Proceedings of the Edinburgh Mathematical Society (2016) **59**, 193–222
DOI:10.1017/S0013091514000340

ON THE LOCAL Tb THEOREM: A DIRECT PROOF UNDER THE DUALITY ASSUMPTION

MICHAEL T. LACEY[1] AND ANTTI V. VÄHÄKANGAS[2]

[1]*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA* (lacey@math.gatech.edu)

[2]*Department of Mathematics and Statistics, PO Box 68, 00014, University of Helsinki, Finland* (antti.vahakangas@helsinki.fi)

(Received 7 September 2012)

Abstract We give a new direct proof of the local Tb theorem in the Euclidean setting and under the assumption of dual exponents. This theorem provides a flexible framework for proving the boundedness of a Calderón–Zygmund operator, supposing the existence of systems of local accretive functions. We assume that the integrability exponents on these systems of functions are of the form $1/p + 1/q \leqslant 1$, the 'dual case' $1/p + 1/q = 1$ being the most difficult one. Our proof is direct: it avoids a reduction to the perfect dyadic case unlike some previous approaches. The principal point of interest is in the use of random grids and the corresponding construction of the corona. We also use certain twisted martingale transform inequalities.

Keywords: local Tb theorem; $T1$ theorem; corona; twisted martingale transform; stopping cubes

2010 *Mathematics subject classification:* Primary 42B20
Secondary 42B25; 42B35

1. Introduction

Our subject is the local Tb theorem in the classical Euclidean setting. There are many results under this topic, all of which extend the David–Journé $T1$ theorem [**7**] and the Tb theorem of Christ [**6**] by giving flexible conditions under which an operator T with a Calderón–Zygmund kernel extends to a bounded linear operator on L^2; the lectures of Hofmann [**10**] indicate the range of interest in these types of results. By 'local' we understand that the Tb conditions involve a family of test functions b_Q, one for each cube Q, that should satisfy a non-degeneracy condition on its 'own' Q. Furthermore, both b_Q and Tb_Q are subject to normalized integrability conditions on Q. Symmetric assumptions are imposed on T^*.

The goal of this paper is to give a new direct proof of a known local Tb theorem, Theorem 1.2. This theorem applies, in particular, when the integrability conditions imposed in the hypotheses are those in duality, namely, $1/p_1 + 1/p_2 = 1$. Our argument is direct in the sense that it avoids a reduction to the so-called perfect dyadic case such as that seen in Auscher and Yang [**2**]. A companion paper [**18**] addresses a perfect dyadic variant of Theorem 1.2 for the full range $1 < p_1, p_2 < \infty$; it contains many of the features of the argument in the present paper, with significantly fewer technicalities.

We say that T is a *Calderón–Zygmund operator* if it is a bounded linear operator on $L^2(\mathbb{R}^n)$ with the following representation. For every $f \in L^2(\mathbb{R}^n)$,

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y)\,\mathrm{d}y, \quad x \notin \operatorname{supp}(f),$$

where the kernel $K\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is assumed to satisfy the following estimates for some $\eta > 0$:

$$|K(x,y)| \leqslant |x-y|^{-n}, \quad x \neq y; \tag{1.1}$$

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leqslant \frac{|x-x'|^\eta}{|x-y|^{n+\eta}}, \quad |x-x'| < \tfrac{1}{2}|x-y|. \tag{1.2}$$

We define $\boldsymbol{T}$ to be the norm of T as an operator on $L^2(\mathbb{R}^n)$.

Definition 1.1. Fix $1 < p < \infty$. A collection of functions $\{b_Q\colon Q \subset \mathbb{R}^n \text{ is a cube}\}$ is called *a system of p-accretive functions with constant $\boldsymbol{A} > 1$* if the following conditions hold for each cube Q:

(1) b_Q is supported on Q and $\int_Q b_Q(x)\,\mathrm{d}x = |Q|$;

(2) $\|b_Q\|_p \leqslant \boldsymbol{A}|Q|^{1/p}$.

We aim to prove the following local Tb theorem. Define $p' = p/(p-1)$.

Theorem 1.2. *Fix $1 < p_1, p_2 < \infty$ so that $1/p_1 + 1/p_2 \leqslant 1$. Suppose that T is a Calderón–Zygmund operator for which there are systems $\{b_Q^j\}$ of p_j-accretive functions, $j \in \{1,2\}$, with a constant $\boldsymbol{A}$, satisfying the following testing condition: there is a constant $\boldsymbol{T}_{\mathrm{loc}}$ such that for all cubes Q,*

$$\int_Q |Tb_Q^1|^{p_2'} \leqslant \boldsymbol{T}_{\mathrm{loc}}^{p_2'}|Q|, \qquad \int_Q |T^* b_Q^2|^{p_1'} \leqslant \boldsymbol{T}_{\mathrm{loc}}^{p_1'}|Q|.$$

We then have a quantitative estimate $\boldsymbol{T} \lesssim_{n,\eta,p_1,p_2,\boldsymbol{A}} 1 + \boldsymbol{T}_{\mathrm{loc}}$ for the operator norm of T.

In the case of perfect dyadic operators, the full range $1 < p_1, p_2 < \infty$ of exponents is allowed, as was shown in [**3**, p. 48]. It was also hoped that the result could be lifted to the continuous case. This lifting turned out to be a difficult problem: some of the direct methods [**9**,**13**] to attack it require assumptions that are stronger than the duality assumption. Theorem 1.2 is due to Auscher and Yang [**2**] who provide an indirect argument: a reduction to the perfect dyadic case. The Auscher and Yang paper does not reach the difficult case $1/p_1 + 1/p_2 > 1$, which is also known as 'Hofmann's problem' as it was emphasized by Hofmann in [**10**]. This problem was partially solved by Auscher and Routin [**1**] via the adaptation of the Beylkin–Coifman–Rokhlin (BCR) algorithm (see [**4**, **8**]) as well as the martingale transform inequalities; at the same time, Auscher and Routin obtained a direct proof of Theorem 1.2. An essentially full solution to Hofmann's problem has very recently been obtained by Hytönen and Nazarov [**14**]. By applying perturbation techniques for both the operator and the accretive functions, they obtained a variant of Theorem 1.2 for $1 < p_1, p_2 < \infty$.

Our main contribution is an alternate direct proof of Theorem 1.2. It is desirable to have such proofs from the viewpoint of the extension of the argument to other settings. As an example, in the literature [**13**,**15**,**21**] on the local Tb theorem in the non-homogeneous setting [**21**] one encounters stronger $L^\infty(\mathbb{R}^n)$ (or BMO) conditions on the Tb_Qs, as well as on test functions b_Q. Some of the techniques in the present paper have been subsequently applied in order to relax these conditions in the case of square functions [**16**]. It even seems plausible that a variant of Theorem 1.2 could be recovered in the non-homogeneous setting; see [**17**].

Outline of the proof

Let us turn to a discussion of the proof technique. As is quite common, absorbtion parameters enter into the proof at several stages, permitting us to resort to the assumed finite, but non-quantitative, norm bound on T, provided that it is multiplied by a small absorption parameter. We use the well-known non-homogeneous techniques of [**22**], in particular, the powerful technique of 'good cubes'. In the local Tb setting, there is however a delicate problem with the typical method of restricting to the good cubes, as is pointed out by Hytönen–Martikainen [**13**, Remark 4.1]. An important innovation of the present paper is the corona construction, which enables us to restrict to good cubes in a natural way. This construction depends on two random dyadic grids, $\mathcal{D}^1$ and $\mathcal{D}^2$, that are defined on independent probability spaces Ω^j, $j = 1, 2$. A cube $Q \in \mathcal{D}^1$ is called *bad* if it is close to the boundary of some significantly larger cube in the other grid, $\mathcal{D}^2$. The badness of Q is an event in Ω^2 with probability that can be made arbitrarily small, giving rise to an absorption parameter. A cube Q is *good* if it is not bad.

Let us describe the corona construction in three steps. First, by a $T1$ theorem [**7**] it suffices to consider the bilinear form $\langle T\tilde{f}_1, \tilde{f}_2\rangle$, where $|\tilde{f}_1| = |\tilde{f}_2| = \mathbf{1}_{Q^0}$ for a fixed cube Q^0. One projects $\tilde{f}_1$ onto the good cubes, calling the result f_1, which can be viewed as a function of Ω^1 and Ω^2. This also contributes an error term that is small in all L^p spaces on average and is treated by the first of several absorption arguments. One then makes a standard selection of stopping cubes $\tilde{\mathcal{S}}^j \subset \mathcal{D}^j$ and local testing functions b_S^j for $S \in \mathcal{S}^j$. The stopping cubes $\tilde{\mathcal{S}}^j$ constitute a sparse collection; in particular, $\tilde{\mathcal{S}}^j$ is a Carleson sequence of cubes.

In the next step, we construct functions β_S^1 by projecting b_S^1 *away from* those bad cubes that themselves have S as a parent in $\tilde{\mathcal{S}}^1$. By doing so, we gain the following desirable feature: the twisted martingale difference of f_1, with respect to β_S^1 and over a bad cube Q with $\tilde{\mathcal{S}}^1$ parent S, will typically be 0. On the downside, β_S^1 is now a function of Ω^1 and Ω^2 and the original collection of stopping cubes $\tilde{\mathcal{S}}^1$ is not so well adapted to the β_S^1. On the other hand, favourably to us, β_S^1 can be viewed as a small perturbation of b_S^1.

In the last step, to adopt the usage of perturbed functions β_S^1 in twisted martingale differences, one cannot run the stopping cube selection process again due to the unacceptable dependencies on Ω^1 and Ω^2. Instead one invokes absorbtion, arguing that one can *truncate* the stopping tree $\tilde{\mathcal{S}}^1$ inside a set B^1 that is small on average. The corona construction is now described and its details take up § 2, which is almost half the length of this paper.

There are also tools in § 3 that are useful, namely martingale transform inequalities for twisted martingale differences, and the associated half-twisted inequalities that are *universal* in that they hold in all L^q-spaces. These inequalities also play a crucial role in [**1**, Lemma 5.3] and in [**18**].

Turning to the remaining part of the argument, one is in a familiar situation [**22**] in the sense that only *good* cubes $P \in \mathcal{D}^1$ and $Q \in \mathcal{D}^2$ need to be considered. The double sum over P, Q is reduced, by symmetry, to the case in which $\ell P \geqslant \ell Q$ and this sum is further decomposed into subcases according to the position and size of Q relative to P. The case of Q deeply inside P admits a direct control by using the twisted martingale transform inequalities; this 'inside' case incorporates the paraproduct term. For experts we remark that we do not appeal to Carleson measure arguments at any stage of the argument; in this we follow [**1**, **19**, **20**]. The case of P and Q having the same approximate size and position requires new perturbation inequalities for the twisted martingale transforms. This 'diagonal' case is the hardest one in many existing arguments, including ours. A potentially troublesome case is when $Q \subset 3P \setminus P$ and Q is substantially smaller than P; however, due to goodness, Q is still relatively far from the boundary of P. We address this 'nearby' case by exploiting the smoothness condition on the kernel K, and the universal half-twisted inequalities. The remaining 'far' case depends upon standard off-diagonal estimates for singular integrals and universal martingale transform inequalities.

Notation

For a cube Q, $\langle f\rangle_Q := |Q|^{-1}\int_Q f\,\mathrm{d}x$ and $\ell Q = |Q|^{1/n}$ is the side length of the cube. $A \lesssim B$ means that $A \leqslant CB$, where C is an unspecified constant which need not be tracked. The distances in $\mathbb{R}^n$ are measured in terms of the supremum norm, $|x| = \|x\|_\infty$ for $x \in \mathbb{R}^n$. Given $Q \in \mathcal{D}^j$, we denote by $\mathrm{ch}(Q)$ the 2^n dyadic children of Q. Given $\mathcal{S} \subset \mathcal{D}^j$, we write $\mathrm{ch}_{\mathcal{S}}(S)$ for the $\mathcal{S}$-children of $S \in \mathcal{S}$: these are the maximal elements S' of $\mathcal{S}$ that are strictly contained in S. For a cube $Q \in \mathcal{D}^j$ that is contained in a cube in $\mathcal{S}$, we take $\pi_{\mathcal{S}} Q$ to be the $\mathcal{S}$-parent of Q: this is the minimal element of $\mathcal{S}$ that contains Q.

2. The corona

It is a straightforward consequence of the $T1$ theorem [**7**] that

$$\boldsymbol{T} \lesssim 1 + \sup_{Q\subset\mathbb{R}^n \text{ cube}} |Q|^{-1}\|\mathbf{1}_Q T^*\mathbf{1}_Q\|_{L^1} + \sup_{Q\subset\mathbb{R}^n \text{ cube}} |Q|^{-1}\|\mathbf{1}_Q T\mathbf{1}_Q\|_{L^1}.$$

Without loss of generality, we can assume that the last term dominates. Fix a cube Q^0 for which

$$\boldsymbol{T}|Q^0| \lesssim \|\mathbf{1}_{Q^0} T\mathbf{1}_{Q^0}\|_{L^1}. \tag{2.1}$$

For notational convenience, let us take two functions $\tilde{f}_1$, $\tilde{f}_2$ such that

$$|\tilde{f}_1| = |\tilde{f}_2| = \mathbf{1}_{Q^0} \quad \text{and} \quad \|\mathbf{1}_{Q^0} T\mathbf{1}_{Q^0}\|_{L^1} = \langle T\tilde{f}_1, \tilde{f}_2\rangle.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

The main purpose of the present section is to devise a corona-type decomposition, which helps us to restrict to good cubes, after which it will be straightforward to complete the proof of the following lemma.

Lemma 2.1. *Fix $0 < \upsilon_0 < 1$. There are functions f_1, f_2 and a constant $C > 0$, independent of both $\boldsymbol{T}$ and $\boldsymbol{T}_{\mathrm{loc}}$, such that the following inequalities hold:*

$$\|\tilde{f}_j - f_j\|_2 < \upsilon_0 |Q^0|^{1/2}, \quad j = 1, 2; \tag{2.2}$$

$$|\langle T f_1, f_2 \rangle| < \{C(1 + \boldsymbol{T}_{\mathrm{loc}}) + \upsilon_0 \boldsymbol{T}\} |Q^0|. \tag{2.3}$$

This lemma and an absorption argument complete the proof of Theorem 1.2. The construction of the corona is rather complicated. It will be highly dependent upon certain random constructions and there will be several absorption parameters that lead to the constant υ_0. The main advantage of our corona construction is that it allows us to restrict to the good cubes in a natural manner; this and other useful features admit a straightforward proof of inequality (2.3).

