



**HAL**  
open science

# Higher order Journé commutators and characterizations of multi-parameter BMO

Yumeng Ou, Stefanie Petermichl, Elizabeth Strouse

► **To cite this version:**

Yumeng Ou, Stefanie Petermichl, Elizabeth Strouse. Higher order Journé commutators and characterizations of multi-parameter BMO. *Advances in Mathematics*, In press. hal-01977607

**HAL Id: hal-01977607**

**<https://hal.science/hal-01977607>**

Submitted on 10 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Higher order Journé commutators and characterizations of multi-parameter BMO

Yumeng Ou<sup>a,1</sup> Stefanie Petermichl<sup>b,2,3</sup> Elizabeth Strouse<sup>c</sup>

<sup>a</sup>*Department of mathematics, Brown University, 151 Thayer Street, Providence RI 02912, USA*

<sup>b</sup>*Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France*

<sup>c</sup>*Institut de Mathématiques de Bordeaux, 351 cours de la Libération, F-33405 Talence, France*

---

## Abstract

We characterize  $L^p$  boundedness of iterated commutators of multiplication by a symbol function and tensor products of Riesz and Hilbert transforms. We obtain a two-sided norm estimate that shows that such operators are bounded on  $L^p$  if and only if the symbol belongs to the appropriate multi-parameter BMO class. We extend our results to a much more intricate situation; commutators of multiplication by a symbol function and paraproduct-free Journé operators. We show that the boundedness of these commutators is also determined by the inclusion of their symbol function in the same multi-parameter BMO class. In this sense the tensor products of Riesz transforms are a representative testing class for Journé operators.

Previous results in this direction do not apply to tensor products and only to Journé operators which can be reduced to Calderón-Zygmund operators. Upper norm estimate of Journé commutators are new even in the case of no iterations. Lower norm estimates for iterated commutators only existed when no tensor products were present. In the case of one dimension, lower estimates were known for products of two Hilbert transforms, and without iterations. New methods using Journé operators are developed to obtain these lower norm estimates in the multi-parameter real variable setting.

*Key words:* Iterated commutator, Journé operator, multi-parameter, BMO

---

## 1 Introduction

As dual of the Hardy space  $H^1$ , the classical space of functions of bounded mean oscillation, BMO, arises naturally in many endpoint results in analysis, partial differential equations and probability. When entering a setting with several free parameters, a large variety of spaces are encountered, some of which lose the feature of mean oscillation itself. We are interested in characterizations of multi-parameter BMO spaces through boundedness of commutators.

A classical result of Nehari [26] shows that a Hankel operator with anti-analytic symbol  $b$  mapping analytic functions into the space of anti-analytic functions by  $f \mapsto P_-bf$  is bounded with respect to the  $L^2$  norm if and only if the symbol belongs to BMO. This theorem has an equivalent formulation in terms of the boundedness of the commutator of the multiplication operator with symbol function  $b$  and the Hilbert transform  $[H, b] = Hb - bH$ .

Ferguson-Sadosky in [14] and later Ferguson-Lacey in their groundbreaking paper [13] study the symbols of bounded ‘big’ and ‘little’ Hankel operators on the bidisk through commutators of the tensor product or of the iterated form

$$[H_1H_2, b], \text{ and } [H_1, [H_2, b]].$$

Here  $b = b(x_1, x_2)$  and the  $H_k$  are the Hilbert transforms acting in the  $k^{\text{th}}$  variable. A full characterization of different two-parameter BMO spaces, Cotlar-Sadosky’s little BMO and Chang-Fefferman’s product BMO space, is given through these commutators.

Through the use of completely different real variable methods, in [6] Coifman-Rochberg-Weiss extended Nehari’s one-parameter theory to real analysis in the sense that the Hilbert transform was replaced by Riesz transforms. These one-parameter results in [6] were treated in the multi-parameter setting in Lacey-Petermichl-Pipher-Wick [18]. Both the upper and lower estimate have proofs very different from those in one parameter. In addition, in both cases it is observed that the Riesz transforms are a representative testing class in the sense that BMO also ensures boundedness for (iterated) commutators with more general Calderon-Zygmund operators, a result now known in full generality due to Dalenc-Ou [8]. Notably the Riesz commutator has found

---

*Email address:* [stefanie.petermichl@gmail.com](mailto:stefanie.petermichl@gmail.com) (Stefanie Petermichl).

*URL:* <http://math.univ-toulouse.fr/~petermic> (Stefanie Petermichl).

<sup>1</sup> Research supported in part by NSF-DMS 0901139.

<sup>2</sup> Research supported in part by ANR-12-BS01-0013-02. The author is a member of IUF.

<sup>3</sup> Corresponding author, Tel:+33 5 61 55 76 59, Fax: +33 5 61 55 83 85

striking applications to compensated compactness and div-curl lemmas, [3], [20].

Our extension to the multi-parameter setting is two-fold. On the one hand we replace the Calderon-Zygmund operators by Journé operators  $J_i$  and on the other hand we also iterate the commutator:

$$[J_1, \dots, [J_t, b] \dots].$$

We prove the remarkable fact that a multi-parameter BMO class still ensures boundedness in this situation and that the collection of tensor products of Riesz transforms remains the representative testing class. The BMO class encountered is a mix of little BMO and product BMO that we call a little product BMO. Its precise form depends upon the distribution of variables in the commutator. Our result is new even when no iterations are present: in this case, lower estimates were only known in the case of the double Hilbert transform [14]. The sufficiency of the little BMO class for boundedness of Journé commutators had never been observed.

It is a general fact that two-sided commutator estimates have an equivalent formulation in terms of weak factorization. We find the pre-duals of our little product BMO spaces and prove a corresponding weak factorization result.

Necessity of the little product BMO condition is shown through a lower estimate on the commutator. There is a sharp contrast when tensor products of Riesz transforms are considered instead of multiple Hilbert transforms and when iterations are present.

In the Hilbert transform case, Toeplitz operators with operator symbol arise naturally. Using Riesz transforms in  $\mathbb{R}^d$  as a replacement, there is an absence of analytic structure and tools relying on analytic projection or orthogonal spaces are not readily available. We overcome part of this difficulty through the use of Calderón-Zygmund operators whose Fourier multiplier symbols are adapted to cones. This idea is inspired by [18]. Such operators are also mentioned in [31]. A class of operators of this type classifies little product BMO through two-sided commutator estimates, but it does not allow the passage to a classification through iterated commutators with tensor products of Riesz transforms. In a second step, we find it necessary to consider upper and lower commutator estimates using a well-chosen family of Journé operators that are not of tensor product type. Through geometric considerations and an averaging procedure of zonal harmonics on products of spheres, we construct the multiplier of a special Journé operator that preserves lower commutator estimates and resembles the multiple Hilbert transform: it has large plateaus of constant values and is a polynomial in multiple Riesz transforms. We expect

that this construction allows other applications.

There is an increase in difficulty when the dimension is greater than two, due to the simpler structure of the rotation group on  $\mathbb{S}^1$ . In higher dimension, there is a rise in difficulty when tensor products involve more than two Riesz transforms.

The actual passage to the Riesz transforms requires a stability estimate in commutator norms for certain multi-parameter singular integrals in terms of the mixed BMO class. In this context, we prove a qualitative upper estimate for iterated commutators using paraproduct free Journé operators. We make use of recent versions of  $T(1)$  theorems in this setting. These recent advances are different from the corresponding theorem of Journé [16]. The results we allude to have the additional feature of providing a convenient representation formula for bi-parameter in [22] and even multi-parameter in [28] Calderón-Zygmund operators by dyadic shifts.

## 2 Aspects of Multi-Parameter Theory

This section contains some review on Hardy spaces in several parameters as well as some new definitions and lemmas relevant to us.

### 2.1 Chang-Fefferman BMO

We describe the elements of product Hardy space theory, as developed by Chang and Fefferman as well as Journé. By this we mean the Hardy spaces associated with domains like the poly-disk or  $\mathbb{R}^{\mathbf{d}} := \bigotimes_{s=1}^t \mathbb{R}^{d_s}$  for  $\mathbf{d} = (d_1, \dots, d_t)$ . While doing so, we typically do not distinguish whether we are working on  $\mathbb{R}^{\mathbf{d}}$  or  $\mathbb{T}^{\mathbf{d}}$ . In higher dimensions, the Hilbert transform is usually replaced by the collection of Riesz transforms.

The (real) one-parameter Hardy space  $H_{\text{Re}}^1(\mathbb{R}^d)$  denotes the class of functions with the norm

$$\sum_{j=0}^d \|R_j f\|_1$$

where  $R_j$  denotes the  $j^{\text{th}}$  Riesz transform or the Hilbert transform if the dimension is one. Here and below we adopt the convention that  $R_0$ , the  $0^{\text{th}}$  Riesz transform, is the identity. This space is invariant under the one-parameter family of isotropic dilations, while the product Hardy space  $H_{\text{Re}}^1(\mathbb{R}^{\mathbf{d}})$  is invariant under dilations of each coordinate separately. That is, it is invariant under a  $t$

parameter family of dilations, hence the terminology ‘multi-parameter’ theory. One way to define a norm on  $H_{\text{Re}}^1(\mathbb{R}^d)$  is

$$\|f\|_{H^1} \sim \sum_{0 \leq j_l \leq d_l} \left\| \bigotimes_{l=1}^t R_{l, j_l} f \right\|_1.$$

$R_{l, j_l}$  is the Riesz transform in the  $j_l^{\text{th}}$  direction of the  $l^{\text{th}}$  variable, and the  $0^{\text{th}}$  Riesz transform is the identity operator.

The dual of the real Hardy space  $H_{\text{Re}}^1(\mathbb{R}^d)^*$  is  $\text{BMO}(\mathbb{R}^d)$ , the  $t$ -fold product BMO space. It is a theorem of S.-Y. Chang and R. Fefferman [4], [5] that this space has a characterization in terms of a product Carleson measure.

Define

$$\|b\|_{\text{BMO}(\mathbb{R}^d)} := \sup_{U \subset \mathbb{R}^d} \left( |U|^{-1} \sum_{R \subset U} \sum_{\varepsilon \in \text{sig}_d} |\langle b, w_R^\varepsilon \rangle|^2 \right)^{1/2}. \quad (1)$$

Here the supremum is taken over all open subsets  $U \subset \mathbb{R}^d$  with finite measure, and we use a wavelet basis  $w_R^\varepsilon$  adapted to rectangles  $R = Q_1 \times \cdots \times Q_t$ , where each  $Q_l$  is a cube. The superscript  $\varepsilon$  reflects the fact that multiple wavelets are associated to any dyadic cube, see [18] for details. The fact that the supremum admits all open sets of finite measure cannot be omitted, as Carleson’s example shows [2]. This fact is responsible for some of the difficulties encountered when working with this space.

**Theorem 1 (Chang, Fefferman)** *We have the equivalence of norms*

$$\|b\|_{(H_{\text{Re}}^1(\mathbb{R}^d))^*} \sim \|b\|_{\text{BMO}(\mathbb{R}^d)}.$$

*That is,  $\text{BMO}(\mathbb{R}^d)$  is the dual to  $H_{\text{Re}}^1(\mathbb{R}^d)$ .*

This BMO norm is invariant under a  $t$ -parameter family of dilations. Here the dilations are isotropic in each parameter separately. See also [10] and [12].

## 2.2 Little BMO

Following [7] and [14], we recall some facts about the space little BMO, often written as ‘bmo’, and its predual. A locally integrable function  $b : \mathbb{R}^d = \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_s} \rightarrow \mathbb{C}$  is in bmo if and only if

$$\|b\|_{\text{bmo}} = \sup_{\mathbf{Q} = Q_1 \times \cdots \times Q_s} |\mathbf{Q}|^{-1} \int_{\mathbf{Q}} |b(\mathbf{x}) - b_{\mathbf{Q}}| < \infty$$

Here the  $Q_k$  are  $d_k$ -dimensional cubes and  $b_{\mathbf{Q}}$  denotes the average of  $b$  over  $\mathbf{Q}$ .

It is easy to see that this space consists of all functions that are uniformly in BMO in each variable separately. Let  $\mathbf{x}_{\hat{v}} = (x_1, \dots, x_{v-1}, \cdot, x_{v+1}, \dots, x_s)$ . Then  $b(\mathbf{x}_{\hat{v}})$  is a function in  $x_v$  only with the other variables fixed. Its BMO norm in  $x_v$  is

$$\|b(\mathbf{x}_{\hat{v}})\|_{\text{BMO}} = \sup_{Q_v} |Q_v|^{-1} \int_{Q_v} |b(\mathbf{x}) - b(\mathbf{x}_{\hat{v}})_{Q_v}| dx_v$$

and the little BMO norm becomes

$$\|b\|_{\text{bmo}} = \max_v \left\{ \sup_{\mathbf{x}_{\hat{v}}} \|b(\mathbf{x}_{\hat{v}})\|_{\text{BMO}} \right\}.$$

On the bi-disk, this becomes

$$\|b\|_{\text{bmo}} = \max \left\{ \sup_{x_1} \|b(x_1, \cdot)\|_{\text{BMO}}, \sup_{x_2} \|b(\cdot, x_2)\|_{\text{BMO}} \right\},$$

the space discussed in [14]. Here, the pre-dual is the space  $H^1(\mathbb{T}) \otimes L^1(\mathbb{T}) + L^1(\mathbb{T}) \otimes H^1(\mathbb{T})$ . All other cases are an obvious generalization, at the cost of notational inconvenience.

### 2.3 Little product BMO

In this section we define a BMO space which is in between little BMO and product BMO. As mentioned in the introduction, we aim at characterizing BMO spaces consisting for example of those functions  $b(x_1, x_2, x_3)$  such that  $b(x_1, \cdot, \cdot)$  and  $b(\cdot, \cdot, x_3)$  are uniformly in product BMO in the remaining two variables.

**Definition 1** Let  $b : \mathbb{R}^d \rightarrow \mathbb{C}$  with  $\mathbf{d} = (d_1, \dots, d_t)$ . Take a partition  $\mathcal{I} = \{I_s : 1 \leq s \leq l\}$  of  $\{1, 2, \dots, t\}$  so that  $\dot{\cup}_{1 \leq s \leq l} I_s = \{1, 2, \dots, t\}$ . We say that  $b \in \text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$  if for any choices  $\mathbf{v} = (v_s), v_s \in I_s$ ,  $b$  is uniformly in product BMO in the variables indexed by  $v_s$ . We call a BMO space of this type a ‘little product BMO’. If for any  $\mathbf{x} = (x_1, \dots, x_t) \in \mathbb{R}^d$ , we define  $\mathbf{x}_{\hat{\mathbf{v}}}$  by removing those variables indexed by  $v_s$ , the little product BMO norm becomes

$$\|b\|_{\text{BMO}_{\mathcal{I}}} = \max_{\mathbf{v}} \left\{ \sup_{\mathbf{x}_{\hat{\mathbf{v}}}} \|b(\mathbf{x}_{\hat{\mathbf{v}}})\|_{\text{BMO}} \right\}$$

where the BMO norm is product BMO in the variables indexed by  $v_s$ .

For example, when  $\mathbf{d} = (1, 1, 1) = \mathbf{1}$ , when  $t = 3$  and  $l = 2$  with  $I_1 = (13)$  and  $I_2 = (2)$ , writing  $\mathcal{I} = (13)(2)$  the space  $\text{BMO}_{(13)(2)}(\mathbb{T}^1)$  arises, which consists of those functions that are uniformly in product BMO in the variables (1, 2) and (3, 2) respectively, as described above. Moreover, as degenerate cases, it

is easy to see that  $BMO_{(12\dots t)}$  and  $BMO_{(1)(2)\dots(t)}$  are exactly little BMO and product BMO respectively, the spaces we are familiar with.

