

**CHARACTERIZATION OF MATRIX VALUED BMO BY  
COMMUTATORS AND SPARSE DOMINATION OF OPERATORS**

A Dissertation  
Presented to  
The Academic Faculty

By

Darío Mena Arias

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy in the  
School of Mathematics

Georgia Institute of Technology

May 2018

Copyright © Darío Mena Arias 2018

CHARACTERIZATION OF MATRIX VALUED BMO BY  
COMMUTATORS AND SPARSE DOMINATION OF OPERATORS

Approved by:

Dr. Michael Lacey, Advisor  
School of Mathematics  
*Georgia Institute of Technology*

Dr. Christopher Heil  
School of Mathematics  
*Georgia Institute of Technology*

Dr. Plamen Iliev  
School of Mathematics  
*Georgia Institute of Technology*

Dr. Shahaf Nitzan  
School of Mathematics  
*Georgia Institute of Technology*

Dr. Brett Wick  
School of Mathematics  
*Washington University*

Date Approved: March 27, 2018

## ACKNOWLEDGEMENTS

I would like to extend my gratitude to my advisor Michael Lacey for all the help, support and guidance during the process, for all the interesting projects and problems, and helping me grow as a professional and as a person.

I also thank Brett Wick, for his help and advice as my mentor in the beginning of my PhD program.

I also wish to thank the members of the committee Christopher Heil, Plamen Iliev and Shahaf Nitzan, not only because of their time and dedication in reading my thesis, but also because as their student I benefited a lot from their knowledge and experience.

Finally, thanks to my family, my girlfriend, and friends, for all the support and patience through all these years, and for never letting me feel alone, in spite of the distance.

## TABLE OF CONTENTS

<b>Acknowledgments</b> . . . . .	iii
<b>List of Figures</b> . . . . .	vi
<b>Chapter 1</b> <b>Characterization of two parameter matrix-valued BMO by commutator with the Hilbert transform</b> . . . . .	1
1.1 Introduction . . . . .	1
1.2 Upper bound . . . . .	4
1.2.1 Remark: Logarithmic estimate . . . . .	19
1.3 Lower bound . . . . .	20
<b>Chapter 2</b> <b>Introduction to the sparse theory</b> . . . . .	22
2.1 Definitions and basic concepts . . . . .	22
2.2 Boundedness of the sparse operators . . . . .	23
2.2.1 Weighted inequalities . . . . .	24
<b>Chapter 3</b> <b>The Sparse T1 Theorem</b> . . . . .	25
3.1 Introduction . . . . .	25
3.2 Random Grids . . . . .	28
3.3 The Proof of the Sparse Bound . . . . .	30
3.3.1 Stopping Cubes . . . . .	32

3.3.2	The Inside Terms . . . . .	33
3.3.3	The Near Terms . . . . .	37
3.3.4	The Neighbors . . . . .	37
3.3.5	The Far Term . . . . .	38
3.4	Lemmas . . . . .	38
<b>Chapter 4 Uniform Sparse Bounds for Discrete Quadratic Phase Hilbert Transforms . . . . .</b>		<b>45</b>
4.1	Introduction . . . . .	45
4.2	Preliminaries . . . . .	49
4.3	Decomposition of the multiplier . . . . .	51
4.4	Minor Arcs estimate . . . . .	54
4.5	Major Arcs estimate . . . . .	55
<b>Chapter 5 Sparse Bounds for Bochner-Riesz Multipliers . . . . .</b>		<b>59</b>
5.1	Introduction . . . . .	59
5.2	The Full Statement . . . . .	61
5.3	Background on Sparse Forms . . . . .	66
5.4	Proof of the Sparse Bounds . . . . .	69
5.5	Sharpness of the sparse bounds . . . . .	74
5.6	Weighted Consequences . . . . .	76
<b>Bibliography . . . . .</b>		<b>78</b>

## LIST OF FIGURES

5.1	The trapezoid $\mathbf{R}(2, \delta)$ of Theorem 5.1.3. The Bochner-Riesz bounds are sharp at the indices $\frac{1}{p} = \frac{3}{4}, \frac{1}{4}$ , which corresponds to the Carleson-Sjölin bounds. The sparse bounds for $B_\delta$ hold for all $(\frac{1}{r}, \frac{1}{s})$ inside the dotted trapezoid, and fails outside the trapezoid. We abbreviate $v_{2,\delta,2} = v_2$ , and similarly for $v_3$ . . . . .	61
5.2	The interpolation argument for the sparse bounds. We have the sparse bounds $(\frac{1}{p_0}, \frac{1}{p'_0})$ , $(1, 0)$ , and $(1, 1)$ as well as their duals. The sparse bound at $(\frac{1}{p_0}, \frac{1}{p'_0})$ is, up to logarithmic terms, uniformly bounded in $0 < \tau < \frac{1}{2}$ , while the others are bounded by $\tau^{-\frac{n-1}{2}}$ . Interpolation, along the dotted lines, to the circles, yields sparse bounds with growth $\tau^{-\delta}$ , for $0 < \delta < \frac{n-1}{2}$ . . . . .	74

## SUMMARY

The first part of the thesis, consists on a result in the area of commutators. The classic result by Coifman, Rochberg and Weiss [16], establishes a relation between a BMO function, and the commutator of such a function with the Hilbert transform. More precisely, given a function  $b$ , if  $H$  denotes the Hilbert transform, then

$$\|b\|_{BMO} \sim \|[b, H]\|_{L^2 \rightarrow L^2}.$$

This equivalence of norms, has been carried over to different contexts, for example, multi-parameter theory, multilinear theory and Sobolev spaces. The result obtained for this thesis, is in the two parameters setting (with obvious generalizations to more than two parameters) in the case where the BMO function is matrix valued.

The second part of the thesis corresponds to domination of operators by using a special class called sparse operators. These operators are positive and highly localized, and therefore, allows for a very efficient way of proving weighted and unweighted estimates.

There are three main results regarding sparse operators, present in this thesis: The first one, a joint work with Michael Lacey, is a sparse version of the celebrated  $T1$  theorem of David and Journé. We impose standard  $T1$ -type assumptions on a Calderón-Zygmund operator  $T$ , and deduce that for bounded compactly supported functions  $f, g$  there is a sparse bilinear form  $\Lambda$  so that

$$|\langle Tf, g \rangle| \lesssim \Lambda(f, g).$$

The proof is short and elementary. The sparse bound quickly implies all the standard mapping properties of a Calderón-Zygmund on a (weighted)  $L^p$  space.

The second result, in collaboration with Robert Kesler, considers the discrete

quadratic phase Hilbert Transform acting on  $\ell^2(\mathbb{Z})$  finitely supported functions

$$H^\alpha f(n) := \sum_{m \neq 0} \frac{e^{i\alpha m^2} f(n-m)}{m}.$$

We prove that, uniformly in  $\alpha \in \mathbb{T}$ , there is a sparse bound for the bilinear form  $\langle H^\alpha f, g \rangle$ . The sparse bound implies several mapping properties such as weighted inequalities in an intersection of Muckenhoupt and reverse Hölder classes.

The last result, jointly with Michael Lacey and María Carmen Reguera, expands the sparse domination to the Bochner-Riesz multipliers. We show that these operators satisfy a range of sparse bounds, for all  $0 < \delta < \frac{n-1}{2}$ . The range of sparse bounds increases to the optimal range, as  $\delta$  increases to the critical value,  $\delta = \frac{n-1}{2}$ , even assuming only partial information on the Bochner-Riesz conjecture in dimensions  $n \geq 3$ . In dimension  $n = 2$ , we prove a sharp range of sparse bounds. The method of proof is based upon a ‘single scale’ analysis, and yields the sharpest known weighted estimates for the Bochner-Riesz multipliers in the category of Muckenhoupt weights.

**CHAPTER 1**  
**CHARACTERIZATION OF TWO PARAMETER MATRIX-VALUED**  
**BMO BY COMMUTATOR WITH THE HILBERT TRANSFORM**

**1.1 Introduction**

It is well known, by the work of R. Coifman, R. Rochberg, and G. Weiss [16], that the space of functions of bounded mean oscillation (BMO) can be characterized by commutators with the Hilbert transform (and in general, with the Riesz transforms). Given  $b \in BMO$ , let  $M_b$  represent the multiplication operator  $M_b(f) = bf$ , if  $H$  represents the Hilbert transform, defined as

$$Hf(x) = p.v \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(y)}{x-y} dy,$$

then we have

$$\|b\|_{BMO} \lesssim \|[M_b, H]\|_{L^2 \rightarrow L^2} \lesssim \|b\|_{BMO}.$$

The study of the norm of the commutator has several implications in the characterization of Hankel operators, the problem of factorization and weak factorization of function spaces and the div-curl problem. Several extensions and generalizations have been made in different settings. In the two parameter version of this result, the upper bound was shown by S. Ferguson and C. Sadosky in [29], while the lower bound was proved by S. Ferguson and M. Lacey in [28]. The formulation in this case is the following: If  $H_i$  represents the Hilbert transform in the  $i$ -th variable, then

$$\|b\|_{BMO} \lesssim \|[M_b, H_1], H_2\|_{L^2 \rightarrow L^2} \lesssim \|b\|_{BMO}.$$

Here, we are considering the product BMO of S.Y. Chang and R. Fefferman [14]. These results were later extended to the multi-parameter case by M. Lacey and E. Terwilleger [52].

The idea of the present work, is to obtain the same characterization in the two parameter case, for a matrix-valued BMO function. In the one parameter setting, we have the desired characterization due to S. Petermichl [66], and also F. Nazarov, G. Pisier, S. Treil and A. Volberg [62].

Consider the collection  $\mathcal{D}$  of dyadic intervals, that is

$$\mathcal{D} := \{ [k2^{-j}, (k+1)2^{-j}) : j, k \in \mathbb{Z} \},$$

and the collection of “shifted” dyadic intervals

$$\mathcal{D}^{\alpha, r} = \{ \alpha + r[k2^j, (k+1)2^j) : k, j \in \mathbb{Z} \}, \quad \alpha, r \in \mathbb{R}.$$

Define the dyadic Haar function as  $h_I := \frac{1}{\sqrt{|I|}}(\mathbb{1}_{I_-} - \mathbb{1}_{I_+})$ , where  $I_-$  and  $I_+$  represent the left and right half of the interval  $I$ , respectively. Denote also  $h_I^1 = \frac{\mathbb{1}_I}{\sqrt{|I|}}$  (non-cancellative Haar function). The family  $\{h_I : I \in \mathcal{D}\}$  (or  $I \in \mathcal{D}^{\alpha, r}$ ), is an orthonormal basis for  $L^2(\mathbb{R}; \mathbb{C}^d)$ ; here, for two Banach spaces  $X$  and  $Y$ , we use the notation  $L^p(X; Y)$  to denote the set  $\{f : X \rightarrow Y : \int_X \|f\|_Y^p < \infty\}$ .

Define the dyadic Haar shift by  $\mathbb{H}^{\alpha, r}(h_I) = \frac{1}{\sqrt{2}}(h_{I_-} - h_{I_+})$ , and extend to a general function  $f$  by

$$\mathbb{H}^{\alpha, r}(f) = \sum_{I \in \mathcal{D}^{\alpha, r}} \langle f, h_I \rangle \mathbb{H}^{\alpha, r}(h_I) = \sum_{I \in \mathcal{D}^{\alpha, r}} \langle f, h_I \rangle \frac{1}{\sqrt{2}}(h_{I_-} - h_{I_+}).$$

Note that  $\mathbb{H}^{\alpha, r}$  is bounded from  $L^2(\mathbb{R}; \mathbb{C}^d)$  to  $L^2(\mathbb{R}; \mathbb{C}^d)$ , with operator norm 1. As proven by Petermichl in [66], the kernel for the Hilbert transform can be written

as an average of dyadic shifts, in particular

$$K(t, x) = \lim_{L \rightarrow \infty} \frac{1}{2 \log L} \int_{1/L}^L \lim_{R \rightarrow \infty} \frac{1}{2R} \int_{-R}^R K^{\alpha, r}(t, x) d\alpha \frac{dr}{r}.$$

Where  $K^{\alpha, r}(t, x) = \sum_{I \in \mathcal{D}^{\alpha, r}} h_I(t) \mathbb{H}^{\alpha, r}(h_I(x))$ . Therefore, it is enough to prove the upper bound for the commutator with the shift  $[M_B, \mathbb{H}]$  (the estimates don't depend on  $\alpha$  or  $r$ ).

Let  $B$  be a function with values in the space of  $d \times d$  matrices. We consider the commutator  $[M_B, H]$  acting on a vector-valued function  $f$  by

$$[M_B, H]f = BH(f) - H(Bf).$$

The result obtained by Petermichl is based on a decomposition in paraproducts, and uses the estimates obtained by Katz [39], and Nazarov, Treil and Volberg [63] independently. We have

$$\|[M_B, H]\|_{L^2(\mathbb{R}; \mathbb{C}^d) \rightarrow L^2(\mathbb{R}; \mathbb{C}^d)} \lesssim \log(1 + d) \|B\|.$$

Motivated by this result, we wish to find a generalization in a two parameter setting, with the corresponding definition of the product BMO space (analogous to the one given by Chang and Fefferman in [14]). The main result of the paper can be stated as follows.

**Theorem 1.1.1.** *Let  $B$  be a  $d \times d$  matrix-valued BMO function on  $\mathbb{R}^2$ . If  $M_B$  denotes the operator “multiplication by  $B$ ”, and  $H_i$  represents the Hilbert transform in the  $i$ -th parameter, for  $i = 1, 2$ , then the norm of the iterated commutator  $[[M_B, H_1], H_2]$  satisfies*

$$d^{-2} \|B\|_{BMO} \lesssim \|[M_B, H_1], H_2\|_{L^2(\mathbb{R}^2, \mathbb{C}^d) \rightarrow L^2(\mathbb{R}^2, \mathbb{C}^d)} \lesssim d^3 \|B\|_{BMO}.$$

The paper is organized as follows. Section 2, contains the proof of the upper bound for the norm of the commutator, using a decomposition in paraproducts. Section 3 contains the proof of the lower bound, that relies on the proof for the scalar case by S. Ferguson and M. Lacey in [28]. Throughout the paper, we use the notation  $A \lesssim B$  to indicate that there is a positive constant  $C$ , such that  $A \leq CB$ .

## 1.2 Upper bound

Consider  $\mathcal{R} = \mathcal{D} \times \mathcal{D}$ , the class of rectangles consisting on products of dyadic intervals. Given a subset  $E$  of  $\mathbb{R}^2$ , denote by  $\mathcal{R}(E)$  the family of dyadic rectangles contained in  $E$ .

Consider the wavelet  $w_I$  constructed by Meyer in [57], and the two-parameters wavelet  $v_R(x, y) = w_I(x)w_J(y)$  for  $R = I \times J$ , with all its properties listed in [28]. We start by giving the definitions of product *BMO* and product dyadic *BMO*.

**Definition 1.2.1** (BMO). A function  $B$  is in  $BMO(\mathbb{R}^2)$  if and only if there are constants  $C_1$  and  $C_2$  such that, for any open set  $U \subseteq \mathbb{R}^2$  we have

$$\begin{aligned} 1. & \left( \frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \langle B, v_R \rangle \langle B, v_R \rangle^* \right)^{1/2} \leq C_1 I_d \\ 2. & \left( \frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \langle B, v_R \rangle^* \langle B, v_R \rangle \right)^{1/2} \leq C_2 I_d. \end{aligned}$$

The inequalities are considered in the sense of operators,  $I_d$  is the identity  $d \times d$  matrix. The *BMO*-norm is defined as the smallest constant, denoted by  $\|B\|_{BMO}$ , for which the two inequalities are satisfied simultaneously. If we take the supremum only over rectangles  $U$ , we obtain the rectangular *BMO*-norm, denoted by  $\|B\|_{BMO_{\text{rec}}}$ .

If  $h_I$  represents the Haar function associated to a dyadic interval  $I$ , define

$$h_R(x, y) = h_I(x)h_J(y), \quad \text{for } R = I \times J.$$

That is  $h_R = h_I \otimes h_J$ . The family  $\{h_R\}_{R \in \mathcal{R}}$  is an orthonormal basis for  $L^2(\mathbb{R}^2, \mathbb{C}^d)$ . We have the following definition of dyadic  $BMO$ . Note that it is the same definition, but considering the Haar wavelet instead of the Meyer wavelet.

**Definition 1.2.2** (Dyadic  $BMO$ ). A matrix-valued function  $B$  is in  $BMO_d(\mathbb{R}^2)$  (dyadic  $BMO$ ) if and only if, there are constants  $C_1$  and  $C_2$  such that for any open subset  $U$  of the plane, we have

$$\begin{aligned} 1. & \left( \frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \langle B, h_R \rangle \langle B, h_R \rangle^* \right)^{1/2} \leq C_1 I_d \\ 2. & \left( \frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \langle B, h_R \rangle^* \langle B, h_R \rangle \right)^{1/2} \leq C_2 I_d. \end{aligned}$$

Where the inequality is in the sense of operators. And the corresponding norm  $\|B\|_{BMO_d}$  is, again, the best constant for the two inequalities.

It is known that  $\|B\|_{BMO_d} \leq \|B\|_{BMO}$ ; this fact can be found in [76]. In that paper, the proof of the inequality is given in the multi-parameter setting, for Hilbert space-valued functions, by means of the dual inequality  $\|f\|_{H^1} \leq \|f\|_{H_d^1}$  (Estimate 2.3 in [76]). The duality in the dyadic case is discussed later, in the proof of Proposition 1.2.4. Using this fact, for the proof of the upper bound, it's enough to consider the dyadic version of  $BMO$  for the computations. For the rest of this section, we use  $\widehat{B}(R)$  to denote the Haar coefficient of the function  $B$ , associated to the function  $h_R$ , that is

$$\widehat{f}(R) = \langle f, h_R \rangle = \int_{\mathbb{R}^2} f(x, y) h_R(x, y) dx dy.$$

Since  $\widehat{B}(R)\widehat{B}(R)^*$  is a positive semi-definite matrix, we have

$$\sqrt{\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \|\widehat{B}(R)\|^2} \simeq \sqrt{\text{Tr} \left( \frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \widehat{B}(R)\widehat{B}(R)^* \right)}$$

$$\leq \text{Tr} \sqrt{\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \widehat{B}(R) \widehat{B}(R)^*}.$$

So, if we consider the two inequalities

$$\sqrt{\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \widehat{B}(R) \widehat{B}(R)^*} \leq C I_d, \quad \sqrt{\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \widehat{B}(R)^* \widehat{B}(R)} \leq C I_d,$$

taking the trace on both sides, we get

$$\sqrt{\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} \|\widehat{B}(R)\|^2} \leq C d. \quad (1.2.1)$$

The initial computations are similar to the ones found in [24]. In this, we need simplified versions, since we are dealing only with the biparameter Hilbert transform; differences will arise when we deal with the various paraproducts that result from this process, due to the *BMO* symbol being a matrix (which implies losing commutativity and requiring the use of matrix norms). Similar computations are used in [49], and this ideas can also be implemented in our case. Although we can use some equivalent results from [60, 61] to deal with the boundedness of the paraproducts, the ones arising from our computations can be given self contained proofs of their boundedness.

The dyadic shift operator  $\text{III}(f) = \sum_{I \in \mathcal{D}} \widehat{f}(I) \frac{1}{\sqrt{2}} (h_{I_-} - h_{I_+})$  corresponds to the operator  $S^{1,0}$  described by Dalenc and Ou in [24], given by

$$S^{1,0} f = \sum_{K \in \mathcal{D}} \sum_{I \subseteq K}^{(0)} \sum_{J \subseteq K}^{(1)} a_{IJK} \langle f, h_I \rangle h_J, \quad a_{IJK} = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } J = K_-, \\ -\frac{1}{\sqrt{2}}, & \text{if } J = K_+. \end{cases}$$

Here, the symbol  $\sum_{I \subseteq J}^{(k)}$  represents summing over those dyadic intervals  $I$  such that  $I \subseteq J$ , and  $|I| = 2^{-k}|J|$ . Let  $\tilde{I}$  represent the parent of the dyadic interval  $I$ , that is, the unique dyadic interval containing  $I$  with  $|\tilde{I}| = 2|I|$ , then, the shift can also be

expressed in a simpler way by

$$\mathbb{I}\mathbb{I}(f) = \sum_{I \in \mathcal{D}} a_I \widehat{f}(\tilde{I}) h_I, \quad (1.2.2)$$

where  $a_I = \frac{1}{\sqrt{2}}$  if  $I = \tilde{I}_-$ , and  $-\frac{1}{\sqrt{2}}$  if  $I = \tilde{I}_+$ .

If we write  $B = \sum_{I \in \mathcal{D}} \widehat{B}(I) h_I$ , and  $f = \sum_{J \in \mathcal{D}} \widehat{f}(J) h_J$ , then we can write

$$Bf = \sum_I \sum_J \widehat{B}(I) h_I \widehat{f}(J) h_J.$$

Therefore the commutator

$$[M_B, \mathbb{I}\mathbb{I}](f) = M_B \mathbb{I}\mathbb{I}(f) - \mathbb{I}\mathbb{I}(M_B f) = B \mathbb{I}\mathbb{I}(f) - \mathbb{I}\mathbb{I}(Bf),$$

can be written as

$$\begin{aligned} [M_B, \mathbb{I}\mathbb{I}](f) &= \sum_{I,J} \widehat{B}(I) \widehat{f}(J) h_I \mathbb{I}\mathbb{I}(h_J) - \sum_{I,J} \widehat{B}(I) \widehat{f}(J) \mathbb{I}\mathbb{I}(h_I h_J) \\ &= \sum_{I,J} \widehat{B}(I) \widehat{f}(J) [M_{h_I}, \mathbb{I}\mathbb{I}](h_J). \end{aligned}$$

Note that the terms are non-zero, only when  $I \cap J \neq \emptyset$ , also, if  $J \subsetneq I$ , we have that  $h_I$  is constant in  $I \cap J$ , therefore, for every  $x \in I \cap J$ , we have

$$\begin{aligned} [M_{h_I}, \mathbb{I}\mathbb{I}](h_J) &= h_I(x) \mathbb{I}\mathbb{I}(h_J(x)) - \mathbb{I}\mathbb{I}(h_I(x) h_J(x)) \\ &= h_I(x) \mathbb{I}\mathbb{I}(h_J(x)) - h_I(x) \mathbb{I}\mathbb{I}(h_J(x)) = 0. \end{aligned}$$

Then, the only non-trivial terms are those for which  $I \subset J$ .

