Winter 12-15-2021

A Continuous Wavelet Representation for Single and Bi-Parameter Calder\'on-Zygmund Operators

Tyler Williams
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Mathematics and Statistics

Dissertation Examination Committee:
Brett Wick, Chair
Francesco Di Plinio, Co-Chair
Henri Martikainen
Kabe Moen
Donsub Rim

A Continuous Wavelet Representation for Single and Bi-Parameter Calderón-Zygmund Operators

by

Tyler Williams

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2021
St. Louis, Missouri

Table of Contents

Acknowledgments

First, I would like to thank my advisors Brett Wick and Francesco Di Plinio whose expertise, insights, and encouragements guided me through graduate school and made this thesis possible. Thank you for the countless help, patience, and understanding during my time in school.

Next, thank you to the committee members: Henri Martikainen, Donsub Rim, and Kabe Moen. All of whom are excellent examples of great mathematicians and whose work helped shape this thesis.

Thank you to Washington University in St. Louis and the Department of Mathematics and Statistics for no better place to research and go to graduate school.

I would like to personally thank my fellow graduate students whose friendships made this process possible. Thank you to Ben Castor, my roommate throughout our time in St. Louis, to Chris Felder, Nathan Wagner, and Cody Stockdale whose friendships and collaboration resulted in many memories and succesful mathematics in Tower Grove East. Lastly to Alberto Dayan, Manasa Vempati, Jeet Sampat, Claire Huang, Atzimba Martinez, and Jeff Norton all of whom I have depended on for insights and mathematical discourse.

Thank you to the University of Alabama and its Mathematics Department who helped me lay a foundation for graduate school especially Kabe Moen for introducing me to harmonic analysis and wavelets, Lawrence Roberts, and David Cruz-Uribe.

Thank you to the Homewood City Schools whose amazing educators fostered the importance of learning and my personal joy for mathematics especially Jennifer Clifton, Sherry Putt, Tracy Carpenter, Laura Thuirer, Mark Hellmers, Tim Hurry, and Dr. Bill Cleveland. I would not be the person or mathematician I am today without them.

Thank you to my grandparents Fox, Lou, Mammaw, and Grandaddy. Your love, support, and words of wisdom shaped me into the person I am today. I love all of you.

Thank you to my siblings, cousins, aunts/uncles, and family especially my sister Sarah for opening her home which made moving away and driving home possible.

Finally, thank you to the entire Pappas and Rotton Family for the continued care and love as one of your own. A very special and sincere thank you to my future Father and Mother in law, Dennis and Kellie Pappas both for your daughter and the never ending support.

Tyler Williams

Washington University in St. Louis

December 2021

Dedicated to Maizy Pappas.

Your love and encouragement made this journey possible; I would not have been able to do this without you. I look forward to our life together.

In loving memory of my nephew Oliver Butler,

And my grandaddy, Jack Williams, you made me into the man I am today; I love you and cherished every moment we shared.

ABSTRACT OF THE DISSERTATION

A Continuous Wavelet Representation for Single and Bi-Parameter Calderón-Zygmund Operators

by

Tyler Williams

Doctor of Philosophy in Mathematics

Washington University in St. Louis, 2021

Professor Brett D. Wick, Chair

Professor Francesco Di Plinio, Co-Chair

This thesis develops a novel approach to the representation of singular integral operators of Calderón-Zygmund type in terms of continuous model operators, in both the classical and the bi-parametric setting. The representation is realized as a finite sum of averages of wavelet projections of either cancellative or noncancellative type, which are themselves Calderón-Zygmund operators. Both properties are out of reach for the established dyadic-probabilistic technique. Unlike their dyadic counterparts, this new representation reflects the additional kernel smoothness of the operator being analyzed.

These representation formulas lead naturally to a new family of $T(1)$ theorems on weighted Sobolev spaces whose smoothness index is naturally related to kernel smoothness. In the one parameter case, the Sobolev space analogue of the A_2 theorem is proven; that is, sharp dependence of the Sobolev norm of T on the weight characteristic is obtained in the full range of exponents. In the bi-parametric setting, where local average sparse domination is not generally available, quantitative A_p estimates are established which are best known, and sharp in the range $\max\{p, p'\} \geq 3$ for the fully cancellative case.

Chapter 1

Introduction and Statement of Main Results

In the area of harmonic analysis, singular integral operators and more specifically Calderón-Zygmund theory has been studied extensively for decades with applications in partial differential equations and complex analysis, and connections to Littlewood-Paley theory and wavelets. A singular integral operator acts on a function by integration against a kernel, K, that asymptotically fails to be integrable. Calderón-Zygmund theory places continuity and cancellation conditions on these kernels and studies when and how these operators are bounded on L^2. The $T(1)$ theorem gives such a characterization. The rich theory developed by Alberto Calderón and Antoni Zygmund is the cornerstone of this thesis, specifically the L^p boundedness of these operators, T, for $1 < p < \infty$.

Beginning as the study of frequency localization operators, Littlewood-Paley theory gave a way to decompose functions as a superposition of components localized to dyadic annuli in frequency. These decompositions and the associated square function gave a nice way to study L^2 and L^p behavior of the Calderón-Zygmund operator , T. In fact, this square function addresses the boundedness of T at the end point $p = 1$, where the real Hardy Space H^1 is mapped into L^1. The space H^1 can be characterized as the space of functions the square function maps into L^1. Additionally, square function space characterizations are a crucial tool in the lauded duality of H^1 and BMO of C. Fefferman.

Much like square functions, wavelets are powerful tool used in function space characterizations and sampling theory. A square integrable function ϕ is a wavelet if the collection of functions $\{\phi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}^d}$ is an orthonormal basis for $L^2(\mathbb{R}^d)$ with $\phi_{m,n}(x) := 2^{\frac{md}{2}}\phi(2^m x - n)$. The most widely known wavelet is the Haar basis with the mother wavelet given by $\phi(x) =$

$\mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}$The Haar basis enters the definition of martingale transforms and dyadic square functions, generalized versions of these dyadic operators may in turn be used to represent Calderón-Zygmund operators [28,39]. This thesis takes advantage of square functions defined via wavelets with more smoothness than the Haar wavelet. The idea of smooth wavelet coefficients goes back to the work of David and Journé [13] in their work on $T(1)$ type theorems.

1.1 Single Parameter Calderón-Zygmund Forms

We first begin in the single parameter setting for Calderón-Zygmund theory on $\mathbb{R}^d$ with the Lebesgue measure of a set E given by $|E|$. By a single parameter, we mean we only consider objects invariant under a single scale dilation. It is convenient to employ the Japanese bracket symbol

$$\langle x \rangle = \max\{1, |x|\}, \qquad x \in \mathbb{R}^d.$$

Lastly, as with Fourier analysis, necessary is the Schwartz space, $\mathcal{S}(\mathbb{R}^d)$, of all smooth functions whose derivatives decay faster than any polynomial. We consider the class of Calderón-Zygmund operators invariant under single parameter families of translations and dilations on $\mathbb{R}^d$. In this section, Λ stands for a continuous bilinear form on $\mathcal{S}(\mathbb{R}^d)$ with adjoint form

$$\Lambda^\star : \mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}, \qquad \Lambda^\star(f,g) := \overline{\Lambda(g,f)}.$$

and two adjoint linear continuous operators

$$T, T^\star : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d), \qquad \langle Tf, g \rangle = \Lambda(f,g), \qquad \langle T^\star f, g \rangle = \overline{\Lambda(g,f)}.$$

Below $k \in \mathbb{N}$ and $\delta > 0$ are two parameters quantifying the weak boundedness and off-diagonal kernel smoothness of the form Λ. This quantification is summarized by the norm

$$\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^d,k,\delta)} := \|\Lambda\|_{\mathrm{WB},\delta} + \|\Lambda\|_{\mathrm{K},k,\delta} \tag{1.1.1}$$

with the quantities on the right hand side defined below.

Definition 1.1.1 (Weak boundedness)**.** The form Λ has the *δ-weak boundedness property* if there exists $C > 0$ such that

$$s^d|\Lambda(\varphi_z, \upsilon_z)| \le C$$

uniformly over all $z = (x, s) \in Z^d$, $\varphi_z, \upsilon_z \in \Psi_z^{\delta;1}$ with supp φ_z, supp $\upsilon_z \subset \mathsf{B}_z$. In this case, call $\|\Lambda\|_{\mathrm{WB},\delta}$ the least such constant C.

Definition 1.1.2 (Kernel estimates)**.** For a function $K = K(u, v) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, recall the finite difference notation

$$\Delta_{h|\cdot}K(u,v) = K(u+h,v)-K(u,v), \qquad \Delta_{\cdot|h}K(u,v) = K(u,v+h)-K(u,v), \qquad u,v,h \in \mathbb{R}^d.$$

The continuous bilinear form Λ on $\mathcal{S}(\mathbb{R}^d)$ has the *standard (k, δ)-kernel estimates* if the following holds. There exists a function $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, k-times continuously differentiable away from the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$ such that

$$\Lambda(f, g) = \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} K(u, v)f(v)g(u)\,\mathrm{d}v\mathrm{d}u$$

whenever $f, g \in \mathcal{S}(\mathbb{R}^d)$ are disjointly supported, and satisfying the size and smoothness estimates for all $u \neq v \in \mathbb{R}^d$, $h \in \mathbb{R}^d$ with $0 < |h| \leq \frac{1}{2}|u-v|$:

$$|u-v|^d \left[|\nabla_u^\kappa K(u,v)| + |\nabla_v^\kappa K(u,v)|\right] \leq C, \qquad 0 \leq \kappa \leq k; \tag{1.1.2}$$

$$|u-v|^{d+k} \left[|\Delta_{h|\cdot} \nabla_u^k K(u,v)| + |\Delta_{\cdot|h} \nabla_v^k K(u,v)|\right] \leq C \left(\frac{|h|}{|u-v|}\right)^\delta. \tag{1.1.3}$$

We call $\|\Lambda\|_{\mathrm{K},k,\delta}$ the least constant C such that (1.1.2) and (1.1.3) hold and say that $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ if the constant (1.1.1) is finite. In the case $k = 0$, (1.1.2) and (1.1.3) reduce to the usual size and smoothness estimates for Calderón-Zygmund kernels. The following are two examples of such operators.

Example 1.1.1. In dimension $d = 1$, the Hilbert Transform is a example of such an operator

$$H(f)(x) = \frac{1}{\pi} \text{ p.v.} \int_{\mathbb{R}} \frac{f(y)}{x-y} \, \mathrm{d}y. \tag{1.1.4}$$

Example 1.1.2. In dimension d,

$$R_j(f)(x) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \text{ p.v.} \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x-y|^{d+1}} f(y) \, \mathrm{d}y \tag{1.1.5}$$

A natural question is how do L^2 bounded operators act on $L^1(\mathbb{R}^d)$? Unfortunately, these operators are not bounded here. For $0 < p < \infty$, the *weak* L^p space

$$L^{p,\infty}(\mathbb{R}^d) \coloneqq \{f \text{ is measurable} : \|f\|_{L^{p,\infty}(\mathbb{R}^d)} < \infty\}$$

with,

$$\|f\|_{L^{p,\infty}(\mathbb{R}^d)} \coloneqq \sup_{\alpha > 0} \alpha |\{x \in \mathbb{R}^d : |f(x)| > \alpha\}|^{\frac{1}{p}}$$

is strictly larger than L^p. Notice that $|x|^{\frac{-d}{p}} \in L^{p,\infty}(\mathbb{R}^d) \setminus L^p(\mathbb{R}^d)$. The following well-known Lemma (1.1.1) may be used to prove the weak $(1,1)$ bound $T : L^1(\mathbb{R}^d) \to L^{1,\infty}(\mathbb{R}^d)$ and answers our previous question.

Lemma 1.1.1. Let $f \in L^1_{loc}(\mathbb{R}^d)$ and $\alpha > 0$. Then there exist functions g and b on $\mathbb{R}^d$ and dyadic cubes $\{Q_j\}_j$ with $Q_j \cap Q_k$ when $j \neq k$ such that

$$f = g + b \tag{1.1.6}$$

$$\|g\|_{L^1} \leq \|f\|_{L^1} \text{ and } \|g\|_{L^\infty} \leq 2^d\alpha \tag{1.1.7}$$

$$b = \sum_j b_j, \text{ where each } b_j \text{ is supported in the dyadic cube } Q_j \tag{1.1.8}$$

$$\int\limits_{Q_j} b_j \, \mathrm{d}x = 0 \tag{1.1.9}$$

$$\|b_j\|_{L^1} \leq 2^{d+1}\alpha|Q_j| \tag{1.1.10}$$

$$\sum_j |Q_j| \leq \alpha^{-1}\|f\|_{L^1}. \tag{1.1.11}$$

The thus obtained weak $(1,1)$ bound [8] and L^2 boundedness along with interpolation gives the bound $T : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ for $1 < p < 2$. By symmetry of the assumption when passing to the adjoint, this result is extended to the full range, $1 < p < \infty$.

Definition 1.1.3. For a locally integrable function f on $\mathbb{R}^d$ and

$$\|f\|_{BMO} := \sup_Q \frac{1}{|Q|} \int\limits_Q |f - \langle f \rangle_Q| dx \tag{1.1.12}$$

f is of *bounded mean oscillation* if $\|f\|_{BMO} < \infty$ and $BMO(\mathbb{R}^d)$ is the set of all such locally integrable f with $\|f\|_{BMO} < \infty$.

At the other endpoint, we have $T : L^\infty(\mathbb{R}^d) \to BMO(\mathbb{R}^d)$. The L^2 boundedness of T is equivalent to $T(1), T^\star(1) \in BMO$. Generalizing to a larger set of measures, a weight, w, is a nonnegative locally integrable function. For a Lebesgue measurable set E, w gives $w(E) := \int_E w(x)\,\mathrm{d}x$. Naturally, one can ask what condition on w gives the weighted L^p bound $T : L^p(w) \to L^p(w)$ and that is given by the following

Definition 1.1.4. Let w be a weight on $\mathbb{R}^d$ is said to be an A_p *Muckenhoupt weight* for $1 < p < \infty$ if

$$[w]_{A_p} := \sup_Q \left(\frac{1}{|Q|} \int\limits_Q w\,dx \right) \left(\frac{1}{|Q|} \int\limits_Q w^{-\frac{1}{p-1}} \right)^{p-1} < \infty \tag{1.1.13}$$

and for $p = 1$ if

$$[w]_{A_1} := \sup_Q \left(\frac{1}{|Q|} \int\limits_Q w\,dx \right) \|w^{-1}\|_{L^\infty(Q)} < \infty. \tag{1.1.14}$$

Qualitative weighted bounds are due to Muckenhoupt. The first sharp bound for a singular integral operator was proven when Petermichl represented the Hilbert transform as an average of dyadic shifts [44]. This idea of representation of an operator was extended to the Beurling transform[14], Riesz transform [45], dyadic paraproducts [6], sufficiently nice convolution Calderón-Zygmund operators[50]. These results were cases which confirmed the hypothesized bound in the A_2 conjecture which states: for $f \in L^2(w)$ and $w \in A_2$ the quantitative bound for a Calderón-Zygmund operator is

$$\|T\|_{L^2(w)\to L^2(w)} \lesssim_{d,p,T} [w]_{A_2}.$$

For $1 < p < \infty$, this inequality with Rubio de Francia's extrapolation theorem implies the weighted bound $\|T\|_{L^p(w)\to L^p(w)} \lesssim_{d,p,T} [w]_{A_p}^{\max(1,\frac{p'}{p})}$. The known cases above were brought to full generality in the seminal work of Hytönen [28] where a Calderón-Zygmund operator was represented as a infinite average of shifted dyadic operators over random dyadic grids.

Theorem 1.1.1. If T is a Calderón-Zygmund Operator satisfying the kernel size and smoothness and weak boundedness estimates along with $T\mathbf{1}, T^\star\mathbf{1} \in \mathrm{BMO}(\mathbb{R}^d)$ then

$$\langle Tf, g\rangle = \left[\mathbb{E}_\omega \sum_{u,v=0}^{\infty} 2^{-\varepsilon(u+v)} \langle T_{u,v,D(\omega)} f, g\rangle \right].$$

This one parameter result was extended by Martikainen to the bi-parameter setting [39] and later to include the use of smoother wavelets instead of Haar wavelets [27]. In this thesis we pursue a different representation strategy which is especially advantageous when trying to exploit additional kernel smoothness and obtain estimates in smoothness spaces such as Sobolev and Besov spaces. In addition, the building blocks of our representation will have the same invariance properties enjoyed by Calderón-Zygmund operators.

Theorem. Let T be a linear operator on $\mathbb{R}^d$, satisfying the weak boundedness testing condition, the standard δ-kernel estimates for some $\delta > 0$ and with $T1, T^\star 1 \in \mathrm{BMO}(\mathbb{R}^d)$. Let $0 < \varepsilon < \delta$. Then there exists a family of L^1-adapted, ε-smooth and $(d+\varepsilon)$-decaying cancellative wavelets $\{\upsilon_{(y,t)} : y \in \mathbb{R}^d, t > 0\}$, such that

$$Tf(x) = \int\limits_{\mathbb{R}^d \times (0,\infty)} \langle f, \varphi_{(y,t)} \rangle \upsilon_{(y,t)}(x) \, \frac{\mathrm{d}y\mathrm{d}t}{t} + \Pi_{T1} f(x) + \Pi^\star_{T^\star 1} f(x), \qquad x \in \mathbb{R}^d$$

where $\varphi_{(y,t)}$ is the (y,t)-rescaling of a smooth mother cancellative wavelet with compact support, Π_b are explicitly constructed paraproducts of the form in Definition 2.4.2.

In the body of the thesis, we show how this type of representation leads to sparse bounds and to A_2 bounds for Sobolev spaces. Finally, the representation theorem we obtain can be used to prove sparse bounds, which we now describe.

Definition 1.1.5. A family of cubes $\mathcal{S}$ on $\mathbb{R}^d$ is called *sparse* if for all $Q \in \mathcal{S}$ there is $E_Q \subset Q$ such that $|E_Q| > \delta|Q|$ with $0 < \delta < 1$ and if $Q \neq Q'$ then $E_Q \cap E_{Q'} = \emptyset$.

For a fixed sparse collection $\mathcal{S}$, define a sparse operator $\mathcal{A}_\mathcal{S}$ as

$$\mathcal{A}_\mathcal{S} f(x) := \sum_{Q \in \mathcal{S}} \langle f \rangle_Q \mathbf{1}_Q(x).$$

Lerner first used this technique on bounds for oscillations of dyadic operators [33].

1.2 Multiple Parameter Calderón-Zygmund Operators

Throughout, $\mathbf{d} = (d_1, d_2)$ is used to keep track of dimension in each parameter. The base space is the product Euclidean space

$$x = (x_1, x_2) \in \mathbb{R}^{\mathbf{d}} \coloneqq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}.$$

In this section Λ indicates a generic bilinear continuous form on $\mathcal{S}(\mathbb{R}^{\mathbf{d}}) \times \mathcal{S}(\mathbb{R}^{\mathbf{d}})$. If $f_j \in \mathcal{S}(\mathbb{R}^{d_j})$ for $j = 1, 2$, then $f_1 \otimes f_2 \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$ stands for $(x_1, x_2) \mapsto f_1(x_1)f_2(x_2)$. Define the full adjoint of Λ by

$$\Lambda^\star : \mathcal{S}(\mathbb{R}^{\mathbf{d}}) \times \mathcal{S}(\mathbb{R}^{\mathbf{d}}) \to \mathbb{C}, \qquad \Lambda^\star(f, g) \coloneqq \overline{\Lambda(g, f)},$$

and, when Λ acts on tensor products, the partial adjoints are given by

$$\begin{aligned}
&\Lambda^{\star_1}, \Lambda^{\star_2} : \mathcal{S}(\mathbb{R}^{d_1}) \otimes \mathcal{S}(\mathbb{R}^{d_2}) \times \mathcal{S}(\mathbb{R}^{d_1}) \otimes \mathcal{S}(\mathbb{R}^{d_2}) \to \mathbb{C}, \\
&\Lambda^{\star_1}(f_1 \otimes f_2, g_1 \otimes g_2) \coloneqq \overline{\Lambda(g_1 \otimes f_2, f_1 \otimes g_2)}, \qquad \Lambda^{\star_2}(f_1 \otimes f_2, g_1 \otimes g_2) \coloneqq \overline{\Lambda(f_1 \otimes g_2, g_1 \otimes f_2)}.
\end{aligned}$$

To unify, we write $\Lambda^\circ = \Lambda$ and $\Lambda^{\mathfrak{a}}$, with $\mathfrak{a}$ varying in the set $\vec{\mathfrak{a}} = \{\circ, \star, \star_1, \star_2\}$ At times, the adjoint linear operators $T^{\mathfrak{a}} : \mathcal{S}(\mathbb{R}^{\mathbf{d}}) \to \mathcal{S}'(\mathbb{R}^{\mathbf{d}})$,

$$\langle T^{\mathfrak{a}}(f_1 \otimes f_2), g_1 \otimes g_2 \rangle \coloneqq \Lambda^{\mathfrak{a}}(f_1 \otimes f_2, g_1 \otimes g_2), \qquad \mathfrak{a} \in \vec{\mathfrak{a}}$$

will be considered. A bi-parameter wavelet basis is needed. For $j = 1, 2$ let $\varphi_j \in \mathcal{C}_0^\infty(\mathbb{R}^{d_j})$ be such that $\varphi = \varphi_j$, $d = d_j$ in (2.1.7), and $D \geq 8(d_1 + d_2)$ sufficiently large. Set $\varphi \coloneqq \varphi_1 \otimes \varphi_2 \in$

$\mathcal{C}_0^\infty(\mathbb{R}^{\mathbf{d}})$ as our mother wavelet on $\mathbb{R}^{\mathbf{d}}$, and rescale it by

$$\varphi_z = \varphi_{z_1} \otimes \varphi_{z_2} \coloneqq \mathsf{Sy}_z \varphi = \mathsf{Sy}_{z_1}\varphi_1 \otimes \mathsf{Sy}_{z_2}\varphi_2, \qquad z = (z_1, z_2) \in Z^{\mathbf{d}}. \tag{1.2.1}$$

With this position, $\varphi_z \in c\Psi_z^{(D,D),1;0}$ for all $z \in Z^{\mathbf{d}}$. The boundedness and kernel smoothness properties of bi-parameter singulars are quantified by the parameters $k = (k_1, k_2) \in \mathbb{N}^2$ and $\delta > 0$, and summarized by the norm

$$\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} \coloneqq \|\Lambda\|_{\mathrm{PWB},k,\delta} + \|\Lambda\|_{\mathrm{K,f},k,\delta} \tag{1.2.2}$$

whose summands are described below. Notice that the norm (1.2.2) is stable under full and partial adjoints: this fact will be used without explicit mention from now on.

Definition 1.2.1 (Partial kernel and weak boundedness)**.** For $j = 1, 2$, $z_j \in Z^{d_j}$, and $u_{z_j}, v_{z_j} \in \Psi_{z_j}^{k_j,\delta;1}$ with $\operatorname{supp} u_{z_j}, \operatorname{supp} v_{z_j} \subset \mathsf{B}_{z_j}$, define the forms

$$\begin{aligned} &\Lambda_{1,u_{z_1},v_{z_1}} : \mathcal{S}(\mathbb{R}^{d_2}) \times \mathcal{S}(\mathbb{R}^{d_2}) \to \mathbb{C}, \qquad \Lambda_{1,u_{z_1},v_{z_1}}(f_2, g_2) = s_1^{d_1}\Lambda(u_{z_1} \otimes f_2, v_{z_1} \otimes g_2), \\ &\Lambda_{2,u_{z_2},v_{z_2}} : \mathcal{S}(\mathbb{R}^{d_1}) \times \mathcal{S}(\mathbb{R}^{d_1}) \to \mathbb{C}, \qquad \Lambda_{2,u_{z_2},v_{z_2}}(f_1, g_1) = s_2^{d_2}\Lambda(f_1 \otimes u_{z_2}, g_1 \otimes v_{z_2}). \end{aligned}$$

The form Λ has the (k,δ)-*partial kernel and weak boundedness* properties if there exists $C > 0$ such that

$$\left\|\Lambda_{1,u_{z_1},v_{z_1}}\right\|_{\mathrm{WB},\delta} + \left\|\Lambda_{1,u_{z_2},v_{z_2}}\right\|_{\mathrm{K},k_2,\delta} + \left\|\Lambda_{2,u_{z_2},v_{z_2}}\right\|_{\mathrm{WB},\delta} + \left\|\Lambda_{2,u_{z_2},v_{z_2}}\right\|_{\mathrm{K},k_1,\delta} \leq C$$

uniformly over $z_j \in Z^{d_j}$ and $u_{z_j}, v_{z_j} \in \Psi_{z_j}^{k_j,\delta;1}$ with $\operatorname{supp} u_{z_j}, \operatorname{supp} v_{z_j} \subset \mathsf{B}_{z_j}$, $j = 1, 2$. In this case, $\|\Lambda\|_{\mathrm{PWB},k,\delta}$ is the least such constant C. The $\|\Lambda\|_{\mathrm{PWB},\delta}$ norm is stable under full and partial adjoints, and subsumes all of weak boundedness and partial kernel assumptions of [39], see also [43].

Definition 1.2.2 (Full kernel)**.** For a function $K = K(u,v)$ on $\mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}$, with $u = (u_1, u_2), v = (v_1, v_2)$ again using the finite difference notation

$$\Delta^1_{h_1|\cdot} K(u,v) \coloneqq K((u_1 + h_1, u_2), v) - K(u,v), \qquad \Delta^2_{h_2|\cdot} K(u,v) \coloneqq K((u_1, u_2 + h_2), v) - K(u,v),$$
$$\Delta^1_{\cdot|h_1} K(u,v) \coloneqq K(u, (v_1 + h_1, v_2)) - K(u,v), \qquad \Delta^2_{\cdot|h_2} K(u,v) \coloneqq K(u, (v_1, v_2 + h_2)) - K(u,v),$$

for $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R}^{\mathbf{d}}$, $h_j \in \mathbb{R}^{d_j}, j = 1, 2$. In preparation for (1.2.4), introduce the norms

$$\|K\|_{\kappa} \coloneqq \sup_{(u,v)\in\mathbb{R}^{\mathbf{d}}\times\mathbb{R}^{\mathbf{d}}} \left(\prod_{j=1,2} |u_j - v_j|^{d_j+\kappa_j} \right) |K(u,v)|,$$
$$\|K\|_{\kappa,\delta,\Delta^1_{\square|\cdot}} \coloneqq \sup_{(u,v)\in\mathbb{R}^{\mathbf{d}}\times\mathbb{R}^{\mathbf{d}}} \sup_{0<2|h_1|<|u_1-v_1|} \frac{|u_1 - v_1|^{d_1+\kappa_1+\delta}}{|h_1|^{\delta}} |u_2 - v_2|^{d_2+\kappa_2} \left|(\Delta^1_{h_1|\cdot} K)(u,v)\right|,$$
$$\|K\|_{\kappa,\delta,\Delta^1_{\square|\cdot}\Delta^2_{\square|\cdot}} \coloneqq \sup_{(u,v)\in\mathbb{R}^{\mathbf{d}}\times\mathbb{R}^{\mathbf{d}}} \sup_{\substack{0<2|h_j|<|u_j-v_j| \\ j=1,2}} \left(\prod_{j=1,2} \frac{|u_j - v_j|^{d_j+\kappa_j+\delta}}{|h_j|^{\delta}} \right) \left|(\Delta^1_{h_1|\cdot}\Delta^2_{h_2|\cdot} K)(u,v)\right|,$$

with similar definitions for the other finite difference operators: here $\kappa = (\kappa_1, \kappa_2) \in [0, \infty)^2$ and $\delta > 0$. Then the form Λ satisfies the *full kernel estimates* if the following holds. There exists a $k = (k_1, k_2)$-times continuously differentiable $K(u,v)$ on $\mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}$ such that

$$\Lambda(f_1 \otimes f_2, g_1 \otimes g_2) = \int_{\mathbb{R}^{\mathbf{d}}\times\mathbb{R}^{\mathbf{d}}} K(u,v) f_1(u_1) f_2(u_2) g_1(v_1) g_2(v_2) \, \mathrm{d}u\mathrm{d}v \tag{1.2.3}$$

for all tuples $f_j, g_j \in \mathcal{S}(\mathbb{R}^{d_j})$ such that $\operatorname{supp} f_j \cap \operatorname{supp} g_j = \varnothing$, $j = 1, 2$, with the property that for all $0 \leq \kappa_1 \leq k_1,\ 0 \leq \kappa_2 \leq k_2$,

$$\begin{aligned}
\left\|\nabla_{u_1}^{\kappa_1}\nabla_{u_2}^{\kappa_2}K\right\|_{(\kappa_1,\kappa_2)} + \left\|\nabla_{v_1}^{\kappa_1}\nabla_{v_2}^{\kappa_2}K\right\|_{(\kappa_1,\kappa_2)} &\leq C,\\
\left\|\nabla_{u_1}^{k_1}\nabla_{u_2}^{\kappa_2}K\right\|_{(k_1,\kappa_2),\delta,\Delta^1_{\square|\cdot}} + \left\|\nabla_{v_1}^{k_1}\nabla_{v_2}^{\kappa_2}K\right\|_{(k_1,\kappa_2),\delta,\Delta^1_{\cdot|\square}} &\leq C,\\
\left\|\nabla_{u_1}^{k_1}\nabla_{u_2}^{\kappa_2}K\right\|_{(\kappa_1,k_2),\delta,\Delta^2_{\square|\cdot}} + \left\|\nabla_{v_1}^{k_1}\nabla_{v_2}^{\kappa_2}K\right\|_{(\kappa_1,k_2),\delta,\Delta^2_{\cdot|\square}} &\leq C,\\
\left\|\nabla_{u_1}^{k_1}\nabla_{u_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\square|\cdot}\Delta^2_{\square|\cdot}} + \left\|\nabla_{v_1}^{k_1}\nabla_{v_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\cdot|\square}\Delta^2_{\cdot|\square}} &\leq C,\\
\left\|\nabla_{u_1}^{k_1}\nabla_{v_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\square|\cdot}\Delta^2_{\cdot|\square}} + \left\|\nabla_{v_1}^{k_1}\nabla_{u_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\cdot|\square}\Delta^2_{\square|\cdot}} &\leq C.
\end{aligned} \tag{1.2.4}$$

The least C such that (1.2.3) and (1.2.4) hold for each $\Upsilon \in \mathbf{\Lambda}$ will be denoted by $\|\Lambda\|_{\mathrm{K,f},k,\delta}$. Notice that the latter constant is preserved under full and partial adjoints as well. If $k_1 = k_2 = 0$, these are the usual full kernel estimates of a bi-parameter Calderón-Zygmund operator, see for example [39]. The following is an example of such an operator.

Example 1.2.1. Let H_j be the Hilbert Transforms in the j^{th} direction. Then the tensor product,

$$Tf = (H_1 \otimes H_2)f = \text{ p.v.} \int\limits_{\mathbb{R}^2} \frac{f(v_1, v_2)}{(u_1 - v_1)(u_2 - v_2)}\,\mathrm{d}v_1\mathrm{d}v_2 \tag{1.2.5}$$

is a bi-parameter Calderón-Zygmund operator. In fact when T_1, T_2 are Calderón-Zygmund operators on $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}$ repectively then $T_1 \otimes T_2$ is a Calderón-Zygmund operator on $\mathbb{R}^{\mathbf{d}}$.

Although R. Fefferman proved that convolution type bi-parameter operators, a smaller class than the one above, can be extended to an operator bounded on $L^2(\mathbb{R}^{\mathbf{d}})$ [18], these operators were first described in the sense of vector-valued Calderón-Zygmund theory by Journé [29]. A one parameter operator T maps $T : L^1 \to L^{1,\infty}$, but the proof requires a stopping time argument using the Hardy-Littlewood maximal function. Given the additional geometric freedom in the bi-parameter setting, we must use the strong maximal function,

$$M_s f(x) := \sup_{x \in R} \frac{1}{|R|} \int_R |f(y)| \, \mathrm{d}y,$$

where the supremum is taken over all rectangles. The lack of martingale structure of the strong maximal operator does not allow for similar stopping time arguments. In order to map boundedly into *weak* $L^1(\mathbb{R}^{\mathbf{d}})$, the Orlicz Space $L \log^+ L(\mathbb{R}^{\mathbf{d}})$ is the required space, and along with the Marcinkiewicz Multiplier theorem the L^p bound is established for $1 < p < \infty$ [18]. At the other endpoint, we have the following change to the BMO space.

