Off-Diagonal Estimates for Bilinear Commutators

Tuomas Oikari

Received: 18 August 2021 / Accepted: 22 March 2022 / Published online: 15 April 2022
© The Author(s) 2022

Abstract
We find a minimal notion of non-degeneracy for bilinear singular integral operators $T$ and identify testing conditions on the multiplying function $b$ that characterize the $L^p \times L^q \rightarrow L^r$, $1 < p, q < \infty$ and $r > \frac{1}{2}$, boundedness of the bilinear commutator $[b, T](f, g) = bT(f, g) - T(bf, g)$. Our arguments cover almost all arrangements of the integrability exponents $p, q, r$ with a single open problem presented in the end. Additionally, the arguments extend to the multilinear setting.

Keywords Calderón–Zygmund operators · Singular integrals · Commutators · Multilinear analysis

Mathematics Subject Classification (2010) 42B20

1 Introduction

The study of commutator estimates have their roots in the work of Nehari [20], where the boundedness of the commutator of the Hilbert transform and a multiplying function $b$,

$$[b, H]f(x) = b(x)Hf(x) - H(bf)(x), \quad Hf(x) = \text{p.v.} \int_{\mathbb{R}} f(x - y) \frac{dy}{y},$$

was characterized through Hankel operators. Later, Coifman and Rochberg and Weiss [6] extended Nehari’s result by real analytic methods and showed that

$$\|b\|_{\text{BMO}} \lesssim \sum_{j=1}^{d} \|[b, \mathcal{R}_j]\|_{L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)} \lesssim \|b\|_{\text{BMO}} := \sup_{I} \int_{I} |b - \langle b \rangle_I|, \quad p \in (1, \infty),$$

was characterized through Hankel operators. Later, Coifman and Rochberg and Weiss [6] extended Nehari’s result by real analytic methods and showed that

$$\|b\|_{\text{BMO}} \lesssim \sum_{j=1}^{d} \|[b, \mathcal{R}_j]\|_{L^p(\mathbb{R}^d) \rightarrow L^p(\mathbb{R}^d)} \lesssim \|b\|_{\text{BMO}} := \sup_{I} \int_{I} |b - \langle b \rangle_I|, \quad p \in (1, \infty),$$

(1.1)
where the supremum is taken over all cubes $I \subset \mathbb{R}^d$ and $\langle b \rangle_I = \frac{1}{|I|} \int_I b$. The upper bound in Eq. 1.1 was proved for many Calderón-Zygmund operators (CZO) $T$, i.e. for bounded singular integral operators (SIOs), while with the lower bounds they had to employ all the $d$ Riesz transforms simultaneously; commutator upper bounds are usually valid for all bounded SIOs, while the lower bounds require some non-degeneracy. The lower bound in Eq. 1.1 was improved separately by both Janson [13] and Uchiyama [22] to $\|b\|_{BMO} \lesssim \|\langle b, T \rangle\|_{L^p (\mathbb{R}^d) \rightarrow L^q (\mathbb{R}^d)}$ for a wide class of singular integrals $T$ that incudes any single Riesz transform. Janson’s proof also gives the following off-diagonal characterization of the boundedness of the commutator

$$\|\langle b, T \rangle\|_{L^p (\mathbb{R}^d) \rightarrow L^q (\mathbb{R}^d)} \sim \|b\|_{\hat{c}^{0,0}} \sim \sup_{Q} \ell(Q)^{-\alpha} \int_Q |b - \langle b \rangle_Q|, \quad \alpha = d \left( \frac{1}{p} - \frac{1}{q} \right),$$

when $1 < p < q < \infty$ and where the supremum is taken over all cubes. The off-diagonal characterizations in the case $1 < q < p < \infty$ turned out to be harder and was only recently solved by an application of the approximate weak factorization (awf) argument in Hytönen [11],

$$\|\langle b, T \rangle\|_{L^p (\mathbb{R}^d) \rightarrow L^q (\mathbb{R}^d)} \sim \|b\|_{L^q}, \quad \frac{1}{q} = \frac{1}{s} + \frac{1}{p}, \quad 1 < q < p < \infty.$$  (1.3)

Each of the above cases Eqs. 1.1, 1.2 and 1.3 have applications and it is often important to have a full characterization, i.e. both commutator upper and lower bounds. Commutator estimates imply factorization results for Hardy spaces, see [6], they have applications in partial differential equations by compensated compactness, div-curl lemmas, see [5], and they have been crucial in the recent investigations to the Jacobian problem, see Lindberg [18] and [11].

The awf argument is strong in that it gives a unified approach to all of the three cases, in that it works for many singular integrals with kernels satisfying only minimum non-degeneracy assumptions, both variable and rough kernels, and in that it is flexible enough to grant e.g. multi-parameter and multilinear extensions (as we will shortly see). For the multi-parameter variants of the awf argument see Airta, Hytönen, Li, Martikainen and Oikari [1], and Oikari [21], where respectively the commutators

$$[T_2, [T_1, b]], [b, T] : L^{p_1} (\mathbb{R}^{d_1}; L^{p_2} (\mathbb{R}^{d_2})) \rightarrow L^{p_1} (\mathbb{R}^{d_1}; L^{p_2} (\mathbb{R}^{d_2}))$$

were treated. On the line (1.4) $1 < p_1, p_2, q_1, q_2 < \infty$, $T_i$ is a one-parameter CZO on $\mathbb{R}^{d_i}$, for $i = 1, 2$, and $T$ is a bi-parameter CZO on $\mathbb{R}^{d_1 + d_2}$. The adaptability of the awf argument to the bi-parameter settings was not perfect and for both commutators on the line (1.4) the characterization of boundedness in terms of testing conditions on the symbol $b$ is for some arrangements of the integrability exponents still open.

In this article we extend the awf argument to the bilinear setting and study the two bilinear commutators

$$[b, T]_1 (f, g) = bT(f, g) - T(bf, g), \quad [b, T]_2 (f, g) = bT(f, g) - T(f, bg)$$
as mappings $L^p \times L^q \rightarrow L^r$ for $r > \frac{1}{2}$ and $1 < p, q < \infty$. Our cases separate accordingly to the following three conditions

$$\frac{1}{r} < \frac{1}{p} + \frac{1}{q}, \quad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \quad \frac{1}{r} > \frac{1}{p} + \frac{1}{q}. \quad (1.5)$$

If $r > 1$ then we are in the Banach range of exponents and if $r \leq 1$ then we are in the quasi Banach range of exponents. In Chaffee [4] the necessity of $b \in BMO$ on the diagonal in the
Banach range of exponents was shown with kernels whose multiplicative inverses are locally expandable as a Fourier series. A unified approach to the diagonal and sub-diagonal cases was given in Guo, Lian and Wu [10], which covers the diagonal in the whole quasi Banach range, however on the sub-diagonal they only treat the linear case. In addition to involving the new super-diagonal case in the bilinear setting, our results generalize previous work in the diagonal and sub-diagonal cases: the definition of non-degeneracy is weaker than those supposed in [4, 10, 12, 15, 17]; the awf argument allows us to consider complex valued functions \( b \), whereas [17] was limited to the real valued case; the full quasi-Banach range is reached in the diagonal and sub-diagonal cases, whereas [4] is limited to the Banach range; and in that the awf argument encompasses bilinear singular integrals with both variable and rough kernels. Lastly, due to us studying the quasi Banach range, the arguments involve additional twists absent from previous research articles. Our full results are recorded as Theorems 4.1, 4.2, 4.15 and 5.21, the following being a condensed version.

Theorem 1.6 Let \( b \in L^1_{\text{loc}}(\mathbb{R}^d; \mathbb{C}) \), let \( T \) be a non-degenerate bilinear Calderón-Zygmund operator, let \( \frac{1}{2} < r < \infty \) and \( 1 < p, q < \infty \). Then, for \( i = 1, 2 \), there holds that

\[
\| [b, T]_i \|_{L^p \times L^q \to L^r} \sim \begin{cases} \| b \|_{\text{BMO}}, & \text{if} \quad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \\ \| b \|_{C^{0,0}}, & \text{if} \quad \frac{1}{r} < \frac{1}{p} + \frac{1}{q}, \\ \| b \|_{L^r}, & \text{if} \quad \frac{1}{r} > \frac{1}{p} + \frac{1}{q}, \end{cases}
\]

\[ \alpha = d \left( \frac{1}{p} + \frac{1}{q} - \frac{1}{r} \right), \]

\[ \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \quad r \geq 1. \]

2 Definitions and Preliminaries

2.1 Basic Notation

We let \( \Sigma = \Sigma(\mathbb{R}^d) \) denote the linear span of indicator functions of cubes on \( \mathbb{R}^d \). Similarly we denote \( L^1_{\text{loc}}(\mathbb{R}^d) = L^1_{\text{loc}} \mathbb{R}^d = \mathcal{I} \), and so on, mostly leaving out the ambient space if this information is obvious. We denote averages with \( \langle f \rangle_A = \frac{1}{|A|} \int_A f \), where \( |A| \) denotes the Lebesgue measure of the set \( A \). The indicator function of a set \( A \) is denoted by \( 1_A \).

In this article we study the \( L^p \times L^q \to L^r \) boundedness and hence it is useful to denote \( \sigma(p, q)^{-1} = p^{-1} + q^{-1} \); then, Hölder’s inequality writes as \( \| fg \|_{L^r(p,q)} \leq \| f \|_{L^p} \| g \|_{L^q} \).

Lastly, we denote \( A \lesssim B \), if \( A \leq CB \) for some constant \( C > 0 \) depending only on the dimension of the underlying space, on integration exponents and on other absolute constants appearing in the assumptions that we do not care about. Then \( A \sim B \), if \( A \lesssim B \) and \( B \lesssim A \). Subscripts on constants \( (C_{a,b,c,...}) \) and quantifiers \( (\lesssim_{a,b,c,...}) \) signify their dependence on those subscripts.

2.2 Bilinear Singular Integrals

We denote the diagonal with

\[ \Delta = \{(x, y, z) \in (\mathbb{R}^d)^3 : x = y = z\} \]

and say that a mapping \( K : (\mathbb{R}^d)^3 \setminus \Delta \to \mathbb{C} \) is a bilinear Calderón-Zygmund kernel if it satisfies the size estimate

\[
|K(x, y, z)| \leq C_K (|x - y| + |x - z|)^{-2d}, \quad (2.1)
\]
and the regularity estimate
\[ |G(x, y, z) - G(x', y, z)| \leq \omega \left( \frac{|x - x'|}{|x - y| + |x - z|} \right)(|x - y| + |x - z|)^{-2d}, \]  \hspace{1cm} (2.2)

for \( G \in \{ K, K^{1*}, K^{2*} \} \), whenever \( |x - x'| \leq \frac{1}{2} \max(|x - y|, |x - z|) \). Here the function \( \omega \) is increasing, subadditive, and such that \( \omega(0) = 0 \) and \( \| \omega \|_{\text{Dini}} = \int_0^1 \omega(t) \frac{dt}{t} < \infty \). We also assume that the appearing constants \( C_K \), \( \| \omega \|_{\text{Dini}} \) are the best possible. The class of all such kernels we denote with \( \text{CZ}(2, d, \omega) \) and the related norm is \( \| K \|_{\text{CZ}(2, d, \omega)} = C_K + \| \omega \|_{\text{Dini}} \).

**Definition 2.3** A bilinear operator \( T : \Sigma^2 \rightarrow L^1_{\text{loc}} \) is said to be a (variable kernel) bilinear singular integral, if there exists a bilinear kernel \( K \in \text{CZ}(2, d, \omega) \) so that
\[
\langle T(f_1, f_2), g \rangle = \int \int \int K(x, y, z) f_1(y) f_2(z) g(x) \, dy \, dz \, dx
\]
for all triples \( f_1, f_2, g \in \Sigma \) satisfying \( \cap_{i=1}^2 \text{spt}(f_i) \cap \text{spt}(g) = \emptyset \).

**Definition 2.4** A bilinear operator \( T_\Omega : \Sigma^2 \rightarrow L^1_{\text{loc}} \) is said to be a rough bilinear singular integral if
\[
T_\Omega(f_1, f_2)(x) = \lim_{\varepsilon \to 0} \int \int \max(|x-y|, |x-z|) > \varepsilon K_\Omega(x, y, z) f_1(y) f_2(z) \, dy \, dz,
\]
where
\[
K_\Omega(x, y, z) = \frac{\Omega((x-y, x-z)')}{|(x-y, x-z)|^{2d}}, \quad \Omega(h') = \Omega\left(\frac{h}{|h|}\right), \quad \Omega \in L^1(\mathbb{S}^{d-1}), \quad \int_{\mathbb{S}^{d-1}} \Omega = 0.
\]

### 2.3 Truncations

For \( K \in \text{CZ}(2, d, \omega) \) define the truncated operator \( T_\varepsilon \) as
\[
T_\varepsilon(f, g)(x) = \int \int \max(|x-y|, |x-z|) > \varepsilon K(x, y, z) f(y) g(z) \, dy \, dz.
\] \hspace{1cm} (2.5)

Then, a particular case of Cotlar’s inequality in the bilinear setting states that
\[
T^*(f, g) = \sup_{\varepsilon > 0} |T_\varepsilon(f, g)| \lesssim \|T(f, g)| + Mf Mg, \quad K \in \text{CZ}(2, d, \cdot), \quad \delta \in (0, 1), \] \hspace{1cm} (2.6)

where \( M \) is the Hardy-Littlewood maximal operator. For Eq. 2.6, see e.g. Grafakos, Torres [9]. Since \( T, M \) are bounded, it follows from Eq. 2.6 that \( \sup_{\varepsilon > 0} \| T_\varepsilon \|_{L^p \times L^q \rightarrow L^s} < \infty \), for \( 1 < p, q < \infty \). We have use for such uniform upper bounds for the sub-diagonal commutator upper bounds, see the point (3) in Definition 2.7 below.