We consider the two parameter commutator  $[[M_B, H_1], H_2]$  acting on a vector-

valued function  $f$  by

$$\begin{aligned} [[M_B, H_1], H_2]f &= BH_1(H_2(f)) - H_1(B(H_2(f))) \\ &\quad - H_2(BH_1(f)) + H_2(H_1(Bf)). \end{aligned}$$

Where  $H_1$  and  $H_2$  represent the Hilbert transform, on the first and second variable respectively. That is,

$$H_1 f(x, y) = p.v \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(z, y)}{x - z} dz, \quad H_2 f(x, y) = p.v \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(x, z)}{y - z} dz.$$

The main result we want to prove in this section is the following

**Theorem 1.2.3.** *Let  $B$  be a matrix-valued  $BMO_d(\mathbb{R}^2)$  function and  $f$  in  $L^2(\mathbb{R}^2; \mathbb{C}^d)$ , then*

$$\|[[M_B, H_1], H_2]\|_{L^2(\mathbb{R}^2; \mathbb{C}^d) \rightarrow L^2(\mathbb{R}^2; \mathbb{C}^d)} \lesssim \|B\|_{BMO_d}.$$

**Proof:** Let  $\mathbb{I}\mathbb{I}\mathbb{I}_1$  and  $\mathbb{I}\mathbb{I}\mathbb{I}_2$  represent the dyadic shift operator in the first and second variable respectively, that is,  $\mathbb{I}\mathbb{I}\mathbb{I}_1(h_R) = \mathbb{I}\mathbb{I}(h_I) \otimes h_J$ , and  $\mathbb{I}\mathbb{I}\mathbb{I}_2(h_R) = h_I \otimes \mathbb{I}\mathbb{I}(h_J)$ , for  $R = I \times J$ , and extending to a function  $f$  by

$$\mathbb{I}\mathbb{I}\mathbb{I}_j(f) = \sum_{R \in \mathcal{R}} \widehat{f}(R) \mathbb{I}\mathbb{I}\mathbb{I}_j(h_R), \quad j = 1, 2.$$

Or in the notation of (1.2.2),

$$\mathbb{I}\mathbb{I}\mathbb{I}_1(f) = \sum_{I, J \in \mathcal{D}} a_I \widehat{f}(\tilde{I} \times J) h_I \otimes h_J, \quad \mathbb{I}\mathbb{I}\mathbb{I}_2(f) = \sum_{I, J \in \mathcal{D}} a_J \widehat{f}(I \times \tilde{J}) h_I \otimes h_J.$$

Again, due to the representation of  $H$  as an average of shifts, it is enough to prove the result for the commutator  $[[M_B, \mathbb{I}\mathbb{I}\mathbb{I}_1], \mathbb{I}\mathbb{I}\mathbb{I}_2]$ . By an iteration of the computation for the one parameter case, using the Haar expansion of the functions  $B$  and  $f$  and

taking their formal product, we obtain that  $[M_B, \mathbb{I}\mathbb{I}_1](f)$  is equal to

$$\begin{aligned}
& \sum_{R,S \in \mathcal{R}} \widehat{B}(R) \widehat{f}(S) (h_R \mathbb{I}\mathbb{I}_1(h_S) - \mathbb{I}\mathbb{I}_1(h_R h_S)) \\
&= \sum_{R,S \in \mathcal{R}} \widehat{B}(R) \widehat{f}(S) [M_{h_R}, \mathbb{I}\mathbb{I}_1](h_S) \\
&= \sum_{I,J,K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) (h_I \mathbb{I}\mathbb{I}_1 h_K - \mathbb{I}\mathbb{I}_1(h_I h_K)) \otimes h_J h_L.
\end{aligned}$$

Repeating the same computations, we get that the two parameters commutator  $[[M_B, \mathbb{I}\mathbb{I}_1], \mathbb{I}\mathbb{I}_2](f)$  is equal to

$$\begin{aligned}
& \sum_{I,J \in \mathcal{D}} \sum_{K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) h_I \mathbb{I}\mathbb{I}_1 h_K \otimes h_J \mathbb{I}\mathbb{I}_2 h_L \\
&- \sum_{I,J \in \mathcal{D}} \sum_{K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) \mathbb{I}\mathbb{I}_1(h_I h_K) \otimes h_J \mathbb{I}\mathbb{I}_2 h_L \\
&- \sum_{I,J \in \mathcal{D}} \sum_{K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) h_I \mathbb{I}\mathbb{I}_1 h_K \otimes \mathbb{I}\mathbb{I}_2(h_J h_L) \\
&+ \sum_{I,J \in \mathcal{D}} \sum_{K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) \mathbb{I}\mathbb{I}_1(h_I h_K) \otimes \mathbb{I}\mathbb{I}_2(h_J h_L) \\
&= T_1 f - T_2 f - T_3 f + T_4 f \\
&= \sum_{I,J \in \mathcal{D}} \sum_{K,L \in \mathcal{D}} \widehat{B}(I \times J) \widehat{f}(K \times L) [M_{h_I}, \mathbb{I}\mathbb{I}_1](h_K) \otimes [M_{h_J}, \mathbb{I}\mathbb{I}_2](h_L).
\end{aligned}$$

If either  $I \cap K = \emptyset$ ,  $J \cap L = \emptyset$ ,  $K \subsetneq I$  or  $L \subsetneq J$ , then we have that  $[M_{h_I}, \mathbb{I}\mathbb{I}_1](h_K) \otimes [M_{h_J}, \mathbb{I}\mathbb{I}_2](h_L) = 0$ ; therefore, the terms are non-trivial only when  $I \subseteq K$  and  $J \subseteq L$ . We have four different cases, that can be analyzed independently for each term in the sum. The computations for the four terms are similar, only the complete details for the term  $T_2$  will be provided, and at the end of the proof of the proposition we mention briefly how to deal with the other cases. Let  $\widetilde{T}_j$  represent  $T_j$  restricted to

the case  $I \subseteq K$  and  $J \subseteq L$ , then we have.

$$\tilde{T}_2 f = \mathbb{I}\mathbb{I}_1 \left( \sum_K \sum_L \sum_{I \subseteq K} \sum_{J \subseteq L} \widehat{B}(I \times J) \widehat{f}(K \times L) h_I h_K \otimes h_J \mathbb{I}\mathbb{I}_2 h_L \right).$$

To analyze each of the four cases, we need the following proposition.

**Proposition 1.2.4.** *Consider the following paraproducts*

$$(i) P_B^1(f) = \sum_{I, J \in \mathcal{D}} \pm \widehat{B}(I \times \tilde{J}) \langle f, h_I \otimes h_J \rangle h_I^1 \otimes h_J |I|^{-1/2} |\tilde{J}|^{-1/2}.$$

$$(ii) P_B^2(f) = \sum_{I, J} \pm \widehat{B}(I \times \tilde{J}) \langle f, h_I^1 \otimes h_J \rangle h_I \otimes h_J |I|^{-1/2} |\tilde{J}|^{-1/2}.$$

$$(iii) P_B^3(f) = \sum_{I, J \in \mathcal{D}} \widehat{B}(I \times J) \langle f, h_I^1 \otimes h_J^1 \rangle h_I \otimes h_J |I|^{-1/2} |J|^{-1/2}.$$

$$(iv) P_B^4(f) = \sum_{I, J \in \mathcal{D}} \widehat{B}(I \times J) \langle f, h_I \otimes h_J^1 \rangle h_I^1 \otimes h_J |I|^{-1/2} |J|^{-1/2}.$$

$$(v) P_B^5(f) = \sum_{I, J \in \mathcal{D}} \widehat{B}(I \times J) \langle f, h_I^1 \otimes h_J \rangle h_I \otimes h_J^1 |I|^{-1/2} |J|^{-1/2}.$$

We have that for  $i = 1, 2, 3, 4$ ,

$$\|P_B^i(f)\|_{L^2(\mathbb{R}^2; \mathbb{C}^d)} \lesssim d \|B\|_{BMO_d} \|f\|_{L^2(\mathbb{R}^2; \mathbb{C}^d)}.$$

**Proof of proposition:** In the following computations, for simplification we will write  $L^2(Y) = L^2(\mathbb{R}^2; Y)$ , since all the functions that we consider are defined on  $\mathbb{R}^2$ .

(i) We make use of a well known result, which is discussed in [13] for the bidisc case, but it is easily extended to the plane.

**Theorem 1.2.5** (Carleson Embedding Theorem). *Let  $\{a_R\}_{R \in \mathcal{R}}$  be a sequence of non-negative numbers, indexed by the grid of dyadic rectangles. Then the following are equivalent:*

$$(i) \sum_{R \in \mathcal{R}} a_R \langle f \rangle_R^2 \leq C_1 \|f\|_{L^2}^2, \text{ for all } f \in L^2.$$

(ii)  $\frac{1}{|U|} \sum_{R \in \mathcal{R}(U)} a_R \leq C_2$ , for all connected open sets  $U \subseteq \mathbb{R}^2$ .

Moreover,  $C_1 \simeq C_2$ .

We have the following basic estimates

$$\begin{aligned}
& \left| \langle P_B^1 f, g \rangle_{L^2} \right| = \left| \int_{\mathbb{R}^2} \langle P_B^1 f, g \rangle_{\mathbb{C}^d} dx dy \right| \\
&= \left| \int_{\mathbb{R}^2} \left\langle \sum_{I,J} \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J) \mathbb{1}_I |I|^{-1} \otimes h_J |\tilde{J}|^{-1/2}, g \right\rangle_{\mathbb{C}^d} dx dy \right| \\
&= \left| \int_{\mathbb{R}^2} \sum_{I,J} \left\langle \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J), g \mathbb{1}_I |I|^{-1} \otimes h_J |\tilde{J}|^{-1/2} \right\rangle_{\mathbb{C}^d} dx dy \right| \\
&= \left| \sum_{I,J} \int_{\mathbb{R}^2} \left\langle \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J), g \mathbb{1}_I |I|^{-1} \otimes h_J |\tilde{J}|^{-1/2} \right\rangle_{\mathbb{C}^d} dx dy \right| \\
&= \left| \sum_{I,J} \left\langle \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J), \int_{\mathbb{R}^2} \frac{1}{\sqrt{2}} g \mathbb{1}_I |I|^{-1} \otimes h_J |J|^{-1/2} dx dy \right\rangle_{\mathbb{C}^d} \right| \\
&= \frac{1}{\sqrt{2}} \left| \sum_{I,J} \left\langle \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J), \langle g, \mathbb{1}_I |I|^{-1} \otimes h_J |J|^{-1/2} \rangle \right\rangle_{\mathbb{C}^d} \right| \\
&\leq \frac{1}{\sqrt{2}} \sum_{I,J} \left| \left\langle \pm \widehat{B}(I \times J^{(1)}) \widehat{f}(I \times J), \langle g, \mathbb{1}_I |I|^{-1} \otimes h_J |J|^{-1/2} \rangle \right\rangle_{\mathbb{C}^d} \right| \\
&\leq \frac{1}{\sqrt{2}} \sum_{I,J} \left\| \widehat{B}(I \times J^{(1)}) \right\| \left\| \widehat{f}(I \times J) \right\|_{\mathbb{C}^d} \left\| \langle g, \mathbb{1}_I |I|^{-1} \otimes h_J |J|^{-1/2} \rangle \right\|_{\mathbb{C}^d} \\
&\leq \frac{1}{\sqrt{2}} \sum_{I,J} \left\| \widehat{f}(I \times J) \right\|_{\mathbb{C}^d} \left\| \widehat{B}(I \times J^{(1)}) \right\| \langle \|g\|_{\mathbb{C}^d} \rangle_{I \times J} \\
&\leq \frac{1}{\sqrt{2}} \left( \sum_{I,J} \left\| \widehat{f}(I \times J) \right\|_{\mathbb{C}^d}^2 \right)^{\frac{1}{2}} \left( \sum_I \sum_J \left\| \widehat{B}(I \times J^{(1)}) \right\|^2 \langle \|g\|_{\mathbb{C}^d} \rangle_{I \times J}^2 \right)^{\frac{1}{2}} \\
&\leq \frac{1}{\sqrt{2}} \|f\|_{L^2(\mathbb{C}^d)} \left( \sum_{I,J} \left\| \widehat{B}(I \times J^{(1)}) \right\|^2 \langle \|g\|_{\mathbb{C}^d} \rangle_{I \times J}^2 \right)^{\frac{1}{2}} \\
&\lesssim \|f\|_{L^2(\mathbb{C}^d)} d \|B\|_{BMO_d} \|g\|_{\mathbb{C}^d} \|L^2(\mathbb{R})\| = d \|B\|_{BMO_d} \|f\|_{L^2(\mathbb{C}^d)} \|g\|_{L^2(\mathbb{C}^d)}.
\end{aligned}$$

Here, we used the fact that since  $B \in BMO_d$ , then by (1.2.1), the second condition in

Theorem 1.2.5 is satisfied with  $a_R = \|\widehat{b}(R)\|^2$ . Note, that we have a linear dependence on the dimension of the matrix, due to the use of the trace. Note also that the same computations allow us to replace each individual  $I$  and  $J$  for a parent or “great parent” of  $I$  and  $J$ , in which case, the implied constant will depend also on complexity (level of relation with its ancestor); we will use  $P_B^1$  to denote any of these kind of paraproducts.

(ii) A direct computation shows that  $(P_B^2)^*$  is of the type  $P_{B^*}^1$ , therefore, by the symmetry of the definition of  $BMO_d$ -norm, the boundedness for  $P_B^2$  follows from that of  $P_B^1$ .

(iii) Denote by  $S_2^d$  the space of  $d \times d$  complex matrices, equipped with the norm derived from the inner product  $\langle A, B \rangle_{\text{Tr}} = \text{tr}(AB^*)$ , that is  $\|A\|_{S_2^d}^2 = \text{tr}(AA^*)$ . To estimate the  $L^2$ -norm of this operator, we compute  $\langle P_B^3(f), g \rangle$ .

$$\begin{aligned}
&= \int_{\mathbb{R}^2} \left\langle \sum_{I,J} \widehat{B}(I \times J) \langle f, h_I^1 \otimes h_J^1 \rangle \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}}, g \right\rangle_{\mathbb{C}^d} dx dy \\
&= \sum_{I,J} \int_{\mathbb{R}^2} \left\langle \widehat{B}(I \times J) \langle f, h_I^1 \otimes h_J^1 \rangle, g \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{\mathbb{C}^d} dx dy \\
&= \sum_{I,J} \left\langle \widehat{B}(I \times J) \langle f, h_I^1 \otimes h_J^1 \rangle, \langle g, h_I \otimes h_J \rangle \frac{1}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{\mathbb{C}^d} \\
&= \sum_{I,J} \left\langle \widehat{B}(I \times J), \langle g, h_I \otimes h_J \rangle \langle f, h_I^1 \otimes h_J^1 \rangle^* \frac{1}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} \\
&= \sum_{I,J} \int_{\mathbb{R}^2} \left\langle B h_I \otimes h_J, \langle g, h_I \otimes h_J \rangle \langle f, h_I^1 \otimes h_J^1 \rangle^* \frac{1}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} dx dy \\
&= \int_{\mathbb{R}^2} \left\langle B, \sum_{I,J} \langle g, h_I \otimes h_J \rangle \langle f, h_I^1 \otimes h_J^1 \rangle^* \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} dx dy \\
&= \left\langle B, \sum_{I,J} \langle g, h_I \otimes h_J \rangle \langle f, h_I^1 \otimes h_J^1 \rangle^* \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\rangle_{L^2(S_2^d)} \\
&= \langle B, \Pi_1(f, g) \rangle.
\end{aligned}$$

Define the space  $H_d^1$  to be the space of  $d \times d$  matrix-valued functions  $\Phi$  such that

$\|\Phi\|_{H_d^1} = \|S\Phi\|_{L^1}$ , where  $S$  is the square function defined by

$$S^2\Phi(x, y) := \sum_{I \in \mathcal{D}} \sum_{J \in \mathcal{D}} \|\langle \Phi, h_I \otimes h_J \rangle\|_{S_2^d}^2 \frac{\mathbb{1}_I(x) \mathbb{1}_J(y)}{|I| |J|}.$$

Note that if  $\Phi$  is in  $H_d^1$ , then all of its components are in scalar  $H^1$ , and for  $1 \leq i, j \leq d$ , we have  $\|\Phi_{i,j}\|_{H^1} \leq \|\Phi\|_{H_d^1}$ . Also, if  $B$  is a matrix-valued  $BMO_d$  function, then all of its components are in scalar dyadic  $BMO$ , and an easy computation shows that for  $1 \leq i, j \leq d$ ,  $\|B_{i,j}\|_{BMO} \leq d\|B\|_{BMO_d}$ . Using these facts, we can easily verify the following duality statement:

**Lemma 1.2.6 (BMO<sub>d</sub> – H<sub>d</sub><sup>1</sup> duality).** *Let  $B$  in  $BMO_d$  and  $\Phi$  in  $H_d^1$ , then*

$$\langle B, \Phi \rangle_{L^2(S_2^d)} \lesssim d^3 \|B\|_{BMO_d} \|\Phi\|_{H_d^1}.$$

Using this result, it is enough to prove that

$$\|\Pi_1(f, g)\|_{H_d^1} \simeq \|S(\Pi_1(f, g))\|_{L^1} \lesssim \|f\|_{L^2} \|g\|_{L^2}.$$

We compute  $[S(\Pi_1(f, g))(x, y)]^2$  to get

$$\begin{aligned} &= \sum_{I, J} \|\langle g, h_I \otimes h_J \rangle \langle f, h_I^1 \otimes h_J^1 \rangle^* |I|^{-1/2} |J|^{-1/2}\|_{S_2^d}^2 \frac{\mathbb{1}_I(x) \mathbb{1}_J(y)}{|I| |J|} \\ &= \sum_{I, J} \|\langle g, h_I \otimes h_J \rangle\|_{\mathbb{C}^d}^2 \left\| \left\langle f, \frac{h_I \otimes h_J}{|I|^{1/2} |J|^{1/2}} \right\rangle \right\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_{I \times J}(x, y)}{|I \times J|} \\ &\leq \sup_{(x, y) \in I \times J} \left\| \left\langle f, \frac{h_I \otimes h_J}{|I|^{1/2} |J|^{1/2}} \right\rangle \right\|_{\mathbb{C}^d}^2 \sum_{I, J} \|\langle g, h_I \otimes h_J \rangle\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_{I \times J}(x, y)}{|I \times J|} \\ &\leq \sup_{(x, y) \in I \times J} \left\langle \|f\|_{\mathbb{C}^d}, \frac{h_I \otimes h_J}{|I|^{1/2} |J|^{1/2}} \right\rangle^2 \sum_{I, J} \|\langle g, h_I \otimes h_J \rangle\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_{I \times J}(x, y)}{|I \times J|} \\ &\leq [\mathcal{M}(\|f\|_{\mathbb{C}^d})(x, y)]^2 [S(g)(x, y)]^2. \end{aligned}$$

Here,  $\mathcal{M}$  represents the strong maximal function. Using the  $L^2$ -boundedness of the maximal and square functions, we conclude

$$\|\Pi_1(f, g)\|_{H_d^1} \lesssim \|S(\Pi_1(f, g))\|_{L^1} \lesssim \|\mathcal{M}(\|f\|_{\mathbb{C}^d})S(g)\|_{L^1} \lesssim \|f\|_{L^2}\|g\|_{L^2}.$$

(iv) As in the previous case, we compute  $\langle P_B^4(f), g \rangle$

$$\begin{aligned} &= \int_{\mathbb{R}^2} \left\langle \sum_{I, J} \widehat{B}(I \times J) \langle f, h_I \otimes h_J^1 \rangle \frac{h_I^1 \otimes h_J}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}}, g \right\rangle_{\mathbb{C}^d} dx dy \\ &= \sum_{I, J} \int_{\mathbb{R}^2} \left\langle \widehat{B}(I \times J) \langle f, h_I \otimes h_J^1 \rangle, g \frac{h_I^1 \otimes h_J}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{\mathbb{C}^d} dx dy \\ &= \sum_{I, J} \left\langle \widehat{B}(I \times J) \langle f, h_I \otimes h_J^1 \rangle, \langle g, h_I^1 \otimes h_J \rangle \frac{1}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{\mathbb{C}^d} \\ &= \sum_{I, J} \left\langle \widehat{B}(I \times J), \langle g, h_I^1 \otimes h_J \rangle \langle f, h_I \otimes h_J^1 \rangle^* \frac{1}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} \\ &= \sum_{I, J} \int_{\mathbb{R}^2} \left\langle B h_I \otimes h_J, \langle g, h_I^1 \otimes h_J \rangle \langle f, h_I \otimes h_J^1 \rangle^* \frac{1}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} dx dy \\ &= \int_{\mathbb{R}^2} \left\langle B, \sum_{I, J} \langle g, h_I^1 \otimes h_J \rangle \langle f, h_I \otimes h_J^1 \rangle^* \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{\text{Tr}} dx dy \\ &= \left\langle B, \sum_{I, J} \langle g, h_I^1 \otimes h_J \rangle \langle f, h_I \otimes h_J^1 \rangle^* \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}}|J|^{\frac{1}{2}}} \right\rangle_{L^2(S_2^d)} \\ &= \langle B, \Pi_2(f, g) \rangle. \end{aligned}$$

Therefore, by duality, it is enough to prove that

$$\|\Pi_2(f, g)\|_{H_d^1} \lesssim \|f\|_{L^2}\|g\|_{L^2}.$$

For this, we proceed again to find a pointwise estimate for the square function.

We compute  $[S(\Pi_2(f, g))]^2$

$$\begin{aligned}
&= \sum_{I, J} \left\| \langle g, h_I^1 \otimes h_J \rangle \langle f, h_I \otimes h_J^1 \rangle^* \frac{1}{|I|^{\frac{1}{2}} |J|^{\frac{1}{2}}} \right\|_{S_2^d}^2 \frac{\mathbb{1}_{I \times J}}{|I \times J|} \\
&= \sum_{I, J} \|\langle g, h_J \rangle_I\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_J}{|J|} \|\langle f, h_I \rangle_J\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_I}{|I|} \\
&\leq \sum_{I, J} \|\langle g, h_J \rangle_I\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_J}{|J|} \|\langle f, h_I \rangle_J\|_{\mathbb{C}^d}^2 \frac{\mathbb{1}_I}{|I|} \\
&\leq \left( \sum_I (\mathcal{M}_2 \|\langle f, h_I \rangle\|_{\mathbb{C}^d})^2 \frac{\mathbb{1}_I}{|I|} \right) \left( \sum_J (\mathcal{M}_1 \|\langle g, h_J \rangle\|_{\mathbb{C}^d})^2 \frac{\mathbb{1}_J}{|J|} \right).
\end{aligned}$$

Where  $\mathcal{M}_1$  and  $\mathcal{M}_2$  represent the maximal function in the first and second variable, respectively. These last two factors are symmetric to each other, so it is enough to prove the  $L^2$ -boundedness for the operator

$$\tilde{S}f(x, y) = \left( \sum_I (\mathcal{M}_2 \|\langle f, h_I \rangle\|_{\mathbb{C}^d}(y))^2 \frac{\mathbb{1}_I(x)}{|I|} \right)^{1/2}.$$

But this is easy, since

$$\begin{aligned}
\int_{\mathbb{R}^2} (\tilde{S}f(x, y))^2 dx dy &= \sum_I \int_{\mathbb{R}} (\mathcal{M}_2 \|\langle f, h_I \rangle\|_{\mathbb{C}^d}(y))^2 dy \\
&\lesssim \sum_I \int_{\mathbb{R}} \|\langle f(\cdot, y), h_I(\cdot) \rangle\|_{\mathbb{C}^d}^2 dy = \|f\|_{L^2}^2.
\end{aligned}$$

(v) The computations are symmetric to those for (iv), exchanging the roles of  $I$  and  $J$ . ■

We proceed now to prove the upper bound for the four different cases. In each of them, the idea is to reduce the term to an expression of the form  $\mathbb{III}_1 \circ P_B^i \circ \mathbb{III}_2$ , therefore, by Proposition 1.2.4 and the boundedness of the shifts, we get the desired result. The estimates for the rest of the terms are similar, since they are reduced to find an upper bound for the norm of the four variants of paraproduct studied above. More

specifically, they correspond to expressions of the form  $\mathbb{I}_i(P_B(\mathbb{I}_j f))$ ,  $\mathbb{I}_i(\mathbb{I}_j(P_B f))$  and  $\mathbb{I}_i(\mathbb{I}_j(P_B f))$ ,  $\mathbb{I}_i(P_B f)$ , or duals of operators of the form  $\mathbb{I}_i(P_{B^*}(\mathbb{I}_j f))$ ,  $\mathbb{I}_i(\mathbb{I}_j(P_{B^*} f))$ ,  $\mathbb{I}_i(\mathbb{I}_j(P_{B^*} f))$  and  $\mathbb{I}_i(P_{B^*} f)$ .

**Case  $\mathbf{I} = \mathbf{K}, \mathbf{J} = \mathbf{L}$ .** In this case, using the definition of the shift, we have

$$\begin{aligned} & \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \widehat{f}(I \times J) h_I^2 h_J \mathbb{I}_2 h_J \right) \\ &= \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times \tilde{J}) \widehat{f}(I \times \tilde{J}) h_I^2 \otimes h_J a_J h_J \right). \end{aligned}$$

Since  $\mathbb{I}_2 \langle f, h_I \rangle = \sum_L a_L \widehat{f}(I \times L^{(1)}) h_L$ , then,  $\langle \mathbb{I}_2 \langle f, h_I \rangle, h_J \rangle = a_J \widehat{f}(I \times J^{(1)})$ .

So, the previous expression is equal to

$$\begin{aligned} & \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times \tilde{J}) \langle \mathbb{I}_2 \langle f, h_I \rangle, h_J \rangle h_I^2 \otimes h_J h_J \right) \\ &= \mathbb{I}_1 \left( \sum_I \sum_J \pm \widehat{B}(I \times \tilde{J}) \langle \mathbb{I}_2 \langle f, h_I \rangle, h_J \rangle \mathbb{1}_I |I|^{-1} \otimes h_J |\tilde{J}|^{-1/2} \right) \\ &= \mathbb{I}_1 \left( \sum_I \sum_J \pm \widehat{B}(I \times \tilde{J}) \langle \mathbb{I}_2 f, h_I \otimes h_J \rangle h_I^1 \otimes h_J |I|^{-1/2} |\tilde{J}|^{-1/2} \right). \end{aligned}$$

$\mathbb{I}_1(P_B^1(\mathbb{I}_2 f))$ .

