UMD-Extensions of Calderón–Zygmund Operators with Mild Kernel Regularity

Emil Airta $^1$ · Henri Martikainen $^{1,2}$ · Emil Vuorinen $^1$

Received: 30 January 2022 / Accepted: 15 June 2022 / Published online: 13 July 2022 © The Author(s) 2022

Abstract
Armed with new methods we revisit a result of Figiel concerning the UMD-extensions of linear Calderón–Zygmund operators with mild kernel regularity and then extend our new proof to the multilinear setting improving recent UMD-valued estimates of multilinear singular integrals.

Keywords  Singular integrals · Kernel regularity · UMD spaces · Multilinear analysis

Mathematics Subject Classification  42B20

1 Introduction

The UMD—unconditional martingale differences—property of a Banach space $X$ is a well-known necessary and sufficient condition for the boundedness of various singular integral operators (SIOs)
$$T f (x) = \int_{\mathbb{R}^d} K(x, y) f(y) \, dy$$
onumber

on $L^p(\mathbb{R}^d; X) = L^p(X)$. A Banach space $X$ has the UMD property if $X$-valued martingale difference sequences converge unconditionally in $L^p$ for some $p \in (1, \infty)$.

By Burkholder [2] and Bourgain [1] we have that $X$ is a UMD space if and only if a particular singular integral operator—the Hilbert transform

$$H f(x) = \text{p. v.} \int_{\mathbb{R}} \frac{f(y) \, dy}{x-y}$$

admits an $L^p(\mathbb{X})$-bounded extension. This theory quickly advanced up to the vector-valued $T_1$ theorem of Figiel [12]. It is important to understand that the fundamental result that all scalar-valued $L^2$ bounded SIOs can be extended to act boundedly on $L^p(X)$, $p \in (1, \infty)$, goes through this key theorem.

The deep work of Figiel also contains estimates for the required kernel regularity in terms of some key characteristics of the UMD space $X$. This concerns the required regularity of the continuity-moduli $\omega$ appearing in the various kernel estimates, such as,

$$|K(x, y) - K(x', y)| \leq \omega \left( \frac{|x - x'|}{|x - y|} \right) \frac{1}{|x - y|^d}, \quad |x - x'| \leq |x - y|/2.$$ 

Recently, Grau de la Herrán and Hytönen [16] proved that the modified Dini condition

$$\|\omega\|_{\text{Dini}_\alpha} := \int_0^1 \omega(t) \left( 1 + \log \frac{1}{t} \right)^{\alpha} \frac{dt}{t}$$

with $\alpha = \frac{1}{2}$ is sufficient to prove a scalar-valued $T_1$ theorem even with an underlying measure $\mu$ that can be non-doubling. This matches the best known [6, 12] sufficient condition for the classical homogeneous $T_1$ theorem [5]. The exponent $\alpha = \frac{1}{2}$ has a fundamental feeling in all of the existing arguments—it seems very difficult to achieve a $T_1$ theorem with a weaker assumption.

It turns out that for completely general UMD spaces the threshold $\alpha = 1/2$ needs to be replaced by a more complicated expression, while in the simpler function lattice case $\alpha = 1/2$ suffices by a much more elementary argument. We revisit these fundamental linear results of Figiel using new optimized dyadic methods and obtain a modern proof of the following theorem. In our terminology, a Calderón–Zygmund operator (CZO) is an SIO that satisfies the $T_1$ assumptions (equivalently, $T : L^2 \to L^2$ boundedly).

**Theorem 1.1** Let $T$ be a linear $\omega$-CZO and $X$ be a UMD space with type $r \in (1, 2]$ and cotype $q \in [2, \infty)$. If $\omega \in \text{Dini}_{1/\min(r,q')}$, we have

$$\|T f\|_{L^p(X)} \lesssim \|f\|_{L^p(X)}, \quad p \in (1, \infty).$$

See the main text for the exact definitions. If $X$ is a Hilbert space, then $r = q = 2$ and we get the usual $\alpha = 1/2$. Again, this theory is relevant for UMD spaces that go beyond the function lattices, such as, non-commutative $L^p$ spaces.
A major part of our arguments has to do with the extension of this result to the multilinear setting. A basic model of an $n$-linear SIO $T$ in $\mathbb{R}^d$ is obtained by setting
\[
T(f_1, \ldots, f_n)(x) = U(f_1 \otimes \cdots \otimes f_n)(x, \ldots, x), \quad x \in \mathbb{R}^d, \quad f_i : \mathbb{R}^d \to \mathbb{C},
\]
where $U$ is a linear singular integral operator in $\mathbb{R}^d$. See e.g. Grafakos–Torres [15] for the basic theory. Multilinear SIOs appear in applications ranging from partial differential equations to complex function theory and ergodic theory. For example, $L^p$ estimates for the homogeneous fractional derivative $D^\alpha f = \mathcal{F}^{-1}(|\xi|^\alpha \hat{f}(\xi))$ of a product of two or more functions—the fractional Leibniz rules—are used in the area of dispersive equations. Such estimates descend from the multilinear Hörmander–Mihlin multiplier theorem of Coifman and Meyer [3]—See e.g. Kato and Ponce [23] and Grafakos and Oh [14].

The multilinear analogue of Theorem 1.1 extending the recent work [9] goes as follows.

**Theorem 1.2** Let $\{X_1, \ldots, X_{n+1}\}$ be a UMD Hölder tuple as in Definition 3.21, and denote the cotype of $X_j$ by $s_j$. Suppose that $T$ is an $n$-linear $\omega$-CZO with \( \omega \in \text{Dini}_\alpha \), where
\[
\alpha = \frac{1}{\min \left\{ \frac{(n+1)}{n}, s'_1, \ldots, s'_{n+1} \right\}}.
\]
Then for all exponents $1 < p_1, \ldots, p_n \leq \infty$ and $1/r = \sum_{j=1}^n 1/p_j > 0$ we have
\[
\|T(f_1, \ldots, f_n)\|_{L^r(\prod_{j=1}^n X_j^*)} \lesssim \prod_{j=1}^n \|f_j\|_{L^{p_j}(X_j^*)}.
\]

Until recently, vector-valued extensions of multilinear SIOs had mostly been studied in the framework of $\ell^p$ spaces and function lattices, rather than general UMD spaces—see e.g. [4, 13, 24, 25, 28]. Taking the work [7] much further, the paper [9] finally established $L^p$ bounds for the extensions of $n$-linear SIOs with the usual Hölder modulus of continuity to tuples of UMD spaces tied by a natural product structure, such as, the composition of operators in the Schatten–von Neumann subclass of the algebra of bounded operators on a Hilbert space. In [10] the bilinear case of [9] was applied to prove UMD-extensions for modulation invariant singular integrals, such as, the bilinear Hilbert transform. See also [8] for the related operator-valued theory. With new and refined methods, we are able to prove the above Figiel type result, Theorem 1.2, in the multilinear setting. The proofs of $T1$ theorems display a fundamental structural decomposition of SIOs into their cancellative parts and so-called paraproducts. It is this structure that is extremely important for obtaining further estimates beyond the initial scalar-valued $L^p$ boundedness. The original dyadic representation theorem of Hytönen [18, 19] (extending an earlier special case of Petermichl [29]) provides a decomposition of the cancellative part of an SIO into so-called dyadic shifts. In [16] a new type of representation theorem appears, where the key difference to the original representation theorems [18, 19] is that the decomposition of the cancellative part is
in terms of different operators that package multiple dyadic shifts into one and offer more efficient bounds when it comes to kernel regularity. Aiming to prove Theorem 1.2 we develop these tools in the multilinear setting and present a useful, explicit and clear exposition of the appearing new dyadic model operators.

**Notation**

Throughout this paper \( A \lesssim B \) means that \( A \leq CB \) with some constant \( C \) that we deem unimportant to track at that point. We write \( A \sim B \) if \( A \lesssim B \lesssim A \).

Given a dyadic grid \( D \), \( I \in D \) and \( k \in \mathbb{Z}, k \geq 0 \), we use the following notation:

1. \( \ell(I) \) is the side length of \( I \).
2. \( I(k) \in D \) is the \( k \)th parent of \( I \), i.e., \( I \subset I(k) \) and \( \ell(I(k)) = 2^k \ell(I) \).
3. \( \text{ch}(I) \) is the collection of the children of \( I \), i.e., \( \text{ch}(I) = \{ J \in D : J^{(1)} = I \} \).
4. \( E_I f = \langle f \rangle_I I \) is the averaging operator, where \( \langle f \rangle_I = \frac{1}{|I|} \int_I f \).
5. \( E_{I,k} f \) is defined via
   \[
   E_{I,k} f = \sum_{J \in D \atop J^{(k)} = I} E_J f.
   \]
6. \( \Delta_I f \) is the martingale difference \( \Delta_I f = \sum_{J \in \text{ch}(I)} E_J f - E_I f \).
7. \( \Delta_{I,k} f \) is the martingale difference block
   \[
   \Delta_{I,k} f = \sum_{J \in D \atop J^{(k)} = I} \Delta_J f.
   \]
8. \( P_{I,k} f \) is the following sum of martingale difference blocks
   \[
   P_{I,k} f = \sum_{j=0}^{k} \Delta_{I,j} f = \sum_{J \in D \atop J \subset I \atop \ell(J) \geq 2^k \ell(I)} \Delta_J f.
   \]

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 \( h^0_J = |J|^{-1/2} 1_J \) and \( h^1_J = |J|^{-1/2} (1_{J_l} - 1_{J_r}) \). Let now \( I = I_1 \times \cdots \times I_d \subset \mathbb{R}^d \) be a cube, and define the Haar function \( h^\eta_I, \eta = (\eta_1, \ldots, \eta_d) \in \{0, 1\}^d \), by setting
\[
h^\eta_I = h^\eta_{I_1} \otimes \cdots \otimes h^\eta_{I_d}.
\]

If \( \eta \neq 0 \) the Haar function is cancellative: \( \int h^\eta_I = 0 \). We exploit notation by suppressing the presence of \( \eta \), and write \( h_I \) for some \( h^\eta_I, \eta \neq 0 \). Notice that for \( I \in D \) we have \( \Delta_I f = \langle f, h_I \rangle h_I \) (where the finite \( \eta \) summation is suppressed), \( \langle f, h_I \rangle := \int f h_I \).
2 Singular Integrals

Let $\omega$ be a modulus of continuity: an increasing and subadditive function with $\omega(0) = 0$. A relevant quantity is the modified Dini condition

$$\|\omega\|_{\text{Dini}_\alpha} := \int_0^1 \omega(t) \left(1 + \log \frac{1}{t}\right)^\alpha \frac{dt}{t}, \quad \alpha \geq 0.$$ (2.1)

In practice, the quantity (2.1) arises as follows:

$$\sum_{k=1}^{\infty} \omega(2^{-k})k^\alpha = \sum_{k=1}^{\infty} \frac{1}{\log 2} \int_{2^{-k}}^{2^{-k+1}} \omega(t) \left(1 + \log \frac{1}{t}\right)^\alpha \frac{dt}{t} \lesssim \int_0^1 \omega(t) \left(1 + \log \frac{1}{t}\right)^\alpha \frac{dt}{t}.$$ (2.2)

For many standard arguments $\alpha = 0$ is enough. For the $T1$ type arguments we will—at the minimum—always need $\alpha = 1/2$. When we do UMD-extensions beyond function lattices, we will need a bit higher $\alpha$ depending on the so-called type and cotype constants of the underlying UMD space $X$.

Multilinear singular integrals

A function

$$K : \mathbb{R}^{(n+1)d} \setminus \Delta \to \mathbb{C}, \quad \Delta = \{x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{(n+1)d} : x_1 = \cdots = x_{n+1}\},$$

is called an $n$-linear $\omega$-Calderón–Zygmund kernel if it holds that

$$|K(x)| \leq \frac{C_K}{\left(\sum_{m=1}^{n} |x_{n+1} - x_m|\right)^{dn}},$$

and for all $j \in \{1, \ldots, n+1\}$ it holds that

$$|K(x) - K(x')| \leq \omega \left(\frac{|x_j - x'_j|}{\sum_{m=1}^{n} |x_{n+1} - x_m|}\right) \frac{1}{\left(\sum_{m=1}^{n} |x_{n+1} - x_m|\right)^{dn}}$$

whenever $x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{(n+1)d} \setminus \Delta$ and $x' = (x_1, \ldots, x_{j-1}, x'_j, x_{j+1}, \ldots x_{n+1}) \in \mathbb{R}^{(n+1)d}$ satisfy

$$|x_j - x'_j| \leq 2^{-1} \max_{1 \leq m \leq n} |x_{n+1} - x_m|.$$
Definition 2.3 An $n$-linear operator $T$ defined on a suitable class of functions—e.g. on the linear combinations of cubes—is an $n$-linear $\omega$-SIO with an associated kernel $K$, if we have

$$\langle T(f_1, \ldots, f_n), f_{n+1} \rangle = \int_{\mathbb{R}^{(n+1)d}} K(x_{n+1}, x_1, \ldots, x_n) \prod_{j=1}^{n+1} f_j(x_j) \, dx$$

whenever $\text{spt} \ f_i \cap \text{spt} \ f_j = \emptyset$ for some $i \neq j$.

Definition 2.4 We say that $T$ is an $n$-linear $\omega$-CZO if the following conditions hold:

- $T$ is an $n$-linear $\omega$-SIO.
- We have that $\| T^{m*}(1, \ldots, 1) \|_{\text{BMO}} := \sup_D \sup I \in D \left( \frac{1}{|I|} \sum_{f_j \in I} |\langle T^{m*}(1, \ldots, 1), h_j \rangle|^2 \right)^{1/2} < \infty$ for all $m \in \{0, \ldots, n\}$. Here $T^{0*} := T$, $T^{m*}$ denotes the $m$th, $m \in \{1, \ldots, n\}$, adjoint

$$\langle T(f_1, \ldots, f_n), f_{n+1} \rangle = \langle T^{m*}(f_1, \ldots, f_{m-1}, f_{n+1}, f_{m+1}, \ldots, f_n), f_m \rangle$$

of $T$, and the pairings $\langle T^{m*}(1, \ldots, 1), h_j \rangle$ have a standard $T \, 1$ type definition with the aid of the kernel $K$.
- We have that $\| T \|_{\text{WBP}} := \sup_D \sup I \in D |I|^{-1} |\langle T(1, \ldots, 1), 1_I \rangle| < \infty$.