For rough kernels Cotlar’s inequality was not found. However, to achieve a uniform bound on the truncations we need less. It was very recently shown in Theorem 1.1. of [8] that under the assumptions \( \Omega \in L^q(\mathbb{S}^{d-1}) \), for some \( q > \frac{4}{3} \), there holds that \( \| T_\Omega^* \|_{L^2 \times L^2 \rightarrow L^1} \lesssim \| \Omega \|_{L^q(\mathbb{S}^{d-1})} \) and this implies a uniform bound of the desired type.

### 2.4 Boundedness Assumptions on \( T \)

The majority of this article is devoted to proving commutator lower bounds and there we do not need any boundedness assumptions on the operator \( T \) – only non-degeneracy assumptions on the kernel \( K \) of \( T \) and some very weak regularity conditions, see Section 2.5 below.
For the upper bounds we require some boundedness on $T$ and this will vary depending whether we are on the sub-diagonal, diagonal or the super-diagonal case.

**Definition 2.7** A bilinear Calderón-Zygmund operator refers to a bilinear singular integral $T$ that enjoys some boundedness properties, and in this article this will be one of the following:

1. $\Omega \in L^{\infty}(\mathbb{S}^{2d-1})$,
2. $\|T\|_{L^p \times L^q \rightarrow L^s(p,q)} < \infty$ for a single tuple of exponents $p, q \in (1, \infty)$,
3. $\sup_{t > 0} \|T_t\|_{L^p \times L^q \rightarrow L^s(p,q)} < \infty$ for a single tuple of exponents $p, q \in (1, \infty)$.

We assume (1v) and (1Ω) respectively for the diagonal upper bound for bilinear SIOs with variable and rough kernels. The points (2) and (3) we assume for bilinear SIOs of both variable and rough kernels, (2) for the super-diagonal upper bound and (3) for the sub-diagonal upper bound.

**Remark 2.8** We do not believe that the above listed conditions are necessary optimal, but these are chosen to give a sense of what is sufficient, without going too much into details. After all, we are in this article mainly interested in commutator lower bounds.

### 2.5 Bilinear Non-Degeneracy

We first recall the definition of non-degeneracy for kernels of linear singular integrals with variable and rough kernels. Originally, this is from [11].

**Definition 2.9** A variable kernel $K : \mathbb{R}^d \times \mathbb{R}^d \setminus \Delta \rightarrow \mathbb{C}$ is said to be non-denegerate, if given $y \in \mathbb{R}^d$ and $r > 0$ there exists a point $x \in \mathbb{R}^d$ so that $|x - y| > r$, $|K(x, y)| \gtrsim r^{-d}$.

A kernel of a rough linear SIO $K_\Omega$ is said to be non-degenerate, provided that $\Omega$ is not the zero function, see also Definition 2.4 below.

For bilinear singular integrals with variable kernel we set the following.

**Definition 2.10** A kernel $K : (\mathbb{R}^d)^3 \setminus \Delta \rightarrow \mathbb{C}$ is said to be non-degenerate if both of the following items hold:

1. For all $y \in \mathbb{R}^d$ and $r > 0$ there exist $x, z \in \mathbb{R}^d$ such that
   $\max_{a, b \in \{x, y, z\}} |a - b| > r$, $|K(x, y, z)| \gtrsim r^{-2d}$.
2. For all $z \in \mathbb{R}^d$ and $r > 0$ there exist $x, y \in \mathbb{R}^d$ such that
   $\max_{a, b \in \{x, y, z\}} |a - b| > r$, $|K(x, y, z)| \gtrsim r^{-2d}$.

It is immediate from the size estimate (2.1) that if $y, r$ and then $x, z$ are given as in the item (1) of Definition 2.10 then $\max_{a, b \in \{x, y, z\}} |a - b| \sim r$. Indeed, to see this, we simply check that $r^{-2d} \lesssim |K(x, y, z)| \lesssim (|x - y| + |x - z|)^{-2d} \lesssim (\max_{a, b \in \{x, y, z\}} |a - b|)^{-2d} \lesssim r^{-2d}$, (2.11)
which shows the claim. Similarly for the point (2).

Remark 2.12 We will use the assumption (1) to prove Theorem 1.6 for the index \( i = 1 \) and the assumption (2) for the index \( i = 2 \). It follows that we will run the proofs of all of our results through with the assumption (1) of Definition 2.10 and it is clear how to modify them to get the case \( i = 2 \).

For kernels of rough singular integrals we set the following.

Definition 2.13 A kernel \( K_\Omega \) is non-degenerate if \( \Omega \neq 0 \), i.e. it has at least one non-zero Lebesgue point \( \theta = (\theta_1, \theta_2) \in S^{2d-1} \).

Remark 2.14 The bilinear Riesz transforms, one of which is

\[
\mathcal{R}_j(f, g)(x) = p.v. \int \int \frac{x_i - y_i}{(|x - y| + |x - z|)^{2d+1}} f(y)g(z) \, dy \, dz,
\]

satisfy each of the above boundedness properties (1), (2), (3) as in Definition 2.7, their kernels satisfy both items (1) and (2) as in Definition 2.10, and are also non-degenerate when considered as rough bilinear SIOs, as in Definition 2.13. Consequently, Theorem 1.6 is valid as stated for \( T = \mathcal{R}_j \).

In Li [17] the following definition of non-degeneracy is given; to contrast it with the non-degeneracy we name it the strong non-degeneracy.

Definition 2.15 A kernel \( K : (\mathbb{R}^d)^3 \setminus \Delta \to \mathbb{C} \) is said to be strongly non-degenerate if for each given point \( y \in \mathbb{R}^d \) and \( r > 0 \) there exists a point \( x \in B(y, r)^c \) such that

\[
|K(x, y, y)| \gtrsim r^{-2d}.
\]

It is immediate from the definitions that strong non-degeneracy is stronger than non-degeneracy.

Proposition 2.16 Let \( K \) be a strongly non-degenerate kernel. Then, the kernel \( K \) is non-degenerate.

Proof We only show the point (1) from Definition 2.10. Fix a point \( y \in \mathbb{R}^d \), then by strong non-degeneracy there exists a point \( x \in B(y, r)^c \) so that \( |K(x, y, y)| \gtrsim r^{-2d} \). Consequently, it remains to write the previous estimate as \( |K(x, y, z)| \gtrsim r^{-2d} \), where \( z = y \) and to notice that \( x \in B(y, r)^c \).

Definition 2.17 We say that a bilinear SIO \( T \) is non-degenerate if its kernel \( K \) is non-degenerate. Similarly, a bilinear CZO is non-degenerate, if it is a bilinear non-degenerate SIO that satisfies at least some of the properties (1\( \nu \)), (1\( \Omega \)), (2), (3) as in Definition 2.7.

3 Bilinear Approximate Weak Factorization

When proving the commutator lower bounds we do not need the full strength of the kernel assumption (2.2) and we will replace this with the following weaker assumption: the
function $\omega$ satisfies $\omega(0) = 0$, is increasing, is subadditive and such that

$$|G(x, y, z) - G(x', y, z)| \leq \omega \left( \frac{|x - x'|}{|x - y| + |x - z|} \right) |x - y| \leq |x - z|^{2d} \quad (3.1)$$

for $G \in \{ K, K^{1*}, K^{2*} \}$, whenever $|x - x'| \leq \frac{1}{2} \max(|x - y|, |x - z|)$. Another strengthening is in that the awf argument only requires us to consider the following off-support information on the kernel $K$,

$$T(f \mathcal{Q}^0, g \mathcal{Q}^2)(y) = \int_{\mathcal{Q}^2} \int_{\mathcal{Q}^0} K(x, y, z) f(x) g(z) \, dx \, dz, \quad y \in \mathcal{Q}^1,$$

where $\mathcal{Q}^0, \mathcal{Q}^1, \mathcal{Q}^2$ are cubes of the same size such that $\max_{i=0,2} \text{dist}(\mathcal{Q}^i, \mathcal{Q}^1) \sim \ell(\mathcal{Q}^1)$. To press the point, there is no reference whatsoever to the operator $T$ and everything is defined with the kernel $K$ only.

We move to prove the main technical Propositions 3.2 and 3.14. Recall that we only need the assumption (1) from Definition 2.10 to show the lower bounds for the commutator $[b, T]_i$. In the following we always work with three cubes $\mathcal{Q}^0, \mathcal{Q}^1, \mathcal{Q}^2$ and variables are reserved to be used as follows, $x \in \mathcal{Q}^0, y \in \mathcal{Q}^1, z \in \mathcal{Q}^2$.

**Proposition 3.2** Let $K$ either

1. be a non-degenerate bilinear kernel that satisfies the estimates (2.1), (3.1), or
2. be a rough non-degenerate bilinear kernel.

Let $\mathcal{Q}^1 \subset \mathbb{R}^d$ be a cube with centre point $c_{\mathcal{Q}^1}$ and let $\mathcal{D}^0$ and $\mathcal{D}^2$ be arbitrary dyadic grids. Then, there exists a constant $A \geq 3$, cubes $\mathcal{Q}^i \in \mathcal{D}^i$ and points $c_{\mathcal{Q}^i} \in \mathcal{Q}^i$, $i = 0, 2$, so that the following items hold.

(i) The cubes are separated and have size as follows

$$\max_{a \in \{0,2\}} |c_{\mathcal{Q}^a} - c_{\mathcal{Q}^i}| \sim A \ell(\mathcal{Q}^1), \quad \ell(\mathcal{Q}^0) \sim \ell(\mathcal{Q}^1) \sim \ell(\mathcal{Q}^2). \quad (3.3)$$

(ii) There holds that

$$|K(c_{\mathcal{Q}^0}, c_{\mathcal{Q}^1}, c_{\mathcal{Q}^2})| \sim A^{-2d} |\mathcal{Q}^1|^{-2}. \quad (3.4)$$

(iii) For all $y \in \mathcal{Q}^1$, there holds that

$$\int_{\mathcal{Q}^0} \int_{\mathcal{Q}^2} |K(x, y, z) - K(c_{\mathcal{Q}^0}, c_{\mathcal{Q}^1}, c_{\mathcal{Q}^2})| \, dx \, dz \lesssim \omega(A^{-1}) A^{-2d}, \quad (3.5)$$

where $\omega(A^{-1}) \to 0$ as $A \to \infty$, and

$$\int_{\mathcal{Q}^0} \int_{\mathcal{Q}^2} K(x, y, z) \, dx \, dz \sim \int_{\mathcal{Q}^0} \int_{\mathcal{Q}^2} K(x, y, z) \, dx \, dz \sim A^{-2d}. \quad (3.6)$$

Moreover, the symmetric estimates to Eqs. 3.5 and 3.6 where we always integrate over any two of the cubes $\mathcal{Q}^0, \mathcal{Q}^1, \mathcal{Q}^2$ with the corresponding variables $x, y, z, h$ hold.

**Remark 3.7** We only need to set up $\mathcal{Q}^i \in \mathcal{D}^i$, $i = 0, 2$, for the study of the super-diagonal case $r = 1$.

**Proof of the Case (1):** As we will mostly manage without the property $\mathcal{Q}^i \in \mathcal{D}^i$, we first find any two cubes $\mathcal{Q}^i$, $i = 1, 2$, satisfying the rest of the claims.
We fix a cube $Q^1 \subset \mathbb{R}^d$ and denote its centre point with $c_{Q^1}$. Let $r = A \text{diam}(Q^1)/2$ and by non-degeneracy find two points $c_{Q^0}, c_{Q^2}$ so that $c_{Q^2} \in B(c_{Q^1}, r)^c$ (the case $c_{Q^0} \in B(c_{Q^1}, r)^c$ is completely symmetric) and

$$|K(c_{Q^0}, c_{Q^1}, c_{Q^2})| \sim r^{-2d} \sim A^{-2d}|Q^0|^{-2}. \quad (3.8)$$

The fact that we have $\sim$ above where indicated by $\ast$ follows from the discussion after Definition 2.10, see line (2.11). Hence the claim (3.4) holds. Then, we let

$$Q^0 = (c_{Q^0} - c_{Q^1}) + Q^1, \quad Q^2 = (c_{Q^2} - c_{Q^1}) + Q^1$$

be the cubes respectively with the centre points $c_{Q^0}$ and $c_{Q^2}$. Then, it is clear that the claims on the line (3.3) hold.

