ZYGMUND DILATIONS: BILINEAR ANALYSIS AND COMMUTATOR ESTIMATES

EMIL AIRTA, KANGWEI LI, AND HENRI MARTIKAINEN

ABSTRACT. We develop both bilinear theory and commutator estimates in the context of entangled dilations, specifically Zygmund dilations $(x_1, x_2, x_3) \mapsto (\delta_1 x_1, \delta_2 x_2, \delta_1 \delta_2 x_3)$ in $\mathbb{R}^3$. We construct bilinear versions of recent dyadic multiresolution methods for Zygmund dilations and apply them to prove a paraproduct free $T1$ theorem for bilinear singular integrals invariant under Zygmund dilations. Independently, we prove linear commutator estimates even when the underlying singular integrals do not satisfy weighted estimates with Zygmund weights. This requires new paraproduct estimates.

1. INTRODUCTION

"Entangled" systems of dilations, see Nagel-Wainger [22], in the m-parameter product space $\mathbb{R}^d = \prod_{i=1}^m \mathbb{R}^{d_i}$ have the general form

$$(x_1, \ldots, x_m) \mapsto (\delta_1^{\lambda_{11}} \cdots \delta_k^{\lambda_{1k}} x_1, \ldots, \delta_1^{\lambda_{m1}} \cdots \delta_k^{\lambda_{mk}} x_m), \qquad \delta_1, \ldots, \delta_k > 0,$$

and appear naturally throughout analysis. For instance, in $\mathbb{R}^3$ the *Zygmund dilations* $(x_1, x_2, x_3) \mapsto (\delta_1 x_1, \delta_2 x_2, \delta_1 \delta_2 x_3)$ are compatible with the group law of the Heisenberg group, see e.g. Müller–Ricci–Stein [21]. Even these simplest entangled dilations are not completely understood, especially when it comes to the associated Calderón–Zygmund type singular integral operators (SIOs).

Until recently, multiresolution methods were still missing in the Zygmund dilations setting, as pointed out in [5]. This was a big restriction on how to go about developing singular integral theory. However, the last two authors together with T. Hytönen and E. Vuorinen recently developed this missing Zygmund multiresolution analysis in [14]. Such dyadic representation theorems and related multiresolution techniques had been highly influential in recent advances on SIOs and their applications (see e.g. [12, 13, 20, 23]), but developing them in the entangled situation required new ideas. These tools then yielded very delicate weighted norm inequalities $L^p(w) \to L^p(w)$ for general non-convolution form Zygmund singular integrals in the optimal generality of Zygmund

weights (introduced by Fefferman–Pipher [6])

$$[w]_{A_{p,Z}} := \sup_{I \in \mathcal{R}_Z} \Big(\frac{1}{|I|} \int_I w(x) \, \mathrm{d}x \Big) \Big(\frac{1}{|I|} \int_I w^{-1/(p-1)}(x) \, \mathrm{d}x \Big)^{p-1} < \infty, \qquad 1 < p < \infty,$$

where the supremum is over *Zygmund rectangles* $I = I_1 \times I_2 \times I_3$, $\ell(I_3) = \ell(I_1)\ell(I_2)$.

In fact, there is a precise threshold: if the kernel decay in terms of the deviation of $z \in \mathbb{R}^3$ from the "Zygmund manifold" $|z_1 z_2| = |z_3|$ is not fast enough, singular integrals invariant under Zygmund dilations fail to be bounded with Zygmund weights. We constructed counterexamples and showed the delicate positive result in the optimal range using the new multiresolution analysis. Previous results include [5, 6, 11, 24].

This rather striking threshold for weighted estimates means that it is, in particular, unclear in what generality natural estimates for commutators $[b, T] = bT - T(b \cdot)$ hold. Of course, ever since the classical one-parameter result of Coifman–Rochberg–Weiss [2], stating that $\|[b, T]\|_{L^p \to L^p} \sim \|b\|_{\mathrm{BMO}}$, commutator estimates have been a large and fundamental part of the theory of SIOs and their applications. Commutator estimates in the Zygmund dilation setting were previously considered in [5] using the so-called Cauchy integral trick. That method requires weighted bounds with Zygmund weights – this is because it uses the fact that natural Zygmund adapted BMO functions generate Zygmund weights. But we now know [14] that such weighted bounds are quite delicate – and it turns out that the commutator bounds are true even in the regime where weighted estimates fail. We prove the following.

1.1. **Theorem.** *Let* $b \in L^1_{loc}$ *and* T *be a linear paraproduct free Calderón-Zygmund operator adapted to Zygmund dilations as in* [14]. *Let* $\theta \in (0, 1]$ *be the kernel exponent measuring the decay in terms of the Zygmund ratio*

$$D_\theta(x) := \left(\frac{|x_1 x_2|}{|x_3|} + \frac{|x_3|}{|x_1 x_2|} \right)^{-\theta}.$$

Then for all such θ *we have*

$$\|[b, T]\|_{L^p \to L^p} \lesssim \|b\|_{\mathrm{bmo}_Z}, \qquad 1 < p < \infty.$$

As weighted estimates only hold with $\theta = 1$, this requires a proof based on the multiresolution decomposition [14] and a new family of "Zygmund paraproducts". Studying paraproducts is also interesting from the technical viewpoint that, generally, proofs of $T1$ theorems display a structural decomposition of SIOs into their cancellative parts and paraproducts. The new Zygmund theory in [14] is designed for the fully cancellative case leaving out paraproducts and BMO considerations, so this is the first paper, as far as we know, where paraproducts are considered in the Zygmund situation. They are tricky objects in the entangled situation. However, while this is also a step forward towards a full $T1$ theorem in the Zygmund setting, the commutator theory that we develop does not require so-called partial paraproducts, and so the paraproduct tools developed here are not yet sufficient to prove a $T1$ theorem in the non-cancellative case. We also mention that during our proof we include some results of independent interest, mainly, a new, extremely short proof of the A_∞ extrapolation theorem [3].

Moving to a different direction, we push the Zygmund multiresolution methods [14] to the multilinear setting and study bilinear SIOs invariant under Zygmund dilations. A

classical model of an n-linear SIO T in $\mathbb{R}^d$ is obtained by setting

$$T(f_1, \ldots, f_n)(x) = U(f_1 \otimes \cdots \otimes f_n)(x, \ldots, x), \qquad x \in \mathbb{R}^d, \ f_i \colon \mathbb{R}^d \to \mathbb{C},$$

where U is a linear SIO in $\mathbb{R}^{nd}$. See e.g. Grafakos–Torres [9] for the basic theory. Estimates for classical multilinear SIOs play a fundamental role in pure and applied analysis – for example, L^p estimates for the homogeneous fractional derivative $D^\alpha f = \mathcal{F}^{-1}(|\xi|^\alpha \widehat{f}(\xi))$ of a product of two or more functions, the fractional Leibniz rules, are used in the area of dispersive equations, see e.g. Kato–Ponce [15] and Grafakos–Oh [8]. We do not otherwise attempt to summarize the massive body of literature here and simply mention that the closest existing result is perhaps [18], which develops multiresolution methods in the non-entangled multilinear bi-parameter case.

In this paper we prove the following "paraproduct free" $T1$ theorem for bilinear Zygmund SIOs.

1.2. **Theorem.** *Let T be a bilinear paraproduct free Calderón-Zygmund operator adapted to Zygmund dilations as in Definition 3.5. Let $1 < p_1, p_2 < \infty$ and $\frac{1}{2} < p < \infty$ with $\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2}$. Then we have*

$$\|T(f_1, f_2)\|_{L^p} \lesssim \|f_1\|_{L^{p_1}} \|f_2\|_{L^{p_2}}.$$

Notice that we can conclude the full bilinear range, including the quasi-Banach range, just from the paraproduct free $T1$ type assumptions. Also relevant is the fact that e.g. the appearing weak boundedness condition only involves Zygmund rectangles – that is, the $T1$ assumptions of Definition 3.5 are Zygmund adapted and in this respect weaker than the corresponding tri-parameter assumptions.

It would also be very interesting to develop weighted theory with suitable kernel assumptions like in the linear case [14]. That is, to generalize our recent paper [19] from the standard multi-parameter setting to this entangled Zygmund setting. Recall that it would be key to deal with "genuine" multilinear weights, i.e., only impose a *joint A_p* condition on the associated tuple of weights $\vec{w} = (w_1, \ldots, w_n)$. While such multilinear weighted estimates had been known for one-parameter SIOs for over 10 years by the influential paper [16], the multi-parameter version was only recently solved in [19]. The entangled situation is very difficult, though, and we do not achieve such estimates in this paper. Indeed, we are splitting our operators in a way that is sufficient for the unbounded estimates, but not for the weighted estimates. In fact, already the unweighted estimates are surprisingly delicate and the only way we found to achieve them was with using this additional decomposition and even some sparse domination tools.

Here is an outline of the paper. In Section 2 we develop the fundamental Zygmund adapted multiresolution methods in the bilinear setting. Section 3 introduces the singular integrals and the corresponding testing conditions, and Section 4 uses the kernel estimates to bound the various coefficients arising in the multiresolution analysis. Section 5 contains a further decomposition of our dyadic model operators – this is then required in Section 6, where the L^p estimates of these model operators are proved. Section 6 concludes with the proof of Theorem 1.2. Section 7 contains the proof of the linear commutator estimates, Theorem 1.1, and the corresponding theory of product and little BMO commutators in the Zygmund setting. Appendix A considers bilinear variants of the multipliers studied by Fefferman-Pipher [6] – this is motivation for the abstract definitions of Section 3.

2. BILINEAR ZYGMUND MULTIRESOLUTION ANALYSIS

2.A. **Dyadic intervals, Zygmund rectangles and basic randomization.** Given a dyadic grid $\mathcal{D}$, $I \in \mathcal{D}$ and $k \in \mathbb{Z}$, $k \geq 0$, we use the following notation:

(1) $\ell(I)$ is the side length of I.
(2) $I^{(k)} \in \mathcal{D}$ is the kth parent of I, i.e., $I \subset I^{(k)}$ and $\ell(I^{(k)}) = 2^k \ell(I)$.
(3) $\mathrm{ch}(I)$ is the collection of the children of I, i.e., $\mathrm{ch}(I) = \{J \in \mathcal{D}: J^{(1)} = I\}$.
(4) $E_I f = \langle f \rangle_I 1_I$ is the averaging operator, where $\langle f \rangle_I = \fint_I f = \frac{1}{|I|} \int_I f$.
(5) $\Delta_I f$ is the martingale difference $\Delta_I f = \sum_{J \in \mathrm{ch}(I)} E_J f - E_I f$.
(6) $\Delta_{I,k} f$ or $\Delta_I^k f$ is the martingale difference block

$$\Delta_{I,k} f = \Delta_I^k f = \sum_{\substack{J \in \mathcal{D} \\ J^{(k)} = I}} \Delta_J f.$$

We will have use for randomization soon. While often the grids are fixed and we suppress the dependence on the random parameters, it will be important to understand the definitions underneath. So we go ahead and introduce the related notation and standard results now. Let $\mathcal{D}_0$ be the standard dyadic grid in $\mathbb{R}$. For $\omega \in \{0,1\}^{\mathbb{Z}}$, $\omega = (\omega_i)_{i \in \mathbb{Z}}$, we define the shifted lattice

$$\mathcal{D}(\omega) := \Big\{L + \omega := L + \sum_{i:\, 2^{-i} < \ell(L)} 2^{-i} \omega_i \colon L \in \mathcal{D}_0\Big\}.$$

Let $\mathbb{P}_\omega$ be the product probability measure on $\{0,1\}^{\mathbb{Z}}$. We recall the following notion of a good interval from [10]. We say that $G \in \mathcal{D}(\omega, k)$, $k \geq 2$, if $G \in \mathcal{D}(\omega)$ and

(2.1) $$d(G, \partial G^{(k)}) \geq \frac{\ell(G^{(k)})}{4} = 2^{k-2} \ell(G).$$

Notice that for all $L \in \mathcal{D}_0$ and $k \geq 2$ we have

(2.2) $$\mathbb{P}_\omega(\{\omega \colon L + \omega \in \mathcal{D}(\omega, k)\}) = \frac{1}{2}.$$

The key implication (of practical use later) of $G \in \mathcal{D}(\omega, k)$ is that for $n \in \mathbb{Z}$ with $|n| \leq 2^{k-2}$ we have

(2.3) $$(G \dotplus n)^{(k)} = G^{(k)}, \qquad G \dotplus n := G + n\ell(G).$$

In fact, we will not need much more of randomization – it only remains to move the notation to our actual setting of $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R}^2$. We define for

$$\sigma = (\sigma^1, \sigma^2, \sigma^3) \in \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}} \times \{0,1\}^{\mathbb{Z}}$$

that

$$\mathcal{D}(\sigma) := \mathcal{D}(\sigma^1) \times \mathcal{D}(\sigma^2) \times \mathcal{D}(\sigma^3).$$

Let

$$\mathbb{P}_\sigma := \mathbb{P}_{\sigma^1} \times \mathbb{P}_{\sigma^2} \times \mathbb{P}_{\sigma^3}.$$

For $k = (k^1, k^2, k^3)$, $k^1, k^2, k^3 \geq 2$, we define

$$\mathcal{D}(\sigma, k) = \mathcal{D}(\sigma^1, k^1) \times \mathcal{D}(\sigma^2, k^2) \times \mathcal{D}(\sigma^3, k^3).$$

We also e.g. write

$$\mathcal{D}(\sigma, (k^1, 0, k^3)) = \mathcal{D}(\sigma^1, k^1) \times \mathcal{D}(\sigma^2) \times \mathcal{D}(\sigma^3, k^3),$$

that is, a 0 will designate that we do not have goodness in that parameter.

As for most of the argument σ is fixed, it makes sense to mainly suppress it from the notation and abbreviate, whenever possible, that

$$\mathcal{D}^m = \mathcal{D}(\sigma^m), \qquad \mathcal{D}(\sigma^m, k^m) = \mathcal{D}^m(k^m), \qquad m = 1, 2, 3.$$

Then also

$$\mathcal{D} = \mathcal{D}(\sigma) = \prod_{m=1}^{3} \mathcal{D}^m, \qquad \mathcal{D}(k) = \prod_{m=1}^{3} \mathcal{D}^m(k^m).$$

We define the Zygmund rectangles $\mathcal{D}_Z \subset \mathcal{D}$ by setting

(2.4) $$\mathcal{D}_Z = \Big\{ I = \prod_{m=1}^{3} I^m \in \mathcal{D} \colon \ell(I^1)\ell(I^2) = \ell(I^3) \Big\}.$$

Obviously, $\mathcal{D}_Z(k)$ is defined similarly as above but also requires $\prod_{m=1}^{3} I^m \in \mathcal{D}(k)$.

2.B. **Zygmund martingale differences.** Given $I = \prod_{m=1}^{3} I^m$ we define the Zygmund martingale difference operator

$$\Delta_{I,Z} f := \Delta_{I^1} \Delta_{I^2 \times I^3} f.$$

2.5. *Remark.* We highlight that the martingale difference $\Delta_{I^2 \times I^3}$ is the **one-parameter** (and not the bi-parameter) martingale difference on the rectangle $I^2 \times I^3$:

$$\Delta_{I^2 \times I^3} = \Delta_{I^2} \Delta_{I^3} + E_{I^2} \Delta_{I^3} + \Delta_{I^2} E_{I^3} \neq \Delta_{I^2} \Delta_{I^3}.$$

Moreover, the above operators really act on the full product space but only on the given parameters – for instance, $\Delta_{I^1} f(x_1, x_2, x_3) = \Delta_{I^1}^1 f(x_1, x_2, x_3) = (\Delta_{I^1} f(\cdot, x_2, x_3))(x_1)$.

We recall the following facts from [14]. For a dyadic $\lambda > 0$ define the dilated lattices

$$\mathcal{D}_\lambda^{2,3} = \{ I^{2,3} \in \mathcal{D}^{2,3} := \mathcal{D}^2 \times \mathcal{D}^3 \colon \ell(I^3) = \lambda \ell(I^2) \}.$$

The basic Zygmund expansion goes as follows:

(2.6) $$f = \sum_{I^1 \in \mathcal{D}^1} \Delta_{I^1} f = \sum_{I^1 \in \mathcal{D}^1} \sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}^{2,3}} \Delta_{I^1} \Delta_{I^{2,3}} f = \sum_{I \in \mathcal{D}_Z} \Delta_{I,Z} f.$$

However, the way we split our operators will not be this simple.

The following basic results hold for the martingale differences. For $I, J \in \mathcal{D}_Z$ we have

$$\Delta_{I,Z} \Delta_{J,Z} f = \begin{cases} \Delta_{I,Z} & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases}$$

Notice also that the Zygmund martingale differences satisfy

$$\int_{\mathbb{R}} \Delta_{I,Z} f \, \mathrm{d}x_1 = 0 \qquad \text{and} \qquad \int_{\mathbb{R}^2} \Delta_{I,Z} f \, \mathrm{d}x_2 \, \mathrm{d}x_3 = 0.$$

Moreover, we have

$$\int (\Delta_{I,Z} f) g = \int f \Delta_{I,Z} g.$$

2.C. **Haar functions.** For an interval $J \subset \mathbb{R}$ we denote by J_l and J_r the left and right halves of J, respectively. We define

$$h_J^0 = |J|^{-1/2} 1_J \qquad \text{and} \qquad h_J^1 = h_J = |J|^{-1/2}(1_{J_l} - 1_{J_r}).$$

The reader should carefully notice that h_I^0 is the non-cancellative Haar function for us and that in some other papers a different convention is used.

As we mostly work on $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R}^2$ we require some Haar functions on $\mathbb{R}^2$ as well. For $I^2 \times I^3 \subset \mathbb{R}^2$ and $\eta = (\eta_2, \eta_3) \in \{0, 1\}^2$ define

$$h_{I^2 \times I^3}^\eta = h_{I^2}^{\eta_2} \otimes h_{I^3}^{\eta_3}.$$

Similarly, as h_{I^1} denotes a cancellative Haar function on $\mathbb{R}$, we let $h_{I^2 \times I^3}$ denote a cancellative **one-parameter** Haar function on $I^2 \times I^3$. This means that

$$h_{I^2 \times I^3} = h_{I^2 \times I^3}^\eta$$

for some $\eta = (\eta_2, \eta_3) \in \{0, 1\}^2 \setminus \{(0, 0)\}$. We only use a 0 to denote a non-cancellative Haar function: $h_{I^2 \times I^3}^0 = h_{I^2 \times I^3}^{(0,0)}$.

We suppress this η dependence in all that follows in the sense that a finite η summation is not written. For example, given $I = I^1 \times I^2 \times I^3 \in \mathcal{D}_Z \subset \prod_{m=1}^3 \mathcal{D}^m$ decompose

$$\Delta_{I,Z} f = \Delta_{I^1} \Delta_{I^2 \times I^3} f = \langle f, h_{I^1} \otimes h_{I^2 \times I^3} \rangle h_{I^1} \otimes h_{I^2 \times I^3} =: \langle f, h_{I,Z} \rangle h_{I,Z}.$$

2.D. **Bilinear Zygmund shifts.** In preparation for defining the shifts, we define the following notation. Let I_1, I_2, I_3 be rectangles, $I_j = I_j^1 \times I_j^2 \times I_j^3 = I_j^1 \times I_j^{2,3}$, and f_1, f_2, f_3 be functions defined on $\mathbb{R}^3$. For $j_1, j_2 \in \{1, 2, 3\}$ define

$$A_{I_1,I_2,I_3}^{j_1,j_2} = A_{I_1,I_2,I_3}^{j_1,j_2}(f_1, f_2, f_3) := \prod_{j=1}^3 \langle f_j, v_{I_j} \rangle,$$

where

$$v_{I_j} = \widetilde{h}_{I_j^1} \otimes \widetilde{h}_{I_j^{2,3}};$$

$$\widetilde{h}_{I_{j_1}^1} = h_{I_{j_1}^1} \qquad \text{and} \qquad \widetilde{h}_{I_j^1} = h_{I_j^1}^0, \ j \neq j_1;$$

$$\widetilde{h}_{I_{j_2}^{2,3}} = h_{I_{j_2}^{2,3}} \qquad \text{and} \qquad \widetilde{h}_{I_j^{2,3}} = h_{I_j^{2,3}}^0, \ j \neq j_2.$$

For a dyadic $\lambda > 0$ define

$$\mathcal{D}_\lambda = \{K = K^1 \times K^2 \times K^3 \in \mathcal{D} \colon \lambda \ell(K^1) \ell(K^2) = \ell(K^3)\}.$$

Moreover, for a rectangle $I = I^1 \times I^2 \times I^3$ and $k = (k^1, k^2, k^3)$ define

$$I^{(k)} = I_1^{(k^1)} \times I_2^{(k^2)} \times I_3^{(k^3)}.$$

2.7. Definition. Let $k = (k^1, k^2, k^3)$, $k^i \in \{0, 1, 2, \dots\}$, be fixed. A bilinear Zygmund shift $Q = Q_k$ of complexity k has the form

$$\langle Q_k(f_1, f_2), f_3 \rangle$$

$$= \sum_{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}} \sum_{\substack{I_1,I_2,I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} a_{K,(I_j)} \Big[A_{I_1,I_2,I_3}^{j_1,j_2} - A_{I_{j_1}^1 \times I_1^{2,3}, I_{j_1}^1 \times I_2^{2,3}, I_{j_1}^1 \times I_3^{2,3}}^{j_1,j_2}$$

$$- A^{j_1,j_2}_{I_1^1 \times I_{j_2}^{2,3}, I_2^1 \times I_{j_2}^{2,3}, I_3^1 \times I_{j_2}^{2,3}} + A^{j_1,j_2}_{I_{j_1}^1 \times I_{j_2}^{2,3}, I_{j_1}^1 \times I_{j_2}^{2,3}, I_{j_1}^1 \times I_{j_2}^{2,3}} \Big]$$

for some $j_1, j_2 \in \{1, 2, 3\}$. The coefficients $a_{K,(I_j)}$ satisfy

$$|a_{K,(I_j)}| \leq \frac{|I_1|^{1/2}|I_2|^{1/2}|I_3|^{1/2}}{|K|^2} = \frac{|I_1|^{3/2}}{|K|^2}.$$

Now, the game is to represent bilinear singular integrals using the operators Q_k and also – independently – bound the operators Q_k suitably. We start with the representation part and deal with bounding the operators later. We have not defined our singular integrals carefully yet, however, a lot of the required decomposition can be formally carried out for an arbitrary operator T. The singular integral part is later required to get sufficient decay for the appearing scalar coefficients and to handle the paraproducts.

2.E. **Zygmund decomposition of** $\langle T(f_1, f_2), f_3 \rangle$**.** For now, we focus on the multiresolution part and start formally decomposing a general bilinear operator. We begin by writing $\langle T(f_1, f_2), f_3 \rangle$ as

$$\sum_{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$= \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1), \ell(I_2^1) > \ell(I_3^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1), \ell(I_3^1) > \ell(I_2^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_2^1), \ell(I_3^1) > \ell(I_1^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1) > \ell(I_2^1) = \ell(I_3^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_2^1) > \ell(I_1^1) = \ell(I_3^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_3^1) > \ell(I_1^1) = \ell(I_2^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$$

$$+ \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1) = \ell(I_2^1) = \ell(I_3^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle.$$

We collapse the first six sums, which are not already diagonal sums, into diagonal sums

$$\sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1) = \ell(I_2^1) = \ell(I_3^1)}} .$$

This has the effect that whenever we have an inequality $\ell(I_i^1) > \ell(I_j^1)$, the martingale difference operator $\Delta_{I_i^1}$ corresponding with the larger cube is changed to the averaging operator $E_{I_i^1}$. Thus, in the first three sums we now have two averaging operators, and in the next three we have one averaging operator. The more averaging operators we have, the less cancellation we have, and thus the main challenge are the first three sums with the least cancellation. We mainly focus on the first three sums for this reason.