Model Operators

Let $i = (i_1, \ldots, i_{n+1}), i_j \in \{0, 1, \ldots\}$, and let $D$ be a dyadic lattice in $\mathbb{R}^d$. An operator $S_i$ is called an $n$-linear dyadic shift if it has the form

$$\langle S_i(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{K \in D} \langle A_K(f_1, \ldots, f_n), f_{n+1} \rangle,$$  \hspace{1cm} (2.5)

where

$$\langle A_K(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{I_1, \ldots, I_{n+1} \in D} a_{K,(i_j)} \prod_{j=1}^{n+1} (f_j, \tilde{h}_{I_j}).$$

\[ \text{Birkhäuser} \]
Here $a_{K,(I_j)} = a_{K,I_1,...,I_{n+1}}$ is a scalar satisfying the normalization

$$|a_{K,(I_j)}| \leq \frac{\prod_{j=1}^{n+1} |I_j|^{1/2}}{|K|^n},$$

and there exist two indices $j_0, j_1 \in \{1, \ldots, n+1\}$, $j_0 \neq j_1$, so that $\tilde{h}_{I_{j_0}} = h_{I_{j_0}}$, $\tilde{h}_{I_{j_1}} = h_{I_{j_1}}$ and for the remaining indices $j \notin \{j_0, j_1\}$ we have $\tilde{h}_{I_j} \in \{h_{I_j}^0, h_{I_j}^1\}$.

A modified $n$-linear shift $Q_k$, $k \in \{1, 2, \ldots\}$, has the form

$$\langle Q_k(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{K \in \mathcal{D}} \langle B_K(f_1, \ldots, f_n), f_{n+1} \rangle,$$

where

$$\langle B_K(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{I_1^{(k)} = \ldots = I_{n+1}^{(k)} = K} a_{K,(I_j)} \left[ \prod_{j=1}^{n} \langle f_j, h_{I_j}^0 \rangle - \prod_{j=1}^{n} \langle f_j, h_{I_{n+1}}^0 \rangle \right] \langle f_{n+1}, h_{I_{n+1}} \rangle, \quad (2.6)$$

or $B_K$ has one of the other symmetric forms, where the role of $f_{n+1}$ is replaced by some other $f_j$. The coefficients satisfy the same (but now $|I_1| = \ldots = |I_{n+1}|$) normalization

$$|a_{K,(I_j)}| \leq \frac{|I_1|^{(n+1)/2}}{|K|^n}.$$

An $n$-linear dyadic paraproduct $\pi = \pi_D$ also has $n + 1$ possible forms, but there is no complexity associated to them. One of the forms is

$$\langle \pi(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{I \in \mathcal{D}} a_I \prod_{j=1}^{n} \langle f_j, I \rangle \langle f_{n+1}, h_I \rangle,$$

where the coefficients satisfy the usual BMO condition

$$\sup_{I_0 \in \mathcal{D}} \left( \frac{1}{|I_0|} \sum_{I \subseteq I_0} |a_I|^2 \right)^{1/2} \leq 1. \quad (2.7)$$

In the remaining $n$ alternative forms the cancellative Haar function $h_I$ is in a different position.

When we represent a CZO we will have modified dyadic shifts $Q_k$, standard dyadic shifts of the very special form $S_k,...,k$ and paraproducts $\pi$. Dyadic shifts $S_k,...,k$ are simply easier versions of the operators $Q_k$. Paraproducts do not involve a complexity parameter and are thus inherently not even relevant for the kernel regularity considerations (we just need their boundedness).
Remark 2.8  At least in the linear situation, we can easily unify the study of shifts \( S_k \) and modified shifts \( Q_k \). This viewpoint could work in the multilinear generality also (with some tensor product formalism), but we did not pursue it. We can understand a modified linear shift to have the more general form \( Q_k \), \( k = 0, 1, \ldots \), where

\[
\langle Q_k f, g \rangle = \sum_{K \in D} \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle f, h_I \rangle \langle g, H_{I,J} \rangle
\]

(2.9)

or

\[
\langle Q_k f, g \rangle = \sum_{K \in D} \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle f, H_{I,J} \rangle \langle g, h_J \rangle,
\]

(2.10)

the constants \( a_{IJK} \) satisfy the usual normalization and the functions \( H_{I,J} \) satisfy

1. \( H_{I,J} \) is supported on \( I \cup J \) and constant on the children of \( I \) and \( J \), i.e., we have

\[
H_{I,J} = \sum_{L \in \text{ch}(I) \cup \text{ch}(J)} b_L 1_L, \quad b_L \in \mathbb{R},
\]

2. \( |H_{I,J}| \leq |I|^{-1/2} \) and
3. \( \int H_{I,J} = 0 \).

2.1 The Threshold \( \alpha = 1/2 \)

We quickly explain the role of the regularity threshold \( \alpha = 1/2 \), which appears naturally in the fundamental scalar-valued theory.

Lemma 2.11  Let \( p \in (1, \infty) \). There holds that

\[
\left\| \left( \sum_{K \in D} |P_{K,k} f|^2 \right)^{1/2} \right\|_{L^p} \sim \sqrt{k + 1} \| f \|_{L^p}, \quad k \in \{0, 1, 2, \ldots \}.
\]

Proof  If \( f_i \in L^p \) then a basic vector-valued square function estimate says that

\[
\left\| \left( \sum_{i=0}^\infty \sum_{I \in D} |\Delta_I f_i|^2 \right)^{1/2} \right\|_{L^p} \sim \left\| \left( \sum_{i=0}^\infty |f_i|^2 \right)^{1/2} \right\|_{L^p}.
\]

(2.12)

Let \( K \in D \). We have that

\[
\sum_{I \in D} |\Delta_I P_{K,k} f|^2 = \sum_{j=0}^k |\Delta_{K,j} f|^2.
\]
Thus, (2.12) gives that

\[
\left\| \left( \sum_{K \in D} |P_{K} f|^2 \right)^{1/2} \right\|_{L^p} \sim \left\| \left( \sum_{K \in D} \sum_{j=0}^{k} |\Delta_{K,j} f|^2 \right)^{1/2} \right\|_{L^p}
\]

\[
= \left\| \left( \sum_{j=0}^{k} \sum_{I \in D} |\Delta_{I} f|^2 \right)^{1/2} \right\|_{L^p} \sim \sqrt{k+1} \|f\|_{L^p}.
\]

Proposition 2.13 Suppose that \(Q_k\) is an \(n\)-linear modified shift. Let \(1 < p_j < \infty\) with \(\sum_{j=1}^{n+1} 1/p_j = 1\). Then we have

\[
|\langle Q_k(f_1, \ldots, f_n), f_{n+1} \rangle| \lesssim (k + 1)^{1/2} \prod_{j=1}^{n+1} \|f_j\|_{L^{p_j}}.
\]  

(2.14)

Proof We may assume that \(Q_k\) has the form (2.6). Notice that if \(I^{(k)} = K\) then we have

\[
\langle f, h_0^I \rangle = \langle E_{K} f, h_0^I \rangle = \langle E_{K} f + P_{K,k-1} f, h_0^I \rangle.
\]

Using this we have for \(I_1^{(k)} = \cdots = I_{n+1}^{(k)} = K\) that

\[
\prod_{j=1}^{n} \langle f_j, h_0^I \rangle = \langle P_{K,k-1} f_1, h_0^I \rangle \prod_{j=2}^{n} \langle f_j, h_0^I \rangle
\]

\[
+ \langle E_{K} f_1, h_0^I \rangle \langle P_{K,k-1} f_2, h_0^I \rangle \prod_{j=3}^{n} \langle f_j, h_0^I \rangle
\]

\[
+ \cdots + \prod_{j=1}^{n-1} \langle E_{K} f_j, h_0^I \rangle \langle P_{K,k-1} f_n, h_0^I \rangle
\]

\[
+ |I_{n+1}|^{n/2} \prod_{j=1}^{n} \langle f_j \rangle_K
\]

and

\[
\prod_{j=1}^{n} \langle f_j, h_0^{I_{n+1}} \rangle = \langle P_{K,k-1} f_1, h_0^{I_{n+1}} \rangle \prod_{j=2}^{n} \langle f_j, h_0^{I_{n+1}} \rangle
\]

\[
+ \langle E_{K} f_1, h_0^{I_{n+1}} \rangle \langle P_{K,k-1} f_2, h_0^{I_{n+1}} \rangle \prod_{j=3}^{n} \langle f_j, h_0^{I_{n+1}} \rangle
\]
\[
+ \cdots + \prod_{j=1}^{n-1} \left( E_{K \cdot f_j, h_{I_{n+1}^0}} \right) \left( P_{K, k-1 \cdot f_n, h_{I_{n+1}^0}} \right) + \left| I_{n+1} \right|^{n/2} \prod_{j=1}^{n} \left( f_j \right)_K .
\]

We see that the last terms of these expansions cancel out in the difference \( \prod_{j=1}^{n} \left( f_j, h_{I_{j}^0}^0 \right) - \prod_{j=1}^{n} \left( f_j, h_{I_{j}^0}^1 \right) \). It remains to estimate the other terms one by one.

We pick the concrete (but completely representative) term \( \left( E_{K \cdot f_1, h_{I_{1}^0}^1} \right) \left( P_{K, k-1 \cdot f_2, h_{I_{2}^0}^0} \right) \prod_{j=3}^{n} \left( f_j, h_{I_{j}^0}^0 \right) \) from the expansion of \( \prod_{j=1}^{n} \left( f_j, h_{I_{j}^0}^1 \right) \) and look at

\[
\sum_{K} \sum_{I_{j}^{(k)}=\cdots=I_{I_{n+1}^0}=K} \left| a_{K, (I_{j})} \left( E_{K \cdot f_1, h_{I_{I_{n+1}^0}^0}^0} \right) \left( P_{K, k-1 \cdot f_2, h_{I_{I_{n+1}^0}^0}^0} \right) \prod_{j=3}^{n} \left( f_j, h_{I_{I_{j}^0}^0}^0 \right) \left( f_{n+1}, h_{I_{n+1}^0}^1 \right) \right|
\leq \sum_{K} \sum_{I_{j}^{(k)}=\cdots=I_{I_{n+1}^0}=K} \frac{1}{\left| K \right|^n} \int_{I_{1}} \left| E_{K \cdot f_1} \right| \int_{I_{2}} \left| P_{K, k-1 \cdot f_2} \right| \prod_{j=3}^{n} \int_{I_{n+1}^0} \left| f_j \right| \int_{I_{n+1}^0} \left| \Delta_{K, k \cdot f_{n+1}} \right|
\leq \sum_{K} \int_{I_{1}} \prod_{j=1}^{n} \left( f_j \right)_K \left( \left| P_{K, k-1 \cdot f_2} \right| \right)_K \left| \Delta_{K, k \cdot f_{n+1}} \right|
\leq \int_{I_{1}} \prod_{j=1}^{n} \left( f_j \right)_K \left( \sum_{K} \left| M_{P_{K, k-1 \cdot f_2}} \right| \right)^{1/2} \left( \sum_{K} \left| \Delta_{K, k \cdot f_{n+1}} \right| \right)^{1/2}.
\]

It remains to use Hölder’s inequality, maximal function and square function estimates and Lemma 2.11. We remark that the estimate for \( f_{n+1} \) is, indeed, just the usual square function estimate, since

\[
\sum_{K} \left| \Delta_{K, k \cdot f_{n+1}} \right|^2 = \sum_{I} \left| \Delta_{I \cdot f_{n+1}} \right|^2 .
\]

We now pick the corresponding term \( \left( E_{K \cdot f_1, h_{I_{I_{n+1}^0}^0}^0} \right) \left( P_{K, k-1 \cdot f_2, h_{I_{I_{n+1}^0}^0}^0} \right) \prod_{j=3}^{n} \left( f_j, h_{I_{I_{j}^0}^0}^0 \right) \) from the expansion of \( \prod_{j=1}^{n} \left( f_j, h_{I_{I_{j}^0}^0}^1 \right) \) and look at

\[
\sum_{K} \sum_{I_{j}^{(k)}=\cdots=I_{I_{n+1}^0}=K} \left| a_{K, (I_{j})} \left( E_{K \cdot f_1, h_{I_{I_{n+1}^0}^0}^0} \right) \left( P_{K, k-1 \cdot f_2, h_{I_{I_{n+1}^0}^0}^0} \right) \prod_{j=3}^{n} \left( f_j, h_{I_{I_{j}^0}^0}^0 \right) \left( f_{n+1}, h_{I_{n+1}^0}^1 \right) \right|
\times \prod_{j=3}^{n} \left( f_j, h_{I_{I_{j}^0}^0}^0 \right) \left( f_{n+1}, h_{I_{n+1}^0}^1 \right)
\leq \sum_{K} \sum_{I_{j}^{(k)}=\cdots=I_{I_{n+1}^0}=K} \frac{1}{\left| K \right|^n} \int_{I_{n+1}^0} \left| E_{K \cdot f_1} \right| \int_{I_{n+1}^0} \left| P_{K, k-1 \cdot f_2} \right|
\times \prod_{j=3}^{n} \int_{I_{n+1}^0} \left| f_j \right| \int_{I_{n+1}^0} \left| \Delta_{K, k \cdot f_{n+1}} \right| .
\]
Notice that

\[ \sum_{I_j^{(k)} = K} 1 = \frac{1}{|I_{n+1}|} \sum_{I_j^{(k)} = K} |I_j| = \frac{|K|}{|I_{n+1}|}. \]

We are thus left with

\[
\sum_{K} \langle |f_1| \rangle_K \sum_{I_{n+1}^{(k)} = K I_{n+1}} \int \langle |P_{K,k-1} f_2| \rangle_{I_{n+1}} \prod_{j=3}^{n} \langle |f_j| \rangle_{I_{n+1}} |\Delta_{K,k} f_{n+1}| \\
\leq \sum_{K} \langle |f_1| \rangle_K \int_K MP_{K,k-1} f_2 \prod_{j=3}^{n} Mf_j \cdot |\Delta_{K,k} f_{n+1}| \\
\leq \int \prod_{j=1}^{n} Mf_j \left( \sum_{K} |MP_{K,k-1} f_2|^2 \right)^{1/2} \left( \sum_{K} |\Delta_{K,k} f_{n+1}|^2 \right)^{1/2}.
\]

This is the same upper bound as in the first case, and thus handled with in the same way. \(\square\)

**Remark 2.17** Proposition 2.13 considers only the Banach range boundedness of \(Q_k\). We can, in any case, upgrade the boundedness to the full range with standard methods when we consider CZOs.