Towards the remaining two claims, we first estimate

$$|K(x, y, z) - K(c_{Q^0}, c_{Q^1}, c_{Q^2})| \leq |K(x, y, z) - K(c_{Q^0}, y, z)| + |K(c_{Q^0}, y, z) - K(c_{Q^0}, c_{Q^1}, z)| + |K(c_{Q^0}, c_{Q^1}, z) - K(c_{Q^0}, c_{Q^1}, c_{Q^2})|.$$ 

Then, as for all points $x \in Q^0, y \in Q^1, z \in Q^2$, we have

$$|x - c_{Q^0}| \leq \frac{1}{2}|c_{Q^0} - z| \leq \frac{1}{2}\max(|c_{Q^0} - y|, |c_{Q^0} - z|),$$

(which follows immediately by $c_{Q^2} \in B(x, Ar)^c, z \in Q^2$ and that $A \geq 3$), the regularity estimate (3.1) is applicable and we estimate first of the three intermediate terms as

$$|K(x, y, z) - K(c_{Q^0}, y, z)| \lesssim \omega \left(\frac{|x - c_{Q^0}|}{|c_{Q^0} - z| + |c_{Q^0} - y|}\right)(|c_{Q^0} - z| + |c_{Q^0} - y|)^{-2d}$$

$$\lesssim \omega \left(\frac{1/2 \text{diam}(Q^0)}{A/3 \text{diam}(Q^0)}\right)(A \text{diam}(Q^0))^{-2d} \lesssim \omega(A^{-1})A^{-2d}|Q^1|^{-2},$$

where in the last estimate we used the sub-additivity of $\omega$.

The remaining two terms estimate similarly and consequently we find that

$$|K(x, y, z) - K(c_{Q^0}, c_{Q^1}, c_{Q^2})| \lesssim \omega(A^{-1})A^{-2d}|Q^1|^{-2}. \quad (3.9)$$

Now, by choosing $A$ large enough, subtracting and adding $K(c_{Q^0}, c_{Q^1}, c_{Q^2})$ and using Eqs. 3.8 and 3.9 we actually find that

$$|K(x, y, z)| \sim A^{-2d}|Q^1|^{-2}, \quad (3.10)$$

which is an improvement of Eq. 3.4. Similarly, by using the estimates (3.8) and (3.9), the claims (3.5) and (3.6) follow immediately.

We still need to argue that we can arrange $Q^i \in D^i$, for $i = 0, 2$. Assume that we have shown the claims for the triple of cubes $\tilde{Q}^0, \tilde{Q}^1, \tilde{Q}^2$ with centre points $c_{\tilde{Q}^0}, c_{\tilde{Q}^1}, c_{\tilde{Q}^2}$. Then, we let $Q^i_d \in D^i$ be the largest dyadic cube such that $Q^i_d \subset \tilde{Q}^i$. Now the cubes $Q^i_d$ clearly satisfy the claims on the line (3.3), and as Eq. 3.10 is valid especially with the triple of points $(c_{Q^0_d}, c_{Q^1_d}, c_{Q^2_d})$, we find

$$|K(c_{Q^0_d}, c_{Q^1_d}, c_{Q^2_d})| \sim A^{-2d}|Q^1_d|^{-2} \sim A^{-2d}|Q^1_d|^{-2} \quad (3.11)$$

and hence (3.4) is checked. The remaining claims are similarly immediate (with $Q^i_d$ in place of $Q^i$) and follow as before. \qed
Off-Diagonal Estimates...  

Proof of the Case (2): We first check the claims with balls in place of cubes. By the non-degeneracy assumption let \( \theta = (\theta_0, \theta_2) \in \mathbb{S}^{2d-1} \) be a non-zero Lebesgue point of \( \Omega \). Then, fix a ball \( B^i \) with centre \( c_{B^i} \) and radius \( r \). Let the points \( c_{B^0}, c_{B^2} \) be defined by the following identities

\[
c_{B^0} - c_{B^1} = r A \theta_0, \quad c_{B^0} - c_{B^2} = r A \theta_2,
\]

and let \( B^i \) be a ball with centre \( c_{B^i} \) and radius \( r \). It is then clear that Eq. 3.3 holds and that

\[
K_\Omega(c_{B^0}, c_{B^1}, c_{B^2}) = \frac{\Omega(c_{B^0} - c_{B^1}, c_{B^0} - c_{B^2})}{|c_{B^0} - c_{B^1}, c_{B^0} - c_{B^2}|^{2d}} = \frac{\Omega((r A \theta_0, r A \theta_2))}{|(r A \theta_0, r A \theta_2)|^{2d}} \sim A^{-2d} |B^1|^{-2d} |\Omega(\theta_0, \theta_1)|,
\]

hence (3.4) holds.

It remains to check (3.5) and (3.6). Let \( x \in B^0, y \in B^1, z \in B^2 \) be arbitrary and write

\[
x = c_{B^0} + ru_x, \quad y = c_{B^0} - r A \theta_0 + ru_y, \quad z = c_{B^0} - r A \theta_2 + u_z,
\]

for a specific \( u_a \in B(0, 1) \) depending on the parameter \( a \in \{x, y, z\} \). To relax notation we write \( \Omega(h^i) = \Omega(h) \) and \( K_\Omega = K \). Then, we have

\[
K(x, y, z) - K(c_{B^0}, c_{B^1}, c_{B^2}) = \frac{\Omega(x - y, x - z)}{|(x - y, x - z)|^{2d}} - \frac{\Omega(c_{B^0} - c_{B^1}, c_{B^0} - c_{B^2})}{|(c_{B^0} - c_{B^1}, c_{B^0} - c_{B^2})|^{2d}} = \frac{\Omega(r A \theta_0 + r (u_x - u_y), r A \theta_2 + r (u_x - u_z))}{|(r A \theta_0 + r (u_x - u_y), r A \theta_2 + r (u_x - u_z))|^{2d}} - \frac{\Omega(r A \theta_0, A r \theta_2)}{|r A \theta_0, A r \theta_2|^{2d}} = (r A)^{-2d} \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \left| \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \right|^{-2d} - \Omega(\theta_0, \theta_2) = (r A)^{-2d} (I + II),
\]

where

\[
I = \left( \Omega \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) - \Omega(\theta_0, \theta_2) \right) \left| \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \right|^{-2d}
\]

and

\[
II = \Omega(\theta_0, \theta_2) \left( \left| \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \right|^{-2d} - 1 \right).
\]

With a choice of \( A \) large enough we find that

\[
|II| \leq \Omega(\theta_0, \theta_2) \left| 1 - \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \right|^{2d} \leq \Omega(\theta_0, \theta_2) \left| \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \right|^{2d} \leq \left( \frac{u_x - u_y}{A}, \frac{u_x - u_z}{A} \right) \leq A^{-1},
\]

where as indicated by * the mean value theorem was applied with \( x \mapsto x^{2d} \) and we used the estimate \( |u_x - u_y| + |u_x - u_z| \leq 1 \). Hence, we find that

\[
\int_{B^0} \int_{B^2} (r A)^{-2d} |II| \, dx \, dz \leq \omega_{II}(A^{-1}) A^{-2d}, \quad \omega_{II}(A^{-1}) = A^{-1}.
\]
With a fixed \( y \), the point \( u_x - u_y \) varies over \( B(0, 2) \) and with a fixed \( y \), \( x \) the point \( u_x - u_z \) varies over \( B(0, 2) \). Hence, we estimate

\[
\int_{B^0} \int_{B^2} (Ar)^{-2d} |I| \, dx \, dz
\]

\[
\leq A^{-2d} \int_{B^0} \int_{B^2} |\Omega(0_1 + \frac{u_x - u_y}{A}, 0_2 + \frac{u_x - u_z}{A}) - \Omega(0_1, 0_2)| \, dx \, dz
\]

\[
\leq A^{-2d} \int_{B(0,2)} \int_{B(0,2)} |\Omega(0_1 + \frac{u_x - u_y}{A}, 0_2 + \frac{t}{A}) - \Omega(0_1, 0_2)| \, dx \, dz
\]

\[
\leq A^{-2d} \int_{B(0,2)} \int_{B(0,2)} |\Omega(0_1 + s, 0_2 + t) - \Omega(0_1, 0_2)| \, ds \, dt
\]

\[
= A^{-2d} \omega_I(A^{-1}), \quad (3.13)
\]

where we denote

\[
\omega_I(A^{-1}) = \int_{B(0, \frac{2}{A})} \int_{B(0, \frac{2}{A})} |\Omega(0_1 + s, 0_2 + t) - \Omega(0_1, 0_2)| \, dx \, ds \, dt
\]

and clearly \( \omega_I(A^{-1}) \to 0 \) as \( A \to \infty \), by \( \theta = (0_1, 0_2) \) being a Lebesgue point of \( \Omega \). Having the preceding estimate together with Eq. 3.12 shows Eq. 3.5,

\[
\int_{B^0} \int_{B^2} |K(x, y, z) - K(c_{B^0}, c_{B^1}, c_{B^2})| \, dy \, dz \leq \int_{B^0} \int_{B^2} (Ar)^{-2d} |I + II| \, dx \, dz
\]

\[
\leq (\omega_I(A^{-1}) + \omega_{II}(A^{-1})) A^{-2d} = \omega(A^{-1}) A^{-2d}.
\]

As before, Eq. 3.6 follows from Eqs. 3.4 and 3.5.

Lastly, we replace the balls with the desired cubes. Let \( Q^1 \) be a cube with centre point \( c_{Q^1} = c_{B^1} \) such that \( Q^1 \subset B^1 \) for a minimal ball \( B^1 \) with centre point \( c_{B^1} \), and let \( Q_d^i \) be the maximal dyadic cubes in \( \mathcal{D}^i \) such that \( c_{B^i} \in Q_d^i \subset B^i \), for \( i = 1, 2 \). We define \( c_{Q_d^i} = c_{B^i} \) (these are not necessarily the centre-points). It is then clear from the setup that the triple of cubes \( Q_d^0, Q_d^1, Q_d^2 \) and the points \( (c_{Q_d^0}, c_{Q_d^1}, c_{Q_d^2}) \) satisfy the claims (3.3) and (3.4). Of the remaining claims, the claim Eq. 3.5 follows, for example, by using the just shown result for balls,

\[
\int_{Q_d^0} \int_{Q_d^2} |K(x, y, z) - K(c_{Q_d^0}, c_{Q_d^1}, c_{Q_d^2})| \, dx \, dz
\]

\[
\leq \int_{B^0} \int_{B^2} |K(x, y, z) - K(c_{B^0}, c_{B^0}, c_{B^2})| \, dx \, dz \leq \omega(A^{-1}) A^{-2d},
\]

and Eq. 3.5 together with Eq. 3.4 implies (3.6). The last claim (we can integrate over any two of the cubes) follows by noting that the estimate for the term \( II \) was point-wise and inspecting the estimate (3.13) for the term \( I \).

\[ \square \]

From now on whenever we fix a cube \( Q^1 \), the associated cubes \( Q^0 \) and \( Q^2 \) will stand for the cubes generated through Proposition 3.2. If a function has support in the cube \( Q^i \) then it has the subscript \( i \) or \( Q^i \), e.g. if \( \text{spt}(g) \subset Q^i \), then we write \( g = g_i = g_{Q^i} \).
Proposition 3.14 Suppose that $K$ is a non-degenerate bilinear kernel. Then, there exists a large parameter $A$ so that supposing the following items:

(i) let $Q^1$ be a cube and let $Q^0, Q^2$ stand for the cubes generated by Proposition 3.2 above,
(ii) let $f$ be a bounded function with zero mean supported on the cube $Q^1$, and
(iii) let $g_i$ be functions such that $\text{spt}(g_i) \subset Q^i$ and $\langle g_i \rangle_{Q^i} \sim \|g_i\|_\infty \gtrsim 1$

hold, then the function $f$ can be written as

$$f = [h_1 T^{1*}(g_0, g_2) - g_0 T(h_1, g_2)] + [h_0 T(g_1, g_2) - g_1 T^{1*}(h_0, g_2)] + \tilde{f} \quad (3.15)$$

and we have the following size and support localization information

$$|h_1| \lesssim A^{2d} |f|, \quad |h_0| \lesssim A^{2d}\omega(A^{-1}) \|f\|_\infty |g_0|, \quad |\tilde{f}| \lesssim \omega(A^{-1}) \|f\|_\infty |g_1| \quad (3.16)$$

where the implicit constants on the line (3.16) depend only on the implicit constants present in the point (iii) and are otherwise independent of the functions $g_i$, $i = 0, 1, 2$. Moreover, there holds that $\int_{Q^1} \tilde{f} = 0$.

Remark 3.17 If we were dealing only with the integrability exponents $p, q, r \in (1, \infty)$ then we could choose the appearing functions $g_i$ simply as $1_{Q^i}$, however, due to the fact that we allow $r \leq 1$ quite arbitrary functions $g_i$ have to be allowed, see the point (iii) in the statement.