Definition 1.2.3. (Product BMO) For a locally integrable function f on $\mathbb{R}^{\mathbf{d}}$ and

$$\|f\|_{BMO} := \sup_{\Omega} \frac{1}{|\Omega|} \sum_{R \subset \Omega} |\langle f, \psi_R \rangle|^2 \tag{1.2.6}$$

with the sum above over all dyadic rectangles R contained in the open set Ω. f is of *bounded mean oscillation* if $\|f\|_{BMO} < \infty$ and $BMO(\mathbb{R}^{\mathbf{d}})$ is the set of all such locally integrable f with $\|f\|_{BMO} < \infty$.

It is important to note that supremum above is taken over all open sets because a supremum over only rectangles would give the larger class of functions BMO_{rect} as shown by Carleson's well known counterexample [9]. This rectangle BMO space contains the smaller yet space, *bmo* which is all functions satisfying (1.1.12) with cubes replaced with rectangles. Contrasting to the issues with BMO, the product Muckenhoupt weights, A_p, are the following classes of weights.

Definition 1.2.4. The nonnegative locally integrable function w on $\mathbb{R}^{\mathbf{d}}$, a weight, is said to be an A_p *Muckenhoupt weight* for $1 < p < \infty$ if

$$[w]_{A_p} := \sup_R \left(\frac{1}{|R|} \int_R w \, dx \right) \left(\frac{1}{|R|} \int_R w^{-\frac{1}{p-1}} \right)^{p-1} < \infty \tag{1.2.7}$$

and for $p = 1$ if

$$[w]_{A_1} := \sup_R \left(\frac{1}{|R|} \int_R w \, dx \right) \|w^{-1}\|_{L^\infty(R)} < \infty. \tag{1.2.8}$$

These changes allow for the desired weighted L^p bounds $T : L^p(w) \to L^p(w)$ for bi-parameter operators due to R. Fefferman [16] as one would hope given the one parameter case. The following is Martikainen [39] dyadic representation theorem for bi-parameter operators.

Theorem 1.2.1. If T is a bi-parameter Calderón-Zygmund operator satisfying the partial and full kernel size and smoothness and weak boundedness estimates then

$$\langle Tf, g \rangle = \left[\mathbb{E}_{\omega_n} E_{\omega_m} \sum_{\substack{(i_1,i_2)\in\mathbb{Z}^2_+ \\ (j_1,j_2)\in\mathbb{Z}^2_+}} 2^{-\max(i_1,i_2)\delta/2} 2^{-\max(j_1,j_2)\delta/2} \langle S^{i_1 i_2 j_1 j_2}_{\mathcal{D}_{\omega_n}\mathcal{D}_{\omega_m}} f, g \rangle \right].$$

Using [39] as a black box, some weighted type results of quantitative nature have been obtained by Barron and Pipher via a weighted quantification of the bounds for the dyadic shift operator $S^{i_1 i_2 j_1 j_2}_{\mathcal{D}_{\omega_n}\mathcal{D}_{\omega_m}}$ [3]. These estimates are far from sharp. One reason for this is that [3] derives these weighted bounds from sparse estimates involving square function averages, $\langle Sf \rangle_R$, instead of $\langle f \rangle_R$. The main bi-parameter result of this thesis significantly improves the quantification of [3] and obtains the sharp result in a certain range of exponents. Theorem 3.4.1 contains the precise statement.

Theorem. Let T be a linear operator satisfying the hypotheses of a bi-parameter δ-

Calderón-Zygmund operator as in Chapter 3. Let $0 < \varepsilon < \delta \leq 1$. Then there exists a family of L^1-adapted, ε-smooth and $(d_j + \varepsilon)$-decaying in the j-th parameter, product cancellative wavelets

$$\{\upsilon_{((y_1,t_1),(y_2,t_2))} : y_j \in \mathbb{R}^{d_j}, t_j > 0, j = 1, 2\},$$

such that for $(x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$,

$$Tf(x_1, x_2) = \int\limits_{\mathbb{R}^{d_1}\times(0,\infty)} \int\limits_{\mathbb{R}^{d_2}\times(0,\infty)} \langle f, \varphi_{(y_1,t_1)} \otimes \varphi_{(y_2,t_2)} \rangle \upsilon_{((y_1,t_1),(y_2,t_2))}(x_1, x_2) \frac{\mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}t_1 \mathrm{d}t_2}{t_1 t_2}$$

$$+ \text{ four paraproduct terms} + \text{four partial paraproduct terms.}$$

As a corollary, in the range $\max\{p, p'\} \geq 3$, this main theorem applied to cancellative operators results in the sharp weighted bound $\|T\|_{L^p(w)\to L^p(w)} \lesssim_{d,p,T} [w]_{A_p}^2$.

Chapter 2

Wavelet Representation of One Parameter Calderón-Zygmund Operators and Sparse Bounds

This thesis sets forth a new technique for analyzing singular integral operators based on rank 1 wavelet projections

$$f \mapsto s^d \langle f, \varphi_{(x,s)} \rangle \varphi_{(x,s)}, \qquad (x,s) \in \mathbb{R}^d \times (0,\infty)$$

where $\varphi_{(x,s)}$ are L^1-normalized wavelets living at scale s near the point x. The method used is to instead take a weighted average of these wavelet projections with respect to the operator wavelet coefficients $\langle T\varphi_{(x,s)}, \varphi_{(y,t)} \rangle$, with $(y,t) \in \mathbb{R}^d \times (0,\infty)$. The simplest version of this principle is the resolution of the identity operator (2.1.11) below, widely known as the Calderón reproducing formula. Certainly, the Calderón reproducing formula (2.1.11) and wavelet coefficients have been used countless times in the proof of $T(1)$ type estimates, beginning with the works of David and Journé [13] and Journé in the bi-parametric setting [29]. Our approach takes these seminal ideas one step further, in that we aim for equalities, rather than inequalities, and employ the wavelet coefficients of T in a wavelet averaging procedure instead of estimating them, see Lemma 2.3.1, essentially turning the original wavelet basis into another wavelet family adapted to the operator being analyzed. Our approach takes advantage of the fact that a Calderón-Zygmund operator applied to a smooth wavelet basis with compact support yields again a collection of wavelets, though possibly rougher and with smeared out support.

When analyzing one parameter Calderón-Zygmund operators, this results in a representation formula involving a single, *complexity zero* cancellative operator, a single paraproduct and a single adjoint paraproduct, all of which are Calderón-Zygmund operators themselves, in contrast to the dyadic expansion of e.g. [28], involving probabilistic averaging of countably many dyadic shifts. The following is the main result of this section.

Theorem 2.0.1. Let $k \in \mathbb{N}$, $0 < \varepsilon < \delta \leq 1$. There exists an absolute constant $C = C_{k,\delta,\varepsilon,d}$ such that the following holds. Let Λ be a standard (k,δ)-CZ form, satisfying the weak boundedness condition, the kernel estimates and having paraproducts with normalization $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} \leq 1$. Then, there exists a family $\{\upsilon_z \in C\Psi_z^{k,\varepsilon;0} : z \in Z^d\}$, such that for all $f, g \in \mathcal{S}(\mathbb{R}^d)$

$$\Lambda(f,g) = \int_{Z^d} \langle f, \varphi_z \rangle \langle \upsilon_z, g \rangle \, \mathrm{d}\mu(z) + \sum_{0 \leq |\gamma| \leq k} \Pi_{b_\gamma,\gamma}(f,g) + \Pi_{b_\gamma^\star,\gamma}(g,f), \tag{2.0.1}$$

where, for all $0 \leq |\gamma| \leq k$, $\Pi_{b_\gamma,\gamma}$ and $\Pi_{b_\gamma^\star,\gamma}$ are explicitly constructed paraproducts of the form in Definition 2.4.2, with $\|b_\gamma\|_{\mathrm{BMO}(\mathbb{R}^d)}, \|b_\gamma^\star\|_{\mathrm{BMO}(\mathbb{R}^d)} \leq \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)}$ as in Definition 2.4.3.

2.1 Wavelet Representation

We must discuss the relevant theory and necessary notation before we introduce the wavelets we consider in our representation theorem. Our analysis of these forms is based on a symmetry parameter space description. In the classical, single-parameter setting, the parameter space and its associated natural measure μ are

$$z = (x,s) \in Z^d := \mathbb{R}^d \times (0,\infty), \qquad \int_{Z^d} f(z) \, \mathrm{d}\mu(z) := \int_{\mathbb{R}^d \times (0,\infty)} f(x,s) \, \frac{\mathrm{d}x \mathrm{d}s}{s}.$$

Points of Z^d conveniently parametrize open balls in $\mathbb{R}^d$ by

$$\mathsf{B}_z = \{y \in \mathbb{R}^d : |y - x| < s\}, \qquad z = (x, s) \in Z^d.$$

When two points, or families of points, of Z^d appear in the same statement and the context allows for it, the notation $z = (x, s)$ and the corresponding Greek version $\zeta = (\xi, \sigma)$ are used; for instance, see (2.1.5) below. For each $\zeta = (\xi, \sigma) \in Z^d$ it is convenient to refer to the following partition of Z^d:

$$\begin{aligned} Z^d_+(\zeta) &:= \{z = (x, s) \in Z^d : s \geq \sigma\} = F_+(\zeta) \sqcup S(\zeta) \sqcup A(\zeta), \\ F_+(\zeta) &:= \big\{z = (x, s) \in Z^d_+(\zeta) : |x - \xi| > 3s\big\}, \\ S(\zeta) &:= \big\{z = (x, s) \in Z^d_+(\zeta) : s \in [\sigma, 3\sigma], |x - \xi| < 3s\big\}, \\ A(\zeta) &:= \big\{z = (x, s) \in Z^d_+(\zeta) : s > 3\sigma, |x - \xi| < 3s\big\}. \end{aligned} \tag{2.1.1}$$

A fixed $z \in Z^d$ gives rise to the one parameter family of symmetries on $\phi \in \mathcal{S}(\mathbb{R}^d)$ via the following:

$$\begin{aligned} &\mathsf{Tr}_x \phi(\cdot) = \phi(\cdot - x), \qquad \mathsf{Dil}^p_s \phi(\cdot) = \frac{1}{s^{d/p}} \phi\left(\frac{\cdot}{s}\right), \qquad x \in \mathbb{R}^d,\ s > 0,\ p \in (0, \infty]; \\ &\mathsf{Sy}_z \phi = \mathsf{Dil}^1_s \circ \mathsf{Tr}_x, \qquad z = (x, s) \in Z^d. \end{aligned}$$

If we choose $\alpha \in \mathcal{C}^\infty(\mathbb{R}^d)$, radial and with $\alpha = 1$ on $\mathsf{B}_{(0,2)}$, supp $\alpha \subset \mathsf{B}_{(0,4)}$, then for $z \in Z^d$ define the cutoffs

$$\alpha_z := \mathsf{Tr}_x \mathsf{Dil}^\infty_s \alpha, \qquad \beta_z = 1 - \alpha_z, \qquad z = (x, s) \in Z^d. \tag{2.1.2}$$

Note that supp $\alpha_z \subset 4\mathsf{B}_z = \mathsf{B}_{(x,4s)}$ supp $\beta_z \subset \mathbb{R}^d \setminus 2\mathsf{B}_z$, and unlike most other functions parametrized by $z \in Z^d$ the cutoffs α_z, β_z will always be ∞-normalized. For a parameter

$\nu > 0$, in order to measure decay in Z^d define the function

$$[\cdot]_\nu : Z^d \to (0,1], \qquad [(x,s)]_\nu := \frac{(\min\{1,s\})^\nu}{(\max\{1,s,|x|\})^{d+\nu}} \tag{2.1.3}$$

with the useful property that

$$\int_{Z^d} [z]_\nu \, \mathrm{d}\mu(z) \lesssim_\nu 1, \qquad \nu > 0. \tag{2.1.4}$$

Next, the geometric separation in the parameter space Z^d is described by the function

$$[z,\zeta]_\nu = \frac{1}{s^d}\left[\left(\frac{\xi - x}{s}, \frac{\sigma}{s}\right)\right]_\nu = \frac{(\min\{s,\sigma\})^\nu}{(\max\{s,\sigma,|x-\xi|\})^{d+\nu}}, \qquad z=(x,s),\ \zeta=(\xi,\sigma) \in Z^d. \tag{2.1.5}$$

We now gather the needed properties and tools involving the Fourier transform. Throughout this thesis, for $k \in \mathbb{N}$, a multi-index $\gamma \in \mathbb{R}^d$ with $|\gamma| = k$ and a Schwartz function f with $\widehat{f}(\xi) = O(|\xi|^k)$ as $\xi \to 0$, we denote

$$\partial^{-\gamma} f(x) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \frac{(i\xi)^\gamma}{|\xi|^{2k}} \mathrm{e}^{ix\cdot\xi} \, \mathrm{d}\xi. \tag{2.1.6}$$

Then $\partial^{-\gamma} f \in \mathcal{S}(\mathbb{R}^d)$ and Plancherel's theorem implies the equality

$$\langle f, g\rangle = \sum_{|\gamma|=k} \left\langle \partial^{-\gamma} f, \partial^\gamma g\right\rangle =: \langle \nabla^{-k} f, \nabla^k g\rangle, \qquad g \in \mathcal{S}(\mathbb{R}^d).$$

As denoted above, let $\kappa \in \mathbb{R}$, $u \in \{1,\dots,d\}$ and $|\nabla|^\kappa$, R_u be respectively the κ-th order Riesz potential and u-th Riesz transform on $\mathbb{R}^d$,

$$|\nabla|^{-\kappa} f(x) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} |\xi|^{-\kappa} \widehat{f}(\xi) \mathrm{e}^{ix\cdot\xi} \, \mathrm{d}\xi, \qquad R_u f(x) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \frac{i\xi_u}{|\xi|} \mathrm{e}^{ix\cdot\xi} \, \mathrm{d}\xi.$$

For a multi-index $\gamma = (\gamma_1, \ldots, \gamma_d)$, let $R^\gamma = R_1^{\gamma_1} \circ \cdots \circ R_d^{\gamma_d}$. With this notation $\partial^{-\gamma} = |\nabla|^{-|\gamma|} R^\gamma$ up to a multiplicative constant depending on $d, |\gamma|$ only. This multiplicative constant will be ignored in the subsequent uses of this fact. Finally, we arrive at our chosen mother wavelet that will act as a model for our wavelet classes. Let $\Phi \in \mathcal{C}^\infty(\mathbb{R}^d)$ be radial and supported on $\mathsf{B}_{(0,1)}$, $D \in \mathbb{N}$ a fixed large parameter,[1] and $a = a(d, D) > 0$ chosen so that (2.1.11) below holds. Define the mother wavelet for a radial φ by,

$$\varphi \coloneqq a\Delta^{4D}\Phi \in \mathcal{C}^\infty(\mathbb{R}^d), \qquad \operatorname{supp}\varphi \subset \mathsf{B}_{(0,1)}, \qquad \int_{\mathbb{R}^d} |\varphi| \, \mathrm{d}x = C(d, D). \tag{2.1.7}$$

This definition implies

$$\partial^{-\alpha}\varphi = \partial^\alpha \Delta^{4D-|\alpha|}\Phi \in \mathcal{S}(\mathbb{R}^d), \qquad \operatorname{supp}\partial^{-\alpha}\varphi \subset \mathsf{B}_{(0,1)}, \qquad \forall 0 \leq |\alpha| \leq D, \tag{2.1.8}$$

and in particular

$$\int_{\mathbb{R}^d} x^\gamma \psi(x) \, \mathrm{d}x = 0 \tag{2.1.9}$$

holds for all $\psi \in \{\partial^{-\alpha}\varphi : 0 \leq |\alpha| \leq D\}$ and all $0 \leq |\gamma| \leq D$. The translated, rescaled functions

$$\varphi_z = \mathsf{Sy}_z\varphi \qquad z \in Z^d \tag{2.1.10}$$

yield the Schwartz version of the Calderón reproducing formula [5, 23, 53]

$$h = \int_{Z^d} \langle h, \varphi_\zeta \rangle \varphi_\zeta \, \mathrm{d}\mu(\zeta), \qquad h \in \mathcal{S}(\mathbb{R}^d). \tag{2.1.11}$$

[1] For instance, when proving Theorems 2.0.1, 3.4.1 below, any $D \geq 8(\max\{k_1, k_2\} + d_1 + d_2)$ will suffice.

2.1.1 Wavelet Classes

For $\nu > 0$, $0 < \delta \le 1$ define the norm on $\mathcal{S}(\mathbb{R}^d)$

$$\|\phi\|_{\star,\nu,\delta} = \sup_{x\in\mathbb{R}^d} \langle x\rangle^{d+\nu} |\phi(x)| + \sup_{x\in\mathbb{R}^d} \sup_{\substack{h\in\mathbb{R}^d \\ 0<|h|\le 1}} \langle x\rangle^{d+\nu} \frac{|\phi(x+h)-\phi(x)|}{|h|^\delta}.$$

Using this norm and Sy_z, adapted classes are defined. For $k \in \mathbb{N}$, $0 < \delta \le 1$ set

$$\Psi_z^{k,\delta;1} = \left\{\phi \in \mathcal{S}(\mathbb{R}^d) : s^{|\gamma|} \left\|(\mathsf{Sy}_z)^{-1}\partial^\gamma \phi\right\|_{\star,k+\delta,\delta} \le 1 : 0 \le |\gamma| \le k\right\}, \qquad z=(x,s)\in Z^d.$$

The membership $\phi \in \Psi_z^{k,\delta;1}$, for a fixed $z=(x,s)\in Z^d$, yields the following quantitative decay and smoothness conditions: for each multi-index γ on $\mathbb{R}^d$ with $0 \le |\gamma| \le k$, there holds

$$|\partial^\gamma\phi(y)| \le \frac{1}{s^{d+|\gamma|}} \left\langle \frac{y-x}{s}\right\rangle^{-(d+k+\delta)}, \qquad y \in \mathbb{R}^d; \tag{2.1.12}$$

$$|\partial^\gamma\phi(y+h) - \partial^\gamma\phi(y)| \le \frac{|h|^\delta}{s^{d+|\gamma|+\delta}} \left\langle \frac{y-x}{s}\right\rangle^{-(d+k+\delta)}, \qquad y\in\mathbb{R}^d, h\neq 0. \tag{2.1.13}$$

Then set

$$\Psi_z^{k,\delta;0} := \left\{\psi \in \Psi_z^{k,\delta;1} : \text{ (2.1.9) holds } \forall\, 0 \le |\gamma| \le k\right\}. \tag{2.1.14}$$

When $k=0$, the notation is simplified by writing $\Psi_z^{\delta;1}, \Psi_z^{\delta;0}$ in place of $\Psi_z^{0,\delta;1}, \Psi_z^{0,\delta;0}$. We point out that φ_z defined in (2.1.10) belongs to $C\Psi_z^{D,1;0}$. More generally if $0 \le |\gamma| \le D$

$$\varphi_{\gamma,z} := \mathsf{Sy}_z[\partial^{-\gamma}\varphi] = s^{-|\gamma|}\partial^{-\gamma}\mathsf{Sy}_z\varphi \in C\Psi_z^{D,1;0}, \qquad \operatorname{supp}\varphi_{\gamma,z} \subset \mathsf{B}_z. \tag{2.1.15}$$

Limited decay wavelets enjoy the following almost-orthogonality estimates.

Lemma 2.1.1. Let $0 < \eta < \delta \leq 1$, $0 \leq k \leq D$, $z = (x, s), \zeta = (\xi, \sigma) \in Z^d$ with $s \leq \sigma$. Then

$$\sup_{\psi \in \Psi_z^{k,\delta;0}} \sup_{\phi \in \Psi_\zeta^{k,\delta;1}} |\langle \psi, \phi \rangle| \lesssim_\eta [z, \zeta]_{k+\eta}, \qquad \sup_{\phi \in \Psi_\zeta^{k,\delta;1}} |\langle \varphi_z, \phi \rangle| \lesssim_{k,\delta} [z, \zeta]_{k+\delta}.$$

Proof. Consider the first estimate. By scale invariance and symmetry one may reduce to the case $\zeta = (0, 1)$, $z = (x, s)$ with $s \leq 1$. Further, assume $|x| \geq 1$ as the case $|x| \leq 1$ is strictly easier. In this case $[z, \zeta]_{k+\eta} = s^{k+\eta}|x|^{-(d+\eta)}$. Thanks to the vanishing moment properties of ψ, one can subtract $T_x(y)$, the Taylor polynomial of ϕ of order k centered at x. Then one has

$$|\langle \psi, \phi \rangle| \leq \int_{|y-x|<1} |\phi(y) - T_x(y)|\, |\psi(y)| \, \mathrm{d}y + \int_{|y-x|\geq 1} |T_x(y)||\psi(y)| \, \mathrm{d}y + \int_{|y-x|\geq 1} |\phi(y)||\psi(y)| \, \mathrm{d}y.$$

Using (2.1.13) for $\nabla^k \phi$ and (2.1.12) for ψ,

$$\int_{|y-x|<1} |\phi(y) - T_x(y)|\, |\psi(y)| \, \mathrm{d}y \lesssim \frac{1}{|x|^{(d+\delta+k)}} \int_{|y-x|<1} \frac{|y-x|^{\delta+k}}{\left(1 + \frac{|y-x|}{s}\right)^{d+k+\delta}} \frac{\mathrm{d}y}{s^d}$$
$$\lesssim \frac{s^{k+\delta} |\log s|}{|x|^{(d+k+\delta)}} \lesssim_\eta [z, \zeta]_{k+\eta}.$$

Using (2.1.12) for $\nabla^k \phi$ instead gives,

$$\int_{|y-x|\geq 1} |T_x(y)||\psi(y)| \, \mathrm{d}y \lesssim \frac{s^{k+\delta}}{|x|^{d+k+\delta}} \int_{|y-x|>1} |y-x|^{-(d+\delta)} \mathrm{d}y \lesssim [z, \zeta]_\delta,$$
$$\int_{\substack{|y-x|\geq 1 \\ 2|y|<|x|}} |\phi(y)||\psi(y)| \, \mathrm{d}y \leq \int_{2|y-x|>|x|} |\psi(y)| \, \mathrm{d}y \lesssim \int_{\frac{|x|}{2s}}^{\infty} t^{-(k+\delta+1)} \mathrm{d}t \lesssim [z, \zeta]_\delta,$$
$$\int_{\substack{|y-x|\geq 1 \\ 2|y|>|x|}} |\phi(y)||\psi(y)| \, \mathrm{d}y \leq \frac{1}{|x|^{(d+\delta+k)}} \int_{|y-x|>1} |\psi(y)| \, \mathrm{d}y \lesssim [z, \zeta]_\delta.$$

Assembling the last two displays yields the claimed first estimate.

The second estimate is proved similarly. Again, renormalize to have $\zeta = (0,1)$, $z = (x,s)$ with $s \le 1$. Using equation (2.1.15) to rely on the vanishing mean of $\varphi_{\gamma,z}$, and using the decay (2.1.13) for $\phi \in \Psi_{\zeta}^{k,\delta;1}$ gives

$$|\langle \phi, \varphi_z \rangle| \le s^k \sum_{|\gamma|=k} \int_{\mathsf{B}_z} |\partial^\gamma \phi(y) - \partial^\gamma \phi(x)| |\varphi_{\gamma,z}(y)| \,\mathrm{d}y \lesssim \frac{s^{k+\delta}}{\max\{1,|x|\}^{d+k+\delta}} = [z,\zeta]_{k+\delta},$$

and the proof is complete. □

We include a necessary family of functions along with some properties. Choose a collection of functions $\phi_\gamma \in \mathcal{S}(\mathbb{R}^d)$, indexed by multi-indices $0 \le |\gamma| \le 2D$, supported in the unit cube of $\mathbb{R}^d$ with the properties

$$\int_{\mathbb{R}^d} x^\alpha \phi_\gamma(x) \,\mathrm{d}x = \delta_{\gamma\alpha} \qquad \forall 0 \le |\alpha| \le 2D. \tag{2.1.16}$$

The collection ϕ_γ has been explicitly constructed by Alpert [1], see also the extension to general measures in [47, Theorem 1.1]. The folowing lemma 2.1.17 involves the Alpert basis and is needed in the content of the main proof of this section.

Lemma 2.1.2. Let $z = (x,s), \zeta = (\xi,\sigma) \in Z^d$, and with reference to (2.1.1), $z \in A(\zeta)$. Define

$$P_{z,\zeta}(v) := \sum_{0 \le |\gamma| \le k} \langle \varphi_z, \mathsf{Sy}_\zeta \phi_\gamma \rangle \left(\frac{v-\xi}{\sigma} \right)^\gamma, \qquad \chi_{z,\zeta}(v) := \varphi_z(v) - P_{z,\zeta}(v), \qquad v \in \mathbb{R}^d.$$

Then

$$|\chi_{z,\zeta}(v)| \lesssim \frac{1}{s^d} \left(\frac{|v-\xi|}{s} \right)^k \min\left\{1, \frac{\max\{|v-\xi|,\sigma\}}{s}\right\}, \qquad v \in \mathbb{R}^d. \tag{2.1.17}$$

Proof. Let

$$T_\xi \varphi_z(v) = \sum_{0\le|\gamma|\le k} q_\gamma \left(\frac{v-\xi}{\sigma}\right)^\gamma, \qquad q_\gamma := \frac{\sigma^{|\gamma|}\partial^\gamma \varphi_z(\xi)}{\gamma!} = \langle T_\xi\varphi_z, \mathsf{Sy}_\zeta \phi_\gamma\rangle$$

be the degree k Taylor polynomial of φ_z centered at ξ; the equality involving q_γ is due to (2.1.16). By Taylor's theorem, as $\operatorname{supp} \mathsf{Sy}_\zeta\phi_\gamma \subset \mathsf{B}_\zeta$,

$$\left|\langle \varphi_z, \mathsf{Sy}_\zeta\phi_\gamma\rangle - q_\gamma\right| = \left|\langle \varphi_z - T_\xi\varphi_z, \mathsf{Sy}_\zeta\phi_\gamma\rangle\right| \lesssim \|\varphi_z - T_\xi\varphi_z\|_{L^\infty(\mathsf{B}_\zeta)} \lesssim \frac{\sigma^{k+1}}{s^{d+k+1}}.$$

It follows that

$$|\chi_{z,\zeta}(v)| \lesssim |\varphi_z(v) - T_\xi\varphi_z(v)| + \frac{1}{s^d}\sum_{0\le|\gamma|\le k} \frac{\sigma^{k+1-|\gamma|}}{s^{k+1-|\gamma|}}\left(\frac{|v-\xi|}{s}\right)^{|\gamma|}.$$

The first summand of the last display complies with the estimate in the right hand side of (2.1.17), by Taylor's theorem and the fact that $\varphi_z \in \Psi_z^{k,1;0}$. The second summand is also bounded by the right hand side of (2.1.17): this is easily seen by checking the cases $|v-\xi| \le \sigma, \sigma < |v-\xi| \le s, |v-\xi| > s$ separately. The latter remark completes the proof of the Lemma. □

2.2 Intrinsic Forms and Sparse Estimates

Lemma 2.1.1 leads to the L^2-boundedness of an intrinsic square function associated to the classes $\Psi_z^{\delta;0}$. This square function will now be defined. For $f \in L^p(\mathbb{R}^d)$, $1 \le p \le \infty$, $\iota \in \{0,1\}$ and $z \in Z^d$ define the intrinsic wavelet coefficients

$$\Psi_z^{\delta;\iota} f := \sup_{\psi\in\Psi_z^{\delta;\iota}} |\langle f,\psi\rangle|. \tag{2.2.1}$$

The $\iota = 0$ coefficients enter the intrinsic square function

$$\mathrm{S}_\delta f(x) = \left(\int_0^\infty \left(\Psi^{\delta;0}_{(x,s)} f \right)^2 \frac{\mathrm{d}s}{s} \right)^{\frac{1}{2}}. \tag{2.2.2}$$

The value $\delta > 0$ is fixed but arbitrary and, whenever possible, it will be omitted from the notation in (2.2.1) and (2.2.2), writing for instance $\mathrm{S}f$ instead of $\mathrm{S}_\delta f$. Lemma 2.1.1 implies easily the L^2 estimate of the following proposition.

Proposition 2.2.1. Let $f \in L^2(\mathbb{R}^d)$. Then $\|\mathrm{S}_\delta f\|_2 \lesssim_\delta \|f\|_2$.

Proof. It suffices to work with $f \in L^2(\mathbb{R}^d)$ of unit norm. Standard considerations and a change of variable reduce the claimed bound to the estimate

$$\begin{aligned} &\int_{(x,s)\in\mathbb{R}^d\times(0,\infty)} \int_{(\alpha,\beta)\in\mathbb{R}^d\times(0,1)} |\langle f, \psi_{(x,s)}\rangle| |\langle s^d \psi_{(x,s)}, \psi_{(x+\alpha s,\beta s)}\rangle| |\langle f, \psi_{(x+\alpha s,\beta s)}\rangle| \frac{\mathrm{d}x\mathrm{d}s\mathrm{d}\alpha\mathrm{d}\beta}{s\beta} \\ &\lesssim \int_{z\in Z^d} |\langle f, \psi_z\rangle|^2 \,\mathrm{d}\mu(z) \end{aligned} \tag{2.2.3}$$

with implied constant uniform over the choice of $\psi_z \in \Psi^{\delta;0}_z$ and $z \in Z^d$. Lemma 2.1.1 then yields

$$|\langle s^d \psi_{(x,s)}, \psi_{(x+\alpha s,\beta s)}\rangle| \lesssim [(\alpha,\beta)]_{\frac{\delta}{2}}$$

so that (2.2.3) follows by an application of Cauchy-Schwarz and (2.1.4). □

Next, consider the intrinsic bisublinear form

$$\Psi^\delta(f,g) = \int_{Z^d} \Psi^\delta_z f \cdot \Psi^\delta_z g \,\mathrm{d}\mu(z) \tag{2.2.4}$$

acting on pairs $f, g \in \bigcup_{1\le p\le\infty} L^p(\mathbb{R}^d)$. Note that the sum in (2.2.4) is of nonnegative terms, therefore issues of convergence in the definition may be disregarded. The intrinsic form of

(2.2.4) models cancellative operators of Calderón-Zygmund type. Analogous intrinsic forms modeling paraproducts are also needed in the analysis. Referring to the wavelet coefficients (2.2.1), define on triples $f_j \in L^1_{\text{loc}}(\mathbb{R}^d)$ the intrinsic forms

$$\Pi^{j,\delta}(f_1, f_2, f_3) = \int_{Z^d} \Psi_z^{\delta;1} f_j \left(\prod_{\iota \in \{1,2,3\} \setminus j} \Psi_z^{\delta;0} f_\iota \right) \mathrm{d}\mu(z). \tag{2.2.5}$$

The index j in the notation (2.2.5) identifies the noncancellative index of the paraproduct form. As these are modeling bilinear, one parameter Calderón-Zygmund forms, the case where $j \in \{1, 2\}$ and $f_3 \in \mathrm{BMO}(\mathbb{R}^d)$ is of particular interest. In this case, the simplified notation

$$\pi_{f_3}^{\delta}(f_1, f_2) = \Pi^{1,\delta}(f_1, f_2, f_3) \tag{2.2.6}$$

is adopted. The analogous result to Proposition 2.2.1 for paraproducts follows. As for (2.2.4), when δ is fixed and not important in that context, write π_{f_3} in place of $\pi_{f_3}^{\delta}$.

Theorem 2.2.1. For each pair $f, g \in L^1(\mathbb{R}^d)$ there exists a sparse collection $\mathcal{S}$ of cubes of $\mathbb{R}^d$ with the property that

$$\Psi^{\delta}(f, g) \lesssim_{\delta} \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_Q \langle g \rangle_Q, \qquad \langle f \rangle_Q := \frac{1}{|Q|} \int |f| \mathbf{1}_Q \, dx. \tag{2.2.7}$$

The sparse domination algorithm we employ, originating in [11], revolves around an iterated Calderón-Zygmund decomposition and the decay estimates of the lemma below. It is convenient to introduce some terminology to simplify statements and later proofs.

Let $I, Q \subset \mathbb{R}^d$ be any two cubes. We say that the pair (Q, I) is *floating* if $Q \cap (9I)^c \neq \varnothing$. Notice that (Q, I) might be floating without (I, Q) being floating, namely the property of floating is not a symmetric relation. If (Q, I) is a floating pair, we say that it is of type I if $\ell_Q > \ell_I$ and of type II otherwise. If (Q, I) is a floating pair of dyadic cubes. Then

$Q \cap (9I)^c \neq \varnothing$ forces either of the following properties to hold.

