MULTILINEAR SINGULAR INTEGRALS ON NON-COMMUTATIVE L^p SPACES

FRANCESCO DI PLINIO, KANGWEI LI, HENRI MARTIKAINEN, AND EMIL VUORINEN

Abstract. We prove L^p bounds for the extensions of standard multilinear Calderón-Zygmund operators to tuples of UMD spaces tied by a natural product structure. The product can, for instance, mean the pointwise product in UMD function lattices, or the composition of operators in the Schatten-von Neumann subclass of the algebra of bounded operators on a Hilbert space. We do not require additional assumptions beyond UMD on each space – in contrast to previous results, we e.g. show that the Rademacher maximal function property is not necessary. The obtained generality allows for novel applications. For instance, we prove new versions of fractional Leibniz rules via our results concerning the boundedness of multilinear singular integrals in non-commutative L^p spaces. Our proof techniques combine a novel scheme of induction on the multilinearity index with dyadic-probabilistic techniques in the UMD space setting.

1. Introduction

A Banach space X has the UMD property if any X-valued martingale difference sequence converges unconditionally in L^p for some (equivalently, all) $p \in (1, \infty)$. Standard examples of UMD spaces are provided by the reflexive L^p function spaces, as well as the reflexive Schatten-von Neumann subclasses S^p of the algebra of bounded operators on a Hilbert space. The works by Burkholder [2] and Bourgain [1] yield an alternative characterization: X is a UMD space if and only if singular integrals, in particular the Hilbert transform, admit an $L^p(X)$-bounded extension. Such equivalence, albeit striking, is not so surprising when viewed from the modern dyadic-probabilistic perspective on singular integral operators. Indeed, Petermichl [43, 44] realized that the Hilbert transform lies in the convex hull of certain dyadic operators akin to martingale transforms (the so-called dyadic shifts), while Hytönen [28] extended this representation to general singular integral operators of Calderón-Zygmund type, relying on a probabilistic construction. These

2010 *Mathematics Subject Classification.* 42B20 (primary), 46E40, 46L52 (secondary).

Key words and phrases. Calderón–Zygmund operators, singular integrals, multilinear analysis, non-commutative spaces, representation theorems, UMD spaces.

F. Di Plinio has been partially supported by the National Science Foundation under the grants NSF-DMS-1650810 and NSF-DMS-1800628.

K. Li was supported by Juan de la Cierva - Formación 2015 FJCI-2015-24547, by the Basque Government through the BERC 2018-2021 program and by Spanish Ministry of Economy and Competitiveness MINECO through BCAM Severo Ochoa excellence accreditation SEV-2017-0718 and through project MTM2017-82160-C2-1-P funded by (AEI/FEDER, UE) and acronym "HAQMEC".

H. Martikainen was supported by the Academy of Finland through the grants 294840 and 306901, and by the three-year research grant 75160010 of the University of Helsinki. He is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research.

E. Vuorinen was supported by the Academy of Finland through the grant 306901, by the Finnish Centre of Excellence in Analysis and Dynamics Research, and by Jenny and Antti Wihuri Foundation.

results have roots in the pioneering work of Figiel [13] and on the probabilistic approach of Nazarov–Treil–Volberg to non-homogeneous Tb theorems [41].

The theory of *linear* singular integrals on Banach spaces, beyond its intrinsic interest, has historically been motivated by its interplay with several related areas, such as geometry of Banach spaces [31, 32], elliptic and parabolic regularity theory [3, 47], the theory of quasiconformal mappings [15]. Furthermore, vector-valued bounds may often be used in the pursuit of their multi-parameter analogs [22, 27].

In this article, we are concerned with Banach-valued extensions of *multilinear* singular integral operators. A linear singular integral takes the general form

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\,\mathrm{d}y,$$

where different assumptions on the *kernel* K lead to important classes of linear transformations arising across pure and applied analysis. The term singular integral refers just to the underlying kernel structure – a Calderón-Zygmund operator is a bounded singular integral operator. A heuristic model of an n-linear singular integral operator T in $\mathbb{R}^d$ is then obtained by setting

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

where U is a linear singular integral operator in $\mathbb{R}^{nd}$. For the basic theory see e.g. Grafakos–Torres [18].

Multilinear singular integrals arise naturally from applications to partial differential equations, complex function theory and ergodic theory, among others. Focusing on the results of greater significance for the present work, we mention that L^p estimates for the fractional derivative of a product, often referred to as *fractional Leibniz rules*, are widely employed in the study of dispersive equations starting from the work of Kato and Ponce [33], descend from the multilinear Hörmander-Mihlin multiplier theorem of Coifman-Meyer [4]. The bilinear Hilbert transform is a prime example of a *modulation invariant* bilinear Calderón-Zygmund operator. It rose to prominence with Calderón's first commutator program, and has been featured as a model operator in the study of bilinear ergodic averages; the latter connection is expounded in e.g. [11]. Proving L^p estimates for the bilinear Hilbert transform in the Lacey-Thiele framework [34, 35] involves a decomposition into *single trees*, which are essentially modulated bilinear Calderón-Zygmund operators.

Vector-valued extensions of multilinear Calderón-Zygmund operators have mostly been studied within the more restrictive framework of ℓ^p spaces and function lattices. Boundedness of these extensions is classically obtained through weighted norm inequalities, more recently in connection with localized techniques such as sparse domination: see [16] and the more recent [6, 37, 42] for a non-exhaustive overview of their interplay. The paper [10], by Y. Ou and one of us, contains a bilinear multiplier theorem which applies to certain non-lattice UMD spaces. The approach of [10] is based on a localization of the UMD-valued tent space norms, see for instance [23], within the Carleson embedding framework of Do and Thiele [12]. The tent space techniques lead to the additional assumption of L^p estimates for a certain analogue of the Hardy-Littlewood maximal operator obtained by replacing uniform bounds with randomized, or $\mathcal{R}$-bounds, see e.g. [47] for a definition. This assumption, usually referred to as the RMF property of X, dates back to the work of Hytönen, McIntosh and Portal on the vector-valued Kato square root

problem [21], and is in fact necessary for the X-valued Carleson embedding theorem to hold [20].

In this article, we obtain vector-valued extensions of multilinear singular integrals to tuples of UMD spaces tied by a natural product structure, such as that of pointwise product in UMD function lattices or, more generally in fact, that of composition within the Schatten-von Neumann classes. We do not require additional conditions on the spaces involved – in particular, we do not require the RMF property. Thus, we are able to extend multilinear Calrerón–Zygmund operators to natural tuples of non-commutative L^p spaces – a result which does not seem attainable via abstract theorems involving multilinear RMF type assumptions. A motivating corollary is a version of the fractional Leibniz rule for products of Schatten-von Neumann class-valued functions.

In contrast to [10, 21, 23], our techniques are dyadic-probabilistic: a multilinear version of the representation theorem of Hytönen [28], which appeared in the bilinear case in [39] by Y. Ou and three of us, reduces the problem to the boundedness of the extensions of a class of multilinear dyadic model operators, namely *paraproducts* and *multilinear dyadic shifts* of arbitrary complexity. The novelty lies in how we treat these operators – multilinearity poses significant problems in the vector-valued setup.

We note that UMD-valued extensions of bilinear, complexity zero dyadic shifts have implicitly been treated in the work by Hytönen, Lacey and Parissis on the UMD dyadic model of the bilinear Hilbert transform [30, Section 6]. The simple approach of [30] does not extend to either the higher complexity or the multilinear cases. We tackle the n-linear case by inducting suitably on the linearity, which is made possible by associating to our n-tuples of UMD spaces a collection of related m-tuples, $m < n$. The framework is carefully designed to allow us to treat non-commutative theory. Moreover, bilinear theory would not reveal all the difficulties and is, in fact, strictly easier – a feature that is also present in our followup paper [9] involving operator-valued multilinear analysis. Before providing further insights on the novelty of our proof techniques, and comparisons to previous approaches, we give the statements of our main results.

1.1. **Main results.** We start by discussing a simpler question, where the current literature already has some restrictions that we can lift. If X is a Banach space and T is an n-linear integral operator on $\mathbb{R}^d$ acting on n-tuples of functions in $L^\infty_c(\mathbb{R}^d)$, we may let T act on $(L^\infty_c(\mathbb{R}^d) \otimes X) \times L^\infty_c(\mathbb{R}^d) \times \cdots \times L^\infty_c(\mathbb{R}^d)$ by

$$T(f_1, f_2, \dots, f_n)(x) = \sum e_{1,j} T(f_{1,j}, f_2, \dots, f_n)(x), \qquad x \in \mathbb{R}^d,$$
$$f_1 = \sum e_{1,j} f_{1,j}, \qquad f_{1,j} \in L^\infty_c(\mathbb{R}^d),\ e_{1,j} \in X.$$

A basic thing implied by our methods is that n-linear Calderón-Zygmund operators extend boundedly when applied to one UMD-valued function and $n-1$ scalar functions, without any additional assumption on the UMD space. We send to Subsection 2.4 for the precise definition of an n-linear Calderón-Zygmund operator. This is the simplest complete multilinear analogue of Bourgain's UMD Hörmander-Mihlin multiplier theorem from [1]; see also Weis [47] and Hytönen-Weis [26] for the operator-valued, non-translation invariant case.

In the bilinear, translation invariant, operator-valued setting, a related result appeared in [10, Corollary 1.2] under the assumption, known to be rather restrictive, that X is a UMD space with the non-tangential Rademacher maximal function property [21]. Theorem 1.1

shows, in particular, that the latter assumption is unnecessary. However, we formulate the following more general version to facilitate the discussion below regarding the somewhat special nature of bilinear theory.

1.1. **Theorem.** *Let X_1, X_2, Y_3 be* UMD *spaces with an associated product (a bounded bilinear operator)*

$$X_1 \times X_2 \to Y_3 \colon (x_1, x_2) \mapsto x_1 x_2, \qquad |x_1 x_2|_{Y_3} \leq |x_1|_{X_1} |x_2|_{X_2}.$$

Let $n \geq 2$ and T be an n-linear Calderón-Zygmund operator on $\mathbb{R}^d$. The n-linear operator

$$T(f_1, f_2, \dots, f_n)(x) = \sum_{j_1, j_2} e_{1,j_1} e_{2,j_2} T(f_{1,j_1}, f_{2,j_2}, f_3, \dots, f_n)(x), \qquad x \in \mathbb{R}^d,$$
$$f_1 = \sum_{j_1} e_{1,j_1} f_{1,j_1}, \ f_2 = \sum_{j_2} e_{2,j_2} f_{2,j_2} \qquad f_{1,j}, f_{2,j} \in L^\infty_c(\mathbb{R}^d), \ e_{1,j_1} \in X_1, \ e_{2,j_2} \in X_2,$$

extends to a bounded operator

$$T \colon L^{p_1}(\mathbb{R}^d; X_1) \times L^{p_2}(\mathbb{R}^d; X_2) \times \prod_{k=3}^{n} L^{p_j}(\mathbb{R}^d) \to L^{q_{n+1}}(\mathbb{R}^d; Y_3),$$
$$1 < p_k \leq \infty, \ \frac{1}{n} < q_{n+1} < \infty, \ \tfrac{1}{q_{n+1}} = \sum_{k=1}^{n} \tfrac{1}{p_k}.$$

The proof of this model case is an adaptation of the proof of Theorem 3.31 with some additional observations regarding the bilinear case – see Remark 4.13. This simpler result also showcases why the genuine n-linear theory that we formulate next is harder than bilinear theory: the n-linear theory requires us to exploit a more careful product setting so that we can run our inductive proof. We also note that at least in the basic case $X_1 = Y_3 = X$ and $X_2 = \mathbb{C}$, Theorem 1.1 can also be seen as a corollary of Theorem 3.31 using Example 3.17. It is simpler to just look at the proof, however.

Our main theorem concerns extensions of n-linear CZO operators T to an n-tuple $X_1, \dots, X_n$ of UMD Banach spaces lying in an enveloping algebra $\mathcal{A}$, allowing for a standard definition of (associative, not necessarily abelian) product $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$. We refer to these configurations as UMD Hölder tuples if certain conditions are in place, in particular, if the n-tuples are associated with suitable collections of related m-tuples, $m < n$. If each X_j is a subspace of $\mathcal{A}$, and $f_k \in L^\infty_c(\mathbb{R}^d) \otimes X_k$ for $1 \leq k \leq n$, we may define the extension of a scalar integral operator by

$$T(f_1, \dots, f_n)(x) = \sum_{j_1, \dots, j_n} T(f_{1,j_1}, \dots, f_{n,j_n})(x) \prod_{k=1}^{n} e_{k,j_k}, \qquad x \in \mathbb{R}^d, \tag{1.2}$$
$$f_k = \sum_{j_k} e_{k,j_k} f_{k,j_k}, \qquad f_{k,j_k} \in L^\infty_c(\mathbb{R}^d), \ e_{k,j_k} \in X_k.$$

The abstract setup is developed in Section 3. For expository purposes, herein we provide a statement in a rather general concrete case of a UMD Hölder tuple. In the statement, we denote by $L^p(\mathcal{M})$ the non-commutative L^p spaces associated to a von Neumann algebra $\mathcal{M}$ endowed with a normal, semifinite, faithful trace τ.

1.3. **Theorem.** *Let $\mathcal{M}$ be a von Neumann algebra endowed with a normal, semifinite, faithful trace. For $s = 1, \dots, S$, let (M_s, μ_s) be measure spaces and for $s = 0, \dots, S$ let*

$$1 < p_1^s, \dots p_n^s, q_{n+1}^s < \infty, \qquad \frac{1}{q_{n+1}^s} = \sum_{k=1}^n \frac{1}{p_k^s}$$

be Banach Hölder tuples. Let

$$\begin{aligned} X_k &= L^{p_k^S}(M_S, \mu_s; L^{p_k^{S-1}}(M_{S-1}, \mu_{S-1}; \cdots L^{p_k^1}(M_1, \mu_1; L^{p_k^0}(\mathcal{M}))\cdots), \qquad k = 1, \dots, n, \\ Y_{n+1} &= L^{q_{n+1}^S}(M_S, \mu_S; L^{q_{n+1}^{S-1}}(M_{S-1}, \mu_{S-1}; \cdots L^{q_{n+1}^1}(M_1, \mu_1; L^{q_{n+1}^0}(\mathcal{M}))\cdots). \end{aligned} \tag{1.4}$$

The n-linear operator (1.2) extends to a bounded operator

$$T : \prod_{k=1}^n L^{p_k}(\mathbb{R}^d; X_k) \to L^{q_{n+1}}(\mathbb{R}^d; Y_{n+1}), \quad 1 < p_k \le \infty, \ \frac{1}{n} < q_{n+1} < \infty, \ \tfrac{1}{q_{n+1}} = \sum_{k=1}^n \tfrac{1}{p_k},$$
$$T : \prod_{k=1}^n L^1(\mathbb{R}^d; X_k) \to L^{\frac{1}{n},\infty}(\mathbb{R}^d; Y_{n+1}).$$

In fact, we have the stronger estimate

$$\begin{aligned} |\langle T(f_1, \dots, f_n), f_{n+1}\rangle| &\lesssim \left\| \mathrm{M}\left(|f_1|_{X_1}, \dots, |f_n|_{X_n}, |f_{n+1}|_{Y_{n+1}^*}\right)\right\|_1, \\ \mathrm{M}(g_1, \dots, g_{n+1})(x) &= \sup_{x \in Q} \prod_{j=1}^{n+1} \langle |g_j| \rangle_Q, \qquad \langle g \rangle_Q := \frac{1}{|Q|}\int_Q g. \end{aligned} \tag{1.5}$$

The estimate (1.5) is equivalent to a certain sparse bound, see Remark 3.29.

We send to Subsection 3.3 and to the references [7, 8] for more details on sparse bounds and to [37, 38] for a survey of the weighted inequalities that may be derived as a consequence.

Theorem 1.3 is obtained as a corollary of Theorem 3.31 using Example 3.21. However, we remark that, at least to the best of the authors' knowledge, the spaces (1.4) encompass all known *examples* of UMD Banach spaces. We further remark that the mixed norm structure of the spaces (1.4) prevents from using purely non-commutative tools, as (1.4) may be interpreted as semi-commutative spaces only if p_k^s does not vary with s for all $1 \le k \le n$; on the other hand, (1.4) are not UMD lattices, so that Theorem 1.3 is out of reach of purely lattice-type techniques.

Theorems 1.1 and 1.3 can be used to deduce certain weighted multilinear Leibniz rules in the UMD-valued and non-commutative setting. For simplicity of notation, we particularize the statements to the bilinear, unweighted, non-endpoint case for the homogeneous fractional derivative $D^s f = \mathcal{F}^{-1}(|\xi|^s \widehat{f}(\xi))$, in the setting of Theorem 1.1. A variety of formulations may be found e.g. in the article by Grafakos and Oh [17].

1.6. **Corollary** (Fractional Leibniz rules in UMD spaces)**.** *Let X_1, X_2, Y_3 be* UMD *spaces as in the statement of Theorem 1.1. For all sufficiently smooth $f_1 : \mathbb{R}^d \to X_1$, $f_2 : \mathbb{R}^d \to X_2$ there holds*

$$\left\|D^s(f_1 f_2)\right\|_{L^{q_3}(\mathbb{R}^d;Y_3)} \lesssim \left\|D^s f_1\right\|_{L^{p_1}(\mathbb{R}^d;X_1)} \left\|f_2\right\|_{L^{p_2}(\mathbb{R}^d;X_2)} + \left\|f_1\right\|_{L^{r_1}(\mathbb{R}^d;X_1)} \left\|D^s f_2\right\|_{L^{r_2}(\mathbb{R}^d;X_2)}$$

whenever $s > d$ and

$$1 < p_1, p_2, r_1, r_2 \le \infty,\ \frac{1}{2} < q_3 < \infty,\ \tfrac{1}{q_3} = \tfrac{1}{p_1} + \tfrac{1}{p_2} = \tfrac{1}{r_1} + \tfrac{1}{r_2}.$$

Corollary 1.6 appears to be the first instance of a Leibniz type rule in the full vector-valued setting, with no additional assumptions on the UMD spaces involved. We have not strived for optimality of the range for the fractional exponent s. While the range obtained in Corollary 1.6 is wider than what would follow from results of Coifman-Meyer type, see [17, Remark 1], the extension to the sharp range $s > \max\left\{0, d\left(\frac{1}{q_3} - 1\right)\right\}$ requires bilinear estimates for kernels which fail to be of the standard CZ type considered herein. Such estimates are carried out e.g. in [17]: their extension to the full vector-valued setting is left for future work.