Our representation of \(T\) will involve \(\sum_{k=0}^{\infty} \omega(2^{-k}) Q_k(f_1, \ldots, f_n)\), and thus by (2.2) and Proposition 2.13 we will always need Dini1/2. The above proof readily generalises to so-called UMD function lattices. In Sect. 3 we tackle the much deeper case of general UMD spaces.

**Modified Shifts are Sums of Standard Shifts**

The standard linear shifts satisfy the *complexity free* bound

\[ \|S_{i_1,i_2} f\|_{L^p} \lesssim \|f\|_{L^p}, \quad p \in (1, \infty). \]

Similar estimates hold in the multilinear generality—for example, the following complexity free bilinear estimate is true

\[
\|S_{i_1,i_2,i_3} (f_1, f_2)\|_{L^{q_3}} \lesssim \|f_1\|_{L^{p_1}} \|f_2\|_{L^{p_2}}, \quad \forall 1 < p_1, p_2 \leq \infty, \quad \frac{1}{2} < q_3 < \infty, \quad \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{q_3}.
\]

We need roughly \(nk\) shifts to represent an \(n\)-linear modified shift \(Q_k\) as a sum of standard shifts, see Lemma 2.18 below. Therefore, these estimates lose to estimates like (2.14), and would lead to Dini1.
The following Lemma 2.18 is of philosophical importance. It should not be resorted to when more efficient estimates can be obtained by the direct study of the operators $Q_k$. If one can accept some loss of regularity, then it can be practical. We present it for completeness and for the big picture.

**Lemma 2.18** Let $Q_k$, $k \in \{1, 2, \ldots\}$, be a modified $n$-linear shift of the form (2.6). Then for some $C \lesssim 1$ we have

$$Q_k = \sum_{m=1}^{n} \sum_{i=0}^{k-1} S_{0, \ldots, 0, i, \ldots, k} - C \sum_{m=1}^{n} \sum_{i=0}^{k-1} S_{0, \ldots, 0, 1, \ldots, 1, i},$$

where in the first sum there are $m - 1$ zeroes and in the second sum $m$ zeroes in the complexity of the shift.

**Remark 2.19** In the proof below we decompose various martingale differences using Haar functions, which strictly speaking leads to the fact that there is an implicit dimensional summation in the above decomposition.

**Proof of Lemma 2.18** The underlying decomposition is, in part, more sophisticated than the one in the beginning of the proof of Proposition 2.13. See the multilinear collapse (2.22) and (2.24). This feels necessary for this result—moreover, we will later use this decomposition strategy when we do general UMD-valued estimates.

Write $b_{K,(I_j)} = |I_1|^{n/2} a_{K,(I_j)}$ so that

$$a_{K,(I_j)} \left[ \prod_{j=1}^{n} \langle f_j, h_{I_j}^0 \rangle - \prod_{j=1}^{n} \langle f_j, h_{I_{n+1}}^0 \rangle \right] = b_{K,(I_j)} \left[ \prod_{j=1}^{n} \langle f_j \rangle_{I_j} - \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}} \right].$$

We then write

$$\prod_{j=1}^{n} \langle f_j \rangle_{I_j} - \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}} = \left[ \prod_{j=1}^{n} \langle f_j \rangle_{I_j} - \prod_{j=1}^{n} \langle f_j \rangle_{K} \right] + \left[ \prod_{j=1}^{n} \langle f_j \rangle_{K} - \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}} \right].$$

We start working with the first term. Notice that

$$\langle f_j \rangle_{I_j} = \langle E_{K,k} f_j \rangle_{I_j} = \langle P_{K,k-1} f_j \rangle_{I_j} + \langle f_j \rangle_{K}.$$ 

Using this we can write

$$\prod_{j=1}^{n} \langle f_j \rangle_{I_j} - \prod_{j=1}^{n} \langle f_j \rangle_{K} = \langle P_{K,k-1} f_1 \rangle_{I_1} \prod_{j=2}^{n} \langle f_j \rangle_{I_j} + \langle f_1 \rangle_{K} \langle P_{K,k-1} f_2 \rangle_{I_2} \prod_{j=3}^{n} \langle f_j \rangle_{I_j} + \cdots + \prod_{j=1}^{n} \langle f_j \rangle_{K} \langle P_{K,k-1} f_n \rangle_{I_n}. \quad (2.20)$$
Consider now, for $m \in \{1, \ldots, n\}$, the following part of the modified shift

$$\langle A_m(f_1, \ldots, f_n), f_{n+1} \rangle \equiv \sum_{k} \sum_{l_1^{(k)} = \cdots = l_{m+1}^{(k)} = K}^{m-1} b_{K,(t_j)} \prod_{j=1}^{m-1} \langle f_j \rangle_K \cdot \langle P_{K,k} f_m \rangle_{I_m} \cdot \prod_{j=m+1}^{n} \langle f_j \rangle_{I_j} \cdot \langle f_{n+1}, h_{I_{n+1}} \rangle.$$  

(2.21)

Next, write

$$\langle P_{K,k-1} f_m \rangle_{I_m} = \sum_{i=0}^{k-1} \sum_{L^{(i)} = K} \langle \Delta_{L} f_m \rangle_{I_m} = \sum_{i=0}^{k-1} \sum_{L^{(i)} = K} \langle f_m, h_L \rangle_{fL_{I_m}}.$$  

We can write $\langle A_m(f_1, \ldots, f_n), f_{n+1} \rangle$ in the form

$$\sum_{i=0}^{k-1} \sum_{L^{(i)} = K} \sum_{l_1^{(k)} = \cdots = l_{m+1}^{(k)} = K}^{m-1} b_{K,(t_j)} |K|^{-(m-1)/2} |h_L|_{I_{n+1}} |h_{I_{n+1}}|^{-(n-m)/2}$$

$$\times \prod_{j=1}^{m-1} \langle f_j, h_{K}^{0} \rangle \cdot \langle f_m, h_L \rangle \cdot \prod_{j=m+1}^{n} \langle f_j, h_{K}^{0} \rangle \cdot \langle f_{n+1}, h_{I_{n+1}} \rangle.$$  

Notice the normalization estimate

$$\sum_{l_1^{(k)} = \cdots = l_{m+1}^{(k)} = K}^{m-1} |b_{K,(t_j)}| |K|^{-(m-1)/2} |L|^{1/2} |I_{n+1}|^{-(n-m)/2}$$

$$\leq \prod_{j=1}^{m-1} |K|^{1/2} \cdot |L|^{1/2} \cdot \prod_{j=m+1}^{n+1} |I_{j}|^{1/2}.$$  

We also have two cancellative Haar functions, so for every $i \in \{0, \ldots, k-1\}$, the inner sum in $A_m$ is a standard $n$-linear shift of complexity $(0, \ldots, 0, i, k, \ldots, k)$, where the $i$ is in the $m$th slot:

$$\langle A_m(f_1, \ldots, f_n), f_{n+1} \rangle = \sum_{i=0}^{k-1} \langle S_{0,\ldots,0,i,k,\ldots,k} f_1, \ldots, f_m \rangle_{f_{n+1}}.$$  

We now turn to the part of the modified shift associated with

$$\prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}} - \prod_{j=1}^{n} \langle f_j \rangle_{K} = \sum_{i=0}^{k-1} \left( \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}^{(i)}} - \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}^{(i+1)}} \right).$$  

(2.22)
Further, we write

\[ \prod_{j=1}^{n} (f_{j})_{(i)}_{n+1} - \prod_{j=1}^{n} (f_{j})_{(i+1)}_{n+1} = \langle \Delta f_{(i+1)}^{1} f_{1} \rangle_{n+1} \prod_{j=2}^{n} (f_{j})_{(i)}_{n+1} \]

\[ + \langle f_{1} \rangle_{n+1} \langle \Delta f_{(i+1)}^{2} f_{2} \rangle_{n+1} \prod_{j=3}^{n} (f_{j})_{(i)}_{n+1} \]

\[ + \cdots + \prod_{j=1}^{n-1} (f_{j})_{(i+1)}_{n+1} \cdot \langle \Delta f_{(i+1)} f_{n} \rangle_{n+1}. \tag{2.23} \]

Consider now, for \( m \in \{1, \ldots, n\} \), the following part of the modified shift

\[ \langle U_{m}(f_{1}, \ldots, f_{n}), f_{n+1} \rangle = \sum_{i=0}^{k-1} \sum_{k} \sum_{i_{1}=i_{n+1}=L} \sum_{j} b_{K,(I_{j})} \prod_{j=1}^{m-1} (f_{j})_{(i)}_{n+1} \cdot \prod_{j=m+1}^{n} (f_{j})_{(i)}_{n+1} \cdot \langle f_{n+1}, h_{n_{n+1}} \rangle. \tag{2.24} \]

We can write \( \langle U_{m}(f_{1}, \ldots, f_{n}), f_{n+1} \rangle \) in the form

\[ \sum_{i=0}^{k-1} \sum_{k} \sum_{i_{1}=i_{n+1}=L} \sum_{j} b_{K,(I_{j})} |L^{(1)}|^{-(m-1)/2} |h_{L^{(1)}}| L |L|^{-(n-m)/2} \]

\[ \times \prod_{j=1}^{m-1} \langle f_{j}, h_{L^{(1)}}^{0} \rangle \cdot \langle f_{m}, h_{L^{(1)}} \rangle \cdot \prod_{j=m+1}^{n} \langle f_{j}, h_{L}^{0} \rangle \cdot \langle f_{n+1}, h_{n_{n+1}} \rangle. \]

Notice the normalization estimate

\[ \sum_{l_{1}^{(k)} = \cdots = l_{n}^{(k)} = K} |b_{K,(I_{j})}| \langle L^{(1)} \rangle |L^{(1)}|^{-(m-1)/2} |L^{(1)}|^{1/2} |L^{(1)}|^{-(n-m)/2} \]

\[ \lesssim \frac{|L^{(1)}| m/2 |L|^{(n-m)/2} |I_{n+1}|^{1/2}}{|L^{(1)}|^{n}}. \]

Therefore, for some constant \( C \lesssim 1 \) we get that

\[ \langle U_{m}(f_{1}, \ldots, f_{n}), f_{n+1} \rangle = C \sum_{i=0}^{k-1} \langle S_{0,\ldots,0,1,\ldots,1,i}(f_{1}, \ldots, f_{m}), f_{n+1} \rangle, \]

where there are \( m \) zeroes in \( S_{0,\ldots,0,1,\ldots,1,i} \). \( \square \)
The Optimized Representation Theorem

Let \( \sigma = (\sigma^i)_{i \in \mathbb{Z}} \), where \( \sigma^i \in \{0, 1\}^d \). Let \( \mathcal{D}_0 \) be the standard dyadic grid on \( \mathbb{R}^d \),

\[
\mathcal{D}_0 := \{ 2^{-k}([0, 1)^d + m) : k \in \mathbb{Z}, m \in \mathbb{Z}^d \}.
\]

We define the new dyadic grid

\[
\mathcal{D}_\sigma = \left\{ I + \sum_{i: 2^{-i} < \ell(I)} 2^{-i} \sigma^i : I \in \mathcal{D}_0 \right\} = \{ I + \sigma : I \in \mathcal{D}_0 \},
\]

where we simply have defined

\[
I + \sigma := I + \sum_{i: 2^{-i} < \ell(I)} 2^{-i} \sigma^i.
\]

It is straightforward that \( \mathcal{D}_\sigma \) inherits the key nestedness property of \( \mathcal{D}_0 \): if \( I, J \in \mathcal{D}_\sigma \), then \( I \cap J \in \{ I, J, \emptyset \} \). Moreover, there is a natural product probability measure \( \mathbb{P}_\sigma = \mathbb{P} \) on \( ((0, 1]^d)^\mathbb{Z} \) – this gives us the notion of random dyadic grids \( \sigma \mapsto \mathcal{D}_\sigma \) over which we take the expectation \( \mathbb{E}_\sigma \) below.

Remark 2.25 The assumption \( \omega \in \text{Dini}_{1/2} \) in the theorem below is only needed to have a converging series. The regularity is not explicitly used in the proof of the representation. It is required due to the estimates of the model operators briefly discussed above.

Theorem 2.26 Suppose that \( T \) is an \( n \)-linear \( \omega \)-CZO, where \( \omega \in \text{Dini}_{1/2} \). Then we have

\[
\langle T(f_1, \ldots, f_n), f_{n+1} \rangle = C_T \mathbb{E}_\sigma \sum_{k=0}^\infty \sum_{u=0}^{c_{d,n}} \omega(2^{-k}) \langle V_{k,u,\sigma}(f_1, \ldots, f_n), f_{n+1} \rangle,
\]

where \( V_{k,u,\sigma} \) is always either a standard \( n \)-linear shift \( S_{k,\ldots,k} \), a modified \( n \)-linear shift \( Q_k \) or an \( n \)-linear paraproduct (this requires \( k = 0 \)) in the grid \( \mathcal{D}_\sigma \). Moreover, we have

\[
|C_T| \lesssim \sum_{m=0}^n \|T_{m^*}(1, \ldots, 1)\|_{\text{BMO}} + \|T\|_{\text{WBP}} + C_K + 1.
\]
Proof. We begin with the decomposition
\[
\langle \mathcal{T}(f_1, \ldots, f_n), f_{n+1} \rangle = \mathbb{E}_\sigma \sum_{I_1, \ldots, I_{n+1}} \langle \Delta_{I_1} f_1, \ldots, \Delta_{I_n} f_n, \Delta_{I_{n+1}} f_{n+1} \rangle \\
= \sum_{j=1}^{n+1} \mathbb{E}_\sigma \sum_{I_1, \ldots, I_{n+1} \in \mathcal{D} \text{ for } \ell(I_i) > \ell(I_j) \text{ for } i \neq j} \langle \Delta_{I_1} f_1, \ldots, \Delta_{I_n} f_n, \Delta_{I_{n+1}} f_{n+1} \rangle \\
+ \mathbb{E}_\sigma R_\sigma,
\]
where \( I_1, \ldots, I_{n+1} \in \mathcal{D}_\sigma \) for some \( \sigma \in (\{0, 1\}^d)_{\mathbb{Z}} \). We deal with the remainder term \( R_\sigma \) later, and now focus on dealing with one of the main terms
\[
\Sigma_{j, \sigma} = \sum_{I_1, \ldots, I_{n+1} \in \mathcal{D} \text{ for } \ell(I_i) > \ell(I_j) \text{ for } i \neq j} \langle \Delta_{I_1} f_1, \ldots, \Delta_{I_n} f_n, \Delta_{I_{n+1}} f_{n+1} \rangle,
\]
where \( j \in \{1, \ldots, n+1\} \).