I. If (Q, I) is of type I, there exist $n \geq 1$ and $m \in \mathbb{Z}^d$ with the property that

$$\ell_I = 2^{-n}\ell_Q, \qquad I \subset Q + \ell_Q m;$$

II. If (Q, I) is of type II, there exist $n \geq 1$ and $m \in \mathbb{Z}^d \setminus \{0\}$ with the property that

$$\ell_Q = 2^{-n}\ell_I, \qquad Q \subset I + \ell_I m.$$

We call such m, n the *indices* of the floating pair (Q, I). We record the observations that for each $I, Q \in \mathcal{D}$

$$\begin{aligned} \sum_{I' \in \mathcal{I}(I,m,n)} |I'| &\leq |Q|, \\ \sum_{Q' \in \mathcal{Q}(I,m,n)} |Q'| &\leq |I|, \end{aligned} \tag{2.2.8}$$

where $\mathcal{I}(Q, m, n)$ is the collection of $I' \in \mathcal{D}$ such that (Q, I') is a floating pair of type I with indices m, n, and $\mathcal{Q}(I, m, n)$ is the collection of $Q' \in \mathcal{D}$ such that (Q', I) is a floating pair of type II with indices m, n,

Lemma 2.2.1. Let $(Q, I) \subset \mathbb{R}^d$ be a floating pair with indices m, n. Let $b_I \in L^1(\mathbb{R}^d)$ be supported on I. Referring to (2.2.1), we have

1. If (Q, I) is of type *II*, we have

$$\Phi_Q^{2\delta} b_I \lesssim_\delta \langle b_I \rangle_I \frac{2^{-\delta n}}{(1 + |m|)^{d+\delta}}.$$

2. Suppose in addition that b_I has mean zero. If (Q, I) is of type I, we have

$$\Phi_Q^{2\delta} b_I \lesssim_\delta \langle b_I \rangle_I \frac{2^{-\delta n}}{(1+|m|)^{d+\delta}} \frac{|I|}{|Q|}.$$

Proof. Without loss of generality, we scale out $\langle b_I \rangle_I = 1$, and use $\|b_I\|_1 = |I|$. Suppose (Q, I) is of type II. For $y \in I$, we then have $|y - c_Q| \sim |m|\ell_I$. We can thus use (2.1.12) and estimate, whenever $\phi_Q \in \Phi_Q^{2\delta}$,

$$|\langle b_I, \phi_Q \rangle| \leq \left| \int_I |b_I(y)| \frac{1}{(\ell_Q)^d} \left(\frac{|m|\ell_I}{\ell_Q} \right)^{-(d+\delta)} \mathrm{d}y \right| \lesssim \frac{2^{-\delta n}}{(1+|m|)^{d+\delta}}.$$

Suppose now (Q, I) is of type I and b_I has mean zero. For $y \in I$, we then have $|y - c_Q| \sim (1+|m|)\ell_Q$. We can thus use (2.1.13) and estimate, whenever $\phi_Q \in \Phi_Q^{2\delta}$,

$$|\langle b_I, \psi_Q \rangle| = \left| \int_I b_I(y)[\phi_Q(y) - \phi_Q(c_I)] \, \mathrm{d}y \right| \lesssim \frac{2^{-\delta n}}{(1+|m|)^{d+\delta}} \frac{|I|}{|Q|}.$$

The proof is complete. □

In this proof, $\delta > 0$ is fixed, and we write $\Psi_Q f$ in place of $\Psi_Q^\delta f$. Let $f_1, f_2 \in L^1(\mathbb{R}^d)$ be fixed. First of all

$$\Psi(f_1, f_2) = \sup \Psi_{\mathcal{Q}}(f_1, f_2), \qquad \Psi_{\mathcal{Q}}(f_1, f_2) := \sum_{Q \in \mathcal{Q}} |Q| \prod_{j=1}^{2} \Psi_Q f_j \tag{2.2.9}$$

where the supremum above is taken over all finite subcollections $\mathcal{Q} \subset \mathcal{D}$. Below, for each such $\mathcal{Q}$ we construct a sparse collection $\mathcal{S}(\mathcal{Q})$ with the property that

$$\Psi_{\mathcal{Q}}(f_1, f_2) \lesssim \sum_{Q \in \mathcal{S}(\mathcal{Q})} |Q| \prod_{j=1}^{2} \langle f_j \rangle_Q. \tag{2.2.10}$$

The estimate of Theorem 2.2.1 then follows from (2.2.9) and (2.2.10) by appealing to the *one sparse form rules them all* principle of Lacey and Mena Arias [31, Lemma 4.7], see also [12]. We move to the proof of (2.2.10). From here onwards, we also fix the finite collection $\mathcal{Q}$. For any cube $Q \in \mathcal{D}$ we write

$$f = f^{\mathsf{in},Q} + f^{\mathsf{out},Q}, \qquad f^{\mathsf{in},Q} = f\mathbf{1}_Q, \qquad f^{\mathsf{out},Q} = f\mathbf{1}_{\mathbb{R}^d\setminus Q}. \tag{2.2.11}$$

leading to the definition of the fully localized form, for $h_1, h_2 \in L^1(\mathbb{R}^d)$

$$P_Q(h_1,h_2) = \Psi_{\mathcal{Q}(Q)}(h^{\mathsf{in},7Q}, h^{\mathsf{in},7Q}), \qquad \mathcal{Q}(Q) = \mathcal{Q} \cap \mathcal{D}(Q). \tag{2.2.12}$$

We may choose $Q_0 \in \mathcal{D}$ with the property that each element of $\mathcal{Q}$ is contained in $3Q_0$. Let $Q_\kappa = Q + \kappa\ell_Q$, for $\kappa \in \{1,0,-1\}^d$ be the 3^d dyadic translates of Q, so that $\mathcal{Q}$ splits into the disjoint union of 3^d collections $\mathcal{Q}(Q_\kappa)$. Relying on support considerations, we have

$$\begin{aligned}
\Psi_{\mathcal{Q}}(f_1,f_2) &= \sum_{\kappa\in\{1,0,-1\}^d} \Psi_{\mathcal{Q}(Q_\kappa)}(f_1^{\mathsf{in},5Q_0}, f_2^{\mathsf{in},5Q_0}) + \sum_{(*_1,*_2)\neq(\mathsf{in},\mathsf{in})} \Psi_{\mathcal{Q}}(f_1^{*_1,5Q_0}, f_2^{*_2,5Q_0}) \\
&= \sum_{\kappa\in\{1,0,-1\}^d} P_{Q_\kappa}(f_1^{\mathsf{in},5Q_0}, f_2^{\mathsf{in},5Q_0}) + \sum_{(*_1,*_2)\neq(\mathsf{in},\mathsf{in})} \Psi_{\mathcal{Q}}(f_1^{*_1,5Q_0}, f_2^{*_2,5Q_0}).
\end{aligned} \tag{2.2.13}$$

After the removal of the tail terms of the second sum, (2.2.10) is obtained by iterative application of Lemma 2.2.2 below, starting from $Q = Q_\kappa$ for each κ, with $h_j = f_j^{\mathsf{in},5Q_0}$, $j = 1,2$. The iteration ends after finitely many steps due to the finiteness of the collections $\mathcal{Q}(Q_j)$. We omit the standard details and simply restrict ourselves to carrying out the proof of the tail estimates and of Lemma 2.2.2.

Lemma 2.2.2. Let $Q \in \mathcal{D}$. We may find a collection $\mathcal{I} \subset \mathcal{D}(Q)$ with the property that

$$\sum_{I \in \mathcal{I}} |I| \leq \frac{|Q|}{2}, \tag{2.2.14}$$

$$P_Q(h_1, h_2) \leq C|Q|\langle h_1\rangle_{7Q}\langle h_2\rangle_{7Q} + \sum_{I \in \mathcal{I}} P_I(h_1, h_2). \tag{2.2.15}$$

Tail removal. We present all the details for the case $(*_1, *_2) = (\mathsf{out}, \mathsf{out})$. The other two cases are simpler. Denote by $I^m = Q_0 + m\ell_{Q_0}$. We then have

$$f_j^{\mathsf{out},5Q_0} = \sum_{|m|\geq 3} f_j \mathbf{1}_{I^m}.$$

Note that for $R \in \mathcal{Q}$, which is a dyadic cube contained in $3Q_0$, (R, I^m) is a floating pair of type *II*. Let $\ell_R = 2^{-n}\ell_{Q_0}$. Then by Lemma 2.2.1,

$$\Psi_R(f_j \mathbf{1}_{I^m}) \lesssim_\delta \langle f_j\rangle_{I^m} \frac{2^{\frac{-n\delta}{2}}}{(1+|m|)^{d+\frac{\delta}{2}}}$$

so that

$$|R|\Psi_R(f_1^{\mathsf{out},5Q_0})\Psi_R(f_2^{\mathsf{out},5Q_0}) \lesssim |R|2^{\frac{-n\delta}{2}} \sum_{|m|,|k|\geq 3} \frac{\langle f_1\rangle_{I^m}\langle f_2\rangle_{I^k}}{|m|^{d+\frac{\delta}{2}}|k|^{d+\frac{\delta}{2}}}$$

and summing over all $R \in \mathcal{Q}$,

$$\Psi_{\mathcal{Q}}(f_1^{\mathsf{out},5Q_0}, f_1^{\mathsf{out},5Q_0}) \lesssim |Q_0| \sum_{|m|,|k|\geq 3} \frac{\langle f_1\rangle_{I^m}\langle f_2\rangle_{I^k}}{|m|^d|k|^d}.$$

Now, consider the portion of the above sum when $|m| \geq |k|$,

$$\sum_{|m|\geq|k|} \frac{\langle f_1\rangle_{I^m}\langle f_2\rangle_{I^k}}{|m|^d|k|^d} \lesssim \sum_{n\geq 1} \sum_{2^n\leq|m|<2^{n+1}} \langle f_1\rangle_{I^m} 2^{nd} \sum_{|k|\leq|m|} |I^k|\langle f_2\rangle_{I^k}$$

Notice for a fixed n that $I^m, I^k \subset Q_0^{(n)} = 2^{n+5}Q_0$ and by disjointness,

$$\sum_{2^n \le |m| < 2^{n+1}} \frac{\langle f_1 \rangle_{I^m}}{2^{nd}} \le \sum_{|m| \le 2^{n+1}} \frac{1}{2^{nd}|Q_0|} \int_{I^m} f_1 \lesssim \langle f_1 \rangle_{Q_0^{(n)}}, \qquad \sum_{|k| \le |m|} |I^k| \langle f_2 \rangle_{I^k} \le |Q_0^{(n)}| \langle f_2 \rangle_{Q_0^{(n)}}.$$

Summarizing, we have obtained the estimate

$$\Psi_{\mathcal{Q}}(f_1^{\mathsf{out},5Q_0}, f_1^{\mathsf{out},5Q_0}) \lesssim \sum_n |Q_0^{(n)}| \langle f_1 \rangle_{Q_0^{(n)}} \langle f_2 \rangle_{Q_0^{(n)}}.$$

As the collection $\{Q_0^{(n)} : n \ge 0\}$ is sparse, this contribution is acceptable for (2.2.10). $\square$

Proof of Lemma 2.2.2. There is no loss in generality with replacing h_j by $h_j^{\mathsf{in},7Q}$, $j = 1, 2$, and normalizing

$$\langle h_1 \rangle_{7Q} = \langle h_2 \rangle_{7Q} = 1. \tag{2.2.16}$$

which simplifies notation. We begin the proof by defining

$$O = \left\{ x \in \mathbb{R}^d : \max_{j=1,2} \mathrm{M} h_j(x) > C \right\}.$$

Note that if C is large enough, we have $O \subset 9Q$. Let now

$$\begin{aligned} \mathcal{J} &= \text{maximal } J \in \mathcal{D} \text{ with } 9J \subset O, \\ \mathcal{I} &= \text{maximal elements of } \{I \in \mathcal{D}(Q) : \exists J \in \mathcal{J} \text{ with } I \subset 9J\}. \end{aligned} \tag{2.2.17}$$

Below, we will use the easily verified bounds

$$\sup_{R \in \mathcal{J} \cup \mathcal{I}} \inf_{x \in R} \mathrm{M} h_j(x) \lesssim 1, \qquad j = 1, 2 \tag{2.2.18}$$

and the packing estimates

$$\sum_{J\in\mathcal{J}}|J|+\sum_{I\in\mathcal{I}}|I|\le 2|O|<\frac{|Q|}{2} \tag{2.2.19}$$

which follows from pairwise disjointness and inclusion in O, provided C is chosen sufficiently large. We now decompose $\mathcal{Q}(Q)$ into two disjoint pieces. The first is

$$\mathcal{G}=\{R\in\mathcal{Q}(Q):\nexists J\in\mathcal{J}\text{ with }R\subset 9J\}. \tag{2.2.20}$$

For the remaining cubes $R\in\mathcal{Q}(Q)\backslash\mathcal{G}$ there must be a unique $I\in\mathcal{I}$ with $Q\subset I$. Accordingly,

$$P_Q(f_1,f_2)=\Psi_{\mathcal{G}}(h_1,h_2)+\sum_{I\in\mathcal{I}}P_I(f_1,f_2)+\sum_{I\in\mathcal{I}}\sum_{(*_1,*_2)\neq(\mathsf{in},\mathsf{in})}\Psi_{\mathcal{Q}(I)}(h_1^{*_1,7I},h_2^{*_2,7I}). \tag{2.2.21}$$

The last I-summation term on the right hand side is made of tail terms. Each may be estimated as

$$\sum_{(*_1,*_2)\neq(\mathsf{in},\mathsf{in})}\Psi_{\mathcal{Q}(I)}(h_1^{*_1,7I},h_2^{*_2,7I})\lesssim|I|\prod_{j=1}^{2}\inf_{x\in I}\mathrm{M}h_j(x)\lesssim|I| \tag{2.2.22}$$

with the repeated usage of Lemma 2.2.1 and the help of (2.2.18). The summation in $I\in\mathcal{I}$ is taken care of by the packing estimate (2.2.19). We omit the details, which are similar to the tail removal argument seen above. In view of (2.2.21) and of the same packing estimate that guarantees (2.2.14), Lemma 2.2.2 is completely proved once we establish a suitable bound on $\Psi_{\mathcal{G}}(h_1,h_2)$. We execute a standard Calderón-Zygmund decomposition for h_1,h_2 relative to $J\in\mathcal{J}$, namely set

$$\begin{aligned}&h_j=g_j+b_j,\\&g_j:=h_j\mathbf{1}_{\mathbb{R}^d\backslash O}+\sum_{J\in\mathcal{J}}\left(\frac{1}{|J|}\int_J h_j\right)\mathbf{1}_J,\quad b_j:=\sum_{J\in\mathcal{J}}b_{j,J}=\sum_{J\in\mathcal{J}}\left(h_j-\frac{1}{|J|}\int_J h_j\right)\mathbf{1}_J.\end{aligned} \tag{2.2.23}$$

The triangle inequality then leads to

$$\Psi_{\mathcal{G}}(h_1,h_2) \leq \Psi_{\mathcal{G}}(g_1,g_2) + \Psi_{\mathcal{G}}(g_1,b_2) + \Psi_{\mathcal{G}}(b_1,g_2) + \Psi_{\mathcal{G}}(b_1,b_2) \tag{2.2.24}$$

and each term is estimated separately. Using Proposition 2.2.1 and relying in particular on (2.2.18) for the second step,

$$\Psi_{\mathcal{G}}(g_1,g_2) \leq \|g_1\|_2\|g_2\|_2 \lesssim |Q|.$$

We move to the $(g_1,b_2),(b_1,g_2)$ terms in (2.2.24). By symmetry, we write out the bound for the first pairing only. Using (2.2.18),

$$\Psi_{\mathcal{G}}(g_1,b_2) \leq \left(\sup_{R\in\mathcal{D}} \Psi_R g_1\right) \sum_{R\in\mathcal{G}} |R|\Psi_R b_2 \lesssim \|g_1\|_\infty \sum_{R\in\mathcal{G}} |R|\Psi_R b_2 \lesssim \sum_{R\in\mathcal{G}} |R|\Psi_R b_2.$$

Thus it suffices to prove

$$\sum_{R\in\mathcal{G}} |R|\Psi_R b_2 \lesssim |Q|. \tag{2.2.25}$$

Notice that by construction, for all pairs $(R,J)\in\mathcal{G}\times\mathcal{J}$ there holds $R\not\subset 9J$, namely the pair (R,J) is floating. We can therefore split into type I and II pairs and estimate using Lemma 2.2.1, writing $\eta=\delta/2$ for graphical reasons, as

$$\begin{aligned}
&\sum_{R\in\mathcal{G}} |R|\Psi_R b_1 \leq \sum_{R\in\mathcal{G}} |R| \sum_{J\in\mathcal{J}} \Psi_R b_{2J} \\
&= \sum_{n=0}^{\infty}\sum_{m\in\mathbb{Z}}\sum_{R\in\mathcal{G}} \sum_{\substack{J\in\mathcal{J}\\ J\subset R+m\ell_R\\ \ell_J=2^{-n}\ell_R}} |R|\Psi_R b_{2J} + \sum_{n=0}^{\infty}\sum_{\substack{m\in\mathbb{Z}\\ |m|\geq 4}}\sum_{J\in\mathcal{J}}\sum_{\substack{R\in\mathcal{G}\\ R\subset J+m\ell_J\\ \ell_R=2^{-n}\ell_J}} |R|\Psi_R b_{2J} \\
&\lesssim \sum_{n=0}^{\infty}\sum_{m\in\mathbb{Z}}\sum_{R\in\mathcal{G}} \sum_{\substack{J\in\mathcal{J}\\ J\subset R+m\ell_R\\ \ell_J=2^{-n}\ell_R}} |J|\langle b_{2J}\rangle_J + \sum_{n=0}^{\infty}\sum_{\substack{m\in\mathbb{Z}\\ |m|\geq 4}} \frac{2^{-\eta u}}{\langle m\rangle^{\eta+d}} \sum_{J\in\mathcal{J}}\sum_{\substack{R\in\mathcal{G}\\ R\subset J+m\ell_J\\ \ell_R=2^{-n}\ell_J}} |R|\langle b_{2J}\rangle_J.
\end{aligned} \tag{2.2.26}$$

Fixing (n, m) in the first summation on the last right hand side of (2.2.26), notice that each $J \in \mathcal{J}$ appears at most once, in correspondence with the unique R which is the $-m$ shift of the v-fold dyadic parent of J. Therefore, also in view of (2.2.18) and (2.2.19),

$$\sum_{R \in \mathcal{G}} \sum_{\substack{J \in \mathcal{J} \\ J \subset R + m\ell_R \\ \ell_J = 2^{-n}\ell_R}} |J| \langle b_{2J} \rangle_J \lesssim |Q|. \tag{2.2.27}$$

Summation over (n, m) in (2.2.26) shows that this term complies with (2.2.25). Fixing instead (n, m) in the second summation on the last right hand side of (2.2.26), we have the easy control

$$\sum_{J \in \mathcal{J}} \sum_{\substack{R \in \mathcal{G} \\ R \subset J + m\ell_J \\ \ell_R = 2^{-n}\ell_J}} |R| \langle b_{2J} \rangle_J \leq \sum_{J \in \mathcal{J}} |J| \langle b_{2J} \rangle_J \lesssim |Q|, \tag{2.2.28}$$

so that summing over (n, m) in (2.2.26) finishes the proof of (2.2.25). We now move to handling the (b_1, b_2) term in (2.2.24). We have

$$\Psi_{\mathcal{G}}(b_1, b_2) \leq \left(\sup_{R \in \mathcal{G}} \Psi_R b_1 \right) \sum_{R \in \mathcal{G}} |R| \Psi_R b_2$$

thus, by virtue of (2.2.25), it is enough to prove

$$\Psi_R b_1 \leq \Phi_R b_1 \leq \sum_{\substack{J \in \mathcal{J} \\ \ell_I \geq \ell_R}} \Phi_R b_{1J} + \sum_{\substack{J \in \mathcal{J} \\ \ell_I < \ell_R}} \Phi_R b_{1J} \lesssim 1 \tag{2.2.29}$$

with uniform bound over $R \in \mathcal{G}$. For the first summand in (2.2.29), floating pairs of type I are involved. Fixing $R \in \mathcal{G}$, we rewrite it and estimate using Lemma 2.2.1 as

$$\sum_{n=0}^{\infty} \sum_{m \in \mathbb{Z}} \sum_{\substack{J \in \mathcal{J} \\ J \subset R + m\ell_R \\ \ell_J = 2^{-n}\ell_R}} \Phi_R b_{1J} \lesssim \sum_{n=0}^{\infty} \sum_{m \in \mathbb{Z}} \frac{2^{-\eta n}}{(1 + |m|)^{d+\eta}} \sum_{\substack{J \in \mathcal{J} \\ J \subset R + m\ell_R \\ \ell_J = 2^{-n}\ell_R}} \langle b_{1J} \rangle_J \frac{|J|}{|R|} \lesssim 1$$

having relied on (2.2.18) to estimate $\langle b_{1J}\rangle_J \lesssim 1$. Similarly, if $J(m,n,R)$ is the unique $J \in \mathcal{J}$ with $\ell_R = 2^{-n}\ell_J$ and $R \subset J + m\ell_J$, the second summand in (2.2.29) is rewritten and controlled using Lemma 2.2.1 as

$$\sum_{n=0}^{\infty}\sum_{m\in\mathbb{Z}} \Phi_R b_{1J(m,n,R)} \lesssim \sum_{n=0}^{\infty}\sum_{m\in\mathbb{Z}} \frac{2^{-\eta n}}{(1+|m|)^{d+\eta}} \langle b_{1J(m,n,R)}\rangle_{J(m,n,R)} \lesssim 1.$$

Summarizing, we have proved (2.2.29), and finished the estimation of (2.2.24). The proof of the Lemma is thus complete. □

Theorem 2.2.2. Let $\beta \in \mathrm{BMO}(\mathbb{R}^d)$. For each pair $f_j \in L^1(\mathbb{R}^d)$ there exists a sparse collection $\mathcal{S}$ of cubes of $\mathbb{R}^d$ with the property that

$$\pi_\beta(f_1, f_2) \lesssim_\delta \|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)} \sum_{Q\in\mathcal{S}} |Q| \prod_{j=1}^{2} \langle f_j\rangle_Q.$$

The proof of Theorem 2.2.2 is similar to the proof of the previous theorem so only the necessary changes are presented. It is convenient to isolate the following estimate for the $\Psi_Q^\delta \beta$ coefficients.

Lemma 2.2.3. $\sup_{Q\in\mathcal{D}} \frac{1}{|Q|} \sum_{R\in\mathcal{D}(Q)} |R|(\Psi_Q^\delta \beta)^2 \lesssim \|\beta\|^2_{\mathrm{BMO}(\mathbb{R}^d)}$.

Proof. Notice that, due to the mean zero condition, $\Psi_Q^\delta\beta$ does not depend on the choice of representative within the equivalence class of $\beta \in \mathrm{BMO}(\mathbb{R}^d)$. Thus we may assume that $\int_{3Q} \beta = 0$. Decompose $\beta = \beta^{\mathsf{in},3Q} + \beta^{\mathsf{out},3Q}$ as in (2.2.11). Proposition 2.2.1 and the John-Nirenberg inequality yield

$$\sum_{R\in\mathcal{D}(Q)} |R|(\Psi_Q^\delta \beta^{\mathsf{in},3Q})^2 \leq \left\|S_\delta[\beta^{\mathsf{in},3Q}]\right\|_2^2 \lesssim \left\|\beta^{\mathsf{in},3Q}\right\|_2^2 \lesssim |Q|\|\beta\|^2_{\mathrm{BMO}(\mathbb{R}^d)}. \tag{2.2.30}$$

The crude estimate (2.1.12) may then be used in standard fashion to prove the first inequality in

$$\sum_{R\in\mathcal{D}(Q)} |R|(\Psi^\delta_Q \beta^{\mathsf{out},3Q}) \lesssim \sum_{n\geq 1} 2^{-n(d+\frac{\delta}{2})} \left\|\beta^{\mathsf{in},3^nQ}\right\|_1 \lesssim \|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)} \sum_{n\geq 1} n2^{-n\frac{\delta}{2}}|Q| \lesssim |Q|\|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)}.$$

The estimate of the lemma then follows by combining the last display with the previously obtained (2.2.30). □

Proof of Theorem 2.2.2. Fix $\beta \in \mathrm{BMO}(\mathbb{R}^d)$ of unit norm. We only describe the changes required in the outline of proof of Theorem 2.2.1. Having defined the forms

$$\Pi_{\mathcal{Q}}(h_1,h_2) = \sum_{Q\in\mathcal{Q}} \Phi^\delta_Q h_1 \Psi^\delta_Q h_2 \Psi^\delta_Q \beta$$

for a finite collection $\mathcal{Q} \subset \mathcal{D}$, the localized form replacing (2.2.12) in this context is

$$P_Q(h_1,h_2) = \Pi_{\mathcal{Q}(Q)}(h^{\mathsf{in},7Q}, h^{\mathsf{in},7Q}), \qquad \mathcal{Q}(Q) = \mathcal{Q} \cap \mathcal{D}(Q). \tag{2.2.31}$$

The reduction of Theorem 2.2.2 to the corresponding estimate of Lemma 2.2.2, in particular the tail removal procedure, may be repeated step by step once Lemma 2.2.3 is available. It thus suffices to give a proof of Lemma 2.2.2 for (2.2.31) in place of (2.2.12), and we turn to this task. We enforce once again the normalization (2.2.16), and repeat the construction of the collections $\mathcal{J}, \mathcal{I}, \mathcal{G}$ from (2.2.17) and (2.2.20) leading to the following analogue of (2.2.21):

$$P_Q(f_1,f_2) = \Pi_{\mathcal{G}}(h_1,h_2) + \sum_{I\in\mathcal{I}} P_I(f_1,f_2) + \sum_{I\in\mathcal{I}} \sum_{(*_1,*_2)\neq(\mathsf{in},\mathsf{in})} \Pi_{\mathcal{Q}(I)}(h_1^{*_1,7I}, h_2^{*_2,7I}). \tag{2.2.32}$$

The tail terms in the last summation are controlled via Lemma 2.2.1, Lemma 2.2.3, and relying on help of (2.2.18). We then repeat the Calderón-Zygmund decomposition (2.2.23) and turn to estimating each term in

$$\Pi_{\mathcal{G}}(h_1,h_2)=\Pi_{\mathcal{G}}(g_1,g_2)+\Pi_{\mathcal{G}}(b_1,g_2)+\Pi_{\mathcal{G}}(g_1,b_2)+\Pi_{\mathcal{G}}(b_1,b_2)\lesssim|Q|=|Q|\prod_{j=1}^{2}\langle h_j\rangle_{7Q} \quad (2.2.33)$$

which is where the substantial changes lie. We bound

$$\begin{aligned}\Pi_{\mathcal{G}}(g_1,g_2)&\leq\left(\sup_{R\in\mathcal{D}}\Phi_R g_1\right)\left(\sum_{R\in\mathcal{D}}|R|(\Psi_R g_2)^2\right)^{\frac{1}{2}}\left(\sum_{R\in\mathcal{D}}|R|(\Psi_R\beta)^2\right)^{\frac{1}{2}}\\&\lesssim\|g_1\|_\infty\|g_2\|_2|Q|^{\frac{1}{2}}\|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)}\lesssim|Q|\end{aligned} \quad (2.2.34)$$

using Lemma 2.2.3, Proposition 2.2.1 and subsequently (2.2.18) in the third step. The estimate for the term $\Pi_{\mathcal{G}}(b_1,g_2)$ differs only in the first component, namely

$$\Pi_{\mathcal{G}}(b_1,g_2)\leq\left(\sup_{R\in\mathcal{G}}\Phi_R b_1\right)\left(\sum_{R\in\mathcal{D}}|R|(\Psi_R g_2)^2\right)^{\frac{1}{2}}\left(\sum_{R\in\mathcal{D}}|R|(\Psi_R\beta)^2\right)^{\frac{1}{2}}\lesssim|Q|, \quad (2.2.35)$$

as the first factor has been controlled in (2.2.29). On the other hand,

$$\Pi_{\mathcal{G}}(g_1,b_2)\leq\left(\sup_{R\in\mathcal{D}}\Phi_R g_1\right)\left(\sum_{R\in\mathcal{G}}|R|\Psi_R b_2\right)\left(\sup_{R\in\mathcal{D}}\Psi_R\beta\right)\lesssim\|g_1\|_\infty|Q|\|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)}\lesssim|Q| \quad (2.2.36)$$

with the help of (2.2.18) for the g_1 factor, (2.2.25) for the b_2 factor, and Lemma 2.2.3 for the last factor in the first line Proposition 2.2.1 and subsequently (2.2.18) in the third step. The estimate for the last term is similar, namely

$$\Pi_{\mathcal{G}}(b_1,b_2)\leq\left(\sup_{R\in\mathcal{G}}\Phi_R b_1\right)\left(\sum_{R\in\mathcal{G}}|R|\Psi_R b_2\right)\left(\sup_{R\in\mathcal{D}}\Psi_R\beta\right)\lesssim|Q|\|\beta\|_{\mathrm{BMO}(\mathbb{R}^d)}\lesssim|Q| \quad (2.2.37)$$

where we used (2.2.29), Lemma 2.2.3 and (2.2.25). We have thus achieved the claimed bound in (2.2.33), and in turn, finished the proof of Theorem 2.2.2. □

2.3 Averaging Wavelet Coefficients Lemma

In the representation theorems, the key steps involve a certain averaging of the wavelet φ of (2.1.7) which allows the reduction to single family of wavelets in our representation.

Lemma 2.3.1. Let $\{\varphi_z : z \in Z^d\}$ be as in (2.1.10). Let $0 < \eta < \delta \leq 1$ and $0 \leq k \leq D$. Let $u : Z^d \to \mathbb{C}$ be a Borel measurable function with $|u(z)| \leq 1$. Then, there exists $C \lesssim_{k,\delta,\eta} 1$ such that for all $z = (x, s) \in Z^d$

$$\psi_z(\cdot) \coloneqq \int\limits_{\alpha\in\mathbb{R}^d}\int\limits_{0<\beta\leq 1} \frac{\beta^{k+\delta}u((\alpha,\beta))}{\langle\alpha\rangle^{d+k+\delta}}\varphi_{(x+\alpha s,\beta s)}(\cdot)\frac{\mathrm{d}\beta\mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\delta;0}, \tag{2.3.1}$$

$$\nu_z(\cdot) \coloneqq \int\limits_{\alpha\in\mathbb{R}^d}\int\limits_{\beta>1} \frac{u((\alpha,\beta))}{(\max\{|\alpha|,\beta\})^{d+k+\delta}}\varphi_{(x+\alpha s,\beta s)}(\cdot)\frac{\mathrm{d}\beta\mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\eta;0}. \tag{2.3.2}$$

In particular, with reference to (2.1.3),

$$\upsilon_z \coloneqq \psi_z + \nu_z = \int\limits_{(\alpha,\beta)\in Z^d} [(\alpha,\beta)]_{k+\delta}u((\alpha,\beta))\varphi_{(x+\alpha s,\beta s)}\frac{\mathrm{d}\beta\mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\eta;0}. \tag{2.3.3}$$

The proof of Lemma 2.3.1 is postponed till after the following useful application for our continuous paraproducts.

