PURE and APPLIED ANALYSIS

FRANCESCO DI PLINIO, WALTON GREEN AND BRETT D. WICK

BILINEAR WAVELET REPRESENTATION OF CALDERÓN–ZYGMUND FORMS

msp

vol. 5 no. 1 2023

PURE and APPLIED ANALYSIS

Vol. 5, No. 1, 2023
https://doi.org/10.2140/paa.2023.5.47

msp

BILINEAR WAVELET REPRESENTATION OF CALDERÓN–ZYGMUND FORMS

FRANCESCO DI PLINIO, WALTON GREEN AND BRETT D. WICK

We represent a bilinear Calderón–Zygmund operator at a given smoothness level as a finite sum of cancellative, complexity-zero operators, involving smooth wavelet forms, and continuous paraproduct forms. This representation results in a sparse $T(1)$-type bound, which in turn yields directly new sharp weighted bilinear estimates on Lebesgue and Sobolev spaces. Moreover, we apply the representation theorem to study fractional differentiation of bilinear operators, establishing Leibniz-type rules in weighted Sobolev spaces which are new even in the simplest case of the pointwise product.

1. Introduction

Wavelet decompositions play a central role in the study of singular operators on real variable function spaces. Haar wavelet techniques finding their roots in the works of Figiel [1990], Nazarov, Treil and Volberg [Nazarov et al. 2003], Petermichl [2007] among others, have led to a powerful and comprehensive theory of singular integrals on Lebesgue spaces, most prominently for Calderón–Zygmund operators (CZOs). Smooth wavelets, sometimes called smooth atoms or molecules, have similarly powered the study of mapping properties of linear and multilinear singular operators on smoothness scales such as the Sobolev, Besov and Triebel–Lizorkin scales; see, e.g., [Frazier et al. 1988] and the more recent [Hart and Oliveira 2017]. Our approach herein seeks to unify these two perspectives.

The driving result of this article, continuing the theme from [Di Plinio et al. 2022], is a representation of bilinear Calderón–Zygmund operators in terms of model operators which reflects the eventual additional smoothness of their off-diagonal kernel. This representation is realized as a sum of continuous paraproduct forms and finitely many cancellative forms, which are themselves smooth bilinear Calderón–Zygmund operators. To wit, the cancellative components of our decomposition, which we term *wavelet forms*, are completely diagonalized forms with respect to a suitable wavelet-type basis. Furthermore, each wavelet form should be viewed as a certain approximate projection in the frequency domain. Our prototypical one is

$$U(f,g,h)=\int_0^\infty\int_{\mathbb{R}^d}\langle f\otimes g,\nu_{w,t}\rangle\langle h,\phi_{w,t}\rangle\frac{\mathrm{d}w\,\mathrm{d}t}{t}. \tag{1-1}$$

In this formula, $\phi_{w,t}=t^{-d}\phi((\cdot-w)/t)$ for a smooth wavelet ϕ, while $\nu_{w,t}$ behaves like the tensor product of two smooth wavelets translated by w and dilated by t. We say that $\nu_{w,t}$ belongs to the wavelet class $\Psi^{k,\delta;1,0}_{w,t}$ defined below: the cancellation structure of this class reflects the "low-high-high" component

Di Plinio was partially supported by the National Science Foundation under the grant NSF-DMS-2000510. Wick was partially supported by NSF grant NSF-DMS-1800057 as well as ARC DP190100970.

MSC2020: primary 42B20; secondary 42B25.

Keywords: wavelet representation theorem, bilinear singular integrals, $T(1)$-theorems, sharp weighted bounds, Leibniz rules, fractional differentiation, sparse domination.

of the resolution of the pointwise product. The diagonal nature of the representation of bilinear CZOs is obtained at the expense of converting the compactly supported wavelets occurring in the resolution of the bilinear identity into noncompactly supported wavelets $\nu_{w,t}$ via the *wavelet averaging lemma*, Lemma 2.4 below. Next, we explain the advantages of our representation. The wavelet and paraproduct model operators are dominated by intrinsic localized forms, which in turn satisfy a sharp form of *sparse bounds*. Sparse domination, a technique originating from the early work of Lerner [2013] and then developed by several authors within and beyond Calderón–Zygmund theory, see, e.g., [Beltran and Cladek 2020; Culiuc et al. 2018; Lacey 2017; Lerner 2016], subsumes the full range, and the sharp quantification of the weighted norm inequalities for the operator under sparse control. Thus, in combination with sparse bounds, our representation theorem yields a variety of novel bilinear, weighted and sharply quantified $T(1)$-theorems on smoothness spaces. We exemplify this paradigm by the loosely described weighted Sobolev theorem that follows, summarizing the results of Section 4. Let us informally introduce a few definitions.

A bilinear operator T is a $(0, 0, \delta)$ CZO if its off-diagonal kernel satisfies standard bilinear δ-kernel estimates, while T satisfies both the bilinear weak boundedness property and bilinear $T(1)$ testing conditions. These are standard conditions under which the Lebesgue space mapping properties of T are now well understood. We generalize to (k_1, k_2, δ) CZOs whose kernels are $k_1 + k_2$ times differentiable, with appropriate decay estimates and in addition to the weak boundedness testing condition, satisfy an iterative testing condition on monomials x^γ, producing elements $b^i_\gamma \in \mathrm{BMO}$ for $i = 0, 1, 2$ and $\gamma = (\gamma_1, \gamma_2)$ with $|\gamma_j| \le k_j$. Sections 3A and 3B contain the precise definitions.

Theorem. *Let $k_1, k_2 \in \mathbb{N}$, $\delta > 0$, $0 \le \sigma \le \min\{k_1, k_2\}$, $1 < p_1, p_2 \le \infty$, $\frac{1}{2} < p := p_1 p_2/(p_1 + p_2) < \infty$. Let T be a (k_1, k_2, δ) CZO on $\mathbb{R}^d$ such that $D^{\sigma-|\gamma|} b^0_\gamma \in \mathrm{BMO}$ for all $|\gamma| \le \sigma$. Suppose that the weight vector $\vec{v} = (v_1, v_2, v_3)$ satisfies, with $\vec{p} = (p_1, p_2, p/(p-1))$,*

$$\sigma \in \{0, 1, \dots, d-1\} \cup [d, \infty), \quad \vec{v} \in A_{\vec{p}},$$

or

$$\sigma \in \Big(\frac{d}{p} - d, d\Big), \quad \vec{v} \in A_{\vec{p},\vec{r}}, 1 \le r_i < p_i, \quad i = 1, 2, \quad \frac{1}{r_1} + \frac{1}{r_2} < \frac{\sigma + d}{d}, \quad r_3 = 1.$$

Then there exists $C > 0$ such that

$$\Big\| \frac{D^\sigma T(f_1, f_2)}{v_3} \Big\|_{L^p(\mathbb{R}^d)} \le C \|f_1\|_{W^{\sigma,p_1}(v_1)} \|f_2\|_{W^{\sigma,p_2}(v_2)}. \tag{1-2}$$

The constant C depends on the parameters above, the operator T, and the appropriate weight characteristic of the weight vector $\vec{v}$.

The bilinear Muckenhoupt weight vectors, that is, the classes $A_{\vec{p}}$, $A_{\vec{p},\vec{r}}$ appearing in the statement, are explicitly defined in Section 4A. They were first introduced in [Lerner et al. 2009], and later appeared in [Chaffee et al. 2017; Culiuc et al. 2017; 2018], as the natural multilinear substitute for the role of the classical A_p in linear Calderón–Zygmund theory. In the subsequent articles [Nieraeth 2019; Li et al. 2020; 2021], a complete and useful extrapolation theory for these classes was developed. These references also contain details on the relation between $A_{\vec{p},\vec{r}}$ and the linear classes A_p. The inhomogeneous weighted

Sobolev spaces $W^{\sigma,q}(v)$ are defined by the norm

$$\|f\|_{W^{\sigma,q}(v)} = \|[D^\sigma f]v\|_{L^q(\mathbb{R}^d)} + \sum_{k=0}^{\lfloor\sigma\rfloor} \|[D^k f]v\|_{L^q(\mathbb{R}^d)},$$

where D^σ is the fractional derivative defined by the Fourier multiplier $m(\xi) = |\xi|^\sigma$. The statement above shows how our representation theorem unifies the treatment of smooth CZOs with fractional differentiation. To see this, note that taking T to be the pointwise product operator returns a form of the Coifman–Meyer–Kenig–Stein–Grafakos–Torres fractional Leibniz rule; see the excellent survey [Grafakos 2017].

Bilinear representations of dyadic-probabilistic type, originating from Hytönen's theorem [2012], have been developed more recently [Li et al. 2019] to obtain results of this type in the case $\sigma = 0$: see also the recent works [Di Plinio et al. 2020a; 2020b] for multilinear representations in UMD spaces and previous works [Li et al. 2014; Lerner and Nazarov 2019] on sharp weighted norm inequalities for multilinear operators not reliant on representation formulas. When σ is a positive integer, boundedness on Sobolev, Besov and Triebel–Lizorkin spaces is known for certain cancellative CZOs in the Banach range [Bényi 2003; Maldonado and Naibo 2009]. A complete multilinear, sharp weighted theory on fractional smoothness spaces, and with full treatment of the paraproducts was unknown prior to this work. In the previously given example, the dependence of C in (1-2) upon the characteristic of the weight $\vec{v}$ is sharply quantified [Li et al. 2014; 2020; Nieraeth 2019; Lerner and Nazarov 2019]. See the statement of Theorem B below for the explicit form.

On the other hand, the fractional derivative D^σ applied to $T(f_1, f_2)$ has received renewed interest since its initial study by Kato and Ponce [1988] and Kenig, Ponce and Vega [Kenig et al. 1993] when Grafakos and Oh [2014] and Muscalu and Schlag [2013] independently extended the $L^{p_1} \times L^{p_2} \to L^p$ results to the sharp bilinear range, taking $p < 1$.

Since then, Leibniz rules for Fourier multiplier operators and certain pseudodifferential operators have also been obtained. Prior weighted estimates require the memberships $v_i \in A_{p_i}$ for each single weight [Muscalu and Schlag 2013; Cruz-Uribe and Naibo 2016; Hart et al. 2018; Brummer and Naibo 2018; 2019; Naibo and Thomson 2019]. This condition is strictly more restrictive than membership of the weight vector $\vec{v}$ to the multilinear weighted classes $A_{\vec{p}}$ required in our theorem, so that strictly speaking (1-2) is new even when T is the identity. Furthermore, the class of smooth CZOs which we consider includes both smooth Fourier multipliers and certain classes of pseudodifferential operators.

Organization. The paper is almost entirely self-contained. We only use the well-known principle that sparse domination implies sharp weighted Lebesgue space bounds as a black box. One can consult [Li et al. 2020; Nieraeth 2019] for a precise statement, but we also refer to some of the pioneering works [Lerner 2013; Li et al. 2014; Lerner and Nazarov 2019; Conde-Alonso and Rey 2016] concerning this principle. Otherwise, we do not appeal to dyadic or linear representation or $T(1)$ theorems, abstract sparse domination results, or the Coifman–Meyer multiplier theorem. We will need two technical lemmas from the study of the linear wavelet representation theory in [Di Plinio et al. 2022] on the boundedness of the intrinsic square function and the almost-orthogonality of the wavelet classes.

We begin by recalling the Calderón reproducing formula and extending it to a certain multilinear setting, using high-low cancellation. In the same section, we introduce the linear and multilinear wavelet classes, Ψ_z, and prove the key wavelet averaging Lemma 2.4. This lemma allows us to avoid

wavelet operators with *complexity*, in the language of [Hytönen 2012], completely diagonalizing the CZO. In Section 3, we state the technical definitions of CZOs and higher-order paraproducts, which are smooth testing conditions of $T(1)$-type. After this, we prove the off-diagonal estimates and deduce the representation theorem. Section 4 is devoted to applications of the representation theorem, specifically to obtain the weighted Sobolev and fractional Sobolev space bounds as a consequence of the sparse domination of the intrinsic forms the latter being proved in Section 5. Section 6 contains an asymmetric formulation of the results and the extension from bilinear to m-linear operators. We conclude with some remarks and further questions in Section 7.

2. Wavelets

To facilitate the description of our wavelet system a few pieces of notation need to be introduced. We work with a fixed dimension $d \geq 1$; thus the space of Schwartz functions $\mathcal{S}(\mathbb{R}^d)$ is simply denoted by $\mathcal{S}$ when no confusion arises. The Fourier transform $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is normalized as

$$(\mathcal{F}\phi)(\xi) = \hat{\phi}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \phi(x) \mathrm{e}^{-ix\xi} \, \mathrm{d}x.$$

With the above normalization, $\mathcal{F}\hat{\phi} = \phi(-\,\cdot\,)$. The affine group $Z^d := \mathbb{R}^d \ltimes (0, \infty)$ acts on $\phi \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ unitarily by

$$\mathsf{Sy}_z \phi(\,\cdot\,) = \phi_z(\,\cdot\,) := \frac{1}{t^d} \phi\left(\frac{\cdot - w}{t}\right), \quad z = (w, t) \in Z^d.$$

A function $\phi \in \mathcal{S}$ is *admissible* if

$$\int_0^\infty |\hat{\phi}(\rho\xi)|^2 \, \frac{\mathrm{d}\xi}{\xi} = 1 \tag{2-1}$$

for all $\rho \in \mathbb{S}^{d-1}$. If ϕ is admissible, the *Calderón formula*

$$f = \int_{Z^d} \langle f, \phi_z \rangle \phi_z \, \mathrm{d}\mu(z) \quad \text{for all } f \in \mathcal{S} \tag{2-2}$$

holds; see, e.g., [Frazier et al. 1991]. Here and in what follows, μ is the invariant measure on Z^d given by

$$\int_{Z^d} f(z) \, \mathrm{d}\mu(z) = \int_{\mathbb{R}^d \times (0,\infty)} f(w, t) \, \frac{\mathrm{d}w \, \mathrm{d}t}{t}, \quad f \in C_0(Z^d).$$

The admissibility condition (2-1) implies that ϕ has mean zero. In general, our wavelets are required to have more cancellation. Define

$$\mathcal{S}_j = \left\{ \phi \in \mathcal{S} : \int x^\alpha \phi(x) \, dx = 0 \ \text{ for } \ 0 \leq |\alpha| \leq j \right\}.$$

For functions $\phi \in \mathcal{S}_j$, and $\gamma \in \mathbb{N}^d$, $0 \leq |\gamma| \leq j$, define the antiderivative of order γ to be the Fourier multiplier

$$\partial^{-\gamma} \phi(x) = \int_{\mathbb{R}^d} \frac{\xi^\gamma}{|\xi|^{2|\gamma|}} \hat{\phi}(\xi) \mathrm{e}^{ix\xi} \, \mathrm{d}\xi, \quad x \in \mathbb{R}^d.$$

If $\phi \in \mathcal{S}_j$, then $|\xi|^{-j}\hat{\phi}(\xi)$ is bounded for all ξ (specifically close to zero) so the integral defining $\partial^{-\gamma}\phi$ converges absolutely. Denote by ∂^γ the usual partial differentiation operator. This can also be written, for $f \in \mathcal{S}$,

$$\partial^\gamma f(x) = \int_{\mathbb{R}^d} \xi^\gamma \hat{f}(\xi) e^{ix\xi}\, d\xi, \quad x \in \mathbb{R}^d.$$

By Plancherel's theorem, for $f \in \mathcal{S}$ and $\phi \in \mathcal{S}_j$, the integration-by-parts formula

$$\langle f, \phi\rangle = \sum_{|\gamma|=j} \langle \partial^\gamma f, \partial^{-\gamma}\phi\rangle$$

holds. The symbol D^σ stands for the fractional differentiation operator, namely the Fourier multiplier $m(\xi) = |\xi|^\sigma$ for any σ real. We will also utilize the Japanese bracket $\langle x\rangle = 1 + |x|$ and the fact that it is equivalent to $\max\{1, |x|\}$ and for $x = (x_1, \dots, x_m) \in (\mathbb{R}^d)^m$ we have $\langle x\rangle \sim \max\{1, |x_1|, \dots, |x_m|\}$.

Definition 2.1. Let D be a nonnegative integer. We say $\phi \in \mathcal{S}_{2D}$ is a mother wavelet if ϕ is supported in $B\big(0, \frac{1}{2}\big)$, admissible, and for all $0 \le |\alpha| \le D$

$$\partial^{\pm\alpha}\phi \in \mathcal{S}_D.$$

If a function is radial, Schwartz, and mean-zero, then it only needs to be normalized so that (2-1) holds. So the admissibility condition can more or less be dropped from the definition. Such wavelets can be constructed as $\Delta^{4D}\Phi$, where $\Phi \in C_0^\infty\big(B\big(0, \frac{1}{2}\big)\big)$.

Crucial to our program are the functions which behave like wavelets in their scale and decay, similar to the so-called molecules of Frazier, Jawerth, and Weiss [Frazier et al. 1988; 1991]. Accordingly, introduce the norm

$$\|\varphi\|_{\star,\eta,\delta} := \sup_{x\in\mathbb{R}^d} \langle x\rangle^{d+\eta}|\varphi(x)| + \sup_{x\in\mathbb{R}^d, 0\le|h|\le 1} \langle x\rangle^{d+\eta} \frac{|\varphi(x+h) - \varphi(x)|}{|h|^\delta}. \tag{2-3}$$

Definition 2.2. The wavelet class $\Psi_z^{k,\delta;1}$ is defined by

$$\{\varphi \in \mathcal{C}^k(\mathbb{R}^d) : t^{|\gamma|}\|(\mathsf{Sy}_z)^{-1}\partial^\gamma\varphi\|_{\star,|\gamma|,\delta} \le 1 \text{ for } 0 \le |\gamma| \le k\}$$

and its cancellative subclass is given by

$$\Psi_z^{k,\delta;0} = \left\{\varphi \in \Psi_z^{k,\delta;1} : \int x^\alpha \phi(x) = 0 \ \text{ for } \ 0 \le |\alpha| \le k\right\}.$$

Notice that $\phi_z \in \Psi_z^{k,1;1}$, $\psi_z \in \Psi_z^{k,1;0}$ whenever $\phi \in \mathcal{S}$, $\psi \in \mathcal{S}_k$ are suitably normalized.

The study of bilinear operators requires a suitable tensor-type class $\Psi_z^{k,\delta;1,1}$. First, for functions in $L^1_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^d)$, write Sy_z^j for the action of $z \in Z^d$ on the j-th copy of $\mathbb{R}^d$, $j = 1, 2$. Then, $\Psi_z^{k,\delta;1,1}$ is the collection of all $\phi \in \mathcal{C}^k(\mathbb{R}^{2d})$ which satisfy the estimates

$$t^{|\gamma|}\|(\mathsf{Sy}_z^1\mathsf{Sy}_z^2)^{-1}\partial^\gamma\phi\|_{\star,k,\delta} \le 1$$

for all $\gamma \in \mathbb{N}^{2d}$, $0 \le |\gamma| \le k$. The norm is defined by (2-3) but replacing $\mathbb{R}^d$ with $\mathbb{R}^{2d}$.

This norm is larger than purely tensoring the norm $\|\cdot\|_{\star,\eta,\delta}$ which, in fact, is not enough for the L^p boundedness of our intrinsic form in the full multilinear range of exponents (see Section 5). To

demonstrate the usefulness of this class, we introduce the cancellative intrinsic forms which will be used in our representation. For $\nu_z \in \Psi_z^{0,\delta;1,1}$ and $\phi \in \mathcal{C}_0^\infty(B(0,1))$, define

$$U(f,g,h) = \int_{Z^d} \langle f \otimes g, \nu_z \rangle \langle h, \phi_z \rangle \, \mathrm{d}\mu(z).$$

The form U can be written as $U(f,g,h) = \langle K, f \otimes g \otimes h \rangle$ with kernel

$$K(x_0,x_1,x_2) = \int_{Z^d} \nu_z(x_1,x_2)\phi_z(x_0) \, \mathrm{d}\mu(z).$$

For the size estimate on K fix x_0, x_1, x_2 and divide the integration in t into the two regions $L = \{2t < \max_j |x_j - x_0|\}$ and $L^c = (0,\infty)\backslash L$. Using the fact that

$$\nu_z(x_1,x_2) \le \frac{1}{t^{2d}\max\{1, |x_1-w|/t, |x_2-w|/t\}^{2d+\delta}} = \frac{t^\delta}{\max_{j\neq 0}\{t, |x_j-w|\}}$$

and that ϕ is supported in $B(0,1)$,

$$\begin{aligned} |K(x_0,x_1,x_2)| &\le \int_{t\in L} + \int_{t\in L^c}\int_{w\in B(x_0,t)} \frac{t^\delta}{\max_{j\neq 0}\{t,|x_j-w|\}^{2d+\delta}} t^{-d}\phi\left(\frac{x_0-w}{t}\right)\frac{\mathrm{d}w\,\mathrm{d}t}{t} \\ &\lesssim (\max_j |x_j - x_0|)^{-2d}. \end{aligned}$$

A similar Hölder estimate can also be proved. More precisely, to use terminology of Section 3A below, U is a $(0,\delta)$ singular integral (SI) form. If $\nu_z \in \Psi_z^{k+\delta;1,1}$ then U is a (k,δ) SI form.

