

# THE LOCAL NON-HOMOGENEOUS $Tb$ THEOREM FOR VECTOR-VALUED FUNCTIONS

TUOMAS P. HYTÖNEN AND ANTTI V. VÄHÄKANGAS

*Department of Mathematics and Statistics, Gustaf Hällströmin katu 2b, FI-00014 University of Helsinki, Finland*

*e-mail: tuomas.hytonen@helsinki.fi; e-mail: antti.vahakangas@helsinki.fi*

(Received 7 March 2012; revised 10 September 2013; accepted 14 September 2013; first published online 26 August 2014)

**Abstract.** We extend the local non-homogeneous  $Tb$  theorem of Nazarov, Treil and Volberg to the setting of singular integrals with operator-valued kernel that act on vector-valued functions. Here, ‘vector-valued’ means ‘taking values in a function lattice with the UMD (unconditional martingale differences) property’. A similar extension (but for general UMD spaces rather than UMD lattices) of Nazarov–Treil–Volberg’s global non-homogeneous  $Tb$  theorem was achieved earlier by the first author, and it has found applications in the work of Mayboroda and Volberg on square-functions and rectifiability. Our local version requires several elaborations of the previous techniques, and raises new questions about the limits of the vector-valued theory.

2010 *Mathematics Subject Classification.* 42B20 (Primary); 42B25, 46E40, 60G46 (Secondary)

## 1. Introduction.

**Background and motivation.** This paper is a continuation of [4], where the first author extended the ‘global’ non-homogeneous  $Tb$  theorem of Nazarov, Treil and Volberg [12] to  $L^p$  spaces of vector-valued functions. The goal of the paper at hand is to obtain a similar extension for the ‘local’ version of Nazarov, Treil and Volberg’s result [13].

By ‘local’ we understand that the  $Tb$  conditions involve a family (an ‘accretive system’) of testing functions  $b_Q$ , one for each cube  $Q$ , where  $b_Q$  is only required to satisfy a non-degeneracy condition on its ‘own’  $Q$ ; this contrasts with the ‘global’  $Tb$  conditions, where a single testing function  $b$  should be appropriately non-degenerate over all positions and length-scales. While the two types of  $Tb$  theorems are not strictly comparable, the verification of the local conditions has turned out more approachable in a number of applications.

By ‘vector-valued’ we understand functions taking values in a possibly infinite-dimensional Banach space  $X$ . It is well known that the most general class of Banach spaces in which extensions of deeper results in harmonic analysis can be hoped for consists of the spaces with the UMD property (unconditionality of martingale

---

T.P.H. was supported by the Academy of Finland, grants 1130166, 1133264 and 1218148, and by the European Union through the ERC Starting grant ‘Analytic–probabilistic methods for borderline singular integrals’. A.V.V. was supported by the Academy of Finland, grants 75166001 and 1134757.

differences); see [1, 3]. The quest for vector-valued extensions of theorems in classical analysis has three types of motivation:

First, by revisiting a proof in a more general framework we can often develop new insight into the original argument; in particular, the tools available in an abstract UMD space often lead us into discovering new martingale structure behind the classical scene. In the present case, for example, we are led to study the  $L^p$  estimates for martingale difference expansions adapted to an accretive system of functions, where mainly the Hilbert space  $L^2$  theory for such expansions existed so far. While the  $L^p$  theory for the ‘globally’ adapted martingale differences was developed in [4], the local setting brings several new complications, most prominently the fact that the expansion is no longer with respect to a basis of adapted Haar functions but rather with respect to an overdetermined frame.

Second, new connections between different properties of Banach spaces are revealed, when looking for the minimal conditions under which we can run a given classical analysis argument. In particular, for vector-valued functions, there appears a subtle difference between the square function estimates for the adapted martingale difference operators and for their adjoints, and we are only able to handle the latter case under the additional assumption that our Banach space is a function lattice. Whether this assumption could be eliminated from certain key inequalities raises interesting questions for further investigation.

Finally, the extended scope of the theorem allows for wider applications. The applications of vector-valued singular integrals in general are widespread; for the vector-valued non-homogeneous  $Tb$  theorem [4] in particular, we mention the work of Mayboroda and Volberg [10] on square functions and rectifiability, see [10, p. 1056].

Now, let us turn to a more detailed discussion of the objects of this paper.

**Calderón–Zygmund operators.** Let  $\mu$  be a compactly supported Borel measure on  $\mathbb{R}^N$  which satisfies the upper bound

$$\mu(B(x, r)) \leq r^d, \quad d \in (0, N], \quad (1.1)$$

for any ball  $B(x, r)$  of centre  $x \in \mathbb{R}^N$  and radius  $r > 0$ . A  $d$ -dimensional Calderón–Zygmund kernel is a complex-valued function  $K(x, y)$  of variables  $x, y \in \mathbb{R}^N$ ,  $x \neq y$ , such that

$$|K(x, y)| \leq \frac{1}{|x - y|^d} \quad (1.2)$$

and, if  $2|x - x'| \leq |x - y|$ , then

$$|K(x, y) - K(x', y)| + |K(y, x) - K(y, x')| \leq \frac{|x - x'|^\alpha}{|x - y|^{d+\alpha}} \quad (1.3)$$

for some  $\alpha > 0$ . An operator  $T$  acting on some functions is called a Calderón–Zygmund operator with kernel  $K$  if

$$Tf(x) = \int_{\mathbb{R}^N} K(x, y)f(y) d\mu(y), \quad x \notin \text{supp } f. \quad (1.4)$$

**Testing functions.** Following [13] we say that a collection  $\{b_Q\} \subset L^\infty(\mathbb{R}^N, \mu; \mathbb{C})$  of functions is an  $L^\infty$ -accretive system (supported on cubes) if for every cube  $Q$  in  $\mathbb{R}^N$  there exists a function  $b_Q$  from the system such that

$$\text{supp } b_Q \subset Q, \quad \|b_Q\|_{L^\infty(\mu)} \leq 1, \quad \left| \int_Q b_Q d\mu \right| \geq \delta \mu(Q). \tag{1.5}$$

Here the constant  $\delta \in (0, 1)$  is not allowed to depend on  $Q$ .

We say that  $\{b_Q\}$  is an  $L^\infty$ -accretive system for a Calderón–Zygmund operator  $T$  if, for every cube  $Q$  in  $\mathbb{R}^N$ , there is a function  $b_Q$  from the system such that the conditions (1.5) hold true and

$$\|Tb_Q\|_{L^\infty(\mu)} \leq B, \tag{1.6}$$

where  $B > 0$  is a constant that is independent of  $Q$ .

**Banach spaces.** We want to study the action of  $T$  as in (1.4) on the Bochner space  $L^p(\mathbb{R}^N, \mu; X)$  of functions with values in the Banach space  $X$ . As is well-known, even for the simplest non-trivial case where  $T$  is the Hilbert transform with  $d = N = 1$  and  $\mu$  is the Lebesgue measure, a necessary condition for the boundedness on the Bochner space is that  $X$  be a UMD space [1]. For the more complicated operators as described, we will need to assume some further conditions.

We will make the more restrictive assumption that  $X$  is a UMD function lattice, i.e.,  $X$  is a UMD space whose elements are represented by functions on some measure space, and the norm of  $X$  is compatible with the pointwise comparison of functions in that  $|f| \leq |g|$  pointwise implies that  $\|f\|_X \leq \|g\|_X$ . See [14] for more information on function lattices with the UMD property. We will make use of this assumption both directly, via Theorem 4.16, and through the following consequence established by Hytönen, McIntosh and Portal [7]: such spaces satisfy the so-called RMF property, also introduced in [7], which means the boundedness of the so-called Rademacher maximal function from  $L^p(\mathbb{R}^N, \mu; X)$  to  $L^p(\mathbb{R}^N, \mu)$ . A detailed study of this property can be found in Kempainen [8]. The RMF property is used to estimate the so-called paraproducts arising in the proof of the  $Tb$  theorem; for the same purpose, RMF was also assumed in an earlier version of the global non-homogeneous  $Tb$  theorem [4], but it was subsequently circumvented there. In addition, we make explicit use of the lattice structure at one specific point of the proof to obtain a certain auxiliary square function estimate. We do not know about the necessity of this assumption, so it seems interesting to single out the place where we use it for possible further investigation. Note that many of the concrete UMD spaces appearing in harmonic analysis, like the  $L^p$  and Lorentz spaces, are all lattices; others, like Sobolev spaces, can still be identified with closed subspaces of such lattices, e.g. for  $\Omega \subset \mathbb{C}^N$ , we have  $W^{1,p}(\Omega) \simeq \{(f, g) \in L^p(\Omega) \times L^p(\Omega)^N : g = \nabla f\} \subset L^p(\Omega)^{N+1} \simeq L^p(\bigcup_{i=0}^N \Omega_i)$ , where the  $\Omega_i$ 's are disjoint copies of  $\Omega$ ; but some other examples of UMD spaces like the Schatten ideals  $C_p$  fall outside this class of spaces.

We are ready to formulate our main result.

**THEOREM 1.7.** *Suppose that  $X$  is a UMD function lattice. Assume that  $T$  is a Calderón–Zygmund operator, and that there exists two  $L^\infty$ -accretive systems of complex-valued functions,  $b^1 = \{b^1_Q\}$  for  $T$  and  $b^2 = \{b^2_Q\}$  for  $T^*$ . Then, under the qualitative a*

*priori assumption that  $T \in \mathcal{L}(L^p(\mathbb{R}^N, \mu; X))$  for some  $p \in (1, \infty)$ , we have the quantitative bound*

$$\|T\|_{\mathcal{L}(L^p(\mathbb{R}^N, \mu; X))} \leq C,$$

where the constant  $C = C(N, d, \alpha, \delta, B, p, X) > 0$  is independent of  $T$ .

Having stated this, we should admit two things. First, this result remains valid for general UMD spaces. Second, it follows relatively easily, even in the just mentioned more general form, from a combination of the results of [4], [13] and [12]. Namely, Nazarov, Treil and Volberg's local  $Tb$  theorem [13] states that under the mentioned assumptions we have the scalar valued bound  $\|T\|_{\mathcal{L}(L^2(\mathbb{R}^N, \mu))} \leq C$ . Then, the converse direction of Nazarov, Treil and Volberg's global  $Tb$  theorem [12] tells that  $T$  satisfies the global  $Tb$  (or even  $T1$  conditions)  $\|T1\|_{\text{BMO}(\mathbb{R}^N, \mu)} + \|T^*1\|_{\text{BMO}(\mathbb{R}^N, \mu)} \leq C$ , where  $\text{BMO}(\mathbb{R}^N, \mu)$  is an appropriate bounded mean oscillation space adapted to the non-homogeneous situation. Finally, the vector-valued global  $Tb$  theorem of [4] completes the argument, as we have just checked that its assumptions are satisfied.

What, then, is the point of struggling for a weaker statement, when a stronger one is available for free? Sure, we can still develop some new insight into the proof technique of [13], but there is also a more substantial reason on the level of actual results. Namely, the proof of Theorem 1.7 that we give immediately yields a further generalization to the case of operator-valued kernels  $K$ , i.e., for kernels  $K(x, y) \in \mathcal{L}(X)$ . Then the associated  $\mathcal{L}(X)$ -valued operator  $T$  in (1.4) is genuinely an object of the vector-valued realm, and the above shortcut via the scalar-valued theory is no longer available.

**Rademacher–Calderón–Zygmund operators.** Let us consider an operator  $T$  given by the same formula (1.4) as before, but with  $K(x, y) \in \mathcal{L}(X)$ . The kernel bounds (1.2) and (1.3) will have to be replaced by certain operator-theoretic analogues involving the notion of  $\mathcal{R}$ -boundedness (see definition in (2.2)), and we refer the reader to Section 14 for a precise statement. We then say that  $K$  is a  $d$ -dimensional Rademacher–Calderón–Zygmund kernel, and that  $T$  is an  $\mathcal{L}(X)$ -valued Rademacher–Calderón–Zygmund operator. For further details, we refer to Section 14.

The testing functions are now as follows. We say that  $\{b_Q^1\}$  is an  $L^\infty$ -accretive system for an  $\mathcal{L}(X)$ -valued Rademacher–Calderón–Zygmund operator  $T$  if, for every cube  $Q$  in  $\mathbb{R}^N$ , there is a function  $b_Q^1$  from the system such that the conditions (1.5) hold true for  $b_Q = b_Q^1$  and  $Tb_Q^1 : \mathbb{R}^N \rightarrow Y$  satisfies  $\|Tb_Q^1\|_{L^\infty(\mathbb{R}^N, \mu; Y)} \leq B$ , where  $Y \subset \mathcal{L}(X)$  is an UMD function lattice which has cotype 2 (see definition in (2.3)), and whose unit ball  $\bar{B}_Y$  is an  $\mathcal{R}$ -bounded subset of  $\mathcal{L}(X)$ . In a similar manner, we say that  $\{b_Q^2\}$  is an  $L^\infty$ -accretive system for  $T^*$  if, for every cube  $Q$  in  $\mathbb{R}^N$ , there is a function  $b_Q^2$  from the system such that the conditions (1.5) hold true for  $b_Q = b_Q^2$  and  $T^*b_Q^2 : \mathbb{R}^N \rightarrow Z$  satisfies  $\|T^*b_Q^2\|_{L^\infty(\mathbb{R}^N, \mu; Z)} \leq B$ , where  $Z \subset \mathcal{L}(X^*)$  is an UMD function lattice which has cotype 2 and whose unit ball  $\bar{B}_Z$  is an  $\mathcal{R}$ -bounded subset of  $\mathcal{L}(X^*)$ .

The following local  $Tb$  theorem for operator-valued kernels is obtained by employing the entire power of the proof of Theorem 1.7, with minor necessary adjustments. Unlike Theorem 1.7, it cannot be obtained by a shortcut from the scalar-valued  $Tb$  theorem of Nazarov, Treil and Volberg [13]. We do not know if the function lattice assumption is necessary in the following theorem; we will make some comments on its usage in the course of the proof.

**THEOREM 1.8.** *Suppose that  $X$  is a UMD function lattice. Assume that  $T$  is an  $\mathcal{L}(X)$ -valued Rademacher–Calderón–Zygmund operator, and that there exists two  $L^\infty$ -accretive systems,  $b^1 = \{b_Q^1\}$  for  $T$  and  $b^2 = \{b_Q^2\}$  for  $T^*$ . Then, under the qualitative a priori assumption that  $T \in \mathcal{L}(L^p(\mathbb{R}^N, \mu; X))$  for some  $p \in (1, \infty)$ , we have the quantitative bound*

$$\|T\|_{\mathcal{L}(L^p(\mathbb{R}^N, \mu; X))} \leq C,$$

where the constant  $C = C(N, d, \alpha, \delta, B, p, X, Y, Z) > 0$  is independent of  $T$ .

Concerning the interest and potential applicability of such a result over the simpler Theorem 1.7, we make the following remarks. First, in applying the global vector-valued  $Tb$  theorem from [4], Mayboroda and Volberg [10, p. 1056] specifically use the operator-kernel version [4,  $Tb$  theorem 4]. Second, in the mentioned application, all the Banach spaces are function lattices, so that this assumption is not too restrictive for such purposes. We also recall, although this is not directly connected to the non-homogeneous issues at hand, that the theory of singular integrals with operator-kernel has been a necessary strengthening of the vector-valued scalar-kernel theory in applications like the maximal regularity question for partial differential equations; see in particular the influential paper [17].

**Organization of the paper.** In order to keep the notation somewhat lighter, we will concentrate in the main body of the paper on the proof of Theorem 1.7 about scalar-valued kernels. Mostly, however, this argument goes through without trouble for the operator-kernel version of Theorem 1.8 as well, and we only explain a few necessary modifications in the final Section 14. After collecting some preliminaries in Section 2, the proof of Theorem 1.7 is presented in Sections 3 through 13:

Sections 3 and 4 present a detailed analysis, and related inequalities, of functions  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$  in terms of appropriate adapted martingale difference operators  $D_Q^a$ . In Section 5, these expansions of functions then give a representation of the operator  $T$  in terms of matrix elements  $T_{RQ}$ , where  $R$  and  $Q$  range over dyadic cubes, and the rest of the proof is concerned with the estimation of different parts of this matrix.

Section 6 presents a general martingale decoupling inequality — our best substitute for orthogonality estimates in  $L^2$  —, which will be used several times during the proof. The parts of the matrix  $T_{RQ}$  leading to different types of treatment are as follows: the separated cubes (handled in Section 7), the deeply nested cubes (Sections 8 through 10, where the last one deals with the paraproduct part of the operator), and the nearby cubes of comparable size (Sections 11 and 12). Finally, Section 13 collects the different estimates together, and also takes care of the remaining ‘bad’ cubes which were excluded from the previous cases.

## 2. Preparations.

**Notation.** We denote  $\mathbb{N} = \{1, 2, \dots\}$  and  $\mathbb{N}_0 = \{0, 1, \dots\}$ . All distances in  $\mathbb{R}^N$  are measured in terms of the supremum norm, defined by  $|x| = \|x\|_\infty$  for  $x \in \mathbb{R}^N$ . Accordingly, we henceforth write  $B(x, t)$  for the  $\ell^\infty$  ball in  $\mathbb{R}^N$  centered at  $x$  with radius  $t > 0$ . (Note that the main assumption (1.1) is still true, possibly after scaling  $\mu$  by

a constant.) We assume that  $K$  is a  $d$ -dimensional Calderón–Zygmund kernel in  $\mathbb{R}^N$ , satisfying both (1.2) and (1.3) for some  $\alpha > 0$ . In the sequel  $r > 0$  is a (large) integer which is to be quantified later. We fix a constant  $\gamma$ ,

$$\gamma \in (0, 1), \quad d\gamma/(1 - \gamma) \leq \alpha/4, \quad \gamma \leq \frac{\alpha}{2(d + \alpha)}. \tag{2.1}$$

We denote

$$\theta(j) = \left\lceil \frac{\gamma j + r}{1 - \gamma} \right\rceil \quad \text{for } j = 0, 1, 2, \dots$$

A cube  $Q$  in  $\mathbb{R}^N$  has sides parallel to the coordinate axes, and its side length is denoted by  $\ell(Q)$ . If  $Q, R \subset \mathbb{R}^N$  are cubes, their *long distance*  $D(Q, R)$  is defined by  $D(Q, R) = \ell(Q) + \text{dist}(Q, R) + \ell(R)$ .

**UMD and  $\mathcal{R}$ -boundedness.** A Banach space  $X$  has the UMD property if there is a constant  $C > 0$  so that

$$\left\| \sum_{k=1}^n \varepsilon_k d_k \right\|_{L^p(\mu; X)} \leq C \left\| \sum_{k=1}^n d_k \right\|_{L^p(\mu; X)}$$

whenever  $(d_k)_{k=1}^n$  is a martingale difference sequence in  $L^p(\mu; X)$  and  $\varepsilon_k = \pm 1$ . This property is known to be independent of the parameter  $p \in (1, \infty)$ , and its validity for dyadic martingales with respect to the Lebesgue measure already implies the general condition, [9]. UMD spaces are reflexive.

Let  $(\varepsilon_k)_{k \in \mathbb{Z}}$  be a sequence of Rademacher functions, i.e., a sequence of independent random variables attaining values  $\pm 1$  with an equal probability  $\mathbf{P}(\varepsilon_k = -1) = \mathbf{P}(\varepsilon_k = 1) = 1/2$ . By  $\Omega$  we denote the probability space supporting the distribution of  $(\varepsilon_k)_{k \in \mathbb{Z}}$ . The Khintchine–Kahane inequality says that

$$\left\| \sum_{k=1}^n \varepsilon_k \xi_k \right\|_{L^p(\Omega; X)} \approx \left\| \sum_{k=1}^n \varepsilon_k \xi_k \right\|_{L^2(\Omega; X)}$$

for all  $p \in (0, \infty)$ . For  $X = \mathbb{C}$ , this is called just Khintchine’s inequality, and the right hand side can be written as the quadratic expression  $(\sum_{k=1}^n |\xi_k|^2)^{1/2}$ . Because of this, inequalities for the random series involving the  $\varepsilon_k$  are often referred to as ‘square-function’ estimates even in the vector-valued case, even if no squares explicitly appear.

We recall that an operator family  $\mathcal{T} \subset \mathcal{L}(X_1, X_2)$  is called *Rademacher-bounded*, or  $\mathcal{R}$ -bounded, if there is a constant  $C$  such that for all  $n \in \mathbb{N}$ , all  $\xi_1, \dots, \xi_n \in X_1$  and all  $T_1, \dots, T_n \in \mathcal{T}$ ,

$$\left\| \sum_{k=1}^n \varepsilon_k T_k \xi_k \right\|_{L^2(\Omega; X_2)} \leq C \left\| \sum_{k=1}^n \varepsilon_k \xi_k \right\|_{L^2(\Omega; X_1)}. \tag{2.2}$$

Denote the smallest admissible  $C$  by  $\mathcal{R}(\mathcal{T})$ .

We will often use the following *Stein’s inequality* (more precisely, its vector-valued extension due to Bourgain [2]), which says that an increasing sequence of conditional

expectations  $E_k$  is  $\mathcal{R}$ -bounded on  $L^p(\mathbb{R}^N, \mu; X)$  if  $X$  is a UMD space:

$$\left\| \sum_{k=1}^n \varepsilon_k E_k f_k \right\|_{L^p(\Omega \times \mathbb{R}^N; X)} \leq C \left\| \sum_{k=1}^n \varepsilon_k f_k \right\|_{L^p(\Omega \times \mathbb{R}^N; X)}$$

with constant  $C > 0$  independent of the sequence  $(f_k)_{k=1}^n \subset L^p(\mathbb{R}^N, \mu; X)$ .

A different condition arises by requiring the pointwise (in  $x \in \mathbb{R}^N$ )  $\mathcal{R}$ -boundedness of the sequence of vectors  $E_k f(x) \in X \simeq \mathcal{L}(\mathbb{C}, X)$ , where the last identification is the obvious one:  $\xi \in X$  is identified with the operator  $\lambda \in \mathbb{C} \mapsto \lambda \xi \in X$ . We denote

$$M_R f(x) := \mathcal{R}(\{E_k f(x) : k \in \mathbb{Z}\}),$$

and say that  $X$  has the *RMF (Rademacher maximal function) property*, if  $M_R : L^p(\mathbb{R}^N, \mu; X) \rightarrow L^p(\mathbb{R}^N, \mu)$  boundedly for some (and then all)  $p \in (1, \infty)$ . This notion was introduced in [7]; see [5, 7, 8] for more information.

**Cotype of a Banach space.** A Banach space  $X$  is said to have cotype  $s \in [2, \infty)$ , i.e., there is a constant  $C > 0$  such that for all sequences  $(\xi_j)_{j=1}^n$  in  $X$  we have

$$\left( \sum_{j=1}^n |\xi_j|_X^s \right)^{1/s} \leq C \left\| \sum_{j=1}^n \varepsilon_j \xi_j \right\|_{L^2(\Omega; X)}. \tag{2.3}$$

This leads to an improvement of the contraction principle, [4, Proposition 11.4].

**PROPOSITION 2.4.** *Let  $X$  be a Banach space of cotype  $s \in [2, \infty)$  and suppose that  $\{\rho_j : j \in \mathbb{N}\} \subset L^t(\tilde{\Omega})$  for some  $\sigma$ -finite measure space  $\tilde{\Omega}$  and  $t \in (s, \infty)$ . Then*

$$\left\| \sum_{j=1}^{\infty} \varepsilon_j \rho_j \xi_j \right\|_{L^t(\tilde{\Omega}; L^2(\Omega; X))} \lesssim \sup_j \|\rho_j\|_{L^t(\tilde{\Omega})} \cdot \left\| \sum_{j=1}^{\infty} \varepsilon_j \xi_j \right\|_{L^2(\Omega; X)}$$

if  $\{\xi_j : j \in \mathbb{N}\} \subset X$ .

Some of the subsequent estimates are based on the fact that every UMD has cotype  $s$  for some  $s \in [2, \infty)$ . (See e.g. [14, p. 202] for the stronger property that every UMD has type  $t$  for some  $t \in (1, 2]$ .)

**Generic dyadic systems.** Let  $\hat{\mathcal{D}}$  denote the standard dyadic system, consisting of all of the cubes of the form  $2^k(m + [0, 1]^N)$ , where  $k \in \mathbb{Z}$  and  $m \in \mathbb{Z}^N$ . We also denote

$$\hat{\mathcal{D}}_k = \{Q \in \hat{\mathcal{D}} : \ell(Q) = 2^k\}.$$

A generic dyadic system, parametrised by  $\beta \in (\{0, 1\}^N)^{\mathbb{Z}}$ , is of the form

$$\mathcal{D}(\beta) = \bigcup_{k \in \mathbb{Z}} \mathcal{D}_k(\beta),$$

where

$$\mathcal{D}_k(\beta) = \{\hat{R} + x_k(\beta) : \hat{R} \in \hat{\mathcal{D}}_k\}, \quad x_k(\beta) = \sum_{j < k} \beta_j 2^j.$$

Given  $Q \in \mathcal{D}(\beta)$  and  $n \in \mathbb{N}_0$ , then the expression  $Q^{(n)}$  denotes the dyadic ancestor of  $Q$  of the  $n$ 'th generation, i.e., it is the unique cube such that  $Q \subset Q^{(n)} \in \mathcal{D}(\beta)$  and  $\ell(Q^{(n)}) = 2^n \ell(Q)$ .

**Random dyadic systems.** The generic dyadic systems give rise to random dyadic systems by assigning the complete product probability measure  $\mathbf{P}_\beta$  on the set  $(\{0, 1\}^{\mathbb{Z}})^{\mathbb{Z}}$  so that the coordinate functions  $\beta_j, j \in \mathbb{Z}$ , are independent and  $\mathbf{P}_\beta[\beta_j = \eta] = 2^{-N}$  if  $\eta \in \{0, 1\}^N$ .

REMARK 2.5. As many (but not all) papers in the area, we will use two independent random dyadic systems  $\mathcal{D} = \mathcal{D}(\beta)$  and  $\mathcal{D}' = \mathcal{D}(\beta')$  to insert randomization into the argument. It seems to us that this is a necessary technicality in the present context, see Remark 12.22 for a critical point in the proof.

Altogether, it seems that the most elaborate versions of  $T1$  and  $Tb$  theorems need two independent systems. For example, consider the non-homogeneous  $T1$  theorem, even just the scalar version. The original global  $T1$  (actually,  $Tb$ ) formulation was proved by Nazarov, Treil and Volberg [12] by using two independent systems. Now there is a more recent local  $T1$  theorem by Volberg [16], proven by just one random system, but it appears that this version is not quite as strong as the ‘global  $T1$ ’. Indeed, Volberg makes the assumption

$$\|T1_Q\|_{L^2(\mu)}^2 + \|T^*1_Q\|_{L^2(\mu)}^2 \leq C\mu(Q),$$

where the  $L^2$ -norms are computed over the whole space and estimated by the measure of  $Q$ , whereas the global ‘ $T1 \in \text{BMO}$ ’ conditions of Nazarov, Treil and Volberg can be easily deduced from the local assumptions

$$\|1_Q T1_Q\|_{L^2(\mu)}^2 + \|1_Q T^*1_Q\|_{L^2(\mu)}^2 \leq C\mu(\lambda Q), \quad \lambda > 1.$$

Here the  $L^2$ -norms are local, and estimated by the (possibly much larger) measure of the expanded cube  $\lambda Q$ .

Let  $n \in \mathbb{Z}$ . A cube  $Q \in \mathcal{D}$  is called  $n$ -bad (w.r.t.  $\mathcal{D}'$ ) if there exists  $R \in \mathcal{D}'$  such that

$$\ell(Q) \leq 2^{-(n\vee r)} \ell(R), \quad \text{dist}(Q, \partial R) \leq \ell(Q)^\gamma \ell(R)^{1-\gamma}.$$

If  $Q$  is not  $n$ -bad (w.r.t.  $\mathcal{D}'$ ) then it is  $n$ -good (w.r.t.  $\mathcal{D}'$ ). The set of  $n$ -good cubes in  $\mathcal{D}$  is denoted by

$$\mathcal{D}_{n\text{-good}} = \mathcal{D}_{n\text{-good}(\gamma,r)}.$$

The family of  $n$ -bad cubes in  $\mathcal{D}$  is denoted by  $\mathcal{D}_{n\text{-bad}} = \mathcal{D}_{n\text{-bad}(\gamma,r)}$ . A cube  $Q \in \mathcal{D}_i$  is  $R$ -bad,  $R \in \mathcal{D}'_j$ , if  $Q$  is  $(j - i - 1)$ -bad. A cube  $Q \in \mathcal{D}$  is  $R$ -good if it is not  $R$ -bad. The family of  $R$ -good cubes in  $\mathcal{D}$  is denoted by  $\mathcal{D}_{R\text{-good}} = \mathcal{D}_{R\text{-good}(\gamma,r)}$ . The family of  $R$ -bad cubes in  $\mathcal{D}$  is denoted by  $\mathcal{D}_{R\text{-bad}} = \mathcal{D}_{R\text{-bad}(\gamma,r)}$ .

In a symmetric manner we define the  $n/Q$ -bad and  $n/Q$ -good cubes in  $\mathcal{D}'$ .

REMARK 2.6. Assume that  $Q \in \mathcal{D}$  is  $\bar{R}$ -good, where  $\bar{R} \in \mathcal{D}'$ . Then

$$\text{dist}(Q, \partial \bar{R}) > \ell(Q)^\gamma \ell(\bar{R})^{1-\gamma} \geq \ell(Q)^{\alpha/2(d+\alpha)} \ell(\bar{R})^{1-\alpha/2(d+\alpha)}.$$



for every  $R \in \mathcal{D}'$  satisfying  $\ell(R) \geq 2^{-1}\ell(\bar{R}) \vee 2^r\ell(Q)$ . The second inequality follows from the estimate  $\gamma \leq \alpha/2(d + \alpha)$  in (2.1).

