Wavelet representation of singular integral operators

Francesco Di Plinio[1] · **Brett D. Wick[2]** · **Tyler Williams[2]**

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

This article develops a novel approach to the representation of singular integral operators of Calderón–Zygmund type in terms of continuous model operators, in both the classical and the bi-parametric setting. The representation is realized as a finite sum of averages of wavelet projections of either cancellative or noncancellative type, which are themselves Calderón–Zygmund operators. Both properties are out of reach for the established dyadic-probabilistic technique. Unlike their dyadic counterparts, our representation reflects the additional kernel smoothness of the operator being analyzed. Our representation formulas lead naturally to a new family of $T(1)$ theorems on weighted Sobolev spaces whose smoothness index is naturally related to kernel smoothness. In the one parameter case, we obtain the Sobolev space analogue of the A_2 theorem; that is, sharp dependence of the Sobolev norm of T on the weight characteristic is obtained in the full range of exponents. In the bi-parametric setting, where local average sparse domination is not generally available, we obtain quantitative A_p estimates which are best known, and sharp in the range $\max\{p, p'\} \geq 3$ for the fully cancellative case.

Mathematics Subject Classification. Primary 42B20; Secondary 42B25

✉ Brett D. Wick
  bwick@wustl.edu

  Francesco Di Plinio
  francesco.diplinio@unina.it

  Tyler Williams
  tgwilliams@wustl.edu

[1] Dipartimento di Matematica e Applicazioni Università degli Studi di Napoli "Federico II" Via Cintia, Monte S. Angelo I-80126 Napoli, Italy

[2] Department of Mathematics, Washington University in Saint Louis, 1 Brookings Drive, Saint Louis, MO 63130, USA

1 Introduction

The class of Calderón–Zygmund singular integrals may succinctly be described as consisting of those linear operators bounded on $L^p(\mathbb{R}^d)$ for some $1 < p < \infty$, whose Schwartz kernel off the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$ satisfies the same size and smoothness properties enjoyed by the kernels of the Hilbert ($d = 1$) or Riesz transforms. These properties, and thus the Calderón–Zygmund class, are quantitatively preserved by translations and dilations of $\mathbb{R}^d$.

Calderón–Zygmund singular integrals play a pivotal role in regularity theory of elliptic PDEs, operator theory, and differentiation theory. In particular, their quantified and sharp weighted norms have been the keystone in the solution to two important problems in the theory of distortion by quasiconformal mappings. Sharp weighted bounds for the Ahlfors–Beurling transform have been used by Petermichl and Volberg [52] to prove injectivity at the critical exponent of the solution operator to the Beltrami equation, completing the scheme devised in [2] by Astala, Iwaniec and Saksman. A sharp estimate of the Beurling transform on suitable fractal-dimensional weights has been employed by Lacey, Sawyer and Uriarte-Tuero [34] to prove Astala's conjecture on distortion of Hausdorff measures under quasiconformal maps.

Beginning with Figiel's $T(1)$ theorem [22], through the seminal work of Nazarov, Treil and Volberg [48] on two-weight inequalities for non-homogeneous Haar shifts, the study of singular integrals and their sharp weighted bounds has relied on dyadic-probabilistic techniques. An important change in perspective within this line of investigation was brought by the work of Petermichl [50] where the Hilbert transform was *represented*, instead of *estimated*, as an average of dyadic shift operators. This idea was extended to the Beurling transform, Riesz transform, and other nice convolution type Calderón–Zygmund operators in [15, 51, 59]. A forceful augmentation of this strategy is Hytönen's proof of the A_2 conjecture [29], which relied on the representation of a Calderón–Zygmund operator as a probabilistic average of shifted dyadic operators, the simplest of which is the dyadic martingale transform. Hytönen's representation theorem has since been extended to bi-parameter Calderón–Zygmund operators by Martikainen in [41] and to the multi-parameter setting by Ou [49], resulting in $T(1)$ type theorems in two and higher parameters.

While the significance of dyadic representation theorems cannot be overstated, dyadic-probabilistic realizations of Calderón–Zygmund operators suffer from certain intrinsic drawbacks, originating from the discrete nature of the Haar basis employed. A first one is that the representation formula contains dyadic shifts of arbitrarily large *complexity* parameter: complexity may be described, roughly, as the width of the band of the matrix associated to the dyadic shift in the Haar basis. The representation formula itself involves a rather delicate averaging procedure over shifted dyadic grids, and, for each shifting parameter, a countable collection of shifts of unbounded complexity. Explicitly computing the dyadic components of a generic singular integral operator is thus not feasible.

Secondly, the dyadic representation formula, due to the roughness of the basis, fails to detect additional kernel smoothness. The latter is often relevant for the behavior of a singular integral on smoother spaces, of e.g. Sobolev or Besov type. We note that the recent article [28] partially addresses this question within the framework of

dyadic representation theorems, replacing the Haar basis with a smoother wavelet but keeping the same structure involving shifted grids. The comparison with [28] is elaborated upon in Remark 4.13.

Bi-parameter singular integrals on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ may be informally defined as elements of the closed convex hull of the set of tensor products $T_1 \otimes T_2$, where each T_j is a $\mathbb{R}^{d_j}$-singular integral operators as above. This class arises naturally in connection with the theory of bi-harmonic functions [19, 21] for instance in the weak factorization of functions in the Hardy space on the bi-disk [20]. The L^p and mixed norm estimates for their multilinear analogues are at the root of partial fractional Leibniz rules [46, 47], namely, anisotropic variants of the bilinear estimates popularized e.g. by Kato–Ponce [31] in connection with the Navier–Stokes, Schrödinger and KdV equations.

One specific drawback of dyadic representations in the bi-parameter context, see [41, 49] for instance, is that they do not reduce L^p and weighted estimates for the analyzed operator to a single model operator whose weighted theory is significantly simpler. In the one parameter case, this can be verified directly for shift operators as in [33] and can also be done by means of sparse operators as in [13]. In higher parameters, the approach via direct verification is challenging [27, 40] and domination by local average sparse operators are generally not available as the counterexample of [3] shows. Thus, for instance, one cannot expect precise information on the dependence of $\|T\|_{L^p(w)}$ on the corresponding relevant weight characteristic.

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

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

where $\varphi_{(x,s)}$ are L^1-normalized wavelets living at scale s near the point x. The method used is to instead take a weighted average of these wavelet projections with respect to the operator wavelet coefficients $\langle T\varphi_{(x,s)}, \varphi_{(y,t)} \rangle$, with $(y, t) \in \mathbb{R}^d \times (0, \infty)$. The simplest version of this principle is the resolution of the identity operator (2.11) below, widely known as the Calderón reproducing formula. Certainly, the Calderón reproducing formula (2.11) and wavelet coefficients have been used countless times in the proof of $T(1)$ type estimates, beginning with the works of David and Journé [14] and Journé in the bi-parametric setting [30]. Our approach takes these seminal ideas one step further, in that we aim for equalities, rather than inequalities, and employ the wavelet coefficients of T in a wavelet averaging procedure instead of estimating them, see Lemma 3.3, essentially turning the original wavelet basis into another wavelet family adapted to the operator being analyzed. Our approach takes advantage of the fact that a Calderón–Zygmund operator applied to a smooth wavelet basis with compact support yields again a collection of wavelets, though possibly rougher and with smeared out support.

When analyzing one parameter Calderón–Zygmund operators, this results in a representation formula involving a single, *complexity zero* cancellative operator, a single paraproduct and a single adjoint paraproduct, all of which are Calderón–Zygmund operators themselves, in contrast to the dyadic expansion of e.g. [29], involving probabilistic averaging of countably many dyadic shifts. The following result is a loosely

stated excerpt of Theorem A, Sect. 4, which is referred to for a rigorous and detailed statement, in particular concerning the formal definition of T.

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

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

where $\varphi_{(y,t)}$ is the (y, t)-rescaling of a smooth mother cancellative wavelet with compact support, Π_b are explicitly constructed paraproducts of the form in Definition 4.4.

If additional smoothness of the kernel is available, the cancellative portion will be made of smoother and faster decaying wavelets, provided that additional paraproducts are subtracted. In turn, this refinement may be employed to obtain sharp weighted $T(1)$ bounds on Sobolev spaces. A sample is given in the theorem below, which is a simplified version of Corollary A.1. The case $k = 0$ has no restriction on paraproducts, thus it is yet another proof of the sharp weighted bound for Calderón–Zygmund operators by Hytönen [29]. For the case $k > 0$, the restriction $T(x^\gamma) = 0$ for $0 \le |\gamma| < k$ is necessary, see Remark 4.14 where the relation with [55, Theorem 1.1] is also expounded.

Theorem *Let T be a linear operator on $\mathbb{R}^d$, satisfying the weak boundedness testing condition of Definition 4.1, whose kernel has k-th Hölder continuous derivatives in the sense of Definition 4.2, and whose κ-th order paraproducts are in BMO$(\mathbb{R}^d)$ for $0 \le \kappa \le k$, according to Definition 4.7. If $k > 0$, assume in addition that $T(x^\gamma) = 0$ for all multi-indices γ on $\mathbb{R}^d$ with $0 \le |\gamma| < k$. Then T has the sharp weighted bound on the homogeneous L^2-Sobolev space*

$$\|Tf\|_{\dot{W}^{k,2}(\mathbb{R}^d;w)} \lesssim [w]_{A_2} \|f\|_{\dot{W}^{k,2}(\mathbb{R}^d;w)}.$$

Corollary A.1 may be seen as the A_2-theorem on Sobolev spaces, in analogy with the celebrated A_2 theorem of Hytönen [29]. Past results in the unweighted setting, and their relation with Corollary A.1, are recalled in Remark 4.15. Several unweighted $T(1)$ theorems of this type have been developed and used in connection with regularity of the Beltrami solution operator and smoothness of the Beurling transform on $C^{k,\alpha}$-type domains, see e.g. [10, 11, 53, 54]. In analogy with the application of the Petermichl–Volberg theorem [52] to the scheme of [2] we expect that suitable versions of Corollary A.1 can be employed to simplify and sharpen the higher order Beltrami regularity results of e.g. [54]. This will be carried out in future work.

The proof techniques may be naturally transported to the bi-parametric dilation setting. We again paraphrase the main result and point the reader to Sect. 6, Theorem B, for the precise statement.

Theorem *Let T be a linear operator satisfying the hypotheses of a bi-parameter δ-Calderón–Zygmund operator as in Sect. 6. Let $0 < \varepsilon < \delta \leq 1$. Then there exists a family of L^1-adapted, ε-smooth and $(d_j + \varepsilon)$-decaying in the jth parameter, product cancellative wavelets*

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

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

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

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

Notice that unlike the one parameter results recalled in Remark 4.15, no smooth $T(1)$ type theorems have appeared in the literature before, even in the unweighted case. This result may be contrasted with the bi-parameter dyadic representation theorem of Martikainen [41]. The assumptions on T are of the same nature as the ones appearing in [41], namely weak boundedness, full and partial kernel estimates, paraproducts in product BMO. However, we drop the diagonal BMO conditions appearing in [41] which are subsumed by a combination of the other assumptions. While this article focuses for simplicity on the bi-parameter case, multiple parameter versions can in principle be obtained by imposing a suitable hierarchy of weak-boundedness, full and partial kernel conditions and defining the corresponding partial paraproducts. This will be pursued in future work.

In addition to the simpler, and more computationally feasible nature of the continuous formula, the model operators we obtain have a much simpler weighted theory, which allows for quantitative, and sharp in certain cases, weighted norm inequalities for T.

Theorem *Let T be a (k_1, k_2)-smooth bi-parameter Calderón–Zygmund operator, see Sect. 6. If $(k_1, k_2) \neq (0, 0)$ assume in addition the paraproduct condition (6.18). For all $1 < p < \infty$ product A_p-weights w on $\mathbb{R}^{\mathbf{d}} := \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ there holds*

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

If T is fully cancellative, the improved estimate

$$\|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} Tf\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim_{k,\delta} [w]_{A_p}^{\theta(p)} \|\nabla_{x_1}^{k_1} \nabla_{x_2}^{k_2} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)},$$

$$\theta(p) = \begin{cases} \frac{2}{p-1} & 1 < p \leq \frac{3}{2} \\ \text{see (6.20)} & \frac{3}{2} < p < 3 \\ 2 & p \geq 3 \end{cases}$$

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

The above result, precisely stated in Corollary B.1, generalizes and quantifies R. Fefferman's qualitative weighted bounds for bi-parameter Journé-type operators [17]. While Martikainen's work [41] did not contain weighted $T(1)$-type implications, a simplified proof of Fefferman's result was recently obtained in [27] relying on the representation from [41]. Some quantitative estimates, weaker[1] than the ones of Corollary B.1, have been obtained in [4] by a shifted square function form-type domination for cancellative Journé operators, also relying on [41] within the proof. At present, it does not seem possible to match the quantification obtained in Corollary B.1 using dyadic representation theorems in the vein of [41, 49]. Part of the challenge with this is that one parameter proofs of the quantitative results typically rely upon some variant of stopping time (sparse operators, weak-type (1,1), or Bellman functions) as a key ingredient and not easily adaptable to the bi-parameter setting. Our analysis based on square function methods is able to circumvent this issue at least for $\max\{p, p'\} \geq 3$.

Qualitative estimates for bi-linear, bi-parameter paraproduct free singular integrals have been obtained very recently in [40]. We expect that multilinear variants of our result can quantify and extend the scope of the weighted inequalities of [40]. The method used in this paper has an additional benefit over dyadic representation. The dyadic representation theorems rely heavily upon the Haar basis, making it difficult to work in the settings of other Calderón–Zygmund operators that respect different dilation structures, such as Zygmund-type dilations [18, 44, 45]. Our method is easily adaptable towards representation and weighted $T(1)$ theorems for more general dilation structures; this will be pursued in future work.

Structure

Section 2 contains the definition of the one parameter wavelet classes and introduces the related intrinsic square function. Section 3 provides technical lemmas in the one parameter setting, describing the averaging procedure of the wavelet basis and setting up an auxiliary Alpert basis to handle higher degree paraproducts. Section 4 gives the statement and proof of the one parameter smooth representation theorem and weighted Sobolev $T(1)$ corollary, see Theorem A and Corollary A.1. Sections 5 through 7 set forth the statement and proof of the bi-parameter versions, Theorem B and Corollary B.1. The concluding Sect. 8 is devoted to establishing weighted norm inequalities for the model operators appearing in the bi-parameter representation, via new weighted bounds for intrinsic operators such as the bi-parameter intrinsic square function.

Notation

The symbol $C = C(a_1, \ldots, a_n)$ and the constant implied by almost inequality and comparability symbols $\lesssim_{a_1,\ldots,a_n}, \sim_{a_1,\ldots,a_n}$ are meant to depend on the parameters $a_1, \ldots, a_n$ only and may vary at each occurrence without explicit mention. It is con-

---

[1] The exponent of the A_2 constant obtained in [4, Corollary 3.2] is 10, in contrast with power 3 obtained in (6.20). Therein, it is claimed that tracking the constants in the argument of [27] yields power 8.

venient to employ the Japanese bracket symbol

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

The Fourier transform of $f \in \mathcal{S}(\mathbb{R}^d)$ is normalized as

$$\widehat{f}(\xi) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} \, dx, \quad \xi \in \mathbb{R}^d.$$

2 Wavelet classes and the intrinsic square function

In this section we introduce the normalized classes of wavelets with limited decay that appear in our wavelet representation theorem for Calderón–Zygmund forms on $\mathbb{R}^d$.

2.1 Analysis in the symmetry parameter space

Our analysis of these forms is based on a symmetry parameter space description. In the classical, single-parameter setting of Section 4, the parameter space and its associated natural measure μ are

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

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

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

When two points, or families of points, of Z^d appear in the same statement and the context allows for it, the notation $z = (x, s)$ and the corresponding Greek version $\zeta = (\xi, \sigma)$ are used; for instance, see (2.5) below. For each $\zeta = (\xi, \sigma) \in Z^d$ it is convenient to refer to the following partition of Z^d:

$$
\begin{aligned}
Z_+^d(\zeta) &:= \{z = (x, s) \in Z^d : s \geq \sigma\} = F_+(\zeta) \sqcup S(\zeta) \sqcup A(\zeta), \\
F_+(\zeta) &:= \{z = (x, s) \in Z_+^d(\zeta) : |x - \xi| > 3s\}, \\
S(\zeta) &:= \{z = (x, s) \in Z_+^d(\zeta) : s \in [\sigma, 3\sigma], |x - \xi| < 3s\}, \\
A(\zeta) &:= \{z = (x, s) \in Z_+^d(\zeta) : s > 3\sigma, |x - \xi| < 3s\}.
\end{aligned}
\tag{2.1}
$$

2.1.1 Symmetries parametrized by $z \in Z^d$

Throughout, the one parameter family of symmetries on $\phi \in \mathcal{S}(\mathbb{R}^d)$ is defined as

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

2.1.2 Cutoffs

Choose $\alpha \in \mathcal{C}^\infty(\mathbb{R}^d)$, radial and with $\alpha = 1$ on $\mathsf{B}_{(0,2)}$, supp $\alpha \subset \mathsf{B}_{(0,4)}$, and accordingly define the cutoffs

$$\alpha_z := \operatorname{Tr}_x \operatorname{Dil}_s^\infty \alpha, \quad \beta_z = 1 - \alpha_z, \quad z = (x, s) \in Z^d. \tag{2.2}$$

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

2.1.3 Measuring decay in Z^d

For a decay parameter $\nu > 0$, define the function

$$[\cdot]_\nu : Z^d \to (0, 1], \quad [(x, s)]_\nu := \frac{(\min\{1, s\})^\nu}{(\max\{1, s, |x|\})^{d+\nu}}. \tag{2.3}$$

Throughout the article, the fact that

$$\int_{Z^d} [z]_\nu \, d\mu(z) \lesssim_\nu 1, \quad \nu > 0 \tag{2.4}$$

will be heavily exploited. The geometric separation in the parameter space Z^d is then described by the function

$$[z, \zeta]_\nu = \frac{1}{s^d} \left[\left(\frac{\xi - x}{s}, \frac{\sigma}{s} \right) \right]_\nu$$
$$= \frac{(\min\{s, \sigma\})^\nu}{(\max\{s, \sigma, |x - \xi|\})^{d+\nu}}, \quad z = (x, s), \ \zeta = (\xi, \sigma) \in Z^d. \tag{2.5}$$

2.1.4 Integration by parts

Throughout the paper, for $k \in \mathbb{N}$, a multi-index $\gamma \in \mathbb{R}^d$ with $|\gamma| = k$ and a Schwartz function f with $\widehat{f}(\xi) = O(|\xi|^k)$ as $\xi \to 0$, we denote

$$\partial^{-\gamma} f(x) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \frac{(i\xi)^\gamma}{|\xi|^{2k}} e^{ix \cdot \xi} \, d\xi. \tag{2.6}$$

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

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

Remark 2.1 Let $\kappa \in \mathbb{R}$, $u \in \{1, \ldots, d\}$ and $|\nabla|^{\kappa}$, R_u be respectively the κ-th order Riesz potential and uth Riesz transform on $\mathbb{R}^d$,

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

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

For a multi-index $\gamma = (\gamma_1, \ldots, \gamma_d)$, let $R^{\gamma} = R_1^{\gamma_1} \circ \cdots \circ R_d^{\gamma_d}$. With this notation $\partial^{-\gamma} = |\nabla|^{-|\gamma|} R^{\gamma}$ up to a multiplicative constant depending on $d, |\gamma|$ only. This multiplicative constant will be ignored in the subsequent uses of this remark.