Proof of Corollary 1.6 . We follow the beginning of the proof of [17, Theorem 1]. The estimate we seek is reduced to a bound for the UMD-valued extension of three different bilinear paraproducts (meaning suitable parts of a Littlewood–Paley decomposition of a product of functions – not in the exact sense as we use the word in connection with dyadic model operators). We note that the symbol of the high-low paraproducts Π_1 and Π_2 is of Coifman-Meyer type; therefore Π_1, Π_2 are bilinear CZO operators as defined in Subsection 2.4 and Theorem 1.3 applies directly. The high-high term Π_3 is a bilinear integral operator with kernel

$$K(x, y_1, y_2) = \sum_{m\in\mathbb{Z}} \int_{\mathbb{R}^d} 2^{3md} \phi_s(2^m(u-x))\psi(2^m(u-y_1))\psi(2^m(u-y_2))\,\mathrm{d}u$$

where ψ is a Schwartz function whose Fourier transform Ψ is supported in an annular region around the origin and $\phi_s = D^s\phi$ for some Schwartz function ϕ such that its Fourier transform has compact support containing 0, so that

$$|\phi_s(x)| \lesssim (1 + |x|)^{-(d+s)}, \qquad x \in \mathbb{R}^d.$$

As $s > d$ for us, this implies that Π_3 is a bilinear CZO operator with a kernel K satisfying

$$\|\Pi_3\|_{L^3\times L^3 \to L^{\frac{3}{2}}} + \|K\|_{\mathrm{CZ}_{(s-d)/2}} \lesssim 1,$$

where $\|K\|_{\mathrm{CZ}_\alpha}$ is the kernel constant defined in the beginning of Section 2.4. The required bounds for Π_3 follow from an application of Theorem 1.1. □

1.2. **Proof techniques and novelties.** A basic example of an n-linear dyadic shift operator of complexity zero on $\mathbb{R}$, in adjoint form, is

$$(f_1, \dots, f_{n+1}) \mapsto \sum_{m\in\mathbb{Z}} \varepsilon_m \int \Big(\prod_{k\in C} \Delta_m f_k(x)\Big)\Big(\prod_{k\in N} E_m f_k(x)\Big)\,\mathrm{d}x$$

where ε_m are bounded coefficients, and E_m and Δ_m respectively indicate the conditional expectation on the m-th dyadic filtration and the corresponding martingale difference, $C \cap N = \emptyset$ and $C \cup N = \{1, \dots, n+1\}$, with the key feature that the cardinality of the *cancellative* indices C is always at least 2. We approach UMD-valued extensions of the above forms to $(n+1)$-tuples of UMD spaces via a novel induction argument, aimed at reducing the cardinality of the set of *non-cancellative* indices N and the linearity of the shift n at the same time. The induction relies upon a certain structure of the tuples involved, which is most easily described in the bilinear, $n = 2$, case. Loosely speaking, we consider

UMD spaces X_1, X_2, X_3 endowed with a linear functional τ defined on all products $e_1e_2e_3$, $e_j \in X_j$, with the property that

$$\|e_1\|_{X_1} \sim \sup_{|e_2|_{X_2}=|e_3|_{X_3}=1} |\tau(e_1e_2e_3)|$$

and the same holds for all permutations of X_1, X_2, X_3. In combination with the martingale decoupling inequality of McConnell [40] and Hytönen [29], and Stein's inequality in UMD spaces, this structure allows to reduce a trilinear shift form on X_1, X_2, X_3 where, say, $1 \in C$ and $2 \in N$, to a bilinear shift form on X_1, X_1^*, where both indices are cancellative, and whose boundedness is known from the UMD character of X_1. The induction is crucial in the n-linear case to allow a repeated use of Stein's inequality.

We remark here that the martingale decoupling has been previously used by Hänninen and Hytönen [19] in the proof of a $T1$ theorem for linear singular integrals on UMD spaces with operator-valued kernels, providing among other results a non-translation invariant analogue of Weis's theorem [47]. The multilinear operator-valued theory, together with a related representation theorem, is the object of forthcoming work by the authors [9].

Acknowledgments. The authors would like to warmly thank Yumeng Ou for fruitful discussions on the subject of multilinear UMD-valued singular integrals. F. Di Plinio is grateful to Ben Hayes and Vittorino Pata for enlightening exchanges on factorization in noncommutative L^p spaces.

2. Definitions and preliminaries

2.1. **Vinogradov notation.** We write $A \lesssim B$ if $A \leq CB$ for some absolute constant C. The constant C can at least depend on the dimensions of the appearing Euclidean spaces, on integration exponents, on the degree of linearity of the multilinear operators, and on various Banach space constants. We use the notation $A \sim B$ if $B \lesssim A \lesssim B$.

2.2. **Dyadic notation.** Let $\mathcal{D}_0$ be the dyadic lattice in $\mathbb{R}^d$, defined by

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

We recall the random dyadic grids of Nazarov–Treil–Volberg, see for example [41]. The version we use here is from [29]. Let $\Omega = (\{0,1\}^d)^{\mathbb{Z}}$ and let $\mathbb{P}$ be the natural probability measure on Ω such that the coordinates are independent and uniformly distributed on $\{0,1\}^d$. If $Q \in \mathcal{D}_0$ and $\omega = (\omega_k)_{k\in\mathbb{Z}} \in \Omega$, we set

$$Q + \omega := Q + \sum_{k\colon 2^{-k}<\ell(Q)} \omega_k 2^{-k}.$$

The random dyadic lattice $\mathcal{D}_\omega$ on $\mathbb{R}^d$ is defined by $\mathcal{D}_\omega = \{Q + \omega \colon Q \in \mathcal{D}_0\}$. By a dyadic lattice $\mathcal{D}$ we mean that $\mathcal{D} = \mathcal{D}_\omega$ for some ω.

Let X be a Banach space. If $p \in (0,\infty]$ we denote by $L^p(X) = L^p(\mathbb{R}^d; X)$ the usual Bochner space of X-valued functions $f\colon \mathbb{R}^d \to X$. Let $\mathcal{D}$ be a dyadic lattice. Suppose $Q \in \mathcal{D}$ and $f \in L^1_{\mathrm{loc}}(X)$ (the set of locally integrable functions). We use the following notation:

- The side length of Q is denoted by $\ell(Q)$;
- $\mathrm{ch}(Q)$ consists of those $Q' \in \mathcal{D}$ such that $Q' \subset Q$ and $\ell(Q') = \ell(Q)/2$;
- If $k \in \mathbb{Z}, k \geq 0$, then $Q^{(k)}$ denotes the cube $R \in \mathcal{D}$ such that $Q \subset R$ and $2^k\ell(Q) = \ell(R)$;
- The average of f over Q is $\langle f\rangle_Q = \frac{1}{|Q|}\int_Q f\,\mathrm{d}x$, and we also write $E_Q f = \langle f\rangle_Q 1_Q$;

- The martingale difference $\Delta_Q f$ is $\Delta_Q f = \sum_{Q' \in \mathrm{ch}(Q)} E_{Q'} f - E_Q f$;
- For $k \in \mathbb{Z}$, $k \geq 0$, define

$$\Delta_Q^k f = \sum_{\substack{R \in \mathcal{D} \\ R^{(k)} = Q}} \Delta_R f \quad \text{and} \quad E_Q^k f = \sum_{\substack{R \in \mathcal{D} \\ R^{(k)} = Q}} E_R f.$$

Haar functions. When $Q \in \mathcal{D}$ we denote by h_Q a cancellative L^2 normalized Haar function. This means the following. Writing $Q = I_1 \times \cdots \times I_d$ we can define the Haar function h_Q^η, $\eta = (\eta_1, \ldots, \eta_d) \in \{0,1\}^d$, by setting

$$h_Q^\eta = h_{I_1}^{\eta_1} \otimes \cdots \otimes h_{I_d}^{\eta_d},$$

where $h_{I_i}^0 = |I_i|^{-1/2} 1_{I_i}$ and $h_{I_i}^1 = |I_i|^{-1/2}(1_{I_{i,l}} - 1_{I_{i,r}})$ for every $i = 1, \ldots, d$. Here $I_{i,l}$ and $I_{i,r}$ are the left and right halves of the interval I_i respectively. If $\eta \neq 0$ the Haar function is cancellative: $\int h_Q^\eta = 0$. We usually exploit notation by suppressing the presence of η, and simply write h_Q for some h_Q^η, $\eta \neq 0$.

Notice that if $f \in L^1_{\mathrm{loc}}(X)$, then $\Delta_Q f = \sum_{\eta \neq 0} \langle f, h_Q^\eta \rangle h_Q^\eta$, or suppressing the η summation, $\Delta_Q f = \langle f, h_Q \rangle h_Q$. Here $\langle f, h_Q \rangle = \int f h_Q$.

2.3. **Definitions and properties related to Banach spaces.** An extensive treatment of Banach space theory is given in the books [24, 25] by Hytönen, van Neerven, Veraar and Weis.

We say that $\{\varepsilon_k\}_k$ is a *collection of independent random signs*, where k runs over some index set, if there exists a probability space $(\mathcal{M}, \mu)$ so that $\varepsilon \colon \mathcal{M} \to \{-1, 1\}$, $\{\varepsilon_k\}_k$ is independent and $\mu(\{\varepsilon_k = 1\}) = \mu(\{\varepsilon_k = -1\}) = 1/2$. Below, $\{\varepsilon_k\}_k$ 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 $x_1, \ldots, x_M \in X$ and $p, q \in (0, \infty)$ there holds that

$$\Big(\mathbb{E}\Big|\sum_{m=1}^M \varepsilon_m x_m\Big|_X^p\Big)^{1/p} \sim \Big(\mathbb{E}\Big|\sum_{m=1}^M \varepsilon_m x_m\Big|_X^q\Big)^{1/q}.$$

We also denote

$$\|(x_m)\|_{\mathrm{Rad}(X)} := \Big(\mathbb{E}\Big|\sum \varepsilon_m x_m\Big|_X^2\Big)^{1/2}.$$

The Kahane contraction principle says that if $(a_m)_{m=1}^M$ is a sequence of scalars and $p \in (0, \infty]$, then

$$\Big(\mathbb{E}\Big|\sum_{m=1}^M \varepsilon_m a_m x_m\Big|_X^p\Big)^{1/p} \lesssim \max |a_m| \Big(\mathbb{E}\Big|\sum_{m=1}^M \varepsilon_m x_m\Big|_X^p\Big)^{1/p}. \tag{2.1}$$

Actually, if $p \in [1, \infty]$ and $a_m \in \mathbb{R}$, then (2.1) holds with "$\leq$" in place of "$\lesssim$", see [24] for more details.

A Banach space X is said to be a UMD *space* if for all $p \in (1, \infty)$, all X-valued L^p-martingale difference sequences $(d_j)_{j=1}^k$ and signs $\epsilon_j \in \{-1, 1\}$ there holds that

$$\Big\|\sum_{j=1}^k \epsilon_j d_j\Big\|_{L^p(X)} \lesssim \Big\|\sum_{j=1}^k d_j\Big\|_{L^p(X)}. \tag{2.2}$$

Here the $L^p(X)$-norm is with respect to the measure space where the martingale differences are defined. If the estimate (2.2) holds for one $p_0 \in (1,\infty)$, then it holds for all $p \in (1,\infty)$.

A version for UMD-valued functions of Stein's inequality concerning conditional expectations is due to Bourgain. For a proof, see for example [24, Theorem 4.2.23]. For our purposes we formulate the estimate in the following way. Suppose X is a UMD space and let $\mathcal{D} \subset \mathbb{R}^d$ be a dyadic lattice. Suppose that for each $Q \in \mathcal{D}$ we have a function $f_Q \in L^1_{\mathrm{loc}}(X)$ supported in Q (such that only finitely many of them are non-zero). Then for all $p \in (1,\infty)$ there holds that

$$\mathbb{E}\Big\| \sum_{Q\in\mathcal{D}} \varepsilon_Q \langle f_Q\rangle_Q 1_Q \Big\|_{L^p(X)} \lesssim \mathbb{E}\Big\| \sum_{Q\in\mathcal{D}} \varepsilon_Q f_Q \Big\|_{L^p(X)}. \tag{2.3}$$

The decoupling inequality. We record a special case of the decoupling estimate [19, Theorem 6] by Hänninen–Hytönen. These decoupling estimates originate from McConnell [40], but see also Hytönen [29].

Let $\mathcal{D}$ be a dyadic lattice in $\mathbb{R}^d$ and $Q \in \mathcal{D}$. Let $\mathcal{V}_Q$ be the probability measure space $\mathcal{V}_Q = (Q, \mathrm{Leb}(Q), |Q|^{-1}\,\mathrm{d}x\lfloor Q)$, where $\mathrm{Leb}(Q)$ is the set of Lebesgue measurable subsets of Q and $|Q|^{-1}\,\mathrm{d}x\lfloor Q$ is the normalized Lebesgue measure restricted to Q. Define the product probability space $\mathcal{V} = \prod_{Q\in\mathcal{D}} \mathcal{V}_Q$, and let ν be the related measure. If $y \in \mathcal{V}$, we denote the coordinate related to $Q \in \mathcal{D}$ by y_Q.

Suppose X is a UMD space, $p \in (1,\infty)$ and $f \in L^p(X)$. Let $k \in \{0,1,2,\dots\}$ and $j \in \{0,\dots,k\}$. Define $\mathcal{D}_{j,k} \subset \mathcal{D}$ by

$$\mathcal{D}_{j,k} = \{Q \in \mathcal{D} \colon \ell(Q) = 2^{m(k+1)+j} \text{ for some } m \in \mathbb{Z}\}. \tag{2.4}$$

[19, Theorem 6] implies that

$$\int_{\mathbb{R}^d} \Big| \sum_{Q\in\mathcal{D}_{j,k}} \Delta_Q^l f(x)\Big|_X^p \,\mathrm{d}x \sim \mathbb{E}\int_{\mathbb{R}^d}\int_{\mathcal{V}} \Big| \sum_{Q\in\mathcal{D}_{j,k}} \varepsilon_Q 1_Q(x) \Delta_Q^l f(y_Q)\Big|_X^p \,\mathrm{d}\nu(y)\,\mathrm{d}x \tag{2.5}$$

for any $l \in \{0,1,\dots,k\}$. The point of dividing to the subcollections $\mathcal{D}_{j,k}$ is that now $\Delta_Q^l f$ is constant on every $Q' \in \mathcal{D}_{j,k}$ such that $Q' \subsetneq Q$, which is required by the decoupling theorem (together with the fact that $\int \Delta_Q^l f = 0$ and $\operatorname{spt} \Delta_Q^l f \subset Q$).

2.4. **Multilinear singular integrals and model operators.**

A function

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

is called an n-linear basic kernel if for some $\alpha \in (0,1]$ and $C_K < \infty$ it holds that

$$|K(x)| \le \frac{C_K}{\Big(\sum_{m=2}^{n+1} |x_1 - x_m|\Big)^{dn}},$$

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

$$|K(x) - K(x')| \le C_K \frac{|x_j - x_j'|^\alpha}{\Big(\sum_{m=2}^{n+1} |x_1 - x_m|\Big)^{dn+\alpha}}$$

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

$$|x_j - x_j'| \leq 2^{-1} \max_{2 \leq m \leq n+1} |x_1 - x_m|.$$

The best constant C_K is called $\|K\|_{\mathrm{CZ}_\alpha}$.

An n-linear operator T defined on a suitable class of functions (e.g. on the linear combinations of cubes) is an n-linear *singular integral operator (SIO)* with an associated kernel K, if we have

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

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

We say that T is an n-linear *Calderón–Zygmund operator (CZO)* if the following conditions hold:

- T is an n-linear SIO.
- We have that for all $m \in \{0, \dots, n\}$ there holds that

$$\|T^{m*}(1, \dots, 1)\|_{\mathrm{BMO}} := \sup_{\mathcal{D}} \sup_{K_0 \in \mathcal{D}} \Big(\frac{1}{|K_0|} \sum_{\substack{K \in \mathcal{D} \\ K \subset K_0}} |\langle T^{m*}(1, \dots, 1), h_K \rangle|^2 \Big)^{1/2} < \infty,$$

  where the first supremum is taken over all dyadic lattices $\mathcal{D}$. Here $T^{0*} := T$, T^{m*} denotes the mth adjoint of T for $m \in \{1, \dots, n\}$, and the pairings $\langle T^{m*}(1, \dots, 1), h_K \rangle$ have a standard $T1$ type definition with the aid of the kernel K.
- We have that

$$\|T\|_{\mathrm{WBP}} := \sup_{\mathcal{D}} \sup_{Q \in \mathcal{D}} |Q|^{-1} |\langle T(1_Q, \dots, 1_Q), 1_Q \rangle| < \infty.$$

An SIO T is a CZO if and only if

$$\|T(f_1, \dots, f_n)\|_{L^{q_{n+1}}(\mathbb{R}^d)} \lesssim \prod_{m=1}^{n} \|f_m\|_{L^{p_m}(\mathbb{R}^d)} \tag{2.6}$$

for *some* (equivalently for all) exponents $p_1, \dots, p_n \in (1, \infty)$, $q_{n+1} \in (1/n, \infty)$ satisfying $\sum_{m=1}^{n} 1/p_m = 1/q_{n+1}$. While such a $T1$ theorem is well-known (see e.g. [9, 18, 39]), we will need a very precise version of this called a dyadic representation theorem. To this end, we need some definitions.

