, h_{K_1} \otimes h_{K_3} \rangle_{1,3}, \frac{1_{K_2}}{|K_2|} \right\rangle_2 \\
& \quad \times \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{K_3} h_{K_3} \\
& = \sum_{K_1, K_2, K_3} a_{K_1, K_2} \left[\sum_{Q_2 \subset K_2} |K_2|^{-1} \left\langle b, \frac{1_{K_1}}{|K_1|} \otimes h_{Q_2} \otimes h_{K_3} \right\rangle_{1,2,3} \right. \\
& \quad \times \langle f, h_{K_1} \otimes h_{Q_2} \otimes h_{K_3} \rangle_{1,2,3} \\
& \quad \left. - \sum_{Q_1 \subset K_1} \left\langle b, h_{Q_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes h_{K_3} \right\rangle_{1,2,3} \langle f, h_{K_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes h_{K_3} \rangle_{1,2,3} \otimes h_{Q_1} \right] \\
& \quad \times \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{K_3} h_{K_3},
\end{aligned}$$

where $b_{K_3} = \langle b, 1_{K_3}/|K_3| \rangle_3$. These terms are similar to handle and we deal with the first one. By the L^p duality, we have

$$\left| \int_{\mathbb{R}^{d_4}} \sum_{K_1, K_2, K_3} a_{K_1, K_2} |K_2|^{-1} \sum_{Q_2 \subset K_2} \left\langle b_{x_4}, \frac{1_{K_1}}{|K_1|} \otimes h_{Q_2} \otimes h_{K_3} \right\rangle_{1,2,3} \right. \\ \times \left\langle f_{x_4}, h_{K_1} \otimes h_{Q_2} \otimes h_{K_3} \right\rangle_{1,2,3} \\ \left. \times \left\langle g_{x_4}, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{K_3} h_{K_3} \right\rangle_{1,2,3} dx_4 \right|.$$

Fix $x_4 \in \mathbb{R}^{d_4}$. By Proposition 2.9 we need to show the boundedness of

$$\int_{\mathbb{R}^{d_1+d_2+d_3}} \sum_{K_1, K_2} |a_{K_1, K_2}| \frac{1_{K_1}}{|K_1|} \\ \times \left(\sum_{Q_2, K_3} \langle f_{x_4}, h_{K_1} \otimes h_{Q_2} \otimes h_{K_3} \rangle^2 \langle |g_{x_4}, h_{K_2}\rangle_2 \rangle_{K_1 \times K_3}^2 \frac{1_{Q_2}}{|Q_2|} \otimes \frac{1_{K_3}}{|K_3|} \right)^{\frac{1}{2}} \frac{1_{K_2}}{|K_2|} \nu_{x_4} \\ \leq \int_{\mathbb{R}^{d_1+d_2+d_3}} \sum_{K_1} S^{2,3} \langle f_{x_4}, h_{K_1} \rangle_1 \sum_{K_2} |a_{K_1, K_2}| \frac{1_{K_2}}{|K_2|} \otimes \langle M^3 \langle g_{x_4}, h_{K_2} \rangle_2 \rangle_{K_1, 1} \nu_{x_4}.$$

We begin by writing

$$G = \sum_{I_1, K_2} |a_{I_1, K_2}| h_{I_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes \langle M^3 \langle g_{x_4}, h_{K_2} \rangle_2 \rangle_{I_1, 1}.$$

Hence, by standard estimates we get

$$\int_{\mathbb{R}^{d_1+d_2+d_3}} \sum_{K_1} S^{2,3} \langle f_{x_4}, h_{K_1} \rangle_1 \langle G, h_{K_1} \rangle_1 \frac{1_{K_1}}{|K_1|} \nu_{x_4} \\ \leq \int_{\mathbb{R}^{d_1+d_2+d_3}} \left(\sum_{K_1} S^{2,3} \langle f_{x_4}, h_{K_1} \rangle_1^2 \otimes \frac{1_{K_1}}{|K_1|} \right)^{\frac{1}{2}} \left(\sum_{K_1} \langle G, h_{K_1} \rangle_1^2 \otimes \frac{1_{K_1}}{|K_1|} \right)^{\frac{1}{2}} \nu_{x_4} \\ \leq \|S^{1,2,3} f_{x_4}\|_{L^p(\mu_{x_4})} \|S^1 G\|_{L^{p'}(\lambda_{x_4}^{1-p'})} \\ \lesssim [\mu]_{A_p, [\lambda]_{A_p}} \|f\|_{L^p(\mu_{x_4})} \|g\|_{L^{p'}(\lambda_{x_4}^{1-p'})},$$

and applying Hölder's inequality once more to the integral on \mathbb{R}^{d_4} we have the desired bound.