The main terms are symmetric, and we choose to handle \( \Sigma_\sigma := \Sigma_{n+1, \sigma} \). After collapsing the sums
\[
\sum_{I_i: \ell(I_i) > \ell(I_{n+1})} \Delta_{I_i} f_i = \sum_{I_i: \ell(I_i) = \ell(I_{n+1})} E_{I_i} f_i,
\]
we have
\[
\Sigma_\sigma = \sum_{\ell(I_1) = \cdots = \ell(I_{n+1})} \mathcal{T}(E_{I_1} f_1, \ldots, E_{I_n} f_n, \Delta_{I_{n+1}} f_{n+1}).
\]
Further, we write
\[
\langle T(E_{I_1} f_1, \ldots, E_{I_n} f_n), \Delta_{I_{n+1}} f_{n+1} \rangle \\
= \langle T(h^0_{I_1}, \ldots, h^0_{I_n}), h_{I_{n+1}} \rangle \prod_{j=1}^n \langle f_j, h^0_{I_j} \rangle \langle f_{n+1}, h_{I_{n+1}} \rangle \\
= \langle T(h^0_{I_1}, \ldots, h^0_{I_n}), h_{I_{n+1}} \rangle \left[ \prod_{j=1}^n \langle f_j, h^0_{I_j} \rangle - \prod_{j=1}^n \langle f_j, h^0_{I_{n+1}} \rangle \right] \langle f_{n+1}, h_{I_{n+1}} \rangle \\
+ \langle T(1_{I_1}, \ldots, 1_{I_n}), h_{I_{n+1}} \rangle \prod_{j=1}^n \langle f_j \rangle_{I_{n+1}} \langle f_{n+1}, h_{I_{n+1}} \rangle.
\]

We define the abbreviation
\[
\varphi_{I_1, \ldots, I_{n+1}} := \langle T(h^0_{I_1}, \ldots, h^0_{I_n}), h_{I_{n+1}} \rangle \\
\times \left[ \prod_{j=1}^n \langle f_j, h^0_{I_j} \rangle - \prod_{j=1}^n \langle f_j, h^0_{I_{n+1}} \rangle \right] \langle f_{n+1}, h_{I_{n+1}} \rangle.
\]
If we now sum over \( I_1, \ldots, I_{n+1} \) we may express \( \Sigma_\sigma \) in the form

\[
\Sigma_\sigma = \sum_{\ell(I_1)=\cdots=\ell(I_{n+1})} \varphi_{I_1,\ldots,I_{n+1}} + \sum_I \langle T(1, \ldots, 1), h_I \rangle \prod_{j=1}^n \langle f_j, f_{n+1}, h_I \rangle = \Sigma^1_\sigma + \Sigma^2_\sigma,
\]

where we recognize that the second term \( \Sigma^2_\sigma \) is a paraproduct. Thus, we only need to continue working with \( \Sigma^1_\sigma \).

Since \( \varphi_{I_1,\ldots,I_{n+1}} = 0 \), we have that

\[
\Sigma^1_\sigma = \sum_{m_1,\ldots,m_n \in \mathbb{Z}^d} \sum_{I} \varphi_{I+m_1 \ell(I),\ldots,I+m_n \ell(I),I} = \sum_{k=2}^{\infty} \sum_{\max |m_j| \in (2^{k-3},2^{k-2})} \sum_{I \in D_\sigma, \text{good}(k)} \varphi_{I+m_1 \ell(I),\ldots,I+m_n \ell(I),I}.
\]

As in [16] we say that \( I \) is \( k \)-good for \( k \geq 2 \) – and denote this by \( I \in D_\sigma, \text{good}(k) \) – if \( I \in D_\sigma \) satisfies

\[
d(I, \partial I^{(k)}) \geq \frac{\ell(I^{(k)})}{4} = 2^{k-2} \ell(I).
\]

Notice that for all \( I \in D_0 \) we have

\[
\mathbb{P}(\{ \sigma : I + \sigma \in D_\sigma, \text{good}(k) \}) = 2^{-d}.
\]

Thus, by the independence of the position of \( I \) and the \( k \)-goodness of \( I \) we have

\[
\mathbb{E}_\sigma \Sigma^1_\sigma = 2^d \mathbb{E}_\sigma \sum_{k=2}^{\infty} \sum_{\max |m_j| \in (2^{k-3},2^{k-2})} \sum_{I \in D_\sigma, \text{good}(k)} \varphi_{I+m_1 \ell(I),\ldots,I+m_n \ell(I),I} = C \omega(2^{-k}) \langle Q_k(f_1, \ldots, f_n), f_{n+1} \rangle,
\]

where

\[
\langle Q_k(f_1, \ldots, f_n), f_{n+1} \rangle := \frac{1}{C \omega(2^{-k})} \sum_{\max |m_j| \in (2^{k-3},2^{k-2})} \sum_{I \in D_\sigma, \text{good}(k)} \varphi_{I+m_1 \ell(I),\ldots,I+m_n \ell(I),I}
\]

and \( C \) is large enough.
Next, the key implication of the $k$-goodness is that

$$\left(I + m\ell(I)\right)^{(k)} = I^{(k)} =: K$$

(2.30)

if $|m| \leq 2^{k-2}$ and $I \in D_{\sigma, \text{good}}(k)$. Indeed, notice that e.g. $c_I + m\ell(I) \in [I + m\ell(I)] \cap K$ (so that $[I + m\ell(I)] \cap K \neq \emptyset$ which is enough) as

$$d(c_I + m\ell(I), K^C) \geq d(c_I, K^C) - |m|\ell(I) = d(I, \partial K) - 2^{k-2}\ell(I) - 2^{k-2}\ell(I) = 0.$$ 

Therefore, to conclude that $Q_k$ is a modified $n$-linear shift it only remains to prove the normalization

$$|\langle T(h^0_{I+m\ell(I)} \cdot \cdot \cdot , h^0_{I+m_n\ell(I)}), h_I \rangle|_{\omega(2^{-k})} \lesssim |I|^{(n+1)/2} |K|^n. \quad (2.31)$$

Suppose first that $k \sim 1$. Recall that $(m_1, \ldots, m_n) \neq (0, 0)$ and assume for example that $m_1 \neq 0$. We have using the size estimate of the kernel that

$$|\langle T(h^0_{I+m_1\ell(I)} \cdot \cdot \cdot , h^0_{I+m_n\ell(I)}), h_I \rangle|$$

$$\lesssim \int_{\mathbb{R}^{(n+1)d}} \left( \sum_{m=1}^{n} |x_{n+1} - x_m| \right)^{dn} \frac{dx}{|I|^{(n+1)/2}} \lesssim \frac{1}{|I|^{(n-1)/2}}, \quad (2.32)$$

where in the second step we repeatedly used estimates of the form

$$\int_{\mathbb{R}^d} \frac{dz}{(r + |z_0 - z|)^{d+\alpha}} \lesssim \frac{1}{r^\alpha}. \quad (2.33)$$

Notice that this is the right upper bound (2.31) in the case $k \sim 1$.

Suppose then that $k$ is large enough so that we can use the continuity assumption of the kernel. In this case we have that if $x_{n+1} \in I$ and $x_1 \in I + m_1\ell(I), \ldots, x_n \in I + m_n\ell(I)$...
\( I + m_n \ell(I) \), then \( \sum_{m=1}^n |x_{n+1} - x_m| \sim 2^k \ell(I) = \ell(K) \). Thus, there holds that

\[
\left| \langle T(h_{I+\ell(I)}^0, \ldots, h_{I+m_n \ell(I)}^0, h_1^0) \rangle \right|
\]

\[
= \left| \int_{\mathbb{R}^{(n+1)d}} (K(x_{n+1}, x_1, \ldots, x_n) - K(c_I, x_1, \ldots, x_n)) \times \prod_{j=1}^n h_{I+m_j \ell(I)}^0(x_j) h_I(x_{n+1}) \, dx \right|
\]

\[
\leq \int_{\mathbb{R}^{(n+1)d}} \omega(2^{-k}) \frac{1}{|K|^n} \prod_{j=1}^n h_{I+m_j \ell(I)}^0(x_j) |h_I(x_{n+1})| \, dx
\]

\[
= \omega(2^{-k}) \frac{1}{|K|^n} |I|^{n+1} |I|^{-(n+1)/2} = \omega(2^{-k}) \frac{|I|^{(n+1)/2}}{|K|^n}.
\]

We have proved (2.31). This ends our treatment of \( \mathbb{E}_\sigma \Sigma_\sigma \).

We now only need to deal with the remainder term \( \mathbb{E}_\sigma R_\sigma \). Write

\[ R_\sigma = \sum_{(I_1, \ldots, I_{n+1}) \in \mathcal{I}_\sigma} \langle T(\Delta_{I_1} f_1, \ldots, \Delta_{I_n} f_n), \Delta_{I_{n+1}} f_{n+1} \rangle, \]

where each \((I_1, \ldots, I_{n+1}) \in \mathcal{I}_\sigma\) satisfies that if \( j \in \{1, \ldots, n \} \) is such that \( \ell(I_j) \leq \ell(I_i) \) for all \( i \in \{1, \ldots, n \} \), then \( \ell(I_j) = \ell(I_{i_0}) \) for at least one \( i_0 \in \{1, \ldots, n \} \setminus \{j\} \). The point why the remainder is simpler than the main terms is that we can split this summation so that there are always at least two sums which we do not need to collapse—that means we will readily have two cancellative Haar functions. To give the idea, it makes sense to explain the bilinear case \( n = 2 \). In this case we can, in a natural way, decompose

\[
\sum_{(I_1, I_2, I_3) \in \mathcal{I}_\sigma} = \sum_{\ell(I_2) = \ell(I_3) < \ell(I_1)} + \sum_{\ell(I_1) = \ell(I_2) < \ell(I_3)} + \sum_{\ell(I_1) = \ell(I_2) < \ell(I_3)} + \sum_{\ell(I_1) = \ell(I_2) = \ell(I_3)},
\]

which – after collapsing the relevant sums – gives that

\[
R_\sigma = \sum_{\ell(I_1) = \ell(I_2) = \ell(I_3)} \left( \langle T(E_{I_1} f_1, \Delta_{I_2} f_2), \Delta_{I_3} f_3 \rangle + \langle T(\Delta_{I_1} f_1, E_{I_2} f_2), \Delta_{I_3} f_3 \rangle \right) + \langle T(\Delta_{I_1} f_1, \Delta_{I_2} f_2), E_{I_3} f_3 \rangle + \langle T(\Delta_{I_1} f_1, \Delta_{I_2} f_2), \Delta_{I_3} f_3 \rangle = 4 \sum_{i=1}^4 R_{\sigma_i}.
\]
These are all handled similarly (the point is that there are at least two martingale differences remaining in all of them) so we look for example at

\[
R_\sigma^2 = \sum_{\ell(I_1) = \ell(I_2) = \ell(I_3)} \langle T(\Delta I_1 f_1, E I_2 f_2), \Delta I_3 f_3 \rangle
\]

\[
= \sum_{\ell(I_1) = \ell(I_2) = \ell(I_3)} \langle T(\Delta I_1 f_1, E I_2 f_2), \Delta I_3 f_3 \rangle + \sum_I \langle T(\Delta I_1 f_1, E I f_2), \Delta I f_3 \rangle.
\]

We can represent these terms as sums of standard bilinear shifts of the form \(S_{k,k,k}\). The first term is handled exactly like \(\Sigma_1\) above. The second term is readily a zero complexity shift. To prove the estimate for the coefficient we write \(\langle T(\Delta I_1 f_1, E I_2 f_2), \Delta I_3 f_3 \rangle\) as the sum of

\[
\sum_{I_1, I_2, I_3 \subset I} \langle T(1_{I_1} h_{I_1}, 1_{I_2} h_{I_2}^0), 1_{I_3} h_{I_1} \rangle \tag{2.35}
\]

and

\[
\sum_{I' \subset I} \langle T(1_{I'} h_{I'}, 1_{I'} h_{I'}^0), 1_{I'} h_{I} \rangle. \tag{2.36}
\]

There are \(\lesssim 1\) terms in both sums. In (2.35) we use the size of the kernel of \(T\) and in (2.36) the weak boundedness property \(|\langle T(1_{I_1}, 1_I)\rangle| \lesssim |I|\).

The general \(n\)-linear remainder term \(R_\sigma\) is analogous and only yields standard \(n\)-linear shifts \(S_{k,...,k}\). We are done. \(\square\)

3 UMD-Valued Extensions of Singular Integrals

Preliminaries of Banach Space Theory

An extensive treatment of Banach space theory is given in the books \([20, 21]\) by Hytönen, van Neerven, Veraar and Weis.

We say that \(\{\varepsilon_i\}\) is a collection of independent random signs, where \(i\) runs over some index set, if there exists a probability space \((\mathcal{M}, \mu)\) so that \(\varepsilon_i: \mathcal{M} \to \{-1, 1\}\), \(\{\varepsilon_i\}\) is independent and \(\mu(\{\varepsilon_i = 1\}) = \mu(\{\varepsilon_i = -1\}) = 1/2\). Below, \(\{\varepsilon_i\}\) will always denote a collection of independent random signs.

Suppose \(X\) is a Banach space. We denote the underlying norm by \(|\cdot|_X\). The Kahane-Khintchine inequality says that for all \(x_1, \ldots, x_N \in X\) and \(p, q \in (0, \infty)\) there holds

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

The general \(n\)-linear remainder term \(R_\sigma\) is analogous and only yields standard \(n\)-linear shifts \(S_{k,...,k}\). We are done. \(\square\)
that
\[
\left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i x_i \right\|_X^p \right)^{1/p} \sim \left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i x_i \right\|_X^q \right)^{1/q}.
\] (3.1)

Definitions related to Banach spaces often involve such random sums and the definition may involve some fixed choice of the exponent—but the choice is irrelevant by the Kahane–Khintchine inequality.