While admissible wavelets themselves satisfy the remarkable orthogonality properties which yield the Calderón reproducing formula (2-2), the elements of the wavelet class satisfy the following almost orthogonality estimate.

Lemma 2.3 [Di Plinio et al. 2022, Lemma 2.3]. *Let $0 < \eta < \delta \le 1$, $0 \le k \le D$, and $s \ge t$. Set $z = (w,t)$ and $\zeta = (v,s)$. Then*

$$\sup_{\nu_z \in \Psi_z^{k,\delta;0}} \sup_{\theta_\zeta \in \Psi_\zeta^{k,\delta;1}} |\langle \nu_z, \theta_\zeta \rangle| \lesssim \frac{t^{k+\eta}}{\max\{s, |v-w|\}^{d+k+\eta}}$$

and, for a mother wavelet ϕ,

$$\sup_{\theta_\zeta \in \Psi_\zeta^{k,\delta;1}} |\langle \phi_z, \theta_\zeta \rangle| \lesssim \frac{t^{k+\delta}}{\max\{s, |v-w|\}^{d+k+\delta}}.$$

In the sequel, we will often denote elements of $\Psi_z^{k,\delta;i}$ or $\Psi_z^{k,\delta;i,j}$ by ν_z or θ_z. This means only that the function ν_z is associated to a point $z \in Z^d$, *not* that ν_z is given by the group action $\mathrm{Sy}_z \nu$ for some function ν. Whether the subscript denotes group action or not will be clear from the context; e.g., if ϕ is first introduced and then ϕ_z, of course ϕ_z is the group action. If ν_z is introduced as an element of $\Psi_z^{k,\delta;1}$ then it is a function associated to z. There is even less ambiguity since ϕ_z is of course a function associated to z.

2A. ***Averaging of wavelets.*** We will use the following *wavelet averaging lemma* to diagonalize the wavelet shifts (the continuous analogue of the Haar shifts).

Lemma 2.4. *Let $0 < \eta < \delta$ and $H : \mathbb{R}^{3d} \times (0,\infty)^2 \to \mathbb{C}$ be such that*

$$|H(u, v, w, s, t)| \lesssim \frac{t^\delta}{\max\{s, |u-w|, |v-w|\}^{2d+\delta}}.$$

If $\psi, \phi \in \mathcal{S}$ then

$$\nu_{w,t}(x, y) = \int_t^\infty \int_{\mathbb{R}^{2d}} H(u, v, w, s, t)\psi_{u,s}(x)\phi_{v,s}(y)\,\frac{\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}s}{s}$$

satisfies

$$\|(\mathsf{Sy}_z^1\mathsf{Sy}_z^2)^{-1}\nu_z\|_{\star,\eta,1} \lesssim 1.$$

Proof. Let us fix w, t and simply write ν for $\nu_{w,t}$. Make the changes of variable $\alpha = (u-w)/t$, $\alpha' = (v-w)/t$ and $\beta = s/t$. In this way,

$$(\mathsf{Sy}_{w,t}^1\mathsf{Sy}_{w,t}^2)^{-1}\nu(x, y) = \int_1^\infty \int_{\mathbb{R}^{2d}} H'(\alpha, \alpha', \beta)\psi'(x)\phi'(y)\,\frac{\mathrm{d}\alpha\,\mathrm{d}\alpha'\,\mathrm{d}\beta}{\beta}, \tag{2-4}$$

where $\psi' = \mathsf{Sy}_{\alpha,\beta}\mathsf{Sy}_{u,s}^{-1}\psi_{u,s}$, $\phi' = \mathsf{Sy}_{\alpha',\beta}\mathsf{Sy}_{v,s}^{-1}\phi_{v,s}$ and $H' \lesssim \max\{\beta, |\alpha|, |\alpha'|\}^{-(2d+\delta)}$. We have suppressed the dependence of ψ' and ϕ' on α, α', β.

We first get the size estimate on ν. Since ϕ' and ψ' are not assumed to have compact support, we decompose into annuli and divide the scale parameter β accordingly:

$$L_{j,i} = \{\beta > 1 : \beta \le \max\{2^{-(j+2)}|x|, 2^{-(i+2)}|y|\}\} \quad \text{and} \quad L_{j,i}^c = (1,\infty)\backslash L_{j,i}.$$

Then, for each $\beta > 1$, define the annuli $A_j(x, \beta) = B(x, 2^{j+1}\beta)\backslash B(x, 2^j\beta)$ for $j \ge 1$ and $A_0(x, \beta) = B(x, 2\beta)$. Let $\alpha \in A_j(x, \beta)$ and $\alpha' \in A_i(y, \beta)$. Then

$$\max\{\beta, |\alpha|, |\alpha'|\} \ge \begin{cases} \frac{1}{2}\max\{|x|, |y|\} & \text{if } \beta \in L_{j,i}, \\ \beta & \text{if } \beta \in L_{j,i}^c. \end{cases}$$

We obtain, for $j \ge 1$, that if $\alpha \in A_j(x, \beta)$, $\beta \in L_{j,i}$ then for any $r > 0$

$$|\psi'(x)| \le \beta^{-d}\left(1 + \frac{|x-\alpha|}{\beta}\right)^{-d-r} \le \beta^{-d}(1+2^j)^{-d-r} \le 2^{-jr}(\beta 2^j)^{-d}$$

and similarly for ϕ' and for $\alpha' \in A_i(y, \beta)$. The estimate also holds when i or j is 0 simply because ϕ and ψ are bounded. Thus, for each $j, i \ge 0$,

$$\int_{A_j(x,\beta)}\int_{A_i(y,\beta)} \frac{|\psi'(x)\phi'(y)|\,\mathrm{d}\alpha'\,\mathrm{d}\alpha}{\max\{\beta, |\alpha|, |\alpha'|\}^{2d+\delta}} \lesssim 2^{-r(j+i)} \begin{cases} \max\{|x|, |y|\}^{-(2d+\delta)}, & \beta \in L_{j,i}, \\ \beta^{-(2d+\delta)}, & \beta \in L_{j,i}^c. \end{cases}$$

Therefore,

$$\begin{aligned} |(\mathsf{Sy}_{w,t}^1\mathsf{Sy}_{w,t}^2)^{-1}\nu(x, y)| &\lesssim \sum_{i,j=0}^\infty 2^{-r(j+i)}\left[\int_{L_{j,i}} \max\{|x|, |y|\}^{-(2d+\delta)}\,\frac{\mathrm{d}\beta}{\beta} + \int_{L_{j,i}^c} \beta^{-(2d+\delta)}\,\frac{\mathrm{d}\beta}{\beta}\right] \\ &\lesssim \sum_{i,j=0}^\infty 2^{-r(j+i)}\left[\frac{\log\max\{2^{-j}|x|, 2^{-i}|y|\}}{\max\{|x|, |y|\}^{2d+\delta}} + \max\{2^{-j}|x|, 2^{-i}|y|\}^{-(2d+\delta)}\right] \\ &\lesssim \max\{|x|, |y|\}^{-(2d+\eta)} + \sum_{i,j=0}^\infty 2^{-r(j+i)}2^{\max(j,i)(2d+\delta)}\max\{|x|, |y|\}^{-(2d+\delta)}. \end{aligned}$$

Picking $r > 2d + \delta$ guarantees convergence of the sum and concludes the proof of the size estimate. For the Hölder estimate, let $h = (h_1, h_2) \in \mathbb{R}^d \times \mathbb{R}^d$. By (2-4) above,

$$|(\mathsf{Sy}^1_{w,t}\mathsf{Sy}^2_{w,t})^{-1}\nu(x,y) - (\mathsf{Sy}^1_{w,t}\mathsf{Sy}^2_{w,t})^{-1}\nu(x+h_1,y+h_2)|$$
$$\leq \int_1^\infty \int_{\mathbb{R}^{2d}} \frac{1}{\max\{\beta,|\alpha|,|\alpha'|\}^{2d+\delta}} \big(|\phi'(y)||\psi'(x)-\psi'(x+h_1)| + |\psi'(x+h_1)||\phi'(y)-\phi'(y+h_2)|\big) \frac{\mathrm{d}\alpha\,\mathrm{d}\alpha'\,\mathrm{d}\beta}{\beta}.$$

We will only handle the second term as the first is similar. First, assuming that $|h_1|, |h_2| < 1$, we obtain an estimate analogous to that above: if $\alpha \in A_j(x,\beta)$ and $\alpha' \in A_i(y,\beta)$, then, for any $r > 0$,

$$|\psi'(x+h_1)||\phi'(y) - \phi'(y+h_2)| \lesssim 2^{-r(j+i)} \frac{|h_2|}{(\beta 2^j)^d (\beta 2^i)^d}.$$

Following the remaining steps as above proves the result. If $|h_1|$ or $|h_2|$ are larger than 1, the Hölder estimate follows from the size estimate. □

An immediate corollary follows.

Proposition 2.5. *Let $k \in \mathbb{N}$, $0 < \eta < \delta \leq 1$, and $H : \mathbb{R}^{3d} \times (0,\infty)^2 \to \mathbb{C}$ be such that*

$$|H(u,v,w,s,t)| \lesssim \frac{t^{k+\delta}}{\max\{s,|u-w|,|v-w|\}^{2d+k+\delta}}.$$

Let $\psi, \phi \in \mathcal{S}$ and define

$$\nu_{w,t}(x,y) = \int_t^\infty \int_{\mathbb{R}^{2d}} H(u,v,w,s,t)\psi_{u,s}(x)\phi_{v,s}(y) \frac{\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}s}{s}. \tag{2-5}$$

Then, there exists $C > 0$ such that $\nu_{w,t} \in C\Psi^{k,\eta;1,1}_{w,t}$. If moreover $\phi \in \mathcal{S}_k$, then $\nu_{w,t} \in C\Psi^{k,\eta;1,0}_{w,t}$ and $t^{-|\kappa|}\partial_y^{-\kappa}\nu_{w,t} \in \Psi^{0,\eta;1,0}_{w,t}$ for $|\kappa| \leq k$.

Proof. Applying $t^{|\gamma|}\partial^\gamma$ to the formula (2-5) for $|\gamma| \leq k$, one can see that the symbol

$$H(u,v,w,s,t)\left(\frac{t}{s}\right)^{|\gamma|}$$

satisfies the condition of Lemma 2.4 with δ replaced by $k+\delta$, which proves the first statement. Similarly, if $\phi \in \mathcal{S}_k$, then $\partial^{-\kappa}\phi \in \mathcal{S}$ for $|\kappa| \leq k$ and the symbol obtained by applying $t^{-|\kappa|}\partial^{-\kappa}$ to (2-5) satisfies

$$H(u,v,w,s,t)\left(\frac{s}{t}\right)^{|\kappa|} \lesssim \frac{t^{k-|\kappa|+\delta}}{\max\{s,|u-w|,|v-w|\}^{2d+k-|\kappa|+\delta}} \lesssim \frac{t^\delta}{\max\{s,|u-w|,|v-w|\}^{2d+\delta}}. \qquad \square$$

The next lemma is similar, and will be used to convert a portion of the paraproduct into a wavelet form.

Lemma 2.6. *Let $k \in \mathbb{N}$, $0 < \eta < \delta \leq 1$, and $G : \mathbb{R}^{3d} \times (0,\infty)^2 \to \mathbb{C}$ satisfy*

$$|G(u,v,w,s,t)| \lesssim \frac{s^{k+\delta}}{t^{2d+k+\delta}}.$$

Then, for $\|\psi\|_{\star,d+\delta,\delta}, \|\phi\|_{\star,d+\delta,\delta} \leq 1$, there exists $C > 0$ such that

$$\theta_{w,t}(x,y) := \int_0^{3t} \int_{|u-w|,|v-w|\leq 9t} G(u,v,w,s,t)\psi_{u,s}(x)\phi_{v,s}(y) \frac{\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}s}{s} \in C\Psi^{k,\delta;1,1}_{w,t}.$$

Moreover, if ψ (or ϕ) has vanishing moments up to k, then

$$\theta_{w,t} \in C\Psi_{w,t}^{k,\delta;0,1} (\text{or } C\Psi_{w,t}^{k,\delta;1,0}) \quad \text{and} \quad t^{-|\kappa|}\partial_x^{-\kappa}\theta_{w,t} \in C\Psi_{w,t}^{0,\delta;0,1} \quad (\text{or } t^{-|\kappa|}\partial_y^{-\kappa}\theta_{w,t} \in C\Psi_{w,t}^{0,\delta;1,0})$$

for every $|\kappa| \le k$.

Proof. To check the size estimate of $\theta = \theta_{w,t}$, perform the change of variable as in the previous lemma. Assume $|x| > |y|$ and $|x| > 18$. In this case,

$$|x-\alpha| \ge |x| - |\alpha| \ge |x| - 9 \ge \tfrac{1}{2}|x|$$

so that

$$|\psi'(x)| \le \beta^{-d}\left(1 + \frac{|x-\alpha|}{\beta}\right)^{-2d-\delta} \le \beta^{d+\delta}(\beta + |x|)^{-(2d+\delta)}.$$

Thus,

$$\begin{aligned} |(\mathsf{Sy}_{w,t}^1 \mathsf{Sy}_{w,t}^2)^{-1}\partial^{\pm\kappa}\theta(x,y)| &\le t^{\mp|\kappa|} \int_{\alpha\in B(0,9),\, \alpha'\in B(0,9),\, 0<\beta\le 3} \frac{\beta^{k\mp|\kappa|+2\delta}}{(\beta+|x|)^{2d+\delta}} \frac{\mathrm{d}\alpha\,\mathrm{d}\alpha'\,\mathrm{d}\beta}{\beta} \\ &\lesssim \frac{1}{t^{\pm|\kappa|}|x|^{2d+\delta}}. \end{aligned}$$

Symmetry yields the case when $|y| > |x|$. When both $|x|$ and $|y|$ are smaller than 18, we can check the original formula to see

$$\theta(x,y) \lesssim \int_0^{3t} s^{k-|\kappa|+\delta} t^{-(k+\delta)} \frac{\mathrm{d}s}{s} \|\psi\|_{L^1}\|\phi\|_{L^1} \lesssim t^{\mp|\kappa|}.$$

The same method yields the Hölder estimate. □

2B. ***High-low cancellation of wavelets.*** We return to the Calderón formula (2-2) from the Introduction. In general, it is difficult to analyze operators acting on many different scales at once. It will be helpful in the future to place two functions on the same scale and vary the third. To do so, we use the fact that in the superposition of many wavelets, the smallest scale (highest frequency) dominates. We state this precisely in the following lemma.

Lemma 2.7. *Let ϕ be a radial mother wavelet and $m \ge 0$. There exist functions ψ^j, $j = 1,2,3,4$, satisfying*:

(i) $\operatorname{supp}\psi^j \subset B(0,1)$.

(ii) $\psi^1, \psi^3 \in C^m$.

(iii) $\psi^2, \psi^4 \in \mathcal{S}_D$.

(iv) *For any $s > 0$ and $f \in L^2(\mathbb{R}^d)$,*

$$\int_{r\ge s}\int_{u\in\mathbb{R}^d} \langle f, \phi_{u,r}\rangle \phi_{u,r} \frac{\mathrm{d}u\,\mathrm{d}r}{r} = \int_{\mathbb{R}^d} \langle f, \psi_{u,s}^1\rangle \psi_{u,s}^2 + \langle f, \psi_{u,s}^3\rangle \psi_{u,s}^4 \,\mathrm{d}u.$$

Proof. Define

$$\Phi(w) = \int_1^\infty \phi * \phi\left(\frac{w}{r}\right) \frac{\mathrm{d}r}{r^{d+1}}.$$

Φ retains the vanishing moments properties of ϕ. Since ϕ is radial, so is $\phi * \phi$ and thus Φ. Changing the variables and denoting by p the radial function $p(|x|) = \phi * \phi(x)$ (supported in $[0, 1)$), we can rewrite

$$\Phi(w) = |w|^{-d} \int_0^{|w|} p(\tau)\tau^{d-1}\,\mathrm{d}\tau.$$

Thus, if $w \geq 1$,

$$\Phi(w) = |w|^{-d} \int_0^1 p(\tau)\tau^{d-1}\,\mathrm{d}\tau = |w|^{-d} \int_{\mathbb{R}^d} \phi * \phi = 0.$$

In particular, Φ is supported in $B(0, 1)$. Next, a few changes of variables yield

$$\int_s^\infty \int_{\mathbb{R}^d} \langle f, \phi_{u,r}\rangle \phi_{u,r}(x) \frac{\mathrm{d}u\,\mathrm{d}r}{r} = \langle f, \Phi_{x,s}\rangle$$

for any $f \in L^2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. We will be done if we can decompose Φ as $\Phi = \psi^1 * \psi^2 + \psi^3 * \psi^4$. In that case, changing the variables again,

$$\langle f, \Phi_{x,s}\rangle = \int_{\mathbb{R}^d} \langle f, \psi^1_{u,s}\rangle \psi^2_{u,s}(x) + \langle f, \psi^3_{u,s}\rangle \psi^4_{u,s}(x)\,\mathrm{d}u.$$

Decomposing Φ is difficult if we want the functions to remain Schwartz and compactly supported [Yulmukhametov 1999]. Obviously $\Phi = \delta_0 * \Phi$. We decompose the delta distribution as follows. Set

$$H(x) = \begin{cases} \dfrac{1}{(m+1)!^d} x_1^{m+1} x_2^{m+1} \cdots x_d^{m+1} & x_i \geq 0, \\ 0 & \text{else.} \end{cases}$$

Setting $D = (\partial/\partial x_1)(\partial/\partial x_2)\cdots(\partial/\partial x_d)$, integrating by parts yields $H * D^{m+2}\Phi = \Phi$. Now $H \in C^m$ but is not compactly supported. This can be fixed by taking $g \in C^\infty$ such that $g = H$ for $|x| \geq \frac{1}{2}$. Then, $G := D^{m+2}g \in C_0^\infty(B(0, 1))$. Distributionally, $\delta = D^{m+2}(H - g) + G := D^{m+2}F + G$. Therefore,

$$\Phi = F * D^{m+2}\Phi + G * \Phi =: \psi^1 * \psi^2 + \psi^3 * \psi^4. \qquad \square$$

This allows us to obtain the single-scale variant of the bilinear Calderón formula.

Lemma 2.8.

$$f \otimes g = \int_0^\infty \int_{\mathbb{R}^{2d}} \langle f, \psi^1_{u,s}\rangle \langle g, \phi_{v,s}\rangle \psi^2_{u,s} \otimes \phi_{v,s} + \langle f, \psi^3_{u,s}\rangle \langle g, \phi_{v,s}\rangle \psi^4_{u,s} \otimes \phi_{v,s}$$
$$+ \langle f, \phi_{u,s}\rangle \langle g, \psi^1_{v,s}\rangle \phi_{u,s} \otimes \psi^2_{v,s} + \langle f, \phi_{u,s}\rangle \langle g, \psi^3_{v,s}\rangle \phi_{u,s} \otimes \psi^4_{v,s}\, \frac{\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}s}{s}.$$

Proof. Use the Calderón formula (2-2) on f and g to obtain

$$f \otimes g = \int_{Z^d} \int_{Z^d} \langle f, \phi_{u,r}\rangle \langle g, \phi_{v,s}\rangle \phi_{u,r} \otimes \phi_{v,s} \frac{\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}v\,\mathrm{d}s}{rs}.$$

Split the integral into $r \geq s$ and $s > r$. On the first one, apply Lemma 2.7 with f and on the second, apply it to g. $\square$

3. Representation formula

3A. ***Singular integrals.*** Let $\vec{1}_d = (1, \dots, 1) \in \mathbb{R}^d$. Given $k \in \mathbb{N}$, a function $K \in L^1_{\mathrm{loc}}(\mathbb{R}^{3d} \setminus \mathbb{R}\mathbf{1}_{3d})$ is a $(\vec{\ell}, \delta)$ SI (singular integral) kernel if there exist $C, \delta > 0$ such that, for all $0 \le |\kappa| \le k$,

$$|\partial_{x_i}^{\kappa} K(x_0, x_1, x_2)| \le \frac{C}{\big(\sum_{j \ne i} |x_i - x_j|\big)^{2d+|\kappa|}},$$

$$|\partial_{x_i}^{\kappa} \Delta_h^i K(x_0, x_1, x_2)| \le \frac{C|h|^{\delta}}{\big(\sum_{j \ne i} |x_i - x_j|\big)^{2d+|\kappa|+\delta}}.$$

Δ_h^i denotes the difference operator in the i-th position. We say Λ is a (k, δ) trilinear SI form if

$$\int K(x_0, x_1, x_2) f(x_1) g(x_2) h(x_0)\, \mathrm{d}x = \Lambda(f, g, h)$$

for all $f, g, h \in \mathcal{S}$ with $\operatorname{supp} f \cap \operatorname{supp} g \cap \operatorname{supp} h = \varnothing$ and for a (k, δ) SI kernel K. Notice that a (k, δ) SI form is a (k', δ') form for any $k' \le k$ and $\delta' \le \delta$.