Next, we deal with the second term in (5.1) with $i_1 = 2$. More precisely, it is enough to consider the term

$$\begin{aligned}
 & \sum_{K_1, K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle h_{K_1} \otimes [\langle b_{K_1} \rangle_{J_3, 3} - \langle b_{K_1} \rangle_{I_3 \times K_4, 3, 4}] \\
 & \quad \times \left\langle f, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle_{1, 3, 4} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|} \\
 = & \sum_{K_1, K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle h_{K_1} \\
 & \quad \otimes \left[\sum_{Q_4 \subset K_4} \Delta_{Q_4}^4 \langle b_{K_1} \rangle_{J_3, 3} + \sum_{t_3=1}^{l_3} \langle \Delta_{J_3^{(t_3)}}^3 b_{K_1} \rangle_{J_3 \times K_4, 3, 4} \right. \\
 & \quad \left. - \sum_{s_3=1}^{k_3} \langle \Delta_{I_3^{(s_3)}}^3 b_{K_1} \rangle_{I_3 \times K_4, 3, 4} \right] \\
 & \quad \times \left\langle f, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle_{1, 3, 4} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|} =: \sigma_1 + \sigma_2 - \sigma_3,
 \end{aligned}$$

where $b_{K_1} := \langle b, h_{K_1} \rangle_1$. Terms σ_2 and σ_3 are handled similarly. Therefore, we show the boundedness of σ_1 and σ_2 . We begin with the dual form of the first one

$$\begin{aligned}
 (5.2) \quad & \left| \int_{\mathbb{R}^{d_2}} \sum_{K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle |K_4|^{-1} \sum_{\substack{K_1 \\ Q_4 \subset K_4}} \langle \langle b, h_{K_1} \otimes h_{Q_4} \rangle_{1, 4} \rangle_{J_3, 3} \right. \\
 & \quad \left. \times \left\langle f, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle_{1, 3, 4} \langle g, h_{K_1} \otimes h_{J_3} \otimes h_{Q_4} \rangle_{1, 3, 4} \right|.
 \end{aligned}$$

Fix $x_2 \in \mathbb{R}^{d_2}$. By Proposition 2.9 it is enough to estimate

$$\begin{aligned}
 & \int_{\mathbb{R}^{d_1+d_3+d_4}} \sum_{K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| \left(\sum_{K_1, Q_4} \left\langle f_{x_2}, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle^2 \right. \\
 & \quad \left. \times \langle g_{x_2}, h_{K_1} \otimes h_{J_3} \otimes h_{Q_4} \rangle^2 \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{Q_4}}{|Q_4|} \right)^{\frac{1}{2}} \frac{1_{J_3}}{|J_3|} \otimes \frac{1_{K_4}}{|K_4|} \nu_{x_2}.
 \end{aligned}$$

First, we estimate

$$\begin{aligned} \left| \left\langle f_{x_2}, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle \right| &\leq \langle M^1 \Delta_{K_3 \times K_4, (k_3, 0)} f_{x_2} \rangle_{I_3 \times K_4, 3, 4} |I_3|^{\frac{1}{2}} |K_4|^{\frac{1}{2}} \\ &=: \langle F_{x_2, K_3, K_4} \rangle_{I_3 \times K_4, 3, 4} |I_3|^{\frac{1}{2}} |K_4|^{\frac{1}{2}}. \end{aligned}$$

Hence, we can estimate

$$\begin{aligned} &\left(\sum_{K_1, Q_4} \langle g_{x_2}, h_{K_1} \otimes h_{J_3} \otimes h_{Q_4} \rangle^2 \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{Q_4}}{|Q_4|} \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{K_1, Q_4} \langle M^3 \Delta_{K_1 \times K_3 \times Q_4, (0, l_3, 0)} g_{x_2} \rangle_{K_1 \times Q_4}^2 1_{K_1} \otimes 1_{Q_4} \right)^{\frac{1}{2}} |J_3|^{\frac{1}{2}} \\ &=: G_{x_2, K_3} |J_3|^{\frac{1}{2}}. \end{aligned}$$

Using these estimates we get

$$\begin{aligned} &\int_{\mathbb{R}^{d_1+d_3+d_4}} \sum_{K_3, K_4} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| \left(\sum_{K_1, Q_4} \left\langle f_{x_2}, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle^2 \right. \\ &\quad \times \langle g_{x_2}, h_{K_1} \otimes h_{J_3} \otimes h_{Q_4} \rangle^2 \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{Q_4}}{|Q_4|} \Big)^{\frac{1}{2}} \frac{1_{J_3}}{|J_3|} \otimes \frac{1_{K_4}}{|K_4|} \nu_{x_2} \\ &\leq \int_{\mathbb{R}^{d_1+d_3}} \sum_{K_3} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} |I_3|^{\frac{1}{2}} \frac{1_{J_3}}{|J_3|^{\frac{1}{2}}} \\ &\quad \times \sum_{K_4} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| \langle F_{x_2, K_3, K_4} \rangle_{I_3 \times K_4, 3, 4} \langle G_{x_2, K_3} \nu_{x_2} \rangle_{K_4} |K_4|^{\frac{1}{2}} \\ &\leq \int_{\mathbb{R}^{d_1+d_3+d_4}} \sum_{K_3} |K_3|^{-1} M^4 (G_{x_2, K_3} \nu_{x_2}) \\ &\quad \times \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \left(\sum_{K_4} |I_3|^2 \langle F_{x_2, K_3, K_4} \rangle_{I_3 \times K_4, 3, 4}^2 \otimes 1_{K_4} \right)^{\frac{1}{2}} \otimes 1_{J_3} \\ &\stackrel{(*)}{\lesssim} \left\| \left(\sum_{K_3} \left\langle \left(\sum_{K_4} \langle F_{x_2, K_3, K_4} \rangle_{K_4, 4}^2 \otimes 1_{K_4} \right)^{\frac{1}{2}} \right\rangle_{K_3}^2 \otimes 1_{K_3} \right)^{\frac{1}{2}} \right\|_{L^p(\lambda_{x_2})} \\ &\quad \times \left\| \left(\sum_{K_3} M^4 (G_{x_2, K_3} \nu_{x_2})^2 \right)^{\frac{1}{2}} \right\|_{L^{p'}(\lambda_{x_2}^{1-p'})} \\ &\lesssim [\mu]_{A_p} [\lambda]_{A_p} \|f_{x_2}\|_{L^p(\mu_{x_2})} \|g_{x_2}\|_{L^{p'}(\lambda_{x_2}^{1-p'})}, \end{aligned}$$