Let $k = (k_1, \dots, k_{n+1})$, $0 \leq k_i \in \mathbb{Z}$, and let $\mathcal{D}$ be a dyadic lattice in $\mathbb{R}^d$. An operator $S = S_{\mathcal{D}}^k$ is called an n-linear dyadic shift if it has the form

$$S(f_1, \dots, f_n) = \sum_{K \in \mathcal{D}} A_K(f_1, \dots, f_n), \tag{2.7}$$

where

$$A_K(f_1, \dots, f_n) = \sum_{\substack{Q_1, \dots, Q_{n+1} \in \mathcal{D} \\ Q_j^{(k_j)} = K}} a_{K,(Q_j)} \prod_{j=1}^{n} \langle f_j, \widetilde{h}_{Q_j} \rangle \widetilde{h}_{Q_{n+1}}.$$

Here $a_{K,(Q_j)} = a_{K,Q_1,\dots,Q_{n+1}}$ is a scalar satisfying the normalization

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

and there exist two indices $j_0, j_1 \in \{1,\dots,n+1\}$, $j_0 \neq j_1$, so that $\widetilde{h}_{Q_{j_0}} = h_{Q_{j_0}}$, $\widetilde{h}_{Q_{j_1}} = h_{Q_{j_1}}$ and $\widetilde{h}_{Q_j} = h^0_{Q_j}$ if $j \notin \{j_0, j_1\}$.

An n-linear dyadic paraproduct $\pi = \pi_{\mathcal{D}}$ also has $n+1$ possible forms, but there is no complexity (the $k = (k_1,\dots,k_{n+1})$) associated to them. One of the forms is

$$\pi(f_1,\dots,f_n) = \sum_{K\in\mathcal{D}} a_K \prod_{j=1}^{n} \langle f_j\rangle_K h_K,$$

where the coefficients satisfy the BMO condition

$$\sup_{K_0\in\mathcal{D}} \Big(\frac{1}{|K_0|}\sum_{\substack{K\in\mathcal{D}\\ K\subset K_0}} |a_K|^2\Big)^{1/2} \le 1. \tag{2.8}$$

This is the paraproduct associated with the tuple $(1_K/|K|,\dots,1_K/|K|,h_K)$, and in the remaining n alternative forms the h_K is in a different position.

We call shifts and paraproducts *dyadic model operators (DMOs)*. Suppose T is an n-linear Calderón-Zygmund operator in $\mathbb{R}^d$ related to a kernel K. If $f_1,\dots,f_{n+1}$ are, say, $L^{n+1}(\mathbb{R}^d)$ functions, then the representation theorem states that

$$\langle T(f_1,\dots,f_n), f_{n+1}\rangle = C_T \mathbb{E}_\omega \sum_{k_1,\dots,k_{n+1}=0}^{\infty} \sum_u 2^{-\max k_i \alpha/2} \langle U^k_{\mathcal{D}_\omega,u}(f_1,\dots,f_n), f_{n+1}\rangle. \tag{2.9}$$

Here

$$|C_T| \lesssim \sum_{m=0}^{n} \|T^{m*}(1,\dots,1)\|_{\mathrm{BMO}} + \|T\|_{\mathrm{WBP}} + \|K\|_{\mathrm{CZ}_\alpha}$$
$$\lesssim \|T\|_{L^{n+1}\times\cdots\times L^{n+1}\to L^{(n+1)/n}} + \|K\|_{\mathrm{CZ}_\alpha},$$

α is the parameter in the Hölder continuity assumptions of the kernel of T, and the sum over u is finite, say, over $u = 1,2,\dots,C(n,d)$. If $\max k_i > 0$, then $U^k_{\mathcal{D}_\omega,u}$ is some dyadic shift $S^k_{\mathcal{D}_\omega}$ of complexity k with respect to the lattice $\mathcal{D}_\omega$. If $\max k_i = 0$, then $U^k_{\mathcal{D}_\omega,u}$ is a shift of complexity zero or a paraproduct. In this sense, a CZO T can be represented using DMOs. For $n = 2$, a proof of this result is given by three of us and Y. Ou in [39]. The n-linear case for general n, which requires certain modifications, is [9, Theorem 6.3]. The reference [9, Theorem 6.3] is a more general theorem involving operator-valued CZOs. We note that the additional assumptions related to the operator-valued setup, such as the RMF assumption, concern only the estimation of the model operators. They are not needed for the above stated structural theorem, which has essentially the same proof in the scalar-valued and operator-valued settings.

As DMOs satisfy L^p estimates in the full expected range of exponents, the $T1$ theorem follows from the representation theorem. Our main task in this paper will be to prove L^p-bounds for the extensions of n-linear DMOs to suitably defined tuples of UMD spaces, which we term UMD Hölder tuples and define in the subsequent section.

3. UMD Hölder tuples and the boundedness of multilinear SIOs

Throughout this section, and the remainder of the article, we make use of the following notational conventions. For $m \in \mathbb{N}$ we write $\mathcal{J}_m := \{1, \dots, m\}$ and denote the set of permutations of $\mathcal{J} \subset \mathcal{J}_m$ by $\Sigma(\mathcal{J})$. We simply write $\Sigma(m)$ in place of $\Sigma(\mathcal{J}_m)$. We say that $p_1, \dots, p_m$ is a Hölder tuple of exponents if

$$1 < p_1, \dots, p_m < \infty, \qquad \sum_{j=1}^{m} \tfrac{1}{p_j} = 1. \tag{3.1}$$

3.1. **UMD Hölder tuples.** The notion of UMD Hölder tuple involves fixing an associative algebra $\mathcal{A}$ over $\mathbb{C}$. We denote the associative operation $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ by the product notation, that is, we write $(e, f) \mapsto ef$. In the abstract definition, we do not find useful for $\mathcal{A}$ itself to be endowed with a topology; on the other hand, we will work with linear subspaces of $\mathcal{A}$ endowed with a Banach norm.

We assume that there exists a subspace $\mathcal{L}^1$ of $\mathcal{A}$ and a linear functional $\tau : \mathcal{L}^1 \to \mathbb{C}$, which we refer to as *trace*.

Given an m-tuple $(X_1, \dots, X_m)$ of Banach subspaces of $\mathcal{A}$, we construct the seminorm

$$|e|_{Y(X_1,\dots,X_m)} = \sup \left\{ \left| \tau \left(e \prod_{\ell=1}^{m} e_{\sigma(\ell)} \right) \right| : \sigma \in \Sigma(m), |e_j|_{X_j} = 1, j = 1, \dots, m \right\} \tag{3.2}$$

on the subspace

$$Y(X_1, \dots, X_m) = \left\{ e \in \mathcal{A} : e \prod_{\ell=1}^{m} e_{\sigma(\ell)} \in \mathcal{L}^1 \, \forall \sigma \in \Sigma(m), \, e_j \in X_j, \, j = 1, \dots, m \right\} \tag{3.3}$$

of $\mathcal{A}$. The next lemma clarifies the intent of definition (3.2): if $|\cdot|_Z$ is a seminorm such that all $(m+1)$-linear forms on $X_1 \times \cdots \times X_m \times Z$ in (3.5) below are bounded, then the Z-seminorm dominates the seminorm $Y(X_1, \dots, X_m)$.

3.4. **Lemma.** *Let $(X_1, \dots, X_m)$ be a m-tuple of Banach subspaces of $\mathcal{A}$. Suppose that $e \in \mathcal{A}$ belongs to the subspace (3.3). Then*

$$\left| \tau \left(e \prod_{\ell=1}^{m} e_{\sigma(\ell)} \right) \right| \leq |e|_Z \prod_{j=1}^{m} |e_j|_{X_j}, \qquad \forall \sigma \in \Sigma(m), \, e_j \in X_j, \, j = 1, \dots, m, \tag{3.5}$$

holds for $|e|_Z = |e|_{Y(X_1,\dots,X_m)}$. In addition, if $|\cdot|_Z$ is a seminorm on $\mathcal{A}$ such that (3.5) holds, $|e|_{Y(X_1,\dots,X_m)} \lesssim |e|_Z$.

Proof. Immediate from the definitions. □

3.6. **Definition** (Admissible spaces)**.** We say that a Banach subspace X of $\mathcal{A}$ is admissible if $Y(X)$ from (3.3) is a Banach space with respect to $|\cdot|_{Y(X)}$ of (3.2)[1], the map

$$y \in Y(X) \mapsto x^*[y] \in X^*, \qquad x^*[y](x) = \tau(yx), \quad x \in X, \tag{3.7}$$

is onto, and furthermore, for each $x \in X$, $y \in Y(X)$, $xy \in \mathcal{L}^1$ and

$$\tau(xy) = \tau(yx). \tag{3.8}$$

[1]This includes that if $y \in Y(X)$ then $|y|_{Y(X)} < \infty$.

3.9. *Remark.* If X is admissible, then the map (3.7) is an isometric bijection from $Y(X)$ onto X^*. We are thus allowed to identify $Y(X)$ with X^* via (3.7) and we do so without explicit mention from now on. Notice that if X is admissible, then X is a UMD space if and only if $Y(X)$ is.

For our purposes, it is convenient to state the next observation in the form of a lemma.

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

Proof. The reflexivity of X and Remark 3.9 imply that $Y(Y(X))$ is isometrically isomorphic with X. Here we want to show that they are actually equal as sets with equal norms. Denote $Y := Y(X)$ and $Z := Y(Y)$. It follows quite directly from the definitions that X is a subset of Z.

Let $\varphi\colon X^* \to Y$ be the isometric isomorphism from the definition of the admissibility of X. This induces the isometric isomorphism $\phi\colon X^{**} \to Y^*$ defined by

$$\phi(x^{**})(y) := x^{**}(\varphi^{-1}(y)),$$

where $x^{**} \in X^{**}$ and $y \in Y$. Since X is reflexive and Y is admissible, we have the canonical isometric isomorphism $\rho\colon X \to X^{**}$ and the isometric isomorphism $\eta\colon Y^* \to Z$. Now, the composition $\eta \circ \phi \circ \rho\colon X \to Z$ is an isometric isomorphism.

Suppose $x \in X$ and denote $z := \eta \circ \phi \circ \rho(x)$. Let $y \in Y$. Then we have that

$$\tau(zy) = \eta^{-1}(z)(y) = \phi^{-1} \circ \eta^{-1}(z)(\varphi^{-1}(y)) = \varphi^{-1}(y)(\rho^{-1} \circ \phi^{-1} \circ \eta^{-1}(z)) = \tau(xy).$$

Since x and z are both elements of Z, the fact that $\tau(zy) = \tau(xy)$ for all $y \in Y$ implies that $x = z$. Thus, the isometric isomorphism $\eta \circ \phi \circ \rho\colon X \to Z$ is actually the identity map. □

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

We turn to defining UMD Hölder m-tuples relatively to $\mathcal{A}, \tau$. We first do so for $m = 2$.

3.11. **Definition** (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 $X_2 = Y(X_1)$. In view of Remark 3.9 and Lemma 3.10 one can equivalently say that $\{X_1, X_2\}$ is a UMD Hölder pair if X_2 is a UMD space and $X_1 = Y(X_2)$.

For $m \geq 3$ the definition of a UMD Hölder m-tuple is given inductively on m as follows.

3.12. **Definition** (UMD Hölder m-tuple, $m \geq 3$)**.** Let $X_1, \dots, X_m$ be admissible spaces. We say that $\{X_1, \dots, X_m\}$ is a UMD *Hölder m-tuple* if the following properties hold.

P1. For all $j_0 \in \mathcal{J}_m$ there holds

$$X_{j_0} = Y\Big(\big\{X_j : j \in \mathcal{J}_m \setminus \{j_0\}\big\}\Big).$$

P2. If $1 \leq k \leq m-2$ and $\mathcal{J} = \{j_1 < j_2 < \cdots < j_k\} \subset \mathcal{J}_m$, then $Y(X_{j_1}, \dots, X_{j_k})$ is an admissible Banach space with the norm (3.2) and

$$\{X_{j_1}, \dots, X_{j_k}, Y(X_{j_1}, \dots, X_{j_k})\} \tag{3.13}$$

is a UMD Hölder $(k+1)$-tuple.

The following remark is an important consequence of the definition.

3.14. *Remark.* Let $m \geq 3$ and $\{X_1, \dots, X_m\}$ 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 3.11 implies that X_{j_0} and $Y(X_{j_0})$ are UMD spaces. The inductive nature of the definition then ensures that each $Y(X_{j_1}, \dots, X_{j_k})$ appearing in (3.13) is a UMD space.

3.15. *Remark.* Let $m \geq 2$ and $\{X_1, \dots, X_m\}$ be a UMD m-Hölder tuple. Let $e_j \in X_j$ for $j \in \mathcal{J}_m$. For each $\sigma \in \Sigma(m)$, as $X_{\sigma(1)} = Y(X_{\sigma(2)}, \dots, X_{\sigma(m)})$, we necessarily have $\prod_{j=1}^m e_{\sigma(j)} \in \mathcal{L}^1$ and

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

We clarify the extent of our definition with some examples of UMD Hölder tuples.

3.16. **Example.** It is immediate to verify that the m-tuple $X_j = \mathbb{C}$, $j = 1, \dots, m$, is a UMD Hölder m-tuple with respect to the usual product.

The next example is of relevance if one wants to deduce Theorem 1.1 in the basic case $X_1 = Y_3 = X$ and $X_2 = \mathbb{C}$ from Theorem 3.31. However, otherwise we do not need it, and Theorem 1.1 is best seen mimicking our main proofs.

3.17. **Example.** Let $X = X_1$ be a complex UMD space and denote $X_2 = X^*$. The goal of this example is to show that for each $m \geq 2$ the tuple $\{X_1, X_2, \dots, X_m\}$ with $X_j = \mathbb{C}$ for $2 < j \leq m$ is a UMD Hölder tuple. This is conceptually simple but requires some work in order to define a suitable enveloping algebra $\mathcal{A}$. We let $V = X \oplus X^*$, and define $\mathcal{A}$ to be the tensor algebra over V, namely

$$\mathcal{A} = \bigoplus_{k=0}^{\infty} V^{\otimes k}.$$

We let

$$\mathcal{L}^1 = \operatorname{span}\{e \otimes e^* + f^* \otimes f,\ e, f \in X,\ e^*, f^* \in X^*\};$$

notice that this is a linear subspace of $V^{\otimes 2}$. We then define the functional τ by

$$\tau\Big(e \otimes e^* + f^* \otimes f\Big) = \langle f^*, e\rangle + \langle e^*, f\rangle$$

for $e, f \in X$, $e^*, f^* \in X^*$ and extend it to all of $\mathcal{L}^1$ by linearity. We notice that the definition (3.3) yields that

$$Y(X_{j_1}, \dots, X_{j_k}) = \begin{cases} X & 1 \notin \{j_1 \dots, j_k\},\ 2 \in \{j_1 \dots, j_k\}, \\ X^* & 1 \in \{j_1 \dots, j_k\},\ 2 \notin \{j_1 \dots, j_k\}, \\ \mathbb{C} & \{1,2\} \subset \{j_1 \dots, j_k\} \text{ or } \{1,2\} \cap \{j_1 \dots, j_k\} = \varnothing. \end{cases}$$

With this information in hand, we learn that $X, X^*, \mathbb{C}$ are admissible spaces. Proceeding by induction on m, we then easily verify that $\{X_1, X_2, \dots, X_m\}$ is a UMD Hölder tuple.

We now start explaining how non-commutative L^p spaces fit our abstract framework.

3.18. **Example.** Consider a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(H)$, namely a self-adjoint unital subalgebra of the algebra of bounded linear operators on a complex Hilbert space H which is closed in the weak operator topology [45, 46]. Let $\mathcal{M}_+ = \{A \in \mathcal{M} : \langle Ah, h\rangle \geq 0\ \forall h \in H\}$ denote the positive part of $\mathcal{M}$. A *trace* τ is a functional $\mathcal{M}_+ \to [0, \infty]$ satisfying

$$\tau(A + \lambda B) = \tau(A) + \lambda\tau(B), \qquad \forall A, B \in \mathcal{M}_+, \lambda > 0$$

as well as the tracial property

$$\tau(AA^*) = \tau(A^*A)$$

for all $A \in \mathcal{M}$. Following [46], we assume τ is *normal, semifinite, faithful* (n.s.f.) and define the corresponding space of *measurable operators* $\mathcal{A} = L^0(\mathcal{M})$ equipped with convergence in measure: a detailed definition is in [46]. Then $\mathcal{A}$ is a (metrizable) topological $*$-algebra and $\mathcal{M}$ is dense in $\mathcal{A}$. We will also recall the notion of $\mathcal{S}_+, \mathcal{S}$ as introduced in [46, p.1463]: $\mathcal{S}_+$ is the cone of those $A \in \mathcal{M}_+$ such that $\tau(\operatorname{supp} A) < \infty$, where $\operatorname{supp} A$ is the least projection $P \in \mathcal{M}_+$ with $PA = A$, and $\mathcal{S} \subset \mathcal{M}$ is the linear span of $\mathcal{S}_+$. We note [48, Proposition 1.15(ii)] that τ may be extended to a unique linear functional on $\mathcal{S}$, satisfying

$$\tau(A^*) = \overline{\tau(A)}, \qquad \tau(AB) = \tau(BA), \qquad \forall A, B \in \mathcal{S}. \tag{3.19}$$

For $1 \leq p < \infty$, we call *noncommutative L^p space* the Banach subspace of $\mathcal{A}$ obtained by completion of $\mathcal{S}$ with respect to the norm

$$\|A\|_{L^p(\mathcal{M})} = \left[\tau\left(\left(A^\star A\right)^{\frac{p}{2}}\right)\right]^{\frac{1}{p}}, \qquad 1 \leq p < \infty.$$

In fact, we record the characterization

$$L^p(\mathcal{M}) = \left\{A \in \mathcal{A} : \tau\left((A^\star A)^{\frac{p}{2}}\right) < \infty\right\};$$

in the above equality, τ denotes the extension of the trace to the positive part of $\mathcal{A}$ defined via generalized singular numbers [46]. We also point out the Hölder inequality

$$\|\xi_1\xi_2\|_{L^p(\mathcal{M})} \leq \|\xi_1\|_{L^{p_1}(\mathcal{M})}\|\xi_2\|_{L^{p_2}(\mathcal{M})}, \qquad \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$$

valid whenever $1 \leq p_1, p_2, p < \infty$. A suitable substitute holds for $p = \infty$ if the $L^p(\mathcal{M})$-norm is replaced by the $\mathcal{B}(H)$-norm. Furthermore, notice that τ may be extended from $\mathcal{S}$ to a unique linear bounded functional on $L^1(\mathcal{M})$ satisfying

$$|\tau(A)| \leq \|A\|_{L^1(\mathcal{M})}.$$

The tracial property (3.19) extends to the following: if $A, B \in \mathcal{A}$ are such that $A \in L^p(\mathcal{M})$ and $B \in L^{p'}(\mathcal{M})$, then

$$\tau(AB) = \tau(BA). \tag{3.20}$$

This is the concrete equivalent of property (3.8) we assumed in the abstract setup. We refer to [48, Rem. 1.2.11] for the details of (3.20).