2.1. Random grids

We make use of so-called random grids, due to Nazarov *et al.* [**21**]. These turned out to be of fundamental importance; see, for example, [**11**, **12**, **20**, **24**].

We will have a random grid $\mathcal{D}^1$ for the functions $\tilde{f}_1$, f_1 and a random grid $\mathcal{D}^2$ for the functions $\tilde{f}_2$, f_2. These random grids are constructed as follows. Let $\mathcal{D}^0$ be the standard dyadic grid in $\mathbb{R}^n$. For a fixed cube $\hat{Q} \in \mathcal{D}^0$, let us consider the translated cube

$$Q := \hat{Q} \dot{+} \omega^1 := \hat{Q} + \sum_{j:\, 2^{-j} < \ell Q} 2^{-j} \omega^1_j,$$

which is a function of $\omega^1 \in \Omega^1 := (\{0,1\}^n)^{\mathbb{Z}}$. Define $\mathcal{D}^1 = \{\hat{Q} \dot{+} \omega^1 : \hat{Q} \in \mathcal{D}^0\}$. The natural uniform probability measure $\mathbb{P}^1$ is placed upon Ω^1. That is, each component ω^1_j, $j \in \mathbb{Z}$, has an equal probability 2^{-n} of taking any of the 2^n values and all the components are independent of each other. The expectation with respect to $\mathbb{P}^1$ is denoted by $\mathbb{E}^1$. Define Ω^2 in the same manner, with an independent copy of Ω^1. It will be important to distinguish between these two copies, so we write $\omega^j \in \Omega^j$ for the elements of the probability space that define $\mathcal{D}^j$. The product $\mathbb{P}^1 \otimes \mathbb{P}^2$ is denoted by $\mathbb{P}$ and the corresponding expectation $\mathbb{E}^1 \mathbb{E}^2$ is denoted by $\mathbb{E}$.

We need notation. Define the familiar [**11**, **13**, **21**] and convenient number

$$\epsilon := \frac{\eta}{2(\eta + n)}. \tag{2.4}$$

Throughout, $r \geqslant 3/\epsilon$ should be thought of as a large integer, which satisfies condition (3) below and whose exact value is assigned later. We say that a cube $Q \in \mathcal{D}^1$ is *bad* if there is $P \in \mathcal{D}^2$ such that $\ell(P) \geqslant 2^r \ell(Q)$ and $\mathrm{dist}(Q, \partial P) \leqslant (\ell Q)^{\epsilon} (\ell P)^{1-\epsilon}$. Otherwise, Q is *good*. The definitions for $Q \in \mathcal{D}^2$ are similar. The following properties are well known for a cube $Q \in \mathcal{D}^1$:

(1) the goodness/badness of Q is a random variable on Ω^2;

(2) the probability $\pi_{\text{good}} := \mathbb{P}^2(Q \text{ is good})$ is independent of Q;

(3) $\pi_{\text{bad}} := 1 - \pi_{\text{good}} \lesssim 2^{-\epsilon r}$, provided that ϵr is sufficiently large.

Define the good and bad projections by $I = P^j_{\text{good}} + P^j_{\text{bad}}$, where

$$P^j_{\text{good}}\phi := \sum_{Q \in \mathcal{D}^j :\, Q \text{ is good}} D_Q\phi, \quad j = 1, 2.$$

Here $D_Q\phi = \sum_{Q' \in \text{ch}(Q)} \{\langle\phi\rangle_{Q'} - \langle\phi\rangle_Q\}\mathbf{1}_{Q'}$ is the usual martingale difference associated with Q.

We have the following proposition on the bad projections; the constant $0 < c_q < 1$ that appears in the exponent on the right will be a function of p_1 and p_2. In the following, we suppress this dependence in notation, writing only $2^{-c\epsilon r}$.

Proposition 2.2. *If $1 < q < \infty$ and $\{j, k\} = \{1, 2\}$, then there is a constant $c_q > 0$ such that*

$$\mathbb{E}^k \|P^j_{\text{bad}}\phi\|^q_q \lesssim 2^{-c_q\epsilon r}\|\phi\|^q_q. \tag{2.5}$$

Here $\omega^j \in \Omega^j$ is fixed and $\phi \in L^q$ is any function that is independent of sequences $\omega^k \in \Omega^k$.

Proof. The basic idea is to apply the Marcinkiewicz interpolation theorem to the linear operator $P^j_{\text{bad}} : L^q(\mathrm{d}x) \to L^q(\mathbb{P}^k \otimes \mathrm{d}x)$. The projection to bad cubes is a martingale transform [**5**], and hence the following inequality with no decay holds:

$$\mathbb{E}^k \|P^j_{\text{bad}}\phi\|^p_p \leqslant \sup\{\|P^j_{\text{bad}}\phi\|^p_p : \omega^k \in \Omega^k\} \lesssim \|\phi\|^p_p, \quad 1 < p < \infty.$$

Thus, it suffices to verify the claimed decay for $q = 2$. To this end, by independence,

$$\mathbb{E}^k \|P^j_{\text{bad}}\phi\|^2_2 = \mathbb{E}^k \sum_{\substack{Q \in \mathcal{D}^j \\ Q \text{ is bad}}} \|D_Q\phi\|^2_2 = \pi_{\text{bad}} \sum_{Q \in \mathcal{D}^j} \|D_Q\phi\|^2_2 = \pi_{\text{bad}}\|\phi\|^2_2.$$

Indeed, both $\mathcal{D}^j$ and $\|D_Q\phi\|^2_2$ for $Q \in \mathcal{D}^j$ are independent of $\omega^k \in \Omega^k$ and the badness of $Q \in \mathcal{D}^j$ is a random variable on Ω^k, $\{j, k\} = \{1, 2\}$. □

2.2. Selection of f_j

We will prove Lemma 2.1 by averaging over random grids. Fix $j \in \{1, 2\}$. Let $\mathcal{A}^j_*$ denote all (at most 2^n) cubes $Q \in \mathcal{D}^j$ such that $Q \cap Q^0 \neq \emptyset$ and $\ell Q^0 \leqslant \ell Q < 2\ell Q^0$. Let $\mathcal{A}^j$ be all cubes in $\mathcal{D}^j$ that are contained in some $Q \in \mathcal{A}^j_*$. Recall that the function $\tilde{f}_j$ is chosen in connection with (2.1) and that it is equal to $\mathbf{1}_{Q^0}$ in absolute value. We define an approximate f_j of this function to be

$$f_j := \sum_{Q \in \mathcal{A}^j_*} \langle\tilde{f}_j\rangle_Q \mathbf{1}_Q + \sum_{\substack{Q \in \mathcal{A}^j \\ Q \text{ is good}}} D_Q\tilde{f}_j.$$

In view of Proposition 2.2, we have

$$\mathbb{E}\|\tilde{f}_j - f_j\|_2^2 \lesssim 2^{-c\epsilon r}|Q^0|. \tag{2.6}$$

Hence, it suffices to estimate $\mathbb{E}|\langle Tf_1, f_2\rangle|$.

The functions f_j lie in BMO: a dyadic variant associated with the grid $\mathcal{D}^j$. It follows from the associated John–Nirenberg inequality that

$$\|f_j\|_q \lesssim |Q^0|^{1/q}, \quad 1 < q < \infty, \tag{2.7}$$

with the implied constant independent of sequences ω^1 and ω^2. The fact that the functions f_j can nevertheless be unbounded creates a minor set of difficulties for us.

2.3. The setup for stopping cubes construction

In order to accommodate the reduction to good cubes, we will need a significant modification of the usual selection process of stopping trees and local b functions. The following definition will help to explain the end result that we are after; it is convenient to denote $T^1 = T$ and $T^2 = T^*$.

Definition 2.3. Fix constants $0 < \tau$, $\delta < 1$ and let $\{j, k\} = \{1, 2\}$. A collection of integrable functions $\{\beta_S^j\colon S \in \mathcal{S}^j \subset \mathcal{D}^j\}$ is a *stopping datum* (a *perturbed stopping datum*) for a collection $\mathcal{G}^j \subset \mathcal{D}^j$ of cubes if the following conditions hold with $\boldsymbol{A}_j = \frac{1}{2}$, $\boldsymbol{B}_j = \delta^{-1}\boldsymbol{A}^{p_j}$, $\boldsymbol{C}_j = \delta^{-1}\boldsymbol{T}_{\mathrm{loc}}^{p_k'}$ (in the case of perturbed stopping data: $\boldsymbol{A}_j = \frac{1}{4}$, $\boldsymbol{B}_j \lesssim \delta^{-1}\boldsymbol{A}^{p_j}$ and $\boldsymbol{C}_j \lesssim \delta^{-1}\boldsymbol{T}_{\mathrm{loc}}^{p_k'} + \upsilon_1^{p_k'}\boldsymbol{T}^{p_k'}$ for some constant $0 < \upsilon_1 < 1$).

(1) Every $Q \in \mathcal{G}^j$ is contained in some $S \in \mathcal{S}^j$. The same holds for every child $Q' \in \mathrm{ch}(Q)$, whose parent $\pi_{\mathcal{S}^j}Q'$ need not equal $\pi_{\mathcal{S}^j}Q$, even if Q is a minimal cube in $\mathcal{G}^j$.

(2) If $Q \in \mathcal{G}^j$ with $\pi_{\mathcal{S}^j}Q = S$ (or $Q \in \mathrm{ch}(R)$ with $R \in \mathcal{G}^j$ and $\pi_{\mathcal{S}^j}Q = S$), then

 (a) $\langle \beta_S^j\rangle_Q \geqslant \boldsymbol{A}_j$ (do not divide by zero),

 (b) $\langle |M\beta_S^j|^{p_j}\rangle_Q \leqslant \boldsymbol{B}_j$ (local norm of $M\beta_S^j$ controlled) and

 (c) $\langle |T^j\beta_S^j|^{p_k'}\rangle_Q \leqslant \boldsymbol{C}_j$ (local norm of $T^j\beta_S^j$ is controlled).

(3) $\sum_{S'\in \mathrm{ch}_{\mathcal{S}^j}(S)}|S'| \leqslant \tau|S|$ for all $S \in \mathcal{S}^j$, i.e. $\mathcal{S}^j$ is a sparse collection of cubes.

For $Q \in \mathcal{G}^j$ and $\phi \in L^1_{\mathrm{loc}}$, we define a *twisted martingale difference* by

$$\Delta_Q^{\beta^j}\phi := \sum_{Q'\in\mathrm{ch}(Q)} \left\{ \frac{\langle\phi\rangle_{Q'}}{\langle \beta^j_{\pi_{\mathcal{S}^j}Q'}\rangle_{Q'}}\beta^j_{\pi_{\mathcal{S}^j}Q'} - \frac{\langle\phi\rangle_{Q}}{\langle \beta^j_{\pi_{\mathcal{S}^j}Q}\rangle_{Q}}\beta^j_{\pi_{\mathcal{S}^j}Q} \right\}\mathbf{1}_{Q'}. \tag{2.8}$$

This is well defined, as Q has an $\mathcal{S}^j$ parent, and there is no division by zero; see conditions (1) and (2) (a). We also define a *half-twisted martingale difference* by

$$\tilde{D}_Q^{\beta^j}\phi := \left\{ \sum_{\substack{Q'\in\mathrm{ch}(Q)\\ \pi_{\mathcal{S}^j}Q = \pi_{\mathcal{S}^j}Q'}} \frac{\langle\phi\rangle_{Q'}}{\langle \beta^j_{\pi_{\mathcal{S}^j}Q'}\rangle_{Q'}}\mathbf{1}_{Q'} \right\} - \frac{\langle\phi\rangle_{Q}}{\langle \beta^j_{\pi_{\mathcal{S}^j}Q}\rangle_{Q}}\mathbf{1}_{Q}. \tag{2.9}$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Observe that here we do not multiply by a β^j function and that the sum over the children excludes those with a different $\mathcal{S}^j$ parent (in particular, there is no change in the β^j function: $\pi_{\mathcal{S}^j} Q' = \pi_{\mathcal{S}^j} Q$).

The following lemma provides the reduction to good cubes. In particular, it helps us to eliminate the martingale differences that are associated with bad cubes.

Lemma 2.4. *Suppose that $\Lambda > 1$ and $0 < \upsilon_1 < 4^{-1-n}$. Fix $j \in \{1, 2\}$. There is a collection $\mathcal{G}^j \subset \mathcal{D}^j$ of cubes, and a perturbed stopping datum $\{\beta^j_S \colon S \in \mathcal{S}^j\}$ for $\mathcal{G}^j$, so that the following four conditions hold.*

(1) *Every cube $Q \in \mathcal{G}^j$ is good.*

(2) *For all $Q \in \mathcal{G}^j$, we have $\langle |f_j| \rangle_Q \leqslant \Lambda$.*

(3) *Suppose that $Q \in \mathcal{G}^j$ with a child Q' and that $S \in \mathcal{S}^j$ with $\pi_{\mathcal{S}^j} Q \subset S$. Define a constant $\lambda_{Q'}$ by*

$$\lambda_{Q'} \mathbf{1}_{Q'} := \mathbf{1}_{Q'} \sum_{\substack{P \in \mathcal{G}^j \colon P \supset Q \\ \pi_{\mathcal{S}^j} P = S}} \tilde{D}_P^{\beta^j} f_j. \tag{2.10}$$

We then have $|\lambda_{Q'}| \lesssim \Lambda$.

(4) *Assuming $\Lambda^{-1} + \Lambda \upsilon_1^{-1} 2^{-c\epsilon r} < 1$, there holds*

$$\mathbb{E} \Bigg| \langle T f_1, f_2 \rangle - \sum_{P \in \mathcal{G}^1} \sum_{Q \in \mathcal{G}^2} \langle T \Delta_P^{\beta^1} f_1, \Delta_Q^{\beta^2} f_2 \rangle \Bigg| \\ \leqslant C_1 \{1 + \boldsymbol{T}_{\mathrm{loc}} + (\upsilon_1 + \Lambda^{-1} + \Lambda \upsilon_1^{-1} 2^{-c\epsilon r}) \boldsymbol{T}\} |Q^0|. \tag{2.11}$$

Here, $C_1 = C_1(p_1, p_2, n, \boldsymbol{A})$ does not depend upon the absorption parameters υ_1, Λ and r.

Before the lengthy proof of this lemma, let us indicate its usage.