2.1.5 Mother wavelet

Let $\Phi \in \mathcal{C}^{\infty}(\mathbb{R}^d)$ be radial and supported on $\mathsf{B}_{(0,1)}$, $D \in \mathbb{N}$ a fixed large parameter,[2] and $a = a(d, D) > 0$ chosen so that (2.11) below holds. Define the mother wavelet φ by

$$\varphi := a\Delta^{4D}\Phi \in \mathcal{C}^{\infty}(\mathbb{R}^d), \quad \operatorname{supp}\varphi \subset \mathsf{B}_{(0,1)}, \quad \varphi \text{ radial}, \quad \int_{\mathbb{R}^d} |\varphi| \, dx = C(d, D). \tag{2.7}$$

This definition implies

$$\partial^{-\alpha}\varphi = a\partial^{\alpha}\Delta^{4D-|\alpha|}\Phi \in \mathcal{S}(\mathbb{R}^d), \quad \operatorname{supp}\partial^{-\alpha}\varphi \subset \mathsf{B}_{(0,1)}, \quad \forall 0 \leq |\alpha| \leq D, \tag{2.8}$$

and in particular

$$\int_{\mathbb{R}^d} x^{\gamma} \psi(x) \, dx = 0 \tag{2.9}$$

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

$$\varphi_z = \mathsf{Sy}_z\varphi \qquad z \in \mathbb{Z}^d \tag{2.10}$$

---

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

yield the Schwartz version of the Calderón reproducing formula [6, 24, 62]

$$h = \int_{Z^d} \langle h, \varphi_\zeta \rangle \varphi_\zeta \, d\mu(\zeta), \qquad h \in \mathcal{S}(\mathbb{R}^d). \tag{2.11}$$

2.2 Wavelet classes

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

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

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

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

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

$$|\partial^\gamma \phi(y)| \leq \frac{1}{s^{d+|\gamma|}} \left\langle \frac{y-x}{s} \right\rangle^{-(d+k+\delta)}, \quad y \in \mathbb{R}^d; \tag{2.12}$$

$$|\partial^\gamma \phi(y+h) - \partial^\gamma \phi(y)| \leq \frac{|h|^\delta}{s^{d+|\gamma|+\delta}} \left\langle \frac{y-x}{s} \right\rangle^{-(d+k+\delta)}, \quad y \in \mathbb{R}^d, h \neq 0. \tag{2.13}$$

Then set

$$\Psi_z^{k,\delta;0} := \left\{ \psi \in \Psi_z^{k,\delta;1} : (2.9) \text{ holds } \forall 0 \leq |\gamma| \leq k \right\}. \tag{2.14}$$

When $k = 0$, the notation is simplified by writing $\Psi_z^{\delta;1}, \Psi_z^{\delta;0}$ in place of $\Psi_z^{0,\delta;1}, \Psi_z^{0,\delta;0}$.

Remark 2.2 Note that φ_z defined in (2.10) belongs to $C\Psi_z^{D,1;0}$. More generally if $0 \leq |\gamma| \leq D$

$$\varphi_{\gamma,z} := \mathsf{Sy}_z[\partial^{-\gamma}\varphi] = s^{-|\gamma|} \partial^{-\gamma} \mathsf{Sy}_z \varphi \in C\Psi_z^{D,1;0}, \quad \operatorname{supp} \varphi_{\gamma,z} \subset B_z. \tag{2.15}$$

Limited decay wavelets enjoy the following almost-orthogonality estimates.

Lemma 2.3 *Let* $0 < \eta < \delta \leq 1, 0 \leq k \leq D, z = (x, s), \zeta = (\xi, \sigma) \in Z^d$ *with* $s \leq \sigma$. *Then*

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

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

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

Using (2.13) for $\nabla^k \phi$ and (2.12) for ψ,

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

Using (2.12) for $\nabla^k \phi$ instead gives,

$$\int_{|y-x|\geq 1} |T_x(y)||\psi(y)| \, dy \lesssim \frac{s^{k+\delta}}{|x|^{d+k+\delta}} \int_{|y-x|>1} |y-x|^{-(d+\delta)} dy \lesssim [z, \zeta]_{k+\eta},$$

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

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

Assembling the last two displays yields the claimed first estimate.

The second estimate is proved similarly. Again, renormalize to have $\zeta = (0, 1)$, $z = (x, s)$ with $s \leq 1$. Now observe that if $s \leq 1$, $|x| \geq 1$ and $|y - x| \geq 1$, then

$$|\psi(y)| \leq \frac{1}{s^d} \frac{1}{\max\{1, \frac{|y-x|}{s}\}^{d+k+\delta}} = \frac{1}{s^d} \frac{1}{\left(\frac{|x-y|}{s}\right)^{d+k+\delta}}$$

And since $|x - y| \geq \frac{|x|}{2}$ we can further have the estimate

$$|\psi(y)| \leq \frac{s^{k+\delta}}{|x|^{d+k+\delta}}.$$

We also have that $|\phi(y)| \lesssim \frac{1}{(1+|y|)^{d+k+\delta}}$ and to $\|\phi\|_{L^1} \lesssim 1$. With these estimates it is clear that the second estimate above holds and the proof is complete.

Taking advantage of Remark 2.2 to rely on the vanishing mean of $\varphi_{\gamma,z}$, and using the decay (2.13) for $\phi \in \Psi_\zeta^{k,\delta;1}$ gives

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

and the proof is complete. $\square$

2.3 Intrinsic forms and sparse estimates

Lemma 2.3 leads to the L^2-boundedness of an intrinsic square function associated to the classes $\Psi_z^{\delta;0}$. This square function will now be defined. For $f \in L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, $\iota \in \{0, 1\}$ and $z \in Z^d$ define the intrinsic wavelet coefficients

$$\Psi_z^{\delta;\iota} f := \sup_{\psi \in \Psi_z^{\delta;\iota}} |\langle f, \psi \rangle|. \tag{2.16}$$

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

$$\mathsf{S}_\delta f(x) = \left(\int_0^\infty \left(\Psi_{(x,s)}^{\delta;0} f \right)^2 \frac{ds}{s} \right)^{\frac{1}{2}}. \tag{2.17}$$

The value $\delta > 0$ is fixed but arbitrary and, whenever possible, it will be omitted from the notation in (2.16) and (2.17), writing for instance $\mathsf{S}f$ instead of $\mathsf{S}_\delta f$. Lemma 2.3 implies easily the L^2 estimate of the following proposition.

Proposition 2.4 *Let $f \in L^2(\mathbb{R}^d)$. Then $\|\mathsf{S}_\delta f\|_2 \lesssim_\delta \|f\|_2$.*

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

$$\int_{(x,s)\in\mathbb{R}^d\times(0,\infty)} \int_{(\alpha,\beta)\in\mathbb{R}^d\times(0,1)} |\langle f, \psi_{(x,s)} \rangle| |\langle s^d \psi_{(x,s)}, \psi_{(x+\alpha s, \beta s)} \rangle| |\langle f, \psi_{(x+\alpha s, \beta s)} \rangle|$$

$$\times \frac{dx \, ds \, d\alpha \, d\beta}{s\beta} \lesssim \int_{z\in Z^d} |\langle f, \psi_z \rangle|^2 \, d\mu(z) \tag{2.18}$$

with implied constant uniform over the choice of $\psi_z \in \Psi_z^{\delta;0}$ and $z \in Z^d$. Lemma 2.3 then yields

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

so that (2.18) follows by an application of Cauchy–Schwarz and (2.4). □

Next, consider the intrinsic bisublinear forms

$$\Psi^\delta(f_1, f_2) = \int_{Z^d} \Psi_z^{\delta;0} f_1 \cdot \Psi_z^{\delta;0} f_2 \, d\mu(z) \tag{2.19}$$

acting on pairs $f_1, f_2 \in \bigcup_{1 \le p \le \infty} L^p(\mathbb{R}^d)$, modeling cancellative operators of Calderón–Zygmund type. Analogous intrinsic forms modeling paraproducts are also needed in the analysis. Referring to the wavelet coefficients (2.16), define on triples $f_j \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ the intrinsic forms

$$\Pi^{j,\delta}(f_1, f_2, f_3) = \int_{Z^d} \Psi_z^{\delta;1} f_j \left(\prod_{\iota \in \{1,2,3\} \setminus j} \Psi_z^{\delta;0} f_\iota \right) d\mu(z). \tag{2.20}$$

The index j in the notation (2.20) identifies the noncancellative index of the paraproduct form. As these are modeling bilinear, one parameter Calderón–Zygmund forms, the case where $j \in \{1, 2\}$ and $f_3 \in \mathrm{BMO}(\mathbb{R}^d)$ is of particular interest. In this case, the simplified notation

$$\pi_{f_3}^\delta(f_1, f_2) = \Pi^{1,\delta}(f_1, f_2, f_3) \tag{2.21}$$

is adopted. Note that the integrands in (2.19)–(2.20) are nonnegative, therefore issues of convergence in the definition may be disregarded.

The next pair of estimates are of basic nature.

Proposition 2.5 *For all $\delta > 0$, $1 < p < \infty$, there holds*

$$\Psi^\delta(f_1, f_2) \lesssim_\delta \frac{p^2}{p-1} \|f_1\|_p \|f_2\|_{p'}, \tag{2.22}$$

$$\pi_{f_3}^\delta(f_1, f_2) \lesssim_\delta \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)} \frac{p^2}{p-1} \|f_1\|_p \|f_2\|_{p'}. \tag{2.23}$$

Proof By standard limiting arguments, it suffices to prove either estimate for $f_1, f_2 \in L_0^\infty(\mathbb{R}^d)$. We argue for (2.22) first. Having fixed such a pair f_1, f_2, there exists a compact set $K \subset Z^d$ and a z-measurable choice of $\psi_z, \upsilon_z \in \Psi_z^{0,\delta;0}$ so that

$$\Psi^\delta(f_1, f_2) \le \frac{1}{2} \Lambda(f_1, f_2), \quad \Lambda(f_1, f_2) := \int_K \langle f_1, \varphi_z\rangle \langle \upsilon_z, f_2\rangle \, d\mu(z).$$

The form Λ with its adjoint satisfies the standard η-smooth Calderón–Zygmund kernel estimates for all $0 < \eta < \delta$, see Sect. 4.2 below. By standard CZ decomposition,

interpolation and duality, (2.22) follows from

$$\Lambda(f_1, f_2) \lesssim \langle S_\delta f_1, S_\delta f_2 \rangle \leq \|S_\delta f_1\|_2 \|S_\delta f_2\|_2 \lesssim_\delta \|f_1\|_2 \|f_2\|_2,$$

which is a consequence of Proposition 2.4. Similarly, to prove (2.23) it suffices to show that

$$\left| \Pi_{f_3}(f_1, f_2) := \int_K \langle f_3, \varphi_z \rangle \langle f_1, \phi_z \rangle \langle \varphi_z, \upsilon_z \rangle \, \mathrm{d}\mu(z) \right| \lesssim_\delta \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)} \|f_1\|_2 \|f_2\|_2, \tag{2.24}$$

where K, ψ_z, $\upsilon_z \in \Psi_z^{0,\delta;0}$ are as above and $\phi_z \in \Psi_z^{0,\delta;1}$. This is obtained exactly as in [43, Ch. 5, Proposition 2] once the Carleson measure estimate

$$\sup_{(x,s) \in Z^d} \frac{1}{s^d} \int_{z \in \mathsf{T}(x,s)} |\langle f_3, \varphi_z \rangle|^2 \, \mathrm{d}\mu(z) \lesssim_\delta \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)}^2 \tag{2.25}$$

is in place. This estimate is classical, see [43, Ch. 5, (6.7)], but we sketch the proof. First of all, fixing $(x, s) \in Z^d$ and letting $Q = \{y \in \mathbb{R}^d : |y - x| < 3\frac{s}{2}\}$, choose a suitable representative of $f_3 \in \mathrm{BMO}(\mathbb{R}^d)$ with $\int_Q f_3 = 0$. Then using Proposition 2.4

$$\int_{z \in \mathsf{T}(x,s)} |\langle f_3 \mathbf{1}_Q, \varphi_z \rangle|^2 \, \mathrm{d}\mu(z) \lesssim \left\| S_\delta[f_3 \mathbf{1}_Q] \right\|_2^2 \lesssim_\delta \left\| f_3 \mathbf{1}_Q \right\|_2^2 \leq s^d \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)}^2,$$

while a standard tail estimate and the vanishing mean of f_3 on Q yield

$$\int_{z \in \mathsf{T}(x,s)} |\langle f_3 \mathbf{1}_{\mathbb{R}^d \setminus Q}, \varphi_z \rangle|^2 \, \mathrm{d}\mu(z) \lesssim_\delta s^d \left(\inf_{y \in Q} [f_3]^\#(y) \right)^2 \leq s^d \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)}^2,$$

and (2.25) follows by coupling the last two displays. $\qquad\square$

Proposition 2.6 *For each pair $f_1, f_2 \in L^1(\mathbb{R}^d)$ there exists a sparse collection $\mathcal{S}$ of cubes of $\mathbb{R}^d$ with the property that*

$$\Psi^\delta(f_1, f_2) \lesssim_\delta \sum_{Q \in \mathcal{S}} |Q| \langle f_1 \rangle_Q \langle f_2 \rangle_Q, \quad \langle f \rangle_Q := \frac{1}{|Q|} \int |f| \mathbf{1}_Q \, \mathrm{d}x.$$

Proposition 2.7 *Let $f_3 \in \mathrm{BMO}(\mathbb{R}^d)$. For each pair $f_j \in L^1(\mathbb{R}^d)$ there exists a sparse collection $\mathcal{S}$ of cubes of $\mathbb{R}^d$ with the property that*

$$\pi_{f_3}^\delta(f_1, f_2) \lesssim_\delta \|f_3\|_{\mathrm{BMO}(\mathbb{R}^d)} \sum_{Q \in \mathcal{S}} |Q| \prod_{j=1}^2 \langle f_j \rangle_Q.$$

Proof of Propositions 2.6 and 2.7 The two propositions are obtained via identical reasoning based on (2.22)–(2.23) respectively. Arguing for Proposition 2.6 to fix ideas, it has already been noticed that $\Psi^\delta(f_1, f_2)$ is controlled by a suitable L^2-bounded linearization Λ whose kernel is a η-Hölder continuous Calderón–Zygmund kernel form for all $0 < \eta < \delta$. Therefore, sparse bounds for Λ and ultimately for Ψ^δ follow by a direct application of [9, Theorem D]. □

While we have appealed to the result of [9, Theorem D] in the course of the previous proof, the sparse domination principles of [13, 32, 36, 37] among others might have been used just as well. Also, a direct argument along the lines of [9, 13] may also be employed to obtain a more precise version of the above propositions, where local oscillations replace averages in the sparse form for the cancellative arguments. This will appear in forthcoming work.

3 Some technical preliminaries

This section contains a few technical tools that will be used in the proofs of the representation theorems.

3.1 Alpert basis

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

$$\int_{\mathbb{R}^d} x^\alpha \phi_\gamma(x)\, dx = \delta_{\gamma\alpha} \quad \forall 0 \le |\alpha| \le 2D, \tag{3.1}$$

where $\delta_{\gamma\alpha}$ is the Dirac delta function associated to the multi-index γ. The collection ϕ_γ has been explicitly constructed by Alpert [1], see also the extension to general measures in [56, Theorem 1.1]. A first technical lemma involves the Alpert basis.

Lemma 3.1 *Let $z = (x, s)$, $\zeta = (\xi, \sigma) \in Z^d$, and with reference to (2.1), $z \in A(\zeta)$. Define*

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

Then

$$|\chi_{z,\zeta}(v)| \lesssim \frac{1}{s^d} \left(\frac{|v - \xi|}{s} \right)^k \min\left\{ 1, \frac{\max\{|v - \xi|, \sigma\}}{s} \right\}, \quad v \in \mathbb{R}^d. \tag{3.2}$$

Proof Let

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

be the degree k Taylor polynomial of φ_z centered at ξ; the equality involving q_γ is due to (3.1). By Taylor's theorem, as supp $\mathsf{Sy}_\zeta \phi_\gamma \subset \mathsf{B}_\zeta$,

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

It follows that

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

The first summand of the last display complies with the estimate in the right hand side of (3.2), by Taylor's theorem and the fact that $\varphi_z \in \Psi_z^{k,1;0}$. The second summand is also bounded by the right hand side of (3.2): this is easily seen by checking the cases $|v - \xi| \le \sigma, \sigma < |v - \xi| \le s, |v - \xi| > s$ separately. The latter remark completes the proof of the Lemma. □

3.2 Averaging yields rough wavelets

In the representation theorems, the key steps involve a certain averaging of the wavelet φ of (2.7).

Lemma 3.2 *Let $\{\varphi_z : z \in Z^d\}$ be as in (2.10). Let $0 < \eta < \delta \le 1$ and $0 \le k \le D$. Let $u : Z^d \to \mathbb{C}$ be a Borel measurable function with $|u(z)| \le 1$. Then, there exists $C \lesssim_{k,\delta,\eta} 1$ such that for all $z = (x, s) \in Z^d$*

$$\psi_z(\cdot) := \int\limits_{\alpha \in \mathbb{R}^d} \int\limits_{0 < \beta \le 1} \frac{\beta^{k+\delta} u((\alpha, \beta))}{\langle \alpha \rangle^{d+k+\delta}} \varphi_{(x+\alpha s, \beta s)}(\cdot) \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\delta;0}, \tag{3.3}$$

$$\upsilon_z(\cdot) := \int\limits_{\alpha \in \mathbb{R}^d} \int\limits_{\beta > 1} \frac{u((\alpha, \beta))}{(\max\{|\alpha|, \beta\})^{d+k+\delta}} \varphi_{(x+\alpha s, \beta s)}(\cdot) \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\eta;0}. \tag{3.4}$$

In particular, with reference to (2.3),

$$\upsilon_z := \psi_z + \upsilon_z = \int\limits_{(\alpha,\beta) \in Z^d} [(\alpha, \beta)]_{k+\delta} u((\alpha, \beta)) \varphi_{(x+\alpha s, \beta s)} \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta} \in C\Psi_z^{k,\eta;0}. \tag{3.5}$$

The proof of Lemma 3.2 is postponed till after the following useful application.

Lemma 3.3 *Let $\{\varphi_z : z \in Z^d\}$ be as in (2.10). Let $\zeta \in Z^d$ be fixed and $q_\zeta \in \Psi_\zeta^{k,1;1}$ with supp $q_\zeta \subset \mathsf{B}_\zeta$. Then there exists an absolute constant $C = C(d, k)$ and $\vartheta_\zeta \in C\Psi_\zeta^{k,1;1}$ such that*

$$\int\limits_{z \in A(\zeta)} \langle h, \varphi_z \rangle \langle \varphi_z, q_\zeta \rangle \, \mathrm{d}\mu(\zeta) = \langle h, \vartheta_\zeta \rangle \quad \forall h \in \mathcal{S}(\mathbb{R}^d). \tag{3.6}$$

Furthermore,

$$\int_{\mathbb{R}^d} x^\gamma \vartheta_\zeta(x) \, dx = \int_{\mathbb{R}^d} x^\gamma q_\zeta(x) \, dx, \quad \forall 0 \le |\gamma| \le k. \tag{3.7}$$

Proof Write $\zeta = (\xi, \sigma)$ throughout. Formula (2.11) yields that

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

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

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

where $I(\zeta) := \{(x, s) : s \le 3\sigma, |x - \xi| \le 3 \max\{s, \sigma\}\}$. If $(x, s) \in I(\zeta)$, Lemma 2.3 implies that $|\langle \varphi_{(x,s)}, q_\zeta \rangle| \lesssim \sigma^{-d}(s/\sigma)^{k+1}$. A change of variable and an application of Lemma 3.2, (3.3) in particular, shows $\psi_\zeta \in C\Psi_\zeta^{k,1;0}$. The proof is completed by setting $\vartheta_z = q_z - \psi_z$ and deducing (3.7) from Fubini's theorem. $\quad\square$

3.3 Proof of Lemma 3.2

First of all, Fubini's theorem immediately implies that ψ_z and ν_z inherit the moment properties (2.9). The memberships $c\psi_z \in \Psi_z^{k,\delta;1}$, $c\nu_z \in \Psi_z^{k,\eta;1}$ are needed and proved now.