Here is a useful lemma controlling the probability of bad cubes:

LEMMA 2.7. *Let  $n \in \mathbb{Z}$  and  $Q \in \mathcal{D} = \mathcal{D}(\beta)$  be fixed. Then*

$$\mathbf{P}_{\beta'}[Q \in \mathcal{D}_{n\text{-bad}(\gamma,r)}] \leq 2N \frac{2^{-(r \vee n)\gamma}}{1 - 2^{-\gamma}}.$$

*Proof.* Just follow the proof of [13, Lemma 7.1] with  $n \vee r$  in place of  $r$ . □

Various estimates are conducted while keeping the parameters  $\beta, \beta' \in (\{0, 1\}^{\mathbb{N}})^{\mathbb{Z}}$ , and hence also the associated dyadic systems, fixed. During these estimates, we will assume that these fixed dyadic systems satisfy the following condition: there are (fixed) cubes  $Q_0 \in \mathcal{D}(\beta)$  and  $R_0 \in \mathcal{D}(\beta')$  for which

$$\ell(Q_0) = \ell(R_0) = 2^s \text{ and } \text{supp } \mu \subset Q_0 \cap R_0. \tag{2.8}$$

From the probabilistic point-of-view this assumption is justified by the following lemma, when applied to the compact set  $K = \text{supp } \mu$ :

LEMMA 2.9. *Let  $K \subset \mathbb{R}^N$  be a bounded set. Denote by  $A$  the set of parameters  $\sigma \in (\{0, 1\}^{\mathbb{N}})^{\mathbb{Z}}$  for which  $K$  is not contained in any cube  $R \in \mathcal{D}(\sigma)$ . Then  $\mathbf{P}(A) = 0$ .*

*Proof.* By using completeness of  $\mathbf{P}$ , it suffices to show that  $A$  is contained in a set of probability zero. To this end, we use the fact that  $K$  is bounded as follows: there are dyadic cubes  $Q_1, \dots, Q_{2^N} \in \hat{\mathcal{D}}_k$  from the standard dyadic system of sufficiently large generation  $k \in \mathbb{Z}$  such that  $K \subset \cup_{j=1}^{2^N} Q_j$ . Thus, if  $\sigma \in A$  and  $n \in \mathbb{N}, n \geq r$ , there exists an index  $j \in \{1, 2, \dots, 2^N\}$  and a cube  $R \in \mathcal{D}_{k+n}(\sigma)$  so that  $\text{dist}(Q_j, \partial R) = 0$ . Hence  $Q_j \in \hat{\mathcal{D}}_{n\text{-bad}(\gamma,r)}$  with respect to  $\mathcal{D}(\sigma)$ . We have shown that

$$A \subset \bigcap_{n \geq r} \bigcup_{j=1}^{2^N} \{\sigma : Q_j \in \hat{\mathcal{D}}_{n\text{-bad}(\gamma,r)} \text{ w.r.t. } \mathcal{D}(\sigma)\}.$$

Using Lemma 2.7, we see that the probability of the right hand side is zero. □

**Layers of cubes.** Following [13] we will define certain layers of cubes in a given dyadic system  $\mathcal{D}$ . For this purpose, we fix  $\beta, \beta' \in (\{0, 1\}^{\mathbb{N}})^{\mathbb{Z}}$ , and assume that  $Q_0 \in \mathcal{D} = \mathcal{D}(\beta)$  and  $R_0 \in \mathcal{D}' = \mathcal{D}(\beta')$  are cubes such that (2.8) holds true. By  $s \in \mathbb{Z}$  we denote a sufficiently large integer for which  $Q_0 \in \mathcal{D}_s$  and  $R_0 \in \mathcal{D}'_s$ .

Let  $\mathcal{L}_0 = \{Q_0\}$  be the *zeroth layer of cubes*. Assume that the layers  $\mathcal{L}_0, \dots, \mathcal{L}_{j-1}$  of cubes have been defined. We then define the  *$j$ 'th layer of cubes*  $\mathcal{L}_j$  as follows. If  $\mathcal{L}_{j-1} = \emptyset$ , we set  $\mathcal{L}_j = \emptyset$ . Otherwise we consider a cube  $R \in \mathcal{L}_{j-1}$ . We say  $Q \in \mathcal{D}$  is  *$R$ -maximal*, if it is the maximal cube in  $\mathcal{D}$  satisfying the conditions  $Q \subsetneq R$  and

$$\left| \int_Q b_R^1 d\mu \right| < \delta^2 \mu(Q).$$

Then we denote  $\mathcal{L}_j = \cup_{R \in \mathcal{L}_{j-1}} \{Q \in \mathcal{D} : Q \text{ is } R\text{-maximal}\}$ . By analogy, we define the layers  $\mathcal{L}'_j$  associated with the system  $\mathcal{D}'$ .

Assuming that  $\mathcal{D} \ni Q \subset Q_0$ , we denote by  $Q^a$  the smallest cube in  $\cup_{j \geq 0} \mathcal{L}_j$  that contains  $Q$ . Such a cube exists because  $Q \subset Q_0 \in \cup_{j \geq 0} \mathcal{L}_j$ , and it is also unique due to properties of dyadic cubes; hence  $Q^a$  is well defined. If  $Q \not\subset Q_0$  then we denote  $Q^a = Q_0$  for the sake of convenience. Note that, in any case, we have

$$\left| \int_Q b_{Q^a}^1 d\mu \right| \geq \delta^2 \mu(Q). \tag{2.10}$$

In an analogous manner, we define  $R^a \in \cup_{j \geq 0} \mathcal{L}'_j$  for cubes  $R \in \mathcal{D}'$ .

For a fixed  $j \in \mathbb{N}$ , we have the estimate

$$\mu\left(\bigcup_{R \in \mathcal{L}_j: R \subsetneq Q} R\right) \leq (1 - \tau)\mu(Q) \quad \text{for } Q \in \mathcal{L}_{j-1}. \tag{2.11}$$

A proof is in [13, pp. 269–270]. Here  $\tau \in (0, 1)$  is a constant, depending only on  $\delta_b$ . This estimate generalises by simple iteration as follows:

LEMMA 2.12. *Let  $Q \in \mathcal{D}$  and  $Q^a \in \mathcal{L}_M$  for  $M \in \mathbb{N}_0$ . Then, for  $j \geq 1$ , we have the estimate*

$$\mu\left(\bigcup_{S \in \mathcal{L}_{M+j}: S \subsetneq Q} S\right) = \sum_{S \in \mathcal{L}_{M+j}: S \subsetneq Q} \mu(S) \leq (1 - \tau)^j \mu(Q).$$

REMARK 2.13. Lemma 2.12 yields that  $\mu$ -a.e. point  $x \in Q_0$  belongs to at most finitely many cubes in the family  $\cup_{j \in \mathbb{N}_0} \mathcal{L}_j$ . To prove this, let us denote  $f = \sum_{j=0}^\infty \sum_{Q \in \mathcal{L}_j} \chi_Q$ . The cubes in  $\mathcal{L}_j$  for  $j \in \mathbb{N}$  are disjoint, and they are all included in  $Q_0 = Q_0^a \in \mathcal{L}_0$ . Using Lemma 2.12 with  $Q = Q_0$  and  $M = 0$ , we get

$$\|f\|_{L^1(Q_0)} \leq \sum_{j=0}^\infty \sum_{Q \in \mathcal{L}_j} \mu(Q) \leq \mu(Q_0) \left(1 + \sum_{j=0}^\infty (1 - \tau)^j\right) < \infty. \tag{2.14}$$

The claim follows.

**Carleson embeddings.** Let  $\mathcal{D}$  be a generic dyadic system and let  $d_k \in L^1(\mathbb{R}^N, \mu; \mathbb{R})$ ,  $k \in \mathbb{Z}$ , be a sequence of functions, and denote

$$\|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^p(\mathcal{D})} = \sup_{\substack{Q \in \mathcal{D} \\ \mu(Q) \neq 0}} \frac{1}{\mu(Q)^{1/p}} \left\| 1_Q \sum_{k: 2^k \leq \ell(Q)} \varepsilon_k d_k \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; \mathbb{R})}.$$

If  $d_k = E_k d_k$  for all  $k \in \mathbb{Z}$ , then the Carleson norms are equivalent for all  $p \in [1, \infty)$ . For a proof, see Proposition 3.1 in [4]. We recall two Carleson embedding theorems; The following result is Theorem 3.4 in [4].

THEOREM 2.15. *Let  $X$  be a UMD space and  $1 < p < \infty$ . Let  $\{d_k\}_{k \in \mathbb{Z}} \subset L^1(\mathbb{R}^N, \mu; \mathbb{R})$  be a sequence such that  $d_k = E_k d_k$  for every  $k \in \mathbb{Z}$ . Then*

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k d_k E_k f \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; X)} \lesssim \|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

The following embedding result for RMF spaces is essentially Theorem 8.2 in [7], where it is stated for the Lebesgue measure; see also [5] for a general measure  $\mu$  and an interesting converse statement.

**THEOREM 2.16.** *Let  $X$  be an RMF space,  $1 < p < \infty$ , and  $\eta > 0$ . Assuming that  $\{d_k\}_{k \in \mathbb{Z}}$  is a sequence in  $L^1(\mathbb{R}^N, \mu; \mathbb{R})$ , then*

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k d_k E_k f \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; X)} \lesssim \|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^{p+\eta}(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

**3. Adapted martingale decompositions.** Throughout this section we assume that  $X$  is a Banach space and that  $b = b^1$  is an  $L^\infty$ -accretive system. The assumption that  $b = b^1$  is only for notational convenience, and all of the results throughout this section remain valid if we replace  $b^1$  with  $b^2$  and the random dyadic system  $\mathcal{D}$  with  $\mathcal{D}'$ .

**Adapted conditional expectations.** Let  $\mathcal{D} = \bigcup_{k \in \mathbb{Z}} \mathcal{D}_k$  be a generic system of dyadic cubes in  $\mathbb{R}^N$  and  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$ . In what follows we will define various operators acting on this function. First of all, the conditional expectation for  $k \in \mathbb{Z}$  is defined by

$$E_k f := \sum_{Q \in \mathcal{D}_k} 1_Q \langle f \rangle_Q, \quad \langle f \rangle_Q := \frac{1}{\mu(Q)} \int_Q f \, d\mu.$$

If  $\mu(Q) = 0$  for a cube  $Q$ , we agree that  $\langle f \rangle_Q = 0$ . If  $Q \in \mathcal{D}_k$  is a cube, then the local version of this conditional expectation is defined by  $E_Q f := 1_Q E_k f$ . The corresponding martingale difference is defined by  $D_k f := E_{k-1} f - E_k f$  and its local version is  $D_Q f := 1_Q D_k f$ . For the  $L^\infty$ -accretive system  $b = \{b_Q\}$  and  $k \in \mathbb{Z}$ , we define

$$b_k^a := \sum_{Q \in \mathcal{D}_k} 1_Q b_Q^a.$$

The  $b$ -adapted conditional expectation and its local version, for  $k \in \mathbb{Z}$  and  $Q \in \mathcal{D}_k$ , are defined by

$$E_k^a f := b_k^a \frac{E_k f}{E_k b_k^a}, \quad E_Q^a f := 1_Q E_k^a f.$$

The corresponding  $b$ -adapted martingale difference and its local version are

$$D_k^a f := E_{k-1}^a f - E_k^a f, \quad D_Q^a f := 1_Q D_k^a f,$$

where  $k \in \mathbb{Z}$  and  $Q \in \mathcal{D}_k$ .

We agree on the following slightly abusive notation:

$$\begin{aligned} \{b_{k-1}^a = b_k^a\} &:= \bigcup_{\substack{Q \in \mathcal{D}_{k-1} \\ Q^a = (Q^1)^a}} Q, \\ \chi_{k-1} &:= \{b_{k-1}^a \neq b_k^a\} := \mathbb{R}^N \setminus \{b_{k-1}^a = b_k^a\} = \bigcup_{\substack{Q \in \mathcal{D}_{k-1} \\ Q = Q^a \neq Q_0}} Q, \end{aligned} \tag{3.1}$$

where  $k \in \mathbb{Z}$ .

**A representation for  $D_Q^a$ .** Here we compute a useful representation for the adapted martingale differences of  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$ . For this purpose, we let  $Q \in \mathcal{D}_k$  and denote by  $Q_1, \dots, Q_{2^N} \in \mathcal{D}_{k-1}$  the subcubes of  $Q$  (in some order) so that  $\bigcup_{i=1}^{2^N} Q_i = Q$ . Then

$$D_Q^a f = 1_Q(E_{k-1}^a f - E_k^a f) = \sum_{i=1}^{2^N} b_{Q_i^a} \frac{\langle f \rangle_{Q_i}}{\langle b_{Q_i^a} \rangle_{Q_i}} 1_{Q_i} - b_{Q^a} \frac{\langle f \rangle_Q}{\langle b_{Q^a} \rangle_Q} 1_Q.$$

Writing  $\langle f \rangle_Q = \frac{1}{\mu(Q)} \sum_{i=1}^{2^N} \mu(Q_i) \langle f \rangle_{Q_i}$ , we get

$$D_Q^a f = \sum_{i=1}^{2^N} \langle f \rangle_{Q_i} \left( \frac{b_{Q_i^a}}{\langle b_{Q_i^a} \rangle_{Q_i}} 1_{Q_i} - \frac{\mu(Q_i)}{\mu(Q)} \frac{b_{Q^a}}{\langle b_{Q^a} \rangle_Q} 1_Q \right).$$

This computations motivates the following definition: If  $\mu(Q_i) \neq 0$ , we define

$$\varphi_{Q,i}^a := \frac{b_{Q_i^a}}{\langle b_{Q_i^a} \rangle_{Q_i}} 1_{Q_i} - \frac{\mu(Q_i)}{\mu(Q)} \frac{b_{Q^a}}{\langle b_{Q^a} \rangle_Q} 1_Q. \tag{3.2}$$

Otherwise we define  $\varphi_{Q,i} \equiv 0$ .

The following lemma draws conclusions from above, and provides further properties for the resulting decomposition. The proof is straightforward.

**LEMMA 3.3.** *Let  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$  and  $Q \in \mathcal{D}_k$ . Let  $Q_1, \dots, Q_{2^N} \in \mathcal{D}_{k-1}$  denote the subcubes of  $Q$  in some order. Then we can write*

$$D_Q^a f = \sum_{i=1}^{2^N} \langle f \rangle_{Q_i} \varphi_{Q,i}^a.$$

Furthermore, if  $i \in \{1, 2, \dots, 2^N\}$ , then

- a)  $\|\varphi_{Q,i}^a\|_{L^\infty(\mu)} \lesssim 1$ ;
- b)  $\|\varphi_{Q,i}^a\|_{L^1(d\mu)} \lesssim \mu(Q_i)$ ;
- c)  $\text{supp}(\varphi_{Q,i}^a) \subset Q$ ;
- d)  $\int_{\mathbb{R}^N} \varphi_{Q,i}^a d\mu = 0$ .

Within these estimates, the implicit constant depends on  $\delta$  in (1.5).

**A representation for  $(D_k^a)^*$ .** We compute the adjoint of the  $b$ -adapted martingale difference operator  $D_k^a$  for  $k \in \mathbb{Z}$ . For this purpose we fix  $f \in L^p(\mathbb{R}^N; X)$  and  $g \in$

$L^q(\mathbb{R}^N; X^*)$ , where  $1/p + 1/q = 1$ . Recall that

$$D_k^a f = E_{k-1}^a f - E_k^a f = b_{k-1}^a \frac{E_{k-1} f}{E_{k-1} b_{k-1}^a} - b_k^a \frac{E_k f}{E_k b_k^a}. \tag{3.4}$$

The self-adjointness of the expectation operator  $E_k$  yields

$$\langle f, E_k^a g \rangle = \left\langle b_k^a \frac{f}{E_k b_k^a}, E_k g \right\rangle = \left\langle E_k \left( b_k^a \frac{f}{E_k b_k^a} \right), g \right\rangle.$$

As a consequence, we find that

$$A_k f := (E_k^a)^* f = E_k \left( b_k^a \frac{f}{E_k b_k^a} \right) = \frac{E_k(b_k^a f)}{E_k b_k^a}.$$

Substituting this identity to (3.4), we get the representation

$$(D_k^a)^* f = A_{k-1}^a f - A_k f = \frac{E_{k-1}(b_{k-1}^a f)}{E_{k-1} b_{k-1}^a} - \frac{E_k(b_k^a f)}{E_k b_k^a}. \tag{3.5}$$

**A representation for  $(D_k^a)^2$ .** Assuming that  $k, l \in \mathbb{Z}$ ,

$$E_k^a E_l^a f = \frac{b_k^a}{E_k b_k^a} E_k \left( b_l^a \frac{E_l f}{E_l b_l^a} \right) = \begin{cases} \frac{b_k^a}{E_k b_k^a} E_k b_l^a \frac{E_l f}{E_l b_l^a}, & \text{if } l \geq k; \\ E_k^a f, & \text{if } l \leq k. \end{cases}$$

Here we used twice the identity  $E_k = E_k E_l$  if  $l \leq k$ . As a consequence, we have the identity

$$\begin{aligned} (D_k^a)^2 &= (E_{k-1}^a - E_k^a)^2 = (E_{k-1}^a)^2 - E_{k-1}^a E_k^a - E_k^a E_{k-1}^a + (E_k^a)^2 \\ &= (E_{k-1}^a)^2 - E_{k-1}^a E_k^a - \underbrace{E_k^a - E_k^a}_{=0} = E_{k-1}^a D_k^a. \end{aligned}$$

Using the notation (3.1), we write  $D_k^a f$  for a function  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$  as follows:

$$\begin{aligned} D_k^a f &= 1_{\{b_{k-1}^a = b_k^a\}} b_k^a \left( \frac{E_{k-1} f}{E_{k-1} b_{k-1}^a} - \frac{E_k f}{E_k b_k^a} \right) \\ &\quad + 1_{\{b_{k-1}^a \neq b_k^a\}} \left( b_{k-1}^a \frac{E_{k-1} f}{E_{k-1} b_{k-1}^a} - b_k^a \frac{E_k f}{E_k b_k^a} \right). \end{aligned}$$

Using this and the basic properties of the operator  $E_{k-1}$ , we get

$$\begin{aligned}
 E_{k-1}^a D_k^a f &= \frac{b_{k-1}^a}{E_{k-1} b_{k-1}^a} E_{k-1} D_k^a f \\
 &= \frac{b_{k-1}^a}{E_{k-1} b_{k-1}^a} 1_{\{b_{k-1}^a = b_k^a\}} E_{k-1} b_k^a \left( \frac{E_{k-1} f}{E_{k-1} b_k^a} - \frac{E_k f}{E_k b_k^a} \right) \\
 &\quad + \frac{b_{k-1}^a}{E_{k-1} b_{k-1}^a} 1_{\{b_{k-1}^a \neq b_k^a\}} \left( E_{k-1} b_{k-1}^a \frac{E_{k-1} f}{E_{k-1} b_{k-1}^a} - E_{k-1} b_k^a \frac{E_k f}{E_k b_k^a} \right) \\
 &= 1_{\{b_{k-1}^a = b_k^a\}} b_k^a \left( \frac{E_{k-1} f}{E_{k-1} b_k^a} - \frac{E_k f}{E_k b_k^a} \right) \\
 &\quad + 1_{\{b_{k-1}^a \neq b_k^a\}} \left( b_{k-1}^a \frac{E_{k-1} f}{E_{k-1} b_{k-1}^a} - b_k^a \frac{E_k f}{E_k b_k^a} \right) \\
 &\quad + 1_{\{b_{k-1}^a \neq b_k^a\}} E_k f \left( \frac{b_k^a}{E_k b_k^a} - \frac{b_{k-1}^a}{E_{k-1} b_{k-1}^a} \frac{E_{k-1} b_k^a}{E_k b_k^a} \right) \\
 &= D_k^a f + \omega_k^a E_k f,
 \end{aligned}$$

where we have denoted

$$\omega_k^a := 1_{\{b_{k-1}^a \neq b_k^a\}} \left( \frac{b_k^a}{E_k b_k^a} - \frac{b_{k-1}^a}{E_{k-1} b_{k-1}^a} \frac{E_{k-1} b_k^a}{E_k b_k^a} \right).$$

The following lemma draws conclusions from the computations above.

LEMMA 3.6. *Let  $k \in \mathbb{Z}$ . Then*

$$(D_k^a)^2 f = D_k^a f + \omega_k^a E_k f, \quad \text{if } f \in L^1_{\text{loc}}(\mathbb{R}^N; X). \tag{3.7}$$

Furthermore, the functions  $\omega_k^a$  have the following properties a)–c):

- a)  $\omega_k^a(x) = 0$  if  $x \in \mathbb{R}^N \setminus \{b_{k-1}^a \neq b_k^a\}$ ,
- b)  $\|\omega_k^a\|_{L^\infty(\mu)} \lesssim 1$ ,
- c)  $E_{k-1} \omega_k^a = 0$ .

The implicit constant in b) depends on  $\delta$  defined in (1.5).

*Proof.* The identity (3.7) is established above. The property a) is clear; b) follows from (1.5) and (2.10). For c) we notice that

$$E_{k-1} \omega_k^a = 1_{\{b_{k-1}^a \neq b_k^a\}} \left( \frac{E_{k-1} b_k^a}{E_k b_k^a} - \frac{E_{k-1} b_{k-1}^a}{E_{k-1} b_{k-1}^a} \frac{E_{k-1} b_k^a}{E_k b_k^a} \right) = 0.$$

□

We also define the following two local versions of  $\omega_k^a$ . Let  $Q \in \mathcal{D}_k$  and denote by  $Q_1, \dots, Q_{2^N} \in \mathcal{D}_{k-1}$  the subcubes of  $Q$ . Then we define

$$\omega_Q^a := 1_Q \omega_k^a, \quad \omega_{Q,i}^a := 1_{Q_i} \omega_Q^a. \tag{3.8}$$

The following lemma collects the basic properties of these local versions.

LEMMA 3.9. *If  $Q \in \mathcal{D}$  and  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$ , then*

$$(D_Q^a)^2 f = D_Q^a f + \omega_Q^a E_Q f. \tag{3.10}$$

Also,

$$\|\omega_Q^a\|_{L^1(\mu)} \lesssim \mu(Q) \text{ and } \|\omega_{Q_i}^a\|_{L^1(\mu)} \lesssim \mu(Q_i) \text{ for } i = 1, 2, \dots, 2^N. \tag{3.11}$$

*Proof.* Assume that  $Q \in \mathcal{D}_k$ . Then, by using (3.7), we get

$$(D_Q^a)^2 f = 1_Q D_k^a (1_Q D_k^a f) = 1_Q (D_k^a)^2 (1_Q f) = D_k^a f + \omega_Q^a E_Q f.$$

The estimate (3.11) follows from Lemma 3.6. □

**A decomposition of functions.** Recall that  $b$  is an  $L^\infty$ -accretive system. In the sequel we assume that  $b = b^1$  and consider the cube  $Q_0 \in \mathcal{D}_s$  that is defined in (2.8). It is a large cube such that the support of  $\mu$  is contained in it.

We will show that

$$f - b_{Q_0} \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0} \rangle_{\mathbb{R}^N}} = \sum_{j=-\infty}^{\infty} D_j^a f$$

where the convergence takes place both pointwise  $\mu$ -almost everywhere and also in  $L^p(\mathbb{R}^N, \mu; X)$ -norm. We begin with the following lemma.

LEMMA 3.12. *For  $\mu$  almost every point  $x$  in  $Q_0$ , we have*

$$b_{-\infty}^a(x) := \lim_{k \rightarrow -\infty} b_k^a(x) = \lim_{k \rightarrow -\infty} E_k b_k^a(x).$$

Furthermore, the limit satisfies the estimate  $|b_{-\infty}^a(x)| \geq \delta^2$ , where  $\delta > 0$  is defined in connection with (1.5).

*Proof.* By Remark 2.13 we may restrict ourselves to those points  $x \in Q_0$  that belong to at most finitely many cubes in the family  $\cup_{j \in \mathbb{N}_0} \mathcal{L}_j$ . Because the family  $\cup_{j \in \mathbb{N}} \mathcal{L}_j$  is countable, we can also assume that  $\mu(Q_k(x)) \neq 0$  for every  $k \in \mathbb{Z}$  where  $(Q_k(x))_{k \in \mathbb{Z}}$  is the unique sequence of cubes such that  $x \in Q_k(x) \in \mathcal{D}_k$  for every  $k \in \mathbb{Z}$ . Finally, we can also assume that

$$\lim_{k \rightarrow -\infty} \langle b_Q \rangle_{Q_k(x)} = b_Q(x), \quad \text{if } Q \in \bigcup_{j \in \mathbb{N}_0} \mathcal{L}_j. \tag{3.13}$$

Indeed, by martingale convergence, the identity (3.13) holds true for almost every  $x \in Q_0$  if  $Q$  is fixed, and the family  $\cup_{j \in \mathbb{N}} \mathcal{L}_j$  is countable.

Fix a point  $x$  as described above and consider the sequence  $(Q_k(x))_{k \in \mathbb{Z}}$ . Note that  $x \in Q_k(x) \subset (Q_k(x))^a$  for every  $k \leq s$ . In particular, there is an index  $k(x) \leq s$  such that  $(Q_k(x))^a = (Q_{k(x)}(x))^a$  if  $k \leq k(x)$ . As a consequence, for  $k \leq k(x)$ , we can write

$$b_k^a(x) = \sum_{Q \in \mathcal{D}_k} 1_Q(x) b_Q^a(x) = b_{(Q_k(x))^a}(x) = b_{(Q_{k(x)}(x))^a}(x).$$

It follows that  $\lim_{k \rightarrow -\infty} b_k^a(x) = b_{(Q_{k(x)})^a}$ . On the other hand, if  $k \leq k(x)$ , we use the assumption (3.13) for

$$E_k b_k^a(x) = \langle b_k^a \rangle_{Q_k(x)} = \langle b_{(Q_k(x))^a} \rangle_{Q_k(x)} = \langle b_{(Q_{k(x)})^a} \rangle_{Q_k(x)} \xrightarrow{k \rightarrow -\infty} b_{(Q_{k(x)})^a}(x) = \lim_{k \rightarrow -\infty} b_k^a(x). \tag{3.14}$$

This is as required. Finally, since  $\mu(Q_k(x)) \neq 0$  for every  $k$ , we can use (2.10) to conclude that  $|b_{-\infty}^a(x)| = \lim_{k \rightarrow -\infty} |\langle b_{(Q_k(x))^a} \rangle_{Q_k(x)}| \geq \delta^2$ .  $\square$

With the aid of this lemma, we can establish useful convergence results. For this purpose, we fix  $f \in L^1_{\text{loc}}(\mathbb{R}^N; X)$ . By martingale convergence,  $\lim_{k \rightarrow -\infty} E_k f(x) = f(x)$  for  $\mu$ -a.e.  $x \in \mathbb{R}^N$  and, as a consequence of Lemma 3.12, we have

$$E_k^a f(x) = b_k^a(x) \frac{E_k f(x)}{E_k b_k^a(x)} \xrightarrow{k \rightarrow -\infty} f(x) \tag{3.15}$$

for  $\mu$ -a.e.  $x \in \mathbb{R}^N$ . Recall that  $\ell(Q_0) = 2^s$  and  $\text{supp} \mu \subset Q_0$ . Hence, for points  $x$  in  $Q_0 = Q_0^a$ , we have

$$E_s^a f(x) = b_s^a(x) \frac{E_s f(x)}{E_s b_s^a(x)} = b_{Q_0}(x) \frac{\langle f \rangle_{Q_0}}{\langle b_{Q_0} \rangle_{Q_0}} = b_{Q_0}(x) \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0} \rangle_{\mathbb{R}^N}}.$$

Using also (3.15) yields the decomposition

$$\begin{aligned} f - b_{Q_0} \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0} \rangle_{\mathbb{R}^N}} &= \lim_{k \rightarrow -\infty} E_k^a f - E_s^a f \\ &= \lim_{k \rightarrow -\infty} \sum_{j=k+1}^s \underbrace{(E_{j-1}^a f - E_j^a f)}_{=D_j^a f} = \sum_{j=-\infty}^{\infty} D_j^a f \end{aligned} \tag{3.16}$$

that is valid  $\mu$ -almost everywhere in  $\mathbb{R}^N$ . In the last step we used the identity  $E_i^a f = E_s^a f$  if  $i \geq s$  for  $s$  defined in (2.8); hence,  $D_j^a f = 0$   $\mu$ -almost everywhere if  $j > s$ .

Let us then consider the convergence in the  $L^p$ -norm with  $1 < p < \infty$ . In order to do this, we fix  $f \in L^p(\mathbb{R}^N, \mu; X)$ . Let  $j \leq s$ . Using (2.10), we see that  $|E_j b_j^a| \geq \delta^2$  almost everywhere and, by (1.5), we have  $\|b_j^a\|_{L^\infty(\mu)} \leq 1$ . Hence the following norm-estimates are valid pointwise  $\mu$ -almost everywhere

$$\|E_j^a f\|_X = \left\| b_j^a \frac{E_j f}{E_j b_j^a} \right\|_X \leq \delta^{-2} \|E_j f\|_X \leq \delta^{-2} Mf \in L^p(\mathbb{R}^N, \mu; \mathbb{R}),$$

where  $M$  is Doob’s maximal operator. Hence, by the dominated convergence theorem, we see that the decomposition (3.16) converges in  $L^p(\mathbb{R}^N, \mu; X)$ .