A conditional proof of Lemma 2.1. In order to complete the proof of Lemma 2.1, it remains to verify Lemma 2.4 and the following inequality:

$$\Bigg| \sum_{P \in \mathcal{G}^1} \sum_{Q \in \mathcal{G}^2} \langle T \Delta_P^{\beta^1} f_1, \Delta_Q^{\beta^2} f_2 \rangle \Bigg| \leqslant \{C_2 \{1 + \boldsymbol{T}_{\mathrm{loc}}\} + C_3 r \upsilon_1 \Lambda^2 \boldsymbol{T}\} |Q^0|. \tag{2.12}$$

We emphasize that inequality (2.12) is uniform in ω^1 and ω^2 and that it is distinct from (2.11). The constant $C_3 = C_3(p_1, p_2, n, \eta, \boldsymbol{A})$, which is independent of absorption parameters, and the product $r \upsilon_1 \Lambda^2$ of absorption parameters appear on the right. The constant

$$C_2 = C_2(p_1, p_2, n, \eta, \boldsymbol{A}, r, \Lambda, \upsilon_1)$$

is allowed to depend also upon the absorption parameters. Returning to the proof of Lemma 2.1, let us consider (2.6), (2.11) and (2.12). By taking $\Lambda > 1$ sufficiently large and then choosing r large enough and assigning $\upsilon_1 = r^{-2}$, the proof is complete; apart from Lemma 2.4 and inequality (2.12). □

At this stage, let us make several clarifying remarks.

Remark 2.5. Hytönen and Martikainen [**13**, Remark 4.1] have pointed to serious concerns with some existing approaches to the reduction to good cubes in local Tb theorems. The substance of the problem arises from the fact that the twisted martingale differences depend upon the choice of grid and the collection of local b functions, making averaging arguments, such as the one used in the proof of Proposition 2.2, not transparently true. Our corona construction establishes a transparent reduction to good cubes in (2.12) and this is one of our main contributions.

Remark 2.6. The proof of inequality (2.12), taken up in §§ 4 and 5, is now largely standard in nature, following the lines of [**22**, **25**] and including innovations from [**19**, **20**] to avoid auxiliary Carleson measure estimates. However, certain perturbation inequalities are needed when treating cubes that are nearby, both in size and position. There are also advantages for us.

(1) We need only consider good cubes, which is the primary goal of the corona construction.

(2) By normalizing both f_1 and f_2 with a factor Λ^{-1}, the sums (2.10) are bounded by $c \lesssim 1$, which is related to the telescoping property needed in the control of paraproduct terms. This normalization, allowing us to set $\Lambda = 1$, is assumed while proving inequality (2.12) in the beginning of § 3 and thereafter.

Remark 2.7. The dependence of the quantitative estimates on the parameters aside from $\boldsymbol{T}$ and $\boldsymbol{T}_{\mathrm{loc}}$ is not straightforward and typically we do not track it. However, we need to track the dependence of a constant c on absorption parameters r, Λ and υ_1 if it appears in an expression $c \cdot \boldsymbol{T}$.

The rest of this section is taken up with the proof of Lemma 2.4.

2.4. Auxiliary stopping data

Fix $j \in \{1, 2\}$. We construct auxiliary *stopping data* $\{b_S^j : S \in \tilde{\mathcal{S}}^j\}$ for the collection $\mathcal{A}^j$, which was defined when selecting the function f_j. The *perturbed stopping data* will be later constructed by using this auxiliary stopping data. The following construction of $\tilde{\mathcal{S}}^j$ and $\{b_S^j : S \in \tilde{\mathcal{S}}^j\}$ is fairly standard and it only depends upon ω^j.

Initialize $\tilde{\mathcal{S}}^j$ to be $\mathcal{A}_*^j$. For each cube S in this collection, consider the function b_S^j given to us by the local Tb hypothesis (see the formulation of Theorem 1.2). Add to $\tilde{\mathcal{S}}^j$ the maximal dyadic descendants $Q \subset S$, which either fail any of the criteria (a)–(c) in Definition 2.3 with $\beta_S^j := b_S^j$, or fail the condition

$$\inf_{x \in Q} M|b_S^j|^{p_j}(x) \leqslant \delta^{-1} \boldsymbol{A}^{p_j}. \tag{2.13}$$

Concerning these stopping conditions, let E_S be the union of the maximal descendants Q of S such that $\langle b_S^j \rangle_Q < \frac{1}{2}$. We have, using the higher integrability of b_S^j,

$$|S| = \int_S b_S^j \,\mathrm{d}x = \int_{E_S} b_S^j \,\mathrm{d}x + \int_{S \setminus E_S} b_S^j \,\mathrm{d}x \leqslant \tfrac{1}{2}|S| + \boldsymbol{A}|S \setminus E_S|^{1/p_j'} |S|^{1/p_j}.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Hence, $(2\boldsymbol{A})^{-p_j'}|S| \leqslant |S \setminus E_S|$. Next, let us consider the union F_S of the maximal descendants Q of S, failing (2.13) or one of the mentioned criteria (b), (c). By inspection, we have $|F_S| \lesssim \delta|S|$. Therefore, with a choice of $\delta = \delta(p_j, n, \boldsymbol{A})$ and $\tau = \tau(p_j, n, \boldsymbol{A})$, we can continue the construction of $\tilde{\mathcal{S}}^j$ inductively to meet conditions (1)–(3) in Definition 2.3.

Below, we will refer to $\tilde{\mathcal{S}}^j$ and its subsets as collections of *stopping cubes*.

2.5. Perturbation of the b functions

In a departure from standard arguments, we modify the functions b_S^j, $S \in \tilde{\mathcal{S}}^j$, that are already selected. For $S \in \tilde{\mathcal{S}}^j$, we define

$$\beta_S^j := b_S^j - \tilde{\beta}_S^j, \quad \text{where } \tilde{\beta}_S^j := \sum_{\substack{Q \in \mathcal{A}^j :\, \pi_{\tilde{\mathcal{S}}^j} Q = S \\ Q \text{ is bad}}} D_Q b_S^j. \tag{2.14}$$

Note that the sum defining $\tilde{\beta}_S^j$ is formed by using the classical martingale differences that are associated with bad cubes in $\mathcal{A}^j$ that have the same stopping parent. Particular care must be taken with these perturbations β_S^j as they are now functions of both ω^1 and ω^2.

Nevertheless, $\tilde{\beta}_S^j$ is a small function on average.

Lemma 2.8. *For $\{j,k\} = \{1,2\}$ and all $S \in \tilde{\mathcal{S}}^j$, $\|\tilde{\beta}_S^j\|_{\mathrm{BMO}} \lesssim 1$ holds and, moreover,*

$$\mathbb{E}^k \|\tilde{\beta}_S^j\|_q^q \lesssim 2^{-c\epsilon r}|S|, \quad 1 < q < \infty.$$

Proof. Let $Q \in \mathcal{D}^j$ be such that $\pi_{\tilde{\mathcal{S}}^j} Q = S$. Writing $\epsilon_{Q'} := \mathbf{1}_{Q' \subset Q} \mathbf{1}_{\pi_{\tilde{\mathcal{S}}^j} Q' = S} \mathbf{1}_{Q' \text{ is bad}}$, we obtain

$$\begin{aligned}\left(\int_Q |\tilde{\beta}_S^j - \langle \tilde{\beta}_S^j \rangle_Q|^{p_j} \, \mathrm{d}x \right)^{1/p_j} &= \Big\| \sum_{Q' \subset Q} D_{Q'} \tilde{\beta}_S^j \Big\|_{p_j} = \Big\| \sum_{Q' \in \mathcal{D}^j} \epsilon_{Q'} D_{Q'} (\mathbf{1}_Q b_S^j) \Big\|_{p_j} \\ &\lesssim \|\mathbf{1}_Q b_S^j\|_{p_j} \leqslant \langle |M b_S^j|^{p_1} \rangle_Q^{1/p_j} |Q|^{1/p_j} \\ &\lesssim |Q|^{1/p_j}.\end{aligned}$$

Here, we have appealed to the boundedness of martingale transforms and the stopping rules. The remaining cases either reduce to this or are trivial. Hence, the BMO assertion is true.

Concerning the L^q estimate, we apply Proposition 2.2 and the John–Nirenberg inequality:

$$\begin{aligned}\mathbb{E}^k \|\tilde{\beta}_S^j\|_q^q &= \mathbb{E}^k \Big\| P_{\mathrm{bad}}^j \Big[\sum_{Q :\, \pi_{\tilde{\mathcal{S}}^j} Q = S} D_Q b_S^j \Big] \Big\|_q^q \\ &\lesssim 2^{-c\epsilon r} \Big\| \sum_{Q :\, \pi_{\tilde{\mathcal{S}}^j} Q = S} D_Q b_S^j \Big\|_q^q \\ &\lesssim 2^{-c\epsilon r} |S| \Big\| \sum_{Q :\, \pi_{\tilde{\mathcal{S}}^j} Q = S} D_Q b_S^j \Big\|_{\mathrm{BMO}}^q.\end{aligned}$$

By arguing as above, we finish the proof. □

2.6. Truncation of the stopping tree

We will use the functions β^j_S as the basis of perturbed stopping data (see Lemma 2.4) but the path to this is not yet clear for the following reasons:

(A) the functions β^1_S are not necessarily suitable for forming twisted martingale differences;

(B) even if defined, the twisted martingale differences associated with bad cubes need not vanish; and

(C) the functions f_j are unbounded.

A truncation of the stopping tree will address all three of these issues.

Concerning point (B), there is a simple sufficient condition for a twisted martingale difference to be identically 0.

Proposition 2.9. *Assume that $Q \in \mathcal{A}^j$ is bad and that no child of Q is in $\tilde{\mathcal{S}}^j$. Suppose that $\langle \beta^j_S \rangle_Q \neq 0$, where $S = \pi_{\tilde{\mathcal{S}}^j} Q$. Then both $\Delta^{\beta^j}_Q f_j$ and $\tilde{D}^{\beta^j}_Q f_j$ are well defined using $\tilde{\mathcal{S}}^j$ in parent selectors for β^j functions, and $\Delta^{\beta^j}_Q f_j \equiv 0 \equiv \tilde{D}^{\beta^j}_Q f_j$.*

Proof. By assumptions and definitions, the averages of f_j and β^j_S do not change in moving from cube Q to a child of Q. By inspection of (2.8) and (2.9), the ratios in the definition of either martingale difference of f_j are all well defined and equal, and hence they cancel. □

The previous considerations lead to the following three types of undesirable cubes $Q \in \mathcal{A}^j$, where $\Lambda > 1$ and $0 < \upsilon_1 < 4^{-1-n}$ are absorption parameters and $\{j,k\} = \{1,2\}$.

Type A: $\{\langle |M\tilde{\beta}^j_{\pi_{\tilde{\mathcal{S}}^j}Q}|^{p_j}\rangle_Q \geqslant \upsilon_1^{p_j}$ or $\langle |T^j\tilde{\beta}^j_{\pi_{\tilde{\mathcal{S}}^j}Q}|^{p'_k}\rangle_Q \geqslant \upsilon_1^{p'_k}\boldsymbol{T}^{p'_k}\}$ or Q has a child $S \in \tilde{\mathcal{S}}^j$ such that $\{\langle |M\tilde{\beta}^j_S|^{p_j}\rangle_S \geqslant \upsilon_1^{p_j}$ or $\langle |T^j\tilde{\beta}^j_S|^{p'_k}\rangle_S \geqslant \upsilon_1^{p'_k}\boldsymbol{T}^{p'_k}\}$.

Type B: Q is not of Type A and Q has a child in $\tilde{\mathcal{S}}^j$ and Q is bad.

Type C: Q is neither of Type A nor of Type B and $\langle |f_j|\rangle_Q > \Lambda$.

Each of these three types depends upon both ω^1 and ω^2. Let $\mathcal{B}^{j,\alpha}$ be the collection of maximal cubes in $\mathcal{A}^j$ of Type α, $\alpha = A, B, C$, and let $\mathcal{B}^j$ be the maximal cubes in the union of these three collections. Define $B^{j,\alpha} := \bigcup\{Q\colon Q \in \mathcal{B}^{j,\alpha}\}$ and $B^j := B^{j,A} \cup B^{j,B} \cup B^{j,C}$.

Let us verify that the sets B^j are, on average, small in measure. Certain error terms coming from the truncation can then be later absorbed.

Lemma 2.10. *For $\{j,k\} = \{1,2\}$, we have*

$$\mathbb{E}|B^j| = \mathbb{E}^j\mathbb{E}^k|B^j| \lesssim \{\Lambda^{-2p_j} + \upsilon_1^{-p_j}2^{-c\epsilon r}\}|Q^0|.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Proof. We first prove that

$$\mathbb{E}^k|B^{j,A}| \lesssim \upsilon_1^{-p_j} 2^{-c\epsilon r}|Q^0|, \quad \text{where } \{j,k\} = \{1,2\}. \tag{2.15}$$

Recall that the collection $\tilde{\mathcal{S}}^j$ is only a function of ω^j. By sparseness, $\sum_{S\in\tilde{\mathcal{S}}^j}|S| \lesssim |Q^0|/(1-\tau) \lesssim |Q^0|$. A cube is of Type A for four potential reasons. First, fix $S \in \tilde{\mathcal{S}}^j$ and let $\mathcal{B}_S^{j,A_1}$ be the maximal cubes $Q \in \mathcal{A}^j$ with $\pi_{\tilde{\mathcal{S}}^j}Q = S$ and $\langle |M\tilde{\beta}_S^j|^{p_j}\rangle_Q \geqslant \upsilon_1^{p_j}$. By Lemma 2.8,

$$\mathbb{E}^k \sum_{Q\in\mathcal{B}_S^{j,A_1}} |Q| \lesssim \upsilon_1^{-p_j}\mathbb{E}^k \int_S |\tilde{\beta}_S^j|^{p_j} \lesssim \upsilon_1^{-p_j}2^{-c\epsilon r}|S|.$$

Second, let $\mathcal{B}_S^{j,A_2}$ be the maximal cubes $Q \in \mathcal{A}_j$ with $\pi_{\tilde{\mathcal{S}}^j}Q = S$, and $\langle |T^j\tilde{\beta}_S^j|^{p_k'}\rangle_Q \geqslant \boldsymbol{T}^{p_k'}\upsilon_1^{p_k'}$. Then, using the *a priori* norm bound $c\boldsymbol{T}$ for the operator T^j on $L^{p_k'}$ and inequality $p_k' \leqslant p_j$,

$$\mathbb{E}^k \sum_{Q\in\mathcal{B}_S^{j,A_2}} |Q| \lesssim \upsilon_1^{-p_k'}\mathbb{E}^k \int_S |\tilde{\beta}_S^j|^{p_k'} \lesssim \upsilon_1^{-p_j}2^{-c\epsilon r}|S|.$$

Third, let $\mathcal{B}^{j,A_3}$ be the collection of cubes Q in $\mathcal{A}^j$ having a child $S \in \tilde{\mathcal{S}}^j$ with $\langle |M\tilde{\beta}_S^j|^{p_j}\rangle_S \geqslant \upsilon_1^{p_j}$. Then,

$$\mathbb{E}^k \sum_{Q\in\mathcal{B}^{j,A_3}} |Q| \lesssim \upsilon_1^{-p_j} \sum_{S\in\tilde{\mathcal{S}}^j} \mathbb{E}^k \int_S |\tilde{\beta}_S^j|^{p_j} \lesssim \upsilon_1^{-p_j}2^{-c\epsilon r} \sum_{S\in\tilde{\mathcal{S}}^j}|S| \lesssim \upsilon_1^{-p_j}2^{-c\epsilon r}|Q^0|.$$

A similar estimate for the remaining collection $\mathcal{B}^{j,A_4}$ of cubes Q in $\mathcal{A}^j$, having a child $S \in \tilde{\mathcal{S}}^j$ such that $\langle |T^j\tilde{\beta}_S^j|^{p_k'}\rangle_S \geqslant \upsilon_1^{p_k'}\boldsymbol{T}^{p_k'}$, finishes the proof of inequality (2.15).