Proof We write out the function $f$ as

$$f = h_1 T^{1*}(g_0, g_2) - g_0 T(h_1, g_2) + \tilde{w}, \quad h_1 = \frac{f}{T^{1*}(g_0, g_2)}, \quad \tilde{w} = g_0 T(h_1, g_2)$$

and will next check

$$|h_1| \lesssim A^{2d} |f|, \quad |\tilde{w}(x)| \lesssim \omega(A^{-1}) \|f\|_\infty |g_0(x)|, \quad \int \tilde{w} = 0. \quad (3.18)$$

We first check that the function $h_1$ is well-defined. We denote with $K_Q$ the constant $K(c_{Q^0}, c_{Q^1}, c_{Q^2})$, where $c_{Q^i} \in Q^i$ are the points as in Proposition 3.2. Let $y \in \text{spt}(f) \subset Q^1$ and split into two parts,

$$T^{1*}(g_0, g_2)(y) = \left( T^{1*}(g_0, g_2)(y) - K_Q \int_{Q^2} \int_{Q^0} g_0(x) g_2(z) \, dx \, dz \right)$$

$$+ K_Q \int_{Q^2} \int_{Q^0} g_0(x) g_2(z) \, dx \, dz$$

$$= I(y) + II.$$ 

By the lines (3.4) and (3.5), respectively, of Proposition 3.2 we find

$$|I(y)| \lesssim \int_{Q^2} \int_{Q^0} |K(x, y, z) - K(c_{Q^0}, c_{Q^1}, c_{Q^2})| |g_0(x)||g_2(x)| \, dx \, dz$$

$$\lesssim \omega(A^{-1}) A^{-2d} \|g_0\|_\infty \|g_2\|_\infty \quad (3.19)$$

and

$$|II| = |K_Q| \|Q^1\|^2 \|\langle g_0 \rangle_{Q^0}\| \|\langle g_2 \rangle_{Q^2}\| \sim A^{-2d} \|g_0\|_\infty \|g_2\|_\infty. \quad (3.20)$$
Consequently, after choosing \( A \) sufficiently large, the estimates (3.19) and (3.20) imply that

\[
|T^{1^*}(g_1, g_2)(y)| \sim A^{-2d}||g_0||_\infty ||g_2||_\infty \gtrsim A^{-2d},
\]

and hence especially that the function \( h_1 \) is well-defined, and the left claim on the line (3.18) is also clear. Next, we control the term \( \tilde{w} \). We expand

\[
T(h_1, g_2) = T(h_1 - \frac{f}{K_Q \int g_0 g_2}, g_2) + T\left(\frac{f}{K_Q \int g_0 g_2}, g_2\right).
\]

To estimate the first term on the right-hand side fix a point \( y \in \text{spt}(h_1 - \frac{f}{K_Q \int g_0 g_2}) = \text{spt}(f) \subset Q^1 \). By the lines (3.5) and (3.6) of Proposition 3.2 we have

\[
\left|h_1(y) - \frac{f(y)}{K_Q \int g_0 g_2}\right| = \left|f(y) \int_{Q^1} \left(K_Q - K(x, y, z)\right)g_0(x)g_2(z) \, dx \, dz\right|
\times \left|T^{1^*}(g_0, g_2)(y)K_Q \int g_0 g_2\right|^{-1}
\lesssim |f(y)||g_0||_\infty ||g_2||_\infty \omega(A^{-1})A^{-2d}
\times \left(|g_0||g_0||Q^0||g_2|Q^2\right)\left|A^{-2d}||g_0||_Q^0||g_2||_Q^2\right|^{-1}
\lesssim \omega(A^{-1})A^{2d}|f(y)|.
\]

where we used the assumption (iii). Consequently, for \( x \in \text{spt}(g_0) \subset Q^0 \) there holds that

\[
\left|T\left(\frac{f}{K_Q \int g_0 g_2}, g_2\right)(x)\right| \leq \int_{Q^1} \int_{Q^2} |(K(x, y, z) - K_Q)|h_1(y) - \frac{f(y)}{K_Q \int g_0 g_2}|g_2(z)| \, dy \, dz
\lesssim \omega(A^{-1})A^{2d} \int_{Q^1} \int_{Q^2} |(K(x, y, z)||f(y)||g_2(z)| \, dy \, dz
\lesssim \omega(A^{-1})||f||_\infty.
\]

For the remaining term, we use \( \int_{Q^1} f = 0 \) to estimate

\[
\left|T\left(\frac{f}{K_Q \int g_0 g_2}, g_2\right)(x)\right| = \left|\int_{Q^1} \int_{Q^2} \left(K(x, y, z) - K_Q\right)f(y)g_2(z) \, dy \, dz\right|
\times \left|K_Q \int g_0 g_2\right|^{-1}
\lesssim \int_{Q^2} \int_{Q^1} \left|\left(K(x, y, z) - K_Q\right)f(y)g_2(z)\right| \, dy \, dz
\times \left(|g_0||g_0||Q^0||g_2|Q^2\right)\left|A^{-2d}\right|^{-1}
\lesssim \omega(A^{-1})A^{-2d}||f||_\infty ||g_2||_\infty \times \left(|g_0||g_0||Q^0||g_2|Q^2\right)\left|A^{-2d}\right|^{-1}
\lesssim \omega(A^{-1})||f||_\infty.
\]

By having the above two estimates together it follows that \( |\tilde{w}(x)| \lesssim \omega(A^{-1})\|f\|_\infty \|g_0(x)\| \). Hence it remains to check the last claim on the line (3.18):

\[
\int \tilde{w} = \int g_0 T\left(\frac{f}{T^{1^*}(g_0, g_2)}, g_2\right) = \int \frac{f}{T^{1^*}(g_0, g_2)} T^{1^*}(g_0, g_2) = \int f = 0. \quad (3.22)
\]

All the properties of the function \( f \) on the cube \( Q^1 \) that allowed us to run through the first iteration of the decomposition, are enjoyed by the function \( \tilde{w} \) on the cube \( Q^0 \). Also, for the kernel \( K^{1^*} \) of \( T^{1^*} \) we have \( |K^{1^*}(cQ^1, cQ^0, cQ^2)| = |K(cQ^0, cQ^1, cQ^2)| \gtrsim r^{-2d} \).
Exchanging the roles of $T$ and $T^{1*}$ we iterate the above once more and we write out the function $\tilde{\omega}$ as

$$\tilde{\omega} = h_0 T(g_1, g_2) - g_1 T^{1*}(h_0, g_2) + \tilde{f}, \quad h_0 = \frac{\tilde{\omega}}{T(g_1, g_2)}, \quad \tilde{f} = g_1 T^{1*}(h_0, g_2).$$

Repeating the above arguments, we find that

$$|h_0| \lesssim A^{2d} |\tilde{\omega}| \lesssim A^{2d} \omega(A^{-1}) \|f\|_{\infty}[g_0]$$

and

$$|\tilde{f}| \lesssim \omega(A^{-1}) \|\tilde{\omega}\|_{\infty}[g_1] \lesssim \omega(A^{-1})^2 \|f\|_{\infty}[g_1].$$

Consequently, we have checked the remaining claims on the line (3.16) and it remains to check that $\int \tilde{f} = 0$, however, this follows by using the adjoints similarly as it did for the function $\tilde{\omega}$ on the line (3.22).

In the remaining propositions of this section we relate the oscillation

$$\text{osc}(b; \mathcal{Q}) = \int_{\mathcal{Q}} |b - \langle b \rangle_{\mathcal{Q}}|$$

to commutator norms. Also, for a fixed $\gamma \in (0, 1)$, a subset $F' \subset F$ is said to be a $\gamma$-major subset, provided that $|F'| > \gamma |F|$.  

**Proposition 3.23** Suppose that $K$ is a bilinear non-degenerate kernel, $b \in L^1_{\text{loc}}$ and $\gamma \in (0, 1)$. Fix a cube $Q^1$ and let $g_i = 1_{E_{Q^i}}$, for $i = 0, 1, 2$, where $E_{Q^i} \subset Q^i$ is a $\gamma$-major subset. Then, there holds that

$$|Q^1| \, \text{osc}(b; Q^1) \lesssim |\{b, T\}_1(h_1, g_2, g_0)| + |\{b, T\}_1(g_1, g_2, h_0)|, \quad (3.24)$$

and we have the following size and support localization information,

$$|h_1| \lesssim 1_{Q^1}, \quad |h_0| \lesssim \omega(A^{-1})[g_0], \quad (3.25)$$

where the implicit constants depend only on $\gamma$.

**Proof** By $b - \langle b \rangle_{Q^1}$ having zero mean on the cube $Q^1$ and duality find a function $f$ with the properties $\int_{Q^1} f = 0$, $\|f\|_{L^\infty} \leq 2$ such that

$$|Q^1| \, \text{osc}(b; Q^1) = \int_{Q^1} bf.$$

By Proposition 3.14 we write out the function $f$ to arrive at

$$\int_{Q^1} bf = \int h_0 T^{1*}(g_0, g_2) - g_0 T(h_1, g_2) + \int h_0 T(g_1, g_2) - g_1 T^{1*}(h_0, g_2) + \int_{Q^1} b \tilde{f}
= -\{b, T\}_1(h_1, g_2, g_0) - \{b, T\}_1(g_1, g_2, h_0) + \int_{Q^1} b \tilde{f}.$$

The claims on the line (3.25) follow immediately from $\|f\|_{L^\infty} \leq 2$, the choice of the functions $g_i$ and the corresponding information in Proposition 3.14. Then, by $\int \tilde{f} = 0$ and the bound (3.16) on the error term, we estimate

$$\left| \int_{Q^1} b \tilde{f} \right| = \left| \int_{Q^1} (b - \langle b \rangle_{Q^1}) \tilde{f} \right| \leq \|\tilde{f}\|_{L^\infty} |Q^1| \, \text{osc}(b; Q^1) \lesssim \omega(A^{-1})|Q^1| \, \text{osc}(b; Q^1).$$
Consequently, we find that
\[ |Q^1| \text{osc}(b; Q^1) \lesssim \|[(b, T)]_1(h_1, g_2, g_0) + [(b, T)]_1(g_1, g_2, h_0)\| + \omega(A^{-1})|Q^1| \text{osc}(b; Q^1). \]

Now, as \( b \in L^1_{\text{loc}} \), the common term shared on both sides of the estimate is finite, and hence, by choosing \( A \) large enough, we absorb it to the left-hand side and the claim follows. 

The off-support norms that model the commutator norm will be given next. When \( 1 < r < \infty \) we will use the following off-support norm.

**Definition 3.26** Let \( p, q, r \in (1, \infty) \). Let \( K \) be a kernel that is locally bounded outside the diagonal \( \Delta \) and let \( b \in L^1_{\text{loc}} \). Then, we define the off-support norm

\[
O^A_{p,q,r}(b; K) = \sup \left| \frac{1}{Q^0} \int \int \int_{Q^1} (b(x) - b(y))K(x, y, z)f_1(y)f_2(z)f_0(x) dy \, dz \, dx \right|
\]

\[
\times |Q^0|^{-(1/p+1/q+1/r')},
\]

where the supremum is taken over all triples of cubes \( Q^0, Q^1, Q^2 \) of the same size such that

\[
\max_{a,b \in \{0,1,2\}} \text{dist}(Q^a, Q^b) \sim A \text{diam}(Q^0)
\]

and over all functions \( f_a \) such that \( |f_a| \leq 1_{Q^a} \) for \( a \in \{0,1,2\} \).

**Remark 3.27** It is immediate from Hölder’s inequality that

\[
O^A_{p,q,r}(b; K) \leq \|[b, T]_1\|_{L^p \times L^q \rightarrow L^r},
\]

whenever \( K \) is the kernel of \( T \). For example, when \( r = 1 \) and \( r' = \infty \), we have

\[
O^A_{p,q,r}(b; K) = \sup \left| \{f_01_{Q^0}, [b, T]_1(f_1, f_2)\} \right| \times |Q^1|^{-(1/p+1/q)}
\]

\[
\leq \sup \|f_0\|_{L^\infty} \|[b, T]_1(f_1, f_2)\|_{L^1} \times |Q^1|^{-(1/p+1/q)} \leq \|[b, T]_1\|_{L^p \times L^q \rightarrow L^1}.
\]

When \( 0 < r < 1 \) we will use the following off-support norm.

**Definition 3.28** Let \( r \in (0, \infty) \) and \( p, q \in (1, \infty) \), let \( K \) be any kernel that is locally bounded outside the diagonal and let \( b \in L^1_{\text{loc}} \). Then, we define the weak off-support norm

\[
O^A_{p,q,r}(b; K) \text{ to be the smallest constant } C \text{ such that for all triples of cubes } Q^0, Q^1, Q^2 \text{ of the same size and functions } f_1, f_2 \text{ satisfying}
\]

\[
\max_{a,b \in \{0,1,2\}} \text{dist}(Q^a, Q^b) \sim A \text{diam}(Q^0), \quad |f_1| \leq 1_{Q^1}, \quad |f_2| \leq 1_{Q^2},
\]

there exists a major subset \( F' \subset Q^0 \) such that if \( |f_0| \leq 1_{F'} \), then

\[
\left| \int_{Q^0} \int_{Q^1} (b(x) - b(y))K(x, y, z)f_1(y)f_2(z)f_0(x) dy \, dz \, dx \right| \leq C|Q^0|^{1/p+1/q+1/r'}.
\]

We now fix the constant \( A \) to be so large that all the above propositions where it appears are applicable. Hence, we will also drop the superscript \( A \) from the off-support norms of Definitions 3.26 and 3.28 and only write \( O_{p,q,r}(b; K) \). As \( O^\infty_{p,q,r}(b; K) \leq O_{p,q,r}(b; K) \), also \( O^\infty_{p,q,r} \) is a reasonable off-support norm in the Banach range of exponents. Before connecting the off-support norms to the commutator, we remark the following a priori upper bound.
Remark 3.30  If $K$ is a bilinear kernel satisfying the size estimate (2.1), then

$$O_{p,q,r}(b; K) \lesssim \sup_Q \ell(Q)^{\alpha} \int_Q |b - \langle b \rangle_Q|.$$  

This is quickly seen as follows: fix triples $Q^i$, $f_i$ for $i \in \{0, 1, 2\}$ as in the Definition 3.26 and let $\tilde{Q}$ be a minimal cube such that $Q^0, Q^1 \subset \tilde{Q}$. Then, by the triangle inequality we estimate the difference $(b(x) - b(y))$ inside the integral of the line (3.29) as

$$|b(x) - b(y)| \leq |b(x) - \langle b \rangle_{\tilde{Q}}| + |\langle b \rangle_{\tilde{Q}} - b(y)|$$  

and see that it is enough to control two symmetric terms of which the other is controlled as

$$\hat{c} \int_{Q^0} \int_{Q^1} |b(x) - \langle b \rangle_{\tilde{Q}}| |K(x, y, z) f_1(y) f_2(z) f_0(x)| \, dy \, dz \, dx \lesssim \int_{Q^0} |Q^1|^{-(1/p + 1/q + 1/r')} \int_{Q^1} |b(x) - \langle b \rangle_{\tilde{Q}}| \int_{Q^0} |b(x) - \langle b \rangle_{\tilde{Q}}| \, dx \lesssim \int_{Q^0} |b(x) - \langle b \rangle_{\tilde{Q}}| \, dx \sim \ell(\tilde{Q})^{-\alpha} \int_{Q^0} |b - \langle b \rangle_{\tilde{Q}}|.$$  

The second term is handled identically.