In addition, the first three sums are symmetric, so we may focus on only one of them, and choose to look at

$$\sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1), \ell(I_2^1) > \ell(I_3^1)}} \langle T(\Delta_{I_1^1} f_1, \Delta_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle = \sum_{\substack{I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1 \\ \ell(I_1^1) = \ell(I_2^1) = \ell(I_3^1)}} \langle T(E_{I_1^1} f_1, E_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle.$$

Now, we fix $I_1^1, I_2^1, I_3^1 \in \mathcal{D}^1$ with $\ell(I_1^1) = \ell(I_2^1) = \ell(I_3^1)$ and repeat the argument for $\langle T(E_{I_1^1} f_1, E_{I_2^1} f_2), \Delta_{I_3^1} f_3 \rangle$ using the lattice $\mathcal{D}_{\ell(I^1)}^{2,3}$, where recall that for a dyadic $\lambda > 0$ we have

$$\mathcal{D}_\lambda^{2,3} = \{I^2 \times I^3 \in \mathcal{D}^{2,3} := \mathcal{D}^2 \times \mathcal{D}^3 \colon \ell(I^3) = \lambda\ell(I^2)\}.$$

This produces seven terms, and we again focus on

$$\sum_{\substack{I_1^2 \times I_1^3, I_2^2 \times I_2^3, I_3^2 \times I_3^3 \in \mathcal{D}_{\ell(I^1)}^{2,3} \\ \ell(I_1^2) = \ell(I_2^2) = \ell(I_3^2)}} \langle T(E_{I_1^1} E_{I_1^2 \times I_1^3} f_1, E_{I_2^1} E_{I_2^2 \times I_2^3} f_2), \Delta_{I_3^1} \Delta_{I_3^2 \times I_3^3} f_3 \rangle.$$

Altogether, our focus, for now, is on the key term

(2.8)
$$\sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ \ell(I_1) = \ell(I_2) = \ell(I_3)}} \langle T(E_{I_1} f_1, E_{I_2} f_2), \Delta_{I_3, Z} f_3 \rangle,$$

where $\ell(I_1) = \ell(I_2) = \ell(I_3)$ means that

$$\ell(I_1^m) = \ell(I_2^m) = \ell(I_3^m), \qquad m = 1, 2, 3.$$

This was completely generic – we now go a step further to the direction of Zygmund shifts and start introducing Haar functions into the mix.

2.F. **Further decomposition of** (2.8). Write

$$\langle T(E_{I_1} f_1, E_{I_2} f_2), \Delta_{I_3, Z} f_3 \rangle = \langle T(h_{I_1}^0, h_{I_2}^0), h_{I_3, Z} \rangle \langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle \langle f_3, h_{I_3, Z} \rangle.$$

Now, we perform a rather complicated decomposition of the product $\langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle$. To this end, start by writing

$$\langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle$$
$$= \left[\langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle - \langle f_1, h_{I_3^1}^0 h_{I_1^{2,3}}^0 \rangle \langle f_2, h_{I_3^1}^0 h_{I_2^{2,3}}^0 \rangle \right] + \langle f_1, h_{I_3^1}^0 h_{I_1^{2,3}}^0 \rangle \langle f_2, h_{I_3^1}^0 h_{I_2^{2,3}}^0 \rangle$$
$$=: A_1 + A_2.$$

We then further decompose A_1 as follows

$$A_1 = \Big[\langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle - \langle f_1, h_{I_3^1}^0 h_{I_1^{2,3}}^0 \rangle \langle f_2, h_{I_3^1}^0 h_{I_2^{2,3}}^0 \rangle$$
$$- \langle f_1, h_{I_1^1}^0 h_{I_3^{2,3}}^0 \rangle \langle f_2, h_{I_2^1}^0 h_{I_3^{2,3}}^0 \rangle + \langle f_1, h_{I_3^1}^0 h_{I_3^{2,3}}^0 \rangle \langle f_2, h_{I_3^1}^0 h_{I_3^{2,3}}^0 \rangle \Big]$$

$$+ \left\{ \langle f_1, h^0_{I^1_1} h^0_{I_{2,3}} \rangle \langle f_2, h^0_{I^1_2} h^0_{I_{2,3}} \rangle - \langle f_1, h^0_{I^1_3} h^0_{I_{2,3}} \rangle \langle f_2, h^0_{I^1_3} h^0_{I_{2,3}} \rangle \right\}.$$

When we later specialize to singular integrals, we will in particular make the following assumption. We say that T is a paraproduct free operator, if for all cancellative Haar functions h_{I^1} and $h_{I^{2,3}}$ we have

$$\langle T(1 \otimes 1_{J^{2,3}_1}, 1 \otimes 1_{J^{2,3}_2}), h_{I^1} \otimes 1_{J^{2,3}_3} \rangle = \langle T^{*,j}_1(1 \otimes 1_{J^{2,3}_1}, 1 \otimes 1_{J^{2,3}_2}), h_{I^1} \otimes 1_{J^{2,3}_3} \rangle$$

$$= \langle T(1_{I^1_1} \otimes 1, 1_{I^1_2} \otimes 1), 1_{I^1_3} \otimes h_{I^{2,3}} \rangle = \langle T^{*,j}_{2,3}(1_{I^1_1} \otimes 1, 1_{I^1_2} \otimes 1), 1_{I^1_3} \otimes h_{I^{2,3}} \rangle = 0$$

for all the adjoints $j \in \{1, 2\}$. With this assumption in the full summation (2.8) everything else vanishes except

$$\sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ \ell(I_1) = \ell(I_2) = \ell(I_3)}} \langle T(h^0_{I_1}, h^0_{I_2}), h_{I_3, Z} \rangle \Big[\langle f_1, h^0_{I_1} \rangle \langle f_2, h^0_{I_2} \rangle - \langle f_1, h^0_{I^1_3 \times I^{2,3}_1} \rangle \langle f_2, h^0_{I^1_3 \times I^{2,3}_2} \rangle$$

$$- \langle f_1, h^0_{I^1_1 \times I^{2,3}_3} \rangle \langle f_2, h^0_{I^1_2 \times I^{2,3}_3} \rangle + \langle f_1, h^0_{I_3} \rangle \langle f_2, h^0_{I_3} \rangle \Big] \langle f_3, h_{I_3, Z} \rangle.$$

So we eliminated the paraproducts by assumption, and now we have to manipulate this remaining term to a suitable form involving shifts.

In the above sum we will relabel $I_3 = I = I^1 \times I^2 \times I^3 = I^1 \times I^{2,3}$. Then, for $n_1 = (n^1_1, n^2_1, n^3_1) = (n^1_1, n^{2,3}_1)$ we write

$$I_1 = I \dotplus n_1 = (I^1 + n^1_1 \ell(I^1)) \times (I^2 + n^2_1 \ell(I^2)) \times (I^3 + n^3_1 \ell(I^3)) = (I^1 \dotplus n^1_1) \times (I^{2,3} \dotplus n^{2,3}_1).$$

We write I_2 similarly as $I_2 = I \dotplus n_2$. Notice that if $n^1_1 = n^1_2 = 0$, then the term inside the summation vanishes. Similarly, if $n^{2,3}_1 = n^{2,3}_2 = (0, 0)$, the term inside the summation vanishes. So we need to study

$$\sum_{\substack{n_1, n_2 \in \mathbb{Z}^3 \\ \max(|n^1_1|, |n^1_2|) \neq 0 \\ \max(|n^2_1|, |n^2_2|) \neq 0 \text{ or } \max(|n^3_1|, |n^3_2|) \neq 0}} \sum_{I \in \mathcal{D}_Z} c_{I, n_1, n_2},$$

where

$$c_{I, n_1, n_2}$$

$$= \langle T(h^0_{I \dotplus n_1}, h^0_{I \dotplus n_2}), h_{I, Z} \rangle \Big[\langle f_1, h^0_{I \dotplus n_1} \rangle \langle f_2, h^0_{I \dotplus n_2} \rangle - \langle f_1, h^0_{I^1 \times (I^{2,3} \dotplus n^{2,3}_1)} \rangle \langle f_2, h^0_{I^1 \times (I^{2,3} \dotplus n^{2,3}_2)} \rangle$$

$$- \langle f_1, h^0_{(I^1 \dotplus n^1_1) \times I^{2,3}} \rangle \langle f_2, h^0_{(I^1 \dotplus n^1_2) \times I^{2,3}} \rangle + \langle f_1, h^0_I \rangle \langle f_2, h^0_I \rangle \Big] \langle f_3, h_{I, Z} \rangle.$$

We write

$$\sum_{\substack{n_1, n_2 \in \mathbb{Z}^3 \\ \max_{j=1,2} |n^1_j| \neq 0 \\ \max_{j=1,2} |n^2_j| \neq 0 \text{ or } \max_{j=1,2} |n^3_j| \neq 0}} \sum_{I \in \mathcal{D}_Z} c_{I, n_1, n_2}$$

$$= \sum_{k^1, k^2, k^3 = 2}^{\infty} \sum_{\substack{n_1, n_2 \in \mathbb{Z}^3 \\ \max_{j=1,2} |n^m_j| \in (2^{k^m - 3}, 2^{k^m - 2}] \\ m=1,2,3}} \sum_{I \in \mathcal{D}_Z} c_{I, n_1, n_2}$$

$$+ \sum_{\substack{k^1,k^2=2 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2 \\ n_1^3 = n_2^3 = 0}}^{\infty} \sum_{n_1,n_2 \in \mathbb{Z}^3} \sum_{I \in \mathcal{D}_Z} c_{I,n_1,n_2}$$

$$+ \Sigma_{sym},$$

where Σ_{sym} is symmetric to the second term and has $n_1^2 = n_2^2 = 0$.

Recall how everything implicitly depends on the random parameter σ, so that we can average over it. By independence, like in [14], we have by (2.2) that

(2.9)
$$\mathbb{E}_\sigma \sum_{\substack{k^1,k^2,k^3=2 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2,3}}^{\infty} \sum_{n_1,n_2 \in \mathbb{Z}^3} \sum_{I \in \mathcal{D}_Z} c_{I,n_1,n_2}$$

$$= 8\mathbb{E}_\sigma \sum_{\substack{k^1,k^2,k^3=2 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2,3}}^{\infty} \sum_{n_1,n_2 \in \mathbb{Z}^3} \sum_{I \in \mathcal{D}_Z(k)} c_{I,n_1,n_2}, \qquad k = (k^1, k^2, k^3).$$

For the other two terms, where $n_j^2 = 0$ or $n_j^3 = 0$, we perform the above but do not add goodness to the second and third parameters, respectively. For example, we have

$$\mathbb{E}_\sigma \sum_{\substack{k^1,k^2=2 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2 \\ n_1^3 = n_2^3 = 0}}^{\infty} \sum_{n_1,n_2 \in \mathbb{Z}^3} \sum_{I \in \mathcal{D}_Z} c_{I,n_1,n_2}$$

$$= 4\mathbb{E}_\sigma \sum_{\substack{k^1,k^2=2 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2 \\ n_1^3 = n_2^3 = 0}}^{\infty} \sum_{n_1,n_2 \in \mathbb{Z}^3} \sum_{I \in \mathcal{D}_Z(k^1,k^2,0)} c_{I,n_1,n_2}.$$

Continuing with (2.9), we write it as

$$C 8\mathbb{E}_\sigma \sum_{k^1,k^2,k^3=2}^{\infty} (|k|+1)^2 \varphi(k)$$

$$\sum_{\substack{K \in \mathcal{D}_\lambda \\ I^{(k)}=K}} \sum_{\substack{I \in \mathcal{D}_Z(k) }} \sum_{\substack{n_1,n_2 \in \mathbb{Z}^3 \\ \max_{j=1,2} |n_j^m| \in (2^{k^m-3}, 2^{k^m-2}] \\ m=1,2,3}} \frac{c_{I,n_1,n_2}}{C(|k|+1)^2 \varphi(k)},$$

where

$$\mathcal{D}_\lambda = \{ K = K^1 \times K^2 \times K^3 \in \mathcal{D} : \lambda \ell(K^1)\ell(K^2) = \ell(K^3) \}, \qquad \lambda = 2^{k^3-k^1-k^2},$$

and C is some suitably large constant depending on T. Recall that by (2.3) we also have

$$(I \dotplus n_1)^{(k)} = (I \dotplus n_2)^{(k)} = I^{(k)} = K.$$

We have arrived to a point where we cannot go further without talking about singular integrals. Indeed, we need kernel estimates to control the coefficients. But on a structural level (with the paraproduct free assumption), we have obtained a reasonable representation of the main term (2.8) in terms of sums of bilinear Zygmund shifts.

3. BILINEAR ZYGMUND SINGULAR INTEGRALS

We begin by defining the required kernel estimates and cancellation conditions for bilinear singular integrals T invariant under Zygmund dilations. For motivation for the form of the kernel estimates, see Appendix A for kernel bounds of bilinear multipliers. This viewpoint makes the kernel estimates natural – on the other hand, they are also of the right form so that we will be able to bound the coefficients from the multiresolution decomposition and obtain reasonable decay.

3.A. **Full kernel representation.** Our bilinear singular integral T invariant under Zygmund dilations is related to a full kernel K in the following way. The kernel K is a function

$$K \colon (\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3) \setminus \Delta \to \mathbb{C},$$

where

$$\Delta = \{(x, y, z) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \colon x_i = y_i = z_i \text{ for at least one } i = 1, 2, 3\}.$$

We look at the action of T on rectangles like $I^1 \times I^2 \times I^3 =: I^1 \times I^{2,3}$ in $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R} \times \mathbb{R}^2$. So let $I_i = I_i^1 \times I_i^2 \times I_i^3$ be rectangles, $i = 1, 2, 3$. Assume that there exists $i_1, i_2, j_1, j_2 \in \{1, 2, 3\}$ so that $I_{i_1}^1$ and $I_{i_2}^1$ are disjoint and also $I_{j_1}^{2,3}$ and $I_{j_2}^{2,3}$ are disjoint. Then we have the full kernel representation

$$\langle T(1_{I_1}, 1_{I_2}), 1_{I_3} \rangle = \iiint K(x, y, z) 1_{I_1}(x) 1_{I_2}(y) 1_{I_3}(z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

The kernel K satisfies the following estimates.

First, we define the decay factor

$$D_\theta(x, y) = \left(\frac{\prod_{i=1}^2 (|x_i| + |y_i|)}{|x_3| + |y_3|} + \frac{|x_3| + |y_3|}{\prod_{i=1}^2 (|x_i| + |y_i|)} \right)^{-\theta}, \qquad \theta \in (0, 2],$$

and the tri-parameter bilinear size factor

$$S(x, y) = \prod_{i=1}^3 \frac{1}{\left(|x_i| + |y_i|\right)^2}.$$

We demand the following size estimate

$$(3.1) \qquad |K(x, y, z)| \lesssim D_\theta(x - z, y - z) S(x - z, y - z).$$

Let now $c = (c_1, c_2, c_3)$ be such that $|c_i - x_i| \leq \max(|x_i - z_i|, |y_i - z_i|)/2$ for $i = 1, 2, 3$. We assume that K satisfies the mixed size and Hölder estimates

$$
\begin{aligned}
&|K((c_1, x_2, x_3), y, z) - K(x, y, z)| \\
(3.2) \qquad &\lesssim \Big(\frac{|c_1 - x_1|}{|x_1 - z_1| + |y_1 - z_1|} \Big)^{\alpha_1} D_\theta(x - z, y - z) S(x - z, y - z),
\end{aligned}
$$

and

(3.3)
$$
\begin{aligned}
&|K((x_1, c_2, c_3), y, z) - K(x, y, z)| \\
&\lesssim \Big(\frac{|c_2 - x_2|}{|x_2 - z_2| + |y_2 - z_2|} + \frac{|c_3 - x_3|}{|x_3 - z_3| + |y_3 - z_3|} \Big)^{\alpha_{23}} D_\theta(x - z, y - z) S(x - z, y - z),
\end{aligned}
$$

where $\alpha_1, \alpha_{23} \in (0, 1]$. Finally, we assume that K satisfies the Hölder estimate

$$
\begin{aligned}
&|K(c, y, z) - K((c_1, x_2, x_3), y, z) - K((x_1, c_2, c_3), y, z) + K(x, y, z)| \\
(3.4) \qquad &\lesssim \Big(\frac{|c_1 - x_1|}{|x_1 - z_1| + |y_1 - z_1|} \Big)^{\alpha_1} \Big(\frac{|c_2 - x_2|}{|x_2 - z_2| + |y_2 - z_2|} + \frac{|c_3 - x_3|}{|x_3 - z_3| + |y_3 - z_3|} \Big)^{\alpha_{23}} \\
&\quad \times D_\theta(x - z, y - z) S(x - z, y - z).
\end{aligned}
$$

We also demand the symmetrical mixed size and Hölder estimates and Hölder estimates. For $j = 1, 2$, define the adjoint kernels $K^{*,j}$, $K_1^{*,j}$ and $K_{2,3}^{*,j}$ via the natural formulas, e.g.,

$$
K^{*,1}(x, y, z) = K(z, y, x), \quad K_1^{*,2}(x, y, z) = K(x, (z_1, y_2, y_3), (y_1, z_2, z_3)).
$$

We assume that each adjoint kernel satisfies the same estimates as the kernel K.

3.B. **Partial kernel representations.** Let $\widetilde{\theta} \in (0, 1]$. For every interval I^1 we assume that there exists a kernel

$$
K_{I^1} \colon (\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2) \setminus \{(x_{2,3}, y_{2,3}, z_{2,3}) \colon x_i = y_i = z_i \text{ for } i = 2 \text{ or } i = 3\} \to \mathbb{C},
$$

so that if $I_{j_1}^{2,3}$ and $I_{j_2}^{2,3}$ are disjoint for some $j_1, j_2 \in \{1, 2, 3\}$, then

$$
\begin{aligned}
&\langle T(1_{I^1} \otimes 1_{I_1^{2,3}}, 1_{I^1} \otimes 1_{I_2^{2,3}}), 1_{I^1} \otimes 1_{I_3^{2,3}} \rangle \\
&= \iiint K_{I^1}(x_{2,3}, y_{2,3}, z_{2,3}) 1_{I_1^{2,3}}(x_{2,3}) 1_{I_2^{2,3}}(y_{2,3}) 1_{I_3^{2,3}}(z_{2,3}) \, \mathrm{d}x_{2,3} \, \mathrm{d}y_{2,3} \, \mathrm{d}z_{2,3}.
\end{aligned}
$$

We demand the following estimates for the kernel K_{I^1}: The size estimate

$$
\begin{aligned}
&|K_{I^1}(x_{2,3}, y_{2,3}, z_{2,3})| \\
&\lesssim \Big(\frac{|I^1|(|x_2 - z_2| + |y_2 - z_2|)}{|x_3 - z_3| + |y_3 - z_3|} + \frac{|x_3 - z_3| + |y_3 - z_3|}{|I^1|(|x_2 - z_2| + |y_2 - z_2|)} \Big)^{-\widetilde{\theta}} \frac{|I^1|}{\prod_{i=2}^3 (|x_i - z_i| + |y_i - z_i|)^2}
\end{aligned}
$$

and the continuity estimate

$$
\begin{aligned}
&|K_{I^1}(c_{2,3}, y_{2,3}, z_{2,3}) - K_{I^1}(x_{2,3}, y_{2,3}, z_{2,3})| \\
&\lesssim \Big(\frac{|c_2 - x_2|}{|x_2 - z_2| + |y_2 - z_2|} + \frac{|c_3 - x_3|}{|x_3 - z_3| + |y_3 - z_3|} \Big)^{\alpha_{23}} \\
&\quad \times \Big(\frac{|I^1|(|x_2 - z_2| + |y_2 - z_2|)}{|x_3 - z_3| + |y_3 - z_3|} + \frac{|x_3 - z_3| + |y_3 - z_3|}{|I^1|(|x_2 - z_2| + |y_2 - z_2|)} \Big)^{-\widetilde{\theta}} \frac{|I^1|}{\prod_{i=2}^3 (|x_i - z_i| + |y_i - z_i|)^2}
\end{aligned}
$$

whenever $c_{2,3} = (c_2, c_3)$ is such that $|c_i - x_i| \leq \max(|x_i - z_i|, |y_i - z_i|)/2$ for $i = 2, 3$. We also assume the symmetrical continuity estimates.

We assume similar one-parameter conditions for the other partial kernel representation. That is, for every rectangle $I^{2,3}$, there exists a standard bilinear Calderón-Zygmund kernel $K_{I^{2,3}}$ so that if $I_{j_1}^1$ and $I_{j_2}^1$ are disjoint for some $j_1, j_2 \in \{1, 2, 3\}$, then

$$\langle T(1_{I_1^1} \otimes 1_{I^{2,3}}, 1_{I_2^1} \otimes 1_{I^{2,3}}), 1_{I_3^1} \otimes 1_{I^{2,3}} \rangle$$
$$= \iiint K_{I^{2,3}}(x_1, y_1, z_1) 1_{I_1^1}(x_1) 1_{I_2^1}(y_1) 1_{I_3^1}(z_1) \, dx_1 \, dy_1 \, dz_1.$$

The kernel $K_{I^{2,3}}$ satisfies the standard estimates

$$|K_{I^{2,3}}(x_1, y_1, z_1)| \leq C_{K_{I^{2,3}}} \frac{1}{(|x_1 - z_1| + |y_1 - z_1|)^2},$$

$$|K_{I^{2,3}}(x_1, y_1, z_1) - K_{I^{2,3}}(c_1, y_1, z_1)| \leq C_{K_{I^{2,3}}} \frac{|x_1 - c_1|^{\alpha_1}}{(|x_1 - z_1| + |y_1 - z_1|)^{2+\alpha_1}}$$

whenever $|x_1 - c_1| \leq \max(|x_1 - z_1|, |y_1 - z_1|)/2$, and the symmetric continuity estimates. The smallest possible constant $C_{K_{I^{2,3}}}$ in these inequalities is denoted by $\|K_{I^{2,3}}\|_{\mathrm{CZ}_{\alpha_1}}$. We then assume that

$$\|K_{I^{2,3}}\|_{\mathrm{CZ}_{\alpha_1}} \lesssim |I^{2,3}|.$$

3.C. **Cancellation assumptions: paraproduct free operators.** We say that T is a paraproduct free operator, if for all cancellative Haar functions h_{I^1} and $h_{I^{2,3}}$ we have

$$\langle T(1 \otimes 1_{J_1^{2,3}}, 1 \otimes 1_{J_2^{2,3}}), h_{I^1} \otimes 1_{J_3^{2,3}} \rangle = \langle T_1^{*,j}(1 \otimes 1_{J_1^{2,3}}, 1 \otimes 1_{J_2^{2,3}}), h_{I^1} \otimes 1_{J_3^{2,3}} \rangle$$
$$= \langle T(1_{I_1^1} \otimes 1, 1_{I_2^1} \otimes 1), 1_{I_3^1} \otimes h_{I^{2,3}} \rangle = \langle T_{2,3}^{*,j}(1_{I_1^1} \otimes 1, 1_{I_2^1} \otimes 1), 1_{I_3^1} \otimes h_{I^{2,3}} \rangle = 0$$

for all the adjoints $j \in \{1, 2\}$. We always assume that all bilinear Zygmund operators in this article satisfy this cancellation condition. The intention of this condition is to guarantee that our operator is representable using cancellative shifts only.