Let us then consider the set $B^{j,B}$. The collection $\tilde{\mathcal{S}}^j$ is only a function of ω^j and, holding that variable fixed, the event that $S \in \tilde{\mathcal{S}}^j$ has a bad parent is an event in Ω^k. And so,

$$\mathbb{E}^k|B^{j,B}| \leqslant 2^n\mathbb{E}^k \sum_{S\in\tilde{\mathcal{S}}^j} |S|\mathbf{1}_{\pi S \text{ is bad}} \lesssim \frac{1}{1-\tau}2^{-\epsilon r}|Q^0| \lesssim 2^{-\epsilon r}|Q^0|. \tag{2.16}$$

For the remaining set $B^{j,C}$, recall that f_j is a dyadic BMO function, uniformly over ω^1 and ω^2. More precisely, by Chebyshev's inequality and (2.7), we have

$$|B^{j,C}| = \sum_{Q\in\mathcal{B}^{j,C}} |Q| \leqslant |\{Mf_j > \Lambda\}| \leqslant \Lambda^{-2p_j}\|Mf_j\|_{2p_j}^{2p_j} \lesssim \Lambda^{-2p_j}|Q^0|. \tag{2.17}$$

The proof is completed by combining inequalities (2.15)–(2.17). □

Next we define the collection $\mathcal{G}^j$ and the perturbed stopping data for $\mathcal{G}^j$, whose existence is stated in Lemma 2.4. This is done by truncating the stopping tree $\tilde{\mathcal{S}}^j$ at $\mathcal{B}^j$.

Definition 2.11. Take $\mathcal{G}^j$ to be all *good* cubes in $\mathcal{A}^j$ that are *not contained* in any cube in $\mathcal{B}^j$. Set $\mathcal{S}^j$ to be $\tilde{\mathcal{S}}^j$ minus all cubes that are *strictly* contained in some $Q \in \mathcal{B}^j$. For convenience, we also denote by $\mathcal{R}^j \supset \mathcal{G}^j$ all cubes in $\mathcal{A}^j$, *both good and bad*, not contained in any cube in $\mathcal{B}^j$. Take the data for $\mathcal{G}^j$ to be $\{\beta_S^j : S \in \mathcal{S}^j\}$.

Let us emphasize the fact that $Q \in \mathcal{R}^j$ is not of any Type α, $\alpha = $ A, B, C. In the remaining part of this section, we will check all the assertions in Lemma 2.4.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Verification of the perturbed stopping data

First we show that $\{\beta^j_S \colon S \in \mathcal{S}^j\}$ is indeed a perturbed stopping data for $\mathcal{G}^j$, as claimed. By construction,

$$\pi_{\mathcal{S}^j} Q = \pi_{\tilde{\mathcal{S}}^j} Q, \qquad \pi_{\mathcal{S}^j} Q' = \pi_{\tilde{\mathcal{S}}^j} Q' \tag{2.18}$$

if $Q \in \mathcal{R}^j$ and $Q' \in \mathrm{ch}(Q)$. Accordingly, $\{\beta^j_S \colon S \in \mathcal{S}^j\}$ satisfies property (1) in the Definition 2.3 of perturbed stopping data. Another consequence of (2.18) is that we can compute the martingale differences $\Delta^{\beta^j}_P$ and $\tilde{D}^{\beta^j}_P$ for the case of $P \in \mathcal{R}^j$ by using freely either $\mathcal{S}^j$ or $\tilde{\mathcal{S}}^j$ in the parent selectors for β^j functions.

The sparseness property (3) is trivial for $\mathcal{S}^j$, since $\tilde{\mathcal{S}}^j$ satisfies it and $\mathcal{S}^j \subset \tilde{\mathcal{S}}^j$. The remaining properties (2) (a)–(c) of the perturbed stopping data follow from the next lemma.

Lemma 2.12. *Fix $j \in \{1, 2\}$ and a cube $S \in \mathcal{S}^j$. The following conditions then hold.*

(1) $\langle \beta^j_S \rangle_S = 1$.

(2) $\langle |\beta^j_S|^{p_j} \rangle_S \lesssim \boldsymbol{A}^{p_j}$.

(3) *Suppose that $Q \in \mathcal{R}^j$ and $\pi_{\mathcal{S}^j} Q = S$ (or Q is a child of a cube in $\mathcal{R}^j$ and $\pi_{\mathcal{S}^j} Q = S$). Then the following hold:*

   (a) $\langle \beta^j_S \rangle_Q \geqslant \frac{1}{4}$;

   (b) $\langle |M \beta^j_S|^{p_j} \rangle_Q \lesssim \delta^{-1} \boldsymbol{A}^{p_j}$;

   (c) $\langle |T^j \beta^j_S|^{p'_k} \rangle_Q \lesssim \delta^{-1} \boldsymbol{T}^{p'_k}_{\mathrm{loc}} + \upsilon_1^{p'_k} \boldsymbol{T}^{p'_k}$, *where $\{j, k\} = \{1, 2\}$.*

Proof. By Definition (2.14), $|S| = \int_S b^j_S \, \mathrm{d}x = \int_S \beta^j_S \, \mathrm{d}x$, so property (1) holds. The boundedness of martingale transforms implies property (2):

$$\int_S |\beta^j_S|^{p_j} \, \mathrm{d}x \lesssim \int_S |b^j_S|^{p_j} \, \mathrm{d}x \leqslant \boldsymbol{A}^{p_j} |S|.$$

Properties (3) (a)–(c) are a consequence of (2.18) and the failure of the condition defining Type A cubes. First let us consider property (3) (a). If $Q \in \mathcal{R}^j$ and $\pi_{\mathcal{S}^j} Q = S$, then

$$\langle \beta^j_S \rangle_Q \geqslant \langle b^j_S \rangle_Q - \langle |M \tilde{\beta}^j_S|^{p_j} \rangle_Q^{1/p_j} \geqslant \tfrac{1}{2} - \upsilon_1,$$

which is greater than $\frac{1}{4}$ (recall that *stopping data* is slightly stronger on this point). If Q is a child of a cube in $\mathcal{R}^j$ and $\pi_{\mathcal{S}^j} Q = S$, then either $Q \in \mathcal{S}^j$, in which case $\langle \beta^j_S \rangle_Q = \langle \beta^j_Q \rangle_Q = 1$, or property (3) (a) follows as above by first comparing the average of $|\tilde{\beta}^j_S|$ on Q to its average on πQ. Let us then consider (3) (b) and (3) (c) for $Q \in \mathcal{R}^j$. By sub-linearity and stopping rules,

$$\langle |M \beta^j_S|^{p_j} \rangle_Q \lesssim \langle |M b^j_S|^{p_j} \rangle_Q + \langle |M \tilde{\beta}^j_S|^{p_j} \rangle_Q \leqslant \delta^{-1} \boldsymbol{A}^{p_j} + \upsilon_1^{p_j} \lesssim \delta^{-1} \boldsymbol{A}^{p_j}.$$

Likewise,

$$\langle |T^j \beta^j_S|^{p'_k} \rangle_Q \lesssim \langle |T^j b^j_S|^{p'_k} \rangle_Q + \langle |T^j \tilde{\beta}^j_S|^{p'_k} \rangle_Q \leqslant \delta^{-1} \boldsymbol{T}^{p'_k}_{\mathrm{loc}} + \upsilon_1^{p'_k} \boldsymbol{T}^{p'_k}.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

These properties for a child Q of a cube in $\mathcal{R}^j$ follow by comparing the average on Q to that on πQ for the case in which $Q \notin \mathcal{S}^j$, and by the stopping rules for the case in which $Q \in \mathcal{S}^j$. □

Verification of conditions (1)–(3) in Lemma 2.4

Every cube $Q \in \mathcal{G}^j$ is good by definition and, by construction, $\langle |f_j| \rangle_Q \leqslant \Lambda$ (recall Type C cubes). Let us then consider property (3), concerning the sum of half-twisted differences in (2.10). For a fixed $Q \in \mathcal{G}^j$ with a child Q' and $S \in \mathcal{S}^j$ with $\pi_{\mathcal{S}^j} Q \subset S$, let us consider the constant $\lambda_{Q'}$ defined by

$$\lambda_{Q'} \mathbf{1}_{Q'} := \mathbf{1}_{Q'} \sum_{\substack{P \in \mathcal{A}^j :\, P \supset Q \\ \pi_{\mathcal{S}^j} P = S}} \tilde{D}_P^{\beta^j} f_j.$$

In contrast to the series in (2.10), the series above extends over *all cubes* with the same $\mathcal{S}^j$ parent. Nevertheless, we are not redefining $\lambda_{Q'}$. Indeed, if P is a bad cube in the series above, then $P \in \mathcal{R}^j$ and it has no stopping children in $\tilde{\mathcal{S}}^j$ due to the construction; by (2.18) and Proposition 2.9, we find that $\tilde{D}_P^{\beta^j} f_j \equiv 0$ so the two series, in fact, coincide.

Then, by inspection of (2.9), the series above on Q' is telescoping to the difference of two ratios (or to a single ratio). On the numerator of the ratios are averages of f_j, which are bounded by the definition of Type C cubes. The denominator of the ratios is an average of β_S^j, which is bounded below by $\frac{1}{4}$ because of Lemma 2.12 (3) (a). All in all, we find that $|\lambda_{Q'}| \lesssim \Lambda$.

2.7. Completion of the proof of Lemma 2.4

The proof of (2.11) remains and we need an appropriate representation formula for the f_js so that we can compute the difference in (2.11). We begin with certain preparations for the representation Lemma 2.14.

Define $\phi^j := \sum_{Q \in \mathcal{B}^j} \phi_Q^j$, where $\phi_Q^j = f_j \mathbf{1}_Q$ if $Q \in \mathcal{B}^j \cap \mathcal{A}_*^j$ and, otherwise,

$$\phi_Q^j := f_j \mathbf{1}_Q - \frac{\langle f_j \rangle_Q}{\langle \beta^j_{\pi_{\mathcal{S}^j} Q} \rangle_Q} \beta^j_{\pi_{\mathcal{S}^j} Q} \mathbf{1}_Q. \tag{2.19}$$

For the following lemma, recall that the set B^j is a function of both ω^1 and ω^2, and it is of small measure in expectation.

Lemma 2.13. *We have* $\|\phi^j\|_{p_j}^{p_j} \lesssim \Lambda^{p_j} |B^j|$ *for* $j \in \{1, 2\}$.

Proof. If $Q \in \mathcal{B}^j \cap \mathcal{A}_*^j$, then, by (2.7), $\|\phi_Q^j\|_{p_j} \lesssim |Q^0|^{1/p_j} \lesssim |B^j|^{1/p_j}$. There are at most 2^n such cubes. For the remaining terms we notice that, since f_j is in BMO and the

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

average values of f_j are controlled,

$$\begin{aligned}\Big\|\sum_{Q\in\mathcal{B}^j\setminus\mathcal{A}^j_*}\phi^j_Q\Big\|^{p_j}_{p_j} &= \sum_Q\Big\|f_j\mathbf{1}_Q-\frac{\langle f_j\rangle_Q}{\langle\beta^j_{\pi_{\mathcal{S}^j}Q}\rangle_Q}\beta^j_{\pi_{\mathcal{S}^j}Q}\mathbf{1}_Q\Big\|^{p_j}_{p_j}\\ &\lesssim\sum_Q\Big\{\|f_j\mathbf{1}_Q-\langle f_j\rangle_Q\mathbf{1}_Q\|^{p_j}_{p_j}+\Big\|\langle f_j\rangle_Q\mathbf{1}_Q-\frac{\langle f_j\rangle_Q}{\langle\beta^j_{\pi_{\mathcal{S}^j}Q}\rangle_Q}\beta^j_{\pi_{\mathcal{S}^j}Q}\mathbf{1}_Q\Big\|^{p_j}_{p_j}\Big\}\\ &\lesssim\Lambda^{p_j}\sum_{Q\in\mathcal{B}^j\setminus\mathcal{A}^j_*}|Q|\leqslant\Lambda^{p_j}|B^j|.\end{aligned}\tag{2.20}$$

We used the definition of Type C cubes and Lemma 2.12, along with the observation that the parent of Q is in $\mathcal{R}^j$ if $Q\in\mathcal{B}^j\setminus\mathcal{A}^j_*$. □

Concerning the representation of f_j, we have the following lemma.

Lemma 2.14. *Fix $j\in\{1,2\}$. Then the following equality holds both pointwise almost everywhere and in L^{p_j}:*

$$f_j=\sum_{Q\in\mathcal{A}^j_*\setminus\mathcal{B}^j}\langle f_j\rangle_Q\beta^j_Q+\sum_{Q\in\mathcal{G}^j}\Delta^{\beta^j}_Q f_j+\phi^j.\tag{2.21}$$

Proof. Let Q be any bad cube, which is not contained in a cube in $\mathcal{B}^j$. By construction and Proposition 2.9, $\Delta^{\beta^j}_Q f_j\equiv 0$. It follows that for any $x\in B^j$, the sum above is in fact finite, and telescoping. By inspection, it is equal to $f_j(x)$.