Proof of (3.3) By invariance of assumptions and conclusions under the family Sy_z, it suffices to work in the case $z = (0, 1)$. As z is thus fixed below, it is omitted from the subscript notation. We turn to showing that $\|\partial^\gamma \psi\|_{\star, k+\delta, \delta} \lesssim 1$ for each γ with $0 \le |\gamma| = \kappa \le k$. Fix $w \in \mathbb{R}^d$, and let $\phi = \partial^\gamma \varphi$ locally. Then,

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

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

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

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

$$
\begin{aligned}
|\partial^\gamma \psi(w)| &\lesssim \langle w \rangle^{-(d+k+\delta)} \int\limits_{\substack{\alpha \in Q_w \\ 0 < \beta \leq 1}} \beta^\delta \left| \phi \left(\frac{w-\alpha}{\beta} \right) \right| \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta^{d+1}} \\
&= \langle w \rangle^{-(d+k+\delta)} \int\limits_{\substack{v \in \mathbb{R}^d \\ 0 < \beta \leq 1}} \beta^{\delta-1} |\phi(v)| \, \mathrm{d}v \mathrm{d}\beta \lesssim_\delta \langle w \rangle^{-(d+k+\delta)}
\end{aligned}
\tag{3.8}
$$

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

$$
\left| \partial^\gamma \psi(w+h) - \partial^\gamma \psi(w) \right| \lesssim |h|^\delta \langle w \rangle^{-(d+k+\delta)}, \quad h \in \mathbb{R}^d. \tag{3.9}
$$

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

$$
\begin{aligned}
&\int\limits_{\substack{\alpha \in Q_w \\ 0 < \beta \leq \frac{|h|}{2}}} \beta^\delta \left[\left| \phi \left(\frac{w-\alpha}{\beta} \right) \right| + \left| \phi \left(\frac{w+h-\alpha}{\beta} \right) \right| \right] \frac{\mathrm{d}\alpha \mathrm{d}\beta}{\beta^{d+1}} \\
&+ \int\limits_{\substack{\alpha \in Q_w \\ \frac{|h|}{2} < \beta \leq 1}} \beta^\delta \left| \phi \left(\frac{w+h-\alpha}{\beta} \right) - \phi \left(\frac{w-\alpha}{\beta} \right) \right| \frac{\mathrm{d}\alpha \mathrm{d}\beta}{\beta^{d+1}}.
\end{aligned}
\tag{3.10}
$$

A change of variable shows that both summands in the first integral of (3.10) are

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

Notice that in the α-support of the second integral in (3.10), that $|h| \leq 2\beta$ and

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

because of the support property of ϕ. Therefore such support has diameter $\lesssim \beta$. Using this fact and the mean value theorem, the second integral in (3.10) is

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

This completes the proof that $\psi \in C\Psi_{(0,1)}^{\delta;0}$ as desired. $\qquad \square$

Proof of (3.4). Again normalize $z = (0, 1)$. Fixing $0 \leq \kappa \leq k$, and using the local notation $f := \nabla^\kappa v_z$, it must be shown that $\|f\|_{\star,k+\eta,\eta} \lesssim 1$. Note that

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

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

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

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

$$|f(w)| \lesssim \langle w \rangle^{-(d+k+\delta)} \int\limits_{1}^{\max\{\frac{|w|}{4}, 1\}} \frac{\mathrm{d}\beta}{\beta} + \int\limits_{\max\{\frac{|w|}{4}, 1\}}^{\infty} \frac{\mathrm{d}\beta}{\beta^{d+k+\delta+1}}$$

$$\lesssim \langle w \rangle^{-(d+k+\delta)} \log\langle w \rangle \lesssim_\eta \langle w \rangle^{-(d+k+\eta)}.$$

Using the mean value theorem for φ and the previous observations

$$|f(w + h) - f(w)|$$

$$\leq \int\limits_{\beta > 1} \int\limits_{\alpha \in Q_\beta} \left|\phi\left(\frac{w + h - \alpha}{\beta}\right) - \phi\left(\frac{w - \alpha}{\beta}\right)\right| \frac{\mathrm{d}\beta \mathrm{d}\alpha}{(\max\{|\alpha|, \beta\})^{d+k+\delta} \beta^{d+1}}$$

$$\lesssim |h| \int_1^\infty \frac{\mathrm{d}\beta}{\max\{|w|, 4\beta\}^{d+k+\delta} \beta^2} \lesssim |h|\langle w \rangle^{-(d+k+\delta)},$$

and collecting the last two estimates is more than enough to show that $\|f\|_{\star,k+\eta,\eta} \lesssim 1$. This also completes the proof of Lemma 3.2. $\qquad\qquad\square$

4 Wavelet representation of one parameter Calderón–Zygmund operators

This section provides a representation theorem for one parameter Calderón–Zygmund forms Λ involving wavelet coefficients with vanishing moments. Throughout the section, Λ stands for a continuous bilinear form on $\mathcal{S}(\mathbb{R}^d)$ with adjoint form

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

and two adjoint linear continuous operators

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

4.1 Weak boundedness, kernel estimates and paraproducts

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

$$\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^d, k, \delta)} := \|\Lambda\|_{\mathrm{WB}, \delta} + \|\Lambda\|_{\mathrm{K}, k, \delta} \tag{4.1}$$

with the quantities on the right hand side defined below.

Definition 4.1 (*Weak boundedness*). The form Λ has the δ-*weak boundedness property* if there exists $C > 0$ such that.

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

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

Definition 4.2 (*Kernel estimates*). For a function $K = K(u, v) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, recall the finite difference notation.

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

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

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

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

$$|u - v|^d \left[|\nabla_u^\kappa K(u, v)| + |\nabla_v^\kappa K(u, v)| \right] \leq C, \quad 0 \leq \kappa \leq k; \tag{4.2}$$

$$|u - v|^{d+k} \left[|\Delta_{h|\cdot} \nabla_u^k K(u, v)| + |\Delta_{\cdot|h} \nabla_v^k K(u, v)| \right] \leq C \left(\frac{|h|}{|u - v|} \right)^\delta. \tag{4.3}$$

Call $\|\Lambda\|_{\mathrm{K}, k, \delta}$ the least constant C such that (4.2) and (4.3) hold.

Say that $\Lambda \in \text{SI}(\mathbb{R}^d, k, \delta)$ if the constant (4.1) is finite. The next two examples of forms in $\text{SI}(\mathbb{R}^d, k, \delta)$ are the fundamental building blocks of Calderón–Zygmund forms with higher degree smoothness.

Definition 4.3 (Wavelet form). Let $\{\beta_z, \upsilon_z \in \Psi_z^{k,\delta;0} : z \in Z^d\}$ be two families of cancellative wavelets. The form

$$\Lambda(f, g) = \int_{Z^d} \langle f, \beta_z \rangle \langle \upsilon_z, g \rangle \, \mathrm{d}\mu(z) \tag{4.4}$$

belongs to $\text{SI}(\mathbb{R}^d, k, \delta)$ and $\|\Lambda\|_{\text{SI}(\mathbb{R}^d, k, \delta)} \lesssim 1$. The weak boundedness property is contained in Proposition 2.6 while the (k, δ) kernel estimate is obtained via a standard computation reliant on (2.12)–(2.13).

Definition 4.4 (*Paraproduct forms*). Let $0 \le |\gamma| \le D$ be a multi-index. Call the family $\{\vartheta_{\gamma,z} \in C\Psi_z^{D,1;1} : z \in Z^d\}$ a γ-family if

$$\int_{\mathbb{R}^d} x^\alpha \vartheta_{\gamma,z}(x) \, \mathrm{d}x = t^{|\alpha|} \delta_{\gamma\alpha}, \quad \forall 0 \le |\alpha| \le |\gamma|. \tag{4.5}$$

Here, we recall that a point $z \in Z^d$ is written as $z = (x_0, t) \in \mathbb{R}^d \times (0, \infty)$, with an associated position and scale and the t above in (4.5) is coming from the scale of the function $\vartheta_{\gamma,z}$.

As an example of such a family, for instance, if ϕ_γ satisfies (3.1), then $\{\vartheta_{\gamma,z} := \text{Sy}_z \phi_\gamma : z \in Z^d\}$ is a γ-family. For a function $b \in \text{BMO}(\mathbb{R}^d)$, and multi-indices γ, α, referring to (2.15) for $\varphi_{\alpha,z}$ define

$$\Pi_{b,\gamma,\alpha}(f, g) = \int_{Z^d} \langle b, \varphi_{\alpha,z} \rangle \langle f, \vartheta_{\gamma,z} \rangle \langle \varphi_z, g \rangle \, \mathrm{d}\mu(z). \tag{4.6}$$

If $\gamma = \alpha$, simply write $\Pi_{b,\gamma}$. It is important to stress, see Remark 2.2, that $\varphi_{\gamma,z} \in C\Psi_z^{D,1;0}$ for all $z \in Z^d$. Absolute convergence of the above integral for $f, g \in L^1(\mathbb{R}^d)$ is granted by the easily verified intrinsic estimate

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

referring to (2.21). In particular $\Pi_{b,\gamma,\alpha}$ has the $(1, 1)$-sparse bound, which implies $L^2(\mathbb{R}^d)$ estimates and *a fortiori* weak boundedness property of $\Pi_{b,\gamma,\alpha}$, with $\|\Pi_{b,\gamma,\alpha}\|_{\text{WB},\delta} \lesssim \|b\|_{\text{BMO}(\mathbb{R}^d)}$. Standard calculations show that $\|\Pi_{b,\gamma,\alpha}\|_{\text{K},k,1} \lesssim_k \|b\|_{\text{BMO}(\mathbb{R}^d)}$ for all $0 \le k \le D$, so that

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

Remark 4.5 The weak boundedness property of Definition 4.1 tests Λ on smooth functions. The recent literature related to $T(1)$ and representation theorems, see for instance [29, 39, 41] and references therein, favors testing conditions on indicator functions. When the form Λ also satisfies kernel estimates, the weak boundedness

condition employed in this paper actually follows from indicator-type conditions and is therefore less restrictive. More precisely, suppose that the bilinear form Λ is well defined on $L_0^\infty(\mathbb{R}^d) \times L_0^\infty(\mathbb{R}^d)$ and satisfies

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

in addition to the δ-kernel estimates (4.2) and (4.3). Then $\|\Lambda\|_{\mathrm{WB},\delta} \lesssim 1$, namely Λ has the weak boundedness property of Definition 4.1. A proof of this implication is found in [57, Lecture Notes 7, Lemma 1.3].

4.2 Calderón–Zygmund forms of class (k, δ)

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

$$\mathbf{1}_{\mathsf{B}_{(0,1)}} \leq \phi \leq \mathbf{1}_{\mathsf{B}_{(0,2)}}, \tag{4.7}$$

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

$$p_R^\alpha \in \mathcal{S}(\mathbb{R}^d), \quad p_R^\alpha(x) = x^\alpha \mathrm{Dil}_R^\infty \phi(x), \quad x \in \mathbb{R}^d. \tag{4.8}$$

Let $\mathcal{S}_D(\mathbb{R}^d)$ be the subspace of functions $\psi \in \mathcal{S}(\mathbb{R}^d)$ with the vanishing moment property (2.9) for all multi-indices $0 \leq |\alpha| \leq D$. If $0 \leq |\alpha| \leq k < D$ and $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$, the limits

$$\Lambda(x^\alpha, \psi) = \lim_{R \to \infty} \Lambda \left(p_R^\alpha, \psi \right), \quad \psi \in \mathcal{S}_D(\mathbb{R}^d) \tag{4.9}$$

exist, do not depend on the particular choice of ϕ, and define linear continuous functionals on $\mathcal{S}_D(\mathbb{R}^d)$ see [23, Lemma 1.91] for a proof.

Remark 4.6 If $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ is a wavelet form of the type (4.4), then the functionals $\Lambda(x^\alpha, \cdot)$ vanish for all $0 \leq |\alpha| \leq k$. This is easily verified by appealing to the cancellation properties of the families $\{\beta_z, \upsilon_z \in \Psi_z^{k,\delta;0} : z \in Z^d\}$.

With (4.9) in hand, it is possible to ask whether $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ admits κ-th order paraproducts for $0 \leq \kappa \leq k$.

Definition 4.7 (Λ *has paraproducts of κth order*). Say that $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ has *paraproducts of 0th order* if there exists $b_0, b_0^\star \in \mathrm{BMO}(\mathbb{R}^d)$ with the property that

$$\Lambda(\mathbf{1}, \psi) = \langle b_0, \psi \rangle, \quad \Lambda^\star(\mathbf{1}, \psi) = \langle \psi, b_0^\star \rangle \quad \forall \psi \in \mathcal{S}_D(\mathbb{R}^d) \tag{4.10}$$

If this is the case, referring to (4.6), define the 0-*th order cancellative part* of Λ as

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

We now define inductively the property of having paraproducts of order κ for $1 \leq \kappa \leq k$. Suppose Λ has paraproducts of order $0 \leq \kappa < k$ and the *κth order cancellative*

part of Λ has been defined. Then Λ has *paraproducts of* $(\kappa + 1)$*th order* if for each multi-index α with $|\alpha| = \kappa + 1$ there exists $b_\alpha, b_\alpha^\star \in \mathrm{BMO}(\mathbb{R}^d)$ with the property that

$$\Lambda_\kappa(x^\alpha, \psi) = (-1)^{\kappa+1} \langle b_\alpha, \partial^{-\alpha} \psi \rangle, \quad \Lambda_\kappa^\star(x^\alpha, \psi) = (-1)^{\kappa+1} \langle \partial^{-\alpha} \psi, b_\alpha^\star \rangle \tag{4.11}$$

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

$$\Lambda_{\kappa+1}(f, g) = \Lambda_\kappa(f, g) - \sum_{|\alpha|=\kappa+1} \left[\Pi_{b_\alpha, \alpha}(f, g) + \Pi_{b_\alpha^\star, \alpha}(g, f) \right]. \tag{4.12}$$

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

Remark 4.8 Observe that (4.12) is equivalent to

$$\Lambda_\kappa(f, g) = \Lambda(f, g) - \sum_{0 \le |\alpha| \le \kappa} \Pi_{b_\alpha, \alpha}(f, g) + \Pi_{b_\alpha^\star, \alpha}(g, f). \tag{4.13}$$

The inductive procedure of the proof of Theorem A reduces to the case $\Lambda(f, g) = \Lambda_\kappa(f, g)$.

Remark 4.9 The 0-th order condition (4.10) is equivalent to the familiar assumption

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

For $0 \le \kappa \le k - 1$, let $T_\kappa, T_\kappa^\star$ be the adjoint operators to Λ_κ. In view of Remark 2.1, as R^γ preserves $\mathrm{BMO}(\mathbb{R}^d)$, the condition may be reformulated as

$$|\nabla|^\kappa T_{\kappa-1}(x^\alpha) = a_\alpha \in \mathrm{BMO}(\mathbb{R}^d), \quad |\nabla|^\kappa T_{0,\kappa-1}^\star(x \mapsto x^\alpha) = a_\alpha^\star \in \mathrm{BMO}(\mathbb{R}^d), \tag{4.14}$$

in the sense of $\mathcal{S}_D'(\mathbb{R}^d)$, where $a_\alpha := R^\alpha b_\alpha$ and similarly for $a_\alpha^\star$.

Remark 4.10 Using (4.5) and Remark 2.2, one directly computes

$$\Pi_{b,\gamma,\alpha}(x^\beta, g) = (-1)^{|\alpha|} \delta_{\gamma\beta} \langle b, \partial^{-\alpha} g \rangle, \quad \Pi_{b,\gamma,\alpha}^\star(x^\beta, f) = 0, \quad 0 \le |\beta| \le |\gamma|. \tag{4.15}$$

Thus $\Pi_{b,\gamma}$ has paraproducts of order $|\gamma|$ according to Definition 4.7, with $b_\beta = \delta_{\gamma\beta} b$ and $b_\beta^\star = 0$ for all $0 \le |\beta| \le |\gamma|$.

Definition 4.11 The continuous bilinear form Λ belongs to the class $\mathrm{CZ}(\mathbb{R}^d, k, \delta)$ of (k, δ)-*Calderón–Zygmund (CZ) forms* if $\Lambda \in \mathrm{SI}(\mathbb{R}^d, k, \delta)$ and Λ has paraproducts of order k. For further use, define the norm

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)} := \|\Lambda\|_{\mathrm{SI}(\mathbb{R}^d, k, \delta)} + \sum_{0 \leq |\alpha| \leq k} \left(\|b_\alpha\|_{\mathrm{BMO}(\mathbb{R}^d)} + \|b_\alpha^\star\|_{\mathrm{BMO}(\mathbb{R}^d)} \right).$$

(4.16)

The statement of Theorem A below is the representation and (sparse) $T(1)$-theorem for (k, δ)-CZ forms. Its proof is postponed to Sect. 4.3. The weighted $T(1)$ result is stated separately in Corollary A.1 with the deduction of the corollary given at the end of this subsection.

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

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

(4.17)

where, for all $0 \leq |\gamma| \leq k$, $\Pi_{b_\gamma, \gamma}$ and $\Pi_{b_\gamma^\star, \gamma}$ are explicitly constructed paraproducts of the form in Definition 4.4, with $\|b_\gamma\|_{\mathrm{BMO}(\mathbb{R}^d)}, \|b_\gamma^\star\|_{\mathrm{BMO}(\mathbb{R}^d)} \leq \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)}$ as in Definition 4.7.

Remark 4.12 The first term on the right hand side of (4.17) is not symmetric with respect to taking adjoints: there is one point in the proof, see (4.30) where the symmetry between f and g is broken by choosing on which side the averaging Lemma 3.2 is applied. A representation where the φ_z-family is kept on the g-side may be obtained by reversing this choice.

Remark 4.13 The recent article [28] devises a dyadic representation theorem for L^2-bounded Calderón–Zygmund operators T whose kernel obey estimates (4.2) for some $k \geq 1$ and satisfying special cancellation assumptions, namely $T(x^\gamma), T^\star(x^\gamma)$ is a polynomial of degree $\leq k$ for all multi-indices $0 \leq |\gamma| \leq k$, which is equivalent to assuming (4.14) holds with $a_\gamma = a_\gamma^\star = 0$ for all γ. Their representation is of dyadic nature in the sense that it involves shifted dyadic grids and dyadic shift operators of arbitrary complexity, just like the classical one of [29], however a wavelet orthonormal basis is employed as a building block of the dyadic shifts instead of the Haar basis. The useful gain in comparison with [29] is a better decay rate of the dyadic shift coefficients in the representation.

Theorem A differs from [28, Theorem 1.1] in several aspects. First, it is of $T(1)$ type and the vanishing assumptions on the higher order paraproducts is replaced by the more general BMO assumption (4.14). Furthermore, our model operators have

complexity zero and no averaging procedure over shifted grids is needed, resulting in a compact and computationally feasible expansion. This feature allows for a seamless and powerful extension to the bi-parameter case, which we perform in Theorem B below.

We come to the $T(1)$ theorem. Notice that (4.18) below is a vacuous assumption when $k = 0$, whence Theorem A has a sparse, sharp weighted version of the classical $T(1)$ theorem as a corollary. Also notice that no assumption is being made on the adjoint paraproducts $b_\gamma^\star$.