3.D. **Weak boundedness property.** We say that T satisfies the weak boundedness property if

$$|\langle T(1_I, 1_I), 1_I \rangle| \lesssim |I|$$

for all Zygmund rectangles $I = I^1 \times I^2 \times I^3$.

3.5. **Definition.** We say that a bilinear operator T is a paraproduct free Calderón-Zygmund operator adapted to Zygmund dilations (CZZ operator) if T has the full kernel representation, the partial kernel representations, is paraproduct free and satisfies the weak boundedness property.

4. ESTIMATES FOR THE SHIFT COEFFICIENTS

We now move to consider the shift coefficients that appeared in the decomposition of T in Section 2.F. When T is a CZZ operator, we can estimate them. Without loss of generality, we estimate

$$\langle T(h_{I \dotplus n_1}^0, h_{I \dotplus n_2}^0), h_{I,Z} \rangle$$

for $I \in \mathcal{D}_Z$ and different values of $n_1, n_2 \in \mathbb{Z}^3$, and without loss of generality we assume $\theta = \tilde{\theta} < 1$. The coefficients related to the other terms of the decomposition (other than the main term (2.8)) may have a different set of Haar functions, but they are treated similarly.

We show that

$$(4.1) \qquad |\langle T(h^0_{I+n_1}, h^0_{I+n_2}), h_{I,Z}\rangle| \lesssim (|k|+1)^2 \varphi(k) \frac{|I|^{\frac{3}{2}}}{|K|^2},$$

where

$$\varphi(k) := 2^{-k^1 \alpha_1 - k^2 \min\{\alpha_{23}, \theta\} - \max\{k^3 - k^1 - k^2, 0\}\theta}.$$

For terms of this particular form, we would not actually need to analyze some of the diagonal cases (see Section 2.F). However, these diagonal terms would appear in some other forms, so it makes sense to deal with them here (even though in the real situation the Haar functions might be permuted differently, this does not matter, and the calculations we present apply). It is very helpful to study the linear case [14], since the kernel estimates are relatively involved and we will not repeat every detail when they are similar.

Let $m^i := \max_{j=1,2} |n^i_j|$. The analysis of the coefficients splits into combinations of

$$\begin{cases} m^1 \in (2^{k^1-3}, 2^{k^1-2}], & k^1 = 3, 4, \dots, & \text{(Separated)} \\ m^1 = 1, & & \text{(Adjacent)} \\ m^1 = 0, & & \text{(Identical)} \end{cases}$$

and

$$\begin{cases} m^i \in (2^{k^i-3}, 2^{k^i-2}], & i = 2, 3, \ k^i = 3, 4, \dots, & \text{(Separated)} \\ m^2 < 2 \text{ and } m^3 \in (2^{k^3-3}, 2^{k^3-2}], & k^3 = 3, 4, \dots, & \text{(Separated)} \\ m^2 \in (2^{k^2-3}, 2^{k^2-2}] \text{ and } m^3 < 2 & k^2 = 3, 4, \dots, & \text{(Separated)} \\ m^2 = 1 \text{ and } m^3 \leq 1 & & \text{(Adjacent)} \\ m^2 = 0 \text{ and } m^3 = 1 & & \text{(Adjacent)} \\ m^2 = 0 = m^3. & & \text{(Identical)} \end{cases}$$

It is enough to consider $m^i = n^i_1$ since the case $m^i = n^i_2$ is symmetrical. We will not go through explicitly every combination – rather, we choose some illustrative examples.

Separated/Separated. We begin with the case $|n^i_1| \geq 2$ for all $i = 1, 2, 3$. Hence,

$$|x_i - z_i| \geq |n^i_1| \ell(I^i) \geq 2^{k^i-3} \ell(I^i)$$

and

$$|x_i - z_i| \leq |n^i_1| \ell(I^i) + 2\ell(I^i) \leq 2^{k^i-1} \ell(I^i)$$

for $i = 1, 2, 3$. Moreover, $|x_i - z_i| \geq |y_i - z_i|/2 \geq 0$ for $i = 1, 2, 3$. Thus, we have the estimate

$$\left(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{(|x_3 - z_3| + |y_3 - z_3|)} + \frac{|x_3 - z_3| + |y_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)} \right)^{-\theta}$$

$$\sim \left(\frac{\prod_{i=1}^2 |x_i - z_i|}{|x_3 - z_3|} + \frac{|x_3 - z_3|}{\prod_{i=1}^2 |x_i - z_i|} \right)^{-\theta}$$

$$\sim \left(\frac{\prod_{i=1}^2 2^{k^i} \ell(I^i)}{2^{k^3} \ell(I^3)} + \frac{2^{k^3} \ell(I^3)}{\prod_{i=1}^2 2^{k^i} \ell(I^i)} \right)^{-\theta} = (2^{k^1+k^2-k^3} + 2^{k^3-k^1-k^2})^{-\theta}.$$

Let c_{I^i} denote the center of the interval I^i. Furthermore, notation like c_I then refers to the corresponding tuple $(c_{I^1}, c_{I^2}, c_{I^3})$. Using the cancellation of the Haar function we then have

$$\Big| \iiint K(x,y,z) h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h_{I,Z}(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \Big|$$

$$= \Big| \iiint \Big(K(x,y,z) - K(x,y,(c_{I^1}, z_{2,3})) - K(x,y,(z_1, c_{I^{2,3}})) + K(x,y,c_I) \Big)$$

$$\times h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h_{I,Z}(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \Big|$$

$$\lesssim \iiint 2^{-k^1 \alpha_1} (2^{-k^2} + 2^{-k^3})^{\alpha_{23}} \frac{(2^{k^1 + k^2 - k^3} + 2^{k^3 - k^1 - k^2})^{-\theta}}{|K|^2} h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h^0_I(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z$$

$$= 2^{-k^1 \alpha_1} (2^{-k^2} + 2^{-k^3})^{\alpha_{23}} (2^{k^1 + k^2 - k^3} + 2^{k^3 - k^1 - k^2})^{-\theta} \frac{|I|^{\frac{3}{2}}}{|K|^2} \leq \varphi(k) \frac{|I|^{\frac{3}{2}}}{|K|^2}.$$

Let us then consider the case, where we have separation in the parameter 3 but not in the parameter 2 – that is, $|n_1^2| < 2 \leq |n_1^3|$. Then

$$(4.2) \qquad \Big(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{|x_3 - z_3| + |y_3 - z_3|} + \frac{|x_3 - z_3| + |y_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)} \Big)^{-\theta}$$

$$\sim \Big(\frac{|x_2 - z_2| + |y_2 - z_2|}{2^{k^3 - k^1}|I^2|} + \frac{2^{k^3 - k^1}|I^2|}{|x_2 - z_2| + |y_2 - z_2|} \Big)^{-\theta}$$

$$\lesssim \Big(\frac{|x_2 - z_2|}{2^{k^3 - k^1}|I^2|} + \frac{2^{k^3 - k^1}|I^2|}{|x_2 - z_2|} \Big)^{-\theta} + \Big(\frac{|y_2 - z_2|}{2^{k^3 - k^1}|I^2|} + \frac{2^{k^3 - k^1}|I^2|}{|y_2 - z_2|} \Big)^{-\theta},$$

and so using the mixed estimates

$$\Big| \iiint K(x,y,z) h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h_{I,Z}(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \Big|$$

$$= \Big| \iiint \Big(K(x,y,z) - K(x,y,(c_{I^1}, z_{2,3})) \Big) h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h_{I,Z}(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \Big|$$

$$\lesssim \iiint 2^{-k^1 \alpha_1} |K^1|^{-2} |K^3|^{-2} \frac{\Big(\frac{|x_2 - z_2| + |y_2 - z_2|}{2^{k^3 - k^1}|I^2|} + \frac{2^{k^3 - k^1}|I^2|}{|x_2 - z_2| + |y_2 - z_2|} \Big)^{-\theta}}{\big(|x_2 - z_2| + |y_2 - z_2| \big)^2}$$

$$\times h^0_{I \dotplus n_1}(x) h^0_{I \dotplus n_2}(y) h^0_I(z) \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z$$

$$= 2^{-k^1 \alpha_1} \frac{|I^1|^{\frac{3}{2}} |I^3|^{\frac{3}{2}}}{|K^1|^2 |K^3|^2} \iiint \frac{\Big(\frac{|x_2 - z_2| + |y_2 - z_2|}{2^{k^3 - k^1}|I^2|} + \frac{2^{k^3 - k^1}|I^2|}{|x_2 - z_2| + |y_2 - z_2|} \Big)^{-\theta}}{\big(|x_2 - z_2| + |y_2 - z_2| \big)^2}$$

$$\times h^0_{I^2 \dotplus n_1^2}(x_2) h^0_{I^2 \dotplus n_2^2}(y_2) h^0_{I^2}(z_2) \,\mathrm{d}x_2 \,\mathrm{d}y_2 \,\mathrm{d}z_2$$

$$\lesssim \varphi(k) \frac{|I|^{\frac{3}{2}}}{|K|^2}.$$

We note that the last inequality requires a case study (see also [14, Lemma 8.5]) and we used the standard estimate

$$(4.3) \qquad \int_{\mathbb{R}^d} \frac{\mathrm{d}u}{(r + |u_0 - u|)^{d+\alpha}} \lesssim r^{-\alpha}.$$

Symmetrical estimates hold if $|n_1^2| \geq 2 > |n_1^3|$.

Adjacent/Separated. We look at the example case $|n_1^2| \geq 2 > |n_1^3|$ and $|n_1^1| = 1$. By the size estimate we have

$$|\langle T(h_{I+n_1}^0, h_{I+n_2}^0), h_{I,Z} \rangle|$$

$$\lesssim \frac{|I^2|^{3/2}}{|I^{1,3}|^{3/2}|K^2|^2} \iiint \frac{\left(\frac{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)}{|x_3-z_3|+|y_3-z_3|} + \frac{|x_3-z_3|+|y_3-z_3|}{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)} \right)^{-\theta}}{\left(|x_1-z_1| + |y_1-z_1| \right)^2 \left(|x_3-z_3| + |y_3-z_3| \right)^2}$$

$$\times 1_{I^{1,3}+n_1^{1,3}}(x_{1,3}) 1_{I^{1,3}+n_2^{1,3}}(y_{1,3}) 1_{I^{1,3}}(z_{1,3}) \, \mathrm{d}x_{1,3} \, \mathrm{d}y_{1,3} \, \mathrm{d}z_{1,3}.$$

Similarly as (4.2), we can split the integral into two terms. Then by (4.3) we reduce the problem to estimating

$$\iiint \frac{\left(\frac{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)}{|x_3-z_3|} + \frac{|x_3-z_3|}{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)} \right)^{-\theta}}{\left(|x_1-z_1| + |y_1-z_1| \right)^2 |x_3-z_3|}$$

$$\times 1_{I^{1,3}+n_1^{1,3}}(x_{1,3}) 1_{I^1+n_2^1}(y_1) 1_{I^{1,3}}(z_{1,3}) \, \mathrm{d}x_{1,3} \, \mathrm{d}y_1 \, \mathrm{d}z_{1,3}$$

$$+ \iiint \frac{\left(\frac{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)}{|y_3-z_3|} + \frac{|y_3-z_3|}{(|x_1-z_1|+|y_1-z_1|)2^{k^2}\ell(I^2)} \right)^{-\theta}}{(|x_1-z_1| + |y_1-z_1|)^2 |y_3-z_3|}$$

$$\times 1_{I^{1,3}+n_1^{1,3}}(x_{1,3}) 1_{I^{1,3}+n_2^{1,3}}(y_{1,3}) 1_{I^1}(z_1) \, \mathrm{d}x_{1,3} \, \mathrm{d}y_{1,3} \, \mathrm{d}z_1.$$

Since they are similar, we only bound the first one. Note that

$$\left(\frac{(|x_1-z_1| + |y_1-z_1|)2^{k^2}\ell(I^2)}{|x_3-z_3|} + \frac{|x_3-z_3|}{(|x_1-z_1| + |y_1-z_1|)2^{k^2}\ell(I^2)} \right)^{-\theta}$$

$$\times (|x_1-z_1| + |y_1-z_1|)^{-2}$$

$$\leq \left(\frac{(|x_1-z_1| + |y_1-z_1|)2^{k^2}\ell(I^2)}{|x_3-z_3|} \right)^{-\theta} (|x_1-z_1| + |y_1-z_1|)^{-2} \chi_{\{|x_1-z_1|2^{k^2}\ell(I^2) \geq |x_3-z_3|\}}$$

$$+ \left(\frac{|x_3-z_3|}{(|x_1-z_1| + |y_1-z_1|)2^{k^2}\ell(I^2)} \right)^{-\theta} (|x_1-z_1| + |y_1-z_1|)^{-2} \chi_{\{|x_1-z_1|2^{k^2}\ell(I^2) < |x_3-z_3|\}}.$$

Then apply (4.3) to the integral over y_1, then by following the linear case [14, Lemma 8.11] we get that the above integral is bounded by $|I^{1,3}|k^2 2^{-k^2\theta}$. Thus, we get

$$|\langle T(h_{I+n_1}^0, h_{I+n_2}^0), h_{I,Z} \rangle| \lesssim \frac{|I^2|^{3/2}}{|I^{1,3}|^{1/2}|K^2|^2} k^2 2^{-k^2\theta} \lesssim k^2 \varphi(k) \frac{|I|^{\frac{3}{2}}}{|K|^2}.$$

Adjacent/Adjacent. We again have no major changes to the linear case but in order to use the estimate

$$(4.4) \qquad \int_{\mathbb{R}} \frac{\left(\frac{t}{|u|} + \frac{|u|}{t}\right)^{-\theta}}{t|u|} |f(u)| \, \mathrm{d}u \lesssim t^{-1} M f(0)$$

we need to first use (4.3) repeatedly. For example, consider $|n_1^1| = 1$ and $|n_1^2| = 1, |n_1^3| \leq 1$. By the size estimate of the kernel, we need to control

$$\frac{\left(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{|x_3 - z_3| + |y_3 - z_3|} + \frac{|x_3 - z_3| + |y_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}\right)^{-\theta}}{\prod_{i=1}^3 \left(|x_i - z_i| + |y_i - z_i|\right)^2} h_{I \dot{+} n_1}^0(x) h_{I \dot{+} n_2}^0(y) h_I^0(z).$$

As before, we split this into two terms, one of them is

$$\frac{\left(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{|x_3 - z_3|} + \frac{|x_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}\right)^{-\theta}}{\prod_{i=1}^3 \left(|x_i - z_i| + |y_i - z_i|\right)^2} h_{I \dot{+} n_1}^0(x) h_{I \dot{+} n_2}^0(y) h_I^0(z).$$

We then apply (4.3) to the integral over y_3, and then use the previous trick repeatedly. That is, we write

$$\left(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{|x_3 - z_3|} + \frac{|x_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}\right)^{-\theta}$$

$$\leq \left(\frac{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}{|x_3 - z_3|}\right)^{-\theta} \chi_{\{|x_1 - z_1|(|x_2 - z_2| + |y_2 - z_2|) \geq |x_3 - z_3|\}}$$

$$+ \left(\frac{|x_3 - z_3|}{\prod_{i=1}^2 (|x_i - z_i| + |y_i - z_i|)}\right)^{-\theta} \chi_{\{|x_1 - z_1|(|x_2 - z_2| + |y_2 - z_2|) < |x_3 - z_3|\}}$$

and apply (4.3) to the integral over y_1. Then, after a similar argument on y_2, we finally arrive at

$$\frac{1}{|I|^{\frac{1}{2}}} \iint \frac{\left(\frac{\prod_{i=1}^2 |x_i - z_i|}{|x_3 - z_3|} + \frac{|x_3 - z_3|}{\prod_{i=1}^2 |x_i - z_i|}\right)^{-\theta}}{\prod_{i=1}^3 |x_i - z_i|} h_{I \dot{+} n_1}^0(x) h_I^0(z) \, \mathrm{d}x \, \mathrm{d}z$$

$$\lesssim \frac{1}{|I|^{\frac{1}{2}}} \lesssim \frac{|I|^{\frac{3}{2}}}{|K|^2}.$$

Adjacent/Identical. We consider the case $|n_1^1| = 1$ and $n_j^2 = n_j^3 = 0, j = 1, 2$. We write

$$\sum_{Q_1^{2,3}, Q_2^{2,3}, Q_3^{2,3} \in \mathrm{ch}(I^{2,3})} \langle T(h_{I \dot{+} n_1}^0 1_{Q_1^{2,3}}, h_{I \dot{+} n_2}^0 1_{Q_2^{2,3}}), h_{I,Z} 1_{Q_3^{2,3}} \rangle.$$

It is enough to consider $Q_1^{2,3} = Q_2^{2,3} = Q_3^{2,3}$ since otherwise we have adjacent intervals, and we are back in the Adjacent/Adjacent case. Hence, the partial kernel representation 3.B yields that

$$\left| \pm |I^{2,3}|^{-\frac{3}{2}} \iiint K_{Q_1^{2,3}} h_{I^1 \dot{+} n_1^1}^0 h_{I^1 \dot{+} n_2^1}^0 h_{I^1} \right|$$

$$\lesssim \frac{|I^{2,3}|^{\frac{3}{2}}}{|K^{2,3}|^2} \iiint \frac{1}{(|x_1 - z_1| + |y_1 - z_1|)^2} h_{I^1 \dot{+} n_1^1}^0(x_1) h_{I^1 \dot{+} n_2^1}^0(y_1) h_{I^1}(z_1) \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}z_1.$$

Then, first using (4.3) and then standard integration methods we get the following inequality

$$\iiint \frac{1}{(|x_1 - z_1| + |y_1 - z_1|)^2} h^0_{I^1 \dotplus n^1_1}(x_1) h^0_{I^1 \dotplus n^1_2}(y_1) h_{I^1}(z_1) \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}z_1$$

$$\lesssim \frac{1}{|I^1|^{\frac{1}{2}}} \iint \frac{1}{|x_1 - z_1|} h^0_{I^1 \dotplus n^1_1}(x_1) h_{I^1}(z_1) \, \mathrm{d}x_1 \, \mathrm{d}z_1$$

$$\lesssim \frac{1}{|I^1|^{\frac{1}{2}}} \sim \frac{|I^1|^{\frac{3}{2}}}{|K^1|^2}$$

as desired.

Identical/Identical. Just like in above we split the pairing to

$$\sum_{Q^1_1, Q^1_2, Q^1_3 \in \mathrm{ch}(I^1)} \sum_{Q^{2,3}_1, Q^{2,3}_2, Q^{2,3}_3 \in \mathrm{ch}(I^{2,3})} \langle T(h^0_{I \dotplus n_1}(1_{Q^1_1} \otimes 1_{Q^{2,3}_1}), h^0_{I \dotplus n_2}(1_{Q^1_2} \otimes 1_{Q^{2,3}_2})), h_{I,Z}(1_{Q^1_3} \otimes 1_{Q^{2,3}_3}) \rangle.$$

The cases when $Q^1_i \neq Q^1_j$ for some $i, j = 1, 2, 3, i \neq j$ are essentially included in the cases of the two previous subsections. Hence, we consider $Q^1_1 = Q^1_2 = Q^1_3$. Then there are two cases left, that is, either $Q^{2,3}_i \neq Q^{2,3}_j$ for some $i, j = 1, 2, 3, i \neq j$, or $Q^{2,3}_1 = Q^{2,3}_2 = Q^{2,3}_3$. Beginning from the latter one, similarly as in [14], by splitting Q^3_1 into sub-intervals we get

$$|\langle T(1_{Q^1_1} \otimes 1_{Q^{2,3}_1}, 1_{Q^1_1} \otimes 1_{Q^{2,3}_1}), 1_{Q^1_1} \otimes 1_{Q^{2,3}_1} \rangle| \lesssim |Q^1_1||Q^{2,3}_1|$$

by the weak boundedness property 3.D and the Identical/Adjacent case. Hence, we get the desired bound

$$|\langle T(h^0_{I \dotplus n_1}(1_{Q^1_1} \otimes 1_{Q^{2,3}}), h^0_{I \dotplus n_2}(1_{Q^1_1} \otimes 1_{Q^{2,3}})), h_{I,Z}(1_{Q^1_1} \otimes 1_{Q^{2,3}}) \rangle| \lesssim \frac{|Q_1|}{|I|^{\frac{3}{2}}} \leq \frac{|I|^{\frac{3}{2}}}{|K|^2}.$$

We handle the remaining case $Q^{2,3}_i \neq Q^{2,3}_j$ for some $i, j = 1, 2, 3, i \neq j$. By the partial kernel representation and its size estimate we get

$$\Big| \pm |I|^{-\frac{3}{2}} \iiint K_{Q^1_1} 1_{Q^{2,3}_1} 1_{Q^{2,3}_2} 1_{Q^{2,3}_3} \Big|$$

$$\lesssim \frac{1}{|I^1|^{\frac{1}{2}}} \frac{1}{|I^{2,3}|^{\frac{3}{2}}} \iiint \Big(\frac{|I^1|(|x_2 - z_2| + |y_2 - z_2|)}{|x_3 - z_3| + |y_3 - z_3|} + \frac{|x_3 - z_3| + |y_3 - z_3|}{|I^1|(|x_2 - z_2| + |y_2 - z_2|)} \Big)^{-\theta}$$

$$\times \prod_{i=2}^3 \frac{1}{(|x_i - z_i| + |y_i - z_i|)^2} 1_{Q^{2,3}_1} 1_{Q^{2,3}_2} 1_{Q^{2,3}_3} \, \mathrm{d}x_{2,3} \, \mathrm{d}y_{2,3} \, \mathrm{d}z_{2,3}.$$

Then using similar arguments as in the Adjacent/Adjacent case and (4.4) gives us the desired bound.

5. Structural decomposition of Zygmund shifts

In this section we decompose the bilinear Zygmund shifts (see Section 2.D) as a sum of operators with simpler cancellation properties. The decomposition is not optimal (in the sense that weighted estimates with Zygmund weights cannot be obtained via this) – however, it is sufficient for unweighted boundedness in the full range that we later obtain via tri-parameter theory. Recall that $k = (k^1, k^2, k^3)$ is the complexity of the bilinear Zygmund shift.

5.1. **Definition.** Bilinear operators of the form

$$(5.2) \qquad S_{(l_1,l_2,l_3)}(f_1, f_2) = \sum_{L \in \mathcal{D}_\lambda} \sum_{I_j^{(\ell_j)} = L} a_{L,(I_j)} \langle f_1, h_{I_1^1} \otimes h_{I_1^{2,3}}^0 \rangle \langle f_2, h_{I_2^1}^0 \otimes h_{I_2^{2,3}} \rangle h_{I_3},$$

where $\lambda = 2^n, n \in \mathbb{Z}, |n| \le 3 \max(k^i)$ and

$$|a_{L,(I_j)}| \le \frac{|I_j|^{\frac{3}{2}}}{|L|^2},$$

are tri-parameter bilinear shifts of Zygmund nature if at least one rectangle $I_{i_1}^1 \times I_{i_2}^{2,3}, i_1 = 1, 3, i_2 = 2, 3$ is a Zygmund rectangle and

(1) $\ell_j^i \le k^i$ for all $i, j = 1, 2, 3$;
(2) $(\ell_j^3 - \ell_j^2)_+ \le (k^3 - k^2)_+$ for all $j = 1, 2, 3$.

Moreover, any adjoint

$$S_{(l_1,l_2,l_3)}^{j_1^*, j_{2,3}^*}, \qquad j_1, j_{2,3} \in \{0, 1, 2\},$$

is also considered to be a tri-parameter bilinear shift of Zygmund nature. Here, the adjoint $j_{2,3}^*$ means that, for example, in case $j_{2,3} = 1$ functions $h_{I_1^{2,3}}^0$ and $h_{I_3^{2,3}}$ switch places.