The Kahane contraction principle says that if \((a_i)_{i=1}^N\) is a sequence of scalars, \(x_1, \ldots, x_N \in X\) and \(p \in (0, \infty]\), then
\[
\left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i a_i x_i \right\|_X^p \right)^{1/p} \lesssim \max |a_i| \left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i x_i \right\|_X^p \right)^{1/p}.
\]

**Definition 3.2** Let \(X\) be a Banach space, let \(r \in [1, 2]\) and \(q \in [2, \infty]\).

1. The space \(X\) has type \(r\) if there exists a finite constant \(\tau \geq 0\) such that for all finite sequences \(x_1, \ldots, x_N \in X\) we have
\[
\left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i x_i \right\|_X^r \right)^{1/r} \leq \tau \left( \sum_{i=1}^{N} |x_i|_X^r \right)^{1/r}.
\]

2. The space \(X\) has cotype \(q\) if there exists a finite constant \(c \geq 0\) such that for all finite sequences \(x_1, \ldots, x_N \in X\) we have
\[
\left( \sum_{i=1}^{N} |x_i|_X^q \right)^{1/q} \leq c \left( \mathbb{E} \left\| \sum_{i=1}^{N} \varepsilon_i x_i \right\|_X^q \right)^{1/q}.
\]

For \(q = \infty\) the usual modification is used.

The least admissible constants are denoted by \(\tau_{r,X}\) and \(c_{q,X}\)—they are the type \(r\) constant and cotype \(q\) constant of \(X\).

In [21, Sect. 7] the reader can find the basic theory of types and cotypes. We only need a few basic facts, however.

If \(X\) has type \(r\) (cotype \(q\)), then it also has type \(u\) for all \(u \in [1, r]\) (cotype \(v\) for all \(v \in [q, \infty]\)), and we have \(\tau_{u,X} \leq \tau_{r,X}\) \((c_{v,X} \leq c_{q,X})\). It is also trivial that always \(\tau_{1,X} = c_{\infty,X} = 1\). We say that \(X\) has non-trivial type if \(X\) has type \(r\) for some \(r \in (1, 2]\) and finite cotype if it has cotype \(q\) for some \(q \in [2, \infty)\).

For the types and cotypes of \(L^p\) spaces we have the following: if \(X\) has type \(r\), then \(L^p(X)\) has type \(\min(r, p)\), and if \(X\) has cotype \(q\), then \(L^p(X)\) has cotype \(\max(q, p)\).

The UMD property is a necessary and sufficient condition for the boundedness of various singular integral operators on \(L^p(\mathbb{R}^d, X) = L^p(X)\), see [20, Sect. 5.2.c and the Notes to Sect. 5.2].
Definition 3.3 A Banach space $X$ is said to be a UMD space, where UMD stands for unconditional martingale differences, if for all $p \in (1, \infty)$, all $X$-valued $L^p$-martingale difference sequences $(d_i)_{i=1}^N$ and all choices of fixed signs $\epsilon_i \in \{-1, 1\}$ we have

$$\left\| \sum_{i=1}^N \epsilon_i d_i \right\|_{L^p(X)} \lesssim \left\| \sum_{i=1}^N d_i \right\|_{L^p(X)}.$$  \hspace{1cm} (3.4)

The $L^p(X)$-norm is with respect to the measure space where the martingale differences are defined.

A standard property of UMD spaces is that if (3.4) holds for one $p_0 \in (1, \infty)$ it holds for all $p \in (1, \infty)$ [20, Theorem 4.2.7]. Moreover, if $X$ is UMD then so is the dual space $X^*$ [20, Proposition 4.2.17]. Importantly, UMD spaces have non-trivial type and a finite cotype.

Stein’s inequality says that for a UMD space $X$ we have

$$E \left\| \sum_{I \in \mathcal{D}} \epsilon_I \langle f_I \rangle I \right\|_{L^p(X)} \lesssim E \left\| \sum_{I \in \mathcal{D}} \epsilon_I f_I \right\|_{L^p(X)}, \quad p \in (1, \infty).$$

This UMD-valued version of Stein’s inequality is by Bourgain, for a proof see e.g. Theorem 4.2.23 in the book [20].

We now introduce some definitions related to the so called decoupling estimate. For $K \in \mathcal{D}$ denote by $Y_K$ the measure space $(K, \text{Leb}(K), \nu_K)$. Here $\text{Leb}(K)$ is the collection of Lebesgue measurable subsets of $K$ and $\nu_K = dx|K/|K|$, where $dx|K$ is the $d$-dimensional Lebesgue measure restricted to $K$. We then define the product probability space

$$(Y, \mathcal{A}, \nu) := \prod_{K \in \mathcal{D}} Y_K.$$

If $y \in Y$ and $K \in \mathcal{D}$, we denote by $y_K$ the coordinate related to $Y_K$.

In our upcoming estimates, it will be important to separate scales using the following subgrids. For $k \in \{0, 1, \ldots \}$ and $l \in \{0, \ldots, k\}$ define

$$\mathcal{D}_{k,l} := \{ K \in \mathcal{D} : \ell(K) = 2^{m(k+1)+l} \text{ for some } m \in \mathbb{Z} \}. \hspace{1cm} (3.5)$$

The following proposition concerning decoupling is a special case of Theorem 3.1 in [17]. It is a result that can be stated in the generality of suitable filtrations, but we prefer to only state the following dyadic version.

Proposition 3.6 Let $X$ be a UMD space, $p \in (1, \infty)$, $k \in \{0, 1, \ldots \}$ and $l \in \{0, \ldots, k\}$. Suppose $f_K$, $K \in \mathcal{D}_{k,l}$, are functions such that

1. $f_K = 1_K f_K$,
2. $\int f_K = 0$ and
(3) $f_K$ is constant on those $K' \in D_{k,l}$ for which $K' \subsetneq K$.

Then we have

$$
\left| \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} f_K(x) \right|^p \ dx \sim \mathbb{E} \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K f_K(y) \right|^p \ dx \ dv(y),
$$

(3.7)

where the implicit constant is independent of $k, l$.

Hilbert spaces are the only Banach spaces with both type 2 and cotype 2. Below we will prove estimates for modified shifts like

$$
\| Q_k f \|_{L^2(X)} \lesssim (k + 1)^{1/\min(r,q')} \| f \|_{L^2(X)},
$$

where the UMD space $X$ has type $r$ and cotype $q$. Therefore, in the Hilbert space case—and thus in the scalar-valued case—these estimates recover the best possible regularity $\alpha = 1/2$. The presented estimates are efficient in completely general UMD spaces—however, in UMD function lattices it is more efficient to mimic the scalar-valued theory (Proposition 2.13) and use square function estimates instead.

The Linear Case

We feel that it is too difficult to jump directly into the multilinear estimates, as they are quite involved. Thus, we first study the linear case. We show that the framework of modified dyadic shifts gives a modern and convenient proof of the results of Figiel [11, 12] concerning UMD-extensions of CZOs with mild kernel regularity.

Before moving to the main $X$-valued estimate for $Q_k$, we state the following result for paraproducts. We have

$$
\| \pi f \|_{L^p(X)} \lesssim \| f \|_{L^p(X)},
$$

(3.8)

whenever $p \in (1, \infty)$ and $X$ is UMD. We understand that this is usually attributed to Bourgain— in any case, a simple proof can now be found in [17].

Remark 3.9 The estimate in Proposition 3.10 below is best used for $p = 2$, since then e.g. $\min(r, p) = r$, if $r \in (1, 2]$ is an exponent such that $X$ has type $r$. Indeed, it is efficient to only move the $p = 2$ estimate for the CZO $T$, and then interpolate to get the $L^p$ boundedness under a modified Dini type assumption that is independent of $p$. On the $Q_k$ level improving an $L^2$ estimate into an $L^p$ estimate with good dependency on the complexity does not seem so simple. Interpolation would introduce some additional complexity dependency, since the weak $(1, 1)$ inequality of $Q_k$ is not complexity free.

Proposition 3.10 Let $p \in (1, \infty)$ and $X$ be a UMD space. If $Q_k$ is a modified shift of the form (2.10), then

$$
\| Q_k f \|_{L^p(X)} \lesssim (k + 1)^{1/\min(r,p)} \| f \|_{L^p(X)},
$$

Birkhäuser
where \( r \in (1, 2] \) is an exponent such that \( X \) has type \( r \). If \( Q_k \) is a modified shift of the form (2.9), we have

\[
\| Q_k f \|_{L^p(X)} \lesssim (k + 1)^{1/\min(q', p')} \| f \|_{L^p(X)},
\]

where \( q \in [2, \infty) \) is an exponent such that \( X \) has cotype \( q \).

**Proof** We assume that \( Q_k \) has the form (2.10)—the other result follows by duality. This uses that if the UMD space \( X \) has cotype \( q \), then the dual space \( X^* \) has type \( q' \)—see [21, Proposition 7.4.10].

If \( K \in D \) we define

\[
B_K f := \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle f, H_{I,J} \rangle h_J.
\]

Recall the lattices \( D_{k,l} \) from (3.5) and write \( Q_k f = \sum_{l=0}^{k} \sum_{K \in D_{k,l}} B_K f \). By using the UMD property of \( X \) and the Kahane–Khintchine inequality we have for all \( s \in (0, \infty) \) that

\[
\| Q_k f \|_{L^p(X)} \sim \left( \mathbb{E} \left\| \sum_{l=0}^{k} \varepsilon_l \sum_{K \in D_{k,l}} B_K f \right\|_{L^p(X)}^s \right)^{1/s}.
\]

We use this with the choice \( s := \min(r, p) \), since \( L^p(X) \) has type \( s \). Using this we have

\[
\left( \mathbb{E} \left\| \sum_{l=0}^{k} \varepsilon_l \sum_{K \in D_{k,l}} B_K f \right\|_{L^p(X)}^s \right)^{1/s} \lesssim \left( \sum_{l=0}^{k} \left\| \sum_{K \in D_{k,l}} B_K f \right\|_{L^p(X)}^s \right)^{1/s}.
\]

To end the proof, it remains to show that

\[
\left\| \sum_{K \in D_{k,l}} B_K f \right\|_{L^p(X)} \lesssim \| f \|_{L^p(X)}
\]

uniformly on \( l \).

We have that \( \langle f, H_{I,J} \rangle = \langle P_{K,k} f, H_{I,J} \rangle \) and so

\[
B_K f = \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle P_{K,k} f, H_{I,J} \rangle h_J = \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle P_{K,k} f, 1_J H_{I,J} \rangle h_J
\]

\[
+ \sum_{I^{(k)} = J^{(k)} = K} a_{IJK} \langle P_{K,k} f, 1_I H_{I,J} \rangle h_J.
\]
Accordingly, this splits the estimate of (3.11) into two parts.

We consider first the part related to \( \langle P_{K,k} f, 1_I H_{I,J} \rangle \). By the UMD property and the Kahane–Khintchine inequality we have

\[
\mathbb{E} \left\| \sum_{K \in D_{k,l}} \sum_{I \neq J} a_{IJK} \langle P_{K,k} f, 1_I H_{I,J} \rangle h_J \right\|_{L^p(X)}^p \sim \left( \mathbb{E} \left\| \sum_{K \in D_{k,l}} \sum_{I \neq J} a_{IJK} \langle P_{K,k} f, 1_I H_{I,J} \rangle h_J \right\|_{L^p(X)}^p \right)^{1/p}.
\]

(3.12)

Notice then that for

\[
a_K(x,y) := |K| \sum_{I \neq J} a_{IJK} 1_J(y) H_{I,J}(y) h_J(x)
\]

we have

\[
\sum_{I \neq J} a_{IJK} \langle P_{K,k} f, 1_I H_{I,J} \rangle h_J(x) = \frac{1}{|K|} \int K a_K(x,y) P_{K,k} f(y) \, dy = \int YK a_K(x,yK) P_{K,k} f(yK) \, dv_K(yK) = \int Y a_K(x,yK) P_{K,k} f(yK) \, dv(y).
\]

Using this we have by Hölder’s inequality (recalling that \( v \) is a probability measure) that

\[
\mathbb{E} \left\| \sum_{K \in D_{k,l}} \sum_{I \neq J} a_{IJK} \langle P_{K,k} f, 1_I H_{I,J} \rangle h_J \right\|_{L^p(X)}^p \leq \mathbb{E} \int \int \left\| \sum_{K \in D_{k,l}} \sum_{I \neq J} a_K(x,yK) P_{K,k} f(yK) \right\|_X^p \, dv(y) \, dx.
\]

Notice now that \( |a_K(x,y)| \leq 1_K(x) \). Thus, the Kahane contraction principle implies that for fixed \( x \) and \( y \) there holds that

\[
\mathbb{E} \left\| \sum_{K \in D_{k,l}} \sum_{I \neq J} \epsilon_K a_K(x,yK) P_{K,k} f(yK) \right\|_X^p \leq \mathbb{E} \sum_{K \in D_{k,l}} \epsilon_K 1_K(x) P_{K,k} f(yK) \right\|_X^p.
\]
Using this we are left with

\[ \mathbb{E} \int_{\mathbb{R}^d} \int_{\mathcal{D}_{k,l}} \sum_{K \in \mathcal{D}_{k,l}} \| P_{K,k,f} \|_X^p \, d\nu(y) \, dx \sim \int_{\mathbb{R}^d} \sum_{K \in \mathcal{D}_{k,l}} \| P_{K,k,f} \|_X^p \, dx = \| f \|_{L^p(X)}, \]

where we used the decoupling estimate (3.7) and noticed that \( \sum_{K \in \mathcal{D}_{k,l}} P_{K,k,f} = f \).