Corollary A.1 *Suppose that Λ is a standard (k, δ)-CZ form with*

$$b_\gamma = 0 \quad \forall 0 \le |\gamma| < k. \tag{4.18}$$

Then referring to (2.19) and (2.21), for each $|\alpha| = k$ the following estimate is true

$$\left| \Lambda(f, \partial^\alpha g) \right| \lesssim_\eta \sum_{|\beta|=k} \left[\Psi^\eta(\partial^\beta f, g) + \sum_{|\gamma|=k} \pi_{b_\gamma}(\partial^\beta f, g) + \sum_{0 \le |\gamma| \le k} \pi_{b_\gamma^\star}(g, \partial^\beta f) \right]. \tag{4.19}$$

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

$$\|Tf\|_{\dot{W}^{k,p}(\mathbb{R}^d; w)} \lesssim [w]_{A_p}^{\max\left\{1, \frac{1}{p-1}\right\}} \|f\|_{\dot{W}^{k,p}(\mathbb{R}^d; w)}, \quad p \in (1, \infty). \tag{4.20}$$

Remark 4.14 Condition (4.18) is also necessary for (4.20) to hold, i.e. Corollary A.1 is a characterization of (4.20). This generalizes the case $\Omega = \mathbb{R}^d$ of [55, Theorem 1.1] to the non-convolution case; in fact, a scaling argument shows that when $\Omega = \mathbb{R}^d$, condition b. in [55, Theorem 1.1] is equivalent to (4.18). To see the necessity, suppose that (4.20) holds for some exponent p_0 and all weights $w \in A_{p_0}$. Extrapolation of weighted norm inequalities [12, 16] then implies that (4.20) holds for $p = 2d$ and w equals Lebesgue measure. The content of Corollary A.1 also allows to assume that the adjoint T to Λ equals

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

Fix $0 \le |\gamma| = \kappa < k$ and let $\varepsilon := k - (\kappa + \frac{1}{2}) > 0$. Then define $f_R(x) := R^\varepsilon x^\gamma \alpha_{(0,R)}(x)$ where α_z is the cutoff from (2.2) and $R > 1$ is arbitrary. It is immediate to show that $\|f_R\|_{\dot{W}^{k,2d}(\mathbb{R}^d)} \sim 1$, so Tf_R is a bounded sequence in $\dot{W}^{k,2d}(\mathbb{R}^d)$. Also, using the properties (4.5) followed by (2.11), $R^{-\varepsilon} Tf_R \to \partial^{-\gamma} b_\gamma = T(x^\gamma)$ in the sense of Definition 4.7. These two properties entail $T(x^\gamma) = 0$ in $\dot{W}^{k,2d}(\mathbb{R}^d)$. Thus $T(x^\gamma)$ is a polynomial of degree $\le k$. Appealing to Definition 4.7 again reveals that $b_\gamma = 0$ as claimed.

Remark 4.15 Testing type theorems for smooth singular integral operators have previously appeared in several works: a non-exhaustive list includes [23, 26, 42, 58, 60] as well as the already mentioned [28, 55] and references therein. In particular, [60, Theorem 1, cases (6,7)] is essentially equivalent to the unweighted version of Corollary A.1. Corollary A.1 appears to be the first weighted $T(1)$ theorem of this type. A sparse bound in the vein of Corollary A.1 was proved in [5] for the case $k = 1$ using techniques from [36]. However, the result of [5] is not of testing type and was obtained under the stronger assumption that T is a priori bounded on the Sobolev space $\dot{W}^{1,2}(\mathbb{R}^d)$.

Proof of Corollary A.1 Before the actual proof, make the following observations referring to the wavelets φ_z, υ_z in the representation (4.17): for $z = (x, s) \in Z$, $|\alpha| = |\beta| = |\nu| = k, 0 \leq |\gamma| \leq k$

$$s^k \partial^\alpha \varphi_z, \ s^k \partial^\alpha \upsilon_z \in C\Psi_z^{\varepsilon;0}, \quad s^{-k} \partial^{-\beta} \vartheta_{\nu,z}, \ s^k \partial^\alpha \vartheta_{\gamma,z} \in C\Psi_z^{1;1}. \tag{4.21}$$

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

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

Integrating by parts and using Remark 2.2 gives

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

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

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

The $b_\gamma^\star$, $|0| \leq k \leq \gamma$ type paraproduct is controlled similarly: with reference to Remark 2.2 for $\varphi_{\beta,z}$,

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

$$\lesssim \sum_{|\beta|=k} \pi_{b_\gamma}(\partial^\beta f, g).$$

This completes the proof of (4.19). The weighted norm inequality then follows as a consequence of the sparse estimates

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

obtained by combining (4.19) with the Propositions 2.6 and 2.7. □

4.3 Proof of Theorem A

Start by normalizing $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)} = 1$. Throughout the proof, the properties (2.7)–(2.10) and (3.1) will be referred to frequently. Recall that $\varepsilon \in (0, \delta)$ is fixed but arbitrary, and let $\eta = \frac{\varepsilon + \delta}{2}$. Throughout the proof, for $z, \zeta \in Z^d$ we write

$$\chi_{z,\zeta} := \varphi_z - P_{z,\zeta} \mathbf{1}_{A(\zeta)}(z) \tag{4.22}$$

referring to (2.1) and Lemma 3.1; note that this does not override the definition of Lemma 3.1.

Lemma 4.16 $\left|\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z})\right| \lesssim [z, \zeta]_{k+\eta} \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)}.$

Proof It suffices by symmetry to work in the region $z \in Z_+^d(\zeta)$, see (2.1). The estimates are then verified by case analysis.

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

Case $z \in A(\zeta)$. Let $\alpha_\zeta, \beta_\zeta$ as in (2.2). Then

$$\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) = \Lambda(\chi_{z,\zeta}, \varphi_\zeta) = \Lambda(\Theta, \varphi_\zeta) + \Lambda(\Xi, \varphi_\zeta), \quad \Theta := \chi_{z,\zeta} \alpha_\zeta,$$
$$\Xi := \chi_{z,\zeta} \beta_\zeta \tag{4.23}$$

and one seeks estimates for each of the summands in the last right hand side. For the first, apply the weak boundedness property for the point $\tilde{\zeta} = (\xi, 4\sigma)$, so that $B_{\tilde{\zeta}} = 4B_\zeta$, and use (3.2) to estimate $\|\Theta\|_\infty$, obtaining

$$|\Lambda(\Theta, \varphi_\zeta)| \leq \|\Lambda\|_{\mathrm{WB}, \delta} \|\Theta\|_\infty \lesssim \frac{1}{s^d} \left(\frac{\sigma}{s}\right)^{k+1} \lesssim \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)}[z, \zeta]_{k+1}. \tag{4.24}$$

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

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

$$
\begin{aligned}
|\Lambda(\Xi, \varphi_\zeta)| &= \left| \int_{\mathbb{R}^d \backslash 2B_\zeta} \Xi(v) \int_{B_\zeta} K(u, v) \varphi_\zeta(u) \, du dv \right| = \left| \int_{\mathbb{R}^d \backslash 2B_\zeta} \Xi(v) \langle F_v, \varphi_\zeta \rangle \, dv \right| \\
&\leq \sigma^k \sum_{|\gamma|=k} \left| \int_{\mathbb{R}^d \backslash 2B_\zeta} \Xi(v) \langle \partial_u^\gamma F_v, \sigma^{-k} \partial^{-\gamma} \varphi_\zeta \rangle \right| \\
&\leq \sigma^k \sum_{|\gamma|=k} \int_{\mathbb{R}^d \backslash 2B_\zeta} |\Xi(v)| \sup_{u \in B_\zeta} \left| \Delta_{u-\xi} \cdot \partial_u^\gamma K(\xi, v) \right| \|\varphi_{\gamma,\zeta}\|_1 \, dv \\
&\lesssim \sigma^{k+\delta} \|\Lambda\|_{K,k,\delta} \int_{\mathbb{R}^d \backslash 2B_\zeta} \frac{|\Xi(v)|}{|v - \xi|^{d+k+\delta}} \, dv.
\end{aligned}
\tag{4.25}
$$

Here, the passage to the second line is obtained by using supp φ_ζ, supp $\varphi_{\gamma,\zeta} \subset B_\zeta$, consult (2.8), and integrating by parts. The subsequent (in)equality follows from the mean zero property of $\varphi_{\gamma,\zeta}$, see Remark 2.2, and the next step is obtained via the kernel estimates (4.3). Bound the last right hand side by splitting the integral on $\mathbb{R}^d \backslash 2B_\zeta$ into the pieces

$$
\sigma^{k+\delta} \int_{\sigma < |v-\xi| \leq s} \frac{|\Xi(v)|}{|v - \xi|^{d+k+\delta}} \, dv \lesssim \frac{1}{s^d} \left(\frac{\sigma}{s}\right)^{k+\delta} \int_\sigma^s \frac{1}{t} \, dt \lesssim \frac{1}{s^d} \left(\frac{\sigma}{s}\right)^{k+\delta} \log\left(\frac{s}{\sigma}\right)
\tag{4.26}
$$

where the δ-geometric mean of the estimates in (3.2) is used, and

$$
\sigma^{k+\delta} \int_{|y-u|>s} \frac{|\Xi(y)|}{|y - \xi|^{d+k+\delta}} \, dy \lesssim \frac{1}{s^d} \frac{\sigma^{k+\delta}}{s^k} \int_s^\infty \frac{1}{t^{1+\delta}} \, dt \lesssim \frac{1}{s^d} \left(\frac{\sigma}{s}\right)^{k+\delta}. \tag{4.27}
$$

Putting together (4.23), (4.25), (4.26) and (4.27) provides

$$
\begin{aligned}
\left| \Lambda(\chi_{z,\zeta}, \varphi_\zeta) \right| &\lesssim \frac{\|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)}}{s^d} \left(\frac{\sigma}{s}\right)^{k+\delta} \log\left(\frac{\sigma}{s}\right) \lesssim_\eta \frac{1}{\|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)} s^d} \left(\frac{\sigma}{s}\right)^{k+\eta} \\
&= \|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)}[z, \zeta]_{k+\eta}
\end{aligned}
\tag{4.28}
$$

as claimed. *Case $z \in F_+(\zeta)$.* In this case $\Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) = \Lambda(\varphi_z, \varphi_\zeta)$ and the supports of φ_z and φ_ζ are separated. Thus, the kernel representation of Λ, the cancellation of $\varphi_{\gamma,\zeta}$ and the kernel estimates can be used. Proceeding exactly like in (4.25) with φ_z in place of Ξ,

$$
\begin{aligned}
|\Lambda(\varphi_z, \varphi_\zeta)| &\lesssim \sigma^{k+\delta} \|\Lambda\|_{K,k,\delta} \int_{B_z} \frac{|\varphi_z(v)|}{|v - \xi|^{d+k+\delta}} \, dv \lesssim \frac{\|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)} \sigma^{k+\delta}}{s^d |x - \xi|^{k+\delta}} \\
&= \|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)}[z, \zeta]_{k+\delta} \leq \|\Lambda\|_{CZ(\mathbb{R}^d, k, \delta)}[z, \zeta]_{k+\eta}.
\end{aligned}
\tag{4.29}
$$

This completes the proof of the lemma. □

The main line of proof of Theorem A now begins. First of all, notice that

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

The proof uses induction on $0 \leq \kappa \leq k$ in a subtle way.

4.3.1 Base case and main part of inductive step

The proof of the representation Theorem A first begins under the additional assumption

$$a(k): b_\gamma, b_\gamma^\star \neq 0 \implies |\gamma| = k$$

namely, all paraproducts vanish except those of highest order. Notice that $a(0)$ is not an extra assumption. Let now $f, g \in \mathcal{S}(\mathbb{R}^d)$. Use (2.11), bilinearity, $\mathcal{S}(\mathbb{R}^d)$-continuity of Λ and definition (4.22) to expand $\Lambda(f, g)$ as

$$\int\limits_{Z^d \times Z^d} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\varphi_z, \varphi_\zeta) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta)$$

$$= \int\limits_{Z^d \times Z^d} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\chi_{z,\zeta}, \chi_{\zeta,z}) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta)$$

$$+ \int\limits_{Z_\zeta^d} \int\limits_{A(\zeta)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(P_{z,\zeta}, \varphi_\zeta) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta)$$

$$+ \int\limits_{Z_z^d} \int\limits_{A(z)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\varphi_z, P_{\zeta,z}) \, \mathrm{d}\mu(\zeta) \mathrm{d}\mu(z).$$

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

$$\int\limits_{Z^d} \langle f, \varphi_z \rangle \langle \upsilon_z, g \rangle \, \mathrm{d}\mu(z), \quad \upsilon_{(x,s)} := \int\limits_{(\alpha,\beta) \in Z^d} [(\alpha, \beta)]_{k+\eta} u_{(x,s)}(\alpha, \beta) \varphi_{(x+\alpha s, \beta s)} \frac{\mathrm{d}\beta \mathrm{d}\alpha}{\beta}$$

(4.30)

where

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

is uniformly bounded by Lemma 4.16. Thus $\upsilon_z \in C\Psi_z^{k,\varepsilon;0}$ by Lemma 3.2 applied with $u = u_{(x,s)}$. This constructs the first expression in the right hand side of (4.17). With reference to Remark 4.12, an alternative form of the term in (4.30) with roles of

f and g exchanged up to conjugation, may be obtained by making instead the change of variable $x = \xi + \alpha\sigma$, $s = \beta\sigma$ and applying Lemma 3.2 accordingly.

It remains to identify the second and third summand of the main decomposition as a sum of paraproduct terms. Turning to this task for the first term, begin by noticing that due to assumption $a(k)$ used twice, and Remark 2.2

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

Therefore, applying Lemma 3.3 to $h = f$, $q_\zeta = \mathsf{Sy}_\zeta \phi_\gamma$ to obtain the last equality

$$\int\limits_{Z_\zeta^d} \int\limits_{z \in A(\zeta)} \langle f, \varphi_z\rangle\langle \varphi_\zeta, g\rangle\Lambda(P_{z,\zeta}, \varphi_\zeta)\, \mathrm{d}\mu(z)\mathrm{d}\mu(\zeta)$$

$$= \sum_{|\gamma|=k} \int\limits_{Z_\zeta^d} \int\limits_{z \in A(\zeta)} \langle f, \varphi_z\rangle\langle \varphi_z, \mathsf{Sy}_\zeta\phi_\gamma\rangle\langle \varphi_\zeta, g\rangle\langle b_\gamma, \varphi_{\gamma,\zeta}\rangle\, \mathrm{d}\mu(z)\mathrm{d}\mu(\zeta)$$

$$= \sum_{|\gamma|=k} \int_{Z^d} \langle f, \vartheta_{\gamma,\zeta}\rangle\langle \varphi_\zeta, g\rangle\langle b_\gamma, \varphi\rangle\, \mathrm{d}\mu(\zeta) := \Pi_{b_\gamma, \gamma}(f, g).$$

Notice that Lemma 3.3 together with (2.7)–(2.10) and (3.1) ensure that $\vartheta_{\gamma,\zeta}$, the output of Lemma 3.3 corresponding to $q_\zeta = \mathsf{Sy}_\zeta\phi_\gamma$, belongs to $\Psi_\zeta^{D,1;1}$ and is a γ-family. A totally symmetric argument deals with the third summand in the main decomposition, and completes the proof of (4.17) under the assumption $a(k)$.

4.3.2 Induction

It remains to explain how to obtain (4.17) without the assumption $a(k)$. In fact, it will be shown that Λ satisfies an instance of representation (4.17) for all $0 \leq \kappa \leq k$. This is done by induction on κ. Before starting the induction, observe that $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, \kappa, \delta)} \leq \|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)} \leq 1$. For $\kappa = 0$, (4.17) is achieved in the previous step.

Assume that Λ has been represented in the form (4.17) for an integer $0 < \kappa < k$. Taking advantage of Definition 4.4, and in particular of Remark 4.10, the κ-cancellative part of Λ defined in (4.13) satisfies $\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^d, k, \delta)} \lesssim 1$ and the equality

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

for some family $\{\upsilon_z \in C\Psi_z^{\kappa, \varepsilon; 0} : z \in Z^d\}$. The last equality of the above display tells us that all paraproducts of Λ_κ having order less than or equal to κ equal zero, cf. Remark 4.6. Therefore, $\Lambda_\kappa(f, g)$ satisfies the assumptions of the theorem, and in

addition $a(\kappa + 1)$. Apply the previous part of the proof to $\Lambda_\kappa(f, g)$, and obtain

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

with $\{\upsilon_z \in C\Psi_z^{\kappa+1, \eta; 0} : z \in Z^d\}$. This equality, rearranged, yields a representation of Λ in the form (4.17) for the value $\kappa + 1$, completing the inductive step and the proof of Theorem A.

5 Wavelets in the bi-parameter setting

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

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

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

$$\langle \phi, F \rangle_1 = \int_{\mathbb{R}^{d_1}} \phi(x_1, \cdot) F(x_1) \, dx_1 \in \mathcal{S}(\mathbb{R}^{d_2})$$

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

$$\begin{aligned}
\phi^{[1,x_1]} &: \mathbb{R}^{d_2} \to X, \quad \phi^{[1,x_1]} := \phi(x_1, \cdot), \\
\phi^{[2,x_2]} &: \mathbb{R}^{d_1} \to X, \quad \phi^{[2,x_2]} := \phi(\cdot, x_2).
\end{aligned} \tag{5.1}$$

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

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

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

composition,

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

5.1 Wavelet classes and the intrinsic square function

This is the bi-parameter analogue of Sect. 2.2. The notation now overrides what was previously laid out in the one parameter case; unfortunately, due to the different dilation structure and working with borderline sufficient decay in each of the parameters, definitions may not be re-used.

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

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

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

is finite. In the bi-parameter case, the relevant cancellation properties of ψ are encoded by requiring (2.9) to hold in the variable x_ι for each slice $\psi^{[\hat{\iota}, x_{\hat{\iota}}]}$ and each $\iota = 1, 2$. This necessitates the introduction of the bi-parameter analogue of the classes $\Psi^{k,\delta;\iota}_z$. Hereafter, $\gamma_j \in \mathbb{N}^{d_j}$ for either $j = 1, 2$ is a multi-index on $\mathbb{R}^{d_j}$. For $k = (k_1, k_2) \in \mathbb{N}^2$, $0 < \delta \leq 1$ define

$$\Psi^{k,\delta;1,1}_z := \left\{ \phi \in \mathcal{S}(\mathbb{R}^{\mathbf{d}}) : s^{|\gamma_1|}_1 s^{|\gamma_2|}_2 \left\| (\mathsf{Sy}_z)^{-1} \partial^{\gamma_1} \partial^{\gamma_2} \phi \right\|_{\star,k+\delta,\delta} \leq 1 \right\},$$

$$\Psi^{k,\delta;0,1}_z := \left\{ \phi \in \Psi^{k,\delta;1,1}_z : (2.9) \text{ holds for } \psi = \psi^{[2,x_2]}, d = d_1, \gamma = \gamma_1, \right.$$

$$\left. \forall x_2 \in \mathbb{R}^{d_2}, \forall 0 \leq |\gamma_1| \leq k_1 \right\},$$

$$\Psi^{k,\delta;1,0}_z := \left\{ \phi \in \Psi^{k,\delta;1,1}_z : (2.9) \text{ holds for } \psi = \psi^{[1,x_1]}, d = d_2, \gamma = \gamma_2, \right.$$

$$\left. \forall x_1 \in \mathbb{R}^{d_1}, \forall 0 \leq |\gamma_2| \leq k_2 \right\},$$

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

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

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

$$(5.2)$$

The analogue of the almost orthogonality Lemma 2.5 in the bi-parameter setting is the following lemma.