Little product BMO spaces on  $\mathbb{T}^d$  can be defined in the same way. Now we find the predual of  $BMO_{(13)(2)}$ , which is a good model for other cases. We choose the order of variables most convenient for us.

**Theorem 2** *The pre-dual of the space  $BMO_{(13)(2)}(\mathbb{T}^1)$  is equal to the space*

$$\begin{aligned} & H_{Re}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T}) + L^1(\mathbb{T}) \otimes H_{Re}^1(\mathbb{T}^{(1,1)}) \\ & := \{f + g : f \in H_{Re}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T}) \text{ and } g \in L^1(\mathbb{T}) \otimes H_{Re}^1(\mathbb{T}^{(1,1)})\}. \end{aligned}$$

*Proof.* The space

$$H_{Re}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T}) = \{f \in L^1(\mathbb{T}^3) : H_1 f, H_2 f, H_1 H_2 f \in L^1(\mathbb{T}^3)\}$$

equipped with the norm  $\|f\| = \|f\|_1 + \|H_1 f\|_1 + \|H_2 f\|_1 + \|H_1 H_2 f\|_1$  is a Banach space. Let  $W^1 = L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3)$  equipped with the norm

$$\|(f_1, f_2, f_3, f_4)\|_{W^1} = \|f_1\|_1 + \|f_2\|_1 + \|f_3\|_1 + \|f_4\|_1.$$

Then we see that  $H_{Re}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T})$  is isomorphically isometric to the closed subspace

$$V = \{(f, H_1(f), H_2(f), H_1 H_2(f)) : f \in H^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T})\}$$

of  $W^1$ . Now, the dual of  $W^1$  is equal to  $W^\infty = L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3)$  equipped with the norm  $\|(g_1, g_2, g_3, g_4)\|_\infty = \max\{\|g_i\|_\infty : 1 \leq i \leq 4\}$  so the dual space of  $V$  is equal to the quotient of  $W^\infty$  by the annihilator  $U$  of the subspace  $V$  in  $W^\infty$ . But, using the fact that the Hilbert transforms are self-adjoint up to a sign change, we see that

$$U = \{(g_1, g_2, g_3, g_4) : g_1 + H_1 g_2 + H_2 g_3 + H_1 H_2 g_4 = 0\}$$

and so:

$$V^* \cong W^\infty / U \cong \text{Im}(\theta)$$

where

$$\theta(g_1, g_2, g_3, g_4) = g_1 + H_1 g_2 + H_2 g_3 + H_1 H_2 g_4$$

since  $U = \ker(\theta)$ . But

$$\text{Im}(\theta) = L^\infty(\mathbb{T}^3) + H_1(L^\infty(\mathbb{T}^3)) + H_2(L^\infty(\mathbb{T}^3)) + H_1(H_2(L^\infty(\mathbb{T}^3)))$$

is equal to the functions that are uniformly in product BMO in variables 1 and 2.

Using the same reasoning we see that the dual of  $L^1(\mathbb{T}) \otimes H_{Re}^1(\mathbb{T}^{(1,1)})$  is equal to  $L^\infty(\mathbb{T}^3) + H_2(L^\infty(\mathbb{T}^3)) + H_3(L^\infty(\mathbb{T}^3)) + H_2 H_3(L^\infty(\mathbb{T}^3))$ , which is equal to



the space of functions that are uniformly in product BMO in variables 2 and 3.

Now, we consider the ‘ $L^1$  sum’ of the spaces  $H_{\text{Re}}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T})$  and  $L^1(\mathbb{T}) \otimes H_{\text{Re}}^1(\mathbb{T}^{(1,1)})$ ; that is

$$M_{(13)(2)} = \{(f, g) : f \in H_{\text{Re}}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T}); g \in L^1(\mathbb{T}) \otimes H_{\text{Re}}^1(\mathbb{T}^{(1,1)})\}$$

equipped with the norm

$$\|(f, g)\| = \|f\|_{H_{\text{Re}}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T})} + \|g\|_{L^1(\mathbb{T}) \otimes H_{\text{Re}}^1(\mathbb{T}^{(1,1)})}.$$

We see that, if  $\phi : M_{(13)(2)} \rightarrow L^1(\mathbb{T}^3)$  is defined by  $\phi(f, g) = f + g$ , then the image of  $\phi$  is isometrically isomorphic to the quotient of  $M_{(13)(2)}$  by the space

$$\begin{aligned} N &= \{(f, g) \in M_{(13)(2)} : f + g = 0\} \\ &= \{(f, -f) : f \in H_{\text{Re}}^1(\mathbb{T}^{(1,1)}) \otimes L^1(\mathbb{T}) \cap L^1(\mathbb{T}) \otimes H_{\text{Re}}^1(\mathbb{T}^{(1,1)})\}. \end{aligned}$$

Now, recall that the dual of the quotient  $M/N$  is equal to the annihilator of  $N$ . It is easy to see that the annihilator of  $N$  is equal to the set of ordered pairs  $(\phi, \phi)$  with  $\phi$  in the intersection of the duals of the two spaces. Thus the dual of the image of  $\theta$  is equal to  $\text{BMO}_{(13)(2)}$ . The norm of an element in the predual is equal to its norm as an element of the double dual which is easily computed. *QED*

Following this example, the reader may easily find the correct formulation for the predual of other little product BMO spaces as well those in several variables, replacing the Hilbert transform by all choices of Riesz transforms. For instance, one can prove that the predual of the space  $\text{BMO}_{(13)(2)}(\mathbb{R}^d)$  is equal to  $H_{\text{Re}}^1(\mathbb{R}^{(d_1, d_2)}) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H_{\text{Re}}^1(\mathbb{R}^{(d_2, d_3)})$ .

### 3 The Hilbert transform case

In this section, we characterize the boundedness of commutators of the form  $[H_2, [H_3 H_1, b]]$  as operators on  $L^2(\mathbb{T}^3)$ . In the case of the Hilbert transform, this case is representative of the general case and provides a starting point that is easier to read because of the simplicity of the expression of products and sums of projection onto orthogonal subspaces. Its general form can be found at the beginning of Section 4.

Now let  $b \in L^1(\mathbb{T}^n)$  and let  $P$  and  $Q$  denote orthogonal projections onto subspaces of  $L^2(\mathbb{T}^n)$ . We shall describe relationships between functions in the little product BMOs and several types of projection-multiplication operators.

These will be Hilbert transform-type operators of the form  $P - P^\perp$ ; and iterated Hankel or Toeplitz type operators of the form  $Q^\perp b Q$  (Hankel),  $P b P$  (Toeplitz),  $P Q^\perp b Q P$  (mixed), where  $b$  means the (not a priori bounded) multiplication operator  $M_b$  on  $L^2(\mathbb{T}^n)$ .

We shall use the following simple observation concerning Hilbert transform type operators again and again:

**Remark 1** *If  $H = P - P^\perp$  and  $T : L^2(\mathbb{T}^n) \rightarrow L^2(\mathbb{T}^n)$  is a linear operator then*

$$[H, T] = 2PTP^\perp - 2P^\perp TP$$

*and  $H$  is bounded if and only if  $PTP^\perp$  and  $P^\perp TP$  are.*

*Proof.*

$$\begin{aligned} (P - P^\perp)T - T(P - P^\perp) &= (P - P^\perp)T(P + P^\perp) - (P + P^\perp)T(P - P^\perp) \\ &= 2PTP^\perp - 2P^\perp TP. \end{aligned}$$

*QED*

We state the main result of this section.

**Theorem 3** *Let  $b \in L^1(\mathbb{T}^3)$ . Then the following are equivalent with linear dependence on the respective norms*

- (1)  $b \in BMO_{(13)(2)}$
- (2) *The commutators  $[H_2, [H_1, b]]$  and  $[H_2, [H_3, b]]$  are bounded on  $L^2(\mathbb{T}^3)$*
- (3) *The commutator  $[H_2, [H_3 H_1, b]]$  is bounded on  $L^2(\mathbb{T}^3)$ .*

**Corollary 1** *We have the following two-sided estimate*

$$\|b\|_{BMO_{(13)(2)}} \lesssim \|[H_2, [H_3 H_1, b]]\|_{L^2(\mathbb{T}^3) \rightarrow L^2(\mathbb{T}^3)} \lesssim \|b\|_{BMO_{(13)(2)}}.$$

It will be useful to denote by  $Q_{13}$  orthogonal projection on the subspace of functions which are either analytic or anti-analytic in the first and third variables;  $Q_{13} = P_1 P_3 + P_1^\perp P_3^\perp$ . Then the projection  $Q_{13}^\perp$  onto the orthogonal of this subspace is defined by  $Q_{13}^\perp = P_1^\perp P_3 + P_1 P_3^\perp$ . We reformulate properties (2) and (3) in the statement of Theorem 3 in terms of Hankel Toeplitz type operators.

**Lemma 1** *We have the following algebraic facts on commutators and projection operators.*

- (1) *The commutators  $[H_2, [H_1, b]]$  and  $[H_2, [H_3, b]]$  are bounded on  $L^2(\mathbb{T}^3)$  if and only if the operators  $P_i P_2 b P_i^\perp P_2^\perp$ ,  $P_i^\perp P_2 b P_i P_2^\perp$ ,  $P_i P_2^\perp b P_i^\perp P_2$ ,  $P_i^\perp P_2^\perp b P_i P_2$  with  $i \in \{1, 3\}$  are bounded on  $L^2(\mathbb{T}^3)$ .*

(2) The commutator  $[H_2, [H_3 H_1, b]]$  is bounded on  $L^2(\mathbb{T}^3)$  if and only if all four operators  $P_2 Q_{13} b Q_{13}^\perp P_2^\perp, P_2^\perp Q_{13}^\perp b Q_{13} P_2, P_2 Q_{13}^\perp b Q_{13} P_2^\perp, P_2^\perp Q_{13} b Q_{13}^\perp P_2$  are bounded on  $L^2(\mathbb{T}^3)$ .

*Proof.* Using Remark 1 it is easy to see that

$$[H_2, [H_1, b]] = 4((P_2 P_1 b P_1^\perp P_2^\perp - P_2 P_1^\perp b P_1 P_2^\perp) - (P_2^\perp P_1 b P_1^\perp P_2 - P_2^\perp P_1^\perp b P_1 P_2))$$

and that the corresponding equation for  $[H_2, [H_3, b]]$  is also true. This, along with the observation that the ranges of all arising summands are mutually orthogonal, gives assertion (1). To prove (2) we just notice that  $H_1 H_3 = Q_{13} - Q_{13}^\perp$  is a Hilbert transform type operator which permits us to repeat the above argument replacing  $P_1$  by  $Q_{13}$ . QED

The following lemma will allow us to insert an additional Hilbert transform into the commutator without reducing the norm.

**Lemma 2**  $\|P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3\|_{L^2 \rightarrow L^2} = \|P_1^\perp P_2^\perp b P_1 P_2\|_{L^2 \rightarrow L^2}$ .

*Proof.*

The inequality  $\leq$  is trivial, since  $P_3$  is a projection which commutes with  $P_1^\perp$  and  $P_2^\perp$ . To see  $\geq$ , notice that  $P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3$  is a Toeplitz operator with symbol  $P_1^\perp P_2^\perp b P_1 P_2$ . So  $\|P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3\| = \sup_{x_3} \|P_1^\perp P_2^\perp b(\cdot, \cdot, x_3) P_1 P_2\|$ . The latter is just  $\|P_1^\perp P_2^\perp b P_1 P_2\|$ . For convenience we include a sketch of the facts about Toeplitz operators we use. Let  $W_3$  be the operator of multiplication by  $z_3$ ,  $W_3(f) = z_3 f$ , acting on  $L^2(\mathbb{T}^3)$ . If we define  $B = P_1^\perp P_2^\perp b P_1 P_2$  as well as

$$A_n = W_3^{*n} (P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3) W_3^n \text{ and } C_n = W_3^n (P_3^\perp P_1^\perp P_2^\perp b P_1 P_2 P_3^\perp) W_3^{*n}$$

as operators acting on  $L^2(\mathbb{T}^3)$  then the sequences  $A_n$  and  $C_n$  converge to  $B$  in the strong operator topology: it is easy to see that  $W_3, W_3^*$ ; and  $P_3$  commute with  $P_1, P_2, P_1^\perp$  and  $P_2^\perp$ . The multiplier  $b$  satisfies the equation  $W_3^{*n} b W_3^n = b$  and  $W_3^n W_3^{*n} = Id$ . So we see that

$$A_n = P_1^\perp P_2^\perp (W_3^{*n} P_3 W_3^n) b P_1 P_2 (W_3^{*n} P_3 W_3^n).$$

But if  $f \in L^2(\mathbb{T}^3)$ , then, since  $W_3^n$  is a unitary operator:

$$\|W_3^{*n} P_3 W_3^n(f) - f\| = \|P_3 W_3^n(f) - W_3^n(f)\| = \|(P_3 - I)(W_3^n)(f)\| \rightarrow 0 \quad (n \rightarrow \infty),$$

as tail of a convergent Fourier series. This means that  $W_3^{*n} P_3 W_3^n$  converges to the identity in the strong operator topology. Thus, for each  $f \in L^2(\mathbb{T}^3)$  we

have  $\|(A_n - B)(f)\| \rightarrow 0$ . So

$$\begin{aligned} \|P_1^\perp P_2^\perp b P_1 P_2\| &\leq \sup_{n \in \mathbb{N}} \|W_3^{*n} (P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3) W_3^n\| \\ &\leq \|P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3\|, \end{aligned}$$

QED

Now, we are ready to proceed with the proof of the main theorem of this section.

*Proof.* (of Theorem 3) We show (1)  $\Leftrightarrow$  (2) and (2)  $\Leftrightarrow$  (3).

(1)  $\Leftrightarrow$  (2). Consider  $f = f(x_1, x_2)$  and  $g = g(x_3)$ . Then  $[H_2, [H_1, b]](fg) = g \cdot [H_2, [H_1, b]](f)$ . So  $\|[H_2, [H_1, b]](fg)\|_{L^2(\mathbb{T}^3)}^2 = \|Fg\|_{L^2(\mathbb{T})}^2$  where  $F(x_2) = \|[H_2, [H_1, b]](f)\|_{L^2(\mathbb{T}^2)}$ . The map  $g \mapsto Fg$  has  $L^2(\mathbb{T})$  operator norm  $\|F\|_\infty$ . Now change the roles of  $x_1$  and  $x_3$ . The Ferguson-Lacey equivalences  $\|[H_2, [H_i, b]]\| \sim \|b\|_{\text{BMO}}$  give the desired result.

(2)  $\Rightarrow$  (3). Boundedness of the commutators  $[H_2, [H_1, b]]$  and  $[H_2, [H_3, b]]$  implies the boundedness of the mixed commutator  $[H_2, [H_1 H_3, b]]$  by the identity  $[H_2, [H_1 H_3, b]] = H_1 [H_2, [H_3, b]] + [H_2, [H_1, b]] H_3$ .