where in step (*), along with obvious Hölder's inequalities, we used the following application of Kahane–Khintchine's inequality

$$\begin{aligned}
 |K|^{-1} \sum_{I^{(k)}=K} \left(\sum_i |I|^2 \langle \varphi_i \rangle_I^2 \right)^{\frac{1}{2}} &\sim |K|^{-1} \sum_{I^{(k)}=K} \mathbb{E} \left| \sum_i \epsilon_i |I| \langle \varphi_i \rangle_I \right| \\
 &= |K|^{-1} \sum_{I^{(k)}=K} \mathbb{E} \left| \sum_i \epsilon_i \int_I \varphi_i(x) dx \right| \\
 &\leq |K|^{-1} \sum_{I^{(k)}=K} \mathbb{E} \int_I \left| \sum_i \epsilon_i \varphi_i(x) \right| dx \\
 &\sim \left\langle \left(\sum_i |\varphi_i|^2 \right)^{\frac{1}{2}} \right\rangle_K.
 \end{aligned}$$

After applying Hölder's inequality to the integral on \mathbb{R}^{d_2} , we get the desired bound for σ_1 .

Then take σ_2 with fixed $t_3 \in [1, l_3]$. By duality and Proposition 2.9, the term that we are estimating equals to

$$\begin{aligned}
 &\int_{\mathbb{R}^d} \sum_{K_3} \left(\sum_{\substack{K_1 \\ P_3^{(t_3-t_3)}=K_3}} \right. \\
 &\quad \times \left[\sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(t_3)}=P_3}} \sum_{K_4} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| |P_3|^{-\frac{1}{2}} \left| \left\langle f, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle_{1,3,4} \right| \right. \\
 &\quad \left. \times \left| \left\langle g, h_{K_1} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|} \right\rangle_{1,3,4} \right| \otimes \frac{1_{K_4}}{|K_4|} \right]^2 \otimes \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{P_3}}{|P_3|} \Bigg)^{\frac{1}{2}} \nu.
 \end{aligned}$$

Begin by fixing $x_2 \in \mathbb{R}^{d_2}$. Then by sparse domination of bilinear paraproducts (see e.g. Lemma 6.7 in [33]) we can deduce

$$\begin{aligned}
 &\int_{\mathbb{R}^{d_4}} \sum_{K_4} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| \left| \left\langle f_{x_2}, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \otimes h_{K_4} \right\rangle \right| \\
 &\quad \times \left| \left\langle g_{x_2}, h_{K_1} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|} \right\rangle \right| \frac{1_{K_4}}{|K_4|} \nu_{x_2} \\
 &\lesssim_{[\nu]_{A_\infty}} \frac{|I_3|^{\frac{1}{2}} |J_3|^{\frac{1}{2}}}{|K_3|} \int_{\mathbb{R}^{d_4}} M^4 \left\langle f_{x_2}, \frac{1_{K_1}}{|K_1|} \otimes h_{I_3} \right\rangle_{1,3} M^4 \langle g_{x_2}, h_{K_1} \otimes h_{J_3} \rangle_{1,3} \nu_{x_2} \\
 &\leq \frac{|I_3| |J_3| |K_1|^{\frac{1}{2}}}{|K_3|} \int_{\mathbb{R}^{d_4}} \langle M^4 \Delta_{K_3, k_3}^3 f_{x_2} \rangle_{K_1 \times I_3, 1,3} \langle M^4 \Delta_{K_1, K_3, (0, l_3)}^{1,3} g_{x_2} \rangle_{K_1 \times J_3, 1,3} \nu_{x_2}.
 \end{aligned}$$

Using the previous estimates and A_∞ -extrapolation we get that our term is bounded by

$$\begin{aligned}
& \sum_{K_3} \int_{\mathbb{R}^{d_1+d_3+d_4}} \left(\sum_{\substack{K_1 \\ P_3^{(l_3-t_3)}=K_3}} \langle M^4 \Delta_{K_3, k_3}^3 f_{x_2} \rangle_{K_1 \times K_3, 1, 3}^2 \right. \\
& \quad \left. \times \langle M^4 \Delta_{K_1, K_3, (0, l_3)}^{1, 3} g_{x_2} \rangle_{K_1 \times P_3, 1, 3}^2 \otimes 1_{K_1} \otimes 1_{P_3} \right)^{\frac{1}{2}} \nu_{x_2} \\
& \leq \int_{\mathbb{R}^{d_1+d_3+d_4}} \left(\sum_{K_3} \langle M^{1, 4} \Delta_{K_3, k_3}^3 f_{x_2} \rangle_{K_3, 3}^2 \otimes 1_{K_3} \right)^{\frac{1}{2}} \\
& \quad \times \left(\sum_{K_1, K_3} \langle M^{3, 4} \Delta_{K_1, K_3, (0, l_3)}^{1, 3} g_{x_2} \rangle_{K_1, 1, 3}^2 \otimes 1_{K_1} \right)^{\frac{1}{2}} \nu_{x_2} \\
& \lesssim [\mu]_{A_p, [\lambda]_{A_p}} \|f\|_{L^p(\mu_{x_2})} \|g\|_{L^{p'}(\lambda_{x_2}^{1-p'})}.
\end{aligned}$$

Apply Hölder's inequality to the integral on \mathbb{R}^{d_2} to get the desired boundness.