3B. ***Calderón–Zygmund forms.*** Our representation formula will be built using the following intrinsic singular integral forms.

3B1. *Wavelet forms.*

Definition 3.1. A trilinear form U is called a (k, δ)-smooth wavelet form if for each $z \in Z^d$ there exists $\nu_z \in \Psi_z^{k,\delta;1,0}$ such that

$$U(\pi(f, g, h)) = \int_{Z^d} \langle f \otimes g, \nu_z \rangle \langle h, \phi_z \rangle \, \mathrm{d}\mu(z)$$

for some permutation $\pi \in S^3$ and a mother wavelet ϕ.

3B2. *Paraproducts.* Let $\{\theta_z^{\gamma} \in \Psi_z^{D,\delta;1}\}_{z \in Z^d}$ be a γ-family, which means

$$\int x^{\beta} \theta_z^{\gamma}(x)\, \mathrm{d}x = t^{|\beta|} \delta_{\beta,\gamma}$$

for each $|\beta| \le |\gamma| \le D$. These can be constructed by taking a single function ϑ^{γ}, smooth and compactly supported such that

$$\int_{\mathbb{R}^d} \vartheta^{\gamma}(x) x^{\beta}\, \mathrm{d}x = \delta_{\beta,\gamma}, \quad 0 \le \beta \le \gamma,$$

and then acting on ϑ^{γ} with the affine group Z^d, yielding ϑ_z^{γ}. Such functions ϑ^{γ} do indeed exist; see [Alpert 1993; Rahm et al. 2021].

Given a function $b \in \mathrm{BMO}$ and multi-index $\gamma \in \mathbb{N}^{2d}$, define the γ-order paraproduct form

$$\Pi_{b,\gamma}(f, g, h) = \int_{Z^d} \langle b, (\partial^{-\gamma_1-\gamma_2}\phi)_z \rangle \langle f, \vartheta_z^{\gamma_1} \rangle \langle g, \vartheta_z^{\gamma_2} \rangle \langle h, \phi_z \rangle \, \mathrm{d}\mu(z), \tag{3-1}$$

where ϕ is a mother wavelet and $\vartheta_z^{\gamma_\ell}$ are compactly supported γ_ℓ families. $\Pi_{b,\gamma}$ is a (M, δ) SI form for any $M > 0$ up to the smoothness of ϑ^{γ} and ϕ and any $0 < \delta \le 1$. This can be verified using the same

reasoning as in the discussion after Definition 2.2 of the wavelet classes, only it is simpler since the wavelets are compactly supported.

We will use the partial ordering on multi-indices $\vec{j}, \vec{k} \in \mathbb{N}^m$: $\vec{j} = (j_1, j_2, \dots, j_m) \le \vec{k} = (k_1, k_2, \dots, k_m)$ if $j_\ell \le k_\ell$ and $\vec{j} < \vec{k}$ if $\vec{j} \le \vec{k}$ but $\vec{j} \ne \vec{k}$. In this way, for all $\kappa \le \gamma$ and ϕ with vanishing moments up to $|\gamma|$,

$$\Pi_{b,\gamma}(x^{\kappa_1}, y^{\kappa_2}, \phi) = \delta_{\kappa,\gamma}\langle b, \partial^{-\gamma_1-\gamma_2}\phi\rangle,$$
$$\Pi^{i*}_{b,\gamma}(x^{\kappa_1}, y^{\kappa_2}, \phi) = 0,$$

where $\Pi^{*1}_b(f,g,h) = \Pi_b(h,g,f)$ and $\Pi^{*2}_b(f,g,h) = \Pi_b(f,h,g)$. In general, this is difficult to compute for $\kappa > \gamma$ (see Section 7 below). Now for $\gamma \in \mathbb{N}^{2d}$ we will iteratively define the γ-th order paraproducts of a form Λ.

Recall $\mathcal{S}_j = \{\psi \in \mathcal{S} : \int x^\gamma \psi(x)\,dx = 0 \text{ for } |\gamma| \le j\}$. When $\gamma = 0$, we say a $(0,\delta)$ SI form Λ has 0-th order paraproducts if there exist BMO functions b^i_0, $i = 0, 1, 2$, such that, for all $\psi \in \mathcal{S}_0$,

$$\Lambda(1,1,\psi) = \langle b^0_0, \psi\rangle, \quad \Lambda(\psi,1,1) = \langle b^1_0, \psi\rangle, \quad \Lambda(1,\psi,1) = \langle b^2_0, \psi\rangle.$$

This is the standard bilinear $T(1,1)$ condition [Christ and Journé 1987; Li et al. 2019]. Now, for $(k_1,k_2) > 0$, we define the (k_1,k_2)-th order paraproducts inductively. Suppose Λ has paraproducts b^i_γ for all $(|\gamma_1|,|\gamma_2|) < (k_1,k_2)$. Then, we say Λ has (k_1,k_2)-th order paraproducts if for each $|\gamma_1| = k_1$, $|\gamma_2| = k_2$ there exist $b^i_\gamma \in \mathrm{BMO}$ such that, for all $\psi \in \mathcal{S}_{k_1+k_2}$,

$$\Lambda_{k_1,k_2} := \Lambda - \sum_{i=0}^{2} \sum_{(|\kappa_1|,|\kappa_2|)<(k_1,k_2)} \Pi^{i*}_{b^i_\kappa,\kappa}$$

satisfies

$$\Lambda_{k_1,k_2}(x^{\gamma_1}, y^{\gamma_2}, \psi) = \langle b^0_\gamma, \partial^{-\gamma_1-\gamma_2}\psi\rangle,$$
$$\Lambda_{k_1,k_2}(\psi, y^{\gamma_2}, x^{\gamma_1}) = \langle b^1_\gamma, \partial^{-\gamma_1-\gamma_2}\psi\rangle,$$
$$\Lambda_{k_1,k_2}(x^{\gamma_1}, \psi, y^{\gamma_2}) = \langle b^2_\gamma, \partial^{-\gamma_1-\gamma_2}\psi\rangle.$$

Under this definition, one can verify by induction that Λ_{k_1,k_2} has vanishing paraproducts of all orders $< (k_1,k_2)$. The action of (k,δ) SI forms on polynomials of degree (k_1,k_2) with $k_1+k_2 \le k$ can be defined as elements of the dual space of $\mathcal{S}_{k_1+k_2}$, see [Frazier et al. 1988; Bényi 2003].

Definition 3.2. Let $k_1+k_2 \le k$. A (k,δ) SI form Λ is called a (k_1,k_2,δ) Calderón–Zygmund (CZ) form if it has paraproducts up to order (k_1,k_2) and satisfies the weak boundedness property (WBP), which means

$$t^{2d}|\Lambda(\phi_z, \psi_z, \vartheta_z)| \le C \tag{3-2}$$

for all $\phi_z, \psi_z, \vartheta_z \in \Psi^{0,\delta;1,1}_z$ supported in $B(w,t)$ and $(w,t) = z \in Z^d$. We also say T is a (k_1,k_2,δ) Calderón–Zygmund operator (CZO) if $\Lambda(f,g,h) = \langle T(f,g), h\rangle$ is a (k_1,k_2,δ) CZ form.

3C. *Smooth representation theorem.*

Theorem A. *Let Λ be a (k_1,k_2,δ) CZ form and $\eta < \delta$. There exists (j,η)-smooth wavelet forms U^i_j for $j = \min\{k_1,k_2\}, \dots, k_1+k_2$, $i = 1,\dots,6$, and paraproduct forms $\Pi^{i*}_{b^i_\gamma,\gamma}$, $i = 0,1,2$, such that*

$$\Lambda(f,g,h) = \sum_{i=1}^{6} \sum_{j=\min\{k_1,k_2\}}^{k_1+k_2} U^i_j(f,g,h) + \sum_{i=0}^{2} \sum_{|\gamma_\ell|\le|k_\ell|} \Pi^{i*}_{b^{i*}_\gamma,\gamma}(f,g,h).$$

The region of interest $Z(w,t) := \mathbb{R}^{2d} \times (t,\infty)$ will be partitioned into the following regions (far, near, high-low):

$$\begin{aligned} F(w,t) &:= \{\max\{|u-w|,|v-w|\} \geq 3s,\ s \geq t\},\\ S(w,t) &:= \{\max\{|u-w|,|v-w|\} \leq 3s,\ s \leq 3t\},\\ A(w,t) &:= \{\max\{|u-w|,|v-w|\} \leq 3s,\ s \geq 3t\}. \end{aligned} \tag{3-3}$$

We will also use the region $I(w,t) := \{\max\{|u-w|,|v-w|\} \leq 3\max\{s,t\},\ s \leq 3t\}$.

3C1. *Kernel estimates.*

Lemma 3.3. *Let Λ be a (k_1,k_2,δ) CZ form, ϕ be a mother wavelet, and $\psi \in \mathcal{C}_0^\infty(B(0,1))$. For each $0 \leq |\gamma| \leq \max\{k_1,k_2\}$, let $\vartheta_{w,t}^\gamma$ be a γ-family. Define*

$$P_k(x) = \sum_{0\leq|\gamma|\leq k} \langle \psi_{u,s}, \vartheta_{w,t}^\gamma\rangle \left(\frac{x-w}{t}\right)^\gamma$$

and $\widetilde{P}_k$, replacing $\psi_{u,s}$ with $\phi_{v,s}$. For $s \geq t$, define

$$\begin{aligned} &\Upsilon(u,v,w,s,t)\\ &\quad = \Lambda(\psi_{u,s},\phi_{v,s},\phi_{w,t}) - 1_{A(w,t)}(u,v,s)[\Lambda(P_{k_1},\widetilde{P}_{k_2},\phi_{w,t})\\ &\qquad\qquad + \Lambda(P_{k_1-1},\phi_{v,s}-\widetilde{P}_{k_2},\phi_{w,t}) - \Lambda(\psi_{u,s}-P_{k_1},\widetilde{P}_{k_2-1},\phi_{w,t})]. \end{aligned} \tag{3-4}$$

Then, for any $\eta < \delta$,

$$|\Upsilon(u,v,w,s,t)| \lesssim_\eta \frac{t^{k_1+k_2+\eta}}{\max\{s,|u-w|,|v-w|\}^{2d+k_1+k_2+\eta}}.$$

Proof. Region 1: When $(u,v,s) \in F(w,t)$, $\max\{|v-w|,|u-w|\} \geq 3s$. Integrating by parts k_1+k_2 times, rewrite $\Lambda(\psi,\phi,\phi) = \sum_{|\gamma|=k_1+k_2} \langle \partial_{x_0}^\gamma K,\ \psi\otimes\phi\otimes\partial^{-\gamma}(\phi)\rangle$. Use the Hölder estimate of the kernel and the fact that $\partial^{-\gamma}(\phi_{w,t}) = t^{|\gamma|}(\partial^{-\gamma}\phi)_{w,t}$ to get

$$t^{k_1+k_2}\left|\int_{x_1\in B(u,s)}\int_{x_2\in B(v,s)} \frac{t^\delta \psi_{u,s}(x_1)\phi_{v,s}(x_2)\,dx_1\,dx_2}{(|x_1-w|+|v-w|)^{2d+k_1+k_2+\delta}}\right| \leq \frac{t^{k_1+k_2+\delta}}{\max\{|u-w|,|v-w|\}^{2d+k_1+k_2+\delta}}.$$

It is also important here that $\partial^{-\gamma}\phi$ is still mean zero.

Region 2: In the region $S(w,t)$, appealing to the WBP gives the estimate of t^{-2d}.

Region 3: In the final region, $A(w,t)$, $|u-w|,|v-w| \leq 3s$ and $s \geq 3t$. Here we rewrite

$$\begin{aligned} &\Lambda(\psi_{u,s},\phi_{v,s},\phi_{w,t}) - \Lambda(P_{k_1},\widetilde{P}_{k_2},\phi_{w,t})\\ &\qquad = \Lambda(\psi_{u,s}-P_{k_1},\phi_{v,s}-\widetilde{P}_{k_2},\phi_{w,t}) + \Lambda(P_{k_1},\phi_{v,s}-\widetilde{P}_{k_2},\phi_{w,t}) + \Lambda(\psi_{u,s}-P_{k_1},\widetilde{P}_{k_2},\phi_{w,t}). \end{aligned} \tag{3-5}$$

P and $\widetilde{P}$ satisfy the estimate

$$|\psi_{u,s}(x)-P_k(x)|,\ |\phi_{v,s}(x)-\widetilde{P}_k(x)| \leq \frac{1}{s^d}\left(\frac{|x-w|}{s}\right)^k \min\left\{1,\frac{\max\{|x-w|,t\}}{s}\right\}.$$

See [Di Plinio et al. 2022, Lemma 3.1], where this is derived using Taylor polynomials in conjunction with Lemma 2.3.

Let α be a cutoff function around $B(w,t)$ and set $\Xi_u = \psi_{u,s} - P$ and $\Xi_v = \phi_{v,s} - \widetilde{P}$. We decompose the first term in (3-5), $\Lambda(\Xi_u, \Xi_v, \phi_{w,t})$, as

$$T(\Xi_u, \Xi_v) = T(\alpha\Xi_u, \alpha\Xi_v) + T(\alpha\Xi_u, (1-\alpha)\Xi_v) + T(1-\alpha)\Xi_u, \Xi_v).$$

On the first term, use WBP around (w,t) to get

$$t^d t^{-d} \|\psi - P\|_{L^\infty(B(w,t))} \|\phi - \widetilde{P}\|_{L^\infty(B(w,t)} \sim \frac{t^{k_1+k_2+2}}{s^{2d+k_1+k_2+2}} \leq \frac{t^{k_1+k_2+\delta}}{s^{2d+k_1+k_2+\delta}}$$

for any $\delta \leq 2$. For the third (and second) terms, use the Hölder kernel estimate as above after integrating by parts $k_1 + k_2$ times to get

$$\int_{\substack{x_1 \in B(w,t)^c \\ x_2 \in \mathbb{R}^d}} \frac{t^{k_1+k_2+\delta} |\Xi_u(x_1)\Xi_v(x_2)|}{(|x_1 - w| + |x_2 - w|)^{2d+k+\delta}} \, dx_1 \, dx_2 \lesssim \frac{t^{k_1+k_2+\delta}}{s^d} \int_{x_1 \in B(w,t)^c} \frac{|\Xi(x_1)| \, dx_1}{|x_1 - w|^{d+k_1+k_2+\delta}}, \tag{3-6}$$

where we have used

$$\begin{aligned}\int_{\mathbb{R}^d} \frac{|\Xi_v(x_2)|}{(|x_1 - w| + |x_2 - w|)^{2d+k_1+k_2+\delta}} \, dx_2 &\leq s^{-d-k_2} \int_0^\infty (\tau + |x_1 - w|)^{-(2d+k_1+k_2+\delta)} \tau^{d+k_2-1} \, d\tau \\ &\leq s^{-d-k_2} c_d |x_1 - w|^{-(d+k_1+\delta)}.\end{aligned}$$

Break up the remaining integral in (3-6) into $t \leq |x_1 - w| \leq s$ and $s < |x_1 - w|$. In the first case, we have the estimate $|\Xi_u(x_1)| \leq 1^{1-\delta} |x_1 - w|^{k_1+\delta} s^{-d-k_1-\delta}$. Thus,

$$\int_{t \leq |x_1 - w| \leq s} \frac{|\Xi_u(x_1)| \, dx_1}{|x_1 - w|^{d+k_1+\delta}} \leq \frac{1}{s^{d+k_1+\delta}} \int_t^s \tau^{-1} \, d\tau \leq \frac{1}{s^{d+k+\delta}} \log\left(\frac{s}{t}\right).$$

On the other hand,

$$\int_{s < |x_1 - w|} \Xi_u(x_1) |x_1 - w|^{-(d+k_1+\delta)} \, dx_1 \leq s^{-d-k_1} \int_s^\infty \tau^{-(1+\delta)} \, d\tau \leq s^{-(d+k_1+\delta)}.$$

The remaining terms in (3-5) are $\Lambda(\psi_{u,s} - P_{k_1}, \widetilde{P}_{k_2}, \phi)$ and $\Lambda(P_{k_1}, \phi_{v,s} - \widetilde{P}_{k_2}, \phi)$. Comparing (3-5) with (3-4), we see that we only need to estimate two terms of the form

$$\Lambda(\Xi, \widetilde{P}, \phi_{w,t}) = \sum_{|\gamma| = k_2} \langle \psi_{u,s}, \vartheta^\gamma_{w,t} \rangle \Lambda(\Xi, p^\gamma_{w,t}, \phi_{w,t}), \tag{3-7}$$

where $p^\gamma_{w,t}(x) = ((x-w)/t)^\gamma$. We will need estimates for each summand in the future for all γ, so we will estimate the general form $\Lambda(\Xi, p^\gamma, \phi)$. As before, we take the decomposition

$$\Lambda(\Xi, p, \phi_{w,t}) = \Lambda(\alpha\Xi, \alpha p, \phi_{w,t}) + \Lambda(\alpha\Xi, (1-\alpha)p, \phi_{w,t}) + \Lambda((1-\alpha)\Xi, p, \phi_{w,t}),$$

where α is a smooth cutoff around $B(w,t)$.

For the first term, use WBP to get

$$t^d t^{-d} \|p\|_{L^\infty(B(w,t))} \|\psi - P\|_{L^\infty(B(w,t))} \lesssim \frac{t^{k_1+1}}{s^{d+k_1+1}}.$$

For the second term, follow the outline above, integrating by parts $k_1+|\gamma|$ times and applying the Hölder kernel estimate to get

$$|\Lambda(\Xi,\, p^\gamma,\, \phi)| \lesssim t^{k_1+|\gamma|+1} \int_{x_1\in B(w,t)} \int_{x_2\in B(w,t)^c} \frac{|\Xi(x_1)\, p^\gamma(x_2)|}{(|x_1-w|+|x_2-w|)^{2d+k_1+|\gamma|+\delta}}\, \mathrm{d}x_1\, \mathrm{d}x_2$$
$$\le \frac{t^{k_1+|\gamma|+k_1+1+\delta}}{s^{d+k_1+\delta}} \int_{x_2\in B(w,t)^c} \frac{|p^\gamma(x_2)|\, \mathrm{d}x_2}{|x_2-w|^{d+k_1+|\gamma|+\delta}} \le \frac{t^{k_1+|\gamma|+k_1+2\delta}}{s^{d+k_1+\delta}} \frac{1}{t^{k_1+\delta}} = \frac{t^{k_1+1}}{s^{d+k_1+1}}.$$

For the second inequality, we used the estimate $|\Xi(y)| \le t^{k_1+1}/s^{d+k_1}$ when $|y-w| \le t$. The third term in the decomposition is similar, but more closely follows the line of proof used on $\Lambda(\Xi,\, \Xi,\, \phi)$ above. Thus, the summands in (3-7) have the estimates

$$|\langle \psi_{u,s},\, \vartheta^\gamma_{w,t}\rangle \Lambda(p^\gamma_{w,t},\, \phi_{v,s} - P_{k_2},\, \phi_{w,t})| \le \frac{t^{|\gamma|+k_1+\eta}}{s^{2d+|\gamma|+k_1+\eta}} \tag{3-8}$$

since the coefficients satisfy $|\langle \psi_{u,s},\, \vartheta^\gamma_{w,t}\rangle| \lesssim t^{|\gamma|}/s^{d+|\gamma|}$ (see Lemma 2.3 using the fact that $\vartheta^\gamma_{w,t}$ has vanishing moments up to $|\gamma|$ and Hölder exponent $\delta = 1$). Taking the case $|\gamma| = k_2$, we see that the remaining terms in Υ, (3-7), satisfy the estimates claimed in the lemma. □

In the proof of the representation theorem, we will still have to deal with the error terms subtracted off of Υ in the region $A(w,t)$. $\Lambda(P,\, P,\, \phi)$ is controlled by the paraproducts assumption, but the so-called half-paraproducts $\Lambda(P_{k_1-1},\, \phi - P,\, \phi)$, satisfy the worse estimates (3-8) with $|\gamma| \le k_2 - 1$.