**4. Norm estimates for adapted martingales.** We prove ‘square-function’ estimates for the adapted martingale differences and their adjoints, see Theorems 4.1 and 4.16. The first result is true for general UMD spaces, whereas the UMD function lattice property is needed in proof of the second theorem. This dichotomy between the square-function estimates for the  $D_j^a$  and their adjoints,  $(D_j^a)^*$ , seems somewhat unexpected: In the original scalar-valued argument of Nazarov, Treil and Volberg [13, Section 3], only



the estimate for  $D_j^a$  is proven explicitly, while the dual case is just stated, suggesting that it should follow in a similar way. To some extent it does, but this similarity seems to break down in the vector-valued realm, and we will give a careful consideration of both estimates in this section. We begin with:

**THEOREM 4.1.** *Let  $X$  be a UMD space and  $1 < p < \infty$ . Then*

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j D_j^a f \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)} \tag{4.2}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

In what follows we prove Theorem 4.1. For this purpose we first prove various lemmata; the following is a consequence of Theorem 2.15.

**LEMMA 4.3.** *Let  $X$  be a UMD space and  $1 < p < \infty$ . Then*

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a \neq b_j^a\}} E_{j-1} f \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)} \tag{4.4}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

*Proof.* If  $j \in \mathbb{Z}$ , then  $\chi_{j-1} := 1_{\{b_{j-1}^a \neq b_j^a\}} = E_{j-1} 1_{\{b_{j-1}^a \neq b_j^a\}}$  belongs to  $L^1(\mathbb{R}^N, \mu; \mathbb{R})$ . Therefore we can invoke the Carleson embedding theorem 2.15. Hence we can bound the left hand side of (4.4) by a constant multiple of  $\|\{\chi_j\}_{j \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \|f\|_p$ . The first factor is estimated as follows:

$$\|\{\chi_j\}_{j \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \leq \sup_{\substack{Q \in \mathcal{D} \\ \mu(Q) \neq 0}} \frac{1}{\mu(Q)} \sum_{j: 2^j \leq \ell(Q)} \|1_Q \chi_j\|_{L^1}.$$

Fix  $Q \in \mathcal{D}$  with  $\mu(Q) \neq 0$ . Fix  $r \in \mathbb{N}$  such that  $Q^a \in \mathcal{L}_r$ . Using the definition (3.1) and Lemma 2.12, we obtain

$$\begin{aligned} \sum_{j: 2^j \leq \ell(Q)} \|1_Q \chi_j\|_1 &\leq \mu(Q) + \sum_{k=1}^{\infty} \sum_{S \in \mathcal{L}_{r+k}: S \subsetneq Q} \mu(S) \\ &\leq \mu(Q) + \sum_{k=1}^{\infty} (1 - \tau)^{k-1} \mu(Q) \lesssim \mu(Q). \end{aligned}$$

Taking the supremum over  $Q \in \mathcal{D}$  as above, we have  $\|\{\chi_j\}_{j \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \lesssim 1$ . □

Another useful estimate is the following.

**LEMMA 4.5.** *Let  $X$  be a UMD space and  $1 < p < \infty$ . Then*

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a = b_j^a\}} \left( \frac{E_{j-1} f}{E_{j-1} b_{j-1}^a} - \frac{E_j f}{E_j b_j^a} \right) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)} \tag{4.6}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

*Proof.* Using (2.10) and that

$$1_{\{b_{j-1}^a=b_j^a\}}E_{j-1}b_{j-1}^a = 1_{\{b_{j-1}^a=b_j^a\}}E_{j-1}b_j^a, \tag{4.7}$$

we see that the following identities hold pointwise  $\mu$ -almost everywhere in  $\{b_{j-1}^a = b_j^a\}$ :

$$\begin{aligned} \frac{E_{j-1}f}{E_{j-1}b_{j-1}^a} - \frac{E_jf}{E_jb_j^a} &= \frac{E_{j-1}f}{E_{j-1}b_j^a} - \frac{E_jf}{E_jb_j^a} \\ &= E_{j-1}f \left( \frac{1}{E_{j-1}b_j^a} - \frac{1}{E_jb_j^a} \right) + (E_{j-1}f - E_jf) \frac{1}{E_jb_j^a} \\ &= E_{j-1}f \frac{-D_jb_j^a}{E_{j-1}b_j^a E_jb_j^a} + D_jf \frac{1}{E_jb_j^a}. \end{aligned}$$

Hence, the left hand side of (4.6) is dominated by

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a=b_j^a\}} \frac{D_jb_j^a}{E_{j-1}b_j^a E_jb_j^a} E_{j-1}f \right\|_{L^p} + \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a=b_j^a\}} \frac{1}{E_jb_j^a} D_jf \right\|_{L^p}. \tag{4.8}$$

Observe that by (2.10) we have  $|E_jb_j^a| \geq \delta^2$  for  $\mu$ -almost every point. Therefore the contraction principle gives the following estimate for the second term in (4.8):

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a=b_j^a\}} \frac{1}{E_jb_j^a} D_jf \right\|_{L^p} \lesssim \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j D_jf \right\|_{L^p}.$$

The UMD-property of  $X$  allows us to dominate the right hand side by a constant multiple of  $\|f\|_p$ .

For the first term in (4.8) we use the Carleson embedding theorem. Using (4.7) and (2.10) we see that  $1_{\{b_{j-1}^a=b_j^a\}}|E_{j-1}b_j^a|$  and  $|E_jb_j^a|$  are bounded from below by  $\delta^2$  in  $\mu$ -almost every point. Thus, the contraction principle gives the estimate

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a=b_j^a\}} \frac{D_jb_j^a}{E_{j-1}b_j^a E_jb_j^a} E_{j-1}f \right\|_{L^p} \lesssim \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j D_jb_j^a E_{j-1}f \right\|_{L^p}. \tag{4.9}$$

If  $j \in \mathbb{Z}$ , then  $d_{j-1} := D_jb_j^a \in L^1(\mathbb{R}^N, \mu; \mathbb{R})$  and  $D_jb_j^a = E_{j-1}D_jb_j^a$ . Therefore the Carleson embedding Theorem 2.15 applies, and it gives the estimate

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j D_jb_j^a E_{j-1}f \right\|_{L^p} \lesssim \|f\|_{L^p} \sup_{\substack{Q \in \mathcal{D} \\ \mu(Q) \neq 0}} \frac{1}{\mu(Q)} \left\| 1_Q \sum_{k: 2^k \leq \ell(Q)} \varepsilon_k d_k \right\|_{L^1(\mathbb{R}^N \times \Omega, \mu \otimes \mathbb{P}; \mathbb{R})}.$$

In order to estimate the right hand side, we fix a cube  $Q \in \mathcal{D}$  for which  $\mu(Q) \neq 0$ . We have

$$\left\| 1_Q \sum_{k: 2^k \leq \ell(Q)} \varepsilon_k d_k \right\|_{L^1} \leq \left\| 1_Q \sum_{k: 2^k \leq \ell(Q)} \varepsilon_{k-1} D_k b_k^a \right\|_{L^1} + C\mu(Q).$$

The first term on the right hand side is

$$\begin{aligned} & \left\| 1_Q \sum_{k: 2^k \leq \ell(Q)} \varepsilon_{k-1} \sum_{R \in \mathcal{D}_k: R \subset Q} 1_R D_k b_k^a \right\|_{L^1} \\ &= \left\| \sum_{k: 2^k \leq \ell(Q)} \varepsilon_{k-1} \sum_{R \in \mathcal{D}_k: R \subset Q} D_R b_{R^a} \right\|_{L^1} = \left\| \sum_{R \in \mathcal{D}: R \subset Q} \varepsilon_R D_R b_{R^a} \right\|_{L^1} =: \Sigma_Q. \end{aligned}$$

Assume that  $Q^a \in \mathcal{L}_u$ , where  $u \in \mathbb{N}$ . We write  $\Sigma_Q$  in terms of the layers of cubes as follows

$$\Sigma_Q = \left\| \left( \sum_{R \subset Q: R^a = Q^a} + \sum_{j=1}^{\infty} \sum_{S \subsetneq Q: S \in \mathcal{L}_{u+j}} \sum_{R \subset S: R^a = S} \right) \varepsilon_R D_R b_{R^a} \right\|_{L^1}.$$

The triangle inequality gives

$$\begin{aligned} \Sigma_Q &\leq \left\| 1_Q \sum_{R \subset Q: R^a = Q^a} \varepsilon_R D_R b_{Q^a} \right\|_{L^1} \\ &\quad + \sum_{j=1}^{\infty} \sum_{S \subsetneq Q: S \in \mathcal{L}_{u+j}} \left\| 1_S \sum_{R \subset S: R^a = S} \varepsilon_R D_R b_S \right\|_{L^1} =: \Sigma_{Q,1} + \Sigma_{Q,2}. \end{aligned} \tag{4.10}$$

Observe that  $D_R b_{Q^a} = D_R(1_Q b_{Q^a})$  if  $R \subset Q$ . Hence, by applying Hölder’s inequality and (1.5),

$$\Sigma_{Q,1} \leq \|1_Q\|_{L^2} \left\| \sum_{R \subset Q: R^a = Q^a} \varepsilon_R D_R(1_Q b_{Q^a}) \right\|_{L^2} \lesssim \mu(Q)^{1/2} \|1_Q b_{Q^a}\|_{L^2} \lesssim \mu(Q).$$

The second term  $\Sigma_{Q,2}$  is first estimated in a similar manner. Then we use Lemma 2.12 as follows:

$$\Sigma_{Q,2} \lesssim \sum_{j=1}^{\infty} \sum_{S \subsetneq Q: S \in \mathcal{L}_{u+j}} \mu(S) \leq \mu(Q) \sum_{j=1}^{\infty} (1 - \tau)^{j-1} \lesssim \mu(Q). \tag{4.11}$$

We have shown that  $\Sigma_Q \leq \Sigma_{Q,1} + \Sigma_{Q,2} \lesssim \mu(Q)$ . Collecting the estimates above, we see that the left hand side of (4.9) is bounded by a constant multiple of  $\|f\|_{L^p}$ .  $\square$

We are now ready for the proof of Theorem 4.1.

*Proof of Theorem 4.1* We decompose  $D_j^a f$  as follows:

$$\begin{aligned}
 D_j^a f &= E_{j-1}^a f - E_j^a f = b_{j-1}^a \frac{E_{j-1} f}{E_{j-1} b_{j-1}^a} - b_j^a \frac{E_j f}{E_j b_j^a} \\
 &= \mathbf{1}_{\{b_{j-1}^a = b_j^a\}} b_j^a \underbrace{\left( \frac{E_{j-1} f}{E_{j-1} b_{j-1}^a} - \frac{E_j f}{E_j b_j^a} \right)}_{:=I(j)} \\
 &\quad + \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} \underbrace{\left( b_{j-1}^a \frac{E_{j-1} f}{E_{j-1} b_{j-1}^a} - b_j^a \frac{E_j f}{E_j b_j^a} \right)}_{:=II(j)}.
 \end{aligned}
 \tag{4.12}$$

First, using the contraction principle followed by Lemma 4.5 yields

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a = b_j^a\}} b_j^a I(j) \right\|_{L^p} \leq \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a = b_j^a\}} I(j) \right\|_{L^p} \lesssim \|f\|_{L^p}.
 \tag{4.13}$$

In order to estimate the remaining quantity, we fix  $j \in \mathbb{Z}$  and use the identity  $E_j f = E_{j-1} f - D_j f$  for

$$II(j) = \left( \frac{b_{j-1}^a}{E_{j-1} b_{j-1}^a} - \frac{b_j^a}{E_j b_j^a} \right) E_{j-1} f + \frac{b_j^a}{E_j b_j^a} D_j f.$$

This representation leads to the estimate

$$\begin{aligned}
 &\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} II(j) \right\|_{L^p} \\
 &\leq \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} \left( \frac{b_{j-1}^a}{E_{j-1} b_{j-1}^a} - \frac{b_j^a}{E_j b_j^a} \right) E_{j-1} f \right\|_{L^p} \\
 &\quad + \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} \frac{b_j^a}{E_j b_j^a} D_j f \right\|_{L^p}.
 \end{aligned}
 \tag{4.14}$$

The last term above is estimated by using first (1.5) and (2.10) with the contraction principle, and then followed by the UMD-property of  $X$ . This results in the required upper bound  $c\|f\|_{L^p}$  for the term in question.

Applying the contraction principle to the first term in the right hand side of (4.14) yields the estimate

$$\begin{aligned}
 &\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} \left( \frac{b_{j-1}^a}{E_{j-1} b_{j-1}^a} - \frac{b_j^a}{E_j b_j^a} \right) E_{j-1} f \right\|_{L^p} \\
 &\lesssim \left\| \sum_{j=-\infty}^{\infty} \varepsilon_j \mathbf{1}_{\{b_{j-1}^a \neq b_j^a\}} E_{j-1} f \right\|_{L^p}.
 \end{aligned}$$

Using Lemma 4.3, we see that the last term can be dominated by  $c\|f\|_{L^p}$ .

By collecting the estimates beginning from (4.14), we get

$$\left\| \sum_{j=-\infty}^{\infty} \varepsilon_j 1_{\{b_{j-1}^a \neq b_j^a\}} \left( b_{j-1}^a \frac{E_{j-1} f}{E_{j-1} b_{j-1}^a} - b_j^a \frac{E_j f}{E_j b_j^a} \right) \right\|_{L^p} \lesssim \|f\|_{L^p}. \tag{4.15}$$

The required estimate (4.2) follows now by combining the identity (4.12) with the estimates (4.13) and (4.15).  $\square$

**Estimate for the adjoints  $(D_k^a)^*$ .** Here we prove a norm estimate under the UMD function lattice assumption. The need for this assumption was somewhat unexpected to us, but with our present techniques, we were unable to avoid it. Proving (or disproving) the dual square-function estimate in the absence of the lattice assumption would be an interesting question for a deeper understanding of the vector-valued theory.

**THEOREM 4.16.** *Let  $X$  be a UMD function lattice and  $1 < p < \infty$ . Then*

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k (D_k^a)^* f \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)} \tag{4.17}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

In order to prove Theorem 4.16, we first prove the following Carleson embedding for UMD function lattices.

**PROPOSITION 4.18.** *Let  $X$  be a UMD function lattice and  $1 < p < \infty$ . Let*

$$\{d_k \in L^1(\mathbb{R}^N, \mu; \mathbb{R})\}_{k \in \mathbb{Z}}, \quad \{c_k \in L^\infty(\mathbb{R}^N, \mu; \mathbb{R})\}_{k \in \mathbb{Z}},$$

be such that  $d_k = E_k d_k$  and  $\|c_k\|_{L^\infty} \leq 1$  for every  $k \in \mathbb{Z}$ . Then

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k d_k E_k(c_k f) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)} \tag{4.19}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

*Proof.* Since  $X$  is a lattice of functions, for  $\xi \in X$ , we can consider its pointwise absolute value  $|\xi| \in X$ , which satisfies  $\|\xi\|_X = \||\xi|\|_X$ . Moreover, we have the following inequalities, which can be seen by the argument on [14, p. 212]:

$$\int_{\Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \xi_k \right|_X^p d\mathbf{P}(\varepsilon) \approx \left| \left( \sum_{k \in \mathbb{Z}} |\xi_k|^2 \right)^{1/2} \right|_X^p \approx \int_{\Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k |\xi_k| \right|_X^p d\mathbf{P}(\varepsilon).$$

Since  $E_k d_k = d_k$ , we have  $E_k |d_k| = |d_k|$ . Hence, by the Carleson embedding Theorem 2.15, we can estimate the left hand side of (4.19) as follows

$$\begin{aligned} LHS(4.19) &\approx \left\| \left( \sum_{k \in \mathbb{Z}} |d_k E_k(c_k f)|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^N, \mu, X)} \\ &\lesssim \left\| \left( \sum_{k \in \mathbb{Z}} (|d_k| E_k |f|)^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^N, \mu, X)} \\ &\approx \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k |d_k| E_k |f| \right\|_{L^p(\mathbb{R}^N, \mu, X)} \lesssim \|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \| |f| \|_{L^p(\mathbb{R}^N, \mu, X)}. \end{aligned}$$

The required estimate now follows because  $\| |f| \|_p = \|f\|_p$ . □

We also need the following notation and representation formulae.

If  $k \in \mathbb{Z}$ , we write  $Q_k(x)$  for the unique cube in  $\mathcal{D}_k$  containing the point  $x \in \mathbb{R}^N$ .

For  $n \in \mathbb{N}_0$  and  $x \in \mathbb{R}^N$ , we denote by  $Q^n(x) \in \mathcal{L}_n$  the cube in the  $n$ th layer that contains the point  $x$  (if such a cube exists), see Section 2. We also denote  $\sigma_n(x) = \log_2(\ell(Q^n(x)))$  if  $x \in Q^n(x) \in \mathcal{L}_n$  and  $\sigma_n(x) = -\infty$  if there is no cube in  $\mathcal{L}_n$  which contains the point  $x$ .

If  $x \in \cup_{Q \in \mathcal{L}_n} Q$ , we denote  $b^n(x) = b_{Q^n(x)}(x)$ . Note that, for  $k \leq s$  (recall that  $Q_0 \in \mathcal{D}_s$ ) and  $x \in Q_0$ ,

$$\sigma_{n+1}(x) < k \leq \sigma_n(x) \iff (Q_k(x))^a = Q^n(x) \in \mathcal{L}_n. \tag{4.20}$$

In particular, if (4.20) is valid, then  $b_k^a(x) = b_{(Q_k(x))^a}(x) = b^n(x)$ .

We also denote

$$E_{\sigma_n} = \sum_{Q \in \mathcal{L}_n} E_Q, \quad n \in \{0, 1, \dots\}.$$

LEMMA 4.21. *Let  $x \in \mathbb{R}^N$  and  $p \in (1, \infty)$ . Then*

$$\begin{aligned} &\int_{\Omega} \left| \sum_{n=0}^{\infty} \varepsilon_n E_{\sigma_n}(b^n f)(x) \right|^p d\mathbf{P}(\varepsilon) \\ &\lesssim \int_{\Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\{b_{k-1}^a \neq b_k^a\}}(x) E_{k-1}(b_{k-1}^a f)(x) \right|^p d\mathbf{P}(\varepsilon) + |E_{Q_0}(b_{Q_0} f)(x)|^p \end{aligned}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

*Proof.* Denote  $A_k = \mathcal{D}_k \cap (\cup_{n=1}^{\infty} \mathcal{L}_n)$ . Let  $\varepsilon \in \Omega$  and  $x \in \mathbb{R}^N$ . By (3.1),

$$\begin{aligned} \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\{b_{k-1}^a \neq b_k^a\}}(x) E_{k-1}(b_{k-1}^a f)(x) &= \sum_{k \in \mathbb{Z}} \sum_{Q \in A_{k-1}} \varepsilon_k E_Q(b_{Q^a} f)(x) \\ &= \sum_{n=1}^{\infty} \varepsilon_{\sigma_n(x)+1} \sum_{Q \in \mathcal{L}_n} E_Q(b_{Q^a} f)(x). \end{aligned} \tag{4.22}$$

If  $x \in Q \in \mathcal{L}_n$ , then  $Q = Q^a = Q^n(x)$ . Hence,  $1_Q b_{Q^a} f = 1_Q b^n f$ , and applying the last identity to the right hand side of (4.22), we get

$$LHS(4.22) = \sum_{n=1}^{\infty} \varepsilon_{\sigma_n(x)+1} \sum_{Q \in \mathcal{L}_n} E_Q(b^n f)(x) = \sum_{n=1}^{\infty} \varepsilon_{\sigma_n(x)+1} E_{\sigma_n}(b^n f)(x).$$

The required identity follows by taking  $p$ -absolute values, integrating, and relabeling the random variables  $\varepsilon_{\sigma_n(x)+1}$ . □

LEMMA 4.23. *Let  $x \in \mathbb{R}^N$  and  $p \in (1, \infty)$ . Then*

$$\int_{\Omega} \left| \sum_{n=0}^{\infty} \varepsilon_n E_{\sigma_{n+1}}(b^n f)(x) \right|^p d\mathbf{P}(\varepsilon) = \int_{\Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\{b_{k-1} \neq b_k\}}(x) E_{k-1}(b_k^a f)(x) \right|^p d\mathbf{P}(\varepsilon)$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$ .

*Proof.* Let  $A_k$  be as in the previous proof. Let  $x \in \mathbb{R}^N$  and  $\varepsilon \in \Omega$ . Then, if  $x \in Q \in \mathcal{L}_n$  with  $n \geq 1$ , we have  $x \in (Q^{(1)})^a \in \mathcal{L}_{n-1}$ , and therefore

$$b^{n-1}(x) = b_{Q^{n-1}(x)}(x) = b_{(Q^{(1)})^a}(x).$$

Using this identity, we get

$$\begin{aligned} \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\{b_{k-1} \neq b_k\}}(x) E_{k-1}(b_k^a f)(x) &= \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{Q \in A_{k-1}} E_Q(b_{(Q^{(1)})^a} f)(x) \\ &= \sum_{n=1}^{\infty} \varepsilon_{\sigma_n(x)+1} \sum_{Q \in \mathcal{L}_n} E_Q(b^{n-1} f)(x) = \sum_{n=0}^{\infty} \varepsilon_{\sigma_{n+1}(x)+1} E_{\sigma_{n+1}}(b^n f)(x). \end{aligned}$$

The required estimate follows by taking  $p$ -absolute values, integrating, and relabeling the random variables  $\sigma_{n+1}(x) + 1$ . □

We are ready for the proof of Theorem 4.16.

*Proof of Theorem 4.16* By (3.5), we get

$$\begin{aligned} (D_k^a)^* f &= \frac{E_{k-1}(b_{k-1}^a f)}{E_{k-1} b_{k-1}^a} - \frac{E_k(b_k^a f)}{E_k b_k^a} \\ &= \frac{E_k b_k^a - E_{k-1} b_{k-1}^a}{E_k b_k^a E_{k-1} b_{k-1}^a} E_{k-1}(b_{k-1}^a f) + \frac{E_{k-1}(b_{k-1}^a f) - E_k(b_k^a f)}{E_k b_k^a}. \end{aligned} \tag{4.24}$$

Denote  $d_k = E_k b_k^a - E_{k-1} b_{k-1}^a$  for  $k \in \mathbb{Z}$ . Then  $|d_k| \leq 2$  and  $E_{k-1} d_k = d_k$ . By (2.10), we have  $|E_k b_k^a E_{k-1} b_{k-1}^a| \geq \delta^4$  for  $\mu$ -almost every point in  $\mathbb{R}^N$ . Hence, by the contraction principle and Proposition 4.18, we obtain

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \frac{E_k b_k^a - E_{k-1} b_{k-1}^a}{E_k b_k^a E_{k-1} b_{k-1}^a} E_{k-1}(b_{k-1}^a f) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \| \{d_k\}_{k \in \mathbb{Z}} \|_{\text{Car}^1(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)}. \tag{4.25}$$

Let us prove that  $\|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \lesssim 1$ . Let  $Q \in \mathcal{D}$  be such that  $\mu(Q) \neq 0$ . Then  $1_{\{b_{k-1}^a = b_k^a\}} d_k = -1_{\{b_{k-1}^a = b_k^a\}} D_k b_k^a$ , and therefore

$$\left\| 1_Q \sum_{2^k \leq \ell(Q)} \varepsilon_k d_k \right\|_{L^1(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}; \mathbb{R})} \lesssim \left\| 1_Q \sum_{2^k \leq \ell(Q)} \varepsilon_k 1_{\{b_{k-1}^a \neq b_k^a\}} d_k \right\|_1 + \left\| \sum_{P \subset Q} \varepsilon_P D_P b_{P^a} \right\|_1. \tag{4.26}$$

The first term on the right hand side of (4.26) is first estimated by using contraction principle. Then proceeding as in the proof of Lemma 4.3 gives the upper bound  $c\mu(Q)$  for that term.

The second term on the right hand side of (4.26) is estimated as in connection with (4.10), yielding the same upper bound  $c\mu(Q)$ . Combining the estimates above, we find that  $\|\{d_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \lesssim 1$ .

It remains to prove that

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \frac{E_{k-1}(b_{k-1}^a f) - E_k(b_k^a f)}{E_k b_k^a} \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)}. \tag{4.27}$$

By (2.10), we have  $|E_k b_k^a| \geq \delta^2$  for  $\mu$ -almost every point in  $\mathbb{R}^N$ . Using the contraction principle we eliminate the terms  $1/E_k b_k^a$  from the left hand side of (4.27). Then we consider the following decomposition, where  $k \in \mathbb{Z}$ ,

$$\begin{aligned} E_{k-1}(b_{k-1}^a f) - E_k(b_k^a f) &= E_{k-1}(\chi_{k-1}(b_{k-1}^a - b_k^a)f) + (E_{k-1} - E_k)(b_k^a f) \\ &= \chi_{k-1} E_{k-1}((b_{k-1}^a - b_k^a)f) + D_k(b_k^a f), \end{aligned} \tag{4.28}$$

where we have denoted  $\chi_{k-1} = 1_{\{b_{k-1}^a \neq b_k^a\}}$ .

Using Proposition 4.18, we obtain the following norm-estimate involving the first term on the right hand side of (4.28)

$$\begin{aligned} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \chi_{k-1} E_{k-1}((b_{k-1}^a - b_k^a)f) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \\ \lesssim \|\{\chi_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)}. \end{aligned} \tag{4.29}$$

On the other hand, the proof of Lemma 4.3 shows that  $\|\{\chi_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \lesssim 1$ .

In order to complete the proof of (4.27), we still need to prove the following estimate involving the second term on the right hand side of (4.28),

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k D_k(b_k^a f) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)}. \tag{4.30}$$

For this purpose, we introduce independent Rademacher variables  $\tilde{\varepsilon} \in (\tilde{\Omega}, \tilde{\mathbf{P}})$ . For  $x \in Q_0$  and  $k \leq s$ , we denote  $\tilde{\varepsilon}_k^a(x) = \tilde{\varepsilon}_n$  if  $n$  is such that  $\sigma_{n+1}(x) < k \leq \sigma_n(x)$ . By the fact that  $\mu(\mathbb{R}^N \setminus Q_0) = 0$  and (4.20), we find that the functions  $\tilde{\varepsilon}_k^a$ , for  $k \leq s$ , are defined  $\mu$ -almost everywhere and they are  $\mathcal{D}_k$ -measurable.

Then, for every  $x \in Q_0$  and  $\tilde{\varepsilon} \in \tilde{\Omega}$ , we have

$$\int_{\tilde{\Omega}} \left| \sum_{k \leq s} \varepsilon_k D_k(b_k^a f)(x) \right|_X^p d\mathbf{P}(\varepsilon) = \int_{\tilde{\Omega}} \left| \sum_{k \leq s} \varepsilon_k D_k(\tilde{\varepsilon}_k^a b_k^a f)(x) \right|_X^p d\mathbf{P}(\varepsilon).$$



Recall that  $D_k b_k^a = 0$   $\mu$ -almost everywhere if  $k > s$  and  $\mu(\mathbb{R}^N \setminus Q_0) = 0$ . Hence, by integrating, and using the UMD-property of  $X$  and measurability of  $\tilde{\varepsilon}_k^a$ , we obtain

$$\begin{aligned} LHS(4.30) &= \left\| \sum_{k \leq s} \varepsilon_k D_k(\tilde{\varepsilon}_k^a b_k^a f) \right\|_{L^p(\mathbb{R}^N \times \Omega, \mu \otimes \mathbf{P}, X)} \\ &\lesssim \left\| \sum_{k \leq s} D_k(\tilde{\varepsilon}_k^a b_k^a f) \right\|_{L^p(\mathbb{R}^N, \mu; X)} = \left\| \sum_{k \leq s} \tilde{\varepsilon}_k^a D_k(b_k^a f) \right\|_{L^p(\mathbb{R}^N, \mu; X)}. \end{aligned}$$

Reindexing the last sum gives

$$\begin{aligned} LHS(4.30) &\lesssim \left\| \sum_{n=0}^{\infty} \sum_{k \leq s} 1_{\sigma_{n+1} < k \leq \sigma_n} \tilde{\varepsilon}_k^a D_k(b_k^a f) \right\|_{L^p(\mathbb{R}^N, \mu; X)} \\ &\lesssim \left\| \sum_{n=0}^{\infty} \tilde{\varepsilon}_n \sum_{k \leq s} 1_{\sigma_{n+1} < k \leq \sigma_n} D_k(b^n f) \right\|_{L^p(\mathbb{R}^N, \mu; X)}. \end{aligned}$$

In the last inequality we used the fact that the indicators  $x \mapsto 1_{\sigma_{n+1}(x) < k \leq \sigma_n(x)}$  are  $\mathcal{D}_k$ -measurable by (4.20). Taking the expectation over  $\tilde{\varepsilon} \in \tilde{\Omega}$ , we find that

$$LHS(4.30) \lesssim \left\| \sum_{n=0}^{\infty} \tilde{\varepsilon}_n \sum_{k \leq s} 1_{\sigma_{n+1} < k \leq \sigma_n} D_k(b^n f) \right\|_{L^p(\mathbb{R}^N \times \{0,1\}^{\mathbb{N}_0}, \mu \otimes \tilde{\mathbf{P}}, X)}. \tag{4.31}$$

By (4.20) and martingale convergence,

$$\begin{aligned} \sum_{k \leq s} 1_{\sigma_{n+1} < k \leq \sigma_n} D_k &= \sum_{Q: Q^a \in \mathcal{L}_n} D_Q = \sum_{Q: Q^a \in \cup_{m \geq n} \mathcal{L}_m} D_Q - \sum_{Q: Q^a \in \cup_{m > n} \mathcal{L}_m} D_Q \\ &= (1_{\cup \mathcal{L}_n} - E_{\sigma_n}) - (1_{\cup \mathcal{L}_{n+1}} - E_{\sigma_{n+1}}) = E_{\sigma_{n+1}} - E_{\sigma_n} + 1_{\cup \mathcal{L}_n \setminus \cup \mathcal{L}_{n+1}}. \end{aligned}$$

Apply this operator identity to  $b^n f \in L^p(\mathbb{R}^N, \mu; X)$  and substitute the resulting function to the right hand side of (4.31). Using the triangle inequality results in three terms; one of them can be estimated (using that  $\cup \mathcal{L}_n \supset \cup \mathcal{L}_{n+1}$  and  $\|b^n\|_{\infty} \lesssim 1$  if  $n \in \{0, 1, \dots\}$ ) as follows

$$\left\| \underbrace{\left( \sum_{n=0}^{\infty} \tilde{\varepsilon}_n 1_{\cup \mathcal{L}_n \setminus \cup \mathcal{L}_{n+1}} b^n \right) f}_{\lesssim 1} \right\|_{L^p(\mathbb{R}^N \times \{0,1\}^{\mathbb{N}_0}, \mu \otimes \tilde{\mathbf{P}}, X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)}.$$

The two remaining terms can be first estimated by using lemmata 4.21 and 4.23, and then invoking Proposition 4.18. This leads to the upper bound

$$c \|\{1_{\{b_{k-1}^a \neq b_k^a\}}\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \|f\|_{L^p(\mathbb{R}^N, \mu; X)} + c \|E_{Q_0}(b_{Q_0} f)\|_{L^p(\mathbb{R}^N, \mu; X)} \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)}.$$

for these two terms.