The third and fourth terms in (5.1) are similar. Thus, we only need to take care of the last term

$$\sum_{\substack{K_1, K_2 \\ K_3, K_4}} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} a_{K_1, K_2} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle A_{K_1, K_2, I_3, J_3, K_4}^{f, b} \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|},$$

where

$$A_{K_1, K_2, I_3, J_3, K_4}^{f, b} = \langle B \langle f, h_{K_1} \otimes h_{I_3} \otimes h_{K_4} \rangle_{1, 3, 4} \rangle_{K_2, 2}$$

and

$$B = \langle b \rangle_{K_1 \times J_3, 1, 3} - \langle b \rangle_{K_1 \times I_3 \times K_4, 1, 3, 4} - \langle b \rangle_{K_2 \times J_3, 2, 3} + \langle b \rangle_{K_2 \times I_3 \times K_4, 2, 3, 4}.$$

We write

$$\begin{aligned}
B &= \sum_{\substack{Q_2 \subset K_2 \\ Q_4 \subset K_4}} \Delta_{Q_2 \times Q_4} \langle b \rangle_{K_1 \times J_3, 1, 3} + \sum_{t_3=1}^{l_3} \sum_{Q_2 \subset K_2} \langle \Delta_{J_3^{(t_3)} \times Q_2} \langle b \rangle_{K_1 \times K_4, 1, 4} \rangle_{J_3, 3} \\
&\quad - \sum_{s_3=1}^{k_3} \sum_{Q_2 \subset K_2} \langle \Delta_{I_3^{(s_3)} \times Q_2} \langle b \rangle_{K_1 \times K_4, 1, 4} \rangle_{I_3, 3} - \sum_{\substack{Q_1 \subset K_1 \\ Q_4 \subset K_4}} \Delta_{Q_1 \times Q_4} \langle b \rangle_{K_2 \times J_3, 2, 3}
\end{aligned}$$

$$\begin{aligned}
& - \sum_{t_3=1}^{l_3} \sum_{Q_1 \subset K_1} \langle \Delta_{J_3^{(t_3)} \times Q_1} \langle b \rangle_{K_2 \times K_4, 2, 4} \rangle_{J_3, 3} \\
& + \sum_{s_3=1}^{k_3} \sum_{Q_1 \subset K_1} \langle \Delta_{I_3^{(s_3)} \times Q_1} \langle b \rangle_{K_2 \times K_4, 2, 4} \rangle_{I_3, 3} \\
& =: \sum_{i=1}^6 B_i.
\end{aligned}$$

We consider the terms associated to B_1 and B_2 since the rest can be estimated similarly.

The dual form of the term associated to B_1 equals to

$$\begin{aligned}
& \left| \sum_{\substack{K_1, K_2 \\ K_3, K_4}} \sum_{\substack{I_3^{(k_3)} = K_3 \\ J_3^{(l_3)} = K_3}} a_{K_1, K_2} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle |K_2|^{-1} |K_4|^{-1} \right. \\
& \quad \times \sum_{\substack{Q_2 \subset K_2 \\ Q_4 \subset K_4}} \left\langle b, \frac{1_{K_1}}{|K_1|} \otimes h_{Q_2} \otimes \frac{1_{J_3}}{|J_3|} \otimes h_{Q_4} \right\rangle \langle f, h_{K_1} \otimes h_{Q_2} \otimes h_{I_3} \otimes h_{K_4} \rangle \\
& \quad \left. \times \left\langle g, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{J_3} \otimes h_{Q_4} \right\rangle \right|.
\end{aligned}$$

By Proposition 2.9 it is enough to estimate the following term

$$\begin{aligned}
& \int_{\mathbb{R}^d} \left(\sum_{\substack{K_1, K_2 \\ K_3, K_4}} \sum_{\substack{I_3^{(k_3)} = K_3 \\ J_3^{(l_3)} = K_3}} |a_{K_1, K_2}| |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{K_2}}{|K_2|} S^2 \right. \\
& \quad \left. \times \langle f, h_{K_1} \otimes h_{I_3} \otimes h_{K_4} \rangle_{1, 3, 4} \otimes \frac{1_{J_3}}{|J_3|} \otimes \frac{1_{K_4}}{|K_4|} S^4 \left\langle g, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{J_3} \right\rangle_{1, 2, 3} \right) \nu.
\end{aligned}$$

First, we write

$$G := \sum_{Q_1, K_2, Q_3} |a_{Q_1, K_2}| h_{Q_1} \otimes \frac{1_{K_2}}{|K_2|} \otimes h_{Q_3} \otimes S^4 \left\langle g, \frac{1_{Q_1}}{|Q_1|} \otimes h_{K_2} \otimes h_{Q_3} \right\rangle_{1, 2, 3}.$$