For $1 < p < \infty$, we then have $L^p(\mathcal{M})^* = L^{p'}(\mathcal{M})$ with isometric isomorphism given by the Riesz representation map

$$\lambda \in L^p(\mathcal{M})^* \mapsto B_\lambda \in L^{p'}(\mathcal{M}), \qquad \lambda(A) = \tau(B_\lambda A) \quad \forall A \in L^p(\mathcal{M}).$$

A fortiori, $L^p(\mathcal{M})$ is reflexive for $1 < p < \infty$. For our purposes, it is also important to observe that $L^p(\mathcal{M})$ is a UMD space in the same range [46, Corollary 7.7]. We detail below two concrete examples of von Neumann algebras equipped with a n.s.f. trace.

If $\mathcal{M}$ is an abelian von Neumann algebra, then $\mathcal{M} = L^\infty(M, \mu)$ for some measure space (M, μ), a n.s.f. trace is obtained by integration with respect to the measure μ, and $\mathcal{A} = L^0(M, \mu)$, the topological $*$-algebra of measurable functions on M with respect to convergence in measure. Then $L^p(\mathcal{M}) = L^p(M, \mu)$ for $1 \leq p < \infty$.

If $\mathcal{M} = \mathcal{B}(H)$, the bounded linear operators over a separable Hilbert space H and

$$\tau(A) = \sum_{j=1}^{\infty} \langle Ae_i, e_i \rangle$$

where e_i is any orthonormal basis of H [46, Example (ii), p. 1465], then the spaces $L^p(\mathcal{M})$ are referred to as *Schatten-von Neumann classes* and denoted by S^p.

Let now p_j, $j = 1, \dots, m$ be a Hölder tuple as in (3.1). We claim that $X_j = L^{p_j}(\mathcal{M})$ is a UMD Hölder tuple relative to the algebra $\mathcal{A} = L^0(\mathcal{M})$, with trace τ. This can be proved by induction on m, relying on the equality

$$L^{p(\mathcal{J})}(\mathcal{M}) = Y(\{L^{p_j}(\mathcal{M}) : j \in \mathcal{J}\}), \qquad \frac{1}{p(\mathcal{J})} = 1 - \sum_{j \in \mathcal{J}} \frac{1}{p_j}$$

valid for each $\varnothing \subsetneq \mathcal{J} \subsetneq \mathcal{J}_m$, whose verification is immediate and left to the reader.

3.21. **Example.** In Appendix A, we prove that if $\{p_j^s : 1 \le j \le m\}$ are Hölder tuples of exponents as in (3.1) for $s = 0, \dots, S$, $\mathcal{M}$ is a von Neumann algebra with n.s.f. trace τ as in Example 3.18, and (M_s, μ_s) are σ-finite Borel measure spaces for $s = 1, \dots, S$, the tuple of spaces

$$X_j = L^{p_j^S}(M_S, \mu_S; L^{p_j^{S-1}}(M_{S-1}, \mu_{S-1}; \cdots L^{p_j^1}(M_1, \mu_1; L^{p_j^0}(\mathcal{M})) \cdots)$$

is a UMD Hölder m-tuple relative to the trace

$$f \mapsto \int_{M_1 \times \cdots \times M_S} \tau(f(t_1, \dots, t_S))\, \mathrm{d}\mu_1 \times \cdots \times \mu_S(t_1, \dots, t_S).$$

A precise statement is provided in Proposition A.1.

3.2. **Extensions of CZOs.** If X is a Banach space we will use the notation $L_c^\infty \otimes X$ for functions of the type $\sum_{i=1}^N f_i e_i$, where $N \in \mathbb{N}$, $f_i \in L_c^\infty(\mathbb{R}^d) =: L_c^\infty$ and $e_i \in X$.

Let $\{X_1, \dots, X_{n+1}\}$ be a UMD Hölder tuple where $n \ge 1$. Suppose T_0 is an n-linear CZO with a kernel K_0 as defined in Section 2.4. Since we know that T_0 is a bounded operator, see (2.6), we know that $\langle T_0(f_1, \dots, f_n), f_{n+1} \rangle$ makes sense for $f_j \in L_c^\infty$. We define the corresponding $(n+1)$-linear form

$$\begin{aligned} &\Lambda_{T_0} : L_c^\infty \otimes X_1 \times \cdots \times L_c^\infty \otimes X_{n+1} \to \mathbb{C}, \\ &\Lambda_{T_0}(f_1, \dots, f_{n+1}) = \sum_{a_1, \dots, a_{n+1}} \langle T_0(f_{1,a_1}, \dots, f_{n,a_n}), f_{n+1,a_{n+1}} \rangle \tau\Big(\prod_{j=1}^{n+1} e_{j,a_j} \Big), \end{aligned} \tag{3.22}$$

where $f_j = \sum_{a=1}^{N_j} f_{j,a} e_{j,a}$. If U is a dyadic model operator as in Section 2.4 we define the form Λ_U in the corresponding way. We can also make sense of Λ_U more directly. For example, if U is a dyadic shift as in (2.7), then

$$\Lambda_U(f_1, \dots, f_{n+1}) = \sum_{K \in \mathcal{D}} \sum_{\substack{Q_1, \dots, Q_{n+1} \in \mathcal{D} \\ Q_j^{(k_j)} = K}} a_{K,(Q_j)} \tau\Big(\prod_{j=1}^{n+1} \langle f_j, \widetilde{h}_{Q_j} \rangle \Big). \tag{3.23}$$

3.24. *Remark.* We chose to utilize the identity permutation in $\Sigma(n+1)$ for the product appearing in (3.22). However, the notion of being a UMD Hölder tuple is clearly invariant under reordering of $\{X_1, \dots, X_{n+1}\}$.

Let $p_j \in (1,\infty)$ for $j \in \mathcal{J}_{n+1}$ be such that $\sum_{j=1}^{n+1} 1/p_j = 1$. From Theorem 3.31 it will follow among other things that

$$|\Lambda_{T_0}(f_1, \dots, f_{n+1})| \lesssim \prod_{j=1}^{n+1} \|f_j\|_{L^{p_j}(X_j)}. \tag{3.25}$$

Based on this boundedness one can define as usual $n+1$ adjoint operators. Let us describe how the adjoints look like in our Hölder tuple set up.

Fix $j_0 \in \mathcal{J}_{n+1}$ and $f_j \in L^{p_j}(X_j)$ for $j \in \mathcal{J}_{n+1} \setminus \{j_0\}$. Consider the linear functional

$$f_{j_0} \in L^{p_{j_0}}(X_{j_0}) \mapsto \Lambda_{T_0}(f_1, \dots, f_{n+1}), \tag{3.26}$$

which is bounded because of (3.25). Recall that $L^{p_{j_0}}(X_{j_0})^*$ is identified with $L^{(p_{j_0})'}(Y(X_{j_0}))$ with duality pairing

$$\langle g, f_{j_0} \rangle = \int_{\mathbb{R}^d} \tau(g(x) f_{j_0}(x))\, \mathrm{d}x.$$

Therefore, there exists a function

$$T^{*j_0}(f_j : j \in \mathcal{J}_{n+1} \setminus \{j_0\}) := T^{j_0 *}(f_1, \dots, f_{j_0-1}, f_{j_0+1}, \dots, f_{n+1}) \in L^{(p_{j_0})'}(Y(X_{j_0}))$$

so that

$$\Lambda_{T_0}(f_1, \dots, f_{n+1}) = \int_{\mathbb{R}^d} \tau(T^{*j_0}(f_j : j \in \mathcal{J}_{n+1} \setminus \{j_0\})(x) f_{j_0}(x))\, \mathrm{d}x.$$

The n-linear bounded operator

$$T^{j_0 *} \colon L^{p_1}(X_1) \times \cdots \times L^{p_{j_0-1}}(X_{j_0-1}) \times L^{p_{j_0+1}}(X_{j_0+1}) \times \cdots \times L^{p_{n+1}}(X_{n+1}) \to L^{(p_{j_0})'}(Y(X_{j_0}))$$

is one of the adjoint operators. In the same way one can define the adjoint $T_0^{j_*}$ of T_0 so that

$$\langle T_0^{j_0 *}(g_1, \dots, g_{j_0-1}, g_{j_0+1}, \dots, g_{n+1}), g_{j_0} \rangle = \langle T_0(g_1, \dots, g_n), g_{n+1} \rangle,$$

where $g_j \in L^{p_j}$.

Suppose $f_j = \sum_{a=1}^{N_j} f_{j,a} e_{j,a} \in L^\infty_c \otimes X_j$ for $j \in \mathcal{J}_{n+1} \setminus \{j_0\}$. A calculation involving the invariance of τ under cyclic permutations yields that

$$\begin{aligned} &T^{*j_0}(f_j : j \in \mathcal{J}_{n+1} \setminus \{j_0\}) \\ &\quad = \sum T_0^{j_0 *}(f_{j,a_j} : j \in \mathcal{J}_{n+1} \setminus \{j_0\}) e_{j_0+1,a_{j_0+1}} \cdots e_{n+1,a_{n+1}} \cdots e_{j_0-1,a_{j_0-1}}. \end{aligned}$$

3.3. **Sparse domination of dyadic operators.** The following basic sparse domination result, Lemma 3.27, was first proved by Culiuc, Ou and one of us in the linear scalar-valued setting in [6, 7] and recast by Y. Ou and three of us in the multilinear scalar-valued case [39]. The proof in our current Banach-valued setting is completely analogous.

Let $\eta \in (0,1)$. We say that a collection $\mathcal{S}$ of cubes in $\mathbb{R}^d$ (not necessarily dyadic) is η-sparse (or just sparse) if for every $Q \in \mathcal{S}$ there exists a set $E_Q \subset Q$ with $|E_Q| > \eta |Q|$ so that the sets E_Q, $Q \in \mathcal{S}$, are pairwise disjoint.

3.27. **Lemma.** *Let $n \geq 1$, $\{X_1, \dots, X_{n+1}\}$ be a* UMD *Hölder tuple, $\mathcal{D}$ be a dyadic grid, $k = (k_1, \dots, k_{n+1})$, $0 \leq k_i \in \mathbb{Z}$. Suppose that the scalars $a_{K,(Q_j)}$ satisfy the normalization*

$$|a_{K,(Q_j)}| \leq A_1 \prod_{j=1}^{n+1} |Q_j|^{1/2} |K|^{-n}$$

and we are given scalar functions $u_{j,Q} = \sum_{Q' \in \mathrm{ch}(Q)} c_{j,Q'} 1_{Q'}$ satisfying $|u_{j,Q}| \leq |Q|^{-1/2}$.

If there exists a Hölder tuple $p_1, \dots, p_{n+1}$ as in (3.1) such that the forms

$$U_{\mathcal{D}'}(g_1, \dots, g_{n+1}) := \sum_{K \in \mathcal{D}'} \sum_{\substack{Q_1, \dots, Q_{n+1} \in \mathcal{D} \\ Q_i^{(k_i)} = K}} a_{K,(Q_i)} \tau\Big(\prod_{j=1}^{n+1} \langle g_j, u_{j,Q_j} \rangle \Big), \qquad \mathcal{D}' \subset \mathcal{D},$$

satisfy

$$\sup_{\mathcal{D}' \subset \mathcal{D}} |U_{\mathcal{D}'}(g_1, \dots, g_{n+1})| \leq A_2 \prod_{j=1}^{n+1} \|g_j\|_{L^{p_j}(\mathbb{R}^d; X_j)}, \qquad g_j \in L^\infty_c(\mathbb{R}^d; X_j),\ j = 1, \dots, n+1,$$

then for each tuple $f_j \in L^\infty_c(X_j)$, $j = 1, \dots, n+1$, and $\eta > 0$ there exists an η-sparse collection $\mathcal{S} = \mathcal{S}((f_j), \eta) \subset \mathcal{D}$ such that

$$|\langle U_{\mathcal{D}}(f_1, \dots, f_n), f_{n+1} \rangle| \lesssim_\eta (A_1 + A_1 \kappa + A_2) \sum_{Q \in \mathcal{S}} |Q| \prod_{j=1}^{n+1} \langle |f_j|_{X_j} \rangle_Q,$$

where $\kappa = \max k_m$.

In the previous lemma the sparse collection is in the same grid where the dyadic operator is defined. The result can be updated to involve a universal sparse set, which is explained in Remark 3.28. This is important when we move the sparse estimate from DMOs to CZOs via the representation theorem, which involves a family of dyadic grids.

3.28. *Remark.* There exist dyadic grids $\mathcal{D}_i$, $i = 1, \dots, 3^d$, with the following property, see Lacey–Mena [36], [39], or [8] for a simple proof. Let $g_m \in L^1_{\mathrm{loc}}$, $m = 1, \dots, n+1$, be scalar-valued and let $\eta_1, \eta_2 \in (0, 1)$. Then for some i there exists an η_2-sparse collection $\mathcal{U} = \mathcal{U}((g_m), \eta_2) \subset \mathcal{D}_i$, so that for all η_1-sparse collections of cubes $\mathcal{S}$ we have

$$\sum_{Q \in \mathcal{S}} |Q| \prod_{m=1}^{n+1} \langle |g_m| \rangle_Q \lesssim_{\eta_1, \eta_2} \sum_{Q \in \mathcal{U}} |Q| \prod_{m=1}^{n+1} \langle |g_m| \rangle_Q.$$

3.29. *Remark.* In [8], it is noted that the sparse domination estimate for an $n+1$-linear form Λ on $\mathbb{R}^d$, acting on scalar functions

$$|\Lambda(f_1, \dots, f_{n+1})| \lesssim \sum_{Q \in \mathcal{S}} |Q| \prod_{j=1}^{n+1} \langle |f_j| \rangle_Q,$$

is equivalent to the estimate in terms of the multilinear maximal operator M

$$|\Lambda(f_1, \dots, f_{n+1})| \lesssim \|\mathrm{M}(f_1, \dots, f_{n+1})\|_1, \qquad \mathrm{M}(f_1, \dots, f_{n+1})(x) = \sup_{x \in Q} \prod_{j=1}^{n+1} \langle |f_j| \rangle_Q.$$

Vector-valued versions of this principle may be formulated in a totally analogous way. We have used this equivalence to state the sparse bounds in our main results; this is particularly convenient as the formulation in terms of the multilinear maximal function may be given without defining what a sparse collection is.

Next, we discuss the well known fact that the sparse domination of an operator implies boundedness in the full range: for more details and weighted corollaries see [8, 39] and references therein.

Let $X_1, \dots, X_{n+1}$ be Banach spaces, $n \geq 1$. Assume that Λ is an $(n+1)$-linear form initially defined on $L^\infty_c(\mathbb{R}^d) \otimes X_1 \times \cdots \times L^\infty_c(\mathbb{R}^d) \otimes X_{n+1}$ such that if $f_j \in L^\infty_c(\mathbb{R}^d) \otimes X_j$, then there exists a dyadic lattice $\mathcal{D}$ and a sparse collection $\mathcal{S} \subset \mathcal{D}$ so that

$$|\Lambda(f_1, \dots, f_{n+1})| \lesssim \sum_{Q \in \mathcal{S}} |Q| \prod_{j=1}^{n+1} \langle |f_j|_{X_j} \rangle_Q. \tag{3.30}$$

This easily implies that if $p_j \in (1, \infty)$ for $j \in \mathcal{J}_{n+1}$ are such that $\sum_{j=1}^{n+1} 1/p_j = 1$ then Λ can be extended to a bounded form $\Lambda \colon L^{p_1}(X_1) \times \cdots \times L^{p_{n+1}}(X_{n+1}) \to \mathbb{C}$. Indeed, just use Hölder's inequality and then Carleson embedding theorem in the right hand side of (3.30).

We estimate the adjoints T^{j*} of Λ, which are defined in the usual way based on the functional as in (3.26). By symmetry it will suffice to tackle the case $j = n+1$ and simply write T in place of $T^{(n+1)*}$.