3C2. *Proof of Theorem A.* First, we represent Λ at the (k_1, k_2)-th level under the assumption it has vanishing paraproducts of all orders less than (k_1, k_2). We decompose $\Lambda(f, g, h)$ by applying the results of Section 2B to f, g, and h. By Calderón's formula, we obtain

$$\Lambda(f,g,h) = \int_{Z^d\times Z^d\times Z^d} \langle f,\phi_{u,r}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle \Lambda(\phi_{u,r},\phi_{v,s},\phi_{w,t})\, \mathrm{d}\mu(w,t)\, \mathrm{d}\mu(v,s)\, \mathrm{d}\mu(u,r)$$
$$= \int_{u,v,w} \left(\int_{r,s\ge t>0} + \int_{s,t\ge r>0} + \int_{t,r\ge s>0} \right).$$

Split the first integral as $\int_{r\ge s\ge t} + \int_{s\ge r\ge t}$ and use Lemma 2.7 above to get

$$\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ r,s\ge t>0}} = \int_{\substack{(v,w)\in\mathbb{R}^{2d}\\ s\ge t>0}} \Lambda\left(\int_{\substack{u\in\mathbb{R}^d\\ r\ge s}} \langle f,\phi_{u,r}\rangle \phi_{u,r}\, \mathrm{d}\mu(u,r), \phi_{v,s}, \phi_{w,t} \right) \langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle \mathrm{d}\mu(v,s)\, \mathrm{d}\mu(w,t)$$
$$+ \int_{\substack{(u,w)\in\mathbb{R}^{2d}\\ r\ge t}} \Lambda\left(\phi_{u,r}, \int_{\substack{v\in\mathbb{R}^d\\ s\ge r}} \langle g,\phi_{v,s}\rangle \phi_{v,s}\, \mathrm{d}\mu(v,s), \phi_{w,t} \right) \langle f,\phi_{u,r}\rangle\langle h,\phi_{w,t}\rangle\, \mathrm{d}\mu(u,r)\, \mathrm{d}\mu(w,t)$$
$$= \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ s\ge t>0}} \langle f,\psi^1_{u,s}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle \Lambda(\psi^2_{u,s},\phi_{v,s},\phi_{w,t})\, \frac{\mathrm{d}s\, \mathrm{d}t\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w}{st}$$
$$+ \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ s\ge t>0}} \langle f,\psi^3_{u,s}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle \Lambda(\psi^4_{u,s},\phi_{v,s},\phi_{w,t})\, \frac{\mathrm{d}s\, \mathrm{d}t\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w}{st}$$
$$+ \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ r\ge t>0}} \langle f,\phi_{u,r}\rangle\langle g,\psi^1_{v,r}\rangle\langle h,\phi_{w,t}\rangle \Lambda(\phi_{u,r},\psi^2_{v,r},\phi_{w,t})\, \frac{\mathrm{d}r\, \mathrm{d}t\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w}{rt}$$
$$+ \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ r\ge t>0}} \langle f,\phi_{u,r}\rangle\langle g,\psi^3_{v,r}\rangle\langle h,\phi_{w,t}\rangle \Lambda(\phi_{u,r},\psi^4_{v,r},\phi_{w,t})\, \frac{\mathrm{d}r\, \mathrm{d}t\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}w}{rt} =: I.$$

Recall that only ψ^1 and ψ^3 are noncancellative. Split each of the three terms in the same way, obtaining a decomposition $\Lambda(f,g,h) = I + II + III$, where I is given above, while II and III are defined by

$$\begin{aligned}
II &= \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ t\geq r>0}} \langle f,\phi_{u,r}\rangle\langle g,\psi^1_{v,t}\rangle\langle h,\phi_{w,t}\rangle \Lambda^{1*}(\phi_{w,t},\psi^2_{v,t},\phi_{u,r})\,\frac{\mathrm{d}t\,\mathrm{d}r\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{tr}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ t\geq r>0}} \langle f,\phi_{u,r}\rangle\langle g,\psi^3_{v,t}\rangle\langle h,\phi_{w,t}\rangle \Lambda^{1*}(\phi_{w,t},\psi^4_{v,t},\phi_{u,r})\,\frac{\mathrm{d}t\,\mathrm{d}r\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{tr}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ t\geq s>0}} \langle f,\phi_{u,t}\rangle\langle g,\phi_{v,s}\rangle\langle h,\psi^1_{w,t}\rangle \Lambda^{1*}(\psi^2_{w,t},\phi_{v,s},\phi_{u,t})\,\frac{\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{ts}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ t\geq s>0}} \langle f,\phi_{u,t}\rangle\langle g,\phi_{v,s}\rangle\langle h,\psi^3_{w,t}\rangle \Lambda^{1*}(\psi^4_{w,t},\phi_{v,s},\phi_{u,t})\,\frac{\mathrm{d}t\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{ts},\\
III &= \int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ s\geq r>0}} \langle f,\psi^1_{u,r}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,s}\rangle \Lambda^{2*}(\psi^2_{w,s},\phi_{v,s},\phi_{u,r})\,\frac{\mathrm{d}s\,\mathrm{d}r\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{sr}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ s\geq r>0}} \langle f,\psi^3_{u,r}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,s}\rangle \Lambda^{2*}(\psi^4_{w,s},\phi_{v,s},\phi_{u,r})\,\frac{\mathrm{d}s\,\mathrm{d}r\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{sr}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ r\geq s>0}} \langle f,\phi_{u,r}\rangle\langle g,\phi_{v,s}\rangle\langle h,\psi^1_{w,r}\rangle \Lambda^{2*}(\phi_{u,r},\psi^2_{w,r},\phi_{v,s})\,\frac{\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{rs}\\
&\quad+\int_{\substack{(u,v,w)\in\mathbb{R}^{3d}\\ r\geq s>0}} \langle f,\phi_{u,r}\rangle\langle g,\phi_{v,s}\rangle\langle h,\psi^3_{w,r}\rangle \Lambda^{2*}(\phi_{u,r},\psi^4_{w,r},\phi_{v,s})\,\frac{\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w}{rs}.
\end{aligned}$$

Due to the apparent symmetry, it is enough to handle only the first summand in I; let us call it σ_1. The remaining 11 terms are handled almost exactly the same. Recalling Υ from (3-4) and the different regions of Z^d from (3-3), we can write

$$\begin{aligned}
\sigma_1 &= \int_{(w,t)\in Z^d}\int_{(u,v,s)\in Z(w,t)} \Upsilon(u,v,w,s,t)\langle f,\psi^1_{u,s}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle\,\frac{\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}w}{st}\\
&\quad+\left(\int_{Z^d}\int_{(u,v,s)\in Z(w,t)} - \int_{Z^d}\int_{(u,v,s)\in Z(w,t)\setminus A(w,t)}\right)\Lambda(P_{k_1},\widetilde{P}_{k_2},\phi_{w,t})\\
&\qquad\qquad\times\langle f,\psi^1_{u,s}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle\frac{\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}w}{st}\\
&\quad+\int_{Z^d}\int_{(u,v,s)\in A(w,t)} [\Lambda(P_{k_1-1},\phi_{v,s}-\widetilde{P}_{k_2},\phi_{w,t})+\Lambda(\psi^1-P_{k_1-1},\widetilde{P}_{k_2},\phi_{w,t})]\\
&\qquad\qquad\times\langle f,\psi^1_{u,s}\rangle\langle g,\phi_{v,s}\rangle\langle h,\phi_{w,t}\rangle\frac{\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}w}{st}\\
&= \sigma_{1,0}+\sigma_{1,1}+\sigma_{1,2}+\sigma_{1,3}+\sigma_{1,4}.
\end{aligned}\tag{3-9}$$

Therefore, using Proposition 2.5 and the kernel estimates on Υ (Lemma 3.3), we obtain $\nu_z \in \Psi^{k_1+k_2,\delta;1,0}_z$ such that

$$\sigma_{1,0} = \int_{Z^d}\langle f\otimes g,\nu_z\rangle\langle h,\phi_z\rangle\,\mathrm{d}\mu(z) =: U^1_{k_1+k_2}(f,g,h).$$

Recalling the estimate (3-8), and again applying Proposition 2.5, we obtain $\nu_z^j \in C\Psi_z^{j,\delta;1,0}$ such that

$$\sigma_{1,3}+\sigma_{1,4}=\sum_{j=\min\{k_1,k_2\}}^{k_1+k_2-1}\int_{Z^d}\langle f\otimes g,\nu_z^j\rangle\langle h,\phi_z\rangle\,d\mu(z).$$

In this way we have constructed the remaining wavelet forms U_j^1 for $j=\min\{k_1,k_2\},\dots,k_1+k_2-1$. The vanishing paraproducts assumption allows us to compute

$$\Lambda(P,\widetilde{P},\phi_{w,t})=\sum_{|\gamma_i|=k_i}\langle\psi_{u,s},\vartheta_{w,t}^{\gamma_1}\rangle\langle\phi_{v,s},\vartheta_{w,t}^{\gamma_2}\rangle\langle b_\gamma^0,(\partial^{-\gamma_1-\gamma_2}\phi)_{w,t}\rangle.$$

By considering the supports of ϕ, ψ, and ϑ^{γ_i},

$$G(u,v,w,s,t):=\langle\phi_{v,s},\vartheta_{w,t}^{\gamma_1}\rangle\langle\psi_{u,s},\vartheta_{w,t}^{\gamma_2}\rangle$$

vanishes whenever $|u-w|\ge 3\max\{s,t\}$ or $|v-w|\ge 3\max\{s,t\}$. Moreover, $|G|\lesssim s^M/t^{2d+M}$ for any M up to which ψ and ϕ have vanishing moments and remain smooth (see Lemma 2.3). Therefore, for $\sigma_{1,2}$, the integration region $Z(w,t)\backslash A(w,t)$ can be replaced by $I(w,t)$ and by Lemma 2.6, there exists $\theta_z\in\Psi_z^{M,\delta;1,0}$ such that

$$\sigma_{1,2}=\sum_{|\gamma_\ell|=k_\ell}\int_{Z^d}\langle f\otimes g,\theta_z\rangle\langle h,\phi_z\rangle\langle b_\gamma,(\partial^{-\gamma_1-\gamma_2}\phi)_z\rangle\,d\mu(z).$$

Furthermore, since $b_\gamma^0\in\mathrm{BMO}$ and $\partial^{-\gamma_1-\gamma_2}\phi$ has mean zero, each summand in $\sigma_{1,2}$ is a wavelet form with the wavelet

$$\nu_z=\langle b_\gamma,(\partial^{-\gamma_1-\gamma_2}\phi)_z\rangle\theta_z.$$

We break up each σ_i as $\sigma_i=\sigma_{i,0}+\sigma_{i,1}+\sigma_{i,2}+\sigma_{i,3}+\sigma_{1,4}$, $i=1,2,\dots,12$. Each $\sigma_{i,0}$, $\sigma_{i,2}$, $\sigma_{i,3}$, and $\sigma_{i,4}$ is handled similarly, giving the wavelet forms U_j^i.

We now deal with the remaining terms, $\sigma_{i,1}$. We reassemble $\sigma_{i,1}+\sigma_{i+1,1}+\sigma_{i+2,1}+\sigma_{i+3,1}$ for $i=1,5,9$ and use Lemma 2.8 (the expanded tensor Calderón reproducing formula) to obtain

$$\sum_{\ell=1}^4\sigma_{\ell,1}=\sum_{|\gamma_i|=k_i}\int_{Z^d}\langle b_\gamma,(\partial^{-\gamma_1-\gamma_2}\phi)_z\rangle\langle f,\vartheta_z^{\gamma_1}\rangle\langle g,\vartheta_z^{\gamma_2}\rangle\langle h,\phi_z\rangle\,d\mu(z)=:\sum_{|\gamma_\ell|=k_\ell}\Pi_{b_\gamma,\gamma}(f,g,h).$$

Doing so similarly for $\sum_{\ell=5}^8\sigma_{\ell,1}$ and $\sum_{\ell=9}^{12}\sigma_{\ell,1}$ and using the vanishing paraproducts assumption on Λ^{1*} and Λ^{2*} yields

$$\Lambda(f,g,h)=\sum_{i=1}^6\sum_{j=\min\{k_1,k_2\}}^{k_1+k_2}U_j^i(f,g,h)+\sum_{|\gamma_\ell|=k_\ell}\sum_{i=0}^2\Pi_{b_\gamma^{i*},\gamma}^{i*}(f,g,h).\tag{3-10}$$

To remove the vanishing paraproducts assumption, we first recall the definition of (k_1,k_2) paraproducts. This means that

$$\Lambda_{k_1,k_2}:=\Lambda-\sum_{i=0}^2\sum_{(|\gamma_1|,|\gamma_2|)<(k_1,k_2)}\Pi_{b_\gamma^{i*},\gamma}^{i*}$$

is a (k_1,k_2,δ) CZ form and has vanishing paraproducts of orders $<(k_1,k_2)$. Thus the theorem is proved by applying (3-10) to Λ_{k_1,k_2}.

4. Sobolev space bounds

The cancellation structure of the forms U and Π is important for the results below. To reflect this, we introduce the intrinsic subtrilinear form

$$\Pi(f,g,h)=\int_{Z^d}\Psi_z^{0,\delta;1,0}(f,g)\,\Psi_z^{\mathcal{S}_0}(h)\,\mathrm{d}\mu(z), \tag{4-1}$$

where $\Psi_z^{\mathcal{S}_0}f$ and $\Psi_z^{k,\delta;i,j}(f,g)$ are the intrinsic wavelet coefficients defined by

$$\Psi_z^{0,\delta;i,j}(f,g)=\sup_{\nu_z\in\Psi_z^{0,\delta;i,j}}|\langle f\otimes g,\nu_z\rangle|,\qquad \Psi_z^{\mathcal{S}_0}(f)=\sup_{\substack{\theta_z\in\Psi_z^{0,\delta;0},\\ \|\mathsf{Sy}_z^{-1}\theta_z\|_{\star,2d+1,\delta}\le 1}}|\langle f,\theta_z\rangle|.$$

It is important to distinguish among the three arguments, since the first one is noncancellative, the first two have limited decay, and the third has rapid decay. In this first application, Section 4B, the collection of wavelets used in $\Psi^{\mathcal{S}_0}$ will actually be compactly supported, but we will need to consider rapidly decaying ones in Section 4C.

For the paraproducts, we define the intrinsic paraproduct form for $b\in\mathrm{BMO}$

$$\pi_b(f,g,h)=\int_{Z^d}\Psi_z^{\mathcal{S}_0}(b)\Psi_z^{0,\delta;1,1}(f,g)\Psi_z^{\mathcal{S}_0}(h)\,\mathrm{d}\mu(z). \tag{4-2}$$

Estimates for Λ are achieved using the representation theorem and then by appealing to estimates for Π and π_b. In particular, these forms have sparse $(1,1,1)$ bounds, which we will now define.

Definition 4.1. A collection Q of cubes $Q\subset\mathbb{R}^d$ is sparse if there is a disjoint collection of sets $\{E_Q:Q\in\mathsf{Q}\}$ such that

$$E_Q\subset Q\quad\text{and}\quad |E_Q|>\tfrac{1}{2}|Q|.$$

Above, $|\cdot|$ is the Lebesgue measure. A subtrilinear form $\mathcal{S}$ has sparse (p_1,p_2,p_3) bounds if for each triple $f_j\in L^\infty(\mathbb{R}^d)$ with compact support, $j=1,2,3$, there is a sparse collection $\mathsf{Q}=\mathsf{Q}(f_1,f_2,f_3)$ such that

$$\mathcal{S}(f,g,h)\le C\sum_{Q\in\mathsf{Q}}|Q|\langle f_1\rangle_{p_1,Q}\langle f_2\rangle_{p_2,Q}\langle f_1\rangle_{p_3,Q},\qquad \langle f\rangle_{p,Q}:=|Q|^{-1/p}\|f\mathbf{1}_Q\|_p. \tag{4-3}$$

The fact that π_b and Π have sparse bounds can be achieved through standard approaches; see for example [Conde-Alonso et al. 2017; Lacey 2017; Lerner 2013; Barron 2017], since they are more or less Calderón–Zygmund forms. However, in Section 5, Proposition 5.1 below, a direct proof is given. Such proof also applies to more general forms which do not necessarily satisfy kernel estimates.

4A. ***Weight classes.*** Sparse bounds are naturally related to weighted norm inequalities. Accordingly, the definition of the multilinear Muckenhoupt $A_{\vec{p},\vec{r}}$ weights, first appearing in [Lerner et al. 2009], is recalled below. We choose to employ the normalization of [Li et al. 2020, pp. 101–102] and stick to trilinear weight vectors, but the extension to higher linearities is a mere matter of changing the notation.

Throughout this discussion, unless otherwise specified, a weight vector $\vec{v}=(v_1,v_2,v_3)$ refers to a triple of positive measurable functions on $\mathbb{R}^3$ such that

$$1=\prod_{j=1}^3 v_j(x),\qquad x\in\mathbb{R}^d. \tag{4-4}$$

For $\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in (0, \infty]^3$, and a weight vector $\vec{v}$, define the characteristic

$$[\vec{v}]_{\vec{\varepsilon}} = \sup_Q \prod_{j=1}^{d} \left\langle \frac{1}{v_j} \right\rangle_{\varepsilon_j, Q},$$

where Q is allowed to vary among all cubes in $\mathbb{R}^d$. Notice that if $\varepsilon_j = \infty$, the corresponding local norm simply indicates the essential supremum on Q. We work with the extended simplex S and with the set of generalized Hölder tuples P

$$S = \left\{ \vec{\alpha} \in \left[-\tfrac{1}{2}, 1\right]^3 : \sum_{j=1}^{3} \alpha_j = 1 \right\}, \quad P = \left\{ \vec{p} = (p_1, p_2, p_3) \in (-\infty, \infty]^3 : \left(\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p_3}\right) \in S \right\}.$$

Say that the tuple $\vec{r} = (r_1, r_2, r_3) \in [1, \infty)^3$ satisfies $\vec{r} \prec \vec{p}$ for $\vec{p} \in P$ if

$$\varepsilon_j := \frac{p_j r_j}{p_j - r_j} > 0, \quad j = 1, 2, 3.$$

Above, we mean that $\varepsilon_j = r_j$ if $p_j = \infty$, in natural agreement with taking limits in the definitions. If $\vec{r} \prec \vec{p}$, writing $\vec{\varepsilon}(\vec{p}, \vec{r})$ for $\vec{\varepsilon}$ defined above, the weight vector $\vec{v}$ belongs to the class $A_{\vec{p},\vec{r}}$ if $[\vec{v}]_{A_{\vec{p},\vec{r}}} := [\vec{v}]_{\vec{\varepsilon}(\vec{p},\vec{r})} < \infty$. This definition, unlike that of [Li et al. 2020], is completely symmetric with respect to matching permutations of $\vec{v}$, $\vec{p}$, $\vec{r}$. However, for our purpose of studying bilinear operators acting on Lebesgue and Sobolev spaces whose integrability exponents p_1, p_2 are ≥ 1, it is convenient to break the symmetry and work in the corresponding portion of P. To wit, define

$$P_\circ = \{\vec{p} \in P : 1 < \min\{p_1, p_2\} < \infty\}, \quad p(\vec{p}) := \frac{p_3}{p_3 - 1} = \frac{p_1 p_2}{p_1 + p_2}.$$

Notice that $\frac{1}{2} < p(\vec{p}) < \infty$ is automatic from the definition of $P_\circ$. On the other hand, also observe that (at most) one of p_1, p_2 may be $= \infty$ when $\vec{p} \in P_\circ$. For comparison with [Li et al. 2020, pp. 101–102], when $\vec{v} \in A_{\vec{p},\vec{r}}$, one may single out the dual weight

$$w = w(\vec{v}) = \prod_{j=1}^{2} v_j = \frac{1}{v_3}$$

corresponding to the weight w associated to the pair (v_1, v_2) therein. We will not make use of the notation $w(\vec{v})$ in our statements to minimize redundancy. The most important classes for the study of bilinear Calderón–Zygmund operators correspond to the choice $r_1 = r_2 = r_3 = 1$. In that case, we simply write $[\vec{v}]_{A_{\vec{p}}}$ in place of $[\vec{v}]_{A_{\vec{p},\vec{r}}}$.

A consequence of the sparse bounds of Proposition 5.1 below for π_b and Π is the following weighted Lebesgue space result.