Thus we get

$$\begin{aligned}
& \int_{\mathbb{R}^{d_1+d_2+d_3}} \sum_{K_1, K_3} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{J_3}}{|J_3|} \\
& \quad \times \sum_{K_4} |\langle a_{K_3, I_3, J_3}, h_{K_4} \rangle| S^2 \langle f, h_{K_1} \otimes h_{I_3} \otimes h_{K_4} \rangle_{1,3,4} \langle \langle G, h_{K_1} \otimes h_{J_3} \rangle_{1,3} \nu \rangle_{K_4,4} \\
& \lesssim \int_{\mathbb{R}^d} \sum_{K_1, K_3} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{J_3}}{|J_3|} \frac{|I_3|^{\frac{1}{2}} |J_3|^{\frac{1}{2}}}{|K_3|} \\
& \quad \times \left(\sum_{K_4} S^2 \langle f, h_{K_1} \otimes h_{I_3} \otimes h_{K_4} \rangle_{1,3,4}^2 \langle \langle G, h_{K_1} \otimes h_{J_3} \rangle_{1,3} \nu \rangle_{K_4,4}^2 \frac{1_{K_4}}{|K_4|} \right)^{\frac{1}{2}}.
\end{aligned}$$

Next, we estimate

$$|\langle \langle G, h_{K_1} \otimes h_{J_3} \rangle_{1,3} \nu \rangle_{K_4,4}| \leq M^4 [\langle M^3 \Delta_{K_1 \times K_3, (0, l_3)}^{1,3} G \rangle_{K_1, 1} \nu] |J_3|^{\frac{1}{2}} |K_1|^{\frac{1}{2}}$$

and

$$\begin{aligned}
& \left(\sum_{K_4} S^2 \langle f, h_{K_1} \otimes h_{I_3} \otimes h_{K_4} \rangle_{1,3,4}^2 \otimes \frac{1_{K_4}}{|K_4|} \right)^{\frac{1}{2}} \\
& \lesssim \langle S^{2,4} \Delta_{K_1 \times K_3, (0, k_3)}^{1,3} f \rangle_{K_1 \times I_3, 1, 3} |I_3|^{\frac{1}{2}} |K_1|^{\frac{1}{2}}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
& \int_{\mathbb{R}^d} \sum_{K_1, K_3} M^4 [M^3 \langle \Delta_{K_1 \times K_3, (0, l_3)}^{1,3} G \rangle_{K_1, 1} \nu] \\
& \quad \times \langle S^{2,4} \Delta_{K_1 \times K_3, (0, k_3)}^{1,3} f \rangle_{K_1 \times K_3, 1, 3} \otimes 1_{K_1} \otimes 1_{K_3} \\
& \lesssim \left\| \left(\sum_{K_1, K_3} M^4 [M^3 \langle \Delta_{K_1 \times K_3, (0, l_3)}^{1,3} G \rangle_{K_1, 1} \nu]^2 1_{K_1} \otimes 1_{K_3} \right)^{\frac{1}{2}} \right\|_{L^{p'}(\mu^{1-p'})} \\
& \quad \times \left\| \left(\sum_{K_1, K_3} \langle S^{2,4} \Delta_{K_1 \times K_3, (0, k_3)}^{1,3} f \rangle_{K_1 \times K_3, 1, 3}^2 \otimes 1_{K_1} \otimes 1_{K_3} \right)^{\frac{1}{2}} \right\|_{L^p(\mu)} \\
& \lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|g\|_{L^{p'}(\lambda^{1-p'})} \|f\|_{L^p(\mu)}.
\end{aligned}$$

Next, fix the integer $t_3 \in [1, l_3]$ and consider the dual form of the term B_2

$$\begin{aligned} & \left| \sum_{\substack{K_1, K_2 \\ K_3, K_4}} \sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(l_3)}=K_3}} a_{K_1, K_2} \langle a_{K_3, I_3, J_3}, h_{K_4} \rangle \right. \\ & \quad \times \sum_{Q_2 \subset K_2} |K_2|^{-1} |J_3^{(t_3)}|^{-\frac{1}{2}} \left\langle b, \frac{1_{K_1}}{|K_1|} \otimes h_{Q_2} \otimes h_{J_3^{(t_3)}} \otimes \frac{1_{K_4}}{|K_4|} \right\rangle \\ & \quad \times \langle f, h_{K_1} \otimes h_{Q_2} \otimes h_{I_3} \otimes h_{K_4} \rangle \left. \left\langle g, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{J_3} \otimes \frac{1_{K_4}}{|K_4|} \right\rangle \right|. \end{aligned}$$

Thus, by Proposition 2.9 we are estimating the following term

$$\begin{aligned} & \int_{\mathbb{R}^d} \sum_{\substack{K_1, K_2 \\ K_3}} |a_{K_1, K_2}| \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{K_2}}{|K_2|} \\ & \times \left(\sum_{\substack{Q_2 \\ P_3^{(l_3-t_3)}=K_3}} \left(\sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(t_3)}=P_3}} |P_3|^{-\frac{1}{2}} \pi^4(f_{K_1, Q_2, I_3}, g_{K_1, K_2, J_3}) \right)^2 \frac{1_{Q_2}}{|Q_2|} \otimes \frac{1_{P_3}}{|P_3|} \right)^{\frac{1}{2}} \nu, \end{aligned}$$