Consider $x\notin B^j$. Then, by Proposition 2.9, for any cube $P\ni x$,

$$\sum_{Q\in\mathcal{A}^j_*\setminus\mathcal{B}^j}\langle f_j\rangle_Q\beta^j_Q(x)+\sum_{Q\in\mathcal{G}^j:\,P\subsetneq Q}\Delta^{\beta^j}_Q f_j(x)=\frac{\langle f_j\rangle_P}{\langle\beta^j_{\pi_{\mathcal{S}_j}P}\rangle_P}\beta^j_{\pi_{\mathcal{S}^j}P}(x).$$

Now, since $\mathcal{S}^j$ is sparse, almost every x is in only a finite number of cubes $S\in\mathcal{S}^j$. Hence, the proof is finished by appealing to a straightforward modification of [**13**, Lemma 3.5]. □

We also need a Hardy inequality. For a proof, we refer the reader to [**1**, § 9].

Lemma 2.15. *Let Q be any cube in $\mathbb{R}^n$ and let $\kappa>1$. For every $1<p<\infty$,*

$$\int_{\kappa Q\setminus Q}\int_Q\frac{|g_1(y)g_2(x)|}{|x-y|^n}\,\mathrm{d}y\,\mathrm{d}x\lesssim\|g_1\|_p\|g_2\|_{p'},\quad 1/p+1/p'=1,\tag{2.22}$$

holds. The implied constant depends upon κ, p, n.

Proof of (2.11). When expanding $\langle Tf_1,f_2\rangle$ by using (2.21), there are a number of error terms. They are treated by the following estimates and their duals, as applicable, which we do not directly state. For $P\in\mathcal{A}^1_*\setminus\mathcal{B}^1$, the cubes P and Q^0 are roughly of the same size so that $|\langle f_1\rangle_P| \lesssim 1$ by inequality (2.7). Furthermore, using the local Tb hypothesis, the definition of Type A cubes and the Hardy inequality stated in Lemma 2.15,

$$|\langle T\beta_P^1, f_2\rangle| \leqslant |\langle T\beta_P^1, f_2\mathbf{1}_P\rangle| + |\langle T\beta_P^1, f_2\mathbf{1}_{6P\setminus P}\rangle| \lesssim \{1+\boldsymbol{T}_{\mathrm{loc}}+\upsilon_1\boldsymbol{T}\}|Q^0|.$$

And for $P \in \mathcal{A}_*^1\setminus\mathcal{B}^1$ and $Q \in \mathcal{A}_*^2\setminus\mathcal{B}^2$, likewise, we have $|\langle T\beta_P^1, \beta_Q^2\rangle| \lesssim \{1+\boldsymbol{T}_{\mathrm{loc}}+\upsilon_1\boldsymbol{T}\}|Q^0|$. Next, for a cube P as above, there holds, by the assumed norm inequality on T and Lemmas 2.13 and 2.10,

$$\mathbb{E}\{|\langle T\beta_P^1, \phi^2\rangle| + |\langle T\phi^1, f_2\rangle|\} \lesssim \boldsymbol{T}\{\varLambda^{-1} + \varLambda\upsilon_1^{-1}2^{-c\epsilon r}\}|Q^0|.$$

Lastly, when $\varLambda^{-1} + \varLambda\upsilon_1^{-1}2^{-c\epsilon r} < 1$, we have $\mathbb{E}|\langle T\phi^1, \phi^2\rangle| \lesssim \boldsymbol{T}\{\varLambda^{-1} + \varLambda\upsilon_1^{-1}2^{-c\epsilon r}\}|Q^0|$. When combined with (2.7), these inequalities, and their duals, complete the proof of (2.11). □

The proof of Lemma 2.4 and the corona construction are both complete.

3. Useful inequalities

3.1. The martingale transform inequalities

We recall essential tools that we will need. Fix a function b supported on a dyadic* cube S_0, satisfying $\int b\,\mathrm{d}x = |S_0|$ and $\|b\|_p \leqslant \boldsymbol{B}|S_0|^{1/p}$, where $1 < p < \infty$ is fixed. We will consider a fixed but arbitrary collection $\mathcal{T}$ of disjoint dyadic cubes inside S_0, the 'terminal cubes'. Let $\mathcal{Q}$ be all dyadic cubes contained in S_0 but not contained in any terminal cube $T \in \mathcal{T}$. We require that there is $\sigma \in (0,1)$ such that, for all $Q \in \mathcal{Q}$,

$$\left|\int_Q b\,\mathrm{d}x\right| \geqslant 4^{-1}|Q| \quad\text{and}\quad \int_Q |b|^p\,\mathrm{d}x \leqslant \sigma^{-1}\boldsymbol{B}^p|Q|. \tag{3.1}$$

For each terminal cube T, we have a function b_T supported on T and satisfying $\int b_T\,\mathrm{d}x = |T|$ and $\|b_T\|_p \leqslant \boldsymbol{B}|T|^{1/p}$. If the conditions above are met, then we say that the collection, comprised of functions b and b_T, $T \in \mathcal{T}$, is *admissible*. We will not keep track of the constants σ and $\boldsymbol{B}$, and the implied constants will depend upon them.

For $Q \in \mathcal{Q}$ we define the *(half-) twisted martingale differences*

$$D_Q^b f := \sum_{Q'\in\mathrm{ch}(Q)\setminus\mathcal{T}} \left\{\frac{\langle f\rangle_{Q'}}{\langle b\rangle_{Q'}} - \frac{\langle f\rangle_Q}{\langle b\rangle_Q}\right\}\mathbf{1}_{Q'},$$

$$\tilde{D}_Q^b f := \left\{\sum_{Q'\in\mathrm{ch}(Q)\setminus\mathcal{T}} \frac{\langle f\rangle_{Q'}}{\langle b\rangle_{Q'}}\mathbf{1}_{Q'}\right\} - \frac{\langle f\rangle_Q}{\langle b\rangle_Q}\mathbf{1}_Q,$$

$$\Delta_Q^b f := \sum_{Q'\in\mathrm{ch}(Q)} \left\{\frac{\langle f\rangle_{Q'}}{\langle b_{Q'}\rangle_{Q'}}b_{Q'} - \frac{\langle f\rangle_Q}{\langle b\rangle_Q}b\right\}\mathbf{1}_{Q'},$$

where we set $b_{Q'} = b$ if $Q' \notin \mathcal{T}$ and otherwise $b_{Q'}$ is defined as above.

The following theorem is proved in [**1**, Lemma 5.3] and [**18**, § 2].

* In our applications, the underlying dyadic grid will be $\mathcal{D}^j$, $j \in \{1,2\}$.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Theorem 3.1. *Suppose that b and b_T, $T \in \mathcal{T}$, constitutes an admissible collection. Then the following inequalities hold for all selections of constants $|\varepsilon_Q| \leqslant 1$ indexed by $Q \in \mathcal{Q}$:*

$$\Big\| \sum_{Q\in\mathcal{Q}} \varepsilon_Q \tilde{D}_Q^b f \Big\|_q + \Big\| \sum_{Q\in\mathcal{Q}} \varepsilon_Q D_Q^b f \Big\|_q \lesssim \|f\|_q, \quad f \in L^q,\ 1<q<\infty;$$

$$\Big\| \sum_{Q\in\mathcal{Q}} \varepsilon_Q \Delta_Q^b f \Big\|_p \lesssim \|f\|_p, \quad f \in L^p,$$

where $1<p<\infty$ is the exponent associated with the admissible function b.

We will recourse to the following theorem several times. As well as Theorem 3.1, it depends upon the sparseness of the stopping tree $\mathcal{S}^j$.

Theorem 3.2. *Fix $j \in \{1,2\}$. For each cube Q in $\mathbb{R}^n$ and any selection of coefficients $|\varepsilon_P| \lesssim 1$,*

$$\Big\| \sum_{P\in\mathcal{G}^j\colon P\subset Q} \varepsilon_P \Delta_P^{\beta^j} f_j \Big\|_{p_j} \lesssim |Q|^{1/p_j}. \tag{3.2}$$

The same statement also holds true with $\Delta_P^{\beta^j}$ replaced by $\Delta_P^{b^j}$.

Before the proof of this theorem, let us make the following instructive remark.

Remark 3.3. Of particular importance later will be the following assignments. For a fixed $S_0 \in \mathcal{S}^j$ that is not contained in a cube in $\mathcal{B}^j$, we set $\mathcal{T} \subset \mathcal{D}^j$ to be maximal cubes in the collection

$$\mathrm{ch}_{\mathcal{S}^j}(S_0) \cup \{T\colon T\subset S_0,\ T\in \mathrm{ch}(R),\ R\in\mathcal{B}^j\}.$$

By construction of our perturbed stopping data, it is straightforward to verify that the assignments $\beta := \beta^j_{S_0}$ and

$$\beta_T := \begin{cases} b_T^j & T\in\mathrm{ch}(R) \text{ for some } R\in\mathcal{B}^j,\\ \beta_T^j & \text{otherwise}\end{cases}$$

yield an admissible collection with $p=p_j$ and constants $\sigma \simeq \delta$ and $\boldsymbol{B} \simeq \boldsymbol{A}$. Likewise, setting $b := b^j_{S_0}$ and $b_T := b_T^j$ if $T \in \mathcal{T}$ yields admissible functions. Observe also that $P\in\mathcal{Q}$ if $P \in \mathcal{G}^j$ satisfies $\pi_{\mathcal{S}^j}P = S_0$. Under the same assumption, we also have $\Delta_P^{\beta} = \Delta_P^{\beta^j}$ and $\Delta_P^b = \Delta_P^{b^j}$. Here, the right-hand sides are defined in (2.8). Observe that the terminal functions β_T and b_T for $T\in\mathcal{T}\cap\mathrm{ch}(R)$, $R\in\mathcal{B}^j$, do not play any role in these last identities.

Proof of Theorem 3.2. By considering the disjoint collection of those maximal cubes in $\mathcal{G}^j$ that are contained in Q, we are reduced to the case of $Q\in\mathcal{G}^j$. By Theorem 3.1 and Remark 3.3, we first obtain a weaker inequality. Indeed, letting $S=\pi_{\mathcal{S}^j}Q$, we have

$$\Big\| \sum_{\substack{P\colon \pi_{\mathcal{S}^j}P=S\\ P\subset Q}} \varepsilon_P \Delta_P^{\beta^j} f_j \Big\|_{p_j} \lesssim \|f_j \mathbf{1}_Q\|_{p_j} \lesssim |Q|^{1/p_j}. \tag{3.3}$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

We have the last inequality due to the construction of functions f_j: compare to inequalities in (2.20) and recall the normalization of f_j by Λ^{-1}.

We apply (3.3) recursively for the remaining terms, for which $\pi_{\mathcal{S}^j} P \subsetneq Q$. Using the sparseness of the collection $\mathcal{S}^j$, one can easily complete the proof. □

We need a variant of the q-universal inequality for the half-twisted differences to control several error terms that arise. For $P \in \mathcal{G}^j$, let us define

$$\square_P^{\beta^j} f_j := |\tilde{D}_P^{\beta^j} f_j| + \tilde{\chi}_P, \tag{3.4}$$

where

$$\tilde{\chi}_P := \begin{cases} \mathbf{1}_P & \text{if a child of } P \text{ is in } \mathcal{S}^j, \\ 0 & \text{otherwise.} \end{cases}$$

The functions $\square_P^{b^j} f_j$ are defined analogously. If the applied function β^j or b^j is clear from the context, we omit the superscripts. Now, the following q-universal inequality is a consequence of the sparseness of the stopping cubes $\mathcal{S}^j$ and the half-twisted inequality, Theorem 3.1,

$$\bigg\| \bigg[\sum_{P \in \mathcal{G}^j :\, P \subset Q} |\square_P^{\beta^j} f_j|^2 \bigg]^{1/2} \bigg\|_q \lesssim |Q|^{1/q}, \quad 1 < q < \infty. \tag{3.5}$$

Here Q is any cube in $\mathbb{R}^n$ and the corresponding inequality is also true if we use b^j-functions. For further details concerning the proof of (3.5), we refer the reader to [**18**, § 5].

3.2. An estimate for perturbations of b

For a later discussion of the diagonal term in § 5.3, we need a novel *perturbation inequality* for the twisted martingale differences. The estimate is general in nature, but in the interest of brevity we state it in the form needed.

Theorem 3.4. *For $j \in \{1, 2\}$ and $0 < \upsilon_1 < 4^{-1-n}$, we have the following inequality:*

$$\bigg\| \bigg[\sum_{Q \in \mathcal{G}^j} \Big| \{\Delta_Q^{\beta^j} - \Delta_Q^{b^j}\} f_j \Big|^2 \bigg]^{1/2} \bigg\|_{p_j} \lesssim \upsilon_1 |Q^0|^{1/p_j}. \tag{3.6}$$

Proof. Let $\mathcal{R}_0 = \mathcal{A}_*^j$ and inductively set $\mathcal{R}_{k+1}$ to be the maximal cubes $S' \in \mathcal{S}^j$ strictly contained in any $S \in \mathcal{R}_k$. Since $\mathcal{S}^j$ is sparse, we have $\sum_{S \in \mathcal{R}^k} |S| \lesssim \tau^k |Q^0|$ if $k \geqslant 0$. By disjointness of each collection $\mathcal{R}^k$, the left-hand side of (3.6) is bounded by

$$\sum_{k=0}^{\infty} \bigg[\sum_{S \in \mathcal{R}^k} \bigg\| \bigg[\sum_{\substack{Q \in \mathcal{G}^j \\ \pi_{\mathcal{S}^j} Q = S}} |\{\Delta_Q^{\beta^j} - \Delta_Q^{b^j}\} f_j|^2 \bigg]^{1/2} \bigg\|_{p_j}^{p_j} \bigg]^{1/p_j}.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Thus, it suffices to show that

$$\bigg\| \bigg[\sum_{\substack{Q\in\mathcal{G}^j \\ \pi_{\mathcal{S}^j} Q = S}} |\{\Delta_Q^{\beta^j} - \Delta_Q^{b^j}\} f_j|^2 \bigg]^{1/2} \bigg\|_{p_j} \lesssim \upsilon_1 |S|^{1/p_j}, \quad S \in \mathcal{R}^k,\ k \geqslant 0, \tag{3.7}$$

for then sparseness will complete the proof. Below, we regard $S \in \mathcal{S}^j$ as fixed and set $p = p_j$ and $f = f_j$. The basic reduction to a square function involving summation over a (possibly larger) family of cubes $Q \in \mathcal{Q}$ and differences Δ_Q^β and Δ_Q^b is described in Remark 3.3 with $S_0 := S$.