Proposition 4.2. *Let $\vec{p} \in P_\circ$, $\vec{v} \in A_{\vec{p}}$, $p = p(\vec{p})$, $b \in \mathrm{BMO}$. Then, denoting by T the bilinear operator defined by either $\langle T(f_1, f_2), h\rangle = \pi_b(f_1, f_2, h)$ or $\Pi(f_1, f_2, h)$, there holds*

$$\left\| \frac{T(f_1, f_2)}{v_3} \right\|_{L^p(\mathbb{R}^d)} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_1', p_2', p\}} \prod_{j=1}^{2} \|f_j v_j\|_{L^{p_j}(\mathbb{R}^d)}.$$

Proof. Let $\vec{q} = (3, 3, 3)$. Fix f_1, f_2 and h. Proposition 5.1 yields the existence of a sparse collection Q such that

$$|\langle T(f_1, f_2)v_3^{-1}, h\rangle\rangle| \lesssim \sum_{Q\in\mathsf{Q}} |Q|\,\langle f_1\rangle_{1,Q}\,\langle f_2\rangle_{1,Q}\,\langle hv_3^{-1}\rangle_{1,Q} \lesssim [\vec{v}]_{A_{\vec{q}}}^{3/2}\|f_1v_1\|_3\,\|f_2v_2\|_3\,\|h\|_3.$$

The sharp $A_{\vec{q}}$ weighted norm inequality for the sparse forms, see [Culiuc et al. 2018, Lemma 6.1], has been used for the last bound. The a priori estimate we obtained is in particular the case $\vec{p} = (3, 3, 3)$ of the proposition. The general case is now obtained by invoking the extrapolation result [Li et al. 2020, Theorem 2.1]. □

4B. ***Classical weighted Sobolev spaces.*** Recall that the weighted Sobolev norm $W^{k,p}(v)$, for $k \in \mathbb{N}$, $0 < p \le \infty$, and a weight v, is given by

$$\|f\|_{\dot{W}^{k,p}(v)} = \sum_{|\alpha|=k} \|[\partial^\alpha f]v\|_{L^p(\mathbb{R}^d)}, \qquad \|f\|_{W^{k,p}(v)} = \sum_{j=0}^{k} \|f\|_{\dot{W}^{j,p}(v)}.$$

Theorem B. *Let $k_0, k_1, k_2 \in N$ with $k_0 \le \min\{k_1, k_2\}$ and $\delta > 0$. Let Λ be a (k_1, k_2, δ) CZ form such that*

$$D^{k_0-|\gamma|}b_\gamma^0 \in \mathrm{BMO}, \quad |\gamma| < k_0. \tag{4-5}$$

Then,

$$\begin{aligned}\sum_{|\kappa|=k_0} \Lambda(f, g, \partial^\kappa h) \lesssim{}& \sum_{|\gamma|=k_0} \Pi(f, \partial^\gamma g, h) + \Pi(g, \partial^\gamma f, h)\\ &+ \sum_{|\gamma_\ell|\le k_\ell} \sum_{|\beta|=\max\{0,k_0-|\gamma|\}} \sum_{|\alpha|=k_0-|\beta|} \pi_{\partial^\beta b_\gamma^0}(\partial^{\alpha_1} f, \partial^{\alpha_2} g, h)\\ &+ \sum_{|\gamma_\ell|\le k_\ell} \sum_{|\beta|=\max\{0,k_0-|\gamma|\}} \sum_{|\alpha|=k_0-|\beta|} \pi_{\partial^\beta b_\gamma^0}(\partial^{\alpha_1} f, \partial^{\alpha_2} g, h)\\ &+ \sum_{|\gamma_\ell|\le k_\ell} \sum_{|\alpha|=k_0} \pi_{b_\gamma^1}(h, g, \partial^\alpha f) + \pi_{b_\gamma^2}(f, h, \partial^\alpha g).\end{aligned} \tag{4-6}$$

Proposition 4.2 then leads to the following bounds. For $\vec{p} \in P_\circ$, $\vec{v} \in A_{\vec{p}}$, $p = p(\vec{p})$,

$$\|T(f, g)\|_{\dot{W}^{k_0,p}(1/v_3)} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_1', p_2', p\}} \sum_{0\le i+j\le k_0} \|f\|_{\dot{W}^{i,p_1}(v_1)}\|g\|_{\dot{W}^{j,p_2}(v_2)}, \tag{4-7}$$

$$\|T(f, g)\|_{W^{k_0,p}(1/v_3)} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_1', p_2', p\}} \|f\|_{W^{k_0,p_1}(v_1)}\|g\|_{W^{k_0,p_2}(v_2)}. \tag{4-8}$$

If T has vanishing paraproducts, i.e., $b_\gamma^0 = 0$ for $(|\gamma_1|, |\gamma_2|) < (k_1, k_2)$, then the corresponding terms vanish from (4-6) so that (4-7) becomes

$$\|T(f, g)\|_{\dot{W}^{k_0,p}(1/v_3)} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_1', p_2', p\}} (\|f\|_{\dot{W}^{k_0,p_1}(v_1)}\|gv_2\|_{L^{p_2}(\mathbb{R}^d)} + \|fv_1\|_{L^{p_1}(\mathbb{R}^d)}\|g\|_{\dot{W}^{k_0,p_2}(v_2)}).$$

Theorem B is sharp in a couple of ways, but the precise sense must be explained. First, the appearance of the many norms on the right-hand side is necessary when considering the entire class of (k_1, k_2, δ) CZOs. In other words, for each pair (i, j) with $0 \le i + j \le k_0$, we can exhibit a (k_1, k_2, δ) CZO which

only maps $\dot{W}^{i,p_1} \times \dot{W}^{j,p_2}$ into $\dot{W}^{k_0,p}$. In fact, such an operator is

$$\langle T(f,g),h\rangle = \Pi_{b,\gamma}(f,g,h), \quad |\gamma_1| = i,\ |\gamma_2| = j,\ D^{k_0-|\gamma|}b \in \mathrm{BMO}\,.$$

Taken in a similar sense, both the exponent of the weight characteristic and the condition $D^{k_0-|\gamma|}b^0_\gamma \in \mathrm{BMO}$ are sharp. Concerning the exponent, the following sharpness result holds.

Proposition 4.3. *Fix $\vec{p} \in P_\circ$, $k_0, k_1, k_2 \in \mathbb{N}$, with $k_0 \le \min\{k_1, k_2\}$. For any weight vector $\vec{v}$, define*

$$\|T\|_{k_0,\vec{p},\vec{v}} := \sup_{f,g} \frac{\|T(f,g)\|_{\dot{W}^{k_0,p}(1/v_3)}}{\displaystyle\sum_{0\le i+j\le k_0} \|f\|_{\dot{W}^{i,p_1}(v_1)}\|g\|_{\dot{W}^{j,p_2}(v_2)}}.$$

Then, for each $M, \delta > 0$ and φ satisfying $\varphi(t) = o(t^{\max\{p_1',p_2',p\}})$ as $t \to \infty$,

$$\sup_{\vec{v},\,T} \frac{\|T\|_{\vec{p},k_0,\vec{v}}}{\varphi([\vec{v}]_{A_{\vec{p}}})} = \infty,$$

where the supremum is taken over all $\vec{v} \in A_{\vec{p}}$ and all T which are (k_1,k_2,δ) CZOs satisfying $\|T\|_{\vec{p},k_0,\vec{w}} \le M$, with $\vec{w} = (\mathbf{1}_{\mathbb{R}^d}, \mathbf{1}_{\mathbb{R}^d}, \mathbf{1}_{\mathbb{R}^d})$.

The analogous statement switching the input spaces also holds. When $k_0 = 0$, this result is contained in [Li et al. 2014, pp. 763–764]. Below, in Section 4B2, we will give a slight modification of their proof, adapted to our smooth operators.

As regards the paraproduct assumptions (4-5), notice that we are imposing additional requirements on the derivatives of b^0_γ, but not b^1_γ and b^2_γ. One may compare to the linear case, where one can obtain $T : \dot{W}^{k,p} \to \dot{W}^{k,p}$ if and only if b^γ_0 *vanishes* for $|\gamma| < k_0$. This phenomenon persists in the multilinear setting in the following sense.

Proposition 4.4. *Let T be a $(j+i,\delta)$ SI operator satisfying*

$$\|T\|_{\dot{W}^{j,p_1}\times\dot{W}^{i,p_2}\to\dot{W}^{k,p}} \lesssim 1$$

for some $\vec{p} \in P_\circ$ and $p = p(\vec{p}) \in (d,\infty)$. Then, for $(|\gamma_1|,|\gamma_2|) \le (j,i)$,

$$\begin{aligned} D^k T(x_1^{\gamma_1}, x_2^{\gamma_2}) &\in \mathrm{BMO} \quad \text{for } |\gamma_1| = j,\ |\gamma_2| = i,\\ D^k T(x_1^{\gamma_1}, x_2^{\gamma_2}) &= 0 \qquad \text{for } |\gamma| < j+i. \end{aligned}$$

As in the above discussion, such a result does not apply to our multilinear operators directly since they do not, as a whole, preserve homogeneous Sobolev spaces. However, if some portion of the operator — as seen from the proof, $\Pi_{b^0_\gamma}$ — does map accordingly, then we recover the conditions on b^0_γ. We present the proof of Proposition 4.4 in Section 4B2.

4B1. *Proof of Theorem B.* We first deal with U^1_j for $j \ge k_0$. Integrating by parts twice, we have, with $z = (w,t)$,

$$\begin{aligned} \sum_{|\kappa|=k_0} \langle f\otimes g, \nu_z\rangle\langle \partial^\kappa h, \phi_z\rangle &= \sum_{|\kappa|=k_0} \langle f\otimes g, \nu_z\rangle t^{-k_0} t^{k_0}\langle h, \partial^\kappa \phi_z\rangle\\ &= \sum_{|\kappa|=k_0}\sum_{|\gamma|=k_0} \langle f\otimes \partial^\gamma g, t^{-k_0}\partial_y^{-\gamma}\nu_z\rangle\langle h, t^{k_0}\partial^\kappa\phi_z\rangle. \end{aligned}$$

Noting that $t^{|\gamma|}\partial_y^{-\gamma}\Psi_z^{j,\delta;1,0} \subset \Psi_z^{0,\delta;1,0}$ for $|\gamma| \leq j$ (see Proposition 2.5), we obtain $\nu_z' \in \Psi_z^{0,\delta;1,0}$ and $\phi' = \partial^\kappa \phi$ with mean zero such that

$$\sum_{|\kappa|=k_0} \langle f \otimes g, \nu_z\rangle \langle \partial^\kappa h, \phi_z\rangle = \sum_{|\kappa|=k_0}\sum_{|\gamma|=k_0} \langle f \otimes \partial^\gamma g, \nu_z'\rangle\langle h, \phi_z'\rangle.$$

In this way,

$$\sum_{|\kappa|=k_0} |U(f, g, \partial^\kappa h)| \lesssim \sum_{|\gamma|=k_0} \Pi(f, \partial^\gamma g, h).$$

The same argument is applied to each U_j^2, placing all the derivatives on f and bounding above by $\sum_{|\gamma|=k_0}\Pi(g, \partial^\gamma f, h)$ since in this case g is in the noncancellative position. For the remaining wavelet forms, $U_j^3, \dots, U_j^6$ — the "adjoint" ones — we have more freedom with the derivatives. In fact, for any $|\kappa| = k^* \leq k_1 + k_2$, we can unwind the wavelet form and integrate by parts there. Let us only do U_3^j. Integrating by parts twice as before,

$$\begin{aligned} &U_j^3(f, g, \partial^\kappa h)\\ &\quad= \int_{\substack{(w,t)\in Z^d\\(u,v,s)\in Z(w,t)}} \Upsilon_j(u, v, w, s, t)\langle f, \phi_{w,t}\rangle\langle g, \psi_{v,s}^1\rangle\langle \partial^\kappa h, \phi_{u,s}\rangle \frac{\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}w}{st}\\ &\quad= \sum_{|\gamma|=k^*}\int_{\substack{(w,t)\in Z^d\\(u,v,s)\in Z(w,t)}} \frac{\Upsilon_j(u, v, w, s, t)t^{|\kappa|}}{s^{|\kappa|}}\langle \partial^\gamma f, (\partial^{-\gamma}\phi)_{w,t}\rangle\langle g, \psi_{v,s}^1\rangle\langle h, (\partial^\kappa\phi)_{u,s}\rangle \frac{\mathrm{d}s\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}t\,\mathrm{d}w}{st}. \end{aligned}$$

The new symbol $\Upsilon_j(u, v, w, s, t)(t/s)^{|\gamma|}$ satisfies much better estimates than needed in Proposition 2.5, and in fact supplies $\nu_z' \in \Psi_z^{j+|\gamma|,\delta;0,1} \subset \Psi_z^{0,\delta;0,1}$, which by the same argument as before gives

$$\sum_{|\kappa|=k^*} |U_j^i(f, g, \partial^\kappa h)| \lesssim \sum_{|\gamma|=k^*} \Pi(g, \partial^\gamma f, h).$$

The argument used on U_j^1, U_j^2 also applies to $\Pi_{b_\gamma^i,\gamma}^{i*}$ for $i = 1, 2$ since they have the same cancellation structure in the first two arguments. In this way,

$$\sum_{i=1}^2\sum_{|\kappa|=k^*} \Pi_{b_\gamma^i,\gamma}^{i*}(f, g, \partial^\kappa h) \lesssim \sum_{|\gamma|=k^*} \pi_{b_\gamma^1}(h, g, \partial^\gamma f) + \pi_{b_\gamma^2}(f, h, \partial^\gamma g).$$

However, we can actually see that the number of derivatives, k^*, can be taken all the way up to $k_1 + k_2$ since we placed the extra ones in the fully cancellative position.

The final term to estimate is the paraproduct $\Pi_{b_\gamma^0,\gamma}$. We are restricted here since the γ_ℓ-family $\vartheta_z^{\gamma_\ell}$ only has vanishing moments up to γ_ℓ. So we can only place $|\gamma_1|$ derivatives on f and $|\gamma_2|$ on g, and $k^* - |\gamma|$ derivatives remain on h. These must go on the symbol b_γ. We follow the above reasoning to obtain

$$\begin{aligned} \sum_{|\kappa|=k^*} \Pi_{b_\gamma,\gamma}(f, g, \partial^\kappa h) &= \sum_{|\kappa|=k^*}\sum_{|\alpha_\ell|=|\gamma_\ell|}\sum_{|\beta|=k^*-|\gamma|}\int \langle \partial^\beta b, (\partial^{-\beta-\gamma_1-\gamma_2}\phi)_{w,t}\rangle \frac{\mathrm{d}t\,\mathrm{d}w}{t}\\ &\qquad\times \langle \partial^{\alpha_1} f, (\partial^{-\alpha_1}\vartheta^{\gamma_2})_{w,t}\rangle\langle \partial^{\alpha_2} g, (\partial^{-\alpha_2}\vartheta^{\gamma_2})_{w,t}\rangle\langle h, (\partial^\kappa\phi)_{w,t}\rangle\\ &\leq \sum_{|\alpha_\ell|=|\gamma_\ell|}\sum_{|\beta|=k^*-|\gamma|} \pi_{\partial^\beta b}(\partial^{\alpha_1} f, \partial^{\alpha_2} g, h). \end{aligned}$$

□

From the proof we can see that all that is holding us back from taking k_0 all the way up to k_1+k_2 (represented by k^* in the proof) are the wavelet forms U_j^1 and U_j^2 for $j<k_1+k_2$. Thus, if these vanish or have some hidden regularity, we obtain the better result.

Corollary B.1. *Let Λ be a (k_1,k_2,δ) CZ form. Assume that $k^*\le k_1+k_2$, $D^{k^*-|\gamma|}b_\gamma^0\in\mathrm{BMO}$ for $|\gamma|\le k^*$, and that U_j^1 and U_j^2 are zero for $j<k^*$. Then, for any $\vec p\in P_\circ$, $p=p(\vec p)$, and $\vec v\in A_{\vec p}$,*

$$\|T(f,g)\|_{W^{k^*,p}(\mathbb{R}^d,1/v_3)}\lesssim [\vec v]_{A_{\vec p}}^{\max\{p_1',p_2',p\}}\|f\|_{W^{k^*,p_1}(v_1)}\|g\|_{W^{k^*,p_2}(v_2)}.$$

The assumption that U_j^1 and U_j^2 vanish is the same as checking that Λ has *vanishing half paraproducts* of appropriate orders. However, this condition can be difficult to check, and it is not even clear that the product operator satisfies this assumption. One could replace this with the assumption that U_j^1, U_j^2 are smoothing in some appropriate sense, similar to the assumption that $D^{k^*-|\gamma|}b_\gamma\in\mathrm{BMO}$, which makes the paraproduct forms smoothing operators. This is essentially what we do in assuming Λ has *more* than k_0 smoothness in Theorem B.

4B2. *Proofs of sharpness.* Let us now return to the proofs of Propositions 4.3 and 4.4, which demonstrate the sharpness of the exponent on the weight characteristic and the paraproduct assumptions (4-5) in Theorem B.

Proof of Proposition 4.3. From [Li et al. 2014], the bilinear operator given by

$$R(f,g)(x_0)=\int K(x_0-x_1,x_0-x_2)f(x_1)g(x_2),\qquad K(y_1,y_2)=\frac{y_1^1y_2^1}{(|y_1|^2+|y_2|^2)^{(2d+1)/2}},$$

satisfies

$$\sup_{\vec v\in A_{\vec p}}\frac{\|R\|_{0,\vec p,\vec v}}{\varphi([\vec v]_{A_{\vec p}})}=\infty.$$

R is a $(k_1,k_2,1)$ CZO for any $k_1,k_2\in\mathbb{N}$ and for any α,

$$\partial^\alpha R(f,g)=\sum_{\gamma\le\alpha}R(\partial^\gamma f,\partial^{\alpha-\gamma}g),$$

which implies that $\|R\|_{k_0,\vec p,\vec v}\sim\|R\|_{0,\vec p,\vec v}$ for any weight $\vec v$. □

Proof of Proposition 4.4. Let $|\alpha|=k$, $|\gamma_1|=j$, and $|\gamma_2|=i$. First we will show for each cube Q, setting $p_Q(x_1,x_2)=(x_1-x_Q)^{\gamma_1}(x_2-x_Q)^{\gamma_2}=:p_Q^{\gamma_1}(x_1)p_Q^{\gamma_2}(x_2)$,

$$\fint_Q|\partial^\alpha Tp_Q-c_Q|\mathrm{d}x\lesssim 1, \tag{4-9}$$

where $c_Q=\int\partial_{x_0}^\alpha K(x_1,x_2,x_Q)(1-\phi_Q)p_Q(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$ and x_Q is the center of the cube Q. Take the decomposition $p_Q=\phi_Qp_Q+(1-\phi_Q)p_Q$, where

$$\phi_Q(x_1,x_2)=\phi\left(\frac{x_1-x_Q}{\ell(Q)}\right)\phi\left(\frac{x_2-x_Q}{\ell(Q)}\right)$$

and $\phi\in C_0^\infty(B(0,2))$, with $\phi=1$ on $B(0,1)$. We estimate

$$\fint_Q|\partial^\alpha T\phi_Qp_1|\mathrm{d}x\le|Q|^{-1}|Q|^{1-1/p}\|D^j\phi_Qp_Q^{\gamma_1}\|_{L^{p_1}(Q)}\|D^i\phi_Qp_Q^{\gamma_2}\|_{L^{p_2}(Q)}\lesssim 1.$$

On the other hand, for $x_0 \in Q$,

$$\left| \int_{Q^c} \partial_{x_0}^{\alpha} K(x_1, x_2, x_0) - \partial_{x_0}^{\alpha} K(x_1, x_2, x_Q)(1 - \phi_Q(x_1, x_2))(x_1 - x_Q)^{\gamma_1}(x_2 - x_Q)^{\gamma_2} \, dx_1 \, dx_2 \right|$$
$$\lesssim \ell(Q)^{\delta} \int_{Q^c} \frac{|x_1 - x_Q|^j |x_2 - x_Q|^i}{(|x_Q - x_1| + |x_Q - x_2|)^{2d+j+i+\delta}} \, dx_1 \, dx_2 \lesssim 1.$$

Thus (4-9) is established. To complete the proposition, we must replace Tp_Q by T applied to the untranslated polynomial. To do so, we will show that the mapping properties of T imply that $T(x_1^{\kappa_1}, x_2^{\kappa_2})$ is a polynomial of degree $< k$ for $(|\kappa_1|, |\kappa_2|) < (j, i)$. Indeed, $\|x_1^{\kappa_1}\phi(x_1 R^{-1})\|_{\dot{W}^{j,p}} \lesssim R^{|\kappa_1|-j+d/p_1}$ and similarly for κ_2, i, and p_2. Thus, for any $\psi \in \mathcal{S}_k$,

$$\begin{aligned} |\langle T(x_1^{\kappa_1}, x_2^{\kappa_2}), \psi \rangle| &= \lim_{R\to\infty} |\langle T(x_1^{\kappa_1}\phi(x_1 R^{-1}), x_2^{\kappa_2}\phi(x_2 R^{-1})), \psi \rangle| \\ &\lesssim \lim_{R\to\infty} \|T(x_1^{\kappa_1}\phi(x_1 R^{-1}), x_2^{\kappa_2}\phi(x_2 R^{-1}))\|_{\dot{W}^{k,p}} \\ &\lesssim \lim_{R\to\infty} \|x_1^{\kappa_1}\phi(x_1 R^{-1})\|_{\dot{W}^{j,p_1}} \|x_2^{\kappa_2}\phi(x_2 R^{-1})\|_{\dot{W}^{i,p_2}} = 0, \end{aligned}$$

which implies $T(x_1^{\kappa_1}, x_2^{\kappa_2})$ is a polynomial of degree strictly less than k. In particular,

$$\partial^{\alpha} T(x_1^{\kappa_1}, x_2^{\kappa_2}) = 0.$$

For each cube Q and $(|\gamma_1|, |\gamma_2|) = (j, i)$,

$$T(x_1^{\gamma_1}, x_2^{\gamma_2}) = \sum_{\kappa<\gamma} c_{\kappa,\gamma} x_Q^{\gamma-\kappa} T(x_1^{\kappa_1}, x_2^{\kappa_2}) + Tp_Q.$$

Therefore $\partial^{\alpha} T(x_1^{\gamma_1}, x_2^{\gamma_2}) = \partial^{\alpha} Tp_Q$. □

4C. ***Weighted fractional Sobolev space estimates.*** Until now D^{σ}, the Fourier multiplier by $|\xi|^{\sigma}$, has only been used when σ is a positive integer. We now generalize to any σ positive. Denote by $W^{\sigma,p}(v)$ the weighted inhomogeneous fractional Sobolev space on $\mathbb{R}^d$ with norm defined by

$$\|f\|_{W^{\sigma,p}(v)} = \|[D^{\sigma} f]v\|_{L^p(\mathbb{R}^d)} + \sum_{j=0}^{\lfloor\sigma\rfloor} \|[D^j f]v\|_{L^p(\mathbb{R}^d)}.$$

Two new features enter here, which will require some modification of the intrinsic forms. First, D^{σ} applied to a mother wavelet will no longer be a mother wavelet; however, it (and many of its derivatives) will still have rapid decay. Second, when D^{σ} is applied to a noncancellative wavelet, its decay will only be $\langle \cdot \rangle^{d+\sigma}$ (see Lemma 4.6 below). This motivates the introduction of limited decay wavelets and intrinsic forms, π_b^{σ} and Π^{σ}, which we will define later. For now, know that they are defined the same as π_b and Π above, but the wavelet class corresponding to the first argument has worse decay, depending on σ. Furthermore, as we will show in Proposition 5.1 below, these forms have sparse $(1, p_2, p_3)$ bounds for

$$\frac{1}{p_2} + \frac{1}{p_3} < \frac{\sigma + d}{d}.$$

This implies the following weighted Lebesgue space bounds.