(3)  $\Rightarrow$  (2). This part relies on Lemma 2. We wish to conclude from the boundedness of  $[H_2, [H_3 H_1, b]]$  the boundedness of  $[H_2, [H_1, b]]$  and  $[H_2, [H_3, b]]$ . To see boundedness of  $[H_2, [H_1, b]]$ , let us look at one of the Hankels from Lemma 1. Lemma 2 shows that  $P_2^\perp P_1^\perp b P_2 P_1$  is bounded if and only if the operator  $P_3 P_1^\perp P_2^\perp b P_1 P_2 P_3$  is. And the latter is an operator found in the list from part (2) of Lemma 1. The analogous reasoning shows that all eight Hankels in 1 are bounded and so (2) is proved. QED

## 4 Real variables: lower bounds

In this section, we are again in  $\mathbb{R}^d$  with  $\mathbf{d} = (d_1, \dots, d_t)$  and a partition  $\mathcal{I} = (I_s)_{1 \leq s \leq l}$  of  $\{1, \dots, t\}$ . It is our aim to prove the following characterization theorem of the space  $\text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$ .

**Theorem 4** *The following are equivalent with linear dependence of the respective norms.*

(1)  $b \in \text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$

- (2) All commutators of the form  $[R_{k_1, j_{k_1}}, \dots, [R_{k_l, j_{k_l}}, b] \dots]$  are bounded in  $L^2(\mathbb{R}^d)$  where  $k_s \in I_s$  and  $R_{k_s, j_{k_s}}$  is the one-parameter Riesz transform in direction  $j_{k_s}$ .
- (3) All commutators of the form  $[\mathbf{R}_{1, \mathbf{j}^{(1)}}, \dots, [\mathbf{R}_{l, \mathbf{j}^{(l)}}, b] \dots]$  are bounded in  $L^2(\mathbb{R}^d)$  where  $\mathbf{j}^{(s)} = (j_k)_{k \in I_s}$ ,  $1 \leq j_k \leq d_k$  and the operators  $\mathbf{R}_{s, \mathbf{j}^{(s)}}$  are a tensor product of Riesz transforms  $\mathbf{R}_{s, \mathbf{j}^{(s)}} = \bigotimes_{k \in I_s} R_{k, j_k}$ .

Such two-sided estimates also hold in  $L^p$  for  $1 < p < \infty$ . Remarks will be made in section 7. From the inductive nature of our arguments, it will also be apparent that the characterization holds when we consider intermediate cases, meaning commutators with any fixed number of Riesz transforms in each iterate. Below we state our most general two-sided estimate through Riesz transforms.

**Theorem 5** *Let  $1 < p < \infty$ . Under the same assumptions as Corollary 2 and for any fixed  $\mathbf{n} = (n_s)$  where  $1 \leq n_s \leq |I_s|$ , we have the two-sided estimate*

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)} \lesssim \sup_{\mathbf{j}} \|[ \mathbf{R}_{1, \mathbf{j}^{(1)}}, \dots, [\mathbf{R}_{l, \mathbf{j}^{(l)}}, b] \dots ]\|_{L^p(\mathbb{R}^d) \curvearrowright} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)}$$

where  $\mathbf{j}^{(s)} = (j_k)_{k \in I_s}$ ,  $0 \leq j_k \leq d_k$  and for each  $s$ , there are  $n_s$  non-zero choices. A Riesz transform in direction 0 is understood as the identity.

For  $p = 2$  and  $\mathbf{n} = \mathbf{1}$  this is the equivalence (1)  $\Leftrightarrow$  (2) and for  $\mathbf{n} = (|I_1|, \dots, |I_l|)$  it is the equivalence (1)  $\Leftrightarrow$  (3) from Theorem 4.

Our main focus is of course on a two-sided estimate when  $\mathbf{n} = (|I_1|, \dots, |I_l|)$  when the tensor product is a paraproduct-free Journé operator:

**Corollary 2** *Let  $\mathbf{j} = (j_1, \dots, j_t)$  with  $1 \leq j_k \leq d_k$  and let for each  $1 \leq s \leq l$ ,  $\mathbf{j}^{(s)} = (j_k)_{k \in I_s}$  be associated a tensor product of Riesz transforms  $\mathbf{R}_{s, \mathbf{j}^{(s)}} = \bigotimes_{k \in I_s} R_{k, j_k}$ ; here the  $R_{k, j_k}$  are  $j_k^{\text{th}}$  Riesz transforms acting on functions defined on the  $k^{\text{th}}$  variable. We have the two-sided estimate*

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)} \lesssim \sup_{\mathbf{j}} \|[ \mathbf{R}_{1, \mathbf{j}^{(1)}}, \dots, [\mathbf{R}_{t, \mathbf{j}^{(t)}}, b] \dots ]\|_{L^p(\mathbb{R}^d) \curvearrowright} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)}.$$

The statements above also serve as the statement of the general case for products of Hilbert transforms. In fact, when any  $d_k = 1$  just replace the Riesz transforms by the Hilbert transform in that variable. In this section, we consider the case  $d_k \geq 2$  for  $1 \leq k \leq t$  and thus iterated commutators with tensor products of Riesz transforms only. The special case when  $d_k = 1$  for some  $k$  is easier but requires extra care for notation, which is why we omit it here.

The proof in the Hilbert transform case relied heavily on analytic projections and orthogonal spaces, a feature that we do not have when working with

Riesz transforms. We are going to simulate the one-dimensional case by a two-step passage via intermediary Calderón-Zygmund operators whose multiplier symbols are adapted to cones.

In dimension  $d \geq 2$ , a cone  $C \subset \mathbb{R}^d$  with cubic base is given by the data  $(\xi, Q)$  where  $\xi \in \mathbb{S}^{d-1}$  is the direction of the cone and the cube  $Q \subset \xi^\perp$  centered at the origin is its aperture. The cone consists of all vectors  $\theta$  that take the form  $(\theta_\xi \xi, \theta_\perp)$  where  $\theta_\xi = \langle \theta, \xi \rangle$  and  $\theta_\perp \in \theta_\xi Q$ . By  $\lambda C$  we mean the dilated cone with data  $(\xi, \lambda Q)$ .

A cone  $D$  with ball base has data  $(\xi, r)$  for  $0 < r < \pi/2$  and  $\xi \in \mathbb{S}^{d-1}$  and consists of the vectors  $\{\eta \in \mathbb{R}^d : d(\xi, \eta/\|\eta\|) \leq r\}$  where  $d$  is the geodesic distance (with distance of antipodal points being  $\pi$ ).

Given any cone  $C$  or  $D$ , we consider its Fourier projection operator defined via  $\widehat{P_C} f = \chi_C \widehat{f}$ . When the apertures are cubes, such operators are combinations of Fourier projections onto half spaces and as such admit uniform  $L^p$  bounds. Among others, this fact made cubic cones necessary in the considerations in [18] and [9] that we are going to need. For further technical reasons in the proof these operators are not quite good enough, mainly because they are not of Calderón-Zygmund type. For a given cone  $C$ , consider a Calderón-Zygmund operator  $T_C$  with a kernel  $K_C$  whose Fourier symbol  $\widehat{K_C} \in C^\infty$  and satisfies the estimate  $\chi_C \leq \widehat{K_C} \leq \chi_{(1+\tau)C}$ . This is accomplished by mollifying the symbol  $\chi_C$  of the cone projection associated to cone  $C$  on  $\mathbb{S}^{d-1}$  and then extending radially. We use the same definition for  $T_D$ .

Given a collection of cones  $\mathbf{C} = (C_k)$  we denote by  $T_{\mathbf{C}}, P_{\mathbf{C}}$  the corresponding tensor product operators.

In [18] it has been proved that Calderón-Zygmund operators adapted to certain cones of cubic aperture classify product BMO via commutators. As part of the argument, it was observed that test functions with opposing Fourier supports made the commutator large. In [9] a refinement was proven, that will be helpful to us. We prefer to work with cones with round base. Lower bounds for such commutators can be deduced from the assertion of the main theorem in [9], but we need to preserve the information on the Fourier support of the test function in order to succeed with our argument. Information on this test function is instrumental to our argument: it reduces the terms arising in the commutator to those resembling Hankel operators. We have the following lemma, very similar to that in [18] section 7 and [9] section 3, the only difference being that the cones are based on balls instead of cubes.

**Lemma 3** *For every parameter  $1 \leq k \leq t$  there exist a finite set of directions  $\Upsilon_k \in \mathbb{S}^{d_k-1}$  and an aperture  $0 < r_k < \pi/2$  so that, for every symbol  $b$  belonging to product BMO, there exist cones  $D_k = D(\xi_k, r_k)$  with  $\xi_k \in \Upsilon_k$  as well as a normalised test function  $f = \bigotimes_{k=1}^t f_k$  whose components have Fourier support*

in the opposing cones  $D(-\xi_k, r_k)$  such that

$$\|[T_{1,D_1}, \dots, [T_{t,D_t}, b] \dots] f\|_2 \gtrsim \|b\|_{BMO_{(1)\dots(t)}(\mathbb{R}^d)}.$$

The stress is on the fact that the collection is finite, somewhat specific and serves all admissible product BMO functions.

*Proof.* The lemma in [9] supplies us with the sets of directions  $\Upsilon_k$  as well as cones of cubic aperture  $Q_k$  and a test function  $f$  supported in the opposing cones. Now choose the aperture  $r_k$  large enough so that  $(1 + \tau)C(\xi_k, Q_k) \subset D(\xi_k, r_k)$ . Then we have the commutator estimate

$$\|[T_{1,D_1}, \dots, [T_{t,D_t}, b] \dots] f\|_2 \gtrsim \|b\|_{BMO_{(1)\dots(t)}(\mathbb{R}^d)}.$$

In fact, both commutators with cones  $C$  and  $D$  are  $L^2$  bounded and reduce to  $\|T_{\mathcal{D}}(bf)\|_2$  or  $\|T_{\mathcal{C}}(bf)\|_2$  respectively thanks to the opposing Fourier support of  $f$ . Observe that  $T_{\mathcal{C}}(bf) = T_{\mathcal{D}}(T_{\mathcal{C}}(bf)) = T_{\mathcal{C}}(T_{\mathcal{D}}(bf))$ . With  $\|T_{\mathcal{C}}\|_{2 \rightarrow 2} \leq 1$ , we see that  $\|T_{\mathcal{D}}(bf)\|_2 \geq \|T_{\mathcal{C}}(bf)\|_2$ . QED

Using this a priori lower estimate, we are going to prove the lemma below.

**Lemma 4** *Let us suppose we are in  $\mathbb{R}^d$  with  $\mathbf{d} = (d_1, \dots, t)$  and a partition  $\mathcal{I} = (I_s)_{1 \leq s \leq l}$ . For every  $1 \leq k \leq t$  there exists a finite set of directions  $\Upsilon_k \subset \mathbb{S}^{d_k-1}$  and an aperture  $r_k$  so that the following hold for all  $b \in BMO_{\mathcal{I}}(\mathbb{R}^d)$ :*

(1) *For every  $1 \leq s \leq l$  there exists a coordinate  $v_s \in I_s$  and a direction  $\xi_{v_s} \in \Upsilon_{v_s}$  and so that with the choice of cone  $D_{v_s} = D(\xi_{v_s}, r_{v_s})$  and arbitrary  $D_k$  for coordinates  $k \in I_s \setminus \{v_s\}$  and if  $\mathbf{D}_s$  denotes their tensor product, then we have*

$$\|[T_{1,\mathbf{D}_1}, \dots, [T_{l,\mathbf{D}_l}, b] \dots]\|_{2 \rightarrow 2} \gtrsim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)},$$

(2) *The test function  $f = \bigotimes_{k=1}^t f_k$  which gives us a large  $L^2$  norm in (1) has Fourier supports of the  $f_k$  contained in  $D(-\xi_k, r_k)$  when  $k = v_s$  and in  $D_k$  otherwise.*

Before we can begin with the proof of Lemma 4, we will need a real variable version of the facts on Toeplitz operators used earlier.

**Lemma 5** *Let  $D_k$  for  $1 \leq k \leq t$  denote any cones with respect to the  $k^{\text{th}}$  variable. Let  $T_{D_k}$  denote the adapted Calderón-Zygmund operators. Let  $K$  be any proper subset of  $\{k : 1 \leq k \leq t\}$ , let  $\mathbf{D}_K = \bigotimes_{k \in K} D_k$  and  $T_{\mathbf{D}_K}$  the associated tensor product of Calderón-Zygmund operators. Let  $P_{\mathbf{D}_K}^\sigma$  be a tensor product of projection operators on cones  $D(\xi_k, r_k)$  or opposing cones  $D(-\xi_k, r_k)$ . Let  $j \notin K$ . Then*

$$\|T_{\mathbf{D}_K} T_{D_j} b P_{\mathbf{D}_K}^\sigma P_{D_j}\|_{L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)} = \|T_{\mathbf{D}_K} b P_{\mathbf{D}_K}^\sigma\|_{L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)}.$$

*Proof.*

We will establish this by composing some unilateral shift operators and studying their Fourier transform in the  $j$  variable. Let  $\xi_j$  denote the direction of the cone  $D_j$ , for any  $l$  define the shift operator

$$S_l g(x_j) = \int_{\mathbb{R}^{d_j}} \hat{g}(\eta_j) e^{2\pi i(l\xi_j + \eta_j)x_j} d\eta_j.$$

$S_l$  is a translation operator on the Fourier side along the direction  $\xi_j$  of the cone  $D_j$ . It is not hard to observe that  $S_l^* = S_{-l}$ . Now define

$$A_l = S_{-l} T_{\mathcal{D}_K} T_{D_j} b P_{\mathcal{D}_K}^\sigma P_{D_j} S_l, \text{ and } B = T_{\mathcal{D}_K} b P_{\mathcal{D}_K}^\sigma.$$

We will prove that as  $l \rightarrow +\infty$ ,  $A_l \rightarrow B$  in the strong operator topology. As in the argument in Lemma 2, this together with the fact that  $S_l$  is an isometry will complete the proof. To see the convergence, let's first remember that  $S_l$  only acts on the  $j$  variable, and one always has the identities

$$S_l S_{-l} = Id \quad \text{and} \quad S_{-l} b S_l = b.$$

This implies

$$\begin{aligned} A_l &= T_{\mathcal{D}_K} (S_{-l} T_{D_j} S_l) (S_{-l} b S_l) P_{\mathcal{D}_K}^\sigma (S_{-l} P_{D_j} S_l) \\ &= T_{\mathcal{D}_K} (S_{-l} T_{D_j} S_l) b P_{\mathcal{D}_K}^\sigma (S_{-l} P_{D_j} S_l). \end{aligned}$$

We claim that both  $S_{-l} T_{D_j} S_l$  and  $S_{-l} P_{D_j} S_l$  converge to the identity operator in the strong operator topology, which then implies that  $A_l \rightarrow B$  as  $l \rightarrow \infty$ . We will only prove  $S_{-l} T_{D_j} S_l \rightarrow Id$  as the second limit is almost identical. Observe that  $\|S_{-l} T_{D_j} S_l f - f\| = \|(T_{D_j} - I) S_l f\|$ . Given any  $L^2$  function  $f$  and any fixed large  $l \geq 0$ . Consider the  $f$  with frequencies supported in  $\mathbb{R}^{d_1} \times \dots \times (D_j - l\xi_j) \times \dots \times \mathbb{R}^{d_t}$ . In this case,  $S_l f$  has Fourier support in  $\mathbb{R}^{d_1} \times \dots \times D_j \times \dots \times \mathbb{R}^{d_t}$  where the symbol of  $T_{D_j}$  equals 1. Thus, for such  $f$ , we have  $S_{-l} T_{D_j} S_l f = f$ . The sets  $\mathbb{R}^{d_1} \times \dots \times (D_j - l\xi_j) \times \dots \times \mathbb{R}^{d_t}$  exhaust the frequency space. With  $\|T_{D_j} - I\|_{2 \rightarrow 2} \leq 1$  the operators  $S_{-l} T_{D_j} S_l$  converge to the Identity in the strong operator topology, and the lemma is proved. Observe that the aperture of the cone  $D_j$  is not relevant to the proof. *QED*

We proceed with the proof of the lower estimate for cone transforms.