**Case  $\mathbf{I} \subsetneq \mathbf{K}, \mathbf{J} \subsetneq \mathbf{L}$ .** Here we have

$$\begin{aligned} & \mathbb{I}_1 \left( \sum_K \sum_{I \subsetneq K} \sum_L \sum_{J \subsetneq L} \widehat{B}(I \times J) \widehat{f}(K \times L) h_I h_K \otimes h_J \mathbb{I}_2 h_L \right) \\ &= \mathbb{I}_1 \left( \sum_{J,K} \sum_{I \subsetneq K} \widehat{B}(I \times J) h_I h_K \otimes \left( \sum_{L \supseteq J} \langle \langle f, h_K \rangle, h_L \rangle \mathbb{I}_2 h_L \mathbb{1}_J \right) h_J \right). \end{aligned}$$

By using the definition of the shift, and the known average identity  $\langle f, h_J^1 \rangle |J|^{-1/2} =$

$\sum_{I \supseteq J} \widehat{f}(I) h_I \mathbb{1}_J$ , we have

$$\begin{aligned}
& \sum_{L \supseteq J} \langle \langle f, h_K \rangle, h_L \rangle \mathbb{I}_2 h_L \mathbb{1}_J = \mathbb{I}_2 \left( \sum_{L \supseteq J} \langle \langle f, h_K \rangle, h_L \rangle h_L \right) \mathbb{1}_J \\
& = \sum_{L \supseteq J} a_L \langle \langle f, h_K \rangle, h_{\bar{L}} \rangle h_L \mathbb{1}_J \\
& = a_J \langle \langle f, h_K \rangle, h_{\bar{J}} \rangle h_J + \sum_{L \supseteq J} a_L \langle \langle f, h_K \rangle, h_{\bar{L}} \rangle h_L \mathbb{1}_J \\
& = \langle \mathbb{I}_2 \langle f, h_K \rangle, h_J^1 \rangle |J|^{-1/2} \mathbb{1}_J + \langle \mathbb{I}_2 \langle f, h_K \rangle, h_J \rangle h_J.
\end{aligned}$$

This divides the original sum into two sums  $S_1 + S_2$ . The first one,  $S_1$ , is equal to

$$\begin{aligned}
& \mathbb{I}_1 \left( \sum_K \sum_{I \subsetneq K} \sum_J \widehat{B}(I \times J) \langle \mathbb{I}_2 \langle f, h_K \rangle, h_J^1 \rangle h_I h_K \otimes \frac{h_J}{|J|^{1/2}} \right) \\
& = \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \left( \sum_{K \supseteq I} \langle \langle \mathbb{I}_2 f, h_J^1 \rangle, h_K \rangle h_K \mathbb{1}_I \right) h_I \otimes \frac{h_J}{|J|^{1/2}} \right) \\
& = \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \langle \mathbb{I}_2 f, h_J^1 \rangle, h_I^1 \rangle \frac{h_I \otimes h_J}{|I|^{1/2} |J|^{1/2}} \right) \\
& = \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \mathbb{I}_2 f, h_I^1 \otimes h_J^1 \rangle \frac{h_I \otimes h_J}{|I|^{1/2} |J|^{1/2}} \right).
\end{aligned}$$

Which has the form  $\mathbb{I}_1(P_B^3(\mathbb{I}_2 f))$ . And with similar computations, we get

$$\begin{aligned}
S_2 & = \mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \mathbb{I}_2 f, h_I^1 \otimes h_J \rangle h_I \otimes h_J^1 |I|^{-1/2} |J|^{-1/2} \right) \\
& = \mathbb{I}_1(P_B^5(\mathbb{I}_2 f)).
\end{aligned}$$

**Case  $\mathbf{I} = \mathbf{K}, \mathbf{J} \subsetneq \mathbf{L}$ .** In this case we get

$$\mathbb{I}_1 \left( \sum_I \sum_L \sum_{J \subsetneq L} \widehat{B}(I \times J) \widehat{f}(I \times L) h_I^2 \otimes h_J \mathbb{I}_2 h_L \right)$$

$$\begin{aligned}
&= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) h_I^2 \otimes \left( \sum_{L \supseteq J} \langle \langle f, h_I \rangle, h_L \rangle \mathbb{I}\mathbb{I}_2 h_L \mathbb{1}_J \right) h_J \right) \\
&= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) h_I^2 \otimes \langle \mathbb{I}\mathbb{I}_2 \langle f, h_I \rangle, h_J^1 \rangle h_J |J|^{-1/2} \right) \\
&\quad + \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) h_I^2 \otimes \langle \mathbb{I}\mathbb{I}_2 \langle f, h_I \rangle, h_J \rangle h_J \right) \\
&= S_1 + S_2.
\end{aligned}$$

Again, by the definition of the shift

$$\begin{aligned}
S_1 &= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) h_I^2 \otimes \langle \mathbb{I}\mathbb{I}_2 \langle f, h_I \rangle, \mathbb{1}_J |J|^{-1} \rangle h_J \right) \\
&= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \mathbb{I}\mathbb{I}_2 f, h_I \otimes \mathbb{1}_J |J|^{-1} \rangle \mathbb{1}_I |I|^{-1} \otimes h_J \right) \\
&= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \mathbb{I}\mathbb{I}_2 f, h_I \otimes h_J^1 \rangle h_I^1 \otimes h_J |I|^{-1/2} |J|^{-1/2} \right).
\end{aligned}$$

Which has the form  $\mathbb{I}\mathbb{I}_1(P_B^4(\mathbb{I}\mathbb{I}_2 f))$ . And similarly

$$\begin{aligned}
S_2 &= \mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \mathbb{I}\mathbb{I}_2 f, h_I \otimes h_J \rangle h_I^1 \otimes h_J^1 |I|^{-1/2} |J|^{-1/2} \right) \\
&= \mathbb{I}\mathbb{I}_1((P_{B^*}^3)^*(\mathbb{I}\mathbb{I}_2 f)).
\end{aligned}$$

**Case  $\mathbf{I} \subsetneq \mathbf{K}, \mathbf{J} = \mathbf{L}$ .** last case we have

$$\begin{aligned}
&\mathbb{I}\mathbb{I}_1 \left( \sum_K \sum_J \sum_{I \subsetneq K} \widehat{B}(I \times J) \widehat{f}(K \times J) h_I h_K \otimes h_J \mathbb{I}\mathbb{I}_2 h_J \right) \\
&\mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \left( \sum_{K \supseteq I} \langle \langle f, h_J \rangle, h_K \rangle h_K \mathbb{1}_I \right) h_I \otimes h_J \mathbb{I}\mathbb{I}_2 h_J \right) \\
&\mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \widehat{B}(I \times J) \langle \langle f, h_J \rangle, h_I^1 \rangle h_I |I|^{-1/2} \otimes (h_{J_-} - h_{J_+}) |J|^{-1/2} \right).
\end{aligned}$$

This is a sum of two terms of the form

$$\mathbb{I}\mathbb{I}_1 \left( \sum_I \sum_J \pm \widehat{B}(I \times \tilde{J}) \langle f, h_I^1 \otimes h_{\tilde{J}} \rangle \frac{h_I \otimes h_J}{|I|^{\frac{1}{2}} |\tilde{J}|^{\frac{1}{2}}} \right) = \mathbb{I}\mathbb{I}_1(P_B^2(f)).$$

This concludes the proof of the estimate for the term  $\widetilde{T}_2$ .

### 1.2.1 Remark: Logarithmic estimate

Note that, because of (1.2.1), the previous estimates for the upper bound depend on a dimensional constant. Using a slightly different ordering of the terms in the formal Haar expansion of the product  $Bf$ , we obtain a decomposition in paraproducts of the form

$$\begin{aligned} & \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,0)} \rangle \langle f, h_R^{(0,0)} \rangle h_R^{(1,1)} + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,0)} \rangle \langle f, h_R^{(0,1)} \rangle h_R^{(1,0)} \\ & + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,1)} \rangle \langle f, h_R^{(0,0)} \rangle h_R^{(1,0)} + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(1,0)} \rangle \langle f, h_R^{(0,0)} \rangle h_R^{(0,1)} \\ & + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(1,0)} \rangle \langle f, h_R^{(0,1)} \rangle h_R^{(0,0)} + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(1,1)} \rangle \langle f, h_R^{(0,0)} \rangle h_R^{(0,0)} \\ & + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,0)} \rangle \langle f, h_R^{(1,0)} \rangle h_R^{(0,1)} + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,0)} \rangle \langle f, h_R^{(1,1)} \rangle h_R^{(0,0)} \\ & + \sum_{R \in \mathcal{D}^2} \langle B, h_R^{(0,1)} \rangle \langle f, h_R^{(1,0)} \rangle h_R^{(0,0)} \\ & = (T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8 + T_9)(f). \end{aligned}$$

Here,  $h_R^{(\varepsilon,\delta)} = h_I^\varepsilon h_J^\delta$ , with  $\varepsilon, \delta \in \{0, 1\}$ , and  $h_I^0 = h_I$ ,  $h_I^1 = |I|^{-1/2} \mathbb{1}_I$ . Then,

$$[[M_B, \mathbb{I}\mathbb{I}_1], \mathbb{I}\mathbb{I}_2](f) = [[T_1, \mathbb{I}\mathbb{I}_1], \mathbb{I}\mathbb{I}_2](f) + \cdots [[T_9, \mathbb{I}\mathbb{I}_1], \mathbb{I}\mathbb{I}_2](f).$$

Therefore, to find an upper bound for the commutator, it suffices to find upper bounds for the different paraproducts in the above expansion. By the previous section, this

upper bound depends also on a dimensional constant, however, it is possible for the terms  $T_1$ ,  $T_6$ , and  $T_8$  (by duality), to find a better estimate of order  $\log^2(1+d)$ . This is possible due to a generalization of the results obtained by Pisier in [69] for the one parameter case, combined with the characterization by two index martingales given by Bernard in [4].

With the rest of the terms, it's still not clear how to find this improved dimensional bound for the paraproduct, since we don't have a representation in two-index martingales in these cases, or the appropriate embedding theorem.

### 1.3 Lower bound

The lower bound can be proved by using the result in the scalar case (proved by Ferguson and Lacey in [28]). That, is, there is a constant  $C > 0$  such that

$$\|b\|_{BMO} \leq C \|[[M_b, H_1], H_2]\|_{L^2 \rightarrow L^2},$$

for all scalar functions  $b$  in  $BMO(\mathbb{R}^2)$ . Let us recall the definition of  $BMO$  given in 1.2.1. The lower bound estimate in the matrix-valued setting is the following

**Theorem 1.3.1** (Lower bound). *Let  $B$  be a matrix-valued function on  $\mathbb{R}^2$ , then*

$$d^{-2} \|B\|_{BMO} \lesssim \|[[M_B, H_1], H_2]\|_{L^2(\mathbb{C}^d) \rightarrow L^2(\mathbb{C}^d)}.$$

**Proof:** Denote by  $\widehat{B}(R)$  the wavelet coefficient  $\langle B, v_R \rangle$ . Consider the functions  $f, g \in L^2(\mathbb{C})$ . Let  $\{\vec{e}_1, \dots, \vec{e}_d\}$  represent the canonical basis of  $\mathbb{R}^d$ , then, for  $1 \leq i, j \leq d$ , the functions  $\tilde{f} = f\vec{e}_i$  and  $\tilde{g} = g\vec{e}_j$  both belong to  $L^2(\mathbb{C}^d)$ . If  $B = (b_{ij})$ , an easy computation shows that

$$\left\langle [[M_B, H_1], H_2] \tilde{f}, \tilde{g} \right\rangle_{L^2(\mathbb{C}^d)} = \left\langle [[M_{b_{ji}}, H_1], H_2] f, g \right\rangle_{L^2(\mathbb{C})}$$

Therefore, for every  $i, j \in \{1, \dots, d\}$ , we have

$$\|[[M_{b_{ji}}, H_1], H_2]\|_{L^2(\mathbb{C}) \rightarrow L^2(\mathbb{C})} \leq \|[[M_B, H_1], H_2]\|_{L^2(\mathbb{C}^d) \rightarrow L^2(\mathbb{C}^d)}. \quad (1.3.2)$$

Let  $\{E_{ij} : 1 \leq i, j \leq d\}$  be the canonical basis for the  $d \times d$  matrices, that is,  $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ . We can write  $B = \sum_{i,j} b_{ij}E_{ij}$ , and proceed to find an estimate for the  $BMO$  norm of the matrices  $\tilde{B}_{ij} = b_{ij}E_{ij}$ .

Note that  $\widehat{\tilde{B}_{ij}}(R)\widehat{\tilde{B}_{ij}}(R)^* = \widehat{\tilde{B}_{ij}}(R)^*\widehat{\tilde{B}_{ij}}(R) = \widehat{b_{ij}}(R)E_{ij}\widehat{b_{ij}}(R)E_{ji} = |\widehat{b_{ij}}(R)|^2E_{ii}$ .

Then, for any open set  $U \subseteq \mathbb{R}^2$ , we have

$$\begin{aligned} \frac{1}{|U|} \sum_{R \subseteq U} \widehat{\tilde{B}_{ij}}(R)\widehat{\tilde{B}_{ij}}(R)^* &= \frac{1}{|U|} \sum_{R \subseteq U} |\widehat{b_{ij}}(R)|^2 E_{ii} \\ &\leq \frac{1}{|U|} \sum_{R \subseteq U} |\widehat{b_{ij}}(R)|^2 I_d \leq \|b_{ij}\|_{BMO} I_d. \end{aligned}$$

Using the one parameter result, and equation 1.3.2, we get

$$\begin{aligned} \frac{1}{|U|} \sum_{R \subseteq U} \widehat{\tilde{B}_{ij}}(R)\widehat{\tilde{B}_{ij}}(R)^* &\lesssim \|[[M_{b_{ji}}, H_1], H_2]\|_{L^2(\mathbb{C}) \rightarrow L^2(\mathbb{C})} I_d \\ &\leq \|[[M_B, H_1], H_2]\|_{L^2(\mathbb{C}^d) \rightarrow L^2(\mathbb{C}^d)}. \end{aligned}$$

That is,  $\|\tilde{B}_{ij}\|_{BMO} \lesssim \|[[M_B, H_1], H_2]\|_{L^2(\mathbb{C}^d) \rightarrow L^2(\mathbb{C}^d)}$ . Therefore,

$$\|B\|_{BMO} \leq \sum_{i,j} \|\tilde{B}_{ij}\|_{BMO} \lesssim d^2 \|[[M_B, H_1], H_2]\|_{L^2(\mathbb{C}^d) \rightarrow L^2(\mathbb{C}^d)}.$$

Which is the desired lower bound. ■

**CHAPTER 2**  
**INTRODUCTION TO THE SPARSE THEORY**

**2.1 Definitions and basic concepts**

**Definition 2.1.1.** Let  $0 < c < 1$ , a collection of cubes  $\mathcal{S}$  (usually taken to be dyadic) is said to be  $c$ -sparse (or just sparse, when the particular value of  $c$  is not relevant), if for every cube  $S \in \mathcal{S}$  there is a subset  $E_S \subseteq S$  such that:

1.  $|E_S| > c|S|$  for every  $S \in \mathcal{S}$ .
2.  $\|\sum_{S \in \mathcal{S}} \mathbb{1}_{E_S}\|_\infty < c^{-1}$ .

Here,  $|S|$  represents the Lebesgue measure of  $S$ . The second condition is often made stronger by requiring that the sets  $E_S$  are pairwise disjoint instead.

An equivalent formulation of a sparse collection, and the way in which most of the times this collections are constructed, is the following: For a cube  $S \in \mathcal{S}$ , let  $\text{Ch}_\mathcal{S}(S)$  to be the collection of maximal cubes in  $\mathcal{S}$  that are strictly contained in  $S$ . For a fixed  $0 < c < 1$ , a collection  $\mathcal{S}$  is said to be  $c$ -sparse if for every cube  $S \in \mathcal{S}$  we have

$$\sum_{P \in \text{Ch}_\mathcal{S}(S)} |P| \leq c|S|. \tag{2.1.2}$$

Note that if for every cube  $S$  in  $\mathcal{S}$  we consider the set  $E_S = S \setminus \bigcup_{P \in \text{Ch}_\mathcal{S}(S)} P$ , the collection  $\mathcal{S}$  satisfies the conditions of Definition 2.1.1.

**Definition 2.1.3.** A sparse operator, is an operator of the form

$$\Lambda_\mathcal{S} f(x) = \sum_{S \in \mathcal{S}} \langle f \rangle_S \mathbb{1}_S(x) \tag{2.1.4}$$

Where  $\mathcal{S}$  is a sparse collection, and  $\langle f \rangle_S$  represents the average of  $|f|$  over  $S$ , that is  $\langle f \rangle_S = |S|^{-1} \int_S |f(x)| dx$ .

Related to the sparse operators, we also consider the following bilinear forms:

**Definition 2.1.5.** Let  $\mathcal{S}$  be a sparse collection, and  $r, s \geq 1$  real numbers, we define the sparse  $(r, s)$  form by

$$\Lambda_{\mathcal{S}, r, s}(f, g) = \sum_{S \in \mathcal{S}} \langle f \rangle_{S, r} \langle g \rangle_{S, s} |S| \quad (2.1.6)$$

Here,  $\langle f \rangle_{S, r} = \langle |f|^r \rangle_S^{1/r}$ . We say that an operator  $T$  is in  $\text{Sparse}(r, s)$  if there is a sparse form  $\Lambda_{\mathcal{S}, r, s}$  such that for every  $f, g$  compactly supported, we have

$$\langle Tf, g \rangle \lesssim \Lambda_{\mathcal{S}, r, s}(f, g).$$

In a further chapter, this will be stated in terms of a sparse norm.

## 2.2 Boundedness of the sparse operators

By using a Calderón-Zygmund decomposition of the function  $f$ , it is straightforward to prove that a sparse operator 2.1.4 satisfies a weak 1-1 inequality. These operators are also (strongly) bounded on  $L^p$ , for  $p > 1$ : Let  $f \in L^p$  and  $g \in L^{p'}$ , then,

$$\begin{aligned} \langle \Lambda_{\mathcal{S}} f, g \rangle &= \sum_{S \in \mathcal{S}} \langle f \rangle_S \langle g \rangle_S |S| \leq c \sum_{S \in \mathcal{S}} \langle f \rangle_S \langle g \rangle_S |E_S| = \int \sum_{S \in \mathcal{S}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_{E_S}(x) dx \\ &\leq c \int \sup_Q (\langle f \rangle_Q \langle g \rangle_Q \mathbb{1}_Q(x)) \sum_{S \in \mathcal{S}} \mathbb{1}_{E_S}(x) dx \leq \int \mathcal{M}f(x) \cdot \mathcal{M}g(x) dx \\ &\leq \|\mathcal{M}f\|_{L^p} \|\mathcal{M}g\|_{L^{p'}} \lesssim p p' \|f\|_{L^p} \|g\|_{L^{p'}}. \end{aligned}$$

Here,  $\mathcal{M}$  represents the maximal function, which is known to be bounded in  $L^p$ , for  $p > 1$  and weakly bounded for  $p = 1$ . This dependence upon the index  $p$  is sharp.

Note also, that from the very first line in the previous computation, it follows that if an operator  $T$  is in  $\text{Sparse}(1, 1)$ , then it is bounded on  $L^p$  for  $p > 1$ . By using restricted weak type estimates, it can be proved that domination by a sparse bilinear form implies weak 1-1 estimates; a proof of this fact can be found in [21].

In general, a modification of the previous computation can be used to prove that if  $T$  is in  $\text{Sparse}(r, s)$  with  $1 \leq r < s'$ , then  $T$  is bounded on  $L^p$  for every  $p \in (r, s')$ .

### 2.2.1 Weighted inequalities

A function  $w$  is called a weight, if it is nonnegative and locally integrable. A weight  $w$  is in the class  $A_p$  if

$$[w]_{A_p} = \sup_Q \langle w \rangle_Q \left\langle w^{1-p'} \right\rangle_Q^{p-1} < \infty,$$

where the supremum is taken over all cubes  $Q$  in  $\mathbb{R}^n$ . The quantity  $[w]_{A_p}$  is called the  $A_p$  characteristic of the weight  $w$ .

The space  $L^p(w)$  is defined as the space of functions  $f$  that satisfy

$$\|f\|_{L^p(w)}^p = \int |f(x)|^p w(x) dx < \infty.$$

With a similar proof as in the unweighted case, we can verify that if an operator  $T$  is in  $\text{Sparse}(r, s)$ , for  $1 \leq r < s'$ , then, for every  $p \in (r, s')$ , then  $T$  is bounded on  $L^p(w)$ , for weights  $w$  belonging to an intersection of a special class  $A_{p_0}$  and a reverse-Hölder class. For a more precise statement, and a detailed proof of this boundedness, the reader can check section 6 in [5].

**CHAPTER 3**  
**THE SPARSE T1 THEOREM**

(Joint work with Michael Lacey)

**3.1 Introduction**

We recast the statement of the  $T1$  theorem of David and Journé [25], replacing the conclusion that the operator  $T$  admits a quantitative bound on its  $L^2$ -norm, with the conclusion that  $T$  admits a quantitative sparse bound. From the sparse bound, one can quickly derive a wide range of (weighted)  $L^p$  type inequalities for  $T$ . That is, the theory devoted to deriving these properties for  $T$  can be replaced by the much simpler approach via sparse operators.

We say that an operator  $T$  is a Calderón-Zygmund operator on  $\mathbb{R}^d$  if (a) it is bounded on  $L^2$ , (b) there is a kernel  $K(x, y) : \mathbb{R}^d \times \mathbb{R}^d \setminus \{(x, x) : x \in \mathbb{R}^d\} \rightarrow \mathbb{R}$  so that for functions  $f, g$  smooth, compactly supported, have disjoint closed supports,

$$B_T(f, g) = \langle Tf, g \rangle = \iint K(x, y) f(y) g(x) dy dx.$$

(c) For some constant  $\mathcal{K}_T$ , the kernel  $K(x, y)$  satisfies

$$|K(x, y)| \leq \frac{\mathcal{K}_T}{|x - y|}, \quad x \neq y \in \mathbb{R}^d, \quad (3.1.1)$$

$$|K(x, y) - K(x', y)| < \mathcal{K}_T \frac{|x - x'|^\eta}{|x - y|^{d+\eta}}, \quad 0 < 2|x - x'| < |x - y|. \quad (3.1.2)$$

And, the same condition with the roles of  $x$  and  $y$  reversed. Above,  $\eta > 0$  is a fixed small constant.

A sparse bilinear form  $\Lambda(f, g)$  is defined this way: There is a collection of cubes

$\mathfrak{S}$ , so that for each  $S \in \mathfrak{S}$ , there is an  $E_S \subset S$  so that (a)  $|E_S| > c|S|$ , and (b)  $\|\sum_{S \in \mathfrak{S}} \mathbb{1}_{E_S}\|_\infty \leq c^{-1}$ . Then, set

$$\Lambda(f, g) = \sum_{S \in \mathfrak{S}} \langle f \rangle_S \langle g \rangle_S |S|,$$

where  $\langle f \rangle_S = |S|^{-1} \int_S f(x) dx$ . Here, we will not focus on the role of the constant  $0 < c < 1$ , and remark that many times it is assumed that the sets  $E_S$  being pairwise disjoint, that is  $\|\sum_{S \in \mathfrak{S}} \mathbb{1}_{E_S}\|_\infty = 1$ .

Our generalization does not affect the outlines of the theory, and makes some arguments somewhat simpler.

It is very useful to think of  $\Lambda(f, g)$  as a *positive bilinear Calderón-Zygmund form*. In particular, all the standard inequalities can be quickly proved for  $\Lambda$ . And, for weighted inequalities, it is easy to derive bounds that are sharp in the  $A_p$  characteristic.

Our formulation of the  $T1$  theorem considers the usual  $L^1$  testing condition on  $T$ , phrased in bilinear language.

**Theorem 3.1.3.** *Suppose that  $T$  is a Calderón-Zygmund operator on  $\mathbb{R}^d$ , and moreover there is a constant  $\mathbf{T}$  so that for all cubes  $Q$  and functions  $|\phi| < \mathbb{1}_Q$ , there holds*

$$|B_T(\mathbb{1}_Q, \phi)| + |B_T(\phi, \mathbb{1}_Q)| \leq \mathbf{T}|Q|. \quad (3.1.4)$$

*Then there is a constant  $C = C(\mathcal{K}_T, \mathbf{T}, d, \eta)$  so that for all bounded compactly supported functions  $f, g$ , there is a sparse operator  $\Lambda$  so that*

$$|B_T(f, g)| < C\Lambda(|f|, |g|). \quad (3.1.5)$$

The proof is elementary, using (a) facts about averages and conditional expectations; (b) random dyadic grids as a convenient tool to reduce the complexity of

the argument; (c) orthogonality of martingale transforms, and the most sophisticated fact (d) a sparse bound for a certain bilinear square function, with complexity, detailed in Lemma 3.4.6. In addition, the testing condition (3.1.4) appears solely in the construction of the stopping times. The proof is carried out in §3.3. There are many terms, organized so that there is one crucial term, in §3.3.2. Almost all of the remaining cases use standard off-diagonal considerations, and the simple argument to prove the sparse bound for a martingale transform. This is detailed in §3.4.

The consequences of the sparse bound (3.1.5) are:

1. The weak type  $(1, 1)$  inequality, and the  $L^p$  inequalities, for  $1 < p < \infty$ . These hold with the sharp dependence upon  $p$ . To wit, using  $\|M : L^p \mapsto L^p\| \lesssim p' = \frac{p}{p-1}$ , we have

$$\begin{aligned} \Lambda(f, g) &= \int \sum_{S \in \mathfrak{S}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_S \, dx \lesssim \int \sum_{S \in \mathfrak{S}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_{E_S} \, dx \\ &\leq \int Mf \cdot Mg \, dx \leq \|Mf\|_p \|Mg\|_{p'} \lesssim p \cdot p' \|f\|_p \|g\|_{p'}. \end{aligned}$$

2. The weighted version of the same, relative to  $A_p$  weights. The dependence upon the  $A_p$  characteristic is sharp, for  $1 < p < \infty$ , and the best known for the case of  $p = 1$ . See the arguments in [54].
3. The exponential integrability results of Karagulyan [38, 65].

Our statement of the  $T1$  theorem follows the ‘testing inequality’ approach of the Sawyer two weight theorems [70, 71], and the statement in Stein’s monograph [73]. Our approach is a descendant of the radically dyadic approach of Figiel [30], further influenced by the martingale approach of Nazarov-Treil-Volberg [64]. (Also see [34].) Our use of the stopping cubes follows that of the proof of the two weight Hilbert transform estimate [51].

The bound by sparse operators has been an active and varied recent research

topic. It had a remarkable success in Lerner’s approach to the  $A_2$  bound [54], which cleverly bounded on the weighted norm of a Calderón-Zygmund by a the norm of a sparse operator. The pointwise approach first established in [17], with a somewhat different approach in [45]. The latter approach has been studied from several different points of view [5, 26, 55, 78]. The form approach used here, is however successful in settings where the pointwise approach would fail, most notably the setting of the bilinear Hilbert transform [21], Bochner Riesz multipliers [3], and oscillatory singular integrals [50]. The interested reader can consult the papers above for more information and references.