Next, we relate the weak off-support norm $O_{p,q,r}^\infty$ to the commutator. For this, recall that a function $f$ belongs to the space $L^{s,\infty}(\mathbb{R}^d)$, $0 < s < \infty$, if

$$\|f\|_{L^{s,\infty}(\mathbb{R}^d)} := \sup_{\lambda > 0} \lambda \{ x \in \mathbb{R}^d : |f(x)| > \lambda \}^{1/s} < \infty.$$  

Also, recall that $\|f\|_{L^{s,\infty}} \leq \|f\|_{L^{s,f}}$, $s > 0$. The following Lemma 3.31 is standard, see e.g. Section 2.4. Dualization of quasi-norms in the book [19] of Muscalu and Schlag.

Lemma 3.31  Let $s \in (0, \infty)$ and fix a constant $C > 0$. Then, the following are equivalent:

1. There holds that $\|\psi\|_{L^{s,\infty}(\mathbb{R}^d)} \lesssim C$.

2. For each set $F$ with $|F| \in (0, \infty)$, there exists a major subset $F' \subset F$ such that for all functions $|g| \leq 1_{F'}$, there holds that $|\langle \psi, g \rangle| \lesssim C|F|^{1/s'}$.

Taken together, the following two propositions control the oscillation with the commutator norm.

Proposition 3.32  Let $p, q, r \in (0, \infty)$ be arbitrary exponents. Then, there holds that

$$O_{p,q,r}^{\infty}(b; K) \lesssim \|b, T\|_{L^p \times L^q \rightarrow L^{r,\infty}},$$  

whenever $T$ has the kernel $K$ and the commutator is well-defined.

Proof  Consider a triple $Q^0$, $Q^1$, $Q^2$ and functions $f_1$, $f_2$ as in the Definition 3.28 of $O_{p,q,r}^{\infty}(b; K)$. Clearly we may assume that the right-hand side of Eq. 3.33 is finite and hence that $\|[b, T](f_1, f_2)\|_{L^{r,\infty}} < \infty$. Then, denote $F = Q^0$ and let $F' \subset F$ be the major subset given by the item (2) in Lemma 3.31 such that for all functions $|f_0| \leq 1_{F'}$ there holds that

$$\left| \int_{Q^0} \int_{Q^1} \int_{Q^2} (b(x) - b(y)) K(x, y, z) f_1(y) f_2(z) f_0(x) \, dy \, dz \, dx \right| \lesssim \|b, T\|_{L^p \times L^q \rightarrow L^{r,\infty}} \|F\|^{1/r'}$$  

which implies the claim.
Proposition 3.34 Let $p, q, r \in (0, \infty)$ be arbitrary exponents and let $K$ be a bilinear non-degenerate kernel. Then, for all cubes $Q^1 \subset \mathbb{R}^d$ there holds that
\[
\operatorname{osc}(b; Q^1) \lesssim \mathcal{O}_{p,q,r}^\infty(b; K) |Q^1|^{1/p+1/q-1/r}.
\]

Proof Fix a cube $Q^1$ and let $Q^0, Q^2$ be the cubes given by Proposition 3.14 and let $g_i = 1_{Q^i}, i = 1, 2$. Then, according to the definition of $\mathcal{O}_{p,q,r}^\infty(b; K)$ let $F' \subset Q^0$ be a major subset and define $g_0 = 1_{F'}$. Then, by Proposition 3.23, we find that
\[
|Q^1| \operatorname{osc}(b; Q^1) \lesssim |[b, T]_1(h_1, g_2, g_0)| + |[b, T]_1(g_1, g_2, h_0)| \lesssim \mathcal{O}_{p,q,r}^\infty(b; K) |Q^1|^{1/p+1/q+1/r},
\]
which is the claim rearranged. \qed

4 The Cases $r^{-1} \leq \sigma(p, q)^{-1}$

In this section we will be either on the diagonal, meaning that $r^{-1} = \sigma(p, q)^{-1}$, or on the sub-diagonal, meaning that $r^{-1} < \sigma(p, q)^{-1}$. In both cases the lower bounds formulate simultaneously in Theorem 4.1 and the upper bounds in Theorems 4.2 and 4.15.

Theorem 4.1 Let $b \in L^1_{\text{loc}}$ and let $0 < r, p, q < \infty$ be such that $r^{-1} \leq \sigma(p, q)^{-1}$ and let $\alpha := d\{\sigma(p, q)^{-1} - r^{-1}\}$. Suppose that $K$ is a bilinear non-degenerate CZ-kernel, then
\[
\sup_{Q} \ell(Q)^{-\alpha} \int_Q |b - (b)_Q| \lesssim \mathcal{O}_{p,q,r}^\infty(b; K) \lesssim \|[b, T]_1\|_{L^p \times L^q \rightarrow L^{r,\infty}}
\]

Proof Follows immediately from Propositions 3.32 and 3.34. \qed

Next, we record the already known diagonal upper bound, see the remark after for citations and a sketch of the proof.

Theorem 4.2 Let $\frac{1}{2} < r < \infty$ and $1 < p, q < \infty$ be such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ and let $T$ be a bilinear SIO that satisfies the conditions $(1v)$ or $(1\Omega)$ depending on whether $T$ has a variable or a rough kernel. Then,
\[
\|[b, T]_1\|_{L^p \times L^q \rightarrow L^r} \lesssim \|b\|_{\text{BMO}}.
\]

Remark 4.3 The upper bound in the case of the variable kernel is well-known and is recorded e.g. in [17]. For the rough kernel, the upper bound is also most certainly known. We were however unable to find a direct reference in the desired form, hence we reason as follows. Provided that $\Omega \in L^\infty$, we know that with respect to weighted estimates $T_\Omega$ admits a sufficiently strong sparse domination of the corresponding trilinear form, see [3] and especially Theorem 1. and Lemma 5.1. therein. Hence, by the Cauchy integral trick, see e.g. [6], or Theorem 6.8. in [2] for a modern exposition in the multilinear bi-parameter setting, the bilinear commutator $[b, T_\Omega]$ is bounded as desired.

The sub-diagonal upper bound in Theorem 4.15 requires some preparation consisting of extending parts from the linear theory to the bilinear setting. We refer the reader to Grafakos [7] for a complete account of the corresponding linear theory.
**Proposition 4.4** Let $U, T : \Sigma \times \Sigma \to L^1_{\text{loc}}$ be bilinear singular integrals with the same kernel. Then, there exists a function $m \in L^1_{\text{loc}}$ so that $(U - T)(f_1, f_2) = mf_1 f_2$ for all $f_1, f_2 \in \Sigma$.

In addition, if $U, T$ are CZOs (bounded), then the identity $(U - T)(f_1, f_2) = mf_1 f_2$ extends to all functions $f_1, f_2 \in L^\infty_\mathbb{C}$ and the function $m$ is bounded.

**Proof** We will first show the so-called consistency condition. Let $Q \subset \mathbb{R}^d$ be a cube and $f_1, f_2 \in \Sigma$, then almost everywhere

$$(U - T)(1_Q f_1, 1_Q f_2) = 1_Q(U - T)(f_1, f_2). \tag{4.5}$$

We reduce this to two parts; clearly (4.5) follows if we show that

$$(U - T)(f_1, 1_Q f_2) = 1_Q(U - T)(f_1, f_2), \quad (U - T)(1_Q f_1, f_2) = 1_Q(U - T)(f_1, f_2). \tag{4.6}$$

We only show the right identity on the line (4.6) with the left being similar.

As the operators $U, T$ share the kernel $K$, i.e. $U_x = T_x$, the claim (4.6) follows if we show that for $H \in \{U, T\}$ and all points $x \in \mathbb{R}^d$, there exists $\varepsilon > 0$ such that

$$(H - H_\varepsilon)(1_Q f_1, f_2)(x) = 1_Q(x)(H - H_\varepsilon)(f_1, f_2)(x). \tag{4.7}$$

Assume first that $x \in (Q \cup \partial Q)^c$ (the claim is made modulo sets of measure zero and hence we remove the boundary). Then, choose $\varepsilon = \frac{1}{2}\text{dist}(x, \partial Q)$ so that for all points $y \in Q$ there holds that $\max(|x - y|, |x - z|) \geq |x - y| > \varepsilon$. Consequently, as $x \notin \text{spt}(1_Q f_1)$, it follows by the definition of $K$ being the kernel of $H$ that

$$H(1_Q f_1, f_2)(x) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x, y, z)(1_Q f_1)(y) f_2(z) \, dy \, dz$$

$$= \int \int \max\{|x - y|, |x - z|\} > \frac{1}{2} K(x, y, z)(1_Q f_1)(y) f_2(z) \, dy \, dz = H_\varepsilon(1_Q f_1, f_2)(x).$$

and we find both sides of Eq. 4.7 to be zero. Then, let $x \in Q \setminus \partial Q$ and again fix $\varepsilon = \frac{1}{2}\text{dist}(x, \partial Q)$. Then, as above, we see that $(H - H_\varepsilon)(1_Q c f_1, f_2)(x) = 0$, and consequently, for $x \in Q \setminus \partial Q$ there holds that

$$(H - H_\varepsilon)(1_Q f_1, f_2)(x) = (H - H_\varepsilon)(1_{\mathbb{R}^d} f_1, f_2)(x) - (H - H_\varepsilon)(1_Q c f_1, f_2)(x)$$

$$= (H - H_\varepsilon)(f_1, f_2)(x) - 0$$

$$= 1_Q(x)(H - H_\varepsilon)(f_1, f_2)(x).$$

Hence, we have checked the right identity on the line (4.6) to hold almost everywhere.

Let $Q_N = [-N, N]^d$, for $N \geq 0$. Then, we define the function $m$ by

$$m(x) = 1_Q N(x)(U - T)(1_{Q_N}, 1_{Q_N})(x), \quad x \in Q_N; \tag{4.8}$$

the property (4.5) shows in particular that this definition does not depend on the choice of $N$, as long as $x \in Q_N$, because $Q_N \cap Q_{N'} = Q_{\text{min}(N, N')}$. Then, let $f_i, i = 1, 2$, be simple and let $x \in \mathbb{R}^d$. Fix a cube $Q_N$ such that $\text{spt}(f_1) \cup \text{spt}(f_2) \cup \{x\} \subset Q_N$. Then, by Eq. 4.5 and linearity there holds that

$$(U - T)(1_{Q_N} f_1, 1_{Q_N} f_2)(x) = f_1(x)(U - T)(1_{Q_N}, 1_{Q_N} f_2)(x)$$

$$= f_1(x) f_2(x)(U - T)(1_{Q_N}, 1_{Q_N})(x) = f_1(x) f_2(x)m(x).$$

Consequently, we have shown that $U - T = m$ on $\Sigma \times \Sigma$, and this also gives $m \in L^1_{\text{loc}}$ by testing against simple functions.
If $U, T$ are bounded operators (say $L^4 \times L^4 \to L^2$) the identity $(U - T)(f_1, f_2) = mf_1 f_2$ follows by approximating $L^4$ functions with those in the class $\Sigma$ (for which the identity holds) and as $L^\infty \subset L^4$ the desired identity follows. Also, by testing against simple functions, it follows by the boundedness of $U, T$ that necessarily $\|m\|_{\infty} \leq \|U\|_{L^4 \times L^4 \to L^2} + \|T\|_{L^4 \times L^4 \to L^2}$.

\[\text{Proposition 4.9} \quad \text{Let } K \text{ be a kernel such that } \sup_{\epsilon > 0} \|T_{\epsilon}\|_{L^p \times L^q \to L^{\sigma(p,q)}} < \infty \text{ for some exponents satisfying } 1 < p, q < \infty \text{ and } 1 \leq \sigma(p,q). \text{ Then, there exists a bounded bilinear operator } T_0 : L^p \times L^q \to L^{\sigma(p,q)} \text{ with the kernel } K \text{ and a sequence } \epsilon_k \to 0 \text{ such that}
\]
\[
\lim_{\epsilon_k \to 0} \left\langle T_{\epsilon_k}(f_1, f_2), f_3 \right\rangle = \left\langle T_0(f_1, f_2), f_3 \right\rangle, \tag{4.10}
\]

for all $f_1, f_2, f_3 \in L^\infty_c$. In addition, if $T$ is a CZO with the kernel $K$, then (by Proposition 4.4) there exists a bounded function $m$ such that $T_0 = T + m$.