Note that these operators share a 'weaker' Zygmund structure. Ideally, we would want to have $I_3 \in \mathcal{D}_Z$ and $I_1^1 \times I_2^{2,3} \in \mathcal{D}_Z$.

5.3. **Proposition.** *Let $Q_k, k = (k^1, k^2, k^3)$, be a bilinear Zygmund shift operator as defined in Section 2.D. Then*

$$Q_k = C \sum_{u=1}^{c} \sum_{l^1=0}^{k^1-1} \sum_{l^{2,3}=0}^{k^{2,3}-1} S^u,$$

where S^u is a bilinear operator as in Definition 5.1 with complexity depending on l and k, and

$$\sum_{l^{2,3}=0}^{k^{2,3}-1} := \begin{cases} \displaystyle\sum_{0 \le l^2 = l^3 \le k^2 - 1} + \sum_{\substack{l^2 = k^2 \\ k^2 \le l^3 \le k^3 - 1}}, & \text{if } k^3 \ge k^2 \\ \displaystyle\sum_{0 \le l^2 = l^3 \le k^3 - 1} + \sum_{\substack{k^3 \le l^2 \le k^2 - 1 \\ l^3 = k^3}}, & \text{if } k^3 < k^2. \end{cases}$$

Proof. The argument is similar in spirit to the purely bi-parameter decomposition in [1]. For notational convenience, we consider a shift Q_k of the particular form

$$\langle Q_k(f_1, f_2), f_3 \rangle$$

$$
\begin{aligned}
= \sum_{\substack{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}}} \sum_{\substack{I_1,I_2,I_3 \in \mathcal{D}_Z \\ I_j^{(k)}=K}} a_{K,(I_j)} \Big[& A^{3,3}_{I_1,I_2,I_3} - A^{3,3}_{I_3^1 \times I_1^{2,3}, I_3^1 \times I_2^{2,3}, I_3} \\
& - A^{3,3}_{I_1^1 \times I_3^{2,3}, I_2^1 \times I_3^{2,3}, I_3} + A^{3,3}_{I_3,I_3,I_3} \Big]
\end{aligned}
$$

$$
\begin{aligned}
= \sum_{\substack{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}}} \sum_{\substack{I_1,I_2,I_3 \in \mathcal{D}_Z \\ I_j^{(k)}=K}} a_{K,(I_j)} \langle f_3, h_{I_3} \rangle \Big[& \langle f_1, h_{I_1}^0 \rangle \langle f_2, h_{I_2}^0 \rangle - \langle f_1, h_{I_3^1}^0 h_{I_1^{2,3}}^0 \rangle \langle f_2, h_{I_3^1}^0 h_{I_2^{2,3}}^0 \rangle \\
& - \langle f_1, h_{I_1^1}^0 h_{I_3^{2,3}}^0 \rangle \langle f_2, h_{I_2^1}^0 h_{I_3^{2,3}}^0 \rangle + \langle f_1, h_{I_3}^0 \rangle \langle f_2, h_{I_3}^0 \rangle \Big].
\end{aligned}
$$

There is no essential difference in the general case. Let us also use the usual abbreviation $\mathcal{D}_{2^{-k^1-k^2+k^3}} = \mathcal{D}_\lambda$.

We define
$$
b_{K,(I_j)} = |I_1| a_{K,(I_j)}
$$
and
$$
B^{3,3}_{I_1,I_2,I_3} = \langle f_1 \rangle_{I_1} \langle f_2 \rangle_{I_2} \langle f_3, h_{I_3} \rangle.
$$

We can write the shift Q_k using these by replacing a with b and A with B.

Recall the notation
$$
\Delta_{K^1}^{l^1} f = \sum_{\substack{L^1 \in \mathcal{D}^1 \\ (L^1)^{(l^1)}=K^1}} \Delta_{L^1} f, \qquad P_{K^1}^{k^1} f = \sum_{l^1=0}^{k^1-1} \Delta_{K^1}^{l^1} f,
$$
$$
E_{K^1} f = \langle f \rangle_{K^1} 1_{K^1}, \qquad E_{K^1}^{k^1} f = \sum_{\substack{L^1 \in \mathcal{D}^1 \\ (L^1)^{(k^1)}=K^1}} \langle f \rangle_{L^1} 1_{L^1}.
$$

Let us define

$$
(5.4) \qquad P_{K^{2,3}}^{k^{2,3}} f := \sum_{l^{2,3}=0}^{k^{2,3}-1} \Delta_{K^{2,3}}^{(l^2,l^3)} f := \begin{cases} \displaystyle\sum_{l^2=0}^{k^2-1} \Delta_{K^{2,3}}^{l^2} f + \sum_{l^3=k^2}^{k^3-1} E_{K^2}^{k^2} \Delta_{K^3}^{l^3} f, & \text{if } k^3 \geq k^2 \\ \displaystyle\sum_{l^3=0}^{k^3-1} \Delta_{K^{2,3}}^{l^3} f + \sum_{l^2=k^3}^{k^2-1} \Delta_{K^2}^{l^2} E_{K^3}^{k^3} f, & \text{if } k^3 < k^2, \end{cases}
$$

where we have the standard one-parameter definition
$$
\Delta_{K^{2,3}}^{l^i} f = \sum_{\substack{L^{2,3} \in \mathcal{D}^{2,3} \\ (L^2)^{(l^i)} \times (L^3)^{(l^i)}=K^2 \times K^3}} \Delta_{L^{2,3}} f.
$$

We also use a similar shorthand for the expanded martingale blocks
$$
\sum_{l^{2,3}=0}^{k^{2,3}-1} \Delta_{K^{2,3}}^{(l^2,l^3)} f = \sum_{l^{2,3}=0}^{k^{2,3}-1} \sum_{(L^{2,3})^{(l^{2,3})}=K^{2,3}} \langle f, h_{L^{2,3}} \rangle h_{L^{2,3}},
$$

where we allow, for example, that $h_{L^{2,3}} = h_{L^2}^0 \otimes h_{L^3}$ when $k^3 > k^2$ and $l^2 = k^2$.

Using this notation we define the following. For a cube I and integers $l, j_0 \in \{1, 2, \dots\}$ we define

$$(5.5) \qquad D_{I,l}(j, j_0) = \begin{cases} E_I, & \text{if } j \in \{1, \dots, j_0 - 1\}, \\ P_I^l, & \text{if } j = j_0, \\ \text{id}, & \text{if } j \in \{j_0 + 1, j_0 + 2, \dots\}, \end{cases}$$

where id denotes the identity operator, and if we have a rectangle $I^{2,3}$ and a tuple $l^{2,3}$ we use the modified $P_{I^{2,3}}^{l^{2,3}}$.

Let I_1, I_2, I_3 be as in the summation of Q_k. We use the above notation in parameter one $D_{I^1, l^1}(j, j_0)$ and for the other two parameters we use $D_{I^{2,3}, l^{2,3}}(j, j_0)$. Thus, expanding to the martingale blocks leads us to

$$B_{I_1, I_2, I_3}^{3,3}$$
$$= \sum_{m_1, m_2 = 1}^{3} \prod_{j=1}^{2} \langle D_{K^1, k^1}^1(j, m_1) D_{K^{2,3}, k^{2,3}}^{2,3}(j, m_2) f_j \rangle_{I_j} \langle f_3, h_{I_3} \rangle.$$

Hence, we may write

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} B_{I_1, I_2, I_3}^{3,3} =: \sum_{m_1, m_2 = 1}^{3} \Sigma_{m_1, m_2}^1.$$

Also, we have that

$$B_{I_3^1 \times I_1^{2,3}, I_3^1 \times I_2^{2,3}, I_3^1 \times I_3^{2,3}}^{3,3} = \sum_{m_2 = 1}^{3} \prod_{j=1}^{2} \langle D_{K^{2,3}, k^{2,3}}^{2,3}(j, m_2) f_j \rangle_{I_3^1 \times I_j^{2,3}} \langle f_3, h_{I_3} \rangle$$

and

$$B_{I_1^1 \times I_3^{2,3}, I_2^1 \times I_3^{2,3}, I_3^1 \times I_3^{2,3}}^{3,3} = \sum_{m_1 = 1}^{3} \prod_{j=1}^{2} \langle D_{K^1, k_1}^1(j, m_1) f_j \rangle_{I_j^1 \times I_3^{2,3}} \langle f_3, h_{I_3} \rangle,$$

which gives that

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} B_{I_3^1 \times I_1^{2,3}, I_3^1 \times I_2^{2,3}, I_3^1 \times I_3^{2,3}}^{3,3} =: \sum_{m_2 = 1}^{3} \Sigma_{m_2}^2$$

and

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} B_{I_1^1 \times I_3^{2,3}, I_2^1 \times I_3^{2,3}, I_3^1 \times I_3^{2,3}}^{3,3} =: \sum_{m_1 = 1}^{3} \Sigma_{m_1}^3.$$

Finally, we just set

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} B_{I_3, I_3, I_3}^{3,3} =: \Sigma^4.$$

Thus, we have the following decomposition

$$\langle Q_k(f_1, f_2), f_3 \rangle = \sum_{m_1, m_2 = 1}^{2} \Sigma_{m_1, m_2}^1 + \sum_{m_2=1}^{2} (\Sigma_{3, m_2}^1 - \Sigma_{m_2}^2)$$
$$+ \sum_{m_1=1}^{2} (\Sigma_{m_1, 3}^1 - \Sigma_{m_1}^3) + (\Sigma_{3,3}^1 - \Sigma_3^2 - \Sigma_3^3 + \Sigma^4).$$

First, we take one Σ_{m_1, m_2}^1 with $m_1, m_2 \in \{1, 2\}$. For notational convenience, we choose the case $m_1 = m_2 = 2$. Recall that

$$\Sigma_{2,2}^1 = \sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} b_{K, (I_j)} \langle f_1 \rangle_K \langle P_{K^1}^{k^1} P_{K^{2,3}}^{k^{2,3}} f_2 \rangle_{I_2} \langle f_3, h_{I_3} \rangle.$$

We expand

$$\langle P_{K^1}^{k^1} P_{K^{2,3}}^{k^{2,3}} f_2 \rangle_{I_2} = \sum_{l^1=0}^{k^1-1} \sum_{l^{2,3}=0}^{k^{2,3}-1} \sum_{\substack{(L^1)^{(l^1)}=K^1 \\ (L^{2,3})^{(l^{2,3})}=K^{2,3}}} \langle f_2, h_{L^1} \otimes h_{L^{2,3}} \rangle \langle h_{L^1} \otimes h_{L^{2,3}} \rangle_{I_2}$$

and note that L is not necessarily a Zygmund rectangle. It holds that

$$\Sigma_{2,2}^1 = \sum_{l^1=0}^{k^1-1} \sum_{l^{2,3}=0}^{k^{2,3}-1} \sum_{K \in \mathcal{D}_\lambda} \sum_{L^{(l^1, l^2, l^3)}=K} \sum_{\substack{I_3 \in \mathcal{D}_Z \\ I_3^{(k)}=K}} \Big(\sum_{\substack{I_1 \\ I_1^{(k)}=K}} \sum_{\substack{I_2 \subset L \\ I_2^{(k)}=K}} \frac{b_{K, (I_j)} \langle h_L \rangle_{I_2}}{|K|^{\frac{1}{2}}} \Big)$$
$$\langle f_1, h_K^0 \rangle \langle f_2, h_L \rangle \langle f_3, h_{I_3} \rangle.$$

Now, since we can easily check that

$$\Big| \sum_{\substack{I_1 \\ I_1^{(k)}=K}} \sum_{\substack{I_2 \subset L \\ I_2^{(k)}=K}} \frac{b_{K, (I_j)} \langle h_L \rangle_{I_2}}{|K|^{\frac{1}{2}}} \Big| \leq \frac{|K|^{\frac{1}{2}} |L|^{\frac{1}{2}} |I_3|^{\frac{1}{2}}}{|K|^2},$$

we get a sum of operators we wanted

$$\Sigma_{2,2}^1 = \sum_{l^1=0}^{k^1-1} \sum_{l^{2,3}=0}^{k^{2,3}-1} \langle S_{(0, (l^1, l^2, l^3), k)}(f_1, f_2), f_3 \rangle,$$

where $S_{(0, (l^1, l^2, l^3), k)}$ is a type of the shift (5.2). The general case Σ_{m_1, m_2}^1 is analogous.

We turn to the terms $\Sigma_{3, m_2}^1 - \Sigma_{m_2}^2$. Let us take, for example, the case $m_2 = 1$. After expanding $P_{K^{2,3}}^{k^{2,3}}$ in the first slot, $\Sigma_{3,1}^1 - \Sigma_1^2$ can be written as

$$\sum_{l^{2,3}=0}^{k^{2,3}-1} \sum_{K \in \mathcal{D}_\lambda} \sum_{(L^{2,3})^{(l^{2,3})}=K^{2,3}} \sum_{\substack{I_1, I_2, I_3 \\ I_j^{(k)}=K}} b_{K, (I_j)} \langle h_{L^{2,3}} \rangle_{I_1^{2,3}} \Big[\Big\langle f_1, \frac{1_{K^1}}{|K_1|} \otimes h_{L^{2,3}} \Big\rangle \langle f_2 \rangle_{K^1 \times I_2^{2,3}}$$
$$- \Big\langle f_1, \frac{1_{I_3^1}}{|I_3^1|} \otimes h_{L^{2,3}} \Big\rangle \langle f_2 \rangle_{I_3^1 \times I_2^{2,3}} \Big] \langle f_3, h_{I_3} \rangle.$$

For the moment, we fix one $l^{2,3}$ and write $g_1 = \langle f_1, h_{L^2} \rangle$ and $g_2 = \langle f_2 \rangle_{I_2^{2,3}}$. We write inside the brackets

$$\prod_{j=1}^{2} \langle g_j \rangle_{K^1} - \prod_{j=1}^{2} \langle g_j \rangle_{I_3^1} = -\sum_{l^1=0}^{k^1-1} \Big(\prod_{j=1}^{2} \langle g_j \rangle_{(I_3^1)^{(l^1)}} - \prod_{j=1}^{2} \langle g_j \rangle_{(I_3^1)^{(l^1+1)}} \Big)$$

and then expand $\prod_{j=1}^{2} \langle g_j \rangle_{(I_3^1)^{(l^1)}} - \prod_{j=1}^{2} \langle g_j \rangle_{(I_3^1)^{(l^1+1)}}$ as

$$\langle \Delta_{(I_3^1)^{(l^1+1)}} g_1 \rangle_{I_3^1} \langle g_2 \rangle_{(I_3^1)^{(l^1)}} + \langle g_1 \rangle_{(I_3^1)^{(l^1+1)}} \langle \Delta_{(I_3^1)^{(l^1+1)}} g_2 \rangle_{I_3^1}.$$

We get

$$\prod_{j=1}^{2} \langle g_j \rangle_{K^1} - \prod_{j=1}^{2} \langle g_j \rangle_{I_3^1}$$

$$= -\sum_{l^1=0}^{k^1-1} \Big(\langle \Delta_{(I_3^1)^{(l^1+1)}} g_1 \rangle_{I_3^1} \langle g_2 \rangle_{(I_3^1)^{(l^1)}} + \langle g_1 \rangle_{(I_3^1)^{(l^1+1)}} \langle \Delta_{(I_3^1)^{(l^1+1)}} g_2 \rangle_{I_3^1} \Big),$$

where we can expand

$$\langle \Delta_{(I_3^1)^{(l^1+1)}} g_j \rangle_{I_3^1} = \langle g_j, h_{(I_3^1)^{(l^1+1)}} \rangle \langle h_{(I_3^1)^{(l^1+1)}} \rangle_{I_3^1}.$$

For fixed l^1 and $l^{2,3}$ the expansion leads to the term

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{(L^{2,3})^{(l^{2,3})}=K^{2,3}} \sum_{\substack{I_1,I_2,I_3 \\ I_j^{(k)}=K}} b_{K,(I_j)} \langle h_{(I_3^1)^{(l^1+1)}} \otimes h_{L^{2,3}} \rangle_{I_3^1 \times I_1^{2,3}}$$

$$\Big\langle f_1, h_{(I_3^1)^{(l^1+1)}} \otimes h_{L^{2,3}} \Big\rangle \langle f_2 \rangle_{(I_3^1)^{(l^1)} \times I_2^{2,3}} \langle f_3, h_{I_3} \rangle,$$

and to the symmetrical one, where the cancellation $h_{(I_3^1)^{(l^1+1)}}$ is paired with the second function and f_1 is averaged over $(I_3^1)^{(l^1+1)}$. Again, we want to reorganize the summations and verify the correct normalization for the shifts of the form (5.2). In the first parameter we will now take $(I_3^1)^{(l^1+1)}$ as the new top cube, that is,

(5.6)
$$\sum_{K^1} \sum_{(L^1)^{(k^1-l^1)}=K^1} \sum_{K^{2,3} \in \mathcal{D}_{2^{-l^1}-k^2+k^3 \ell(L^1)}} \sum_{(I_3^1)^{(l^1)}=L^1}$$

$$\sum_{(L^{2,3})^{(l^{2,3})}=K^{2,3}} \sum_{\substack{I_2^{2,3},I_3^{2,3} \\ (I_j^i)^{(k^i)}=K^i}} c_{K^1,L^1,I_3^1,K^{2,3},L^{2,3},I_2^{2,3},I_3^{2,3}} \Big\langle f_1, h_{(L^1)^{(1)}} \otimes h_{L^{2,3}} \Big\rangle \langle f_2 \rangle_{L^1 \times I_2^{2,3}} \langle f_3, h_{I_3} \rangle,$$

where

$$c_{K^1,L^1,I_3^1,K^{2,3},L^{2,3},I_2^{2,3},I_3^{2,3}}$$

$$= \sum_{\substack{I_1^1,I_2^1 \\ (I_j^1)^{(k^1)}=K^1}} \sum_{\substack{I_1^{2,3} \subset L^{2,3} \\ (I_1^i)^{(k^i)}=K^i}} b_{K,(I_j)} \langle h_{(L^1)^{(1)} \times L^{2,3}} \rangle_{I_3^1 \times I_1^{2,3}}.$$

Moreover, we have

$$|c_{K^1,L^1,I_3^1,K^{2,3},L^{2,3},I_2^{2,3},I_3^{2,3}}| \leq \frac{|(L^1)^{(1)}|^{\frac{3}{2}}|I_3^1|^{\frac{1}{2}}}{|(L^1)^{(1)}|^2} \times \frac{|L^{2,3}|^{\frac{1}{2}}|I_2^{2,3}||I_3^{2,3}|^{\frac{1}{2}}}{|K^{2,3}|^2}.$$

Notice that this is the right normalization for (5.2), since f_2 is related to L^1 and $|(L^1)^{(1)}| = 2|L^1|$, and we can change the averages into pairings against non-cancellative Haar functions.

We conclude that for some $C \geq 1$ we have

$$C^{-1}(5.6) = \langle S_{((0,l^{2,3}),(1,k^{2,3}),(l^1+1,k^{2,3}))}(f_1,f_2), f_3 \rangle,$$

where $S_{((0,l^{2,3}),(1,k^{2,3}),(l^1+1,k^{2,3}))}$ is an operator of the desired type and of complexity

$$(0, l^{2,3}), (1, k^{2,3}), (l^1 + 1, k^{2,3}).$$

The other term and the other case of $\Sigma_{3,2}^1 - \Sigma_2^2$ are analogous.

The cases $\Sigma_{m_1,3}^1 - \Sigma_{m_1}^3$ are handled almost identically, however, we need to treat

$$\prod_{j=1}^2 \langle g_j \rangle_{K^{2,3}} - \prod_{j=1}^2 \langle g_j \rangle_{I_3^{2,3}}$$

slightly differently. We expand the rectangles $I_3^{2,3}$ in the one-parameter fashion until we reach the smaller of the cubes K^2, K^3. Then we continue with one-parameter expansion with only one of the cubes until we reach the bigger of the cubes K^2, K^3. For example, if $k^3 > k^2$, we expand as

$$\prod_{j=1}^2 \langle g_j \rangle_{K^{2,3}} - \prod_{j=1}^2 \langle g_j \rangle_{I_3^{2,3}}$$

$$= -\sum_{l^2=0}^{k^2-1} \Big[\langle \Delta_{(I_3^{2,3})^{(l^2+1,l^2+1)}} g_1 \rangle_{(I_3^{2,3})^{(l^2,l^2)}} \langle g_2 \rangle_{(I_3^{2,3})^{(l^2,l^2)}}$$

$$+ \langle g_1 \rangle_{(I_3^{2,3})^{(l^2+1,l^2+1)}} \langle \Delta_{(I_3^{2,3})^{(l^2+1,l^2+1)}} g_2 \rangle_{(I_3^{2,3})^{(l^2,l^2)}} \Big]$$

$$- \sum_{l^3=k^2}^{k^3-1} \Big[\langle E_{K^2} \Delta_{(I_3^3)^{(l^3+1)}} g_1 \rangle_{K^2 \times (I_3^3)^{(l^3)}} \langle g_2 \rangle_{K^2 \times (I_3^3)^{(l^3)}}$$

$$+ \langle g_1 \rangle_{K^2 \times (I_3^3)^{(l^3+1)}} \langle E_{K^2} \Delta_{(I^3)^{(l^3+1)}} g_2 \rangle_{K^2 \times (I_3^3)^{(l^3)}} \Big],$$

The case $k^3 \leq k^2$ can be expanded similarly. Similarly as in the previous cases, we can now write the terms in the particular form (5.2). For example, related to the latter term,

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{L^{2,3} \in \mathcal{D}_{\lambda,k}^{2,3} \ell(K^1) \\ L^2 = K^2 \\ (L^3)^{(k^3-l^3)} = K^3}} \sum_{(L^1)^{(l^1)} = K^1} \sum_{(I_3^1)^{(k^1)} = K^1} \sum_{\substack{(I_3^2)^{(k^2)} = K^2 \\ (I_3^3)^{(l^3)} = L^3}}$$

$$c_{K,L,I_3} \Big\langle f_1, \frac{1_{K^1}}{|K^1|} \otimes h_{(L^{2,3})^{(0,1)}} \Big\rangle \langle f_2, h_{L^1} \otimes \frac{1_{L^{2,3}}}{|L^{2,3}|} \rangle \langle f_3, h_{I_3} \rangle,$$

where $l^3 \in \{k^2, \ldots, k^3 - 1\}$, $\lambda_{l,k} = 2^{-k^1 - k^2 + l^3}$ and

$$|c_{K,L,I_3}| = \Big| \sum_{\substack{I_1, I_2 \\ (I_j)^{(k)} = K \\ I_2^1 \subset L^1}} a_{K,(I_j)} |I_1| \langle h_{L^1} \otimes h_{K^2 \times (L^3)^{(1)}} \rangle_{I_2^1 \times K^2 \times L^3} \Big|$$

$$\leq \sum_{\substack{I_1, I_2 \\ (I_j)^{(k)} = K \\ I_2^1 \subset L^1}} \frac{|I_3|^{\frac{1}{2}} |I_1| |I_2|}{|K|^2} |K^2|^{-\frac{1}{2}} |(L^3)^{(1)}|^{-\frac{1}{2}} |L^1|^{-\frac{1}{2}}$$

$$= \frac{|L^1|^{\frac{1}{2}}}{|K^1|} \cdot \frac{|I_3|^{\frac{1}{2}} |K^2 \times (L^3)^{(1)}|^{\frac{3}{2}}}{|K^2 \times (L^3)^{(1)}|^2}.$$

This normalization is an absolute constant away from the correct one since we consider that $K^2 \times (L^3)^{(1)}$ is the top rectangle in parameters 2 and 3.