Lemma 2.3.2. Let $\{\varphi_z : z \in Z^d\}$ be as in (2.1.10). Let $\zeta \in Z^d$ be fixed and $q_\zeta \in \Psi_\zeta^{k,1;1}$ with $\operatorname{supp} q_\zeta \subset \mathsf{B}_\zeta$. Then there exists an absolute constant $C = C(d,k)$ and $\vartheta_\zeta \in C\Psi_\zeta^{k,1;1}$ such that

$$\int\limits_{z\in A(\zeta)} \langle h,\varphi_z\rangle\langle\varphi_z,q_\zeta\rangle\,\mathrm{d}\mu(\zeta) = \langle h,\vartheta_\zeta\rangle \quad \forall h \in \mathcal{S}(\mathbb{R}^d). \tag{2.3.4}$$

Furthermore,

$$\int\limits_{\mathbb{R}^d} x^\gamma\vartheta_\zeta(x)\,\mathrm{d}x = \int\limits_{\mathbb{R}^d} x^\gamma q_\zeta(x)\,\mathrm{d}x, \qquad \forall 0 \leq |\gamma| \leq k. \tag{2.3.5}$$

Proof. Write $\zeta = (\xi, \sigma)$ throughout. Formula (2.1.11) yields that

$$\int_{A(\zeta)} \langle h, \varphi_z \rangle \langle \varphi_z, q_\zeta \rangle \, \mathrm{d}\mu(z) = \langle h, q_\zeta \rangle - \int_{Z^d \setminus A(\zeta)} \langle h, \varphi_z \rangle \langle \varphi_z, q_\zeta \rangle \, \mathrm{d}\mu(z).$$

Support considerations show that $\langle \varphi_z, q_\zeta \rangle = 0$ for $z = (x, s)$ with $|x - \xi| > 3 \max\{\sigma, s\}$. An application of Fubini's theorem leads to the equality

$$\int_{Z^d \setminus A(\zeta)} \langle h, \varphi_z \rangle \langle \varphi_z, q_\zeta \rangle \, \mathrm{d}\mu(z) = \langle h, \psi_\zeta \rangle, \qquad \psi_\zeta \coloneqq \int_{I(\zeta)} \langle \varphi_z, q_\zeta \rangle \varphi_z \, \mathrm{d}\mu(z),$$

where $I(\zeta) \coloneqq \{(x, s) : s \le 3\sigma, |x - \xi| \le 3 \max\{s, \sigma\}\}$. If $(x, s) \in I(\zeta)$, Lemma 2.1.1 implies that $|\langle \varphi_{(x,s)}, q_\zeta \rangle| \lesssim \sigma^{-d} (s/\sigma)^{k+1}$. A change of variable and an application of Lemma 2.3.1, (2.3.1) in particular, shows $\psi_\zeta \in C\Psi_\zeta^{k,1;0}$. The proof is completed by setting $\vartheta_z = q_z - \psi_z$ and deducing (2.3.5) from Fubini's theorem. □

2.3.1 Proof of Averaging Lemma

First of all, Fubini's theorem immediately implies that ψ_z and ν_z inherit the moment properties (2.1.9). The memberships $c\psi_z \in \Psi_z^{k,\delta;1}$, $c\nu_z \in \Psi_z^{k,\eta;1}$ are needed and proved now.

Proof of (2.3.1). By invariance of assumptions and conclusions under the family Sy_z, it suffices to work in the case $z = (0, 1)$. As z is thus fixed below, it is omitted from the subscript notation. We turn to showing that $\|\partial^\gamma \psi\|_{\star, k+\delta, \delta} \lesssim 1$ for each γ with $0 \le |\gamma| = \kappa \le k$. Fix $w \in \mathbb{R}^d$, and let $\phi = \partial^\gamma \varphi$ locally. Then,

$$\partial^\gamma \psi(w) = \int_{\alpha \in \mathbb{R}^d} \int_{0 < \beta \le 1} \frac{\beta^{k-\kappa+\delta} u((\alpha, \beta))}{\langle \alpha \rangle^{d+k+\delta}} \phi\left(\frac{w - \alpha}{\beta}\right) \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta^{d+1}}.$$

Due to the support properties of φ, one observes that the functions

$$\alpha \mapsto \phi\left(\frac{w-\alpha}{\beta}\right), \quad \alpha \mapsto \phi\left(\frac{w+h-\alpha}{\beta}\right),$$

are supported in the cube $Q_w = w + [-3,3]^d$, and $\langle \alpha \rangle \sim \langle w \rangle$ for $\alpha \in Q_w$. Hence,

$$|\partial^\gamma \psi(w)| \lesssim \langle w \rangle^{-(d+k+\delta)} \int\limits_{\substack{\alpha \in Q_w \\ 0<\beta\leq 1}} \beta^\delta \left|\phi\left(\tfrac{w-\alpha}{\beta}\right)\right| \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta^{d+1}} = \int\limits_{\substack{v \in \mathbb{R}^d \\ 0<\beta\leq 1}} \beta^{\delta-1} |\phi(v)| \, \mathrm{d}v \mathrm{d}\beta \lesssim_\delta 1 \tag{2.3.6}$$

by Fubini's theorem and the change of variable $v = \frac{w-\alpha}{\beta}$. Hence $\sup_{x \in \mathbb{R}^d} \langle x \rangle^{k+\delta} |\partial^\gamma \psi(x)| \lesssim 1$. We turn to the Hölder continuity estimate

$$|\partial^\gamma \psi(w+h) - \partial^\gamma \psi(w)| \lesssim |h|^\delta \langle w \rangle^{-(d+k+\delta)}, \qquad h \in \mathbb{R}^d. \tag{2.3.7}$$

This is stronger than $\|\partial^\gamma \psi\|_{\star,k+\delta,\delta} \lesssim 1$ only in the range $|h| \leq \frac{1}{2}$, which will now be assumed. Proceeding as before, two integrals must be controlled

$$\int\limits_{\substack{\alpha \in Q_w \\ 0<\beta\leq \frac{|h|}{2}}} \beta^\delta \left[\left|\phi\left(\tfrac{w-\alpha}{\beta}\right)\right| + \left|\phi\left(\tfrac{w+h-\alpha}{\beta}\right)\right|\right] \frac{\mathrm{d}\alpha \mathrm{d}\beta}{\beta^{d+1}} + \int\limits_{\substack{\alpha \in Q_w \\ \frac{|h|}{2}<\beta\leq 1}} \beta^\delta \left|\phi\left(\tfrac{w+h-\alpha}{\beta}\right) - \phi\left(\tfrac{w-\alpha}{\beta}\right)\right| \frac{\mathrm{d}\alpha \mathrm{d}\beta}{\beta^{d+1}}. \tag{2.3.8}$$

A change of variable shows that both summands in the first integral of (2.3.8) are

$$\lesssim \int_0^{\frac{|h|}{2}} \beta^{\delta-1} \, \mathrm{d}\beta \lesssim |h|^\delta.$$

Notice that in the α-support of the second integral in (2.3.8), that $|h| \leq 2\beta$ and

$$\min\{|w-\alpha|, |w+h-\alpha|\} \leq \beta$$

because of the support property of ϕ. Therefore such support has diameter $\lesssim \beta$. Using this fact and the mean value theorem, the second integral in (2.3.8) is

$$\lesssim |h| \int_{\frac{|h|}{2}}^{1} \beta^{\delta-2}\, \mathrm{d}\beta \lesssim |h|^{\delta}.$$

This completes the proof that $\psi \in C\Psi^{\delta;0}_{(0,1)}$ as desired. □

Proof of (2.3.2). Again normalize $z = (0,1)$. Fixing $0 \leq \kappa \leq k$, and using the local notation $f := \nabla^{\kappa}\nu_z$, it must be shown that $\|f\|_{\star,k+\eta,\eta} \lesssim 1$. Note that

$$f(\cdot) = \int_{\alpha\in\mathbb{R}^d} \int_{\beta>1} \frac{u((\alpha,\beta))}{(\max\{|\alpha|,\beta\})^{d+k+\delta}} \phi\left(\frac{\cdot-\alpha}{\beta}\right) \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta^{d+\kappa+1}},$$

where $\phi = \nabla^{\kappa}\varphi$ locally. Bound the factor $\beta^{-\kappa}$ below by 1, even if it may improve certain estimates slightly. Fix $w, h \in \mathbb{R}^d$ with $|h| \leq \frac{1}{2}$. First, observe that for each $\beta > 1$, the set

$$Q_\beta = \left\{\alpha \in \mathbb{R}^d : \phi\left(\frac{w-\alpha}{\beta}\right) \neq 0\right\} \cup \left\{\alpha \in \mathbb{R}^d : \phi\left(\frac{w+h-\alpha}{\beta}\right) \neq 0\right\}$$

has diameter $\lesssim \beta$ due to the support condition on φ, whence $|Q_\beta| \lesssim \beta^d$. Furthermore, if $|w| \geq 4\beta$ and $\alpha \in Q_\beta$ then $|\alpha| \geq \frac{|w|}{2} \geq 2\beta$. This provides

$$|f(w)| \lesssim \langle w\rangle^{-(d+k+\delta)} \int_{1}^{\max\{\frac{|w|}{4},1\}} \frac{\mathrm{d}\beta}{\beta} + \int_{\max\{\frac{|w|}{4},1\}}^{\infty} \frac{\mathrm{d}\beta}{\beta^{d+k+\delta+1}} \lesssim \langle w\rangle^{-(d+k+\delta)} \log\langle w\rangle \lesssim_\eta \langle w\rangle^{-(d+k+\eta)}.$$

Using the mean value theorem for φ and the previous observations

$$|f(w+h)-f(w)| \leq \int\limits_{\beta>1}\int\limits_{\alpha\in Q_\beta} \left|\phi\left(\frac{w+h-\alpha}{\beta}\right)-\phi\left(\frac{w-\alpha}{\beta}\right)\right| \frac{\mathrm{d}\beta\mathrm{d}\alpha}{(\max\{|\alpha|,\beta\})^{d+k+\delta}\beta^{d+1}}$$
$$\lesssim |h| \int\limits_1^\infty \frac{\mathrm{d}\beta}{\max\{|w|,4\beta\}^{d+k+\delta}\beta^2} \lesssim |h|\langle w\rangle^{-(d+k+\delta)},$$

and collecting the last two estimates is more than enough to show that $\|f\|_{\star,k+\eta,\eta} \lesssim 1$. This also completes the proof of Lemma 2.3.1. □

2.4 Calderón-Zygmund Forms of Class (k,δ)

Given the developed framework and notation we refer back to Definition 1.1.1 and Definition 1.1.2 for $\mathrm{SI}(\mathbb{R}^d,k,\delta)$ the space of Calderón-Zygmund Forms of Class (k,δ). The weak boundedness property of Definition 1.1.1 tests Λ on smooth functions. The recent literature related to $T(1)$ and representation theorems, see for instance [28, 37, 39] and references therein, favors testing conditions on indicator functions. When the form Λ also satisfies kernel estimates, the weak boundedness condition employed in this paper actually follows from indicator-type conditions and is therefore less restrictive. More precisely, suppose that the bilinear form Λ is well defined on $L^\infty_0(\mathbb{R}^d)\times L^\infty_0(\mathbb{R}^d)$ and satisfies

$$s^{-d}\left|\Lambda\left(\mathbf{1}_{\mathsf{B}_z},\mathbf{1}_{\mathsf{B}_z}\right)\right| \leq 1 \qquad \forall z=(x,s)\in Z^d$$

in addition to the δ-kernel estimates (1.1.2) and (1.1.3). Then $\|\Lambda\|_{\mathrm{WB},\delta} \lesssim 1$, namely Λ has the weak boundedness property of Definition 1.1.1. A proof of this implication is found in [48].

We first give two relevant examples of (k,δ) forms which relate to previous results.

Definition 2.4.1 (Wavelet form)**.** Let $\{\beta_z, \upsilon_z \in \Psi_z^{k,\delta;0} : z \in Z^d\}$ be two families of cancellative wavelets. The form

$$\Lambda(f,g) = \int_{Z^d} \langle f, \beta_z\rangle \langle \upsilon_z, g\rangle \,\mathrm{d}\mu(z) \tag{2.4.1}$$

belongs to $\mathrm{SI}(\mathbb{R}^d, k, \delta)$ and $\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^d,k,\delta)} \lesssim 1$.

The weak boundedness property is contained in Proposition 2.2.1 while the (k,δ) kernel estimate is obtained via a standard computation reliant on (2.1.12)-(2.1.13). If $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ is a wavelet form of the type (2.4.1), then the functionals $\Lambda(x^\alpha, \cdot)$ vanish for all $0 \leq |\alpha| \leq k$. This is easily verified by appealing to the cancellation properties of the families $\{\beta_z, \upsilon_z \in \Psi_z^{k,\delta;0} : z \in Z^d\}$.

Definition 2.4.2 (Paraproduct forms)**.** Let $0 \leq |\gamma| \leq D$ be a multi-index. Call the family $\{\vartheta_{\gamma,z} \in C\Psi_z^{D,1;1} : z \in Z^d\}$ a γ-family if

$$\int_{\mathbb{R}^d} x^\alpha \vartheta_{\gamma,z}(x)\,\mathrm{d}x = t^{|\alpha|}\delta_{\gamma\alpha}, \qquad \forall 0 \leq |\alpha| \leq |\gamma|. \tag{2.4.2}$$

For instance, if ϕ_γ satisfies (2.1.16), then $\{\vartheta_{\gamma,z} \coloneqq \mathsf{Sy}_z\phi_\gamma : z \in Z^d\}$ is a γ-family. For a function $b \in \mathrm{BMO}(\mathbb{R}^d)$, and multi-indices γ, α, referring to (2.1.15) for $\varphi_{\alpha,z}$ define

$$\Pi_{b,\gamma,\alpha}(f,g) = \int_{Z^d} \langle b, \varphi_{\alpha,z}\rangle \langle f, \vartheta_{\gamma,z}\rangle \langle \varphi_z, g\rangle \,\mathrm{d}\mu(z). \tag{2.4.3}$$

If $\gamma = \alpha$, simply write $\Pi_{b,\gamma}$. It is important to highlight that $\varphi_{\gamma,z} \in C\Psi_z^{D,1;0}$ for all $z \in Z^d$ via equation (2.1.15). Absolute convergence of the above integral for $f, g \in L^1(\mathbb{R}^d)$ is granted by the easily verified intrinsic estimate

$$|\Pi_{b,\gamma,\alpha}(f,g)| \lesssim \pi_b(f,g)$$

referring to (2.2.6). In particular $\Pi_{b,\gamma,\alpha}$ has the $(1,1)$-sparse bound, which implies $L^2(\mathbb{R}^d)$ estimates and *a fortiori* weak boundedness property of $\Pi_{b,\gamma,\alpha}$, with $\|\Pi_{b,\gamma,\alpha}\|_{\mathrm{WB},\delta} \lesssim \|b\|_{\mathrm{BMO}(\mathbb{R}^d)}$. Standard calculations show that $\|\Pi_{b,\gamma,\alpha}\|_{\mathrm{K},k,1} \lesssim_k \|b\|_{\mathrm{BMO}(\mathbb{R}^d)}$ for all $0 \leq k \leq D$, so that

$$\|\Pi_{b,\gamma,\alpha}\|_{\mathrm{SI}(\mathbb{R}^d,k,1)} \lesssim \|b\|_{\mathrm{BMO}(\mathbb{R}^d)}, \qquad 0 \leq k \leq D.$$

Let $\phi \in \mathcal{S}(\mathbb{R}^d)$ be an auxiliary function with

$$\mathbf{1}_{\mathsf{B}_{(0,1)}} \leq \phi \leq \mathbf{1}_{\mathsf{B}_{(0,2)}}, \tag{2.4.4}$$

and introduce the notation, for each multi-index $0 \leq |\alpha| \leq k$ and $R > 0$

$$p_R^\alpha \in \mathcal{S}(\mathbb{R}^d), \qquad p_R^\alpha(x) = x^\alpha \mathsf{Dil}_R^\infty \phi(x), \quad x \in \mathbb{R}^d. \tag{2.4.5}$$

Let $\mathcal{S}_D(\mathbb{R}^d)$ be the subspace of functions $\psi \in \mathcal{S}(\mathbb{R}^d)$ with the vanishing moment property (2.1.9) for all multi-indices $0 \leq |\alpha| \leq D$. If $0 \leq |\alpha| \leq k < D$ and $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$, the limits

$$\Lambda(x^\alpha, \psi) = \lim_{R\to\infty} \Lambda\left(p_R^\alpha, \psi\right), \qquad \psi \in \mathcal{S}_D(\mathbb{R}^d) \tag{2.4.6}$$

exist, do not depend on the particular choice of ϕ, and define linear continuous functionals on $\mathcal{S}_D(\mathbb{R}^d)$ see [22, Lemma 1.91] for a proof. With (2.4.6) in hand, it is possible to ask whether $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ admits κ-th order paraproducts for $0 \leq \kappa \leq k$.

Definition 2.4.3 (Λ has paraproducts of κ-th order)**.** Say that $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ has *paraproducts of* 0*-th order* if there exists $b_0, b_0^\star \in \mathrm{BMO}(\mathbb{R}^d)$ with the property that

$$\Lambda(\mathbf{1}, \psi) = \langle b_0, \psi\rangle, \qquad \Lambda^\star(\mathbf{1}, \psi) = \langle \psi, b_0^\star\rangle \qquad \forall \psi \in \mathcal{S}_D(\mathbb{R}^d) \tag{2.4.7}$$

If this is the case, referring to (2.4.3), define the 0-*th order cancellative part* of Λ as

$$\Lambda_0(f,g) = \Lambda(f,g) - \left[\Pi_{b_0,0}(f,g) + \Pi_{b_0^\star,0}(g,f)\right]$$

We now define inductively the property of having paraproducts of order κ for $1 \leq \kappa \leq k$. Suppose Λ has paraproducts of order $0 \leq \kappa < k$ and the κ-*th order cancellative part* of Λ has been defined. Then Λ has *paraproducts of* $(\kappa+1)$-*th order* if for each multi-index α with $|\alpha| = \kappa + 1$ there exists $b_\alpha, b_\alpha^\star \in \mathrm{BMO}(\mathbb{R}^d)$ with the property that

$$\Lambda_\kappa(x^\alpha, \psi) = (-1)^{\kappa+1}\langle b_\alpha, \partial^{-\alpha}\psi\rangle, \qquad \Lambda_\kappa^\star(x^\alpha, \psi) = (-1)^{\kappa+1}\langle \partial^{-\alpha}\psi, b_\alpha^\star\rangle \tag{2.4.8}$$

for all $\psi \in \mathcal{S}_D(\mathbb{R}^d)$.

Notice that the pairings on the right hand sides are well defined, as $\partial^{-\alpha}\psi \in H^1(\mathbb{R}^d)$ whenever $\psi \in \mathcal{S}_D(\mathbb{R}^d)$ and $|\alpha| < D$. If this is the case, we define the k-*th order cancellative part* of Λ by

$$\Lambda_{\kappa+1}(f,g) = \Lambda_\kappa(f,g) - \sum_{|\alpha|=\kappa+1}\left[\Pi_{b_\alpha,\alpha}(f,g) + \Pi_{b_\alpha^\star,\alpha}(g,f)\right]. \tag{2.4.9}$$

Here we set $\Lambda_{-1}(f,g) = \Lambda(f,g)$ to be consistent with the definition of $\Lambda_0(f,g)$ given above. Observe that (2.4.9) is equivalent to

$$\Lambda_\kappa(f,g) = \Lambda(f,g) - \sum_{0\leq|\alpha|\leq\kappa} \Pi_{b_\alpha,\alpha}(f,g) + \Pi_{b_\alpha^\star,\alpha}(g,f). \tag{2.4.10}$$

The following are important implications of our choice of definition. The inductive procedure of the proof of Theorem 2.0.1 reduces to the case $\Lambda(f,g) = \Lambda_\kappa(f,g)$; furthermore, the 0-th order condition (2.4.7) is equivalent to the familiar assumption

$$T(\mathbf{1}) = b \in \mathrm{BMO}(\mathbb{R}^d), \qquad T^\star(\mathbf{1}) = b_\star \in \mathrm{BMO}(\mathbb{R}^d).$$

For $0 \leq \kappa \leq k-1$, let $T_\kappa, T_\kappa^\star$ be the adjoint operators to Λ_κ. As R^γ preserves $\mathrm{BMO}(\mathbb{R}^d)$, the condition may be reformulated as

$$|\nabla|^\kappa T_{\kappa-1}(x^\alpha) = a_\alpha \in \mathrm{BMO}(\mathbb{R}^d), \qquad |\nabla|^\kappa T^\star_{0,\kappa-1}(x \mapsto x^\alpha) = a^\star_\alpha \in \mathrm{BMO}(\mathbb{R}^d), \tag{2.4.11}$$

in the sense of $\mathcal{S}'_D(\mathbb{R}^d)$, where $a_\alpha \coloneqq R^\alpha b_\alpha$ and similarly for $a^\star_\alpha$. Lastly, using equations (2.1.15) and (2.4.2), one directly computes

$$\Pi_{b,\gamma,\alpha}(x^\beta, g) = (-1)^{|\alpha|}\delta_{\gamma\beta}\langle b, \partial^{-\alpha} g\rangle, \qquad \Pi^\star_{b,\gamma,\alpha}(x^\beta, f) = 0, \qquad 0 \leq |\beta| \leq |\gamma|. \tag{2.4.12}$$

Thus $\Pi_{b,\gamma}$ has paraproducts of order $|\gamma|$ according to Definition 2.4.3, with $b_\beta = \delta_{\gamma\beta} b$ and $b^\star_\beta = 0$ for all $0 \leq |\beta| \leq |\gamma|$. We now characterize CZ-forms with higher order paraproducts.

Definition 2.4.4. The continuous bilinear form Λ belongs to the class $\mathrm{CZ}(\mathbb{R}^d, k, \delta)$ of (k,δ)-*Calderón-Zygmund (CZ) forms* if $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ and Λ has paraproducts of order k. For further use, define the norm

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} \coloneqq \|\Lambda\|_{\mathrm{SI}(\mathbb{R}^d,k,\delta)} + \sum_{0 \leq |\alpha| \leq k} \left(\|b_\alpha\|_{\mathrm{BMO}(\mathbb{R}^d)} + \|b^\star_\alpha\|_{\mathrm{BMO}(\mathbb{R}^d)} \right). \tag{2.4.13}$$

The statement of Theorem 2.0.1 is the representation and (sparse) $T(1)$-theorem for (k,δ)-CZ forms. The weighted $T(1)$ result is stated separately in Corollary 2.6.1 with the deduction of the corollary given in Subsection 2.6. We now move to proof of Theorem 2.0.1.

2.5 Proof of Single Parameter Representation Theorem

Start by normalizing $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} = 1$. Throughout the proof, the properties (2.1.7)-(2.1.10) and (2.1.16) will be referred to frequently. Recall that $\varepsilon \in (0,\delta)$ is fixed but arbitrary, and let $\eta = \frac{\varepsilon+\delta}{2}$. Throughout the proof, for $z, \zeta \in Z^d$ we write

$$\chi_{z,\zeta} \coloneqq \varphi_z - P_{z,\zeta}\mathbf{1}_{A(\zeta)}(z) \tag{2.5.1}$$

referring to (2.1.1) and Lemma 2.1.2; note that this does not override the definition of Lemma 2.1.2. First, we prove a needed lemma in order to apply the averaging lemma.

Lemma 2.5.1. $|\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z})| \lesssim [(z,\zeta)]_{k+\eta}\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)}$.

Proof. It suffices by symmetry to work in the region $z \in Z^d_+(\zeta)$, see (2.1.1). The estimates are then verified by case analysis.

Case $z \in S(\zeta)$. Estimate $\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) = \Lambda(\varphi_z, \varphi_\zeta)$ appealing to the weak boundedness property. The details are standard and omitted.

Case $z \in A(\zeta)$. Let $\alpha_\zeta, \beta_\zeta$ as in (2.1.2). Then

$$\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) = \Lambda(\chi_{z,\zeta}, \varphi_\zeta) = \Lambda(\Theta, \varphi_\zeta) + \Lambda(\Xi, \varphi_\zeta), \qquad \Theta \coloneqq \chi_{z,\zeta}\alpha_\zeta, \qquad \Xi \coloneqq \chi_{z,\zeta}\beta_\zeta \tag{2.5.2}$$

and one seeks estimates each of the summands in the last right hand side. For the first, apply the weak boundedness property for the point $\tilde{\zeta} = (\xi, 4\sigma)$, so that $\mathsf{B}_{\tilde{\zeta}} = 4\mathsf{B}_\zeta$, and use (2.1.17) to estimate $\|\Theta\|_\infty$, obtaining

$$|\Lambda(\Theta, \varphi_\zeta)| \le \|\Lambda\|_{\mathrm{WB},\delta}\|\Theta\|_\infty \lesssim \frac{1}{s^d}\left(\frac{\sigma}{s}\right)^{k+1} \lesssim [(z,\zeta)]_{k+1}. \tag{2.5.3}$$

Continue now to estimate the second summand. The functions Ξ and φ_ζ have disjoint support, and thus the kernel representation of the form Λ can be used. For each fixed

$v \in \mathbb{R}^d \setminus 2\mathsf{B}_\zeta$, consider the function $F_v \in \mathcal{C}^k(\overline{\mathsf{B}_\zeta})$, $F_v(u) := K(u,v)$ for $u \in \overline{\mathsf{B}_\zeta}$. Then

$$\begin{aligned}
|\Lambda(\Xi,\varphi_\zeta)| &= \left| \int_{\mathbb{R}^d\setminus 2\mathsf{B}_\zeta} \Xi(v) \int_{\mathsf{B}_\zeta} K(u,v)\varphi_\zeta(u)\,\mathrm{d}u\mathrm{d}v \right| = \left| \int_{\mathbb{R}^d\setminus 2\mathsf{B}_\zeta} \Xi(v)\langle F_v, \varphi_\zeta\rangle \,\mathrm{d}v \right| \\
&\le \sigma^k \sum_{|\gamma|=k} \left| \int_{\mathbb{R}^d\setminus 2\mathsf{B}_\zeta} \Xi(v)\langle \partial_u^\gamma F_v, \sigma^{-k}\partial^{-\gamma}\varphi_\zeta\rangle \right| \\
&\le \sigma^k \sum_{|\gamma|=k} \int_{\mathbb{R}^d\setminus 2\mathsf{B}_\zeta} |\Xi(v)| \sup_{u\in\mathsf{B}_\zeta} \left|\Delta_{u-\xi|\cdot}\partial_u^\gamma K(\xi,v)\right| \|\varphi_{\gamma,\zeta}\|_1\,\mathrm{d}v \\
&\lesssim \sigma^{k+\delta}\|\Lambda\|_{\mathrm{K},k,\delta} \int_{\mathbb{R}^d\setminus 2\mathsf{B}_\zeta} \frac{|\Xi(v)|}{|v-\xi|^{d+k+\delta}}\,\mathrm{d}v.
\end{aligned} \tag{2.5.4}$$

Here, the passage to the second line is obtained by using supp φ_ζ, supp $\varphi_{\gamma,\zeta} \subset \mathsf{B}_\zeta$, consult (2.1.8), and integrating by parts. The subsequent (in)equality follows from the mean zero property of $\varphi_{\gamma,\zeta}$, and the next step is obtained via the kernel estimates (1.1.3). Bound the last right hand side by splitting the integral on $\mathbb{R}^d \setminus 2\mathsf{B}_\zeta$ into the pieces

$$\sigma^{k+\delta} \int_{\sigma<|v-\xi|\le s} \frac{|\Xi(v)|}{|v-\xi|^{d+k+\delta}}\,\mathrm{d}v \lesssim \frac{1}{s^d}\left(\frac{\sigma}{s}\right)^{k+\delta} \int_\sigma^s \frac{1}{t}\,\mathrm{d}t \lesssim \frac{1}{s^d}\left(\frac{\sigma}{s}\right)^{k+\delta} \log\left(\frac{s}{\sigma}\right) \tag{2.5.5}$$

where the δ-geometric mean of the estimates in (2.1.17) is used, and

$$\sigma^{k+\delta} \int_{|y-u|>s} \frac{|\Xi(y)|}{|y-\xi|^{d+k+\delta}}\,\mathrm{d}y \lesssim \frac{1}{s^d}\frac{\sigma^{k+\delta}}{s^k} \int_s^\infty \frac{1}{t^{1+\delta}}\,\mathrm{d}t \lesssim \frac{1}{s^d}\left(\frac{\sigma}{s}\right)^{k+\delta}. \tag{2.5.6}$$

Putting together (2.5.2), (2.5.4), (2.5.5) and (2.5.6) provides

$$|\Lambda(\chi_{z,\zeta},\varphi_\zeta)| \lesssim \frac{\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)}}{s^d}\left(\frac{\sigma}{s}\right)^{k+\delta}\log\left(\frac{\sigma}{s}\right) \lesssim_\eta \frac{1}{s^d}\left(\frac{\sigma}{s}\right)^{k+\eta} = [(z,\zeta)]_{k+\eta} \tag{2.5.7}$$

as claimed.

Case $z \in F_+(\zeta)$. In this case $\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) = \Lambda(\varphi_z, \varphi_\zeta)$ and the supports of φ_z and φ_ζ are separated. Thus, the kernel representation of Λ, the cancellation of $\varphi_{\gamma,\zeta}$ and the kernel estimates can be used. Proceeding exactly like in (2.5.4) with φ_z in place of Ξ,

$$|\Lambda(\varphi_z, \varphi_\zeta)| \lesssim \sigma^{k+\delta} \|\Lambda\|_{\mathrm{K},k,\delta} \int_{\mathsf{B}_z} \frac{|\varphi_z(v)|}{|v-\xi|^{d+k+\delta}} \, \mathrm{d}v \lesssim \frac{\sigma^{k+\delta}}{s^d |x-\xi|^{k+\delta}} = [z,\zeta]_{k+\delta} \le [z,\zeta]_{k+\eta}. \tag{2.5.8}$$

This completes the proof of the lemma. □

From here, we complete the main content of the proof by first noticing that

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d),\kappa,\delta} \le \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} = 1, \qquad 0 \le \kappa \le k.$$

First, we begin by induction on $0 \le \kappa \le k$ with the additional assumption

$$a(k)\colon\ b_\gamma, b_\gamma^\star \ne 0 \implies |\gamma| = k$$

namely, all paraproducts vanish except those of highest order. Notice that $a(0)$ is not an extra assumption. Let now $f, g \in \mathcal{S}(\mathbb{R}^d)$. Use (2.1.11), bilinearity, $\mathcal{S}(\mathbb{R}^d)$-continuity of Λ and definition (2.5.1) to expand $\Lambda(f,g)$ as

$$\begin{aligned} \int_{Z^d \times Z^d} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\varphi_z, \varphi_\zeta) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) &= \int_{Z^d \times Z^d} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) \\ + \int_{Z^d_\zeta} \int_{A(\zeta)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(P_{z,\zeta}, \varphi_\zeta) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) &+ \int_{Z^d_z} \int_{A(z)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\varphi_z, P_{\zeta,z}) \, \mathrm{d}\mu(\zeta) \mathrm{d}\mu(z). \end{aligned}$$

Making the change of variable $\xi = x + \alpha s, \sigma = \beta s$ and using Fubini's theorem in the inner variable of $\langle \varphi_{(x+\alpha s, \beta s)}, g \rangle$, the first summand in the last right hand side equals

$$\int_{Z^d} \langle f, \varphi_z \rangle \langle v_z, g \rangle \, \mathrm{d}\mu(z), \quad v_{(x,s)} := \int_{(\alpha,\beta) \in Z^d} [(\alpha,\beta)]_{k+\eta} u_{(x,s)}(\alpha,\beta) \varphi_{(x+\alpha s, \beta s)} \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta} \tag{2.5.9}$$

where

$$u_{(x,s)}(\alpha,\beta) := \frac{\Lambda\left(\chi_{(x,s),(x+\alpha s,\beta s)}, \chi_{(x+\alpha s,\beta s),(x,s)}\right)}{[(\alpha,\beta)]_{k+\eta}}$$

is uniformly bounded by Lemma 2.5.1. Thus $\upsilon_z \in C\Psi_z^{k,\varepsilon;0}$ by Lemma 2.3.1 applied with $u = u_{(x,s)}$. This constructs the first expression in the right hand side of (2.0.1). An alternative form of the term in (2.5.9) with roles of f and g exchanged up to conjugation, may be obtained by making instead the change of variable $x = \xi + \alpha\sigma, s = \beta\sigma$ and applying Lemma 2.3.1 accordingly.

It remains to identify the second and third summand of the main decomposition as a sum of paraproduct terms. Turning to this task for the first term, begin by noticing that due to assumption $a(k)$ used twice, and equation (2.1.15)

$$\Lambda\left(y \mapsto \left(\frac{y-\xi}{\sigma}\right)^{\gamma}, \varphi_\zeta\right) = \Lambda\left(y \mapsto y^{\gamma}, \sigma^{-|\gamma|}\varphi_\zeta\right) = \begin{cases} \langle b_\gamma, \varphi_{\gamma,\zeta}\rangle & |\gamma| = k \\ 0 & |\gamma| \neq k. \end{cases}$$

Therefore, applying Lemma 2.3.2 to $h = f$, $q_\zeta = \mathsf{Sy}_\zeta \phi_\gamma$ to obtain the last equality

$$\begin{aligned} &\int_{Z^d_\zeta} \int_{z \in A(\zeta)} \langle f, \varphi_z\rangle\langle\varphi_\zeta, g\rangle \Lambda(P_{z,\zeta}, \varphi_\zeta)\, \mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) \\ &= \sum_{|\gamma|=k} \int_{Z^d_\zeta} \int_{z\in A(\zeta)} \langle f, \varphi_z\rangle\langle\varphi_z, \mathsf{Sy}_\zeta\phi_\gamma\rangle\langle\varphi_\zeta, g\rangle\langle b_\gamma, \varphi_{\gamma,\zeta}\rangle\, \mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) \\ &= \sum_{|\gamma|=k} \int_{Z^d} \langle f, \vartheta_{\gamma,\zeta}\rangle\langle\varphi_\zeta, g\rangle\langle b_\gamma, \varphi\rangle\, \mathrm{d}\mu(\zeta) := \Pi_{b_\gamma,\gamma}(f,g). \end{aligned}$$

Notice that Lemma 2.3.2 together with (2.1.7)-(2.1.10) and (2.1.16) ensure that $\vartheta_{\gamma,\zeta}$, the output of Lemma 2.3.2 corresponding to $q_\zeta = \mathsf{Sy}_\zeta\phi_\gamma$, belongs to $\Psi_\zeta^{D,1;1}$ and is a γ-family. A totally symmetric argument deals with the third summand in the main decomposition, and completes the proof of (2.0.1) under the assumption $a(k)$. It remains to explain how to obtain (2.0.1) without the assumption $a(k)$. In fact, it will be shown that Λ satisfies an instance of representation (2.0.1) for all $0 \le \kappa \le k$. This is done by induction on κ. Before starting the induction, observe that $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d),\kappa,\delta} \le \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} \le 1$. For $\kappa = 0$, (2.0.1) is achieved in the previous step.