We record useful observations. Without loss of generality, we may assume that the cube S_0 is not contained in any cube in $\mathcal{B}^j$. Thus, a case study using the definition of Type C cubes shows that $\|f\mathbf{1}_{S_0}\|_p \lesssim |S_0|^{1/p}$ and that there is a constant $0 < \lambda \lesssim 1$, independent of the absorption parameters, such that $|\langle f\rangle_Q| \leqslant \lambda$ if $Q \in \mathcal{Q} \cup \mathcal{T}$. (Recall that f is in dyadic BMO and the normalization by $\Lambda = 1$ takes place, $\Lambda \neq \lambda$.)

Let us then quantify that b and β, and the corresponding terminal functions, are 'close'; we define $\upsilon := 2^{n/p}\upsilon_1 < \frac{1}{8}$. That S_0 is not contained in any cube in $\mathcal{B}^j$, combined with the definition of Type A cubes, yields

$$\int_Q |b - \beta|^p = \int_Q |\tilde{\beta}^j_{S_0}|^p \leqslant 2^n \upsilon_1^p |Q| = \upsilon^p |Q|, \quad Q \in \mathcal{Q}. \tag{3.8}$$

We turn to the cubes $T \in \mathcal{T}$; see Remark 3.3. If T is a child of a cube in $\mathcal{B}^j$, then $b_T = \beta_T$. In the complementary case, $T \in \mathrm{ch}_{\mathcal{S}^j}(S_0)$, and its parent is not contained in any cube in $\mathcal{B}^j$. Thus, $\int_T |b_T - \beta_T|^p = \int_T |\tilde{\beta}^j_T|^p \leqslant \upsilon_1^p |T|$ by the definition of Type A cubes. In any case,

$$\int_T |b_T - \beta_T|^p \leqslant \upsilon_1^p |T| \leqslant \upsilon^p |T|, \quad T \in \mathcal{T}. \tag{3.9}$$

This concludes the 'close'-type estimates.

We then proceed to estimate the square function in (3.7), but summed over a possibly larger collection $\mathcal{Q}$. For a cube Q in this collection, we write $\Delta_Q^{\beta^j} f - \Delta_Q^{b^j} f = \Delta_Q^\beta f - \Delta_Q^b f$ as

$$\{D_Q^\beta f - D_Q^b f\} b + D_Q^\beta f \cdot (\beta - b) + \sum_{Q' \in \mathrm{ch}(Q) \cap \mathcal{T}} \{F^1_{Q'} - F^2_{Q'} - F^3_{Q'}\} \mathbf{1}_{Q'},$$

where we have defined

$$F^1_{Q'} := \langle f\rangle_{Q'} \{\beta_{Q'} - b_{Q'}\},$$

$$F^2_{Q'} := \langle f\rangle_Q \bigg\{ \frac{1}{\langle\beta\rangle_Q} - \frac{1}{\langle b\rangle_Q} \bigg\} b, \qquad F^3_{Q'} := \langle f\rangle_Q \frac{1}{\langle\beta\rangle_Q} \{\beta - b\}.$$

The main difficulty will be obtaining the following auxiliary estimate without the b-function,

$$\bigg\| \bigg[\sum_{Q\in\mathcal{Q}} |\{D_Q^\beta - D_Q^b\} f|^2 \bigg]^{1/2} \bigg\|_p \lesssim \upsilon \{ \|f\mathbf{1}_{S_0}\|_p + \lambda |S_0|^{1/p} \}. \tag{3.10}$$

We postpone its proof and first finish the proof of the theorem assuming (3.10). Having martingale difference inequalities, we can proceed as in [**18**, § 2]. Let us first consider the square function of $D_Q^\beta f \cdot (\beta - b)$; to this end, we define

$$Sf := \biggl[\sum_{Q\in\mathcal{Q}} |D_Q^\beta(f\mathbf{1}_{S_0})|^2\biggr]^{1/2}$$

and consider the events $E_t := \{|Sf| \geqslant t\} \subset S_0$ for $t > 0$. It is important to realize that we can compare Lebesgue measure estimates and estimates with respect to $|\beta - b|^p \,\mathrm{d}x$. Namely, by inequality (3.8), the Lebesgue differentiation theorem, and the fact that Sf is constant on terminal cubes $T \in \mathcal{T}$, we obtain $\int_{E_t} |\beta - b|^p \,\mathrm{d}x \leqslant 2^n \upsilon^p |E_t|$. Hence, by the Lebesgue measure estimates in Theorem 3.1,

$$\int_{S_0} |Sf|^p |\beta - b|^p \,\mathrm{d}x = p\int_0^\infty t^{p-1} \int_{E_t} |\beta - b|^p \,\mathrm{d}x \,\mathrm{d}t \lesssim \upsilon^p \|f\mathbf{1}_{S_0}\|_p^p.$$

The square function of $\{D_Q^\beta f - D_Q^b f\}b$ is estimated analogously by using (3.10), which also contributes the constant υ. The remaining square functions associated with $F_{Q'}^i$, $i = 1, 2, 3$, are estimated by using (3.8) and (3.9), and the facts that $\mathcal{T}$ is a disjoint collection and that the Lebesgue measure is doubling. For the case in which $i = 2$, we also use expansion (3.12) and the first inequality in (3.11) below. This concludes the proof of the theorem except for (3.10).

Let us continue with the following preparations for proving (3.10). Fix $Q \in \mathcal{Q}$ and $Q' \in \mathrm{ch}(Q) \setminus \mathcal{T}$. Set $\tilde\beta = b - \beta$ and write $\beta_{k,Q} := (\langle\tilde\beta\rangle_Q / \langle\beta\rangle_Q)^k$. Define $\beta_{k,Q'}$ analogously. Observe that the following inequalities hold for every $k \geqslant 1$:

$$|\beta_{k,Q'}| + |\beta_{k,Q}| \leqslant 2(4\upsilon)^k, \qquad |\beta_{k,Q'} - \beta_{k,Q}| \leqslant |\beta_{1,Q'} - \beta_{1,Q}| k(8\upsilon)^{k-1}. \tag{3.11}$$

Indeed, these follow from (3.1), (3.8) and the fact that $Q, Q' \in \mathcal{Q}$. For the latter inequality above, one also applies the mean value theorem.

We then write

$$\frac{1}{\langle\beta\rangle_Q} - \frac{1}{\langle b\rangle_Q} = \frac{1}{\langle\beta\rangle_Q}\biggl\{1 - \frac{\langle\beta\rangle_Q}{\langle\beta\rangle_Q + \langle\tilde\beta\rangle_Q}\biggr\} = \frac{1}{\langle\beta\rangle_Q}\sum_{k=1}^\infty (-1)^{k+1}\beta_{k,Q}. \tag{3.12}$$

Using the same expansion with Q replaced by Q' yields

$$|\{D_Q^\beta - D_Q^b\}f|\mathbf{1}_{Q'} \leqslant \sum_{k=1}^\infty \biggl|\frac{\langle f\rangle_{Q'}}{\langle\beta\rangle_{Q'}}\beta_{k,Q'} - \frac{\langle f\rangle_Q}{\langle\beta\rangle_Q}\beta_{k,Q}\biggr|\mathbf{1}_{Q'}.$$

Then, for a fixed k, write the summand on the right-hand side as

$$\begin{aligned}&\biggl|\beta_{k,Q} D_Q^\beta f \cdot \mathbf{1}_{Q'} + \frac{\langle f\rangle_{Q'}}{\langle\beta\rangle_{Q'}}\{\beta_{k,Q'} - \beta_{k,Q}\}\mathbf{1}_{Q'}\biggr| \\ &\qquad\leqslant 2(4\upsilon)^k |D_Q^\beta f|\mathbf{1}_{Q'} + 4\lambda k(8\upsilon)^{k-1}|\beta_{1,Q'} - \beta_{1,Q}|\mathbf{1}_{Q'}.\end{aligned}$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

Here we used assumptions and (3.11). Observe that $|\beta_{1,Q'} - \beta_{1,Q}|\mathbf{1}_{Q'} = |D^{\beta}_{Q}\tilde{\beta}|\mathbf{1}_{Q'}$. By summing the series over k and then summing resulting estimates over $Q' \in \mathrm{ch}(Q) \setminus \mathcal{T}$,

$$|\{D^{\beta}_{Q} - D^{b}_{Q}\}f| \lesssim \upsilon|D^{\beta}_{Q}f| + \lambda|D^{\beta}_{Q}\tilde{\beta}| = \upsilon|D^{\beta}_{Q}(f\mathbf{1}_{S_0})| + \lambda|D^{\beta}_{Q}(\tilde{\beta}\mathbf{1}_{S_0})|.$$

Inequality (3.10) follows from (3.8) and universal martingale transform inequalities with $q = p$; see Theorem 3.1. □

4. The inner product and the main term

During the course of the remaining sections we prove (2.12), namely,

$$\left|\sum_{P\in\mathcal{G}^1}\sum_{Q\in\mathcal{G}^2}\langle T\Delta^{\beta^1}_{P}f_1, \Delta^{\beta^2}_{Q}f_2\rangle\right| \leqslant \{C_2\{1 + \boldsymbol{T}_{\mathrm{loc}}\} + C_3 r\upsilon_1 \varLambda^2\boldsymbol{T}\}|Q^0|, \tag{4.1}$$

where C_3 is a constant not allowed to depend upon the absorption parameters. This inequality completes the proof of Lemma 2.1 which, in turn, implies our main result. Let us recall that the functions f_j have been normalized, allowing us to assume that $\varLambda = 1$.

The sum above is split into dual triangular sums, one of which is the sum over $(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2$ such that $\ell P \geqslant \ell Q$. By using goodness this triangular sum is split into different collections:

$$\begin{aligned}\mathcal{P}_{\mathrm{far}} &:= \{(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2 \colon 3P \cap Q = \emptyset,\ \ell Q \leqslant \ell P\};\\ \mathcal{P}_{\mathrm{diagonal}} &:= \{(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2 \setminus \mathcal{P}_{\mathrm{far}} \colon 2^{-r}\ell P \leqslant \ell Q \leqslant \ell P\};\\ \mathcal{P}_{\mathrm{nearby}} &:= \{(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2 \setminus \mathcal{P}_{\mathrm{diagonal}} \colon Q \subset 3P \setminus P\};\\ \mathcal{P}_{\mathrm{inside}} &:= \{(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2 \setminus \mathcal{P}_{\mathrm{diagonal}} \colon Q \subset P\}.\end{aligned}$$

The sums over these collections are handled separately and, aside from the ‘inside’ and ‘diagonal’ terms, one can sum over the absolute value of the inner products. The main tools to control these terms include the twisted martingale transform inequalities combined with the local Tb hypothesis. All of the cubes are good, which is a point used systematically. This useful fact is frequently combined with the smoothness condition on the kernel to conclude that certain maximal functions applied to the β functions appear. That these maximal functions are controlled will be a consequence of the corona construction combined with the universal half-twisted martingale inequalities. In the analysis of the diagonal term, perturbation inequalities in § 3 play a key role.

In this section, we concentrate on the ‘inside’ term, which is the main term. The conditions for $(P,Q) \in \mathcal{P}_{\mathrm{inside}}$ are $Q \subset P$, $2^r\ell Q < \ell P$ and $(P,Q) \in \mathcal{G}^1 \times \mathcal{G}^2$; these conditions are abbreviated to $Q \Subset P$ below. Even though Q is in a different grid from that of P, a child of P contains Q because of goodness and we denote that child by P_Q. We will write $\Delta_P := \Delta^{\beta^1}_{P}$ (likewise for Q) and $\tilde{\Delta}_P f_1 := \tilde{D}_P f_1 \cdot \beta^1_{\pi_{\mathcal{S}^1}P}$, where the half-twisted martingale difference $\tilde{D}_P = \tilde{D}^{\beta^1}_{P}$ of (2.9) does not sum over the children of P that have a different stopping parent from that of P.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

In order to control the inside term, it suffices to bound the sum over $S \in \mathcal{S}^1$ of the terms

$$\left|\mathbf{1}_{\{\pi S \in \mathcal{G}^1\}} \sum_{Q\colon Q \Subset \pi S} \langle f_1\rangle_S \langle T\beta_S^1, \Delta_Q f_2\rangle\right| + \Big|\underbrace{\sum_{P\colon \pi_{\mathcal{S}^1} P = S} \sum_{Q\colon Q\Subset P} \langle T\tilde{\Delta}_P f_1, \Delta_Q f_2\rangle}_{=:B_S(f_1,f_2)}\Big|. \qquad (4.2)$$

The point of this step is that, in the left-hand side, the argument of T depends only on β_S^1. And a sufficient cube-wise inequality is

$$(4.2) \lesssim \tilde{\boldsymbol{T}}_{\mathrm{loc}}|S|, \quad \tilde{\boldsymbol{T}}_{\mathrm{loc}} := \boldsymbol{T}_{\mathrm{loc}} + \upsilon_1 \boldsymbol{T},$$

where the implied constant is not allowed to depend upon the absorption parameters. Since the collection $\mathcal{S}^1$ is sparse, this upper bound is summable over $S \in \mathcal{S}^1$ to a multiple of $\tilde{\boldsymbol{T}}_{\mathrm{loc}}|Q^0|$.

The left-hand side of (4.2) is easy to control. First of all, by the local Tb properties stated in Lemma 2.12 and the twisted martingale inequality (3.2),

$$\left|\mathbf{1}_{\{\pi S\in\mathcal{G}^1\}} \sum_{\substack{Q\colon Q\Subset \pi S\\ Q\subset S}} \langle f_1\rangle_S \langle T\beta_S^1, \Delta_Q f_2\rangle\right| \lesssim \tilde{\boldsymbol{T}}_{\mathrm{loc}}|S|^{1/p_2'} \left\| \sum_{Q\colon Q\Subset\pi S} \langle f_1\rangle_S \Delta_Q f_2\right\|_{p_2} \lesssim \tilde{\boldsymbol{T}}_{\mathrm{loc}}|S|.$$

The remaining part of the left-hand side is a sum over cubes $Q \Subset \pi S$ for which $Q\cap S = \emptyset$. This part is conveniently estimated by using Hardy's inequality in Lemma 2.15 and the inequality $p_2' \leqslant p_1$.