*Proof.* (of Lemma 4) For a given symbol  $b \in \text{BMO}_{\mathcal{I}}$ , there exist for all  $1 \leq s \leq l$  coordinates  $\mathbf{v} = (v_s), v_s \in I_s$  and a choice of variables not indexed by  $v_s, \mathbf{x}_{\mathbf{v}}^0$  so that up to an arbitrarily small error

$$\|b\|_{\text{BMO}_{\mathcal{I}}} = \|b(\mathbf{x}_{\mathbf{v}}^0)\|_{\text{BMO}_{(v_1)\dots(v_l)}}.$$

By Lemma 3, there exist cones  $D_{v_s} = D(\xi_{v_s}, r_{v_s})$  with directions  $\xi_{v_s} \in \Upsilon_{v_s}$  and a normalised test function  $f_H$  in variables  $v_s$  with opposing Fourier support



such that we have the lower estimate

$$\|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b(\mathbf{x}_{\hat{\mathbf{v}}}^0)] \dots](f_H)\|_{L^2(\mathbb{R}^{d_{\mathbf{v}}})} \gtrsim \|b(\mathbf{x}_{\hat{\mathbf{v}}}^0)\|_{\text{BMO}_{(v_1) \dots (v_l)}}$$

where  $\mathbb{R}^{d_{\mathbf{v}}} = \mathbb{R}^{d_{v_1}} \times \dots \times \mathbb{R}^{d_{v_l}}$ .

We now consider the commutator with the same cones but with full symbol  $b = b(\cdot, \dots, \cdot)$ . Due to the lack of action on the variables not indexed by  $v_s$ , in the commutator, we have

$$[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f_H g) = g \cdot [T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f_H)$$

for  $g$  that only depends upon variables not indexed by  $v_s$ . Again using that multiplication operators in  $L^2$  have norms equal to the  $L^\infty$  norm of their symbol, for the ‘worst’  $L^2$ -normalised  $g$  we have

$$\begin{aligned} & \|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f_H g)\|_{L^2(\mathbb{R}^d)} \\ &= \sup_{\mathbf{x}_{\hat{\mathbf{v}}}} \|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b(\mathbf{x}_{\hat{\mathbf{v}}}^0)] \dots](f_H)\|_{L^2(\mathbb{R}^{d_{\mathbf{v}}})} \\ &\geq \|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b(\mathbf{x}_{\hat{\mathbf{v}}}^0)] \dots](f_H)\|_{L^2(\mathbb{R}^{d_{\mathbf{v}}})} \\ &\gtrsim \|b(\mathbf{x}_{\hat{\mathbf{v}}}^0)\|_{\text{BMO}_{(v_1) \dots (v_l)}(\mathbb{R}^{d_{\mathbf{v}}})} = \|b\|_{\text{BMO}_{\mathcal{I}}(\mathbb{R}^d)}. \end{aligned}$$

Note that the test function  $g$  can be chosen with well distributed Fourier transform. Take any cones in the variables not indexed by  $v_s$  and let  $\mathbf{D}$  denote the tensor product of their projections.  $f_T = P_{\mathbf{D}}g$ . Notice that

$$\|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f_H f_T)\| \gtrsim \|[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f_H g)\|$$

with constants depending upon how small the aperture of the chosen cones is. Notice that the test function  $f := f_H f_T$  has the Fourier support as required in part (2) of the statement of Lemma 4.

Now build cones  $\mathbf{D}_s$  from the  $D_{v_s}$  and the other chosen cones  $D_k$  as well as operators  $T_{s, \mathbf{D}_s}$ . Notice that the commutators  $[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots]$  and  $[T_{1, \mathbf{D}_1}, \dots, [T_{l, \mathbf{D}_l}, b] \dots]$  reduce significantly when applied to a test function  $f$  with Fourier support like ours. When the operators  $T_{v_s, D_{v_s}}$  or any tensor product  $T_{s, \mathbf{D}_s}$  fall directly on  $f$ , the contribution is zero due to opposing Fourier supports of the test function and the symbols of the operators. The only terms left in the commutators  $[T_{1, \mathbf{D}_1}, \dots, [T_{l, \mathbf{D}_l}, b] \dots](f)$  and  $[T_{v_1, D_{v_1}}, \dots, [T_{v_l, D_{v_l}}, b] \dots](f)$  have the form  $\bigotimes_s T_{s, \mathbf{D}_s}(bf)$  and  $\bigotimes_s T_{v_s, D_{v_s}}(bf)$  respectively.

By repeated use of Lemma 5 we have the operator norm estimates for any symbol  $b$ , valid on the subspace of functions with Fourier support as described for  $f$ :  $\|\bigotimes_s T_{s, \mathbf{D}_s} b\|_{2 \rightarrow 2} = \|\bigotimes_s T_{v_s, D_{v_s}} b\|_{2 \rightarrow 2}$ . We conclude that a normalised test function  $f$  with Fourier support as described in the statement (2) of Lemma 4

exists, so that  $\|\otimes_s T_{s, \mathcal{D}_s}(bf)\|_2 \gtrsim \|b\|_{\text{BMO}_{\mathcal{I}}(\mathbb{R}^d)}$ . In particular, we get the desired estimate in (1). QED

It does not seem possible to pass directly to a lower commutator estimate for tensor products of Riesz transforms from that for tensor products of cone operators. Just using tensor products of operators adapted to cones merely gives us *some* lower bound where we are unable to control that a Riesz transform does appear in every variable such as required in (3) of Theorem 4. The reason for this will become clear as we advance in the argument. Instead of using operators  $T_{s, \mathcal{D}_s}$  directly, we will build upon them more general multi-parameter Journé type cone operators not of tensor product type that we now describe.

Let us explain the multiplier we need for  $i$  copies of  $\mathbb{S}^{d-1}$  when all dimensions are the same. We will explain how to pass to the case of  $i$  copies of varying dimension  $d_k$  below. A picture illustrating a base case, a product of two 1-spheres, can be found at the end of this section.

For  $0 < b < a < 1$ , let  $\varphi : [-1, 1] \rightarrow [-1, 1]$  be a smooth function with  $\varphi(x) = 1$  when  $a \leq x \leq 1$  and  $\varphi(x) = 0$  when  $b \geq x \geq 0$ . And let  $\varphi$  be odd, meaning antisymmetric about  $t = 0$ . The function  $\varphi$  gives rise to a zonal function with pole  $\xi_1$  on the first copy of  $\mathbb{S}^{d-1}$ , denoted by  $C_1(\xi_1; \eta_1)$ . This is the multiplier of a one-parameter Calderón-Zygmund operator adapted to a cone  $D(\xi_1, r)$  for  $r = \pi/2(1 - a)$ . For  $i > 1$  we define  $C_k(\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_k)$  for  $1 < k \leq i$  inductively. In what follows, expectation is taken with respect to traces of surface measure. When  $\eta_i = \pm \xi_i$ , then conditional expectation is over a one-point set.

$$\begin{aligned} & C_k(\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_k) \\ &= \mathbb{E}_{a_{k-1}}(C_{k-1}(\xi_1, \dots, \xi_{k-1}; \eta_1, \dots, \eta_{k-1}) \mid d(a_{k-1}, \xi_{k-1}) = d(\eta_k, \xi_k)). \end{aligned}$$

If the dimensions are not equal take  $d = \max d_j$  and imbed  $\mathbb{S}^{d_j-1}$  into  $\mathbb{S}^{d-1}$  by the map  $\xi = (\xi_1, \dots, \xi_{d_j}) \mapsto (\xi_1, \dots, \xi_{d_j}, 0, \dots, 0)$ . Obtain in this manner the function  $C_i$  and then restrict to the original number of variables when the dimension is smaller than  $d$ .

The multiplier  $\mathbf{J} = C_i(\boldsymbol{\xi}; \cdot)$  gives rise to a multi-parameter Calderón-Zygmund operator of convolution type (but not of tensor product type),  $\mathbf{T}_{\mathbf{J}} = \mathbf{T}_{C_i(\boldsymbol{\xi}; \cdot)}$ . In fact, it is defined through principal value convolution against a kernel  $\mathbf{K}_{\mathbf{J}} = \mathbf{K}_{C_i(\boldsymbol{\xi}; \cdot)}(x_1, \dots, x_i)$  such that

$$\forall l : \int_{\alpha < |x_l| < \beta} \mathbf{K}_{\mathbf{J}}(x_1, \dots, x_i) dx_l = 0, \forall 0 < \alpha < \beta, x_j \in \mathbb{R}^{d_j} \text{ fixed } \forall j \neq l,$$

$$\left| \frac{\partial^{|\mathbf{n}|}}{\partial x_1^{n_1} \dots \partial x_i^{n_i}} \mathbf{K}_{\mathbf{J}}(x_1, \dots, x_i) \right| \leq A_{\mathbf{n}} |x_1|^{-d_1-n_1} \dots |x_i|^{-d_i-n_i}, n_j \geq 0.$$

This kind of operator is a special case of the more general, non-convolution type discussed in Section 5. It has many other nice features that will facilitate our passage to Riesz transforms. One of them is its very special representation in terms of homogeneous polynomials, the other one a lower commutator estimate in terms of the  $BMO_{\mathcal{I}}$  norm.

**Lemma 6** *Let  $C_i$  be a multiplier in  $\bigotimes_{k=1}^i \mathbb{R}^{d_k}$  as described above, with any fixed direction and aperture. Let  $m$  be an integer of order  $d = \max d_k$ . For any  $\delta > 0$ , the function  $C_i$  has an approximation by a polynomial  $C_i^N$  in the  $\prod_{k=1}^i d_k$  variables  $\{\prod_{k:1 \leq k \leq i} \eta_{k,j_k} \mid 1 \leq j_k \leq d_k\}$  so that  $\|C_i - C_i^N\|_{\mathcal{C}^m(\mathbb{S}^{d_k-1})} < \delta$  in each variable separately.*

$\mathcal{C}^m$  indexes the norm of uniform convergence on functions that are  $m$  times continuously differentiable. On the space side,  $C_i^N$  corresponds to an operator that is a polynomial in Riesz transforms of the variables  $\bigotimes_k R_{k,j_k}$ .

**Lemma 7** *We are in  $\mathbb{R}^d$  with partition  $\mathcal{I} = (I_s)_{1 \leq s \leq l}$ . Let  $\Upsilon$  consist of vectors  $\xi = (\xi_k)_{k=1}^l$  with  $\xi_k \in \Upsilon_k$ . Let  $\Upsilon^{(s)}$  consist of  $\xi^{(s)} = (\xi_k)_{k \in I_s}$ . Let us consider the class of Journé type cone multipliers  $\mathbf{J}_s = C_{i_s}(\xi^{(s)}; \cdot)$  of aperture  $r_s$  with associated multi-parameter Calderón-Zygmund operators  $\mathbf{T}_{s,\mathbf{J}_s}$ . Then we have the two-sided estimate*

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)} \lesssim \sup_{\xi \in \Upsilon} \|\mathbf{T}_{1,\mathbf{J}_1}, \dots, \mathbf{T}_{l,\mathbf{J}_l}, b\|_{L^2(\mathbb{R}^d)} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)}.$$

In order to proceed with the proof of these lemmas, we will use some well known facts about zonal harmonics. Fix a pole  $\xi \in \mathbb{S}^{d-1}$ . The zonal harmonic with pole  $\xi$  of degree  $n$  is written as  $Z_{\xi}^{(n)}(\eta)$ . With  $t = \langle \xi, \eta \rangle \in [-1, 1]$ , one writes  $Z_{\xi}^{(n)}(\eta) = P_n(t)$  where  $P_n$  is the Legendre polynomial of degree  $n$ . It is common to suppress the dependence on  $d$  in the notation for  $Z_{\xi}^{(n)}$  and  $P_n$ .

$Z_{\xi}^{(n)}$  are reproducing for spherical harmonics of degree  $n$ ,  $Y^{(n)}$ . When  $Y^{(n)}$  is harmonic and homogeneous of degree  $n$  with  $Y^{(n)}(\xi) = 1$  and  $Y^{(n)}(R\eta) = Y^{(n)}(\eta)$  for any rotation  $R \in \mathcal{O}(d)$  with  $R\xi = \xi$ , then  $Y^{(n)} = Z_{\xi}^{(n)}$ .

The lemma below will aid us in understanding the special form of the functions  $C_i$ .

**Lemma 8** *Let  $\xi_1, \xi_2 \in \mathbb{S}^{d-1}$ . We have*

$$\begin{aligned} Z_{\xi_1}^{(n)}(\eta_1)Z_{\xi_2}^{(n)}(\eta_2) &= \mathbb{E}_{a_1}(Z_{\eta_1}^{(n)}(a_1) \mid d(\xi_1, a_1) = d(\xi_2, \eta_2)) \\ &= \mathbb{E}_{a_2}(Z_{\eta_2}^{(n)}(a_2) \mid d(\xi_2, a_2) = d(\xi_1, \eta_1)). \end{aligned}$$

*Proof.* The first equality is a change of variable, thanks to symmetry of the zonal harmonic in its variables and invariance with respect to action of the measure preserving elements of the orthogonal group fixing poles  $\xi_1$  or  $\xi_2$ , that we now detail. By a rotation in one of the spheres, assume  $\xi_1 = \xi_2 = \xi$ . Take a small ball

$$B_{\xi, \eta_1}(a_2^0; \varepsilon_2) = \{a_2 : d(a_2, a_2^0) < \varepsilon_2\} \cap \{a_2 : d(a_2, \xi) = d(\eta_1, \xi)\}.$$

Note  $\{a_2 : d(a_2, \xi) = d(\eta_1, \xi)\} \sim \mathbb{S}^{d-2}$ . Every  $a_2 \in B_{\xi, \eta_1}(a_2^0; \varepsilon_2)$  gives rise to a canonical orthogonal map  $\sigma_{a_2}$  along geodesics in a scaled copy of  $\mathbb{S}^{d-2}$ . Lifted to  $\mathbb{S}^{d-1}$ , these are orthogonal maps fixing  $\xi$ . Let  $\sigma^0$  fix  $\xi$  and map  $a_2^0$  to  $\eta_1$ . Let  $a_1^0 = \sigma^0(\eta_2)$ . We observe that  $\{\sigma^0 \sigma_{a_2}(\eta_2) : a_2 \in B_{\xi, \eta_1}(a_2^0; \varepsilon_2)\} = B_{\xi, \eta_2}(a_1^0; \varepsilon_1)$  with  $\varepsilon_1$  so that

$$\mathbb{P}(d(a_2, a_2^0) < \varepsilon_2 \mid d(\xi, a_2) = d(\xi, \eta_1)) = \mathbb{P}(d(a_1, a_1^0) < \varepsilon_1 \mid d(\xi, a_1) = d(\xi, \eta_2)).$$

Together with the symmetry and the rotation property  $Z_\eta^{(n)}(a) = Z_a^{(n)}(\eta) = Z_{\sigma(a)}^{(n)}(\sigma(\eta))$ , we obtain the first equality.