Finally, we consider $\Sigma_{3,3}^1 - \Sigma_3^2 - \Sigma_3^3 + \Sigma^4$ that equals to

(5.7)
$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I_1, I_2, I_3 \in \mathcal{D}_Z \\ I_j^{(k)} = K}} b_{K,(I_j)}$$

$$\Big[\prod_{j=1}^2 \langle f_j \rangle_K - \prod_{j=1}^2 \langle f_j \rangle_{I_3^1 \times K^{2,3}} - \prod_{j=1}^2 \langle f_j \rangle_{K^1 \times I_3^{2,3}} + \prod_{j=1}^2 \langle f_j \rangle_{I_3} \Big] \langle f_3, h_{I_3} \rangle.$$

As we already showed, we can expand

$$\prod_{j=1}^2 \langle f_j \rangle_K - \prod_{j=1}^2 \langle f_j \rangle_{I_3^1 \times K^{2,3}}$$

$$= - \sum_{l^1=0}^{k^1-1} \Big(\langle \Delta_{(I_3^1)^{(l^1+1)}} g_1 \rangle_{I_3^1} \langle g_2 \rangle_{(I_3^1)^{(l^1)}} + \langle g_1 \rangle_{(I_3^1)^{(l^1+1)}} \langle \Delta_{(I_3^1)^{(l^1+1)}} g_2 \rangle_{I_3^1} \Big),$$

where $g_j = \langle f_j \rangle_{K^{2,3}}$, and similarly for

$$\prod_{j=1}^n \langle f_j \rangle_{I_3} - \prod_{j=1}^2 \langle f_j \rangle_{K^1 \times I_3^{2,3}}$$

we get same expansion with the positive sign and $g_j = \langle f_j \rangle_{I_3^{2,3}}$.

Then we sum the two expansions together and expand in the parameters 2 and 3. That is, we will expand

$$\sum_{l^1=0}^{k^1-1} \langle h_{(I_3^1)^{(l^1+1)}} \rangle_{(I_3^1)^{(l^1)}} \Big\langle f_1, h_{(I_3^1)^{(l^1+1)}} \otimes \frac{1_{K^{2,3}}}{|K^{2,3}|} \Big\rangle \langle f_2 \rangle_{(I_3^1)^{(l^1)} \times K^{2,3}}$$

$$- \Big\langle f_1, h_{(I_3^1)^{(l^1+1)}} \otimes \frac{1_{I_3^{2,3}}}{|I_3^{2,3}|} \Big\rangle \langle f_2 \rangle_{(I_3^1)^{(l^1)} \times I_3^{2,3}}.$$

Thus, we get, for example when $k^2 < k^3$, that

$$
\sum_{l^1=0}^{k^1-1} \sum_{l^2=0}^{k^2-1} \sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{L^1 \in \mathcal{D}^1 \\ (L^1)^{(k^1-l^1)}=K^1}} \sum_{\substack{L^{2,3} \in \mathcal{D}^{2,3}_{2^{-l^1}\ell(L^1)} \\ (L^{2,3})^{(k^2-l^2,k^3-l^2)}=K^{2,3}}} \sum_{\substack{I_3 \in \mathcal{D}_Z \\ (I_3)^{(l^1,l^2,l^2)}=L}}
$$
$$
\times c_{K,L,I_3} \Big\langle f_1, h_{(L^1)^{(1)}} \otimes h^0_{(L^{2,3})^{(1,1)}} \Big\rangle \Big\langle f_2, h^0_{L^1} \otimes h_{(L^{2,3})^{(1,1)}} \Big\rangle \langle f_3, h_{I^3} \rangle
$$
$$
+ \sum_{l^1=0}^{k^1-1} \sum_{l^3=k^2}^{k^3-1} \sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{L^1 \in \mathcal{D}^1 \\ (L^1)^{(k^1-l^1)}=K^1}} \sum_{\substack{L^{2,3} \in \mathcal{D}_{2^{-l^1-k^2+l^3}\ell(L^1)} \\ L^2=K^2 \\ (L^3)^{(k^3-l^3)}=K^3}} \sum_{\substack{I_3 \in \mathcal{D}_Z \\ (I_3)^{(l^1,k^2,l^3)}=L}}
$$
$$
\times c_{K,L,I_3} \Big\langle f_1, h_{(L^1)^{(1)}} \otimes h^0_{(L^{2,3})^{(0,1)}} \Big\rangle \Big\langle f_2, h^0_{L^1} \otimes h_{(L^{2,3})^{(0,1)}} \Big\rangle \langle f_3, h_{I^3} \rangle.
$$

Here

$$
|c_{K,L,I_3}| = \Big| \sum_{\substack{I_1,I_2 \in \mathcal{D}_Z \\ I_j^k=K}} a_{K,(I_j)} |I_1| |L^1|^{-\frac{1}{2}} |(L^{2,3})^{(1)}|^{-\frac{1}{2}} \langle h_{(L^1)^{(1)}} \otimes h_{(L^{2,3})^{(1)}} \rangle_{L^1 \times L^{2,3}} \Big|
$$
$$
\leq \frac{|I_3|^{\frac{1}{2}} |(L^1)^{(1)}|^{\frac{3}{2}}}{|(L)^{(1)}|^2} |L^1|^{-\frac{1}{2}} |(L^{2,3})^{(1)}|^{-\frac{1}{2}} \sim \frac{|I_3|^{\frac{1}{2}} |(L^1)^{(1)}|^{\frac{1}{2}} |L^1|^{\frac{1}{2}} |(L^{2,3})^{(1)}|^{\frac{3}{2}}}{|(L^1)^{(1)}|^2 |(L^{2,3})^{(1)}|^2}.
$$

We abused notation slightly by $(L^{2,3})^{(1)}$ meaning both $(L^{2,3})^{(1,1)}$ and $(L^{2,3})^{(0,1)}$. The other terms are handled analogously. $\qquad \square$

6. BOUNDEDNESS OF ZYGMUND SHIFTS

In this section we prove the boundedness of Zygmund shifts. We first prove the following. A collection $\mathscr{S}$ is called γ-sparse if there are pairwise disjoint subsets $E(S) \subset S$, $S \in \mathscr{S}$, with $|E(S)| \geq \gamma|S|$. Often the precise value of γ is not important and we just talk about sparse collections.

6.1. **Proposition.** *Let* $\lambda = 2^k$ *for some* $k \in \mathbb{Z}$ *and*

$$
\Lambda(f_1, f_2, f_3) = \sum_{K \in \mathcal{D}^{2,3}_\lambda} \sum_{(I_j)^{(\ell_j)}=K} \frac{\prod_{j=1}^3 |I_j|^{\frac{1}{2}}}{|K|^2} |\langle f_1, h^0_{I_1} \rangle| \cdot |\langle f_2, h_{I_2} \rangle| \cdot |\langle f_3, h_{I_3} \rangle|.
$$

Then there exists a sparse collection $\mathcal{S} \subset \mathcal{D}^{2,3}_\lambda$ *such that*

$$
\Lambda(f_1, f_2, f_3) \lesssim \max\{k^2, k^3\} \sum_{S \in \mathcal{S}} |S| \prod_{j=1}^3 \langle |f_j| \rangle_S.
$$

Proof. The proof is an easy adaptation of the sparseness argument in [17, Section 5]. In fact, we only need to check the validity of

$$
\Lambda(f_1, f_2, f_3) \lesssim \|f_1\|_{L^p} \|f_2\|_{L^q} \|f_3\|_{L^r},
$$

where $p, q, r \in (1, \infty)$ and $1/p + 1/q + 1/r = 1$. This can be done by direct computation:

$$
\begin{aligned}
\Lambda(f_1, f_2, f_3) &\leq \int f_1 \sum_{K \in \mathcal{D}_\lambda^{2,3}} \langle |\Delta_K^{\ell_2} f_2| \rangle_K \langle |\Delta_K^{\ell_3} f_3| \rangle_K 1_K \\
&\leq \|f_1\|_{L^p} \Big\| \Big(\sum_{K \in \mathcal{D}_\lambda^{2,3}} [M_{\mathcal{D}_\lambda^{2,3}} |\Delta_K^{\ell_2} f_2|]^2 \Big)^{\frac{1}{2}} \Big\|_{L^q} \Big\| \Big(\sum_{K \in \mathcal{D}_\lambda^{2,3}} [M_{\mathcal{D}_\lambda^{2,3}} |\Delta_K^{\ell_3} f_3|]^2 \Big)^{\frac{1}{2}} \Big\|_{L^r} \\
&\lesssim \|f_1\|_{L^p} \|f_2\|_{L^q} \|f_3\|_{L^r}.
\end{aligned}
$$

$\square$

6.2. Proposition. *Let Q_k, $k = (k^1, k^2, k^3)$, be a bilinear Zygmund shift as in Section 2.D, and let $1 < p_1, p_2 < \infty$ and $\frac{1}{2} < p < \infty$ with $\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2}$. Let*

$$
w_1, w_2 \in A_p(\mathbb{R} \times \mathbb{R} \times \mathbb{R}), \qquad \text{and} \qquad w := w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}}.
$$

Then, for every $\eta \in (0, 1)$ we have

$$
\|Q_k(f_1, f_2)\|_{L^p(w)} \lesssim \max_i \{k^i\}^2 2^{k^1 \eta} \|f_1\|_{L^{p_1}(w_1)} \|f_2\|_{L^{p_2}(w_2)}.
$$

Proof. We prove the weighted boundedness $L^4(w_1) \times L^4(w_2) \to L^2(w)$, of the tri-parameter bilinear shifts of Zygmund nature (5.2). We do this with tri-parameter weights $w_i \in A_4$. We then extrapolate the result to the full bilinear range using the traditional multilinear extrapolation by Grafakos–Martell (and Duoandikoetxea) [4, 7]. Our result then follows from Proposition 5.3.

Note that if we have $I_3 \in \mathcal{D}_Z$ in (5.2), then the related λ in Proposition 6.1 is

$$
2^{\ell_3^3 - \ell_3^2 - \ell_3^1} |L^1|.
$$

(For other cases, for instance if $I_1^1 \times I_2^{2,3} \in \mathcal{D}_Z$, then $\lambda = 2^{\ell_2^3 - \ell_2^2 - \ell_1^1} |L^1|$). Assume $v \in A_{4,\lambda}(\mathbb{R}^2)$; recall that $A_{p,\lambda}(\mathbb{R}^2)$ is defined similarly as $A_p(\mathbb{R}^2)$ except that the supremum is taken over rectangles $R = I \times J$ with $|J| = \lambda |I|$. Then

$$
\sum_{S \in \mathcal{S}} |S| \prod_{j=1}^3 \langle |f_j| \rangle_S = \sum_{S \in \mathcal{S}} \langle |f_1| \rangle_S \langle |f_2| \rangle_S \langle |f_3| v^{-1} \rangle_S^v v(S).
$$

Since for any $R \in \mathcal{S}$,

$$
\sum_{\substack{S \subset R \\ S \in \mathcal{S}}} v(S) = \sum_{\substack{S \subset R \\ S \in \mathcal{S}}} \frac{v(S)}{|S|} |S| \lesssim \sum_{\substack{S \subset R \\ S \in \mathcal{S}}} \frac{v(S)}{|S|} |E_S| \leq \int_R M_{\mathcal{D}_\lambda^{2,3}}(v 1_R) \lesssim_{[v]_{A_{4,\lambda}(\mathbb{R}^2)}} v(R),
$$

by the Carleson embedding theorem we have

$$
(6.3) \qquad \sum_{S \in \mathcal{S}} |S| \prod_{j=1}^3 \langle |f_j| \rangle_S \lesssim_{[v]_{A_{4,\lambda}(\mathbb{R}^2)}} \int_{\mathbb{R}^2} M_{\mathcal{D}_\lambda^{2,3}} |f_1| M_{\mathcal{D}_\lambda^{2,3}} |f_2| M_{\mathcal{D}_\lambda^{2,3}}^v(|f_3| v^{-1}) v.
$$

Now, given weights $w_j \in A_4(\mathbb{R}^3)$, $j = 1, 2$, we know that $w = w_1^{1/2} w_2^{1/2} \in A_4(\mathbb{R}^3)$. We have

$$|\langle S(f_1, f_2), f_3 \rangle| = \sum_{L^1} \sum_{(I_j^1)^{(\ell_j^1)} = L^1} \frac{\prod_{j=1}^3 |I_j^1|^{\frac{1}{2}}}{|L^1|^2} \Lambda(\langle f_1, h_{I_1^1} \rangle, \langle f_2, h_{I_2^1}^0 \rangle, \langle f_3, h_{I_3^1} \rangle).$$

Note that $\langle w \rangle_{L^1} \in A_{4,\lambda}(\mathbb{R}^2)$ with $[\langle w \rangle_{L^1}]_{A_{4,\lambda}(\mathbb{R}^2)} \le [w]_{A_4}$ for any λ. Thus, applying (6.3) with $v = \langle w \rangle_{L^1}$ we have

$|\langle S(f_1, f_2), f_3 \rangle|$

$$\lesssim \max_i \{k^i\} \sum_{L^1} \sum_{(I_j^1)^{(\ell_j^1)} = L^1} \frac{\prod_{j=1}^3 |I_j^1|^{\frac{1}{2}}}{|L^1|^2} \int_{\mathbb{R}^2} M_{\mathcal{D}_\lambda^{2,3}} \langle f_1, h_{I_1^1} \rangle M_{\mathcal{D}_\lambda^{2,3}} \langle f_2, h_{I_2^1}^0 \rangle M_{\mathcal{D}_\lambda^{2,3}}^v (\langle f_3, h_{I_3^1} \rangle v^{-1}) v$$

$$= \max_i \{k^i\} \sum_{L^1} \int_{\mathbb{R}^3} \langle M_{\mathcal{D}} |\Delta_{L^1}^{\ell_1^1} f_1| \rangle_{L^1} \langle M_{\mathcal{D}} |f_2| \rangle_{L^1} \sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{1}{2}} M_{\mathcal{D}_\lambda^{2,3}}^{\langle w \rangle_{L^1}} (\langle f_3, h_{I_3^1} \rangle \langle w \rangle_{L^1}^{-1}) \frac{1_{L^1}}{|L^1|} w$$

$$\le \max_i \{k^i\} \left\| \Big(\sum_{L^1} \big[M_{\mathcal{D}^1} M_{\mathcal{D}} |\Delta_{L^1}^{\ell_1^1} f_1| \big]^2 \Big)^{\frac{1}{2}} \right\|_{L^4(w_1)} \| M_{\mathcal{D}^1} M_{\mathcal{D}} |f_2| \|_{L^4(w_2)}$$

$$\times \left\| \Big(\sum_{L^1} \big[\sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{1}{2}} M_{\mathcal{D}_\lambda^{2,3}}^{\langle w \rangle_{L^1}} (\langle f_3, h_{I_3^1} \rangle \langle w \rangle_{L^1}^{-1}) |L^1|^{-1} \big]^2 1_{L^1} \Big)^{\frac{1}{2}} \right\|_{L^2(w)}.$$

By the well-know square function and maximal function estimates we have

$$\left\| \Big(\sum_{L^1} \big[M_{\mathcal{D}^1} M_{\mathcal{D}} |\Delta_{L^1}^{\ell_1^1} f_1| \big]^2 \Big)^{\frac{1}{2}} \right\|_{L^4(w_1)} \lesssim \| f_1 \|_{L^4(w_1)}$$

and

$$\| M_{\mathcal{D}^1} M_{\mathcal{D}} |f_2| \|_{L^4(w_2)} \lesssim \| f_2 \|_{L^4(w_2)}.$$

The estimate of the last term is a bit tricky. By the (one parameter)vector-valued estimates of $M_{\mathcal{D}_\lambda^{2,3}}^{\langle w \rangle_{L^1}}$ (see e.g. [19, Proposition 4.3] for a bi-parameter version (the proof easily adapts to the one-parameter case)), we have

$$\left\| \Big(\sum_{L^1} \big[\sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{1}{2}} M_{\mathcal{D}_\lambda^{2,3}}^{\langle w \rangle_{L^1}} (\langle f_3, h_{I_3^1} \rangle \langle w \rangle_{L^1}^{-1}) |L^1|^{-1} \big]^2 1_{L^1} \Big)^{\frac{1}{2}} \right\|_{L^2(w)}$$

$$\le 2^{\ell_3^1 \eta} \left\| \Big(\sum_{L^1} \big[\sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{s}{2}} M_{\mathcal{D}_\lambda^{2,3}}^{\langle w \rangle_{L^1}} (\langle f_3, h_{I_3^1} \rangle \langle w \rangle_{L^1}^{-1})^s |L^1|^{-\frac{s}{2}} \big]^{\frac{2}{s}} \Big)^{\frac{1}{2}} \right\|_{L^2(\langle w \rangle_{L^1})}$$

$$\lesssim 2^{\ell_3^1 \eta} \left\| \Big(\sum_{L^1} \big[\sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{s}{2}} |\langle f_3, h_{I_3^1} \rangle \langle w \rangle_{L^1}^{-1}|^s |L^1|^{-\frac{s}{2}} \big]^{\frac{2}{s}} \Big)^{\frac{1}{2}} \right\|_{L^2(\langle w \rangle_{L^1})}$$

$$\le 2^{\ell_3^1 \eta} \left\| \Big(\sum_{L^1} \big[\sum_{(I_3^1)^{(\ell_3^1)} = L^1} |I_3^1|^{\frac{1}{2}} |\langle f_3, h_{I_3^1} \rangle| \langle w \rangle_{L^1}^{-1} |L^1|^{-\frac{1}{2}} \big]^2 \Big)^{\frac{1}{2}} \right\|_{L^2(\langle w \rangle_{L^1})}$$

$$\lesssim 2^{\ell_3^1 \eta} \|f_3\|_{L^2(w^{-1})},$$

where $s = (1/\eta)'$ and in the last step we have used [19, Proposition 5.8]. Thus,

$$\|S(f_1, f_2)\|_{L^2(w)} \lesssim \max_i \{k^i\} 2^{k^1 \eta} \|f_1\|_{L^4(w_1)} \|f_2\|_{L^4(w_2)}.$$

$\square$

Now we are able to conclude the proof of Theorem 1.2.

Proof of Theorem 1.2. By the representation formula discussed in Sections 2.E and 2.F, the coefficient estimates in Section 4 (in particular (4.1)) we get that

$$\langle T(f_1, f_2), f_3 \rangle = C \mathbb{E}_\sigma \sum_{k^1, k^2, k^3 = 2}^\infty (|k| + 1)^2 \varphi(k) \sum_{I \in \mathcal{D}_Z(k)} \frac{\langle Q_{(k^1, k^2, k^3)}(f_1, f_2), f_3 \rangle}{C(|k| + 1)^2 \varphi(k)}.$$

Thus, for $p_1, p_2 \in (1, \infty)$ so that $p \in (1, \infty)$, we conclude by Proposition 6.2 that

$$\|T(f_1, f_2)\|_{L^p(w)} \lesssim \sum_{k^1, k^2, k^3 = 2}^\infty (|k| + 1)^2 \varphi(k) \max_i \{k^i\} 2^{k^1 \eta} \|f_1\|_{L^{p_1}(w_1)} \|f_2\|_{L^{p_2}(w_2)}$$

$$\lesssim \|f_1\|_{L^{p_1}(w_1)} \|f_2\|_{L^{p_2}(w_2)},$$

where we need to take $\eta < \alpha_1$. Consequently, we can now pass the result to the full bilinear range using the traditional multilinear extrapolation [4,7]. $\square$

7. LINEAR COMMUTATORS IN THE ZYGMUND DILATION SETTING

In this section we return to the linear theory and complete the following commutator estimate left open by previous results. This requires new and interesting paraproduct estimates. For the context, see the explanation below.

7.1. Theorem. *Let $b \in L^1_{loc}$ and T be a linear CZZ operator as in [14]. Let $\theta \in (0, 1]$ be the kernel exponent measuring the decay in terms of the Zygmund ratio*

$$D_\theta(x) := \Big(\frac{|x_1 x_2|}{|x_3|} + \frac{|x_3|}{|x_1 x_2|} \Big)^{-\theta}.$$

Then

$$\|[b, T]\|_{L^p \to L^p} \lesssim \|b\|_{\mathrm{bmo}_Z}$$

whenever $p \in (1, \infty)$.

Here the definition of the little BMO is given by

$$\|b\|_{\mathrm{bmo}_Z} := \sup_{\mathcal{D}_Z} \sup_{R \in \mathcal{D}_Z} \frac{1}{|R|} \int_R |b(x) - \langle b \rangle_R| \, \mathrm{d}x < \infty,$$

where the supremum is over all different collections of Zygmund rectangles $\mathcal{D}_Z$ and then over all $R \in \mathcal{D}_Z$.

This theorem was previously considered in [5] using the so-called Cauchy trick. That method requires weighted bounds with Zygmund weights. But we now know [14] how delicate such weighted bounds are – weighted bounds with Zygmund weights do not in general hold if $\theta < 1$. However, the commutator bounds are still true – but we need a different proof, presented here. It suffices to prove the boundedness of commutators $[b, Q_k]$ for any linear shift Q_k of the Zygmund dilation type.

For $\theta = 1$ we could use the Cauchy trick and the weighted bounds from [14] – this would give weighted commutator estimates with Zygmund weights.

We begin by recording lemmas that we need for the main proofs of this section.

7.2. Lemma. *Let b be a locally integrable function. Then the following are equivalent*

(1) $b \in \mathrm{bmo}_{\mathcal{D}_Z}$,

(2)
$$\max\left\{ \sup_{I^1 \in \mathcal{D}^1} \|\langle b \rangle_{I^1,1}\|_{\mathrm{BMO}_{\mathcal{D}^{2,3}_{\ell(I^1)}}}, \ \operatorname*{ess\,sup}_{(x_2,x_3)\in\mathbb{R}^2} \|b(\cdot, x_2, x_3)\|_{\mathrm{BMO}} \right\} < \infty,$$

(3)
$$\max\left\{ \sup_{I^2 \in \mathcal{D}^2} \|\langle b \rangle_{I^2,2}\|_{\mathrm{BMO}_{\mathcal{D}^{2,3}_{\ell(I^2)}}}, \ \operatorname*{ess\,sup}_{(x_1,x_3)\in\mathbb{R}^2} \|b(x_1, \cdot, x_3)\|_{\mathrm{BMO}} \right\} < \infty.$$

For completeness, we give the proof.