In the right-hand side of (4.2) the argument of T is written as follows. If $Q \Subset P$ and $\pi_{\mathcal{S}^1} P = S$, then

$$\begin{aligned}\tilde{\Delta}_P f_1 &= \langle \tilde{D}_P f_1\rangle_{P_Q} \beta_S^1 \mathbf{1}_S - \langle \tilde{D}_P f_1\rangle_{P_Q}\beta_S^1 \mathbf{1}_{S\setminus P_Q} + \tilde{\Delta}_P f_1 \cdot \mathbf{1}_{P\setminus P_Q}\\ &=: \Delta_P^{\mathrm{para}} f_1 - \Delta_P^{\mathrm{stop}} f_1 + \Delta_P^{\mathrm{error}} f_1,\end{aligned}$$

where we treat $\tilde{\Delta}_P f_1 \mathbf{1}_{P_Q}$ as the main contribution and write $\mathbf{1}_{P_Q} = \mathbf{1}_S - \mathbf{1}_{S\setminus P_Q}$. This decomposition of $\tilde{\Delta}_P f_1$ leads to a corresponding decomposition of $B_S(f_1, f_2)$, by which we denote the second term on (4.2) without the absolute values, into the paraproduct term, the stopping term and the error term, written as

$$B_S(f_1,f_2) = B_S^{\mathrm{para}}(f_1,f_2) - B_S^{\mathrm{stop}}(f_1,f_2) + B_S^{\mathrm{error}}(f_1,f_2),$$

where $S \in \mathcal{S}^1$ is fixed. The terminology is drawn from [**23, 25**].

4.1. Control of the paraproduct term

This brief argument is in fact the core of the proof. Consider $B_S^{\mathrm{para}}(f_1, f_2)$. In this term, the argument of T is a certain multiple of $\beta_S^1 \mathbf{1}_S = \beta_S^1$. For the cubes $Q \in \mathcal{G}^2$, let us define

$$\varepsilon_Q := \sum_{\substack{P\colon \pi_{\mathcal{S}^1} P = S\\ Q\Subset P}} \langle \tilde{D}_P f_1\rangle_{P_Q}.$$

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

The condition* (2.10) of Lemma 2.4, was designed so that the numbers ε_Q are uniformly bounded. We can therefore estimate

$$\begin{aligned}|B_S^{\text{para}}(f_1,f_2)| &= \bigg|\sum_{P:\,\pi_{\mathcal{S}^1}P=S}\ \sum_{Q:\,Q\Subset P}\langle\tilde{D}_P f_1\rangle_{P_Q}\langle T\beta_S^1,\Delta_Q f_2\rangle\bigg| \\ &= \bigg|\bigg\langle T\beta_S^1,\sum_{Q:\,Q\Subset S}\varepsilon_Q\Delta_Q f_2\bigg\rangle\bigg| \\ &\leqslant \|\mathbf{1}_S T\beta_S^1\|_{p_2'}\bigg\|\sum_{Q:\,Q\Subset S}\varepsilon_Q\Delta_Q f_2\bigg\|_{p_2} \\ &\lesssim \tilde{\boldsymbol{T}}_{\text{loc}}|S|,\end{aligned}$$

where we appealed to the local Tb hypothesis, Lemma 2.12 (3) (c) and the martingale transform inequality (3.2). This completes the analysis of the paraproduct term.

4.2. The stopping term

Recall that

$$|B_S^{\text{stop}}(f_1,f_2)| = \bigg|\sum_{P:\,\pi_{\mathcal{S}^1}P=S}\ \sum_{Q:\,Q\Subset P}\langle\tilde{D}_P f_1\rangle_{P_Q}\langle T(\beta_S^1\mathbf{1}_{S\setminus P_Q}),\Delta_Q f_2\rangle\bigg|.$$

We will bound this by a constant multiple of $|S|$ by appealing to the fact that

(a) $\int\Delta_Q f_2 = 0$ and the kernel of T has smoothness and

(b) the universal half-twisted inequality (3.5) is valid.

For integers $s>r$, we restrict the side length of Q so that $2^s\ell Q=\ell P$ and thereby obtain a geometric decay in s. To accommodate this, let us define

$$B_{S,s}^{\text{stop}}(f_1,f_2) := \sum_{P:\,\pi_{\mathcal{S}^1}P=S}\ \sum_{\substack{Q:\,Q\Subset P\\ 2^s\ell Q=\ell P}}\langle\tilde{D}_P f_1\rangle_{P_Q}\langle T(\beta_S^1\mathbf{1}_{S\setminus P_Q}),\Delta_Q f_2\rangle.$$

By goodness, $\operatorname{dist}(S\setminus P_Q,Q)\geqslant(\ell Q)^\epsilon(\ell P_Q)^{1-\epsilon}$. Therefore, by the smoothness condition on the kernel and the mean zero property of $\Delta_Q f_2$, we can estimate the inner product as follows; let x_Q be the centre of Q and recall also definition (3.4). Then,

$$\begin{aligned}|\langle T(\beta_S^1\mathbf{1}_{S\setminus P_Q}),\Delta_Q f_2\rangle| &\lesssim \int_Q\int_{S\setminus P_Q}\frac{(\ell Q)^\eta}{|x-y|^{n+\eta}}|\beta_S^1(y)\Delta_Q f_2(x)|\,\mathrm{d}y\,\mathrm{d}x \\ &\lesssim 2^{-\eta' s}\inf_{x\in Q}M\beta_S^1(x)\int_Q\square_Q f_2\,\mathrm{d}x.\end{aligned}$$

This is a standard off-diagonal estimate, obtained by using (1.1) with $\int\Delta_Q f_2=0$ and splitting the region of integration into annuli, combined with the goodness of Q and the

* The condition applies to the minimal cube in $\mathcal{G}^1$, subject to the summation conditions, instead of Q.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

properties of our corona construction. Observe that we gained a geometric decay in s with $\eta' = (1-\epsilon)\eta > 0$.

Since cubes Q with same side length, specified by P, are disjoint, there is a simple appeal to the Cauchy–Schwarz inequality. Following that, we use the trilinear form of Hölder's inequality, with indices p_1, $2p_1'$ and $2p_1'$, and the universal half-twisted inequality (3.5). By doing so, we obtain

$$\begin{aligned} |B^{\mathrm{stop}}_{S,s}(f_1,f_2)| &\lesssim 2^{-\eta' s} \sum_{P\,:\,\pi_{\mathcal{S}^1}P=S} \sum_{\substack{Q\,:\,Q\Subset P\\ 2^s\ell Q=\ell P}} \langle |\tilde{D}_P f_1|\rangle_P \int_Q M\beta^1_S \cdot \square_Q f_2 \,\mathrm{d}x \\ &\lesssim 2^{-\eta' s} \int_S M\beta^1_S \bigg[\sum_{P\,:\,\pi_{\mathcal{S}^1}P=S} \langle |\tilde{D}_P f_1|\rangle_P^2 \mathbf{1}_P\bigg]^{1/2} \bigg[\sum_{Q\,:\,Q\Subset S} |\square_Q f_2|^2\bigg]^{1/2} \mathrm{d}x \\ &\lesssim 2^{-\eta' s}|S|^{1/p_1} \bigg\| \bigg[\sum_{P\,:\,\pi_{\mathcal{S}^1}P=S} |M\tilde{D}_P f_1|^2\bigg]^{1/2}\bigg\|_{2p_1'} \bigg\|\bigg[\sum_{Q\,:\,Q\Subset S} |\square_Q f_2|^2\bigg]^{1/2}\bigg\|_{2p_1'}. \end{aligned} \tag{4.3}$$

The last term is dominated by a constant multiple of $2^{-\eta' s}|S|$ and this completes the analysis of the stopping term.

4.3. The error term

Here we need to control

$$|B^{\mathrm{error}}_S(f_1,f_2)| = \bigg| \sum_{s=r+1}^{\infty} \sum_{P\,:\,\pi_{\mathcal{S}^1}P=S} \sum_{\substack{Q\,:\,Q\Subset P\\ 2^s\ell Q=\ell P}} \langle T(\tilde{\Delta}_P f_1 \cdot \mathbf{1}_{P\setminus P_Q}), \Delta_Q f_2\rangle \bigg|.$$

For a fixed $s > r$, we call the inner double series in the display above $B^{\mathrm{error}}_{S,s}(f_1,f_2)$. We will obtain a geometric decay in s by using essentially the same argument as in the treatment of the stopping term.

Indeed,

$$\begin{aligned} |\langle T(\tilde{\Delta}_P f_1 \cdot \mathbf{1}_{P\setminus P_Q}), \Delta_Q f_2| &= \bigg| \int_Q \int_{P\setminus P_Q} \{K(x,y) - K(x_Q,y)\} \tilde{\Delta}_P f_1(y) \Delta_Q f_2(x) \,\mathrm{d}y\,\mathrm{d}x \bigg| \\ &\lesssim \int_Q \int_{P\setminus P_Q} \frac{(\ell Q)^\eta}{|x-y|^{n+\eta}} |\tilde{\Delta}_P f_1(y) \Delta_Q f_2(x)| \,\mathrm{d}y\,\mathrm{d}x \\ &\lesssim 2^{-\eta' s} \langle |\tilde{D}_P f_1|\rangle_P \inf_{x\in Q} M\beta^1_S \int_Q \square_Q f_2 \,\mathrm{d}x. \end{aligned}$$

Repeating the inequalities starting from (4.3) gives $|B^{\mathrm{error}}_{S,s}(f_1,f_2)| \lesssim 2^{-\eta' s}|S|$, and this suffices.

5. The remaining terms

In this section we estimate all the remaining terms 'nearby', 'far' and 'diagonal'.

5.1. The nearby term

The nearby term concerns pairs of cubes $(P,Q) \in \mathcal{P}_{\mathrm{nearby}}$, that is, cubes in $\mathcal{G}^1 \times \mathcal{G}^2$ with the properties $2^r \ell Q < \ell P$ and $Q \subset 3P \backslash P$. This term can be written as a sum over $S \in \mathcal{S}^1$ of terms

$$\mathbf{1}_{\{\pi S \in \mathcal{G}^1\}} \sum_{\substack{Q:\, Q \subset 3\pi S \backslash \pi S \\ 2^r \ell Q < \ell \pi S}} \langle f_1 \rangle_S \langle T\beta^1_S, \Delta_Q f_2 \rangle + \sum_{P:\, \pi_{\mathcal{S}^1} P = S} \sum_{\substack{Q:\, Q \subset 3P \backslash P \\ 2^r \ell Q < \ell P}} \langle T\tilde{\Delta}_P f_1, \Delta_Q f_2 \rangle, \quad (5.1)$$

where we tacitly assume that $P \in \mathcal{G}^1$ and $Q \in \mathcal{G}^2$. For a fixed $S \in \mathcal{S}^1$, the absolute value of the double series above is estimated by

$$\sum_{s > r} \sum_{P:\, \pi_{\mathcal{S}^1} P = S} \sum_{\substack{Q:\, Q \subset 3P \backslash P \\ 2^s \ell Q = \ell P}} |\langle T\tilde{\Delta}_P f_1, \Delta_Q f_2 \rangle|. \quad (5.2)$$

By using Lemma 5.1 and following the arguments in (4.3) with obvious changes, we find that the inner double series in (5.2), with a fixed $s > r$, is dominated by

$$2^{-s\eta'} \sum_{P:\, \pi_{\mathcal{S}^1} P = S} \sum_{\substack{Q:\, Q \subset 3P \backslash P \\ 2^s \ell Q = \ell P}} \inf_{x \in Q} M\beta^1_S(x) \langle |\tilde{D}_P f_1| \rangle_P \int_Q \square_Q f_2(x) \,\mathrm{d}x \lesssim 2^{-s\eta'} |S|.$$

The right-hand side is summable in s to a constant multiple of $|S|$. Consequently, by applying the sparseness of $\mathcal{S}^1$, we find that (5.2) summed over $S \in \mathcal{S}^1$ is bounded by a constant multiple of $|Q^0|$. The same method of proof controls the first term in (5.1). Alternatively, one may apply the Hardy inequality (see Lemma 2.15).

We now turn to a lemma that is used above. Its proof is a standard off-diagonal argument using the smoothness condition (1.2) and goodness of Q. We omit the easy proof.

Lemma 5.1. *Let $(P,Q) \in \mathcal{P}_{\mathrm{nearby}}$ with $\pi_{\mathcal{S}^1} P = S$. Then, with $\eta' = \eta(1-\epsilon) > 0$, we have*

$$|\langle T\tilde{\Delta}_P f_1, \Delta_Q f_2 \rangle| \lesssim (\ell Q/\ell P)^{\eta'} \inf_{x \in Q} M\beta^1_S(x) \langle |\tilde{D}_P f_1| \rangle_P \int_Q \square_Q f_2 \,\mathrm{d}x. \quad (5.3)$$

5.2. The far term

The far term concerns pairs of cubes $(P,Q) \in \mathcal{P}_{\mathrm{far}}$ satisfying $\ell Q \leqslant \ell P$ and $3P \cap Q = \emptyset$, in particular. The goodness of these cubes is irrelevant here. The absolute value of the far term is bounded by the sum over integers $s \geqslant 0$ and $t \geqslant 1$ of terms

$$\sum_P \sum_{\substack{Q:\, 2^{t-1}\ell P \leqslant \mathrm{dist}(P,Q) < 2^t \ell P \\ 2^s \ell Q = \ell P}} \mathbf{1}_{(P,Q) \in \mathcal{P}_{\mathrm{far}}} |\langle T\Delta_P f_1, \Delta_Q f_2 \rangle|. \quad (5.4)$$

By Lemma 5.2, we obtain the following upper bounds for the term (5.4):

$$\begin{aligned}2^{-\eta s-(n+\eta)t}\int_{\mathbb{R}^n}\sum_{P}\sum_{\substack{Q\colon 2^{t-1}\ell P\leqslant \operatorname{dist}(P,Q)<2^t\ell P\\ 2^s\ell Q=\ell P}}\langle \Box_P f_1\rangle_P \mathbf{1}_Q(x)\Box_Q f_2(x)\,\mathrm{d}x \\ \lesssim 2^{-\eta s-\eta t}\bigg\|\bigg[\sum_{P\in\mathcal{G}^1}|M\Box_P f_1|^2\bigg]^{1/2}\bigg\|_2\bigg\|\bigg[\sum_{Q\in\mathcal{G}^2}|\Box_Q f_2|^2\bigg]^{1/2}\bigg\|_2 \\ \lesssim 2^{-\eta(s+t)}|Q^0|. \qquad (5.5)\end{aligned}$$

Observe that in the first estimate we lose a factor $2^{nt/2}$ twice because of additional summation associated with both of the square functions. In order to see this for the first square function, one changes the order of summation and integration and then applies inequality $|P|^{-1}\sum_Q |Q| \lesssim 2^{tn}$ for each P inside the P-summation.

The last bound in (5.5) is still summable in s and t so that we are left with the following lemma. We omit the easy proof that is of standard off-diagonal nature.