For fixed  $a_1$ , the function  $Z_{\eta_1}^{(n)}(a_1) = Z_{a_1}^{(n)}(\eta_1)$  is a function harmonic in  $\mathbb{R}^d$ ,  $n$ -homogeneous. These properties are preserved when taking expectation in  $a_1$ . So the expression  $\mathbb{E}(Z_{\eta_1}^{(n)}(a_1) \mid d(\xi_1, a_1) = d(\xi_2, \eta_2))$  remains harmonic (regarded as a function in  $\mathbb{R}^d$ ),  $n$ -homogeneous. From the form  $\mathbb{E}(Z_{\eta_2}^{(n)}(a_2) \mid d(\xi_2, a_2) = d(\xi_1, \eta_1))$  we learn that its restriction to  $\mathbb{S}^{d-1}$  depends only upon  $d(\xi_1, \eta_1)$ . This implies that it is a constant multiple of the zonal harmonic with pole  $\xi_1$ . Exchanging the roles of  $\eta_1$  and  $\eta_2$  gives

$$\mathbb{E}(Z_{\eta_1}^{(n)}(a_1) \mid d(\xi_1, a_1) = d(\xi_2, \eta_2)) = c_n Z_{\xi_1}^{(n)}(\eta_1) Z_{\xi_2}^{(n)}(\eta_2).$$

When assuming the normalization  $Z_\xi^{(n)}(\xi) = 1$  then  $c_n = 1$ .

This is a generalisation of the classical symmetrising of the cosinus sum formula  $1/2(\cos(x+y) + \cos(x-y)) = \cos(x)\cos(y)$ .

*QED*

*Proof.* (of Lemma 6) It is well known that zonal harmonic series have convergence properties when representing smooth zonal functions similar to that of the Fourier transform. For any given  $m$  and sufficiently smooth  $\varphi$  of the type described above, then

$$C_1(\xi_1; \eta_1) = \sum_n \varphi_n Z_{\xi_1}^{(n)}(\eta_1)$$

where the convergence is  $\mathcal{C}^m$ -uniform. The degree of smoothness required for  $\varphi$  to obtain convergence in the  $\mathcal{C}^m$  in the above expression depends upon  $m$  and the dimension  $d$ . For our purpose, we choose  $m \geq d$ .

Let us denote this function's representation of degree  $N$  by a series of zonal harmonics by  $C_1^{(N)}(\xi_1; \eta_1)$ .

$$C_1^{(N)}(\xi_1; \eta_1) = \sum_{n \leq N} \varphi_n Z_{\xi_1}^{(n)}(\eta_1).$$

For every  $\delta > 0$  there exists  $N$  so that we have the estimate

$$\|C_1^{(N)}(\xi_1; \eta_1) - C_1(\xi_1; \eta_1)\|_{\mathcal{C}^m(\mathbb{S}^{d_1-1})} < \delta.$$

In the case of  $i$  copies of spheres, we define  $C_i^{(N)}$  inductively in the same manner as  $C_i$ . Let us for the moment make all dimensions equal using the argument discussed above. So we set

$$\begin{aligned} C_k^{(N)}(\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_k) \\ = \mathbb{E}_{a_{k-1}}(C_{k-1}^{(N)}(\xi_1, \dots, a_{k-1}; \eta_1, \dots, \eta_{k-1}) \mid d(a_{k-1}, \xi_{k-1}) = d(\eta_k, \xi_k)). \end{aligned}$$

We claim the identity

$$C_i^{(N)}(\xi; \eta_1, \eta_2, \dots, \eta_i) = \sum_{n \leq N} \varphi_n \prod_{k=1}^i Z_{\xi_k}^{(n)}(\eta_k). \quad (2)$$

This is trivially true for  $i = 1$ . For  $i > 1$  induct on the number of parameters:

$$\begin{aligned} C_i^{(N)}(\xi; \eta_1, \dots, \eta_i) \\ = \mathbb{E}_{a_{i-1}}(C_{i-1}(\xi_1, \xi_2, \dots, a_{i-1}; \eta_1, \dots, \eta_{i-1}) \mid d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i)) \\ = \mathbb{E}_{a_{i-1}} \left( \sum_{n \leq N} \varphi_n \prod_{k=1}^{i-1} Z_{\xi_k}^{(n)}(\eta_k) \mid d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i) \right) \\ = \sum_{n \leq N} \varphi_n \prod_{k=1}^{i-2} Z_{\xi_k}^{(n)}(\eta_k) \mathbb{E}_{a_{i-1}}(Z_{\xi_{i-1}}^{(n)} \mid d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i)) \\ = \sum_{n \leq N} \varphi_n \prod_{k=1}^i Z_{\xi_k}^{(n)}(\eta_k). \end{aligned}$$

The first equality is the definition of  $C_i^{(N)}$ , the second one is the induction hypothesis and the last an application of Lemma 8.

It follows that neither  $C_i$  nor  $C_i^{(N)}$  depend on the order chosen in their definition and

$$C_i(\xi; \eta_1, \dots, \eta_i) = \sum_n \varphi_n \prod_{k=1}^i Z_{\xi_k}^{(n)}(\eta_k)$$

where the convergence is in  $\mathcal{C}^m$  in each variable.

Next, we study the terms arising in multipliers of the form  $C_i^{(N)}$ . When all dimensions are equal, indeed,  $\prod_{k=1}^i Z_{\xi_k}^{(n)}(\eta_k)$  has the important property that, as a product of  $n$ -homogeneous polynomials, has only terms of the form

$$\prod_{k=1}^i \eta_k^{\alpha_k} = \prod_{k=1}^i \left( \prod_{j_k=1}^d \eta_{k,j_k}^{\alpha_{k,j_k}} \right)$$

where  $\eta_k \in \mathbb{S}^{d-1}$  and  $\alpha_k = (\alpha_{k,j_k})$  are multi-indices with  $|\alpha_k| = \sum_{j_k} \alpha_{k,j_k} = n$  for all  $k$ . This form is inherited by  $C_i^{(N)}$  with varying  $n$ . It shows that  $C_i^{(N)}$  is indeed a polynomial in the variables  $\prod_{k=1}^i \eta_{k,j_k}$ . When the dimensions  $d_k$  are not equal, observe that by restricting back to the original number of variables, we certainly lose harmonicity of the polynomials, but not  $n$ -homogeneity or the required form of our polynomials. *QED*

*Proof.* (of Lemma 7) By Lemma 4 we know that for each parameter  $1 \leq s \leq l$  there exists a tensor product of cones  $\mathbf{D}_s = \bigotimes_{k \in I_s} D(\xi_k, r_k)$  with  $r_s := \sum_{k \in I_s} r_k < \pi/2$  and  $\xi_k \in \Upsilon_k$  and test functions  $f_s$  supported as described in Lemma 4 part (2) so that

$$\|[T_{1,\mathbf{D}_1}, \dots, [T_{l,\mathbf{D}_l}, b] \dots](f)\|_2 \gtrsim \|b\|_{\text{BMO}_{\mathcal{T}}(\mathbb{R}^d)}$$

where  $f = \bigotimes_{s=1}^l f_s$ . We make a remark about the apertures  $r_s$ . Let  $d(\cdot, \cdot)$  denote geodesic distance on  $\mathbb{S}^{d-1}$ , where antipodal points have distance  $\pi$ . Let  $\xi^{(s)}$  be the set of directions of  $\mathbf{D}_s$ . Remember that according to Lemma 4, one component had a specific direction  $\xi_v^{(s)} \in \Upsilon_v$  and possibly large aperture with  $(1 + \tau)r_v^{(s)} < \pi/2$ . Let us choose the other directions arbitrarily but with apertures  $r_v^{(s)}$  small enough so that  $(1 + \tau)(r_v^{(s)} + (i-1)r_v^{(s)}) < \pi/2$ . Now choose an aperture  $r_s < \pi/2$  so that  $(1 + \tau)(r_v^{(s)} + (i-1)r_v^{(s)}) < r_s < \pi/2$ .

Writing  $i_s = |I_s|$ , we find Journé type cone multipliers  $\mathbf{J}_s = C_{i_s}(\xi^{(s)}; \cdot)$  according to the construction above with center  $\xi^{(s)}$  and aperture  $r_s$ .

We are going to observe that  $\mathbf{J}_s \equiv 1$  on  $\text{spt}(\mathbf{D}_s)$  and  $\mathbf{J}_s \equiv -1$  on the Fourier support of  $f_s$ . Let us drop the dependence on  $s$  for the moment. We see in an inductive manner that  $C_i(\xi; \cdot)$  takes the value 1 in a certain  $\ell^1$  ball of radius  $r < \pi/2$  centered at  $\xi$ . We show that

$$\sum_k d(\xi_k, \eta_k) < r \Rightarrow C_i(\xi, \eta_1, \dots, \eta_i) = 1.$$

When  $i = 1$ , the assertion is obviously true:  $d(\xi_1, \eta_1) < r \Rightarrow C_1(\xi_1; \eta_1) = 1$  by the choice of  $\varphi, r$  and definition of  $C_1$ . For  $i > 1$ , we proceed by induction. Assume that  $\sum_{k=1}^{i-1} d(\xi_k, \eta_k) < r$  implies  $C_{i-1}(\xi_1, \dots, \xi_{i-1}; \eta_1, \dots, \eta_{i-1}) = 1$ . Let us assume that  $\sum_{k=1}^i d(\xi_k, \eta_k) < r$ . Remembering the definition of  $C_i(\xi; \cdot)$  we assume  $d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i)$ . By the triangle inequality  $\sum_{k=1}^{i-2} d(\xi_k, \eta_k) +$

$d(a_{i-1}, \eta_{i-1}) \leq \sum_{k=1}^{i-2} d(\xi_k, \eta_k) + d(a_{i-1}, \xi_{i-1}) + d(\xi_{i-1}, \eta_{i-1}) = \sum_{k=1}^i d(\xi_k, \eta_k) < r$ .  
So

$$C_{i-1}(\xi_1, \xi_2, \dots, a_{i-1}; \eta_1, \dots, \eta_{i-1}) = 1$$

for all  $a_{i-1}$  relevant to the conditional expectation in the definition of  $C_i(\boldsymbol{\xi}; \cdot)$ . The statement for  $i$  follows.

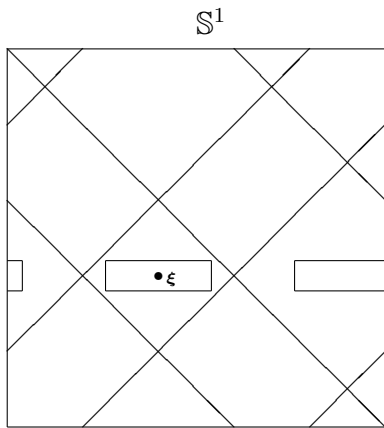
Since  $C_i(\boldsymbol{\xi}; \cdot)$  does not depend upon the order of the variables in its construction, we are also able to see exactly as done above that when  $\sigma_k = -1$  for exactly one choice of  $k$ , then  $\sum_k d(\sigma_k \xi_k, \eta_k) < r \Rightarrow C_i(\boldsymbol{\xi}; \eta_1, \dots, \eta_i) = -1$ .

Consider associated multi-parameter Calderón-Zygmund operators  $\mathbf{T}_{s, \mathbf{J}_s}$  and  $\mathbf{I}d_s = \bigotimes_{k \in I_s} \text{Id}_k$  and  $\text{Id}_k$  the identity on the  $k^{\text{th}}$  variable. Now

$$\begin{aligned} [\mathbf{T}_{1, \mathbf{J}_1}, \dots, [\mathbf{T}_{l, \mathbf{J}_l}, b] \dots](f) &= [\mathbf{T}_{1, \mathbf{J}_1} + \mathbf{I}d_1, \dots, [\mathbf{T}_{l, \mathbf{J}_l} + \mathbf{I}d_l, b] \dots](f) \\ &= \bigotimes_{s=1}^l (\mathbf{T}_{s, \mathbf{J}_s} + \mathbf{I}d_s)(bf) \end{aligned}$$

With  $\|\bigotimes_{s=1}^l (\mathbf{T}_{s, \mathbf{J}_s} + \mathbf{I}d_s)(bf)\|_2 \geq \|\bigotimes_{s=1}^l T_{s, \mathbf{D}_s}(bf)\|_2$  and  $\bigotimes_{s=1}^l T_{s, \mathbf{D}_s}(bf) = [T_{1, \mathbf{D}_1}, \dots, [T_{l, \mathbf{D}_l}, b] \dots](f)$  we get the desired lower bound on the Journé commutator as claimed. QED

Let us illustrate the base case  $(\mathbb{S}^1)^2$  by a picture. The picture is simplified in the sense that the odd function  $\varphi$  above is replaced by an indicator function of an interval.



Cone functions based on the oblique strips containing  $\boldsymbol{\xi}$  are averaged. In the two-dimensional case,  $\mathbb{S}^1$ , expectation is over a one or two point set only. The rectangle around  $\boldsymbol{\xi}$  with sides parallel to the axes representing  $\mathbb{S}^1$  illustrates the support of the tensor product of cone operators with direction  $\boldsymbol{\xi}$ . The longer side is the  $\mathbb{S}^1$  aperture that arises from the Hankel part. The short sides can be chosen freely as they arise from the Toeplitz part and are chosen small so that the rectangle fits into the oblique square. The other small rectangle corresponds to the Fourier support of the test function  $f$ .

*Proof.* (of Theorem 4)

In contrast to the Hilbert transform case, both lower bounds require separate proofs. This is a notable difference that stems from the loss of orthogonal

subspaces in conjunction with the special form of the Hilbert transform only seen in one variable. It does not seem possible to get a lower estimate (3) $\Rightarrow$ (2) directly.

(1) $\Leftrightarrow$ (2). The upper bound (1) $\Rightarrow$ (2) is an easy consequence of the upper estimates of iterated commutators of single Riesz transforms. The lower bound (2) $\Rightarrow$ (1) follows from a standard fact on multipliers in combination with the main result in [18] section 1, the two-sided estimate for iterated commutators with Riesz transforms, similar to the first arguments used in 4.

(1) $\Leftrightarrow$ (3). The upper bound (1) $\Rightarrow$ (3) follows from the tensor product structure and use of the little product BMO norm (see also the remarks in section 7). The lower bound (3) $\Rightarrow$ (1) uses the considerations leading up to this proof: Suppressing again the dependence on  $s$ , we see that the multiplier  $C_i$  is an odd, smooth, bounded function in each  $\eta_k$  when the other variables are fixed. Furthermore, since  $\varphi$ , written as a function of  $t = \langle \xi, \eta \rangle$  is odd with respect to  $t = 0$ , the above series has  $\varphi_n \neq 0$  at most when  $n$  is odd and thus  $Z_\xi^{(n)}$  is odd. So  $C_i^{(N)}$  is as a sum of odd functions odd.

We are now also ready to see that  $\mathbf{T}_J$ , the Journé operator associated to the cone  $J = C_i(\xi; \cdot)$  as well as the operator associated to  $C_i^{(N)}(\xi; \cdot)$  are paraproduct free. In fact, applied to a test function  $f = \otimes_k f_k$  with  $f_k$  acting on the  $k^{\text{th}}$  variable and where  $f_l \equiv 1$  for some  $l$  gives  $\mathbf{T}_J(f) = 0$ . To see this, apply the multiplier  $C_i^{(N)}(\xi; \cdot)$  in the  $l$  variable (acting on 1) first, leaving the other Fourier variables fixed. The multiplier function

$$\eta_l \mapsto C_i^{(N)}(\xi; \eta_1, \dots, \eta_i) = \sum_{n \leq N} \varphi_n Z_{\xi_l}^{(n)}(\eta_l) \prod_{k \neq l, k=1}^i Z_{\xi_k}^{(n)}(\eta_k)$$

is, as a sum of odd functions, odd on  $\mathbb{S}^{d_l-1}$ , bounded by 1 and uniformly smooth for all choices of  $\eta_k$  with  $k \neq l$ . As such it gives rise to a paraproduct free convolution type Calderón-Zygmund operator in the  $l$ th variable whose values are multi-parameter multiplier operators.