where

$$\begin{aligned} & \pi^4(f_{K_1, Q_2, I_3}, g_{K_1, K_2, J_3}) \\ & := \pi_{K_3, I_3, J_3}^4 \left(\langle f, h_{K_1} \otimes h_{Q_2} \otimes h_{I_3} \rangle_{1,2,3}, \left\langle g, \frac{1_{K_1}}{|K_1|} \otimes h_{K_2} \otimes h_{J_3} \right\rangle_{1,2,3} \right) \end{aligned}$$

is a bilinear one-parameter paraproduct where $a_{K_4} \langle \varphi, h_{K_4} \rangle \langle \phi \rangle_{K_4}$ is replaced by $|a_{K_4} \langle \varphi, h_{K_4} \rangle \langle \phi \rangle_{K_4}|$. As previously, by sparse domination and A_∞ -extrapolation we get

$$\begin{aligned} & \int_{\mathbb{R}^d} \sum_{\substack{K_1, K_2 \\ K_3}} |a_{K_1, K_2}| \frac{1_{K_1}}{|K_1|} \otimes \frac{1_{K_2}}{|K_2|} \\ & \times \left(\sum_{\substack{Q_2 \\ P_3^{(l_3-t_3)}=K_3}} \left(\sum_{\substack{I_3^{(k_3)}=K_3 \\ J_3^{(t_3)}=P_3}} |P_3|^{-\frac{1}{2}} \pi^4(f_{K_1, Q_2, I_3}, g_{K_1, K_2, J_3}) \right)^2 \frac{1_{Q_2}}{|Q_2|} \otimes \frac{1_{P_3}}{|P_3|} \right)^{\frac{1}{2}} \nu \\ & \lesssim_{[\nu]_{A_\infty}} \int_{\mathbb{R}^d} \sum_{\substack{K_1, K_2 \\ K_3}} |a_{K_1, K_2}| \frac{1_{K_1}}{|K_1|^{\frac{1}{2}}} \otimes \frac{1_{K_2}}{|K_2|^{\frac{1}{2}}} \otimes 1_{K_3} \langle M^{3,4} \Delta_{K_2 \times K_3, (0, l_3)}^{2,3} g \rangle_{K_1 \times K_2, 1,2} \\ & \times \left(\sum_{Q_2} \langle M^4 \Delta_{K_1 \times Q_2 \times K_3, (0,0,k_3)}^{1,2,3} f \rangle_{K_1 \times Q_2 \times K_3, 1,2,3}^2 \otimes 1_{Q_2} \right)^{\frac{1}{2}} \nu \end{aligned}$$

$$\begin{aligned}
&\lesssim \int_{\mathbb{R}^d} \sum_{K_3} 1_{K_3} \left(\sum_{K_1, K_2} \langle M^{3,4} \Delta_{K_2 \times K_3, (0, l_3)}^{2,3} g \rangle_{K_1 \times K_2, 1,2}^2 \right. \\
&\quad \times \left. \left\langle \left(\sum_{Q_2} \langle M^4 \Delta_{K_1 \times Q_2 \times K_3, (0,0,k_3)}^{1,2,3} f \rangle_{K_1 \times Q_2 \times K_3, 1,2,3}^2 \otimes 1_{Q_2} \right)^{\frac{1}{2}} \nu \right\rangle_{K_1 \times K_2, 1,2}^2 1_{K_1} \otimes 1_{K_2} \right)^{\frac{1}{2}} \\
&\leq \int_{\mathbb{R}^d} \left(\sum_{K_2, K_3} \langle M^{1,3,4} \Delta_{K_2 \times K_3, (0, l_3)}^{2,3} g \rangle_{K_2, 2}^2 \otimes 1_{K_2} \right)^{\frac{1}{2}} \left(\sum_{K_1, K_3} M^2 \right. \\
&\quad \times \left. \left\langle \left(\sum_{Q_2} \langle M^4 \Delta_{K_1 \times Q_2 \times K_3, (0,0,k_3)}^{1,2,3} f \rangle_{K_1 \times Q_2 \times K_3, 1,2,3}^2 \otimes 1_{Q_2} \right)^{\frac{1}{2}} \nu \right\rangle_{K_1, 1}^2 1_{K_1 \times K_3} \right)^{\frac{1}{2}} \\
&\lesssim_{[\mu]_{A_p}, [\lambda]_{A_p}} \|g\|_{L^{p'}(\lambda^{1-p'})} \|f\|_{L^p(\mu)}. \quad \square
\end{aligned}$$

We return to consider the space $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_m}$.

Theorem 5.2. *Let $\nu = \mu^{\frac{1}{p}} \lambda^{-\frac{1}{p}}$, where $\mu, \lambda \in A_p(\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_m})$ in \mathbb{R}^d and $1 < p < \infty$. In addition, let $\mathcal{I} = \{\mathcal{I}_i\}_{i=1}^k$ be a partition of $\{1, 2, \dots, m\}$. Given CZO $T^{\bar{v}_i}$, where $\bar{v}_i = (j)_{j \in \mathcal{I}_i}$, suppose that at least one of the following conditions holds:*

- (1) *the CZO T_i is paraproduct free, or*
- (2) *$\#\mathcal{I}_i \leq 2$*

for all $i = 1, \dots, k$. Then we have

$$\| [T^{\bar{v}_1}, [T^{\bar{v}_2}, \dots [b, T^{\bar{v}_k}]]] \|_{L^p(\mu) \rightarrow L^p(\lambda)} \lesssim \|b\|_{\text{bmo}^{\mathcal{I}}(\nu)}.$$

Here, even with the case $k = 3$ with bi-parameter operators, we have a collection 27 commutators. Actually, even more, when counting different forms of paraproducts. We can use the same strategy as in the case $k = 2$ also here and essentially nothing really changes. Clearly, the number of paraproduct coefficients increase but techniques used in the case $k = 2$ also apply to these situations. The previous theorem is not stated for paraproduct free CZOs. However, if we combine techniques of Theorem 4.4, we can allow paraproduct free CZOs of arbitrary parameters. We omit the details. Furthermore, we remark that the case $k = 1$ is proven in [34] for the bi-parameter CZOs.