Proof. Let us begin showing that $\mathrm{bmo}_Z \implies$ (2) (and by symmetry also (3)). Clearly, for all Zygmund rectangles $I = I^1 \times I^2 \times I^3 \in \mathcal{D}_Z$ we have

(7.3)
$$\|b\|_{\mathrm{bmo}_Z} \geq \frac{1}{|I|} \int_I |b - \langle b \rangle_I| \geq \frac{1}{|I^{2,3}|} \int_{I^{2,3}} |\langle b \rangle_{I^1,1} - \langle b \rangle_I|.$$

So by uniform boundedness we immediately get

$$\|\langle b \rangle_{I^1,1}\|_{\mathrm{BMO}_{\mathcal{D}^{2,3}_{\ell(I^1)}}} := \sup_{I^{2,3}\in\mathcal{D}^{2,3}_{\ell(I^1)}} \frac{1}{|I^{2,3}|} \int_{I^{2,3}} |\langle b \rangle_{I^1,1} - \langle \langle b \rangle_{I^1,1}\rangle_{I^{2,3}}| \leq \|b\|_{\mathrm{bmo}_Z} < \infty.$$

We move on to proving the second assertion inside (2). For fixed $I^1 \in \mathcal{D}^1$ we define $f_{I^1}(x^2, x^3) := \int_{I^1} |b(x^1, x^2, x^3) - \langle b \rangle_{I^1}(x^2, x^3)| \, dx^1$. Then for every $I^{2,3} \in \mathcal{D}^{2,3}_{\ell(I^1)}$ we have

$$\langle f_{I^1}\rangle_{I^{2,3}} \leq \frac{1}{|I^{2,3}|} \int_{I^{2,3}} \int_{I^1} |b - \langle b \rangle_I| + \frac{1}{|I^{2,3}|} \int_{I^{2,3}} \int_{I^1} |\langle b \rangle_{I^1,1} - \langle b \rangle_I| \leq 2|I^1| \|b\|_{\mathrm{bmo}_Z},$$

where last inequality holds by definition and the above estimate (7.3). Now, by the Lebesgue differentiation theorem we get for $(x^2, x^3) \in \mathbb{R}^2 \setminus N(I^1)$, where $N(I^1)$ is a null set depending on I^1, that

$$f_{I^1}(x^2, x^3) \leq 2|I^1| \|b\|_{\mathrm{bmo}_Z}.$$

It is then easy to conclude that

$$\|b(\cdot, x_2, x_3)\|_{\mathrm{BMO}} \leq 2\|b\|_{\mathrm{bmo}_Z}$$

for almost every $(x^2, x^3) \in \mathbb{R}^2$.

Conversely,

$$\int_I |b - \langle b \rangle_I| \leq \int_I |b - \langle b \rangle_{I^1,1}| + \int_I |\langle b \rangle_{I^1,1} - \langle b \rangle_I|$$

$$\leq |I^1| \int_{I^{2,3}} \|b(\cdot, x^2, x^3)\|_{\mathrm{BMO}} + |I| \|\langle b \rangle_{I^1,1}\|_{\mathrm{BMO}_{\ell(I^1)}} \leq |I|(C_1 + C_2),$$

where $C^1 := \operatorname*{ess\,sup}_{(x^2,x^3)\in\mathbb{R}^2} \|b(\cdot, x^2, x^3)\|_{\mathrm{BMO}}$ and $C_2 := \sup_{I^1} \|\langle b \rangle_{I^1,1}\|_{\mathrm{BMO}_{\ell(I^1)}}$.   □

Then the usual duality results imply the following.

7.4. Corollary. *If $b \in \mathrm{bmo}_Z$ and I^1 is fixed, then*

$$\sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}^{2,3}} \langle \langle b \rangle_{I^1}, h_{I^{2,3}} \rangle \varphi_{I^{2,3}} \lesssim \|b\|_{\mathrm{bmo}_Z} \left\| \left(\sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}^{2,3}} |\varphi_{I^{2,3}}|^2 \frac{1_{I^{2,3}}}{|I^{2,3}|} \right)^{\frac{1}{2}} \right\|_{L^1}.$$

Also, for fixed (x_2, x_3), we have

$$\sum_{I^1 \in \mathcal{D}^1} \langle b, h_{I^1} \rangle_1 \varphi_{I^1} \lesssim \|b\|_{\mathrm{bmo}_Z} \left\| \left(\sum_{I^1 \in \mathcal{D}^1} |\varphi_{I^1}|^2 \frac{1_{I^1}}{|I^1|} \right)^{\frac{1}{2}} \right\|_{L^1}.$$

Using the duality type estimates we can use the square function lower bounds to prove the inclusion of product type spaces.

7.5. Definition. Given a lattice of Zygmund rectangles $\mathcal{D}_Z$ and a sequence of scalars $B = (b_I)_{I \in \mathcal{D}_Z}$ we define

$$\|B\|_{\mathrm{BMO}_{\mathrm{prod}}} := \sup_{\Omega} \left(\frac{1}{|\Omega|} \sum_{\substack{I \in \mathcal{D}_Z \\ I \subset \Omega}} |b_I|^2 \right)^{\frac{1}{2}}.$$

The inclusion of the little BMO space can be easily seen from the duality estimate

$$(7.6) \qquad \|B\|_{\mathrm{BMO}_{\mathrm{prod}}} \sim \sup \left\{ \sum_{I \in \mathcal{D}_Z} |a_I| |b_I| : \left\| \left(\sum_{I \in \mathcal{D}_Z} |a_I|^2 \frac{1_I}{|I|} \right)^{\frac{1}{2}} \right\|_{L^1} \leq 1 \right\}.$$

7.A. Paraproduct expansions. Here the correct expansions style is the Zygmund martingale expansion similar to [14, Equation (5.22)]. This gives

$$(7.7) \qquad bf = \sum_{I \in \mathcal{D}_Z} \Big[\Delta_{I,Z} b \Delta_{I,Z} f + \Delta_{I,Z} b \Delta_{I^1} E_{I^{2,3}} f + \Delta_{I^1} E_{I^{2,3}} b \Delta_{I,Z} f$$

$$+ \Delta_{I,Z} b E_{I^1} \Delta_{I^{2,3}} f + \Delta_{I,Z} b E_{I^1} E_{I^{2,3}} f + \Delta_{I^1} E_{I^{2,3}} b E_{I^1} \Delta_{I^{2,3}} f$$

$$+ E_{I^1} \Delta_{I^{2,3}} b \Delta_{I,Z} f + E_{I^1} \Delta_{I^{2,3}} b \Delta_{I^1} E_{I^{2,3}} f + E_{I^1} E_{I^{2,3}} b \Delta_{I,Z} f \Big]$$

$$=: \sum_{i,j=1}^{3} a_{i,j}(b, f),$$

where, for example, $a_{1,1} = \sum_{I \in \mathcal{D}_Z} \Delta_{I,Z} b \Delta_{I,Z} f$ and

$$a_{1,2} = \sum_{I \in \mathcal{D}_Z} \Delta_{I,Z} b \Delta_{I^1} E_{I^{2,3}} f,$$

i.e., interpret so that rows correspond to the first index i and columns correspond with the second index j.

7.8. Lemma. *If $b \in \mathrm{bmo}_Z$, then the paraproducts $a_{i,j}$ such that $(i,j) \neq (3,3)$ are bounded. That is,*

$$\|a_{i,j}(b, f)\|_{L^p} \lesssim \|b\|_{\mathrm{bmo}_Z} \|f\|_{L^p}, \qquad 1 < p < \infty.$$

Proof. Case 1: product type $i \neq 3 \neq j$. We begin with the paraproducts where it would suffice to have a product BMO type assumption (but recall that little BMO is a subset). The symmetry $\Pi = a_{1,1}$ is essentially trivial. By (7.6) we have

$$
|\langle \Pi f, g \rangle| \lesssim \Big\| \Big(\sum_{I \in \mathcal{D}_Z} |\langle f, h_{I,Z} \rangle|^2 \langle |g| \rangle_I^2 \frac{1_I}{|I|} \Big)^{\frac{1}{2}} \Big\|_{L^1}
$$

$$
\leq \Big\| \Big(\sum_{I \in \mathcal{D}_Z} \langle |\Delta_{I,Z} f| \rangle_I^2 1_I \Big)^{\frac{1}{2}} \Big\|_{L^p} \| M_Z g \|_{L^{p'}}
$$

$$
\lesssim \Big\| \Big(\sum_{I \in \mathcal{D}_Z} M_Z(\Delta_{I,Z} f)^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \| g \|_{L^{p'}}
$$

$$
\lesssim \| S_Z f \|_{L^p} \| g \|_{L^{p'}} \lesssim \| f \|_{L^p} \| g \|_{L^{p'}}.
$$

The 'twisted' case $\Pi = a_{1,2}$ (and the symmetrical $a_{2,1}$) is trickier. Indeed, to decouple f and g we cannot blindly take maximal functions only in some parameters – this would break the Zygmund structure. In any case, we begin with the application of (7.6) to get

$$
|\langle \Pi f, g \rangle| \lesssim \Big\| \Big(\sum_{I \in \mathcal{D}_Z} \Big| \Big\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \Big\rangle \Big\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \Big\rangle \Big|^2 \frac{1_I}{|I|} \Big)^{\frac{1}{2}} \Big\|_{L^1}.
$$

The above is an L^1 norm, while L^2 would be nice. This is where A_∞ extrapolation comes in. We fix $\nu \in A_{\infty,Z}$, and move to estimate

$$
\Big\| \Big(\sum_{I \in \mathcal{D}_Z} \Big| \Big\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \Big\rangle \Big\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \Big\rangle \Big|^2 \frac{1_I}{|I|} \Big)^{\frac{1}{2}} \Big\|_{L^2(\nu)}.
$$

We will soon show that

(7.9)
$$
\Big\| \Big(\sum_{I \in \mathcal{D}_Z} \Big| \Big\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \Big\rangle \Big\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \Big\rangle \Big|^2 \frac{1_I}{|I|} \Big)^{\frac{1}{2}} \Big\|_{L^2(\nu)}
$$
$$
\lesssim \Big\| M_Z f \Big(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} g)^2 \Big)^{1/2} \Big\|_{L^2(\nu)}.
$$

The A_∞ extrapolation, Theorem 7.10, then implies that this inequality holds also in $L^p(\nu)$, $p \in (0, \infty)$, $\nu \in A_{\infty,Z}$. We take $p = 1$ and $\nu \equiv 1$ to get that

$$
|\langle \Pi f, g \rangle| \lesssim \Big\| M_Z f \Big(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} g)^2 \Big)^{1/2} \Big\|_{L^1}
$$

$$
\leq \| M_Z f \|_{L^p} \Big\| \Big(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} g)^2 \Big)^{1/2} \Big\|_{L^{p'}} \lesssim \| f \|_{L^p} \| g \|_{L^{p'}}.
$$

It remains to prove (7.9). We write

$$
\Big\| \Big(\sum_{I \in \mathcal{D}_Z} \Big| \Big\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \Big\rangle \Big\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \Big\rangle \Big|^2 \frac{1_I}{|I|} \Big)^{\frac{1}{2}} \Big\|_{L^2(\nu)}^2
$$

$$
= \sum_{I^1 \in \mathcal{D}^1} \sum_{I^2 \times I^3 \in \mathcal{D}_{\ell(I^1)}^{2,3}} \Big| \Big\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \Big\rangle \Big|^2 \Big| \Big\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \Big\rangle \Big|^2 \langle \nu \rangle_I.
$$

Fix some $I^1 \in \mathcal{D}^1$. Let $I_0^2 \times I_0^3 \in \mathcal{D}^{2,3}_{\ell(I^1)}$ and suppose φ_1, φ_2 and φ_3 are locally integrable functions in $\mathbb{R}^2$. Then, there exists a sparse collection $\mathcal{S} = \mathcal{S}(I_0^2 \times I_0^3, \varphi_1, \varphi_2, \varphi_3) \subset \mathcal{D}^{2,3}_{\ell(I^1)}(I_0^2 \times I_0^3)$ so that

$$\sum_{\substack{I^2 \times I^3 \in \mathcal{D}^{2,3}_{\ell(I^1)} \\ I^2 \times I^3 \subset I_0^2 \times I_0^3}} |\langle \varphi_1, h_{I^2 \times I^3} \rangle|^2 |\langle \varphi_2 \rangle_{I^2 \times I^3}|^2 \langle \varphi_3 \rangle_{I^2 \times I^3} \lesssim \sum_{Q \in \mathcal{S}} \langle |\varphi_1| \rangle_Q^2 \langle |\varphi_2| \rangle_Q^2 \langle |\varphi_3| \rangle_Q |Q|.$$

We use this with the functions $\varphi_1 = \langle f \rangle_{I^1}$, $\varphi_2 = \langle g, h_{I^1} \rangle$ and $\varphi_3 = \langle \nu \rangle_{I^1}$ to have that for some sparse collection $\mathcal{S} = \mathcal{S}(I^1, I_0^2 \times I_0^3, f, g, \nu) \subset \mathcal{D}^{2,3}_{\ell(I^1)}$ there holds that

$$\sum_{\substack{I^2 \times I^3 \in \mathcal{D}^{2,3}_{\ell(I^1)} \\ I^2 \times I^3 \subset I_0^2 \times I_0^3}} \left| \left\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \right\rangle \right|^2 \left| \left\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \right\rangle \right|^2 \langle \nu \rangle_I$$

$$\lesssim \sum_{Q \in \mathcal{S}} \langle |\langle f \rangle_{I^1}| \rangle_Q^2 \langle |\langle g, h_{I^1} \rangle| \rangle_Q^2 \langle \nu \rangle_{I^1}(Q)$$

$$\leq \sum_{Q \in \mathcal{S}} \left(\left\langle \left(M^{2,3}_{\ell(I^1)} \langle f \rangle_{I^1} \right) \left(M^{2,3}_{\ell(I^1)} \langle g, h_{I^1} \rangle \right) \right\rangle_Q^{\langle \nu \rangle_{I^1}} \right)^2 \langle \nu \rangle_{I^1}(Q)$$

$$\lesssim \int_{\mathbb{R}^2} \left(M^{2,3}_{\ell(I^1)} \langle f \rangle_{I^1} \right)^2 \left(M^{2,3}_{\ell(I^1)} \langle g, h_{I^1} \rangle \right)^2 \langle \nu \rangle_{I^1},$$

where in the last step we used the fact that $\langle \nu \rangle_{I^1} \in A_{\infty, \ell(I^1)}(\mathbb{R}^2)$ and the Carleson embedding theorem.

Since the last estimate holds uniformly for every $I_0^2 \times I_0^3 \in \mathcal{D}^{2,3}_{\ell(I^1)}$, we get that

$$\sum_{I^1 \in \mathcal{D}^1} \sum_{I^2 \times I^3 \in \mathcal{D}^{2,3}_{\ell(I^1)}} \left| \left\langle f, \frac{1_{I^1}}{|I^1|} \otimes h_{I^2 \times I^3} \right\rangle \right|^2 \left| \left\langle g, h_{I^1} \otimes \frac{1_{I^2 \times I^3}}{|I^2 \times I^3|} \right\rangle \right|^2 \langle \nu \rangle_I$$

$$\lesssim \sum_{I^1 \in \mathcal{D}^1} \int_{\mathbb{R}^2} \left(M^{2,3}_{\ell(I^1)} \langle f \rangle_{I^1} \right)^2 \left(M^{2,3}_{\ell(I^1)} \langle g, h_{I^1} \rangle \right)^2 \langle \nu \rangle_{I^1}$$

$$\leq \sum_{I^1 \in \mathcal{D}^1} \int_{\mathbb{R}^3} \left(M^{2,3}_{\ell(I^1)} \langle f \rangle_{I^1} \right)^2 \left(M^{2,3}_{\ell(I^1)} \langle |\Delta_{I^1} g| \rangle_{I^1} \right)^2 1_{I^1} \nu$$

$$\leq \int_{\mathbb{R}^2} [M_Z f]^2 \sum_{I^1 \in \mathcal{D}^1} M_Z (\Delta_{I^1} g)^2 \nu.$$

Thus, (7.9) is proved.

Case 2: little BMO paraproducts ($i = 3, j = 1, 2$ or $i = 1, 2, j = 3$). Actually, now we only have "trivial" type cases with different twist. Symmetries $a_{1,3}$ and $a_{3,1}$ are similar as well as $a_{2,3}$ and $a_{3,2}$. Let us choose for example $\Pi = a_{1,3}$ first. By Corollary 7.4 we have

$$|\langle \Pi(b, f), g \rangle| \lesssim \left\| \left(\sum_{I^1 \in \mathcal{D}^1} \left(\sum_{I^{2,3} \in \mathcal{D}^{2,3}_{\ell(I^1)}} |\langle f, h_{I,Z} \rangle| |\langle g, h_{I^1} h_{I^1} \otimes h_{I^{2,3}} \rangle| \frac{1_{I^{2,3}}}{|I^{2,3}|} \right)^2 \frac{1_{I^1}}{|I^1|} \right)^{\frac{1}{2}} \right\|_{L^1}.$$

Now we again can use similar sparse method as above and for fixed I^1 prove

$$\int \sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}} |\langle f, h_{I,Z}\rangle||\langle g, h_{I^1}h_{I^1} \otimes h_{I^{2,3}}\rangle| \frac{1_{I^{2,3}}}{|I^{2,3}|} \langle \nu \rangle_{I^1}$$

$$\lesssim \int M^{2,3}_{\ell(I^1)}(\langle |\Delta_{I^1} f|\rangle_{I^1}) M^{2,3}_{\ell(I^1)} \langle |g|\rangle_{I^1} 1_{I^1} \nu.$$

The above estimate together with vector-valued version of Theorem 7.10 (proven in [3] for general Muckenhoupt basis) yields

$$\left\| \left(\sum_{I^1 \in \mathcal{D}^1} \left(\sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}} |\langle f, h_{I,Z}\rangle||\langle g, h_{I^1}h_{I^1} \otimes h_{I^{2,3}}\rangle| \frac{1_{I^{2,3}}}{|I^{2,3}|} \right)^2 \frac{1_{I^1}}{|I^1|} \right)^{\frac{1}{2}} \right\|_{L^1}$$

$$\lesssim \left\| \left(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} f)^2 \frac{1_{I^1}}{|I^1|} \right)^{\frac{1}{2}} M_Z g \right\|_{L^1}$$

$$\leq \left\| \left(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} f)^2 \right)^{1/2} \right\|_{L^p} \|M_Z g\|_{L^{p'}} \lesssim \|f\|_{L^p} \|g\|_{L^{p'}}.$$

Moving to the symmetry $\Pi = a_{3,2}$ we first get

$$|\langle \Pi(b, f), g\rangle|$$

$$= \Big| \sum_{I \in \mathcal{D}_Z} \langle \langle b \rangle_{I^1}, h_{I^{2,3}}\rangle \Big\langle f, h_{I^1} \otimes \frac{1_{I^{2,3}}}{|I^{2,3}|} \Big\rangle \langle g, h_{I,Z}\rangle \Big|$$

$$\lesssim \|b\|_{\mathrm{bmo}_Z} \Big\| \sum_{I^1 \in \mathcal{D}^1} \Big(\sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}} |\langle f, h_{I^1} \otimes \frac{1_{I^{2,3}}}{|I^{2,3}|}\rangle|^2 |\langle g, h_{I,Z}\rangle|^2 \frac{1_{I^{2,3}}}{|I^{2,3}|} \Big)^{\frac{1}{2}} \frac{1_{I^1}}{|I^1|} \Big\|_{L^1},$$

where we use the other estimate in Corollary 7.4. Like above, we continue as follows

$$\Big\| \sum_{I^1 \in \mathcal{D}^1} \Big(\sum_{I^{2,3} \in \mathcal{D}_{\ell(I^1)}} |\langle f, h_{I^1} \otimes \frac{1_{I^{2,3}}}{|I^{2,3}|}\rangle|^2 |\langle g, h_{I,Z}\rangle|^2 \frac{1_{I^{2,3}}}{|I^{2,3}|} \Big)^{\frac{1}{2}} \frac{1_{I^1}}{|I^1|} \Big\|_{L^1}$$

$$\lesssim \Big\| \sum_{I^1 \in \mathcal{D}^1} M^{2,3}_{\ell(I^1)} \langle |\Delta_{I^1} f|\rangle_{I^1} M^{2,3}_{\ell(I^1)} \langle |\Delta_{I^1} g|\rangle_{I^1} 1_{I^1} \Big\|_{L^1}$$

$$\leq \Big\| \Big(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} f)^2 \Big)^{1/2} \Big\|_{L^p} \Big\| \Big(\sum_{I^1 \in \mathcal{D}^1} M_Z(\Delta_{I^1} g)^2 \Big)^{1/2} \Big\|_{L^{p'}}$$

$$\lesssim \|f\|_{L^p} \|g\|_{L^{p'}}.$$

$\square$

In above proof we needed the A_∞ extrapolation with Zygmund A_∞ weights. In fact, we give a very simple proof of A_∞ extrapolation [3] in general.

7.10. **Theorem.** *Let (f, g) be a pair of non-negative functions. Assume that there is some $0 < p_0 < \infty$ such that for all $w \in A_{\infty,Z}$ there holds*

$$\int f^{p_0} w \leq C([w]_{A_{\infty,Z}}) \int g^{p_0} w,$$

where C is an increasing function. Then for all $0 < p < \infty$ and all $w \in A_{\infty,Z}$ there holds

$$\int f^p w \leq C([w]_{A_{\infty,Z}}) \int g^p w.$$

Proof. We have for all $1 < r < \infty$ and all $w \in A_{r,Z}$ that

$$\int (f^{p_0/r})^r w \leq C([w]_{A_{r,Z}}) \int (g^{p_0/r})^r w.$$

Thus, by the classical extrapolation with $A_{p,Z}$ weights we have

$$(7.11) \qquad \int (f^{p_0/r})^s w \leq C([w]_{A_{s,Z}}) \int (g^{p_0/r})^s w$$

for all $1 < s < \infty$ and $w \in A_{s,Z}$.

Finally, let $0 < p < \infty$ and $w \in A_{\infty,Z}$. Then, there exists some $1 < s_0 < \infty$ such that $w \in A_{s_0,Z}$. Choose some $1 < r < \infty$ and $s_0 \leq s < \infty$ such that

$$sp_0/r = p.$$

For example, we can take

$$s = \frac{s_0 p}{p_0}\Big(\frac{p_0}{p} + 1\Big) = s_0\Big(\frac{p}{p_0} + 1\Big), \quad r = s_0\Big(\frac{p_0}{p} + 1\Big).$$

Since $A_{s_0,Z} \subset A_{s,Z}$, we can use (7.11) with the exponents s and r to get the claim. $\qquad\square$

7.B. Zygmund shift commutators. Let $k = (k^1, k^2)$, $k^i \in \{0, 1, 2, \ldots\}$, be fixed. A Zygmund shift $Q = Q_k$ of complexity k, see [14], has the form

$$\langle Q_k f, g \rangle$$
$$= \sum_{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}} \sum_{\substack{I, J \in \mathcal{D}_Z \\ I^{(k)} = K = J^{(k)}}} a_{IJK} \langle f, h_{I^1} \otimes H_{I^{2,3}, J^{2,3}} \rangle \langle g, H_{I^1, J^1} \otimes h_{J^{2,3}} \rangle$$

or

$$\langle Q_k f, g \rangle$$
$$= \sum_{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}} \sum_{\substack{I, J \in \mathcal{D}_Z \\ I^{(k)} = K = J^{(k)}}} a_{IJK} \langle f, h_{I^1} \otimes h_{I^{2,3}} \rangle \langle g, H_{I^1, J^1} \otimes H_{I^{2,3}, J^{2,3}} \rangle,$$

where $H_{I,J}$

   (1) is supported on $I \cup J$ and constant on children:

$$H_{I,J} = \sum_{L \in \mathrm{ch}(I) \cup \mathrm{ch}(J)} b_L 1_L$$

   (2) is L^2 normalized: $|H_{I,J}| \leq |I|^{-\frac{1}{2}}$, and
   (3) has zero average: $\int H_{I,J} = 0$.

We will be focusing on the mixed type form since it is the most interesting one. Usually the other type is much easier and the method is easily recovered from this case.