We use the so-called A_∞ extrapolation from Cruz-Uribe–Martell–Pérez [5]. Let $A_\infty(\mathbb{R}^d)$ be the class of A_∞ weights in $\mathbb{R}^d$, see [5] for a definition. Suppose $v \in A_\infty(\mathbb{R}^d)$ and $f_j \in L^\infty_c(X_j)$ for $j \in \mathcal{J}_n$. Taking $f_{n+1}(x) = \xi(x) v(x)$ for a suitably chosen $\xi \in L^\infty_c(X_{n+1})$ there holds that

$$\begin{aligned}
\int_{\mathbb{R}^d} |T(f_1, \dots, f_n)|_{X^*_{n+1}} v \sim \Lambda(f_1, \dots, f_{n+1}) &\lesssim \sum_{Q \in \mathcal{S}} \prod_{j=1}^{n} \langle |f_j|_{X_j} \rangle_Q v(Q) \\
&\leq \sum_{Q \in \mathcal{S}} \Big(\Big\langle M^n_{\mathcal{D}}(|f_1|_{X_1}, \dots, |f_n|_{X_n})^{1/2} \Big\rangle^v_Q \Big)^2 v(Q) \\
&\lesssim \int_{\mathbb{R}^d} M^n_{\mathcal{D}}(|f_1|_{X_1}, \dots, |f_n|_{X_n}) v,
\end{aligned}$$

where $\langle h \rangle^v_Q = v(Q)^{-1} \int_Q hv$ and $M^n_{\mathcal{D}}(g_1, \dots, g_n) := \sup_{Q \in \mathcal{D}} \prod_{m=1}^n \langle |g_m| \rangle_Q 1_Q$ is the dyadic maximal function and in the last step we used the Carleson embedding theorem. Now, the A_∞ extrapolation result, Theorem 2.1 in [5], gives that

$$\int_{\mathbb{R}^d} |T(f_1, \dots, f_n)|^p_{X^*_{n+1}} v \lesssim \int_{\mathbb{R}^d} M^n_{\mathcal{D}}(|f_1|_{X_1}, \dots, |f_n|_{X_n})^p v$$

for all $p \in (0, \infty)$ and $v \in A_\infty(\mathbb{R}^d)$. Using this with $v = 1$ the boundedness of the maximal function gives that

$$\|T(f_1, \dots, f_n)\|_{L^{q_{n+1}}(X^*_{n+1})} \lesssim \prod_{j=1}^{n} \|f_j\|_{L^{p_j}(X_j)},$$

where $p_j \in (1, \infty]$ are such that $1/q_{n+1} := \sum_{j=1}^n 1/p_j > 0$. Notice that the boundedness of $M^n_{\mathcal{D}}$ follows from Hölder's inequality and the boundedness of $M^1_{\mathcal{D}}$, since there holds

that $M^n_{\mathcal{D}}(g_1,\dots,g_n) \le \prod_{m=1}^n M^1_{\mathcal{D}}(g_m)$. As is clear, multilinear weighted bounds also follow from this argument and the corresponding results of $M^n_{\mathcal{D}}$.

3.4. **Proof of the main theorem.** In this section we state and prove our main theorem assuming the estimates for model operators from Section 4 and Section 5.

3.31. **Theorem.** *Let $n \ge 1$, T_0 be an n-linear CZO with kernel K_0 and $\{X_1,\dots,X_{n+1}\}$ be a* UMD *Hölder tuple. The $(n+1)$-linear form Λ_{T_0} defined in (3.22) can be extended to act on functions $f_j \in L^\infty_c(X_j)$, and given $\eta \in (0,1)$ there exists an η-sparse collection of cubes $\mathcal{S} = \mathcal{S}((f_m),\eta)$ so that*

$$|\Lambda_{T_0}(f_1,\dots,f_{n+1})| \lesssim_\eta \Big[\|K_0\|_{\mathrm{CZ}_\alpha} + \|T_0\|_{\mathrm{WBP}} + \sum_{j=1}^{n+1} \|(T_0)^{j*}(1,\dots,1)\|_{\mathrm{BMO}}\Big] \times \sum_{Q\in\mathcal{S}} |Q| \prod_{j=1}^{n+1} \langle |f_j|_{X_j}\rangle_Q.$$

Consequently, we for instance have

$$\|T_0(f_1,\dots,f_n)\|_{L^{q_{n+1}}(X^*_{n+1})} \lesssim \prod_{j=1}^n \|f_j\|_{L^{p_j}(X_j)}$$

whenever $p_j \in (1,\infty]$ are such that $1/q_{n+1} := \sum_{j=1}^n 1/p_j > 0$. See Section 3.3 for a full discussion of the corollaries of the sparse estimate.

Proof. Let $f_j \in L^\infty_c \otimes X_j$ for $j \in \mathcal{J}_{n+1}$ be of the form $f_j = \sum_{a=1}^{N_j} f_{j,a} e_{j,a}$. Then, we have by the dyadic representation (2.9) that

$$\begin{aligned} &\Lambda_{T_0}(f_1,\dots,f_{n+1}) \\ &= C_T \sum_{a_1,\dots,a_{n+1}} \mathbb{E}_\omega \sum_{k_1,\dots,k_{n+1}=0}^\infty \sum_u 2^{-\frac{\alpha\max k_i}{2}} \langle U^k_{\mathcal{D}_\omega,u}(f_{1,a_1},\dots,f_{n,a_n}), f_{n+1,a_{n+1}}\rangle \tau\Big(\prod_{j=1}^{n+1} e_{j,a_j}\Big) \\ &= C_T \mathbb{E}_\omega \sum_{k_1,\dots,k_{n+1}=0}^\infty \sum_u 2^{-\frac{\alpha\max k_i}{2}} \Lambda_{U^k_{\mathcal{D}_\omega,u}}(f_1,\dots,f_{n+1}). \end{aligned} \tag{3.32}$$

In Section 4 and Section 5 it is shown that if U is a dyadic model operator then

$$|\Lambda_U(g_1,\dots,g_{n+1})| \lesssim \prod_{j=1}^{n+1} \|g_j\|_{L^{p_j}(X_j)} \tag{3.33}$$

holds for any $p_j \in (1,\infty)$ and $g_j \in L^\infty_c(X_j)$, $j \in \mathcal{J}_{n+1}$, such that $\sum_{j=1}^{n+1} 1/p_j = 1$; if U is a shift, then the estimate depends polynomially on the complexity. This implies that Λ_{T_0} can be extended to act on functions $f_j \in L^\infty_c(X_j)$ and that (3.32) holds for such functions.

The estimate (3.33) implies via Lemma 3.27 and Remark 3.28 that if $f_j \in L^\infty_c(X_j)$ for $j \in \mathcal{J}_{n+1}$ then there exist a dyadic grid and an η-sparse collection $\mathcal{S} = \mathcal{S}((f_j)) \subset \mathcal{D}$ so that all the model operators appearing in (3.32) satisfy

$$|\Lambda_{U^k_{\mathcal{D}_\omega,u}}(f_1,\dots,f_{n+1})| \lesssim \sum_{Q\in\mathcal{S}} |Q| \prod_{j=1}^{n+1} \langle |f_j|_{X_j}\rangle_Q,$$

where the estimate depends polynomially on the complexity. This combined with (3.32) finishes the proof. □

In Section 6, we show that the UMD Hölder tuples enjoy a suitable maximal property among tuples of spaces admitting L^p-bounded extensions of n-linear CZO operators and dyadic shifts.

4. Boundedness of multilinear shifts in UMD Hölder tuples

This section is dedicated to the proof of the boundedness of multilinear shifts. Before starting the main argument, we record a randomized bound for UMD Hölder tuples in the following lemma.

4.1. **Lemma.** *Let $\{X_1,\dots,X_{n+1}\}$ be a* UMD *Hölder tuple, $n\ge 2$, and let $K\in\mathbb{Z}_+$. For each $k=1,\dots,K$ let a_k be a scalar such that $|a_k|\le 1$ and for each $j\in\mathcal{J}_n$ assume that we are given $e_{j,k}\in X_j$. Then we have*

$$\Big|\sum_{k=1}^K a_k\prod_{j=1}^n e_{j,k}\Big|_{Y(X_{n+1})}\le\prod_{j=1}^n\|(e_{j,k})_{k=1}^K\|_{\mathrm{Rad}(X_j)}.$$

Proof. Fix K, $|a_k|\le 1$ and $e_{j,k}\in X_j$ as in the assumptions. Let $\{\varepsilon_k^i\}_{k=1}^K$, $i\in\mathcal{J}_{n-1}$, be collections of independent random signs. We denote the expectation with respect to the random variables $\{\varepsilon_k^i\}_{k=1}^K$ by $\mathbb{E}^i$, and write $\mathbb{E}=\mathbb{E}^1\cdots\mathbb{E}^{n-1}$. We have the identity

$$\begin{aligned}\sum_{k=1}^K a_k\prod_{j=1}^n e_{j,k}&=\mathbb{E}\sum_{k_1,\dots,k_n=1}^K\varepsilon_{k_1}^1\varepsilon_{k_2}^1\varepsilon_{k_2}^2\varepsilon_{k_3}^2\cdots\varepsilon_{k_{n-1}}^{n-1}\varepsilon_{k_n}^{n-1}a_{k_1}\prod_{j=1}^n e_{j,k_j}\\&=\mathbb{E}\Big(\sum_{k_1=1}^K\varepsilon_{k_1}^1a_{k_1}e_{1,k_1}\Big)\Big(\sum_{k_2=1}^K\varepsilon_{k_2}^1\varepsilon_{k_2}^2e_{2,k_2}\Big)\cdots\Big(\sum_{k_n=1}^K\varepsilon_{k_n}^{n-1}e_{n,k_n}\Big).\end{aligned}$$

We can dominate this with

$$\mathbb{E}\Big\|\sum_{k_1=1}^K\varepsilon_{k_1}^1a_{k_1}e_{1,k_1}\Big\|_{X_1}\Big(\prod_{i=2}^{n-1}\Big\|\sum_{k_i=1}^K\varepsilon_{k_i}^{i-1}\varepsilon_{k_i}^ie_{i,k_i}\Big\|_{X_i}\Big)\Big\|\sum_{k_n=1}^K\varepsilon_{k_n}^{n-1}e_{n,k_n}\Big\|_{X_n},$$

which is further controlled by

$$\begin{aligned}&\Big(\mathbb{E}\Big\|\sum_{k_1=1}^K\varepsilon_{k_1}^1a_{k_1}e_{1,k_1}\Big\|_{X_1}^2\Big)^{1/2}\\&\times\Big[\mathbb{E}\Big(\prod_{i=2}^{n-1}\Big\|\sum_{k_i=1}^K\varepsilon_{k_i}^{i-1}\varepsilon_{k_i}^ie_{i,k_i}\Big\|_{X_i}^2\Big)\Big\|\sum_{k_n=1}^K\varepsilon_{k_n}^{n-1}e_{n,k_n}\Big\|_{X_n}^2\Big]^{1/2}.\end{aligned}\tag{4.2}$$

The first factor is less than $\|(e_{1,k})_{k=1}^K\|_{\mathrm{Rad}(X_1)}$ by Kahane's contraction principle. We now consider the second factor. We see that the variables ε_k^1 appear only inside the norm X_2, and moreover there holds that

$$\mathbb{E}^1\Big\|\sum_{k_2=1}^K\varepsilon_{k_2}^1\varepsilon_{k_2}^2e_{2,k_2}\Big\|_{X_2}^2=\|(e_{2,k})_{k=1}^K\|_{\mathrm{Rad}(X_2)}^2.$$

After using this identity, the variables ε_k^1 do not appear anymore, and the variables ε_k^2 appear only inside the norm X_3. Repeating the same reasoning, we deduce that the second factor in (4.2) is equal to the product $\prod_{j=2}^n \|(e_{j,k})_{k=1}^K\|_{\mathrm{Rad}(X_j)}$. □

Now, we turn to the actual proof of boundedness of shifts. We assume that $n \geq 1$ and that $\{X_1, \dots, X_{n+1}\}$ is a UMD Hölder tuple. Let $k = (k_1, \dots, k_{n+1})$, $0 \leq k_i \in \mathbb{Z}$, and let $\mathcal{D}$ be a dyadic lattice in $\mathbb{R}^d$. Suppose $S^k := S^k_{\mathcal{D}}$ is an n-linear dyadic shift as described in Equation (2.7). We consider its related $(n+1)$-linear form Λ_{S^k} which acts on locally integrable functions $f_j \colon \mathbb{R}^d \to X_j$ by

$$\Lambda_{S^k}(f_1, \dots, f_{n+1}) = \sum_{K \in \mathcal{D}} \Lambda_K(f_1, \dots, f_{n+1}), \tag{4.3}$$

where

$$\Lambda_K(f_1, \dots, f_{n+1}) = \sum_{\substack{Q_1, \dots, Q_{n+1} \in \mathcal{D} \\ Q_j^{(k_j)} = K}} a_{K,(Q_j)_{j \in \mathcal{J}_{n+1}}} \tau\Big(\prod_{j=1}^{n+1} \langle f_j, \widetilde{h}_{Q_j} \rangle \Big).$$

Here $a_{K,(Q_j)_{j \in \mathcal{J}_{n+1}}}$ is a scalar satisfying $|a_{K,(Q_j)_{j \in \mathcal{J}_{n+1}}}| \leq \prod_{j=1}^{n+1} |Q_j|^{1/2} |K|^{-n}$, and there exist two indices $j_0, j_1 \in \mathcal{J}_{n+1}$, $j_0 \neq j_1$, so that $\widetilde{h}_{Q_{j_0}} = h_{Q_{j_0}}$, $\widetilde{h}_{Q_{j_1}} = h_{Q_{j_1}}$ and $\widetilde{h}_{Q_j} = h^0_{Q_j}$ if $j \in \mathcal{J}_{n+1} \setminus \{j_0, j_1\}$.

The sparse domination lemma 3.27 reduces the problem to the following theorem.

4.4. **Theorem.** *Suppose $p_j \in (1, \infty)$ for $j \in \mathcal{J}_{n+1}$ are such that $\sum_{j=1}^{n+1} 1/p_j = 1$. The dyadic shift form from (4.3) satisfies the estimate*

$$|\Lambda_{S^k}(f_1, \dots, f_{n+1})| \lesssim \prod_{j=1}^{n+1} \|f_j\|_{L^{p_j}(X_j)}$$

for $f_j \in L^\infty_c(X_j)$, where the estimate depends polynomially on $\kappa := \max_j k_j$.

Proof. Let $f_j \in L^\infty_c(X_j)$ for $j \in \mathcal{J}_n$ and consider (4.3). Recall the lattices $\mathcal{D}_{i,\kappa}$ from (2.4), where $\kappa := \max_j k_j$. We first divide the sum over the cubes $K \in \mathcal{D}$ as $\sum_{i=0}^{\kappa} \sum_{K \in \mathcal{D}_{i,\kappa}}$. We fix one i and consider the corresponding term.

Let $\widetilde{\mathcal{J}}$ be the set of those indices such that the corresponding Haar functions are non-cancellative, that is, $\widetilde{h}_{Q_j} = h^0_{Q_j}$. Suppose $j \in \widetilde{\mathcal{J}}$ is such that $k_j > 0$. We use that fact that $\langle f_j, \widetilde{h}_{Q_j} \rangle = \langle E_K^{k_j} f_j, h^0_{Q_j} \rangle$ and split

$$E_K^{k_j} f_j = \sum_{l_j=0}^{k_j - 1} \Delta_K^{l_j} f_j + E_K f_j. \tag{4.5}$$

There holds that

$$\langle E_K f_j, h^0_{Q_j} \rangle = \langle f_j, h^0_K \rangle \langle h^0_K, h^0_{Q_j} \rangle \tag{4.6}$$

and

$$\langle \Delta_K^{l_j} f_j, h^0_{Q_j} \rangle = \langle f_j, h_{Q_j^{(k_j - l_j)}} \rangle \langle h_{Q_j^{(k_j - l_j)}}, h^0_{Q_j} \rangle, \tag{4.7}$$

where as usual we suppressed the summation over the different Haar functions.

We use (4.5) to split $\sum_{K\in\mathcal{D}_{i,\kappa}} \Lambda_K(f_1,\dots,f_{n+1})$ into at most $(1+\kappa)^{n-1}$ terms of the form

$$\sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{Q_1,\dots,Q_{n+1}\in\mathcal{D}\\ Q_j^{(k_j)}=K}} a_{K,(Q_j)_{j\in\mathcal{J}_{n+1}}} \tau\Big(\prod_{j=1}^{n+1}\langle P_{K,j}^{l_j} f_j, \widetilde{h}_{Q_j}\rangle\Big), \tag{4.8}$$

where $l_j\in\mathbb{Z}$, $0\le l_j\le k_j$. For $j\in\mathcal{J}_{n+1}\setminus\widetilde{\mathcal{J}}$ we have that $P_{K,j}^{l_j}$ is the identity operator, and below we write $l_j=k_j$. If $j\in\widetilde{\mathcal{J}}$ and $l_j>0$ then $P_{K,j}^{l_j}=\Delta_K^{l_j}$, and if $j\in\widetilde{\mathcal{J}}$ and $l_j=0$ then $P_{K,j}^0$ is either E_K or Δ_K (but does not change with K). We write

$$\sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{Q_1,\dots,Q_{n+1}\in\mathcal{D}\\ Q_j^{(k_j)}=K}} = \sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} \sum_{\substack{Q_1,\dots,Q_{n+1}\in\mathcal{D}\\ Q_j^{(k_j-l_j)}=L_j}}$$

and notice that by (4.6) and (4.7) we always have that

$$\langle P_{K,j}^{l_j} f_j, \widetilde{h}_{Q_j}\rangle = \langle f_j, h'_{L_j}\rangle \gamma(Q_j,L_j),$$

where $h'_{L_j}\in\{h^0_{L_j}, h_{L_j}\}$ and

$$|\gamma(Q_j,L_j)| = \frac{|Q_j|^{1/2}}{|L_j|^{1/2}}.$$

We can now write (4.8) further as

$$\sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \tau\Big(\prod_{j=1}^{n+1}\langle f_j, h'_{L_j}\rangle\Big), \tag{4.9}$$

where

$$b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} = \sum_{\substack{Q_1,\dots,Q_{n+1}\in\mathcal{D}\\ Q_j^{(k_j-l_j)}=L_j}} a_{K,(Q_j)_{j\in\mathcal{J}_{n+1}}} \prod_{j=1}^{n+1}\gamma(Q_j,L_j).$$

There exists $\mathcal{J}\subset\mathcal{J}_{n+1}$ with $\#\mathcal{J}\ge 2$ so that $h'_{L_j}=h_{L_j}$ for $j\in\mathcal{J}$ and if $j\in\mathcal{J}_{n+1}\setminus\mathcal{J}$ then $h'_{L_j}=h^0_{L_j}$ and $l_j=0$. Also, we have the normalization $|b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}}|\le\prod_{j=1}^{n+1}|L_j|^{1/2}|K|^{-n}$.