Proposition 4.5. *Let T be the bilinear operator defined by either $\langle T(f_1, f_2), h\rangle = \pi_b^\sigma(h, f_1, f_2)$ or $\Pi^\sigma(h, f_1, f_2)$, defined by (4-11) and (4-13) with $b \in \mathrm{BMO}$. Let $\vec{p} \in P_\circ$, $\vec{v} \in A_{\vec{p}}$, $p = p(\vec{p}) > d/(\sigma + d)$. For $\vec{r} = (r_1, r_2, 1)$ with $1 \le r_i < p_i$ and $1/r_1 + 1/r_2 < (\sigma + d)/d$,*

$$\left\| \frac{T(f_1, f_2)}{v_3} \right\|_{L^p(\mathbb{R}^d)} \lesssim [\vec{v}]_{A_{\vec{p},\vec{r}}}^{\max\{p_i/(p_i - r_i), p\}} \prod_{j=1}^{2} \|f_j v_j\|_{L^{p_j}(\mathbb{R}^d)}.$$

This proposition is proved in the same manner as Proposition 4.2.

Theorem C. *Let $k_1, k_2 \in \mathbb{N}$ and $\sigma, \delta > 0$, with $\sigma \le \min\{k_1, k_2\}$. If Λ is a (k_1, k_2, δ) CZ form satisfying*

$$D^{\sigma - |\gamma|} b_\gamma^0 \in \mathrm{BMO}, \quad |\gamma| < \sigma,$$

then,

$$\begin{aligned} |\Lambda(f, g, D^\sigma h)| \lesssim{}& \Pi^\sigma(h, g, D^\sigma f) + \Pi^\sigma(h, f, D^\sigma g) + \Pi(f, D^\sigma g, h) + \Pi(g, D^\sigma f, h) \\ &+ \sum_{|\gamma_\ell| \le k_\ell} \pi_{b_\gamma^1}^\sigma(h, g, D^\sigma f) + \pi_{b_\gamma^2}^\sigma(h, f, D^\sigma g) \\ &+ \sum_{\substack{|\gamma_\ell| \le k_\ell \\ |\gamma| < \sigma}} \pi_{D^{\sigma - |\gamma|} b_\gamma^0}(D^{|\gamma_1|} f, D^{|\gamma_2|} g, h) \\ &+ \sum_{\substack{|\gamma_\ell| \le k_\ell \\ |\gamma| > \sigma}} \pi_{b_\gamma^0}(D^{\min\{\sigma, |\gamma_1|\}} f, D^{\sigma - \min\{\sigma, |\gamma_1|\}} g, h). \end{aligned} \tag{4-10}$$

Applying the estimates on Π^σ, Π, π_b^σ, and π_b from Propositions 4.2 and 4.5, we obtain the following fractional weighted Sobolev space estimate. Let $\vec{p} \in P_\circ$, $p = p(\vec{p}) > (\sigma + d)/d$, and $\vec{r} = (r_1, r_2, 1)$ satisfying $1 \le r_i < p_i$ and $1/r_1 + 1/r_2 < (\sigma + d)/d$. Then, for any $\vec{v} \in A_{\vec{p},\vec{r}}$,

$$\|T(f, g)\|_{W^{\sigma,p}(1/v_3)} \lesssim [\vec{v}]_{A_{\vec{p},\vec{r}}}^{\max\{p_i/(p_i - r_i), p\}} \|f\|_{W^{\sigma,p_1}(v_1)} \|g\|_{W^{\sigma,p_2}(v_2)}.$$

As in the classical case, if $b_\gamma^0 = 0$ then the corresponding terms vanish in (4-10) and we obtain the following simplification for the homogeneous norms.

Corollary C.1. *If in addition to the assumptions of Theorem C,*

$$b_\gamma^0 = 0, \quad (|\gamma_1|, |\gamma_2|) < (k_1, k_2),$$

then, for any $\vec{v} \in A_{\vec{p},\vec{r}}$,

$$\begin{aligned} &\left\| \frac{D^\sigma T(f, g)}{v_3} \right\|_{L^p(\mathbb{R}^d)} \\ &\qquad \lesssim [\vec{v}]_{A_{\vec{p},\vec{r}}}^{\max\{p_i/(p_i - r_i), p\}} (\|[D^\sigma f] v_1\|_{L^{p_1}(\mathbb{R}^d)} \|g v_2\|_{L^{p_2}(\mathbb{R}^d)} + \|f v_1\|_{L^{p_1}(\mathbb{R}^d)} \|[D^\sigma g] v_2\|_{L^{p_2}(\mathbb{R}^d)}). \end{aligned}$$

These new forms will only be needed to represent U_j^3, U_j^6, and $\Pi_{b_\gamma^i, \gamma}^{i*}$ for $i = 1, 2$ — the ones with h in a noncancellative position. So first we give the proof representing all the other terms using the original intrinsic forms Π and π_b.

4C1. *Proof of terms involving* Π *and* π_b. For the remaining wavelet forms, $U_j^1, U_j^2, U_j^4, U_j^5$, we use the fact that $D^{\pm\sigma}\phi_{u,s} = s^{\mp\sigma}(D^{\pm\sigma}\phi)_{u,s}$ and, since $\hat{\phi}$ has a zero of order larger than σ at the origin, $\|D^{\pm\sigma}\phi\|_{\star,M,1} \lesssim 1$ for any $M > 0$. Let us focus on U_j^1. For any $\sigma \le \min\{k_1, k_2\}$, we unwind the wavelet form

$$U_j^1(f, g, D^\sigma h) = \int_{(w,t)\in Z^d} \int_{(u,v,s)\in Z(w,t)} \Upsilon_j(u, v, w, s, t) s^\sigma t^{-\sigma} \langle D^\sigma f, (D^{-\sigma}\phi)_{u,s}\rangle \\ \times \langle g, \psi^1_{v,s}\rangle \langle h, (D^\sigma \phi)_{w,t}\rangle \, \frac{\mathrm{d}s\, \mathrm{d}u\, \mathrm{d}v\, \mathrm{d}t\, \mathrm{d}w}{st}.$$

By Proposition 2.5, $\Upsilon_j(u, v, w, s, t)(s/t)^\sigma$ supplies $\nu'_z \in \Psi_z^{\lfloor j-\sigma \rfloor, \delta; 0, 1} \subset \Psi_z^{0,\delta;0,1}$.

The paraproduct term $\Pi_{b^0_\gamma, \gamma}$ follows this same outline since we will not have to apply D^σ to a noncancellative wavelet.

4C2. *Intrinsic forms* Π^σ *and* π_b^σ. Let us begin by constructing the intrinsic form to represent U_j^3 (and similarly U_j^6). As before,

$$\langle D^\sigma h \otimes g, \nu_z\rangle \langle f, \phi_z\rangle = \langle h \otimes g, t^\sigma D_x^\sigma \nu_z\rangle \langle D^\sigma f, (D^{-\sigma}\phi)_z\rangle.$$

Setting $\nu_z^\sigma = t^\sigma D_y^\sigma \nu_z$ and $\phi' = D^{-\sigma}\phi$, we obtain

$$U_j^3(f, g, D^\sigma h) = \int_{Z^d} \langle h \otimes g, \nu_z^\sigma\rangle \langle D^\sigma f, \phi'_z\rangle \, \mathrm{d}\mu(z),$$

which is dominated by the intrinsic form

$$\Pi^\sigma(h, g, f) = \int_{Z^d} \sup_{\nu_z^\sigma \in \Psi_z^{\sigma,k,\delta;1,0}} \langle h \otimes g, \nu_z^\sigma\rangle \Psi_z^{\mathcal{S}_0}(D^\sigma f) \, \mathrm{d}\mu(z) \tag{4-11}$$

for some new wavelet class $\Psi_z^{\sigma,k,\delta;1,0}$ which we now define. The decay of ν_z^σ can be computed, but it will be asymmetric in the two variables, so let us introduce a new norm

$$\|\psi\|_{\sigma,\eta,\delta} = \sup_{x\in\mathbb{R}^{2d}} \langle x\rangle^{d+\sigma}(1+|x_2|)^{d+\eta-\sigma}|\psi(x)| + \sup_{x\in\mathbb{R}^{2d}} \langle x\rangle^{d+\sigma}(1+|x_2|)^{d+\eta-\sigma} \frac{|\psi(x)-\psi(x+h)|}{|h|^\delta} \tag{4-12}$$

if $\sigma \le d+\eta$. Otherwise $\|\cdot\|_{\sigma,\eta,\delta} = \|\cdot\|_{\star,\eta,\delta}$ from (2-3). Define the associated wavelet class by

$$\Psi_z^{\sigma,k,\delta;1,0} = \{\psi \in C^k : \|(\mathsf{Sy}_z^1 \mathsf{Sy}_z^2)^{-1}\psi\|_{\sigma,k+\delta,\delta} \le 1\}.$$

Notice that if $k > \sigma > d$, then $\Psi_z^{\sigma,k,\delta;i,j} = \Psi_z^{0,\delta;i,j}$. The next lemma will establish that this norm and wavelet class are the correct ones for the modified intrinsic forms.

Lemma 4.6. *Let* $0 < \sigma < k$. *There exists* $C > 0$ *such that*

$$t^\sigma \|(\mathsf{Sy}_z^1 \mathsf{Sy}_z^2)^{-1} D_{x_1}^\sigma \nu_z\|_{\sigma,k+\delta,\delta} \le C$$

for all $\nu_z \in \Psi_z^{k,\delta;1,0}$.

In this way, $U_j^3(f, g, h) \lesssim \Pi^\sigma(h, g, D^\sigma f)$, $U_j^6(f, g, h) \lesssim \Pi^\sigma(h, f, D^\sigma g)$, and similarly for the paraproducts, defining

$$\pi_b^\sigma(f, g, h) = \int_{Z^d} \Psi_z^{\mathcal{S}_0}(b) \Psi_z^{\sigma,k,\delta;1,1}(f, g) \Psi_z^{\mathcal{S}_0}(h) \, \mathrm{d}\mu(z). \tag{4-13}$$

Proof of Lemma 4.6. If ν_z is Schwartz, then this is well known; see [Grafakos and Oh 2014]. Our starting point is from the book [Grafakos 2008, Definition 2.4.5 and Theorem 2.4.6, pp. 127–130], which states that, for all ϕ Schwartz and $\sigma > 0$,

$$\int |\xi|^\sigma \hat{\phi}(\xi)\,\mathrm{d}\xi = c_{d,\sigma} \int |u|^{-d-\sigma}\phi(u)\,\mathrm{d}u.$$

The right-hand side is well-defined once understood as

$$\int_{|u|<1} \frac{\phi(u) - \sum_{|\alpha|\le k}(\phi^{(\alpha)}(0)/\alpha!)u^\alpha}{|u|^{d+\sigma}} + \int_{|u|>1} \frac{\phi(u)}{|u|^{d+\sigma}} + \sum_{|\alpha|\le k} b(\sigma,k,\alpha)\phi^{(\alpha)}(0)$$

for any $k > \sigma$. Written this way the integrals are all absolutely convergent with bounds depending on d, σ, k and linearly on the Hölder constant for $\partial^k\phi$ near 0 and $\|\phi^{(\alpha)}(0)\|_\infty$. By density, the definition can be extended to functions in $\Psi^{k,\delta;1}_{(0,1)}$. Moreover, replacing ϕ by $\phi(\,\cdot - x)$ for some x, we obtain that the first and last terms decay as well as $(1+|x|)^{-(k+d+\delta)}$. The remaining (middle) term is estimated by splitting the integral into the region $|u| > |x|/2$ and $|u| \le |x|/2$. The first region is controlled by the decay of ϕ, let us say it is $M > d$, using the fact that the kernel is bounded away from zero since $|u+x| > 1$. Then, for any $\eta > 0$,

$$\int_{|u|>|x|/2|} |\phi(u)|\,\mathrm{d}u \lesssim (1+|x|)^{d+\eta-M} \int (1+|u|)^{M-d-\eta}|\phi(u)|\,\mathrm{d}u.$$

The other range is limited by σ using the fact that in this range $|u+x| \ge |x| - |u| \ge |x|/2$ so that

$$\int_{|u|<|x|/2} \frac{|\phi(u)|}{|u+x|^{d+\sigma}}\,\mathrm{d}u \lesssim |x|^{d+\sigma}.$$

To complete the proof we use the translation invariance and homogeneity of the kernel to obtain

$$(\mathsf{Sy}^1_z)^{-1}(D^\sigma_{x_1}\nu_z)(x_1,x_2) = t^{-\sigma} D^\sigma_{x_1}(\mathsf{Sy}^1_z)^{-1}\nu_z(x_1,x_2)$$

for any $\nu_z \in \Psi^{k,\delta;1,1}_z$ and applying the previous estimate to $(\mathsf{Sy}^1_z\mathsf{Sy}^2_z)^{-1}\nu_z(\,\cdot\,,x_2)$ to get that the decay is either

$$\max\{1,|x_1|,|x_2|\}^{-(\sigma+d)}(1+|x_2|)^{-[(k-\sigma)+d+\delta]}$$

or

$$\max\{1,|x_1|,|x_2|\}^{-M+d}(1+|x_2|)^{-d},$$

with $M = k+2d+\delta$, whichever is worse. □

5. Sparse bounds for intrinsic forms

Proposition 5.1. *Let* $\Pi, \pi_b, \Pi^\sigma, \pi^\sigma_b$ *be defined by* (4-1), (4-2), (4-11), *and* (4-13). *Let* $b \in \mathrm{BMO}$.

(1) *The forms* Π *and* π_b *have sparse* $(1,1,1)$ *bounds.*

(2) *The forms* Π^σ *and* π^σ_b *have sparse* $(1, p_2, p_3)$ *bounds for any* $1 \le p_2, p_3 \le \infty$, *with* $1/p_2 + 1/p_3 < (\sigma+d)/d$.

Before beginning the proof, we collect some estimates in the following simple lemma.

Lemma 5.2. *Let $\eta \geq \delta > 0$, and $Q \subset \mathbb{R}^d$ be a cube with center $c(Q)$ side length $\ell(Q)$. For all $z = (w, t) \in Z^d$ and $\|\mathsf{Sy}_z^{-1}\theta_z\|_{\star,\eta,\delta} \leq 1$, if $w \notin 3Q$ then*

$$|\langle f 1_Q, \theta_{w,t}\rangle| \lesssim \langle f\rangle_Q \frac{\ell(Q)^d t^\eta}{\max\{t, |w - c(Q)|\}^{d+\eta}}. \tag{5-1}$$

If in addition f has mean zero, then

$$|\langle f 1_Q, \theta_{w,t}\rangle| \lesssim \langle f\rangle_Q \frac{\ell(Q)^{d+\delta} t^{\eta-\delta}}{\max\{t, |w - c(Q)|\}^{d+\eta}}. \tag{5-2}$$

On the other hand, if $w \in Q$, then

$$|\langle f 1_{(3Q)^c}, \theta_z\rangle| \lesssim \left(\frac{t}{\ell(Q)}\right)^\eta \inf_{u\in Q} M(f)(u). \tag{5-3}$$

Proof. We begin with (5-1). If $x \in Q$ and $w \notin 3Q$, then $|x - w| \geq \frac{1}{2}|w - c(Q)|$. Since

$$\theta_{w,t}(x) \leq t^{-d}\left(1 + \frac{|x - w|}{t}\right)^{-(d+\eta)},$$

this implies

$$\int_Q |f(x)\theta_{w,t}(x)|\,dx \lesssim \langle f\rangle_Q |Q| \frac{t^\eta}{\max\{t, |w - c(Q)|\}^{d+\eta}}.$$

If f has mean zero, then

$$\begin{aligned}|\langle f 1_Q, \theta_{w,t}\rangle| &= \left|\int_Q f(x)[\theta_{w,t}(x) - \theta_{w,t}(c(Q))]\,dx\right| \\ &\lesssim \ell(Q)^\delta \int_Q |f(x)| t^{-d-\delta}\left(1 + \frac{|c(Q) - w|}{t}\right)^{-(d+\eta)} dx \leq \langle f\rangle_Q \frac{\ell(Q)^{d+\delta} t^{\eta-\delta}}{\max\{t, |w - c(Q)|\}^{d+\eta}}.\end{aligned}$$

For (5-3), decompose the integral into dyadic annuli $A_j = B(u, 2^{j+1}\ell(Q))\backslash B(u, 2^j\ell(Q))$ for any $u \in Q$. We can skip the ball $B(u, \ell(Q))$ since $B(u, \ell(Q)) \subset 3Q$. Therefore,

$$|\langle f, \theta_{w,t}\rangle| \leq \sum_{j=0}^\infty \int_{A_j} |f(x)\theta_{w,t}(x)|\,dx \lesssim \sum_{j=0}^\infty |A_j| \langle f\rangle_{1,A_j} t^\eta (2^j\ell(Q))^{-(d+\eta)} \lesssim \left(\frac{t}{\ell(Q)}\right)^\eta Mf(u). \quad \square$$

There are only two points at which we will need to distinguish among the four forms, so we will use $\mathcal{A}$ to represent any one of them except at these two crucial points. The key property of $\mathcal{A}$ — the shared property of Π, π_b, Π^σ, and π_b^σ — is the bound

$$|\mathcal{A}(f_1, f_2, f_3)| \lesssim \int_{Z^d} \Psi_z^{0,\delta;1}(f_1)\Psi_z^{0,\delta;1}(f_2)\Psi_z^{0,\delta;1}(f_3)\,d\mu(z), \quad \Psi_z^{0,\delta;1}(f) := \sup_{\theta_z \in \Psi_z^{0,\delta;1}} |\langle f, \theta_z\rangle|. \tag{5-4}$$

All four forms also satisfy the bound $\mathcal{A}(f_1, f_2, f_3) \lesssim \prod_{i=1}^3 \|f_i\|_{q_i}$, where $1 < q_i < \infty$ and $\sum q_i^{-1} = 1$, though for different reasons. This is the first point at which we consider each form separately, since the bound in (5-4) above is not L^p bounded. For Π and Π^σ, they can be bounded by $\|Mf_1\|_{L^{q_1}} \|Sf_2\|_{L^{q_2}} \|Sf_3\|_{L^{q_3}}$ where M and S are modified maximal and intrinsic square functions for which the *linear* L^p-mapping

properties are well known [Di Plinio et al. 2022]. For the paraproducts, it is a little more complicated, yet still within the realm of the standard linear theory, so we do not prove it here.

Proof of Proposition 5.1. In the course of this proof, we will use the following notation. First, to unify, the tuple $\vec{r} = (r_1, r_2, r_3)$ is set as

$$\vec{r} := \begin{cases} (1,1,1), & \mathcal{A} \in \{\Pi^\sigma, \pi_b^\sigma\}, \\ (1, p_2, p_3), & \mathcal{A} \in \{\Pi\, \pi_b\}. \end{cases}$$

There is no loss in generality with the assumption $1 \le p_2, p_3 \le 2$ since $(\sigma+d)/d > 1$ and $\vec{p}$ sparse bounds imply $\vec{p}'$ sparse bounds for any $\vec{p} \le \vec{p}'$. Let $\mathcal{D}$ be the standard dyadic system on $\mathbb{R}^d$. For a cube $Q \in \mathcal{D}$, $T(Q)$ is the Carleson box $T(Q) = (0, \ell(Q)] \times Q \subset (0,\infty) \times \mathbb{R}^d$. If $\mathcal{E} \subset \mathcal{D}$ is a pairwise disjoint cover of $E \subset \mathbb{R}^d$, take $T(E) := \bigcup \{T(Q) : Q \in \mathcal{E}\}$. Finally, if $F \subset (0,\infty) \times \mathbb{R}^d$, the truncated operators $\mathcal{A}_F$ are defined for any $F \subset Z^d$ by integrating over only F in (5-4).