Combining the estimates above yields  $LHS(4.30) \lesssim \|f\|_{L^p(\mathbb{R}^N, \mu; X)}$ . □

**5. Decomposition of a Calderón–Zygmund operator.** Let  $T \in \mathcal{L}(L^p(\mathbb{R}^N, \mu; X))$  be a Calderón–Zygmund operator as in Theorem 1.7. We establish the following estimate:

$$|\langle g, Tf \rangle| \lesssim \|g\|_q \|f\|_p + \left| \sum_{Q \in \mathcal{D}, R \in \mathcal{D}'} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right|, \tag{5.1}$$

where  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu, X^*)$ .

Estimate (5.1) is uniform over all dyadic systems  $\mathcal{D} = \mathcal{D}(\beta)$  and  $\mathcal{D}' = \mathcal{D}(\beta')$ , and it is based on decomposition of functions, treated in Section 3. In the subsequent sections various good parts of the series on the right hand side of (5.1) will be estimated. In Section 13, we finish the proof of Theorem 1.7 by collecting estimates of good parts, and also performing an estimate for the remaining bad part.

In order to prove (5.1), we recall the basic cubes  $Q_0 \in \mathcal{D}$  and  $R_0 \in \mathcal{D}'$  satisfying (2.8). The  $L^\infty$ -accretive systems for  $T$  and  $T^*$ , respectively, are denoted by  $\{b_Q^1\}_{Q \in \mathcal{D}}$  and  $\{b_R^2\}_{R \in \mathcal{D}'}$ . Let  $q \in (1, \infty)$  be such that  $p^{-1} + q^{-1} = 1$  and let  $f \in L^p(\mathbb{R}^N, \mu; X)$ ,  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ .

According to (3.16), and the reasoning therein, we have the decompositions

$$\begin{aligned} f - b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}} &= \sum_{j=-\infty}^{\infty} D_j^{a,1} f = \sum_{Q \in \mathcal{D}} D_Q^{a,1} f; \\ g - b_{R_0}^2 \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}} &= \sum_{j=-\infty}^{\infty} D_j^{a,2} g = \sum_{R \in \mathcal{D}'} D_R^{a,2} g, \end{aligned} \tag{5.2}$$

which converge in  $L^p$  and  $L^q$ , respectively. As a consequence, we see that

$$\begin{aligned} \langle g, Tf \rangle &= \left\langle g, T \left( \sum_{Q \in \mathcal{D}} D_Q^{a,1} f \right) \right\rangle + \left\langle g, T \left( b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}} \right) \right\rangle = \left\langle \sum_{R \in \mathcal{D}'} D_R^{a,2} g, T \left( \sum_{Q \in \mathcal{D}} D_Q^{a,1} f \right) \right\rangle \\ &+ \left\langle T^* \left( b_{R_0}^2 \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}} \right), \sum_{Q \in \mathcal{D}} D_Q^{a,1} f \right\rangle + \left\langle g, T \left( b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}} \right) \right\rangle. \end{aligned} \tag{5.3}$$

Recall that  $b_{R_0}^2$  is scalar-valued while  $g$  is  $X^*$ -valued. Moreover, we have

$$T^*(b_{R_0}^2 \otimes \xi^*) = T^*(b_{R_0}^2) \otimes \xi^* =: x \mapsto (T^*(b_{R_0}^2)(x))\xi^*$$

for every  $\xi^* \in X^*$ . Thus, by Hölder’s inequality and inequality (1.6) for  $T^*$ , we have

$$\begin{aligned} \left\| T^* \left( b_{R_0}^2 \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}} \right) \right\|_q &= \left( \int_{\mathbb{R}^N} \|T^* \left( b_{R_0}^2 \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}} \right)\|_{X^*}^q d\mu \right)^{1/q} \\ &= \frac{|\langle g \rangle_{\mathbb{R}^N}|_{X^*}}{|\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}|} \|T^*(b_{R_0}^2)\|_q \\ &\lesssim \frac{1}{\mu(R_0)} \int_{R_0} |g(x)|_{X^*} d\mu(x) \cdot \|T^*(b_{R_0}^2)\|_\infty \cdot \mu(R_0)^{1/q} \\ &\lesssim B\mu(R_0)^{1/q} \left( \frac{1}{\mu(R_0)} \int_{R_0} |g(x)|_{X^*}^q d\mu(x) \right)^{1/q} \lesssim B\|g\|_q. \end{aligned} \tag{5.4}$$

In a similar manner, we have

$$\|b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}}\|_p \lesssim \|f\|_p. \tag{5.5}$$

As a consequence of (5.4), (5.5), and (5.2), we have

$$|\langle T^*(b_{R_0}^2 \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0}^2 \rangle_{\mathbb{R}^N}}), \sum_{Q \in \mathcal{D}} D_Q^{a,1} f \rangle| \lesssim B \|g\|_q \|f\| - b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}}\|_p \lesssim \|g\|_q \|f\|_p.$$

Computing as above, we also find that

$$|\langle g, T(b_{Q_0}^1 \frac{\langle f \rangle_{\mathbb{R}^N}}{\langle b_{Q_0}^1 \rangle_{\mathbb{R}^N}}) \rangle| \lesssim \|g\|_q \|f\|_p.$$

Combining the estimates above gives us (5.1). Within the summation on its right hand side we can tacitly assume that the summation varies over cubes for which  $Q \subset Q_0$  and  $R \subset R_0$ . Indeed, otherwise  $D_R^{a,2} g = 0$  or  $D_Q^{a,1} f = 0$ .

**6. Decoupling estimates.** We begin with the following *tangent martingale trick* originating from [11], and formulated in a way convenient for our purposes in [4]. Let  $(E, \mathcal{M}, \mu)$  be a  $\sigma$ -finite measure space having a refining sequence of partitions as follows: For each  $k \in \mathbb{Z}$ , let  $\mathcal{A}_k$  be a countable partition of  $E$  into sets of finite positive measure so that  $\sigma(\mathcal{A}_k) \subset \sigma(\mathcal{A}_{k-1}) \subset \mathcal{M}$ , and let  $\mathcal{A} = \cup_{k \in \mathbb{Z}} \mathcal{A}_k$ .

For each  $A \in \mathcal{A}$ , let  $\nu_A$  denote the probability measure  $\mu(A)^{-1} \cdot \mu|_A$ . Let  $(F, \mathcal{N}, \nu)$  be the space  $\prod_{A \in \mathcal{A}} A$  with the product  $\sigma$ -algebra and measure. Its points will be denoted by  $y = (y_A)_{A \in \mathcal{A}}$ . By [4, Theorem 6.1], the following norm equivalence holds:

**THEOREM 6.1.** *Suppose that  $X$  is a UMD space and  $p \in (1, \infty)$ . Then*

$$\begin{aligned} & \iint_{E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} f_A(x) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x) \\ & \approx \iiint_{F \times E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} 1_A(x) f_A(y_A) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x) d\nu(y). \end{aligned}$$

As a consequence, we obtain the following extension of [4, Corollary 6.3].

**THEOREM 6.2.** *Let  $X$  be a UMD space and  $p \in (1, \infty)$ . For each  $A \in \mathcal{A}$ , let*

$$k_A : A \times A \rightarrow \mathcal{L}(X)$$

*be a jointly measurable function for which there is a constant  $C > 0$  such that*

$$\mathcal{R}(\{k_A(x, y_A) : x \in A \in \mathcal{A}\}) \leq C < \infty, \quad \text{if } x \in E \text{ and } y \in F. \tag{6.3}$$

Suppose also that, for each  $A \in \mathcal{A}_k$  with  $k \in \mathbb{Z}$  we are given a  $\sigma(\mathcal{A}_{k-1})$ -measurable function  $f_A : E \rightarrow X$ , supported on  $A$ . Then

$$\begin{aligned} & \iint_{E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} \frac{1_A(x)}{\mu(A)} \int_A k_A(x, z) f_A(z) d\mu(z) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x) \\ & \lesssim C \iint_{E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} f_A(x) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x). \end{aligned} \tag{6.4}$$

*Proof.* Observe that the left hand side of (6.4) can be written as

$$\iint_{E \times \Omega} \left| \int_F \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} 1_A(x) k_A(x, y_A) f_A(y_A) dv(y) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x).$$

By Hölder’s inequality and Fubini’s theorem, this quantity is bounded by

$$\iiint_{F \times E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} 1_A(x) k_A(x, y_A) f_A(y_A) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x) dv(y).$$

By the  $\mathcal{R}$ -boundedness assumption (6.3) and the fact that each  $\mathcal{A}_k, k \in \mathbb{Z}$ , is a countable partition of  $E$ , we can further bound (6.4) by

$$C \iiint_{F \times E \times \Omega} \left| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{A \in \mathcal{A}_k} 1_A(x) f_A(y_A) \right|_X^p d\mathbf{P}(\varepsilon) d\mu(x) dv(y).$$

The proof is finished by using Theorem 6.1. □

Let  $Q_v$  and  $R_u, u, v \in \{1, 2, \dots, 2^N\}$ , denote the son cubes of  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$  in a fixed order. Fix  $u$  and  $v$  as above, and assume that the elements of a matrix

$$\{T_{RQ} \in \mathbb{C} : R \in \mathcal{D}', Q \in \mathcal{D}_{R\text{-good}}, \ell(Q) \leq \ell(R)\}$$

satisfy the estimate

$$\frac{|T_{RQ}|}{\mu(R_u)\mu(Q_v)} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{d+\alpha}}. \tag{6.5}$$

Recall that  $D(Q, R) = \ell(Q) + \text{dist}(Q, R) + \ell(R)$ .

Assume that  $\{f_k \in L^1_{\text{loc}}(\mathbb{R}^N, \mu; X)\}_{k \in \mathbb{Z}}$  and  $\{g_k \in L^1_{\text{loc}}(\mathbb{R}^N, \mu; X^*)\}_{k \in \mathbb{Z}}$  are such that  $E_{k-1}f_k = f_k$  and  $E_{k-1}g_k = g_k$  for every  $k \in \mathbb{Z}$ . If  $Q \in \mathcal{D}_k$ , we denote  $f_Q = 1_Q f_k$  and  $g_R = 1_R g_k$  if  $R \in \mathcal{D}'_k$ .

LEMMA 6.6. *Assume that  $E_{k-1}f_k = f_k$  and  $E_{k-1}g_k = g_k$ , where  $f_k$  and  $g_k, k \in \mathbb{Z}$ , are as above. Assume also that the estimate (6.5) holds. Then*

$$\begin{aligned} & \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} \langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v} \right| \\ & \lesssim \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \end{aligned} \tag{6.7}$$

Here and throughout the paper in what follows, we use the following convention in order not to burden the notation too much: The duality pairing between elements  $\varphi \in X^*$  and  $\xi \in X$  is written in the simple product notation as  $\varphi\xi$ . Thus, above, the expression  $\langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v}$  is the duality action of  $\langle g_R \rangle_{R_u} \in X^*$  on  $T_{RQ} \langle f_Q \rangle_{Q_v} \in X$ , where this latter term is in turn the product of  $T_{RQ} \in \mathbb{C}$  and  $\langle f_Q \rangle_{Q_v} \in X$ . (We keep the scalar  $T_{RQ}$  in the middle to anticipate the operator-kernel case in which  $T_{RQ} \in \mathcal{L}(X)$ , in which case  $\langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v}$  is the only logical order of the ‘product’.)

*Proof.* Consider first the part of the series where the ratio  $\ell(R)/\ell(Q)$  is a fixed number  $2^n$ ,  $n \in \mathbb{N}_0$ , and  $2^j < D(Q, R)/\ell(R) \leq 2^{j+1}$  for a momentarily fixed  $j \in \mathbb{N}_0$ . The last double inequality will be abbreviated as  $D(Q, R)/\ell(R) \sim 2^j$ . If moreover  $R \in \mathcal{D}'_k$ , the estimate (6.5) reads as

$$\frac{|T_{RQ}|}{\mu(R_u)\mu(Q_v)} \lesssim \frac{2^{(k-n)\alpha/2} 2^{k\alpha/2}}{2^{(k+j)(d+\alpha)}} = 2^{-n\alpha/2} 2^{-j\alpha} 2^{-(k+j)d}. \tag{6.8}$$

First of all, we have

$$\begin{aligned} & \left| \sum_{k \in \mathbb{Z}} \sum_{R \in \mathcal{D}'_k} \sum_{\substack{Q \in \mathcal{D}'_{k-n} \\ D(Q,R)/\ell(R) \sim 2^j}} \langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v} \right| \\ &= \left| \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathcal{D}_{k-n}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q,R)/\ell(R) \sim 2^j}} \langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v} \right| \\ &= \left| \iint_{\Omega \times \mathbb{R}^N} \sum_{\ell \in \mathbb{Z}} \sum_{S \in \mathcal{D}_{\ell-n}} \varepsilon_S f_S(x) \right. \\ & \quad \cdot \left. \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathcal{D}_{k-n}} \varepsilon_Q \frac{1_{Q_v}(x)}{\mu(Q_v)} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q,R)/\ell(R) \sim 2^j}} T_{RQ} \langle g_R \rangle_{R_u} d\mathbf{P}(\varepsilon) d\mu(x) \right| \\ &\leq \left\| \sum_{S \in \mathcal{D}} \varepsilon_S f_S \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{Q \in \mathcal{D}_{k-n}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q,R)/\ell(R) \sim 2^j}} 1_{Q_v} \frac{T_{RQ}}{\mu(Q_v)} \langle g_R \rangle_{R_u} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \end{aligned} \tag{6.9}$$

Reorganizing the summation, we have

$$\left\| \sum_{S \in \mathcal{D}} \varepsilon_S f_S \right\|_{L^p(\mathbf{P} \otimes \mu; X)} = \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}$$

and we are left with estimating the quantity

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{Q \in \mathcal{D}_{k-n}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q,R)/\ell(R) \sim 2^j}} 1_{Q_v} \frac{T_{RQ}}{\mu(Q_v)} \langle g_R \rangle_{R_u} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \tag{6.10}$$

In order to estimate this quantity, we first prove that if  $Q \in \mathcal{D}_{k-n}$  and  $R \in \mathcal{D}'_k$  are such that  $Q$  is  $R$ -good and  $D(Q, R)/\ell(R) \sim 2^j$ , then

$$Q \subset R^{(j+\theta(j+n))} \in \mathcal{D}'_{k+j+\theta(j+n)}, \quad \text{where } \theta(j) = \left\lceil \frac{\gamma j + r}{1 - \gamma} \right\rceil. \tag{6.11}$$

For this purpose we first show the following: for  $t \in \mathbb{N}_0$ , we have

$$r \leq n + t \Rightarrow \text{either } Q \subset R^{(t)} \text{ or } Q \subset \mathbb{R}^N \setminus R^{(t)}. \tag{6.12}$$

The condition on the left hand side is equivalent to that  $\ell(Q) \leq 2^{-r}\ell(R^{(t)})$ . In order to prove (6.12) we assume the opposite. Then there are points  $x \in Q \cap R^{(t)}$  and  $y \in Q \cap (\mathbb{R}^N \setminus R^{(t)})$ . In particular, we find that  $Q \cap \partial R^{(t)} \neq \emptyset$ , so that  $\text{dist}(Q, \partial R^{(t)}) = 0$ . However, by using Remark 2.6 and that  $Q$  is  $R$ -good, we have

$$\text{dist}(Q, \partial R^{(t)}) > \ell(Q)^\gamma \ell(R^{(t)})^{1-\gamma} > 0$$

because  $\ell(R^{(t)}) \geq \ell(R) > 2^{-1}\ell(R)$  and  $\ell(R^{(t)}) \geq 2^r\ell(Q)$ . This leads to a contradiction, and (6.12) follows.

We are now ready to prove (6.11). Assume the opposite, that is,

$$Q \not\subset R^{(j+\theta(j+n))}. \tag{6.13}$$

Note that  $t := j + \theta(j + n) \geq r$ , so that  $r \leq n + t$ . Hence (6.12) applies, and it implies that  $Q \subset \mathbb{R}^N \setminus R^{(t)}$ . Using this relation and Remark 2.6, we find that

$$\begin{aligned} 2^{\gamma(k-n)}2^{(1-\gamma)(k+t)} &= \ell(Q)^\gamma \ell(R^{(t)})^{1-\gamma} < \text{dist}(Q, \partial R^{(t)}) \\ &= \text{dist}(Q, R^{(t)}) \leq \text{dist}(Q, R) \leq D(Q, R) \leq 2^{j+1}\ell(R) = 2^{j+1+k}. \end{aligned}$$

Simplifying this inequality leads to the estimate

$$t \leq \frac{1}{1 - \gamma}(j + 1 + \gamma n) = j + \frac{\gamma(j + n) + 1}{1 - \gamma} < j + \theta(j + n),$$

which is a contradiction by definition of  $t$ . It follows that (6.13) fails, so that (6.11) holds true as desired.

Using (6.11) we can now reorganise the summation over  $Q$  in (6.10) so that

$$\sum_{Q \in \mathcal{D}_{k-n}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q, R)/\ell(R) \sim 2^j}} = \sum_{S \in \mathcal{D}'_{k+j+\theta(j+n)}} \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \subset S}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q, R)/\ell(R) \sim 2^j}}$$

For  $S, Q, R$  as in the last summation, we denote

$$\frac{T_{RQ}}{\mu(R_u)\mu(Q_v)} =: 2^{-n\alpha/2}2^{-j\alpha} \frac{t_{RQ}}{2^{(k+j)d}} =: 2^{-(n+j)\alpha/4} \frac{\tilde{t}_{RQ}}{\mu(S)},$$

where  $|\tilde{t}_{RQ}| \lesssim |t_{RQ}| \lesssim 1$  by (6.8) and the following estimates:  $2^{-j\alpha/2} \leq 1$  and

$$\mu(S) \leq 2^{d(k+j+\theta(j+n))} \leq 2^{d(1+r/(1-\gamma))}2^{d(k+j)+(j+n)\alpha/4}.$$

In the first inequality above we used (1.1), and in the last inequality we used the assumption  $d\gamma/(1 - \gamma) \leq \alpha/4$ , see (2.1).

For each  $S \in \mathcal{D}'_{k+j+\theta(j+n)}$ , define a kernel

$$K_S(x, y) := \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \subset S}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ D(Q, R)/\ell(R) \sim 2^j}} 1_Q(x) \tilde{t}_{RQ} 1_{R_u}(y).$$

Then  $K_S$  is supported on  $S \times S$  (notice that  $R^{(j+\theta(j+n))} = S$  because both of these cubes from  $\mathcal{D}'_{k+j+\theta(j+n)}$  contain  $Q$ ) and  $|K_S(x, y)| \lesssim 1$  since there is at most one non zero term in the double sum for any given pair of points  $(x, y)$ .

The quantity inside the norm in (6.10) is  $2^{-(n+j)\alpha/4}$  times

$$\sum_{k_0=0}^{j+\theta(j+n)} \sum_{\substack{k \in \mathbb{Z}; k \equiv k_0 \\ \text{mod } j+\theta(j+n)+1}} \varepsilon_k \sum_{S \in \mathcal{D}'_{k+j+\theta(j+n)}} \frac{1_S(x)}{\mu(S)} \int_S K_S(x, y) 1_S(y) g_k(y) d\mu(y), \tag{6.14}$$

where the fact that  $1_{R_u} g_k = 1_{R_u} g_R$  for  $R \in \mathcal{D}'_k$  was also used. For a fixed  $k_0$ , the series over  $k \equiv k_0 \pmod{j + \theta(j + n) + 1}$  is of the form considered in Theorem 6.2. Indeed,  $1_S g_k$  is supported on  $S \in \mathcal{D}'_{k+j+\theta(j+n)}$ , and it is constant on cubes  $Q' \in \mathcal{D}'_{k-1} = \mathcal{D}'_{k'+j+\theta(j+n)}$ , where  $k' = k - (j + \theta(j + n) + 1)$ . By Theorem 6.2 and the contraction principle, the  $L^q(\mathbf{P} \otimes \mu; X^*)$ -norm of the quantity (6.14), for a fixed  $k_0$ , is dominated by a constant multiple of

$$\left\| \sum_{k \equiv k_0} \varepsilon_k \sum_{S \in \mathcal{D}'_{k+j+\theta(j+n)}} 1_S g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \lesssim \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}.$$

The full series over  $k \in \mathbb{Z}$  consists of  $j + \theta(j + n) \lesssim n + j + 1$  subseries like this, which implies that the quantity in (6.10) is dominated by

$$C 2^{-(n+j)\alpha/4} (n + j + 1) \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}$$

Since this is summable over  $n \in \mathbb{N}_0$  and  $j \in \mathbb{N}$ , this proves the goal (6.7). □

**7. Separated cubes.** This section begins the case by case analysis of different subseries of the series (5.1) to be estimated. We start by dealing with cubes well separated from each other, and more precisely we prove the following proposition.

**PROPOSITION 7.1.** *Under the assumptions of Theorem 1.7, we have*

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \lesssim \|g\|_q \|f\|_p \tag{7.2}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ . Here  $1/p + 1/q = 1$ .

For the following lemma, we denote  $\langle g_R \rangle_{R_j} = \langle 1_R g_k \rangle_{R_j} = \langle g_k \rangle_{R_j}$  if  $R \in \mathcal{D}'_k$ . Recall that the auxiliary functions below are defined in (3.2) and (3.8).

LEMMA 7.3. *The left hand side of (7.2) is bounded (up to a constant) by sum of four terms of the following form:*

$$\sum_{i,j=1}^{2^N} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R) \wedge \text{dist}(Q,R)}} \langle g_R \rangle_{R_j} \langle \psi_{R,j} \rangle_{T\varphi_{Q,i}} \langle f_Q \rangle_{Q_i} \right|, \tag{7.4}$$

where, for fixed  $i$  and  $j$ ,

$$(g_k, \psi_{R,j}) = \begin{cases} (E_{k-1} D_k^{a,2} g, \varphi_{R,j}^{a,2}) & \forall k \in \mathbb{Z} \wedge R \in \mathcal{D}'_k \text{ or} \\ (1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} E_k g, \omega_{R,j}^{a,2}) & \forall k \in \mathbb{Z} \wedge R \in \mathcal{D}'_k \end{cases}$$

and

$$(f_k, \varphi_{Q,i}) = \begin{cases} (E_{k-1} D_k^{a,1} f, \varphi_{Q,i}^{a,1}) & \forall k \in \mathbb{Z} \wedge Q \in \mathcal{D}_k \text{ or} \\ (1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_k f, \omega_{Q,i}^{a,1}) & \forall k \in \mathbb{Z} \wedge Q \in \mathcal{D}_k. \end{cases}$$

In every case, these satisfy  $E_{k-1} g_k = g_k$  and  $E_{k-1} f_k = f_k$ .

*Proof.* Using Lemma 3.9 and Lemma 3.3, we get

$$\begin{aligned} D_Q^{a,1} f &= (D_Q^{a,1})^2 f - \omega_Q^{a,1} E_Q f = \sum_{i=1}^{2^N} (\langle D_Q^{a,1} f \rangle_{Q_i} \varphi_{Q,i}^{a,1} - \omega_{Q,i}^{a,1} E_Q f), \\ D_R^{a,2} g &= (D_R^{a,2})^2 g - \omega_R^{a,2} E_R g = \sum_{j=1}^{2^N} (\langle D_R^{a,2} g \rangle_{R_j} \varphi_{R,j}^{a,2} - \omega_{R,j}^{a,2} E_R g). \end{aligned} \tag{7.5}$$

Because  $\text{supp}(\omega_{Q,i}^{a,1}) \subset Q_i \subsetneq Q$ , we have

$$\omega_{Q,i}^{a,1} E_Q f = \omega_{Q,i}^{a,1} 1_Q \langle f \rangle_Q = \omega_{Q,i}^{a,1} \langle f \rangle_Q.$$

Similarly  $\omega_{R,j}^{a,2} E_R g = \omega_{R,j}^{a,2} \langle g \rangle_R$ . As a consequence, for every  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$ , we can write  $\langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle$  as

$$\begin{aligned} \sum_{i,j=1}^{2^N} \left\{ \langle D_R^{a,2} g \rangle_{R_j} \langle \varphi_{R,j}^{a,2}, T\varphi_{Q,i}^{a,1} \rangle \langle D_Q^{a,1} f \rangle_{Q_i} - \langle g \rangle_R \langle \omega_{R,j}^{a,2}, T\varphi_{Q,i}^{a,1} \rangle \langle D_Q^{a,1} f \rangle_{Q_i} \right. \\ \left. - \langle D_R^{a,2} g \rangle_{R_j} \langle \varphi_{R,j}^{a,2}, T\omega_{Q,i}^{a,1} \rangle \langle f \rangle_Q + \langle g \rangle_R \langle \omega_{R,j}^{a,2}, T\omega_{Q,i}^{a,1} \rangle \langle f \rangle_Q \right\}. \end{aligned} \tag{7.6}$$

Thus, to conclude the proof, it suffices to consider the following computations and their symmetric counterparts for the function  $f$ .

First, if  $(g_k, \psi_{R,j}) = (E_{k-1} D_k^{a,2} g, \varphi_{R,j}^{a,2})$ , we have

$$\langle g_R \rangle_{R_j} = \langle 1_R g_k \rangle_{R_j} = \langle E_{k-1} (1_R D_k^{a,2} g) \rangle_{R_j} = \langle D_R^{a,2} g \rangle_{R_j}.$$

Hence,  $\langle g_R \rangle_{R_j} \psi_{R,j} = \langle D_R^{a,2} g \rangle_{R_j} \varphi_{R,j}^{a,2}$ .



Next we assume that  $(g_k, \psi_{R,j}) = (1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} E_k g, \omega_{R,j}^{a,2})$ . Now

$$\langle g_R \rangle_{R_j} = \langle g_k \rangle_{R_j} = \langle 1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} \rangle_{R_j} \langle E_k g \rangle_{R_j} = \langle 1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} \rangle_{R_j} \langle g \rangle_R \tag{7.7}$$

but also

$$\langle 1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} \rangle_{R_j} \omega_{R,j}^{a,2} = 1_{R_j} \langle 1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} \omega_k^{a,2} \rangle = 1_{R_j} \omega_k^{a,2} = \omega_{R,j}^{a,2}. \tag{7.8}$$

Combining the identities (7.7) and (7.8) above, we get  $\langle g_R \rangle_{R_j} \psi_{R,j} = \langle g \rangle_R \omega_{R,j}^{a,2}$ . □

To proceed further we need two lemmata.

LEMMA 7.9. *Let  $Q \in \mathcal{D}$ ,  $R \in \mathcal{D}'$  satisfy  $\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)$ . Assume that  $\varphi_Q, \psi_R \in L^1(\mathbb{R}^N, \mu; \mathbb{C})$  are such that  $\text{supp}(\varphi_Q) \subset Q$ ,  $\text{supp}(\psi_R) \subset R$ , and*

$$\int \varphi_Q d\mu = 0.$$

Then

$$|\langle \psi_R, T\varphi_Q \rangle| \leq \frac{\ell(Q)^\alpha}{\text{dist}(Q, R)^{d+\alpha}} \|\varphi_Q\|_{L^1(\mu)} \|\psi_R\|_{L^1(\mu)}.$$

*Proof.* See [13, Lemma 4.1]. □

LEMMA 7.10. *Suppose that  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}} \cup \mathcal{D}_{R^{(1)}\text{-good}}$  are cubes such that  $\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)$ . Then*

$$\frac{\ell(Q)^\alpha}{\text{dist}(Q, R)^{d+\alpha}} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{d+\alpha}}. \tag{7.11}$$

*Proof.* See [13, Lemma 4.2]. □

We are ready for the proof of Proposition 7.1.