We have reduced to considering the new shift type operator (4.9). The coefficients satisfy the usual normalization of shifts, but the number $\#\mathcal{J}$ of indices with cancellative Haar functions may be bigger than 2. What is essential is that the complexity related to the non-cancellative indices is zero – that is, if $j\in\mathcal{J}_{n+1}\setminus\mathcal{J}$ then $l_j=0$. We now start estimating (4.9). Also, the separation of scales, $K\in\mathcal{D}_{i,\kappa}$, will allow us to use the decoupling estimate (2.5).

Case 1. Assume that $\mathcal{J} = \mathcal{J}_{n+1}$. Let $q_{n+1} \in (1,\infty)$ be the exponent determined by $1/q_{n+1} = \sum_{j=1}^n 1/p_j$. We need to estimate

$$\Big\| \sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \prod_{j=1}^n \langle f_j, h_{L_j}\rangle h_{L_{n+1}} \Big\|_{L^{q_{n+1}}(Y(X_{n+1}))}$$

$$\sim \Big(\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathcal{V}} \Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x) \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \times \prod_{j=1}^n \langle f_j, h_{L_j}\rangle h_{L_{n+1}}(y_K)\Big|^{q_{n+1}}_{Y(X_{n+1})} \mathrm{d}\nu(y)\,\mathrm{d}x\Big)^{1/q_{n+1}},$$

where we used the decoupling estimate. Notice that since by assumption $X_{n+1} = Y(X_1,\dots,X_n)$, there holds also that $Y(X_{n+1}) = Y(Y(X_1,\dots,X_n))$, so we could also use the norm $|\cdot|_{Y(Y(X_1,\dots,X_n))}$ instead. Write

$$\sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \prod_{j=1}^n \langle f_j, h_{L_j}\rangle h_{L_{n+1}}(y_K)$$

$$= \frac{1}{|K|^n}\int_{K^n} b_K(y_K,z)\prod_{j=1}^n \Delta_K^{l_j} f_j(z_j)\,\mathrm{d}z = \int_{\mathcal{V}^n} b_K(y_K,z_K)\prod_{j=1}^n \Delta_K^{l_j} f_j(z_{j,K})\,\mathrm{d}\nu_n(z),$$

where ν_n is the product measure $\nu\times\cdots\times\nu$ on the product space $\mathcal{V}^n$ and

$$b_K(y_K,z_K) = |K|^n \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \prod_{j=1}^n h_{L_j}(z_{j,K}) h_{L_{n+1}}(y_K).$$

We can now continue the estimate by using Hölder's inequality related to the integral $\int_{\mathcal{V}^n}$. We end up with

$$\Big(\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathcal{V}}\int_{\mathcal{V}^n} \Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x) b_K(y_K,z_K)\prod_{j=1}^n \Delta_K^{l_j} f_j(z_{j,K})\Big|^{q_{n+1}}_{Y(X_{n+1})} \mathrm{d}\nu_n(z)\,\mathrm{d}\nu(y)\,\mathrm{d}x\Big)^{1/q_{n+1}}. \tag{4.10}$$

Suppose $n\geq 2$. Notice that $|b_K(y_K,z_K)|\leq 1$ and use Lemma 4.1 to get that

$$\Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x) b_K(y_K,z_K)\prod_{j=1}^n \Delta_K^{l_j} f_j(z_{j,K})\Big|_{Y(X_{n+1})}$$

$$\leq \prod_{j=1}^n \|(1_K(x)\Delta_K^{l_j} f_j(z_{j,K}))_{K\in\mathcal{D}_{i,\kappa}}\|_{\mathrm{Rad}(X_j)}.$$

Using first Hölder's inequality, then Kahane-Khintchine inequality and finally the decoupling estimate, we conclude that

$$\begin{aligned}(4.10) &\lesssim \prod_{j=1}^{n} \Big(\int_{\mathbb{R}^d} \int_{\mathcal{V}} \|(1_K(x)\Delta_K^{l_j} f_j(z_K))_{K\in\mathcal{D}_{i,\kappa}}\|_{\mathrm{Rad}(X_j)}^{p_j} \,\mathrm{d}\nu(z)\,\mathrm{d}x\Big)^{1/p_j} \\ &\sim \prod_{j=1}^{n} \Big(\mathbb{E} \int_{\mathbb{R}^d} \int_{\mathcal{V}} \Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x)\Delta_K^{l_j} f_j(z_K)\Big|_{X_j}^{p_j} \,\mathrm{d}\nu(z)\,\mathrm{d}x\Big)^{1/p_j} \\ &\lesssim \prod_{j=1}^{n} \|f_j\|_{L^{p_j}(X_j)}.\end{aligned}$$

Suppose then $n = 1$. In this case we have that $q_2 = p_1$ and $Y(X_2) = X_1$. We use Kahane-Khintchine inequality to move the expectation inside of the exponent p_1. Then, we use Kahane's contraction principle and move the expectation out again. This gives that

$$(4.10) \lesssim \Big(\mathbb{E} \int_{\mathbb{R}^d} \int_{\mathcal{V}} \Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x)\Delta_K^{l_1} f_1(z_K)\Big|_{X_1}^{p_1} \,\mathrm{d}\nu(z)\,\mathrm{d}x\Big)^{1/p_1} \lesssim \|f_1\|_{L^{p_1}(X_1)},$$

where the last step used the decoupling estimate. Linear estimates for shifts have appeared e.g. in [19, 29].

Case 2. Assume now that $\mathcal{J} \subsetneq \mathcal{J}_{n+1}$. Since $\#\mathcal{J} \geq 2$, this implies that $n \geq 2$. Let $j_0 \in \mathcal{J}_{n+1} \setminus \mathcal{J}$ be an index such that $j_0 + 1 \in \mathcal{J}$; by $(n+1)+1$ we mean 1. Let $\sigma \in \Sigma(n+1)$ be the cyclic permutation such that $\sigma(n) = j_0$. Then $\sigma(n+1) \in \mathcal{J}$. If $e_j \in X_j$ for $j \in \mathcal{J}_{n+1}$ then from Remark 3.15 one sees that $\prod_{j=1}^n e_j \in Y(X_{n+1})$ and therefore the cyclic invariance of the trace (3.8) gives that $\tau(e_1 \cdots e_{n+1}) = \tau(e_{n+1}e_1 \cdots e_n)$. Repeating this we have that (4.9) is equal to

$$\sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{L_1,\dots,L_{n+1}\in\mathcal{D} \\ L_j^{(l_j)}=K}} b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}} \tau\Big(\prod_{j=1}^{n+1} \langle f_{\sigma(j)}, h'_{L_{\sigma(j)}}\rangle\Big).$$

Having made this important observation, we may now assume, for small notational convenience, that $j_0 = n$ and $\sigma = \mathrm{id}$. Under this assumption $n \in \mathcal{J}_{n+1} \setminus \mathcal{J}$, which implies that $l_n = 0$. Thus, the coefficient $b_{K,(L_j)_{j\in\mathcal{J}_{n+1}}}$ depends only on the cubes $L_1, \dots, L_{n-1}, L_{n+1}$ and K. Below we will write the coefficient as $b_{K,(L_j)}$.

We need to estimate

$$\begin{aligned}&\Big\| \sum_{K\in\mathcal{D}_{i,\kappa}} \sum_{\substack{L_1,\dots,L_{n-1},L_{n+1}\in\mathcal{D} \\ L_j^{(l_j)}=K}} b_{K,(L_j)} \prod_{j=1}^{n-1} \langle f_j, h'_{L_j}\rangle \langle f_n\rangle_K |K|^{1/2} h_{L_{n+1}}\Big\|_{L^{q_{n+1}}(Y(X_{n+1}))} \\ &\quad\sim \Big(\mathbb{E} \int_{\mathbb{R}^d} \int_{\mathcal{V}} \Big| \sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x)\langle \varphi_{K,y}\rangle_K\Big|_{Y(X_{n+1})}^{q_{n+1}} \,\mathrm{d}\nu(y)\,\mathrm{d}x\Big)^{1/q_{n+1}},\end{aligned}$$

where we used the decoupling estimate, and for $K \in \mathcal{D}_{i,\kappa}$ and $y \in \mathcal{V}$ we defined the function $\varphi_{K,y}\colon \mathbb{R}^d \to Y(X_{n+1})$ by setting $\varphi_{K,y}(x)$ to equal

$$|K|^{1/2} \sum_{\substack{L_1,\dots,L_{n-1},L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} b_{K,(L_j)} \prod_{j=1}^{n-1} \langle f_j, h'_{L_j}\rangle f_n(x) h_{L_{n+1}}(y_K).$$

After using Stein's inequality (2.3) with respect to $x \in \mathbb{R}^d$ with fixed $y \in \mathcal{V}$ we are left with

$$\Big(\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathcal{V}} \Big|\sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x)\varphi_{K,y}(x)\Big|^{q_{n+1}}_{Y(X_{n+1})}\,\mathrm{d}\nu(y)\,\mathrm{d}x\Big)^{1/q_{n+1}}. \tag{4.11}$$

Recall that $n \geq 2$ in Case 2. From Remark 3.15 we can deduce that if $e_n \in X_n$ and $e_{n+1} \in X_{n+1}$, then $e_n e_{n+1} \in Y(X_1,\dots,X_{n-1})$ and $|e_n e_{n+1}|_{Y(X_1,\dots,X_{n-1})} \leq |e_n|_{X_n}|e_{n+1}|_{X_{n+1}}$. Also, since $\{X_1,\dots,X_{n-1},Y(X_1,\dots,X_{n-1})\}$ is a UMD Hölder tuple, we see from Remark 3.15 again that if $e_j \in X_j$ for $j \in \mathcal{J}_{n-1}$, then $\prod_{j=1}^{n-1} e_j \in Y(Y(X_1,\dots,X_{n-1}))$. Suppose now that $e_{j,k} \in X_j$ for $j \in \mathcal{J}_{n-1}$, $k = 1,\dots,K$, and $e_n \in X_n$. Then the above consideration implies that the key inequality

$$\Big|\sum_{k=1}^{K}\prod_{j=1}^{n-1} e_{j,k}e_n\Big|_{Y(X_{n+1})} \leq \Big|\sum_{k=1}^{K}\prod_{j=1}^{n-1} e_{j,k}\Big|_{Y(Y(X_1,\dots,X_{n-1}))}|e_n|_{X_n} \tag{4.12}$$

holds. Write $Z := Y(Y(X_1,\dots,X_{n-1}))$ for the moment. Using this in (4.11) and then Hölder's inequality we have that (4.11) is dominated by $\|f_n\|_{L^{p_n}(X_n)}$ multiplied by

$$\Big(\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathcal{V}} \Big|\sum_{K\in\mathcal{D}_{i,\kappa}} \varepsilon_K 1_K(x) \sum_{\substack{L_1,\dots,L_{n-1},L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} \widetilde{b}_{K,(L_j)} \prod_{j=1}^{n-1}\langle f_j, h'_{L_j}\rangle h_{L_{n+1}}(y_K)\Big|^{p(\mathcal{J}_{n-1})}_{Z}\,\mathrm{d}\nu(y)\,\mathrm{d}x\Big)^{\frac{1}{p(\mathcal{J}_{n-1})}}$$

$$\sim \Big\|\sum_{K\in\mathcal{D}_{i,\kappa}}\sum_{\substack{L_1,\dots,L_{n-1},L_{n+1}\in\mathcal{D}\\ L_j^{(l_j)}=K}} \widetilde{b}_{K,(L_j)} \prod_{j=1}^{n-1}\langle f_j, h'_{L_j}\rangle h_{L_{n+1}}\Big\|_{L^{p(\mathcal{J}_{n-1})}(Z)'}$$

where we defined $1/p(\mathcal{J}_{n-1}) := \sum_{j=1}^{n-1} 1/p_j$, $\widetilde{b}_{K,(L_j)} := |K|^{1/2} b_{K,(L_j)}$ and used the decoupling inequality. Notice that

$$|\widetilde{b}_{K,(L_j)}| \leq \frac{\prod_{j=1}^{n-1}|L_j|^{1/2}|L_{n+1}|^{1/2}}{|K|^{n-1}}.$$

We see that we have reduced the estimate to the boundedness of an $(n-1)$-linear shift type operator as in (4.9). Now, we have two possibilities. If all the Haar functions h'_{L_j} for $j \in \mathcal{J}_{n-1}$ are cancellative then we are in a position to apply Case 1 from above to finish the estimate. If some of them is non-cancellative, then we dualize with a function $g \in L^{p(\mathcal{J}_{n-1})'}(Y(X_1,\dots,X_{n-1}))$. This leads us to a corresponding situation as the beginning of Case 2 above but now the form is n-linear and the underlying UMD Hölder n-tuple is $\{X_1,\dots,X_{n-1},Y(X_1,\dots,X_{n-1})\}$. We see that we can repeat the argument in Case 2 until we can apply Case 1. This finishes the proof.

□

4.13. *Remark.* We discuss why Theorem 1.1 works without any UMD Hölder tuple assumptions on the spaces X_1, X_2 and Y_3, and why we can't allow more UMD spaces in Theorem 1.1. The key point is that for $e_{1,k} \in X_1$ and $e_2 \in X_2$ the estimate

$$\Big|\sum_{k=1}^{K} e_{1,k}e_2\Big|_{Y_3} \le \Big|\sum_{k=1}^{K} e_{1,k}\Big|_{X_1} |e_2|_{X_2}, \tag{4.14}$$

which corresponds to (4.12), holds without any further assumptions on the spaces. Using this kind of estimates one can prove Theorem 1.1 with similar techniques as in the proof of Theorem 4.4.

Suppose then we have UMD spaces $X_1, \dots, X_n$ and Y_{n+1}, where $n \ge 3$, and we have a product $X_1 \times \cdots \times X_n \to Y_{n+1}$ – a bounded n-linear operator. Of course, an estimate corresponding to (4.14) holds, namely

$$\Big|\sum_{k=1}^{K} e_{1,k}\prod_{j=2}^{n} e_j\Big|_{Y_{n+1}} \le \Big|\sum_{k=1}^{K} e_{1,k}\Big|_{X_1}\prod_{j=2}^{n} |e_j|_{X_j}.$$

However, in the above proof for shifts, when we use Stein's inequality, we need to reduce the linearity before we can use it again. That is why we need the product structure of UMD Hölder tuples rather than just a product $X_1 \times \cdots \times X_n \to Y_{n+1}$ on the top level.

5. Boundedness of multilinear paraproducts in UMD Hölder tuples

In this section we prove the boundedness of multilinear paraproducts. Let us first recall a result for paraproducts acting on UMD-valued functions. If X is a UMD space, $\mathcal{D}$ is a dyadic lattice and $\{a_Q\}_{Q\in\mathcal{D}}$ is a collection of scalars satisfying the BMO condition (2.8), then

$$\Big\|\sum_{Q\in\mathcal{D}} a_Q\langle f\rangle_Q h_Q\Big\|_{L^p(X)} \lesssim \|f\|_{L^p(X)}, \tag{5.1}$$

where $p \in (1,\infty)$. This result goes back to Bourgain, see Figiel–Wojtaszczyk [14]. Another nice proof is obtained by adapting the argument of Hänninen–Hytönen [19], who consider paraproducts with operator coefficients.

Let $n \ge 1$ and let $\{X_1, \dots, X_{n+1}\}$ be a UMD Hölder tuple. Suppose that $\mathcal{D}$ is a dyadic lattice and that $\pi := \pi_{\mathcal{D}}$ is a paraproduct as described in Section 2.4. Let $j_0 \in \mathcal{J}_{n+1}$ be the index related to the cancellative Haar functions of π and let $\sigma \in \Sigma(n+1)$ be the cyclic permutation such that $\sigma(n+1) = j_0$. We consider the $(n+1)$-linear form Λ_π acting on functions $f_j \in L^\infty_c(X_j)$ by

$$\Lambda_\pi(f_1, \dots, f_{n+1}) = \sum_{Q\in\mathcal{D}} a_Q\tau\left(\Big[\prod_{j=1}^{n}\langle f_{\sigma(j)}\rangle_Q\Big]\langle f_{\sigma(n+1)}, h_Q\rangle\right), \tag{5.2}$$

where the scalars $\{a_Q\}_{Q\in\mathcal{D}}$ satisfy the BMO condition (2.8). The following theorem combined with Lemma 3.27 proves the desired estimate.

5.3. **Theorem.** *Suppose that $p_j \in (1,\infty)$ for $j \in \mathcal{J}_{n+1}$ are such that $\sum_{j=1}^{n+1} 1/p_j = 1$. If $f_j \in L^\infty_c(X_j)$ for $j \in \mathcal{J}_{n+1}$ then the form Λ_π from (5.2) satisfies the estimate*

$$|\Lambda_\pi(f_1,\dots,f_{n+1})| \lesssim \prod_{j=1}^{n+1} \|f_j\|_{L^{p_j}(X_j)}.$$

Proof. For $m \in \mathcal{J}_n$ we let $p(\mathcal{J}_m)$ be the exponent defined by $1/p(\mathcal{J}_m) = \sum_{j=1}^m 1/p_j$. For convenience of notation we may assume that $j_0 = n+1$, so that $\sigma = \mathrm{id}$. In this case we need to estimate the term

$$\Big\| \sum_{Q\in\mathcal{D}} a_Q \prod_{j=1}^n \langle f_j\rangle_Q h_Q \Big\|_{L^{p(\mathcal{J}_n)}(Y(X_{n+1}))}.$$

The case $n=1$ is the known estimate (5.1). Therefore, we assume that $n \geq 2$.