Due to the convergence properties proved above, the difference

$$C_i(\xi; \cdot) - C_i^{(N)}(\xi; \cdot)$$

gives rise to a paraproduct free Journé operator with Calderón-Zygmund norm depending on  $N$ . This is seen by an application of an appropriate version of the Marcinkiewicz multiplier theorem.

By our stability result on Journé commutators in section 5, Corollary 3, there exist for all  $1 \leq s \leq l$  integers  $N_s$  so that  $C_s^{(N_s)}(\xi_s; \cdot)$  with  $\xi_k \in \Upsilon_k$  is a characterizing set of operators via commutators for  $\text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$ . This is a finite set of possibilities because of the universal choice of the  $r_s$  and finiteness of



the set  $\Upsilon$ . Using the multi-parameter analog of the observation  $[AB, b] = A[B, b] + [A, b]B$  and the special form of the  $C_s^{(N_s)}(\xi; \cdot)$ , leaves us with the desired lower bound: Observe that when  $[AB, b]$  has large  $L^2$  norm then either  $[A, b]$  or  $[B, b]$  has a fair share of the norm. We use this argument finitely many times in a row for operators that are polynomials in tensor products of Riesz transforms  $\bigotimes_{k \in I_s} R_{k, j_k}$ . This finishes  $(\beta) \Rightarrow (1)$ . *QED*

We remark that there are two cases of dimension greater than 1, where the proof simplifies. In the case of arbitrarily many copies of  $\mathbb{R}^2$ , one can work with the multiplicative structure of complex numbers and avoid the symmetrizing procedure to obtain cone functions with the appropriate polynomial approximations. If the dimensions are arbitrary, but only tensor products of two Riesz transforms arise, one can avoid part of the construction above by using the addition formula for zonal harmonics.

## 5 Real variables: upper bounds

In this section, we are interested in upper bounds for commutator norms by means of little product BMO norms of the symbol. In the case of the Hilbert transform, we have seen that these estimates, even in the iterated case, are straightforward. Other streamlined proofs exist for Hilbert or Riesz transforms when considering dyadic shifts of complexity one, see [29], [30] and [19]. When considering more general Calderón-Zygmund operators, the arguments required are more difficult, in each case. The first classical upper bound goes back to [6], where an estimate for one-parameter commutators with convolution type Calderón-Zygmund operators is given. Next, the text [18] includes a technical estimate for the multi-parameter case for such Calderón-Zygmund operators with a high enough degree of smoothness. This smoothness assumption was removed in [8] thanks to an approach using the representation formula for Calderón-Zygmund operators by means of infinite complexity dyadic shifts [15]. This last proof also gives a control on the norm of the commutators which depends on the Calderón-Zygmund norm of the operators themselves, a fact we will employ later. Below, we give an estimate by little product BMO when the Calderón-Zygmund operators are of Journé type and cannot be written as a tensor product. While this estimate is interesting in its own right, remember that it is also essential for our characterization result, the lower estimate, in section 4. The first generation of multi-parameter singular integrals that are not of tensor product type goes back to Fefferman [11] and was generalized by Journé in [16] to the non-convolution type in the framework of his  $T(1)$  theorem in this setting. Much later, Journé's  $T(1)$  theorem was revisited, for example in [22], [27], [28]. See also [23] for some difficulties related to this subject. The references [22] in the bi-parameter case and [28] in the

general multi-parameter case include a representation formula by means of adapted, infinite complexity dyadic shifts. While these representation formulae look complicated, they have a feature very useful to us. ‘Locally’, in a dyadic sense, they look as if they were of tensor product type, a feature we will exploit in the argument below. We start with the simplest bi-parameter case with no iterations and make comments about the generalization at the end of this section.

The class of bi-parameter singular integral operators treated in this section is that of any paraproduct free Journé type operator (not necessarily a tensor product and not necessarily of convolution type) satisfying a certain weak boundedness property, which we define as follows:

**Definition 2** *A continuous linear mapping  $T : C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m) \rightarrow [C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m)]'$  is called a paraproduct free bi-parameter Calderón-Zygmund operator if the following conditions are satisfied:*

1.  *$T$  is a Journé type bi-parameter  $\delta$ -singular integral operator, i.e. there exists a pair  $(K_1, K_2)$  of  $\delta CZ$ - $\delta$ -standard kernels so that, for all  $f_1, g_1 \in C_0^\infty(\mathbb{R}^n)$  and  $f_2, g_2 \in C_0^\infty(\mathbb{R}^m)$ ,*

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2 \rangle = \int f_1(y_1) \langle K_1(x_1, y_1) f_2, g_2 \rangle g_1(x_1) dx_1 dy_1$$

*when  $\text{spt}f_1 \cap \text{spt}g_1 = \emptyset$ ;*

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2 \rangle = \int f_2(y_2) \langle K_2(x_2, y_2) f_1, g_1 \rangle g_2(x_2) dx_2 dy_2$$

*when  $\text{spt}f_2 \cap \text{spt}g_2 = \emptyset$ .*

2.  *$T$  satisfies the weak boundedness property  $|\langle T(\chi_I \otimes \chi_J), \chi_I \otimes \chi_J \rangle| \lesssim |I||J|$ , for any cubes  $I \subset \mathbb{R}^n, J \in \mathbb{R}^m$ .*

3.  *$T$  is paraproduct free in the sense that  $T(1 \otimes \cdot) = T(\cdot \otimes 1) = T^*(1 \otimes \cdot) = T^*(\cdot \otimes 1) = 0$ .*

Recall that a  $\delta CZ$ - $\delta$ -standard kernel is a vector valued standard kernel taking values in the Banach space consisting of all Calderón-Zygmund operators. It is easy to see that an operator defined as above satisfies all the characterizing conditions in Martikainen [22] section 2, hence is  $L^2$  bounded and can be represented as an average of bi-parameter dyadic shift operators together with dyadic paraproducts. Moreover, since  $T$  is paraproduct free, one can conclude from observing the proof of Martikainen’s theorem, that all the dyadic shifts in the representation are cancellative.

The base case from which we pass to the general case below, is the following:

**Theorem 6** *Let  $T$  be a paraproduct free bi-parameter Calderón-Zygmund operator, and  $b$  be a little bmo function, there holds*

$$\|[b, T]\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim \|b\|_{bmo(\mathbb{R}^n \times \mathbb{R}^m)},$$

where the underlying constant depends only on the characterizing constants of  $T$ .

*Proof.*

According to the discussion above, for any sufficiently nice functions  $f, g$ , one has the following representation:

$$\langle Tf, g \rangle = C \mathbb{E}_{\omega_1} \mathbb{E}_{\omega_2} \sum_{i_1, j_1=0}^{\infty} \sum_{i_2, j_2=0}^{\infty} 2^{-\max(i_1, j_1)} 2^{-\max(i_2, j_2)} \langle S^{i_1 j_1 i_2 j_2} f, g \rangle, \quad (3)$$

where expectation is with respect to a certain parameter of the dyadic grids. The dyadic shifts  $S^{i_1 j_1 i_2 j_2}$  are defined as

$$\begin{aligned} & S^{i_1 j_1 i_2 j_2} f \\ & := \sum_{\substack{K_1 \in \mathcal{D}_1 \\ \ell(I_1)=2^{-i_1} \ell(K_1) \\ \ell(J_1)=2^{-j_1} \ell(K_1)}} \sum_{\substack{I_1, J_1 \subset K_1, I_1, J_1 \in \mathcal{D}_1}} \sum_{\substack{K_2 \in \mathcal{D}_2 \\ \ell(I_2)=2^{-i_2} \ell(K_2) \\ \ell(J_2)=2^{-j_2} \ell(K_2)}} \sum_{\substack{I_2, J_2 \subset K_2, I_2, J_2 \in \mathcal{D}_2}} a_{I_1 J_1 K_1 I_2 J_2 K_2} \langle f, h_{I_1} \otimes h_{I_2} \rangle h_{J_1} \otimes h_{J_2} \\ & = \sum_{K_1} \sum_{\substack{(i_1, j_1) \\ I_1, J_1 \subset K_1}} \sum_{K_2} \sum_{\substack{(i_2, j_2) \\ I_2, J_2 \subset K_2}} a_{I_1 J_1 K_1 I_2 J_2 K_2} \langle f, h_{I_1} \otimes h_{I_2} \rangle h_{J_1} \otimes h_{J_2}. \end{aligned}$$

The coefficients above satisfy  $a_{I_1 J_1 K_1 I_2 J_2 K_2} \leq \frac{\sqrt{|I_1| |J_1| |I_2| |J_2|}}{|K_1| |K_2|}$ , which also guarantees the normalization  $\|S^{i_1 j_1 i_2 j_2}\|_{L^2 \rightarrow L^2} \leq 1$ . Moreover, since  $T$  is paraproduct free, all the Haar functions appearing above are cancellative.

It thus suffices to show that for any dyadic grids  $\mathcal{D}_1, \mathcal{D}_2$  and fixed  $i_1, j_1, i_2, j_2 \in \mathbb{N}$ , one has

$$\|[b, S^{i_1 j_1 i_2 j_2}]f\|_{L^2} \lesssim (1 + \max(i_1, j_1))(1 + \max(i_2, j_2)) \|b\|_{bmo} \|f\|_{L^2}, \quad (4)$$

as the decay factor  $2^{-\max(i_1, j_1)}, 2^{-\max(i_2, j_2)}$  in (3) will guarantee the convergence of the series.

To see (4), one decomposes  $b$  and a  $L^2$  test function  $f$  using Haar bases:

$$[b, S^{i_1 j_1 i_2 j_2}]f = \sum_{I_1, I_2} \sum_{J_1, J_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1} \otimes h_{I_2}, S^{i_1 j_1 i_2 j_2}] h_{J_1} \otimes h_{J_2}.$$

A similar argument to that in [8] subsection 3.1 implies that  $[h_{I_1} \otimes h_{I_2}, S^{i_1 j_1 i_2 j_2}] h_{J_1} \otimes$

$h_{J_2}$  is nonzero only if  $I_1 \subset J_1^{(i_1)}$  or  $I_2 \subset J_2^{(i_2)}$ , where  $J_1^{(i_1)}$  denotes the  $i_1$ -th dyadic ancestor of  $J_1$ , similarly for  $J_2^{(i_2)}$ . Hence, the sum can be decomposed into three parts:  $I_1 \subset J_1^{(i_1)}$  and  $I_2 \subset J_2^{(i_2)}$  (regular),  $I_1 \subset J_1^{(i_1)}$  and  $I_2 \not\subset J_2^{(i_2)}$ ,  $I_1 \not\subset J_1^{(i_1)}$  and  $I_2 \subset J_2^{(i_2)}$  (mixed).

1) *Regular case:*

Following [8] subsection 4.1 one can decompose the arising sum into sums of classical bi-parameter dyadic paraproducts  $B_0(b, f)$  and its slightly revised version  $B_{kl}(b, f)$ : for any integers  $k, l \geq 0$ ,  $B_{kl}$  is the bi-parameter dyadic paraproduct defined as

$$B_{kl}(b, f) = \sum_{I, J} \beta_{IJ} \langle b, h_{I^{(k)}} \otimes u_{J^{(l)}} \rangle \langle f, h_I^{\varepsilon_1} \otimes u_J^{\varepsilon_2} \rangle h_I^{\varepsilon_1'} \otimes u_J^{\varepsilon_2'} |I^{(k)}|^{-1/2} |J^{(l)}|^{-1/2},$$

where  $\beta_{IJ}$  is a sequence satisfying  $|\beta_{IJ}| \leq 1$ . When  $k > 0$ , all Haar functions in the first variable are cancellative, while when  $k = 0$ , there is at most one of  $h_I^{\varepsilon_1}, h_I^{\varepsilon_1'}$  being noncancellative. The same assumption goes for the second variable. Observe that when  $k = l = 0$ ,  $B_{kl}$  becomes the classical paraproduct  $B_0$ . It is proved in [8] in Lemma 4.1 that

$$\|B_{kl}(b, f)\|_{L^2} \lesssim \|b\|_{\text{BMO}} \|f\|_{L^2}$$

with a constant independent of  $k, l$  and the product BMO norm on the right hand side.

Then since little bmo functions are contained in product BMO, this part can be controlled. More specifically, write

$$\begin{aligned} [b, S^{i_1 j_1 i_2 j_2}] f &= \sum_{I_1, I_2} \sum_{J_1, J_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{I_1} \otimes h_{I_2} S^{i_1 j_1 i_2 j_2} (h_{J_1} \otimes h_{J_2}) \\ &\quad - \sum_{I_1, I_2} \sum_{J_1, J_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle S^{i_1 j_1 i_2 j_2} (h_{I_1} h_{J_1} \otimes h_{I_2} h_{J_2}) \\ &=: I + II, \end{aligned}$$

then one can estimate term I and II separately. According to the definition of

dyadic shifts, term I is equal to

$$\begin{aligned}
& \sum_{J_1, J_2} \sum_{I_1: I_1 \subset J_1^{(i_1)}} \sum_{I_2: I_2 \subset J_2^{(i_2)}} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{I_1} \otimes h_{I_2} \cdot \\
& \quad \left( \sum_{\substack{J'_1: J'_1 \subset J_1^{(i_1)} \\ \ell(J'_1) = 2^{i_1 - j_1} \ell(J_1)}} \sum_{\substack{J'_2: J'_2 \subset J_2^{(i_2)} \\ \ell(J'_2) = 2^{i_2 - j_2} \ell(J_2)}} a_{J_1 J'_1 J_1^{(i_1)} J_2 J'_2 J_2^{(i_2)}} h_{J'_1} \otimes h_{J'_2} \right) \\
& = \sum_{K_1, K_2} \sum_{J_1: J_1 \subset K_1}^{(i_1)} \sum_{J_2: J_2 \subset K_2}^{(i_2)} \sum_{I_1: I_1 \subset K_1} \sum_{I_2: I_2 \subset K_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{I_1} \otimes h_{I_2} \cdot \\
& \quad \left( \sum_{J'_1: J'_1 \subset K_1}^{(j_1)} \sum_{J'_2: J'_2 \subset K_2}^{(j_2)} a_{J_1 J'_1 K_1 J_2 J'_2 K_2} h_{J'_1} \otimes h_{J'_2} \right) \\
& = \sum_{I_1, I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle h_{I_1} \otimes h_{I_2} \sum_{\substack{K_1 \supset I_1 \\ K_2 \supset I_2}} \sum_{J_1, J'_1 \subset K_1}^{(i_1, j_1)} \sum_{J_2, J'_2 \subset K_2}^{(i_2, j_2)} a_{J_1 J'_1 K_1 J_2 J'_2 K_2} \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{J'_1} \otimes h_{J'_2} \\
& = \sum_{I_1, I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle h_{I_1} \otimes h_{I_2} \sum_{J'_1: J'_1 \supset I_1} \sum_{J'_2: J'_2 \supset I_2} \langle S^{i_1 j_1 i_2 j_2} f, h_{J'_1} \otimes h_{J'_2} \rangle h_{J'_1} \otimes h_{J'_2}.
\end{aligned}$$