*Proof of Proposition 7.1* By Lemma 7.3, it suffices to estimate (7.4). To this end, we fix  $i, j \in \{1, 2, \dots, 2^N\}$  and denote  $(\psi_R, \varphi_Q) = (\psi_{R,j}, \varphi_{Q,i})$ . Combining lemmata 7.9 and 7.10 and using the properties of functions  $\psi_R$  and  $\varphi_Q$  that are described in lemmata 3.3 and 3.9, we have

$$|\langle \psi_R, T\varphi_Q \rangle| \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{d+\alpha}} \mu(R_j) \mu(Q_i)$$

if  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$  satisfy  $\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)$ . Invoking Lemma 6.6 with a matrix whose elements are defined by

$$T_{RQ} = \langle \psi_R, T\varphi_Q \rangle 1_{\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)} 1_{Q \in \mathcal{D}_{R\text{-good}}}, \tag{7.12}$$

we see that the quantity (7.4) can be dominated by a constant multiple of

$$\left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \tag{7.13}$$

To estimate these quantities, consider first the case  $f_k = 1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_k f$ . In this case, we have

$$f_k = -1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} (E_{k-1} f - E_k f) + 1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_{k-1} f.$$

Using the contraction principle, UMD-property of  $X$ , and Lemma 4.3 we get the estimate

$$\begin{aligned} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)} &\leq \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k D_k f \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \\ &+ \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k 1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_{k-1} f \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \lesssim \|f\|_p. \end{aligned} \tag{7.14}$$

Next consider the case  $f_k = E_{k-1} D_k^{a,1} f$ . Invoking Stein’s inequality and then using Theorem 4.1, we obtain the estimate

$$\left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \lesssim \|f\|_p.$$

in this case. Combining these estimates with analogous estimates for  $g$ , we obtain the upper bound  $C \|g\|_q \|f\|_p$  for (7.13), and therefore also for (7.4). □

**8. Preparations for deeply contained cubes.** In the analysis of (5.1), we move on from the separated cubes to ones contained inside another one. To streamline the actual analysis, we start with some preparations. We will be summing over cubes of the following type:

*LEMMA 8.1.* *Let  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$  be such that  $Q \subset R$  and  $\ell(Q) < 2^{-r} \ell(R)$ . Then  $Q \subset R_1$  for some child, denoted by  $R_1$ , of  $R$ .*

*Proof.* Denote by  $R_1$  any child of  $R$  for which  $R_1 \cap Q \neq \emptyset$ . It suffices to show that  $Q \subset R_1$ . Note that  $\ell(R_1) = 2^{-1} \ell(R)$  and  $\ell(R_1) \geq 2^r \ell(Q)$ . Because  $Q$  is  $R$ -good, we can invoke Remark 2.6 in order to see that

$$\text{dist}(Q, \partial R_1) > \ell(Q)^\gamma \ell(R_1)^{1-\gamma} > 0.$$

Because  $Q \cap R_1 \neq \emptyset$ , it follows that  $Q \subset R_1$ . □

Let  $Q$  and  $R$  be such cubes that are considered in Lemma 8.1. The children of  $R$  are denoted by  $R_1, \dots, R_{2^N}$ . However, we choose the indexing such that  $Q \subset R_1$ , see Lemma 8.1. The indexing of children depends on  $Q$ ; in particular,  $R_1 = R_1(Q)$  depends on  $Q$ . We will not indicate this dependence explicitly.

The children of  $Q$  are denoted by  $Q_1, \dots, Q_{2^N}$  in some order.

Let  $u, v \in \{1, 2, \dots, 2^N\}$  be fixed. Here we consider a matrix  $\{T_{RQ}^u\}$  satisfying the estimate

$$\frac{|T_{RQ}^u|}{\mu(R_u) \mu(Q_v)} \lesssim \left( \frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \cdot \begin{cases} \mu(R)^{-1} & \text{if } u \neq 1, \\ \mu(R_1)^{-1} & \text{if } u = 1, \end{cases} \tag{8.2}$$

LEMMA 8.3. Let  $\{f_k \in L^1_{\text{loc}}(\mathbb{R}^N, \mu; X)\}_{k \in \mathbb{Z}}$  and  $\{g_k \in L^1_{\text{loc}}(\mathbb{R}^N, \mu; X^*)\}_{k \in \mathbb{Z}}$  be such that  $E_{k-1}f_k = f_k$  and  $E_{k-1}g_k = g_k$  for every  $k \in \mathbb{Z}$ . Then, under the assumption (8.2), we have

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R_1 \\ \ell(Q) < 2^{-r}\ell(R)}} \langle g_R \rangle_{R_u} T^u_{RQ} \langle f_Q \rangle_{Q_v} \right| \tag{8.4}$$

$$\lesssim \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \cdot \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}.$$

*Proof.* Consider first part of the series where the ratio  $\ell(Q)/\ell(R)$  is a fixed number  $2^{-n}$  with  $n \in \{r+1, r+2, \dots\}$ . If  $R \in \mathcal{D}'_k$ , the estimate (8.2) reads as

$$\frac{|T^u_{RQ}|}{\mu(R_u)\mu(Q_v)} \lesssim 2^{-n\alpha/2} \cdot \begin{cases} \mu(R)^{-1} & \text{if } u \neq 1, \\ \mu(R_1)^{-1} & \text{if } u = 1. \end{cases} \tag{8.5}$$

Adapting (6.9) to the present situation yields the estimate

$$\left| \sum_{k \in \mathbb{Z}} \sum_{R \in \mathcal{D}'_k} \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \text{ is } R\text{-good} \\ Q \subset R_1}} \langle g_R \rangle_{R_u} T^u_{RQ} \langle f_Q \rangle_{Q_v} \right|$$

$$\leq \left\| \sum_{S \in \mathcal{D}} \varepsilon_S f_S \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{Q \in \mathcal{D}_{k-n}} \sum_{\substack{R \in \mathcal{D}'_k \\ Q \text{ is } R\text{-good} \\ Q \subset R_1}} 1_{Q_v} \frac{T^u_{RQ}}{\mu(Q_v)} \langle g_R \rangle_{R_u} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}.$$

Reorganizing the summation, we have

$$\left\| \sum_{S \in \mathcal{D}} \varepsilon_S f_S \right\|_{L^p(\mathbf{P} \otimes \mu; X)} = \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}$$

so that we are left with estimating the quantity

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{R \in \mathcal{D}'_k} \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \text{ is } R\text{-good} \\ Q \subset R_1}} 1_{Q_v} \frac{T^u_{RQ}}{\mu(Q_v)} \langle g_R \rangle_{R_u} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \tag{8.6}$$

For each  $R \in \mathcal{D}'_k$  and  $m \in \{2, \dots, 2^N\}$ , define kernel

$$K^m_R(x, y) := 2^{n\alpha/2} \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \text{ is } R\text{-good} \\ Q \subset R_1}} \mu(R) \underbrace{1_R(x) 1_{Q_v}(x)}_{=1_{Q_v}(x)} \frac{T^m_{RQ}}{\mu(Q_v)\mu(R_m)} 1_{R_m}(y).$$

For  $S \in \mathcal{D}'_{k-1}$ , we define

$$K_S^1(x, y) := 2^{n\alpha/2} \sum_{\substack{Q \in \mathcal{D}_{k-n} \\ Q \text{ is } S^{(1)\text{-good}} \\ Q \subset S}} \mu(S) \underbrace{1_S(x)1_{Q_v}(x)}_{=1_{Q_v}(x)} \frac{T_{S^{(1)Q}}^1}{\mu(Q_v)\mu(S)} 1_S(y).$$

We have

$$|K_S^1(x, y)| + \sum_{m=2}^{2^N} |K_R^m(x, y)| \lesssim 1$$

by using (8.5) and the fact that there is at most one non zero term in the sums above for any given pair of points  $(x, y)$ . In the sequel we will use one of these kernels, depending on the value of  $u$ . If  $u \neq 1$ , then  $K_R^u$  is supported on  $R \times R$ . If  $u = 1$ , then  $K_S^u = K_S^1$  is supported on  $S \times S$ .

The quantity inside the  $L^p$ -norm in (8.6) is  $2^{-n\alpha/2} \Sigma_u$ , where

$$\Sigma_u := \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{R \in \mathcal{D}'_k} \frac{1_R(x)}{\mu(R)} \int_R K_R^u(x, y) 1_R(y) g_k(y) d\mu(y), \quad \text{if } u \neq 1; \tag{8.7}$$

and

$$\Sigma_u := \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{S \in \mathcal{D}'_{k-1}} \frac{1_S(x)}{\mu(S)} \int_S K_S^1(x, y) 1_S(y) g_k(y) d\mu(y), \quad \text{if } u = 1. \tag{8.8}$$

Here the fact that  $1_{R_u} g_k = 1_{R_u} g_R$  for  $R \in \mathcal{D}'_k$  was also used.

Then we do a case study; assume first that  $u \neq 1$ . Then  $1_{Rg_k}$  is supported on  $R \in \mathcal{D}'_k$ , and it is constant on cubes  $R' \in \mathcal{D}_{k-1}$ . The tangent martingale trick (see Theorem 6.2) implies that the  $L^q(\mathbf{P} \otimes \mu; X^*)$ -norm of the quantity (8.7) is dominated by a constant multiple of

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{R \in \mathcal{D}'_k} g_R \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} = \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \tag{8.9}$$

Then we assume that  $u = 1$ . In this case  $1_S g_k$  is supported on  $S \in \mathcal{D}'_{k-1}$ , and it is constant on cubes  $R' \in \mathcal{D}'_{k-2}$ . The tangent martingale trick (Theorem 6.2) implies that  $L^q(\mathbf{P} \otimes \mu; X^*)$ -norm of the quantity (8.8) is dominated by a constant multiple of

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \sum_{S \in \mathcal{D}'_{k-1}} 1_S g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \lesssim \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \tag{8.10}$$

Combining the estimates (8.9) and (8.10), we find that the quantity in (8.6) is dominated by

$$C 2^{-n\alpha/2} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}$$

This is summable over  $n \in \{r + 1, r + 2, \dots\}$ , and therefore we obtain (8.4). □

**9. Deeply contained cubes.** During the course of Section 9 and Section 10 we establish the following estimate for the part of the summation in (5.1) involving deeply contained cubes.

PROPOSITION 9.1. *Under the assumptions of Theorem 1.7, we have*

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R \\ \ell(Q) < 2^{-r} \ell(R)}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \lesssim \|f\|_p \|g\|_q \tag{9.2}$$

for every  $f \in L^p(X)$  and  $g \in L^q(X^*)$ . Here  $1/p + 1/q = 1$ .

Let  $R$  and  $Q$  be as in (9.2). Recall from beginning of Section 8 that  $R_1, \dots, R_{2^N}$  are children of  $R$  such that  $Q \subset R_1 \subsetneq R$ . By the proof of Lemma 8.1, we get

$$2^{r(1-\gamma)} \ell(Q) \leq \ell(Q)^\gamma \ell(R_m)^{1-\gamma} < \text{dist}(Q, \partial R_m), \quad m \in \{1, 2, \dots, 2^N\}. \tag{9.3}$$

This is a useful inequality later on.

Writing  $1_R = \sum_{m=1}^{2^N} 1_{R_m}$  and using that  $\text{supp}(D_R^{a,2} g) \subset R$  yields

$$\langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle = \langle 1_{R_1} D_R^{a,2} g, T(D_Q^{a,1} f) \rangle + \sum_{m=2}^{2^N} \langle 1_{R_m} D_R^{a,2} g, T(D_Q^{a,1} f) \rangle. \tag{9.4}$$

The point is that  $Q$  is contained in  $R_1$ , so  $Q$  is separated from the children  $R_2, \dots, R_m$ . Hence, arguments developed in Section 7 can be applied to these terms. Treating the main part of the term associated with the child  $R_1$  requires so called paraproducts; these are discussed in the following section.

Let us sketch what are the estimates that are performed in the remaining part of this section. First we will show that

$$\sum_{m=2}^{2^N} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R_1 \\ \ell(Q) < 2^{-r} \ell(R)}} \langle 1_{R_m} D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \lesssim \|f\|_p \|g\|_q. \tag{9.5}$$

Then, in order to treat the remaining (first) term on the right hand side of (9.4), we write  $R_1^c = \mathbb{R}^N \setminus R_1$  and

$$1_{R_1} D_R^{a,2} g = 1_{R_1} \left( b_{R_1^c} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^c} \rangle_{R_1}} - b_{R^c} \frac{\langle g \rangle_R}{\langle b_{R^c} \rangle_R} \right) = (1 - 1_{R_1^c}) \left( b_{R_1^c} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^c} \rangle_{R_1}} - b_{R^c} \frac{\langle g \rangle_R}{\langle b_{R^c} \rangle_R} \right).$$

In this section we establish the estimate

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R_1 \\ \ell(Q) < 2^{-r} \ell(R)}} \left\langle 1_{R_1^c} \left( b_{R_1^c} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^c} \rangle_{R_1}} - b_{R^c} \frac{\langle g \rangle_R}{\langle b_{R^c} \rangle_R} \right), T(D_Q^{a,1} f) \right\rangle \right| \lesssim \|f\|_p \|g\|_q. \tag{9.6}$$

The remaining term is treated in Section 10 by using paraproducts.

**Proving estimate (9.5).** Proceeding as in the proof of Lemma 7.3, we see that the left hand side of (9.5) is dominated by a series of four terms, each of them being of the form

$$\sum_{m=2}^{2^N} \sum_{i,j=1}^{2^N} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R_1 \\ \ell(Q) < 2^{-r} \ell(R)}} \langle g_R \rangle_{R_j} \langle \psi_{R,j,m}, T\varphi_{Q,i} \rangle \langle f_Q \rangle_{Q_i} \right|, \tag{9.7}$$

where we denote  $\langle g_R \rangle_{R_j} = \langle 1_R g_k \rangle_{R_j} = \langle g_k \rangle_{R_j}$  if  $R \in \mathcal{D}'_k$  (similarly for  $f$ ), and the four summands are determined by the following possibilities:

$$(g_k, \psi_{R,j,m}) \in \{(E_{k-1} D_k^{a,2} g, 1_{R_m} \varphi_{R,j}^{a,2}), (1_{\{b_k^{a,2} \neq b_{k-1}^{a,2}\}} E_k g, 1_{R_m} \omega_{R,j}^{a,2})\} \tag{9.8}$$

and

$$(f_k, \varphi_{Q,i}) \in \{(E_{k-1} D_k^{a,1} f, \varphi_{Q,i}^{a,1}), (1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_k f, \omega_{Q,i}^{a,1})\}. \tag{9.9}$$

Note that, in any case,  $E_{k-1} g_k = g_k$  and  $E_{k-1} f_k = f_k$ .

**LEMMA 9.10.** *Assume that  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$ ,  $Q \subset R_1$  and  $\ell(Q) < 2^{-r} \ell(R)$ . Let  $\psi_{R,j,m}$ ,  $m \geq 2$ , and  $\varphi_{Q,i}$  be any of those functions that are quantified in (9.8) and (9.9), respectively. Then  $T_{RQ}^j := \langle \psi_{R,j,m}, T\varphi_{Q,i} \rangle$ ,  $i, j \in \{1, \dots, 2^N\}$ , satisfies*

$$\frac{|T_{RQ}^j|}{\mu(R_j)\mu(Q_i)} \lesssim \left( \frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \mu(R)^{-1}.$$

*Proof.* Because  $m \geq 2$ , we have  $Q \cap R_m \subset R_1 \cap R_m = \emptyset$  so that

$$\ell(Q) < \text{dist}(Q, \partial R_m) = \text{dist}(Q, R_m)$$

by (9.3) and the assumption that  $1 \leq 2^{r(1-\gamma)}$ , see (2.1). We also have  $\ell(Q) \leq \ell(R_m)$ . Hence, by using lemmata 7.9 and 7.10, the properties of functions  $\psi_{R,j,m}$  and  $\varphi_{Q,i}$  that follow from lemmata 3.3 and 3.6, and (1.1), we obtain

$$\begin{aligned} |\langle \psi_{R,j,m}, T\varphi_{Q,i} \rangle| &\lesssim \frac{\ell(Q)^{\alpha/2} \ell(R_m)^{\alpha/2}}{D(Q, R_m)^{d+\alpha}} \|\psi_{R,j,m}\|_{L^1(\mu)} \|\varphi_{Q,i}\|_{L^1(\mu)} \\ &\lesssim \left( \frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \frac{\|\psi_{R,j,m}\|_{L^1(\mu)} \|\varphi_{Q,i}\|_{L^1(\mu)}}{\ell(R)^d} \lesssim \left( \frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \frac{\mu(R_j)\mu(Q_i)}{\mu(R)}. \end{aligned} \tag{9.11}$$

This is the desired estimate. □

Combining lemmata 8.3 and 9.10 and then estimating as in the end of Section 7, we see that the quantity (9.7) can be dominated by a constant multiple of  $\|f\|_p \|g\|_q$ . As a consequence, we see that the left hand side of (9.5) is dominated by a constant multiple of  $\|f\|_p \|g\|_q$ .

**Proving estimate (9.6).** Let  $R \in \mathcal{D}'_k$ . We write

$$1_{R_1^c} \left( b_{R_1^a} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^a} \rangle_{R_1}} - b_{R^a} \frac{\langle g \rangle_R}{\langle b_{R^a} \rangle_R} \right) = 1_{R_1^c} b_{R^a} \langle s_k \rangle_{R_1} + 1_{R_1^c} b_{R_1^a} \langle h_k \rangle_{R_1} + 1_{R_1^c} b_{R^a} \langle u_k \rangle_{R_1},$$

where

$$s_k = 1_{\{b_{k-1}^{a,2} = b_k^{a,2}\}} \left( \frac{E_{k-1}g}{E_{k-1}b_{k-1}^{a,2}} - \frac{E_k g}{E_k b_k^{a,2}} \right),$$

and

$$h_k = 1_{\{b_{k-1}^{a,2} \neq b_k^{a,2}\}} \frac{E_{k-1}g}{E_{k-1}b_{k-1}^{a,2}}, \quad u_k = -1_{\{b_{k-1}^{a,2} \neq b_k^{a,2}\}} \frac{E_k g}{E_k b_k^{a,2}}.$$

By (7.5), we see that the left hand side of (9.6) can be dominated from above by a sum of six terms, each of them being of the form

$$\sum_{i=1}^{2^N} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{\text{good}} \\ Q \subset R_1 \\ \ell(Q) < 2^{-r} \ell(R)}} \langle g_R \rangle_{R_1} \langle \psi_R, T\varphi_{Q,i} \rangle \langle f_Q \rangle_{Q_i} \right|, \tag{9.12}$$

where  $\langle g_R \rangle_{R_i} = \langle 1_R g \rangle_{R_i} = \langle g \rangle_{R_i}$  if  $R \in \mathcal{D}'_k$  (similarly for  $f$ ), and the six terms are determined by the following choices:

$$(g_k, \psi_R) \in \{(s_k, 1_{R_1^c} b_{R^a}), (h_k, 1_{R_1^c} b_{R_1^a}), (u_k, 1_{R_1^c} b_{R^a})\} \tag{9.13}$$

and

$$(f_k, \varphi_{Q,i}) \in \{(E_{k-1} D_k^{a,1} f, \varphi_{Q,i}^{a,1}), (1_{\{b_k^{a,1} \neq b_{k-1}^{a,1}\}} E_k f, \omega_{Q,i}^{a,1})\}. \tag{9.14}$$

Note that, in any case,  $E_{k-1}g_k = g_k$  and  $E_{k-1}f_k = f_k$ .

**LEMMA 9.15.** *Let  $\psi_R$  and  $\varphi_{Q,i}$  be any of those functions that are quantified in (9.13) and (9.14) for  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$  satisfying  $Q \subset R_1$  and  $\ell(Q) < 2^{-r} \ell(R)$ . Then  $T_{RQ}^1 := \langle \psi_R, T\varphi_{Q,i} \rangle$  satisfies the estimate*

$$\frac{|T_{RQ}^1|}{\mu(Q_i)} \lesssim \left( \frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2}.$$

*Proof.* Denote by  $y_Q$  the midpoint of  $Q$ . Let  $x \in R_1^c$  and  $y \in Q$ . By (2.1), (9.3) and the fact that  $Q \subset R_1$ , we have

$$2|y - y_Q| \leq 2^{r(1-\gamma)} \ell(Q) < \text{dist}(Q, \partial R_1) = \text{dist}(Q, R_1^c) \leq |x - y_Q|.$$

Using the kernel estimate (1.3) and the facts  $\int \varphi_{Q,i} = 0$  and  $\text{supp}(\varphi_{Q,i}) \subset Q$ , we get

$$\begin{aligned} & |\langle \psi_R, T\varphi_{Q,i} \rangle| \\ &= \left| \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \psi_R(x) 1_{R_1^c}(x) (K(x, y) - K(x, y_Q)) \varphi_{Q,i}(y) d\mu(y) d\mu(x) \right| \\ &\lesssim \int_{R_1^c} \int_{\mathbb{R}^N} \frac{|y - y_Q|^\alpha}{|x - y_Q|^{d+\alpha}} |\varphi_{Q,i}(y)| d\mu(y) d\mu(x) \lesssim \|\varphi_{Q,i}\|_1 \int_{R_1^c} \frac{\ell(Q)^\alpha}{|x - y_Q|^{d+\alpha}} d\mu(x). \end{aligned} \tag{9.16}$$

Denoting  $A_k = \{x : 2^k \text{dist}(R_1^c, Q) \leq |x - y_Q| < 2^{k+1} \text{dist}(R_1^c, Q)\}$ , we can estimate the last integral as follows

$$\begin{aligned} \int_{R_1^c} \frac{\ell(Q)^\alpha}{|x - y_Q|^{d+\alpha}} d\mu(x) &\leq \sum_{k=0}^\infty \int_{A_k} \frac{\ell(Q)^\alpha}{(2^k \text{dist}(R_1^c, Q))^{d+\alpha}} d\mu(x) \\ &\leq \sum_{k=0}^\infty \frac{\ell(Q)^\alpha \mu(B(y_Q, 2^{k+1} \text{dist}(R_1^c, Q)))}{(2^k \text{dist}(R_1^c, Q))^{d+\alpha}} \\ &\lesssim \frac{\ell(Q)^\alpha}{\text{dist}(R_1^c, Q)^\alpha} \sum_{k=0}^\infty \frac{1}{2^{\alpha k}} \lesssim \left( \frac{\ell(Q)}{\text{dist}(R_1^c, Q)} \right)^\alpha. \end{aligned} \tag{9.17}$$

This can be further estimated by using that  $\gamma \leq \alpha(2(d + \alpha))^{-1} < 2^{-1}$ , see (2.1). Combining this with (9.3) yields the estimate

$$\ell(Q)^{1/2} \ell(R_1)^{1/2} \leq \ell(Q)^\gamma \ell(R_1)^{(1-\gamma)} \leq \text{dist}(Q, \partial R_1) = \text{dist}(Q, R_1^c).$$

Substituting this into (9.16), we find that

$$|\langle \psi_R, T\varphi_{Q,i} \rangle| \lesssim \left( \frac{\ell(Q)}{\text{dist}(R_1^c, Q)} \right)^\alpha \|\varphi_{Q,i}\|_1 \lesssim \left( \frac{\ell(Q)}{\ell(R_1)} \right)^{\alpha/2} \|\varphi_{Q,i}\|_1.$$

This is as required because  $\|\varphi_{Q,i}\|_1 \lesssim \mu(Q_i)$  and  $\ell(R_1) = 2^{-1} \ell(R)$ . □

Combining lemmata 9.15 and 8.3 we find that each of the six terms of the form (9.12) are bounded (up to a constant) by

$$\left\| \sum_{k=-\infty}^\infty \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \cdot \left\| \sum_{k=-\infty}^\infty \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}.$$

At the end of Section 7 we verified that the second factor above can be dominated by  $\|f\|_p$ . Hence, it remains to verify the following estimate,

$$\left\| \sum_{k=-\infty}^\infty \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \lesssim \|g\|_q. \tag{9.18}$$

The cases  $g_k \in \{h_k, u_k\}$  have been cleared in connection with the separated cubes: (9.18) follows from the contraction principle and (7.14) if we recall that  $|E_k b_k^{a,2}| \geq \delta^2 \mu$ -almost



everywhere. The remaining case

$$g_k = s_k = 1_{\{b_{k-1}^{a,2} = b_k^{a,2}\}} \left( \frac{E_{k-1}g}{E_{k-1}b_{k-1}^{a,2}} - \frac{E_k g}{E_k b_k^{a,2}} \right).$$

is treated by Lemma 4.5.

This concludes the proof of estimate (9.6).

**10. Paraproducts.** In order to finish the proof of Proposition 9.1, we still need to establish the following estimate

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R \\ \ell(Q) < 2^{-r}\ell(R)}} \left\langle b_{R_1^q} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^q} \rangle_{R_1}} - b_{R^a} \frac{\langle g \rangle_R}{\langle b_{R^a} \rangle_R}, T(D_Q^{a,1}f) \right\rangle \right| \lesssim \|f\|_p \|g\|_q. \tag{10.1}$$

We will draw inspiration from the work of Hytönen and Martikainen [6], and the following standing assumptions in Theorem 1.7 are crucial while proving (10.1):

- $X^*$  is an RMF-space;
- $\|T^*b_R^2\|_{L^\infty(\mathbb{R}^N, \mu; \mathbb{C})} \leq 1$  if  $R$  is a cube in  $\mathbb{R}^N$ .

For  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$ , we denote

$$\chi_{Q,R} = \begin{cases} 1, & \text{if } Q \text{ is } R\text{-good, } Q \subset R, \text{ and } \ell(Q) < 2^{-r}\ell(R); \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that  $\chi_{Q,R} = 1$ . Then we write

$$G_{Q,R} := b_{R_1^q} \frac{\langle g \rangle_{R_1}}{\langle b_{R_1^q} \rangle_{R_1}} - b_{R^a} \frac{\langle g \rangle_R}{\langle b_{R^a} \rangle_R}$$

for a quantity that depends on  $Q$  and  $R$ , as  $R_1$  stands for the child of  $R$  for which  $Q \subset R$ . Using the notation above, we can rewrite the left hand side of (10.1) as follows

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ \chi_{Q,R} = 1}} \langle G_{Q,R}, T(D_Q^{a,1}f) \rangle \right| = \left| \sum_{Q \in \mathcal{D}} \left\langle \sum_{\substack{R \in \mathcal{D}' \\ \chi_{Q,R} = 1}} G_{Q,R}, T(D_Q^{a,1}f) \right\rangle \right|. \tag{10.2}$$

It is straightforward to verify that, if  $\chi_{Q,R} = 1$ , then  $\chi_{Q,R^{(m)}} = 1$  for every  $m \in \mathbb{N}_0$ . It follows that, if  $Q \in \mathcal{D}$  and the inner sum on the right hand side is nonempty, there exists a unique cube  $S = S(Q) \in \mathcal{D}'$  containing  $Q$  such that  $\chi_{Q,R} = 1$  if, and only if,  $S \subsetneq R \in \mathcal{D}'$ . If the inner sum in question is empty, we let  $S = S(Q) = \emptyset$ . As a consequence, if  $S(Q) \neq \emptyset$ ,

$$\sum_{\substack{R \in \mathcal{D}' \\ \chi_{Q,R} = 1}} G_{Q,R} = \sum_{\substack{R \in \mathcal{D}' \\ S \subsetneq R}} G_{Q,R} = b_{S^a} \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} - b_{R_0} \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0} \rangle_{\mathbb{R}^N}}.$$

Substituting this identity to the right hand side of (10.2), we get

$$\left| \sum_{Q \in \mathcal{D}} \left\langle \sum_{\substack{R \in \mathcal{D}' \\ \chi_{Q,R}=1}} G_{Q,R}, T(D_Q^{a,1}f) \right\rangle \right| \leq |\langle \Pi g, f \rangle| + \left| \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \left\langle T^* b_{R_0} \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0} \rangle_{\mathbb{R}^N}}, D_Q^{a,1}f \right\rangle \right|, \tag{10.3}$$

where the *paraproduct operator*  $g \mapsto \Pi g$  is defined by

$$\Pi g := \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} (D_Q^{a,1})^* (T^* b_{S^a}) = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \frac{\langle g \rangle_R}{\langle b_{R^a} \rangle_R} (D_Q^{a,1})^* (T^* b_{R^a}). \tag{10.4}$$

Throughout the rest of this section, we will prove the following estimates:

**PROPOSITION 10.5.** *Under the standing assumptions, the paraproduct just defined satisfies*

$$|\langle \Pi g, f \rangle| \lesssim \|f\|_p \|g\|_q, \tag{10.6}$$

and we also have the estimate

$$\left| \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \left\langle T^* b_{R_0} \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0} \rangle_{\mathbb{R}^N}}, D_Q^{a,1}f \right\rangle \right| \lesssim \|f\|_p \|g\|_q. \tag{10.7}$$

Observe that these estimates imply (10.1) which in turn, combined with estimates in Section 9, implies Proposition 9.1.

**Proving estimate (10.6).** Here we will concentrate on paraproducts, and begin with the following lemma.

**LEMMA 10.8.** *Suppose that  $t > q \vee s$ , where  $X^*$  has cotype  $s$ . Assume that a sequence  $\{d_j\}_{j \in \mathbb{Z}}$  of functions  $\mathbb{R}^N \rightarrow L^t(\Omega; \mathbb{C})$  satisfies  $d_j \in L^1(\mathbb{R}^N; L^t(\Omega; \mathbb{C}))$ , then*

$$\left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* d_j E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N; L^t(\Omega; X^*))} \lesssim \|\{d_j(\cdot)\}_{j \in \mathbb{Z}}\|_{\text{Car}^t(\mathcal{D}')} \cdot \|g\|_{L^q(\mathbb{R}^N; X^*)}.$$

*Proof.* This will be a special case of Theorem 3.5 in [4], which says that

$$\left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* d_j E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N; X_3)} \lesssim \|\{d_j(\cdot)\}_{j \in \mathbb{Z}}\|_{X_2} \|\text{Car}^t(\mathcal{D}')\| \cdot \|g\|_{L^q(\mathbb{R}^N; X_1)},$$

whenever  $X_1, X_2, X_3$  are three Banach spaces with  $X_1$  having the RMF property, and  $X_2 \subseteq \mathcal{L}(X_1, X_3)$  embedded in such a way that the unit-ball  $\bar{B}_{X_2}$  is  $R$ -bounded.