References

- [1] A. BARRON, J. M. CONDE-ALONSO, Y. OU, AND G. REY, Sparse domination and the strong maximal function, *Adv. Math.* **345** (2019), 1–26. DOI: 10.1016/j.aim.2019.01.007.
- [2] A. BARRON AND J. PIPHER, Sparse domination for bi-parameter operators using square functions, Preprint (2017). [arXiv:1709.05009](https://arxiv.org/abs/1709.05009).
- [3] S. BLOOM, A commutator theorem and weighted BMO, *Trans. Amer. Math. Soc.* **292**(1) (1985), 103–122. DOI: 10.2307/2000172.

- [4] S.-Y. A. CHANG AND R. FEFFERMAN, A continuous version of duality of H^1 with BMO on the bidisc, *Ann. of Math. (2)* **112**(1) (1980), 179–201. DOI: 10.2307/1971324.
- [5] S.-Y. A. CHANG AND R. FEFFERMAN, Some recent developments in Fourier analysis and H^p -theory on product domains, *Bull. Amer. Math. Soc. (N.S.)* **12**(1) (1985), 1–43. DOI: 10.1090/S0273-0979-1985-15291-7.
- [6] R. R. COIFMAN, R. ROCHBERG, AND G. WEISS, Factorization theorems for Hardy spaces in several variables, *Ann. of Math. (2)* **103**(3) (1976), 611–635. DOI: 10.2307/1970954.
- [7] D. CRUZ-URIBE, J. M. MARTELL, AND C. PÉREZ, Extrapolation from A_∞ weights and applications, *J. Funct. Anal.* **213**(2) (2004), 412–439. DOI: 10.1016/j.jfa.2003.09.002.
- [8] L. DALENC AND Y. OU, Upper bound for multi-parameter iterated commutators, *Publ. Mat.* **60**(1) (2016), 191–220. DOI: 10.5565/PUBLMAT-60116-07.
- [9] G. DAVID AND J.-L. JOURNÉ, A boundedness criterion for generalized Calderón–Zygmund operators, *Ann. of Math. (2)* **120**(2) (1984), 371–397. DOI: 10.2307/2006946.
- [10] X. T. DUONG, J. LI, Y. OU, J. PIPHER, AND B. D. WICK, Commutators of multiparameter flag singular integrals and applications, *Anal. PDE* **12**(5) (2019), 1325–1355. DOI: 10.2140/apde.2019.12.1325.
- [11] C. FEFFERMAN AND E. M. STEIN, Some maximal inequalities, *Amer. J. Math.* **93**(1) (1971), 107–115. DOI: 10.2307/2373450.
- [12] R. FEFFERMAN AND E. M. STEIN, Singular integrals on product spaces, *Adv. in Math.* **45**(2) (1982), 117–143. DOI: 10.1016/S0001-8708(82)80001-7.
- [13] S. H. FERGUSON AND M. T. LACEY, A characterization of product BMO by commutators, *Acta Math.* **189**(2) (2002), 143–160. DOI: 10.1007/BF02392840.
- [14] S. H. FERGUSON AND C. SADOSKY, Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures, *J. Anal. Math.* **81** (2000), 239–267. DOI: 10.1007/BF02788991.
- [15] A. GRAU DE LA HERRÁN, Comparison of $T1$ conditions for multi-parameter operators, *Proc. Amer. Math. Soc.* **144**(6) (2016), 2437–2443. DOI: 10.1090/proc/12891.
- [16] I. HOLMES, M. T. LACEY, AND B. D. WICK, Bloom’s inequality: commutators in a two-weight setting, *Arch. Math. (Basel)* **106**(1) (2016), 53–63. DOI: 10.1007/s00013-015-0840-8.
- [17] I. HOLMES, M. T. LACEY, AND B. D. WICK, Commutators in the two-weight setting, *Math. Ann.* **367**(1–2) (2017), 51–80. DOI: 10.1007/s00208-016-1378-1.
- [18] I. HOLMES, S. PETERMICHL, AND B. D. WICK, Weighted little bmo and two-weight inequalities for Journé commutators, *Anal. PDE* **11**(7) (2018), 1693–1740. DOI: 10.2140/apde.2018.11.1693.
- [19] I. HOLMES AND B. D. WICK, Two weight inequalities for iterated commutators with Calderón–Zygmund operators, *J. Operator Theory* **79**(1) (2018), 33–54. DOI: 10.7900/jot.
- [20] T. P. HYTÖNEN, The sharp weighted bound for general Calderón–Zygmund operators, *Ann. of Math. (2)* **175**(3) (2012), 1473–1506. DOI: 10.4007/annals.2012.175.3.9.
- [21] T. P. HYTÖNEN, The Holmes–Wick theorem on two-weight bounds for higher order commutators revisited, *Arch. Math. (Basel)* **107**(4) (2016), 389–395. DOI: 10.1007/s00013-016-0956-5.