We turn to the part related to the terms \( \langle P_{K,k,f}, 1_J H_{I,J} \rangle \). We begin with

\[ \| \sum_{K \in \mathcal{D}_{k,l}} \sum_{I(k) = J(k) = K} a_{IJK} \langle P_{K,k,f}, 1_J H_{I,J} \rangle h_J \|_{L^p(X)}^p \sim \mathbb{E} \| \sum_{K \in \mathcal{D}_{k,l}} \sum_{I(k) = J(k) = K} \| a_{IJK} \langle P_{K,k,f}, 1_J H_{I,J} \rangle 1_J \|_{L^p(X)}^p, \tag{3.13} \]

where we used the UMD property to introduce the random signs and then the fact that we can clearly replace \( h_J \) by \( |h_J| = 1_J/|J|^{1/2} \) due to the random signs (the relevant random variables are identically distributed). We write the inner sum as

\[ \sum_{I(k) = J(k) = K} \| a_{IJK} \langle P_{K,k,f}, 1_J H_{I,J} \rangle 1_J \|_{L^p(X)}^p, \]

Notice also that

\[ \left| \sum_{I(k) = K} a_{IJK} |J|^{1/2} H_{I,J} \right| \leq \frac{1}{|K|} \sum_{I(k) = K} |I| = 1. \]

We continue from (3.13). Applying Stein’s inequality we get that the last term in (3.13) is dominated by

\[ \mathbb{E} \| \sum_{K \in \mathcal{D}_{k,l}} \sum_{I(k) = K} \| a_{IJK} \langle P_{K,k,f}, 1_J H_{I,J} \rangle 1_J \|_{L^p(X)}^p \leq \mathbb{E} \| \sum_{K \in \mathcal{D}_{k,l}} \sum_{I(k) = K} \| a_{IJK} \langle P_{K,k,f}, 1_J \rangle 1_J \|_{L^p(X)}^p, \]

\( \text{Birkhäuser} \)
where we used the Kahane contraction principle. From here the estimate is easily concluded by

\[
\mathbb{E}\left\| \sum_{K \in D, l} \sum_{J^{(k)}=K} \varepsilon J \mathcal{P}_{K,k} f 1_{J} \right\|_{L^p(X)}^p = \mathbb{E}\left\| \sum_{K \in D, l} \sum_{J^{(k)}=K} \varepsilon K \mathcal{P}_{K,k} f 1_{J} \right\|_{L^p(X)}^p
= \mathbb{E}\left\| \sum_{K \in D, l} \varepsilon K \mathcal{P}_{K,k} f \right\|_{L^p(X)}^p \sim \| f \|_{L^p(X)}^p.
\]

Here we first changed the indexing of the random signs (using that for a fixed \( x \) for every \( K \) there is at most one \( J \) as in the sum for which \( 1_{J}(x) \neq 0 \)) and then applied the UMD property. This finishes the proof.

**Theorem 3.14** Let \( T \) be a linear \( \omega \)-CZO and \( X \) be a UMD space with type \( r \in (1, 2] \) and cotype \( q \in [2, \infty) \). If \( \omega \in \text{Dini}_{1/\min(r,q')} \), we have

\[
\| Tf \|_{L^p(X)} \lesssim \| f \|_{L^p(X)}
\]

for all \( p \in (1, \infty) \).

**Proof** Apply Theorem 2.26 to simple vector-valued functions. By Proposition 3.10 and the \( X \)-valued boundedness of the paraproducts (3.8) we conclude that

\[
\| Tf \|_{L^2(X)} \lesssim \| f \|_{L^2(X)}.
\]

As the weak type \((1,1)\) follows from this even with just the assumption \( \omega \in \text{Dini}_0 \), we can conclude the proof by the standard interpolation and duality method.

**The Multilinear Case**

Let \( (X_1, \ldots, X_{n+1}) \) be UMD spaces and \( Y^+_{n+1} = X_{n+1} \). Assume that there is an \( n \)-linear mapping \( X_1 \times \cdots \times X_n \to Y_{n+1} \), which we denote with the product notation \( (x_1, \ldots, x_n) \mapsto \prod_{j=1}^n x_j \), so that

\[
\left| \prod_{j=1}^n x_j \right|_{Y_{n+1}} \leq \prod_{j=1}^n |x_j|_{X_j}.
\]

With just this setup it makes sense to extend an \( n \)-linear SIO (or some other suitable \( n \)-linear operator) \( T \) using the formula

\[
T(f_1, \ldots, f_n)(x) = \sum_{j_1, \ldots, j_n} T(f_{1,j_1}, \ldots, f_{n,j_n})(x) \prod_{k=1}^n e_{k,j_k}, \quad x \in \mathbb{R}^d,
\]

\[
f_k = \sum_{j_k} e_{k,j_k} f_{k,j_k}, \quad f_{k,j_k} \in L^\infty_c, \quad e_{k,j_k} \in X_k.
\]

(3.15)
In the bilinear case \( n = 2 \) the existence of such a product is the only assumption that we will need. The bilinear case is somewhat harder than the linear case, but the \( n \geq 3 \) case is by far the most subtle. Indeed, for \( n \geq 3 \) we will need a more complicated setting for the tuple of spaces \((X_1, \ldots, X_{n+1})\)—the idea is to model the Hölder type structure typical of concrete examples of Banach \( n \)-tuples, such as that of non-commutative \( L^p \) spaces with the exponents \( p \) satisfying the natural Hölder relation. We will borrow this setting from [9]. In [9] it is shown in detail how natural tuples of non-commutative \( L^p \) spaces fit to this abstract framework. While we borrow the setting, the proof is significantly different. First, we have to deal with the more complicated modified shifts. Second, even in the standard shift case the proof in [9] is—by its very design—extremely costly on its complexity dependency. To circumvent this we need a new strategy.

For \( m \in \mathbb{N} \) we write \( J_m := \{1, \ldots, m\} \) and denote the set of permutations of \( J \subset J_m \) by \( \Sigma(J) \). We write \( \Sigma(m) = \Sigma(J_m) \).

Next, we fix an associative algebra \( A \) over \( \mathbb{C} \), and denote the associative operation \( A \times A \to A \) by \( (e, f) \mapsto ef \). We assume that there exists a subspace \( L^1 \) of \( A \) and a linear functional \( \tau : L^1 \to \mathbb{C} \), which we refer to as \textit{trace}. Given an \( m \)-tuple \((X_1, \ldots, X_m)\) of Banach subspaces \( X_j \subset A \), we construct the seminorm

\[
|e|_{Y(X_1,\ldots,X_m)} = \sup \left\{ \left| \tau \left( e \prod_{\ell=1}^{m} e_{\sigma(\ell)} \right) \right| : \sigma \in \Sigma(m), |e_j|_{X_j} = 1, j = 1, \ldots, m \right\}
\]

on the subspace

\[
Y(X_1,\ldots,X_m) = \{ e \in A : e \prod_{\ell=1}^{m} e_{\sigma(\ell)} \in L^1 \forall \sigma \in \Sigma(m), e_j \in X_j, j = 1, \ldots, m \}.
\] (3.17)

For a Banach subspace \( X \subset A \) and \( y \in Y(X) \) we can define the mapping \( \Lambda_y \in X^* \) by the formula \( \Lambda_y(x) := \tau(yx) \), since by the definition \( yx \in L^1 \) and \( |\tau(yx)| \leq |\tau(y)|_{Y(X)} |x|_X \). We say that a Banach subspace \( X \subset A \) is admissible if the following holds.

1. \( Y(X) \) is a Banach space with respect to \( |\cdot|_{Y(X)} \).
2. The mapping \( y \mapsto \Lambda_y \) from \( Y(X) \) to \( X^* \) is surjective.
3. For each \( x \in X \), \( y \in Y(X) \), we also have \( xy \in L^1 \) and

\[
\tau(xy) = \tau(yx).
\] (3.18)

If \( X \) is admissible, then the map \( y \mapsto \Lambda_y \) is an isometric bijection from \( Y(X) \) onto \( X^* \), and we identify \( Y(X) \) with \( X^* \). The following is [9, Lemma 3.10].

**Lemma 3.19** Let \( X \) be admissible and reflexive (for instance, \( X \) is admissible and UMD). If \( Y(X) \) is also admissible, then \( Y(Y(X)) = X \) as sets and \( |x|_{Y(Y(X))} = |x|_X \) for all \( x \in X \).

\( \text{Birkhäuser} \)
If $X, X_1, \ldots, X_m$ are Banach spaces we write $X = Y(X_1, \ldots, X_m)$ to mean that $X$ and $Y(X_1, \ldots, X_m)$ coincide as sets, $Y(X_1, \ldots, X_m)$ is a Banach space with the norm $\| \cdot \|_{Y(X_1,\ldots,X_m)}$, and that the norms are equivalent, that is, $|x|_X \sim |x|_{Y(X_1,\ldots,X_m)}$ for all $x \in X$.

**Definition 3.20** [UMD Hölder pair] Let $X_1, X_2$ be admissible spaces. We say that \{X_1, X_2\} is a UMD Hölder pair if $X_1$ is a UMD space and $X_2 = Y(X_1)$.

**Definition 3.21** [UMD Hölder $m$-tuple, $m \geq 3$] Let $X_1, \ldots, X_m$ be admissible spaces. We say that \{X_1, \ldots, X_m\} is a UMD Hölder $m$-tuple if the following properties hold.

**P1.** For all $j_0 \in J_m$ there holds

$$X_{j_0} = \left( \{ X_j : j \in J_m \setminus \{j_0\} \} \right).$$

**P2.** If $1 \leq k \leq m - 2$ and $J = \{ j_1 < j_2 < \cdots < j_k \} \subset J_m$, then $Y(X_{j_1}, \ldots, X_{j_k})$ is an admissible Banach space with the norm (3.16) and

$$\{X_{j_1}, \ldots, X_{j_k}, Y(X_{j_1}, \ldots, X_{j_k})\} \quad (3.22)$$

is a UMD Hölder $(k + 1)$-tuple.

The following is a key consequence of the definition. Let $m \geq 3$ and \{X_1, \ldots, X_m\} be a UMD Hölder $m$-tuple. Then according to P2 the pair \{X_{j_0}, Y(X_{j_0})\} is a UMD Hölder pair, which by Definition 3.20 implies that $X_{j_0}$ and $Y(X_{j_0})$ are UMD spaces. The inductive nature of the definition then ensures that each $Y(X_{j_1}, \ldots, X_{j_k})$ appearing in (3.22) is a UMD space.

Notice also the following. Let $m \geq 2$ and \{X_1, \ldots, X_m\} be a UMD $m$-Hölder tuple. Let $e_j \in X_j$ for $j \in J_m$. For each $\sigma \in \Sigma(m)$, as $X_{\sigma(1)} = Y(X_{\sigma(2)}, \ldots, X_{\sigma(m)})$, we necessarily have $\prod_{j=1}^{m} e_{\sigma(j)} \in L^1$ and

$$|\tau(e_{\sigma(1)} \cdots e_{\sigma(m)})| \leq |e_{\sigma(1)}|_{Y(X_{\sigma(2)},\ldots,X_{\sigma(m)})} \prod_{j=2}^{m} |e_{\sigma(j)}|_{X_{\sigma(j)}} \sim \prod_{j=1}^{m} |e_j|_{X_j}. \quad (3.18)$$

Moreover, by (3.18) we have

$$\tau(e_1 \cdots e_m) = \tau(e_m e_1 \cdots e_{m-1}) = \tau(e_{m-1} e_m e_1 \cdots e_{m-2}) = \cdots. \quad (3.23)$$

We have the following Hölder type inequality.

**Lemma 3.24** Let \{X_1, \ldots, X_m\} be a UMD Hölder tuple. Then we have

$$|e v|_{Y(X_m)} \lesssim |e|_{Y(\bar{Y}(X_1,\ldots,X_k))} |v|_{Y(\bar{Y}(X_{k+1},\ldots,X_{m-1}))}. \quad (3.22)$$

**Proof** To estimate $|e v|_{Y(X_m)}$ we need to estimate

$$|\tau(e v e_m)|$$
with an arbitrary $e_m$ with $|e_m|_{X_m} = 1$. First, we bound this with

$$|e|_{Y(Y(X_1,\ldots,X_k))} |ve_m|_{Y(X_1,\ldots,X_k)}.$$  

To estimate $|ve_m|_{Y(X_1,\ldots,X_k)}$ we need to estimate

$$|\tau(ve_m u_{\sigma_1(1)} \cdots u_{\sigma_1(k)})|$$

with arbitrary $|u_j|_{X_j} = 1$ and $\sigma_1 \in \Sigma(k)$. We estimate this with

$$|v|_{Y(Y(X_{k+1},\ldots,X_{m-1}))} |e_m u_{\sigma_1(1)} \cdots u_{\sigma_1(k)}|_{Y(X_{k+1},\ldots,X_{m-1})}.$$  

To estimate $|e_m u_{\sigma_1(1)} \cdots u_{\sigma_1(k)}|_{Y(X_{k+1},\ldots,X_{m-1})}$ we need to estimate

$$|\tau(e_m u_{\sigma_1(1)} \cdots u_{\sigma_1(k)} u_{\sigma_2(k+1)} \cdots u_{\sigma_2(m-1)})|$$  

with an arbitrary permutation $\sigma_2$ of $\{k+1, \ldots, m-1\}$ and $|u_j|_{X_j} = 1$. Finally, we estimate this with

$$|e_m|_{Y(X_1,\ldots,X_{m-1})} \sim |e_m|_{X_m} = 1.$$  

\[\square\]

In all of the statements below an arbitrary UMD Hölder tuple $\{X_1, \ldots, X_{n+1}\}$ is given.

**Proposition 3.25** Suppose that $Q_k$ is an $n$-linear modified shift and $f_j : \mathbb{R}^d \to X_j$. Let $1 < p_j < \infty$ with $\sum_{j=1}^{n+1} 1/p_j = 1$. Then we have

$$|\langle Q_k(f_1, \ldots, f_n), f_{n+1} \rangle| \lesssim (k+1)^{\alpha} \prod_{j=1}^{n+1} \|f_j\|_{L^{p_j}(X_j)},$$

where

$$\alpha = \frac{1}{\min(p_1', \ldots, p_{n+1}', s_1', \ldots, s_{n+1}')},$$

and $X_j$ has cotype $s_j$.

**Proof** We will assume that $Q_k$ is of the form (2.6)—the other cases follow by duality using the property (3.23). We follow the ideas of the decomposition from the proof of Lemma 2.18: we will estimate the terms $A_m$ and $U_m$, $m \in \{1, \ldots, n\}$, from there separately.