This paper proves the sparse bound without appealing to any structural theory of Calderón-Zygmund operators such as boundedness of maximal truncations, which is the approach started in [45]. The other prominent structural fact one could use is the Hytönen structure theorem [33]. This is the approach followed by Culiuc-Di Plinio-Ou [22] also using bilinear forms. They show that this approach has further applications to the matricial setting, avoiding difficulties for the pointwise approach in this setting.

### 3.2 Random Grids

All the proofs here will use Hytönen’s random dyadic grids from [33]. Recall again, that the standard dyadic grid in  $\mathbb{R}^d$  is

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

For a binary sequence  $\omega := (\omega_j)_{j \in \mathbb{Z}} \in (\{0, 1\}^d)^{\mathbb{Z}}$  we define a general dyadic system by

$$\mathcal{D}^\omega := \{ Q \dot{+} \omega : Q \in \mathcal{D}^0 \},$$

where  $Q \dot{+} \omega = Q + \sum_{j:2^{-j} < \ell Q} 2^{-j} \omega_j$ . We consider the standard uniform probability measure on  $\{0, 1\}^d$ , that is, it assigns  $2^{-d}$  to every point. We place on  $\omega$ , the probability measure  $\mathbb{P}$ , the corresponding product measure on  $(\{0, 1\}^d)^{\mathbb{Z}}$ . This way, we can see  $(\mathcal{D}_\omega)$  as a collection of grids with a random set of parameters  $\omega$ . For every  $\omega$ , these dyadic grids satisfy the required properties, namely

1. For  $P, Q \in \mathcal{D}^\omega$ ,  $P \cap Q \in \{P, Q, \emptyset\}$ .
2. For fixed  $k \in \mathbb{Z}$ , the collection  $\mathcal{D}_k^\omega = \{Q \in \mathcal{D}^\omega : \ell Q = 2^{-k}\}$  partitions  $\mathbb{R}^d$ .

**Definition 3.2.1** (Good-bad intervals). Let  $0 < \gamma < 1$  and a positive integer  $r$  such that  $r \geq (1 - \gamma)^{-1}$ . We say that  $Q \in \mathcal{D}_k^\omega$  is  $r$ -bad, if there is an integer  $s \geq r$ , and a choice of coordinate, so that the vectors

$$\omega_{k+\lfloor(1-\gamma)s\rfloor}, \omega_{k+\lfloor(1-\gamma)s\rfloor+1}, \dots, \omega_{k+s} \in \{0, 1\}^d,$$

all agree in that one coordinate. If  $Q$  is not  $r$ -bad, then it is called  $r$ -good.

From now on, we are going to omit the dependence on  $r$ , and we will refer to the cubes as only good or bad. The following lemmas are well known.

**Lemma 3.2.1.** *If  $Q$  is good, then for any cube  $P$  with  $2^r \ell Q < \ell P$  we have*

$$\text{dist}(Q, \partial P) \gtrsim (\ell Q)^\gamma (\ell P)^{1-\gamma},$$

where the implied constant is absolute.

**Lemma 3.2.2.** *Fix  $0 < \gamma < 1$  and  $r > \gamma^{-1}$ , then, there is a constant  $C_d$  such that*

$$\mathbb{P}(Q \text{ is good}) \geq 1 - C_d \gamma^{-1} 2^{-\gamma r}.$$

For an arbitrary dyadic grid  $\mathcal{D}^\omega$ , every function  $f \in L^2(\mathbb{R}^d)$  admits an orthogonal

decomposition

$$f = \sum_{Q \in \mathcal{D}^\omega} \Delta_Q f.$$

Given a dyadic grid  $\mathcal{D}^\omega$ , we define the good and bad projections as

$$P_\omega^{\text{bad}} f := \sum_{\substack{Q \in \mathcal{D}^\omega \\ Q \text{ is bad}}} \Delta_Q f, \quad P_\omega^{\text{good}} f := \sum_{\substack{Q \in \mathcal{D}^\omega \\ Q \text{ is good}}} \Delta_Q f.$$

The following lemma says that in average, the bad projections tend to be small.

**Lemma 3.2.3.** *For all  $1 < p < \infty$  there is an  $\epsilon_p > 0$  such that for all  $0 < \gamma < 1$  and  $r > \gamma^{-1}$  we have*

$$\mathbb{E}_\omega \|P_\omega^{\text{bad}} f\|_{L^p}^p \lesssim 2^{-\epsilon_p r} \|f\|_{L^p}^p.$$

Using this lemma, we can prove that it is enough to estimate bounds only for *good* functions, in the following sense

**Lemma 3.2.4.** *Let  $1 < p < \infty$ . If  $T : L^p \mapsto L^p$  is a bounded operator. If  $0 < \gamma < 1$  is fixed and  $r > C(1 + \log \frac{1}{\gamma})$ , then*

$$\|T : L^p \mapsto L^p\| \leq 4M,$$

where  $M$  is the best constant in the inequality

$$\mathbb{E}_\omega |\langle T P_\omega^{\text{good}} f, P_\omega^{\text{good}} g \rangle| \leq M \|f\|_{L^p} \|g\|_{L^{p'}}.$$

### 3.3 The Proof of the Sparse Bound

As a consequence of Lemma 3.2.4, it is enough for the remainder of the argument to show this: There is a choice of constant  $C > 1$ , so that for all  $f$  and  $g$  compactly

supported, and almost all grids  $\mathcal{D}^\omega$ , there is a sparse operator  $\Lambda = \Lambda_{f,g,\mathcal{D}^\omega}$ , so that

$$|\langle TP^{\text{good}} f, P^{\text{good}} g \rangle| \leq C\Lambda(|f|, |g|). \quad (3.3.1)$$

In view of the Lemma 3.4.9, the random sparse operator above can be replaced by a deterministic one. Averaging over choices of grid will complete the proof.

Almost all random dyadic grids have the property that the functions  $f, g$  are supported on a single good dyadic cube. And, hence, on a sequence of dyadic cubes which exhaust  $\mathbb{R}^n$ . This fact and goodness are the only facts about random grids utilized, so we suppress the  $\omega$  dependence below. The inner product in (3.3.1) is expanded

$$\langle TP^{\text{good}} f, P^{\text{good}} g \rangle = \sum_{\substack{P \in \mathcal{D} \\ P \text{ is good}}} \sum_{\substack{Q \in \mathcal{D} \\ Q \text{ is good}}} \langle T\Delta_P f, \Delta_Q g \rangle. \quad (3.3.2)$$

We will further only consider the case of  $\ell P \geq \ell Q$ , the reverse case being addressed by duality. The fact that  $P$  and  $Q$  are good will be suppressed, but always referenced when it is used. And, by  $Q \Subset P$  we will mean that  $Q \subset P$  and  $2^r \ell Q \leq \ell P$ . Goodness of  $Q$  then implies that

$$\text{dist}(Q, \text{skel}P) \geq (\ell Q)^\epsilon (\ell P)^{1-\epsilon}, \quad (3.3.3)$$

where  $\text{skel}P$  is the union of  $\partial P'$ , where  $P'$  is a child of  $P$ . We will likewise suppress the role of the dyadic grid in our notation.

As just mentioned, the two functions  $f, g$  are supported on a single good cube  $P_0 \in \mathcal{D}$ , which we can take to be very large. Therefore, we can restrict the sum in (3.3.2) to only cubes  $P, Q \subset P_0$ . The bound we obtain will be independent of the

choice of  $P_0$ . The sum we consider is then broken into several subcases.

$$\sum_{P: P \subset P_0} \sum_{\substack{Q: Q \subset P_0 \\ \ell P \geq \ell Q}} \langle T \Delta_P f, \Delta_Q g \rangle \quad (3.3.4)$$

$$= \sum_{P: P \subset P_0} \sum_{Q: Q \in P} \langle T \Delta_P f, \Delta_Q g \rangle \quad (\text{inside}) \quad (3.3.5)$$

$$+ \sum_{P: P \subset P_0} \sum_{\substack{Q: 2^r \ell Q \leq \ell P \\ Q \subset 3P \setminus P}} \langle T \Delta_P f, \Delta_Q g \rangle \quad (\text{near}) \quad (3.3.6)$$

$$+ \sum_{P: P \subset P_0} \sum_{\substack{Q: \ell Q \leq \ell P \\ Q \cap 3P = \emptyset}} \langle T \Delta_P f, \Delta_Q g \rangle \quad (\text{far}) \quad (3.3.7)$$

$$+ \sum_{P: P \subset P_0} \sum_{\substack{Q: \ell Q \leq \ell P \leq 2^r \ell Q \\ Q \cap 3P \neq \emptyset}} \langle T \Delta_P f, \Delta_Q g \rangle. \quad (\text{neighbors}) \quad (3.3.8)$$

### 3.3.1 Stopping Cubes

We define a sparse collection  $\mathcal{S}$  of stopping cubes, and associated stopping values in the following way: Add  $P_0$  to the collection  $\mathcal{S}$ , and set  $\sigma_f(P_0) = \langle |f| \rangle_{P_0}$ , and similarly for  $g$ . In the recursive stage of the construction, for minimal  $S \in \mathcal{S}$ , define three sets

- $F_S^1 = \bigcup \{ S' \in \mathcal{D}(S) : \langle |f| \rangle_{S'} > C_0 \sigma_f(S), S' \text{ maximal} \}.$
- $F_S^2 = \bigcup \{ S' \in \mathcal{D}(S) : \langle |g| \rangle_{S'} > C_0 \sigma_g(S), S' \text{ maximal} \}.$
- $F_S^3 = \bigcup \{ S' \in \mathcal{D}(S) : \langle |T \mathbb{1}_S| \rangle_{S'} > C_0 \mathbf{T}, S' \text{ maximal} \}.$

Let  $F_S = F_S^1 \cup F_S^2 \cup F_S^3$ , and  $\mathcal{F}_S$  be the family of dyadic components of  $F_S$ . The weak-type bound for the dyadic maximal function and the testing condition (3.1.4) implies that there exists  $C_0$  big enough, such that  $|F_S| < \frac{1}{2}|S|$ . Recursively, add, every  $\mathcal{F}_S$  to the collection  $\mathcal{S}$  to form a sparse collection.

We set  $P^\sigma$  to be the smallest stopping cube  $S$  that contains  $P$ . And we set  $Q^\tau$  to be the smallest stopping cube  $S$  such that  $Q \Subset S$ . The Haar projection associated to  $S$  is  $\Pi_S g = \sum_{Q: Q^\tau = S} \Delta_Q g$ .

### 3.3.2 The Inside Terms

We turn our attention to the main term, that of (3.3.5), for which there are three subcases. The argument of  $T$  is  $\Delta_P f$ , which we write as

$$\Delta_P f = \Delta_P f \mathbb{1}_{P \setminus P_Q} + \mathbb{1}_{P_Q} \Delta_P f \quad (3.3.9)$$

$$= \Delta_P f \mathbb{1}_{P \setminus P_Q} + \langle \Delta_P f \rangle_{P_Q} \cdot \begin{cases} \mathbb{1}_S - \mathbb{1}_{S \setminus P_Q} & S = Q^\tau \supset P_Q \\ \mathbb{1}_S + \mathbb{1}_{P_Q \setminus S} & S = Q^\tau \subsetneq P_Q \end{cases} \quad (3.3.10)$$

where  $Q \in P$ , and  $P_Q$  is the child of  $P$  that contains  $Q$ .

#### *First Subcase*

Control the first term on the right in (3.3.10) by off-diagonal considerations. Central to all of these off-diagonal arguments are the class of forms  $B^{u,v}$  defined in (3.4.1), which are in turn bounded by Lemma 3.4.6.

Since  $Q$  is a good cube, the inequality (3.3.3) holds: That is  $Q$  is a long way from the skeleton of  $P$ . By (3.4.12), we have

$$|\langle T(\Delta_P f \mathbb{1}_{P \setminus P_Q}), \Delta_Q g \rangle| \lesssim P_\eta(\Delta_P f \mathbb{1}_{P \setminus P_Q})(Q) \|\Delta_Q g\|_1 \quad (3.3.11)$$

$$\lesssim [\ell Q / \ell P]^{\eta'} \langle |\Delta_P f| \rangle_P \|\Delta_Q g\|_1. \quad (3.3.12)$$

Using the notation of (3.4.1), for integers  $v \geq r$ , we have

$$\sum_P \sum_{\substack{Q: Q \subset P \\ 2^v \ell Q = \ell P}} |\langle T(\Delta_P f \mathbb{1}_{P \setminus P_Q}), \Delta_Q g \rangle| \lesssim 2^{-\eta' v} B^{0,v}(f, g)$$

and by Lemma 3.4.6, this is in turn dominated by a choice of sparse form. Sparse forms are again dominated by a fixed form. We can sum this estimate over  $v \geq r$ , so this case is complete.

*Second Subcase*

We turn attention to the second term in (3.3.10), in which we have  $\langle \Delta_P f \rangle_{P_Q} \mathbb{1}_S$ . This is the most intricate step, in that we combine several elementary steps. The bound we prove is uniform over a choice of  $S \in \mathcal{S}$ . Namely,

$$\left| \sum_{Q: Q^\tau=S} \sum_{P: Q \in P} \langle T(\Delta_P f \cdot \mathbb{1}_S), \Delta_Q g \rangle \right| \lesssim \langle |f| \rangle_S \langle |g| \rangle_S |S| \quad (3.3.13)$$

This is the one point in the argument in which the implied constant depends upon the testing constant  $\mathbf{T}$  in (3.1.4).

For each cube  $Q$  with  $Q^\tau = S$ , define  $\epsilon_Q$  by

$$\epsilon_Q \langle |f| \rangle_S := \sum_{P \in \mathcal{D}, Q \in P_Q} \langle \Delta_P f \rangle_{P_Q}. \quad (3.3.14)$$

By the first stopping condition, corresponding to the control of the averages of  $f$ ,  $\{\epsilon_Q\}_{Q \in \mathcal{D}}$  is uniformly bounded. In particular, this operator is a martingale transform.

$$\Pi_S^\epsilon g = \sum_{Q: Q^\tau=S} \epsilon_Q \Delta_Q g.$$

We make the following observation about the second stopping condition, corresponding to the control of the averages of  $g$ . Setting a conditional expectation on  $S$  to be

$$\mathbb{E}(\phi \mid \mathcal{F}_S) = \begin{cases} \phi(x) & x \in S \setminus F_s \\ \langle \phi \rangle_{S'} & x \in S', S' \in \mathcal{F}_S \end{cases}$$

Then,  $\|\mathbb{E}(g \mathbb{1}_S \mid \mathcal{F}_S)\|_\infty \lesssim \langle |g| \rangle_S$ . We also have  $\Pi_S^\epsilon g = \Pi_S^\epsilon \mathbb{E}(g \mathbb{1}_S \mid \mathcal{F}_S)$ . Therefore, by the  $L^2$  bound for martingale transforms,

$$\|\Pi_S^\epsilon g\|_2 \leq \|\mathbb{E}(g \mathbb{1}_S \mid \mathcal{F}_S)\|_2 \lesssim \langle |g| \rangle_S |S|^{1/2}. \quad (3.3.15)$$

The point of our third stopping condition, corresponding to the control of the average of  $T\mathbb{1}_S$ , is that  $\mathbb{E}(T\mathbb{1}_S \mid \mathcal{F}_S)$  is bounded in  $L^\infty$  by a constant multiple of  $\mathbf{T}$ . Collecting these observations, we can rewrite our sum as below, in which in the first step we use the definition (3.3.14) to collapse the sum over  $P$ .

$$\text{LHS of (3.3.13)} = |\langle |f| \rangle_S \langle T\mathbb{1}_S, \Pi_S^\epsilon g \rangle| \quad (3.3.16)$$

$$= |\langle |f| \rangle_S \langle T\mathbb{1}_S, \mathbb{E}(\Pi_S^\epsilon g \mid \mathcal{F}_S) \rangle| \quad (3.3.17)$$

$$= |\langle |f| \rangle_S \langle \mathbb{E}(T\mathbb{1}_S \mid \mathcal{F}_S), \Pi_S^\epsilon g \rangle| \quad (3.3.18)$$

$$\lesssim \langle |f| \rangle_S \|\mathbb{E}(T\mathbb{1}_S \mid \mathcal{F}_S)\|_2 \|\Pi_S^\epsilon g\|_2 \lesssim \langle |f| \rangle_S \langle |g| \rangle_S |S|. \quad (3.3.19)$$

This completes this case.

### *Third Subcase*

We address the top alternative in (3.3.10), namely we bound

$$\sum_S \sum_{Q: Q^\tau=S} \sum_{\substack{P: Q \in P \\ P_Q \subset S}} \langle \Delta_P f \rangle_{P_Q} \langle T\mathbb{1}_{S \setminus P_Q}, \Delta_Q g \rangle \quad (3.3.20)$$

This is similar to the first subcase, since  $\mathbb{1}_{S \setminus P_Q}$  is supported in  $(2Q)^c$ , then the off-diagonal estimates also imply

$$|\langle T\mathbb{1}_{S \setminus P_Q}, \Delta_Q g \rangle| \lesssim P_\eta(\mathbb{1}_{S \setminus P_Q})(Q) \|\Delta_Q g\|_1 \lesssim \left[ \frac{\ell Q}{\ell P} \right]^{\eta'} \|\Delta_Q g\|_1.$$

Holding the relative lengths of  $Q$  and  $P$  fixed, we then have for integers  $v \geq r$ ,

$$\sum_S \sum_{Q: Q^\tau=S} \sum_{\substack{P: Q \in P \\ P_Q \subset S, 2^v \ell Q = \ell P}} |\langle T(\Delta_P f \mathbb{1}_{S \setminus P_Q}), \Delta_Q g \rangle| \lesssim 2^{-v\eta'} B^{0,v}(f, g).$$

We use the notation (3.4.1), and Lemma 3.4.6 to complete this case.

*Fourth Subcase*

We address the bottom alternative in (3.3.10), namely the case in which  $S = Q^\tau \subsetneq P_Q$ . The point here is to gain geometric decay in the degree to which  $Q$  and  $P_Q$  are separated in the stopping tree  $\mathfrak{S}$ .

Given  $S \in \mathfrak{S}$ , let  $S = S^{(0)} \subsetneq S^{(1)} \subsetneq \cdots \subsetneq P_0$  be the maximal chain of stopping cubes which contain  $S$ , and continue up to  $P_0$ . For each  $S_0 \in \mathfrak{S}$ , and integer  $t \geq 1$ , we bound

$$\left| \sum_{S : S^{(t)}=S_0} \sum_{P : S^{(t-1)} \in P_Q \subset S_0} \langle \Delta_P f \rangle_{P_Q} \langle T \mathbb{1}_{P_Q \setminus S^{(t-1)}}, \Pi_S g \rangle \right| \lesssim 2^{-ct} \langle |f| \rangle_{S_0} \langle |g| \rangle_{S_0} |S_0|. \quad (3.3.21)$$

The point is to use the off-diagonal estimates, but there is a complication in that the stopping cubes are not good. To address this, we let  $\mathcal{Q}(S)$  be the maximal good cubes with  $Q^\tau = S$ , and set

$$\tilde{\Pi}_{Q^*} g = \sum_{Q : Q^\tau=S, Q \subset Q^*} \Delta_Q g, \quad Q^* \in \mathcal{Q}(S).$$

The goodness of the cubes implies that  $\text{dist}(Q^*, \partial S^{(t-1)}) \geq (\ell Q^*)^\epsilon (\ell S^{(t-1)})^{1-\epsilon} \geq 2^{t/2} \ell Q^*$ , by (3.3.3).

The second point is that we have

$$\left\| \sum_{P : S^{(t-1)} \in P_Q \subset S_0} \langle \Delta_P f \rangle_{P_Q} \mathbb{1}_{P_Q \setminus S^{(t-1)}} \right\|_\infty \lesssim \langle |f| \rangle_S.$$

Combining these last two observations with (3.4.14), we see that for each  $Q^* \in \mathcal{Q}(S)$ ,

$$\begin{aligned} \left| \sum_{S : S^{(t)}=S_0} \sum_{P : S^{(t-1)} \in P_Q \subset S_0} \langle \Delta_P f \rangle_{P_Q} \langle T \mathbb{1}_{P_Q \setminus S^{(t-1)}}, \tilde{\Pi}_{Q^*} g \rangle \right| &\lesssim 2^{-t/2} \langle |f| \rangle_{S_0} \|\tilde{\Pi}_{Q^*} g\|_1 \\ &\lesssim 2^{-t/2} \langle |f| \rangle_{S_0} \langle |g| \rangle_S |Q^*|. \end{aligned}$$

Here we have used the stopping condition to dominate  $\tilde{\Pi}_{Q^*}g$ . To conclude, we simply observe that

$$\sum_{S : S^{(\ell)}=S_0} \langle |g| \rangle_S \sum_{Q^* \in \mathcal{Q}(S)} |Q^*| \leq \sum_{S : S^{(\ell)}=S_0} \langle |g| \rangle_S |S| \lesssim \langle |g| \rangle_{S_0} |S_0|.$$

Our proof of (3.3.21) is complete.

### 3.3.3 The Near Terms

We address the term in (3.3.6). Fix an integer  $v \geq r$ , and consider  $Q \subset 3P \setminus P$  with  $2^v \ell Q = \ell P$ . The cube  $Q$  is good, so that by (3.3.3) and (3.4.12), we have

$$|\langle T \Delta_P f, \Delta_Q g \rangle| \lesssim 2^{-v\eta'} \langle |\Delta_P f| \rangle_P \|\Delta_Q g\|_1.$$

But, then, we have

$$|(3.3.6)| \lesssim 2^{-v\eta'} B^{0,v}(f, g),$$

where the latter bilinear form is defined in (3.4.1). It follows from (3.4.1) that the near term is dominated by a sparse bilinear form.

### 3.3.4 The Neighbors

We bound the term in (3.3.8). For  $P$ , let  $P', P''$  be choices children of  $P$ . There are at most  $O(1)$  such choices. For integers  $0 \leq v \leq r$ , we bound

$$\sum_{P : P \subset P_0} \sum_{Q : \ell Q \leq \ell P = 2^v \ell Q, Q \cap 3P \neq \emptyset} \langle T(\Delta_P f \cdot \mathbb{1}_{P'}), \mathbb{1}_{P''} \Delta_Q g \rangle. \quad (3.3.22)$$

The case of  $P' \neq P''$  is straight forward. The function  $\Delta_P f \cdot \mathbb{1}_{P'}$  is constant, so that the Hardy inequality immediately implies that

$$|\langle T(\Delta_P f \cdot \mathbb{1}_{P'}), \mathbb{1}_{P''} \Delta_Q g \rangle| \lesssim |\langle \Delta_P \rangle_{P'}| |P'|^{1/2} \|\mathbb{1}_{P''} \Delta_Q g\|_2$$

$$\lesssim |\langle \Delta_P \rangle_{P'}| \cdot \|\Delta_Q g\|_1.$$

And this can be summed to the bound we want. Namely, it is dominated by  $B^{0,v}(f, g)$ , where the last term is defined in (3.4.1).

The case of  $P' = P''$  reduces to the testing inequality, and we have the same bound as above.

### 3.3.5 The Far Term

We address the terms in (3.3.7). For integers  $u, v \geq 1$ , we impose additional restrictions on  $P$  and  $Q$ , and obtain a sparse bound with geometric decay in these parameters. From this, the required bound follows. Namely, we have for

$$\ell P = \ell P', \quad P' \subset 3^{u-1}P, \quad 2^v \ell Q = \ell P, \quad Q \subset 3^{u+1}P \setminus 3^u P, \quad (3.3.23)$$

we have from (3.4.12) the estimate below.

$$|\langle T \Delta_{P'} f, \Delta_Q g \rangle| \lesssim 2^{-\eta'(u+v)} \langle |\Delta_{P'} f| \rangle_{P'} \|\Delta_Q g\|_1.$$

Therefore, appealing to the definition in (3.4.1)

$$\sum_P \sum_{(P', Q) \text{ satisfy (3.3.23)}} |\langle T \Delta_{P'} f, \Delta_Q g \rangle| \lesssim 2^{-\eta'(u+v)} B^{u,v}(f, g).$$

By Lemma 3.4.6, this case is complete.

## 3.4 Lemmas

We collect three separate groups of Lemma, (a) the sparse domination of a class of dyadic forms; (b) standard off-diagonal estimates; and (c) a Hardy inequality.

## Sparse Domination

We define a class of (sub) bilinear forms that are basic to the proof. For a cube  $P$ , let  $i_P = \log_2(\ell P)$ . Let  $D_k f = \sum_{P: \ell P=2^k} \Delta_P f$ , and define

$$B^{u,v}(f, g) = \sum_P \langle |D_{i_P-u} f| \rangle_{3P} \langle |D_{i_P-v} g| \rangle_{3P} |P| \quad (3.4.1)$$

Above,  $u, v \geq 0$  are fixed integers, so that we are taking the martingale differences that are somewhat smaller, over the triple of  $P$ . We comment that this is a dyadic operator of complexity  $u + v$ , in the language of [33].