- [22] T. P. HYTÖNEN, The two-weight inequality for the Hilbert transform with general measures, *Proc. Lond. Math. Soc.* (3) **117**(3) (2018), 483–526. DOI: 10.1112/plms.12136.
- [23] T. P. HYTÖNEN, The L^p -to- L^q boundedness of commutators with applications to the Jacobian operator, Preprint (2018). [arXiv:1804.11167](https://arxiv.org/abs/1804.11167).
- [24] J.-L. JOURNÉ, Calderón–Zygmund operators on product spaces, *Rev. Mat. Iberoamericana* **1**(3) (1985), 55–91. DOI: 10.4171/RMI/15.
- [25] I. KUNWAR AND Y. OU, Two-weight inequalities for multilinear commutators, *New York J. Math.* **24** (2018), 980–1003.
- [26] M. T. LACEY, Two-weight inequality for the Hilbert transform: a real variable characterization, II, *Duke Math. J.* **163**(15) (2014), 2821–2840. DOI: 10.1215/00127094-2826799.
- [27] M. T. LACEY, S. PETERMICHL, J. C. PIPHER, AND B. D. WICK, Multiparameter Riesz commutators, *Amer. J. Math.* **131**(3) (2009), 731–769. DOI: 10.1353/ajm.0.0059.
- [28] M. T. LACEY, S. PETERMICHL, J. C. PIPHER, AND B. D. WICK, Iterated Riesz commutators: a simple proof of boundedness, in: “*Harmonic Analysis and Partial Differential Equations*”, Contemp. Math. **505**, Amer. Math. Soc., Providence, RI, 2010, pp. 171–178. DOI: 10.1090/conm/505/09922.
- [29] M. T. LACEY, S. PETERMICHL, J. C. PIPHER, AND B. D. WICK, Multi-parameter Div-Curl lemmas, *Bull. Lond. Math. Soc.* **44**(6) (2012), 1123–1131. DOI: 10.1112/blms/bds037.
- [30] M. T. LACEY, E. T. SAWYER, C.-Y. SHEN, AND I. URIARTE-TUERO, Two-weight inequality for the Hilbert transform: a real variable characterization, I, *Duke Math. J.* **163**(15) (2014), 2795–2820. DOI: 10.1215/00127094-2826690.
- [31] A. K. LERNER, S. OMBROSI, AND I. P. RIVERA-RÍOS, On pointwise and weighted estimates for commutators of Calderón–Zygmund operators, *Adv. Math.* **319** (2017), 153–181. DOI: 10.1016/j.aim.2017.08.022.
- [32] A. K. LERNER, S. OMBROSI, AND I. P. RIVERA-RÍOS, Commutators of singular integrals revisited, *Bull. Lond. Math. Soc.* **51**(1) (2019), 107–119. DOI: 10.1112/blms.12216.
- [33] K. LI, H. MARTIKAINEN, Y. OU, AND E. VUORINEN, Bilinear representation theorem, *Trans. Amer. Math. Soc.* **371**(6) (2019), 4193–4214. DOI: 10.1090/tran/7505.
- [34] K. LI, H. MARTIKAINEN, AND E. VUORINEN, Bloom-type inequality for bi-parameter singular integrals: efficient proof and iterated commutators, *Int. Math. Res. Not. IMRN* (2019). DOI: 10.1093/imrn/rnz072.
- [35] K. LI, H. MARTIKAINEN, AND E. VUORINEN, Bilinear Calderón–Zygmund theory on product spaces, *J. Math. Pures Appl.* (9) **138** (2020), 356–412. DOI: 10.1016/j.matpur.2019.10.007.
- [36] K. LI, H. MARTIKAINEN, AND E. VUORINEN, Bloom type upper bounds in the product BMO setting, *J. Geom. Anal.* **30** (2020), 3181–3203. DOI: 10.1007/s12220-019-00194-3.
- [37] H. MARTIKAINEN, Representation of bi-parameter singular integrals by dyadic operators, *Adv. Math.* **229**(3) (2012), 1734–1761. DOI: 10.1016/j.aim.2011.12.019.
- [38] F. NAZAROV, S. TREIL, AND A. VOLBERG, The T_b -theorem on non-homogeneous spaces, *Acta Math.* **190**(2) (2003), 151–239. DOI: 10.1007/BF02392690.

- [39] Y. OU, Multi-parameter singular integral operators and representation theorem, *Rev. Mat. Iberoam.* **33(1)** (2017), 325–350. DOI: 10.4171/RMI/939.
- [40] Y. OU, S. PETERMICHL, AND E. STROUSE, Higher order Journé commutators and characterizations of multi-parameter BMO, *Adv. Math.* **291** (2016), 24–58. DOI: 10.1016/j.aim.2015.12.029.
- [41] S. PETERMICHL, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, *C. R. Acad. Sci. Paris Sér. I Math.* **330(6)** (2000), 455–460. DOI: 10.1016/S0764-4442(00)00162-2.

Department of Mathematics and Statistics, University of Helsinki, P.O.B. 68, FI-00014 University of Helsinki, Finland
E-mail address: `emil.airta@helsinki.fi`

Primera versió rebuda el 28 de juny de 2019,
darrera versió rebuda el 19 de desembre de 2019.