Because of the supports of Haar functions, the inner sum above can be further decomposed into four parts, where

$$\begin{aligned}
I &= \sum_{I_1, I_2} \sum_{J'_1 \supseteq I_1} \sum_{J'_2 \supseteq I_2}, \quad II = \sum_{I_1, I_2} \sum_{J'_1 \supseteq I_1} \sum_{J'_2: J'_2 \subset I_2 \subset J_2^{(j_2)}} \\
III &= \sum_{I_1, I_2} \sum_{J'_1: J'_1 \subset I_1 \subset J_1^{(j_1)}} \sum_{J'_2 \supseteq I_2}, \quad IV = \sum_{I_1, I_2} \sum_{J'_1: J'_1 \subset I_1 \subset J_1^{(j_1)}} \sum_{J'_2: J'_2 \subset I_2 \subset J_2^{(j_2)}}.
\end{aligned}$$

Hence, using the same technique as in [8], one has

$$I = \sum_{I_1, I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle S^{i_1 j_1 i_2 j_2} f, h_{J'_1}^1 \otimes h_{J'_2}^1 \rangle h_{I_1} \otimes h_{I_2} |I_1|^{-1/2} |I_2|^{-1/2},$$

which is a bi-parameter paraproduct  $B_0(b, f)$ . Moreover, one has

$$\begin{aligned}
II &= \sum_{I_1, I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle h_{I_1} \otimes h_{I_2} \sum_{J'_2: J'_2 \subset I_2 \subset J_2^{(j_2)}} \langle S^{i_1 j_1 i_2 j_2} f, h_{I_1}^1 \otimes h_{J'_2} \rangle |I_1|^{-1/2} h_{J'_2} \\
&= \sum_{l=0}^{j_2} \sum_{I_1, J'_2} \beta_{J'_2} \langle b, h_{I_1} \otimes h_{J_2^{(l)}} \rangle \langle S^{i_1 j_1 i_2 j_2} f, h_{I_1}^1 \otimes h_{J'_2} \rangle h_{I_1} \otimes h_{J'_2} |I_1|^{-1/2} |J_2^{(l)}|^{-1/2} \\
&= \sum_{l=0}^{j_2} B_{0l}(b, S^{i_1 j_1 i_2 j_2} f),
\end{aligned}$$

where constants  $\beta_{J'_2} \in \{1, -1\}$ , and  $B_{0l}$  are the generalized bi-parameter paraproducts of type  $(0, l)$  whose  $L^2 \rightarrow L^2$  operator norm is uniformly bounded

by  $\|b\|_{\text{BMO}}$  product BMO. Similarly, one can show that

$$III = \sum_{k=0}^{j_1} B_{k0}(b, S^{i_1 j_1 i_2 j_2} f), \quad IV = \sum_{k=0}^{j_1} \sum_{l=0}^{j_2} B_{kl}(b, S^{i_1 j_1 i_2 j_2} f).$$

Since  $\|b\|_{\text{BMO}} \lesssim \|b\|_{\text{bmo}}$ , all the forms above are  $L^2$  bounded. This completes the discussion of term I.

To get an estimate of term II, we need to decompose it into finite linear combinations of  $S^{i_1 j_1 i_2 j_2}(B_{kl}(b, f))$ . By linearity, one can write  $S^{i_1 j_1 i_2 j_2}$  on the outside from the beginning, and we will only look at the inside sum. One splits for example the sum regarding the first variable into three parts:  $I_1 \subsetneq J_1$ ,  $I_1 = J_1$ ,  $J_1 \subsetneq I_1 \subset J_1^{(i_1)}$ . If we split the second variable as well, there are nine mixed parts, and it's not hard to show that each of them can be represented as a finite sum of  $B_{kl}(b, f)$ . We omit the details.

2) *Mixed case.* Let's call the second and the third 'mixed' parts, and as the two are symmetric, it suffices to look at the second one, i.e.  $I_1 \subset J_1^{(i_1)}$ ,  $I_2 \supsetneq J_2^{(i_2)}$ . In the first variable, we still have the old case  $I_1 \subset J_1^{(i_1)}$  that appeared in [8] subsection 3.1, so morally speaking, we only need to nicely play around with the stronger little bmo norm to handle the second variable. For any fixed  $I_1, J_1, I_2, J_2$ , since  $I_2 \supsetneq J_2^{(i_2)}$ , the definition of dyadic shifts implies that

$$h_{I_1} \otimes h_{I_2} S^{i_1 j_1 i_2 j_2}(h_{J_1} \otimes h_{J_2}) = h_{I_1} S^{i_1 j_1 i_2 j_2}(h_{J_1} \otimes h_{I_2} h_{J_2})$$

and

$$S^{i_1 j_1 i_2 j_2}(h_{I_1} h_{J_1} \otimes h_{I_2} h_{J_2}) = h_{I_2} S^{i_1 j_1 i_2 j_2}(h_{I_1} h_{J_1} \otimes h_{J_2}).$$

Hence, we still have cancellation in the second variable, which converts the mixed case to

$$\begin{aligned} & \sum_{I_1 \subset J_1^{(i_1)}} \sum_{I_2 \supsetneq J_2^{(i_2)}} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1}, S^{i_1 j_1 i_2 j_2}](h_{J_1} \otimes h_{I_2} h_{J_2}) \\ &= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1}, S^{i_1 j_1 i_2 j_2}](h_{J_1} \otimes \sum_{I_2 \supsetneq J_2^{(i_2)}} \langle b, h_{I_1} \otimes h_{I_2} \rangle h_{I_2} h_{J_2}) \\ &= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1}, S^{i_1 j_1 i_2 j_2}](h_{J_1} \otimes \langle b, h_{I_1} \otimes h_{J_2^{(i_2)}}^1 \rangle h_{J_2^{(i_2)}}^1 h_{J_2}) \\ &= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle b, h_{I_1} \otimes h_{J_2^{(i_2)}}^1 \rangle |J_2^{(i_2)}|^{-1/2} \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1}, S^{i_1 j_1 i_2 j_2}](h_{J_1} \otimes h_{J_2}) \\ &= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \langle f, h_{J_1} \otimes h_{J_2} \rangle [h_{I_1}, S^{i_1 j_1 i_2 j_2}](h_{J_1} \otimes h_{J_2}), \end{aligned}$$

where  $\langle b \rangle_{J_2^{(i_2)}}$  denotes the average value of  $b$  on  $J_2^{(i_2)}$ , which is a function of only the first variable.

In the following, we will once again estimate the first term and second term of the commutator separately, and the  $L^2$  norm of each of them will be proved to be bounded by  $\|b\|_{\text{bmo}}\|f\|_{L^2}$ .

a) First term.

By definition of the dyadic shift, the first term is equal to

$$\sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 h_{I_1} \langle f, h_{J_1} \otimes h_{J_2} \rangle \cdot \left( \sum_{\substack{J_1' \subset J_1^{(i_1)} \\ \ell(J_1') = 2^{i_1 - j_1} \ell(J_1)}} \sum_{\substack{J_2' \subset J_2^{(i_2)} \\ \ell(J_2') = 2^{i_2 - j_2} \ell(J_2)}} a_{J_1 J_1' J_1^{(i_1)} J_2 J_2' J_2^{(i_2)}} h_{J_1'} \otimes h_{J_2'} \right),$$

which by reindexing  $K_1 := J_1^{(i_1)}$  is the same as

$$\begin{aligned} & \sum_{I_1, J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 h_{I_1} \cdot \\ & \cdot \sum_{K_1: K_1 \supset I_1} \sum_{J_1 \subset K_1}^{(i_1)} \sum_{J_1' \subset K_1}^{(j_1)} \sum_{J_2' \subset J_2^{(i_2)}}^{(j_2)} a_{J_1 J_1' K_1 J_2 J_2' J_2^{(i_2)}} \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{J_1'} \otimes h_{J_2'} \\ & = \sum_{I_1, J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 h_{I_1} \sum_{J_1': J_1'^{(j_1)} \supset I_1} h_{J_1'} \otimes \langle S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J_1'} \rangle_1, \end{aligned}$$

where the inner sum is the orthogonal projection of the image of  $\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}$  under  $S^{i_1 j_1 i_2 j_2}$  onto the span of  $\{h_{J_1'}\}$  such that  $J_1'^{(j_1)} \supset I_1$ . Taking into account the supports of the Haar functions in the first variable, one can further split the sum into two parts where

$$I := \sum_{J_2} \sum_{I_1 \not\subset J_1'} , \quad II := \sum_{J_2} \sum_{J_1' \subset I_1 \subset J_1'^{(j_1)}} .$$

Summing over  $J_1'$  first implies that

$$\begin{aligned} I & = \sum_{J_2} \sum_{I_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 h_{I_1} (h_{I_1}^1 \otimes \langle S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{I_1}^1 \rangle_1) \\ & =: \sum_{J_2} B_0(\langle b \rangle_{J_2^{(i_2)}}, S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2})) \end{aligned}$$

where  $B_0(b, f) := \sum_I \langle b, h_I \rangle \langle f, h_I^1 \rangle h_I |I|^{-1/2}$  is a classical one-parameter para-product in the first variable. Note that its  $L^2$  norm is bounded by  $\|b\|_{\text{BMO}}\|f\|_{L^2}$ .

Moreover, according to the definition of  $S^{i_1 j_1 i_2 j_2}$ , for any fixed  $J_2$

$$S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}) = \sum_{J'_2: J_2^{(i_2)} = J_2^{(i_2)}} \langle S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2 \otimes h_{J'_2}.$$

In other words,  $S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2})$  only lives on the span of  $\{h_{J'_2} : J_2^{(j_2)} = J_2^{(i_2)}\}$ . Hence, by linearity there holds

$$\begin{aligned} I &= \sum_{J_2} \sum_{J'_2: J_2^{(j_2)} = J_2^{(i_2)}} B_0(\langle b \rangle_{J_2^{(i_2)}}, \langle S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2) \otimes h_{J'_2} \\ &= \sum_{J'_2} \left( B_0(\langle b \rangle_{J_2^{(j_2)}}, \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2) \right) \otimes h_{J'_2}. \end{aligned}$$

Thus, orthogonality in the second variable implies that

$$\begin{aligned} &\|I\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 \\ &= \sum_{J'_2} \|B_0(\langle b \rangle_{J_2^{(j_2)}}, \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2)\|_{L^2(\mathbb{R}^n)}^2 \\ &\lesssim \sum_{J'_2} \|\langle b \rangle_{J_2^{(j_2)}}\|_{\text{BMO}(\mathbb{R}^n)}^2 \|\langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2\|_{L^2(\mathbb{R}^n)}^2. \end{aligned}$$

Observing that  $\|\langle b \rangle_{J_2^{(j_2)}}\|_{\text{BMO}(\mathbb{R}^n)} \leq \|\langle b \rangle_{\text{BMO}(\mathbb{R}^n)}\|_{J_2^{(j_2)}} \leq \|b\|_{\text{bmo}}$ , one has

$$\begin{aligned} &\leq \|b\|_{\text{bmo}}^2 \sum_{J'_2} \|\langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2\|_{L^2(\mathbb{R}^n)}^2 \\ &= \|b\|_{\text{bmo}}^2 \sum_{J'_2} \|\langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2 \otimes h_{J'_2}\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \end{aligned}$$

Note that the sum in the  $L^2$  norm is in fact very simple:

$$\begin{aligned} &\sum_{J'_2} \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2: J_2^{(i_2)} = J_2^{(j_2)}} \langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2 \otimes h_{J'_2} \\ &= \sum_{J_2} \sum_{J'_2: J_2^{(j_2)} = J_2^{(i_2)}} \langle S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_2} \rangle_2 \otimes h_{J'_2} \\ &= \sum_{J_2} S^{i_1 j_1 i_2 j_2}(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}) = S^{i_1 j_1 i_2 j_2}(f). \end{aligned}$$

Hence, the uniform boundedness of the  $L^2 \rightarrow L^2$  operator norm of dyadic shifts implies that

$$\|I\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 \lesssim \|b\|_{\text{bmo}}^2 \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.$$

In order to handle II, we split it into a finite sum depending on the levels of



$I_1$  upon  $J'_1$ , which leads to

$$\begin{aligned}
II &= \sum_{k=0}^{j_1} \sum_{J_2} \sum_{J'_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}} \rangle_1 h_{J_1^{(k)}} h_{J'_1} \otimes \langle S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_1} \rangle_1 \\
&= \sum_{k=0}^{j_1} \sum_{J_2} \sum_{J'_1} \beta_{J'_1, k} |J_1^{(k)}|^{-1/2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}} \rangle_1 h_{J'_1} \otimes \langle S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}), h_{J'_1} \rangle_1 \\
&=: \sum_{k=0}^{j_1} \sum_{J_2} B_k(\langle b \rangle_{J_2^{(i_2)}}, S^{i_1 j_1 i_2 j_2} (\langle f, h_{J_2} \rangle_2 \otimes h_{J_2})),
\end{aligned}$$

where  $B_k(b, f) := \sum_I \beta_{I, k} \langle b, h_{I^{(k)}} \rangle \langle f, h_I \rangle h_I |I^{(k)}|^{-1/2}$  is a generalized one-parameter paraproduct studied in [8] Lemma 3.6, whose  $L^2$  norm is uniformly bounded by  $\|b\|_{\text{BMO}} \|f\|_{L^2}$ , independent of  $k$  and the coefficients  $\beta_{I, k} \in \{1, -1\}$ . Then one can proceed as in part I to conclude that

$$\|II\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim (1 + j_1) \|b\|_{\text{bmo}} \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)},$$

which together with the estimate for part I implies that

$$\|\text{First term}\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim (1 + j_1) \|b\|_{\text{bmo}} \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}.$$

b) Second term.