We remark that a standard argument would write

$$B^{u,v}(f, g) = \int \sum_P \langle |D_{i_P-u} f| \rangle_{3P} \langle |D_{i_P-v} g| \rangle_{3P} \mathbb{1}_P(x) dx \quad (3.4.2)$$

It is clear that we would dominate this last integral by a product of square functions  $\int S_u f \cdot S_v g dx$ , with the square functions defined by

$$(S_u f)^2 = \sum_P \langle |D_{i_P-u} f| \rangle_{3P}^2 \mathbb{1}_P. \quad (3.4.3)$$

The deepest fact needed in our proof of the  $T1$  theorem is this: The square functions  $S_u$  are weakly bounded, with constant linear in  $u$ .

**Lemma 3.4.4.** *We have the inequality below, valid for all integers  $u \geq 0$*

$$\|S_u f : L^1 \mapsto L^{1,\infty}\| \lesssim (1 + u). \quad (3.4.5)$$

*Proof.* The square function  $S_u$  is bounded on  $L^2$ , with constant independent of  $u$ , by the orthogonality of martingale differences. To prove the weak-type inequality, we take  $f \in L^1$ , and apply the Calderón-Zygmund decomposition at height 1. Thus,

$f = g + b$ , where  $\|g\|_2 \lesssim \|f\|_1^{1/2}$ , and we have

$$b = \sum_{B \in \mathcal{B}} b_B,$$

where  $\mathcal{B}$  consists of disjoint dyadic cubes with  $\sum_{B \in \mathcal{B}} |B| \lesssim \|f\|_1$ , and  $b_B$  is supported on  $B$ , has integral zero, and  $\|b_B\|_1 \lesssim |B|$ .

We do not estimate  $S_u f$  on the set  $E = \bigcup_{B \in \mathcal{B}} 3B$ . And estimate

$$|\{x \notin E : S_u f(x) > 2\}| \leq |\{S_u g > 1\}| + |\{x \notin E : S_u b(x) > 1\}|$$

The first term is controlled by the  $L^2$  bound and the fact that  $\|g\|_2^2 \leq \|f\|_1$ .

Concerning the function  $b$ , observe that for  $P \not\subset E$ , that we have  $\langle |D_{i_P - u} f| \rangle_{3P} \neq 0$  only if there is some  $B \in \mathcal{B}$  with  $B \subset 3P$ , and  $2^u \ell B \geq \ell P$ . For a fixed  $B$ , there are only  $3^d(1+u)$  such choices of  $P$ . Therefore, we will estimate

$$\begin{aligned} |\{x \notin E : S_u b(x) > 1\}| &\lesssim \sum_{P: P \not\subset E} \int_P |\Delta b| \, dx \\ &\lesssim \sum_{v=1}^u \sum_{P: P \not\subset E} \sum_{\substack{B \in \mathcal{B} \\ 2^v \ell B = \ell P}} \int_P |\Delta b_B| \, dx \\ &\lesssim \sum_{v=1}^u \sum_{P: P \not\subset E} \sum_{\substack{B \in \mathcal{B} \\ 2^v \ell B = \ell P}} |B| \lesssim u \sum_{B \in \mathcal{B}} |B| \lesssim u \|f\|_1. \end{aligned}$$

Our proof is complete. □

The previous estimate is the principal tool in this sparse bound, which we use repeatedly in our proof of the sparse result.

**Lemma 3.4.6.** *For all  $u, v \geq 0$ , all bounded compactly supported functions  $f, g$ , there is a sparse collection  $\mathcal{S}$  so that*

$$B^{u,v}(f, g) \lesssim (1+u)(1+v)\Lambda(f, g).$$

It is an easy corollary from the conclusion above for  $u, v = 0$  that martingale transforms satisfy a sparse bound. And, we also comment that the linear dependence of the constant above presents no difficulty in application, as we will always have a term that decreases geometrically in  $u + v$ .

*Proof.* Note that from the equality for  $B^{u,v}$  in (3.4.2), we have

$$B^{u,v}(f, g) \lesssim \int S_u f \cdot S_v g \, dx$$

with the square functions defined by (3.4.3). But, we localize this familiar argument.

Define

$$(S_{u,P_0} f)^2 = \sum_{P: P \subset P_0} \langle |D_{i_P - u} f| \rangle_{3P}^2 \mathbb{1}_P,$$

we have for an absolute constant  $C$ , and all choices of  $u \geq 0$ ,

$$|\{x \in 3P_0 : S_{u,P_0} f > C(1+u)\langle |f| \rangle_{3P_0}\}| \leq \frac{1}{8}|P_0|. \quad (3.4.7)$$

Moreover, the set on the left is contained in  $P_0$ .

We construct the sparse bound this way. Fix a large (non-dyadic) cube  $P_0$  that  $\frac{1}{2}P_0$  contains the support of both  $f$  and  $g$ . The sparse cubes outside of  $P_0$  can be taken to  $3^k P_0$ , for  $k \in \mathbb{N}$ . We need to construct the sparse collection inside of  $P_0$ . Consider the restricted sum

$$I(P_0) := \int \sum_{P: P \subset P_0} \langle |D_{i_P - u} f| \rangle_{3P} \langle |D_{i_P - v} g| \rangle_{3P} \mathbb{1}_P \, dx. \quad (3.4.8)$$

Using (3.4.7), set

$$E_0 = \{S_{u,P_0} f > C(1+u)\langle |f| \rangle_{3P_0}\} \cup \{S_{v,P_0} g > C(1+v)\langle |g| \rangle_{3P_0}\}.$$

This set is contained in  $P_0$ , and has measure at most  $\frac{1}{4}|P_0|$ . Let  $\mathcal{E}_0$  be the maximal

dyadic components of  $E_0$ . We have by Cauchy-Schwartz and construction,

$$I(P_0) \leq C^2 \langle |f| \rangle_{3P_0} \langle |g| \rangle_{3P_0} |P_0| + \sum_{Q \in \mathcal{E}_0} I(Q).$$

The first term on the right is the first term in our sparse bound. We recurse on the second terms. This completes the proof.  $\square$

A very general fact about sparse forms is that they admit a ‘universal domination.’

**Lemma 3.4.9.** *Given  $f, g$ , there is a sparse operator  $\Lambda_0$ , and constant  $C > 1$  so that for any other sparse operator  $\Lambda$ , we have  $\Lambda(f, g) < C\Lambda_0(f, g)$ .*

*Proof.* Recall that shifted dyadic grids are a collection  $\mathbf{G}$  of at most  $3^d$  dyadic grids  $\mathcal{G} \in \mathbf{G}$ , so that every cube  $Q \subset \mathbb{R}^d$  can be approximated by some cube in a dyadic grid  $\mathcal{G} \in \mathbf{G}$ . Namely, for each cube  $Q$ , there is a  $\mathcal{G}$  and a cube  $P \in \mathcal{G}$  so that  $\frac{1}{6}\ell(P) \leq \ell(Q)$  and  $Q \subset 6P$ . See [35, Lemma 2.5] for an explicit proof.

Shifted grids permit us to construct a universal sparse operator for each grid  $\mathcal{G} \in \mathbf{G}$ . We show this: For any dyadic grid  $\mathcal{G}$ , let  $\mathcal{S} \subset \mathcal{G}$  be such that for  $S \in \mathcal{S}$ , there is a set  $E_S \subset S$  so that  $|E_S| > c|S|$  and  $\|\sum_{S \in \mathcal{S}} \mathbb{1}_{E_S}\|_\infty \leq c^{-1}$ . Given non-negative  $f, g$  bounded and compactly supported, we construct  $\mathcal{U}_\mathcal{G} \subset \mathcal{G}$  so that there are pairwise disjoint exceptional sets  $\{E_Q : Q \in \mathcal{U}_\mathcal{G}\}$  so that  $E_Q \subset Q$  and  $|E_Q| \geq \frac{1}{2}|Q|$ , and moreover,

$$\sum_{S \in \mathcal{S}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_S \leq 16^d c^{-2} \sum_{U \in \mathcal{U}_\mathcal{G}} \langle f \rangle_U \langle g \rangle_U \mathbb{1}_U. \quad (3.4.10)$$

To complete the proof of the Lemma, we remark that the collection  $\{\mathcal{U}_\mathcal{G} : \mathcal{G} \in \mathbf{G}\}$  is sparse. It dominates every sparse operator formed from some  $\mathcal{G} \in \mathbf{G}$ , hence is universal for all sparse operators.

For integers  $k$ , let  $\mathcal{U}_k$  be the maximal cubes  $Q \in \mathcal{G}$  so that  $\langle f \rangle_Q \langle g \rangle_Q \geq 8^{2dk}$ . Then, the product is at most  $8^{2dk+2d/3}$ . The cubes  $Q \in \mathcal{U}_k$  are pairwise disjoint, by maximality. We check that the children are small in measure. Setting  $\mathcal{C}(Q) = \{P \in$

$\mathcal{U}_{k+1} : P \subsetneq Q$ , we can write  $\mathcal{C}(Q) = \mathcal{C}_f(Q) \cup \mathcal{C}_g(Q)$ , where  $P \in \mathcal{C}_f(Q)$  if  $P \in \mathcal{C}(Q)$  and  $\langle f \rangle_P > 4^d \langle f \rangle_Q$ , and similarly for  $\mathcal{C}_g(Q)$ . But, then it is clear that

$$\sum_{P \in \mathcal{C}_f(Q)} |P| \leq 4^{-d} |Q| \leq \frac{1}{4} |Q|.$$

We set  $E_Q = Q \setminus \bigcup_{P \in \mathcal{C}(Q)} P$ . This set has measure at least  $\frac{1}{2} |Q|$ .

Set  $\mathcal{U}_g = \bigcup_k \mathcal{U}_k$ . The sets  $\{E_Q : Q \in \mathcal{U}\}$  are pairwise disjoint. Now, given the sparse collection as above, provided  $\langle f \rangle_S \langle g \rangle_S \neq 0$ , each  $S \in \mathcal{S}$  has a parent  $S^u \in \mathcal{U}$ , namely the smallest element of  $\mathcal{U}$  that contains  $S$ . Then,

$$\begin{aligned} \sum_{S \in \mathcal{S}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_S &= \sum_{U \in \mathcal{U}_g} \sum_{\substack{S \in \mathcal{S} \\ S^u = U}} \langle f \rangle_S \langle g \rangle_S \mathbb{1}_S \\ &\leq 16^d \sum_{U \in \mathcal{U}_g} \langle f \rangle_U \langle g \rangle_U \sum_{\substack{S \in \mathcal{S} \\ S^u = U}} \mathbb{1}_S \leq 16^d c^{-2} \sum_{U \in \mathcal{U}_g} \langle f \rangle_U \langle g \rangle_U \mathbb{1}_U. \end{aligned}$$

This verifies (3.4.10), so completes the proof. □

### Off-Diagonal Estimates

We begin with the very common off-diagonal estimate. For  $\eta > 0$  consider the Poisson-like operator

$$P_\eta \Phi(Q) := \int_{\mathbb{R}^d} \frac{(\ell Q)^\eta \Phi(y)}{(\ell Q)^{d+\eta} + \text{dist}(y, Q)^{d+\eta}} dy.$$

**Lemma 3.4.11** (Off-diagonal estimate). *Let  $g$  be a function with  $\int g dx = 0$ , supported on a cube  $Q$ , and  $f \in L^2$  supported on  $(2Q)^c$ , then we have*

$$|\langle Tf, g \rangle| \lesssim P_\eta |f|(Q) \|g\|_1 \leq P_\eta |f|(Q) |Q|^{1/2} \|g\|_2. \quad (3.4.12)$$

**Proof:** Let  $x_Q$  be the center of  $Q$ , then we have

$$\begin{aligned} |\langle Tf, g \rangle| &= \left| \int_Q \int_{(2Q)^c} K(x, y) f(y) g(x) dy dx \right| = \left| \int_{(2Q)^c} \int_Q (K(x, y) - K(x_Q, y)) f(y) g(x) dx dy \right| \\ &\leq \mathcal{K}_T \int_{(2Q)^c} \int_Q \frac{|x - x_Q|^\eta}{|x - y|^{d+\eta}} |f(y) g(x)| dx dy \lesssim \mathcal{K}_T P_\eta \|f\|_1 \|g\|_1. \end{aligned}$$

And the second inequality follows from Cauchy-Schwarz.  $\square$

**Lemma 3.4.13.** *Suppose that  $Q \in P$  and  $Q$  is good, then there is  $\eta' = \eta'(\eta, \gamma) > 0$ , such that*

$$P_\eta \mathbb{1}_{\mathbb{R}^d \setminus P}(Q) \lesssim \left[ \frac{\ell Q}{\ell P} \right]^{\eta'}. \quad (3.4.14)$$

*Proof.* Let  $\lambda = (\ell P / \ell Q)^{1-\gamma}$ . By goodness of  $Q$ , Lemma 3.2.1 implies

$$\begin{aligned} P_\eta \mathbb{1}_{\mathbb{R}^d \setminus P}(Q) &= \int_{\mathbb{R}^d \setminus P} \frac{(\ell Q)^\eta}{(\ell Q)^{d+\eta} + \text{dist}(y, Q)^{d+\eta}} dy \\ &\leq \int_{\mathbb{R}^d} \frac{(\ell Q)^\eta}{((\ell Q)^\gamma (\ell P)^{1-\gamma})^{d+\eta} + \text{dist}(y, Q)^{d+\eta}} dy \\ &\leq \left[ \frac{\ell Q}{\ell P} \right]^{\eta(1-\gamma)} P_\eta \mathbb{1}_{\mathbb{R}^d}(\lambda Q). \end{aligned}$$

So, the result follows.  $\square$

### Hardy's Inequality

This is the version of Hardy's inequality that we need. It can be proved from the one dimensional version. In point of fact, we only need this in the case where the function  $f$  below is constant.

**Lemma 3.4.15.** *For any cube,  $P$ , and  $1 < p < \infty$ , we have*

$$\int_{3P \setminus P} \int_P \frac{f(x)g(y)}{|x - y|^n} dx dy \lesssim \|f\|_p \|g\|_{p'}. \quad (3.4.16)$$

**CHAPTER 4**  
**UNIFORM SPARSE BOUNDS FOR DISCRETE QUADRATIC PHASE**  
**HILBERT TRANSFORMS**

(Joint work with Robert Kesler)

**4.1 Introduction**

Let  $e(t) = e^{2\pi it}$  and  $\alpha \in \mathbb{T}$ . We consider the operator  $H^\alpha$  acting on finitely supported functions  $f$  on  $\mathbb{Z}$ , defined by

$$H^\alpha f(n) := \sum_{m \neq 0} \frac{e(\alpha m^2) f(n - m)}{m}.$$

This can be regarded as a discrete oscillatory Hilbert transform with a quadratic phase. As such it satisfies a range of  $\ell^p$  estimates which are uniform in  $\alpha$ . In particular, the result below holds. Indeed, the work of Arkhipov and Oskolkov [1] in the case of  $p = 2$ , and of Pierce [67] in the case of  $1 < p < \infty$ , prove much more than the result below.

**Theorem 4.1.A.** *For  $1 < p < \infty$ , there holds*

$$\sup_{\alpha} \|H^\alpha : \ell^p \rightarrow \ell^p\| < \infty.$$

In this paper we give a further quantification of the uniform boundedness of  $H^\alpha$ , by proving a sparse bound. We set notation for the sparse bound. Let a discrete interval (or just an interval) be a set of the form  $I = \mathbb{Z} \cap [a, b]$ , for  $a, b \in \mathbb{R}$ , and define its length  $|I|$  as its cardinality. For  $1 \leq r < \infty$ , the  $L^r$ -average of a function  $f$  on the

interval  $I$  is defined by

$$\langle f \rangle_{I,r} := \left[ \frac{1}{|I|} \sum_{x \in I} |f(x)|^r \right]^{1/r}.$$

A collection of intervals  $\mathcal{S}$  is called  $\rho$ -sparse if for each  $S \in \mathcal{S}$ , there is a subset  $E_S$  of  $S$  such that (a)  $|E_S| > \rho|S|$ , and (b)  $\|\sum_{S \in \mathcal{S}} \mathbb{1}_{E_S}\|_\infty \leq \rho^{-1}$ . For a sparse collection  $\mathcal{S}$ , a sparse bilinear form  $\Lambda$  is defined by

$$\Lambda_{\mathcal{S},r,s}(f,g) := \sum_{S \in \mathcal{S}} \langle f \rangle_{S,r} \langle g \rangle_{S,r} |S|$$

When  $r = s$ , we write  $\Lambda_{\mathcal{S},r,s} = \Lambda_{\mathcal{S},r}$ . The dependence on  $\rho$  is not relevant, so it can be omitted. We also omit sometimes the dependence on the sparse collection  $\mathcal{S}$  and just write  $\Lambda_{r,s}$  or  $\Lambda_r$ .

To simplify some of the arguments, we make use of the following definition: For an operator  $T$  acting on finitely supported functions on  $\mathbb{Z}$ , and  $1 \leq r, s < \infty$  define its sparse norm

$$\|T : \text{Sparse}(r, s)\| = \|T : (r, s)\|, \tag{4.1.1}$$

as the infimum over the constants  $C > 0$  such that for all finitely supported functions  $f, g$  on  $\mathbb{Z}$  we have

$$|\langle Tf, g \rangle| \leq C \sup \Lambda_{r,s}(f, g).$$

Here, the supremum is taken over all sparse forms.

With this notation, we can state the main result of this paper as follows,

**Theorem 4.1.2.** *There exists  $1 < r < 2$  such that*

$$\sup_{\alpha \in \mathbb{T}} \|H^\alpha : (r, r)\| < \infty.$$

Given the useful structure of the sparse forms, we can derive a variety of mapping properties. For instance, we obtain the following immediate result

**Corollary 4.1.3.** *There exists  $1 < r < 2$  such that for all weights  $w$  that satisfy  $w, w^{-1} \in A_2 \cap RH_r$  we have*

$$\|H^\alpha : \ell^2(w) \mapsto \ell^2(w)\| \lesssim 1.$$

The weights above are in the intersection of the the standard Muckenhoupt class  $A_2$  and some Reverse Hölder class  $RH_r$ . Here and through all the paper, the notation  $A \lesssim B$  means that there is a constant  $C$  such that  $A \leq CB$ ; the dependence of the constant will be indicated when necessary.

The domination by sparse operators has been an active topic initiated by Lerner [54] in his simple proof of the  $A_2$  conjecture, by providing sparse control over the norm of a Calderón-Zygmund operator. This was improved to a pointwise estimate in [17] and following a stopping time argument in [45]. The latter approach has been used in different contexts [5, 37, 55]. The sparse bilinear form approach that we use here, has proven to be successful where the pointwise approach is not convenient or to avoid the use of maximal truncations, for example, the bilinear Hilbert transform [21], Bochner-Riesz multipliers [3] and oscillatory singular integrals [44, 50].

The study of oscillatory singular integrals is motivated by the work of Stein, who in [74] proves the boundedness on  $L^p$ , for  $1 < p < \infty$ , of the following operator,

$$\sup_{\alpha \in \mathbb{R}} \left| \int_{\mathbb{R}} f(x-y) \frac{e(\alpha y^2)}{y} dy \right|. \tag{4.1.4}$$

In the setting of discrete norm inequalities it is important to mention the remarkable work of Bourgain on ergodic theorems regarding polynomial averages [6, 7]. More recent results include the work of Krause [42] which have been extended in different directions by Mirek, Stein and Trojan [58, 59]. A first result in which similar discrete operator can be controlled by sparse forms can be found in [23], and in the case of random discrete operators in [44, 50], where the sparse bound follows from simpler

arguments.

Our main result, and the proof, is a model case for a wider range of results in the discrete setting. Some of the many possible extensions to the main result of this paper are as follows.

1. Extend the result to a general polynomial and kernel. That is, given a polynomial  $P$  and a Calderón-Zygmund Kernel  $K$ , find sparse bounds for the operator

$$T_P f(n) = \sum_{m \neq 0} e(P(m)) K(m) f(n - m),$$

that only depend on the degree of  $P$  and the kernel. More ambitious claims suggest themselves, such as obtaining sparse bounds for discrete Radon transforms, even in the quasi-translation invariant setting. See [67, 68].

2. Sparse version of Krause and Lacey's result [43], that is, find sparse bounds for the following restricted maximal operator, for  $A$  satisfying a certain Minkowski dimension condition,

$$\sup_{\alpha \in A} |H^\alpha f(n)| = \sup_{\alpha \in A} \left| \sum_{m \neq 0} \frac{e(\alpha m^2) f(n - m)}{m} \right|.$$

3. Sparse control over the maximal truncations of the operators above. This would entail extra difficulties.

The paper is organized as follows: In §2, we provide some preliminary results regarding sparse forms and specific operators bounded by them, that are key to our proof. In §3, following techniques from the Hardy-Littlewood circle method, we give a decomposition for the Fourier multiplier of the operator into major and minor arc components, and obtain some estimates for the different parts. We prove the sparse bounds for the minor and major arcs in §4 and §5 respectively to conclude the main

theorem. Of particular interest is the method to bound the major arcs, as it depends upon the sparse bound in Theorem 4.2.C.

## 4.2 Preliminaries

One useful fact about sparse operators is that, in some sense, they admit an universal domination. A version of the following lemma can be found in [47] and has a similar proof.

**Lemma 4.2.1.** *Given finitely supported functions  $f, g$  and  $1 \leq r, s < \infty$ , there is a sparse form  $\Lambda_{r,s}^*$  and a constant  $C > 0$  such that for any other sparse operator  $\Lambda_{r,s}$  we have*

$$\Lambda_{r,s}(f, g) \leq C\Lambda_{r,s}^*(f, g).$$

The Hardy-Littlewood maximal function is defined by

$$\mathcal{M}_{\text{HL}}f(n) := \sup_{N \geq 0} \frac{1}{2N+1} \sum_{j=-N}^N |f(n-j)|, \quad n \in \mathbb{Z}. \quad (4.2.2)$$

A well known result is the following.

**Theorem 4.2.B.** *The Hardy-Littlewood maximal function satisfies  $(1, 1)$  sparse bounds. That is,*

$$\|\mathcal{M}_{\text{HL}} : \text{Sparse}(1, 1)\| \lesssim 1.$$

If  $\mathcal{H}$  is a Hilbert space, we extend the definition of sparse forms to vector valued functions  $f$ , by setting  $\langle f \rangle_I = |I|^{-1} \sum_{x \in I} \|f(x)\|_{\mathcal{H}}$ . It is then straightforward to extend some sparse domination results to Hilbert space valued functions. One of this results, in the continuous setting of oscillatory singular integrals, is the following theorem, that is going to be an important part of our proof.

**Theorem 4.2.C.** [50] *Let  $K$  be a Calderón-Zygmund kernel and  $P$  a polynomial of*

degree  $d$  on  $\mathbb{R}^n$ . Define the operator

$$T_P f(x) = \int_{\mathbb{R}} e(P(y))K(y)f(x - y) dy.$$

For each  $1 < r < 2$  and compactly supported, Hilbert space valued functions  $f, g$ , there is a constant  $C = C(K, d, n, r)$  and a bilinear sparse form  $\Lambda_r$  such that

$$\|T_P f : \text{Sparse}(r, r)\| \leq C.$$

Recall that a Calderón-Zygmung kernel  $K : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{C}$  satisfies

$$\sup_{y \neq 0} |yK(y)| + \left| y^2 \frac{d}{dy} K(y) \right| < \infty,$$

and the corresponding convolution operator is  $L^2(\mathbb{R}^n)$ -bounded. In particular, we are going to apply this result with the Hilbert Transform kernel  $K(y) = 1/y$ . It is important to note that the previous estimate depends on the polynomial only through its degree.

In the subsequent sections,  $\varepsilon > 0$  will denote a small fixed constant. We use the standard notations for the Fourier transform and its inverse:

$$\begin{aligned} \hat{f}(\beta) &= \mathcal{F}f(\beta) = \sum_{n \in \mathbb{Z}} f(n)e(-\beta n), \\ \check{g}(n) &= \mathcal{F}^{-1}g(n) = \int_{\mathbb{T}} g(\beta)e(\beta n) d\beta. \end{aligned}$$

### 4.3 Decomposition of the multiplier

The Fourier multiplier associated to the transformation  $H^\alpha$  is

$$M^\alpha(\beta) := \sum_{m \neq 0} \frac{e(\alpha m^2 - \beta m)}{m}. \quad (4.3.1)$$

The goal of this section is to describe a decomposition of the multiplier  $M^\alpha$  into terms, with uniform control in the variable  $\alpha$ . Let  $\{\psi_j\}_{j \geq 0}$  be a dyadic resolution of the function  $\frac{1}{t}$ , with  $\psi_j(t) = 2^{-j}\psi(2^{-j}t)$ , and  $\psi$  is a odd smooth function satisfying  $\psi(t) \leq \mathbb{1}_{[1/4,1]}(|t|)$ . Then, for  $|t| \geq 1$ , we have  $\frac{1}{t} = \sum_{j \geq 0} \psi_j(t)$ , and in the support of  $\psi_j$ , we have  $2^{j-2} \leq |t| \leq 2^j$ . Using this, we can decompose the multiplier as a sum of terms of the form

$$M_j^\alpha(\beta) := \sum_{m \neq 0} \frac{e(\alpha m^2 - \beta m)}{m} \psi_j(m).$$

That way, we can write  $M^\alpha = \sum_j M_j^\alpha$ .