\[\text{Proof} \quad \text{We will show the argument with the exponents } p = q = 3 \text{ and } \sigma(p,q) = \frac{3}{2}. \text{ Let } F \text{ be a countable dense subset of } L^3. \text{ By the bound } \sup_{\epsilon > 0} \|T_{\epsilon}\|_{L^3 \times L^3 \to L^3} < \infty, \text{ H"{o}lder’s inequality and a diagonalization argument, we find a sequence } \epsilon_k \to 0 \text{ such that for all } f_1, f_2 \in F,
\]
\[
\Lambda_{f_1, f_2}(f_3) = \lim_{\epsilon_k \to 0} \left\langle T_{\epsilon_k}(f_1, f_2), f_3 \right\rangle \tag{4.11}
\]

defines a bounded linear functional on $L^3 \cap F$ with norm
\[
\|\Lambda_{f_1, f_2}\|_{L^3 \cap F \to \mathbb{C}} \leq \sup_{\epsilon > 0} \|T_{\epsilon}\|_{L^3 \times L^3 \to L^3} \|f_1\|_{L^3} \|f_2\|_{L^3}.
\]

By Cauchy sequences this extends as a bounded linear functional to the whole of $L^3$ and then the Riesz representation theorem gives a function $\psi(f_1, f_2) \in L^{\frac{3}{2}} = \left( L^3 \right)^*$ such that
\[
\Lambda_{f_1, f_2}(f_3) = \int \psi(f_1, f_2) f_3, \quad \|\psi(f_1, f_2)\|_{L^{\frac{3}{2}}} \leq \|\Lambda_{f_1, f_2}\|_{L^3 \cap F \to \mathbb{C}}.
\]

Then, we define the operator $T_0(f_1, f_2) = \psi(f_1, f_2)$ in the dense class $F \times F$ and clearly $T_0 : L^3 \cap F \times L^3 \cap F \to L^3$ is a bounded bilinear operator with the kernel $K$ that satisfies Eq. 4.10 for functions $f_1, f_2 \in F$ and $f_3 \in L^3$. Again, by Cauchy sequences $T_0$ extends as a bounded bilinear functional to the whole $L^3 \times L^3$ and then it remains to argue that $T_0$ has the kernel $K$ and that Eq. 4.10 holds for $f_1, f_2, f_3 \in L^3$. That $T_0$ has the kernel $K$ follows by how $T_0$ was extended to $L^3 \times L^3$ via Cauchy sequences, the kernel representation being valid in $F \times F$ and the dominated convergence theorem. Similarly we find that Eq. 4.10 holds for $f_1, f_2, f_3 \in L^3$. As $L^\infty \subset L^3$, we are done.

\[\text{Proposition 4.12} \quad \text{Let } T \text{ be a SIO with a kernel } K \text{ such that } \sup_{\epsilon > 0} \|T_{\epsilon}\|_{L^p \times L^q \to L^{\sigma(p,q)}} < \infty \text{ for some exponents satisfying } 1 < p, q < \infty \text{ and } 1 \leq \sigma(p,q), \text{ and let } f_1, f_2 \in L^\infty_c \text{ and } b \in C^{0,\alpha}. \text{ Then, there holds that}
\]
\[
[b, T]_1(f, g)(x) = \int \int (b(x) - b(y))K(x, y, z)f(y)g(z) \, dy \, dz. \tag{4.13}
\]
Proof As $b \in \dot{C}^{0,\alpha} \subset L^\infty_{\text{loc}}$, $bf_1 \in L^\infty_C$. Then by Proposition 4.9 we have

$$\{b, T\} f_1, f_2, f_3 = \{b, T_{\epsilon_k}\} f_1, f_2, f_3$$

where the last step marked with * follows by the dominated convergence theorem after the following estimate (uniform in $\epsilon_k$)

$$\lim_{\epsilon_k \to 0} \int (b(x) - b(y)) K(x, y, z) f_1(y) f_2(z) dy dz$$

where the finiteness follows simply by the fact that $f, g \in L^\infty$ and that the appearing singularity is weak enough to be locally integrable (see also the last estimate in the proof of Theorem 4.15). Now as Eq. 4.14 holds for all test functions $f_3$ the claim on the line (4.13) follows.

Theorem 4.15 Let $b \in L^1_{\text{loc}}$, let $\frac{1}{2} < r, p, q < \infty$ be such that $r^{-1} < \sigma(p, q)^{-1}$, let $\alpha := d\sigma(p, q)^{-1} - r^{-1}$, and let $T$ be a bilinear SIO such that $\sup_{\epsilon > 0} \|T_{\epsilon}\|_{L^p \times L^q} \to L^{\sigma(p, q)} < \infty$ for one tuple of exponents $1 < p_0, q_0 < \infty$ with $\sigma(p_0, q_0) \geq 1$. Then,

$$\|\{b, T\}\|_{L^p \times L^q} \to L^r \lesssim \sup Q \ell(Q)^{-\alpha} f_1(y) f_2(z) dy dz < \infty,$$

where the finiteness follows simply by the fact that $f, g \in L^\infty$ and that the appearing singularity is weak enough to be locally integrable (see also the last estimate in the proof of Theorem 4.15). Now as Eq. 4.14 holds for all test functions $f_3$ the claim on the line (4.13) follows.

Proof Clearly we may assume that $\|b\|_{C^{0,\alpha}} < \infty$, as otherwise the claimed estimate is immediate. By density it is enough to prove the claim for functions $f_1, f_2 \in L^\infty_{\text{loc}}$. Then, by Proposition 4.12 we write the commutator in a closed form and estimate it as

$$\|\{b, T\}\|_{L^p \times L^q} \to L^r \lesssim \sup Q \ell(Q)^{-\alpha} f_1(y) f_2(z) dy dz$$

The operator $\ell^\alpha$ is the multilinear fractional integral of Kenig and Stein, see [14], where its boundedness is fully characterized: it satisfies exactly the claimed estimates.

5 The Case $r^{-1} > \sigma(p, q)^{-1}$

Now we are on the super-diagonal and we define the exponent $s$ by $r^{-1} = \sigma(s, p, q)^{-1}$. The following proposition shows that the membership of $b \in L^s$ is sufficient for commutator boundedness.

Proposition 5.1 Let $0 < r, s, p, q < \infty$ be such that $r^{-1} = \sigma(s, p, q)^{-1}$. Also, let $T$ be bounded as

$$T : L^p \times L^q \to L^{\sigma(p, q)}, \quad T : L^{\sigma(s, p)} \times L^q \to L^r.$$
Then, there holds that $\| [b, T]_1 \|_{L^p \times L^q \to L^r} \lesssim \| b \|_{L^r}$.

**Proof** We first estimate

$$\| [b, T]_1 (f, g) \|_{L^r} = \| (b - c, T)_1 (f, g) \|_{L^r} \leq \| (b - c) T (f, g) \|_{L^r} + \| T ((b - c) f, g) \|_{L^r}.$$  

By Hölder’s inequality and the boundedness of $T$ we find that

$$\| (b - c) T (f, g) \|_{L^r} \leq \| b - c \|_{L^s} \| T (f, g) \|_{L^s(p, q)} \| b - c \|_{L^s} \| T \|_{L^p \times L^q \to L^s(p, q)} \| f \|_{L^p} \| g \|_{L^q}.$$  

By the boundedness of $T$ and Hölder’s inequality we find that

$$\| T ((b - c) f, g) \|_{L^r} \leq \| T \|_{L^s(p, q) \times L^q \to L^r} \| (b - c) f \|_{L^s(p, q)} \| g \|_{L^q} \leq \| T \|_{L^s(p, q) \times L^q \to L^r} \| b - c \|_{L^s} \| f \|_{L^p} \| g \|_{L^q}.$$  

Taking the infimum over all $c \in \mathbb{C}$ shows the claim.

**Remark 5.2** Suppose that $\frac{1}{2} < r < \infty$ and $1 < p, q < \infty$ and $\sigma(s, p) > 1$. Then any $T \in CZO(2, d, \omega)$ is bounded as in the assumptions of Proposition 5.1. Moreover, if $r \geq 1$, then automatically $p, q, \sigma(s, p) > 1$. It is not however the case that if $r < 1$ and $s, p, q > 1$ that automatically $\sigma(s, p) > 1$, indeed, if $q$ is very large and $r$ is very close to $1/2$, then necessarily $\sigma(s, p)$ is close to $1/2$.

**Definition 5.3** We say that a collection of sets $\mathcal{S}$ is $\gamma$-sparse, if there exists a pairwise disjoint collection of $\gamma$-major subsets $\mathcal{S}_E = \{ E_Q \subset Q : Q \in \mathcal{S} \}$.

Our sparse collections will be built by splitting into dyadic scales. For a cube and a dyadic grid $Q \in \mathcal{D}$, we denote $\mathcal{D}_Q = \{ P \in \mathcal{D} : P \subset Q \}$. Let $f$ be a locally integrable function and let $Q \in \mathcal{D}$ be a cube, then we set

$$S(f; Q) = \{ P \in \mathcal{D}_Q : P \text{ is a maximal cube such that } \langle |f| \rangle_P > 2 \langle |f| \rangle_Q \},$$

and form the principal stopping time family $\mathcal{S} \subset \mathcal{D}_Q$ by

$$\mathcal{S} = \bigcup_k \mathcal{S}_k, \quad \mathcal{S}_{k+1} = \bigcup_{P \in \mathcal{S}_k} S(f; P), \quad \mathcal{S}_0 = \{ Q \}.$$  

For an arbitrary collection $\mathcal{S} \subset \mathcal{D}$ of dyadic cubes and for each $Q \in \mathcal{S}$ we let $\text{ch}_{\mathcal{S}}(Q)$ consist of the maximal cubes $P \in \mathcal{S}$ such that $P \subset Q$. For a given cube $Q \in \mathcal{S}$ we denote $E_Q = Q \setminus \bigcup_{P \in \text{ch}_{\mathcal{S}}(Q)} P$ and for each $P \in \mathcal{D}$ we let $\Pi P = \Pi_{\mathcal{S}} P$ denote the minimal cube $Q$ in $\mathcal{S}$ such that $P \subset Q$ (on the condition that it exists). Then

$$\text{ch}_{\mathcal{S}}(Q) = \{ P \in \mathcal{S} : P \subset Q, \Pi P = Q \}.$$

We also denote $\text{ch}^0_{\mathcal{S}}(Q) = \text{ch}_{\mathcal{S}}(Q)$ and $\text{ch}^{k+1}_{\mathcal{S}}(Q) = \bigcup_{P \in \text{ch}^k_{\mathcal{S}}(Q)} \text{ch}_{\mathcal{S}} P$. Lastly, given a cube $Q \in \mathcal{D}$, we denote

$$\Delta_Q f = \sum_{P \in \text{ch}_{\mathcal{D}}(Q)} (\langle f \rangle_P - \langle f \rangle_Q)^1_P.$$  

The following Lemma is recorded e.g. in [11].

**Lemma 5.4** Fix a cube $Q \in \mathcal{D}$ and a function $f \in L^1_{\text{loc}}$ supported on the cube $Q$. Then, the principal stopping time family $\mathcal{S} \subset \mathcal{D}_Q$ is $\frac{1}{2}$-sparse.

If $f \in L^\infty(Q)$ and $\int_Q f = 0$, then we split the function $f$ according to the partition

$$\mathcal{S} = \bigcup_{k=0}^N \mathcal{S}_k, \quad \mathcal{S}_k = \text{ch}^k_{\mathcal{S}}(Q).$$
where the number $N$ is finite and depends only on $\|f\|_{L^{\infty}(\mathcal{Q})}$, as

$$f = \sum_{k=0}^{N} \sum_{P \in \mathcal{A}_k} f_P, \quad f_P = \sum_{\Pi \supseteq P} \Delta_Q f,$$

and the functions $f_P$ satisfy:

1. $\int f_P = 0$.
2. $\|f_P\|_{\infty} \leq \|f\|_P$ and
3. $\sum_{k=0}^{N} \sum_{P \in \mathcal{A}_k} \|f_P\|_1 \|P\| \leq (Mf)^s$, $s > 0$.

**Lemma 5.6** Let $\mathcal{S}$ be a sparse collection, let $\gamma > 0$ and let $\mathcal{D}$ be a dyadic grid. To each cube $Q \in \mathcal{S}$ associate another cube $\tilde{Q} \in \mathcal{D}$ such that $\text{dist}(Q, \tilde{Q}) \leq \gamma \ell(Q)$ and $\ell(\tilde{Q}) \sim \ell(Q)$. Then, the collection $\tilde{\mathcal{S}} = \{\tilde{Q} : Q \in \mathcal{S}\}$ is sparse.