Lemma 5.1 *Let $m \in \mathbb{N}$, $0 < 2\eta < \delta \leq 1$, $z, \zeta \in Z^{\mathbf{d}}$, $\psi_z \in \Psi_z^{(2m,2m),\delta;0,0}$, $\psi_\zeta \in \Psi_\zeta^{(2m,2m),\delta;0,0}$. Then*

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

Proof Let ι be either 1 or 2. Applying (5.2) together with the first estimate of Lemma 2.3 and integrating,

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

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

5.2 Intrinsic bi-parameter wavelet coefficients

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

$$\Psi_z^{\delta;(\iota_1,\iota_2)} f = \sup_{\psi \in \Psi_z^{\delta;\iota_1,\iota_2}} |\langle f, \psi \rangle|. \tag{5.3}$$

A standard argument based on (5.2) shows that if $f \in L^1_{\text{loc}}(\mathbb{R}^{\mathbf{d}})$,

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

where M_{d_1,d_2} is the bi-parameter maximal function on $\mathbb{R}^{\mathbf{d}}$. In particular the wavelet coefficients of $f \in L^p(\mathbb{R}^{\mathbf{d}})$ are finite for $f \in L^p(\mathbb{R}^{\mathbf{d}})$, $p > 1$, as $M_{d_1,d_2} f$ is finite a. e. in that case. The remainder of this section contains a basic L^2 estimate for the intrinsic square function

$$SS_\delta f(x_1, x_2) = \left(\int_{(0,\infty)^2} \left[\Psi_{(x_1,t_1),(x_2,t_2)}^{\delta;(0,0)} f \right]^2 \frac{dt_1 \, dt_2}{t_1 t_2} \right)^{\frac{1}{2}}. \tag{5.4}$$

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

Proposition 5.2 $\|SS f\|_2 \lesssim_\delta \|f\|_2$. *As a consequence,*

$$\int_{Z^{\mathbf{d}}} \left[\Psi_z^{\delta;(0,0)} f \right] \left[\Psi_z^{\delta;(0,0)} g \right] d\mu(z) \lesssim \|f\|_2 \|g\|_2. \tag{5.5}$$

Proof Notice that (5.5) follows from the square function estimate via two application of Cauchy–Schwarz inequality. The argument for the square function estimate is analogous to the one employed for (2.18). Working with $f \in L^2(\mathbb{R}^d)$ of unit norm, and fixing $\psi_z \in \Psi_z^{\delta;0,0}$, $z \in Z^{\mathbf{d}}$, it suffices to estimate

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

$$\frac{d\alpha_1 d\beta_1 d\alpha_2 d\beta_2}{\beta_1 \beta_2} \lesssim \int_{z \in Z^{\mathbf{d}}} |\langle f, \psi_z \rangle|^2 \, d\mu(z)$$

as well as three more integrals covering all possible relationships between the scales of $z = ((x_1, s_1), (x_2, s_2))$ and $\zeta(\alpha_1, \beta_1, \alpha_2, \beta_2) = ((x_1 + \alpha_1 s_1, \beta_1 s_1), (x_2 + \alpha_2 s_2, \beta_2 s_2))$, which are estimated in an analogous fashion. In this specific case, Lemma 5.1 entails the bound

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

and the required control is again obtained via a combination of two instances of (2.4) and Cauchy–Schwarz. □

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

6 Wavelet representation of bi-parameter Calderón–Zygmund operators

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

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

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

$$\Lambda^{\star 1}, \Lambda^{\star 2} : \mathcal{S}(\mathbb{R}^{d_1}) \otimes \mathcal{S}(\mathbb{R}^{d_2}) \times \mathcal{S}(\mathbb{R}^{d_1}) \otimes \mathcal{S}(\mathbb{R}^{d_2}) \to \mathbb{C},$$
$$\Lambda^{\star 1}(f_1 \otimes f_2, g_1 \otimes g_2) := \overline{\Lambda(g_1 \otimes f_2, f_1 \otimes g_2)},$$
$$\Lambda^{\star 2}(f_1 \otimes f_2, g_1 \otimes g_2) := \overline{\Lambda(f_1 \otimes g_2, g_1 \otimes f_2)}.$$

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

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

will be considered. A bi-parameter wavelet basis is needed. For $j = 1, 2$ let $\varphi_j \in \mathcal{C}_0^\infty(\mathbb{R}^{d_j})$ be such that $\varphi = \varphi_j, d = d_j$ in (2.7), and $D \geq 8(d_1 + d_2)$ sufficiently large. Set $\varphi := \varphi_1 \otimes \varphi_2 \in \mathcal{C}_0^\infty(\mathbb{R}^{\mathbf{d}})$ as our mother wavelet on $\mathbb{R}^{\mathbf{d}}$, and rescale it by

$$\varphi_z = \varphi_{z_1} \otimes \varphi_{z_2} := \mathsf{Sy}_z \varphi = \mathsf{Sy}_{z_1} \varphi_1 \otimes \mathsf{Sy}_{z_2} \varphi_2, \quad z = (z_1, z_2) \in Z^{\mathbf{d}}. \qquad (6.1)$$

With this position, $\varphi_z \in c\Psi_z^{(D,D),1;0}$ for all $z \in Z^{\mathbf{d}}$.

6.1 Bi-parameter singular integrals

The boundedness and kernel smoothness properties of bi-parameter singulars are quantified by the parameters $k = (k_1, k_2) \in \mathbb{N}^2$ and $\delta > 0$, and summarized by the norm

$$\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)} := \|\Lambda\|_{\mathrm{PWB}, k, \delta} + \|\Lambda\|_{\mathrm{K,f}, k, \delta} \qquad (6.2)$$

whose summands are described below. Notice that the norm (6.2) is stable under full and partial adjoints: this fact will be used without explicit mention from now on.

Definition 6.1 (*Partial kernel and weak boundedness*). For $j = 1, 2$, $z_j \in Z^{d_j}$, and $u_{z_j}, v_{z_j} \in \Psi_{z_j}^{k_j, \delta; 1}$ with supp u_{z_j}, supp $v_{z_j} \subset B_{z_j}$, define the forms

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

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

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

uniformly over $z_j \in Z^{d_j}$ and $u_{z_j}, v_{z_j} \in \Psi_{z_j}^{k_j, \delta; 1}$ with supp u_{z_j}, supp $v_{z_j} \subset B_{z_j}$, $j = 1, 2$. In this case, $\|\Lambda\|_{\mathrm{PWB}, k, \delta}$ is the least such constant C. The $\|\Lambda\|_{\mathrm{PWB}, \delta}$ norm is stable under full and partial adjoints, and subsumes all of weak boundedness and partial kernel assumptions of [41], see also [49].

Definition 6.2 (*Full kernel*). For a function $K = K(u, v)$ on $\mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}$, with $u = (u_1, u_2)$, $v = (v_1, v_2)$ again using the finite difference notation

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

for $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R}^{\mathbf{d}}$, $h_j \in \mathbb{R}^{d_j}$, $j = 1, 2$. In preparation for (6.4), introduce the norms

$$\|K\|_{\kappa} := \sup_{(u,v) \in \mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}} \left(\prod_{j=1,2} |u_j - v_j|^{d_j + \kappa_j} \right) |K(u, v)|,$$

$$\|K\|_{\kappa, \delta, \Delta^1_{\square|.}} := \sup_{(u,v) \in \mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}} \sup_{0 < 2|h_1| < |u_1 - v_1|}$$
$$\frac{|u_1 - v_1|^{d_1 + \kappa_1 + \delta}}{|h_1|^{\delta}} |u_2 - v_2|^{d_2 + \kappa_2} \left| (\Delta^1_{h_1|.} K)(u, v) \right|,$$

$$\|K\|_{\kappa, \delta, \Delta^1_{\square|.} \Delta^2_{\square|.}} := \sup_{\substack{(u,v) \in \mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}} \\ }} \sup_{\substack{0 < 2|h_j| < |u_j - v_j| \\ j=1,2}} \left(\prod_{j=1,2} \frac{|u_j - v_j|^{d_j + \kappa_j + \delta}}{|h_j|^{\delta}} \right)$$
$$\left| (\Delta^1_{h_1|.} \Delta^2_{h_2|.} K)(u, v) \right|,$$

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

$$\Lambda(f_1 \otimes f_2, g_1 \otimes g_2) = \int_{\mathbb{R}^{\mathbf{d}} \times \mathbb{R}^{\mathbf{d}}} K(u, v) f_1(u_1) f_2(u_2) g_1(v_1) g_2(v_2) \, du dv \quad (6.3)$$

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

$$\left\| \nabla^{\kappa_1}_{u_1} \nabla^{\kappa_2}_{u_2} K \right\|_{(\kappa_1, \kappa_2)} + \left\| \nabla^{\kappa_1}_{v_1} \nabla^{\kappa_2}_{v_2} K \right\|_{(\kappa_1, \kappa_2)} \le C,$$

$$\left\| \nabla^{k_1}_{u_1} \nabla^{\kappa_2}_{u_2} K \right\|_{(k_1, \kappa_2), \delta, \Delta^1_{\square|.}} + \left\| \nabla^{k_1}_{v_1} \nabla^{\kappa_2}_{v_2} K \right\|_{(k_1, \kappa_2), \delta, \Delta^1_{.|\square}} \le C,$$

$$\left\| \nabla^{k_1}_{u_1} \nabla^{\kappa_2}_{u_2} K \right\|_{(\kappa_1, k_2), \delta, \Delta^2_{\square|.}} + \left\| \nabla^{k_1}_{v_1} \nabla^{\kappa_2}_{v_2} K \right\|_{(\kappa_1, k_2), \delta, \Delta^2_{.|\square}} \le C,$$

$$\left\| \nabla^{k_1}_{u_1} \nabla^{k_2}_{u_2} K \right\|_{(k_1, k_2), \delta, \Delta^1_{\square|.} \Delta^2_{\square|.}} + \left\| \nabla^{k_1}_{v_1} \nabla^{k_2}_{v_2} K \right\|_{(k_1, k_2), \delta, \Delta^1_{.|\square} \Delta^2_{.|\square}} \le C,$$

$$\left\|\nabla_{u_1}^{k_1}\nabla_{v_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\square|},\Delta^2_{\cdot|\square}} + \left\|\nabla_{v_1}^{k_1}\nabla_{u_2}^{k_2}K\right\|_{(k_1,k_2),\delta,\Delta^1_{\cdot|\square}\Delta^2_{\square|}} \le C. \qquad (6.4)$$

The least C such that (6.3) and (6.4) hold for each $\Upsilon \in \mathbf{3}$ will be denoted by $\|\Lambda\|_{K,f,k,\delta}$. Notice that the latter constant is preserved under full and partial adjoints as well. If $k_1 = k_2 = 0$, these are the usual full kernel estimates of a bi-parameter Calderón–Zygmund operator, see for example [41].

Say that $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ if the constant (6.2) is finite. The next definitions establish the main examples of forms in $\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$.

Definition 6.3 (*Bi-parameter wavelet form*). Associate to the two families of bi-parameter cancellative wavelets $\left\{\beta_z, \upsilon_z \in C\Psi_z^{k,\delta;0,0} : z \in Z^{\mathbf{d}}\right\}$ the bi-parameter wavelet form

$$\Lambda(f, g) = \int_{Z^{\mathbf{d}}} \langle f, \beta_z\rangle\langle \upsilon_z, g\rangle \, \mathrm{d}\mu(z). \qquad (6.5)$$

This form belongs to $\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ and $\|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)} \lesssim 1$. The weak boundedness property is a particular case of estimate (5.5) from Proposition 5.2. The partial kernel and full kernel estimates may be obtained by repeatedly employing (5.2) and standard computations.

In what comes next we will need several different BMO spaces. Some will be one-parameter and some will be two parameter, product BMO spaces of Chang and Fefferman, [8]. We will denote $\mathrm{BMO}_{\mathrm{CF}}(\mathbb{R}^{\mathbf{d}})$ as the product BMO space and $\mathrm{BMO}(\mathbb{R}^{d_\iota})$ will denote the usual one-parameter BMO space in the specified coordinate.

Definition 6.4 (*Bi-parameter paraproduct forms*). Paralleling our treatment of the one parameter case, three types of bi-parameter paraproducts will be defined. For a pair of multi-indices (γ_1, γ_2) on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, a function $b \in \mathrm{BMO}_{\mathrm{CF}}(\mathbb{R}^{\mathbf{d}})$, $(\iota_1, \iota_2) \in \{0, 1\}^2$ and $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$, define d

$$\Pi_{(0,0),b,(\gamma_1,\gamma_2)}(f, g) := \int_{Z^{\mathbf{d}}} \langle b, \varphi_{\gamma_1,z_1} \otimes \varphi_{\gamma_2,z_2}\rangle\langle f, \vartheta_{\gamma_1,z_1} \otimes \vartheta_{\gamma_2,z_2}\rangle\langle \varphi_z, g\rangle \, \mathrm{d}\mu(z),$$

$$\Pi_{(0,1),b,(\gamma_1,\gamma_2)}(f, g) := \int_{Z^{\mathbf{d}}} \langle b, \varphi_{\gamma_1,z_1} \otimes \varphi_{\gamma_2,z_2}\rangle\langle f, \vartheta_{\gamma_1,z_1} \otimes \varphi_{z_2}\rangle\langle \varphi_{z_1} \otimes \vartheta_{\gamma_2,z_2}, g\rangle \, \mathrm{d}\mu(z),$$

$$(6.6)$$

where $\varphi_{\gamma_j,z_j} = \mathsf{Sy}_{z_j}\partial^{-\gamma_j}\varphi_j$, see (2.8) and (2.15), and the family $\{\vartheta_{\gamma_j,z_j} \in C\Psi_z^{D,1;0} : z \in Z^{d_j}\}$ is a γ_j-family as in (4.5).

The first form is usually termed *full paraproduct*, while the second is usually referred to as a *partial paraproduct*. The paraproducts defined below are related as partial adjoints, namely

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

For this reason, the notation $\Pi_{b,\gamma}$ will be used in place of $\Pi_{(0,0),b,\gamma}$ throughout this section. Standard computations relying on the smoothness and compact support of the wavelets appearing in (6.6) lead to the following controls on the partial kernel and weak boundedness, and full kernel constants of the paraproduct forms: for any multi-index $\gamma = (\gamma_1, \gamma_2)$, there holds

$$\left\| \Pi_{b,\gamma} \right\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} \lesssim_k \|b\|_{\mathrm{BMO}_{\mathrm{CF}}(\mathbb{R}^{\mathbf{d}})}. \tag{6.7}$$

A third family of paraproducts, which are termed *half-paraproducts*, are constructed using the definitions (4.6) in each parameter $\iota \in \{1, 2\}$. Let $\kappa_{\hat{\iota}} \in \mathbb{N}$, $\eta > 0$, which are kept implicit in the notation, γ_ι be a multi-index on $\mathbb{R}^{d_\iota}$ and $\mathbf{a}$ be a continuous map on $Z^{d_{\hat{\iota}}} \times Z^{d_{\hat{\iota}}}$ taking values in $\mathrm{BMO}(\mathbb{R}^{d_\iota})$. Define, a priori on $\mathcal{S}(\mathbb{R}^{\mathbf{d}}) \times \mathcal{S}(\mathbb{R}^{\mathbf{d}})$, the form

$$\Pi^\iota_{\mathbf{a},\gamma_\iota,\alpha_\iota}(f, g) = \int\limits_{Z^{d_{\hat{\iota}}} \times Z^{d_{\hat{\iota}}}} \Pi_{\mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}}),\gamma_\iota,\alpha_\iota} \left(\langle f, \varphi_{z_{\hat{\iota}}} \rangle_{\hat{\iota}}, \langle g, \varphi_{\zeta_{\hat{\iota}}} \rangle_{\hat{\iota}} \right) [z_{\hat{\iota}}, \zeta_{\hat{\iota}}]_{\kappa_{\hat{\iota}}+\eta} \mathrm{d}\mu(z_{\hat{\iota}}) \mathrm{d}\mu(\zeta_{\hat{\iota}})$$
$$\tag{6.8}$$

where $\Pi_{\mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}}),\gamma_\iota,\alpha_\iota}$ refers to (4.6) for $b = \mathbf{a}(z_{\hat{\iota}},\zeta_{\hat{\iota}})$, $\gamma = \gamma_\iota$, $\alpha = \alpha_\iota$, $d = d_\iota$. Arguing in a similar fashion to (6.7), we record the estimate

$$\left\| \Pi^\iota_{\mathbf{a},\gamma_\iota,\alpha_\iota} \right\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}},k,\delta)} \lesssim \sup_{(z_{\hat{\iota}},\zeta_{\hat{\iota}}) \in Z^{d_{\hat{\iota}}} \times Z^{d_{\hat{\iota}}}} \|\mathbf{a}(z_{\hat{\iota}}, \zeta_{\hat{\iota}})\|_{\mathrm{BMO}(\mathbb{R}^{d_\iota})},$$
$$k = (k_1, k_2), \; k_\iota = |\gamma_\iota|, \; k_{\hat{\iota}} = \kappa_{\hat{\iota}}. \tag{6.9}$$

Furthermore, it is a particularly useful observation that, in view of (4.15) and referring to the notation introduced in (4.9), which is legitimately used whenever $f_2, g_2 \in \mathcal{S}(\mathbb{R}^{d_2})$ are fixed

$$\Pi^1_{\mathbf{a},\gamma_1,\alpha_1}(x_1^{\beta_1} \otimes f_2, g_1 \otimes g_2) = 0 \quad \forall 0 \le |\beta_1| \le |\gamma_1|, \; \beta_1 \ne \gamma_1 \tag{6.10}$$

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

The next technical lemma shows that the norm (6.2) controls certain symbols obtained by partial testing of Λ against monomials. For this, the modified wavelets

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

introduced in (4.22) will be needed. The proof of the next Lemma is postponed to the end of this section, see Sect. 6.3.

Lemma 6.5 *Let $\iota \in \{1, 2\}$, γ_ι be a multi-index in $\mathbb{R}^{d_\iota}$ and $k = (k_1, k_2)$ be such that $|\gamma_\iota| \le k_\iota$. For $\iota \in \{1, 2\}$, $z_{\hat{\iota}}, \zeta_{\hat{\iota}} \in Z^{d_{\hat{\iota}}}$, multi-indices γ_ι in $\mathbb{R}^{d_\iota}$, and $\mathfrak{a} \in \{\circ, \star\}$ define the*

functionals $\mathbf{q}_{\gamma_\iota}^{\iota,\mathfrak{a}}[\Lambda]$ *by*

$$
\begin{aligned}
\langle \mathbf{q}_{\gamma_1}^{1,\mathfrak{a}}[\Lambda](z_2, \zeta_2), f_1 \rangle &:= \Lambda^{\mathfrak{a}}(x_1^{\gamma_1} \otimes \chi_{z_2,\zeta_2}, |\nabla|^{|\gamma_1|} f_1 \otimes \chi_{\zeta_2,z_2}), \\
\langle \mathbf{q}_{\gamma_2}^{2,\mathfrak{a}}[\Lambda](z_1, \zeta_1), f_2 \rangle &:= \Lambda^{\mathfrak{a}}(\chi_{z_1,\zeta_1} \otimes x_2^{\gamma_2}, \chi_{\zeta_1,z_1} \otimes |\nabla|^{|\gamma_2|} f_2)
\end{aligned}
\tag{6.11}
$$

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

$$
\mathbf{a}_{\gamma_\iota,\alpha_\iota}^{\iota,\mathfrak{a}}[\Lambda](z_{\hat{\iota}}, \zeta_{\hat{\iota}}) := [z_{\hat{\iota}}, \zeta_{\hat{\iota}}]_{k_{\hat{\iota}}+\eta}^{-1} R^{\alpha_\iota} \mathbf{q}_{\gamma_\iota}^{\iota,\mathfrak{a}}(z_{\hat{\iota}}, \zeta_{\hat{\iota}}),
\tag{6.12}
$$

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

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

6.2 Bi-parameter Calderón–Zygmund forms of class (k, δ)

If α_ι is a multi-index on $\mathbb{R}^{d_\iota}$, the notation $p_R^{\alpha_\iota}$ refers to p_R^{α} from (4.8), with $d = d_\iota, \alpha = \alpha_\iota$. Below, $\mathcal{S}_D(\mathbb{R}^{\mathbf{d}})$ be the subspace of functions $\psi \in \mathcal{S}(\mathbb{R}^d)$ with the same bi-parameter vanishing moment property of the functions of $\Psi_z^{D,\delta;0,0}$, for some $z \in Z^{\mathbf{d}}$. Then, if $0 \leq |\alpha_\iota| \leq k_\iota$ for $\iota \in \{1, 2\}$, and $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$, and $\mathfrak{a} \in \vec{\mathfrak{a}}$, the limits

$$
\Lambda^{\mathfrak{a}}(x^{\alpha_1} \otimes x^{\alpha_2}, \psi) = \lim_{R \to \infty} \Lambda \left(p_R^{\alpha_1} \otimes p_R^{\alpha_2}, \psi \right), \qquad \psi \in \mathcal{S}_D(\mathbb{R}^{\mathbf{d}})
\tag{6.13}
$$

exist and define linear continuous functionals on $\mathcal{S}_D(\mathbb{R}^d)$: with the full kernel estimates at one's disposal, the proof presented in [23, Lemma 1.91] extends to the bi-parameter case without essential changes.