First, we estimate the part $A_m$ defined in (2.21). We have that

$$A_m(f_1, \ldots, f_n) = \sum_K A_{m,K}(f_1, \ldots, f_n),$$

\[\bowtie\] Birkhäuser
where

\[
A_{m,K}(f_1, \ldots, f_n) := \sum_{I_m^{(k)} = \ldots = I_{n+1}^{(k)} = K} b_{m,K,(I_j)} \prod_{j=1}^{m-1} \langle f_j \rangle_K \cdot \langle P_{K,k-1} f_m \rangle_{I_m} \cdot \\
\prod_{j=m+1}^{n} \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}},
\]

\[
b_{m,K,(I_j)} = b_{m,K,I_m, \ldots, I_{n+1}} := \sum_{I_1^{(k)} = \ldots = I_{m-1}^{(k)} = K} b_{K,(I_j)}
\]

and

\[
b_{K,(I_j)} = b_{K,I_1, \ldots, I_{n+1}} = |I_{n+1}|^{n/2} a_{K,(I_j)}.
\]

Here we have the normalization

\[
|b_{m,K,(I_j)}| \leq \frac{|I_{n+1}|^{n+1/2}}{|K|^n} \sum_{I_1^{(k)} = \ldots = I_{m-1}^{(k)} = K} 1 = \frac{|I_{n+1}|^{n-m+3/2}}{|K|^{n-m+1}}.
\]

Recall the grids \(D_{k,l}, l \in \{0, 1, \ldots, k\}\), from (3.5). Let

\[
Z_{1,\ldots,n} := Y(X_{n+1}) = Y(Y(X_1, \ldots, X_n)).
\]

Recalling that \(X_{n+1}^*\) is identified with \(Y(X_{n+1})\), we have that \(L_{p_{n+1}}'(Z_{1,\ldots,n})\) has type \(s := \min(p_{n+1}', s_{n+1}')\). Thus, there holds that

\[
\|A_m(f_1, \ldots, f_n)\|_{L_{p_{n+1}}'(Z_{1,\ldots,n})} \sim \left(\sum_{l=0}^{k} \mathbb{E} \sum_{K \in D_{k,l}} A_{m,K}(f_1, \ldots, f_n) \right)^{s} \left(\sum_{K \in D_{k,l}} A_{m,K}(f_1, \ldots, f_n) \right)^{1/s}
\]

\[
\lesssim \left(\sum_{l=0}^{k} \sum_{K \in D_{k,l}} A_{m,K}(f_1, \ldots, f_n) \right)^{s} \left(\sum_{K \in D_{k,l}} A_{m,K}(f_1, \ldots, f_n) \right)^{1/s}.
\]

In the first step we used the UMD property of \(Z_{1,\ldots,n}\) and the Kahane–Khintchine inequality. We see that to prove the claim it suffices to show the uniform bound

\[
\| \sum_{K \in D_{k,l}} A_{m,K}(f_1, \ldots, f_n) \|_{L_{p_{n+1}}'(Z_{1,\ldots,n})} \lesssim \prod_{j=1}^{n} \| f_j \|_{L_{p_j}(X_j)}.
\]

We turn to prove (3.28). To avoid confusion with the various \(Y\) spaces, we denote the decoupling space by \((W, \nu)\) in this proof. The decoupling estimate (3.7) gives that
the left hand side of (3.28) is comparable to
\[
\left( \mathbb{E} \int_W \int_{\mathbb{R}^d} \left| \sum_{K \in \mathcal{D}_{k,l}} e_K 1_K(x) \sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} b_{m,K,(I_j)} \prod_{j=1}^{m-1} \langle f_j \rangle_K \cdot \langle P_{K,k-1} f_m \rangle_{I_m} \cdot \prod_{j=m+1}^n \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}}(w_K) \right|_{Z_1,\ldots,n}^{p'_{n+1}} \right)^{1/p'_{n+1}}.
\]

Suppose that \( m > 1 \); if \( m = 1 \) one can start directly from (3.31) below. Now, with a fixed \( w \in W \) can use Stein’s inequality with respect to the function \( f_1 \) to have that the previous term is dominated by
\[
\left( \mathbb{E} \int_W \int_{\mathbb{R}^d} f_1(x) \sum_{K \in \mathcal{D}_{k,l}} e_K 1_K(x) \sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} b_{m,K,(I_j)} \prod_{j=1}^{m-1} \langle f_j \rangle_K \cdot \langle P_{K,k-1} f_m \rangle_{I_m} \cdot \prod_{j=m+1}^n \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}}(w_K) \right|_{Z_1,\ldots,n}^{p'_{n+1}} \right)^{1/p'_{n+1}}.
\]

To move forward, notice that by Lemma 3.24 we have
\[
\left| e_1 \sum_{k} \prod_{j=2}^n e_{j,k} \right|_{Z_1,\ldots,n} \lesssim \left| e_1 \sum_{k} \prod_{j=2}^n e_{j,k} \right|_{Y(X_{n+1})} \lesssim \left| e_1 \right|_{X_1} \sum_{k} \prod_{j=2}^n e_{j,k} \right|_{Z_2,\ldots,n},
\]

where \( Z_{2,\ldots,n} = Y(Y(X_2,\ldots,X_n)) \). Having established (3.30) we can now dominate (3.29) with
\[
\| f_1 \|_{L^{p_1}(X_1)} \left( \mathbb{E} \int_W \int_{\mathbb{R}^d} \left| \sum_{K \in \mathcal{D}_{k,l}} e_K 1_K(x) \sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} b_{m,K,(I_j)} \prod_{j=1}^{m-1} \langle f_j \rangle_K \cdot \langle P_{K,k-1} f_m \rangle_{I_m} \cdot \prod_{j=m+1}^n \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}}(w_K) \right|_{Z_2,\ldots,n}^{q_{2,\ldots,n}} \right)^{1/q_{2,\ldots,n}},
\]

where \( q_{2,\ldots,n} \) is defined by \( 1/q_{2,\ldots,n} = \sum_{j=2}^n 1/p_j \). We can continue this process. In the next step we argue as above but using the Hölder tuple \( (X_2,\ldots,X_n, Y(X_2,\ldots,X_n)) \). Below we write \( Z_{k_1,\ldots,k_2} = Y(Y(X_{k_1},\ldots,X_{k_2})) \) and \( 1/q_{k_1,\ldots,k_2} = \sum_{j=k_1}^{k_2} 1/p_j \).
Iterating this we arrive at $\prod_{j=1}^{m-1} \|f_j\|_{L^p_j(X_j)}$ multiplied by

$$
\left( \mathbb{E} \int_W \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K 1_K(x) \right| \sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} \prod_{j=m+1}^n \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}}(w_K) \bigg| \frac{q_m, \ldots, n}{Z_{m, \ldots, n}} \bigg| dx \, dv(w) \right)^{1/q_m, \ldots, n}.
$$

(3.31)

We assume that $m < n$; if $m = n$ one is already at (3.35) below. Let $K \in D_{k,l}$. We define the kernel

$$
b_{m,K}(x, y_m, \ldots, y_n) := \sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} |K|^{n-m+1} b_{m,K, (I_j)}
$$

$$
\times \prod_{j=m}^n \frac{1_{I_j}(y_j)}{|I_j|} h_{I_{n+1}}(x)
$$

(3.32)

so that

$$
\sum_{I_m^{(k)} = \cdots = I_{n+1}^{(k)} = K} b_{m,K, (I_j)} \langle P_{K,k-1} f_m \rangle_{I_m} \cdot \prod_{j=m+1}^n \langle f_j \rangle_{I_j} \cdot h_{I_{n+1}}(w_K)
$$

$$
= \frac{1}{|K|^{n-m+1}} \int_{K^{n-m+1}} b_{m,K}(w_K, y_m, \ldots, y_n) P_{K,k-1} f_m(y_m)
$$

$$
\prod_{j=m+1}^n f_j(y_j) \, dy.
$$

Using this representation in (3.31) we can use Stein’s inequality with respect to the function $f_n$ to have that the term in (3.31) is dominated by

$$
\left( \mathbb{E} \int_W \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K \frac{1}{|K|^{n-m}} \int_{K^{n-m}} b_{m,K}(w_K, y_m, \ldots, y_{n-1}, x) P_{K,k-1} f_m(y_m) \prod_{j=m+1}^{n-1} f_j(y_j) \, dy \right| \right)^{1/q_m, \ldots, n},
$$

which is (again by Lemma 3.24) dominated by $\|f_n\|_{L^p_n(X_n)}$ multiplied with

$$
\left( \mathbb{E} \int_W \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K \frac{1}{|K|^{n-m}} \int_{K^{n-m}} b_{m,K}(w_K, y_m, \ldots, y_{n-1}, x) P_{K,k-1} f_m(y_m) \right| \right)^{1/q_m, \ldots, n}.
$$
Next, we fix the point $w \in W$ and consider the term

\[
\mathbb{E} \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K \frac{1}{|K|^{n-m}} \int_{K^{n-m}} b_{m,K}(w_K, y_m, \ldots, y_{n-1}, x) P_{K,k-1} f_m(y_m) \prod_{j=m+1}^{n-1} f_j(y_j) \, dy \right|_{q_m,\ldots,n-1}^{q_m,\ldots,n-1} \, dx.
\]

To finish the estimate of (3.31) it is enough to dominate (3.33) by

\[
\prod_{j=m}^{n-1} \| f_j \|_{L^{p_j}(X_j)}
\]

uniformly in $w \in W$. Recalling (3.32) we can write (3.33) as

\[
\mathbb{E} \int_{\mathbb{R}^d} \left| \sum_{K \in D_{k,l}} \varepsilon_K \frac{1}{|K|^{n-m}} \sum_{I_m^{(k)} = \ldots = I_{n+1}^{(k)} = K} |P_{K,k-1} f_m| I_m \prod_{j=m+1}^{n-1} \langle f_j \rangle_{I_j} \frac{1}{|I_n|} h_{I_{n+1}}(w_K) \right|_{q_m,\ldots,n-1}^{q_m,\ldots,n-1} \, dx,
\]

which is further comparable to

\[
\mathbb{E} \int_{\mathbb{R}^d} \sum_{K \in D_{k,l}} \varepsilon_K \sum_{I_m^{(k)} = \ldots = I_{n+1}^{(k)} = K} \left( \sum_{I_m^{(k)} = \ldots = I_{n+1}^{(k)} = K} \frac{|K|}{|I_n|^{1/2}} b_{m,K,(I_j)} h_{I_{n+1}}(w_K) \right) \prod_{j=m+1}^{n-1} \langle f_j \rangle_{I_j} h_{I_n}(x) \right|_{q_m,\ldots,n-1}^{q_m,\ldots,n-1} \, dx.
\]

In the last step we were able to replace $1_{I_n}/|I_n|^{1/2}$ with $h_{I_n}$ because of the random signs, after which we removed the signs using UMD. Recalling the size of the coefficients $b_{m,K,(I_j)}$ from (3.26) we see that

\[
\sum_{I_m^{(k)} = \ldots = I_{n+1}^{(k)} = K} \frac{|K|}{|I_n|^{1/2}} b_{m,K,(I_j)} h_{I_{n+1}}(w_K) \leq \frac{|I_n|^{n-m+1/2}}{|K|^{n-m}}.
\]

since there is only one $I_{n+1}$ such that $h_{I_{n+1}}(w_K) \neq 0$. It is seen that after applying decoupling (3.34) is like (3.31) but the degree of linearity is one less. Therefore,
iterating this we see that (3.33) satisfies the desired bound if we can estimate

$$
\mathbb{E} \left| \int_{\mathbb{R}^d} \sum_{K \in D_{k,l}} \varepsilon K \sum_{j_m^{(k)}=j_{m+1}^{(k)}=K} b_{K, I_m, I_{m+1}} (P_k, k-1 f_m) I_m h_{I_{m+1}}(x) \left| \frac{p_m}{X_m} \right| \right| \, dx,
$$

(3.35)

where

$$
|b_{K, I_m, I_{m+1}}| \leq \frac{|I_m|^{3/2}}{|K|},
$$

by $\|f_m\|_{L_{p_m}^{p_m}(X_m)}$. This is a linear estimate and bounded exactly like the right hand side of (3.12). Therefore, we get the desired bound $\|f_m\|_{L_{p_m}^{p_m}(X_m)}$. This finally finishes our estimate for the term $A_m$.

We turn to estimate the parts $U_m$ from (2.24). Recall that $U_m(f_1, \ldots, f_n)$ is by definition

$$
\sum_{K} \sum_{I_1^{(k)} = \ldots = I_{n+1}^{(k)} = K} b_{K, (I_j)} \sum_{k=1}^{k-1} \sum_{m=0}^{m-1} \left( \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}^{(i+1)}} \cdot \langle \Delta f_{n+1}^{(i+1)} f_m \rangle_{I_{n+1}^{(i)}} \cdot \prod_{j=m+1}^{n} \langle f_j \rangle_{I_{n+1}^{(i)}} \right) h_{I_{n+1}}.
$$

(3.36)

Similarly as with the operators $A_m$, we use the fact that $L_{p_n}^{p_n}(Z_1, \ldots, n)$ has type $s = \min(p_n^{p_n'}, s_n^{p_n'})$ to reduce to controlling the term

$$
\sum_{K \in D_{k,l}} \sum_{I_1^{(k)} = \ldots = I_{n+1}^{(k)} = K} b_{K, (I_j)} \sum_{k=1}^{k-1} \sum_{m=0}^{m-1} \left( \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}^{(i+1)}} \cdot \langle \Delta f_{n+1}^{(i+1)} f_m \rangle_{I_{n+1}^{(i)}} \cdot \prod_{j=m+1}^{n} \langle f_j \rangle_{I_{n+1}^{(i)}} \right) h_{I_{n+1}}.
$$

(3.36)

uniformly on $l$. For every $I_{n+1}$ such that $I_{n+1}^{(k)} = K$ there holds that

$$
\left| \sum_{I_1^{(k)} = \ldots = I_{n}^{(k)} = K} b_{K, (I_j)} \right| \leq |I_{n+1}|^{1/2}.
$$

Therefore, using the UMD property and the Kahane contraction principle the $L_{p_n}^{p_n}(Z_1, \ldots, n)$-norm of (3.36) is dominated by

$$
\mathbb{E} \left\| \sum_{K \in D_{k,l}} \varepsilon K \sum_{I_1^{(k)} = K} \sum_{i=0}^{k-1} \sum_{j=1}^{m-1} \left( \prod_{j=1}^{n} \langle f_j \rangle_{I_{n+1}^{(i+1)}} \cdot \langle \Delta f_{n+1}^{(i+1)} f_m \rangle_{I_{n+1}^{(i)}} \cdot \prod_{j=m+1}^{n} \langle f_j \rangle_{I_{n+1}^{(i)}} \right) 1_{I} \right\|_{L_{p_n}^{p_n}(Z_1, \ldots, n)}.
$$
If all the averages were on the “level \( i + 1 \)” we could estimate this directly. Since there are the averages on the “level \( i \)” we need to further split this. There holds that
\[
\prod_{j=m+1}^{n} (f_{j})_{I^{(i)}} = (f_{m+1})_{I^{(i+1)}} \prod_{j=m+2}^{n} (f_{j})_{I^{(i)}} + (\Delta f_{m})_{I^{(i)}} \prod_{j=m+2}^{n} (f_{j})_{I^{(i)}}.
\]

Then, both of these are expanded in the same way related to \( f_{m+2} \) and so on. This gives that
\[
\prod_{j=m+1}^{n} (f_{j})_{I^{(i)}} 1_I = \sum_{\varphi} \prod_{j=m+1}^{n} D_{I^{(i+1)}}^\varphi(f_{j}) 1_I.
\]

Here the summation is over functions \( \varphi : \{ m+1, \ldots, n \} \to \{0, 1\} \) and for a cube \( I \in D \) we defined \( D_{I}^0 = E_I \) and \( D_{I}^1 = \Delta_I \). We also used the fact that \( \langle \Delta f_{j} \rangle_{I^{(i)}} 1_I = \Delta f_{j} \).