**Proof** Let $\tilde{P}, \tilde{H} \in \tilde{\mathcal{S}}$ be such that $\tilde{H} \subseteq \tilde{P}$. Then, from that $\text{dist}(H, \tilde{H}) \leq \gamma \ell(H)$ and $\ell(H) \leq \ell(\tilde{H})$, it follows that there exists a constant $\beta \sim \gamma$ so that $H \subset \beta \tilde{P}$. Consequently, we find that

$$\sum_{H \in \mathcal{A}} |\tilde{H}| \leq \sum_{H \in \mathcal{A}} |\tilde{H}| \leq \sum_{H \in \mathcal{A}} |H| \leq \sum_{H \in \mathcal{A}} |E_H| \leq |\beta \tilde{P}| \leq |\tilde{P}|,$$

where we used $\ell(\tilde{H}) \leq \ell(H)$ in the second estimate and the sparseness of $\mathcal{S}$ in the third and the fourth estimates. We have shown that the collection $\tilde{\mathcal{S}}$ is Carleson and as the Carleson condition is equivalent with sparseness for dyadic collections, for this fact see e.g. the book of Lerner and Nazarov [16], the claim follows. □

**Lemma 5.7** Let $p \in (1, \infty)$ and $\mathcal{S}$ be a sparse collection. Then, for any constants $a_Q$ there holds that

$$\| \sum_{Q \in \mathcal{S}} a_Q 1_Q \|_{L^p} \leq \| \sum_{Q \in \mathcal{S}} |a_Q| 1_{E_Q} \|_{L^p}. $$

**Proof** The claim follows by duality and the following estimate

$$\{ \sum_{Q \in \mathcal{S}} a_Q 1_Q, g \} \leq \sum_{Q \in \mathcal{S}} |Q| |a_Q| |(|g|)_Q \leq \sum_{Q \in \mathcal{S}} |E_Q| |a_Q| |(|g|)_Q \leq \int M g \sum_{Q \in \mathcal{S}} \sum_{Q \in \mathcal{S}} |a_Q| 1_{E_Q} \|_{L^p} \|g\|_{L^{p'}}.$$

□

**Definition 5.8** Let $b \in L^1_{\text{loc}}$, $1 \leq r, p, q < \infty$ and let $K$ be locally bounded away from the diagonal. Then, we define the super-diagonal off-support norm

$$O_{p, q, r} (b; K) = \sup_{k=0}^{N} \int \int \int (b(x) - b(y)) K(x, y, z) f_{1,k}(y) f_{2,k}(z) f_{0,k}(x) dy \ dz \ dx$$

$\times N_{p, q, r} (\hat{f})^{-1}$. 

© Springer
where
\[
N_{p,q,r}(\tilde{f}) = \left\| \sum_{k=0}^{N} f_{0,k} \right\|_{L^p(\mathbb{R}^d)} \left\| \sum_{k=0}^{N} f_{1,k} \right\|_{L^q(\mathbb{R}^d)} \times \left\| \sum_{k=0}^{N} f_{2,k} \right\|_{L^r(\mathbb{R}^d)},
\]
and the supremum is taken over all finite collections of triples of cubes of the same diameter such that \(\max_{i,j \in \{0,1,2\}} \text{dist}(Q_i^k, Q_j^k) \sim A \text{ diam}(Q_0^k)\) and over all functions such that \(\|f_{i,k}\|_{\infty} < \infty\) and \(|\text{spt}(f_{i,k})| > 0\).

**Remark 5.9** If \(r > 1\), then we can replace the entries \(1_{\text{spt}(f_{i,k})}\) in the three terms of \(N_{p,q,r}(\tilde{f})\), with \(1_{Q_i^k}\), for \(i = 0, 1, 2\).

**Remark 5.10** For bilinear operators \(U\) there holds that
\[
\sum_{i=1}^{N} \langle U(f_i, g_i), h_i \rangle = \mathbb{E} \mathbb{E}^i \left\{ U \left( \sum_{i=1}^{N} \varepsilon_i \varepsilon_i^g f_i, \sum_{j=1}^{N} \varepsilon_j^g g_j \right), \sum_{i=1}^{N} \varepsilon_i h_i \right\},
\]
where \(\varepsilon_i, \varepsilon_i^g\) are independent random signs, over some probability spaces with expectations denoted respectively as \(\mathbb{E}, \mathbb{E}^i\), meaning that \(\mathbb{E} \varepsilon_i \varepsilon_j = \mathbb{E}^i \varepsilon_i^g \varepsilon_j^g = 1 (i = j) (i, j)\). Then, H"older's inequality shows that for \(r \geq 1\) we have \(O^r_{p,q,r}(b; K) \leq \| [b, T]_1 \|_{L^p \times L^q \rightarrow L^r}\), and consequently, that \(O^r_{p,q,r}\) is a reasonable off-support constant for \(r \geq 1\).

**Proposition 5.11** Suppose that \(K\) is a non-degenerate bilinear kernel and \(b \in L^1_{\text{loc}}\). Suppose that \(0 < r < \infty\) and \(1 < s, p, q < \infty\) are such that \(r^{-1} = \sigma(s, p, q)^{-1}\). Then, there holds that
\[
\|b\|_{L^r} \lesssim O^r_{p,q,r}(b; K).
\]

**Proof** Let \(Q^1\) be an arbitrary cube. Fix a constant \(M > 0\) and let \(f\) be a function such that
\[
1_{Q^1} f = f, \quad \int_{Q^1} f = 0, \quad \|f\|_{\infty} \leq M, \quad \|f\|_{L^r} \leq 1.
\]
Note that \(s > r\) and hence \(s, s' > 1\) are both in the Banach range of exponents. Let \(\mathcal{S} \subset \mathcal{D}_Q\) denote the sparse collection of cubes we obtain through Lemma 5.4. Write the function \(f\) as on the line (5.5) and by Proposition 3.14 factorize each of the terms \(f_{p,i}, P^i \in \mathcal{S}\), as in Eq. 3.15, to arrive at
\[
f_{p,i} = [h_{p,i} T_{1s}^1 (g_{p_0}, g_{p_2}) - g_{p_0} T(h_{p_1}, g_{p_2})] + [h_{p_0} T(g_{p_1}, g_{p_2}) - g_{p_1} T_{1s}^1 (h_{p_0}, g_{p_2})] + f_{p,i},
\]
where we have written \(h_i = h_{p,i}\) and \(g_i = g_{p,i}\). Next, we will specify how the cubes and functions \(P^i, g_{p,i}\), for \(i = 0, 2\), are chosen.

By Proposition 3.2 we can assume \(P^0, P^2 \in \mathcal{D}\). Then, by Lemma 5.6 the collection \(\mathcal{S}^0 = \{P^0 : P^1 \in \mathcal{S}\} \subset \mathcal{D}\) is sparse and we will denote the pairwise disjoint major subsets with \(\mathcal{E}_{p_0}\). By Proposition 3.14 we are free to choose the functions \(g_{p,i}\) under the condition \(\|g_i\|_{Q_i} \gtrsim \|g_i\|_{\infty} \geq 1\) and clearly the following choices suffice,
\[
g_{p_0} = 1_{E_{p_0}}, \quad g_{p_1} = 1_{P^1}, \quad g_{p_2} = 1_{P^2}.
\]
Now, the off-support norm only controls finite sums, but the collection $\mathcal{S}$ is potentially infinite. Hence, we empty the collection $\mathcal{S}$ through an increasing chain of finite subcollections $\mathcal{S}^1 \subseteq \mathcal{S}^2 \subseteq \cdots \subseteq \mathcal{S}$. Then, we have

$$
\left| \int_Q b f \right| \leq \lim_{n \to \infty} \sum_{P^1 \in \mathcal{S}_n^1} |\{ [b, T]_1 (h_{P^1}, g_{P^2}), g_{P^0}\}| + \sum_{P^1 \in \mathcal{S}_n^2} |\{ [b, T]_1 (g_{P^1}, g_{P^2}), h_{P^0}\}|
+ \int b \, f_{\Sigma},
$$

(5.14)

where we denote $\tilde{f}_{\Sigma} = \sum_{P^1 \in \mathcal{S}_n} \tilde{f}_{P^1}$ and the implicit change of integration and summation is easily checked by the dominated convergence theorem after the subsequent estimates. We first analyse the second sum on the right-hand side of Eq. 5.14, the first one being similar. By trilinearity we write

$$
\{ [b, T]_1 (g_{P^1}, g_{P^2}), h_{P^0}\} = \{ [b, T]_1 (\alpha_{P^1} g_{P^1}, \alpha_{P^2} g_{P^2}, \alpha_{P^0} h_{P^0}\}
$$

for any constants with $\alpha_{P^0} \alpha_{P^1} \alpha_{P^2} = 1$. Hence, by the relation $r^{-1} = \sigma(s, p, q)^{-1}$, we take

$$
\alpha_{P^0} = \| f_{P^1} \|_{\infty}^{s'}, \quad \alpha_{P^1} = \| f_{P^1} \|_{\infty}^{e'}, \quad \alpha_{P^2} = \| f_{P^1} \|_{\infty}^{e'}, \quad 0 = \left( \frac{e'}{p'} - 1 \right) + \frac{e'}{p} + \frac{e'}{q}.
$$

(5.16)

Then, with the choices (5.16), also using $|h_{P^0}| \lesssim \| f_{P^1} \|_{\infty} |g_{P^0}| = \| f_{P^1} \|_{\infty} E_{P^0}$, we find that

$$
\sum_{P^1 \in \mathcal{S}_n^1} |\{ [b, T]_1 (\alpha_{P^1} g_{P^1}, \alpha_{P^2} g_{P^2}, \alpha_{P^0} h_{P^0}\}|
\leq O_{p, q, r} (b, K) \| f_{P^1} \|_{\infty}^{e'} \| f_{P^2} \|_{\infty}^{e'} \| f_{P^1} \|_{\infty}^{e'} \| f_{P^2} \|_{\infty}^{e'} \|_{L^p(\mathbb{R}^d)} \| f_{P^1} \|_{\infty}^{e'} \| f_{P^2} \|_{\infty}^{e'} \|_{L^q(\mathbb{R}^d)}
\times \| f_{P^1} \|_{\infty}^{e'} \| f_{P^1} \|_{\infty}^{e'} \|_{L^p(\mathbb{R}^d)} = RHS.
$$

Now the proof splits into the cases $r > 1$ and $r = 1$.

We first consider the case $r > 1$ in which all the three terms of $RHS$ are estimated similarly. From that dist$(P^1, P^i) \lesssim \ell(P^i)$ and $\ell(P^1) \lesssim \ell(P^i)$ it follows that there exists an absolute constant $C > 0$ such that $CP^1 \supset P^i$. This means that the collections $\{CP^1 : P^1 \in \mathcal{S}\}$ are sparse with the major subsets $E_{P^1}$. Hence, by Lemma 5.7, for $i \in \{0, 1, 2\}$ and $v \in (0, \infty)$ and $u \in (0, \infty)$, there holds that

$$
\| \sum_{P^1 \in \mathcal{S}_n^1} \| f_{P^1} \|_{L^v(\mathbb{R}^d)} \|_{L^u(\mathbb{R}^d)} \| \leq \| \sum_{P^1 \in \mathcal{S}_n^1} \| f_{P^1} \|_{L^v(\mathbb{R}^d)} \|_{L^u(\mathbb{R}^d)} \| \leq \| \sum_{P^1 \in \mathcal{S}_n^1} \| f_{P^1} \|_{L^v(\mathbb{R}^d)} \|_{L^u(\mathbb{R}^d)} \| \leq \| \sum_{P^1 \in \mathcal{S}_n^1} \| f_{P^1} \|_{L^v(\mathbb{R}^d)} \|_{L^u(\mathbb{R}^d)} \| \leq (Mf)^{u'} \|_{L^v(\mathbb{R}^d)},
$$

where in the last estimate we used the point-wise estimate (3) from Lemma 5.4. Now, we find that

$$
RHS \lesssim O_{p, q, r} (b, K) \| Mf \|_{L^p(\mathbb{R}^d)} \| Mf \|_{L^q(\mathbb{R}^d)} \| Mf \|_{L^{\infty}(\mathbb{R}^d)}
\lesssim O_{p, q, r} (b, K) \| f \|_{L^p(\mathbb{R}^d)} \| f \|_{L^q(\mathbb{R}^d)} \| f \|_{L^\infty(\mathbb{R}^d)}
\leq O_{p, q, r} (b, K),
$$

where we used $s' > 1$ in the second estimate for the boundedness of the maximal function.
In the case $r = 1$ the first two terms of $\mathcal{RHS}$ estimate the same as in the case $r > 1$ and the last term estimates differently

\[ \mathcal{RHS} \lesssim \mathcal{O}_{p,q,r}^{\Sigma}(b; K) \left\| \sum_{p^1 \in \mathbb{P}} \| f_{p^1} \|_{L^\infty} 1_{E_{p^1}} \right\|_{L^r(\mathbb{R}^d)} = \mathcal{O}_{p,q,r}^{\Sigma}(b; K) \left\| \sum_{p^1 \in \mathbb{P}} 1_{E_{p^1}} \right\|_{L^\infty(\mathbb{R}^d)} \leq \mathcal{O}_{p,q,r}^{\Sigma}(b; K) \left\| 1 \right\|_{L^\infty(\mathbb{R}^d)} = \mathcal{O}_{p,q,r}^{\Sigma}(b; K), \]

the crucial step here was the disjointness of the sets $E_{p^1}$.