Applying the UMD property of $Y(X_{n+1})$ we are led to

$$\Big(\mathbb{E}\int_{\mathbb{R}^d} \Big| \sum_{Q\in\mathcal{D}} \varepsilon_Q a_Q \prod_{j=1}^n \langle f_j\rangle_Q |h_Q(x)|\Big|^{p(\mathcal{J}_n)}_{Y(X_{n+1})}\,\mathrm{d}x\Big)^{1/p(\mathcal{J}_n)}, \tag{5.4}$$

where to pass from h_Q to $|h_Q|$ we used that for fixed $x \in \mathbb{R}^d$ the families $\{\varepsilon_Q h_Q(x)\}$ and $\{\varepsilon_Q |h_Q(x)|\}$ are identically distributed. Since $|h_Q| = 1_Q/|Q|^{1/2}$, we can use Stein's inequality to have that

$$(5.4) \lesssim \Big(\mathbb{E}\int_{\mathbb{R}^d} \Big| \sum_{Q\in\mathcal{D}} \varepsilon_Q a_Q \prod_{j=1}^{n-1} \langle f_j\rangle_Q f_n(x) |h_Q(x)|\Big|^{p(\mathcal{J}_n)}_{Y(X_{n+1})}\,\mathrm{d}x\Big)^{1/p(\mathcal{J}_n)}.$$

Now, we use the same inequality we used in the shift proof, Equation (4.12), and Hölder's inequality to have that the last term is less than $\|f_n\|_{L^{p_n}(X_n)}$ multiplied by

$$\Big(\mathbb{E}\int_{\mathbb{R}^d} \Big| \sum_{Q\in\mathcal{D}} \varepsilon_Q a_Q \prod_{j=1}^{n-1} \langle f_j\rangle_Q |h_Q(x)|\Big|^{p(\mathcal{J}_{n-1})}_{Y(Y(X_1,\dots,X_{n-1}))}\,\mathrm{d}x\Big)^{1/p(\mathcal{J}_{n-1})}.$$

Since $\{X_1,\dots,X_{n-1},Y(X_1,\dots,X_{n-1}\}$ is a UMD Hölder n- tuple, we see that we have reduced to a situation as in (5.4) but now the degree of linearity is $n-1$. We can repeat the argument until we end up with a linear operator, and then we apply (5.1). □

6. Maximality of UMD Hölder tuples

In this brief section, we show that UMD Hölder tuples are in a suitable sense maximal for L^p-boundedness of extensions of n-linear CZO operators and dyadic shifts via an associative product as in Section 3. The precise statement is in Proposition 6.3 below.

Therefore, we fix an associative algebra $\mathcal{A}$ and a functional τ as in Section 3. We begin with a lemma.

6.1. **Lemma.** *Let $(X_1,\dots,X_n)$ be a n-tuple of admissible spaces. If X_{n+1} is an admissible space such that for all $(n+1)$-linear shift forms (3.23) and functions $f_j \in C^1(\mathbb{R}^d)\otimes X_j$, $j=1,\dots,n+1$*

$$\Big|\Lambda_{\mathcal{U}^k_{\mathcal{D}_\omega,u}}(f_1,\dots,f_n,f_{n+1})\Big| \lesssim \Big(\prod_{\ell=1}^{n+1} \|f_j\|_{L^{n+1}(\mathbb{R}^d;X_j)}\Big) \tag{6.2}$$

with implicit constant depending possibly on the complexity k, then (3.5) *holds for $m = n$, and in particular $X_{n+1} \hookrightarrow Y(X_1, \dots, X_n)$.*

Proof. Test (6.2) on a suitable trivial shift and appeal to Lemma 3.4. □

To make our maximality claim precise, we need an additional definition. We say that the tuple $\{X_1, \dots, X_{n+1}\}$ of admissible spaces *is an n-linear shift extension* if (6.2) holds for all $(n+1)$-linear shift forms (3.23). If in addition, whenever Z is an admissible space such that for some $j_0 \in \mathcal{J}_{n+1}$ the tuple $\{X_1, \dots X_{j_0-1}, Z, X_{j_0+1}, \dots X_{n+1}\}$ is an n-linear shift extension, it must be $Z \hookrightarrow X_{j_0}$, we say that $\{X_1, \dots, X_{n+1}\}$ is a *maximal* n-linear shift extension.

6.3. **Proposition.** *Let $\{X_1, \dots, X_{n+1}\}$ be a* UMD *Hölder tuple. Then*

- *$\{X_1, \dots, X_{n+1}\}$ is a maximal n-linear shift extension;*
- *whenever $1 \le k \le n-1$ and $\#\mathcal{J} = k$, $\{X_j : j \in \mathcal{J}\} \cup \{Y(\{X_j : j \in \mathcal{J}\})\}$ is a maximal k-linear shift extension.*

Proof. Theorem 4.4 shows that if $\{X_1, \dots, X_{n+1}\}$ is a UMD Hölder tuple, then it is an n-linear shift extension. As $X_{j_0} = Y(\{X_j : j \in \mathcal{J}_0\})$ by definition of UMD Hölder tuple, we learn from Lemma 6.1 that $\{X_1, \dots, X_{n+1}\}$ is in fact a maximal n-linear shift extension. This proves the first point.

By the inductive definition of UMD Hölder tuple, for each $1 \le k \le n-1$ and $\#\mathcal{J} = k$, $\{X_j : j \in \mathcal{J}\} \cup \{Y(\{X_j : j \in \mathcal{J}\})\}$ is a UMD Hölder tuple. Then this tuple must be a maximal k-linear shift extension because of the first point. The second point is also proved. □

Appendix A. Iterated mixed-norm non-commutative L^p spaces

Let $\mathcal{M}$ be a von Neumann algebra equipped with a n.s.f. trace as described in Example 3.18. Recall in particular that $\mathcal{A} = L^0(\mathcal{M})$ is an associative $*$-algebra endowed with a compatible complete metrizable topology, induced by the metric $d_{\mathcal{A}}$ of convergence in measure. For an integer $S \ge 1$, let (M_s, μ_s), $s = 1, \dots, S$, be σ-finite measure spaces and (Ω_S, ω_S) the product measure space

$$\Omega_S = \prod_{s=1}^{S} M_s, \qquad \omega_S = \prod_{s=1}^{S} \mu_s.$$

Let $\mathscr{A}_{0,S}$ be the vector space of *simple functions* $f : \Omega_S \to \mathcal{A}$, namely

$$f(t) = \sum_{j=1}^{J} A_j \mathbf{1}_{E^j}(t), \qquad t = (t_1, \dots, t_s) \in \Omega_S,$$

with $A_j \in \mathcal{A}$, $E^j \subset \Omega_S$ with $\omega_S(E^j) < \infty$. Then $\mathscr{A}_{0,S}$ is an associative algebra with respect to the pointwise product: for $f, g \in \mathscr{A}_{0,S}$, the function fg defined by $(fg)(t) = f(t)g(t)$, where the latter is the strong product in $\mathcal{A}$, belongs to $\mathscr{A}_{0,S}$. We denote by

$$\mathscr{A}_S := \text{ closure of } \mathscr{A}_0 \text{ w.r.t. sequential } d_{\mathcal{A}}\text{-pointwise convergence}$$

namely, $f \in \mathscr{A}_S$ if there exists a sequence $f_n \in \mathscr{A}_{0,S}$ such that

$$\lim_n d_{\mathcal{A}}(f(t), f_n(t)) = 0 \qquad a.e.\ t \in \Omega_S.$$

Then $\mathscr{A}_S$, the class of strongly measurable $\mathcal{A}$-valued functions on Ω_S, is an associative algebra with respect to the same product. Furthermore, $\mathscr{A}_S$ is complete with respect to the topology of convergence in measure, namely $f_n \to f$ if for all $\varepsilon > 0$

$$\lim_n \mu\left(\{t \in \Omega_S : d_{\mathcal{A}}(f(t), f_n(t)) > \varepsilon\}\right) = 0,$$

and the product operation is continuous. Note that the latter topology is also metrizable, proceeding in an analogous way to [24, Proposition A.2.4].

Recall that $\mathcal{M}$ is equipped with the n.s.f. trace τ, which is a linear bounded functional on $L^1(\mathcal{M})$. Then the functional

$$\tau_S(f) \coloneqq \int_{\Omega_S} \tau(f(t))\,\mathrm{d}\omega_S(t)$$

is linear and bounded on the Bochner space $L^1(\Omega_S, \omega_S; L^1(\mathcal{M}))$, which is a subspace of $\mathscr{A}_S$. With this definition, $\mathscr{A}_S$ is endowed with the trace τ_S. Under these assumptions, we have the following proposition.

A.1. **Proposition.** *For a Hölder tuple $\{p_j^0 : 1 \le j \le m\}$ as in (3.1), let*

$$X_j^0 = L^{p_j^0}(\mathcal{M}).$$

Let $\{p_j^s : 1 \le j \le m\}$ be further Hölder tuples of exponents, for $1 \le s \le S$. Then the Banach subspaces of $\mathscr{A}_s$

$$X_j^s = L^{p_j^s}(M_s, \mu_s; X_j^{s-1}), \qquad s = 1, \dots, S, \tag{A.2}$$

are a UMD *Hölder m-tuple.*

Before the proof proper, we need to set some notation, and develop suitable auxiliary lemmata. For $1 \le k \le m-1$, $\mathcal{J} = \{j_1 < j_2 < \cdots < j_k\} \subset \mathcal{J}_m$, and $0 \le s \le S$ we write

$$\frac{1}{q_{\mathcal{J}}^s} = \sum_{u=1}^k \frac{1}{p_{j_u}^s}, \qquad \frac{1}{p_{\mathcal{J}}^s} = 1 - \frac{1}{q_{\mathcal{J}}^s}.$$

It will be convenient to introduce the auxiliary mixed norm spaces

$$\begin{aligned} E_j^1 &= L^{p_j^1}(M_1, \mu_1), \\ E_j^s &= L^{p_j^s}(M_s, \mu_s; E_j^{s-1}), \qquad s = 2, \dots, S, \end{aligned}$$

for $j = 1, \dots, m$ and similarly

$$\begin{aligned} E_{\mathcal{J}}^0 &= \mathbb{C}, \\ E_{\mathcal{J}}^s &= L^{q_{\mathcal{J}}^s}(M_s, \mu_s; E_{\mathcal{J}}^{s-1}), \qquad s = 1, \dots, S. \end{aligned}$$

In general we write $S(X)$ for the unit sphere in the Banach space X.

A.3. **Lemma.** *Let $\mathcal{J} = \{j_u : 1 \le u \le k\}$. There exists maps $B_u^s : S(E_{\mathcal{J}}^s) \to S(E_{j_u}^s)$ such that*

$$f = \prod_{u=1}^k B_u^s(f) \qquad \forall f \in S(E_{\mathcal{J}}^s)$$

and

$$\begin{aligned}&\|f_n - f\|_{E^s_{\mathcal{J}}} \to 0,\ \|f_n(t_s) - f(t_s)\|_{E^{s-1}_{\mathcal{J}}} \to 0 \text{ a.e. } t_s \in M_s \implies \\ &\|B^s_u(f) - B^s_u(f_n)\|_{E^s_{j_u}} \to 0,\ \|B^s_u(f_n)(t_s) - B^s_u(f_n)(t_s)\|_{E^{s-1}_{j_u}} \to 0 \text{ a.e. } t_s \in M_s, \quad 1 \le u \le k.\end{aligned} \tag{A.4}$$

Proof. We deal with the case $j_u = u, u = 1, \dots, k$ which is generic. We prove the statement by induction on s. If $s \ge 2$, assume inductively that maps B^{s-1}_u as in the statement have been constructed; for the base case $s = 1$, we run the argument below with B^0_u the identity map. In both cases, we need to define $B^s_u : S(E^s_{\mathcal{J}}) \to S(E^s_u)$. We use that each $f \in S(E^s_{\mathcal{J}})$ is $E^{s-1}_{\mathcal{J}}$-valued. So for each $t_s \in M_s$, write

$$f(t_s) = |f(t_s)|_{E^{s-1}_{\mathcal{J}}} g(t_s) = \prod_{u=1}^{k} \left(|f(t_s)|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}} g_u(t_s) \right) =: \prod_{u=1}^{k} B^s_u(f)(t_s)$$

where g is $S(E^{s-1}_{\mathcal{J}})$-valued, so that each $g_u = B^{s-1}_u(g)$ is $S(E^{s-1}_u)$-valued. Notice that each $f_u = B^s_u(f)$ is (strongly) μ_s-measurable with values in E^{s-1}_u: in fact $|f(\cdot)|_{E^{s-1}_{\mathcal{J}}}$ is μ_s-measurable and each g_u is μ_s-measurable, as B^{s-1}_u is (norm) continuous from $E^{s-1}_{\mathcal{J}} \to E^{s-1}_u$ and g is μ_s-measurable with values in $E^{s-1}_{\mathcal{J}}$. A direct calculation reveals that

$$\|f_u\|_{E^s_u} = 1, \qquad 1 \le u \le k.$$

It remains to show that the thus defined maps B^s_u are continuous in the sense of (A.4) by assuming the same properties hold for the maps B^{s-1}_u. Let f_n, f be as in the first line of (A.4) and write $f_n(t_s) = |f_n(t_s)|_{E^{s-1}_{\mathcal{J}}} g_n(t_s)$. We first show the pointwise convergence: for each we have

$$\begin{aligned}\|B^s_u(f)(t_s) - B^s_u(f_n)(t_s)\|_{E^{s-1}_u} \le{}& \|f(t_s)\|_{E^{s-1}_{\mathcal{J}}} \|B^{s-1}_u(g(t_s)) - B^{s-1}_u(g_n(t_s))\|_{E^{s-1}_u} \\ &+ \quad \|B^{s-1}_u(g_n(t_s))\|_{E^{s-1}_u} \left| \|f(t_s)\|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}} - \|f_n(t_s)\|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}} \right|\end{aligned}$$

Relying on the norm continuity of B^{s-1}_u we obtain that both summands in the previous display converge to zero for each t_s such that $\|f_n(t_s)\|_{E^{s-1}_{\mathcal{J}}} \to \|f(t_s)\|_{E^{s-1}_{\mathcal{J}}}, \|g_n(t_s) - g(t_s)\|_{E^{s-1}_{\mathcal{J}}} \to 0$; this is a set of full μ_s measure, so that this part of the proof is complete. We come to the norm continuity in (A.4). We have

$$\begin{aligned}\|B^s_u(f) - B^s_u(f_n)\|^{p^s_u}_{E^s_u} \lesssim{}& \int_{M_s} |f(t_s)|^{q^s_{\mathcal{J}}}_{E^{s-1}_{\mathcal{J}}} |B^{s-1}_u(g(t_s)) - B^{s-1}_u(g_n(t_s))|^{p^s_u}_{E^{s-1}_u} \, \mathrm{d}\mu_s(t_s) \\ &+ \int_{M_s} \left| |f(t_s)|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}} - |f_n(t_s)|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}} \right|^{p^s_u} |B^{s-1}_u(g_n(t_s))|^{p^s_u}_{E^{s-1}_u} \, \mathrm{d}\mu_s(t_s)\end{aligned}$$

The first integrand converges to zero pointwise a.e. and is dominated by $|f(t_s)|^{q^s_{\mathcal{J}}}_{E^{s-1}_{\mathcal{J}}}$, so the integral converges to zero by dominated convergence. The second integral is equal to

$$\|F - F_n\|^{p^s_u}_{L^{p^s_u}(M_s,\mu_s)}, \qquad F(t_s) = |f(t_s)|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}}, \qquad F_n(t_s) = |f_n(t_s)|^{\frac{q^s_{\mathcal{J}}}{p^s_u}}_{E^{s-1}_{\mathcal{J}}}.$$

Notice that $\|F\|_{p_u^s} = \|f\|_{E_{\mathcal{J}}^s}^{q_{\mathcal{J}}^s/p_u^s}$, $\|F_n\|_{p_u} = \|f_n\|_{E_{\mathcal{J}}^s}^{q_{\mathcal{J}}^s/p_u^s}$. As $F_n \to F$ pointwise, $F_n, F \in L^{p_u^s}(M_s, \mu_s)$ and $\|F_n\|_{p_u^s} \to \|F\|_{p_u^s}$, then $\|F - F_n\|_{p_u^s}$ converges to zero by a well-known variation of the proof of the L^p dominated convergence theorem. □

A.5. **Lemma.** *Let* [2]

$$X_{\mathcal{J}}^0 = L^{q_{\mathcal{J}}^0}(\mathcal{M}), \qquad X_{\mathcal{J},+}^0 = L^{q_{\mathcal{J}}^0}(\mathcal{M})_+,$$
$$X_{\mathcal{J}}^s = L^{q_{\mathcal{J}}^s}(M_s, \mu_s; X_{\mathcal{J}}^{s-1}), \qquad X_{\mathcal{J},+}^s = L^{q_{\mathcal{J}}^s}(M_s, \mu_s; X_{\mathcal{J},+}^{s-1}), \qquad s = 1, \dots, S.$$

Let $f \in X_{\mathcal{J},+}^s$ be a simple function with $\|f\|_{X_{\mathcal{J}}^s} = 1$. Then there exist $f_u \in X_{j_u,+}^s$, $u = 1, \dots, k$ with

$$f = \prod_{u=1}^{k} f_u, \qquad \|f_u\|_{X_{j_u}^s} = 1.$$

Proof. Again we deal with the generic case $j_u = u, u = 1, \dots, k$. First of all, we make a remark about the case $s = 0$. Fix $A \in X_{\mathcal{J},+}^0$ with $\|A\|_{X_{\mathcal{J}}^0} = 1$. Using the Borel functional calculus for positive closed densely defined operators to define A^θ for $\theta > 0$

$$A = \prod_{u=1}^{k} B_u(A), \qquad B_u(A) = A^{\frac{q_{\mathcal{J}}^0}{p_u^0}}, \quad u = 1, \dots, k. \tag{A.6}$$

Trivially

$$\|B_u(A)\|_{X_u^0} = \|A\|_{X_{\mathcal{J}}^0}^{\frac{q_{\mathcal{J}}^0}{p_u^0}} = 1, \quad u = 1, \dots, k.$$

We now prove the main statement. Let $f \in X_{\mathcal{J},+}^s$ be a simple function with $\|f\|_{X_{\mathcal{J}}^s} = 1$. We factor

$$f(t) = F(t)A(t), \qquad F(t) = |f(t)|_{X_{\mathcal{J}}^0}, \qquad t \in \Omega_s.$$

Notice that $F \in E_{\mathcal{J}}^s$ of unit norm, so that using Lemma A.3

$$F = \prod_{u=1}^{k} B_u^s(F), \qquad \|B_u^s(F)\|_{E_u^s} = 1,$$

and we may write, also using (A.6)

$$f = \prod_{u=1}^{k} f_u, \qquad f_u(t) = B_u^s(F)(t) B_u(A(t)),$$

Notice that each f_u is strongly measurable as $B_u(A(\cdot))$ is a simple $X_{u,+}^0$-valued function and $B_u^s(F)$ is a measurable function in E_u^s. Also as $|B_u(A(t))|_{X_u^0} = 1$ for all $t \in \Omega_s$

$$\|f_u\|_{X_u^s} = \|B_u^s(F)\|_{E_u^s} = 1,$$

which completes the proof of the claim. □

[2] Recall that $L^{q_{\mathcal{J}}^0}(\mathcal{M})_+$ denotes the positive cone of $L^{q_{\mathcal{J}}^0}(\mathcal{M})$, namely the positive operators in $L^{q_{\mathcal{J}}^0}(\mathcal{M})$.