Denote  $X_1 = X^*$ ,  $X_2 = L^t(\Omega; \mathbb{C})$ , and  $X_3 = L^t(\Omega; X^*)$ . Then  $X_1$  is an RMF space by assumption. By the result of [4] just stated, it suffices to verify that the closed unit ball of  $X_2$  is a Rademacher-bounded subspace of  $\mathcal{L}(X_1, X_3)$  when the action of  $\rho \in X_2$  is defined by

$$X_1 \ni x \mapsto \rho(x) := \rho \otimes x : \rho \otimes x(\varepsilon) = \rho(\varepsilon)x.$$

To this end, let  $\{\rho_j : j \in \mathbb{N}\}$  be a sequence in  $\bar{B}_{X_2}$  and  $\{x_j : j \in \mathbb{N}\}$  be a sequence in  $X^*$ . By Fubini's theorem and Proposition 2.4,

$$\begin{aligned} \left( \mathbf{E}_{\varepsilon^*} \left\| \sum_{j=1}^{\infty} \varepsilon_j^* \rho_j \otimes x_j \right\|_{L'(\Omega; X^*)}^t \right)^{1/t} &= \left\| \sum_{j=1}^{\infty} \varepsilon_j^* \rho_j \otimes x_j \right\|_{L'(\Omega; L'(\Omega^*; X^*))} \\ &\lesssim \underbrace{\sup_{j \in \mathbb{N}} \|\rho_j\|_{L'(\Omega)}}_{\leq 1} \cdot \left\| \sum_{j=1}^{\infty} \varepsilon_j^* x_j \right\|_{L'(\Omega^*; X^*)} \leq \left( \mathbf{E}_{\varepsilon^*} \left\| \sum_{j=1}^{\infty} \varepsilon_j^* x_j \right\|_{X^*}^t \right)^{1/t}. \end{aligned}$$

By Kahane–Khinchine inequality, this is as required. □

We need further preparations for establishing (10.6).

Recall that  $D_Q^{a,1} = (D_Q^{a,1})^2 - \omega_Q^{a,1} E_Q$  by (3.10). Denote

$$\chi_Q := 1_Q \chi_{k-1} := 1_Q 1_{\{b_{k-1}^{a,1} \neq b_k^{a,1}\}}, \quad \text{if } Q \in \mathcal{D}_k. \tag{10.9}$$

By Lemma 3.6, we have  $\omega_Q^{a,1} = \chi_Q \omega_Q^{a,1}$ . Furthermore,  $\chi_Q E_Q f = E_{k-1}(\chi_Q E_Q f)$  if  $Q \in \mathcal{D}_k$ . Hence, we can write

$$\begin{aligned} \langle \Pi g, f \rangle &= \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} \langle (D_Q^{a,1})^*(T^* b_{S^a}), D_Q^{a,1} f \rangle \\ &\quad - \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} \langle \omega_Q^{a,1} T^* b_{S^a}, E_{\log_2(\ell(Q))-1}(\chi_Q E_Q f) \rangle \\ &= \int_{\Omega} \left\langle \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \varepsilon_Q \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} (D_Q^{a,1})^*(T^* b_{S^a}), \sum_{Q' \in \mathcal{D}} \varepsilon_{Q'} D_{Q'}^{a,1} f \right\rangle d\mathbf{P}(\varepsilon) \\ &\quad - \int_{\Omega} \left\langle \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \varepsilon_Q \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} E_{\log_2(\ell(Q))-1}(\omega_Q^{a,1} T^* b_{S^a}), \sum_{Q' \in \mathcal{D}} \varepsilon_{Q'} \chi_{Q'} E_{Q'} f \right\rangle d\mathbf{P}(\varepsilon). \end{aligned}$$

Taking the absolute values, and using Hölder's inequality, we get

$$\begin{aligned} |\langle \Pi g, f \rangle| &\leq \left\| \sum_{\substack{Q \in \mathcal{D} \\ S= S(Q) \neq \emptyset}} \varepsilon_Q \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} (D_Q^{a,1})^*(T^* b_{S^a}) \right\|_q \left\| \sum_{Q' \in \mathcal{D}} \varepsilon_{Q'} D_{Q'}^{a,1} f \right\|_p \\ &\quad + \left\| \sum_{\substack{Q \in \mathcal{D} \\ S= S(Q) \neq \emptyset}} \varepsilon_Q \frac{\langle g \rangle_S}{\langle b_{S^a} \rangle_S} E_{\log_2(\ell(Q))-1}(\omega_Q^{a,1} T^* b_{S^a}) \right\|_q \left\| \sum_{Q' \in \mathcal{D}} \varepsilon_{Q'} \chi_{Q'} E_{Q'} f \right\|_p. \end{aligned}$$

Using (2.10) and contraction principle, followed by Theorem 4.1 and Lemma 4.3, we see that  $|\langle \Pi g, f \rangle|$  is bounded by a sum of two terms, both of them being (a constant multiple) of the general form

$$\left\| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \langle g \rangle_R \right\|_q \cdot \|f\|_p. \tag{10.10}$$

Here the two terms are determined by the following choices:

$$\pi_{Q,R^a} \in \{(D_Q^{a,1})^*(T^*b_{R^a}), E_{\log_2(\ell(Q))-1}(\omega_Q^{a,1}T^*b_{R^a})\}. \tag{10.11}$$

Observe that, if  $R$  and  $Q$  are as in (10.10), then

$$\pi_{Q,R^a} = 1_R \pi_{Q,R^a}. \tag{10.12}$$

In order to estimate quantities of the form (10.10), we will use the following lemma.

LEMMA 10.13. *Assume that  $U \in \mathcal{D}'$  and  $t \in (1, \infty)$ . Then*

$$\left\| \sum_{\substack{R \in \mathcal{D}': R \subset U \\ R^a = U^a}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q) = R}} \varepsilon_Q \pi_{Q,U^a} \right\|_{L^t(\mathbb{R}^N \times \Omega; \mathbb{C})} \lesssim \mu(U)^{1/t}. \tag{10.14}$$

*Proof.* Denote  $h = T^*b_{U^a}$  and first consider the case  $\pi_{Q,U^a} = (D_Q^{a,1})^*(h)$ . Because  $Q \subset S(Q)$  if  $S(Q) \neq \emptyset$ , we see that the left hand side of (10.14) is

$$\left\| \sum_{\substack{R \in \mathcal{D}': R \subset U \\ R^a = U^a}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q) = R}} \varepsilon_Q (D_Q^{a,1})^*(1_U h) \right\|_t \leq \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q (D_Q^{a,1})^*(1_U h) \right\|_t.$$

Using Theorem 4.16 with  $X = \mathbb{C}$ , followed by (1.6), we find that the last quantity is bounded by a constant multiple of

$$\|1_U h\|_{L^t(\mathbb{R}^N, \mu; \mathbb{C})} \leq \|h\|_{L^\infty(\mathbb{R}^N, \mu; \mathbb{C})} \|1_U\|_{L^t(\mathbb{R}^N, \mu; \mathbb{C})} = \mu(U)^{1/t} \|h\|_{L^\infty(\mathbb{R}^N, \mu; \mathbb{C})} \leq B\mu(U)^{1/t}.$$

This is the required estimate in the present case.

Then consider the case

$$\pi_{Q,U^a} = E_{\log_2(\ell(Q))-1}(\omega_Q^{a,1}h).$$

Recall that the expectation is taken with respect to  $\mathcal{D}_{\log_2(\ell(Q))-1}$ . By the contraction principle and the facts that  $Q \subset S(Q)$  if  $S(Q) \neq \emptyset$  and  $\omega_Q^{a,1} = \chi_Q \omega_Q^{a,1}$ , see (10.9), we get

$$\begin{aligned} \left\| \sum_{\substack{R \in \mathcal{D}': R \subset U \\ R^a = U^a}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q) = R}} \varepsilon_Q \pi_{Q,U^a} \right\|_t &\lesssim \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q \chi_Q E_{\log_2(\ell(Q))-1}(\omega_Q^{a,1}1_U h) \right\|_t \\ &\lesssim \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \chi_{k-1} \sum_{Q \in \mathcal{D}_k} 1_Q E_{k-1}(\omega_Q^{a,1}1_U h) \right\|_t \\ &\lesssim \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \chi_{k-1} E_{k-1}(\omega_k^{a,1}1_U h) \right\|_t. \end{aligned}$$

Here  $\chi_{k-1} = 1_{\{b_{k-1}^{a,1} \neq b_k^{a,1}\}}$  satisfies  $\chi_{k-1} = E_{k-1}\chi_{k-1}$ . Also,  $\sup_{k \in \mathbb{Z}} \|\omega_k^{a,1}\|_{L^\infty(\mu)} \lesssim 1$  by Lemma 3.6. Hence, by Proposition 4.18 with  $X = \mathbb{C}$ ,

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \chi_{k-1} E_{k-1}(\omega_k^{a,1}1_U h) \right\|_t \lesssim \|\{\chi_k\}_{k \in \mathbb{Z}}\|_{\text{Car}^1(\mathcal{D})} \cdot \|1_U h\|_{L^t(\mathbb{R}^N, \mu; \mathbb{C})}.$$

Using (1.6) and reasoning as in the proof of Lemma 4.4, we conclude that the right hand side above is bounded by a constant multiple of  $\mu(U)^{1/t}$ .  $\square$

We finish the proof of (10.6).

Recall that it suffices to estimate (10.10). Fix a real number  $t > q \vee s$ , where  $s \in [2, \infty)$  is such that  $X^*$  has cotype  $s$ . Let us also introduce Rademacher variables  $\varepsilon' = \{\varepsilon'_R\}_{R \in \mathcal{D}'} \in \Omega'$  that are independent of  $\{\varepsilon_Q\}_{Q \in \mathcal{D}}$ . By (10.12)

$$\begin{aligned} & \left\| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \langle g \rangle_R \right\|_{L^q(\mathbb{R}^N \times \Omega; X^*)} \\ &= \left\| \sum_{R \in \mathcal{D}'} \varepsilon'_R \left( \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right) 1_R \langle g \rangle_R \right\|_{L^q(\Omega' \times \mathbb{R}^N \times \Omega; X^*)}. \end{aligned} \tag{10.15}$$

By Hölder’s inequality (10.15) is bounded by

$$\left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* d_j E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N \times \Omega; X^*)} \leq \left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* d_j E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N; L^1(\Omega; X^*))}, \tag{10.16}$$

where  $\varepsilon^* = \{\varepsilon_j^* : j \in \mathbb{Z}\} \in \Omega^*$  are Rademacher random variables and

$$d_j : \mathbb{R}^N \rightarrow L^1(\Omega) : x \mapsto \left( \varepsilon \mapsto \sum_{R \in \mathcal{D}'_j} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a}(x) \right).$$

Concluding from above and using Lemma 10.8, we see that left hand side of (10.15) is bounded by (a constant multiple of)

$$\sup_{\substack{P \in \mathcal{D}' \\ \mu(P) \neq 0}} \frac{1}{\mu(P)^{1/t}} \cdot \underbrace{\left\| 1_P \sum_{j: 2^j \leq \ell(P)} \varepsilon_j^* \|d_j(\cdot)\|_{L^1(\Omega; \mathbb{C})} \right\|_{L^1(\mathbb{R}^N \times \Omega^*; \mathbb{R})}}_{=: \Sigma(P)} \cdot \|g\|_q. \tag{10.17}$$

To estimate the Carleson norm, we fix  $P \in \mathcal{D}'$  for which  $\mu(P) \neq 0$ . By (10.12),

$$\Sigma(P) = \left\| \sum_{\substack{R \in \mathcal{D}' \\ R \subset P}} \varepsilon'_R \right\| \left\| \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right\|_{L^1(\Omega; \mathbb{C})} \left\| \right\|_{L^1(\mathbb{R}^N \times \Omega'; \mathbb{C})}.$$

By Khinchine and Kahane–Khinchine inequalities,

$$\Sigma(P) \lesssim \left\| \left( \sum_{\substack{R \in \mathcal{D}' \\ R \subset P}} \left\| \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right\|_{L^2(\Omega; \mathbb{C})}^2 \right)^{1/2} \right\|_{L^1(\mathbb{R}^N; \mathbb{C})}. \tag{10.18}$$

Since  $L^2(\Omega; \mathbb{C})$  has cotype 2, see (2.3), we obtain

$$\begin{aligned} \Sigma(P) &\lesssim \left\| \left\| \sum_{\substack{R \in \mathcal{D}' \\ R \subset P}} \varepsilon'_R \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right\|_{L^2(\Omega'; L^2(\Omega; \mathbb{C}))} \right\|_{L^1(\mathbb{R}^N; \mathbb{C})} \\ &= \left\| \left\| \sum_{\substack{R \in \mathcal{D}' \\ R \subset P}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right\|_{L^2(\Omega; \mathbb{C})} \right\|_{L^1(\mathbb{R}^N; \mathbb{C})} \lesssim \left\| \sum_{\substack{R \in \mathcal{D}' \\ R \subset P}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,R^a} \right\|_{L^1(\mathbb{R}^N \times \Omega; \mathbb{C})}. \end{aligned}$$

Suppose that  $M \in \mathbb{N}_0$  is such that  $P^a \in \mathcal{L}'_M$ . Because cubes in a fixed layer  $\mathcal{L}'_m$ ,  $m > M$ , are disjoint, we can estimate as follows

$$\begin{aligned} \Sigma(P) &\lesssim \left\| \sum_{\substack{R \in \mathcal{D}': R \subset P \\ R^a = P^a}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,P^a} \right\|_{L^1(\mathbb{R}^N \times \Omega; \mathbb{C})} \\ &\quad + \sum_{m=M+1}^{\infty} \left( \sum_{\substack{U \in \mathcal{L}'_m \\ U \subsetneq P}} \left\| \sum_{\substack{R \in \mathcal{D}': R \subset U \\ R^a = U}} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q,U} \right\|_{L^1(\mathbb{R}^N \times \Omega; \mathbb{C})} \right)^{1/t}. \end{aligned} \tag{10.19}$$

Using Lemma 10.13 and Lemma 2.12, we can estimate the right hand side of (10.19) as follows

$$\begin{aligned} \Sigma(P) &\lesssim \mu(P)^{1/t} + \sum_{m=M+1}^{\infty} \left( \sum_{U \in \mathcal{L}'_m; U \subsetneq P} \mu(U) \right)^{1/t} \\ &\lesssim \mu(P)^{1/t} + \sum_{m=M+1}^{\infty} (1 - \tau)^{(m-M-1)/t} \mu(P)^{1/t} \lesssim \mu(P)^{1/t}. \end{aligned}$$

The proof of (10.6) finishes by substituting the estimate above in (10.17).

**Proving estimate (10.7).** Randomizing and using Hölder’s inequality as in connection with the paraproduct operator, we get

$$\begin{aligned} &\left| \sum_{\substack{Q \in \mathcal{D} \\ S(Q) \neq \emptyset}} \left\langle T^* b_{R_0} \frac{\langle g \rangle_{\mathbb{R}^N}}{\langle b_{R_0} \rangle_{\mathbb{R}^N}}, D_Q^{a,1} f \right\rangle \right| \\ &\lesssim |\langle g \rangle_{\mathbb{R}^N}| \cdot \|f\|_p \cdot \left\{ \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q (D_Q^{a,1})^* (T^* b_{R_0}) \right\|_q + \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q E_{\log_2(\ell(Q)) - 1} (\omega_Q^{a,1} T^* b_{R_0}) \right\|_q \right\}. \end{aligned}$$

Observe that  $|\langle g \rangle_{\mathbb{R}^N}| \lesssim \mu(\mathbb{R}^N)^{-1/q} \|g\|_q$ . Hence, it suffices to show that the quantity inside the parentheses is bounded by a constant multiple of  $\mu(\mathbb{R}^N)^{1/q}$ .

To this end, we first use Theorem 4.16 with  $X = \mathbb{C}$  and (1.6), we get

$$\left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q (D_Q^{a,1})^* (T^* b_{R_0}) \right\|_q \lesssim \|T^* b_{R_0}\|_q \lesssim \mu(\mathbb{R}^N)^{1/q}.$$

On the other hand, since the family  $\{E_k\}_{k \in \mathbb{Z}}$  of operators in  $L^q(\mathbb{R}^N, \mu)$  is  $\mathcal{R}$ -bounded by Stein's inequality [15], we find that

$$\begin{aligned} & \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q E_{\log_2(\ell(Q)) - 1}(\omega_Q^{a,1} T^* b_{R_0}) \right\|_q \\ &= \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k E_{k-1}(\omega_k^{a,1} T^* b_{R_0}) \right\|_q \lesssim \|T^* b_{R_0}\|_\infty \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \omega_k^{a,1} \right\|_q. \end{aligned}$$

By (1.6), we have  $\|T^* b_{R_0}\|_\infty \lesssim 1$ . Because  $|\omega_k^{a,1}| \lesssim 1_{\{b_{k-1}^a \neq b_k^a\}}$   $\mu$ -almost everywhere, see Lemma 3.6, we can use Lemma 4.3 with  $f \equiv 1$  for

$$\left\| \sum_{k \in \mathbb{Z}} \varepsilon_k \omega_k^{a,1} \right\|_q \lesssim \|1\|_q \lesssim \mu(\mathbb{R}^N)^{1/q}.$$

This concludes the proof of estimate (10.7).

**11. Preparations for comparable cubes.** During the course of the present and following section, we prove Proposition 11.2. It controls a part of the summation in (5.1), involving cubes that are close to each other in their position and size.

We write  $Q \sim R$  for  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$  if

$$2^{-r} \ell(R) \leq \ell(Q) \leq \ell(R) \text{ and } \text{dist}(Q, R) < \ell(Q) = \ell(Q) \wedge \ell(R). \tag{11.1}$$

Note that if  $Q \sim R$ , then  $\ell(Q) \leq \ell(R) \leq D(Q, R) \leq (2 + 2^r)\ell(Q)$ , so that all of these quantities are comparable.

A few words about implicit constants: In the previous sections we have performed estimates where the implicit constants can depend on the parameter  $r$ , introduced in Section 2. At this stage we introduce two new auxiliary parameters  $\eta \in (0, 1)$  and  $\nu \in (0, 1)$ . In the sequel we need to keep track of the dependence of estimates on the parameters  $r, \nu$  and  $\eta$  explicitly.

For the following proposition, we recall that all UMD spaces have a finite cotype.

**PROPOSITION 11.2.** *Under the assumptions of Theorem 1.7, we have*

$$\mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \sim R}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \tag{11.3}$$

$$\lesssim (C(r, \eta, \nu) + (C(r, \eta)\nu^{1/t} + C(r)\eta^{1/t})) \|T\|_{\mathcal{L}(L^p(\mu; X))} \|g\|_{L^q(\mu; X^*)} \|f\|_{L^p(\mu; X)}$$

for every  $f \in L^p(X)$  and  $g \in L^q(X^*)$ . Here  $1/p + 1/q = 1$  and  $t > (s \vee q) \vee p$ , where both  $X$  and  $X^*$  have cotype  $s \in [2, \infty)$ .

The strategy of the proof of this proposition is as follows: at the end of this section we consider a separated part of the summation in (11.3), where expectations over dyadic systems are not required. In the following section a (more complicated) intersecting part of the sum in (11.3) is treated, and the expectations are crucial therein.

Here are preparations for the proof of Proposition 11.2: given  $R \in \mathcal{D}'$ , there are at most  $C = C(r, N)$  cubes  $Q \in \mathcal{D}$  satisfying (11.1). Hence, without essential loss of

generality, it suffices consider a finite number of subseries of the general form

$$\mathbf{E}_{\mathcal{D}}\mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right|, \tag{11.4}$$

where  $Q = Q(R) \in \mathcal{D}_{R\text{-good}}$  inside the summation satisfies  $Q \sim R$ . At this stage we fix one series like this, and the convention that  $Q$  is implicitly a function of  $R$  will be maintained without further notice. Furthermore, without loss of generality, it is possible to act as if the map  $R \mapsto Q(R)$  was invertible, so that (11.4) could also be written in terms of the summation variable  $Q \in \mathcal{D}$ .

Proceeding as in Section 9, we find that (11.4) can be dominated from above by a sum of nine terms, each of them being of the general form

$$\sum_{i,j=1}^{2^N} \mathbf{E}_{\mathcal{D}}\mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} \langle 1_{R_j} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \langle f_Q \rangle_{Q_i} \right|, \tag{11.5}$$

where  $\langle g_R \rangle_{R_j} = \langle 1_R g_k \rangle_{R_j} = \langle g_k \rangle_{R_j}$  if  $R \in \mathcal{D}'_k$  (similarly for  $f$ 's), and the summands are determined by the following choices:

$$(g_k, \psi_{R,j}) \in \{(s_k, b_{R^a}), (h_k, b_{R^a}), (u_k, b_{R^a})\} \tag{11.6}$$

and

$$(f_k, \varphi_{Q,i}) \in \{(\bar{s}_k, b_{Q^a}), (\bar{h}_k, b_{Q^a}), (\bar{u}_k, b_{Q^a})\}. \tag{11.7}$$

Here

$$s_k = 1_{\{b_{k-1}^{a,2} = b_k^{a,2}\}} \left( \frac{E_{k-1}g}{E_{k-1}b_{k-1}^{a,2}} - \frac{E_k g}{E_k b_k^{a,2}} \right), \quad \bar{s}_k = 1_{\{b_{k-1}^{a,1} = b_k^{a,1}\}} \left( \frac{E_{k-1}f}{E_{k-1}b_{k-1}^{a,1}} - \frac{E_k f}{E_k b_k^{a,1}} \right);$$

$$h_k = 1_{\{b_{k-1}^{a,2} \neq b_k^{a,2}\}} \frac{E_{k-1}g}{E_{k-1}b_{k-1}^{a,2}}, \quad \bar{h}_k = 1_{\{b_{k-1}^{a,1} \neq b_k^{a,1}\}} \frac{E_{k-1}f}{E_{k-1}b_{k-1}^{a,1}};$$

and

$$u_k = -1_{\{b_{k-1}^{a,2} \neq b_k^{a,2}\}} \frac{E_k g}{E_k b_k^{a,2}}, \quad \bar{u}_k = -1_{\{b_{k-1}^{a,1} \neq b_k^{a,1}\}} \frac{E_k f}{E_k b_k^{a,1}}.$$

Observe that, in any case,  $E_{k-1}g_k = g_k$  and  $E_{k-1}f_k = f_k$ .

Fix  $i, j \in \{1, 2, \dots, N\}$ .

For each cube  $Q$  in  $\mathbb{R}^N$ , define the boundary region

$$\delta_Q^\eta := (1 + \eta)Q \setminus (1 - \eta)Q,$$

where the parameter  $\eta > 0$  is to be chosen later. If  $R \in \mathcal{D}'$  and  $Q = Q(R)$ , we write

$$\begin{aligned} Q_{i,\partial} &:= Q_i \cap \delta_{R_j}^\eta; & Q_{i,\text{sep}} &:= (Q_i \setminus Q_{i,\partial}) \setminus (Q_i \cap R_j); & \Delta_{Q_i} &:= (Q_i \cap R_j) \setminus Q_{i,\partial}; \\ R_{j,\partial} &:= R_j \cap \delta_{Q_i}^\eta; & R_{j,\text{sep}} &:= (R_j \setminus R_{j,\partial}) \setminus (Q_i \cap R_j); & \Delta_{R_j} &:= (Q_i \cap R_j) \setminus R_{j,\partial}. \end{aligned}$$



Observe that the following unions are disjoint:

$$Q_i = \Delta_{Q_i} \cup Q_{i,\text{sep}} \cup Q_{i,\partial}, \quad R_j = \Delta_{R_j} \cup R_{j,\text{sep}} \cup R_{j,\partial}.$$

Hence, we can write the matrix coefficient in (11.5) as

$$\begin{aligned} \langle 1_{R_j} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle &= \langle 1_{R_{j,\text{sep}}} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle + \langle 1_{R_{j,\partial}} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle \\ &\quad + \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}} \varphi_{Q,i}) \rangle \\ &\quad + \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{Q_{i,\partial}} \varphi_{Q,i}) \rangle + \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{Q_{i,\text{sep}}} \varphi_{Q,i}) \rangle \\ &=: M_1(R) + M_2(R) + M_3(R) + M_4(R) + M_5(R). \end{aligned} \tag{11.8}$$

Using these preparations, it suffices to estimate the following quantity:

$$\mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} (M_1(R) + M_2(R) + M_3(R) + M_4(R) + M_5(R)) \langle f_Q \rangle_{Q_i} \right|, \tag{11.9}$$

where  $i, j \in \{1, 2, \dots, 2^N\}$  and  $Q = Q(R) \in \mathcal{D}_{R\text{-good}}$  satisfies the condition  $Q \sim R$  inside the summation.

**The separated part.** Recall that our aim now is to prove Proposition 11.2. We have reduced this to a problem of estimating the sum (11.9) involving, among others, terms of the form

$$M_1(R) + M_5(R) = \langle 1_{R_{j,\text{sep}}} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle + \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{Q_{i,\text{sep}}} \varphi_{Q,i}) \rangle,$$

where  $Q = Q(R) \in \mathcal{D}_{R\text{-good}}$  satisfies  $Q \sim R$ . In both cases,  $M_1$  and  $M_5$ , the two indicators are associated with sets separated from each other. Hence, a decoupling estimate can be used to establish the following lemma.

LEMMA 11.10. *Suppose that  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ . Then*

$$\left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} (M_1(R) + M_5(R)) \langle f_Q \rangle_{Q_i} \right| \lesssim C(r, \eta) \|g\|_q \|f\|_p.$$

*Proof.* We focus on summation over the terms  $M_1(R)$ . The treatment of summation over the terms  $M_5(R)$  is analogous. If  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$  satisfies  $\ell(Q) \leq \ell(R)$ , we write  $T_{RQ} := 1_{Q=Q(R)} \langle 1_{R_{j,\text{sep}}} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle$ . It suffices to estimate the series

$$\Sigma := \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} \langle g_R \rangle_{R_j} T_{RQ} \langle f_Q \rangle_{Q_i}.$$

Assume that  $T_{RQ} \neq 0$  inside the summation. Then  $Q = Q(R)$ , so that  $Q \sim R$  and, by (1.4) and (1.2),

$$\begin{aligned} |T_{RQ}| &= \left| \int_{R_{j,\text{sep}}} \int_{Q_i} \psi_{R,j}(x) K(x, y) \varphi_{Q,i}(y) d\mu(y) d\mu(x) \right| \\ &\leq \frac{\mu(R_{j,\text{sep}}) \mu(Q_i)}{\text{dist}(R_{j,\text{sep}}, Q_i)^d} \lesssim C(\eta) \frac{\mu(R_j) \mu(Q_i)}{\ell(R_j)^d} \simeq C(\eta) \mu(R_j) \mu(Q_i) \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{d+\alpha}}. \end{aligned}$$

Using Lemma 6.6 and then estimating as in the end of Section 9, we find that

$$|\Sigma| \lesssim C(r, \eta) \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbf{P}^{\otimes \mu; X^*})} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P}^{\otimes \mu; X})} \lesssim C(r, \eta) \|g\|_q \|f\|_p.$$

□

**12. Intersecting part of comparable cubes.** In this section we deal with the remaining part of the comparable cubes, finishing the proof of Proposition 11.2. This will be the most technical part of the entire proof: It still involves various further decompositions and case-by-case analysis, until all different pieces are finally estimated.

Recalling the preparations in Section 11, we observe that it remains to estimate a summation like (11.9) but involving only terms of the form  $M_2(R) + M_3(R) + M_4(R)$ . Part of this summation involves boundary terms that are handled by probabilistic methods, e.g. by taking expectations over the random dyadic systems  $\mathcal{D}$  and  $\mathcal{D}'$ , but we will also introduce a third random dyadic system  $\mathcal{D}^*$ . The assumption that there is an  $L^\infty$ -accretive system for  $T^*$  is used to handle the non-boundary terms.

We aim to prove the following lemma.

LEMMA 12.1. *We have*

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} (M_2(R) + M_3(R) + M_4(R)) \langle f_Q \rangle_{Q_i} \right| \\ & \lesssim (C(r, \eta, \nu) + (C(r, \eta) \nu^{1/t} + C(r) \eta^{1/t}) \|T\|_{\mathcal{L}(L^p(\mu; X))}) \|g\|_{L^q(\mu; X^*)} \|f\|_{L^p(\mu; X)}. \end{aligned}$$

The proof of this lemma is a consequence of various lemmata, namely: 12.13, 12.15, and 12.18. Let us briefly indicate the structure of the proof. Since  $M_2(R)$  and  $M_4(R)$  are so called  $\eta$ -boundary terms, the main difficulties lie in estimating summation involving the term

$$M_3(R) = \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}} \varphi_{Q,i}) \rangle = \alpha_1(R) + \alpha_2(R) + \alpha_3(R),$$

where the last decomposition depends on a new random dyadic system  $\mathcal{D}^*$ , see (12.3). The terms  $\alpha_2(R)$  and  $\alpha_3(R)$  are also  $\eta$ -boundary terms.

The term  $\alpha_1(R)$  will further be expanded in (12.4) and Lemma 12.5 as

$$\alpha_1(R) = A_1(R) + A_2(R) + A_3(R)' + (A_3(R) - A_3'(R)),$$

where  $A_1(R)$ ,  $A_2(R)$ , and  $A_3'(R)$  are so called  $\nu$ -boundary terms. Hence, the main obstacle is to estimate  $A_3(R) - A_3'(R)$ ; the assumption that there is an  $L^\infty$ -accretive systems for  $T^*$  will be exploited here.

**Decomposition of  $M_3(R)$ .** In order to decompose  $M_3(R)$ , we first introduce a random dyadic system

$$\mathcal{D}^* = \mathcal{D}(\beta^*)$$

that is independent of both  $\mathcal{D}$  and  $\mathcal{D}'$ . Fix  $j(\eta) \in \mathbb{Z}$  such that  $\eta/64 \leq 2^{j(\eta)} < \eta/32$ . Then, for every  $R \in \mathcal{D}'$ , we define a family

$$\mathcal{G} = \mathcal{G}(R) := \mathcal{D}_{\log_2(s)}^*$$

of cubes with side length

$$s = 2^{j(\eta)} \ell(Q_i) = 2^{j(\eta)} \cdot (\ell(Q_i) \wedge \ell(R_j)), \tag{12.2}$$

where  $Q = Q(R) \in \mathcal{D}$ . More precisely,  $\mathcal{G}$  is a subfamily of  $\mathcal{D}^*$  that depends on  $R$ ,  $Q = Q(R)$ , and  $\eta$ .