For fixed  $s \in \mathbb{N}$ , define

$$\mathcal{R}_s := \left\{ \left( \frac{A}{Q}, \frac{B}{Q} \right) \in \mathbb{T}^2 : A, B, Q \in \mathbb{Z}, (A, Q) = (B, Q) = 1, 2^{s-1} \leq Q \leq 2^s \right\}.$$

Then, the rationals in the torus, can be written as  $\bigcup_{s \in \mathbb{N}} \mathcal{R}_s$ . Given  $(\frac{A}{Q}, \frac{B}{Q}) \in \mathcal{R}_s$ , and  $j \geq s/\varepsilon$ , define the  $j$ -th major arc at  $(\frac{A}{Q}, \frac{B}{Q})$  by

$$\mathfrak{M}_j(A/Q, B/Q) := \{(\alpha, \beta) \in \mathbb{T}^2 : |\alpha - A/Q| \leq 2^{(\varepsilon-2)j}, |\beta - B/Q| \leq 2^{(\varepsilon-1)j}\}. \quad (4.3.2)$$

Collect the major arcs

$$\mathfrak{M}_j := \bigcup_{\substack{(A,Q)=(B,Q)=1 \\ 0 < Q \leq 2^{6\varepsilon j}}} \mathfrak{M}_j(A/Q, B/Q). \quad (4.3.3)$$

As proven in [43], the union above is over disjoint sets for  $\varepsilon$  small enough. For each  $j$ , we define the minor arcs to be the complement of this union of the major arcs.

Let  $\chi$  be a smooth even bump function, such that  $\mathbb{1}_{[-1/10, 1/10]} \leq \chi \leq \mathbb{1}_{[-1/5, 1/5]}$ . For  $s, j \in \mathbb{N}$ , set  $\chi_s(t) := \chi(10^s t)$ , and define the multiplier

$$L_{j,s}^\alpha(\beta) := \sum_{\left(\frac{A}{Q}, \frac{B}{Q}\right) \in \mathcal{R}_s} S(A/Q, B/Q) U_j(\alpha - A/Q, \beta - B/Q) \chi_s(\alpha - A/Q) \chi_s(\beta - B/Q). \quad (4.3.4)$$

Here,  $U_j$  is a continuous analogue of the multiplier  $M_j$ ,

$$U_j(x, y) := \int_{\mathbb{R}} e(xt^2 - yt) \psi_j(t) dt, \quad (4.3.5)$$

and  $S$  is the complete Gauss sum

$$S(A/Q, B/Q) := \frac{1}{Q} \sum_{r=0}^{Q-1} e(A/Q \cdot r^2 - B/Q \cdot r). \quad (4.3.6)$$

Consider also the following definitions.

$$L_j^\alpha(\beta) := \sum_{s \leq \varepsilon j} L_{j,s}^\alpha(\beta), \quad j \geq 1, \quad (4.3.7)$$

$$L^{\alpha,s}(\beta) := \sum_{j \geq s/\varepsilon} L_{j,s}^\alpha(\beta), \quad s \geq 1, \quad (4.3.8)$$

$$L^\alpha(\beta) := \sum_{j=1}^{\infty} L_j^\alpha(\beta) = \sum_{s=1}^{\infty} L^{\alpha,s}(\beta), \quad (4.3.9)$$

$$E_j^\alpha(\beta) := M_j^\alpha(\beta) - L_j^\alpha(\beta), \quad j \geq 1, \quad (4.3.10)$$

$$E^\alpha(\beta) := \sum_{j=1}^{\infty} E_j^\alpha(\beta). \quad (4.3.11)$$

The proof of the following lemmas can be found in [43]. The first one says that on the major arcs,  $M_j$  is well approximated by its continuous analogue.

**Lemma 4.3.12.** For  $1 \leq s \leq \varepsilon j$ ,  $(A/Q, B/Q) \in \mathcal{R}_s$ , and  $(\alpha, \beta) \in \mathfrak{M}_j(A/Q, B/Q)$ , we have the approximation

$$M_j^\alpha(\beta) = S(A/Q, B/Q)U_j(\alpha - A/Q, \beta - B/Q) + O(2^{(3\varepsilon-1)j}).$$

In the minor arcs we have the following estimates.

**Lemma 4.3.13.** There exists  $\delta = \delta(\varepsilon)$  such that uniformly in  $j \geq 1$ ,

$$|M_j^\alpha(\beta)| + |L_j^\alpha(\beta)| \lesssim 2^{-\delta j}, \quad (\alpha, \beta) \notin \mathfrak{M}_j(A/Q, B/Q)$$

Using these results, we obtain the following bounds.

**Theorem 4.3.14.** There is a choice of  $\delta > 0$  such that, uniformly in  $\alpha \in \mathbb{T}$

$$|S(A/Q, B/Q)| \lesssim 2^{-\delta s}, \quad (A/Q, B/Q) \in \mathcal{R}_s, \quad s \geq 1, \quad (4.3.15)$$

$$\|E_j^\alpha(\beta)\|_\infty \lesssim 2^{-\delta j}, \quad j \geq 1, \quad (4.3.16)$$

$$\left\| \frac{\partial^2}{\partial \beta^2} E_j^\alpha(\beta) \right\|_\infty \lesssim 2^{2j}, \quad j \geq 1. \quad (4.3.17)$$

The first estimate can be found in several places in the literature (see, for example, [32]). Given that by construction  $M_j^\alpha(\beta) = L_j^\alpha(\beta) + E_j^\alpha(\beta)$ , the second estimate is a consequence of the previous two lemmas. The derivative estimate comes from straightforward computations.

We prove the main Theorem by showing that there is a choice of  $1 < r < 2$  and  $\eta > 0$  such that for  $j, s \geq 1$  the following estimates hold, uniformly in  $\alpha \in \mathbb{T}$

$$\|T_{\tilde{E}_j^\alpha} : (r, r)\| \lesssim 2^{-\eta j} \quad (\text{Minor arcs estimate}) \quad (4.3.18)$$

$$\|T_{\tilde{L}^{\alpha, s}} : (r, r)\| \lesssim 2^{-\eta s} \quad (\text{Major arcs estimate}) \quad (4.3.19)$$

Since our operator can be written as  $H^\alpha = \sum_{j \geq 1} T_{\tilde{E}_j^\alpha} + \sum_{s \geq 1} T_{\tilde{L}^{\alpha, s}}$ , from the triangle

inequality for the sparse norm it follows that

$$\|H^\alpha : (r, r)\| \leq \sum_{j \geq 1} \|T_{\check{E}_j^\alpha} : (r, r)\| + \sum_{s \geq 1} \|T_{\check{L}^{\alpha, s}} : (r, r)\| \leq \sum_{j \geq 1} 2^{-\eta j} + \sum_{j \geq 1} 2^{-\eta s} < \infty.$$

Since these estimates are independent of  $\alpha$ , the main theorem follows.

#### 4.4 Minor Arcs estimate

Consider the multiplier  $E_j^\alpha$ , defined in (4.3.10). The  $L^\infty$  estimate (4.3.16) and the derivative estimate (4.3.17) imply that

$$|\mathcal{F}^{-1} E_j^\alpha(m)| \lesssim \min \left\{ 2^{-\varepsilon j}, \frac{2^{2j}}{1+m^2} \right\}. \quad (4.4.1)$$

These bounds are independent of  $\alpha$ , since the derivative estimates are. Write  $\mathcal{F}^{-1} E_j^\alpha = \check{E}_{j,1}^\alpha + \check{E}_{j,2}^\alpha$ , where  $\check{E}_{j,1}^\alpha(m) = \mathcal{F}^{-1} E_j^\alpha(m) \mathbb{1}_{[-2^{3j}, 2^{3j}]}(m)$ . We first estimate for  $\check{E}_{j,2}$ , for this, consider the Hardy-Littlewood maximal function defined in (4.2.2), we have

$$\begin{aligned} |T_{\check{E}_{j,2}^\alpha} f(x)| &= |\check{E}_{j,2}^\alpha * f(x)| \leq \sum_{y \in \mathbb{Z}} |K_2(y) f(x-y)| \lesssim 2^{2j} \sum_{|y| \geq 2^{3j}} \frac{|f(x-y)|}{1+|y|^2} \\ &= 2^{2j} \sum_{|k| \geq 3j} \sum_{2^k \leq |y| < 2^{k+1}} \frac{|f(x-y)|}{1+|y|^2} \lesssim 2^{2j} \sum_{|k| \geq 2^{3j}} 2^{-k} \mathcal{M}_{\text{HL}} f(x) = 2^{-j} \mathcal{M}_{\text{HL}} f(x). \end{aligned}$$

Once again, this estimate is independent of  $\alpha$ . Using Theorem 4.2.B we obtain the result for  $\check{E}_{j,2}^\alpha$ .

For  $\check{E}_{j,1}^\alpha$ , we need to use the following result (Proposition 2.4 in [23]).

**Proposition 4.4.2.** *Let  $T_K f(x) = \sum_n K(n) f(x-n)$  be convolution with kernel  $K$ . Assuming that  $K$  is finitely supported on the interval  $[-N, N]$  we have the inequalities*

$$\|T_K : (r, s)\| \lesssim N^{1/r+1/s-1} \|T_K : \ell^r \mapsto \ell^{s'}\|, \quad 1 \leq r, s < \infty.$$

To proof the sparse bound for  $T_{\check{E}_{j,1}^\alpha}$ , we use the proposition with  $N = 2^{3j}$  and  $r = s$ , that is

$$\|T_{\check{E}_{j,1}^\alpha} : (r, r)\| \lesssim 2^{3j(\frac{2}{r}-1)} \|T_{\check{E}_{j,1}^\alpha} : \ell^r \mapsto \ell^{r'}\|, \quad 1 \leq r < \infty.$$

We just need to find an  $r$  such that the operator norm has a summable decay in  $j$ . It is easy to check for the cases  $r = 1$  and  $r = 2$ . For  $r = 1$ , we have by Young's inequality and (4.4.1)

$$\left\| T_{\check{E}_{j,1}^\alpha} f \right\|_\infty \lesssim \|\check{E}_{j,1}^\alpha\|_\infty \|f\|_1 \lesssim 2^{-\delta j} \|f\|_1$$

And for  $r = 2$ , we have by the  $L^\infty$  estimate of the multiplier (4.3.16), and Plancherel,

$$\left\| T_{\check{E}_{j,1}^\alpha} : \ell^2 \mapsto \ell^2 \right\| \lesssim 2^{-\delta j}.$$

We can now interpolate and choose  $1 < r < 2$  such that  $10(2/r - 1) < \delta/2$  to get the desired decay. Combining this with the estimate over the norm of  $T_{\check{E}_{j,2}^\alpha}$  the proof of (4.3.18) is complete.

#### 4.5 Major Arcs estimate

We proceed now to prove the more complicated estimate (4.3.19). Recall the definition of  $U_j$ , given by (4.3.5). For  $s \geq 0$ , define  $U^s$  to be

$$U^s(x, y) = \sum_{j \geq s/\varepsilon} U_j(x, y).$$

Then, we can write the multiplier  $L^{\alpha,s}$  defined in (4.3.8) as

$$L^{\alpha,s}(\beta) = \sum_{\left(\frac{A}{Q}, \frac{B}{Q}\right) \in \mathcal{R}_s} S(A/Q, B/Q) U^s(\alpha - A/Q, \beta - B/Q) \chi_s(\alpha - A/Q) \chi_s(\beta - B/Q).$$

Given that the support of  $\chi_s$  is contained in  $[-2 \cdot 10^{-s-1}, 2 \cdot 10^{-s-1}]$ , for fixed  $\alpha \in \mathbb{T}$ , there is at most one rational  $\alpha_s = A/Q$  with  $(A, Q) = 1, 2^{s-1} \leq Q \leq 2^s$  for which  $\chi_s(\alpha - A/Q)$  is non zero. To simplify the notation, we make use of the following definition

$$\mathcal{R}_s^\alpha = \{ B/Q \in \mathbb{T} : (A/Q, B/Q) \in \mathcal{R}_s, A/Q = \alpha_s \}.$$

It is important to say that the subsequent analysis only depends upon the cardinality of  $\mathcal{R}_s^\alpha$ , which is at most  $2^{2s}$ , and not the value of  $\alpha$ . We can rewrite  $L^{\alpha,s}$  as

$$L^{\alpha,s}(\beta) = \sum_{\frac{B}{Q} \in \mathcal{R}_s^\alpha} S(\alpha_s, B/Q) U^s(\alpha - \alpha_s, \beta - B/Q) \chi_s(\alpha - \alpha_s) \chi_s(\beta - B/Q).$$

As in [23], we will make use of a sparse bound for Hilbert space valued singular integrals. For this, define for fixed  $\alpha \in \mathbb{T}$  the finite dimensional Hilbert space  $\mathcal{H}_s^\alpha = \ell^2(\mathcal{R}_s^\alpha)$ . Given  $f \in \ell^2$ , if  $\text{Mod}_h f(x) = e(hx)f(x)$  represents the standard modulation by  $h$ , set  $f_{s,h} := \mathcal{F}^{-1}(\chi_s^{1/2}) * \text{Mod}_{-h} f$ . Define the  $\mathcal{H}_s^\alpha$ -valued function  $f_s^\alpha$  by

$$f_s^\alpha := \{ f_{s,B/Q} : B/Q \in \mathcal{R}_s^\alpha \}.$$

Note that the Fourier transforms  $\hat{f}_{s,B/Q}(\beta) = \chi_s^{1/2}(\beta) \hat{f}(\beta + B/Q)$  have disjoint supports, so by Bessel's Theorem  $\|f_s^\alpha\|_{\ell^2(\mathcal{H}_s)} \leq \|f\|_{\ell^2}$ . We have the following simplifications,

$$\begin{aligned} \langle T_{L^{\alpha,s}} f, g \rangle &= \sum_{\frac{B}{Q} \in \mathcal{R}_s^\alpha} \sum_{j \geq s/\varepsilon} S(\alpha_s, \frac{B}{Q}) \left\langle U_j(\alpha - \alpha_s, \cdot - \frac{B}{Q}) \chi_s(\alpha - \alpha_s) \chi_s(\cdot - \frac{B}{Q}) \hat{f}(\cdot), \hat{g}(\cdot) \right\rangle \\ &= \sum_{\frac{B}{Q} \in \mathcal{R}_s^\alpha} \sum_{j \geq s/\varepsilon} S(\alpha_s, \frac{B}{Q}) \left\langle U_j(\alpha - \alpha_s, \cdot) \chi_s(\alpha - \alpha_s) \chi_s^{1/2}(\cdot) \hat{f}(\cdot + \frac{B}{Q}), \chi_s^{1/2}(\cdot) \hat{g}(\cdot + \frac{B}{Q}) \right\rangle \\ &= \chi_s(\alpha - \alpha_s) \sum_{\frac{B}{Q} \in \mathcal{R}_s^\alpha} \sum_{j \geq s/\varepsilon} S(\alpha_s, \frac{B}{Q}) \left\langle U_j(\alpha - \alpha_s, \cdot) \hat{f}_{s,B/Q}, \hat{g}_{s,B/Q} \right\rangle \end{aligned}$$

$$= \chi_s(\alpha - \alpha_s) \sum_{\frac{B}{Q} \in \mathcal{R}_s^\alpha} S(\alpha_s, \frac{B}{Q}) \langle T_{\tilde{U}^s} f_{s,B/Q}, g_{s,B/Q} \rangle.$$

For  $B/Q \in \mathcal{R}_s^\alpha$ , take  $\lambda_{B/Q}$  with unit norm and such that  $\lambda_{B/Q} \langle T_{\tilde{U}^s} f_{s,B/Q}, f_{s,B/Q} \rangle \geq 0$ , and set  $\tilde{f}_s^\alpha = \{ \lambda_{B/Q} f_{s,B/Q} : B/Q \in \mathcal{R}_s^\alpha \}$ , then, using the Gauss sum estimate (4.3.15) and summing over  $j \geq s/\varepsilon$  we have

$$| \langle T_{\tilde{L}^{\alpha,s}} f, g \rangle | \lesssim 2^{-\delta s} \langle T_{\tilde{U}^s} \tilde{f}_s^\alpha, g_s^\alpha \rangle.$$

Since  $\|f_s^\alpha\|_{\mathcal{H}_s^\alpha} = \|\tilde{f}_s^\alpha\|_{\mathcal{H}_s^\alpha}$ , then we can replace  $\tilde{f}_s^\alpha$  by  $f_s^\alpha$  in the inner product. The next step is to find a sparse form  $\Lambda_1$  (on Hilbert space valued functions) that dominates the last inner product. For this, first we write

$$U^s(\alpha - \alpha_s, \beta) = \sum_{j \geq s/\varepsilon} \int_{\mathbb{R}} e((\alpha - \alpha_s)t^2) e(-\beta t) \psi_j(t) dt = \int_{\mathbb{R}} e((\alpha - \alpha_s)t^2 - \beta t) \sum_{j \geq s/\varepsilon} \psi_j(t) dt$$

The integrand above is supported on  $|t| \geq 2^{\lfloor s/\varepsilon \rfloor - 2}$  and by explicit computation  $\sum_{j \geq s/\varepsilon} \psi_j(t)$  coincides with  $\frac{1}{t}$  for  $|t| \geq 2^{\lfloor s/\varepsilon \rfloor}$ . Therefore, this kernel corresponds to a Calderón-Zygmund kernel, and we can apply Theorem 4.2.C. As a consequence, for any  $1 < r_1 < 2$  there is a sparse bilinear form  $\Lambda_{r_1}$  such that

$$| \langle T_{\tilde{U}^s} f_s^\alpha, g_s^\alpha \rangle | \lesssim \Lambda_{r_1}(f_s^\alpha, g_s^\alpha). \quad (4.5.1)$$

The implied constant above does not depend on  $\alpha$ .

To end the proof, we need the following result

**Lemma 4.5.2.** *Let  $1 \leq r_1 < 2$  and  $\delta > 0$ . Let  $\Lambda_{r_1}$  be a sparse form over a collection of intervals all of which have length larger than  $10^s$ . Then there exists  $r$  satisfying*

$r_1 < r < 2$  such that for all  $f, g$  there is a sparse form  $\Lambda_r$  for which

$$\Lambda_{r_1}(f_s^\alpha, g_s^\alpha) \lesssim 2^{\delta s/4} \Lambda_r(f, g).$$

The proof of this lemma is a slight modification of the proof of the  $r_1 = 1$  result given at the end of [23] (the value of  $\alpha$  doesn't affect the proof). Ensuring all the sparse intervals in  $\Lambda_{r_1}$  have length at least  $10^s$  is achieved by taking  $\epsilon > 0$  small enough. Combining the estimates, and letting  $\eta = 3\delta/4$ , we have

$$|\langle T_{\tilde{L}^{\alpha, s}} f, g \rangle| \lesssim 2^{-\eta s} \Lambda_r(f, g).$$

Which proves the major arcs estimate, and therefore, the main theorem.

**CHAPTER 5**  
**SPARSE BOUNDS FOR BOCHNER-RIESZ MULTIPLIERS**

(Joint work with Michael Lacey and María Carmen Reguera)

**5.1 Introduction**

We study *sparse bounds* for the Bochner-Riesz multipliers in dimensions  $n \geq 2$ . The latter are Fourier multipliers  $B_\delta$  with symbol  $(1 - |\xi|^2)_+^\delta$ , for  $\delta > 0$ . That is,

$$\mathcal{F}B_\delta f = (1 - |\xi|^2)_+^\delta \mathcal{F}f,$$

where  $\mathcal{F}$  is a choice of Fourier transform. At  $\delta_n = \frac{n-1}{2}$ , the multiplier is borderline Calderón-Zygmund, and one has the very sharp bounds of Conde-Alonso, Culiuc, di Plinio and Ou [18], which we recall in Theorem 5.2.F below. In this paper, we focus on the super-critical range  $0 < \delta < \frac{n-1}{2}$ , the study of which was initiated by Benea, Bernicot and Luque [3]. We supply sparse bound for all  $0 < \delta < \frac{n-1}{2}$ , and prove a sharp range of estimates in dimension  $n = 2$ .

Sparse bounds are a particular quantification of the (weak)  $L^p$ -bounds for an operator, which in particular immediately imply weighted and vector-valued inequalities. The topic has been quite active, with an especially relevant paper being that of Benea, Bernicot and Luque [3], but also see [5, 18, 21, 36, 46, 47] for more information about this topic. We set notation for the sparse bounds. Call a collection of cubes  $\mathcal{S}$  in  $\mathbb{R}^n$  *sparse* if there are sets  $\{E_S : S \in \mathcal{S}\}$  which are pairwise disjoint,  $E_S \subset S$  and satisfy  $|E_S| > \frac{1}{4}|S|$  for all  $S \in \mathcal{S}$ . For any cube  $Q$  and  $1 \leq r < \infty$ , set  $\langle f \rangle_{Q,r}^r = |Q|^{-1} \int_Q |f|^r dx$ .

Then the  $(r, s)$ -sparse form  $\Lambda_{\mathfrak{S}, r, s} = \Lambda_{r, s}$ , indexed by the sparse collection  $\mathfrak{S}$  is

$$\Lambda_{\mathfrak{S}, r, s}(f, g) = \sum_{S \in \mathfrak{S}} |S| \langle f \rangle_{S, r} \langle g \rangle_{S, s}. \quad (5.1.1)$$

Given a sublinear operator  $T$ , and  $1 \leq r, s < \infty$ , we set  $\|T : (r, s)\|$  to be the infimum over constants  $C$  so that for all bounded compactly supported functions  $f, g$ ,

$$|\langle Tf, g \rangle| \leq C \sup \Lambda_{r, s}(f, g), \quad (5.1.2)$$

where the supremum is over all sparse forms. It is essential that the sparse form be allowed to depend upon  $f$  and  $g$ . But the point is that the sparse form itself varies over a class of operators with very nice properties.

The study of sparse bounds for the Bochner-Riesz multipliers was initiated by Benea, Bernicot and Luque [3], who established sparse bounds for a restricted range of parameters  $\delta, r$  and  $s$  below. We extend their results, using an alternate, less complicated method of proof, yielding results for all  $\delta > 0$ . In two dimensions our main result is as follows.

**Theorem 5.1.3.** *Let  $n = 2$ , and  $0 < \delta < \frac{1}{2}$ . Let  $\mathbf{R}(2, \delta)$  be the open trapezoid with vertices*

$$\begin{aligned} v_{2, \delta, 1} &= \left(\frac{1-2\delta}{4}, \frac{3+2\delta}{4}\right), & v_{2, \delta, 2} &= \left(\frac{1+6\delta}{4}, \frac{3+2\delta}{4}\right), \\ v_{2, \delta, 3} &= \left(\frac{3+2\delta}{4}, \frac{1+6\delta}{4}\right), & v_{2, \delta, 4} &= \left(\frac{3+2\delta}{4}, \frac{1-2\delta}{4}\right). \end{aligned}$$

(See Figure 5.1.) *There holds*

$$\|B_\delta : (r, s)\| < \infty, \quad \left(\frac{1}{r}, \frac{1}{s}\right) \in \mathbf{R}(2, \delta). \quad (5.1.4)$$

*Moreover, the inequality fails for  $\frac{1}{r} + \frac{1}{s} > 1$ , with  $(\frac{1}{r}, \frac{1}{s})$  not in the closure of  $\mathbf{R}(2, \delta)$ .*

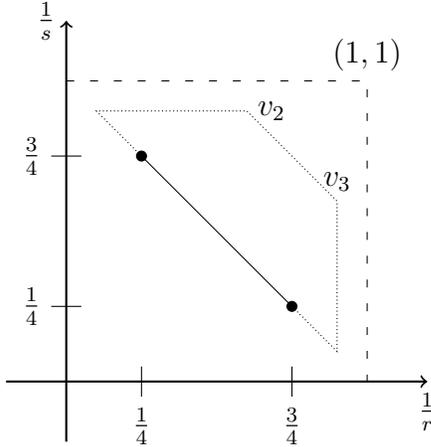


Figure 5.1: The trapezoid  $\mathbf{R}(2, \delta)$  of Theorem 5.1.3. The Bochner-Riesz bounds are sharp at the indices  $\frac{1}{p} = \frac{3}{4}, \frac{1}{4}$ , which corresponds to the Carleson-Sjölin bounds. The sparse bounds for  $B_\delta$  hold for all  $(\frac{1}{r}, \frac{1}{s})$  inside the dotted trapezoid, and fails outside the trapezoid. We abbreviate  $v_{2,\delta,2} = v_2$ , and similarly for  $v_3$ .

As  $\delta$  increases to the critical value of  $\delta = \frac{1}{2}$ , the trapezoid  $\mathbf{R}(2, \delta)$  increases to the upper triangle with vertices  $(1, 0)$ ,  $(0, 1)$  and  $(1, 1)$ . This is the full arrange allowed for the case of  $\delta = \frac{1}{2}$ , see Theorem 5.2.F.

In the next section, we give the full statement of the results in all dimensions. The remainder of the paper is taken up recalling some details about sparse bounds, the (short) proof of the main results, and then drawing out the weighted corollaries.

## 5.2 The Full Statement

In dimensions 3 and higher, we only have partial information about the Bochner-Riesz conjecture. Nevertheless, we show that from this partial information one can obtain sparse bounds as a consequence.