We turn to the proof of the proposition. Namely we need to show that the tuple X_j^s from (A.2) is a UMD Hölder tuple for each $s = 1, \dots, S$. In proving this, by virtue of the case $s = 0$ being already established in Example 3.18 we may argue inductively and assume the claim has been proved in the cases of $0, \dots, s-1$.

Clearly each X_j^s is a subspace of $\mathscr{A}_s$. Denoting by $q_j^s, s = 0, \dots, S$ the conjugate exponent of p_j^s, it is convenient to define the spaces

$$\begin{aligned} Y_j^0 &= L^{q_j^0}(\mathcal{M}), \\ Y_j^s &= L^{q_j^s}(M_s, \mu_s; Y_j^{s-1}), \qquad s = 1, \dots, S. \end{aligned}$$

which are Banach subspaces of $\mathscr{A}_s$. Further, as each X_j^s is a reflexive Banach space and enjoys the Radon-Nikodým property [24, Theorem 1.3.21], an inductive argument yields the Riesz representation theorem (cf. [24, Theorem 1.3.10]) then yields that

$$\left(X_j^s\right)^* = Y_j^s, \qquad 1 \le j \le m$$

through the identification

$$\lambda \in (X_j^s)^* \mapsto g_\lambda \in Y_j^s \qquad \lambda(f) = \tau_s(g_\lambda f), \quad f \in X_j^s.$$

We have in particular shown that each X_j^s is an admissible space for the algebra $\mathscr{A}_s$ with trace τ_s and $Y(X_j^s) = Y_j^s$.

We verify that $\{X_j^s : j \in \mathcal{J}_m\}$ is a UMD Hölder tuple by induction on m. The case $m = 2$ is actually immediate by virtue of the observation and the well known fact that each X_j^s, Y_j^s is a UMD space.

To obtain the inductive step, we fix $m \ge 3$ and verify the following equality. For each $1 \le k \le m-1$, $\mathcal{J} = \{j_1 < j_2 < \cdots < j_k\} \subset \mathcal{J}_m$, there holds

$$Y(\{X_j^s : j \in \mathcal{J}\}) \text{ is isometrically isomorphic to } \left(X_{\mathcal{J}}^s\right)^*, \tag{A.7}$$

where we refer to the spaces defined in Lemma A.5. More explicitly, denoting

$$\begin{aligned} Y_{\mathcal{J}}^0 &= L^{p_{\mathcal{J}}^0}(\mathcal{M}), \\ Y_{\mathcal{J}}^s &= L^{p_{\mathcal{J}}^s}(M_s, \mu_s; Y_{\mathcal{J}}^{s-1}), \qquad s = 1, \dots, S, \end{aligned}$$

we have $Y(\{X_j^s : j \in \mathcal{J}\}) = Y_{\mathcal{J}}^s = \left(X_{\mathcal{J}}^s\right)^*$.

Property P1 then corresponds to this equality in the cases $k = m-1$. Verifying property P2 amounts to checking that when $k < m-1$, the tuple $\{X_j^s : j \in \mathcal{J}\} \cup \{Y_{\mathcal{J}}^s\}$ is a UMD Hölder $(k+1)$-tuple. As $k < m-1$, $\{X_j^s : j \in \mathcal{J}\} \cup \{Y_{\mathcal{J}}^s\}$ is a UMD Hölder $(k+1)$-tuple and the exponents $\{p_j^s : j \in \mathcal{J}, p^s(\mathcal{J})\}$ are a Hölder tuple, this check is made by a straightforward appeal to the induction assumption.

We are left with proving (A.7). To do this we will define a linear surjective isometry $\Phi : Y(\{X_j^s : j \in \mathcal{J}\}) \to Y_{\mathcal{J}}^s$. First of all note that

$$\|g\|_{Y(\{X_j^s : j \in \mathcal{J}\})} \le \|g\|_{L^{p_{\mathcal{J}}^s}(M_s, \mu_s; Y_{\mathcal{J}}^{s-1})} = \|g\|_{Y_{\mathcal{J}}^s} \tag{A.8}$$

descends immediately from Hölder's inequality in $L^p(M_s, \mu_s)$-spaces and Lemma 3.4 applied to the UMD Hölder tuple $X_{j_1}^{s-1}, X_{j_2}^{s-1}, \dots, X_{j_k}^{s-1}$. We will use this below.

Fix then $g \in Y(\{X_j^s : j \in \mathcal{J}\})$. We claim that if f is a simple $X^0_{\mathcal{J},+}$-valued function on Ω_s with $\|f\|_{X^s_{\mathcal{J}}} = 1$, then

$$|\tau_s(gf)| \leq \|g\|_{Y(\{X_j^s : j\in\mathcal{J}\})}. \tag{A.9}$$

Indeed, applying Lemma 3.4 we obtain

$$\big|\tau_s(gf)\big| = \left|\tau_s\left(g\prod_{u=1}^k f_u\right)\right| \leq \|g\|_{Y(\{X_j^s : j\in\mathcal{J}\})}\prod_{u=1}^k \|f_u\|_{X^s_{j_u}}, \qquad \|f_u\|_{X^s_{j_u}} = 1, \quad u = 1,\dots,k,$$

which is (A.9). As $X^s_{\mathcal{J}}$ is the $X^s_{\mathcal{J}}$-norm closure of the linear span of simple $X^0_{\mathcal{J},+}$-valued function on Ω_s, the linear bounded functional $f \mapsto \tau_s(gf)$ extends uniquely to an element $\Phi(g)$ of $(X^s_{\mathcal{J}})^* \equiv Y^s_{\mathcal{J}}$ with

$$\|\Phi(g)\|_{Y^s_{\mathcal{J}}} \leq \|g\|_{Y(\{X_j^s : j\in\mathcal{J}\})}.$$

It is easy to see that the map $\Phi : Y(\{X_j^s : j \in \mathcal{J}\}) \to Y^s_{\mathcal{J}}$ is linear. From (A.8) we gather that if $g \in Y^s_{\mathcal{J}}$ then $\Phi(g)$ is well-defined. In this case the linear bounded functionals $g \mapsto \tau_s(gf)$ and $\Phi(g)$ coincide on a dense set, it must be $\Phi(g) = g$. So Φ is obviously surjective. Furthermore using (A.8) again we obtain

$$\|\Phi(g)\|_{Y^s_{\mathcal{J}}} \geq \|\Phi(g)\|_{Y(\{X_j^s : j\in\mathcal{J}\})} = \|g\|_{Y(\{X_j^s : j\in\mathcal{J}\})} \geq \|\Phi(g)\|_{Y^s_{\mathcal{J}}}$$

whence equality must hold throughout. So Φ is a linear isometric isomorphism and the proof of (A.7) is complete.

References

[1] Jean Bourgain, *Some remarks on Banach spaces in which martingale difference sequences are unconditional*, Ark. Mat. **21** (1983), no. 2, 163–168. MR 727340 1, 3

[2] Donald L. Burkholder, *A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions*, Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, pp. 270–286. MR 730072 1

[3] Alejandro J. Castro and Tuomas P. Hytönen, *Bounds for partial derivatives: necessity of UMD and sharp constants*, Math. Z. **282** (2016), no. 3-4, 635–650. MR 3473635 2

[4] Ronald R. Coifman and Yves Meyer, *Nonlinear harmonic analysis, operator theory and P.D.E*, Beijing lectures in harmonic analysis (Beijing, 1984), Ann. of Math. Stud., vol. 112, Princeton Univ. Press, Princeton, NJ, 1986, pp. 3–45. MR 864370 2

[5] David Cruz-Uribe, José María Martell, and Carlos Pérez, *Extrapolation from A_∞ weights and applications*, J. Funct. Anal. **213** (2004), no. 2, 412–439. MR 2078632 19

[6] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *Domination of multilinear singular integrals by positive sparse forms*, J. Lond. Math. Soc. (2) **98** (2018), no. 2, 369–392. MR 3873113 2, 17

[7] ———, *Uniform sparse domination of singular integrals via dyadic shifts*, Math. Res. Lett. **25** (2018), no. 1, 21–42. MR 3818613 5, 17

[8] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *A sparse estimate for multisublinear forms involving vector-valued maximal functions*, Bruno Pini Mathematical Analysis Seminar **8** (2018), no. 1, 168–184. 5, 18, 19

[9] Francesco Di Plinio, Kangwei Li, Henri Martikainen, and Emil Vuorinen, *Multilinear operator-valued Calderón-Zygmund theory*, J. Funct. Anal. **279** (2020), no. 8, 108666, 62. MR 4109091 3, 7, 10, 11

[10] Francesco Di Plinio and Yumeng Ou, *Banach-valued multilinear singular integrals*, Indiana Univ. Math. J. **67** (2018), no. 5, 1711–1763. MR 3875242 2, 3

[11] Yen Do, Richard Oberlin, and Eyvindur A. Palsson, *Variation-norm and fluctuation estimates for ergodic bilinear averages*, Indiana Univ. Math. J. **66** (2017), no. 1, 55–99. MR 3623404 2

[12] Yen Do and Christoph Thiele, *L^p theory for outer measures and two themes of Lennart Carleson united*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 249–296. MR 3312633 2

[13] Tadeusz Figiel, *Singular integral operators: a martingale approach*, Geometry of Banach spaces (Strobl, 1989), London Math. Soc. Lecture Note Ser., vol. 158, Cambridge Univ. Press, Cambridge, 1990, pp. 95–110. MR 1110189 2
[14] Tadeusz Figiel and Przemysław Wojtaszczyk, *Special bases in function spaces*, Handbook of the geometry of Banach spaces, Vol. I, North-Holland, Amsterdam, 2001, pp. 561–597. MR 1863702 27
[15] Stefan Geiss, Stephen Montgomery-Smith, and Eero Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. **362** (2010), no. 2, 553–575. MR 2551497 2
[16] Loukas Grafakos and José María Martell, *Extrapolation of weighted norm inequalities for multivariable operators and applications*, J. Geom. Anal. **14** (2004), no. 1, 19–46. MR 2030573 2
[17] Loukas Grafakos and Seungly Oh, *The Kato-Ponce inequality*, Comm. Partial Differential Equations **39** (2014), no. 6, 1128–1157. MR 3200091 5, 6
[18] Loukas Grafakos and Rodolfo H. Torres, *Multilinear Calderón-Zygmund theory*, Adv. Math. **165** (2002), no. 1, 124–164. MR 1880324 2, 10
[19] Timo S. Hänninen and Tuomas P. Hytönen, *Operator-valued dyadic shifts and the T(1) theorem*, Monatsh. Math. **180** (2016), no. 2, 213–253. MR 3502626 7, 9, 25, 27
[20] Tuomas Hytönen and Mikko Kemppainen, *On the relation of Carleson's embedding and the maximal theorem in the context of Banach space geometry*, Math. Scand. **109** (2011), no. 2, 269–284. MR 2854692 3
[21] Tuomas Hytönen, Alan McIntosh, and Pierre Portal, *Kato's square root problem in Banach spaces*, J. Funct. Anal. **254** (2008), no. 3, 675–726. MR 2381159 3
[22] Tuomas Hytönen and Pierre Portal, *Vector-valued multiparameter singular integrals and pseudodifferential operators*, Adv. Math. **217** (2008), no. 2, 519–536. MR 2370274 (2009e:42027) 2
[23] Tuomas Hytönen, Jan van Neerven, and Pierre Portal, *Conical square function estimates in UMD Banach spaces and applications to H^∞-functional calculi*, J. Anal. Math. **106** (2008), 317–351. MR 2448989 2, 3
[24] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 63, Springer, Cham, 2016. MR 3617205 8, 9, 30, 33
[25] ———, *Analysis in Banach spaces. Vol. II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 67, Springer, Cham, 2017, Probabilistic methods and operator theory. MR 3752640 8
[26] Tuomas Hytönen and Lutz Weis, *A T1 theorem for integral transformations with operator-valued kernel*, J. Reine Angew. Math. **599** (2006), 155–200. MR 2279101 3
[27] Tuomas P. Hytönen, *On operator-multipliers for mixed-norm $L^{\vec{p}}$ spaces*, Arch. Math. (Basel) **85** (2005), no. 2, 151–155. MR 2161804 (2006c:42011) 2
[28] Tuomas P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506. MR 2912709 1, 3
[29] ———, *The vector-valued nonhomogeneous Tb theorem*, Int. Math. Res. Not. IMRN (2014), no. 2, 451–511. MR 3159078 7, 9, 25
[30] Tuomas P. Hytönen, Michael T. Lacey, and Ioannis Parissis, *The vector valued quartile operator*, Collect. Math. **64** (2013), no. 3, 427–454. MR 3084406 3
[31] Marius Junge and Quanhua Xu, *Representation of certain homogeneous Hilbertian operator spaces and applications*, Invent. Math. **179** (2010), no. 1, 75–118. MR 2563760 2
[32] Nigel Kalton and Lutz Weis, *The H^∞-calculus and sums of closed operators*, Math. Ann. **321** (2001), no. 2, 319–345. MR 1866491 2
[33] Tosio Kato and Gustavo Ponce, *Commutator estimates and the Euler and Navier-Stokes equations*, Comm. Pure Appl. Math. **41** (1988), no. 7, 891–907. MR 951744 2
[34] Michael Lacey and Christoph Thiele, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$*, Ann. of Math. (2) **146** (1997), no. 3, 693–724. MR 1491450 2
[35] ———, *On Calderón's conjecture*, Ann. of Math. (2) **149** (1999), no. 2, 475–496. MR 1689336 2
[36] Michael T. Lacey and Darío Mena Arias, *The sparse T1 theorem*, Houston J. Math. **43** (2017), no. 1, 111–127. MR 3647935 18
[37] Kangwei Li, José María Martell, Henri Martikainen, Sheldy Ombrosi, and Emil Vuorinen, *End-point estimates, extrapolation for multilinear Muckenhoupt classes, and applications*, Trans. Amer. Math. Soc., to appear. 2, 5

[38] Kangwei Li, José María Martell, and Sheldy Ombrosi, *Extrapolation for multilinear muckenhoupt classes and applications*, Adv. Math., to appear. 5
[39] Kangwei Li, Henri Martikainen, Yumeng Ou, and Emil Vuorinen, *Bilinear representation theorem*, Trans. Amer. Math. Soc. **371** (2019), no. 6, 4193–4214. MR 3917220 3, 10, 11, 17, 18, 19
[40] Terry R. McConnell, *Decoupling and stochastic integration in UMD Banach spaces*, Probab. Math. Statist. **10** (1989), no. 2, 283–295. MR 1057936 7, 9
[41] Fedor Nazarov, Sergei Treil, and Alexander Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. **190** (2003), no. 2, 151–239. MR 1998349 2, 7
[42] Bas Nieraeth, *Quantitative estimates and extrapolation for multilinear weight classes*, Math. Ann. **375** (2019), no. 1-2, 453–507. MR 4000248 2
[43] Stefanie Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris Sér. I Math. **330** (2000), no. 6, 455–460. MR 1756958 1
[44] ———, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic*, Amer. J. Math. **129** (2007), no. 5, 1355–1375. MR 2354322 1
[45] Gilles Pisier, *Martingales in Banach spaces*, Cambridge Studies in Advanced Mathematics, vol. 155, Cambridge University Press, Cambridge, 2016. MR 3617459 14
[46] Gilles Pisier and Quanhua Xu, *Non-commutative L^p-spaces*, Handbook of the geometry of Banach spaces, Vol. 2, North-Holland, Amsterdam, 2003, pp. 1459–1517. MR 1999201 14, 15, 16
[47] Lutz Weis, *Operator-valued Fourier multiplier theorems and maximal L_p-regularity*, Math. Ann. **319** (2001), no. 4, 735–758. MR 1825406 2, 3, 7
[48] Quanhua Xu, *Noncommutative L^p-spaces and martingale inequalities*, 2007, Book manuscript. 15

(F.D.P.) Department of Mathematics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130-4899, USA
Email address: francesco.diplinio@gmail.com

Kangwei Li, Center for Applied Mathematics, Tianjin University, Weijin Road 92, 300072 Tianjin, China

and

BCAM, Basque Center for Applied Mathematics, Mazarredo 14, 48009 Bilbao, Basque Country, Spain
Email address: kli@tju.edu.cn

(H.M.) Department of Mathematics and Statistics, University of Helsinki, P.O.B. 68, FI-00014 University of Helsinki, Finland
Email address: henri.martikainen@helsinki.fi

(E.V.) Centre for Mathematical Sciences, University of Lund, P.O.B. 118, 22100 Lund, Sweden
Email address: j.e.vuorin@gmail.com