As the second term by linearity is the same as

$$S^{i_1 j_1 i_2 j_2} \left( \sum_{J_2} \sum_{I_1 \subset J_1^{(i_1)}} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{I_1} h_{J_1} \otimes h_{J_2} \right),$$

the  $L^2 \rightarrow L^2$  boundedness of the shift implies that it suffices to estimate the  $L^2$  norm of the term inside the parentheses. Since  $I_1 \cap J_1 \neq \emptyset$ , one can further split the sum into two parts:

$$I := \sum_{J_2} \sum_{I_1 \not\subset J_1}, \quad II := \sum_{J_2} \sum_{J_1 \subset I_1 \subset J_1^{(i_1)}}.$$

Summing over  $J_1$  first implies that

$$\begin{aligned}
I &= \sum_{J_2} \sum_{I_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \langle f, h_{I_1}^1 \otimes h_{J_2} \rangle h_{I_1} h_{I_1}^1 \otimes h_{J_2} \\
&=: \sum_{J_2} B_0(\langle b \rangle_{J_2^{(i_2)}}, \langle f, h_{J_2} \rangle_2) \otimes h_{J_2},
\end{aligned}$$

where  $B_0(b, f) := \sum_I \langle b, h_I \rangle \langle f, h_I^1 \rangle h_I |I|^{-1/2}$  is a classical one-parameter para-

product in the first variable. Hence,

$$\begin{aligned}
\|I\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 &= \sum_{J_2} \|B_0(\langle b \rangle_{J_2^{(i_2)}}, \langle f, h_{J_2} \rangle_2)\|_{L^2(\mathbb{R}^n)}^2 \\
&\lesssim \sum_{J_2} \|\langle b \rangle_{J_2^{(i_2)}}\|_{\text{BMO}(\mathbb{R}^n)}^2 \|\langle f, h_{J_2} \rangle_2\|_{L^2(\mathbb{R}^n)}^2 \\
&\leq \|b\|_{\text{bmo}}^2 \sum_{J_2} \|\langle f, h_{J_2} \rangle_2\|_{L^2(\mathbb{R}^n)}^2 = \|b\|_{\text{bmo}}^2 \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\end{aligned}$$

For part II, note that it can be decomposed as

$$\begin{aligned}
II &= \sum_{k=0}^{i_1} \sum_{J_2} \sum_{J_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}} \rangle_1 \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{J_1^{(k)}} h_{J_1} \otimes h_{J_2} \\
&= \sum_{k=0}^{i_1} \sum_{J_2} \sum_{J_1} \beta_{J_1, k} |J_1^{(k)}|^{-1/2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}} \rangle_1 \langle \langle f, h_{J_2} \rangle_2, h_{J_1} \rangle_1 h_{J_1} \otimes h_{J_2} \\
&=: \sum_{k=0}^{i_1} \sum_{J_2} B_k(\langle b \rangle_{J_2^{(i_2)}}, \langle f, h_{J_2} \rangle_2) \otimes h_{J_2},
\end{aligned}$$

where coefficients  $\beta_{J_1, k} \in \{1, -1\}$  and the  $L^2$  norm of the generalized para-product  $B_k$  is uniformly bounded as mentioned before. Therefore, the same argument as for part I shows that

$$\|II\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim (1 + i_1) \|b\|_{\text{bmo}} \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)},$$

which completes the discussion of the second term, and thus proves that the mixed case is bounded. QED

The upper bound result we just proved can be extended to  $\mathbb{R}^d$ , to arbitrarily many parameters and an arbitrary number of iterates in the commutator. To do this, consider multi-parameter singular integral operators studied in [28], which satisfy a weak boundedness property and are paraproduct free, meaning that any partial adjoint of  $T$  is zero if acting on some tensor product of functions with one of the components being 1. And consider a little product BMO function  $b \in \text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$ . One can then prove

**Theorem 7** *Let us consider  $\mathbb{R}^d$ ,  $\mathbf{d} = (d_1, \dots, d_t)$  with a partition  $\mathcal{I} = (I_s)_{1 \leq s \leq t}$  of  $\{1, \dots, t\}$  as discussed before. Let  $b \in \text{BMO}_{\mathcal{I}}(\mathbb{R}^d)$  and let  $T_s$  denote a multi-parameter paraproduct free Journé operator acting on functions defined on  $\bigotimes_{k \in I_s} \mathbb{R}^{d_k}$ . Then we have the estimate below*

$$\|[T_1, \dots, [T_l, b] \dots]\|_{L^2(\mathbb{R}^d) \ominus} \lesssim \|b\|_{\text{BMO}_{\mathcal{I}}(\mathbb{R}^d)}.$$

The part of the proof that targets the Journé operators proceeds exactly the same as the bi-parameter case with the multi-parameter version of the rep-

resentation theorem proven in [28]. Certainly, as the number of parameters increases, more mixed cases will appear. However, if one follows the corresponding argument above for each variable in each case, it is not hard to check that eventually, the boundedness of the arising paraproducts is implied exactly by the little product BMO norm of the symbol. The difficulty of higher iterates is overcome in observing that the commutator splits into commutators with no iterates, as was done in [8] subsection 4.1. We omit the details.

The assumption that the operators be paraproduct free is sufficient for our lower estimate. The general case is currently under investigation by one of the authors. Important to our arguments for lower bounds with Riesz transforms is the corollary below, which follows from the control on the norm of the estimate in Theorem 7 by an application of triangle inequality. It is a stability result for characterizing families of Journé operators.

**Corollary 3** *Let for every  $1 \leq s \leq l$  be given a collection  $\mathcal{T}_s = \{T_{s,j_s}\}$  of paraproduct free Journé operators on  $\bigotimes_{k \in I_s} \mathbb{R}^{d_k}$  that characterize  $BMO_{\mathcal{I}}(\mathbb{R}^d)$  via a two-sided commutator estimate*

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)} \lesssim \sup_{\mathbf{j}} \|[T_{1,j_1}, \dots, [T_{l,j_l}, b] \dots]\|_{L^2(\mathbb{R}^d)} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^d)}.$$

*Then there exists  $\varepsilon > 0$  such that for any family of paraproduct free Journé operators  $\mathcal{T}'_s = \{T'_{s,j_s}\}$  with characterizing constants  $\|T'_{s,j_s}\|_{CZ} \leq \varepsilon$ , the family  $\{T_{s,j_s} + T'_{s,j_s}\}$  still characterizes  $BMO_{\mathcal{I}}(\mathbb{R}^d)$ .*

## 6 Weak Factorization

It is well known, that theorems of this form have an equivalent formulation in the language of weak factorization of Hardy spaces. We treat the model case  $\mathbb{R}^d = \mathbb{R}^{(d_1, d_2, d_3)}$  and  $BMO_{(13)(2)}(\mathbb{R}^d)$  only for sake of simplicity. The other statements are an obvious generalization. For the corresponding collections of Riesz transforms  $\mathcal{R}_{k,j_k}$  and  $b \in BMO_{(13)(2)}(\mathbb{R}^d)$ ,  $1 \leq s \leq 3$ , by unwinding the commutator one can define the operator  $\Pi_{\mathbf{j}}$  such that

$$\langle [R_{2,j_2}, [R_{1,j_1} R_{3,j_3}, b]]f, g \rangle_{L^2} = \langle b, \Pi_{\mathbf{j}}(f, g) \rangle_{L^2}.$$

Consider the Banach space  $L^2 * L^2$  of all functions in  $L^1(\mathbb{R}^d)$  of the form  $f = \sum_{\mathbf{j}} \sum_i \Pi_{\mathbf{j}}(\phi_i^{\mathbf{j}}, \psi_i^{\mathbf{j}})$  normed by

$$\|f\|_{L^2 * L^2} = \inf \left\{ \sum_{\mathbf{j}} \sum_i \|\phi_i^{\mathbf{j}}\|_2 \|\psi_i^{\mathbf{j}}\|_2 \right\}$$

with the infimum running over all possible decompositions of  $f$ . Applying a

duality argument and the two-sided estimate in Corollary 2 we are going to prove the following weak factorization theorem.

**Theorem 8**  $H_{\text{Re}}^1(\mathbb{R}^{(d_1, d_2)}) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H_{\text{Re}}^1(\mathbb{R}^{(d_2, d_3)})$  coincides with the space  $L^2 * L^2$ . In other words, for any  $f \in H_{\text{Re}}^1(\mathbb{R}^{(d_1, d_2)}) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H_{\text{Re}}^1(\mathbb{R}^{(d_2, d_3)})$  there exist sequences  $\phi_i^j, \psi_i^j \in L^2$  such that  $f = \sum_j \sum_i \Pi_j(\phi_i^j, \psi_i^j)$  and  $\|f\| \sim \sum_j \sum_i \|\phi_i^j\|_2 \|\psi_i^j\|_2$ .

*Proof.* Let's first show that  $L^2 * L^2$  is a subspace of  $H_{\text{Re}}^1(\mathbb{R}^{(d_1, d_2)}) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H_{\text{Re}}^1(\mathbb{R}^{(d_2, d_3)})$ . Recalling the remark after Theorem 2, this is the same as to show  $\forall f \in L^2 * L^2$ ,  $f$  is a bounded linear functional on  $\text{BMO}_{(13)(2)}(\mathbb{R}^d)$ . This follows from the upper bound on the commutators since

$$\langle b, \sum_j \sum_i \Pi_j(\phi_i^j, \psi_i^j) \rangle = \sum_j \sum_i \langle [R_{2, j_2}, [R_{1, j_1} R_{3, j_3}, b]] \phi_i^j, \psi_i^j \rangle.$$

Now we are going to show

$$\sup_{f \in L^2 * L^2} \left\{ \left| \int f b \right| : \|f\|_{L^2 * L^2} \leq 1 \right\} \sim \|b\|_{\text{BMO}_{(13)(2)}}$$

which gives the equivalence of  $H_{\text{Re}}^1(\mathbb{R}^{(d_1, d_2)}) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H_{\text{Re}}^1(\mathbb{R}^{(d_2, d_3)})$  norm and the  $L^2 * L^2$  norm, thus showing that the two spaces are the same.

To see this, note that the direction  $\lesssim$  is trivial, and the direction  $\gtrsim$  is implied by the lower bound of commutators. For any  $b \in \text{BMO}_{(13)(2)}$ , there exists  $\mathbf{j}$  such that  $\|b\|_{\text{BMO}_{(13)(2)}} \lesssim \|[R_{2, j_2}, [R_{1, j_1} R_{3, j_3}, b]]\|$ . Hence, there exist  $\phi, \psi \in L^2$  with norm 1 such that

$$\|b\|_{\text{BMO}_{(13)(2)}} \lesssim |\langle [R_{2, j_2}, [R_{1, j_1} R_{3, j_3}, b]] \phi, \psi \rangle| = |\langle b, \Pi_{\mathbf{j}}(\phi, \psi) \rangle| \leq LHS,$$

which completes the proof. QED

## 7 Remarks about our results in $L^p$

As mentioned before, the two-sided estimates stated in section 4 and in particular Theorem 5 hold for all  $1 < p < \infty$ . The fact that upper estimates hold for the Riesz commutator in  $L^p$  in the case where no tensor products are present is proved in [18] as well as [19]. It stems from the fact that endpoint estimates for multi-parameter paraproducts hold for all  $1 < p < \infty$  [24], [25]. This estimate carries over easily to tensor products of Riesz transforms or any other tensor products of operators for which we have  $L^p$  estimates on the commutator: one uses  $[T_1 T_2, b] = T_1 [T_2, b] + [T_1, b] T_2$  to handle arising tensor products, followed

by a correct use of the little product BMO norm. The argument is left as an exercise.

The lower estimate or the necessity of the BMO condition can be derived from interpolation. In fact, suppose we have uniform boundedness of our commutators with operators running through all choices of Riesz transforms and some symbol  $b$  in  $L^p$ . Then by duality, we have boundedness in  $L^q$  where  $1/p + 1/q = 1$ . In fact,  $[T, b]^* f = -[T^*, \bar{b}]f = -\overline{[T^*, b]f}$  shows that the boundedness of adjoints is inherited. The same reasoning holds for iterated commutators of tensor products. Thus by interpolation, the boundedness holds in  $L^2$  and the symbol function  $b$  necessarily belongs to the required BMO class.

## References

- [1] Arlen Brown and Paul Halmos, *Algebraic properties of Toeplitz operators*, J. Reine Angew. Math. **213** (1964), 89–102. ↑
- [2] Lennart Carleson, *A counterexample for measures bounded on  $H^p$  spaces for the bidisk*, Mittag-Leffler Rep. No. 7, Inst. Mittag-Leffler (1974). ↑5
- [3] Ronald Coifman, Pierre-Louis Lions, Yves Meyer, and Stephen Semmes, *Compensated compactness and Hardy space*, J. Math. Pures Appl. **72** (1993), no. 9, 247–286. ↑3
- [4] Sun-Yung Chang and Robert Fefferman, *Some recent developments in Fourier analysis and  $H^p$ -theory on product domains*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), no. 1, 1–43. MR 86g:42038 ↑5
- [5] ———, *A continuous version of duality of  $H^1$  with BMO on the bidisc*, Ann. of Math. (2) **112** (1980), no. 1, 179–201. ↑5
- [6] Ronald Coifman, Richard Rochberg, and Guido Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635. ↑2, 24
- [7] Misha Cotlar and Cora Sadosky, *The Helson-Szegő theorem in  $L^p$  of the bidimensional torus*, Contemp. Math. **107** (1990), 19–37. ↑5
- [8] Laurent Dalenc and Yumeng Ou, *Upper bound for multi-parameter iterated commutators*, arXiv:1401.5994, 1–18. ↑2, 24, 26, 27, 28, 29, 32, 34
- [9] Laurent Dalenc and Stefanie Petermichl, *A lower bound criterion for iterated commutators*, J. Funct. Anal. **266** (2014), no. 8, 5300–5320. ↑13, 14
- [10] Robert Fefferman, *Bounded mean oscillation on the polydisk*, Ann. of Math. (2) **110** (1979), no. 2, 395–406. ↑5
- [11] ———, *Singular integrals on product domains*, Bull. Amer. Math. Soc. (N.S.) **4** (1981), no. 2, 195–201. ↑24

- [12] ———, *Harmonic analysis on product spaces*, Ann. of Math. (2) **126** (1987), no. 1, 109–130. ↑5
- [13] Sarah Ferguson and Michael Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143–160. ↑2
- [14] Sarah Ferguson and Cora Sadosky, *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures*, J. Anal. Math. **81** (2000), 239–267. ↑2, 3, 5, 6
- [15] Tuomas Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. **175** (2012), no. 3, 1473–1506. ↑24
- [16] Jean-Lin Journé, *Calderón-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), no. 3, 55–91. ↑4, 24
- [17] ———, *A covering lemma for product spaces*, Proc. Amer. Math. Soc. **96** (1986), no. 4, 593–598. ↑
- [18] Michael Lacey, Stefanie Petermichl, Jill Pipher, and Brett Wick, *Multiparameter Riesz commutators*, Amer. J. Math. **131** (2009), no. 3, 731–769. ↑2, 3, 5, 13, 23, 24, 35
- [19] ———, *Iterated Riesz commutators: a simple proof of boundedness.*, Contemp. Math. **505** (2010), 171–178. ↑24, 35
- [20] ———, *Multi-parameter Div-Curl Lemmas*, Bull. London Math. Soc. **10** (2012), 1123–1131. ↑3
- [21] Michael Lacey and Erin Terwilleger, *Hankel operators in several complex variables and product BMO*, Houston J. Math. **35** (2009), no. 1, 159–183. ↑
- [22] Henri Martikainen, *Representation of bi-parameter singular integrals by dyadic operators*, Adv. Math. **229** (2012), no. 3, 1734–1761. ↑4, 24, 25
- [23] Henri Martikainen and Tuomas Orponen, *Some obstacles in characterizing the boundedness of bi-parameter singular integrals*, arXiv:1404.2216, 1–10. ↑24
- [24] Camil Muscalu, Jill Pipher, Terence Tao, and Christoph Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), no. 2, 269–296. ↑35
- [25] ———, *Multi-parameter paraproducts*, Rev. Mat. Iberoam. **22** (2006), no. 3, 963–976. ↑35
- [26] Zeev Nehari, *On bounded bilinear forms*, Ann. of Math. (2) **65** (1957), 153–162. ↑2
- [27] Yumeng Ou, *A  $T(b)$  theorem on product spaces*, arXiv:1305.1691, to appear Trans. Amer. Math. Soc., 1–45. ↑24
- [28] ———, *Multi-parameter singular integral operators and representation theorem*, arXiv:1410.8055, 1–28. ↑4, 24, 33, 34

- [29] Stefanie Petermichl, *Dyadic Shifts and a Logarithmic Estimate for Hankel Operators with Matrix Symbol*, Comptes Rendus Acad. Sci. Paris **1** (2000), no. 1, 455–460. ↑24
- [30] Stefanie Petermichl, Sergei Treil, and Alexander Volberg, *Why the Riesz transforms are averages of the dyadic shifts?*, Publicacions matemàtiques, ISSN 0214-1493 **46** (2002), no. 2, 209–228. ↑24
- [31] Akihito Uchiyama, *A constructive proof of the Fefferman Stein decomposition of  $BMO(\mathbb{R}^n)$* , Acta Math. **148** (1982), 215–241. ↑3