The just shown estimates also hold for the other term, and as the estimates are uniform in $n$, it follows that

\[ | \int b f | \lesssim \mathcal{O}_{p,q,r}^{\Sigma}(b; K) + | \int b \tilde{f}_\Sigma |. \]

By Lemma 5.4 and Proposition 3.14 we have

\[ | \tilde{f}_\Sigma | \leq \sum_{p^1 \in \mathbb{P}} \| \tilde{f}_{p^1} \|_{L^\infty} \| 1_{p^1} \| \lesssim \omega(A^{-1}) \sum_{p^1 \in \mathbb{P}} \| f_{p^1} \|_{L^\infty} \| 1_{p^1} \| \lesssim \omega(A^{-1}) M f \quad (5.17) \]

and as also $1_{Q^i} \tilde{f}_\Sigma = \tilde{f}_\Sigma$ and $\int_{Q^i} \tilde{f}_\Sigma = 0$, the function $\tilde{f}_\Sigma$ satisfies the conditions on the line (5.12) but now with the additional decay $\lesssim \omega(A^{-1})$. Consequently, we find that

\[ \sup_{(5.12)} | \int b f | \lesssim \mathcal{O}_{p,q,r}^{\Sigma}(b; K) + \omega(A^{-1}) \sup_{(5.12)} | \int b f |. \quad (5.18) \]

The common term on both sides of the estimate (5.18) is finite (recall that $b \in L^1_{\text{loc}}$ and for each $f$ as in the supremum $\| f \|_{L^\infty} < M$) and hence by choosing $A$ sufficiently large and absorbing the common term to the left-hand side we find that

\[ \sup_{(5.12)} | \int f_{Q^i} | \lesssim \mathcal{O}_{p,q,r}^{\Sigma}(b; K). \quad (5.19) \]

Then, as $s' > 1$, the proof is concluded with exactly the same argument by Riesz’ representation theorem as in [11]. For the convenience of the reader we give the full details. Denote $L^\infty_{c,0} = \{ \varphi : \int \varphi = 0, \varphi \in L^\infty_c \}$, where $L^\infty_c$ denotes bounded and compactly supported functions. As the right-hand side of Eq. 5.19 is independent of the cube $Q$ and the constant $M$, we find that

\[ \Lambda : L^{s'} \cap L^\infty_{c,0} \to \mathbb{C}, \quad \Lambda f = \int b f, \quad \| \Lambda \|_{L^{s'} \cap L^\infty_{c,0} \to \mathbb{C}} \lesssim \mathcal{O}_{p,q,r}^{\Sigma}(b; K) \]

defines a bounded linear functional in a dense subset of $L^{s'}$. By density and linearity we find a linear extension $\hat{\Lambda} : L^{s'} \to \mathbb{C}$ of $\Lambda$ such that $\| \hat{\Lambda} \|_{L^{s'} \to \mathbb{C}} \leq \| \Lambda \|_{L^{s'} \cap L^\infty_{c,0} \to \mathbb{C}}$. By the Riesz representation theorem there exists a function $a$ satisfying $\| a \|_{L^{s'}} \leq \| \hat{\Lambda} \|_{L^{s'} \to \mathbb{C}}$ and $\hat{\Lambda} f = \int a f$, for all $f \in L^{s'}$. Especially, as $\hat{\Lambda}$ extends $\Lambda$, there holds that

\[ \int b f = \int a f, \quad f \in L^{s'} \cap L^\infty_{c,0}. \quad (5.20) \]

Let $\psi^k_x = \frac{1_{B(x,k^{-1})}}{|B(x,k^{-1})|}$ be an approximation to identity at the point $x$ and define $\varphi^k_{x,y} = \psi^k_x - \psi^k_y$. Then $\varphi^k_{x,y} \in L^{s'} \cap L^\infty_{c,0}$ and we find by Eq. 5.20 and the Lebesgue differentiation theorem that

\[ b(x) - b(y) = \lim_{k \to \infty} \int b \varphi^k_{x,y} = \lim_{k \to \infty} \int a \varphi^k_{x,y} = a(x) - a(y). \]
It follows that \( b = a + c \) for some constant \( c \), and especially that \( \|b\|_{L^r} \lesssim \mathcal{O}^\Sigma_{p,q,r}(b; K) \). We are done.

Having Propositions 5.1 and 5.11 together gives us

**Theorem 5.21** Let \( 1 \leq r, s, p, q < \infty \) be such that \( r^{-1} = \sigma(s, p, q)^{-1} \) and let \( T \) be a non-degenerate bilinear SIO bounded as

\[
T : L^p \times L^q \to L^{\sigma(p,q)}, \quad T : L^{\sigma(s,p)} \times L^q \to L^r.
\]

Then, there holds that

\[
\|[b, T]\|_{L^p \times L^q \to L^r} \sim \|b\|_{L^r}.
\]

**5.1 The Problem of \( 1/2 < r < 1 \)**

As we just saw, in the super-diagonal case we were unable to relax the assumption \( r \geq 1 \) to \( r > 1/2 \) and it is not immediately clear how this could be done. There is however a line of investigation that could begin from here, let us discuss this next.

When \( r < 1 < s \), the expected characterization should still be in terms of \( \dot{L}^s \), especially so as the upper bound remains valid in some cases. Indeed, by Remark 5.2 there exist integrability parameters \( 1/2 < r < 1 \) and \( s, p, q > 1 \) so that Proposition 5.1 does not apply as such, and hence even the upper bound (if valid) would require a new proof in these cases; meanwhile, there exist the tuple \( r = 3/4 \) and \( s = p = q = 9/4 \) (hence \( r^{-1} = \sigma(s, p, q)^{-1} \) satisfying \( \sigma(s, p) > 1 \), and this verifies the upper bound to be valid with \( 1/2 < r = 3/4 < 1 \) (and the other exponents as above) for a wide class of bilinear CZOs.

Concerning the lower bound when \( r < 1 < s \), we can still initiate the proof of Proposition 5.11 by dualizing with functions as on the line (5.12). Proceeding, one might attempt to define a weak super-diagonal commutator off-support norm \( \mathcal{O}^\Sigma_{p,q,r} \) which would be related to the norm \( \mathcal{O}_p^{\Sigma,\Sigma} \) of Definition 5.8 in parallel with how Definitions 3.26 and 3.28 of \( \mathcal{O}_p^{\infty,\Sigma} \) and \( \mathcal{O}_p^{\infty,\Sigma} \) interrelate. The problem is that we should then simultaneously have \( \mathcal{O}_p^{\Sigma,\Sigma}(b; K) \lesssim \|[b, T]\|_{L^p \times L^q \to L^{\infty}} \) and also a left-over produce that allows a termination of the proof. We did not find a way of doing both of these simultaneously.

Taking into account the above discussion, we raise the following

**Conjecture 5.22** Let \( T \) be a non-degenerate bilinear SIO. Taking into account the above discussion of Section 5.1, let \( r = 3/4 \) and \( s = p = q = 9/4 \). Then, do we have

\[
\|b\|_{L^r} \lesssim \|[b, T]\|_{L^p \times L^q \to L^r}?
\]

**6 Extension to Multilinear Setting**

It is straightforward to extend all definitions and results to the multilinear setting. Without going into all the details, we give the relevant definitions and ideas. First of all, a multilinear SIO of either variable or rough kernel acts formally on an \( n \)-tuple of test functions as \( T : \prod_{i=1}^n \Sigma(\mathbb{R}^d) \to L^1_{\text{loc}}(\mathbb{R}^d) \). The relevant off-support and kernel assumptions can be inferred from the bilinear case, see Section 2 above, or see [17] and [8] respectively for the definitions of variable and rough kernel multilinear SIOs. What is novel is the definition of non-degeneracy.
Definition 6.1 A variable kernel $K : \left( \prod_{i=0}^{n} \mathbb{R}^d \right) \setminus \Delta \to \mathbb{C}$ of a $n$-linear singular integral is said to be non-degenerate in the slot $i = 1, \ldots, n$, provided that for all $y_i \in \mathbb{R}^d$ and $r > 0$, there exists $y_j \in \mathbb{R}^d$, $j \neq i$, so that

$$\max_{a,b \in \{0,1,\ldots,n\}} |y_a - y_b| > r, \quad |K(y_0, y_1, \ldots, y_n)| \gtrsim r^{-nd}. \quad (6.2)$$

Further, we say that $K$ is non-degenerate, provided that it is non-degenerate in all the slots $i = 1, \ldots, n$.

Remark 6.3 As before in the bilinear case of the variable kernel, we only need to assume that $K$ is non-degenerate in the slot $i$ to prove the commutator lower bounds for $[b, T]_i$.

Definition 6.4 A kernel $K_\Omega$ of a rough $n$-linear SIO is said to be non-degenerate, provided that $\Omega \neq 0$, i.e. it has at least one non-zero Lebesgue point $(\theta_1, \theta_2, \ldots, \theta_n) \in S^{n-1}$.

The multilinear extension of Theorem 1.6 is the following.

Theorem 6.5 Let $b \in L^1_{\text{loc}}(\mathbb{R}^d; \mathbb{C})$, let $T$ be a non-degenerate $n$-linear Calderón-Zygmund operator, let $\frac{1}{r} < r < \infty$ and $p_i \in (1, \infty)$, for $i = 1, \ldots, n$. Then, there holds that

$$\|[b, T]_i\|_{\prod_{i=1}^{n} L^{p_i} \to L^r} \sim \begin{cases} \|b\|_{\text{BMO}}, & \text{if } \frac{1}{r} = \sum_{i=1}^{n} \frac{1}{p_i} - \frac{1}{r}, \\ \|b\|_{\dot{C}^{\alpha,0}}, & \text{if } \frac{1}{r} < \sum_{i=1}^{n} \frac{1}{p_i} \\ \|b\|_{L^r}, & \text{if } \frac{1}{r} > \sum_{i=1}^{n} \frac{1}{p_i}, \quad r \geq 1. \end{cases}$$

To obtain the lower bounds in the sub-diagonal and diagonal cases one follows the main line of the proof, establishing first multilinear versions of the technical Propositions 3.2 and 3.14, then of the off-support commutator norms of Definitions 3.26 and 3.28, then the required steps leading to the generalization of Propositions 3.32 and 3.34, after which Theorem 4.1 follows immediately, yielding us both lower bounds. The diagonal upper bound in follows with a similar reasoning as in the bilinear case. The sub-diagonal upper bound follows with the same proof as we gave here, in the end bootstrapping to a multilinear fractional operator of Kenig and Stein [14].

To obtain the super-diagonal lower bound we proceed identically, first extend Definition 5.8 to the multilinear setting, note that Lemmas 5.4, 5.6, 5.7 need no alteration, and finally Proposition 5.11 follows with the same proof as in the bilinear case. The upper bound is a similarly immediate generalization of Proposition 5.1. Taken together, a multilinear generalization of 5.21 follows.

Finally, the discussion of Section 5.1 and the Conjecture 5.22 apply with obvious modifications in the multilinear setting.

Acknowledgements I thank Henri Martikainen for helpful discussions and comments.

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

Data availability statement All data generated or analysed during this study are included in this published article.
Off-Diagonal Estimates...

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Airta, E., Hytönen, T., Li, K., Martikainen, H., Oikari, T.: Off-Diagonal Estimates for Bi-Commutators, Int. Math. Res. Not. (202109).
2. Airta, E., Li, K., Martikainen, H., Vuorinen, E.: Some new weighted estimates on product spaces. Indiana Univ. Math. J. 71, 37–63 (2022).
3. Alexander, B.: Weighted estimates for rough bilinear singular integrals via sparse domination. New York J. Math. 23 (2017).
4. Chaffee, L.: Characterizations of bounded mean oscillation through commutators of bilinear singular integral operators. Proc. R. Soc. Edinb. 146(6), 1159–1166 (2016).
5. Coifman, R., Lions, P.-L., Meyer, Y., Semmes, S.: Compensated compactness and Hardy spaces. J. Math. Pures Appl. 9, 247–286 (1993).
6. Coifman, R.R., Rochberg, R., Weiss, G.: Factorization theorems for Hardy spaces in several variables. Ann. of Math. (2) 103(3), 611–635 (1976).
7. Grafakos, L.: Modern Fourier Analysis. In: Graduate Texts in Mathematics. 3rd edn., vol. 250. Springer, New York (2014). (eng)
8. Grafakos, L., He, D., Honzik, P., Park, B.: $L^2 \times \cdots \times L^2 \to L^{2/m}$ boundedness for maximal multilinear operators. arXiv:2012.10837 (2020).
9. Grafakos, L., Torres, R.H.: Multilinear Calderón-Zygmund theory. Adv. Math. 165(1), 124–164 (2002).
10. Guo, W., Lian, J., Wu, H.: The unified theory for the necessity of bounded commutators and applications. J. Geom. Anal. 30 (201709).
11. Hytönen, T.: The $L^p$-to-$L^{q}$ boundedness of commutators with applications to the Jacobian operator. J. Math. Pures Appl. 156, 351–391 (2021)
12. Li, J., Wick, B.: Weak factorizations of the Hardy space $H^1$ in terms of multilinear Riesz transforms. Can. Math. Bull. 60 (201705).
13. Janson, S.: Mean oscillation and commutators of singular integral operators. Ark. Mat. 16, 263–270 (1978).
14. Kenig, C., Stein, E.: Multilinear estimates and fractional integration. Math. Res. Lett. 6, 1–15 (1999).
15. Kuffner, M.-J.: Weak factorization of the Hardy space $H^p$ for small values of $p$, in the multilinear setting. J. Math. Anal. Appl. 485(1), 123711 (2020).
16. Lerner, A.K., Nazarov, F.: Intuitive dyadic calculus: The basics. Expositiones Mathematicae 37(3), 225–265 (2019).
17. Li, K.: Multilinear commutators in the two-weight setting. to appear in Bull. Lond. Math. Soc. arXiv:2006.09071 (2020).
18. Lindberg, S.: On the Hardy space theory of compensated compactness quantities. Arch. Ration. Mech. Anal. 224, 709–742 (2017).
19. Muscalu, C., Schlag, W.: Classical and multilinear harmonic analysis. In: Cambridge Studies in Advanced Mathematics 138, vol. II. Cambridge University Press, Cambridge (2013).
20. Nehari, Z.: On bounded bilinear forms. Ann. of Math. 65, 153–162 (1957).
21. Oikari, T.: Off-diagonal estimates for commutators of bi-parameter singular integrals. arXiv:2010.01538 (2020).
22. Uchiyama, A.: On the compactness of operators of Hankel type. Tohoku Math. J. 30(1), 163–171 (1978). https://doi.org/10.2748/tmj/1178230105.

Publisher’s Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.