Finally, we take one \( \varphi \) and estimate the related term. It can be written as
\[
\mathbb{E}\left\| \sum_{K \in D_{k,l}} \mathcal{E}_{K} \sum_{L \subseteq K} \prod_{j=1}^{m-1} (f_{j})_{L} \cdot \Delta_{L} f_{m} \cdot \prod_{j=m+1}^{n} D_{L}^\varphi(f_{j}) \right\|_{L_{q_{1,\ldots,m}}^{\varphi}(Z_{1,\ldots,m})}^{1/q_{1,\ldots,m}}.
\]

If \( \varphi(j) = 0 \) for all \( j = m+1, \ldots, n \), then this can be estimated by a repeated use of Stein’s inequality (similarly as above) related to the functions \( f_{j}, j \neq m \).

Suppose that \( \varphi(j) \neq 0 \) for some \( j \). Notice that
\[
\sum_{K \in D_{k,l}} \mathcal{E}_{K} \sum_{L \subseteq K} \prod_{j=1}^{m-1} (f_{j})_{L} \cdot \Delta_{L} f_{m} \cdot \prod_{j=m+1}^{n} D_{L}^\varphi(f_{j})
\]
\[
= \mathbb{E}'\left( \sum_{K \in D_{k,l}} \mathcal{E}_{K} \sum_{L \subseteq K} \mathcal{E}'_{L} \prod_{j=1}^{m-1} (f_{j})_{L} \cdot \Delta_{L} f_{m} \right) \left( \sum_{L} \mathcal{E}'_{L} \prod_{j=m+1}^{n} D_{L}^\varphi(f_{j}) \right).
\]

Therefore, by Lemma 3.24 the term in (3.37) is dominated by
\[
\mathbb{E}\left( \mathbb{E}'\left( \sum_{K \in D_{k,l}} \mathcal{E}_{K} \sum_{L \subseteq K} \mathcal{E}'_{L} \prod_{j=1}^{m-1} (f_{j})_{L} \cdot \Delta_{L} f_{m} \right)^{q_{1,\ldots,m}}_{L^{q_{1,\ldots,m}}(Z_{1,\ldots,m})} \right)^{1/q_{1,\ldots,m}} \]

(3.38)
multiplied by
\[
\left( \mathbb{E} \left\| \sum_{L} \varepsilon'_L \prod_{j=m+1}^{n} D_L^{\omega(j)} f_j \right\|_{L^{q_{m+1}, \ldots, n}(Z_{m+1}, \ldots, n)}^{q_{m+1}, \ldots, n} \right)^{1/\delta_{m+1}, \ldots, n}. \tag{3.39}
\]

The term in (3.38) can be estimated by a repeated use of Stein’s inequality. The term in (3.39) is like the term in (3.37). If there is only one martingale difference, then we can again estimate directly with Stein’s inequality. If there is at least two martingale differences, then one can split into two as we did when we arrived at (3.38) and (3.39). This process is continued until one ends up with terms that contain only one martingale difference, and such terms we can estimate.

The proof of Proposition 3.25 is finished. \qed

With the essentially same proof as for the terms $A_m$ above we also have the following result.

**Proposition 3.40** Suppose that $S_{k, \ldots, k}$ is an n-linear shift of complexity $(k, \ldots, k)$ and $f_j : \mathbb{R}^d \to X_j$. Let $1 < p_j < \infty$ with $\sum_{j=1}^{n+1} 1/p_j = 1$. Then we have

\[
|\langle S_{k, \ldots, k}(f_1, \ldots, f_n), f_{n+1} \rangle| \lesssim (k + 1)^{\alpha} \prod_{j=1}^{n+1} \| f_j \|_{L^{p_j}(X_j)},
\]

where

\[
\alpha = \frac{1}{\min(p_1', \ldots, p_{n+1}', s_1', \ldots, s_{n+1}')} \quad \text{and } X_j \text{ has cotype } s_j.
\]

**Remark 3.41** Ordinary shifts obey a complexity free bound in the scalar-valued setting. However, we do not know how to achieve this with general UMD spaces. It seems that a somewhat better dependency could be obtained, but it would not have any practical use.

The following is [9, Theorem 5.3]. This is a significantly simpler argument than the shift proof and consists of repeated use of Stein’s inequality until one is reduced to the linear case (3.8).

**Proposition 3.42** Suppose that $\pi$ is an n-linear paraproduct and $f_j : \mathbb{R}^d \to X_j$. Let $1 < p_j < \infty$ with $\sum_{j=1}^{n+1} 1/p_j = 1$. Then we have

\[
|\langle \pi(f_1, \ldots, f_n), f_{n+1} \rangle| \lesssim \prod_{j=1}^{n+1} \| f_j \|_{L^{p_j}(X_j)}.\]

Finally, we are ready to state our main result concerning the UMD extensions of n-linear $\omega$-CZOs.
Theorem 3.43 Suppose that $T$ is an $n$-linear $\omega$-CZO. Suppose $\omega \in \text{Dini}_\alpha$, where

$$\alpha = \frac{1}{\min((n + 1)/n, s_1', \ldots, s_{n+1}')},$$

and $X_j$ has cotype $s_j$. Then for all exponents $1 < p_1, \ldots, p_n \leq \infty$ and $1/q_{n+1} = \sum_{j=1}^n 1/p_j > 0$ we have

$$\|T(f_1, \ldots, f_n)\|_{L^{q_{n+1}}(X_{n+1}^\ast)} \lesssim \prod_{j=1}^n \|f_j\|_{L^{p_j}(X_j)}.$$

**Proof** The important part is to establish the boundedness with a single tuple of exponents. We may e.g. conclude from the boundedness of the model operators and Theorem 2.26 that

$$|\langle T(f_1, \ldots, f_n), f_{n+1} \rangle| \lesssim \prod_{j=1}^{n+1} \|f_j\|_{L^{n+1}(X_j)}$$

if we choose $\alpha$ as in the statement of the theorem. It is completely standard how to improve this to cover the full range: we can e.g. prove the end point estimate $T : L^1(X_1) \times \cdots \times L^1(X_n) \to L^{1/n, \infty}(X_{n+1}^\ast)$, see [26], and then use interpolation or good-$\lambda$ methods. See e.g. [15, 27]. For such arguments the spaces $X_j$ no longer play any role (the scalar-valued proofs can readily be mimicked). \hfill $\Box$

**Remark 3.44** The exponent $(n + 1)/n$ in the definition of $\alpha$ is slightly annoying, since now the exponent $\alpha = 1/2$ valid in the scalar-valued case $X_1 = \cdots = X_{n+1} = \mathbb{C}$ does not follow from this result, even though then $s_1' = \cdots = s_{n+1}' = 2$. Of course, it is way more simple to prove scalar-valued estimates directly with other methods anyway (see Sect. 2.1).

Notice that it is also clear that $\text{Dini}_{1/2}$ suffices in suitable tuples $(X_1, \ldots, X_{n+1})$ of UMD function lattices. See e.g. [22, Sects. 2.10–2.12] for an account of the well-known square function and maximal function estimates valid in lattices. In lattices the simple approach of Sect. 2.1 is much better, as then in addition to the factor $(n + 1)/n$ we often have $s_j' < 2$.

In interesting and non-trivial situations the presence of $(n + 1)/n$ is not an additional restriction. Suppose each space $X_j$ is a non-commutative $L^p$ space $L^{p_j}(M)$ and $\sum_{j=1}^{n+1} 1/p_j = 1$, $1 < p_j < \infty$. Then the cotype of $X_j$ is $s_j = \max(2, p_j) \geq p_j$ so that

$$1 = \sum_{j=1}^{n+1} \frac{1}{p_j} \geq \sum_{j=1}^{n+1} \frac{1}{s_j} = n + 1 - \sum_{j=1}^{n+1} \frac{1}{s_j'},$$

and so there has to be an index $j$ so that $s_j' \leq (n + 1)/n$ anyway – thus $\min((n + 1)/n, s_1', \ldots, s_{n+1}') = \min(s_1', \ldots, s_{n+1}')$. 

\[\]
Acknowledgements  We thank Tuomas Hytönen and Kangwei Li for useful discussions.

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

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. Bourgain, J.: Some remarks on Banach spaces in which martingale difference sequences are unconditional. Ark. Mat. 21, 163–168 (1983)
2. Burkholder, D.L.: A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions, 1983, Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 270–286
3. Coifman, R.R., Meyer, Y.: Au delà des opérateurs pseudo-différentiels Astérisque. Société Mathématique de France, Paris (1978)
4. Culiuc, A., Di Plinio, F., Ou, Y.: Domination of multilinear singular integrals by positive sparse forms. J. Lond. Math. Soc. 98(2), 369–392 (2018). https://doi.org/10.1112/jlms.12139
5. David, G., Journé, J.-L.: A boundedness criterion for generalized Calderón–Zygmund operators. Ann. Math. (2) 120, 371–397 (1984). https://doi.org/10.2307/2006946
6. Deng, D., Yan, L., Yang, Q.: Blocking analysis and $T(1)$ theorem: ISSN1006-9283. Sci. China Ser. A 41(8), 801–808 (1998). https://doi.org/10.1007/BF02871663
7. Di Plinio, F., Ou, Y.: Banach-valued multilinear singular integrals. Indiana Univ. Math. J. 67(5), 1711–1763 (2018). https://doi.org/10.1512/iumj.2018.67.7466
8. Di Plinio, F., Li, K., Martikainen, H., Vuorinen, E.: Multilinear operator-valued Calderón–Zygmund theory. J. Funct. Anal. 279(8), 108666 (2020). https://doi.org/10.1016/j.jfa.2020.108666
9. Di Plinio, F., Li, K., Martikainen, H., Vuorinen, E.: Multilinear singular integrals on non-commutative $L^p$ spaces. Math. Ann. 378(3–4), 1371–1414 (2020). https://doi.org/10.1007/s00208-020-02068-4
10. Di Plinio, F., Li, K., Vuorinen, E.: Banach-valued multilinear singular integrals with modulation invariance. Int. Math. Res. Not. IMRN 7, 5256–5319 (2022). https://doi.org/10.1093/imrn/rnaa234
11. Figiel, T.: On equivalence of some bases to the Haar system in spaces of vector-valued functions. Bull. Polish Acad. Sci. Math. 36(3–4), 119–131 (1988)
12. Figiel, T.: Singular Integral Operators: A Martingale Approach. Geometry of Banach Spaces (Strobl, 1989). Lecture Note Ser, vol. 158, pp. 95–110. Cambridge Univ. Press, Cambridge (1990)
13. Grafakos, L., Martell, J.M.: Extrapolation of weighted norm inequalities for multivariable operators and applications. J. Geom. Anal. 14(1), 19–46 (2004). https://doi.org/10.1007/BF02921864
14. Grafakos, L., Oh, S.: Kato-Ponce, the inequality. Commun. Partial Differ. Equ. 39(6), 1128–1157 (2014). https://doi.org/10.1080/036050530.2013.822885
15. Grafakos, L., Torres, R.H.: Calderón–Zygmund, multilinear, theory. Adv. Math. 165(1), 124–164 (2002)
16. Grau de la Herrán, A., Hytönen, T.: Dyadic representation and boundedness of nonhomogeneous Calderón–Zygmund operators with mild kernel regularity. Michigan Math. J. 67(4), 757–786 (2018). https://doi.org/10.1307/mmj/1531447374
17. Hämmen, T.S., Hytönen, T.P.: Operator-valued dyadic shifts and the $T(1)$ theorem. Monatsh. Math. 180(2), 213–253 (2016). https://doi.org/10.1007/s00605-016-0891-3
18. Hytönen, T.: Representation of singular integrals by dyadic operators, and the $A_2$ theorem. Expo. Math. 35, 166–205 (2011)
19. Hytönen, T.P.: The sharp weighted bound for general Calderón–Zygmund operators. Ann. Math. (2) 175(5), 1473–1506 (2012)
20. Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach spaces, vol. I. Springer-Verlag, Martingales and Littlewood-Paley theory (2016)
21. Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach spaces, vol. II. Springer-Verlag, Probabilistic methods and operator theory (2017) https://doi.org/10.1007/978-3-19-69808-3
22. Hytönen, T., Martikainen, H., Vuorinen, E.: Multi-parameter estimates via operator-valued shifts. Proc. Lond. Math. Soc. 119(6), 1560–1597 (2019)
23. Kato, T., Ponce, G.: Commutator estimates and the Euler and Navier–Stokes equations. Commun. Pure Appl. Math. 41(7), 891–907 (1988). https://doi.org/10.1002/cpa.3160410704
24. Li, K., Martell, J.M., Ombrosi, S.: Extrapolation for multilinear Muckenhoupt classes and applications. Adv. Math. 373, 107286 (2020). https://doi.org/10.1016/j.aim.2020.107286
25. Li, K., Martell, J.M., Martikainen, H., Ombrosi, S., Vuorinen, E.: End-point estimates, extrapolation for multilinear Muckenhoupt classes, and applications. Trans. Am. Math. Soc. 374(1), 97–135 (2021). https://doi.org/10.1090/tran/8172
26. Lu, G., Zhang, P.: Calderón–Zygmund, multilinear, operators with kernels of Dini’s type and applications. Nonlinear Anal. 107, 92–117 (2014). https://doi.org/10.1016/j.na.2014.05.005
27. Martikainen, H., Vuorinen, E.: Dyadic-probabilistic methods in bilinear analysis. Mem. Am. Math. Soc. 274, 1344 (2021)
28. Nierath, Z.: Quantitative estimates and extrapolation for multilinear weight classes. Math. Ann. 375(1–2), 453–507 (2019). https://doi.org/10.1007/s00208-019-01816-5
29. Petermichl, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical $A_p$ characteristic. Am. J. Math. 129(5), 1355–1375 (2007)

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