Our task is to prove (4-3) holds for all triples $f_j \in L^\infty(\mathbb{R}^d)$ with compact support. Fix such a triple and take $Q_0 \in \mathcal{D}$ with the property that $\operatorname{supp} f_j \subset 3Q_0$ for all $j = 1, 2, 3$. The proof is iterative in nature, and we begin the main step of the iteration by defining an exceptional subset. Let

$$E := \bigcup_{j=1}^{3} \{x \in \mathbb{R}^d : \mathrm{M}_{r_j} f_j(x) > C \langle f_j \rangle_{r_j, 3Q_0}\}.$$

For C large, by the maximal inequality, $|E| \le 2^{-4d}|Q_0|$. Let now $\mathcal{E}$ be the maximal elements of the collection $\{Q \in \mathcal{D} : 9Q \subset E\}$. Clearly $\mathcal{E}$ is a pairwise disjoint cover of E. Moreover, the stopping nature of $Q \in \mathcal{E}$ yields the property

$$\inf_Q \mathrm{M}_{r_j} f_j \lesssim \langle f_j \rangle_{r_j, 3Q_0}$$

uniformly over $Q \in \mathcal{E}$ and $j = 1, 2, 3$. This property will be tacitly used throughout the proof. We use E to induce the decomposition $\mathcal{A} = \mathcal{A}_{T(E)^c} + \mathcal{A}_{T(E)}$, whose terms we estimate separately. Let us begin with $\mathcal{A}_{T(E)}$. Break this up as

$$\mathcal{A}_{T(E)}(f_1, f_2, f_3) = \sum_{Q \in \mathcal{E}} \sum_{\vec{g}} \mathcal{A}_{T(Q)}(g_1, g_2, g_3),$$

where each $\vec{g} = (g_1, g_2, g_3)$ runs over 2^3 possibilities where each g_j is either $f_j 1_{3Q}$ (in) or $f_j 1_{(3Q)^c}$ (out). We leave alone the term consisting entirely of "in" functions. "Out" functions are good so let us assume g_1 is out and the others are in. This is the second point at which we distinguish among the four forms for $\mathcal{A}$. For Π and π_b, $\vec{r} = (1,1,1)$. We obtain, applying (5-3) to g_1 with $\eta = d + \delta$,

$$\begin{aligned} \int_{T(Q)} \prod_{j=1}^{3} |\langle g_j, \theta_{w,t} \rangle| \, \frac{\mathrm{d}w \, \mathrm{d}t}{t} &\lesssim \int_0^{\ell(Q)} \frac{t^{d+\delta}}{\ell(Q)^{d+\delta}} \inf_{u \in Q} \mathrm{M} g_1(u) \Big\langle |f_2|, \int_Q |\theta_{w,t}| \mathrm{d}w \Big\rangle \|f_3\|_{L^1} t^{-d} \, \frac{\mathrm{d}t}{t} \\ &\lesssim |Q| \inf_Q \mathrm{M}(1_{(3Q)^c} f_1) \langle f_2 \rangle_{3Q} \langle f_3 \rangle_{3Q}, \end{aligned}$$

where we used the fact that $\int_Q |\theta_{w,t}(x)| \, \mathrm{d}w \le \|\theta_{x,t}\|_{L^1} \lesssim 1$. Therefore, summing over $Q \in \mathcal{E}$,

$$\sum_{Q \in \mathcal{E}} |Q| \inf_Q \mathrm{M}(1_{(3Q)^c} f_1) \langle f_2 \rangle_{1,3Q} \langle f_3 \rangle_{1,3Q} \lesssim \prod_{j=1}^{3} |Q_0| \langle f_j \rangle_{r_j, 3Q_0}.$$

One can verify the same result for two out functions, and when all three are out, use the improved decay to obtain

$$\int_{T(Q)} \frac{t^{2(d+\delta)}}{\ell(Q)^{2(d+\delta)}} \prod_{j=1}^{3} \inf_{Q} \mathrm{M} g_j \, \frac{\mathrm{d}w \, \mathrm{d}t}{t} \lesssim |Q| \prod_{j=1}^{3} \langle f_j \rangle_{r_j, 3Q_0}.$$

For Π^σ and π_b^σ, consider the case where there is only one out function and it is in the first position; this is the worst case as the other cases actually result in the same situation above with the power $d+\delta$. The second wavelet in Π^σ has decay greater than $d+\delta$; see (4-12) since $k-\sigma > 0$. Thus, for any $q \geq 1$, set

$$\chi_t(x) = t^{-d} \left(1 + \frac{|x|}{t}\right)^{d+\delta}, \quad \|\theta_{w,t}\|_{L^q} \leq \|\chi_t\|_{L^q} \lesssim t^{d(1/q-1)}.$$

For $j = 2, 3$ set

$$z_j = \left(\frac{1}{p_2} + \frac{1}{p_3}\right) p_j \geq p_j, \quad q_j := \left(\frac{1}{z_j} + 1 - \frac{1}{p_j}\right)^{-1} \geq 1$$

and apply Young's inequality to obtain

$$\left(\int_Q |\langle g_j, \theta_{w,t} \rangle|^{z_j} \, \mathrm{d}w\right)^{1/r_j} \leq \| |g_j| * \chi_t \|_{L^{r_j}} \leq \|g_j\|_{L^{p_j}} \|\chi_t\|_{L^{q_j}} \lesssim \|g_j\|_{L^{p_j}} t^{d(1/q_j - 1)}.$$

Applying Hölder's inequality with exponents z_2 and z_3 along with the above estimate gives

$$\int_Q |\langle g_2, \theta_{w,t} \rangle| \, |\langle g_3, \theta_{w,t} \rangle| \, \mathrm{d}w \lesssim t^{d(1-1/p_2-1/p_3)} \prod_{j=2}^{3} \|g_j\|_{L^{p_j}(Q)}.$$

Now we use the fact that $1/p_2 + 1/p_3 < (\sigma + d)/d$. This implies $\sigma + d(1 - 1/p_2 - 1/p_3) > 0$ so that

$$\begin{aligned}
\int_{T(Q)} \prod_{j=1}^{3} |\langle g_i, \theta_{w,t} \rangle| \, \frac{\mathrm{d}w \, \mathrm{d}t}{t} &\lesssim \int_0^{\ell(Q)} \frac{t^\sigma}{\ell(Q)^\sigma} \inf_{u \in Q} \mathrm{M}(1_{(3Q)^c} g_1)(u) t^{d(1-1/p_2-1/p_3)} \|g_2\|_{L^{p_2}} \|g_3\|_{L^{p_3}} \frac{\mathrm{d}t}{t} \\
&\lesssim \ell(Q)^{d(1-1/p_2-1/p_3)} \big(\inf_{Q} \mathrm{M}(1_{(3Q)^c} g_1)\big) \|g_2\|_{L^{p_2}} \|g_3\|_{L^{p_3}} \\
&= |Q| \big(\inf_{Q} \mathrm{M} g_1\big) \langle g_2 \rangle_{p_2, Q} \langle g_3 \rangle_{p_3, Q}.
\end{aligned}$$

Thus far we have shown

$$\begin{aligned}
&\mathcal{A}(f_1 1_{3Q_0}, f_2 1_{3Q_0}, f_3 1_{3Q_0}) \\
&\qquad \lesssim \sum_{Q \in \mathcal{E}} \mathcal{A}_{T(Q)}(f_1 1_{3Q}, f_2 1_{3Q}, f_3 1_{3Q}) + |Q_0| \prod_{j=1}^{3} \langle f_j \rangle_{r_j, 3Q_0} + \mathcal{A}_{T(E)^c}(f_1 1_{3Q_0}, f_2 1_{3Q_0}, f_3 1_{3Q_0})
\end{aligned}$$

We create the sparse collection by applying the same argument to each $Q \in \mathcal{E}$ as if it were Q_0; see [Conde-Alonso et al. 2017] for details. Iterating this, we will be done once we show

$$\mathcal{A}_{T(E)^c}(f_1 1_{3Q_0}, f_2 1_{3Q_0}, f_3 1_{3Q_0}) \lesssim |Q_0| \prod_{j=1}^{3} \langle f_j \rangle_{r_j, 3Q_0}. \tag{5-5}$$

Perform a Calderón–Zygmund decomposition of each

$$f_j = g_j + b_j = g_j + \sum_{Q\in\mathcal{E}} b_j^Q$$

with respect to the collection of cubes $\mathcal{E}$ and at the level $\langle f_j\rangle_{r_j,3Q_0}$. The good functions g_j are estimated using the $L^3\times L^3\times L^3$ boundedness,

$$\mathcal{A}_{T(E)^c}(g_1,g_2,g_3) \le \mathcal{A}(g_1,g_2,g_3) \lesssim |Q_0| \prod_{j=1}^3 \langle f_j\rangle_{r_j,3Q_0}.$$

The remaining terms all have at least one bad term. Let us say it is in the first argument. The functions in the other two arguments will be estimated using

$$|\langle g_j,\theta_{w,t}\rangle|,\ |\langle b_j,\theta_{w,t}\rangle| \lesssim \langle f_j\rangle_{r_j,3Q_0}, \quad (w,t)\notin T(E). \tag{5-6}$$

For the good functions, (5-6) is an obvious consequence of $\|g_j\|_\infty \le \langle f_j\rangle_{r_j,3Q_0}$. The bad one requires some work. Decompose the sum into two regions:

$$I = \{Q\in\mathcal{E} : w\notin 3Q\}, \quad II = \{Q\in\mathcal{E} : 9t > \ell(Q),\ w\in 3Q\}.$$

Since $(w,t)\notin T(E)$, we claim $\mathcal{E} = I\cup II$. Indeed, for each $Q\in\mathcal{E}$, if $w\in 3Q$, then $w\in Q'$ for some $Q'\in\mathcal{E}$ with $\ell(Q')\ge \ell(Q)/9$. Therefore $9t\ge\ell(Q)$. Considering II first,

$$\sum_{Q\in II} |\langle b_j^Q,\theta_{w,t}\rangle| \le \sum_{Q\in II} \frac{|Q|}{t^d}\langle b_i^Q\rangle_Q \le \langle f_j\rangle_{r_j,3Q_0} t^{-d} \sum_{Q\in II}|Q|.$$

But the cubes are disjoint and contained in the cube centered at w with side length $18t$. This means $\sum_{Q\in II}|Q|\lesssim t^d$. For I, the estimate immediately follows from Lemma 5.2 if we can establish

$$\sum_{Q\in I} \frac{|Q|\min\{\ell(Q),t\}^\delta}{|w-c(Q)|^{d+\delta}} \lesssim 1. \tag{5-7}$$

Let us now complete the proof, postponing (5-7) until the end. For the same reason that $\mathcal{E} = I\cup II$ above,

$$T(E)^c \subset \{w\notin 3Q\}\cup\{9t\ge\ell(Q)\} =: T^*(Q)^c$$

for any $Q\in\mathcal{E}$. Let h_j be either g_j or b_j so that by (5-6) $|\langle h_j,\theta_z\rangle|\lesssim\langle f_j\rangle_{3Q_0}$ for $z\in T(E)^c$. Using the first two statements from Lemma 5.2,

$$\begin{aligned}
&\mathcal{A}_{T(E)^c}(b_1,h_2,h_3)\\
&\lesssim \sum_{Q\in\mathcal{E}} \int_{T(E)^c} |\langle b_1^Q,\theta_z\rangle\langle h_2,\theta_z\rangle\langle h_3,\theta_z\rangle|\,\mathrm{d}\mu(z)\\
&\lesssim \langle f_2\rangle_{r_2,3Q_0}\langle f_3\rangle_{r_3,3Q_0} \sum_{Q\in\mathcal{E}} \langle b_i^Q\rangle_Q |Q| \left(\int_{\ell(Q)/9}^\infty \int_{3Q} t^{-d}\,\frac{\mathrm{d}w\,\mathrm{d}t}{t} + \int_{T^*(Q)^c} \frac{\min\{t,\ell(Q)\}^\delta}{\max\{t,|w-c(Q)|\}^{d+\delta}}\,\frac{\mathrm{d}w\,\mathrm{d}t}{t}\right)\\
&\lesssim \langle f_1\rangle_{r_1,3Q_0}\langle f_2\rangle_{r_2,3Q_0}\langle f_3\rangle_{r_3,3Q_0} \sum_{Q\in\mathcal{E}}|Q| \lesssim |Q_0|\prod_{i=1}^3\langle f_i\rangle_{3Q_0}.
\end{aligned}$$

In the third inequality, we used the fact that, for any $\delta > 0$ and Q cube,

$$\begin{aligned}\int_{T^*(Q)^c} &\frac{\min\{\ell(Q), t\}^\delta}{\max\{t, |w - c(Q)|\}^{d+\delta}} \frac{\mathrm{d}w\,\mathrm{d}t}{t} \\ &\lesssim \int_{\ell(Q)/9}^{\infty} \int_{|w-c(Q)|\le t} \frac{\ell(Q)^\delta}{t^{d+\delta}} \frac{\mathrm{d}w\,\mathrm{d}t}{t} + \int_{\ell(Q)/9}^{\infty} \int_{\{|w-c(Q)|>t\}} \frac{\ell(Q)^\delta}{|w-c(Q)|^{d+\delta}} \\ &\qquad + \int_0^{\ell(Q)/9} \int_{(3Q)^c} \frac{t^\delta}{|w-c(Q)|^{d+\delta}} \frac{\mathrm{d}w\,\mathrm{d}t}{t} \\ &\lesssim \int_{\ell(Q)}^{\infty} \frac{\ell(Q)^\delta}{t^\delta} + \int_0^{\ell(Q)} \frac{t^\delta}{\ell(Q)^\delta} \frac{\mathrm{d}t}{t} \lesssim 1.\end{aligned}$$

This is the continuous version of (5-7) so it is established. □

6. General cases

The results and arguments can be almost immediately extended to m-linear operators and the associated $(m+1)$-linear forms. At the same time, we would like to generalize to kernels which have varying degrees of smoothness in each variable. This second generalization is motivated by the fact that the assumptions in the first representation theorem were symmetric in Λ and both its adjoints. However, in the Sobolev mapping theorem, we saw that the conditions were asymmetric, and in fact some of the estimates on the adjoint terms were a bit too good. So, we give a representation theorem which is asymmetric and allows us to prove the Sobolev result under weaker assumptions. We must slightly alter the definitions above.

6A. ***Singular integrals.*** Let $\vec{1}_d = (1, \dots, 1) \in \mathbb{R}^d$. Given $\vec{\ell} = (\ell_0, \ell_1, \dots, \ell_m) \in \mathbb{N}^{m+1}$, a function $K \in L^1_{\mathrm{loc}}(\mathbb{R}^{(m+1)d} \setminus \mathbb{R}\mathbf{1}_{(m+1)d})$ is an $(\vec{\ell}, \delta)$ SI (singular integral) kernel if there exist $C, \delta > 0$ such that, for all $0 \le |\kappa| \le \ell_i$,

$$|\nabla^\kappa_{x_i} K(x_0, x_1, \dots, x_m)| \le \frac{C}{\left(\sum_{j\ne i} |x_i - x_j|\right)^{md+|\kappa|}},$$

$$|\nabla^\kappa_{x_i} \Delta^i_h K(x_0, x_1, \dots, x_m)| \le \frac{C|h|^\delta}{\left(\sum_{j\ne i} |x_i - x_j|\right)^{md+|\kappa|+\delta}}.$$

We say Λ is an $(\vec{\ell}, \delta)$ $(m+1)$-linear SI form if

$$\int_{(\mathbb{R}^d)^{m+1}} K(x_0, x_1, \dots, x_m) \prod_{j=0}^m f_j(x_j)\,\mathrm{d}x = \Lambda(\vec{f})$$

for all $\vec{f} = (f_0, f_1, \dots, f_m) \in \mathcal{S}^{m+1}$ with $\bigcap_{i=0}^m \operatorname{supp} f_i = \varnothing$ and an $(\vec{\ell}, \delta)$ SI kernel K. Notice that an $(\vec{\ell}, \delta)$ SI form is an $(\vec{\ell}', \delta')$ form for any $\vec{\ell}' \le \vec{\ell}$ and $\delta' \le \delta$.

6B. ***Calderón–Zygmund forms.*** It is useful at this point to define the adjoints of an $(m+1)$-linear form. For each $i = 0, 1, \dots, m$,

$$\Lambda^{i*}(\vec{f}) = \Lambda(f_i, f_1, f_2, \dots, f_{i-1}, f_0, f_{i+1} \dots, f_m).$$

In other words, Λ^{i*} permutes f_0 and f_i and it is clear that $\Lambda^{0*} = \Lambda$.

6B1. *Paraproducts.* We say Λ has $\vec{0}$-th order paraproducts if for each $i = 0, 1, \dots, m$, there exists b_0^i in BMO such that

$$\Lambda^{i*}(\psi, 1, 1, \dots, 1) = \langle b_0^i, \psi \rangle$$

for all $\psi \in \mathcal{S}_0$. For $\vec{j} = (j_0, j_1, \dots, j_m) \in (\mathbb{N}^m)^{m+1}$, we define the $\vec{j}$-th order paraproducts inductively. We again use the paraproduct forms, now defined for any $\gamma = (\gamma_1, \dots, \gamma_m) \in (\mathbb{N}^d)^m$ and $b \in \mathrm{BMO}$ by

$$\Pi_{b,\gamma}(\vec{f}) = \int_{Z^d} \langle b, (\partial^{-(\gamma_1+\cdots+\gamma_m)}\phi)_z \rangle \prod_{i=1}^m \langle f_i, \vartheta_z^{\gamma_i} \rangle \langle f_0, \phi_z \rangle \, \mathrm{d}\mu(z).$$

Suppose for each $i = 0, 1, \dots, m$, Λ has paraproducts b_γ^i for all $(|\gamma_1|, \dots, |\gamma_m|) < j_i$. Then, we say Λ has $\vec{j}$-th order paraproducts if for each $(|\gamma_1|, \dots, |\gamma_m|) = j_i$, there exist $b_\gamma^i \in \mathrm{BMO}$ such that, for all $\psi \in \mathcal{S}_{|j_i|}$,

$$\Lambda_{\vec{j}} := \Lambda - \sum_{i=0}^m \sum_{(|\kappa_1|,\dots,|\kappa_m|)<j_i} \Pi^{i*}_{b_\kappa^i,\kappa}$$

satisfies

$$\Lambda_{\vec{j}}^{i*}(\psi, x_1^{\gamma_1}, x_2^{\gamma_2}, \dots, x_m^{\gamma_m}) = \langle b_\gamma^i, \partial^{-(\gamma_1+\cdots+\gamma_m)}\psi \rangle.$$

Under this definition, one can verify by induction that $\Lambda_{\vec{j}}$ has vanishing paraproducts of all orders $< \vec{j}$.

Definition 6.1. Let $\vec{\ell} \in \mathbb{N}^{m+1}$ and $\vec{k} = (k_0, k_1, \dots, k_m) \in (\mathbb{N}^m)^{m+1}$, with

$$|k_i| \le \ell_i, \quad i = 0, 1, \dots, m.$$

An $(\vec{\ell}, \delta)$ SI form Λ is called a $(\vec{k}, \delta)$ CZ (Calderón–Zygmund) form if it has paraproducts up to order $\vec{k}$ and satisfies the weak boundedness property: there exists $C > 0$ such that

$$t^{md}\Lambda(\psi_z^0, \psi_z^1, \dots, \psi_z^m) \le C$$

for all $\psi_z^i \in \Psi_z^{0,\delta;1}$ supported in the ball $B(w, t)$.