Let  $\Delta_{Q_i}^{\mathcal{G}}, \Delta_{R_j}^{\mathcal{G}} \subset Q_i \cap R_j$  be the following adaptations of  $\Delta_{Q_i}$  and  $\Delta_{R_j}$  to  $\mathcal{G}$ . If necessary, we enlarge the latter sets so that, if  $G \in \mathcal{G}$ , either at least one of the two intersections  $G \cap \Delta_{Q_i}^{\mathcal{G}}$  and  $G \cap \Delta_{R_j}^{\mathcal{G}}$  is empty, or we have  $G \cap \Delta_{Q_i}^{\mathcal{G}} = G \cap \Delta_{R_j}^{\mathcal{G}} = G$  and  $5G \subset Q_i \cap R_j$ . Moreover, this is done in such a way that we can write

$$\Delta_{Q_i}^{\mathcal{G}} = \Delta_{Q_i} \cup \Delta_{Q_i}^{\partial}, \quad \Delta_{R_j}^{\mathcal{G}} = \Delta_{R_j} \cup \Delta_{R_j}^{\partial}$$

both as disjoint unions, so that  $\Delta_{Q_i}^{\partial} \subset Q_{i,\partial} \cap R_j$  and  $\Delta_{R_j}^{\partial} \subset R_{j,\partial} \cap Q_i$ . Observe that

$$\begin{aligned} M_3(R) &= \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}} \varphi_{Q,i}) \rangle \\ &= \langle 1_{\Delta_{R_j}^{\mathcal{G}}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\mathcal{G}}} \varphi_{Q,i}) \rangle - \langle 1_{\Delta_{R_j}^{\partial}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\mathcal{G}}} \varphi_{Q,i}) \rangle - \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\partial}} \varphi_{Q,i}) \rangle \\ &=: \alpha_1(R) + \alpha_2(R) + \alpha_3(R). \end{aligned} \tag{12.3}$$

The terms in this decomposition depend on  $\mathcal{D}^*$ .

In order to define  $\nu$ -boundary terms, we let  $R \in \mathcal{D}'$  and write

$$G_\nu = G_\nu(R) = \bigcup_{G \in \mathcal{G}(R)} \delta_G^\nu, \quad \delta_G^\nu = (1 + \nu)G \setminus (1 - \nu)G.$$

We also write  $\tilde{G} = G \setminus G_\nu$  if  $G \in \mathcal{G} = \mathcal{G}(R)$ . Define

$$\Delta'_{Q_i} = \Delta_{Q_i}^{\mathcal{G}} \cap G_\nu, \quad \tilde{\Delta}_{Q_i} = \Delta_{Q_i}^{\mathcal{G}} \setminus G_\nu$$

and similarly for  $\Delta'_{R_j}$ . Then we have the disjoint unions

$$\Delta_{Q_i}^{\mathcal{G}} = \Delta'_{Q_i} \cup \tilde{\Delta}_{Q_i}, \quad \Delta_{R_j}^{\mathcal{G}} = \Delta'_{R_j} \cup \tilde{\Delta}_{R_j}.$$

Hence, we can write

$$\begin{aligned} \alpha_1(R) &= \langle 1_{\Delta_{R_j}^{\mathcal{G}}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\mathcal{G}}} \varphi_{Q,i}) \rangle \\ &= \langle 1_{\Delta'_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\mathcal{G}}} \varphi_{Q,i}) \rangle + \langle 1_{\tilde{\Delta}_{R_j}} \psi_{R,j}, T(1_{\Delta'_{Q_i}} \varphi_{Q,i}) \rangle + \langle 1_{\tilde{\Delta}_{R_j}} \psi_{R,j}, T(1_{\tilde{\Delta}_{Q_i}} \varphi_{Q,i}) \rangle \\ &=: A_1(R) + A_2(R) + A_3(R). \end{aligned} \tag{12.4}$$

**Estimate for a non-boundary part.** We need to extract the non-boundary terms. This is done in the following lemma which gives us a decomposition of  $A_3(R)$ ; therein  $A_3(R) - A'_3(R)$  is a non-boundary term. The proof of the lemma uses the fact that there is an  $L^\infty$ -accretive system for  $T^*$ .

LEMMA 12.5. *Let  $R \in \mathcal{D}'$ . Then  $A_3(R)$  can be written as  $A'_3(R) + (A_3(R) - A'_3(R))$ , where*

$$|A_3(R) - A'_3(R)| \lesssim C(r, \eta, \nu)\mu(Q_i \cap R_j)$$

and there are functions  $b_{R,G,j} : \mathbb{R}^N \rightarrow \mathbb{C}$ , satisfying  $\|b_{R,G,j}\|_{L^\infty(\mu)} \lesssim 1$  if  $R \in \mathcal{D}'$  and  $G \in \mathcal{G}(R)$ , such that

$$A'_3(R) = \sum_{\substack{G \in \mathcal{G}(R) \\ \tilde{G} \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G}} \langle 1_{\tilde{G}} b_{R,G,j}, T(1_{(1+\nu)\tilde{G}} \varphi_{Q,i}) \rangle.$$

Here  $\tilde{G} = G \setminus G_\nu$  for every  $G \in \mathcal{G}(R)$ .

*Proof.* We expand  $A_3(R)$  into a double series, where a typical summand is of the form

$$\langle 1_G 1_{\tilde{\Delta}_{R_j}} \psi_{R,j}, T(1_H 1_{\tilde{\Delta}_{Q_i}} \varphi_{Q,i}) \rangle, \quad G, H \in \mathcal{G}. \tag{12.6}$$

Let us begin with estimating these quantities, and there are two cases to be treated.

First, if  $G \neq H$ , then

$$\ell(Q_i) \lesssim C(\eta, \nu) \text{dist}(G \cap \tilde{\Delta}_{R_j}, H \cap \tilde{\Delta}_{Q_i}).$$

Hence, by (1.4) and (1.2),

$$\begin{aligned} & |\langle 1_G 1_{\tilde{\Delta}_{R_j}} \psi_{R,j}, T(1_H 1_{\tilde{\Delta}_{Q_i}} \varphi_{Q,i}) \rangle| \\ &= \left| \int_{G \cap \tilde{\Delta}_{R_j}} \int_{H \cap \tilde{\Delta}_{Q_i}} \psi_{R,j}(x) K(x, y) \varphi_{Q,i}(y) d\mu(y) d\mu(x) \right| \\ &\lesssim C(\eta, \nu) \frac{\mu(Q_i \cap R_j) \mu(Q_i \cap R_j)}{\ell(Q_i)^d} \lesssim C(\eta, \nu) \mu(Q_i \cap R_j). \end{aligned} \tag{12.7}$$

Here we also used the facts that  $\tilde{\Delta}_{R_j} \cup \tilde{\Delta}_{Q_i} \subset Q_i \cap R_j$  and  $\mu(Q_i) \lesssim \ell(Q_i)^d$ .

Then we consider the case  $G = H$ . By construction,

$$\begin{aligned} & \langle 1_G 1_{\tilde{\Delta}_{R_j}} \psi_{R,j}, T(1_G 1_{\tilde{\Delta}_{Q_i}} \varphi_{Q,i}) \rangle \\ &= \begin{cases} \langle 1_{\tilde{G}} \psi_{R,j}, T(1_{\tilde{G}} \varphi_{Q,i}) \rangle, & \text{if } G = G \cap \Delta_{R_j}^G = G \cap \Delta_{Q_i}^G; \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{12.8}$$

In what follows, we assume that  $G = G \cap \Delta_{R_j}^G = G \cap \Delta_{Q_i}^G$ .

Consider the decomposition

$$\begin{aligned} \langle 1_{\tilde{G}} \psi_{R,j}, T(1_{\tilde{G}} \varphi_{Q,i}) \rangle &= \langle 1_{\tilde{G}} \psi_{R,j}, T(\varphi_{Q,i}) \rangle - \langle 1_{\tilde{G}} \psi_{R,j}, T(1_{\mathbb{R}^N \setminus 5\tilde{G}} \varphi_{Q,i}) \rangle \\ &\quad - \langle 1_{\tilde{G}} \psi_{R,j}, T(1_{5\tilde{G} \setminus (1+\nu)\tilde{G}} \varphi_{Q,i}) \rangle - \langle 1_{\tilde{G}} \psi_{R,j}, T(1_{(1+\nu)\tilde{G}} \varphi_{Q,i}) \rangle. \end{aligned} \tag{12.9}$$

The fourth term in the right hand side will only contribute to  $A'_3(R)$ . The first and third terms in the right hand side are estimated as follows.

Using both (11.6) and (11.7) together with (1.5) and (1.6), we obtain the estimate  $\|T(\varphi_{Q,i})\|_{L^\infty(\mu)} + \|\psi_{R,j}\|_{L^\infty(\mu)} \lesssim 1$ . In particular,

$$|\langle 1_{\tilde{G}}\psi_{R,j}, T(\varphi_{Q,i}) \rangle| \lesssim \mu(\tilde{G}) \leq \mu(Q_i \cap R_j).$$

Then we consider the third term in the right hand side of (12.9). Recall that we are in the case  $G = G \cap \Delta_{R_j}^G = G \cap \Delta_{Q_i}^G$  hence, by construction,  $5\tilde{G} \subset 5G \subset Q_i \cap R_j$ . For this reason we can repeat the argument (12.7) which, in turn, gives

$$|\langle 1_{\tilde{G}}\psi_{R,j}, T(1_{5\tilde{G} \setminus (1+\nu)\tilde{G}}\varphi_{Q,i}) \rangle| \lesssim C(\eta, \nu)\mu(Q_i \cap R_j).$$

It remains to consider the second term in the right hand side of (12.9). Part of it will contribute to  $A_3(R)$ . We begin with certain preparations, and first denote

$$\tau := T(1_{\mathbb{R}^N \setminus 5\tilde{G}}\varphi_{Q,i}).$$

Using (1.3), reasoning as in (9.17) with  $R_1$  replaced by  $5\tilde{G}$ , and observing the fact that  $\|\varphi_{Q,i}\|_{L^\infty(\mu)} \lesssim 1$ , we see that

$$|\tau(x) - \tau(y)| \lesssim 1, \quad x, y \in \tilde{G}. \tag{12.10}$$

We use the fact that there exists an  $L^\infty$ -accretive system for  $T^*$ . Let  $b_{\tilde{G}}$  be a function which is supported on  $\tilde{G}$ , whose average over  $\tilde{G}$  is one, and

$$\|b_{\tilde{G}}\|_{L^\infty(\mu)} + \|T^*(b_{\tilde{G}})\|_{L^\infty(\mu)} \lesssim 1.$$

Let us denote  $\beta_{\tilde{G}} = \langle b_{\tilde{G}}/\mu(\tilde{G}), \tau \rangle$ . By properties of  $b_{\tilde{G}}$  and (12.10),

$$|\tau(x) - \beta_{\tilde{G}}| = |\langle b_{\tilde{G}}/\mu(\tilde{G}), \tau(x) - \tau \rangle| \lesssim 1, \quad x \in \tilde{G}. \tag{12.11}$$

After these preparations, we write

$$\langle 1_{\tilde{G}}\psi_{R,j}, T(1_{\mathbb{R}^N \setminus 5\tilde{G}}\varphi_{Q,i}) \rangle = \langle 1_{\tilde{G}}\psi_{R,j}, \tau - \beta_{\tilde{G}} \rangle + \langle 1_{\tilde{G}}\psi_{R,j}, \beta_{\tilde{G}} \rangle.$$

By (12.11) and (11.6), we have the estimate  $|\langle 1_{\tilde{G}}\psi_{R,j}, \tau - \beta_{\tilde{G}} \rangle| \lesssim \mu(\tilde{G}) \leq \mu(Q_i \cap R_j)$ . To treat the term  $\langle 1_{\tilde{G}}\psi_{R,j}, \beta_{\tilde{G}} \rangle = \beta_{\tilde{G}}\langle 1_{\tilde{G}}\psi_{R,j}, 1 \rangle$ , we write

$$\begin{aligned} \beta_{\tilde{G}} &= \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, \tau \right\rangle = \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, T(\varphi_{Q,i}) \right\rangle - \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, T(1_{5\tilde{G} \setminus (1+\nu)\tilde{G}}\varphi_{Q,i}) \right\rangle \\ &\quad - \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, T(1_{(1+\nu)\tilde{G}}\varphi_{Q,i}) \right\rangle - \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, T(1_{\tilde{G}}\varphi_{Q,i}) \right\rangle. \end{aligned} \tag{12.12}$$

Observe that the first and second term on the right hand side of (12.12) are bounded in absolute value by a constant  $C \lesssim C(\eta, \nu)$ . This follows from the properties of  $b_{\tilde{G}}$  and the fact  $\|T(\varphi_{Q,i})\|_{L^\infty(\mu)} \lesssim 1$  for the first term and, by reasoning as in (12.7), for the second term.

For the last term in the right hand side of (12.12), we use  $\|T^*(b_{\tilde{G}})\|_{L^\infty(\mu)} \lesssim 1$  for

$$\left| \left\langle \frac{b_{\tilde{G}}}{\mu(\tilde{G})}, T(1_{\tilde{G}}\varphi_{Q,i}) \right\rangle \right| = |\langle T^*(b_{\tilde{G}}), 1_{\tilde{G}}\varphi_{Q,i} \rangle|/\mu(\tilde{G}) \lesssim 1.$$

Regrouping the terms shows that  $\langle 1_{\tilde{G}}\psi_{R,j}, T(1_{\mathbb{R}^N \setminus 5\tilde{G}}\varphi_{Q,i}) \rangle$  can be expressed as a sum of two terms, the first one being bounded in absolute value by a constant  $C \lesssim C(\eta, \nu)\mu(Q_i \cap R_j)$ , and the second one being

$$-\langle b_{\tilde{G}}, T(1_{(1+\nu)\tilde{G} \setminus \tilde{G}}\tilde{\varphi}_{Q,i}) \rangle \langle 1_{\tilde{G}}\psi_{R,j}, 1/\mu(\tilde{G}) \rangle.$$

As said in the beginning of the proof, we expand  $A_3(R)$  by using the cubes in  $\mathcal{G}$  as follows:

$$\begin{aligned} A_3(R) &= \langle 1_{\tilde{\Delta}_{R_j}}\psi_{R,j}, T(1_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}) \rangle \\ &= \sum_{\substack{G,H \in \mathcal{G} \\ G \neq H}} \langle 1_G 1_{\tilde{\Delta}_{R_j}}\psi_{R,j}, T(1_H 1_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}) \rangle + \sum_{G \in \mathcal{G}} \langle 1_G 1_{\tilde{\Delta}_{R_j}}\psi_{R,j}, T(1_G 1_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}) \rangle. \end{aligned}$$

In both of the series above, the finite number of summands depends on  $N$  and  $\eta$ . Hence, using the estimates above for a typical summand (12.6), we get

$$A_3(R) = A'_3(R) + (A_3(R) - A'_3(R)),$$

where  $|A_3(R) - A'_3(R)| \lesssim C(\eta, \nu)\mu(Q_i \cap R_j)$  and  $A'_3(R)$  is the following quantity:

$$\sum_{\substack{G \in \mathcal{G} \\ \tilde{G} \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G}} \left[ \langle b_{\tilde{G}}, T(1_{(1+\nu)\tilde{G} \setminus \tilde{G}}\tilde{\varphi}_{Q,i}) \rangle \langle 1_{\tilde{G}}\psi_{R,j}, 1/\mu(\tilde{G}) \rangle - \langle 1_{\tilde{G}}\psi_{R,j}, T(1_{(1+\nu)\tilde{G} \setminus \tilde{G}}\tilde{\varphi}_{Q,i}) \rangle \right].$$

It is straightforward to verify that  $A'_3(R)$  is of the required form. □

The summation involving the non-boundary terms  $A_3(R) - A'_3(R)$ , given by Lemma 12.5, is controlled by the following result. It gives a uniform estimate for the sum with respect to all systems of dyadic cubes  $\mathcal{D}^*$ ; in particular, no expectations over  $\mathcal{D}^*$  are needed.

LEMMA 12.13. *Let  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ . Then estimate*

$$\left| \sum_{R \in \mathcal{D}^*} \langle g_R \rangle_{R_j} (A_3(R) - A'_3(R)) \langle f_Q \rangle_{Q_i} \right| \lesssim C(r, \eta, \nu) \|g\|_q \|f\|_p$$

is valid for every dyadic system  $\mathcal{D}^*$ .

*Proof.* Let us denote  $A_4(R) := A_3(R) - A'_3(R)$ . By randomizing and using Hölder’s inequality,

$$\begin{aligned} &\left| \sum_{R \in \mathcal{D}^*} \langle g_R \rangle_{R_j} A_4(R) \langle f_Q \rangle_{Q_i} \right| \\ &= \left| \iint_{\Omega \times \mathbb{R}^N} \sum_{S \in \mathcal{D}^*} \varepsilon_S 1_{S_j}(x) \langle g_S \rangle_{S_j} \sum_{R \in \mathcal{D}^*} \varepsilon_R 1_{Q_i}(x) \frac{A_4(R)}{\mu(Q_i \cap R_j)} \langle f_Q \rangle_{Q_i} d\mathbf{P}(\varepsilon) d\mu(x) \right| \quad (12.14) \\ &\leq \left\| \sum_{S \in \mathcal{D}^*} \varepsilon_S 1_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbb{R}^N, \mathbf{P} \otimes \mu; X^*)} \left\| \sum_{R \in \mathcal{D}^*} \varepsilon_R 1_{Q_i} \frac{A_4(R)}{\mu(Q_i \cap R_j)} \langle f_Q \rangle_{Q_i} \right\|_{L^p(\mathbb{R}^N, \mathbf{P} \otimes \mu; X)}. \end{aligned}$$

Note that

$$1_{S_j} \langle g_S \rangle_{S_j} = 1_{S_j} \langle g_k \rangle_{S_j} = 1_{S_j} E_{k-1} g_k = 1_{S_j} g_k, \quad S \in \mathcal{D}'_k.$$

Hence, by using also the contraction principle, we have

$$\left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbb{R}^N, \mathbf{P} \otimes \mu; X^*)} \lesssim \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbb{R}^N, \mathbf{P} \otimes \mu; X^*)} \lesssim \|g\|_{L^q(\mathbb{R}^N, \mu; X^*)}.$$

In the last step we reasoned as in the end of Section 9.

Rewrite the second summation in the last line of (12.14) in terms of  $\mathcal{D}$ . Then using the contraction principle and the fact that  $|A_4(R)| \lesssim C(r, \eta, \nu) \mu(Q_i \cap R_j)$ , given by Lemma 12.5, results in estimates

$$\begin{aligned} & \left\| \sum_{R \in \mathcal{D}'} \varepsilon_R 1_{Q_i} \frac{A_4(R)}{\mu(Q_i \cap R_j)} \langle f_Q \rangle_{Q_i} \right\|_{L^p(\mathbb{R}^N, \mathbf{P} \otimes \mu; X)} \\ & \lesssim C(r, \eta, \nu) \left\| \sum_{Q \in \mathcal{D}} \varepsilon_Q 1_{Q_i} \langle f_Q \rangle_{Q_i} \right\|_{L^p(\mathbb{R}^N, \mathbf{P} \otimes \mu; X)} \lesssim C(r, \eta, \nu) \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbb{R}^N, \mathbf{P} \otimes \mu; X)}. \end{aligned}$$

Reasoning as in the end of Section 9 finishes the proof. □

**Estimate for  $\nu$ -boundary terms.** The following lemma controls a summation involving the  $\nu$ -boundary terms

$$A_1(R) + A_2(R) + A'_3(R).$$

Taking the expectations over the dyadic system  $\mathcal{D}^*$  is invaluable here and, on the other hand, this is the only place where these expectations are required. Elsewhere we obtain uniform estimates over these systems.

LEMMA 12.15. *Suppose that  $s \in [2, \infty)$  is such that both  $X$  and  $X^*$  have cotype  $s$ . Let  $t > (s \vee q) \vee p$  be a positive real number. Then*

$$\begin{aligned} \mathbf{E}_{\mathcal{D}^*} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} (A_1(R) + A_2(R) + A'_3(R)) \langle f_Q \rangle_{Q_i} \right| \\ \lesssim C(r, \eta) \nu^{1/t} \|T\|_{\mathcal{L}(L^p(\mu; X))} \|g\|_{L^q(\mu; X^*)} \|f\|_{L^p(\mu; X)} \end{aligned}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ .

*Proof.* First we focus on the sum involving the terms  $A_1$ ; these are defined in (12.4). Randomise and use Hölder’s inequality for the estimate

$$\begin{aligned} & \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} \langle 1_{\Delta'_j} \psi_{R,j}, T(1_{\Delta^g_{Q_i}} \varphi_{Q,i}) \rangle \langle f_Q \rangle_{Q_i} \right| \\ & = \left| \int_{\Omega} \left\langle \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{\Delta'_j} \psi_{S,j} \langle g_S \rangle_{S_j}, T \left( \sum_{R \in \mathcal{D}'} \varepsilon_R 1_{\Delta^g_{Q_i}} \varphi_{Q,i} \langle f_Q \rangle_{Q_i} \right) \right\rangle d\mathbf{P}(\varepsilon) \right| \tag{12.16} \\ & \leq \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{\Delta'_j} \psi_{S,j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \left\| T \left( \sum_{R \in \mathcal{D}'} \varepsilon_R 1_{\Delta^g_{Q_i}} \varphi_{Q,i} \langle f_Q \rangle_{Q_i} \right) \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \end{aligned}$$

First extract the operator norm from the second factor and index the summation in terms of  $\mathcal{D}$ . Then, by using the contraction principle and estimate

$$|1_{\Delta_{Q_i}^g} \varphi_{Q_i}| \leq 1_{Q_i}$$

which is valid  $\mu$ -almost everywhere, we see that the second factor in the last line of (12.16) is bounded (up to a constant multiple) by

$$\|T\|_{\mathcal{L}(L^p(\mu; X))} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \lesssim \|T\|_{\mathcal{L}(L^p(\mu; X))} \|f\|_{L^p(\mu; X)}.$$

In order to estimate the first factor in the last line of (12.16) we let  $S \in \mathcal{D}'_k$ , where  $k \in \mathbb{Z}$ . Due to (11.1) and (12.2), we have

$$\Delta'_{S_j} \subset G_v(S) = \bigcup_{G \in \mathcal{G}} \delta_G^v \subset \bigcup_{m=j(\eta)+k-r-1}^{j(\eta)+k-1} \bigcup_{G \in \mathcal{D}_m^*} \delta_G^v =: \delta^v(k).$$

As a consequence, we have  $|1_{\Delta'_{S_j}} \psi_{S_j}| \leq 1_{\delta^v(k)} 1_{S_j}$  pointwise  $\mu$ -almost everywhere. Using also the contraction principle and the assumption that  $t \geq q$ , we get

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}^*} \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{\Delta'_{S_j}} \psi_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \\ & \lesssim \mathbf{E}_{\mathcal{D}^*} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\delta^v(k)} \sum_{S \in \mathcal{D}'_k} 1_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \\ & \leq \left( \int_{\mathbb{R}^N} \left[ \mathbf{E}_{\mathcal{D}^*} \left\| \sum_{k \in \mathbb{Z}} \varepsilon_k 1_{\delta^v(k)}(x) \sum_{S \in \mathcal{D}'_k} 1_{S_j}(x) \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P}; X^*)}^t \right]^{q/t} d\mu(x) \right)^{1/q}. \end{aligned}$$

If  $x \in \mathbb{R}^N$ , the last integrand evaluated at  $x$  is of the form as in Proposition 2.4 with

$$\xi_k = \sum_{S \in \mathcal{D}'_k} 1_{S_j}(x) \langle g_S \rangle_{S_j} \in X^*.$$

The random variables

$$\rho_k := 1_{\delta^v(k)}(x)$$

as functions of  $\beta^* \in \Omega^*$ , where  $\Omega^*$  is the probability space supporting the distribution of the random dyadic system  $\mathcal{D}^*$ , belong to  $L^t(\Omega^*)$ , and they satisfy

$$\sup_{k \in \mathbb{Z}} \|1_{\delta^v(k)}(x)\|_{L^t(\Omega^*)} = \sup_{k \in \mathbb{Z}} \mathbf{P}_{\beta^*}(1_{\delta^v(k)}(x) = 1)^{1/t} \lesssim C(r, \eta) v^{1/t}.$$

Hence, by Proposition 2.4

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}^*} \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{\Delta'_{S_j}} \psi_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \\ & \lesssim C(r, \eta) v^{1/t} \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{S_j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \lesssim C(r, \eta) v^{1/t} \|g\|_{L^q(\mu; X^*)}. \end{aligned}$$



Combining the estimates above, we obtain the required estimate for summation involving terms  $A_1(R)$ .

Estimate for the sum involving terms  $A_2(R)$ , see (12.4), is similar to the estimate above, involving terms  $A_1(R)$ . We omit the details.

It remains to estimate the following sum involving terms  $A'_3(R)$ , see Lemma 12.5,

$$\mathbf{E}_{\mathcal{D}^*} \left| \sum_{R \in \mathcal{D}^*} \sum_{\substack{G \in \mathcal{G}(R) \\ \tilde{G} \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G}} \langle g_R \rangle_{R_j} \langle 1_{\tilde{G}} b_{R,G,j}, T(1_{(1+\nu)\tilde{G}} \tilde{G} \varphi_{Q,i}) \rangle \langle f_Q \rangle_{Q_i} \right|.$$

Observe that the inner summation involves only finitely many terms for every fixed  $R$  – in fact, the number of terms is bounded by a constant depending on  $\eta$  and  $N$ . Hence, by reindexing these cubes and using the triangle-inequality, we are left with estimating quantities of the form

$$E := \mathbf{E}_{\mathcal{D}^*} \left| \sum_{R \in \mathcal{D}^*} \langle g_R \rangle_{R_j} \langle 1_{\tilde{G}} b_{R,G,j}, T(1_{(1+\nu)\tilde{G}} \tilde{G} \varphi_{Q,i}) \rangle \langle f_Q \rangle_{Q_i} \right|,$$

where  $G = G(R) \in \mathcal{G}(R)$  inside the summation satisfies  $\tilde{G} \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G$ .

At this stage we randomise, apply Hölder’s inequality, and extract the operator norm in order to obtain the estimate

$$E \leq \|T\|_{\mathcal{L}(L^p(\mu; X))} \mathbf{E}_{\mathcal{D}^*} \left\{ \left\| \sum_{S \in \mathcal{D}^*} \varepsilon_S 1_{\tilde{G}} b_{R,G,j} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \cdot \left\| \sum_{R \in \mathcal{D}^*} \varepsilon_R 1_{(1+\nu)\tilde{G}} \tilde{G} \varphi_{Q,i} \langle f_Q \rangle_{Q_i} \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \right\}. \tag{12.17}$$

By lemma 12.5,

$$|1_{\tilde{G}(S)} b_{R,G,j}| \lesssim 1_{\tilde{G}(S)} \leq 1_{\Delta_{S_j}^G} \leq 1_{S_j}$$

pointwise  $\mu$ -almost everywhere. Also,  $(1 + \nu)\tilde{G}(R) \setminus \tilde{G}(R) \subset 5G(R) \subset Q_i$  and

$$(1 + \nu)\tilde{G}(R) \setminus \tilde{G}(R) \subset G_\nu(R) \subset \delta^\nu(k), \quad Q = Q(R) \in \mathcal{D}_k.$$

It follows that  $|1_{(1+\nu)\tilde{G}} \tilde{G} \varphi_{Q,i}| \lesssim 1_{\delta^\nu(k)} 1_{Q_i}$   $\mu$ -almost everywhere if  $Q \in \mathcal{D}_k$ . Hence, by indexing the second summation in the right hand side of (12.17) in terms of  $\mathcal{D}$ , the argument proceeds as above. We omit the details.  $\square$

**Estimate for  $\eta$ -boundary terms.** Here we focus on a summation involving the  $\eta$ -boundary terms

$$M_2(R) + M_4(R) + \alpha_2(R) + \alpha_3(R),$$

see (11.8) and (12.3). Observe that although both  $\alpha_2(R)$  and  $\alpha_3(R)$  depend on the random dyadic system  $\mathcal{D}^*$ , the estimate below are uniform over all such systems.

LEMMA 12.18. *Suppose that  $s \in [2, \infty)$  is such that both  $X$  and  $X^*$  have cotype  $s$ . Let  $t > (s \vee q) \vee p$  be a positive real number. Then*

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} (M_2(R) + M_4(R) + \alpha_2(R) + \alpha_3(R)) \langle f_Q \rangle_{Q_i} \right| \\ & \lesssim C(r) \eta^{1/t} \|T\|_{\mathcal{L}(L^p(\mu; X))} \|g\|_{L^q(\mu; X^*)} \|f\|_{L^p(\mu; X)} \end{aligned}$$

for every  $f \in L^p(\mathbb{R}^N, \mu; X)$  and  $g \in L^q(\mathbb{R}^N, \mu; X^*)$ .

*Proof.* By (11.8) and (12.3),

$$\begin{aligned} M_2(R) + \alpha_2(R) &= \langle 1_{R_{j,\partial}} \psi_{R,j}, T(1_{Q_i} \varphi_{Q,i}) \rangle - \langle 1_{\Delta_{R_j}^{\partial}} \psi_{R,j}, T(1_{\Delta_{Q_i}^g} \varphi_{Q,i}) \rangle; \\ M_4(R) + \alpha_3(R) &= \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{Q_{i,\partial}} \varphi_{Q,i}) \rangle - \langle 1_{\Delta_{R_j}} \psi_{R,j}, T(1_{\Delta_{Q_i}^{\partial}} \varphi_{Q,i}) \rangle. \end{aligned}$$

Observe that

$$\begin{aligned} |1_{R_{j,\partial}} \psi_{R,j}| + |1_{\Delta_{R_j}^{\partial}} \psi_{R,j}| &\lesssim 1_{R_{j,\partial}}, & |1_{Q_i} \varphi_{Q,i}| + |1_{\Delta_{Q_i}^g} \varphi_{Q,i}| &\lesssim 1_{Q_i}; \\ |1_{\Delta_{R_j}} \psi_{R,j}| &\lesssim 1_{R_j}, & |1_{Q_{i,\partial}} \varphi_{Q,i}| + |1_{\Delta_{Q_i}^{\partial}} \varphi_{Q,i}| &\lesssim 1_{Q_{i,\partial}}. \end{aligned} \tag{12.19}$$

pointwise  $\mu$ -almost everywhere. By triangle inequality, it suffices to estimate the following sums: one involving terms  $m(R) \in \{M_2(R), \alpha_2(R)\}$ , and the other involving terms in  $\{M_4(R), \alpha_3(R)\}$ . We focus on the first sum; the second one is estimated in an analogous manner.