Assume that Λ has been represented in the form (2.0.1) for an integer $0 < \kappa < k$. Taking advantage of Definition 2.4.2, the κ-cancellative part of Λ defined in (2.4.10) satisfies $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} \lesssim 1$ and the equality

$$\Lambda_\kappa(f,g) = \Lambda(f,g) - \sum_{0\le|\gamma|\le\kappa} \Pi_{b_\gamma,\gamma}(f,g) + \Pi_{b^\star_\gamma,\gamma}(g,f) = \int_{Z^d} \langle f, \varphi_z\rangle\langle v_z, g\rangle \,\mathrm{d}\mu(z)$$

for some family $\{v_z \in C\Psi_z^{\kappa,\varepsilon;0} : z \in Z^d\}$. The last equality of the above display tells us that all paraproducts of Λ_κ having order less than or equal to κ equal zero. Therefore, $\Lambda_\kappa(f,g)$ satisfies the assumptions of the theorem, and in addition $a(\kappa+1)$. Apply the previous part of the proof to $\Lambda_\kappa(f,g)$, and obtain

$$\Lambda_\kappa(f,g) = \int_{Z^d} \langle f, \varphi_z\rangle\langle v_z, g\rangle \,\mathrm{d}\mu(z) + \sum_{|\gamma|=\kappa+1} \Pi_{b_\gamma,\gamma}(f,g) + \Pi_{b^\star_\gamma,\gamma}(g,f)$$

with $\{v_z \in C\Psi_z^{\kappa+1,\eta;0} : z \in Z^d\}$. This equality, rearranged, yields a representation of Λ in the form (2.0.1) for the value $\kappa+1$, completing the inductive step and the proof of Theorem 2.0.1.

2.6 Weighted Sobolev Estimates

We come to the $T(1)$ theorem. Notice that (2.6.1) below is a vacuous assumption when $k = 0$, whence Theorem 2.0.1 has a sparse, sharp weighted version of the classical $T(1)$ theorem as a corollary. Also notice that no assumption is being made on the adjoint paraproducts $b^\star_\gamma$.

Corollary 2.6.1. Suppose that Λ is a standard (k,δ)-CZ form with

$$b_\gamma = 0 \qquad \forall 0 \le |\gamma| < k. \tag{2.6.1}$$

Then referring to (2.2.4) and (2.2.6), for each $|\alpha| = k$ the following estimate is true

$$|\Lambda(f, \partial^\alpha g)| \lesssim_\eta \sum_{|\beta|=k} \left[\Psi^\eta(\partial^\beta f, g) + \sum_{|\gamma|=k} \pi_{b_\gamma}(\partial^\beta f, g) + \sum_{0\le|\gamma|\le k} \pi_{b_\gamma^\star}(g, \partial^\beta f) \right]. \tag{2.6.2}$$

Furthermore, the sharp weighted bound on the weighted Sobolev space $\dot{W}^{k,p}(\mathbb{R}^d; w)$ holds

$$\|Tf\|_{\dot{W}^{k,p}(\mathbb{R}^d;w)} \lesssim [w]_{A_p}^{\max\left\{1,\frac{1}{p-1}\right\}} \|f\|_{\dot{W}^{k,p}(\mathbb{R}^d;w)}, \qquad p \in (1,\infty). \tag{2.6.3}$$

We first comment on how to arrive at the necessary hypothesis of the Corollary and the relevance of it. Condition (2.6.1) is also necessary for (2.6.3) to hold, i.e. Corollary 2.6.1 is a characterization of (2.6.3). This generalizes the case $\Omega = \mathbb{R}^d$ of [46, Theorem 1.1] to the non-convolution case; in fact, a scaling argument shows that when $\Omega = \mathbb{R}^d$, condition b. in [46, Theorem 1.1] is equivalent to (2.6.1). To see the necessity, suppose that (2.6.3) holds for some exponent p_0 and all weights $w \in A_{p_0}$. Extrapolation of weighted norm inequalities [10, 15] then implies that (2.6.3) holds for $p = 2d$ and w equals Lebesgue measure. The content of Corollary 2.6.1 also allows to assume that the adjoint T to Λ equals

$$Tf = \sum_{0\le|\gamma|<k} \int_{Z^d} \langle b_\gamma, \varphi_{\gamma,z}\rangle \langle f, \vartheta_{\gamma,z}\rangle \varphi_z \, \mathrm{d}\mu(z).$$

Fix $0 \le |\gamma| = \kappa < k$ and let $\varepsilon := k - (\kappa + \frac{1}{2}) > 0$. Then define $f_R(x) := R^\varepsilon x^\gamma \alpha_{(0,R)}(x)$ where α_z is the cutoff from (2.1.2) and $R > 1$ is arbitrary. It is immediate to show that $\|f_R\|_{\dot{W}^{k,2d}(\mathbb{R}^d)} \sim 1$, so Tf_R is a bounded sequence in $\dot{W}^{k,2d}(\mathbb{R}^d)$. Also, using the properties (2.4.2) followed by (2.1.11), $R^{-\varepsilon} Tf_R \to \partial^{-\gamma} b_\gamma = T(x^\gamma)$ in the sense of Definition 2.4.3.

These two properties entail $T(x^\gamma) = 0$ in $\dot{W}^{k,2d}(\mathbb{R}^d)$. Thus $T(x^\gamma)$ is a polynomial of degree $\leq k$. Appealing to Definition 2.4.3 again reveals that $b_\gamma = 0$ as claimed. Testing type theorems for smooth singular integral operators have previously appeared in several works: a non-exhaustive list includes [22, 25, 40, 49, 51] as well as the already mentioned [27, 46] and references therein. In particular, [51, Theorem 1, cases (6,7)] is essentially equivalent to the unweighted version of Corollary 2.6.1. Corollary 2.6.1 appears to be the first weighted $T(1)$ theorem of this type. A sparse bound in the vein of Corollary 2.6.1 was proved in [4] for the case $k = 1$ using techniques from [35]. However, the result of [4] is not of testing type and was obtained under the stronger assumption that T is *a priori* bounded on the Sobolev space $\dot{W}^{1,2}(\mathbb{R}^d)$.

Proof of Corollary 2.6.1. Before the actual proof, make the following observations referring to the wavelets φ_z, υ_z in the representation (2.0.1): for $z = (x, s) \in Z$, $|\alpha| = |\beta| = |\nu| = k$, $0 \leq |\gamma| \leq k$

$$s^k \partial^\alpha \varphi_z,\ s^k \partial^\alpha \upsilon_z \in C\Psi_z^{\varepsilon;0}, \qquad s^{-k} \partial^{-\beta} \vartheta_{\nu,z},\ s^k \partial^\alpha \vartheta_{\gamma,z} \in C\Psi_z^{1;1}. \tag{2.6.4}$$

Applying the representation theorem to Λ, and using the assumptions on b_γ

$$\Lambda(f, \partial^\alpha g) = \int_{Z^d} \langle f, \varphi_z \rangle \langle \upsilon_z, \partial^\alpha g \rangle \, \mathrm{d}\mu(z) + \sum_{|\gamma|=k} \Pi_{b_\gamma,\gamma}(f, \partial^\alpha g) + \sum_{|\gamma| \leq k} \Pi_{b_\gamma^\star,\gamma}(\partial^\alpha g, f).$$

Integrating by parts and using equation (2.1.15) gives

$$\left| \int_{Z^d} \langle f, \varphi_z \rangle \langle \upsilon_z, \partial^\alpha g \rangle \, \mathrm{d}\mu(z) \right| = \left| \sum_{|\beta|=k} \int_{Z^d} \langle \partial^\beta f, \varphi_{\beta,z} \rangle \left\langle s^k \partial^\alpha \upsilon_z, g \right\rangle \mathrm{d}\mu(z) \right| \lesssim_\eta \sum_{|\beta|=k} \Psi^\eta(\partial^\beta f, g).$$

Fixing $|\nu| = k$ in the b_ν-type paraproduct, and integrating by parts

$$|\Pi_{b_\nu,\nu}(f, \partial^\alpha g)| = \left| \sum_{|\beta|=k} \int_{Z^d} \langle b_\nu, \varphi_{\nu,z}\rangle \left\langle \partial^\beta f, s^{-k}\partial^{-\beta}\vartheta_{\nu,z}\right\rangle \langle s^k \partial^\alpha \varphi_z, g\rangle \,\mathrm{d}\mu(z) \right| \lesssim \sum_{|\beta|=k} \pi_{b_\nu}(\partial^\beta f, g).$$

The $b^\star_\gamma$, $|0| \le k \le \gamma$ type paraproduct is controlled similarly,

$$\left|\Pi_{b^\star_\gamma,\gamma}(\partial^\alpha g, f)\right| = \left| \sum_{|\beta|=k} \int_{Z^d} \langle b^\star_\gamma, \varphi_{\gamma,z}\rangle \langle g, s^k \partial^\alpha \vartheta_{\gamma,z}\rangle \langle \varphi_{\beta,z}, \partial^\beta f\rangle \,\mathrm{d}\mu(z) \right| \lesssim \sum_{|\beta|=k} \pi_{b_\gamma}(\partial^\beta f, g).$$

This completes the proof of (2.6.2). The weighted norm inequality then follows as a consequence of the sparse estimates

$$|\langle \partial^\alpha T f, g\rangle| = |\Lambda(f, \partial^\alpha g)| \lesssim \sum_{Q\in\mathcal{S}} |Q| \langle |\nabla^k f|\rangle_Q \langle g\rangle_Q, \qquad |\alpha| = k$$

obtained by combining (2.6.2) with the Propositions 2.2.1 and 2.2.2. □

Chapter 3

Wavelet Representation of Multiple Parameter Calderón-Zygmund Operators and Sparse Bounds

Bi-parameter singular integrals on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ may be informally defined as elements of the closed convex hull of the set of tensor products $T_1 \otimes T_2$, where each T_j is a $\mathbb{R}^{d_j}$-singular integral operators as above. This class arises naturally in connection with the theory of bi-harmonic functions [19, 21] for instance in the weak factorization of functions in the Hardy space on the bi-disk [20]. The L^p and mixed norm estimates for their multilinear analogues are at the root of partial fractional Leibniz rules [41, 42], namely, anisotropic variants of the bilinear estimates popularized e.g. by Kato-Ponce [30] in connection with the Navier-Stokes, Schrödinger and KdV equations.

One specific drawback of dyadic representations in the bi-parameter context, see [39, 43] for instance, is that they do not reduce L^p and weighted estimates for the analyzed operator to a single model operator whose weighted theory is significantly simpler. In the one parameter case, this can be verified directly for shift operators as in [32] and can also be done by means of sparse operators as in [11]. In higher parameters, the approach via direct verification is challenging [26, 38] and domination by local average sparse operators are generally not available as the counterexample of [2] shows. Thus, for instance, one cannot expect precise information on the dependence of $\|T\|_{L^p(w)}$ on the corresponding relevant weight characteristic. The proof techniques may be naturally transported to the bi-parametric dilation setting. We again paraphrase the main result and point the reader to Theorem 3.4.1, for the precise statement.

Theorem. Let T be a linear operator satisfying the hypotheses of a bi-parameter δ-Calderón-Zygmund operator as in Chapter 3. Let $0 < \varepsilon < \delta \le 1$. Then there exists a family of L^1-adapted, ε-smooth and $(d_j + \varepsilon)$-decaying in the j-th parameter, product cancellative wavelets

$$\{\upsilon_{((y_1,t_1),(y_2,t_2))} : y_j \in \mathbb{R}^{d_j}, t_j > 0, j = 1, 2\},$$

such that for $(x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$,

$$Tf(x_1, x_2) = \int\limits_{\mathbb{R}^{d_1}\times(0,\infty)} \int\limits_{\mathbb{R}^{d_2}\times(0,\infty)} \langle f, \varphi_{(y_1,t_1)} \otimes \varphi_{(y_2,t_2)} \rangle \upsilon_{((y_1,t_1),(y_2,t_2))}(x_1, x_2) \frac{\mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}t_1 \mathrm{d}t_2}{t_1 t_2}$$

$$+ \text{ four paraproduct terms} + \text{four partial paraproduct terms.}$$

Unlike the one parameter results, no smooth $T(1)$ type theorems have appeared in the literature before, even in the unweighted case. This result may be contrasted with the bi-parameter dyadic representation theorem of Martikainen [39]. The assumptions on T are of the same nature as the ones appearing in [39], namely weak boundedness, full and partial kernel estimates, paraproducts in product BMO. However, we drop the diagonal BMO conditions appearing in [39] which are subsumed by a combination of the other assumptions. In addition to the simpler, and more computationally feasible nature of the continuous formula, the model operators we obtain have a much simpler weighted theory, which allows for quantitative, and sharp in certain cases, weighted norm inequalities for T.

Theorem. Let T be a (k_1, k_2)-smooth bi-parameter Calderón-Zygmund operator. If $(k_1, k_2) \neq (0, 0)$ assume in addition the paraproduct condition (3.4.12). For all $1 < p < \infty$

product A_p-weights w on $\mathbb{R}^{\mathbf{d}} := \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ there holds

$$\|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} Tf\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim_{k,\delta} [w]_{A_p}^{\max\left\{3,\frac{2p}{p-1}\right\}} \|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}$$

If T is fully cancellative, the improved estimate

$$\|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} Tf\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim_{k,\delta} [w]_{A_p}^{\theta(p)} \|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}, \qquad \theta(p) = \begin{cases} \frac{2}{p-1} & 1 < p \leq \frac{3}{2} \\ \textit{see } (3.4.14) & \frac{3}{2} < p < 3 \\ 2 & p \geq 3 \end{cases}$$

is available. The above estimate is sharp when $\max\{p, p'\} \geq 3$.

The above result, precisely stated in Corollary 3.4.1, generalizes and quantifies R. Fefferman's qualitative weighted bounds for bi-parameter Journé-type operators [16]. While Martikainen's work [39] did not contain weighted $T(1)$-type implications, a simplified proof of Fefferman's result was recently obtained in [26] relying on the representation from [39]. Some quantitative estimates, weaker than the ones of Corollary 3.4.1, have been obtained in [3] by a shifted square function form-type domination for cancellative Journé operators, also relying on [39] within the proof. At present, it does not seem possible to match the quantification obtained in Corollary 3.4.1 using dyadic representation theorems in the vein of [39, 43]. Part of the challenge with this is that one parameter proofs of the quantitative results typically rely upon some variant of stopping time (sparse operators, weak-type (1,1), or Bellman functions) as a key ingredient and not easily adaptable to the bi-parameter setting. Our analysis based on square function methods is able to circumvent this issue at least for $\max\{p, p'\} \geq 3$.

3.1 Wavelet Representation

This section lays out the bi-parameter analogue of the framework we described in Chapter 2, in preparation to a bi-parameter version of the representation Theorem 2.0.1. Throughout, $\mathbf{d} = (d_1, d_2)$ is used to keep track of dimension in each parameter. The base space is the product Euclidean space

$$x = (x_1, x_2) \in \mathbb{R}^{\mathbf{d}} := \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}.$$

If $\phi \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$ and $F \in \mathcal{S}(\mathbb{R}^{d_1})$, denote

$$\langle \phi, F \rangle_1 = \int_{\mathbb{R}^{d_1}} \phi(x_1, \cdot) F(x_1) \, \mathrm{d}x_1 \in \mathcal{S}(\mathbb{R}^{d_2})$$

and similarly with roles of $1, 2$ reversed. If $\phi : \mathbb{R}^{\mathbf{d}} \to X$, $x_1 \in \mathbb{R}^{d_1}, x_2 \in \mathbb{R}^{d_2}$, the corresponding slices will be denoted by

$$\begin{aligned} &\phi^{[1,x_1]} : \mathbb{R}^{d_2} \to X, \qquad \phi^{[1,x_1]} := \phi(x_1, \cdot), \\ &\phi^{[2,x_2]} : \mathbb{R}^{d_1} \to X, \qquad \phi^{[2,x_2]} := \phi(\cdot, x_2). \end{aligned} \tag{3.1.1}$$

Our parameter space is thus the product space $Z^{\mathbf{d}} = Z^{d_1} \times Z^{d_2}$ with product measure $\mathrm{d}\mu(z) = \mathrm{d}\mu(z_1)\mathrm{d}\mu(z_2)$. Vector notation for points of $Z^{\mathbf{d}}$ is not used and instead, we write $z = (z_1, z_2) \in Z^{\mathbf{d}}$. One embeds Z^{d_j}, $j = 1, 2$ into $Z^{\mathbf{d}}$, regarded as a space of symmetries on $\phi \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$, by taking tensor product with the identity transformation in the complementary parameter. Set

$$\begin{aligned} \left(\mathsf{Sy}^1_{z_1}\phi\right)(y_1, y_2) &:= \left(\mathsf{Sy}_{z_1}\phi^{[2,y_2]}\right)(y_1) = \frac{1}{s_1^{d_1}}\phi\left(\frac{y_1 - x_1}{s_1}, y_2\right), \\ \left(\mathsf{Sy}^2_{z_2}\phi\right)(y_1, y_2) &:= \left(\mathsf{Sy}_{z_2}\phi^{[1,y_1]}\right)(y_2) = \frac{1}{s_2^{d_2}}\phi\left(y_1, \frac{y_2 - x_2}{s_2}\right), \end{aligned}$$

for $z_j = (x_j, s_j) \in Z^{d_j}$. Note that $\mathsf{Sy}^1_{z_1}, \mathsf{Sy}^2_{z_2}$ commute since they act on separate variables. The bi-parameter family of symmetries indexed by $z \in Z^{\mathbf{d}}$ are obtained by composition,

$$\mathsf{Sy}_z\phi = \mathsf{Sy}^1_{z_1} \circ \mathsf{Sy}^2_{z_2}, \qquad z = (z_1, z_2) \in Z^{\mathbf{d}}.$$

3.1.1 Wavelet Classes

For $\nu = (\nu_1, \nu_2) \in (0, \infty)^2$, and $\delta > 0$, define $C_{\star,\nu,\delta}$ to be the subspace of the δ-Hölder continuous functions on $\mathbb{R}^{\mathbf{d}}$ whose norm

$$\|\phi\|_{\star,\nu,\delta} = \sup_{x\in\mathbb{R}^{\mathbf{d}}} \left(\prod_{j=1,2} \langle x_j \rangle^{d_j+\nu_j}\right) |\phi(x)| + \sup_{x\in\mathbb{R}^{\mathbf{d}}} \sup_{\substack{h\in\mathbb{R}^{\mathbf{d}}\\ 0<|h|\leq 1}} \left(\prod_{j=1,2} \langle x_j \rangle^{d_j+\nu_j}\right) \frac{|\phi(x+h)-\phi(x)|}{|h|^{\delta}}$$

is finite. In the bi-parameter case, the relevant cancellation properties of ψ are encoded by requiring (2.1.9) to hold in the variable x_ι for each slice $\psi^{[\widehat{\iota}, x_{\widehat{\iota}}]}$ and each $\iota = 1, 2$. This necessitates the introduction of the bi-parameter analogue of the classes $\Psi^{k,\delta;\iota}_z$. Hereafter, $\gamma_j \in \mathbb{N}^{d_j}$ for either $j = 1, 2$ is a multi-index on $\mathbb{R}^{d_j}$. For $k = (k_1, k_2) \in \mathbb{N}^2$, $0 < \delta \leq 1$ define

$$\Psi^{k,\delta;1,1}_z := \left\{\phi \in \mathcal{S}(\mathbb{R}^{\mathbf{d}}) : s_1^{|\gamma_1|} s_2^{|\gamma_2|} \left\|(\mathsf{Sy}_z)^{-1}\partial^{\gamma_1}\partial^{\gamma_2}\phi\right\|_{\star,k+\delta,\delta} \leq 1\right\},$$

$$\Psi^{k,\delta;0,1}_z := \left\{\phi \in \Psi^{k,\delta;1,1}_z : \text{(2.1.9) holds for } \psi = \psi^{[2,x_2]}, d = d_1,\ \gamma = \gamma_1,\ \forall x_2 \in \mathbb{R}^{d_2}, \forall 0 \leq |\gamma_1| \leq k_1\right\},$$

$$\Psi^{k,\delta;1,0}_z := \left\{\phi \in \Psi^{k,\delta;1,1}_z : \text{(2.1.9) holds for } \psi = \psi^{[1,x_1]}, d = d_2,\ \gamma = \gamma_2,\ \forall x_1 \in \mathbb{R}^{d_1}, \forall 0 \leq |\gamma_2| \leq k_2\right\},$$

$$\Psi^{k,\delta;1,1}_z := \Psi^{k,\delta;0,1}_z \cap \Psi^{k,\delta;1,0}_z,$$

where $z = (z_1, z_2) = ((x_1, s_1), (x_2, s_2)) \in Z^{\mathbf{d}}$. The resulting decay, smoothness and cancellation properties satisfied by $\phi \in \Psi^{k,\delta;\theta_1,\theta_2}_z$ are efficiently described by

$$s_\iota^{d_\iota+|\gamma_\iota|} \left\langle \frac{y_\iota - x_\iota}{s_\iota} \right\rangle^{d_\iota+k_\iota+\delta} [\partial^{\gamma_\iota}\phi]^{[\iota,y_\iota]} \in \Psi^{k_{\widehat{\iota}},\delta;\theta_{\widehat{\iota}}}_{z_{\widehat{\iota}}} \qquad \forall y_\iota \in \mathbb{R}^{d_\iota},\ 0 \leq |\gamma_\iota| \leq k_\iota,\ \iota \in \{1, 2\}. \quad (3.1.2)$$

The analogue of the almost orthogonality Lemma 2.1.1 in the bi-parameter setting is the following lemma.

Lemma 3.1.1. Let $m \in \mathbb{N}$, $0 < 2\eta < \delta \le 1$, $z, \zeta \in Z^{\mathbf{d}}$, $\psi_z \in \Psi_z^{(2m,2m),\delta;0,0}$, $\psi_\zeta \in \Psi_\zeta^{(2m,2m),\delta;0,0}$. Then

$$|\langle \psi_z, \psi_\zeta \rangle| \lesssim_{m,\eta} [z_1, \zeta_1]_{m+\eta} [z_2, \zeta_2]_{m+\eta}.$$

Proof. Let ι be either 1 or 2. Applying (3.1.2) together with the first estimate of Lemma 2.1.1 and integrating,

$$|\langle \psi_z, \psi_\zeta \rangle| = \int_{\mathbb{R}^{d_\iota}} \left| \langle \phi_z^{[\iota, y_\iota]}, \psi_\zeta^{[\iota, y_\iota]} \rangle \right| \, \mathrm{d}y_\iota \lesssim [\max\{s_\iota, \sigma_\iota, |x_\iota - \xi_\iota|\}]^{-d_\iota} [z_{\hat{\iota}}, \zeta_{\hat{\iota}}]_{2m+2\eta}$$

and the lemma follows by taking the 1/2-geometric average of the two inequalities. □

3.2 Intrinsic Square Function and Sparse Estimates

The definition of the intrinsic bi-parameter wavelet coefficients is next given. These may be defined in the generality of $f \in \mathcal{S}'(\mathbb{R}^{\mathbf{d}})$. For such f, and $z = (z_1, z_2) \in Z^{\mathbf{d}}$, set

$$\Psi_z^{\delta;(\iota_1,\iota_2)} f = \sup_{\psi \in \Psi_z^{\delta;\iota_1,\iota_2}} |\langle f, \psi \rangle|. \tag{3.2.1}$$

A standard argument based on (3.1.2) shows that if $f \in L^1_{\mathrm{loc}}(\mathbb{R}^{\mathbf{d}})$,

$$\Psi^{\delta;(1,1)}_{(x_1,s_1),(x_2,s_2)} f \lesssim_\delta \mathrm{M}_{d_1,d_2} f(x), \qquad x = (x_1, x_2) \in \mathbb{R}^{\mathbf{d}}, \quad s_1, \, s_2 > 0$$

where M_{d_1,d_2} is the bi-parameter maximal function on $\mathbb{R}^{\mathbf{d}}$. In particular the wavelet coefficients of $f \in L^p(\mathbb{R}^{\mathbf{d}})$ are finite for $f \in L^p(\mathbb{R}^{\mathbf{d}})$, $p > 1$, as $\mathrm{M}_{d_1,d_2} f$ is finite a.e. in that case.

The remainder of this section contains a basic L^2 estimate for the intrinsic square function

$$\mathrm{SS}_\delta f(x_1,x_2) = \left(\int\limits_{(0,\infty)^2} \left[\Psi^{\delta;(0,0)}_{(x_1,t_1),(x_2,t_2)} f\right]^2 \frac{\mathrm{d}t_1 \mathrm{d}t_2}{t_1 t_2} \right)^{\frac{1}{2}}. \tag{3.2.2}$$

Again, the parameter δ will be fixed and play no role and the operator will be represented as SS later in the paper.

Proposition 3.2.1. $\|\mathrm{SS}_\delta f\|_2 \lesssim_\delta \|f\|_2$. As a consequence,

$$\int\limits_{Z^{\mathbf{d}}} \left[\Psi^{\delta;(0,0)}_z f\right] \left[\Psi^{\delta;(0,0)}_z g\right] \mathrm{d}\mu(z) \lesssim \|f\|_2 \|g\|_2. \tag{3.2.3}$$

Proof. Notice that (3.2.3) follows from the square function estimate via two application of Cauchy-Schwarz inequality. The argument for the square function estimate is analogous to the one employed for (2.2.3). Working with $f \in L^2(\mathbb{R}^d)$ of unit norm, and fixing $\psi_z \in \Psi^{\delta;0,0}_z$, $z \in Z^{\mathbf{d}}$, it suffices to estimate

$$\int\limits_{z\in Z^{\mathbf{d}}} \int\limits_{\substack{(\alpha_1,\beta_1)\in\mathbb{R}^{d_1}\times(0,1) \\ (\alpha_2,\beta_2)\in\mathbb{R}^{d_2}\times(0,1)}} |\langle f,\psi_z\rangle| |\langle s_1^{d_1} s_2^{d_2}\psi_z, \psi_{\zeta(\alpha_1,\beta_1,\alpha_2,\beta_2)}\rangle| |\langle f, \psi_{\zeta(\alpha_1,\beta_1,\alpha_2,\beta_2)}\rangle| \,\mathrm{d}\mu(z) \frac{\mathrm{d}\alpha_1 \mathrm{d}\beta_1 \mathrm{d}\alpha_2 \mathrm{d}\beta_2}{\beta_1\beta_2}$$
$$\lesssim \int\limits_{z\in Z^{\mathbf{d}}} |\langle f,\psi_z\rangle|^2 \,\mathrm{d}\mu(z)$$

as well as three more integrals covering all possible relationships between the scales of $z = ((x_1,s_1),(x_2,s_2))$ and $\zeta(\alpha_1,\beta_1,\alpha_2,\beta_2) = ((x_1+\alpha_1 s_1, \beta_1 s_1),(x_2+\alpha_2 s_2,\beta_2 s_2))$, which are estimated in an analogous fashion. In this specific case, Lemma 3.1.1 entails the bound

$$|\langle s_1^{d_1} s_2^{d_2}\psi_z, \psi_{\zeta(\alpha_1,\beta_1,\alpha_2,\beta_2)}\rangle| \lesssim [(\alpha_1,\beta_1)]_{\frac{\delta}{4}} [(\alpha_2,\beta_2)]_{\frac{\delta}{4}}$$

and the required control is again obtained via a combination of two instances of (2.1.4) and

Cauchy-Schwarz. □

The L^p-theory of double square functions is well studied. On the other hand, working with non-compactly supported, non-tensor product wavelets, is non-standard. In this generality, L^p-estimates may be obtained by direct product John-Nirenberg type arguments involving Journé's lemma. For reasons of space, L^p-bounds are obtained as a particular case of the sharp quantitative bound of Theorem 3.6.1 below.

3.3 A Technical Lemma

The following technical lemma shows that the norm (1.2.2) controls certain symbols obtained by partial testing of Λ against monomials. For this, the modified wavelets

$$\chi_{z,\zeta} = \varphi_z - P_{z,\zeta}\mathbf{1}_{A(\zeta)}(z), \qquad z,\zeta \in Z^d$$

introduced in (2.5.1) will be needed.

Lemma 3.3.1. Let $\iota \in \{1,2\}$, γ_ι be a multi-index in $\mathbb{R}^{d_\iota}$ and $k = (k_1,k_2)$ be such that $|\gamma_\iota| \le k_\iota$. For $\iota \in \{1,2\}$, $z_{\hat\iota},\zeta_{\hat\iota} \in Z^{d_{\hat\iota}}$, multi-indices γ_ι in $\mathbb{R}^{d_\iota}$, and $\mathfrak{a} \in \{\circ,\star\}$ define the functionals $\mathbf{q}^{\iota,\mathfrak{a}}_{\gamma_\iota}[\Lambda]$ by

$$\begin{aligned}\left\langle \mathbf{q}^{1,\mathfrak{a}}_{\gamma_1}[\Lambda](z_2,\zeta_2), f_1\right\rangle &:= \Lambda^{\mathfrak{a}}(x_1^{\gamma_1}\otimes\chi_{z_2,\zeta_2}, |\nabla|^{|\gamma_1|}f_1\otimes\chi_{\zeta_2,z_2}),\\ \left\langle \mathbf{q}^{2,\mathfrak{a}}_{\gamma_2}[\Lambda](z_1,\zeta_1), f_2\right\rangle &:= \Lambda^{\mathfrak{a}}(\chi_{z_1,\zeta_1}\otimes x_2^{\gamma_2}, \chi_{\zeta_1,z_1}\otimes|\nabla|^{|\gamma_2|}f_2)\end{aligned} \tag{3.3.1}$$

initially acting on the subspace $\mathcal{S}(\mathbb{R}^{d_\iota})$ of functions f_ι whose frequency support does not contain the origin. For a multi-index α_ι in $\mathbb{R}^{d_\iota}$, and $0<\eta<\delta$, also define

$$\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}[\Lambda](z_{\hat\iota},\zeta_{\hat\iota}) := [z_{\hat\iota},\zeta_{\hat\iota}]^{-1}_{k_{\hat\iota}+\eta} R^{\alpha_\iota}\mathbf{q}^{\iota,\mathfrak{a}}_{\gamma_\iota}(z_{\hat\iota},\zeta_{\hat\iota}), \tag{3.3.2}$$

where R^{α_1} is the Riesz transform associated to α_1. Assume that $\mathbf{q}^{\iota,\mathfrak{a}}_{\gamma'_\iota} \equiv 0$ for all multi-indices on $\mathbb{R}^{d_\iota}$ with $|\gamma'_\iota| < |\gamma_\iota|$. Then

$$\sup_{Z^{d_{\hat\iota}} \times Z^{d_{\hat\iota}}} \|\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}[\Lambda]\|_{\mathrm{BMO}(\mathbb{R}^{d_{\hat\iota}})} \lesssim \|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)}\,.$$

3.3.1 Proof of Lemma 3.3.1

For the sake of definiteness, work in the completely generic case $\iota = 1, \mathfrak{a} = \circ$. Fix $z_2, \zeta_2 \in Z^{d_2}$, and consider without loss of generality the case $z_2 \in Z^{d_2}_+(\zeta_2)$. Invoking the $\mathrm{BMO}(\mathbb{R}^{d_1})$ boundedness of the Riesz transform R^{γ_1}, we realize it must be shown that

$$\|b\|_{\mathrm{BMO}(\mathbb{R}^{d_1})} \lesssim [z_2,\zeta_2]_{k_2+\eta} \qquad b \text{ defined by } \langle b, g_1\rangle = \Lambda(x_1^{\gamma_1} \otimes \chi_{z_2,\zeta_2}, |\nabla|^{k_1} g_1 \otimes \varphi_{\zeta_2}).$$

Let $M = 2^8(1+k_1)$. By $H^1-\mathrm{BMO}$ duality and H^1-density of the latter class of functions, it will be enough to show that whenever $w_1 = (y_1, t_1) \in Z^{d_1}$, $\psi \in \mathcal{S}(\mathbb{R}^{d_1})$ is a Schwartz function such that $\Psi := \mathsf{Sy}^{-1}_{w_1}\psi$ satisfies

$$\|\Psi\|_{\star,4M,1} \leq 1, \qquad \operatorname{supp}\widehat{\Psi} \subset \{y \in \mathbb{R}^{d_j} : 1 \leq |y| \leq 2\}$$

there holds

$$|\langle b, \psi\rangle| \lesssim [z_2,\zeta_2]_{k_2+\eta}. \tag{3.3.3}$$

The frequency support of ψ ensures that $\upsilon := t_1^{k_1}|\nabla|^{k_1}\psi \in \Psi^{3M,1;0}_{w_1}$. Now introduce the local notation, with reference to (2.1.2)

$$p(\cdot) = \left(\frac{\cdot - y_1}{t_1}\right)^{\gamma_1}, \quad \Theta_2 := \chi_{z_2,\zeta_2}\alpha_{\zeta_2}, \quad \Xi_2 := \chi_{z_2,\zeta_2}\beta_{\zeta_2}.$$

There is an additional technical complication brought by the fact that υ is not of compact support. This is dealt with it by introduction of a sequence of smooth functions $q_n \in \mathcal{C}^\infty(\mathbb{R}^{d_1})$ with $\sum_{n=0}^\infty q_n = 1$, $\operatorname{supp} q_0 \subset \mathsf{B}_{(0,t_1)}$, $\operatorname{supp} q_n \subset A_n := \mathsf{B}_{(0,2^{n+1}t_1)} \setminus \mathsf{B}_{(0,2^{n-1}t_1)}$ for $n \geq 0$. Define $p_n := pq_n$, and $\upsilon_n := \upsilon q_n$. With these notations, the definition of b, and the fact that $\mathbf{q}_{\gamma_1'} \equiv 0$ for all $|\gamma_1'| < |\gamma_1|$,

$$\begin{aligned}\langle b, \psi\rangle &= \Lambda(p \otimes \chi_{z_2,\zeta_2}, \upsilon \otimes \varphi_{\zeta_2}) \\ &= \sum_{m\sim n} + \sum_{m\not\sim n} \Lambda(p_n \otimes \chi_{z_2,\zeta_2}, \upsilon_m \otimes \varphi_{\zeta_2}) + \Lambda(p_n \otimes \chi_{z_2,\zeta_2}, \upsilon_m \otimes \varphi_{\zeta_2})\end{aligned} \tag{3.3.4}$$

where $m \sim n$ if $|m-n| < 2^4$ and $m \not\sim n$ otherwise. The next task consists of bound each summand in the last right hand side of (3.3.4).