7.12. Proposition. *Let Q_k be a Zygmund shift of complexity $k = (k^1, k^2, k^3)$. Let $1 < p < \infty$ and $b \in \mathrm{bmo}_Z$. Then we have*

$$\|[b, Q_k]f\|_{L^p} \lesssim \max(k^1, k^2, k^3)(|k| + 1)^2 \|b\|_{\mathrm{bmo}_Z} \|f\|_{L^p}.$$

Proof. We consider the commutator $[b, Q_k]f \colon bQ_k f - Q_k(bf)$ that in the dual form equals to

$$\sum_{\substack{K \in \mathcal{D}_{2^{-k^1-k^2+k^3}}}} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK} \Big[\langle bf, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle$$

$$- \langle f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle bg, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle \Big].$$

Now, expanding both bf and bg with the expansion (7.7) we get the terms

$$\langle Q_k(a_{i,j}(b,f)), g \rangle \qquad \text{and} \qquad \langle Q_k f, a_{i,j}(b,g) \rangle$$

whenever $(i,j) \neq (3,3)$. These terms are directly bounded separately, in particular, we have $Q_k \colon L^p \to L^p$ and $a_{i,j} \colon L^p \to L^p$. Hence, we are left with bounding

$$\sum_{\substack{K \in \mathcal{D}_\lambda}} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK} \Big[\sum_{L \in \mathcal{D}_Z} \langle b \rangle_L \langle \Delta_{L,Z} f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle$$

$$- \sum_{L \in \mathcal{D}_Z} \langle b \rangle_L \langle f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle \Delta_{L,Z} g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle \Big]$$

$$= \sum_{\substack{K \in \mathcal{D}_\lambda}} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK}$$

$$\times \Big[\sum_{\substack{L \in \mathcal{D}_Z \\ \ell(L^1)=2^{-k^1}\ell(K^1) \\ \ell(K^2) \leq 2^{k^2}\ell(L^2) \leq 2^{\max(k^2,k^3)}\ell(K^2)}} \langle b \rangle_L \langle \Delta_{L,Z} f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle$$

$$- \sum_{\substack{Q \in \mathcal{D}_Z \\ Q^1 \subset K^1,\, \ell(Q^1) \geq \ell(I^1) \\ 2^{-k^1}\ell(K^2) \leq 2^{k^2}\ell(Q^2) \leq \ell(K^2)}} \langle b \rangle_Q \langle f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle \Delta_{Q,Z} g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle \Big],$$

where we have abbreviated $2^{-k^1-k^2+k^3}$ by λ. Now, we write

$$\langle f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle = \sum_{\substack{L \in \mathcal{D}_Z \\ \ell(L^1)=2^{-k^1}\ell(K^1) \\ \ell(K^2) \leq 2^{k^2}\ell(L^2) \leq 2^{\max(k^2,k^3)}\ell(K^2)}} \langle \Delta_{L,Z} f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle$$

and

$$\langle g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle = \sum_{\substack{Q \in \mathcal{D}_Z \\ Q^1 \subset K^1,\, \ell(Q^1) \geq \ell(I^1) \\ 2^{-k^1}\ell(K^2) \leq 2^{k^2}\ell(Q^2) \leq \ell(K^2)}} \langle \Delta_{Q,Z} g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle$$

for the unexpanded terms. Thus, we end up with

$$\sum_{\substack{K \in \mathcal{D}_\lambda}} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK} \sum_{\substack{L \in \mathcal{D}_Z \\ \ell(L^1)=2^{-k^1}\ell(K^1) \\ \ell(K^2) \leq 2^{k^2}\ell(L^2) \leq 2^{\max(k^2,k^3)}\ell(K^2)}} \sum_{\substack{Q \in \mathcal{D}_Z \\ Q^1 \subset K^1,\, \ell(Q^1) \geq \ell(I^1) \\ 2^{-k^1}\ell(K^2) \leq 2^{k^2}\ell(Q^2) \leq \ell(K^2)}}$$

$$\times \Big[(\langle b \rangle_L - \langle b \rangle_Q) \langle \Delta_{L,Z} f, h_{I^1} \otimes H_{I^{2,3}, J^{2,3}} \rangle \langle \Delta_{Q,Z} g, H_{I^1, J^1} \otimes h_{J^{2,3}} \rangle \Big].$$

We write explicitly the complexity levels for Q and L. That is, in the above summations we have $(L^2)^{(l^2)} = (K^2)^{(\max(0, k^3 - k^2))}$ for some $l^2 \in \{0, \dots, \max(k^2, k^3)\}$, $(Q^1)^{(q_1)} = K^1$, for some $q_1 \in \{0, \dots, k^1\}$, and $(Q^2)^{(q_2)} = K^2$ for some $q_2 \in \{k^2, \dots, k^2 + k^1\}$. We get

$$\sum_{\substack{K \in \mathcal{D}_\lambda}} \sum_{\substack{I, J \in \mathcal{D}_Z \\ I^{(k)} = K = J^{(k)}}} a_{IJK} \sum_{l^2 = 0}^{\max(k^2, k^3)} \sum_{\substack{q_1 \in \{0, \dots, k^1\} \\ q_2 \in \{k^2, \dots, k^2 + k^1\}}} \sum_{\substack{L \in \mathcal{D}_Z \\ \ell(L^1) = 2^{-k^1} \ell(K^1) \\ (L^2)^{(l^2)} = (K^2)^{(\max(0, k^3 - k^2))}}} \sum_{\substack{Q \in \mathcal{D}_Z \\ (Q^1)^{(q_1)} = K^1 \\ (Q^2)^{(q_2)} = K^2}}$$

$$\times \Big[(\langle b \rangle_L - \langle b \rangle_Q) \langle \Delta_{L,Z} f, h_{I^1} \otimes H_{I^{2,3}, J^{2,3}} \rangle \langle \Delta_{Q,Z} g, H_{I^1, J^1} \otimes h_{J^{2,3}} \rangle \Big].$$

Here we need to notice that $R = R^1 \times R^2 \times R^3 \supset K, L, Q$, where

$$R = K^{(k^1, \max(0, k^3 - k^2), k^1 + \max(k^2 - k^3, 0))} \text{ and } R \in \mathcal{D}_Z.$$

This is a common "Zygmund ancestor" for all of these rectangles.

Let us expand in the difference $\langle b \rangle_L - \langle b \rangle_Q$ in the following way

$$\begin{aligned}
\langle b \rangle_L &= \langle b \rangle_L - \langle b \rangle_{L^{(0,1,1)}} \\
&\quad + \langle b \rangle_{L^{(0,1,1)}} - \langle b \rangle_{L^{(0,2,2)}} \\
&\quad \vdots \\
&\quad + \langle b \rangle_{L^{(0, l^2 - 1, l^2 - 1)}} - \langle b \rangle_{L^{(0, l^2, l^2)}} + \langle b \rangle_{L^{(0, l^2, l^2)}} \\
&= \sum_{r^2 = 0}^{l^2 - 1} \left(\langle b \rangle_{L^{(0, r^2, r^2)}} - \langle b \rangle_{L^{(0, r^2 + 1, r^2 + 1)}} \right) + \langle b \rangle_{L^{(0, l^2, l^2)}}.
\end{aligned}$$

Notice that since $\ell(L^1)\ell(L^2) = \ell(L^3)$, we have $\ell(L^1)\ell((L^2)^{(r^2)}) = \ell((L^3)^{(r^2)})$, i.e. rectangles $(L^2)^{(r^2)} \times (L^3)^{(r^2)} \in \mathcal{D}_{\ell(L^1)}$ which is desirable since we want to use the characterization (2) in Lemma 7.2. We continue with the last term

$$\begin{aligned}
\langle b \rangle_{L^{(0, l^2, l^2)}} &= \langle b \rangle_{L^{(0, l^2, l^2)}} - \langle b \rangle_{L^{(1, l^2, 1 + l^2)}} \\
&\quad + \langle b \rangle_{L^{(1, l^2, 1 + l^2)}} - \langle b \rangle_{L^{(2, l^2, 2 + l^2)}} \\
&\quad \vdots \\
&\quad \langle b \rangle_{L^{(k^1 - 1, l^2, k^1 - 1 + l^2)}} - \langle b \rangle_{L^{(k^1, l^2, k^1 + l^2)}} + \langle b \rangle_G \\
&= \sum_{r^1 = 0}^{k^1 - 1} \left(\langle b \rangle_{L^{(r^1, l^2, r^1 + l^2)}} - \langle b \rangle_{L^{(r^1 + 1, l^2, r^1 + 1 + l^2)}} \right) + \langle b \rangle_R.
\end{aligned}$$

Recall that $(L^2)^{(l^2)} = (K^2)^{(\max(0, k^3 - k^2))} =: R^2$ and observe that since $\ell((L^3)^{(k^1 + l^2)}) = \ell((L^2)^{(l^2)})\ell((L^1)^{(k^1)}) = \ell(R^2)\ell(K^1)$ we get $(L^3)^{(k^1 + l^2)} = R^3$. Thus, we end up with a sum of terms of the forms

(7.13) $$\langle b \rangle_{L^{(0, r^2, r^2)}} - \langle b \rangle_{L^{(0, r^2 + 1, r^2 + 1)}} \quad \text{and} \quad \langle b \rangle_{L^{(r^1, l^2, r^1 + l^2)}} - \langle b \rangle_{L^{(r^1 + 1, l^2, r^1 + 1 + l^2)}},$$

and we have for fixed r^1 and r^2

$$|(7.13)| \lesssim \|b\|_{\mathrm{bmo}_Z}$$

by Lemma 7.2.

By the same argument as above we get

$$\langle b \rangle_Q = \sum_{\rho^2=0}^{\max(0,k^3-k^2)+q^2-1} \langle b \rangle_{Q^{(0,\rho^2,\rho^2)}} - \langle b \rangle_{Q^{(0,\rho^2+1,\rho^2+1)}}$$
$$+ \sum_{\rho^1=0}^{q^1} \langle b \rangle_{Q^{(\rho^1,\tilde{q}^2,\rho^1+\tilde{q}^2)}} - \langle b \rangle_{Q^{(\rho^1+1,\tilde{q}^2,\rho^1+1+\tilde{q}^2)}}$$
$$+ \langle b \rangle_R,$$

where $\tilde{q}^2 = \max(0, k^3 - k^2) + q^2$,

$$(Q^2)^{(\tilde{q}^2)} = (K^2)^{(\max(0,k^3-k^2))} \qquad \text{and} \qquad (Q^3)^{(q^1+\tilde{q}^2)} = (K^3)^{(k^1+\max(k^2-k^3,0))}.$$

Notice that the last term corresponds to the last term in the previous expansion, and hence, their difference equals to zero. Again, here we have

$$\left| \langle b \rangle_{Q^{(0,\rho^2,\rho^2)}} - \langle b \rangle_{Q^{(0,\rho^2+1,\rho^2+1)}} + \langle b \rangle_{Q^{(\rho^1,\tilde{q}^2,\rho^1+\tilde{q}^2)}} - \langle b \rangle_{Q^{(\rho^1+1,\tilde{q}^2,\rho^1+1+\tilde{q}^2)}} \right| \lesssim \|b\|_{\mathrm{bmo}_Z}$$

for fixed ρ^1 and ρ^2.

Now, we can split the commutator into the two terms

$$\mathcal{W}_{K,k}^b f = 1_K \sum_{\substack{L \in \mathcal{D}_Z \\ \ell(L^1)=2^{-k^1}\ell(K^1) \\ \ell(K^2) \le 2^{k^2}\ell(L^2) \le 2^{\max(k^2,k^3)}\ell(K^2)}} b_{L,K} \Delta_{L,Z} f,$$

where

$$|b_{L,K}| \lesssim \max(k^1, k^2, k^3) \|b\|_{\mathrm{bmo}_Z},$$

and

$$\mathcal{V}_{K,k}^b g = \sum_{\substack{Q \in \mathcal{D}_Z \\ Q^1 \subset K^1,\ \ell(Q^1) \ge \ell(I^1) \\ 2^{-k^1}\ell(K^2) \le 2^{k^2}\ell(Q^2) \le \ell(K^2)}} b_{Q,K} \Delta_{Q,Z} g,$$

where

$$|b_{Q,K}| \lesssim \max(k^1, k^2, k^3) \|b\|_{\mathrm{bmo}_Z}.$$

Thus, the last term of the commutator is the sum of

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK} \langle \mathcal{W}_{K,k}^b f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle \mathcal{V}_{K,k} g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle$$

and

$$\sum_{K \in \mathcal{D}_\lambda} \sum_{\substack{I,J \in \mathcal{D}_Z \\ I^{(k)}=K=J^{(k)}}} a_{IJK} \langle \mathcal{W}_{K,k} f, h_{I^1} \otimes H_{I^{2,3},J^{2,3}} \rangle \langle \mathcal{V}_{K,k}^b g, H_{I^1,J^1} \otimes h_{J^{2,3}} \rangle.$$

The boundedness follows via standard methods (adapt proofs of [14, Theorem 6.2 and Lemma 5.20].)   □

APPENDIX A. BILINEAR FEFFERMAN-PIPHER MULTIPLIERS

In this section we consider bilinear variants of multipliers studied by Fefferman-Pipher [6]. These considerations motivate the kernel estimates in Section 3. After the presented calculations, the reader can easily check how everything fits with Section 3. In fact, we will see that the bilinear Fefferman-Pipher multipliers produce kernels which satisfy the the kernel estimates in Section 3 with

$$\theta = 2, \quad \alpha_1 = 1, \quad \alpha_{2,3} = 1,$$

and an extra logarithm factor. In the partial kernel estimates $\widetilde{\theta} = 1$ and there is also a harmless logarithm factor.

We consider the following multi-parameter dilation on $\mathbb{R}^6$ – define

$$\rho_{s,t}(x, y) = (sx_1, tx_2, stx_3, sy_1, ty_2, sty_3), \qquad s, t > 0,$$

and set

$$A^1 := \{(\xi, \eta) \in \mathbb{R}^6 \colon \tfrac{1}{2} < |(\xi_1, \eta_1)| \leq 1, \tfrac{1}{2} < |(\xi_2, \xi_3, \eta_2, \eta_3)| \leq 1\}.$$

In this section we consider the parameter groups $\{1\}$ and $\{2, 3\}$ only. The grouping $\{\{2\}, \{1, 3\}\}$ is similar, for example, we would set

$$A^2 := \{(\xi, \eta) \in \mathbb{R}^6 \colon \tfrac{1}{2} < |(\xi_2, \eta_2)| \leq 1, \tfrac{1}{2} < |(\xi_1, \xi_3, \eta_1, \eta_3)| \leq 1\}.$$

For Schwartz functions f_1, f_2 we define the bilinear multiplier operator

$$T_{m,1}(f_1, f_2)(x) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} m(\xi, \eta) \widehat{f_1}(\xi) \widehat{f_2}(\eta) e^{2\pi i x \cdot (\xi + \eta)} \, d\xi \, d\eta,$$

where the symbol $m \in C^N$ is assumed to satisfy

$$\|m\|_{\mathcal{M}_Z^1} := \sup_{\substack{|\alpha|_\infty \leq N \\ |\beta|_\infty \leq N}} \sup_{s, t > 0} \sup_{(\xi, \eta) \in A^1} |\partial_\xi^\alpha \partial_\eta^\beta (m \circ \rho_{s,t})(\xi, \eta)| < \infty.$$

Thus, if $(\xi, \eta) \in A^1$, then by definition

(A.1)

$$|(\partial_\xi^\alpha \partial_\eta^\beta m)(s\xi_1, t\xi_2, st\xi_3, s\eta_1, t\eta_2, st\eta_3)| \leq \|m\|_{\mathcal{M}_Z^1} s^{-\alpha_1 - \beta_1} t^{-\alpha_2 - \beta_2} (st)^{-\alpha_3 - \beta_3}$$

$$= \|m\|_{\mathcal{M}_Z^1} s^{-(\alpha_1 + \beta_1) + (\alpha_2 + \beta_2)} (st)^{-(\alpha_2 + \beta_2) - (\alpha_3 + \beta_3)}.$$

Now, for $(\zeta_1, \sigma_1) \neq 0$ and $(\zeta_2, \zeta_3, \sigma_2, \sigma_3) \neq 0$ denote

$$s = |(\zeta_1, \sigma_1)|, \qquad st = |(s\zeta_2, \zeta_3, s\sigma_2, \sigma_3)|,$$

$$(\xi_1, \xi_2, \xi_3) = \Big(\frac{\zeta_1}{s}, \frac{\zeta_2}{t}, \frac{\zeta_3}{st}\Big), \qquad (\eta_1, \eta_2, \eta_3) = \Big(\frac{\sigma_1}{s}, \frac{\sigma_2}{t}, \frac{\sigma_3}{st}\Big).$$

Thus, $(\xi, \eta) \in A^1$ and

(A.2)

$$|\partial_\zeta^\alpha \partial_\sigma^\beta m(\zeta, \sigma)| \lesssim \|m\|_{\mathcal{M}_Z^1} (|\zeta_1| + |\sigma_1|)^{-(\alpha_1 + \beta_1) + (\alpha_2 + \beta_2)}$$

$$\times \big(|((|\zeta_1| + |\sigma_1|)\zeta_2, \zeta_3)| + |((|\zeta_1| + |\sigma_1|)\sigma_2, \sigma_3)|\big)^{-(\alpha_2 + \beta_2) - (\alpha_3 + \beta_3)}.$$

We write, with two standard partitions of unity ϕ_1 on $\mathbb{R}^2 \setminus \{0\}$ and $\phi_{2,3}$ on $\mathbb{R}^4 \setminus \{0\}$, that

$$1 = \sum_{j,k \in \mathbb{Z}} \phi_1(2^{-j}\xi_1, 2^{-j}\eta_1)\phi_{2,3}(2^{-k}\xi_2, 2^{-j-k}\xi_3, 2^{-k}\eta_2, 2^{-j-k}\eta_3).$$

Via this identity we obtain

$$m = \sum_{j,k}(\phi_1 \otimes \phi_{2,3} \circ \rho_{2^{-j},2^{-k}}) \cdot m$$

$$= \sum_{j,k}(\phi_1 \otimes \phi_{2,3} \cdot (m \circ \rho_{2^j,2^k})) \circ \rho_{2^{-j},2^{-k}} =: m_{j,k}.$$

Since ϕ_1 and $\phi_{2,3}$ are supported in $\bar{B}(0,2) \setminus B(0,\frac{1}{2})$ in $\mathbb{R}^2$ and $\mathbb{R}^4$, respectively, we know that

$\operatorname{spt} m_{j,k} \subset$

$$\rho_{2^j,2^k}\Big\{(\xi,\eta) : (\xi_1,\eta_1) \in \bar{B}_{\mathbb{R}^2}(0,2) \setminus B_{\mathbb{R}^2}(0,\tfrac{1}{2}), (\xi_{2,3},\eta_{2,3}) \in \bar{B}_{\mathbb{R}^4}(0,2) \setminus B_{\mathbb{R}^4}(0,\tfrac{1}{2})\Big\}.$$

Using this we get

$$\|\partial^\alpha \partial^\beta m_{j,k}\|_{L^\infty} \lesssim 2^{-(j,k,j+k)\cdot(\alpha+\beta)} \quad \text{and} \quad \|\partial^\alpha \partial^\beta m_{j,k}\|_{L^1} \lesssim 2^{(j,k,j+k)\cdot(\mathbf{2}-(\alpha+\beta))},$$

where $\mathbf{2} = (2,2,2)$.

Let $K_{j,k}(y,z) = \check{m}_{j,k}$ and $K(y,z) = \sum_{j,k} K_{j,k}(y,z)$ – then $K(x-y, x-z)$ is the corresponding kernel. Using similar analysis as in [14] we have

$$\|y^\alpha z^{\tilde{\alpha}} \partial_y^\beta \partial_z^\gamma K_{j,k}\|_{L^\infty} \lesssim \|\partial_\xi^\alpha \partial_\eta^{\tilde{\alpha}}(\xi^\beta \eta^\gamma m_{j,k})\|_{L^1}$$

$$\leq \sum_{\substack{l \leq \alpha \\ \tilde{l} \leq \tilde{\alpha}}} \binom{\alpha}{l}\binom{\tilde{\alpha}}{\tilde{l}} \|\partial^l(\xi^\beta)\partial^{\tilde{l}}(\eta^\gamma) \cdot \partial^{\alpha-l}\partial^{\tilde{\alpha}-\tilde{l}} m_{j,k})\|_{L^1}$$

$$\lesssim 2^{(j,k,j+k)\cdot(\mathbf{2}+(\beta+\gamma)-(\alpha+\tilde{\alpha}))}$$

for multi-indices $\alpha, \tilde{\alpha}, \beta, \gamma$. Hence, we get

$$|y^{\beta+\mathbf{1}} z^{\gamma+\mathbf{1}} \partial_y^\beta \partial_z^\gamma K_{j,k}(y,z)| \lesssim 2^{(j,k,j+k)\cdot(\mathbf{2}+(\beta+\gamma)-(\alpha+\tilde{\alpha}))}|y^{\beta+\mathbf{1}-\alpha}| \cdot |z^{\gamma+\mathbf{1}-\tilde{\alpha}}|.$$

Taking $\alpha_i, \tilde{\alpha}_i \in \{0, N\}$ we obtain

$$|y^{\beta+\mathbf{1}} z^{\gamma+\mathbf{1}} \partial_y^\beta \partial_z^\gamma K(y,z)|$$

$$\lesssim \sum_j \min\{(2^j|y_1|)^{\beta_1+1}, (2^j|y_1|)^{\beta_1+1-N}\} \min\{(2^j|z_1|)^{\gamma_1+1}, (2^j|z_1|)^{\gamma_1+1-N}\}$$

$$\times \sum_k \min\{(2^k|y_2|)^{\beta_2+1}, (2^k|y_2|)^{\beta_2+1-N}\} \min\{(2^k|z_2|)^{\gamma_2+1}, (2^k|z_2|)^{\gamma_2+1-N}\}$$

$$\times \min\{(2^{j+k}|y_3|)^{\beta_3+1}, (2^{j+k}|y_3|)^{\beta_3+1-N}\} \min\{(2^{j+k}|z_3|)^{\gamma_3+1}, (2^{j+k}|z_3|)^{\gamma_3+1-N}\}.$$

We can estimate the inner sum either by

$$\sum_{k:\, 2^k < 1/(|y_2|+|z_2|)} (2^k|y_2|)^{\beta_2+1}(2^k|z_2|)^{\gamma_2+1}(2^{j+k}|y_3|)^{\beta_3+1}(2^{j+k}|z_3|)^{\gamma_3+1}$$

$$+ \sum_{k:\, 2^k \geq 1/(|y_2|+|z_2|) \geq 1/(2|y_2|)} (2^k|y_2|)^{\beta_2+1-N}(2^k|z_2|)^{\gamma_2+1}(2^{j+k}|y_3|)^{\beta_3+1}(2^{j+k}|z_3|)^{\gamma_3+1}$$

$$+ \sum_{k:\, 2^k \geq 1/(|y_2|+|z_2|) > 1/(2|z_2|)} (2^k|y_2|)^{\beta_2+1}(2^k|z_2|)^{\gamma_2+1-N}(2^{j+k}|y_3|)^{\beta_3+1}(2^{j+k}|z_3|)^{\gamma_3+1}$$

$$\lesssim \frac{|y_2|^{\beta_2+1}}{(|y_2|+|z_2|)^{\beta_2+1}} \cdot \frac{|z_2|^{\gamma_2+1}}{(|y_2|+|z_2|)^{\gamma_2+1}} \cdot \frac{(2^j|y_3|)^{\beta_3+1}}{(|y_2|+|z_2|)^{\beta_3+1}} \cdot \frac{(2^j|z_3|)^{\gamma_3+1}}{(|y_2|+|z_2|)^{\gamma_3+1}} =: I_1$$

or by

$$\sum_{k:\, 2^k < 2^{-j}/(|y_3|+|z_3|)} (2^k|y_2|)^{\beta_2+1}(2^k|z_2|)^{\gamma_2+1}(2^{j+k}|y_3|)^{\beta_3+1}(2^{j+k}|z_3|)^{\gamma_3+1}$$

$$+ \sum_{k:\, 2^k \geq 2^{-j}/(|y_3|+|z_3|) \geq 2^{-j}/(2|y_3|)} (2^k|y_2|)^{\beta_2+1}(2^k|z_2|)^{\gamma_2+1}(2^{j+k}|y_3|)^{\beta_3+1-N}(2^{j+k}|z_3|)^{\gamma_3+1}$$

$$+ \sum_{k:\, 2^k \geq 2^{-j}/(|y_3|+|z_3|) > 2^{-j}/(2|z_3|)} (2^k|y_2|)^{\beta_2+1}(2^k|z_2|)^{\gamma_2+1}(2^{j+k}|y_3|)^{\beta_3+1}(2^{j+k}|z_3|)^{\gamma_3+1-N}$$

$$\lesssim \frac{|y_2|^{\beta_2+1}}{[2^j(|y_3|+|z_3|)]^{\beta_2+1}} \cdot \frac{|z_2|^{\gamma_2+1}}{[2^j(|y_3|+|z_3|)]^{\gamma_2+1}} \cdot \frac{|y_3|^{\beta_3+1}}{(|y_3|+|z_3|)^{\beta_3+1}} \cdot \frac{|z_3|^{\gamma_3+1}}{(|y_3|+|z_3|)^{\gamma_3+1}} =: I_2,$$

in both cases provided that $\beta_2 + \beta_3 + \gamma_2 + \gamma_3 < N - 4$.