6B2. *Wavelet forms.* The trilinear wavelet forms and wavelet classes must also be extended to the m-linear setting. Extending the norm $\|\cdot\|_{\star,\eta,\delta}$ to functions defined on $(\mathbb{R}^d)^m$, the wavelet classes $\Psi_z^{k,\delta;\iota}$ are the collection of all $\varphi \in C^k(\mathbb{R}^{md})$ such that

$$t^{|\gamma|}\|(\mathsf{Sy}_z^1 \cdots \mathsf{Sy}_z^m)^{-1}\partial^\gamma \varphi\|_{\star,k+\delta,\delta} \lesssim 1 \quad \text{for } \gamma \in \mathbb{N}^{dm},\ |\gamma| \le k,$$

and $\iota \in \{0, 1\}^m$ controls the cancellation in the obvious way. The main case we will need is $\iota = (1, 1, \dots, 1, 0)$, in which case φ satisfy

$$\int_{\mathbb{R}^d} x_m^{\gamma_m} \varphi(x_1, \dots, x_m) \, \mathrm{d}x_m = 0 \quad \text{for } |\gamma| \le k.$$

A (j, δ) wavelet form U_j is now defined, for some $\nu_z \in \Psi_z^{j,\delta;(1,\dots,1,0)}$, by

$$U_j(\vec{f}) = \int_{Z^d} \left\langle \bigotimes_{i=1}^m f_i, \nu_z \right\rangle \langle f_0, \phi_z \rangle \, \mathrm{d}\mu(z).$$

Theorem D. *Let $0 < \eta < \delta$ and $\vec{k} = (k_0, k_1, \dots, k_m) \in (\mathbb{N}^m)^{m+1}$. Let k_i^* be the smallest entry of k_i and let Λ be a $(\vec{k}, \delta)$ CZ form. Then there exists (j, η)-smooth wavelet forms $U_j^{i,\pi}$ and paraproduct forms Π_γ such that*

$$\Lambda(\vec{f}) = \sum_{i=0}^{m} \Bigg[\sum_{j=k_i^*}^{|k_i|} \sum_{\pi \in S^m} U_j^{i,\pi}(\pi(f_0, \dots, f_{i-1}, f_{i+1}, \dots, f_m), f_i) + \sum_{(|\gamma_1|,\dots,|\gamma_m|) \le k_i} \Pi_{b_\gamma^i, \gamma}^{i*}(\vec{f}) \Bigg]$$

for all $\vec{f} = (f_0, f_1, \dots, f_m) \in \mathcal{S}^{m+1}$.

If $m = 2$, the proof is the same as before, only the steps with *II* and *III* are carried out as if (k_1, k_2) is replaced by k_1 and $k_2 \in \mathbb{N}^2$. For larger m, we outline the necessary modifications. Again, decompose $\Lambda(\vec{f})$ using the Calderón formula (2-2) $m+1$ times to get

$$\Lambda(\vec{f}) = \int_{(Z^d)^{m+1}} \Lambda(\phi_{z_0}, \dots, \phi_{z_m}) \prod_{i=0}^{m} \langle f_i, \phi_{z_i} \rangle \, \mathrm{d}\mu(z_i).$$

Split $(Z^d)^{m+1}$ into $m+1$ regions $Z_i = \{z_i = (w_i, t_i) : t_i = \min_\ell t_\ell\}$ and each Z_i again into $Y_{i,j} = \{t_j = \min_{\ell \ne i} t_\ell\}$. On each $Y_{i,j}$ use Lemma 2.7 $m-1$ times to bring the integration down to the two scales $t_j > t_i$. In this way,

$$\Lambda(\vec{f}) = \sum_{i=0}^{m} \sum_{j \ne i} \sum_{\vec{\psi}, \tilde{\psi}} \int_{t_i > 0} \int_{t_j > t_i} \int_{(\mathbb{R}^d)^{m+1}} \Lambda^{i*}(\phi_{w_i,t_i}, \vec{\psi}) \langle f_i, \phi_{w_i,t_i} \rangle \langle f_j, \phi_{w_j,t_j} \rangle \prod_{\ell=0, \ell \ne i,j}^{m} \langle f_\ell, \tilde{\psi}_{w_\ell,t_j} \rangle \frac{\mathrm{d}w \, \mathrm{d}t_j \, \mathrm{d}t_i}{t_j t_i}.$$

Each $\vec{\psi}$ is a vector of m functions where one entry is the mother wavelet ϕ_{w_j,t_j} and the others are either $\psi^1_{w_\ell,t_j}$ or $\psi^3_{w_\ell,t_j}$ — the cancellative functions from Lemma 2.7 — and $\tilde{\psi}$ is either ψ^2 or ψ^4. This gives $(m+1) \times m \times 2^{m-1}$ terms which correspond to the 12 terms $I + II + III$ from the proof of Theorem A. Each summand is handled in the same way as σ_1 in (3-9) above. The kernel estimates and wavelet averaging lemma (Lemmas 2.4 and 3.3) can be easily reproduced in the same way as in the bilinear case.

Our extension to the nonsymmetric case generalizes results of [Frazier et al. 1988; Bényi 2003] to forms whose paraproducts of lower orders do not vanish. In particular, we obtain Sobolev bounds when the kernel only has extra smoothness in one of the $m+1$ variables.

Corollary D.1. *Let $k_0 \in \mathbb{N}^m$, $\vec{k} = (k_0, 0, \dots, 0)$, and Λ be a $(\vec{k}, \delta)$ CZ form with*

$$D^{k_0^* - |\gamma|} b_\gamma^0 \in \mathrm{BMO} \quad \text{for } |\gamma| \le k_0^*,$$

where k_0^* is the minimum entry of k_0. Then, for $\vec{p} \in P_\circ$, $p = p(\vec{p})$, and $\vec{v} \in A_{\vec{p}}$,

$$\|T(f_1, \dots, f_m)\|_{\dot{W}^{k_0^*,p}(1/v_{m+1})} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_i', p\}} \sum_{|\vec{j}| \le k_0^*} \prod_{i=1}^{m} \|f_i\|_{\dot{W}^{j_i,p_i}(v_i)},$$

$$\|T(f_1, \dots, f_m)\|_{W^{k_0^*,p}(1/v_{m+1})} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_i', p\}} \prod_{i=1}^{m} \|f_i\|_{W^{k_0^*,p_i}(v_i)}.$$

If in addition $b_\gamma^0 = 0$ for $(|\gamma_1|, |\gamma_2|, \dots, |\gamma_m|) < k_0$, then

$$\|T(f_1, \dots, f_m)\|_{\dot{W}^{k_0^*,p}(1/v_{m+1})} \lesssim [\vec{v}]_{A_{\vec{p}}}^{\max\{p_i', p\}} \sum_{j=1}^{m} \|f_j\|_{\dot{W}^{k_0^*,p_j}(v_j)} \prod_{i \ne j} \|f_i v_i\|_{\dot{L}^{p_i}(\mathbb{R}^d)}.$$

7. Comments and further questions

We first discuss the laborious definition of the paraproducts introduced here (Section 3B2) and in [Di Plinio et al. 2022]. The reader might object to this definition because, by looking at any SI form, one cannot immediately tell whether it has paraproducts of, let us say, order (1,0), even after constructing b_0 and subtracting Π_{b_0}.

It may be proposed that one may more immediately test $\Lambda(x, 1, \psi)$ than $\Lambda_{1,0}(x, 1, \psi)$. However, we do not know whether $\Lambda(x, 1, \psi)$ has anything to do with the boundedness properties of Λ. A first example is the form $\Pi_{b,0}$. As shown above, it is enough for b, $Db \in \mathrm{BMO}$ for $\Pi_{b,0} : W^{1,4} \times W^{1,4} \to W^{1,2}$. However, using the ideas of Calderón–Toeplitz operators [Rochberg 1990; Nowak 1993], it can be shown that $\Pi_{b,0}(x, 1, \psi) \sim \langle xb, \psi\rangle + \cdots$ and we see no reason why $D(xb) \in \mathrm{BMO}$ should imply b, $Db \in \mathrm{BMO}$ or vice versa. If some real connection could be realized between $\Pi_{b,0}(x, 1, \phi)$ and b, $|\nabla b|$, then we could simplify the definition of paraproducts. We also refer to [Wang 1997] where this iterative definition is avoided; however, one must pay a price in the testing condition, so that $T(x^\gamma)$ is replaced by $T((x-w)^\gamma)$ for infinitely many w.

Secondly, we would like to remark that our results may be extended to the full spectrum of smoothness spaces, say Triebel–Lizorkin and Besov scales, by simply adjusting the procedures of Section 4C to handle the corresponding smoothness norm. In fact, our framework is particularly apt to handle spaces characterized by wavelet coefficient estimates such as those of Besov or Triebel–Lizorkin type. One can obtain some negative Sobolev space results of the type $T : W^{-k,p_1}(v_1) \times W^{k,p_2}(v_2) \to W^{-k,p}(v)$ by applying our theorems to T^{*1}. Using T^{*2} would exchange the two input spaces. However, we do not know how to obtain $T : \prod_{i=1}^m W^{-k,p_i} \to W^{-k,p}$ with our methods, except when $m = 1$.

Finally, the constraint $1/r_1 + 1/r_2 < (\sigma + d)/d$ in the sparse domination result of Proposition 5.1, which was the main ingredient leading to the fractional Sobolev space bound of Theorem C, is sharp up to the equality possibly holding. Indeed, taking $f(x) = \mathrm{e}^{10ix}\phi(x)$ for $\hat{\phi} \in C_0^\infty(B(0,1))$, $D^\sigma(f\bar{f}) \sim (1+|x|)^{d+\sigma}$ for large x. For $g = 1_{B(0,2^{k+1})} - 1_{B(0,2^k)}$,

$$\langle D^\sigma(f\bar{f}), g\rangle \sim 2^{kd}2^{-k(d+\sigma)}.$$

However, if one had a sparse bound of the form $\langle D^\sigma(f\bar{f}), g\rangle \lesssim \sum_Q |Q|\langle D^\sigma f\rangle_{r_1,Q}\langle f\rangle_{r_2,Q}\langle g\rangle_{r_3,Q}$, then $\langle D^\sigma(f\bar{f}), g\rangle$ would be controlled by

$$2^{kd}\langle D^\sigma f\rangle_{r_1,B(0,2^{k+5})}\langle f\rangle_{r_2,B(0,2^{k+5})}\langle g\rangle_{r_3,B(0,2^{k+1})} \lesssim 2^{kd}2^{-kd/r_1}2^{-kd/r_2}$$

so that $d(1/r_1 + 1/r_2) \le d + \sigma$, i.e., $1/r_1 + 1/r_2 \le (d+\sigma)/d$.

References

[Alpert 1993] B. K. Alpert, "A class of bases in L^2 for the sparse representation of integral operators", *SIAM J. Math. Anal.* **24**:1 (1993), 246–262. MR Zbl

[Barron 2017] A. Barron, "Weighted estimates for rough bilinear singular integrals via sparse domination", *New York J. Math.* **23** (2017), 779–811. MR Zbl

[Beltran and Cladek 2020] D. Beltran and L. Cladek, "Sparse bounds for pseudodifferential operators", *J. Anal. Math.* **140**:1 (2020), 89–116. MR Zbl

[Bényi 2003] A. Bényi, "Bilinear singular integral operators, smooth atoms and molecules", *J. Fourier Anal. Appl.* **9**:3 (2003), 301–319. MR Zbl

[Brummer and Naibo 2018] J. Brummer and V. Naibo, "Bilinear operators with homogeneous symbols, smooth molecules, and Kato–Ponce inequalities", *Proc. Amer. Math. Soc.* **146**:3 (2018), 1217–1230. MR Zbl

[Brummer and Naibo 2019] J. Brummer and V. Naibo, "Weighted fractional Leibniz-type rules for bilinear multiplier operators", *Potential Anal.* **51**:1 (2019), 71–99. MR Zbl

[Chaffee et al. 2017] L. Chaffee, R. H. Torres, and X. Wu, "Multilinear weighted norm inequalities under integral type regularity conditions", pp. 193–216 in *Harmonic analysis, partial differential equations and applications*, edited by S. Chanillo et al., Birkhäuser, Cham, 2017. MR Zbl

[Christ and Journé 1987] M. Christ and J.-L. Journé, "Polynomial growth estimates for multilinear singular integral operators", *Acta Math.* **159**:1-2 (1987), 51–80. MR Zbl

[Conde-Alonso and Rey 2016] J. M. Conde-Alonso and G. Rey, "A pointwise estimate for positive dyadic shifts and some applications", *Math. Ann.* **365**:3-4 (2016), 1111–1135. MR Zbl

[Conde-Alonso et al. 2017] J. M. Conde-Alonso, A. Culiuc, F. Di Plinio, and Y. Ou, "A sparse domination principle for rough singular integrals", *Anal. PDE* **10**:5 (2017), 1255–1284. MR Zbl

[Cruz-Uribe and Naibo 2016] D. Cruz-Uribe and V. Naibo, "Kato–Ponce inequalities on weighted and variable Lebesgue spaces", *Differential Integral Equations* **29**:9-10 (2016), 801–836. MR Zbl

[Culiuc et al. 2017] A. Culiuc, F. Di Plinio, and Y. Ou, "A sparse estimate for multisublinear forms involving vector-valued maximal functions", pp. 168–184 in *Bruno Pini Mathematical Analysis Seminar* 2017, Bruno Pini Math. Anal. Semin. **8**, Univ. Bologna, Alma Mater Stud., Bologna, 2017. MR Zbl

[Culiuc et al. 2018] A. Culiuc, F. Di Plinio, and Y. Ou, "Domination of multilinear singular integrals by positive sparse forms", *J. Lond. Math. Soc.* (2) **98**:2 (2018), 369–392. MR Zbl

[Di Plinio et al. 2020a] F. Di Plinio, K. Li, H. Martikainen, and E. Vuorinen, "Multilinear operator-valued Calderón–Zygmund theory", *J. Funct. Anal.* **279**:8 (2020), art. id. 108666. MR Zbl

[Di Plinio et al. 2020b] F. Di Plinio, K. Li, H. Martikainen, and E. Vuorinen, "Multilinear singular integrals on non-commutative L^p spaces", *Math. Ann.* **378**:3-4 (2020), 1371–1414. MR Zbl

[Di Plinio et al. 2022] F. Di Plinio, B. D. Wick, and T. Williams, "Wavelet representation of singular integral operators", *Math. Ann.* (online publication July 2022).

[Figiel 1990] T. Figiel, "Singular integral operators: a martingale approach", pp. 95–110 in *Geometry of Banach spaces* (Strobl, 1989), edited by P. F. X. Müller and W. Schachermayer, London Math. Soc. Lecture Note Ser. **158**, Cambridge Univ. Press, 1990. MR Zbl

[Frazier et al. 1988] 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**:1 (1988), 41–72. MR Zbl

[Frazier et al. 1991] M. Frazier, B. Jawerth, and G. Weiss, *Littlewood–Paley theory and the study of function spaces*, CBMS Regional Conference Series in Mathematics **79**, Amer. Math. Soc., Providence, RI, 1991. MR Zbl

[Grafakos 2008] L. Grafakos, *Classical Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **249**, Springer, 2008. MR Zbl

[Grafakos 2017] L. Grafakos, "Fractional differentiation: Leibniz meets Hölder", pp. 17–33 in *Excursions in harmonic analysis*, *Vol.* 5, edited by R. Balan et al., Appl. Numer. Harmon. Anal., Birkhäuser, Cham, 2017. MR Zbl

[Grafakos and Oh 2014] L. Grafakos and S. Oh, "The Kato–Ponce inequality", *Comm. Partial Differential Equations* **39**:6 (2014), 1128–1157. MR Zbl

[Hart and Oliveira 2017] J. Hart and L. Oliveira, "Hardy space estimates for limited ranges of Muckenhoupt weights", *Adv. Math.* **313** (2017), 803–838. MR Zbl

[Hart et al. 2018] J. Hart, R. H. Torres, and X. Wu, "Smoothing properties of bilinear operators and Leibniz-type rules in Lebesgue and mixed Lebesgue spaces", *Trans. Amer. Math. Soc.* **370**:12 (2018), 8581–8612. MR Zbl

[Hytönen 2012] T. P. Hytönen, "The sharp weighted bound for general Calderón–Zygmund operators", *Ann. of Math.* (2) **175**:3 (2012), 1473–1506. MR Zbl

[Kato and Ponce 1988] T. Kato and G. Ponce, "Commutator estimates and the Euler and Navier–Stokes equations", *Comm. Pure Appl. Math.* **41**:7 (1988), 891–907. MR Zbl

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg–de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR Zbl

[Lacey 2017] M. T. Lacey, "An elementary proof of the A_2 bound", *Israel J. Math.* **217**:1 (2017), 181–195. MR Zbl

[Lerner 2013] A. K. Lerner, "A simple proof of the A_2 conjecture", *Int. Math. Res. Not.* **2013**:14 (2013), 3159–3170. MR Zbl

[Lerner 2016] A. K. Lerner, "On pointwise estimates involving sparse operators", *New York J. Math.* **22** (2016), 341–349. MR Zbl

[Lerner and Nazarov 2019] A. K. Lerner and F. Nazarov, "Intuitive dyadic calculus: the basics", *Expo. Math.* **37**:3 (2019), 225–265. MR Zbl

[Lerner et al. 2009] A. K. Lerner, S. Ombrosi, C. Pérez, R. H. Torres, and R. Trujillo-González, "New maximal functions and multiple weights for the multilinear Calderón–Zygmund theory", *Adv. Math.* **220**:4 (2009), 1222–1264. MR Zbl

[Li et al. 2014] K. Li, K. Moen, and W. Sun, "The sharp weighted bound for multilinear maximal functions and Calderón–Zygmund operators", *J. Fourier Anal. Appl.* **20**:4 (2014), 751–765. MR Zbl

[Li et al. 2019] K. Li, H. Martikainen, Y. Ou, and E. Vuorinen, "Bilinear representation theorem", *Trans. Amer. Math. Soc.* **371**:6 (2019), 4193–4214. MR Zbl

[Li et al. 2020] K. Li, J. M. Martell, and S. Ombrosi, "Extrapolation for multilinear Muckenhoupt classes and applications", *Adv. Math.* **373** (2020), art. id. 107286. MR Zbl

[Li et al. 2021] K. Li, J. M. Martell, H. Martikainen, S. Ombrosi, and E. Vuorinen, "End-point estimates, extrapolation for multilinear Muckenhoupt classes, and applications", *Trans. Amer. Math. Soc.* **374**:1 (2021), 97–135. MR Zbl

[Maldonado and Naibo 2009] D. Maldonado and V. Naibo, "On the boundedness of bilinear operators on products of Besov and Lebesgue spaces", *J. Math. Anal. Appl.* **352**:2 (2009), 591–603. MR Zbl

[Muscalu and Schlag 2013] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis*, *Vol. I*, Cambridge Studies in Advanced Mathematics **137**, Cambridge University Press, 2013. MR Zbl

[Naibo and Thomson 2019] V. Naibo and A. Thomson, "Coifman–Meyer multipliers: Leibniz-type rules and applications to scattering of solutions to PDEs", *Trans. Amer. Math. Soc.* **372**:8 (2019), 5453–5481. MR Zbl

[Nazarov et al. 2003] F. Nazarov, S. Treil, and A. Volberg, "The Tb-theorem on non-homogeneous spaces", *Acta Math.* **190**:2 (2003), 151–239. MR Zbl

[Nieraeth 2019] Z. Nieraeth, "Quantitative estimates and extrapolation for multilinear weight classes", *Math. Ann.* **375**:1-2 (2019), 453–507. MR Zbl

[Nowak 1993] K. Nowak, "On Calderón–Toeplitz operators", *Monatsh. Math.* **116**:1 (1993), 49–72. MR Zbl

[Petermichl 2007] 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**:5 (2007), 1355–1375. MR Zbl

[Rahm et al. 2021] R. Rahm, E. T. Sawyer, and B. D. Wick, "Weighted Alpert wavelets", *J. Fourier Anal. Appl.* **27**:1 (2021), art. id. 1. MR Zbl

[Rochberg 1990] R. Rochberg, "Toeplitz and Hankel operators, wavelets, NWO sequences, and almost diagonalization of operators", pp. 425–444 in *Operator theory*: *operator algebras and applications*, *Part* 1 (Durham, NH, 1988), edited by W. B. Arveson and R. G. Douglas, Proc. Sympos. Pure Math. **51**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[Wang 1997] K. Wang, "The generalization of paraproducts and the full $T1$ theorem for Sobolev and Triebel–Lizorkin spaces", *J. Math. Anal. Appl.* **209**:2 (1997), 317–340. MR Zbl

[Yulmukhametov 1999] R. S. Yulmukhametov, "Solution of the L. Ehrenpreis problem on factorization", *Mat. Sb.* **190**:4 (1999), 123–157. In Russian; translated in *Sb. Math.* **190**:4 (1999), 597–629. MR Zbl

Received 10 Jun 2021. Revised 13 Jul 2022. Accepted 26 Aug 2022.

FRANCESCO DI PLINIO: francesco.diplinio@unina.it
Dipartimento di Matematica e Applicazioni, Università di Napoli "Federico II", Napoli, Italy

WALTON GREEN: awgreen@wustl.edu
Department of Mathematics and Statistics, Washington University in St. Louis, St. Louis, MO, United States

BRETT D. WICK: bwick@wustl.edu
Department of Mathematics and Statistics, Washington University in St. Louis, St. Louis, MO, United States

PURE and APPLIED ANALYSIS

msp.org/paa

Cover image: The figure shows the outgoing scattered field produced by scattering a plane wave, coming from the northwest, off of the (stylized) letters P A A. The total field satisfies the homogeneous Dirichlet condition on the boundary of the letters. It is based on a numerical computation by Mike O'Neil of the Courant Institute.

See inside back cover or msp.org/paa for submission instructions.

The subscription price for 2023 is US $575/year for the electronic version, and $640/year (+$30, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Pure and Applied Analysis (ISSN 2578-5885 electronic, 2578-5893 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

PAA peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PURE and APPLIED ANALYSIS

vol. 5 no. 1 2023