The $m \sim n$ summand

Notice that in this range $\|p_n\|_\infty \lesssim 2^{k_1 n}$, $\|\upsilon_m\|_\infty \lesssim t_1^{-d_1} 2^{-3Mm} \lesssim t_1^{-d_1} 2^{-2Mn}$ and p_n, υ_m are supported on $\mathsf{B}_{(y_1, 2^{n+2^5} t_1)}$. Also note that Θ_2 is supported on $\mathsf{B}_{(\xi_2, 4\sigma_2)}$ and $\|\Theta_2\|_\infty \lesssim [z_2, \zeta_2]_{k_2+\eta}$; this is obvious if $\chi_{z_2,\zeta_2} = \varphi_{z_2}$ and may be read from (2.1.17) otherwise. Applying the weak boundedness property of Λ with balls $\mathsf{B}_{(y_1, 2^{n+2^5} t_1)}$ and $\mathsf{B}_{(\xi_2, 4\sigma_2)}$,

$$|\Lambda(p_n \otimes \Theta_2, \upsilon_m \otimes \varphi_{\zeta_2})| \lesssim \|p_n\|_\infty \|\upsilon_m\|_\infty \lesssim 2^{-Mn} [z_2, \zeta_2]_{k_2+\eta}.$$

Furthermore, Ξ_2 and φ_{z_2} have separated supports. Therefore, using the partial kernel assumptions for the form $(f, g) \mapsto \Lambda(\Theta_1 \otimes f, \upsilon_{w_1} \otimes g)$ and repeating the computations of (2.5.7) for such form

$$|\Lambda(p_n \otimes \Xi_2, \upsilon_m \otimes \varphi_{\zeta_2})| \lesssim 2^{-Mn} [z_2, \zeta_2]_{k_2+\eta}.$$

The last two estimates are summable on the diagonal $m \sim n$ and this completes the bound for the $\sim$ summand in (3.3.4).

The $m \not\sim n$ summand

We now have that p_n and υ_m have separated supports by $\sim t_1 2^{\max\{m,n\}}$. Applying the partial kernel assumptions for the form $(f,g) \mapsto \Lambda(f \otimes \Theta_2, g \otimes \varphi_{\zeta_2})$ and arguing as in (2.5.7) yields

$$|\Lambda(p_n \otimes \Theta_2, \upsilon_m \otimes \varphi_{\zeta_2})| \lesssim \|p_n\|_\infty 2^{-\max\{m,n\}(k_1+\delta)}[z_2,\zeta_2]_{k_2+\eta} \lesssim 2^{-\max\{m,n\}\delta}.$$

Finally, in the term below, the full kernel assumptions may be used due to functions in both parameters having disjoint supports. Standard computations relying on the kernel estimates as in (2.5.7) then lead to

$$|\Lambda(p_n \otimes \Xi_2, \upsilon_m \otimes \varphi_{\zeta_2})| \lesssim 2^{-\max\{m,n\}\delta}[z_2,\zeta_2]_{k_2+\eta}.$$

The above estimates are summable over $m \not\sim n$, which completes both this case and the proof of the Lemma.

We explain the details of the definitions (3.3.1). Using symmetry with respect to adjoints, it suffices to study $\mathbf{q}^{1,\circ}_{\gamma_1}(z_2,\zeta_2)$. The most complicated case is when either one of $\chi_{z_2,\zeta_2}, \chi_{\zeta_2,z_2}$ contains the polynomial summand. To fix ideas, work with $\chi_{z_2,\zeta_2} = \varphi_{z_2} - P_{z_2,\zeta_2}$. Let $\phi_2 = \phi$ as in (2.4.4) with $d = d_2$ and $p^{\gamma_1}_R$ as in (2.4.5) with $d = d_1$ and $\alpha = \gamma_1$, see Subsection 2.4. If $0 \notin \operatorname{supp} \widehat{f_1}$ then $g_1 = |\nabla|^{|\gamma_1|} f_1 \in \mathcal{S}_D(\mathbb{R}^{d_1})$. The partial kernel estimates of Λ readily show that

$$\langle \mathbf{q}^{1,\circ}_{\gamma_1}(z_2,\zeta_2), f_1\rangle = \lim_{R\to\infty} \Lambda(p^{\gamma_1}_R \otimes [\mathsf{Dil}^\infty_R \phi_2]\chi_{z_2,\zeta_2}, g_1 \otimes \varphi_{\zeta_2}),$$

exists and defines a linear continuous functional on the subspace of $\mathcal{S}(\mathbb{R}^{d_1})$ of functions supported off the frequency origin.

3.4 Bi-Parameter Calderón-Zygmund forms of class (k, δ)

Given the developed framework and notation we refer back to Definition 1.2.1 and Definition 1.2.2 for $\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$, the space of bi-parameter Calderón-Zygmund Forms of Class (k, δ). The following are examples of such forms and will be model operators for the analysis of general forms.

Definition 3.4.1 (Bi-parameter wavelet form)**.** Associate to the two families of bi-parameter cancellative wavelets $\left\{\beta_z, \upsilon_z \in C\Psi_z^{k,\delta;0,0} : z \in Z^{\mathbf{d}}\right\}$ the bi-parameter wavelet form

$$\Lambda(f, g) = \int_{Z^{\mathbf{d}}} \langle f, \beta_z \rangle \langle \upsilon_z, g \rangle \, \mathrm{d}\mu(z). \tag{3.4.1}$$

This form belongs to $\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ and $\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)} \lesssim 1$. The weak boundedness property is a particular case of estimate (3.2.3) from Proposition 3.2.1. The partial kernel and full kernel estimates may be obtained by repeatedly employing (3.1.2) and standard computations. If $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ is a wavelet form of the type (3.4.1), an immediate byproduct of the cancellation properties of the families $\beta_z, \upsilon_z \in C\Psi_z^{k,\delta;0,0}$ is that the functionals

$$\Lambda^{\mathfrak{a}}(x^{\alpha_1} \otimes x^{\alpha_2}, \cdot), \qquad \mathbf{a}^{\iota;\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}[\Lambda]$$

vanish for all $0 \leq |\alpha_\iota|, |\gamma_\iota| \leq k_\iota$, $\iota = 1, 2$.

Definition 3.4.2 (Bi-parameter paraproduct forms)**.** Paralleling our treatment of the one parameter case, three types of bi-parameter paraproducts will be defined. For a pair of multi-indices (γ_1, γ_2) on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, a function $b \in \mathrm{BMO}(\mathbb{R}^{\mathbf{d}})$, $(\iota_1, \iota_2) \in \{0, 1\}^2$ and $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$, define d

$$\begin{aligned}\Pi_{(0,0),b,(\gamma_1,\gamma_2)}(f,g) &\coloneqq \int\limits_{Z^{\mathbf{d}}} \langle b, \varphi_{\gamma_1,z_1}\otimes\varphi_{\gamma_2,z_2}\rangle\langle f, \vartheta_{\gamma_1,z_1}\otimes\vartheta_{\gamma_2,z_2}\rangle\langle\varphi_z, g\rangle\,\mathrm{d}\mu(z),\\ \Pi_{(0,1),b,(\gamma_1,\gamma_2)}(f,g) &\coloneqq \int\limits_{Z^{\mathbf{d}}} \langle b, \varphi_{\gamma_1,z_1}\otimes\varphi_{\gamma_2,z_2}\rangle\langle f, \vartheta_{\gamma_1,z_1}\otimes\varphi_{z_2}\rangle\langle\varphi_{z_1}\otimes\vartheta_{\gamma_2,z_2}, g\rangle\,\mathrm{d}\mu(z),\end{aligned} \tag{3.4.2}$$

where $\varphi_{\gamma_j,z_j} = \mathsf{Sy}_{z_j}\partial^{-\gamma_j}\varphi_j$, see (2.1.8) and (2.1.15), and the family $\{\vartheta_{\gamma_j,z_j} \in C\Psi_z^{D,1;0} : z \in Z^{d_j}\}$ is a γ_j-family as in (2.4.2). The first form is usually termed *full paraproduct*, while the second is usually referred to as a *partial paraproduct*. The paraproducts defined below are related as partial adjoints, namely

$$\Pi_{(0,1),b,(\gamma_1,\gamma_2)} = \left(\Pi_{(0,1),b,(\gamma_1,\gamma_2)}\right)^{\star_1}.$$

For this reason, the notation $\Pi_{b,\gamma}$ will be used in place of $\Pi_{(0,0),b,\gamma}$ throughout this section. Standard computations relying on the smoothness and compact support of the wavelets appearing in (3.4.2) lead to the following controls on the partial kernel and weak boundedness, and full kernel constants of the paraproduct forms: for any multi-index $\gamma = (\gamma_1,\gamma_2)$, there holds

$$\|\Pi_{b,\gamma}\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} \lesssim_k \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})}. \tag{3.4.3}$$

A third family of paraproducts, which are termed *half-paraproducts*, are constructed using the definitions (2.4.3) in each parameter $\iota \in \{1,2\}$. Let $\kappa_{\hat{\iota}} \in \mathbb{N}, \eta > 0$, which are kept implicit in the notation, γ_ι be a multi-index on $\mathbb{R}^{d_\iota}$ and $\mathbf{a}$ be a continuous map on $Z^{d_{\hat{\iota}}} \times Z^{d_{\hat{\iota}}}$ taking values in $\mathrm{BMO}(\mathbb{R}^{d_\iota})$. Define, a priori on $\mathcal{S}(\mathbb{R}^{\mathbf{d}}) \times \mathcal{S}(\mathbb{R}^{\mathbf{d}})$, the form

$$\Pi^\iota_{\mathbf{a},\gamma_\iota,\alpha_\iota}(f,g) = \int\limits_{Z^{d_{\hat{\iota}}}\times Z^{d_{\hat{\iota}}}} \Pi_{\mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}}),\gamma_\iota,\alpha_\iota}\left(\langle f,\varphi_{z_{\hat{\iota}}}\rangle_{\hat{\iota}}, \langle g,\varphi_{\zeta_{\hat{\iota}}}\rangle_{\hat{\iota}}\right)[z_{\hat{\iota}},\zeta_{\hat{\iota}}]_{\kappa_{\hat{\iota}}+\eta}\mathrm{d}\mu(z_{\hat{\iota}})\mathrm{d}\mu(\zeta_{\hat{\iota}}) \tag{3.4.4}$$

where $\Pi_{\mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}}),\gamma_\iota,\alpha_\iota}$ refers to (2.4.3) for $b = \mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}}), \gamma = \gamma_\iota, \alpha = \alpha_\iota, d = d_\iota$. Arguing in a similar fashion to (3.4.3), we record the estimate

$$\left\|\Pi^\iota_{\mathbf{a},\gamma_\iota,\alpha_\iota}\right\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} \lesssim \sup_{(z_{\hat{\iota}},\zeta_{\hat{\iota}})\in Z^{d_{\hat{\iota}}}\times Z^{d_{\hat{\iota}}}} \|\mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}})\|_{\mathrm{BMO}(\mathbb{R}^{d_\iota})}, \qquad k=(k_1,k_2),\ k_\iota = |\gamma_\iota|,\ k_{\hat{\iota}} = \kappa_{\hat{\iota}}. \tag{3.4.5}$$

Furthermore, it is a particularly useful observation that, in view of (2.4.12) and referring to the notation introduced in (2.4.6), which is legitimately used whenever $f_2, g_2 \in \mathcal{S}(\mathbb{R}^{d_2})$ are fixed

$$\Pi^1_{\mathbf{a},\gamma_1,\alpha_1}(x_1^{\beta_1}\otimes f_2, g_1\otimes g_2) = 0 \qquad \forall 0\le |\beta_1| \le |\gamma_1|,\ \beta_1\neq\gamma_1 \tag{3.4.6}$$

in the sense of linear functionals acting on $g_1 \in \mathcal{S}_D(\mathbb{R}^{d_1})$, and similarly for adjoints and half-paraproducts in the second parameter.

If α_ι is a multi-index on $\mathbb{R}^{d_\iota}$, the notation $p_R^{\alpha_\iota}$ refers to p_R^{α} from (2.4.5), with $d=d_\iota, \alpha=\alpha_\iota$. Below, $\mathcal{S}_D(\mathbb{R}^{\mathbf{d}})$ be the subspace of functions $\psi\in\mathcal{S}(\mathbb{R}^d)$ with the same bi-parameter vanishing moment property of the functions of $\Psi_z^{D,\delta;0,0}$, for some $z \in Z^{\mathbf{d}}$. Then, if $0 \le |\alpha_\iota| \le k_\iota$ for $\iota\in\{1,2\}$, and $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)$, and $\mathfrak{a}\in\vec{\mathfrak{a}}$, the limits

$$\Lambda^{\mathfrak{a}}(x^{\alpha_1}\otimes x^{\alpha_2},\psi) = \lim_{R\to\infty}\Lambda\left(p_R^{\alpha_1}\otimes p_R^{\alpha_2},\psi\right), \qquad \psi\in\mathcal{S}_D(\mathbb{R}^{\mathbf{d}}) \tag{3.4.7}$$

exist and define linear continuous functionals on $\mathcal{S}_D(\mathbb{R}^d)$: with the full kernel estimates at one's disposal, the proof presented in [22, Lemma 1.91] extends to the bi-parameter case without essential changes.

The next and final definition details our assumptions on the functionals (3.4.7) associated to $\Lambda\in\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)$. We ask whether $\Lambda\in\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)$ has paraproducts of order κ for $0\le\kappa\le\min\{k_1,k_2\}$ and, if that is the case, at the same time define the κ-th order cancellative part of Λ for all $0\le\kappa\le\max\{k_1,k_2\}$.

Definition 3.4.3 (Paraproducts of order $0 \le \kappa \le \min\{k_1, k_2\}$)**.** Say that $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ has *paraproducts of order* 0 if for each $\mathfrak{a} \in \vec{\mathfrak{a}}$ there exists $b_0^{\mathfrak{a}} \in \mathrm{BMO}(\mathbb{R}^{\mathbf{d}})$, the BMO product space, such that

$$\Lambda^{\mathfrak{a}}(\mathbf{1} \otimes \mathbf{1}, \psi) = \langle b_0^{\mathfrak{a}}, \psi \rangle \qquad \forall \psi \in \mathcal{S}_D(\mathbb{R}^{\mathbf{d}}).$$

As customary, we use the $T(1)$ notation and write $b_0^{\mathfrak{a}} = T^{\mathfrak{a}}(\mathbf{1} \otimes \mathbf{1})$. If Λ has paraproducts of order 0 the 0-th order cancellative part of Λ is given by

$$\Lambda_0(f,g) \coloneqq \Lambda(f,g) - \sum_{\mathfrak{a} \in \vec{\mathfrak{a}}} \left[\Pi_{b_0^{\mathfrak{a}}}\right]^{\mathfrak{a}}(f,g) - \sum_{\substack{\mathfrak{a} \in \{\circ, \star\} \\ \iota \in \{1,2\}}} \left[\Pi^{\iota}_{\mathbf{a}^{\iota,\mathfrak{a}}_{0,0}[\Lambda]}\right]^{\mathfrak{a}}(f,g).$$

Suppose Λ has paraproducts of order $0 \le \kappa < \min\{k_1, k_2\}$ and the κ-th order cancellative part of Λ has been defined. Then Λ has *paraproducts of* $(\kappa{+}1)$*-th order* if for each $\gamma = (\gamma_1, \gamma_2)$ with $|\gamma_1| = \kappa + 1 = |\gamma_2|$ and $\mathfrak{a} \in \vec{\mathfrak{a}}$ there exists $b_\gamma^{\mathfrak{a}} \in \mathrm{BMO}(\mathbb{R}^{\mathbf{d}})$ such that

$$[\Lambda_\kappa]^{\mathfrak{a}}\left(x_1^{\gamma_1} \otimes x_2^{\gamma_2}, \psi\right) = \left\langle b_\gamma^{\mathfrak{a}}, \partial^{-\gamma_1}\partial^{-\gamma_2}\psi \right\rangle \qquad \forall \psi \in \mathcal{S}_D(\mathbb{R}^{\mathbf{d}}).$$

If $T_\kappa^{\mathfrak{a}}$ stand for the adjoints to Λ_κ, the corresponding $T(1)$ notation is then

$$b_\gamma^{\mathfrak{a}} = R^{\gamma_1}_{\mathbb{R}^{d_1}} R^{\gamma_2}_{\mathbb{R}^{d_2}} |\nabla_{\mathbb{R}^{d_1}}|^{|\gamma_1|} |\nabla_{\mathbb{R}^{d_2}}|^{|\gamma_2|} T_\kappa^{\mathfrak{a}}(x_1^{\gamma_1} \otimes x_2^{\gamma_2}), \qquad \mathfrak{a} \in \vec{\mathfrak{a}}.$$

The inductive definition is closed by defining the $\kappa + 1$*-th order cancellative part of* Λ as

$$\Lambda_{\kappa+1}(f,g) = \Lambda_\kappa(f,g) - \sum_{\substack{\gamma = (\gamma_1,\gamma_2) \\ |\gamma_1| = |\gamma_2| = \kappa+1}} \sum_{\mathfrak{a} \in \vec{\mathfrak{a}}} \left[\Pi_{b^{\mathfrak{a}}_\gamma, \gamma}\right]^{\mathfrak{a}}(f,g) - \sum_{\substack{\mathfrak{a} \in \{\circ,\star\} \\ \iota \in \{1,2\} \\ |\gamma_\iota| = |\alpha_\iota| = \kappa+1}} \left[\Pi^{\iota}_{\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}[\Lambda_\kappa]}\right]^{\mathfrak{a}}(f,g). \quad (3.4.8)$$

We do not define paraproducts of order $\kappa{+}1$ for $\min\{k_1, k_2\} \le \kappa \le \max\{k_1, k_2\}{-}1$. However, we define the $(\kappa+1)$-th order cancellative part of Λ, also inductively, by

$$\Lambda_{\kappa+1}(f,g) = \Lambda_\kappa(f,g) - \sum_{\substack{\mathfrak{a}\in\{\circ,\star\} \\ |\gamma_{\iota^*}|=|\alpha_{\iota^*}|=\kappa+1}} \left[\Pi^{\iota^*}_{\mathbf{a}^{\iota^*,\mathfrak{a}}_{\gamma_{\iota^*},\alpha_{\iota^*}}[\Lambda_\kappa]}\right]^{\mathfrak{a}}(f,g), \qquad \iota^* = \arg\max\{k_\iota\}. \quad (3.4.9)$$

Definition 3.4.4. The bi-parameter bilinear form Λ belongs to the class $\mathrm{CZ}(\mathbb{R}^{\mathbf{d}},k,\delta)$ of (k,δ)-*Calderón-Zygmund forms* if $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)$ and Λ has paraproducts of order $\min\{k_1,k_2\}$, with norm

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^{\mathbf{d}},k,\delta)} := \|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} + \sum_{\substack{0\le\kappa\le\min\{k_1,k_2\} \\ |\gamma_1|=|\gamma_2|=\kappa \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left\| |\nabla_{\mathbb{R}^{d_1}}|^{|\gamma_1|} |\nabla_{\mathbb{R}^{d_2}}|^{|\gamma_2|} T^{\mathfrak{a}}_\kappa(x_1^{\gamma_1}\otimes x_2^{\gamma_2}) \right\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})}. \quad (3.4.10)$$

Forms in the $\mathrm{CZ}(\mathbb{R}^{\mathbf{d}},k,\delta)$ may be represented as a linear combination of a wavelet form of type (3.4.1) and order k plus a finite linear combination of paraproducts and half-paraproducts. These definitions and notations now allow us to precisely state the main result of this chapter.

Theorem 3.4.1. Let $k=(k_1,k_2)\in\mathbb{N}^2$ with $\max\{k_1,k_2\}+1\le D$, $0<\varepsilon<\delta\le 1$. There exists an absolute constant $C=C_{k,\delta,\epsilon,d}$ such that the following holds. Let Λ be a form of class $\mathrm{CZ}(\mathbb{R}^{\mathbf{d}},k,\delta)$ with normalization $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^{\mathbf{d}}),k,\delta}\le 1$. Then, there exists a family

$$\left\{\upsilon_z \in C\Psi^{k,\varepsilon;0,0}_z : z\in Z^{\mathbf{d}}\right\},$$

and functions $\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}$ on $Z^{d_{\hat\iota}}\times Z^{d_{\hat\iota}}$ taking values in a bounded subset of $\mathrm{BMO}(\mathbb{R}^{d_\iota})$, $\iota=1,2$

such that for all $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$

$$\Lambda(f,g) = \int_{Z^{\mathbf{d}}} \langle f, \varphi_z\rangle\langle \upsilon_z, g\rangle \,\mathrm{d}\mu(z) \\ + \sum_{\substack{\gamma=(\gamma_1,\gamma_2)\\ 0\le|\gamma_1|=|\gamma_2|\le\min\{k_1,k_2\}\\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b^{\mathfrak{a}}_{\gamma},\gamma}\right]^{\mathfrak{a}}(f,g) + \sum_{\substack{\iota\in\{1,2\}\\ 0\le|\gamma_\iota|\le k_\iota\\ |\alpha_\iota|=|\gamma_\iota|\\ \mathfrak{a}\in\{\circ,\star\}}} \left[\Pi^{\iota}_{\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}}\right]^{\mathfrak{a}}(f,g). \tag{3.4.11}$$

The next corollary to Theorem 3.4.1 is easily proved by combining with (3.4.11) the estimates of Propositions 3.6.1, 3.6.2 and 3.6.3, and, for the cases $k \neq (0,0)$, an integration by parts argument akin to the one used in the proof of Corollary 2.6.1.

Corollary 3.4.1. Let $k \in \mathbb{N}^2, \delta > 0$ and $\Lambda \in \mathrm{CZ}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ be as in Theorem 3.4.1. Assume in addition the bi-parameter analogue of (2.6.1)

$$\begin{aligned} &(|\gamma_1|,|\gamma_2|) \neq k \implies b^{\circ}_{\gamma} = 0;\\ &j\in\{1,2\},\ |\gamma_j| < k_j \implies b^{\star_j}_{\gamma} = 0,\ \mathbf{a}^{j,\circ}_{\gamma_j,\alpha_j} = 0 \quad \forall|\alpha_j| = |\gamma_j|. \end{aligned} \tag{3.4.12}$$

Let $1 < p < \infty$ and w be a product A_p-weight in $\mathbb{R}^{\mathbf{d}}$. Then, if T stands for the adjoint to Λ,

$$\|\nabla^{k_1}_{x_1}\nabla^{k_2}_{x_2} Tf\|_{L^p(\mathbb{R}^d;w)} \lesssim_{k,\delta} [w]_{A_p}^{\max\left\{3,\frac{2p}{p-1}\right\}} \|\nabla^{k_1}_{x_1}\nabla^{k_2}_{x_2} f\|_{L^p(\mathbb{R}^d;w)}. \tag{3.4.13}$$

Notice that (3.4.12) is not an additional assumption in the case $k = (0,0)$ and necessary for (3.4.13) to hold otherwise. We point out that among the bounds provided in these propositions, the exponent in (3.4.13) is achieved by the paraproduct estimate of Proposition 3.6.2 and their adjoints. In fact, if Λ is a $((0,0),\delta)$ Calderón-Zygmund form whose paraproduct terms appearing in (3.4.11) all vanish, then the weighted norm bound for its adjoint

$$\|T\|_{L^p(\mathbb{R}^d;w)} \lesssim [w]_{A_p}^{\theta(p)}, \qquad \theta(p) = \begin{cases} \frac{2}{p-1} & 1 < p \leq \frac{3}{2} \\ \frac{2p-1}{p-1} & \frac{3}{2} < p \leq 2 \\ \frac{p+1}{p-1} & 2 < p \leq 3 \\ 2 & p > 3 \end{cases} \tag{3.4.14}$$

may be read by applying Proposition 3.6.1 to the cancellative terms in (3.4.11) for Λ, if $p \geq 2$, or for its full adjoint $\Lambda^\star$ if $1 < p < 2$. Comparing with the one parameter case, see [34, Theorem 2], estimate (3.4.14) is sharp for $\max\{p, p'\} \geq 3$: there seem to be no instances of sharp weighted norm inequalities for bi-parameter operators in previous literature. Notice that the paraproduct free assumption covers, for instance, bi-parameter convolution-type operators.

3.5 Proof of Bi-Parameter Representation Theorem

Before entering the main argument, a series of preparatory lemmas is required that generalize results from the one parameter case. Below $0 < \delta \leq 1$, $0 < \varepsilon < \delta$ are fixed. Set $\eta = \frac{\delta+\varepsilon}{2}$, so that $\varepsilon < \eta < \delta$. First, the two parameter version of Lemma 2.3.1 is provided.

Lemma 3.5.1. Let φ_z be as in (1.2.1) and $u : Z^{\mathbf{d}} \to \mathbb{C}$ be a Borel measurable function with $|u(z)| \leq 1$. Then, there exists $C \lesssim_{k_1,k_2,\varepsilon} 1$ such that for all $z = (z_1, z_2) \in Z^{\mathbf{d}}$ with $z_j = (x_j, s_j)$, $j = 1, 2$

$$\upsilon_z := \int_{\substack{(\alpha_1,\beta_1)\in Z^{d_1} \\ (\alpha_2,\beta_2)\in Z^{d_2}}} \left(\prod_{j=1}^{2} [(\alpha_j, \beta_j)]_{k_j+\eta}\right) u((\alpha_1, \beta_1), (\alpha_2, \beta_2))\varphi_{((x_1+\alpha_1 s_1,\beta_1 s_1),(x_2+\alpha_2 s_2,\beta_2 s_2))} \frac{\mathrm{d}\beta_2 \mathrm{d}\alpha_2 \mathrm{d}\beta_1 \mathrm{d}\alpha_1}{\beta_1\beta_2}$$

belongs to $C\Psi_z^{(k_1,k_2),\varepsilon;0,0}$.

Proof. There is a direct argument along the lines of the one parameter proof. However, an argument that uses Lemma 2.3.1 as a black box will be given. To save space in the notation, this argument is carried out for $x_j = 0, s_j = 1,\ j = 1,2$. Notice that for each fixed $w_1 \in \mathbb{R}^{d_1}$

$$\upsilon_z^{[1,w_1]} = \int\limits_{(\alpha_2,\beta_2)\in Z^{d_2}} [(\alpha_2,\beta_2)]_{k_2+\eta}\upsilon_{\alpha_2,\beta_2}(w_1)(\varphi_2)_{(x_2+\alpha_2 s_2,\beta_2 s_2)}\frac{\mathrm{d}\beta_2\mathrm{d}\alpha_2}{\beta_2},$$

where

$$\upsilon_{\alpha_2,\beta_2} = \int\limits_{(\alpha_1,\beta_1)\in Z^{d_1}} [(\alpha_1,\beta_1)]_{k_1+\eta}u((\alpha_1,\beta_1),(\alpha_2,\beta_2))(\varphi_1)_{(x_1+\alpha_1 s_1,\beta_1 s_1)}\frac{\mathrm{d}\beta_1\mathrm{d}\alpha_1}{\beta_1} \in C\Psi_{z_1}^{k_1,\varepsilon;0}$$

with uniform constant C by an application of (2.3.3) of Lemma 2.3.1 with η in place of δ. In particular the function $u_{w_1}(\alpha_2,\beta_2) := \langle w_1\rangle^{(d_1+k_1+\varepsilon)}\upsilon_{\alpha_2,\beta_2}(w_1)$ is uniformly bounded. Therefore, another application of (2.3.3), with $u = u_{w_1}(\alpha_2,\beta_2)$ entails

$$\langle w_1\rangle^{(d_1+k_1+\varepsilon)}\upsilon_z^{[1,w_1]} = \int\limits_{(\alpha_2,\beta_2)\in Z^{d_2}} [(\alpha_2,\beta_2)]_{k_2+\eta}u_{w_1}(\alpha_2,\beta_2)(\varphi_2)_{(x_2+\alpha_2 s_2,\beta_2 s_2)}\frac{\mathrm{d}\beta_2\mathrm{d}\alpha_2}{\beta_2} \in C\Psi_{z_2}^{k_2,\varepsilon;0}.$$

Repeating for the second parameter and comparing with equation (3.1.2), proves that $\upsilon_z \in C\Psi_z^{(k_1,k_2),\varepsilon;0,0}$ and thus completes the proof of the lemma. □

Along the same lines of the previous lemma, the next is the bi-parameter analogue for Lemma 2.5.1. The notation $\chi_{z_\iota,\zeta_\iota}$ appearing below refers to (2.5.1).