**Conjecture 5.2.1.** *[Bochner-Riesz Conjecture] We have  $B_\delta : L^p(\mathbb{R}^n) \mapsto L^p(\mathbb{R}^n)$  if*

$$n \left| \frac{1}{p} - \frac{1}{2} \right| < \frac{1}{2} + \delta, \quad \delta > 0. \quad (5.2.2)$$

This has an equivalent formulation, in terms of ‘thin annuli,’ which is the form we prefer. Let  $\mathbf{1}_{[-1/4, 1/4]} \leq \chi \leq \mathbf{1}_{[-1/2, 1/2]}$  be a Schwartz function, and set  $S_\tau$  to be the Fourier multiplier with symbol  $\chi((|\xi| - 1)/\tau)$ .

**Conjecture 5.2.3.** *Subject to the condition  $n\left|\frac{1}{p} - \frac{1}{2}\right| < \frac{1}{2}$ , there holds*

$$\|S_\tau\|_{L^p \rightarrow L^p} \lesssim_\epsilon 1, \quad 0 < \tau < 1. \quad (5.2.4)$$

Above, we use the notation  $A(\tau) \lesssim_\epsilon B(\tau)$  to mean that for all  $0 < \epsilon < 1$ , there is a constant  $C_\epsilon$  so that uniformly in  $0 < \tau < 1$ , there holds  $A(\tau) \leq C_\epsilon \tau^{-\epsilon} B(\tau)$ . It is typical in these types of questions that one expects losses in  $\tau$  that are of a logarithmic nature at end points. This issue need not concern us.

The Theorem below takes as input partial information about the Bochner-Riesz Conjecture and deduces a range of sparse bounds. For  $0 < \delta < \frac{n-1}{2}$ , let  $\mathbf{R}(n, p_0, \delta)$  be the open trapezoid with vertexes

$$v_{n,\delta,1} = \left(\frac{1}{p_0} \left(1 - \frac{2\delta}{n-1}\right), \frac{1}{p_0'} + \frac{1}{p_0} \frac{2\delta}{n-1}\right), \quad v_{n,\delta,2} = \left(\frac{1}{p_0} + \frac{1}{p_0'} \frac{2\delta}{n-1}, \frac{1}{p_0'} + \frac{1}{p_0} \frac{2\delta}{n-1}\right), \quad (5.2.5)$$

$$v_{n,\delta,3} = \overline{v_{n,p_0,2}}, \quad v_{n,\delta,4} = \overline{v_{n,p_0,1}}, \quad \text{where } \overline{(a, b)} = (b, a). \quad (5.2.6)$$

**Theorem 5.2.7.** *Assume dimension  $n \geq 2$ . And let  $1 < p_0 < 2$  be such that the estimate (5.2.4) holds. Then, for  $0 < \delta < \frac{n-1}{2}$ , the following sparse bound hold.*

$$\|B_\delta : (r, s)\| < \infty, \quad \left(\frac{1}{r}, \frac{1}{s}\right) \in \mathbf{R}(n, p_0, \delta). \quad (5.2.8)$$

Moreover, for the critical value of  $p = p(\delta)$  given by  $\frac{n}{p\delta} = \frac{n+1}{2} + \delta$ , the inequality above fails for  $\frac{1}{r} + \frac{1}{s} > 1$ , with  $(\frac{1}{r}, \frac{1}{s})$  not in the closure of  $\mathbf{R}(n, p_\delta, \delta)$ .

This Theorem contains Theorem 5.1.3, since the Bochner-Riesz conjecture holds in dimension 2, as was proved by Carleson-Sjölin [11]. (Also see Córdoba [20].) In dimensions  $n \geq 3$ , the best results are currently due to Bourgain-Guth, [8], but also

see Sanghyuk Lee [53]. We summarize the best known information in

**Theorem 5.2.D.** *These positive results hold for the Bochner-Riesz Conjecture.*

1. [11] *In the case of  $n = 2$ , (5.2.2) holds.*
2. [8, Thm 5] *In the case of  $n \geq 3$ , the condition (5.2.2) holds if  $q = \max(p, p')$  satisfies*

$$q > \begin{cases} 2\frac{4n+3}{4n-3} & n \equiv 0 \pmod{3} \\ \frac{2n+1}{n-1} & n \equiv 1 \pmod{3} \\ \frac{4(n+1)}{2n-1} & n \equiv 2 \pmod{3} \end{cases} \quad (5.2.9)$$

Concerning sparse bounds for the Bochner-Riesz multipliers, the general result of Benea, Bernicot and Luque [3] is a bit technical to state in full generality. We summarize it as follows.

**Theorem 5.2.E.** *These two results hold.*

1. [3, Thm 1] *In dimension  $n = 2$ , for  $\delta > \frac{1}{6}$ , we have  $\|B_\delta : (\frac{6}{5}, 2)\| < \infty$ .*
2. [3, Thm 3] *In dimensions  $n > 3$ , for all  $\delta > 0$ , there is a  $1 < p(\delta) < 2$  for which we have  $\|B_\delta : (p(\delta), 2)\| < \infty$ . (Using our notation, the sparse bound holds when the second coordinate of  $v_{n,\delta,2}$  is  $\frac{1}{2}$ .)*

Our result provides sparse bounds for the Bochner-Riesz multipliers, for all  $\delta > 0$ , and all  $p$  in a non-trivial interval around 2. It is a routine exercise to verify that a consequence of Theorem 5.1.3 that we have

$$\|B_\delta : (2, \frac{6}{5})\| < \infty, \quad n = 2, \delta > \frac{1}{6}. \quad (5.2.10)$$

Indeed, using the notation Theorem 5.1.3, we have  $v_{2,1/6,2} = (\frac{1}{2}, \frac{5}{6})$ . This is the two dimensional result of [3].

The interest in sparse bounds, besides their quantification of  $L^p$  bounds, is that they quickly deliver weighted and vector-valued inequalities. In many examples, these estimates are sharp [5, 18, 45, 56], or dramatically simplify existing proofs, and provide weighted inequalities in settings where none were known before [44, 46, 50]. The mechanism to do this is already well represented in the literature [5], and was initiated by Benea, Bernicot and Luque [3] in the setting of Bochner-Riesz multipliers. We point the interested reader there for more information about weighted estimates in the Bochner-Riesz setting.

That our result and that of [3] coincide at the case of  $r = 2$  is not so surprising. They approach the problem by using sharp results about spherical restriction, as there is a close connection between the Bochner-Riesz Conjecture and spherical restriction, subject to an index in the restriction question being 2. Our approach is more direct, working essentially with the 'single scale' version of the Bochner-Riesz Conjecture directly, through Conjecture 5.2.3. In both cases, we use the 'optimal' unweighted estimates, and derive the sparse bounds.

Concerning the critical index  $\delta_n = \frac{n-1}{2}$ , it is well known that the Bochner-Riesz operator is borderline Calderón-Zygmund. Hence one expects much better sparse bounds. The best sharp bound is due to Conde-Alonso, Culiuc, di Plinio and Ou [18]. It shows not only sparse bounds in the upper triangle of the  $(\frac{1}{r}, \frac{1}{s})$  plane, but also a quantitative estimate at the vertex  $(1, 1)$ .

**Theorem 5.2.F.** [18] *In all dimensions  $n \geq 2$ , we have*

$$\|B_{\delta_n} : (1, 1 + \epsilon)\| \lesssim \epsilon^{-1}, \quad 0 < \epsilon < \infty.$$

Note that the trapezoid of our theorem increases to the upper triangle, as  $\delta$  increases to the critical index  $\delta_n = \frac{n-1}{2}$ . In that sense, our results 'interpolates' the better bounds known in the critical case. We do not recover Theorem 5.2.F. Indeed

we can't as the proof is intrinsically multiscale, whereas ours is not.

The sparse bounds imply vector-valued and weighted inequalities for the Bochner-Riesz multipliers. The weights allowed are in the intersection of certain Muckenhoupt and reverse Hölder classes. The inequalities we can deduce are strongest at the vertex  $v_{n,\delta,2}$ , using the notation of (5.2.5). Indeed, the weighted consequence is the strongest known for the Bochner-Riesz multipliers. The method of deduction follows the model of arguments in [3, §7] and [5, §6], and so we suppress the details.

We conclude with these remarks.

1. Seeger [72] proves an endpoint weak-type result for the Bochner-Riesz operators in the plane. The sparse refinement of that is given Kesler and one of us in [40].
2. Extensions of these results to maximal Bochner-Riesz operators is hardly straight forward. For relevant norm inequalities, see [9, 53, 75].
3. Bak [2] proves endpoint estimates for negative index Bochner-Riesz multipliers. (Also see Gutiérrez [31].) Aside from endpoint issues, it would be easy to derive sparse bounds for these operators using the techniques of this paper. The  $A_{p,q}$  weighted consequences would be new, it seems. The endpoint issues would be interesting.
4. It is also of interest to obtain weighted bounds that more explicitly involve the Kakeya maximal function, as is done by Carbery [9] and Carbery and Seeger [10]. This would require substantially new techniques.

**Acknowledgment.** We benefited from conversations with Andreas Seeger and Richard Oberlin, as well as careful readings by referees.

### 5.3 Background on Sparse Forms

We collect some facts concerning sparse bounds. It is a useful fact that given bounded and compactly supported functions, there is basically one form that controls all others.

**Proposition 5.3.1.** [47, §4] *Given  $1 \leq r, s < \infty$ , and bounded and compactly supported functions  $f$ , and  $g$ , there is a single sparse form  $\Lambda_{\mathfrak{S}_0, r, s}$  for which*

$$\sup_{\mathfrak{S}} \Lambda_{\mathfrak{S}, r, s}(f, g) \lesssim \Lambda_{\mathfrak{S}_0, r, s}(f, g).$$

*The implied constant is only a function of dimension.*

Second, closely related sparse forms are also controlled by the sparse forms we defined at the beginning of the paper. For a cube  $Q$ , and  $1 \leq r < \infty$ , set a non-local average to be

$$\langle\langle f \rangle\rangle_{Q, r} = \left[ |Q|^{-1} \int |f(x)|^r [1 + \text{dist}(x, Q)/|Q|]^{-(n+1)} dx \right]^{\frac{1}{r}}. \quad (5.3.2)$$

And then define a sparse form  $\Lambda'_{\mathfrak{S}, r, s}$  using the non-local averages above in place of  $\langle f \rangle_{Q, r}$ . These forms are not essentially larger.

**Proposition 5.3.3.** [23, Lemma 2.8] *For bounded and compactly supported functions  $f, g$ , and  $1 \leq r, s < \infty$ , we have*

$$\sup_{\mathfrak{S}} \Lambda'_{\mathfrak{S}, r, s}(f, g) \lesssim \sup_{\mathfrak{S}} \Lambda_{\mathfrak{S}, r, s}(f, g).$$

A central point is that the selection of the 'optimal' sparse form in Proposition 5.3.1 is certainly non-linear. But at the same time, one would ideally like to interpolate sparse bounds. We do not know how to do this in general, but the analysis of the operators  $S_\tau$ , being 'single scale', places us in a situation where we can interpolate.

Using the notation of (5.3.2), for  $0 < \tau < 1$ , set

$$\tilde{\Lambda}_{\tau,r,s}(f, g) = \sum_{\substack{Q \in \mathcal{D} \\ 1 \leq \ell Q \leq \frac{1}{\tau^{1+\eta}}}} \langle\langle f \rangle\rangle_{Q,r} \langle\langle g \rangle\rangle_{Q,s} |Q|, \quad 0 < \eta < 1. \quad (5.3.4)$$

Above,  $\mathcal{D}$  denotes the dyadic cubes in  $\mathbb{R}^n$ , and  $\ell Q = |Q|^{\frac{1}{n}}$  is the side length of  $Q$ . That is, the sum is over all dyadic subcubes with side length between 1 and  $1/\tau^{1+\eta}$ .

We have this interpolation fact.

**Proposition 5.3.5.** *Let  $1 \leq r_j, s_j \leq \infty$  for  $j = 0, 1$  and fix  $0 < \tau < \infty$ . Suppose that for some linear operator  $T$  we have*

$$|\langle Tf, g \rangle| \leq C_j \tilde{\Lambda}_{\tau,r_j,s_j}(f, g), \quad j = 0, 1, \quad (5.3.6)$$

for all smooth compactly supported functions  $f, g$ . Then, for  $0 < \theta < 1$ , we have

$$|\langle Tf, g \rangle| \leq C_0^\theta C_1^{1-\theta} \tilde{\Lambda}_{\tau,r_\theta,s_\theta}(f, g), \quad (5.3.7)$$

where  $\frac{1}{r_\theta} = \frac{\theta}{r_0} + \frac{1-\theta}{r_1}$ , and similarly for  $s_\theta$ .

The proof is a variant of Riesz-Thorin interpolation, but we include some details, as the proposition is new in this context, as far as we know.

*Proof.* Let us recast the sparse bound in a slightly more general format. For cubes  $Q$ , set

$$\langle\langle f \rangle\rangle_{\lambda,Q,r} = \left[ \frac{1}{|Q|} \int \frac{|f(x)|^r}{(1 + \text{dist}(x, Q)/|Q|)^{n+1}} d\lambda(x) \right]^{\frac{1}{r}}.$$

Above,  $\lambda$  is some Borel measure. Fix a finite collection of cubes  $\mathcal{Q}$ , and consider a ‘sparse form’ given by

$$B_{r,s}(f, g) = B_{\mathcal{Q},\lambda,w,r,s}(f, g) = \sum_{Q \in \mathcal{Q}} w(Q) \{f\}_{\lambda,Q,r} \{g\}_{\lambda,Q,s}. \quad (5.3.8)$$

Above  $w : \mathcal{Q} \mapsto (0, \infty)$  is a non-negative function. The sparse forms that we consider are special instances of these more general forms.

Appeal to Hölder's inequality. Given  $r_0 < r_1$  and  $s_0 < s_1$ , we have

$$B_{r_\theta, s_\theta}(f, g) \leq B_{r_0, s_0}(f, g)^\theta B_{r_1, s_1}(f, g)^{1-\theta}, \quad 0 < \theta < 1$$

where  $\frac{1}{r_\theta} = \frac{\theta}{r_0} + \frac{1-\theta}{r_1}$ , and similarly for  $s_\theta$ .

Let us consider a bilinear form  $\beta$  for which we have

$$|\beta(f, g)| \leq A_j B_{r_j, s_j}(f, g), \quad j = 0, 1.$$

By multiplying  $f, g$  and the measure  $\lambda$  by various constants, we can assume that

$$B_{r_j, s_j}(f, g) = 1, \quad j = 0, 1.$$

For  $0 < \theta < 1$ , consider the holomorphic function  $F(s) = \beta(f_s, g_s)$ , where

$$f_s = \operatorname{sgn}(f) |f|^{(1-s)\frac{r_\theta}{r_0} + s\frac{r_\theta}{r_1}},$$

and similarly for  $g_s$ . The function  $F(s)$  is of at most exponential growth in the strip  $0 \leq \operatorname{Re} s \leq 1$ . Namely,

$$|F(s)| \leq B_{r_1, s_1}(f_s, g_s) \leq C e^{C|s|}, \quad 0 < \operatorname{Re} s \leq 1.$$

for some finite positive constant  $C$ . This is because  $f$  and  $g$  are bounded functions, and we have a finite collection of cubes  $\mathcal{Q}$ . Our deduced bounds are independent of these *a priori* assumptions. It also holds that  $|F(j + i\sigma)| \leq A_j$ , for  $j = 0, 1$ . It follows

from Lindelöf's Theorem that  $F$  is log-convex on  $[0, 1]$ . In particular,

$$|F(\theta)| = |\beta(f, g)| \leq A_0^\theta A_1^{1-\theta}.$$

From this, we conclude our proposition. □

## 5.4 Proof of the Sparse Bounds

The connection between the Bochner-Riesz and the  $S_\tau$  multipliers is well-known, and central to standard papers in the subject like [9, 20]. We briefly recall it here. For each  $0 < \delta < \frac{n-1}{2}$ , we have

$$B_\delta = T_0 + \sum_{k=1}^{\infty} 2^{-k\delta} \text{Dil}_{1-2^{-k}} S_{2^{-k}},$$

where these conditions hold: First,  $T_0$  is a Fourier multiplier, with the multiplier being a Schwartz function supported near the origin. The operator  $\text{Dil}_s f(x) = f(x/s)$  is a dilation operator. And, each  $S_{2^{-k}}$  is a Fourier multiplier with symbol  $\chi_k(2^k||\xi| - 1|)$ , where the  $\chi_k$  satisfy a uniform class of derivative estimates.

The point is then to show this result, in which we exploit the openness of the condition we are seeking to prove.

**Theorem 5.4.1.** *Assume dimension  $n \geq 2$ . And let  $1 < p_0 < 2$  be such that the estimate (5.2.4) holds. Then, the following sparse bounds hold. For all  $(\frac{1}{r}, \frac{1}{s}) \in \mathbf{R}(n, p_0, \delta)$ , there is a  $\kappa = \kappa(r, s) > 0$  so that*

$$\|S_\tau : (r, s)\| \lesssim_\epsilon \tau^{-\delta+\kappa}, \quad 0 < \tau < 1. \tag{5.4.2}$$

The papers of Córdoba [19, 20] also proceeds by analysis of the operators  $S_\tau$ . Also see Duoandikoetxea [27, Chap 8.5]. Write  $S_\tau f = K_\tau * f$ . The basic properties of this operator and kernel that we need are these.

**Lemma 5.4.3.** For  $0 < \tau < \frac{1}{2}$ , these properties hold.

1. We have this estimate for the kernel  $K_\tau$ . For all  $0 < \eta < 1$  and  $N > 1$ ,

$$|K_\tau(x)| \lesssim \tau \cdot \begin{cases} [1 + |x|]^{-\frac{n-1}{2}} & |x| < C\tau^{-1-\eta} \\ |x|^{\frac{1-n}{2}} [\tau|x|]^{-N} & \text{otherwise.} \end{cases} \quad (5.4.4)$$

The implied constants depend upon  $0 < \eta < 1$ , and  $N > 1$ .

2.  $\|S_\tau\|_{1 \rightarrow 1} \lesssim \tau^{-\frac{n-1}{2}}$ .

*Proof.* The second estimate follows from the first. The first is seen this way. Let  $\sigma$  denote normalized Haar measure on the sphere  $S^{n-1} \subset \mathbb{R}^n$ . Then, recall [73, Chap VIII.3], that the Fourier transform of  $\sigma$  has an expression in terms of Bessel functions as follows.

$$\widehat{d\sigma}(x) = \int_{S^{n-1}} e^{-2\pi i x \cdot \xi} d\sigma(\xi) = 2\pi|x|^{\frac{2-n}{2}} J_{\frac{n-2}{2}}(2\pi|x|).$$

The Bessel function has an expansion [73, Chap VIII.1.4]

$$J_{\frac{n-2}{2}}(s) = \sqrt{\frac{2}{\pi s}} \cos\left(s - \pi \frac{n-3}{4}\right) + O(s^{-3/2}), \quad s \rightarrow \infty.$$

It is preferable to write

$$J_{\frac{n-2}{2}}(s) = \sqrt{\frac{1}{s}} [e^{+is} a_+(s) + e^{-is} a_-(s)], \quad s > 0 \quad (5.4.5)$$

where

$$\left| \frac{d^m}{ds^m} a_\pm(s) \right| \lesssim [1 + s]^{-m}, \quad m \in \mathbb{N}, \quad s > 0. \quad (5.4.6)$$

This follows from the asymptotic expansion of the Bessel functions.

Turning to our estimates for  $K_\tau(x)$ , it is clear that  $|K_\tau(x)| \leq \|\chi(\tau^{-1}|x| - 1)\|_1 \lesssim$

$\tau$ , since we are only estimating the volume of a thin annulus. Thus, we only need to consider  $|x| \gtrsim 1$  below. In terms of the Bessel function, we have

$$\begin{aligned}
K_\tau(x) &= \int \chi(\tau^{-1}||\xi| - 1|) e^{ix \cdot \xi} d\xi \\
&= n \int_{S^{n-1}} \int_0^2 \chi(\tau^{-1}|r - 1|) e^{irx \cdot \xi} d\sigma(\xi) r^{n-1} dr \\
&= 2n\pi|x|^{\frac{2-n}{2}} \int_0^2 \chi(\tau^{-1}|r - 1|) r^{\frac{n}{2}} J_{\frac{n-2}{2}}(2\pi r|x|) dr
\end{aligned} \tag{5.4.7}$$

Using (5.4.5), this last expression is the sum of two terms, both of a similar nature. The first term is

$$2n\pi|x|^{\frac{1-n}{2}} \int_0^2 \chi(\tau^{-1}|r - 1|) r^{\frac{n-1}{2}} a_+(r|x|) e^{ir|x|} dr.$$

The integral above is obviously dominated by  $\tau$ , and this is the estimate that we use for  $1 \leq |x| < \tau^{-1-\eta}$ . For  $|x| \geq \tau^{-1-\eta}$ , we can employ the standard integration by parts argument and the derivative conditions in (5.4.6).  $\square$

The decay condition in (5.4.4) reveal that for fixed  $\tau$ , we need not be concerned with the full complexity of the sparse bound. We can rather work with this modified definition. Recall the sparse form  $\tilde{\Lambda}_{\tau,r,s}$  defined in (5.3.4), where we restrict cubes to have side length at least one, and no more than  $\tau^{-1-\eta}$ . And for which we have the interpolation estimate of Proposition 5.3.5. We define  $\|T : (r, s, \tau)\|$  to be the best constant  $C$  in the inequality

$$|\langle Tf, g \rangle| \leq C \tilde{\Lambda}_{\tau,r,s}(f, g), \tag{5.4.8}$$

the inequality holding uniformly over all bounded and compactly supported functions  $f, g$ .

**Lemma 5.4.9.** *Assume dimension  $n \geq 2$ . And let  $1 < p_0 < 2$  be such that the*

estimate (5.2.4) holds. These sparse bounds hold, for all  $0 < \tau, \eta < 1$ .

$$\|S_\tau : (1, 1, \tau)\| \lesssim \tau^{-\frac{n-1}{2}-n\eta}, \quad (5.4.10)$$

$$\|S_\tau : (1, \infty, \tau)\| \lesssim \tau^{-\frac{n-1}{2}-n\eta}, \quad (5.4.11)$$

$$\|S_\tau : (p_0, p'_0, \tau)\| \lesssim \tau^{-\eta} \quad (5.4.12)$$

The implied constants depend upon  $0 < \eta < 1$ .

*Proof.* It is in the last condition (5.4.12) that the hypothesis is important. Note that if  $f, g$  are supported on a cube  $Q$  of side length  $\tau^{-1-\eta}$ , then we have from the assumption that the Bochner-Riesz estimate (5.2.4) holds for  $p = p_0$ ,

$$|\langle S_\tau f, g \rangle| \lesssim \tau^{-\epsilon} \|f\|_{p_0} \|g\|_{p'_0} \lesssim_\epsilon \langle f \rangle_{Q, p_0} \langle g \rangle_{Q, p'_0} |Q|.$$

In view of the decay beyond the scale  $\tau^{-1-\eta}$  in (5.4.4), and the global form of the average in (5.3.4), we can easily complete the proof of the claimed bound. (And, we only need to use the dyadic cubes of scale  $\tau^{-1-\eta}$ , rather than the full range of scales in (5.3.4).)

In a similar way, if  $f$  and  $g$  are supported on a cube of side length  $\tau^{-1-\eta}$ , one can use the kernel decay in (5.4.4) to see that

$$\begin{aligned} |\langle S_\tau f, g \rangle| &\leq \|K_\tau(x)\|_1 \|f\|_1 \|g\|_\infty \\ &\lesssim \tau^{-\frac{n-1}{2}-n\eta} \|f\|_1 \|g\|_\infty \lesssim \tau^{-\frac{n-1}{2}-n\eta} \langle f \rangle_{Q, 1} \langle g \rangle_{Q, \infty} |Q|. \end{aligned}$$

And from this, we see that (5.4.11) holds.

The case of (5.4.10) is a little more involved, and requires that we use all the scales in our modified sparse operator (5.3.4), whereas the previous cases did not. Very briefly, we can dominate  $K_\tau$  by a positive Calderón-Zygmund kernel, with

Calderón-Zygmund norm at most  $\tau^{-\frac{n-1}{2}-n\eta}$ . From this, and the known results for sparse domination of Calderón-Zygmund operators, the bounds (5.4.10) follow. To be more explicit, let  $\varphi = \mathbf{1}_{|x|<2}$ , and set  $\varphi_k(x) = 2^{-kn}\varphi(x2^{-k})$ . Then, we have

$$|K_\tau(x)\mathbf{1}_{|x|<\tau^{-1-\eta}}| \lesssim \tau^{-\frac{n-1}{2}} \sum_{k:1\leq 2^k\leq\tau^{-1-\eta}} \varphi_k(x).$$

Convolution with  $\varphi_k$  is an average on scale  $2^k$ , so that

$$\left| \iint K_\tau(x-y)\mathbf{1}_{|x-y|<\tau^{-1-\eta}}f(y)g(x) dx dy \right| \lesssim \tau^{-\frac{n-1}{2}} \tilde{\Lambda}_{\tau,1,1}(f,g).$$

But, the same bound holds for the remainder of the kernel  $K_\tau$ , due to the decay estimates in (5.4.4). This completes the proof. □

*Proof of Theorem 5.4.1.* We in fact show that for  $0 < \delta < 1$  and  $(\frac{1}{r}, \frac{1}{s}) \in \mathbf{R}(n, p_0, \delta)$ ,

$$\|S_\tau : (r, s, \tau)\| \lesssim \tau^{-\delta-\eta}, \quad 0 < \tau, \eta < 1. \quad (5.4.13)$$

Above  $\delta$  is fixed, but  $\tau$  and  $\eta$  are allowed to vary. (The implied constant depends upon  $\eta$ .) This proves our Theorem, since for fixed  $(\frac{1}{r}, \frac{1}{s}) \in \mathbf{R}(n, p_0, \delta)$ , we have  $(\frac{1}{r}, \frac{1}{s}) \in \mathbf{R}(n, p_0, \delta - \kappa)$ , for a choice of  $0 < \kappa(r, s) < \delta$ .