The outer sum can then be estimated either by

$$\sum_{j:\, 2^j < 1/(|y_1|+|z_1|)} (2^j|y_1|)^{\beta_1+1}(2^j|z_1|)^{\gamma_1+1}I_1$$

$$+ \sum_{j:\, 2^j \geq 1/(|y_1|+|z_1|) \geq 1/(2|y_1|)} (2^j|y_1|)^{\beta_1+1-N}(2^j|z_1|)^{\gamma_1+1}I_1$$

$$+ \sum_{j:\, 2^j \geq 1/(|y_1|+|z_1|) > 1/(2|z_1|)} (2^j|y_1|)^{\beta_1+1}(2^j|z_1|)^{\gamma_1+1-N}I_1$$

$$\lesssim \frac{|y_1|^{\beta_1+1}|z_1|^{\gamma_1+1}}{(|y_1|+|z_1|)^{\beta_1+\gamma_1+2}} \frac{|y_2|^{\beta_2+1}|z_2|^{\gamma_2+1}}{(|y_2|+|z_2|)^{\beta_2+\gamma_2+2}} \frac{|y_3|^{\beta_3+1}|z_3|^{\gamma_3+1}}{[(|y_1|+|z_1|)(|y_2|+|z_2|)]^{\beta_3+\gamma_3+2}}$$

or, if $(|y_1| + |z_1|)(|y_2| + |z_2|) \leq |y_3| + |z_3|$, by

$$\sum_{j:\, 2^j < (|y_2|+|z_2|)/(|y_3|+|z_3|)} (2^j|y_1|)^{\beta_1+1}(2^j|z_1|)^{\gamma_1+1}I_1$$

$$+ \sum_{j:\, \frac{|y_2|+|z_2|}{|y_3|+|z_3|} \leq 2^j \leq \frac{1}{|y_1|+|z_1|}} (2^j|y_1|)^{\beta_1+1}(2^j|z_1|)^{\gamma_1+1}I_2$$

$$+ \sum_{j:\, 2^j > 1/(|y_1|+|z_1|) > 1/(2|z_1|)} (2^j|y_1|)^{\beta_1+1}(2^j|z_1|)^{\gamma_1+1-N}I_2$$

$$+ \sum_{j:\, 2^j > 1/(|y_1|+|z_1|) > 1/(2|y_1|)} (2^j|y_1|)^{\beta_1+1-N}(2^j|z_1|)^{\gamma_1+1}I_2 =: I + II + III + IV.$$

It is straightforward that

$$I \sim \frac{|y_1|^{\beta_1+1}|z_1|^{\gamma_1+1}}{(|y_1|+|z_1|)^{\beta_1+\gamma_1+2}} \frac{|y_2|^{\beta_2+1}|z_2|^{\gamma_2+1}}{(|y_2|+|z_2|)^{\beta_2+\gamma_2+2}} \frac{|y_3|^{\beta_3+1}|z_3|^{\gamma_3+1}}{(|y_3|+|z_3|)^{\beta_3+\gamma_3+2}}$$

$$\times \left(\frac{(|y_1|+|z_1|)(|y_2|+|z_2|)}{|y_3|+|z_3|}\right)^{\beta_1+\gamma_1+2};$$

$$III \sim IV \sim \frac{|y_1|^{\beta_1+1}|z_1|^{\gamma_1+1}}{(|y_1|+|z_1|)^{\beta_1+\gamma_1+2}} \frac{|y_2|^{\beta_2+1}|z_2|^{\gamma_2+1}}{(|y_2|+|z_2|)^{\beta_2+\gamma_2+2}} \frac{|y_3|^{\beta_3+1}|z_3|^{\gamma_3+1}}{(|y_3|+|z_3|)^{\beta_3+\gamma_3+2}}$$
$$\times \left(\frac{(|y_1|+|z_1|)(|y_2|+|z_2|)}{|y_3|+|z_3|} \right)^{\beta_2+\gamma_2+2}.$$

Lastly, we have

$$II \sim \frac{|y_1|^{\beta_1+1}|z_1|^{\gamma_1+1}}{(|y_1|+|z_1|)^{\beta_1+\gamma_1+2}} \frac{|y_2|^{\beta_2+1}|z_2|^{\gamma_2+1}}{(|y_2|+|z_2|)^{\beta_2+\gamma_2+2}} \frac{|y_3|^{\beta_3+1}|z_3|^{\gamma_3+1}}{(|y_3|+|z_3|)^{\beta_3+\gamma_3+2}}$$
$$\times \left(\frac{(|y_1|+|z_1|)(|y_2|+|z_2|)}{|y_3|+|z_3|} \right)^{\min\{\beta_1+\gamma_1,\beta_2+\gamma_2\}+2} L_{\beta_1,\beta_2,\gamma_1,\gamma_2}(y,z),$$

where

$$L_{\beta_1,\beta_2,\gamma_1,\gamma_2}(y,z) := 1 + \log_+\left(\frac{|y_3|+|z_3|}{(|y_1|+|z_1|)(|y_2|+|z_2|)} \right)$$

when $\beta_1+\gamma_1 = \beta_2+\gamma_2$ and $L_{\beta_1,\beta_2,\gamma_1,\gamma_2}(y,z) = 1$ otherwise. In conclusion, we get

$$|\partial_y^\beta \partial_z^\gamma K(y,z)| \lesssim \frac{1}{[(|y_1|+|z_1|)(|y_2|+|z_2|)+|y_3|+|z_3|]^{\beta_3+\gamma_3+4}}$$
$$\times \frac{1}{(|y_1|+|z_1|)^{\beta_1+\gamma_1}(|y_2|+|z_2|)^{\beta_2+\gamma_2}}$$
$$\times \min\left\{ 1, \left(\frac{(|y_1|+|z_1|)(|y_2|+|z_2|)}{|y_3|+|z_3|} \right)^{\min\{\beta_1+\gamma_1,\beta_2+\gamma_2\}} \right\} L_{\beta_1,\beta_2,\gamma_1,\gamma_2}(y,z).$$

Partial kernel estimates. Let $m \in \mathcal{M}_Z^1$. We define truncations of m by setting

$$m_J := \sum_{|j| \le J^1, |k| \le J^2} m_{j,k}, \quad J = (J^1, J^2) \in \mathbb{N}^2.$$

A.3. **Lemma.** *Suppose that $m \in \mathcal{M}_Z^1$. Let m_J be defined as above and let $K_J = \check{m}_J$. Then for $(y_2, z_2) \ne 0 \ne (y_3, z_3)$ we have the estimate*

$$\left| \iiint_{I^1 \times I^1 \times I^1} \partial_{y_2}^{\beta_2} \partial_{y_3}^{\beta_3} \partial_{z_2}^{\gamma_2} \partial_{z_3}^{\gamma_3} K_J(x_1 - y_1, y_2, y_3, x_1 - z_1, z_2, z_3) \, dy_1 \, dz_1 \, dx_1 \right|$$
$$\lesssim \frac{1}{(|y_2|+|z_2|)^{\beta_2+\gamma_2}} \cdot \frac{1}{(|y_3|+|z_3|)^{\beta_3+\gamma_3}} |I^1| \left(\frac{|I^1|(|y_2|+|z_2|)}{|y_3|+|z_3|} + \frac{|y_3|+|z_3|}{|I^1|(|y_2|+|z_2|)} \right)^{-1}$$
$$\times \frac{1}{\prod_{i=2}^3 (|y_i|+|z_i|)^2} \cdot \left(1 + \log_+ \frac{|y_3|+|z_3|}{|I^1|(|y_2|+|z_2|)} \right),$$

where I^1 is an interval and $\beta_2 + \beta_3 + \gamma_2 + \gamma_3 \le 1$.

Proof. Since $m_J(0, \xi_2, \xi_3, 0, \eta_2, \eta_3) = 0$, using the Fourier transform we know that

(A.4) $$\iint_{\mathbb{R}^2} \partial_{y_2}^{\beta_2} \partial_{y_3}^{\beta_3} \partial_{z_2}^{\gamma_2} \partial_{z_3}^{\gamma_3} K_J(y_1, y_2, y_3, z_1, z_2, z_3) \, dy_1 \, dz_1 = 0.$$

Suppose first that $|I^1|(|y_2|+|z_2|) \ge |y_3|+|z_3|$ – by (A.4) we may equivalently estimate the integral over $I^1 \times (\mathbb{R}^2 \setminus (I^1 \times I^1))$ instead of $I^1 \times I^1 \times I^1$. By the kernel estimates we have

$$\left| \iiint_{I^1 \times (\mathbb{R}^2 \setminus (I^1 \times I^1))} \partial_{y_2}^{\beta_2} \partial_{y_3}^{\beta_3} \partial_{z_2}^{\gamma_2} \partial_{z_3}^{\gamma_3} K_J(x_1 - y_1, y_2, y_3, x_1 - z_1, z_2, z_3) \, dy_1 \, dz_1 \, dx_1 \right|$$

$$\lesssim \int_{I^1} \iint_{\mathbb{R}^2 \setminus (I^1 \times I^1)} \frac{1}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2}}$$

$$\times \frac{1 + \log_+ \frac{|y_3| + |z_3|}{(|x_1 - y_1| + |x_1 - z_1|)(|y_2| + |z_2|)}}{[(|x_1 - y_1| + |x_1 - z_1|)(|y_2| + |z_2|) + |y_3| + |z_3|]^{\beta_3 + \gamma_3 + 4}} \, dy_1 \, dz_1 \, dx_1.$$

Note that we have either $y_1 \in \mathbb{R} \setminus I^1$ or $z_1 \in \mathbb{R} \setminus I^1$, and we may without loss of generality assume $y_1 \in \mathbb{R} \setminus I^1$. Then the integral is dominated by

$$\int_{I^1} \iint_{(\mathbb{R} \setminus I^1) \times \mathbb{R}} \frac{1}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2}}$$

$$\times \frac{1 + \log_+ \frac{|y_3| + |z_3|}{|x_1 - y_1|(|y_2| + |z_2|)}}{[(|x_1 - y_1| + |x_1 - z_1|)(|y_2| + |z_2|) + |y_3| + |z_3|]^{\beta_3 + \gamma_3 + 4}} \, dy_1 \, dz_1 \, dx_1$$

$$\lesssim \frac{1}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \int_{I^1} \int_{\mathbb{R} \setminus I^1} \frac{1 + \log_+ \frac{|y_3| + |z_3|}{|x_1 - y_1|(|y_2| + |z_2|)}}{\left(|x_1 - y_1| + \frac{|y_3| + |z_3|}{|y_2| + |z_2|}\right)^{\beta_3 + \gamma_3 + 3}} \, dy_1 \, dx_1.$$

Let $t := \frac{|y_3| + |z_3|}{|y_2| + |z_2|}$. By a change of variables we reduce to

$$\frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \iint_{t^{-1} I^1 \times (\mathbb{R} \setminus t^{-1} I^1)} \frac{1 + \log_+ \frac{1}{|x_1 - y_1|}}{(|x_1 - y_1| + 1)^{\beta_3 + \gamma_3 + 3}} \, dy_1 \, dx_1$$

$$\lesssim \frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \int_{t^{-1} I^1} \frac{1}{\left(d(x_1, \partial(t^{-1} I^1)) + 1\right)^{\beta_3 + \gamma_3 + 2}} \, dx_1$$

$$\lesssim \frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}}$$

$$= \frac{1}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + 3}} \frac{1}{(|y_3| + |z_3|)^{\beta_3 + \gamma_3 + 1}}$$

$$\sim \frac{1}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2}} \cdot \frac{1}{(|y_3| + |z_3|)^{\beta_3 + \gamma_3}} |I^1| \left(\frac{|I^1|(|y_2| + |z_2|)}{|y_3| + |z_3|} + \frac{|y_3| + |z_3|}{|I^1|(|y_2| + |z_2|)}\right)^{-1}$$

$$\times \frac{1}{\prod_{i=2}^{3} (|y_i| + |z_i|)^2}.$$

Assume then that $|I^1|(|y_2| + |z_2|) < |y_3| + |z_3|$. This time we integrate over $I^1 \times I^1 \times I^1$. Proceeding as above we reduce to the integral

$$\iiint_{t^{-1} I^1 \times t^{-1} I^1 \times t^{-1} I^1} \frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \frac{1 + \log_+ \frac{1}{(|x_1 - y_1| + |x_1 - z_1|)}}{[(|x_1 - y_1| + |x_1 - z_1|) + 1]^{\beta_3 + \gamma_3 + 4}} \, dy_1 \, dz_1 \, dx_1$$

$$\leq \iint_{t^{-1} I^1 \times t^{-1} I^1} \frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \frac{1 + \log_+ \frac{1}{|x_1 - y_1|}}{(|x_1 - y_1| + 1)^{\beta_3 + \gamma_3 + 3}} \, dy_1 \, dx_1$$

$$\sim \frac{t^{-\beta_3 - \gamma_3 - 1}}{(|y_2| + |z_2|)^{\beta_2 + \gamma_2 + \beta_3 + \gamma_3 + 4}} \iint_{t^{-1} I^1 \times t^{-1} I^1} \left(1 + \log_+ \frac{1}{|x_1 - y_1|}\right) \, dy_1 \, dx_1$$

$$\lesssim \frac{t^{-\beta_3-\gamma_3-1}}{(|y_2|+|z_2|)^{\beta_2+\gamma_2+\beta_3+\gamma_3+4}}(t^{-1}|I^1|)^2(1+\log_+(t|I^1|^{-1}))$$

$$= \frac{1}{(|y_2|+|z_2|)^{\beta_2+\gamma_2}} \cdot \frac{1}{(|y_3|+|z_3|)^{\beta_3+\gamma_3}}|I^1|\left(\frac{|I^1|(|y_2|+|z_2|)}{|y_3|+|z_3|}+\frac{|y_3|+|z_3|}{|I^1|(|y_2|+|z_2|)}\right)^{-1}$$

$$\times \frac{1}{\prod_{i=2}^3(|y_i|+|z_i|)^2} \cdot \left(1+\log_+\frac{|y_3|+|z_3|}{|I^1|(|y_2|+|z_2|)}\right).$$

Thus, we are done. $\qquad\square$

With (A.2) at hand, similarly as in the linear case we can derive the following.

A.5. **Lemma.** *Let $m \in \mathcal{M}_Z^1$ and denote by T_m the corresponding Fourier multiplier operator.
Let $f_1, g_1 \in L^4(\mathbb{R})$, $f_{2,3}, g_{2,3} \in L^4(\mathbb{R}^2)$ and $h_1 \in L^2(\mathbb{R})$, $h_{2,3} \in L^2(\mathbb{R}^2)$. Then*

$$\langle T_m(f_1 \otimes f_{2,3}, g_1 \otimes g_{2,3}), h_1 \otimes h_{2,3}\rangle = \langle T_{m_{f_{2,3},g_{2,3},h_{2,3}}}(f_1,g_1), h_1\rangle,$$

where $m_{f_{2,3},g_{2,3},h_{2,3}}$ is a standard bilinear Coifman-Meyer multiplier in $\mathbb{R}$ satisfying the estimates

$$|(\,\mathrm{d}/\,\mathrm{d}\xi_1)^\alpha(\,\mathrm{d}/\,\mathrm{d}\eta_1)^\beta m_{f_{2,3},g_{2,3},h_{2,3}}(\xi_1,\eta_1)|$$
$$\lesssim \|m\|_{\mathcal{M}_Z^1}\|f_{2,3}\|_{L^4}\|g_{2,3}\|_{L^4}\|h_{2,3}\|_{L^2}(|\xi_1|+|\eta_1|)^{-\alpha-\beta}.$$

*Thus, $T_{m_{f_{2,3},g_{2,3},h_{2,3}}}$ is a convolution form bilinear Calderón-Zygmund operator. In particular,
there exists a standard bilinear Calderón-Zygmund kernel $K_{m,f_{2,3},g_{2,3},h_{2,3}}$ such that*

$$\|K_{m,f_{2,3},g_{2,3},h_{2,3}}\|_{CZ_1(\mathbb{R}^2)} \lesssim \|f_{2,3}\|_{L^4}\|g_{2,3}\|_{L^4}\|h_{2,3}\|_{L^2}.$$

Moreover, if $\operatorname{spt} f_1 \cap \operatorname{spt} g_1 \cap \operatorname{spt} h_1 = \emptyset$, then

$$\langle T_m(f_1 \otimes f_{2,3}, g_1 \otimes g_{2,3}), h_1 \otimes h_{2,3}\rangle$$
$$= \iiint K_{m,f_{2,3},g_{2,3},h_{2,3}}(x_1,y_1,z_1)f_1(y_1)g_1(z_1)h_1(x_1)\,\mathrm{d}y_1\,\mathrm{d}z_1\,\mathrm{d}x_1.$$

Conclusion. Notice that Lemma A.5 immediately implies that we have the weak bound-
edness property. Therefore, the bilinear multipliers satisfy Definition 3.5.

REFERENCES

[1] E. Airta, H. Martikainen, and E. Vuorinen, *Product space singular integrals with mild kernel regularity*, J. Geom. Anal. **32** (2022), article number 24. ↑19

[2] R. R. Coifman, R. Rochberg, and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. (2) **103** (1976), no. 3, 611–635. MR412721 ↑2

[3] D. Cruz-Uribe, J. M. Martell, and C. Pérez, *Extrapolation from A_∞ weights and applications*, J. Funct. Anal. **213** (2004), no. 2, 412–439. MR2078632 ↑2, 34

[4] J. Duoandikoetxea, *Extrapolation of weights revisited: new proofs and sharp bounds*, J. Funct. Anal. **260** (2011), no. 6, 1886–1901. ↑27, 29

[5] X. T. Duong, J. Li, Y. Ou, J. Pipher, and B. Wick, *Weighted estimates of singular integrals and commutators in the Zygmund dilation setting*, preprint, arXiv:1905.00999 (2019). ↑1, 2, 29

[6] R. Fefferman and J. Pipher, *Multiparameter operators and sharp weighted inequalities*, Amer. J. Math. **119** (1997), no. 2, 337–369. MR1439553 ↑2, 3, 39

[7] L. Grafakos and J. M. Martell, *Extrapolation of weighted norm inequalities for multivariable operators and applications*, J. Geom. Anal. **14** (2004), no. 1, 19–46. MR2030573 ↑27, 29

[8] L. Grafakos and S. Oh, *The Kato-Ponce inequality*, Comm. Partial Differential Equations **39** (2014), no. 6, 1128–1157. MR3200091 ↑3

[9] L. Grafakos and R. H. Torres, *Multilinear Calderón-Zygmund theory*, Adv. Math. **165** (2002), no. 1, 124–164. MR1880324 ↑3

[10] A. Grau de la Herrán and T. Hytönen, *Dyadic representation and boundedness of nonhomogeneous Calderón-Zygmund operators with mild kernel regularity*, Michigan Math. J. **67** (2018), no. 4, 757–786. MR3877436 ↑4

[11] Y. Han, J. Li, C.-C. Lin, and C. Tan, *Singular Integrals Associated with Zygmund Dilations*, J. Geom. Anal. **29** (2019), 2410–2455. ↑2

[12] I. Holmes, S. Petermichl, and B. D. Wick, *Weighted little bmo and two-weight inequalities for Journé commutators*, Anal. PDE **11** (2018), no. 7, 1693–1740. MR3810470 ↑1

[13] T. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. MR2912709 ↑1

[14] T. Hytönen, K. Li, H. Martikainen, and E. Vuorinen, *Multiresolution analysis and Zygmund dilations*, preprint, arXiv:2203.15777 (2022). ↑1, 2, 3, 5, 10, 14, 16, 18, 29, 30, 31, 35, 38, 40

[15] T. Kato and G. Ponce, *Commutator estimates and the Euler and Navier-Stokes equations*, Comm. Pure Appl. Math. **41** (1988), no. 7, 891–907. MR951744 ↑3

[16] A. Lerner, S. Ombrosi, C. Pérez, R. Torres, and R. Trujillo-González, *New maximal functions and multiple weights for the multilinear Calderón–Zygmund theory*, Adv. Math. **220** (2009), no. 4, 1222–1264. ↑3

[17] K. Li, H. Martikainen, Y. Ou, and E. Vuorinen, *Bilinear representation theorem*, Trans. Amer. Math. Soc. **371** (2019), no. 6, 4193–4214. MR3917220 ↑26

[18] K. Li, H. Martikainen, and E. Vuorinen, *Bilinear Calderón-Zygmund theory on product spaces*, J. Math. Pures Appl. (9) **138** (2020), 356–412. MR4098772 ↑3

[19] _____, *Genuinely multilinear weighted estimates for singular integrals in product spaces*, Adv. Math. **393** (2021), 108099. ↑3, 28, 29

[20] H. Martikainen, *Representation of bi-parameter singular integrals by dyadic operators*, Adv. Math. **229** (2012), no. 3, 1734–1761. MR2871155 ↑1

[21] D. Müller, F. Ricci, and E. M. Stein, *Marcinkiewicz multipliers and multi-parameter structure on Heisenberg (-type) groups. I*, Invent. Math. **119** (1995), no. 2, 199–233. MR1312498 ↑1

[22] A. Nagel and S. Wainger, L^2 *boundedness of Hilbert transforms along surfaces and convolution operators homogeneous with respect to a multiple parameter group*, Amer. J. Math. **99** (1977), 761–785. ↑1

[23] F. Nazarov, S. Treil, and A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. **190** (2003), no. 2, 151–239. MR1998349 ↑1

[24] F. Ricci and E. M. Stein, *Multiparameter singular integrals and maximal functions*, Ann. Inst. Fourier (Grenoble) **42** (1992), no. 3, 637–670. MR1182643 ↑2

(E.A.) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF JYVÄSKYLÄ, P.O. BOX 35 (MAD), FI-40014 UNIVERSITY OF JYVÄKYLÄ, FINLAND
Email address: emil.t.airta@jyu.fi

(K.L.) CENTER FOR APPLIED MATHEMATICS, TIANJIN UNIVERSITY, WEIJIN ROAD 92, 300072 TIANJIN, CHINA
Email address: kli@tju.edu.cn

(H.M.) DEPARTMENT OF MATHEMATICS AND STATISTICS, WASHINGTON UNIVERSITY IN ST. LOUIS, 1 BROOKINGS DRIVE, ST. LOUIS, MO 63130, USA
Email address: henri@wustl.edu