Remark 6.6 If $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ is a wavelet form of the type (6.5), an immediate byproduct of the cancellation properties of the families $\beta_z, \upsilon_z \in C\Psi_z^{k,\delta;0,0}$ is that the functionals

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

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

The next and final definition details our assumptions on the functionals (6.13) associated to $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$. We ask whether $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ has paraproducts of order κ for $0 \leq \kappa \leq \min\{k_1, k_2\}$ and, if that is the case, at the same time define the κ-th order cancellative part of Λ for all $0 \leq \kappa \leq \max\{k_1, k_2\}$.

Definition 6.7 (*Paraproducts of order* $0 \leq \kappa \leq \min\{k_1, k_2\}$) . Say that $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ has *paraproducts of order* 0 if for each $\mathfrak{a} \in \vec{\mathfrak{a}}$ there exists $b_0^{\mathfrak{a}} \in$

$\mathrm{BMO}_{\mathrm{CF}}(\mathbb{R}^{\mathbf{d}})$, the BMO product space of Chang and Fefferman [8], such that

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

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

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

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

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

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

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

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

$$\Lambda_{\kappa+1}(f, g) = \Lambda_\kappa(f, g) - \sum_{\substack{\gamma = (\gamma_1, \gamma_2) \\ |\gamma_1| = |\gamma_2| = \kappa+1}} \sum_{\mathfrak{a} \in \vec{\mathfrak{a}}} \left[\Pi_{b_\gamma^{\mathfrak{a}}, \gamma} \right]^{\mathfrak{a}} (f, g)$$

$$- \sum_{\substack{\mathfrak{a} \in \{\circ, \star\} \\ \iota \in \{1, 2\} \\ |\gamma_\iota| = |\alpha_\iota| = \kappa+1}} \left[\Pi_{\mathbf{a}_{\gamma_\iota, \alpha_\iota}^{\iota, \mathfrak{a}}[\Lambda_\kappa]}^{\iota} \right]^{\mathfrak{a}} (f, g). \tag{6.14}$$

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

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

$$\tag{6.15}$$

Definition 6.8 The bi-parameter bilinear form Λ belongs to the class $\mathrm{CZ}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ of (k, δ)-*Calderón–Zygmund forms* if $\Lambda \in \mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ and Λ has paraproducts of order $\min\{k_1, k_2\}$, with norm

$$\|\Lambda\|_{\mathrm{CZ}(\mathbb{R}^{\mathbf{d}}, k, \delta)} := \|\Lambda\|_{\mathrm{SI}(\mathbb{R}^{\mathbf{d}}, k, \delta)}$$

$$+ \sum_{\substack{0 \leq \kappa \leq \min\{k_1, k_2\} \\ |\gamma_1| = |\gamma_2| = \kappa \\ \mathfrak{a} \in \vec{\mathfrak{a}}}} \left\| |\nabla_{\mathbb{R}^{d_1}}|^{|\gamma_1|} |\nabla_{\mathbb{R}^{d_2}}|^{|\gamma_2|} T_\kappa^{\mathfrak{a}} (x_1^{\gamma_1} \otimes x_2^{\gamma_2}) \right\|_{\mathrm{BMO}(\mathbb{R}^{\mathbf{d}})}.$$

(6.16)

Forms in the $\mathrm{CZ}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ may be represented as a linear combination of a wavelet form of type (6.5) and order k plus a finite linear combination of paraproducts and half-paraproducts.

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

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

$\mathrm{BMO}_{\mathrm{CF}}(\mathbb{R}^{\mathbf{d}})$ *functions* $b_\gamma^{\mathfrak{a}}$, *and functions* $\mathbf{a}_{\gamma_\iota, \alpha_\iota}^{\iota, \mathfrak{a}}$ *on* $Z^{d_i} \times Z^{d_i}$ *taking values in a bounded subset of* $\mathrm{BMO}(\mathbb{R}^{d_\iota})$, $\iota = 1, 2$ *such that for all* $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$

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

$$+ \sum_{\substack{\gamma = (\gamma_1, \gamma_2) \\ 0 \leq |\gamma_1| = |\gamma_2| \leq \min\{k_1, k_2\} \\ \mathfrak{a} \in \vec{\mathfrak{a}}}} \left[\Pi_{b_\gamma^{\mathfrak{a}}, \gamma} \right]^{\mathfrak{a}} (f, g) + \sum_{\substack{\iota \in \{1, 2\} \\ 0 \leq |\gamma_\iota| \leq k_\iota \\ |\alpha_\iota| = |\gamma_\iota| \\ \mathfrak{a} \in \{\circ, \star\}}} \left[\Pi_{\mathbf{a}_{\gamma_\iota, \alpha_\iota}^{\iota, \mathfrak{a}}}^{\iota} \right]^{\mathfrak{a}} (f, g).$$

(6.17)

Theorem B will be proved in Sect. 7. The next corollary to Theorem B is easily proved by combining with (6.17) the estimates of Propositions 8.1, 8.2 and 8.3, and, for the cases $k \neq (0, 0)$, an integration by parts argument akin to the one used in the proof of Corollary A.1.

Corollary B.1 *Let* $k \in \mathbb{N}^2$, $\delta > 0$ *and* $\Lambda \in \mathrm{CZ}(\mathbb{R}^{\mathbf{d}}, k, \delta)$ *be as in Theorem B. Assume in addition the bi-parameter analogue of* (4.18)

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

(6.18)

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

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

(6.19)

Notice that (6.18) is not an additional assumption in the case $k = (0, 0)$ and, proceeding as in Remark 4.14, necessary for (6.19) to hold otherwise.

Remark 6.9 For the reader's convenience, we point out that among the bounds provided in these propositions, the exponent in (6.19) is achieved by the paraproduct estimate of Proposition 8.2 and their adjoints. In fact, if Λ is a $((0, 0), \delta)$ Calderón–Zygmund form whose paraproduct terms appearing in (6.17) all vanish, then the weighted norm bound for its adjoint

$$\|T\|_{L^p(\mathbb{R}^d; w)} \lesssim [w]_{A_p}^{\theta(p)}, \quad \theta(p) = \begin{cases} \frac{2}{p-1} & 1 < p \leq \frac{3}{2} \\ \frac{2p-1}{p-1} & \frac{3}{2} < p \leq 2 \\ \frac{p+1}{p-1} & 2 < p \leq 3 \\ 2 & p > 3 \end{cases} \tag{6.20}$$

may be read by applying Proposition 8.1 to the cancellative terms in (6.17) for Λ, if $p \geq 2$, or for its full adjoint $\Lambda^\star$ if $1 < p < 2$. Comparing with the one parameter case, see [35, Theorem 2], estimate (6.20) is sharp for $\max\{p, p'\} \geq 3$: there seem to be no instances of sharp weighted norm inequalities for bi-parameter operators in previous literature. Notice that the paraproduct free assumption covers, for instance, bi-parameter convolution-type operators.

6.3 Proof of Lemma 6.5

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

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

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

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

there holds

$$|\langle b, \psi \rangle| \lesssim [z_2, \zeta_2]_{k_2+\eta}. \tag{6.21}$$

The frequency support of ψ ensures that $\upsilon := t_1^{k_1} |\nabla|^{k_1} \psi \in \Psi_{w_1}^{3M, 1; 0}$. Now introduce the local notation, with reference to (2.2)

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

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

$$
\begin{aligned}
\langle b, \psi \rangle &= \Lambda(p \otimes \chi_{z_2,\zeta_2}, \upsilon \otimes \varphi_{\zeta_2}) \\
&= \sum_{m \sim n} + \sum_{m \nsim n} \Lambda(p_n \otimes \chi_{z_2,\zeta_2}, \upsilon_m \otimes \varphi_{\zeta_2}) + \Lambda(p_n \otimes \chi_{z_2,\zeta_2}, \upsilon_m \otimes \varphi_{\zeta_2})
\end{aligned}
\tag{6.22}
$$

where $m \sim n$ if $|m - n| < 2^4$ and $m \nsim n$ otherwise. The next task consists of bound each summand in the last right hand side of (6.22).

The $m \sim n$ summand

Notice that in this range $\|p_n\|_\infty \lesssim 2^{k_1 n}$, $\|\upsilon_m\|_\infty \lesssim t_1^{-d_1} 2^{-3Mm} \lesssim t_1^{-d_1} 2^{-2Mn}$ and p_n, υ_m are supported on $B_{(y_1,2^{n+2^5}t_1)}$. Also note that Θ_2 is supported on $B_{(\xi_2,4\sigma_2)}$ and $\|\Theta_2\|_\infty \lesssim [z_2, \zeta_2]_{k_2+\eta}$; this is obvious if $\chi_{z_2,\zeta_2} = \varphi_{\zeta_2}$ and may be read from (3.2) otherwise. Applying the weak boundedness property of Λ with balls $B_{(y_1,2^{n+2^5}t_1)}$ and $B_{(\xi_2,4\sigma_2)}$,

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

Furthermore, Ξ_2 and φ_{z_2} have separated supports. Therefore, using the partial kernel assumptions for the form $(f, g) \mapsto \Lambda(\Theta_1 \otimes f, \upsilon_{w_1} \otimes g)$ and repeating the computations of (4.28) for such form

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

The last two estimates are summable on the diagonal $m \sim n$ and this completes the bound for the $\sim$ summand in (6.22).

The $m \nsim n$ summand. We now have that p_n and υ_m have separated supports by $\sim t_1 2^{\max\{m,n\}}$. Applying the partial kernel assumptions for the form $(f, g) \mapsto \Lambda(f \otimes \Theta_2, g \otimes \varphi_{\zeta_2})$ and arguing as in (4.28) yields

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

Finally, in the term below, the full kernel assumptions may be used due to functions in both parameters having disjoint supports. Standard computations relying on the kernel estimates as in (4.28) then lead to

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

The above estimates are summable over $m \sim n$, which completes both this case and the proof of the Lemma.

Remark 6.10 We explain the details of the definitions (6.11). Using symmetry with respect to adjoints, it suffices to study $\mathbf{q}_{\gamma_1}^{1,\circ}(z_2, \zeta_2)$. The most complicated case is when either one of χ_{z_2,ζ_2}, χ_{ζ_2,z_2} contains the polynomial summand. To fix ideas, work with $\chi_{z_2,\zeta_2} = \varphi_{z_2} - P_{z_2,\zeta_2}$. Let $\phi_2 = \phi$ as in (4.7) with $d = d_2$ and $p_R^{\gamma_1}$ as in (4.8) with $d = d_1$ and $\alpha = \gamma_1$, see Sect. 4.2. If $0 \notin \operatorname{supp} \widehat{f_1}$ then $g_1 = |\nabla|^{|\gamma_1|} f_1 \in \mathcal{S}_D(\mathbb{R}^{d_1})$. The partial kernel estimates of Λ readily show that

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

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

7 Proof of Theorem B

Below $0 < \delta \leq 1, 0 < \varepsilon < \delta$ are fixed. Set $\eta = \frac{\delta+\varepsilon}{2}$, so that $\varepsilon < \eta < \delta$.

7.1 Preliminaries

Before entering the main argument, a series of preparatory lemmas is required. First, the two parameter version of Lemma 3.2 is provided.

Lemma 7.1 Let φ_z be as in (6.1) and $u : Z^{\mathbf{d}} \to \mathbb{C}$ be a Borel measurable function with $|u(z)| \leq 1$. Then, there exists $C \lesssim_{k_1, k_2, \varepsilon} 1$ such that for all $z = (z_1, z_2) \in Z^{\mathbf{d}}$ with $z_j = (x_j, s_j)$, $j = 1, 2$

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

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

Proof There is a direct argument along the lines of the one parameter proof. However, an argument that uses Lemma 3.2 as a black box will be given. To save space in the notation, this argument is carried out for $x_j = 0$, $s_j = 1$, $j = 1, 2$. Notice that for each fixed $w_1 \in \mathbb{R}^{d_1}$

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

where

$$v_{\alpha_2,\beta_2} = \int\limits_{(\alpha_1,\beta_1)\in Z^{d_1}} [(\alpha_1,\beta_1)]_{k_1+\eta} u((\alpha_1,\beta_1),(\alpha_2,\beta_2))(\varphi_1)_{(x_1+\alpha_1 s_1, \beta_1 s_1)}$$

$$\times \frac{d\beta_1 d\alpha_1}{\beta_1} \in C\Psi_{z_1}^{k_1,\varepsilon;0}$$

with uniform constant C by an application of (3.5) of Lemma 3.2 with η in place of δ. In particular the function $u_{w_1}(\alpha_2,\beta_2) := \langle w_1 \rangle^{(d_1+k_1+\varepsilon)} v_{\alpha_2,\beta_2}(w_1)$ is uniformly bounded. Therefore, another application of (3.5), with $u = u_{w_1}(\alpha_2,\beta_2)$ entails

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

Repeating for the second parameter and comparing with Eq. (5.2), proves that $v_z \in C\Psi_z^{(k_1,k_2),\varepsilon;0,0}$ and thus completes the proof of the lemma. $\qquad\square$

The notation $\chi_{z_\iota,\zeta_\iota}$ appearing below refers to (4.22).

Lemma 7.2 *For $(z,\zeta) \in Z^{\mathbf{d}} \times Z^{\mathbf{d}}$ with $z = (z_1,z_2)$, $\zeta = (\zeta_1,\zeta_2)$,*

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

Proof By symmetry with respect to the adjoint, it suffices to consider the case where $z_j \in Z_+^{d_j}(\zeta_j)$ for both $j = 1, 2$. Nine different cases according to which of the sets in (2.1) with $\zeta = \zeta_j$ each z_j belongs to need to be considered. Only the case $z_j \in A(\zeta_j)$ for $j = 1, 2$ will be dealt with explicitly; all remaining cases may be dealt with by the same strategy that will be used for the summands appearing in (7.1). In this case, $\chi_{z_j,\zeta_j} = \varphi_{z_j} - P_{z_j,\zeta_j}$ and $\chi_{\zeta_j,z_j} = \varphi_{\zeta_j}$ for $j = 1, 2$, and thus

$$\Lambda \left(\chi_{z_1,\zeta_1} \otimes \chi_{z_2,\zeta_2}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2} \right) = \sum_{(\iota,J)\in\{\text{in},\text{out}^2\}} \Lambda \left(\Theta_{1,\iota} \otimes \Theta_{2,J}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2} \right), \tag{7.1}$$

$$\Theta_{j,\text{in}} := \chi_{z_j,\zeta_j}\alpha_{\zeta_j}, \quad \Theta_{j,\text{out}} := \chi_{z_j,\zeta_j}\beta_{\zeta_j}, \quad j = 1, 2.$$

Each term in (7.1) will be estimated. The key to the first three summands is that for $j = 1, 2$ the function $\Theta_{j,\text{in}}$ is supported on $4B_{\zeta_j}$ and, from (3.2), $\|\Theta_{j,\text{in}}\|_\infty \lesssim [z_j,\zeta_j]_{k_j+\eta}$. For the in, in summand, employ the weak boundedness of Λ with points $\tilde{\zeta}_j = (\xi_j, 4\sigma_j)$ thus obtaining

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

The in, out summand is bounded as follows. Observe that $\Theta_{2,\text{out}}$ and φ_{ζ_2} have separated support. Now, apply the partial kernel/weak boundedness assumption to the form $(f, g) \mapsto \sigma_1^{d_1} \Lambda(\Theta_{1,\text{in}} \otimes f, \varphi_{\zeta_1}, g)$ at point $\tilde{\zeta}_1 = (\xi_1, 4\sigma_1)$, to which estimate (4.28) with $z = z_2$ and $\zeta = \zeta_2$ actually applies. Such estimate returns a factor of $[z_2, \zeta_2]_{k_2+\eta}$ while the factor $[z_1, \zeta_1]_{k_1+\eta}$ is obtained from $\|\Theta_{1,\text{in}}\|_\infty$. The out, in summand is handled in exactly the same way.

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

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

is the tensor product of the steps (4.25)–(4.28) performed in each variable, the only difference being how the corresponding term in (4.25) involving the finite difference of the derivatives of the kernel is controlled. In that case, for a fixed $v = (v_1, v_2) \in (2B_{\zeta_1} \times 2B_{\zeta_2})^c$ one uses the cancellation and L^1-estimate of $\sigma_j^{-k_j} \partial^{-\gamma_j} \varphi_{\zeta_j}$ and bounds

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

using the kernel estimate in the fourth line (6.4). This completes the proof of the Lemma. $\qquad\square$

7.2 Main line of proof of Theorem B

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

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

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

7.2.1 Base case and main part of inductive step

The base case and main part of the inductive step of the proof works under the additional assumption referring to Lemma 6.5

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

namely, all paraproducts $T^{\mathfrak{a}}(x^{\gamma_1} \otimes x^{\gamma_2})$ vanish except possibly those with $|\gamma_1| = |\gamma_2| = \kappa$, and all half-paraproducts vanish except possibly those of highest order.

Clearly, $a(0)$ is not an extra assumption. Moreover, if $\kappa \geq 1$, assumption $a(\kappa)$ implies that Λ coincides with its $(\kappa - 1)$th order cancellative part $\Lambda_{\kappa-1}$, which means we are allowed to conflate the two forms and just write Λ below.

Let now $f, g \in \mathcal{S}(\mathbb{R}^{\mathbf{d}})$. Using the bi-parameter analogue of (2.11), bilinearity and $\mathcal{S}(\mathbb{R}^{\mathbf{d}})$-continuity of Λ, and later the definition of $U(z, \zeta)$ leads to the decomposition

$$
\begin{aligned}
\Lambda(f, g) &= \int_{Z^{\mathbf{d}} \times Z^{\mathbf{d}}} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(\varphi_z, \varphi_\zeta) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) \\
&= \int_{Z^{\mathbf{d}} \times Z^{\mathbf{d}}} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda \left(\chi_{z_1, \zeta_1} \otimes \chi_{z_2, \zeta_2}, \chi_{\zeta_1, z_1} \otimes \chi_{\zeta_2, z_2} \right) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) \\
&\quad + \int_{Z_{\zeta_1}^{d_1} \times Z_{\zeta_2}^{d_2}} \int_{Z_{z_2}^{d_2}} \int_{z_1 \in A(\zeta_1)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(P_{z_1, \zeta_1} \otimes \chi_{z_2, \zeta_2}, \varphi_{\zeta_1} \otimes \chi_{\zeta_2, z_2}) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) + \cdots \\
&\quad + \int_{Z_{\zeta_1}^{d_1} \times Z_{\zeta_2}^{d_2}} \int_{z_1 \in A(\zeta_1)} \int_{z_2 \in A(\zeta_2)} \langle f, \varphi_z \rangle \langle \varphi_\zeta, g \rangle \Lambda(P_{z_1, \zeta_1} \otimes P_{z_2, \zeta_2}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2}) \, \mathrm{d}\mu(z) \mathrm{d}\mu(\zeta) + \cdots.
\end{aligned}
$$
$$\tag{7.2}$$

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

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

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

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

It is possible to turn the first summand in (7.2) into the first summand of (6.17). First, make the change of variable

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

and then use Fubini's theorem in the inner variable of g. The first summand in (7.2) then equals

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

$$\upsilon_z := \int_{\substack{(\alpha_1,\beta_1)\in Z^{d_1} \\ (\alpha_2,\beta_2)\in Z^{d_2}}} \Lambda\left(\chi_{z_1,\zeta_1} \otimes \chi_{z_2,\zeta_2}, \chi_{\zeta_1,z_1} \otimes \chi_{\zeta_2,z_2}\right) \varphi_\zeta \frac{\mathrm{d}\beta_2 \mathrm{d}\alpha_2 \mathrm{d}\beta_1 \mathrm{d}\alpha_1}{\beta_1\beta_2},$$

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

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

is uniformly bounded via Lemma 7.2, and applying Lemma 7.1 yields $\upsilon_z \in C\Psi_z^{(\kappa,\min\{\kappa,k_2\}),\varepsilon;0}$.