The bounds in (5.4.13) are self-dual and can be interpolated, and so it suffices to verify the bounds above at the vertexes  $v_1 = v_{n,\delta,1}$  and  $v_2 = v_{n,\delta,2}$  of  $\mathbf{R}(n, p_0, \delta)$ , as defined in (5.2.5). But this is again an interpolation. To get the point  $v_1$ , interpolate between the sparse bound (5.4.12) and (5.4.10). For the point  $v_2$ , use (5.4.12) and (5.4.11). See Figure 5.2. □

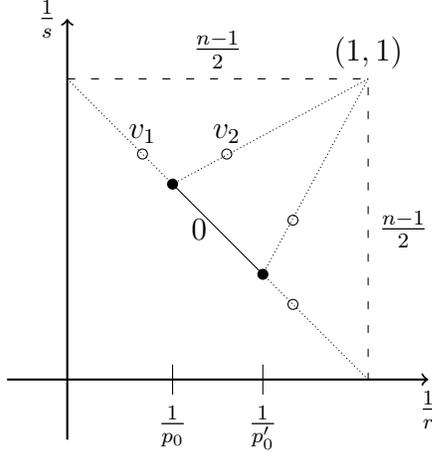


Figure 5.2: The interpolation argument for the sparse bounds. We have the sparse bounds  $(\frac{1}{p_0}, \frac{1}{p'_0})$ ,  $(1, 0)$ , and  $(1, 1)$  as well as their duals. The sparse bound at  $(\frac{1}{p_0}, \frac{1}{p'_0})$  is, up to logarithmic terms, uniformly bounded in  $0 < \tau < \frac{1}{2}$ , while the others are bounded by  $\tau^{-\frac{n-1}{2}}$ . Interpolation, along the dotted lines, to the circles, yields sparse bounds with growth  $\tau^{-\delta}$ , for  $0 < \delta < \frac{n-1}{2}$ .

## 5.5 Sharpness of the sparse bounds

We discuss sharpness of the sparse bounds in Theorem 5.2.7. Recalling that  $p(\delta)$  is the critical index for the Bochner-Riesz operator  $B_\delta$ , we cannot have any sparse bound  $(r, s)$ , where  $1 \leq r < p_\delta$ , as that would imply the boundedness of  $B_\delta$  on  $L^p$ , for  $r < p < p(\delta)$ .

It remains to show sharpness of the  $(r, s)$  sparse bound when  $p_\delta < r, s < p'_\delta$ . This follows from a standard example. We work in dimensions  $n \geq 2$ , Consider the rectangles  $R$  and  $\tilde{R}$  defined by

$$R = \left[ \frac{-1}{\lambda}, \frac{1}{\lambda} \right] \times \left[ \frac{-c}{\sqrt{\lambda}}, \frac{c}{\sqrt{\lambda}} \right]^{n-1}; \quad \tilde{R} = R + \frac{1}{\lambda}(1, 0, \dots, 0).$$

Above  $0 < c < 1$  is a small dimensional constant. Define the functions

$$f(x) = e^{i|x|} \mathbf{1}_R(x), \quad g(x) = e^{-i|x|} \mathbf{1}_{\tilde{R}}(x).$$

Using well known asymptotic estimates for the Bochner Riesz kernel, we have

$$|\langle B_\delta f, g \rangle| \simeq \left| \int_{\tilde{R}} \int_R \frac{e^{i(|x-y|-|x|+|y|)}}{(1+|x-y|)^{\frac{n-1}{2}+\delta}} dy dx \right| \quad (5.5.1)$$

$$\simeq |R|^2 \lambda^{\frac{n-1}{2}+\delta} \simeq \lambda^{-\frac{n+1}{2}+\delta} \quad (5.5.2)$$

The kernel estimates we are referencing are analogs of (5.4.5), which has two exponential terms in it. Above, one can directly verify that the phase function satisfies

$$||x-y|-|x|+|y|| \lesssim c, \quad x \in \tilde{R}, y \in R. \quad (5.5.3)$$

This leads to the estimate above. There is a second exponential term with phase function

$$-|x-y|-|x|+|y| \simeq -2|x-y|, \quad x \in \tilde{R}, y \in R. \quad (5.5.4)$$

So, that term has substantial cancellation.

Recall that the largest value of sparse form  $\Lambda_{s,r,s}(f,g)$  is obtained by a single sparse form. For the functions  $f, g$  above, it is clear that this largest form is obtained by taking  $\mathcal{S}$  to consist of only the smallest cube  $Q$  that contains the support of both  $f$  and  $g$ . That cube has  $\ell Q \simeq \lambda^{-1}$ . And then,

$$|Q| \langle f \rangle_{Q,r} \langle g \rangle_{Q,s} \simeq \lambda^{-n + \frac{n-1}{2}(\frac{1}{r} + \frac{1}{s})}. \quad (5.5.5)$$

We see that the  $(r, s)$  sparse bound for  $B_\delta$  implies that (5.5.2) should be less than (5.5.5) for all small  $\lambda$ . By comparing exponents, we see that

$$-\frac{n+1}{2} + \delta \geq -n + \frac{n-1}{2}(\frac{1}{r} + \frac{1}{s}).$$

The case of equality above is the line that defines the top of the trapezoid  $R(n, p(\delta), \delta)$ , as is verified by inspection.

## 5.6 Weighted Consequences

The sparse bounds imply vector-valued and weighted inequalities for the Bochner-Riesz multipliers. The weights allowed are in the intersection of certain Muckenhoupt and reverse Hölder classes. The inequalities we can deduce are strongest at the vertex  $v_{n,\delta,2}$ , using the notation of (5.2.5). Indeed, the weighted consequence is the strongest known for the Bochner-Riesz multipliers. The method of deduction follows the model of arguments in [3, §7] and [5, §6], as well as [46, §6]. We tread lightly around the details.

It is also of interest to obtain weighted bounds that more explicitly involve the Keakeya maximal function, as is done by Carbery [9] and Carbery and Seeger [10]. We leave to the future to obtain sparse variants of these latter results.

Recall that a weight  $w$  is in the Muckenhoupt  $A_p$  class if it has a density  $w(dx) = w(x)dx$ , with  $w(x) > 0$ , which is locally integrable, and  $\sigma(x) = w(x)^{-\frac{1}{p-1}}$  is also locally integrable, and

$$[w]_{A_p} = \sup_Q \langle w \rangle_Q \langle \sigma \rangle_Q^{p-1} < \infty, \quad (5.6.1)$$

where  $w(Q) = \int_Q w(dx)$ , and the supremum is over all cubes. We use the standard extension to  $p = 1$ , namely

$$[w]_{A_1} = \left\| \frac{Mw}{w} \right\|_\infty$$

We set  $A_p^p = \{w^p : w \in A_p\}$ .

We will set  $B_{\delta,p}$  to be the class of weights  $w$  such that we have the inequality

$$\|B_\delta\|_{L^p(w) \rightarrow L^p(w)}. \quad (5.6.2)$$

Below, we will focus on qualitative results. All results can be made entirely quantitative, but given the incomplete information that we have the Bochner-Riesz conjecture, or even the full range of sparse bounds in two dimensions, we do not pursue the quan-

titative bounds at this time.

The best known results concerning the weighted estimates for the Bochner-Riesz multipliers in the category of  $A_p$  classes, are below. We emphasize that some of these hold for the maximal Bochner-Riesz multiplier, which we are not considering in this paper.

**Theorem 5.6.G.** 1. (*Christ*, [15]) *We have the inclusion below valid in all  $n \geq 2$ .*

$$A_1^{\frac{1+2\delta}{n}} \subset B_{\delta,2}, \quad \frac{n-1}{2(n+1)} < \delta < \frac{n-1}{2}. \quad (5.6.3)$$

2. (*Carro, Duoandikoetxea, Lorente* [12]) *We have the inclusion below, valid in all dimensions  $n \geq 2$ .*

$$A_2^{\frac{2\delta}{n-1}} \subset B_{\delta,2}. \quad (5.6.4)$$

The second result is a consequence of the  $(\frac{1}{2} + \frac{2\delta}{n-1}, \frac{1}{2} + \frac{2\delta}{n-1})$  sparse bound. This latter sparse bound can be deduced from the the (trivial)  $(2, 2, \tau)$  and  $(1, 1, \tau)$  sparse bounds, as defined in Lemma 5.4.9. That is, (5.6.4) has little to do with the Bochner-Riesz operators. (The authors of [12] note a similar argument.)

We are able to deduce this improvement of (5.6.3), in that it applies for all  $0 < \delta < \frac{n-1}{2}$ , and increases the integrability of the Bochner-Riesz bound. Finally, it approximates the known estimate at the critical index, see Theorem 5.2.F, and the earlier result of Vargas [77].

**Theorem 5.6.5.** *In all dimensions  $n \geq 2$ , using the notation of (5.2.5), write the vertex  $v_{n,\delta,2} = (\frac{1}{r}, \frac{1}{s})$ . We have*

$$A_1^{\frac{s-p(s-1)}{s}} \cdot A_1^{1-\frac{p}{r}} \subset B_{\delta',p}, \quad 0 < \delta < \delta' < \frac{n-1}{2}, \quad r < p < s'. \quad (5.6.6)$$

In particular, for  $n = 2$ , we have the explicit value  $v_{2,\delta,2} = (\frac{1+6\delta}{4}, \frac{3+2\delta}{4})$ , and

$$A_1^{1-p\frac{1-2\delta}{4}} \cdot A_1^{1-p\frac{1+6\delta}{4}} \subset B_{\delta',p}, \quad 0 < \delta < \delta' < \frac{1}{2}, \quad \frac{4}{1+6\delta} < p < \frac{4}{1-2\delta}. \quad (5.6.7)$$

This contrasts to [3, Theorem 14], which is restrictive in the range of  $0 < \delta < \frac{n-1}{2}$  that are allowed. (See [3, Corollary 16] for an example of the kind of vector-valued consequences that can be derived.) We use the vertex  $v_{n,\delta,2}$ , as it is the strongest sparse bound we have. The proof is however elementary. We have this known proposition.

**Proposition 5.6.8.** *If a linear operator  $T$  satisfies a  $(r, s)$  sparse bound, with  $1 \leq s < r$ , we then have*

$$\|T : L^p(w^\rho) \mapsto L^p(w^\rho)\| < \infty, \quad w \in A_1^{\frac{s-p(s-1)}{s}} A_1^{1-\frac{p}{r}}, \quad r < p < s'.$$

*Proof of Theorem 5.6.5.* It is a consequence of [5, Prop. 6.4], that the sparse bound assumption implies that

$$\|T : L^p(w) \mapsto L^p(w)\| < \infty, \quad r < p < s'$$

provided the weight  $w$  is in

$$w \in A_{\frac{p}{r}} \cap RH_{(s'/p)'} = A_1^{1/(s'/p)'} \cdot A_1^{1-\frac{p}{r}}$$

Above,  $RH_\rho$  denotes the reverse Hölder class of weights of index  $1 < \rho < \infty$ , and the equality above is classical. By direct calculation,  $1/(s'/p)' = \frac{s-p(s-1)}{s}$ .  $\square$

## BIBLIOGRAPHY

- [1] G. I. Arkhipov and K. I. Oskolkov, *A special trigonometric series and its applications*, Mat. Sb. (N.S.) **134(176)** (1987), no. 2, 147–157, 287 (Russian); English transl., Math. USSR-Sb. **62** (1989), no. 1, 145–155. MR922412
- [2] J.-G. Bak, *Sharp estimates for the Bochner-Riesz operator of negative order in  $\mathbf{R}^2$* , Proc. Amer. Math. Soc. **125** (1997), no. 7, 1977–1986.
- [3] C. Benea, F. Bernicot, and T. Luque, *Sparse bilinear forms for Bochner Riesz multipliers and applications*, ArXiv e-prints (May 2016), available at <http://arxiv.org/abs/1605.06401>.
- [4] A. Bernard, *Espaces  $H^1$  de martingales à deux indices. Dualité avec les martingales de type “BMO”*, Bull. Sci. Math. (2) **103** (1979), no. 3, 297–303.
- [5] F. Bernicot, D. Frey, and S. Petermichl, *Sharp weighted norm estimates beyond Calderón-Zygmund theory*, Analysis & PDE **9** (2016), no. 5, 1079–1113.
- [6] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. **61** (1988), no. 1, 39–72. MR937581 (89f:28037a)
- [7] ———, *On the pointwise ergodic theorem on  $L^p$  for arithmetic sets*, Israel J. Math. **61** (1988), no. 1, 73–84. MR937582 (89f:28037b)
- [8] J. Bourgain and L. Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295.
- [9] A. Carbery, *A weighted inequality for the maximal Bochner-Riesz operator on  $\mathbf{R}^2$* , Trans. Amer. Math. Soc. **287** (1985), no. 2, 673–680.
- [10] A. Carbery and A. Seeger, *Weighted inequalities for Bochner-Riesz means in the plane*, Q. J. Math. **51** (2000), no. 2, 155–167.
- [11] L. Carleson and P. Sjölin, *Oscillatory integrals and a multiplier problem for the disc*, Studia Math. **44** (1972), 287–299. (errata insert).
- [12] M. J. Carro, J. Duoandikoetxea, and M. Lorente, *Weighted estimates in a limited range with applications to the Bochner-Riesz operators*, Indiana Univ. Math. J. **61** (2012), no. 4, 1485–1511.
- [13] S.-Y. A. Chang, *Carleson measure on the bi-disc*, Ann. of Math. (2) **109** (1979), no. 3, 613–620.

- [14] S.-Y. A. Chang and R. Fefferman, *A continuous version of duality of  $H^1$  with BMO on the bidisc*, Ann. of Math. (2) **112** (1980), no. 1, 179–201.
- [15] M. Christ, *On almost everywhere convergence of Bochner-Riesz means in higher dimensions*, Proc. Amer. Math. Soc. **95** (1985), no. 1, 16–20.
- [16] R. R. Coifman, R. Rochberg, and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635.
- [17] J. M. Conde-Alonso and G. Rey, *A pointwise estimate for positive dyadic shifts and some applications*, Math. Ann. **365** (2016), no. 3-4, 1111–1135. MR3521084
- [18] J. M. Conde-Alonso, A. Culiuc, F. Di Plinio, and Y. Ou, *A sparse domination principle for rough singular integrals*, Analysis & PDE **10** (2017), no. 5, 1255–1284.
- [19] A. Córdoba, *The Kakeya maximal function and the spherical summation multipliers*, Amer. J. Math. **99** (1977), no. 1, 1–22.
- [20] ———, *A note on Bochner-Riesz operators*, Duke Math. J. **46** (1979), no. 3, 505–511.
- [21] A. Culiuc, F. Di Plinio, and Y. Ou, *Domination of multilinear singular integrals by positive sparse forms*, ArXiv e-prints (March 2016), available at <http://arxiv.org/abs/1603.05317>.
- [22] A. Culiuc, F. Di Plinio, and Y. Ou, *Uniform sparse domination of singular integrals via dyadic shifts*, ArXiv e-prints (October 2016), available at <http://arxiv.org/abs/1610.01958>.
- [23] A. Culiuc, R. Kesler, and M. T. Lacey, *Sparse Bounds for the discrete cubic Hilbert transform*, ArXiv e-prints (December 2016), available at <http://arxiv.org/abs/1612.08881>.
- [24] L. Dalenc and Y. Ou, *Upper bound for multi-parameter iterated commutators*, Publ. Mat. **60** (2016), no. 1, 191–220.
- [25] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. (2) **120** (1984), no. 2, 371–397. MR763911
- [26] K. Domelevo and S. Petermichl, *A sharp maximal inequality for differentially subordinate martingales under a change of law*, ArXiv e-prints (July 2016), available at <http://arxiv.org/abs/1607.06319>.
- [27] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics, vol. 29, American Mathematical Society, Providence, RI, 2001.

- [28] S. H. Ferguson and M. T. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143–160.
- [29] S. H. Ferguson and C. Sadosky, *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures*, J. Anal. Math. **81** (2000), 239–267.
- [30] T. Figiel, *Singular integral operators: a martingale approach*, Geometry of Banach spaces (Strobl, 1989), 1990, pp. 95–110. MR1110189
- [31] S. Gutiérrez, *A note on restricted weak-type estimates for Bochner-Riesz operators with negative index in  $\mathbf{R}^n$ ,  $n \geq 2$* , Proc. Amer. Math. Soc. **128** (2000), no. 2, 495–501.
- [32] L. K. Hua, *Introduction to number theory*, Springer-Verlag, Berlin-New York, 1982. Translated from the Chinese by Peter Shiu. MR665428 (83f:10001)
- [33] T. P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. MR2912709
- [34] ———, *Vector-valued singular integrals revisited—with random dyadic cubes*, Bull. Pol. Acad. Sci. Math. **60** (2012), no. 3, 269–283. MR2956641
- [35] T. P. Hytönen, M. T. Lacey, and C. Pérez, *Sharp weighted bounds for the  $q$ -variation of singular integrals*, Bull. Lond. Math. Soc. **45** (2013), no. 3, 529–540, DOI 10.1112/blms/bds114. MR3065022
- [36] T. P. Hytönen, S. Petermichl, and A. Volberg, *The sharp square function estimate with matrix weight*, ArXiv e-prints (February 2017), available at <http://arxiv.org/abs/1702.04569>.
- [37] T. P. Hytönen, L. Roncal, and O. Tapiola, *Quantitative weighted estimates for rough homogeneous singular integrals*, ArXiv e-prints (October 2015), available at <http://arxiv.org/abs/1510.05789>.
- [38] G. A. Karagulyan, *Exponential estimates for the Calderón-Zygmund operator and related problems of Fourier series*, Mat. Zametki **71** (2002), no. 3, 398–411. MR1913610
- [39] N. H. Katz, *Matrix valued paraproducts*, Proceedings of the conference dedicated to Professor Miguel de Guzmán (El Escorial, 1996), 1997, pp. 913–921.
- [40] R. Kesler and M. T. Lacey, *Sparse Endpoint Estimates for Bochner-Riesz Multipliers on the Plane*, ArXiv e-prints (July 2017), available at <http://arxiv.org/abs/1707.05844>.

- [41] R. Kesler and D. Mena, *Uniform sparse bounds for discrete quadratic phase Hilbert transforms*, Analysis and Mathematical Physics, posted on September 2017, DOI 10.1007/s13324-017-0195-3.
- [42] B. Krause, *Polynomial Ergodic Averages Converge Rapidly: Variations on a Theorem of Bourgain* (February 2014), available at <http://arxiv.org/abs/1402.1803>.
- [43] B. Krause and M. T. Lacey, *A Discrete Quadratic Carleson Theorem on  $\ell^2$  with a Restricted Supremum*, IMRN, to appear (December 2015), available at <http://arxiv.org/abs/1512.06918>.
- [44] ———, *Sparse Bounds for Random Discrete Carleson Theorems*, ArXiv e-prints (September 2016), available at <http://arxiv.org/abs/1609.08701>.
- [45] M. T. Lacey, *An elementary proof of the  $A_2$  Bound*, Israel J. Math. **217** (2017), no. 1, 181–195.
- [46] ———, *Sparse Bounds for Spherical Maximal Functions*, ArXiv e-prints (February 2017), available at <http://arxiv.org/abs/1702.08594>.
- [47] M. T. Lacey and D. Mena Arias, *The sparse  $T1$  theorem*, Houston J. Math. **43** (2017), no. 1, 111–127. MR3647935
- [48] M. T. Lacey, D. Mena Arias, and M. C. Reguera, *Sparse Bounds for Bochner-Riesz Multipliers*, ArXiv e-prints (May 2017), available at <http://arxiv.org/abs/1705.09375>.
- [49] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, *Iterated Riesz commutators: a simple proof of boundedness*, Harmonic analysis and partial differential equations, Contemp. Math., vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 171–178.
- [50] M. T. Lacey and S. Spencer, *Sparse Bounds for Oscillatory and Random Singular Integrals*, New York Journal of Mathematics, to appear (September 2016), available at <http://arxiv.org/abs/1609.06364>.
- [51] M. T. Lacey, E. T. Sawyer, C.-Y. Shen, and I. Uriarte-Tuero, *Two-weight inequality for the Hilbert transform: a real variable characterization, I*, Duke Math. J. **163** (2014), no. 15, 2795–2820. MR3285857
- [52] M. T. Lacey and E. Terwilleger, *Hankel operators in several complex variables and product BMO*, Houston J. Math. **35** (2009), no. 1, 159–183.
- [53] S. Lee, *Improved bounds for Bochner-Riesz and maximal Bochner-Riesz operators*, Duke Math. J. **122** (2004), no. 1, 205–232.

- [54] A. K. Lerner, *A simple proof of the  $A_2$  conjecture*, Int. Math. Res. Not. IMRN **14** (2013), 3159–3170. MR3085756
- [55] ———, *On pointwise estimates involving sparse operators*, New York J. Math. **22** (2016), 341–349. MR3484688
- [56] A. K. Lerner, S. Ombrosi, and I. P. Rivera-Ríos, *On pointwise and weighted estimates for commutators of Calderón-Zygmund operators*, Adv. Math. **319** (2017), 153–181.
- [57] Y. Meyer, *Wavelets and operators*, Cambridge Studies in Advanced Mathematics, vol. 37, Cambridge University Press, Cambridge, 1992. Translated from the 1990 French original by D. H. Salinger.
- [58] M. Mirek, E. M. Stein, and B. Trojan,  *$L^p(\mathbb{Z}^d)$ -estimates for discrete operators of Radon type: Maximal functions and vector-valued estimates* (December 2015), available at <http://arxiv.org/abs/1512.07518>.
- [59] ———,  *$L^p(\mathbb{Z}^d)$ -estimates for discrete operators of Radon type: Variational estimates* (December 2015), available at <http://arxiv.org/abs/1512.07523>.
- [60] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), no. 2, 269–296.
- [61] ———, *Multi-parameter paraproducts*, Rev. Mat. Iberoam. **22** (2006), no. 3, 963–976.
- [62] F. Nazarov, G. Pisier, S. Treil, and A. Volberg, *Sharp estimates in vector Carleson imbedding theorem and for vector paraproducts*, J. Reine Angew. Math. **542** (2002), 147–171.
- [63] F. Nazarov, S. Treil, and A. Volberg, *Counterexample to the infinite-dimensional Carleson embedding theorem*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 4, 383–388.
- [64] F. Nazarov, S. Treil, and A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. **190** (2003), no. 2, 151–239. MR1998349
- [65] C. Ortiz-Caraballo, C. Pérez, and E. Rela, *Exponential decay estimates for singular integral operators*, Math. Ann. **357** (2013), no. 4, 1217–1243. MR3124931
- [66] S. Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris Sér. I Math. **330** (2000), no. 6, 455–460 (English, with English and French summaries).
- [67] L. B. Pierce, *A note on twisted discrete singular Radon transforms*, Math. Res. Lett. **17** (2010), no. 4, 701–720, DOI 10.4310/MRL.2010.v17.n4.a10. MR2661174

- [68] ———, “Discrete fractional Radon transforms and quadratic forms,” *Duke Math. J.* **161** (2012), 69–106 [[MR2872554](#)], *Duke Math. J.* **162** (2013), no. 6, 1203–1204, DOI 10.1215/00127094-2210146. MR3053569
- [69] G. Pisier, *Notes on Banach space valued  $H^p$ -spaces, non-commutative martingale inequalities and matrix valued Harmonic Analysis*, 2005. Informal Seminar Notes.
- [70] E. T. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, *Studia Math.* **75** (1982), no. 1, 1–11. MR676801
- [71] ———, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, *Trans. Amer. Math. Soc.* **308** (1988), no. 2, 533–545. MR930072
- [72] A. Seeger, *Endpoint inequalities for Bochner-Riesz multipliers in the plane*, *Pacific J. Math.* **174** (1996), no. 2, 543–553.
- [73] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR1232192
- [74] E. M. Stein, *Oscillatory integrals related to Radon-like transforms*, *Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993)*, 1995, pp. 535–551. MR1364908
- [75] T. Tao, *The Bochner-Riesz conjecture implies the restriction conjecture*, *Duke Math. J.* **96** (1999), no. 2, 363–375.
- [76] S. Treil,  *$H^1$  and dyadic  $H^1$*  (2008), available at <http://arxiv.org/pdf/0809.3288>.
- [77] A. M. Vargas, *Weighted weak type  $(1,1)$  bounds for rough operators*, *J. London Math. Soc.* (2) **54** (1996), no. 2, 297–310.
- [78] A. Volberg and P. Zorin-Kranich, *Sparse domination on non-homogeneous spaces with an application to  $A_p$  weights*, *ArXiv e-prints* (June 2016), available at <http://arxiv.org/abs/1606.03340>.