Randomizing, using Hölder’s inequality, extracting the operator norm of  $T$ , and finally applying the contraction principle with (12.19) results in the estimate

$$\begin{aligned} \mathbf{E}_{\mathcal{D}} \left| \sum_{R \in \mathcal{D}'} \langle g_R \rangle_{R_j} m(R) \langle f_Q \rangle_{Q_i} \right| &\lesssim \mathbf{E}_{\mathcal{D}} \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{S_{j,\partial}} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \\ &\quad \cdot \|T\|_{\mathcal{L}(L^p(\mu; X))} \left\| \sum_{R \in \mathcal{D}'} \varepsilon_R 1_{Q_i} \langle f_Q \rangle_{Q_i} \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \end{aligned} \tag{12.20}$$

Indexing the summation in terms of  $\mathcal{D}$  and using the contraction principle, we see that the last factor is bounded  $\|f\|_{L^p(\mu; X)}$ . For the first factor in the right hand side of (12.20), we write

$$\delta^\eta(k) = \bigcup_{m=k-r-1}^{k-1} \bigcup_{Q \in \mathcal{D}_m} \delta_Q^\eta.$$

By (11.1), we have

$$1_{S_{j,\partial}} \leq 1_{S_j} 1_{\delta_{Q_i}^\eta} \leq 1_{S_j} 1_{\delta^\eta(k)}, \quad \text{if } Q = Q(S), S \in \mathcal{D}'_k.$$

Fix  $x \in \mathbb{R}^N$ . The random variables  $\rho_k := 1_{\delta^\eta(k)}(x)$  as functions of  $\beta \in (\{0, 1\}^N)^\mathbb{Z}$ ,  $\mathcal{D} = \mathcal{D}(\beta)$ , belong to  $L^t(\{0, 1\}^N)^\mathbb{Z}$  and they satisfy

$$\sup_{k \in \mathbb{Z}} \|1_{\delta^\eta(k)}(x)\|_{L^t(\{0, 1\}^N)^\mathbb{Z}} = \sup_{k \in \mathbb{Z}} \mathbf{P}_\beta(1_{\delta^\eta(k)}(x) = 1)^{1/t} \lesssim C(r) \eta^{1/t}.$$

Hence, proceeding as in the proof of Lemma 12.15, we find that

$$\mathbf{E}_{\mathcal{D}} \left\| \sum_{S \in \mathcal{D}'} \varepsilon_S 1_{S_{j,0}} \langle g_S \rangle_{S_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \lesssim C(r) \eta^{1/t} \left\| \sum_{R \in \mathcal{D}'} \varepsilon_R 1_{R_j} \langle g_R \rangle_{R_j} \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)}. \tag{12.21}$$

Noticing that the last term is bounded by a constant multiple of  $C(r) \eta^{1/t} \|g\|_{L^q(\mu; X^*)}$  finishes the proof.  $\square$

REMARK 12.22. It seems to us that using two independent dyadic systems  $\mathcal{D}$  and  $\mathcal{D}'$  is necessary for the step (12.21).

**13. Synthesis.** The proof of Theorem 1.7 will be completed. This involves choosing appropriate values for the auxiliary parameters  $r, \eta, \nu$ . Hence, any dependence on these numbers will be indicated explicitly.

*Proof of Theorem 1.7* Let us fix  $f \in L^p(\mu; X)$  and  $g \in L^q(\mu; X^*)$  such that

$$\|T\|_{\mathcal{L}(L^p(\mu; X))} \leq 2|\langle g, Tf \rangle|, \quad \|f\|_p = 1 = \|g\|_q.$$

Taking expectations over estimate (5.1) gives us

$$|\langle g, Tf \rangle| \lesssim \|g\|_q \|f\|_p + \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{Q \in \mathcal{D}, R \in \mathcal{D}'} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right|.$$

Because  $X$  is a UMD function lattice, its dual  $X^*$  is also a UMD function lattice. Hence, by symmetry, it suffices to consider the summation over dyadic cubes  $Q$  and  $R$  for which  $\ell(Q) \leq \ell(R)$ .

We decompose this series further as follows:

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq \ell(R)}} = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} + \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-bad}} \\ \ell(Q) \leq \ell(R)}}. \tag{13.1}$$

Observe that this decomposition to good and bad parts depends on  $\mathcal{D}' = \mathcal{D}(\beta')$ .

Let us first focus on the good summation in the right hand side of (13.1). We denote  $Q \sim R$  if these cubes satisfy (11.1), that is, if they are close to each other both in position and size. Then we have the decomposition

$$\begin{aligned} \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} &= \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \sim R}} + \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R \\ \ell(Q) < 2^{-r} \ell(R)}} \\ &+ \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \not\sim R \\ \ell(Q) < 2^{-r} \ell(R)}} + \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ 2^{-r} \ell(R) \leq \ell(Q) \leq \ell(R) \\ \ell(Q) \leq \text{dist}(Q, R)}}. \end{aligned} \tag{13.2}$$

Let us consider the third double series on the right hand side further. Assume that  $R \in \mathcal{D}'$  and  $Q \in \mathcal{D}_{R\text{-good}}$  satisfies  $Q \not\sim R$  and  $\ell(Q) < 2^{-r} \ell(R)$ . Remark 2.6 implies that

$\text{dist}(Q, R) = \text{dist}(Q, \partial R) > \ell(Q)^\gamma \ell(R)^{1-\gamma} \geq \ell(Q)$ . As a consequence, we can write

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \not\subset R \\ \ell(Q) < 2^{-r} \ell(R)}} = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) < 2^{-r} \ell(R) \\ \ell(Q) \leq \text{dist}(Q, R)}} .$$

Hence, by combining 3rd and 4th term on the right hand side of (13.2), we obtain the identity

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \sim R}} + \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ Q \subset R \\ \ell(Q) < 2^{-r} \ell(R)}} + \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)}} .$$

Invoking Propositions 7.1, 9.1, and 11.2 we are able to estimate all of the summands above, and we reach the estimate

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \\ & \leq C(r, \eta, \nu) + (C(r, \eta) \nu^{1/t} + C(r) \eta^{1/t}) \|T\|_{\mathcal{L}(L^p(\mu; X))}. \end{aligned}$$

Then we concentrate on the remaining bad summation in the right hand side of (13.1). By randomizing, using Hölder’s inequality, and using Theorem 4.1 with the identity  $\|g\|_{L^q(X^*)} = 1$ , we get

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-bad}} \\ \ell(Q) \leq \ell(R)}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \\ & = \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{k=0}^{\infty} \sum_{j \in \mathbb{Z}} \sum_{R \in \mathcal{D}'_j} \sum_{Q \in \mathcal{D}_{j-k}^{(j-(j-k)-1)\text{-bad}}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \tag{13.3} \\ & = \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{k=0}^{\infty} \int_{\Omega} \sum_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} \varepsilon_j \varepsilon_i \sum_{R \in \mathcal{D}'_j} \sum_{Q \in \mathcal{D}_{i-k}^{(k-1)\text{-bad}}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle d\mathbf{P}(\varepsilon) \right| \\ & \lesssim \sum_{k=0}^{\infty} \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left\| T \left( \sum_{i \in \mathbb{Z}} \varepsilon_i \sum_{Q \in \mathcal{D}_{i-k}^{(k-1)\text{-bad}}} D_Q^{a,1} f \right) \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \end{aligned}$$

In order to estimate this series, we fix  $k \geq 0$ . Extracting the operator norm, we see that the  $k$ ’th summand is bounded by

$$\|T\|_{\mathcal{L}(L^p(X))} \cdot \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left\| \sum_{i \in \mathbb{Z}} \varepsilon_i \lambda_{\text{bad},i}^k D_{i-k}^{a,1} f \right\|_{L^p(\mathbf{P} \otimes \mu; X)},$$

where we have denoted

$$\lambda_{\text{bad},i}^k := \sum_{Q \in \mathcal{D}_{i-k}^{(k-1)\text{-bad}}} 1_Q \in L^1(\mathbb{R}^N, \mu; \mathbb{R}).$$

Fix  $t > (s \vee p) \vee q$ , where  $s$  is such that both  $X$  and  $X^*$  have cotype  $s \in [2, \infty)$ . Using Proposition 2.4, we get the estimate

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left\| \sum_{i \in \mathbb{Z}} \varepsilon_i \lambda_{\text{bad},i}^k D_{i-k}^{a,1} f \right\|_{L^p(\mathbf{P} \otimes \mu, X)} \\ & \leq \mathbf{E}_{\mathcal{D}} \left( \int_{\mathbb{R}^N} \left[ \mathbf{E}_{\mathcal{D}'} \left\| \sum_{i \in \mathbb{Z}} \varepsilon_i \lambda_{\text{bad},i}^k(x) D_{i-k}^{a,1} f(x) \right\|_{L^p(\mathbf{P}; X)}^t \right]^{p/t} d\mu(x) \right)^{1/p} \\ & \lesssim \sup_{i,x} \|\lambda_{\text{bad},i}^k(x)\|_{L^t(\mathbf{P}_{\beta^i}; \mathbb{R})} \left\| \sum_{i \in \mathbb{Z}} \varepsilon_i D_{i-k}^{a,1} f \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \end{aligned}$$

Note that, by using Theorem 4.1, we have the estimate

$$\left\| \sum_{i \in \mathbb{Z}} \varepsilon_i D_{i-k}^{a,1} f \right\|_{L^p(\mathbf{P} \otimes \mu; X)} \lesssim 1.$$

On the other hand, if  $x \in \mathbb{R}^N$ , then by Lemma 2.7 we have

$$\begin{aligned} & \sup_i \|\lambda_{\text{bad},i}^k(x)\|_{L^t(\mathbf{P}_{\beta^i}; \mathbb{R})} \\ & = \sup_i \{ \mathbf{P}_{\beta^i}[x \in Q \in \mathcal{D}_{i-k} \text{ and } Q \text{ is } (k-1)\text{-bad}(\gamma, r)]^{1/t} \} \lesssim 2^{-(r \vee (k-1))\gamma/t}, \end{aligned}$$

All in all, we have established the following estimate

$$\begin{aligned} & \mathbf{E}_{\mathcal{D}} \mathbf{E}_{\mathcal{D}'} \left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-bad}} \\ \ell(Q) \leq \ell(R)}} \langle D_R^{a,2} g, T(D_Q^{a,1} f) \rangle \right| \lesssim \|T\|_{\mathcal{L}(L^p(\mu; X))} \sum_{k=0}^{\infty} 2^{-(r \vee (k-1))\gamma/t} \\ & \lesssim r 2^{-r\gamma/t} \|T\|_{\mathcal{L}(L^p(\mu; X))} = \delta(r) \|T\|_{\mathcal{L}(L^p(\mu; X))}. \end{aligned}$$

Here  $\delta(r) \rightarrow 0$  as  $r \rightarrow \infty$ .

Collecting the estimates above, we find that

$$\|T\|_{\mathcal{L}(L^p(\mu; X))} \leq C(r, \eta, \nu) + (C\delta(r) + C(r)\eta^{1/t} + C(r, \eta)\nu^{1/t}) \|T\|_{\mathcal{L}(L^p(\mu; X))}. \tag{13.4}$$

Next we choose  $r$  so large hat  $C\delta(r) < 1/4$ . Then we choose  $\eta > 0$  so small that  $C(r)\eta^{1/t} < 1/4$ . Lastly we choose  $\nu > 0$  so small that  $C(r, \eta)\nu^{1/t} < 1/4$ . This results in the desired estimate

$$\|T\|_{\mathcal{L}(L^p(\mu; X))} \leq C(r, \eta, \nu) + \frac{3}{4} \|T\|_{\mathcal{L}(L^p(\mu; X))}.$$

Indeed, it follows that  $\|T\|_{\mathcal{L}(L^p(\mu; X))} \leq 4C(r, \eta, \nu)$ . □

**14. Operator-valued kernels.** In this section we explain the proof of Theorem 1.8. This proof is a straightforward modification of the proof of Theorem 1.7.

We define a *d-dimensional Rademacher–Calderón–Zygmund kernel* as a function  $K(x, y)$  of variables  $x, y \in \mathbb{R}^N$  with  $x \neq y$  and taking values in  $\mathcal{L}(X)$ , which satisfies

$$\begin{aligned} \mathcal{R}(\{|x - y|^d K(x, y) : x, y \in \mathbb{R}^N, x \neq y\}) &\leq 1; \\ \mathcal{R}\left(\left\{\frac{|x - y|^{d+\alpha}}{|x - x'|^\alpha} [K(x, y) - K(x', y)], \frac{|x - y|^{d+\alpha}}{|x - x'|^\alpha} [K(y, x) - K(y, x')] \right. \right. & \quad (14.1) \\ \left. \left. : x, x', y \in \mathbb{R}^N, 0 < |x - x'| \leq |x - y|/2 \right\}\right) &\leq 1 \end{aligned}$$

for some  $\alpha > 0$ . Recall that  $\mathcal{R}(T)$  designates the Rademacher-bound of an operator family  $T \subset \mathcal{L}(X)$ , as defined after (2.2).

Let  $T : f \mapsto Tf$  be a linear operator acting on some functions  $f : \mathbb{R}^N \rightarrow X$  or  $f : \mathbb{R}^N \rightarrow \mathbb{C}$ , producing new functions  $Tf : \mathbb{R}^N \rightarrow X$  in the former case and  $Tf : \mathbb{R}^N \rightarrow \mathcal{L}(X)$  in the latter. If  $\xi \in X$  and  $F : \mathbb{R}^N \rightarrow \mathbb{C}$  or  $F : \mathbb{R}^N \rightarrow \mathcal{L}(X)$ , define the function  $F \otimes \xi$  by  $(F \otimes \xi)(x) := F(x)\xi$ , where the last expression is the product of a scalar and a vector, or the action of an operator on a vector, respectively. With this notation, suppose that  $T(\varphi \otimes \xi) = (T\varphi) \otimes \xi$  for  $\varphi : \mathbb{R}^N \rightarrow \mathbb{C}$  and  $\xi \in X$ . The adjoint  $T^*$  is defined via duality  $\langle g, f \rangle = \int_{\mathbb{R}^N} \langle g(x), f(x) \rangle d\mu(x)$  between functions  $f : \mathbb{R}^N \rightarrow X$  and  $g : \mathbb{R}^N \rightarrow X^*$ : for  $\varphi, \psi : \mathbb{R}^N \rightarrow \mathbb{C}$ ,  $\xi \in X$  and  $\xi^* \in X^*$ ,

$$\xi^*(\langle \psi, T\varphi \rangle \xi) = \langle \psi \otimes \xi^*, T(\varphi \otimes \xi) \rangle := \langle T^*(\psi \otimes \xi^*), \varphi \otimes \xi \rangle =: (\langle T^* \psi, \varphi \rangle \xi^*)(\xi),$$

and hence  $\langle T^* \psi, \varphi \rangle = (\langle \psi, T\varphi \rangle)^* \in \mathcal{L}(X^*)$  for scalar-valued functions  $\varphi, \psi$ .

Such a  $T$  is called an  *$\mathcal{L}(X)$ -valued Rademacher–Calderón–Zygmund operator* with kernel  $K$  if

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

for points  $x \in \mathbb{R}^N$  outside the support of  $f$ .

We are ready to explain the modifications in the proof of Theorem 1.7. These occur in sections 5–13 and, roughly speaking, they are as follows: one repeats the proof, and the assumed  $\mathcal{R}$ -boundedness conditions ensure that whenever one “pulled out” bounded scalar coefficients from the randomised series, which persist throughout the arguments, the same can be done with the operator coefficients by the very definition (2.2) of  $\mathcal{R}$ -boundedness. Some technicalities arise when treating the paraproducts in Section 10. Our goal is to provide a comprehensive treatment and, at the same time, avoid repeating arguments. To accomplish this task, we have chosen to explain the modifications in sections 6, 7 and 10.

**Operator-valued decoupling estimates.** Let us begin with Section 6. Instead of scalars satisfying (6.5), we now consider the following  $\mathcal{R}$ -bounded families of operators:

$$\mathcal{R}(\{\lambda_{RQ} T_{RQ} \in \mathcal{L}(X) : R \in \mathcal{D}', Q \in \mathcal{D}_{R\text{-good}}, \ell(Q) \leq \ell(R)\}) \lesssim 1, \quad (14.2)$$

where  $T_{RQ} \in \mathcal{L}(X)$  and the scalar coefficients are

$$\lambda_{RQ} := \frac{D(Q, R)^{d+\alpha}}{\mu(R_u)\mu(Q_v)\ell(Q)^{\alpha/2}\ell(R)^{\alpha/2}}.$$

These  $\mathcal{R}$ -bounded families occur in the following counterpart of Lemma 6.6.

LEMMA 14.3. *Assume that  $E_{k-1}f_k = f_k$  and  $E_{k-1}g_k = g_k$ , where  $f_k$  and  $g_k$ ,  $k \in \mathbb{Z}$ , are as in Lemma 6.6. Assume also that the estimate (14.2) holds. Then*

$$\left| \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}_{R\text{-good}} \\ \ell(Q) \leq \ell(R)}} \langle g_R \rangle_{R_u} T_{RQ} \langle f_Q \rangle_{Q_v} \right| \lesssim \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k g_k \right\|_{L^q(\mathbf{P} \otimes \mu; X^*)} \left\| \sum_{k=-\infty}^{\infty} \varepsilon_k f_k \right\|_{L^p(\mathbf{P} \otimes \mu; X)}. \tag{14.4}$$

The proof of this lemma proceeds as the proof of Lemma 6.6 with appropriate modifications. The key fact is that the operators

$$\tilde{t}_{RQ} = 2^{(n+j)\alpha/4} \mu(S) \frac{T_{RQ}}{\mu(R_u)\mu(Q_v)} \in \mathcal{L}(X),$$

where the parameters are clear from the context, belong to an  $\mathcal{R}$ -bounded family. This follows from the normalizations and the condition (14.2).

Then we can proceed to Section 7, where the goal is to prove a counterpart of Proposition 7.1 under the assumptions of Theorem 1.8. For this purpose, we need the following counterpart of Lemma 7.9.

LEMMA 14.5. *Suppose that for every pair of cubes  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$ , satisfying  $\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)$ , we are given functions  $\varphi_Q, \psi_R \in L^1(\mathbb{R}^N, \mu; \mathbb{C})$  such that  $\text{supp}(\varphi_Q) \subset Q$ ,  $\text{supp}(\psi_R) \subset R$ , and*

$$\int \varphi_Q d\mu = 0.$$

Then

$$\mathcal{R}(\{\sigma_{RQ} \langle \psi_R, T\varphi_Q \rangle \in \mathcal{L}(X) : \ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)\}) \leq 1,$$

where the normalizing factors are given by

$$\sigma_{RQ} := \frac{\text{dist}(Q, R)^{d+\alpha}}{\ell(Q)^\alpha \|\varphi_Q\|_{L^1(\mu)} \|\psi_R\|_{L^1(\mu)}}.$$

*Proof.* Suppose that  $Q \in \mathcal{D}$  and  $R \in \mathcal{D}'$  satisfy  $\ell(Q) \leq \ell(R) \wedge \text{dist}(Q, R)$ . Let  $y_Q$  be the center of the cube  $Q$ . Denoting

$$F(x, y) := \frac{|y - y_Q|^\alpha}{|x - y_Q|^{d+\alpha}} \varphi_Q(y) \psi_R(x) \sigma_{RQ},$$

we obtain

$$\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |F(x, y)| d\mu(y) d\mu(x) \leq 1.$$

Hence, by denoting

$$\mathcal{T} = \left\{ \frac{|x - y|^{d+\alpha}}{|y - y'|^\alpha} [K(x, y) - K(x, y')] : y, y', x \in \mathbb{R}^N, |y - y'| \leq |x - y|/2 \right\},$$

we obtain

$$\begin{aligned} \sigma_{RQ}(\psi_R, T\varphi_Q) &= \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} K(x, y)\varphi_Q(y)\psi_R(x)\sigma_{RQ} d\mu(y) d\mu(x) \\ &= \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} [K(x, y) - K(x, y_Q)]\varphi_Q(y)\psi_R(x)\sigma_{RQ} d\mu(y) d\mu(x) \\ &= \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \frac{|x - y_Q|^{d+\alpha}}{|y - y_Q|^\alpha} [K(x, y) - K(x, y_Q)] F(x, y) d\mu(y) d\mu(x) \\ &\in \overline{\text{abs conv}}(\mathcal{T}). \end{aligned}$$

Here the closure is taken in the strong operator topology and the absolute convex hull, denoted by  $\text{abs conv}(\mathcal{T})$ , is the set of all vectors of the form  $\sum_{j=1}^k \lambda_j x_j$  with  $\sum_{j=1}^k |\lambda_j| \leq 1$  and  $x_j \in \mathcal{T}$  for  $j = 1, 2, \dots, k$ . Since,

$$\mathcal{R}(\overline{\text{abs conv}}(\mathcal{T})) = \mathcal{R}(\mathcal{T}),$$

it remains to use the second  $\mathcal{R}$ -boundedness estimate in (14.1) □

Proceeding as in the proof of Proposition 7.1, and using Lemma 14.5 instead of Lemma 7.9, we find that the  $\mathcal{R}$ -boundedness estimate (14.2) holds for the family of operators in  $\mathcal{L}(X)$  defined by the equation (7.12). Hence, after applying Lemma 14.3 instead of Lemma 6.6, the proof of Proposition 7.1 continues as before.

**Operator-valued paraproducts.** We proceed to Section 10. Let us first indicate the modifications in the proof of the estimate (10.6), the boundedness of the paraproduct. The first one comes in the proof of Lemma 10.13: Theorem 4.16 and Proposition 4.18 are used with UMD function lattice  $Z$  instead of  $\mathbb{C}$ .

The step from (10.16) to (10.17) is now established by the following lemma and assumption  $\mathcal{R}(\bar{B}_Z) \lesssim 1$ .

LEMMA 14.6. *Suppose that  $t > q \vee s$ , where  $X^*$  has cotype  $s$ . Then*

$$\begin{aligned} &\left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* d_j E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N \times \Omega; X^*)} \\ &\lesssim \mathcal{R}(\bar{B}_Z) \cdot \|\{ \|d_j(\cdot)\|_{L^t(\Omega; Z)}\}_{j \in \mathbb{Z}}\|_{\text{Car}^t(\mathcal{D}')} \cdot \|g\|_{L^q(\mathbb{R}^N; X^*)}, \end{aligned} \tag{14.7}$$

where  $\varepsilon^* = \{\varepsilon_j^* : j \in \mathbb{Z}\} \in \Omega^*$  are Rademacher random variables and

$$d_j : \mathbb{R}^N \rightarrow L^t(\Omega; Z) : x \mapsto \left( \varepsilon \mapsto \sum_{R \in \mathcal{D}'_j} \sum_{\substack{Q \in \mathcal{D} \\ S(Q)=R}} \varepsilon_Q \pi_{Q, R^a}(x) \right)$$

for functions  $\pi_{Q, R^a} : \mathbb{R}^N \rightarrow Z$  that are determined by (10.11).



*Proof.* Note first that LHS(14.7) can be written as

$$\left( \iiint_{\Omega^* \times \mathbb{R}^N \times \Omega} \left| \sum_{j \in \mathbb{Z}} \varepsilon_j^* \frac{d_j(x, \varepsilon)}{|d_j(x, \varepsilon)|_Z} |d_j(x, \varepsilon)|_Z E_j g(x) \right|_{X^*}^q d\mathbf{P}(\varepsilon) d\mu(x) d\mathbf{P}(\varepsilon^*) \right)^{1/q}.$$

Using Fubini’s theorem and the fact that the closed unit ball of  $Z$  is  $\mathcal{R}$ -bounded, we see that LHS(14.7) can be bounded by a constant multiple of

$$\mathcal{R}(\bar{B}_Z) \left( \iiint_{\Omega \times \mathbb{R}^N \times \Omega^*} \left| \sum_{j \in \mathbb{Z}} \varepsilon_j^* |d_j(x, \varepsilon)|_Z E_j g(x) \right|_{X^*}^q d\mathbf{P}(\varepsilon^*) d\mu(x) d\mathbf{P}(\varepsilon) \right)^{1/q}.$$

Recall that  $t > q$ . Using Fubini’s theorem, followed by the Hölder’s inequality, we find that LHS(14.7) is bounded by a constant multiple of

$$\begin{aligned} & \mathcal{R}(\bar{B}_Z) \left( \iint_{\Omega^* \times \mathbb{R}^N} \left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* |d_j(x)|_Z E_j g(x) \right\|_{L^t(\Omega; X^*)}^q d\mu(x) d\mathbf{P}(\varepsilon^*) \right)^{1/q} \\ &= \mathcal{R}(\bar{B}_Z) \left\| \sum_{j \in \mathbb{Z}} \varepsilon_j^* |d_j(\cdot)|_Z E_j g \right\|_{L^q(\Omega^* \times \mathbb{R}^N; L^t(\Omega; X^*))}. \end{aligned}$$

Let us denote  $\tilde{d}_j(x, \varepsilon) := |d_j(x, \varepsilon)|_Z$ . Then, for a fixed  $x \in \mathbb{R}^N$ ,

$$\|\tilde{d}_j(x)\|_{L^t(\Omega; \mathbb{C})} = \left( \int_{\Omega} |\tilde{d}_j(x, \varepsilon)|^t d\mathbf{P}(\varepsilon) \right)^{1/t} = \|d_j(x)\|_{L^t(\Omega; Z)}.$$

Hence, by using Lemma 10.8, we can conclude that the estimate (14.7) holds. □

In order to estimate the right hand side of (10.18), we use the fact that  $L^2(\Omega, Z)$  has cotype 2 since  $Z$  has it. The described modifications suffice for obtaining estimate (10.6) in the context of Theorem 1.8. Finally, in the proof of estimate (10.7) we use Theorem 4.16, with UMD function lattice  $Z$ , and the fact that the family  $\{E_k\}_{k \in \mathbb{Z}}$  of operators in  $L^q(\mu; Z)$  is  $\mathcal{R}$ -bounded by the UMD-valued Stein’s inequality [2].

This concludes the description of modifications in Section 10.

ACKNOWLEDGEMENTS. This paper benefited from the interaction with a simultaneously on-going (but earlier completed) project of the first author with Henri Martikainen; see [6]. We would like to thank him for this fruitful exchange of ideas. The authors would like to thank the referee for useful remarks.

### REFERENCES

1. J. Bourgain, Some remarks on Banach spaces in which martingale difference sequences are unconditional, *Ark. Mat.* **21**(2) (1983), 163–168.
2. J. Bourgain, Vector-valued singular integrals and the  $H^1$ -BMO duality, in *Probability theory and harmonic analysis (Cleveland, Ohio, 1983)*, volume 98 of *Monogr. Textbooks Pure Appl. Math.* (Dekker, New York, 1986), 1–19.
3. D. L. Burkholder, A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions, in *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., Wadsworth, Belmont, CA (1983, 270–286.

4. T. P. Hytönen, The vector-valued non-homogeneous  $Tb$  Theorem, *Int. Math. Res. Not.* (online, 2012) doi: 10.1093/imrn/rns222.
5. T. Hytönen and M. Kemppainen, On the relation of Carleson's embedding and the maximal theorem in the context of Banach space geometry, *Math. Scand.* **109**(2), 269–284 (2011).
6. T. Hytönen and H. Martikainen, Non-homogeneous  $Tb$  theorem and random dyadic cubes on metric measure spaces, *J. Geom. Anal.* **22**(4) (2012), 1071–1107.
7. T. Hytönen, A. McIntosh and P. Portal, Kato's square root problem in Banach spaces, *J. Funct. Anal.* **254**(3) (2008), 675–726.
8. M. Kemppainen, On the Rademacher maximal function, *Studia Math.* **203**(1) (2011) 1–31.
9. B. Maurey, Système de Haar, in *Séminaire Maurey-Schwartz 1974–1975: Espaces  $L^p$ , applications radonifiantes et géométrie des espaces de Banach, Exp. Nos. I et II* (erratum, p. 1). Centre Math. (École Polytech., Paris, 1975), 26.
10. S. Mayboroda and A. Volberg, Boundedness of the square function and rectifiability, *C. R. Math. Acad. Sci. Paris* **347**(17–18) (2009), 1051–1056.
11. T. R. McConnell, Decoupling and stochastic integration in UMD Banach spaces, *Probab. Math. Statist.* **10**(2) (1989), 283–295.
12. F. Nazarov, S. Treil and A. Volberg, The  $Tb$ -theorem on non-homogeneous spaces, *Acta Math.* **190**(2) (2003), 151–239.
13. F. Nazarov, S. Treil and A. Volberg, Accretive system  $Tb$ -theorems on nonhomogeneous spaces, *Duke Math. J.* **113**(2) (2002) 259–312.
14. José L. Rubio de Francia, Martingale and integral transforms of Banach space valued functions, in *Probability and Banach spaces (Zaragoza, 1985)*, volume 1221 of *Lecture Notes in Math.*, (Springer, Berlin, 1986) 195–222.
15. Elias M. Stein, *Topics in harmonic analysis related to the Littlewood-Paley theory*, Annals of Mathematics Studies, No. 63 (Princeton University Press, Princeton, N.J., 1970).
16. A. Volberg, The proof of non-homogeneous  $T1$  theorem via averaging of dyadic shifts. Preprint, arXiv:1303.0367, 2013.
17. L. Weis, Operator-valued Fourier multiplier theorems and maximal  $L_p$ -regularity, *Math. Ann.* **319**(4) (2001) 735–758.