Lemma 3.5.2. For $(z,\zeta) \in Z^{\mathbf{d}} \times Z^{\mathbf{d}}$ with $z = (z_1,z_2)$, $\zeta = (\zeta_1,\zeta_2)$,

$$|\Lambda\left(\chi_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2},\chi_{\zeta_1,z_1}\otimes\chi_{\zeta_2,z_2}\right)| \lesssim \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^{\mathbf{d}}),k,\delta}\prod_{j=1,2}[z_j,\zeta_j]_{k_j+\eta}.$$

Proof. By symmetry with respect to the adjoint, it suffices to consider the case where $z_j \in Z_+^{d_j}(\zeta_j)$ for both $j = 1, 2$. Nine different cases according to which of the sets in (2.1.1) with $\zeta = \zeta_j$ each z_j belongs to need to be considered. Only the case $z_j \in A(\zeta_j)$ for $j = 1, 2$ will be dealt with explicitly; all remaining cases may be dealt with by the same strategy that will be used for the summands appearing in (3.5.1). In this case, $\chi_{z_j,\zeta_j} = \varphi_{z_j} - P_{z_j,\zeta_j}$ and $\chi_{\zeta_j,z_j} = \varphi_{\zeta_j}$ for $j = 1, 2$, and thus

$$\begin{aligned} &\Lambda\left(\chi_{z_1,\zeta_1} \otimes \chi_{z_2,\zeta_2}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2}\right) = \sum_{(\iota,\jmath)\in\{\mathsf{in},\mathsf{out}^2\}} \Lambda\left(\Theta_{1,\iota} \otimes \Theta_{2,\jmath}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2}\right), \\ &\Theta_{j,\mathsf{in}} := \chi_{z_j,\zeta_j}\alpha_{\zeta_j}, \qquad \Theta_{j,\mathsf{out}} := \chi_{z_j,\zeta_j}\beta_{\zeta_j}, \qquad j = 1, 2. \end{aligned} \tag{3.5.1}$$

Each term in (3.5.1) will be estimated. The key to the first three summands is that for $j = 1, 2$ the function $\Theta_{j,\mathsf{in}}$ is supported on $4\mathsf{B}_{\zeta_j}$ and, from (2.1.17), $\|\Theta_{j,\mathsf{in}}\|_\infty \lesssim [z_j, \zeta_j]_{k_j+\eta}$. For the in, in summand, employ the weak boundedness of Λ with points $\tilde{\zeta}_j = (\xi_j, 4\sigma_j)$ thus obtaining

$$|\Lambda\left(\Theta_{1,\mathsf{in}} \otimes \Theta_{2,\mathsf{in}}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2}\right)| \lesssim \prod_{j=1,2} \|\Theta_{j,\mathsf{in}}\|_\infty \lesssim \prod_{j=1,2} [z_j, \zeta_j]_{k_j+\eta}.$$

The $\mathsf{in}, \mathsf{out}$ summand is bounded as follows. Observe that $\Theta_{2,\mathsf{out}}$ and φ_{ζ_2} have separated support. Now, apply the partial kernel/weak boundedness assumption to the form $(f, g) \mapsto \sigma_1^{d_1}\Lambda(\Theta_{1,\mathsf{in}} \otimes f, \varphi_{\zeta_1}, g)$ at point $\tilde{\zeta}_1 = (\xi_1, 4\sigma_1)$, to which estimate (2.5.7) with $z = z_2$ and $\zeta = \zeta_2$ actually applies. Such estimate returns a factor of $[z_2, \zeta_2]_{k_2+\eta}$ while the factor $[z_1, \zeta_1]_{k_1+\eta}$ is obtained from $\|\Theta_{1,\mathsf{in}}\|_\infty$. The $\mathsf{out}, \mathsf{in}$ summand is handled in exactly the same way.

In the $\mathsf{out}, \mathsf{out}$ summand, note that $\Theta_{j,\mathsf{out}}$ and φ_{ζ_j} both have separated support. The full kernel estimates of Λ are now employed. The calculation leading to the estimate

$$|\Lambda\left(\Theta_{1,\mathsf{out}} \otimes \Theta_{2,\mathsf{out}}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2}\right)| \lesssim \prod_{j=1,2} [z_j, \zeta_j]_{k_j+\eta}$$

is the tensor product of the steps (2.5.4)-(2.5.7) performed in each variable, the only difference being how the corresponding term in (2.5.4) involving the finite difference of the derivatives of the kernel is controlled. In that case, for a fixed $v = (v_1, v_2) \in (2\mathsf{B}_{\zeta_1} \times 2\mathsf{B}_{\zeta_2})^c$ one uses the cancellation and L^1-estimate of $\sigma_j^{-k_j}\partial^{-\gamma_j}\varphi_{\zeta_j}$ and bounds

$$\sup_{u \in \mathsf{B}_{\zeta_1} \times \mathsf{B}_{\zeta_2}} \left|\Delta^1_{u_1-\xi_1|\cdot}\Delta^2_{u_2-\xi_2|\cdot}\partial^{\gamma_1}_{u_1}\partial^{\gamma_1}_{u_1}K(\xi, v)\right| \lesssim \prod_{j=1}^{2} \frac{\sigma_j^{k_j+\delta}}{|v_j - \xi_j|^{d_j+k_j+\delta}}$$

using the kernel estimate in the fourth line (1.2.4). This completes the proof of the Lemma.

□

3.5.1 Main line of proof of Theorem 3.4.1

It is now possible to turn to the main line of proof of Theorem 3.4.1. Notice that

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,(\kappa_1,\kappa_2),\delta)} \leq \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d,k,\delta)} = 1, \qquad 0 \leq \kappa_j \leq k_j,\ j = 1, 2.$$

The proof will be done via two consecutive inductions. The first runs for $0 \leq \kappa \leq \min\{k_1, k_2\}$. The second, if necessary, runs for $\min\{k_1, k_2\} < \kappa \leq \max\{k_1, k_2\}$. The argument is symmetric with respect to interchanging parameters, therefore there is no loss in generality with assuming $k_1 \geq k_2$.

The base case and main part of the inductive step of the proof works under the additional assumption referring to Lemma 3.3.1

$$a(\kappa) : \begin{cases} b^{\mathfrak{a}}_{\gamma} = 0 & \forall \mathfrak{a} \in \vec{\mathfrak{a}},\ \min_{\iota \in \{1,2\}} |\gamma_\iota| < \kappa, \\ \mathbf{q}^{\iota,\mathfrak{a}}_{\gamma_\iota} = 0 & \forall \mathfrak{a} \in \{\circ, \star\}, |\gamma_\iota| < \min\{\kappa, k_\iota\}, \iota \in \{1, 2\}, \end{cases}$$

namely, all paraproducts $T^{\mathfrak{a}}(x^{\gamma_1} \otimes x^{\gamma_2})$ vanish except possibly those with $|\gamma_1| = |\gamma_2| = \kappa$, and all half-paraproducts vanish except possibly those of highest order. Clearly, $a(0)$ is not an extra assumption. Moreover, if $\kappa \geq 1$, assumption $a(\kappa)$ implies that Λ coincides with its $(\kappa-1)$-th order cancellative part $\Lambda_{\kappa-1}$, which means we are allowed to conflate the two forms and just write Λ below.

Let now $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$. Using the bi-parameter analogue of (2.1.11), bilinearity and $\mathcal{S}(\mathbb{R}^{\mathbf{d}})$-continuity of Λ, and later the definition of $U(z,\zeta)$ leads to the decomposition

$$\begin{aligned}
\Lambda(f,g) &= \int_{Z^{\mathbf{d}}\times Z^{\mathbf{d}}} \langle f,\varphi_z\rangle\langle\varphi_\zeta, g\rangle \Lambda(\varphi_z,\varphi_\zeta)\,\mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) \\
&= \int_{Z^{\mathbf{d}}\times Z^{\mathbf{d}}} \langle f,\varphi_z\rangle\langle\varphi_\zeta, g\rangle \Lambda\left(\chi_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2}, \chi_{\zeta_1,z_1}\otimes\chi_{\zeta_2,z_2}\right)\,\mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) \\
&+ \int_{Z^{d_1}_{\zeta_1}\times Z^{d_2}_{\zeta_2}} \int_{Z^{d_2}_{z_2}} \int_{z_1\in A(\zeta_1)} \langle f,\varphi_z\rangle\langle\varphi_\zeta, g\rangle \Lambda(P_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2}, \varphi_{\zeta_1}\otimes\chi_{\zeta_2,z_2})\,\mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) + \cdots \\
&+ \int_{Z^{d_1}_{\zeta_1}\times Z^{d_2}_{\zeta_2}} \int_{z_1\in A(\zeta_1)} \int_{z_2\in A(\zeta_2)} \langle f,\varphi_z\rangle\langle\varphi_\zeta, g\rangle \Lambda(P_{z_1,\zeta_1}\otimes P_{z_2,\zeta_2}, \varphi_{\zeta_1}\otimes\varphi_{\zeta_2})\,\mathrm{d}\mu(z)\mathrm{d}\mu(\zeta) + \cdots .
\end{aligned} \tag{3.5.2}$$

Here, the dots in the third line are hiding three more terms where the integration domain is respectively restricted to $z_2 \in A(\zeta_2)$, $\zeta_1 \in A(z_1)$, $\zeta_2 \in A(z_2)$, and the integrands involve respectively the coefficients

$$\Lambda(\chi_{z_1,\zeta_1}\otimes P_{z_2,\zeta_2}, \chi_{\zeta_1,z_1}\otimes\varphi_{\zeta_2}), \quad \Lambda(\varphi_{z_1}\otimes\chi_{z_2,\zeta_2}, P_{z_1,\zeta_1}\otimes\chi_{\zeta_2,z_2}), \quad \Lambda(\chi_{z_1,\zeta_1}\otimes,\varphi_{z_2}, \chi_{\zeta_1,z_1}\otimes P_{z_2,\zeta_2}),$$

while the dots in the fourth line also hide three more terms where the integration domain is restricted to $\{z_1 \in A(\zeta_1), \zeta_2 \in A(z_2)\}$, $\{\zeta_1 \in A(z_1), z_2 \in A(\zeta_2)\}$, $\{\zeta_1 \in A(z_1), \zeta_2 \in A(z_2)\}$ and the integrands involve respectively the coefficients

$$\Lambda(\varphi_{z_1}\otimes P_{z_2,\zeta_2}, P_{\zeta_1,z_1}\otimes\varphi_{\zeta_2}), \quad \Lambda(P_{z_1,\zeta_1}\otimes\varphi_{z_2}, \varphi_{\zeta_1}\otimes P_{\zeta_2,z_2}), \quad \Lambda(\varphi_{z_1}\otimes\varphi_{z_1}, P_{\zeta_1,z_1}\otimes P_{\zeta_2,z_2}).$$

It is possible to turn the first summand in (3.5.2) into the first summand of (3.4.11). First, make the change of variable

$$\zeta=\zeta(z,(\alpha_1,\beta_1),(\alpha_2,\beta_2))=((x_1+\alpha_1 s_1,\beta_1 s_1),(x_2+\alpha_2 s_2,\beta_2 s_2))$$

and then use Fubini's theorem in the inner variable of g. The first summand in (3.5.2) then equals

$$\int\limits_{Z^{\mathbf{d}}}\langle f,\varphi_z\rangle\langle\upsilon_z,g\rangle\,\mathrm{d}\mu(z),\qquad \upsilon_z:=\int\limits_{\substack{(\alpha_1,\beta_1)\in Z^{d_1}\\(\alpha_2,\beta_2)\in Z^{d_2}}}\Lambda\left(\chi_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2},\chi_{\zeta_1,z_1}\otimes\chi_{\zeta_2,z_2}\right)\varphi_\zeta\frac{\mathrm{d}\beta_2\mathrm{d}\alpha_2\mathrm{d}\beta_1\mathrm{d}\alpha_1}{\beta_1\beta_2},$$

where under the integral sign $\zeta=\zeta(z,(\alpha_1,\beta_1),(\alpha_2,\beta_2))$. With the same convention,

$$u_z\big((\alpha_1,\beta_1),(\alpha_1,\beta_1)\big)=\Lambda\left(\chi_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2},\chi_{\zeta_1,z_1}\otimes\chi_{\zeta_2,z_2}\right)\left(\prod_{j=1}^{2}[(\alpha_j,\beta_j)]_{\min\{\kappa,k_j\}+\eta}\right)^{-1}$$

is uniformly bounded via Lemma 3.5.2, and applying Lemma 3.5.1 yields $\upsilon_z\in C\Psi_z^{(\kappa,\min\{\kappa,k_2\}),\varepsilon;0}$.

It remains to identify the remaining terms in (3.5.2) as a sum of paraproduct terms. Here it is crucial to use assumption $a(\kappa)$, which tells us that

$$|\gamma_\iota|<\min\{\kappa,k_\iota\}\implies\mathbf{q}^{\iota,\mathfrak{a}}_{\gamma_\iota}(z_{\hat\iota},\zeta_{\hat\iota})=0\qquad\forall z_{\hat\iota},\zeta_{\hat\iota}\in Z^{d_{\hat\iota}}.$$

Focus on the term in the third line of (3.5.2) first. The above observation, the definition of P_{z_1,ζ_1} from Lemma 2.1.2, the definition of $\mathbf{q}^{1,\circ}_{\gamma_1}$, the fact that $\partial^{-\alpha_1}=R^{\alpha_1}|\nabla|^{-|\alpha_1|}$ with the definition of $\varphi_{\alpha_1,\zeta_1}$, see (2.1.8), gives

$$\Lambda(P_{z_1,\zeta_1}\otimes\chi_{z_2,\zeta_2},\varphi_{\zeta_1}\otimes\chi_{\zeta_2,z_2})=\sum_{|\gamma_1|=k_1}\langle\varphi_{z_1},\mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle\Lambda(x_1^{\gamma_1}\otimes\chi_{z_2,\zeta_2},{\sigma_1}^{-k_1}\varphi_{\zeta_1}\otimes\chi_{\zeta_2,z_2})$$
$$=\sum_{|\gamma_1|=k_1}\langle\varphi_{z_1},\mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle\langle\mathbf{q}^{1,\circ}_{\gamma_1}(z_2,\zeta_2),\mathsf{Sy}_z[|\nabla|^{-k}\varphi_1]\rangle=\sum_{|\gamma_1|=|\alpha_1|=k_1}\langle\varphi_{z_1},\mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle\langle R^{\alpha_1}\mathbf{q}^{1,\circ}_{\gamma_1}(z_2,\zeta_2),\varphi_{\alpha_1,z_1}\rangle.$$

Finally using Lemma 2.3.2 with $h = \langle f, \varphi_{z_2}\rangle_2$, the summand in the third line of (3.5.2) equals the sum over $|\gamma_1| = |\alpha_1| = \kappa$ of

$$\int_{(Z^{d_2})^2}\int_{Z^{d_1}_{\zeta_1}} \langle \mathbf{a}^{1,\circ}_{\gamma_1,\alpha_1}(z_2,\zeta_2), \varphi_{\alpha_1,z_1}\rangle \left\langle \langle f, \varphi_{z_2}\rangle_2, \vartheta_{\gamma_1,\zeta_1}\right\rangle_1 \left\langle \langle \varphi_{\zeta_2}, g\rangle_2, \varphi_{\zeta_1}\right\rangle_1 \mathrm{d}\mu(\zeta_1)\, [z_2,\zeta_2]_{\kappa_2+\eta}\mathrm{d}\mu(z_2)\mathrm{d}\mu(\zeta_2)$$

$$= \int_{Z^{d_2}\times Z^{d_2}} \Pi_{\mathbf{a}^{1,\circ}_{\gamma_1,\alpha_1}(z_2,\zeta_2),\gamma_1,\alpha_1} \left(\langle f, \varphi_{z_2}\rangle_2, \langle g, \varphi_{\zeta_2}\rangle_2\right) [z_2,\zeta_2]_{\kappa_2+\eta}\mathrm{d}\mu(z_2)\mathrm{d}\mu(\zeta_2) = \Pi_{\mathbf{a}^{1,\circ}_{\gamma_1,\alpha_1},\gamma_1,\alpha_1}(f,g)$$

where $\kappa_2 = \min\{\kappa, k_2\}$, which is one of the summands in the second line of (3.4.11). The three other types of summands in the second and third line of (3.4.11), constructed in exactly the same way, arise from the $\cdots$ terms in the third line of (3.5.2).

It remains to identify the terms of the type appearing in the third line (3.5.2). Using again $a(\kappa)$, these terms will appear only if $\kappa \le k_2$. Lemma 2.1.2 and the definition of the paraproducts of Λ then yield

$$\begin{aligned}
&\Lambda(P_{z_1,\zeta_1}\otimes P_{z_2,\zeta_2}, \varphi_{\zeta_1}\otimes\varphi_{\zeta_2}) \\
&= \sum_{|\gamma_1|=|\gamma_2|=\kappa} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle\langle \varphi_{z_2}, \mathsf{Sy}_{\zeta_2}\phi_{\gamma_2}\rangle \Lambda(x_1^{\gamma_1}\otimes x_2^{\gamma_2}, {\sigma_1}^{-k_1}\varphi_{\zeta_1}\otimes {\sigma_1}^{-k_1}\varphi_{\zeta_1}) \\
&= \sum_{|\gamma_1|=|\gamma_2|=\kappa} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle\langle \varphi_{z_2}, \mathsf{Sy}_{\zeta_2}\phi_{\gamma_2}\rangle \langle b^\circ_\gamma, \varphi_{\gamma_1,\zeta_1}\otimes\varphi_{\gamma_2,\zeta_2}\rangle.
\end{aligned}$$

An application of Lemma 2.3.2 with $h = \langle f, \varphi_{z_1}\rangle_1$ yields

$$F(z_1) := \int_{z_2\in A(\zeta_2)} \langle\langle f, \varphi_{z_1}\rangle_1, \varphi_{z_2}\rangle_2 \langle \varphi_{z_2}, \mathsf{Sy}_{\zeta_2}\phi_{\gamma_2}\rangle \mathrm{d}\mu(z_2) = \langle\langle f, \varphi_{z_1}\rangle_1, \vartheta_{\gamma_2,\zeta_2}\rangle_2 = \langle f, \varphi_{z_1}\otimes\vartheta_{\gamma_2,\zeta_2}\rangle$$

so that the summand in the third line of (3.5.2) equals the sum over $|\gamma_1| = \kappa_1, |\gamma_2| = \kappa_2$ of

$$
\begin{aligned}
&\int\limits_{Z^{d_2}\times Z^{d_2}} \langle b^\circ_\gamma, \varphi_{\gamma_1,\zeta_1} \otimes \varphi_{\gamma_2,\zeta_2}\rangle \langle \varphi_\zeta, g\rangle \int\limits_{z_1\in A(\zeta_1)} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle F(z_1)\mathrm{d}\mu(z_1)\mathrm{d}\mu(\zeta) \\
&= \int\limits_{Z^{d_2}\times Z^{d_2}} \langle b^\circ_\gamma, \varphi_{\gamma_1,\zeta_1} \otimes \varphi_{\gamma_2,\zeta_2}\rangle \langle \varphi_\zeta, g\rangle \int\limits_{z_1\in A(\zeta_1)} \langle \langle f, \vartheta_{\gamma_2,\zeta_2}\rangle_2, \varphi_{z_1}\rangle_1 \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle \mathrm{d}\mu(z_1)\mathrm{d}\mu(\zeta) \\
&= \int\limits_{Z^{d_2}\times Z^{d_2}} \langle b^\circ_\gamma \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \otimes \varphi_{\gamma_2,\zeta_2}\rangle \langle \varphi_\zeta, g\rangle \langle f, \vartheta_{\gamma_1,\zeta_1} \otimes \vartheta_{\gamma_2,\zeta_2}\rangle \,\mathrm{d}\mu(\zeta) = \Pi^\circ_{b^\circ_\gamma,\gamma}(f,g)
\end{aligned}
$$

where another application of Lemma 2.3.2 with $h = \langle f, \vartheta_{\gamma_2,\zeta_2}\rangle_2$ has been carried out and the definition of full paraproduct is finally taken advantage of: this is one of the terms appearing in the fourth line of (3.4.11). This procedure may be repeated for the additional terms in the third line of (3.5.2), thus completing the roster of terms in (3.4.11) under the additional assumption $a(\kappa)$. Namely, under this assumption, we have proved that

$$
\begin{aligned}
\Lambda(f,g) &= \int\limits_{Z^{\mathbf{d}}} \langle f, \varphi_z\rangle \langle \upsilon_z, g\rangle \,\mathrm{d}\mu(z) \\
&+ \begin{cases} \displaystyle\sum_{\substack{|\gamma_1|=|\gamma_2|=\kappa \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b^{\mathfrak{a}}_\gamma,\gamma}\right]^{\mathfrak{a}}(f,g) + \sum_{\substack{\iota\in\{1,2\} \\ \mathfrak{a}\in\{\circ,\star\} \\ |\alpha_\iota|=|\gamma_\iota|=\kappa}} \left[\Pi^\iota_{\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}}\right]^{\mathfrak{a}}(f,g) & \kappa \le k_2 \\ \displaystyle\sum_{\substack{|\alpha_1|=|\gamma_1|=\kappa \\ \mathfrak{a}\in\{\circ,\star\}}} \left[\Pi^1_{\mathbf{a}^{1,\mathfrak{a}}_{\gamma_1,\alpha_1}}\right]^{\mathfrak{a}}(f,g) & k_2 < \kappa \le k_1 \end{cases}
\end{aligned}
\tag{3.5.3}
$$

with families $\{\upsilon_z, \varphi_\zeta \in C\Psi^{(\kappa,\min\{\kappa,k_2\}),\varepsilon;0,0}_z : z \in Z^{\mathbf{d}}\}$ if $\kappa \le k_2$.

The assumption $a(\kappa)$ is then removed by an inductive argument. Recall that $k_1 \ge k_2$. Let $0 \le \kappa < k_1$ and assume that the representation (3.4.11) holds true for $k = (\kappa, \min\{\kappa, k_2\})$. Let $\widetilde{\Lambda}(f,g)$ be the form obtained by subtracting from Λ the second line of (3.4.11). Then

$$
\widetilde{\Lambda}(f,g) = \int\limits_{Z^{\mathbf{d}}} \langle f, \varphi_z\rangle \langle \upsilon_z, g\rangle \,\mathrm{d}\mu(z), \qquad \left\{\upsilon_z, \varphi_z \in C\Psi^{(\kappa,\min\{\kappa,k_2\}),\varepsilon;0,0}_z : z \in Z^{\mathbf{d}}\right\}
$$

coincides with the κ-th order cancellative part of Λ, is a bi-parameter wavelet form of type (3.4.1) and satisfies assumption $a(\kappa+1)$ of having all the relevant paraproducts up to order κ vanishing. We may thus apply the main step to $\widetilde{\Lambda}(f,g)$ with $\kappa+1$ in place of κ, resulting in (3.5.3), and obtain that

$$\Lambda(f,g) = \widetilde{\Lambda}(f,g) + \sum_{\kappa' \le \min\{\kappa, k_2\}} \sum_{\substack{\gamma=(\gamma_1,\gamma_2) \\ |\gamma_1|=|\gamma_2|=\kappa' \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b^{\mathfrak{a}}_{\gamma},\gamma}\right]^{\mathfrak{a}}(f,g) + \sum_{\substack{|\gamma_1|\le\kappa \\ |\gamma_2|\le\min\{\kappa,k_2\}}} \sum_{\substack{\iota\in\{1,2\} \\ |\alpha_\iota|=|\gamma_\iota| \\ \mathfrak{a}\in\{\circ,\star\}}} \left[\Pi^{\iota}_{\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}}\right]^{\mathfrak{a}}(f,g)$$

$$= \int_{Z^{\mathbf{d}}} \langle f, \varphi_z\rangle \langle \widetilde{v}_z, g\rangle \,\mathrm{d}\mu(z)$$

$$+ \sum_{\kappa' \le \min\{\kappa+1, k_2\}} \sum_{\substack{\gamma=(\gamma_1,\gamma_2) \\ |\gamma_1|=|\gamma_2|=\kappa' \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b^{\mathfrak{a}}_{\gamma},\gamma}\right]^{\mathfrak{a}}(f,g) + \sum_{\substack{|\gamma_1|\le\kappa+1 \\ |\gamma_2|\le\min\{\kappa+1,k_2\}}} \sum_{\substack{\iota\in\{1,2\} \\ |\alpha_\iota|=|\gamma_\iota| \\ \mathfrak{a}\in\{\circ,\star\}}} \left[\Pi^{\iota}_{\mathbf{a}^{\iota,\mathfrak{a}}_{\gamma_\iota,\alpha_\iota}}\right]^{\mathfrak{a}}(f,g)$$

with $\{\widetilde{v}_z, \varphi_z \in C\Psi_z^{(\kappa+1,\min\{\kappa+1,k_2\}),\varepsilon;0,0} : z \in Z^{\mathbf{d}}\}$. This achieves (3.4.11) for Λ with $k = (\kappa+1, \min\{\kappa+1, k_2\})$, thus completing the inductive step and the proof of Theorem 3.4.1.

3.6 Weighted Sobolev Estimates for Intrinsic Operators

This section contains the proofs of quantitative, and in some cases sharp, weighted estimates for the four types of summands occurring in the representation (3.4.11): see Propositions 3.6.1 and 3.6.2. Throughout, $[w]_{A_p}$ denotes the standard product weight characteristic on $\mathbb{R}^{\mathbf{d}} = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, see for example [17, 24].

3.6.1 Quantitative Bounds for Bi-Parameter Calderón-Zygmund Model Operators

To begin with, the operator T appearing in the following proposition is the adjoint of the first summand in (3.4.11), in the basic case $k = 0$.

Proposition 3.6.1. For $\delta > 0$ and $\left\{\upsilon_z \in C\Psi_z^{\delta;0,0} : z \in Z^{\mathbf{d}}\right\}$ consider the operator

$$Tf = \int\limits_{Z^{\mathbf{d}}} \langle f, \mathsf{Sy}_{z_1}\varphi_1 \otimes \mathsf{Sy}_{z_2}\varphi_2\rangle \upsilon_z \mathrm{d}\mu(z).$$

Then $\|T\|_{L^p(\mathbb{R}^{d_1}\times\mathbb{R}^{d_2};w)} \lesssim [w]_{A_p}^{\max\left\{2,1+\frac{2}{p-1}\right\}}$ for all $1 < p < \infty$, and this estimate is sharp when $p \geq 3$.

Next, adjoints to the full and partial paraproduct terms in (3.4.11) are treated: compare with the definitions in (3.4.2).

Proposition 3.6.2. Let $D \geq 8(d_1 + d_2)$. Fix $b \in \mathrm{BMO}(\mathbb{R}^{\mathbf{d}})$, $\{\vartheta_{z_j} \in C\Psi_{z_j}^{D,1;1} : z_j \in Z^{d_j}\}$, $\{\upsilon_{z_j}, \psi_{z_j} \in C\Psi_{z_j}^{D,1;0} : z_j \in Z^{d_j}\}$ for $j = 1, 2$. Then, the operators

$$\Pi_{(0,0),b} f := \int\limits_{Z^{\mathbf{d}}} \langle b, \upsilon_{z_1} \otimes \upsilon_{z_2}\rangle \langle f, \vartheta_{z_1} \otimes \vartheta_{z_2}\rangle \psi_{z_1} \otimes \psi_{z_2}\, \mathrm{d}\mu(z),$$

$$\Pi_{(0,1),b} f := \int\limits_{Z^{\mathbf{d}}} \langle b, \upsilon_{z_1} \otimes \upsilon_{z_2}\rangle \langle f, \vartheta_{z_1} \otimes \psi_{z_2}\rangle \psi_{z_1} \otimes \vartheta_{z_2}\, \mathrm{d}\mu(z),$$

satisfy the estimates

$$\|\Pi_{(0,0),b}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{\max\{3,2p\}}{p-1}} \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})}, \qquad 1 < p < \infty,$$

$$\|\Pi_{(0,1),b}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{\max\{2p+3,4p,5p-3\}}{2(p-1)}} \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})}, \qquad 1 < p < \infty.$$

The last proposition concerns adjoints to the half paraproduct terms in (3.4.11), see (3.4.4).

Proposition 3.6.3. Let $0 < \delta < 1$, $\mathbf{a} \in \mathcal{C}(Z^{d_2} \times Z^{d_2}; \mathrm{BMO}(\mathbb{R}^{\mathbf{d}}))$, and fix families

$$\{\vartheta_{z_1} \in C\Psi_{z_1}^{1,1;1} : z_1 \in Z^{d_1}\}, \qquad \{\upsilon_{z_1}, \psi_{z_1} \in C\Psi_{z_1}^{1,1;0} : z_1 \in Z^{d_1}\}, \qquad \{\psi_{z_2} \in C\Psi_{z_2}^{1,1;0} : z_2 \in Z^{d_2}\}.$$

Then, the operator

$$\Pi_{\mathbf{a}} f := \int_{Z^{d_2}_{z_2}} \int_{Z^{d_2}_{\zeta_2}} [z_2, \zeta_2]_\delta \int_{Z^{d_1}_{z_1}} \left\langle \mathbf{a}(z_2,\zeta_2)\right), \upsilon_{z_1} \right\rangle \langle f, \vartheta_{z_1} \otimes \psi_{z_2} \rangle \psi_{z_1} \otimes \psi_{\zeta_2} \, \mathrm{d}\mu(z_1)\mathrm{d}\mu(\zeta_2)\mathrm{d}\mu(z_2),$$

satisfies the estimate

$$\|\Pi_{\mathbf{a}}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{1}{2}+\max\left\{\frac{\max\{2,p\}}{p-1},\frac{3}{2}\right\}} \|\mathbf{a}\|_{\mathcal{C}(Z^{d_2}\times Z^{d_2};\mathrm{BMO}(\mathbb{R}^{\mathbf{d}}))}.$$

The proofs of Propositions 3.6.1, 3.6.2 and 3.6.3 are collected in Subsection 3.6.2. Along the way, we will make use of sharp weighted bounds for the intrinsic square function (3.2.2), as well as the mixed square-maximal operators

$$\mathrm{SM}(x) = \left(\int_{(0,\infty)} \sup_{t_2>0} \sup_{\psi\in\Psi^{\delta;0,1}_{((x_1,t_1),(x_2,t_2))}} |\langle f,\psi\rangle|^2 \frac{\mathrm{d}t_1}{t_1} \right)^{\frac{1}{2}},$$

$$\mathrm{MS}(x) = \sup_{t_1>0} \left(\int_{(0,\infty)} \sup_{\psi\in\Psi^{\delta;1,0}_{((x_1,t_1),(x_2,t_2))}} |\langle f,\psi\rangle|^2 \frac{\mathrm{d}t_2}{t_2} \right)^{\frac{1}{2}}, \qquad x=(x_1,x_2)\in\mathbb{R}^{\mathbf{d}}$$

which enter the L^p and weighted theory of the full and partial paraproduct terms. The square-maximal and maximal-square operators appearing below generalize those introduced in [41, 42]. There seem to be no pre-existing weighted estimate in past literature, thus our results are stated as a theorem.

Theorem 3.6.1. The operator norm bound

$$\|\mathrm{SS}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\max\left\{1,\frac{2}{p-1}\right\}}, \qquad \|\mathrm{SM}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}, \|\mathrm{MS}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{1}{p-1}\max\left\{2,\frac{p+1}{2}\right\}}, \quad (3.6.1)$$

holds for all $0<\delta\le 1$ and $1<p<\infty$. Furthermore, the exponent of (3.6.1) may not be improved for a generic weight w.

The SS bound in Theorem 3.6.1 is the bi-parameter analogue of [34, Theorem 1.2]. Its proof is given in the concluding Subsection 3.6.3 below. Another inequality that will be used a few times in Subsection 3.6.2 is a lower bound for the smaller tensor product square function

$$\mathrm{SS}_{\otimes} f(x_1,x_2) = \left(\int_{(0,\infty)^2} \left| \left\langle f, \mathsf{Sy}_{(x_1,t_1)}\varphi_1 \otimes \mathsf{Sy}_{(x_2,t_2)}\varphi_2 \right\rangle \right|^2 \frac{\mathrm{d}t_1 \mathrm{d}t_2}{t_1 t_2} \right)^{\frac{1}{2}}$$

associated with the wavelets φ_1, φ_2 from (1.2.1). The proof is a simple iteration argument, and is given immediately.

Proposition 3.6.4. $\|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p} \|\mathrm{SS}_{\otimes} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}$ for all $1 < p < \infty$.