It remains to identify the remaining terms in (7.2) as a sum of paraproduct terms. Here it is crucial to use assumption $a(\kappa)$, which tells us that

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

Focus on the term in the third line of (7.2) first. The above observation, the definition of P_{z_1,ζ_1} from Lemma 3.1, the definition of $\mathbf{q}_{\gamma_1}^{1,\circ}$, the fact that $\partial^{-\alpha_1} = R^{\alpha_1}|\nabla|^{-|\alpha_1|}$ with the definition of $\varphi_{\alpha_1,\zeta_1}$, see (2.8) and Remark 2.1, gives

$$\Lambda(P_{z_1,\zeta_1} \otimes \chi_{z_2,\zeta_2}, \varphi_{\zeta_1} \otimes \chi_{\zeta_2,z_2}) = \sum_{|\gamma_1|=k_1} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle \Lambda(x_1^{\gamma_1} \otimes \chi_{z_2,\zeta_2}, \sigma_1^{-k_1}\varphi_{\zeta_1} \otimes \chi_{\zeta_2,z_2})$$

$$= \sum_{|\gamma_1|=k_1} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle \langle \mathbf{q}_{\gamma_1}^{1,\circ}(z_2,\zeta_2), \mathsf{Sy}_z[|\nabla|^{-k}\varphi_1]\rangle$$

$$= \sum_{|\gamma_1|=|\alpha_1|=k_1} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1}\rangle \langle R^{\alpha_1}\mathbf{q}_{\gamma_1}^{1,\circ}(z_2,\zeta_2), \varphi_{\alpha_1,z_1}\rangle.$$

Finally using Lemma 3.3 with $h = \langle f, \varphi_{z_2}\rangle_2$, the summand in the third line of (7.2) equals the sum over $|\gamma_1| = |\alpha_1| = \kappa$ of

$$\int_{(Z^{d_2})^2} \int_{Z_{\zeta_1}^{d_1}} \langle \mathbf{a}_{\gamma_1,\alpha_1}^{1,\circ}(z_2,\zeta_2), \varphi_{\alpha_1,z_1}\rangle \left\langle \langle f, \varphi_{z_2}\rangle_2, \vartheta_{\gamma_1,\zeta_1}\right\rangle_1 \left\langle \langle \varphi_{\zeta_2}, g\rangle_2, \varphi_{\zeta_1}\right\rangle_1$$

$$\times \, \mathrm{d}\mu(\zeta_1) \, [z_2,\zeta_2]_{k_2+\eta} \mathrm{d}\mu(z_2)\mathrm{d}\mu(\zeta_2)$$

$$= \int_{Z^{d_2}\times Z^{d_2}} \Pi_{\mathbf{a}_{\gamma_1,\alpha_1}^{1,\circ}(z_2,\zeta_2),\gamma_1,\alpha_1}\left(\langle f, \varphi_{z_2}\rangle_2, \langle g, \varphi_{\zeta_2}\rangle_2\right)[z_2,\zeta_2]_{k_2+\eta}\mathrm{d}\mu(z_2)\mathrm{d}\mu(\zeta_2)$$

$$= \Pi_{\mathbf{a}_{\gamma_1,\alpha_1}^{1,\circ},\gamma_1,\alpha_1}(f,g)$$

where $\kappa_2 = \min\{\kappa, k_2\}$, which is one of the summands in the second line of (6.17). The three other types of summands in the second and third line of (6.17), constructed in exactly the same way, arise from the $\cdots$ terms in the third line of (7.2).

It remains to identify the terms of the type appearing in the third line (7.2). Using again $a(\kappa)$, these terms will appear only if $\kappa \leq k_2$. Lemma 3.1 and the definition of the paraproducts of Λ then yield

$$\Lambda(P_{z_1,\zeta_1} \otimes P_{z_2,\zeta_2}, \varphi_{\zeta_1} \otimes \varphi_{\zeta_2})$$
$$= \sum_{|\gamma_1|=|\gamma_2|=\kappa} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \rangle \langle \varphi_{z_2}, \mathsf{Sy}_{\zeta_2}\phi_{\gamma_2} \rangle \Lambda(x_1^{\gamma_1} \otimes x_2^{\gamma_2}, \sigma_1^{-k_1}\varphi_{\zeta_1} \otimes \sigma_2^{-k_2}\varphi_{\zeta_2})$$
$$= \sum_{|\gamma_1|=|\gamma_2|=\kappa} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \rangle \langle \varphi_{z_2}, \mathsf{Sy}_{\zeta_2}\phi_{\gamma_2} \rangle \langle b_\gamma^\circ, \varphi_{\gamma_1,\zeta_1} \otimes \varphi_{\gamma_2,\zeta_2} \rangle.$$

An application of Lemma 3.3 with $h = \langle f, \varphi_{z_1} \rangle_1$ yields

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

so that the summand in the third line of (7.2) equals the sum over $|\gamma_1| = \kappa_1$, $|\gamma_2| = \kappa_2$ of

$$\int_{Z^{d_2} \times Z^{d_2}} \langle b_\gamma^\circ, \varphi_{\gamma_1,\zeta_1} \otimes \varphi_{\gamma_2,\zeta_2} \rangle \langle \varphi_\zeta, g \rangle \int_{z_1 \in A(\zeta_1)} \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \rangle F(z_1) d\mu(z_1) d\mu(\zeta)$$
$$= \int_{Z^{d_2} \times Z^{d_2}} \langle b_\gamma^\circ, \varphi_{\gamma_1,\zeta_1} \otimes \varphi_{\gamma_2,\zeta_2} \rangle \langle \varphi_\zeta, g \rangle \int_{z_1 \in A(\zeta_1)} \langle \langle f, \vartheta_{\gamma_2,\zeta_2} \rangle_2, \varphi_{z_1} \rangle_1 \langle \varphi_{z_1}, \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \rangle d\mu(z_1) d\mu(\zeta)$$
$$= \int_{Z^{d_2} \times Z^{d_2}} \langle b_\gamma^\circ \mathsf{Sy}_{\zeta_1}\phi_{\gamma_1} \otimes \varphi_{\gamma_2,\zeta_2} \rangle \langle \varphi_\zeta, g \rangle \langle f, \vartheta_{\gamma_1,\zeta_1} \otimes \vartheta_{\gamma_2,\zeta_2} \rangle d\mu(\zeta) = \Pi_{b_\gamma^\circ,\gamma}^\circ(f, g)$$

where another application of Lemma 3.3 with $h = \langle f, \vartheta_{\gamma_2,\zeta_2} \rangle_2$ has been carried out and the definition of full paraproduct is finally taken advantage of: this is one of the terms appearing in the fourth line of (6.17). This procedure may be repeated for the additional terms in the third line of (7.2), thus completing the roster of terms in (6.17)

under the additional assumption $a(\kappa)$. Namely, under this assumption, we have proved that

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

$$+ \begin{cases} \displaystyle\sum_{\substack{|\gamma_1|=|\gamma_2|=\kappa \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b_\gamma^{\mathfrak{a}},\gamma}\right]^{\mathfrak{a}}(f,g) + \sum_{\substack{\iota\in\{1,2\} \\ \mathfrak{a}\in\{\circ,\star\} \\ |\alpha_\iota|=|\gamma_\iota|=\kappa}} \left[\Pi_{\mathbf{a}_{\gamma_\iota,\alpha_\iota}^{\iota,\mathfrak{a}}}\right]^{\mathfrak{a}}(f,g) & \kappa \le k_2 \\[3em] \displaystyle\sum_{\substack{|\alpha_1|=|\gamma_1|=\kappa \\ \mathfrak{a}\in\{\circ,\star\}}} \left[\Pi_{\mathbf{a}_{\gamma_1,\alpha_1}^{1,\mathfrak{a}}}\right]^{\mathfrak{a}}(f,g) & k_2 < \kappa \le k_1 \end{cases}$$

$$(7.3)$$

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

The assumption $a(\kappa)$ is then removed by an inductive argument. Recall that $k_1 \ge k_2$. Let $0 \le \kappa < k_1$ and assume that the representation (6.17) holds true for $k = (\kappa, \min\{\kappa, k_2\})$. Let $\widetilde{\Lambda}(f,g)$ be the form obtained by subtracting from Λ the second line of (6.17). Then

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

coincides with the κth order cancellative part of Λ, is a bi-parameter wavelet form of type (6.5) and satisfies assumption $a(\kappa+1)$ of having all the relevant paraproducts up to order κ vanishing, see Remark 6.6. We may thus apply the main step to $\widetilde{\Lambda}(f,g)$ with $\kappa+1$ in place of κ, resulting in (7.3), and obtain that

$$\Lambda(f,g) = \widetilde{\Lambda}(f,g) + \sum_{\substack{\kappa'\le\min\{\kappa,k_2\} \\ }} \sum_{\substack{\gamma=(\gamma_1,\gamma_2) \\ |\gamma_1|=|\gamma_2|=\kappa' \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b_\gamma^{\mathfrak{a}},\gamma}\right]^{\mathfrak{a}}(f,g)$$

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

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

$$+ \sum_{\substack{\kappa'\le\min\{\kappa+1,k_2\} \\ }} \sum_{\substack{\gamma=(\gamma_1,\gamma_2) \\ |\gamma_1|=|\gamma_2|=\kappa' \\ \mathfrak{a}\in\vec{\mathfrak{a}}}} \left[\Pi_{b_\gamma^{\mathfrak{a}},\gamma}\right]^{\mathfrak{a}}(f,g)$$

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

with $\{\widetilde{\upsilon}_z, \varphi_z \in C\Psi_z^{(\kappa+1,\min\{\kappa+1,k_2\}),\varepsilon;0,0} : z \in Z^{\mathbf{d}}\}$. This achieves (6.17) for Λ with $k = (\kappa + 1, \min\{\kappa + 1, k_2\})$, thus completing the inductive step and the proof of Theorem B.

8 Weighted estimates for intrinsic operators

This section contains the proofs of quantitative, and in some cases sharp, weighted estimates for the four types of summands occurring in the representation (6.17): see Propositions 8.1 and 8.2. Throughout, $[w]_{A_p}$ denotes the standard product weight characteristic on $\mathbb{R}^{\mathbf{d}} = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, see for example [18, 25].

8.1 Quantitative bounds for bi-parameter Calderón–Zygmund model operators

To begin with, the operator T appearing in the following proposition is the adjoint of the first summand in (6.17), in the basic case $k = 0$.

Proposition 8.1 *For $\delta > 0$ and $\left\{\upsilon_z \in C\Psi_z^{\delta;0,0} : z \in Z^{\mathbf{d}}\right\}$ consider the operator*

$$Tf = \int_{Z^{\mathbf{d}}} \langle f, Sy_{z_1}\varphi_1 \otimes Sy_{z_2}\varphi_2 \rangle \upsilon_z \mathrm{d}\mu(z).$$

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

Next, adjoints to the full and partial paraproduct terms in (6.17) are treated: compare with the definitions in (6.6). We unfortunately do not know whether the exponents of the weighted constants in the estimates below are sharp.

Proposition 8.2 *Let $D \geq 8(d_1+d_2)$. Fix $b \in \mathrm{BMO}(\mathbb{R}^{\mathbf{d}})$, $\{\vartheta_{z_j} \in C\Psi_{z_j}^{D,1;1} : z_j \in Z^{d_j}\}$, $\{\upsilon_{z_j}, \psi_{z_j} \in C\Psi_{z_j}^{D,1;0} : z_j \in Z^{d_j}\}$ for $j = 1, 2$. Then, the operators*

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

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

satisfy the estimates

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

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

The last proposition concerns adjoints to the half paraproduct terms in (6.17), see (6.8).

Proposition 8.3 *Let $0 < \delta < 1$, $\mathbf{a} \in \mathcal{C}(Z^{d_2} \times Z^{d_2}; \mathrm{BMO}(\mathbb{R}^{\mathbf{d}}))$, and fix families*

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

Then, the operator

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

satisfies the estimate

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

The proofs of Propositions 8.1, 8.2 and 8.3 are collected in Sect. 8.2. Along the way, we will make use of sharp weighted bounds for the intrinsic square function (5.4), as well as the mixed square-maximal operators

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

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

which enter the L^p and weighted theory of the full and partial paraproduct terms. The square-maximal and maximal-square operators appearing below generalize those introduced in [46, 47]. There seem to be no pre-existing weighted estimate in past literature, thus our results are stated as a theorem.

Theorem C *The operator norm bound*

$$\|\mathrm{SS}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\max\left\{1, \frac{2}{p-1}\right\}}, \quad \|\mathrm{SM}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}, \|\mathrm{MS}\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p}^{\frac{1}{p-1}\max\left\{2, \frac{p+1}{2}\right\}},$$
$$\tag{8.1}$$

holds for all $0 < \delta \leq 1$ and $1 < p < \infty$. Furthermore, the exponent of (8.1) may not be improved for a generic weight w.

The SS bound in Theorem C is the bi-parameter analogue of [35, Theorem 1.2]. Its proof is given in the concluding Sect. 8.3 below. Another inequality that will be used a few times in Sect. 8.2 is a lower bound for the smaller tensor product square function

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

associated with the wavelets φ_1, φ_2 from (6.1). The proof is a simple iteration argument, and is given immediately.

Proposition 8.4 $\|f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)} \lesssim [w]_{A_p} \|\mathrm{SS}_{\otimes} f\|_{L^p(\mathbb{R}^{\mathbf{d}};w)}$ *for all* $1 < p < \infty$.

Proof Apply the main result of [61], see also [38, Theorem 2.7], first on each x_1-fiber in the second parameter, and subsequently in vector-valued form in the second parameter to see that

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

as claimed. □

8.2 Proofs of Propositions 8.1, 8.2 and 8.3

Proof of Proposition 8.1 Sharpness of $[w]_{A_p}^2$ for $p \geq 3$ follows by taking the tensor product of two counterexamples to sharpness of $[w]_{A_p}$ in one parameter. For the rest of the proof, we claim the pointwise bound

$$
\mathrm{SS}_{\otimes}(Tf) \lesssim \mathrm{SS}f. \tag{8.2}
$$

Assuming (8.2) holds,

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

thanks to an application of Proposition 8.4 in the first step and of Theorem C in the last. This proves Proposition 8.1 up to the verification of claim (8.2), which follows. Fix

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

Using the notation (6.1) for φ_ζ, writing $z = ((x_1, t_1), (x_2, t_2)) \in Z^{\mathbf{d}}$, and making the usual change of variable

$$\langle Tf, \varphi_\zeta \rangle = \int_{z \in Z^{\mathbf{d}}} \langle f, \varphi_z \rangle \langle v_z, \varphi_\zeta \rangle \, \mathrm{d}\mu(z) = \langle f, \psi_\zeta \rangle,$$

$$\psi_\zeta := \int_{\substack{(\alpha_1, \beta_1) \in Z^{d_1} \\ (\alpha_2, \beta_2) \in Z^{d_2}}} \langle v_{((\xi_1+\alpha_1\sigma_1, \beta_1\sigma_1), (\xi_2+\alpha_2\sigma_2, \beta_2\sigma_2))}, \varphi_\zeta \rangle \varphi_{((\xi_1+\alpha_1\sigma_1, \beta_1\sigma_1), (\xi_2+\alpha_2\sigma_2, \beta_2\sigma_2))}$$

$$\frac{\mathrm{d}\beta_2 \mathrm{d}\alpha_2 \mathrm{d}\beta_1 \mathrm{d}\alpha_1}{\beta_1 \beta_2}.$$

Applying Lemma 5.1,

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

whence by Lemma 7.1, $\psi_\zeta \in C\Psi_\zeta^{\frac{\delta}{2}; 0, 0}$, and (8.2) follows immediately from the definition of the intrinsic square function SS. $\qquad\square$

Proof of Proposition 8.2 Let $\sigma := w^{-\frac{1}{p-1}}$ be the dual weight to $w \in A_p$, so that $[\sigma]_{A_{p'}} = [w]_{A_p}^{\frac{1}{p-1}}$. Recall that M_{d_1, d_2} is the bi-parameter maximal function on $\mathbb{R}^{\mathbf{d}}$. The proof for $\Pi_{(0,0),b}$ begins with an appeal to $H^1 - \mathrm{BMO}$ duality, leading to

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

The passage to the second line is justified by a pointwise bound, whose proof is similar to (8.3) below, and is omitted. In the last line, Theorem C has been appealed to, and to the quantitative weighted estimate for the strong maximal function and square functions. The claimed estimate for $\Pi_{(0,0),b}$ then follows by duality.