Proof. Apply the main result of [52], see also [36, Theorem 2.7], first on each x_1-fiber in the second parameter, and subsequently in vector-valued form in the second parameter to see that

$$\begin{aligned} &\|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{1}{2}} \left\| \langle f(x_1,\cdot), \mathsf{Sy}_{(x_2,t_2)}\varphi_2 \rangle \right\|_{L^p(w(x_1,x_2);L^2(\mathrm{d}t_2/t_2))} \\ &\lesssim [w]_{A_p} \left\| \left\langle \langle f, \mathsf{Sy}_{(x_2,t_2)}\varphi_2 \rangle_2, \mathsf{Sy}_{(x_1,t_1)}\varphi_1 \right\rangle_1 \right\|_{L^p(w(x_1,x_2);L^2(\mathrm{d}t_1/t_1;L^2(\mathrm{d}t_2/t_2)))} = [w]_{A_p} \|\mathrm{SS}_{\otimes} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \end{aligned}$$

as claimed. □

3.6.2 Proofs of Propositions 3.6.1, 3.6.2 and 3.6.3

Proof of Proposition 3.6.1. Sharpness of $[w]_{A_p}^2$ for $p \geq 3$ follows by taking the tensor product of two counterexamples to sharpness of $[w]_{A_p}$ in one parameter. For the rest of the proof, we claim the pointwise bound

$$\mathrm{SS}_{\otimes}(Tf) \lesssim \mathrm{SS} f. \tag{3.6.2}$$

Assuming (3.6.2) holds,

$$\|Tf\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p} \|\mathrm{SS}_{\otimes}(Tf)\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p} \|\mathrm{SS} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{1+\max\left\{1,\frac{2}{p-1}\right\}} \|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}$$

thanks to an application of Proposition 3.6.4 in the first step and of Theorem 3.6.1 in the last. This proves Proposition 3.6.1 up to the verification of claim (3.6.2), which follows. Fix

$$\zeta = ((\xi_1, \sigma_1), (\xi_2, \sigma_2)) \in Z^{\mathbf{d}}.$$

Using the notation (1.2.1) for φ_ζ, writing $z = ((x_1, t_1), (x_2, t_2)) \in Z^{\mathbf{d}}$, and making the usual change of variable

$$\begin{aligned}
\langle Tf, \varphi_\zeta \rangle &= \int_{z \in Z^{\mathbf{d}}} \langle f, \varphi_z \rangle \langle \upsilon_z, \varphi_\zeta \rangle \, \mathrm{d}\mu(z) = \langle f, \psi_\zeta \rangle, \\
\psi_\zeta &\coloneqq \int_{\substack{(\alpha_1,\beta_1) \in Z^{d_1} \\ (\alpha_2,\beta_2) \in Z^{d_2}}} \langle \upsilon_{((\xi_1+\alpha_1\sigma_1, \beta_1\sigma_1),(\xi_2+\alpha_2\sigma_2, \beta_2\sigma_2))}, \varphi_\zeta \rangle \varphi_{((\xi_1+\alpha_1\sigma_1, \beta_1\sigma_1),(\xi_2+\alpha_2\sigma_2, \beta_2\sigma_2))} \frac{\mathrm{d}\beta_2 \mathrm{d}\alpha_2 \mathrm{d}\beta_1 \mathrm{d}\alpha_1}{\beta_1 \beta_2}.
\end{aligned}$$

Applying Lemma 3.1.1,

$$\left| \langle \upsilon_{((\xi_1+\alpha_1\sigma_1, \beta_1\sigma_1),(\xi_2+\alpha_2\sigma_2, \beta_2\sigma_2))}, \varphi_\zeta \rangle \right| \lesssim \left(\prod_{j=1}^{2} [(\alpha_j, \beta_j)]_{\frac{\delta}{2}} \right)$$

whence by Lemma 3.5.1, $\psi_\zeta \in C\Psi_\zeta^{\frac{\delta}{2};0,0}$, and (3.6.2) follows immediately from the definition of the intrinsic square function SS. □

Proof of Proposition 3.6.2. Let $\sigma \coloneqq w^{-\frac{1}{p-1}}$ be the dual weight to $w \in A_p$, so that $[\sigma]_{A_{p'}} = [w]_{A_p}^{\frac{1}{p-1}}$. Recall that M_{d_1,d_2} is the bi-parameter maximal function on $\mathbb{R}^{\mathbf{d}}$. The proof for $\Pi_{(0,0),b}$

begins with an appeal to $H^1 - \mathrm{BMO}$ duality, leading to

$$\begin{aligned}
\left|\langle \Pi_{(0,0),b} f, g\rangle\right| &\le \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \left\| \mathrm{SS}_{\otimes}\left(\int_{Z^{\mathbf{d}}} \langle f, \vartheta_{z_1}\otimes \vartheta_{z_2}\rangle \langle \psi_{z_1}\otimes\psi_{z_2}, g\rangle \upsilon_{z_1}\otimes \upsilon_{z_2}\,\mathrm{d}\mu(z)\right)\right\|_{L^1(\mathbb{R}^{\mathbf{d}})} \\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \left\|\mathrm{M}_{d_1,d_2}(f)\mathrm{SS}g\right\|_{L^1(\mathbb{R}^{\mathbf{d}})} \\
&\le \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \left\|\mathrm{M}_{d_1,d_2}(f)\right\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \left\|\mathrm{SS}(g)\right\|_{L^{p'}(\sigma,\mathbb{R}^{\mathbf{d}})} \\
&\lesssim [w]_{A_p}^{\frac{\max\{3,2p\}}{p-1}} \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \|g\|_{L^{p'}(\sigma,\mathbb{R}^{\mathbf{d}})}\,.
\end{aligned}$$

The passage to the second line is justified by a pointwise bound, whose proof is similar to (3.6.3) below, and is omitted. In the last line, Theorem 3.6.1 has been appealed to, and to the quantitative weighted estimate for the strong maximal function and square functions. The claimed estimate for $\Pi_{(0,0),b}$ then follows by duality.

The proof for $\Pi_{(0,1),b}$ is similar. Preliminarily notice that

$$\begin{aligned}
[(\alpha_1,\beta_1)]_{d_1}[(\alpha_2,\beta_2)]_{d_2}|\langle f, \vartheta_{(x_1+\alpha_1 t_1,\beta_1 t_1)}\otimes \psi_{(x_2+\alpha_2 t_2,\beta_2 t_2)}\rangle| &\lesssim \sup_{\psi\in\Psi^{\delta;0,1}_{((x_1,t_1),(x_2,t_2))}} |\langle f,\psi\rangle|, \\
[(\alpha_1,\beta_1)]_{d_1}[(\alpha_2,\beta_2)]_{d_2}|\langle g, \psi_{(x_1+\alpha_1 t_1,\beta_1 t_1)}\otimes \vartheta_{(x_2+\alpha_2 t_2,\beta_2 t_2)}\rangle| &\lesssim \sup_{\psi\in\Psi^{\delta;1,0}_{((x_1,t_1),(x_2,t_2))}} |\langle g,\psi\rangle|.
\end{aligned}$$

As $D \ge 8d_1, 8d_2$, Lemma 2.1.1 applied componentwise to bound $\langle \upsilon_{z_1}\otimes\upsilon_{z_2}, \mathsf{Sy}_{(x_1,t_1)}\varphi_1 \otimes \mathsf{Sy}_{(x_2,t_2)}\varphi_2\rangle$, with $z_j = (x_j+\alpha_j t_j, \beta_j t_j)$, $j = 1,2$ then yields

$$\begin{aligned}
&\left|\left\langle \int_{Z^{\mathbf{d}}} \langle f, \vartheta_{z_1}\otimes\psi_{z_2}\rangle\langle \psi_{z_1}\otimes\theta_{z_2}, g\rangle \upsilon_{z_1}\otimes\upsilon_{z_2}\,\mathrm{d}\mu(z), \mathsf{Sy}_{(x_1,t_1)}\varphi_1\otimes \mathsf{Sy}_{(x_2,t_2)}\varphi_2\right\rangle\right| \\
&\lesssim \left(\sup_{\psi\in\Psi^{\delta;0,1}_{((x_1,t_1),(x_2,t_2))}} |\langle f,\psi\rangle|\right)\left(\sup_{\psi\in\Psi^{\delta;1,0}_{((x_1,t_1),(x_2,t_2))}} |\langle g,\psi\rangle|\right).
\end{aligned} \tag{3.6.3}$$

The proof proper begins now. Using $H^1 - \mathrm{BMO}$ duality again, followed by (3.6.3) and one application of $L^2 - L^\infty$ Hölder inequality in each parameter,

$$\begin{aligned}
\left|\langle \Pi_{(0,1),b} f, g\rangle\right| &\leq \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \left\| \mathrm{SS}_\otimes \left(\int_{Z^{\mathbf{d}}} \langle f, \vartheta_{z_1} \otimes \psi_{z_2}\rangle \langle \psi_{z_1} \otimes \theta_{z_2}, g\rangle \upsilon_{z_1} \otimes \upsilon_{z_2} \, \mathrm{d}\mu(z) \right) \right\|_{L^1(\mathbb{R}^{\mathbf{d}})} \\
&\leq \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \|\mathrm{SM}(f)\mathrm{MS}(g)\|_{L^1(\mathbb{R}^{\mathbf{d}})} \\
&\leq \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \|\mathrm{SM}(f)\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \|\mathrm{MS}(g)\|_{L^{p'}(\sigma,\mathbb{R}^{\mathbf{d}})} \\
&\lesssim [w]_{A_p}^{\frac{\max\{2p+3,4p,5p-3\}}{2(p-1)}} \|b\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})} \|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \|g\|_{L^{p'}(\sigma,\mathbb{R}^{\mathbf{d}})} .
\end{aligned}$$

In the last line, the quantitative weighted estimates of the operators SM and MS from Theorem 3.6.1 have been called upon. By duality, this estimate proves the claimed bound of $\Pi_{(0,1),b}$ on $L^p(w)$ and completes the proof of the proposition. □

Proof of Proposition 3.6.3. This proof relies on the auxiliary operators

$$\mathrm{P}_b h(y_1) = \int_{Z^{d_1}} \langle b, \upsilon_{z_1}\rangle \langle h, \vartheta_{z_1}\rangle \psi_{z_1}(y_1) \, \mathrm{d}\mu(z_1), \qquad y_1 \in \mathbb{R}^{d_1}$$

which is a paraproduct with symbol $b \in \mathrm{BMO}(\mathbb{R}^{d_1})$ in the first parameter, and

$$\mathrm{S}_{(2),(\alpha_2,\beta_2)} h(y_2) = \left(\int_0^\infty \left|\left\langle g, \mathsf{Sy}_{(y_2+\alpha_2 t_2, \beta_2 t_2)} \varphi_2 \right\rangle\right|^2 \frac{\mathrm{d}t_2}{t_2} \right)^{\frac{1}{2}}, \qquad y_2 \in \mathbb{R}^{d_2}, \, (\alpha_2, \beta_2) \in Z^{d_2}$$

which is a shifted square function in the second parameter with smooth, compactly supported mother wavelet φ_2 as in (1.2.1); the simplified notation $\mathrm{S}_{(2)}$ is used in place of $\mathrm{S}_{(2),(0,1)}$. The main results of [7, 34] yield the operator norm bounds

$$\|\mathrm{S}_{(2),(\alpha_2,\beta_2)}\|_{L^p(\mathbb{R}^{d_2};W)} \lesssim_\varepsilon (\min\{1, \beta_2\})^{-\varepsilon} [W]_{A_p}^{\max\{\frac{1}{2}, \frac{1}{p-1}\}} \tag{3.6.4}$$

for all $\varepsilon > 0$, where W is a weight on $\mathbb{R}^{d_2}$ and $[W]_{A_p}$ denotes the corresponding weight characteristic. Then

$$\Pi_{\mathbf{a}} f(u) = \int\limits_{Z^{d_2}_{z_2}} \int\limits_{Z^{d_2}_{\zeta_2}} [z_2,\zeta_2]_\delta \mathrm{P}_{a(z_2,\zeta_2)}(\langle f, \psi_{z_2}\rangle_2)(u_1) \otimes \psi_{\zeta_2}(u_2) \mathrm{d}\mu(\zeta_2)\mathrm{d}\mu(\zeta_1), \qquad u=(u_1,u_2)\in \mathbb{R}^{\mathbf{d}}.$$

A calculation involving Lemma 2.1.1 applied to the inner product $\langle \psi_{\zeta_2}, \mathsf{Sy}_{(y_2,t_2)}\varphi_2\rangle$ followed by the change of variables $z_2 = (y_2 + a_2 t_2, b_2 t_2)$, $\zeta_2 = (y_2 + \alpha_2 t_2, \beta_2 t_2)$ then yields

$$\mathrm{S}_{(2)}[\Pi_{\mathbf{a}} f](y_1,y_2) \lesssim \int\limits_{\substack{\omega_2 := (\alpha_2,\beta_2)\in Z^{d_2}\\ w_2 := (a_2,b_2)\in Z^{d_2}}} \left(\int_0^\infty \left| \mathrm{P}_{\mathbf{a}((y_2+a_2t_2,b_2t_2),(y_2+\alpha_2t_2,\beta_2t_2))} \left(\langle f, \psi_{(y_2+a_2t_2,b_2t_2)}\rangle_2 \right)(y_1) \right|^2 \frac{\mathrm{d}t_2}{t_2} \right)^{\frac12}$$
$$\times [\omega_2]_1 [w_2,\omega_2]_\delta \, \mathrm{d}\mu(w_2)\mathrm{d}\mu(\omega_2).$$

Applying the reverse square function bound of [52] in the second parameter, followed by the sharp weighted estimate for the vector-valued paraproduct $\mathrm{P}_{\mathbf{a}((y_2+a_2t_2,b_2t_2),(y_2+\alpha_2t_2,\beta_2t_2))}$ to pass to the second line, and finally appealing to (3.6.4) with choice $\varepsilon = \frac{\delta}{2}$, we obtain

$$\|\Pi_{\mathbf{a}} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac12} \|\mathrm{S}_{(2)}\Pi_{\mathbf{a}} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}$$
$$\lesssim [w]_{A_p}^{\frac12 + \max\{1,\frac{1}{p-1}\}} \|\mathbf{a}\| \int\limits_{\substack{(\alpha_2,\beta_2)\in Z^{d_2}\\ (a_2,b_2)\in Z^{d_2}}} [(\alpha_2,\beta_2)]_1 [(a_2,b_2),(\alpha_2,\beta_2)]_\delta \left\| \mathrm{S}_{(2),(a_2,b_2)} f \right\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \frac{\mathrm{d}a_2\mathrm{d}b_2}{b_2}\frac{\mathrm{d}\alpha_2\mathrm{d}\beta_2}{\beta_2}$$
$$\lesssim [w]_{A_p}^{\frac12 + \max\{1,\frac{1}{p-1}\} + \max\{\frac12,\frac{1}{p-1}\}} \|\mathbf{a}\| \|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}.$$

For display reasons, above $\|\mathbf{a}\|$ stands for $\|\mathbf{a}\|_{\mathcal{C}(Z^{d_2}\times Z^{d_2};\mathrm{BMO}(\mathbb{R}^{\mathbf{d}}))}$. The proof of Proposition 3.6.3 is thus complete. □

3.6.3 Proof of Theorem 3.6.1

Sharpness of the exponent follows by tensor product of the usual one parameter examples. The one parameter square function example is discussed in [34] and references therein, while the example for the one parameter maximal operator is entirely classical.

The proof of the upper bound is analogous for all three operators, as it proceeds by reduction to iteration of one parameter, vector-valued weighted bounds. To fix ideas, the argument is given for SS, which is the most difficult case.

Fix $f \in L_0^\infty(\mathbb{R}^{\mathbf{d}})$ and let $\left\{\psi_{(x_1,t_1),(x_2,t_2)} \in \Psi^{\delta;0,0}_{(x_1,t_1),(x_2,t_2)}; x_j \in \mathbb{R}^{d_j}, 0 < t_j < \infty, j = 1,2\right\}$ be a family linearizing the supremum in (3.2.2). Throughout this proof, $\eta := \frac{\delta}{16}$. The first step consists of a decomposition of the linearizing family into wavelets with compact frequency support in one of the parameters. Let $\alpha \in \mathcal{S}(\mathbb{R}^{d_1})$ be a radial function with

$$\operatorname{supp}\widehat{\alpha} \subset \mathsf{B}_{(0,2)} \setminus \mathsf{B}_{(0,\frac{1}{2})}, \qquad \int_{-\infty}^{\infty} \widehat{\alpha}(2^s\xi)\,\mathrm{d}s = \mathbf{1}_{\mathbb{R}^{d_1}\setminus\{0\}}(\xi),$$

and also let $\beta \in \mathcal{S}(\mathbb{R}^{d_1})$ satisfy

$$\operatorname{supp}\widehat{\beta} \subset \mathsf{B}_{(0,3)} \setminus \mathsf{B}_{(0,\frac{1}{3})}, \qquad \widehat{\beta}(\xi) = 1 \quad \forall \xi \in \operatorname{supp}\widehat{\alpha}.$$

Set

$$\alpha_s = \mathsf{Dil}^1_{2^s}\alpha, \qquad \beta_s = \mathsf{Dil}^1_{2^s}\beta, \qquad \psi^s_{(x_1,t_1),(x_2,t_2)} := 2^{\eta|s|}\psi_{(x_1,t_1),(x_2,t_2)} *_1 \alpha_{s+\log t_1}$$

so that it is understood that $*_j$ denotes convolution in the j-th parameter only, and note that the scale of the parameter in α_s, β_s is logarithmic. For instance $\alpha_{s+\log t_1}$ below has Fourier support in the annulus $\sim {t_1}^{-1}2^{-s}$.

Lemma 3.6.1. For all $s \in \mathbb{R}$, $x_j \in \mathbb{R}^{d_j}$, $0 < t_j < \infty$, $j = 1,2$ we have $\psi^s_{(x_1,t_1),(x_2,t_2)} \in C\Psi^{\eta;0,0}_{(x_1,t_1),(x_2,t_2)}$.

Proof. By bi-parametric invariance of the assumption and assertion, it suffices to prove the case $x_j = 0_{\mathbb{R}^{d_j}}, t_j = 1$ for $j = 1,2$. For simplicity write ψ in place of $\psi_{(x_1,t_1),(x_2,t_2)}$. Applying

Lemma 2.1.1 for each fixed y_2 gives

$$|\psi *_1 \alpha_s(y_1, y_2)| = |\langle \psi(\cdot, y_2), \mathsf{Tr}_{y_1}\alpha_s\rangle| \lesssim \frac{[(0,1),(y_1,s)]_{8\eta}}{\langle y_2\rangle^{(d_2+8\eta)}} \lesssim \frac{2^{-8\eta|s|}}{\langle y_1\rangle^{(d_1+8\eta)}\langle y_2\rangle^{(d_2+8\eta)}} \qquad (3.6.5)$$

as $\langle y_2\rangle^{d_2+8\eta}\psi^{[2,y_2]}(\cdot) \in \Psi^{\delta;0}_{(0,1)}$ and $\mathsf{Tr}_{y_1}\alpha_s \in \Psi^{\delta;0}_{(y_1,s)}$. The last inequality is best seen by verifying the cases $s \geq 0$, $s < 0$ separately. Using the Fourier support and normalization of α_s, similarly

$$|\nabla_1(\psi *_1 \alpha_s)(y_1, y_2)| = |\psi *_1 \nabla\alpha_s(y_1, y_2)| \lesssim \frac{2^{-s-8\eta|s|}}{\langle y_1\rangle^{(d_1+8\eta)}\langle y_2\rangle^{(d_2+8\eta)}}.$$

If $0 < |h| < 1$ then, by the mean value theorem

$$\begin{aligned}
&|\psi *_1 \alpha_s(y_1 + h, y_2) - \psi *_1 \alpha_s(y_1, y_2)| \\
&\lesssim |h|^\eta \left(\sup_{|u_1|\sim|y_1|} \nabla_1(\psi *_1 \alpha_s)(u_1, y_2)\right)^\eta \left(|\psi *_1 \alpha_s(y_1 + h, y_2)| + |\psi *_1 \alpha_s(y_1, y_2)|\right)^{1-\eta} \qquad (3.6.6)\\
&\lesssim |h|^\eta \frac{2^{|s|[\eta - 8\eta(1-\eta)]}}{\langle y_1\rangle^{(d_1+8\eta)}\langle y_2\rangle^{(d_2+8\eta)}} \leq 2^{-|s|\eta} \frac{|h|^\eta}{\langle y_1\rangle^{(d_1+8\eta)}\langle y_2\rangle^{(d_2+8\eta)}}
\end{aligned}$$

using the elementary inequality $6\eta > 8\eta^2$. The inequality

$$|\psi *_1 \alpha_s(y_1, y_2 + h) - \psi *_1 \alpha_s(y_1, y_2)| \lesssim \frac{|h|^\delta}{\langle y_1\rangle^{(d_1+\delta)}\langle y_2\rangle^{(d_2+\delta)}}$$

is immediate from (3.1.2) and averaging, so that another interpolation with (3.6.5) yields

$$|\psi *_1 \alpha_s(y_1, y_2 + h) - \psi *_1 \alpha_s(y_1, y_2)| \lesssim \frac{|h|^{\frac{\delta}{2}} 2^{-4\eta|s|}}{\langle y_1\rangle^{(d_1+8\eta)}\langle y_2\rangle^{(d_2+8\eta)}} \leq 2^{-\eta|s|} \frac{|h|^\eta}{\langle y_1\rangle^{(d_1+\eta)}\langle y_2\rangle^{(d_2+\eta)}}. \qquad (3.6.7)$$

Collecting (3.6.5), (3.6.6), and (3.6.7), and comparing with (3.1.2), completes the proof. □

The definitions of α_s and β_s lead to the equalities

$$\psi_{(x_1,t_1),(x_2,t_2)} = \int_{-\infty}^{\infty} 2^{-\eta|s|}\psi^s_{(x_1,t_1),(x_2,t_2)}\,\mathrm{d}s, \qquad \langle f, \psi^s_{(x_1,t_1),(x_2,t_2)}\rangle = \langle f *_1 \beta_{s+\log t_1}, \psi^s_{(x_1,t_1),(x_2,t_2)}\rangle.$$

Therefore, in view of Lemma 3.6.1, and using the convergence of the geometric integral, it will be enough to prove the same estimate for the operator

$$O_s f(x) = \left(\int_{(0,\infty)^2} |\langle f *_1 \beta_{s+\log t_1}, \psi^s_{(x_1,t_1),(x_2,t_2)}\rangle|^2 \frac{\mathrm{d}t_1\mathrm{d}t_2}{t_1 t_2} \right)^{\frac{1}{2}}$$

uniformly in the parameter $s \in \mathbb{R}$, which will be kept fixed until the end of the proof. The operator O_s is estimated relying on the auxiliary family of square functions with parameter $t_1 > 0$

$$\mathrm{S}_{t_1} h(x_1, x_2) := \left(\int_0^\infty |\langle h, \psi^s_{(x_1,t_1),(x_2,t_2)}\rangle|^2 \frac{\mathrm{d}t_2}{t_2} \right)^{\frac{1}{2}},$$

which satisfies

$$\|\mathrm{S}_{t_1}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\max\{\frac{1}{p-1},\frac{1}{2}\}}.$$

This can be seen by repeating the sparse domination bound for the Christ-Journé type square function e.g. of [7,34], where the averages in the sparse operators are associated to rectangles with side of fixed length t_1 in the first parameter. The fact that the weight is a product weight ensures that the bound is uniform over all t_1. The weighted bound above upgrades immediately to vector-valued, and may be used in the second step below to yield

$$\begin{aligned}\|O_s f\|_{L^p(w)} &= \|\mathrm{S}_{t_1}(f *_1 \beta_{s+\log t_1})\|_{L^p(w;L^2(\frac{\mathrm{d}t_1}{t_1}))} \lesssim [w]_{A_p}^{\max\{\frac{1}{p-1},\frac{1}{2}\}} \|f *_1 \beta_{s+\log t_1}\|_{L^p(w;L^2(\frac{\mathrm{d}t_1}{t_1}))} \\ &= [w]_{A_p}^{\max\{\frac{1}{p-1},\frac{1}{2}\}} \|f *_1 \beta_{\log t_1}\|_{L^p(w;L^2(\frac{\mathrm{d}t_1}{t_1}))} \lesssim [w]_{A_p}^{\max\{\frac{2}{p-1},1\}} \|f\|_{L^p(w)}.\end{aligned}$$

The very last inequality is obtained by using the straightforward weighted Littlewood-Paley square function bound of [7, 34] in the first parameter and Fubini's theorem. The proof of Theorem 3.6.1 is complete.

Bibliography

[1] Bradley K. Alpert, *A class of bases in L^2 for the sparse representation of integral operators*, SIAM J. Math. Anal. **24** (1993), no. 1, 246–262. MR1199538

[2] Alex Barron, José M. Conde-Alonso, Yumeng Ou, and Guillermo Rey, *Sparse domination and the strong maximal function*, Adv. Math. **345** (2019), 1–26. MR3897437

[3] Alexander Barron and Jill Pipher, *Sparse Domination for Bi-Parameter Operators Using Square Functions*, arXiv e-prints (September 2017), arXiv:1709.05009, available at `1709.05009`.

[4] Cristina Benea and Frédéric Bernicot, *Conservation de certaines propriétés à travers un contrôle épars d'un opérateur et applications au projecteur de Leray-Hopf*, Ann. Inst. Fourier (Grenoble) **68** (2018), no. 6, 2329–2379. MR3897969

[5] Árpád Bényi and Rodolfo H. Torres, *The discrete Calderón reproducing formula of Frazier and Jawerth*, Functional analysis, harmonic analysis, and image processing: a collection of papers in honor of Björn Jawerth, 2017, pp. 79–107. MR3682606

[6] Oleksandra V. Beznosova, *Linear bound for the dyadic paraproduct on weighted lebesgue space l2(w)*, Journal of Functional Analysis **255** (2008), no. 4, 994–1007.

[7] Gianmarco Brocchi, *A sparse quadratic $T1$ theorem*, arXiv e-prints (April 2020), arXiv:2004.05365, available at `2004.05365`.

[8] A. P. Calderon and A. Zygmund, *On the existence of certain singular integrals*, Acta Mathematica **88** (1952), no. none, 85 –139.

[9] Lennart Carleson, *A Counter Example for Measures Bounded on H^p for the Bi-Disc* (1974).

[10] David V. Cruz-Uribe, José Maria Martell, and Carlos Pérez, *Weights, extrapolation and the theory of Rubio de Francia*, Operator Theory: Advances and Applications, vol. 215, Birkhäuser/Springer Basel AG, Basel, 2011. MR2797562

[11] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *Uniform sparse domination of singular integrals via dyadic shifts*, Math. Res. Lett. **25** (2018), no. 1, 21–42. MR3818613

[12] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *A sparse estimate for multisublinear forms involving vector-valued maximal functions*, Bruno Pini Mathematical Analysis Seminar **8** (2018), no. 1, 168–184.

[13] Guy David and Jean-Lin Journé, *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. (2) **120** (1984), no. 2, 371–397. MR763911

[14] Oliver Dragičević and Alexander Volberg, *Sharp estimate of the Ahlfors-Beurling operator via averaging martingale transforms*, Michigan Math. J. **51** (2003), no. 2, 415–435. MR1992955

[15] Javier Duoandikoetxea, *Extrapolation of weights revisited: new proofs and sharp bounds*, J. Funct. Anal. **260** (2011), no. 6, 1886–1901. MR2754896

[16] R. Fefferman, *A^p weights and singular integrals*, Amer. J. Math. **110** (1988), no. 5, 975–987. MR961502

[17] R. Fefferman and J. Pipher, *Multiparameter operators and sharp weighted inequalities*, Amer. J. Math. **119** (1997), no. 2, 337–369. MR1439553

[18] Robert Fefferman, *Singular integrals on product domains*, Bulletin (New Series) of the American Mathematical Society **4** (1981), no. 2, 195 –201.

[19] ______, *Harmonic analysis on product spaces*, Ann. of Math. (2) **126** (1987), no. 1, 109–130. MR898053

[20] Sarah H. Ferguson and Michael T. Lacey, *A characterization of product BMO by commutators*, Acta Math. **189** (2002), no. 2, 143–160. MR1961195

[21] Sarah H. Ferguson and Cora Sadosky, *Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures*, J. Anal. Math. **81** (2000), 239–267. MR1785283

[22] M. Frazier, R. Torres, and G. Weiss, *The boundedness of Calderón-Zygmund operators on the spaces $\dot{F}_p^{\alpha,q}$*, Rev. Mat. Iberoamericana **4** (1988), no. 1, 41–72. MR1009119

[23] Michael Frazier, Björn Jawerth, and Guido Weiss, *Littlewood-Paley theory and the study of function spaces*, CBMS Regional Conference Series in Mathematics, vol. 79, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1991. MR1107300

[24] Paul Hagelstein, Teresa Luque, and Ioannis Parissis, *Tauberian conditions, Muckenhoupt weights, and differentiation properties of weighted bases*, Trans. Amer. Math. Soc. **367** (2015), no. 11, 7999–8032. MR3391907

[25] Jarod Hart and Lucas Oliveira, *Hardy space estimates for limited ranges of muckenhoupt weights*, Advances in Mathematics **313** (2017), 803 –838.

[26] Irina Holmes, Stefanie Petermichl, and Brett D. Wick, *Weighted little bmo and two-weight inequalities for Journé commutators*, Anal. PDE **11** (2018), no. 7, 1693–1740. MR3810470

[27] Tuomas Hytönen and Stefanos Lappas, *Dyadic representation theorem using smooth wavelets with compact support*, arXiv e-prints (March 2020), arXiv:2003.04019, available at `2003.04019`.

[28] Tuomas P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. MR2912709

[29] Jean-Lin Journé, *Calderón-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), no. 3, 55–91. MR836284

[30] Tosio Kato and Gustavo Ponce, *Commutator estimates and the Euler and Navier-Stokes equations*, Comm. Pure Appl. Math. **41** (1988), no. 7, 891–907. MR951744

[31] Michael T. Lacey and Darío Mena Arias, *The sparse T1 theorem*, Houston J. Math. **43** (2017), no. 1, 111–127. MR3647935

[32] Michael T. Lacey, Stefanie Petermichl, and Maria Carmen Reguera, *Sharp A_2 inequality for Haar shift operators*, Math. Ann. **348** (2010), no. 1, 127–141. MR2657437

[33] Andrei Lerner, *A pointwise estimate for the local sharp maximal function with applications to singular integrals*, Bulletin of The London Mathematical Society - BULL LOND MATH SOC **42** (201009).

[34] Andrei K. Lerner, *Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals*, Adv. Math. **226** (2011), no. 5, 3912–3926. MR2770437

[35] ______, *A simple proof of the A_2 conjecture*, Int. Math. Res. Not. IMRN **14** (2013), 3159–3170. MR3085756

[36] ______, *Quantitative weighted estimates for the Littlewood-Paley square function and Marcinkiewicz multipliers*, Math. Res. Lett. **26** (2019), no. 2, 537–556. MR3999554

[37] Kangwei Li, Henri Martikainen, Yumeng Ou, and Emil Vuorinen, *Bilinear representation theorem*, Trans. Amer. Math. Soc. **371** (2019), no. 6, 4193–4214. MR3917220

[38] Kangwei Li, Henri Martikainen, and Emil Vuorinen, *Bilinear Calderón-Zygmund theory on product spaces*, Journal de Mathématiques Pures et Appliquées **138** (2020), 356 –412.

[39] Henri Martikainen, *Representation of bi-parameter singular integrals by dyadic operators*, Adv. Math. **229** (2012), no. 3, 1734–1761. MR2871155

[40] Y. Meyer, *Continuité sur les espaces de Hölder et de Sobolev des opérateurs définis par des intégrales singulières*, Recent progress in Fourier analysis (El Escorial, 1983), 1985, pp. 145–172. MR848145

[41] Camil Muscalu, Jill Pipher, Terence Tao, and Christoph Thiele, *Bi-parameter paraproducts*, Acta Math. **193** (2004), no. 2, 269–296. MR2134868 (2005m:42028)

[42] ______, *Multi-parameter paraproducts*, Rev. Mat. Iberoam. **22** (2006), no. 3, 963–976. MR2320408 (2008b:42037)

[43] Yumeng Ou, *Multi-parameter singular integral operators and representation theorem*, Rev. Mat. Iberoam. **33** (2017), no. 1, 325–350. MR3615454

[44] S. Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic*, Amer. J. Math. **129** (2007), no. 5, 1355–1375. MR2354322

[45] S. Petermichl, S. Treil, and A. Volberg, *Why the Riesz transforms are averages of the dyadic shifts?*, Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000), 2002, pp. 209–228. MR1964822

[46] Martí Prats and Xavier Tolsa, *A $T(P)$ theorem for Sobolev spaces on domains*, J. Funct. Anal. **268** (2015), no. 10, 2946–2989. MR3331790

[47] Robert Rahm, Eric T. Sawyer, and Brett D. Wick, *Weighted Alpert Wavelets*, arXiv e-prints (August 2018), arXiv:1808.01223, available at `1808.01223`.

[48] Terence Tao, *Lecture notes for 247b*, 2007.

[49] Rodolfo H. Torres, *Boundedness results for operators with singular kernels on distribution spaces*, Mem. Amer. Math. Soc. **90** (1991), no. 442, viii+172. MR1048075

[50] Armen Vagharshakyan, *Recovering singular integrals from Haar shifts*, Proc. Amer. Math. Soc. **138** (2010), no. 12, 4303–4309. MR2680056

[51] Kunchuan Wang, *The generalization of paraproducts and the full $T1$ theorem for Sobolev and Triebel-Lizorkin spaces*, J. Math. Anal. Appl. **209** (1997), no. 2, 317–340. MR1474612

[52] J. Michael Wilson, *Weighted norm inequalities for the continuous square function*, Trans. Amer. Math. Soc. **314** (1989), no. 2, 661–692. MR972707

[53] Michael Wilson, *Weighted Littlewood-Paley theory and exponential-square integrability*, Lecture Notes in Mathematics, vol. 1924, Springer, Berlin, 2008. MR2359017