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

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

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

As $D \geq 8d_1, 8d_2$, Lemma 2.3 applied componentwise to bound $\langle \upsilon_{z_1} \otimes \upsilon_{z_2}, \mathsf{Sy}_{(x_1,t_1)}\varphi_1 \otimes \mathsf{Sy}_{(x_2,t_2)}\varphi_2 \rangle$, with $z_j = (x_j + \alpha_j t_j, \beta_j t_j)$, $j = 1, 2$ then yields

$$\left| \left\langle \int_{Z^{\mathbf{d}}} \langle f, \vartheta_{z_1} \otimes \psi_{z_2} \rangle \langle \psi_{z_1} \otimes \theta_{z_2}, g \rangle \upsilon_{z_1} \otimes \upsilon_{z_2} \, \mathrm{d}\mu(z), \mathsf{Sy}_{(x_1,t_1)}\varphi_1 \otimes \mathsf{Sy}_{(x_2,t_2)}\varphi_2 \right\rangle \right|$$

$$\lesssim \left(\sup_{\psi \in \Psi^{\delta;0,1}_{((x_1,t_1),(x_2,t_2))}} |\langle f, \psi \rangle| \right) \left(\sup_{\psi \in \Psi^{\delta;1,0}_{((x_1,t_1),(x_2,t_2))}} |\langle g, \psi \rangle| \right). \tag{8.3}$$

The proof proper begins now. Using $H^1 - \mathrm{BMO}$ duality again, followed by (8.3) and one application of $L^2 - L^\infty$ Hölder inequality in each parameter,

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

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

Proof of Proposition 8.3 This proof relies on the auxiliary operators

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

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

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

which is a shifted square function in the second parameter with smooth, compactly supported mother wavelet φ_2 as in (6.1); the simplified notation $\mathsf{S}_{(2)}$ is used in place of $\mathsf{S}_{(2),(0,1)}$. The main results of [7, 35] yield the operator norm bounds

$$\|\mathsf{S}_{(2),(\alpha_2,\beta_2)}\|_{L^p(\mathbb{R}^{d_2};W)} \lesssim_\varepsilon (\min\{1, \beta_2\})^{-\varepsilon} [W]_{A_p}^{\max\{\frac{1}{2}, \frac{1}{p-1}\}} \tag{8.4}$$

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

$$\Pi_{\mathbf{a}} f(u) = \int\limits_{Z_{z_2}^{d_2}} \int\limits_{Z_{\zeta_2}^{d_2}} [z_2, \zeta_2]_\delta P_{a(z_2, \zeta_2)}(\langle f, \psi_{z_2} \rangle_2)(u_1) \otimes \psi_{\zeta_2}(u_2) \mathrm{d}\mu(\zeta_2) \mathrm{d}\mu(\zeta_1),$$

$$u = (u_1, u_2) \in \mathbb{R}^{\mathbf{d}}.$$

A calculation involving Lemma 2.3 applied to the inner product $\langle \psi_{\zeta_2}, \mathsf{Sy}_{(y_2, t_2)} \varphi_2 \rangle$ followed by the change of variables $z_2 = (y_2 + a_2 t_2, b_2 t_2)$, $\zeta_2 = (y_2 + \alpha_2 t_2, \beta_2 t_2)$ then yields

$$S_{(2)}[\Pi_{\mathbf{a}} f](y_1, y_2)$$

$$\lesssim \int\limits_{\substack{\omega_2 := (\alpha_2, \beta_2) \in Z^{d_2} \\ w_2 := (a_2, b_2) \in Z^{d_2}}} \left(\int\limits_0^\infty \left| P_{\mathbf{a}((y_2 + a_2 t_2, b_2 t_2),(y_2 + \alpha_2 t_2, \beta_2 t_2))} \left(\langle f, \psi_{(y_2 + a_2 t_2, b_2 t_2)} \rangle_2 \right)(y_1) \right|^2 \frac{\mathrm{d}t_2}{t_2} \right)^{\frac{1}{2}}$$

$$\times [\omega_2]_1 [w_2, \omega_2]_\delta \, \mathrm{d}\mu(w_2) \mathrm{d}\mu(\omega_2).$$

Applying the reverse square function bound of [61] in the second parameter, followed by the sharp weighted estimate for the vector-valued paraproduct $P_{\mathbf{a}((y_2 + a_2 t_2, b_2 t_2),(y_2 + \alpha_2 t_2, \beta_2 t_2))}$ to pass to the second line, and finally appealing to (8.4) with choice $\varepsilon = \frac{\delta}{2}$, we obtain

$$\|\Pi_{\mathbf{a}} f\|_{L^p(\mathbb{R}^{\mathbf{d}}; w)} \lesssim [w]_{A_p}^{\frac{1}{2}} \|S_{(2)} \Pi_{\mathbf{a}} f\|_{L^p(\mathbb{R}^{\mathbf{d}}; w)}$$

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

$$\frac{\mathrm{d}a_2 \mathrm{d}b_2}{b_2} \frac{\mathrm{d}\alpha_2 \mathrm{d}\beta_2}{\beta_2}$$

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

For display reasons, above $\|\mathbf{a}\|$ stands for $\|\mathbf{a}\|_{\mathcal{C}(Z^{d_2} \times Z^{d_2}; \mathrm{BMO}(\mathbb{R}^{\mathbf{d}}))}$. The proof of Proposition 8.3 is thus complete. $\qquad\square$

8.3 Proof of Theorem C

Sharpness of the exponent follows by tensor product of the usual one parameter examples. The one parameter square function example is discussed in [35] and references therein, while the example for the one parameter maximal operator is entirely classical.

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

Fix $f \in L_0^\infty(\mathbb{R}^{\mathbf{d}})$ and let $\left\{ \psi_{(x_1,t_1),(x_2,t_2)} \in \Psi_{(x_1,t_1),(x_2,t_2)}^{\delta;0,0}; x_j \in \mathbb{R}^{d_j}, 0 < t_j < \infty, \right.$ $j = 1, 2\}$ be a family linearizing the supremum in (5.4). Throughout this proof, $\eta := \frac{\delta}{16}$. The first step consists of a decomposition of the linearizing family into wavelets with compact frequency support in one of the parameters. Let $\alpha \in \mathcal{S}(\mathbb{R}^{d_1})$ be a radial function with

$$\operatorname{supp} \widehat{\alpha} \subset B_{(0,2)} \backslash B_{(0,\frac{1}{2})}, \quad \int_{-\infty}^{\infty} \widehat{\alpha}(2^s \xi) \, ds = \mathbf{1}_{\mathbb{R}^{d_1} \backslash \{0\}}(\xi),$$

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

$$\operatorname{supp} \widehat{\beta} \subset B_{(0,3)} \backslash B_{(0,\frac{1}{3})}, \quad \widehat{\beta}(\xi) = 1 \quad \forall \xi \in \operatorname{supp} \widehat{\alpha}.$$

Set

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

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

Lemma 8.5 *For all $s \in \mathbb{R}$, $x_j \in \mathbb{R}^{d_j}$, $0 < t_j < \infty$, $j = 1, 2$ we have $\psi_{(x_1,t_1),(x_2,t_2)}^s \in C\Psi_{(x_1,t_1),(x_2,t_2)}^{\eta;0,0}$.*

Proof By bi-parametric invariance of the assumption and assertion, it suffices to prove the case $x_j = 0_{\mathbb{R}^{d_j}}, t_j = 1$ for $j = 1, 2$. For simplicity write ψ in place of $\psi_{(x_1,t_1),(x_2,t_2)}$. Applying Lemma 2.3 for each fixed y_2 gives

$$|\psi *_1 \alpha_s(y_1, y_2)| = \left| \langle \psi(\cdot, y_2), \operatorname{Tr}_{y_1} \alpha_s \rangle \right| \lesssim \frac{[(0, 1), (y_1, s)]_{8\eta}}{\langle y_2 \rangle^{(d_2 + 8\eta)}}$$

$$\lesssim \frac{2^{-8\eta|s|}}{\langle y_1 \rangle^{(d_1+8\eta)} \langle y_2 \rangle^{(d_2+8\eta)}} \tag{8.5}$$

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

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

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

$$
|\psi *_1 \alpha_s (y_1 + h, y_2) - \psi *_1 \alpha_s (y_1, y_2)|
$$

$$
\lesssim |h|^\eta \left(\sup_{|u_1| \sim |y_1|} \nabla_1 (\psi *_1 \alpha_s)(u_1, y_2) \right)^\eta (|\psi *_1 \alpha_s (y_1 + h, y_2)| + |\psi *_1 \alpha_s (y_1, y_2)|)^{1-\eta}
$$

$$
\lesssim |h|^\eta \frac{2^{|s|[\eta - 8\eta(1-\eta)]}}{\langle y_1 \rangle^{(d_1 + 8\eta)} \langle y_2 \rangle^{(d_2 + 8\eta)}} \leq 2^{-|s|\eta} \frac{|h|^\eta}{\langle y_1 \rangle^{(d_1 + 8\eta)} \langle y_2 \rangle^{(d_2 + 8\eta)}}
$$

$$(8.6)$$

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

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

is immediate from (5.2) and averaging, so that another interpolation with (8.5) yields

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

$$
\leq 2^{-\eta|s|} \frac{|h|^\eta}{\langle y_1 \rangle^{(d_1 + \eta)} \langle y_2 \rangle^{(d_2 + \eta)}}. \quad (8.7)
$$

Collecting (8.5), (8.6), and (8.7), and comparing with (5.2), completes the proof. $\quad\square$

The definitions of α_s and β_s lead to the equalities

$$
\psi_{(x_1,t_1),(x_2,t_2)} = \int_{-\infty}^{\infty} 2^{-\eta|s|} \psi^s_{(x_1,t_1),(x_2,t_2)} \, ds,
$$

$$
\langle f, \psi^s_{(x_1,t_1),(x_2,t_2)} \rangle = \langle f *_1 \beta_{s+\log t_1}, \psi^s_{(x_1,t_1),(x_2,t_2)} \rangle.
$$

Therefore, in view of Lemma 8.5, and using the convergence of the geometric integral, it will be enough to prove the same estimate for the operator

$$
O_s f(x) = \left(\int_{(0,\infty)^2} |\langle f *_1 \beta_{s+\log t_1}, \psi^s_{(x_1,t_1),(x_2,t_2)} \rangle|^2 \frac{dt_1 dt_2}{t_1 t_2} \right)^{\frac{1}{2}}
$$

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

$$
S_{t_1} h(x_1, x_2) := \left(\int_0^{\infty} |\langle h, \psi^s_{(x_1,t_1),(x_2,t_2)} \rangle|^2 \frac{dt_2}{t_2} \right)^{\frac{1}{2}},
$$

which satisfies

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

This can be seen by repeating the sparse domination bound for the Christ–Journé type square function e.g. of [7, 35], where the averages in the sparse operators are associated to rectangles with side of fixed length t_1 in the first parameter. The fact that the weight is a product weight ensures that the bound is uniform over all t_1. The weighted bound above upgrades immediately to vector-valued, and may be used in the second step below to yield

$$\begin{aligned}
\|O_s f\|_{L^p(w)} &= \|\mathsf{S}_{t_1}(f *_1 \beta_{s+\log t_1})\|_{L^p(w;L^2(\frac{dt_1}{t_1}))} \\
&\lesssim [w]_{A_p}^{\max\{\frac{1}{p-1},\frac{1}{2}\}} \|f *_1 \beta_{s+\log t_1}\|_{L^p(w;L^2(\frac{dt_1}{t_1}))} \\
&= [w]_{A_p}^{\max\{\frac{1}{p-1},\frac{1}{2}\}} \|f *_1 \beta_{\log t_1}\|_{L^p(w;L^2(\frac{dt_1}{t_1}))} \lesssim [w]_{A_p}^{\max\{\frac{2}{p-1},1\}} \|f\|_{L^p(w)}.
\end{aligned}$$

The very last inequality is obtained by using the straightforward weighted Littlewood-Paley square function bound of [7, 35] in the first parameter and Fubini's theorem. The proof of Theorem C is complete.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Alpert, B.K.: A class of bases in L^2 for the sparse representation of integral operators. SIAM J. Math. Anal. **24**(1), 246–262 (1993)
2. Astala, K., Iwaniec, T., Saksman, E.: Beltrami operators in the plane. Duke Math. J. **107**(1), 27–56 (2001)
3. Barron, A., Conde-Alonso, J.M., Ou, Y., Rey, G.: Sparse domination and the strong maximal function. Adv. Math. **345**, 1–26 (2019)
4. Barron, A., Pipher, J.: Sparse domination for bi-parameter operators using square functions (2017). arXiv:1709.05009
5. Benea, C., Bernicot, F.: Conservation de certaines propriétés à travers un contrôle épars d'un opérateur et applications au projecteur de Leray-Hopf. Ann. Inst. Fourier (Grenoble) **68**(6), 2329–2379 (2018)
6. Bényi, Á., Torres, R.H.: The discrete Calderón reproducing formula of Frazier and Jawerth. Functional analysis, harmonic analysis, and image processing: a collection of papers in honor of Bjorn Jawerth, pp. 79–107 (2017)
7. Brocchi, G., A sparse quadratic T1 theorem, New York. J. Math. **26**, 1232–1272 (2020).

8. Chang, S.-Y.A., Fefferman, R.: A continuous version of duality of H^1 with BMO on the bidisc. Ann. Math. (2) **112**(1), 179–201 (1980)
9. Conde-Alonso, J.M., Culiuc, A., Di Plinio, F., Ou, Y.: A sparse domination principle for rough singular integrals. Anal. PDE **10**(5), 1255–1284 (2017)
10. Cruz, V., Mateu, J., Orobitg, J.: Beltrami equation with coefficient in Sobolev and Besov spaces. Can. J. Math. **65**(6), 1217–1235 (2013)
11. Cruz, V., Tolsa, X.: Smoothness of the Beurling transform in Lipschitz domains. J. Funct. Anal. **262**(10), 4423–4457 (2012)
12. Cruz-Uribe, D.V., Martell, J.M., Pérez, C.: Weights, Extrapolation and the Theory of Rubio de Francia, Operator Theory: Advances and Applications, vol. 215. Birkhäuser/Springer Basel AG, Basel (2011)
13. Culiuc, A., Di Plinio, F., Ou, Y.: Uniform sparse domination of singular integrals via dyadic shifts. Math. Res. Lett. **25**(1), 21–42 (2018)
14. David, G., Journé, J.-L.: A boundedness criterion for generalized Calderón–Zygmund operators. Ann. Math. (2) **120**(2), 371–397 (1984)
15. Dragičević, O., Volberg, A.: Sharp estimate of the Ahlfors–Beurling operator via averaging martingale transforms. Mich. Math. J. **51**(2), 415–435 (2003)
16. Duoandikoetxea, J.: Extrapolation of weights revisited: new proofs and sharp bounds. J. Funct. Anal. **260**(6), 1886–1901 (2011)
17. Fefferman, R.: A^p weights and singular integrals. Am. J. Math. **110**(5), 975–987 (1988)
18. Fefferman, R., Pipher, J.: Multiparameter operators and sharp weighted inequalities. Am. J. Math. **119**(2), 337–369 (1997)
19. Fefferman, R.: Harmonic analysis on product spaces. Ann. Math. (2) **126**(1), 109–130 (1987)
20. Ferguson, S.H., Lacey, M.T.: A characterization of product BMO by commutators. Acta Math. **189**(2), 143–160 (2002)
21. Ferguson, S.H., Sadosky, C.: Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures. J. Anal. Math. **81**, 239–267 (2000)
22. Figiel, T.: Singular Integral Operators: A Martingale Approach. Geometry of Banach Spaces (Strobl, 1989), pp. 95–110 (1990)
23. Frazier, M., Torres, R., Weiss, G.: The boundedness of Calderón–Zygmund operators on the spaces $\dot{F}_p^{\alpha,q}$. Rev. Mat. Iberoam. **4**(1), 41–72 (1988)
24. Frazier, M., Jawerth, B., Weiss, G.: Littlewood–Paley Theory and the Study of Function Spaces. CBMS Regional Conference Series in Mathematics, vol. 79. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence (1991)
25. Hagelstein, P., Luque, T., Parissis, I.: Tauberian conditions, Muckenhoupt weights, and differentiation properties of weighted bases. Trans. Am. Math. Soc. **367**(11), 7999–8032 (2015)
26. Hart, J., Oliveira, L.: Hardy space estimates for limited ranges of Muckenhoupt weights. Adv. Math. **313**, 803–838 (2017)
27. Holmes, I., Petermichl, S., Wick, B.D.: Weighted little bmo and two-weight inequalities for Journé commutators. Anal. PDE **11**(7), 1693–1740 (2018)
28. Hytönen, T., Lappas, S.: Dyadic representation theorem using smooth wavelets with compact support (2020). arXiv:2003.04019
29. Hytönen, T.P.: The sharp weighted bound for general Calderón–Zygmund operators. Ann. Math. (2) **175**(3), 1473–1506 (2012)
30. Journé, J.-L.: Calderón–Zygmund operators on product spaces. Rev. Mat. Iberoam. **1**(3), 55–91 (1985)
31. Kato, T., Ponce, G.: Commutator estimates and the Euler and Navier–Stokes equations. Commun. Pure Appl. Math. **41**(7), 891–907 (1988)
32. Lacey, M.T.: An elementary proof of the A_2 bound. Isr. J. Math. **217**(1), 181–195 (2017)
33. Lacey, M.T., Petermichl, S., Reguera, M.C.: Sharp A_2 inequality for Haar shift operators. Math. Ann. **348**(1), 127–141 (2010)
34. Lacey, M.T., Sawyer, E.T., Uriarte-Tuero, I.: Astala's conjecture on distortion of Hausdorff measures under quasiconformal maps in the plane. Acta Math. **204**(2), 273–292 (2010)
35. Lerner, A.K.: Sharp weighted norm inequalities for Littlewood–Paley operators and singular integrals. Adv. Math. **226**(5), 3912–3926 (2011)
36. Lerner, A.K.: A simple proof of the A_2 conjecture. Int. Math. Res. Not. IMRN **14**, 3159–3170 (2013)
37. Lerner, A.K.: On pointwise estimates involving sparse operators. N. Y. Math. **22**, 341–349 (2016)

38. Lerner, A.K.: Quantitative weighted estimates for the Littlewood-Paley square function and Marcinkiewicz multipliers. Math. Res. Lett. **26**(2), 537–556 (2019)
39. Li, K., Martikainen, H., Ou, Y., Vuorinen, E.: Bilinear representation theorem. Trans. Am. Math. Soc. **371**(6), 4193–4214 (2019)
40. Li, K., Martikainen, H., Vuorinen, E.: Bilinear Calderón–Zygmund theory on product spaces. Journal de Mathématiques Pures et Appliquées **138**, 356–412 (2020)
41. Martikainen, H.: Representation of bi-parameter singular integrals by dyadic operators. Adv. Math. **229**(3), 1734–1761 (2012)
42. Meyer, Y.: Continuité sur les espaces de Hölder et de Sobolev des opérateurs définis par des intégrales singulières. Recent progress in Fourier analysis (El Escorial, 1983), pp. 145–172 (1985)
43. Meyer, Y., Coifman, R.: Wavelets, Cambridge Studies in Advanced Mathematics, vol. 48, Cambridge University Press, Cambridge (1997) **(Calderón-Zygmund and multilinear operators, Translated from the 1990 and 1991 French originals by David Salinger)**
44. Müller, D., Ricci, F., Stein, E.M.: Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups. I. Invent. Math. **119**(2), 199–233 (1995)
45. Müller, D., Ricci, F., Stein, E.M.: Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups. II. Math. Z. **221**(2), 267–291 (1996)
46. Muscalu, C., Pipher, J., Tao, T., Thiele, C.: Bi-parameter paraproducts. Acta Math. **193**(2), 269–296 (2004)
47. Muscalu, C., Pipher, J., Tao, T., Thiele, C.: Multi-parameter paraproducts. Rev. Mat. Iberoam. **22**(3), 963–976 (2006)
48. Nazarov, F., Treil, S., Volberg, A.: The Bellman functions and two-weight inequalities for Haar multipliers. J. Am. Math. Soc. **12**(4), 909–928 (1999)
49. Ou, Y.: Multi-parameter singular integral operators and representation theorem. Rev. Mat. Iberoam. **33**(1), 325–350 (2017)
50. Petermichl, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. Am. J. Math. **129**(5), 1355–1375 (2007)
51. Petermichl, S., Treil, S., Volberg, A.: Why the Riesz transforms are averages of the dyadic shifts? In: Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000), pp. 209–228 (2002)
52. Petermichl, S., Volberg, A.: Heating of the Ahlfors–Beurling operator: weakly quasiregular maps on the plane are quasiregular. Duke Math. J. **112**(2), 281–305 (2002)
53. Prats, M.: Sobolev regularity of the Beurling transform on planar domains. Publ. Mat. **61**(2), 291–336 (2017)
54. Prats, M.: Sobolev regularity of quasiconformal mappings on domains. J. Anal. Math. **138**(2), 513–562 (2019)
55. Prats, M., Tolsa, X.: A T (P) theorem for Sobolev spaces on domains. J. Funct. Anal. **268**(10), 2946–2989 (2015)
56. Rahm, R., Sawyer E.T., Wick, B.D., Weighted Alpert Wavelets. J. Fourier Anal. Appl. **27**(1), 1 (2021)
57. Tao, T.: Lecture notes for 247b (2007). https://math.ucla.edu/~tao/247b.1.07w/notes7.pdf
58. Torres, R.H.: Boundedness results for operators with singular kernels on distribution spaces. Mem. Am. Math. Soc. **90**(442), viii+172 (1991)
59. Vagharshakyan, A.: Recovering singular integrals from Haar shifts. Proc. Am. Math. Soc. **138**(12), 4303–4309 (2010)
60. Wang, K.: The generalization of paraproducts and the full T1 theorem for Sobolev and Triebel–Lizorkin spaces. J. Math. Anal. Appl. **209**(2), 317–340 (1997)
61. Wilson, J.M.: Weighted norm inequalities for the continuous square function. Trans. Am. Math. Soc. **314**(2), 661–692 (1989)
62. Wilson, M.: Weighted Littlewood–Paley Theory and Exponential-Square Integrability. Lecture Notes in Mathematics, vol. 1924. Springer, Berlin (2008)