Lemma 5.2. *Let $(P,Q)\in\mathcal{P}_{\mathrm{far}}$. Then,*

$$|\langle T\Delta_P f_1, \Delta_Q f_2\rangle| \lesssim (\ell Q/\ell P)^{\eta}\left(\frac{\operatorname{dist}(P,Q)}{\ell P}\right)^{-n-\eta}\langle \Box_P f_1\rangle_P\cdot\int_Q \Box_Q f_2.$$

5.3. The diagonal term

The diagonal term is the hardest in many local Tb arguments and this is also true in our situation. The goal is to prove the following inequality:

$$\bigg|\sum_{P\in\mathcal{G}^1}\sum_{\substack{Q\in\mathcal{G}^2,\\ Q\cap 3P\neq\emptyset,\\ 2^{-r}\ell P\leqslant \ell Q\leqslant \ell P}}\langle T\Delta_P f_1,\Delta_Q f_2\rangle\bigg| \lesssim \{C_r(1+\boldsymbol{T}_{\mathrm{loc}})+r\upsilon_1\boldsymbol{T}\}|Q^0|. \qquad (5.6)$$

Note, in particular, that the bound in terms of $\boldsymbol{T}$ has leading absorbing constant $r\upsilon_1$. On the other hand, $\boldsymbol{T}_{\mathrm{loc}}$ has a leading constant C_r that will be exponential in r. The implied constant is independent of the absorption parameters.

The first step in the proof is not so straightforward. Its purpose is to avoid terms $\{2^{cr}\upsilon_1\boldsymbol{T}\}|Q^0|$ that cannot be absorbed. To explain, let us pass back to the heavier notation $\Delta_P f_1 = \Delta_P^{\beta^1} f_1$; the point of the estimate below is that we will replace β^1 in the twisted differences by b^1.

Lemma 5.3. *There holds*

$$\bigg|\sum_{P\in\mathcal{G}^1}\sum_{\substack{Q\in\mathcal{G}^2,\\ Q\cap 3P\neq\emptyset,\\ 2^{-r}\ell P\leqslant \ell Q\leqslant \ell P}}\langle T(\Delta_P^{\beta^1} f_1-\Delta_P^{b^1} f_1),\Delta_Q f_2\rangle\bigg| \lesssim \{r\upsilon_1\boldsymbol{T}\}|Q^0|.$$

Proof. Perturbation inequality is the principal tool here. By introducing independent Rademacher variables $\{\epsilon_P\}_{P\in\mathcal{G}^1}$ that are jointly supported on a probability space $\Omega = \{-1,1\}^{\mathcal{G}^1}$, we have, for integers $0 \leqslant s \leqslant r$,

$$\begin{aligned}
&\left|\sum_{P\in\mathcal{G}^1} \sum_{\substack{Q\in\mathcal{G}^2\\ Q\cap 3P\neq\emptyset,\\ 2^{-s}\ell P=\ell Q}} \langle T(\Delta_P^{\beta^1} f_1 - \Delta_P^{b^1} f_1), \Delta_Q f_2\rangle\right| \\
&\qquad = \left| \int_\Omega \left\langle \sum_{P\in\mathcal{G}^1} \epsilon_P T(\Delta_P^{\beta^1} f_1 - \Delta_P^{b^1} f_1),\ \sum_{R\in\mathcal{G}^1} \epsilon_R \sum_{\substack{Q\in\mathcal{G}^2\\ Q\cap 3R\neq\emptyset,\\ 2^{-s}\ell R=\ell Q}} \Delta_Q f_2 \right\rangle \mathrm{d}\epsilon \right| \\
&\qquad \lesssim \left\{ \int_\Omega \left\| T\left(\sum_{P\in\mathcal{G}^1} \epsilon_P \{\Delta_P^{\beta^1} - \Delta_P^{b^1}\} f_1 \right)\right\|_{p_1}^{p_1} \mathrm{d}\epsilon \right\}^{1/p_1} \\
&\qquad\quad \times \left\{ \int_\Omega \left\| \sum_{P\in\mathcal{G}^1} \sum_{\substack{Q\in\mathcal{G}^2\\ Q\cap 3P\neq\emptyset,\\ 2^{-s}\ell P=\ell Q}} \epsilon_P \Delta_Q f_2 \right\|_{p_1'}^{p_1'} \mathrm{d}\epsilon \right\}^{1/p_1'}.
\end{aligned}$$

Extract the operator norm from the first factor and after that apply Khintchine's inequality and Theorem 3.4. Theorem 3.2 is used to estimate the second factor, but only after having changed the order of summation and having applied Hölder's inequality and inequality $p_1' \leqslant p_2$. Finally, summing the s-series yields the upper bound $rv_1\boldsymbol{T}|Q^0|$. □

It remains to prove Lemma 5.4. Indeed, a straightforward application of (3.5), combined with the two lemmata 5.3 and 5.4, completes the proof of the diagonal estimate (5.6).

Lemma 5.4. *Assume that $3P\cap Q\neq\emptyset$ and that $2^{-r}\ell P \leqslant \ell Q \leqslant \ell P$. Then,*

$$|\langle T\Delta_P^{b^1} f_1, \Delta_Q f_2\rangle| \lesssim \{1+\boldsymbol{T}_{\mathrm{loc}}\}\langle \Box_P^{b^1} f_1\rangle_P \langle \Box_Q f_2\rangle_Q |P|.$$

Proof. The cube P has 2^n children P'. If a child P' is not a stopping cube, $\Delta_P^{b^1} f_1 \cdot \mathbf{1}_{P'}$ is equal to a multiple of $b^1_{S_1}\mathbf{1}_{P'}$, where S_1 is the $\mathcal{S}^1$ parent of P, and the multiple is given by the value of the half-twisted martingale difference $\tilde{D}_P^{b^1} f_1$ on P'. If P' is a stopping cube, then $\Delta_P^{b^1} f_1\cdot\mathbf{1}_{P'}$ in addition involves a bounded multiple of $b^1_{P'}$. In both cases, the constant multiples are bounded in absolute value by $\langle \Box_P^{b^1} f_1\rangle_P$ (cf. definition (3.4)). Similar comments apply to $\Delta_Q f_2 = \Delta_Q^{\beta^2} f_2$ restricted to a child Q'. By these considerations, we need to prove the estimate

$$|\langle T\psi^1, \psi^2\rangle| \lesssim \{1+\boldsymbol{T}_{\mathrm{loc}}\}|P|,$$

where $\psi^1 = b^1_{S_1}\mathbf{1}_{P'}$ and $\psi^2 \in \{\beta^2_{S_2}\mathbf{1}_{Q'}, \beta^2_{Q'}\mathbf{1}_{\{Q'\in\mathcal{S}^2\}}\}$, where S_2 is the $\mathcal{S}^2$ parent of Q. A similar estimate is also required when $\psi^1 = b^1_{P'}$, on the condition that $P'\in\mathcal{S}^1$. An obstruction is that, even though the stopping conditions control the local norm of $Tb^1_{S_1}$, we may have the restriction $b^1_{S_1}\mathbf{1}_{P'}$ *inside* the operator T.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

The case of $\psi^1 = b^1_{P'}$, where we require that $P' \in \mathcal{S}^1$, is especially easy since the obstruction just mentioned does not arise. By the construction of the stopping cubes and the fact that the Lebesgue measure is doubling, $|\langle Tb^1_{P'}, \psi^2 \mathbf{1}_{P'}\rangle| \lesssim \|\mathbf{1}_{P'} Tb^1_{P'}\|_{p_2'} \|\psi^2\|_{p_2} \lesssim \boldsymbol{T}_{\mathrm{loc}}|P|$; here we complied to the stopping rules by restricting ψ_2 to P'. Concerning the contribution outside of P', inequality $p_1' \leqslant p_2$ and Hardy's inequality in Lemma 2.15 together yield $|\langle Tb^1_{P'}, \psi^2 \mathbf{1}_{Q'\setminus P'}\rangle| \lesssim |P|$.

For the case in which $\psi^1 = b^1_{S_1}\mathbf{1}_{P'}$ we must face the obstruction. Again, we write $\psi^2 = \psi^2\mathbf{1}_{P'} + \psi^2\mathbf{1}_{Q'\setminus P'}$. Hardy's inequality controls the second term, giving $|\langle Tb^1_{S_1}\mathbf{1}_{P'}, \psi^2\mathbf{1}_{Q'\setminus P'}\rangle| \lesssim |P|$. For the first term, we return to the local Tb hypothesis and write

$$\psi^2\mathbf{1}_{P'} = \langle\psi^2\rangle_{P'} b^2_{P'} + (\psi^2\mathbf{1}_{P'} - \langle\psi^2\rangle_{P'} b^2_{P'}) =: \langle\psi^2\rangle_{P'} b^2_{P'} + \tilde{\psi}^2.$$

The advantage of the first summand on the right is that the local Tb hypothesis gives us

$$|\langle\psi^2\rangle_{P'}\langle Tb^1_{S_1}\mathbf{1}_{P'}, b^2_{P'}\rangle| = |\langle\psi^2\rangle_{P'}\langle b^1_{S_1}\mathbf{1}_{P'}, T^* b^2_{P'}\rangle| \lesssim \boldsymbol{T}_{\mathrm{loc}}|\langle\psi^2\rangle_{P'}||P| \lesssim \boldsymbol{T}_{\mathrm{loc}}|P|.$$

The advantage of the second summand is that it has integral zero: $\int_{P'} \tilde{\psi}^2 \,\mathrm{d}x = 0$. Note also that $\|\tilde{\psi}^2\|_{p_2} \lesssim |P|^{1/p_2}$. Take $\mathcal{P}$ to be the cubes of the form $P' \dotplus u$, where $u \in \{-1,0,1\}^n \setminus \{(0,0,\dots)\}$. Then,

$$\begin{aligned} |\langle Tb^1_{S_1}\mathbf{1}_{P'}, \tilde{\psi}^2\rangle| &\leqslant |\langle Tb^1_{S_1}, \tilde{\psi}^2\rangle| + |\langle Tb^1_{S_1}\mathbf{1}_{S_1\setminus P'}, \tilde{\psi}^2\rangle| \\ &\leqslant |\langle Tb^1_{S_1}, \tilde{\psi}^2\rangle| + \sum_{R\in\mathcal{P}} |\langle Tb^1_{S_1}\mathbf{1}_R, \tilde{\psi}^2\rangle| + |\langle Tb^1_{S_1}\mathbf{1}_{S_1\setminus 3P'}, \tilde{\psi}^2\rangle|. \end{aligned}$$

The first term is controlled by the stopping rules: $|\langle Tb^1_{S_1}, \tilde{\psi}^2\rangle| \lesssim \boldsymbol{T}_{\mathrm{loc}}|P|$. The second sum is finite and each summand is precisely of the type that appears in Hardy's inequality. Indeed, although $R \in \mathcal{P}$ need not be contained in P, by (2.13) we nevertheless have

$$\langle |b^1_{S_1}|^{p_2'}\rangle_R^{1/p_2'} \leqslant \langle |b^1_{S_1}|^{p_1}\rangle_R^{1/p_1} \leqslant \Big\{4^n \inf_{x\in P} M|b^1_{S_1}|^{p_1}\Big\}^{1/p_1} \lesssim 1.$$

And it follows that

$$\sum_{R\in\mathcal{P}} |\langle Tb^1_{S_1}\mathbf{1}_R, \tilde{\psi}^2\rangle| \lesssim \sum_{R\in\mathcal{P}} \|b^1_{S_1}\mathbf{1}_R\|_{p_2'}|P|^{1/p_2} \lesssim |P|.$$

Finally, by a similar estimate as in the proof of Lemma 5.2 and the stopping rules,

$$|\langle Tb^1_{S_1}\mathbf{1}_{S_1\setminus 3P'}, \tilde{\psi}^2\rangle| \lesssim \inf_{x\in P'} Mb^1_{S_1} \int |\tilde{\psi}^2| \,\mathrm{d}x \lesssim |P|.$$

This completes the proof of Lemma 5.4. □

Acknowledgements. The authors thank the referee for useful comments.

M.T.L. was supported in part by Grant NSF-DMS 0968499 and a grant from the Simons Foundation (Grant #229596). A.V.V. was supported by the School of Mathematics, Georgia Institute of Technology, and by the Finnish Academy of Science and Letters, Vilho, Yrjö and Kalle Väisälä Foundation.

https://doi.org/10.1017/S0013091514000340 Published online by Cambridge University Press

References

1. P. Auscher and E. Routin, Local Tb theorems and Hardy inequalities, *J. Geom. Analysis* **23**(1) (2013), 303–374.
2. P. Auscher and Q. X. Yang, BCR algorithm and the $T(b)$ theorem, *Publ. Mat.* **53**(1) (2009), 179–196.
3. P. Auscher, S. Hofmann, C. Muscalu, T. Tao and C. Thiele, Carleson measures, trees, extrapolation, and $T(b)$ theorems, *Publ. Mat.* **46**(2) (2002), 257–325.
4. G. Beylkin, R. Coifman and V. Rokhlin, Fast wavelet transforms and numerical algorithms, I, *Commun. Pure Appl. Math.* **44**(2) (1991), 141–183.
5. D. L. Burkholder, Explorations in martingale theory and its applications, in *École d'Été de Probabilités de Saint-Flour XIX—1989*, Lecture Notes in Mathematics, Volume 1464, pp. 1–66 (Springer, 1991).
6. M. Christ, A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral, *Colloq. Math.* **60–61**(2) (1990), 601–628.
7. G. David and J.-L. Journé, A boundedness criterion for generalized Calderón–Zygmund operators, *Annals Math. (2)* **120**(2) (1984), 371–397.
8. T. Figiel, Singular integral operators: a martingale approach, in *Geometry of Banach Spaces, Strobl, 1989*, London Mathematical Society Lecture Note Series, Volume 158, pp. 95–110 (Cambridge University Press, 1990).
9. S. Hofmann, A proof of the local Tb Theorem for standard Calderón-Zygmund operators, preprint (arXiv.org/abs/0705.0840, 2007).
10. S. Hofmann, Local $T(b)$ theorems and applications in PDE, in *Harmonic analysis and partial differential equations*, Contemporary Mathematics, Volume 505, pp. 29–52 (American Mathematical Society, Providence, RI, 2010).
11. T. P. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators, *Annals Math. (2)* **175**(3) (2012), 1473–1506.
12. T. Hytönen and H. Martikainen, Non-homogeneous Tb theorem and random dyadic cubes on metric measure spaces, *J. Geom. Analysis* **22**(4) (2012), 1071–1107.
13. T. Hytönen and H. Martikainen, On general local Tb theorems, *Trans. Am. Math. Soc.* **364**(9) (2012), 4819–4846.
14. T. Hytönen and F. Nazarov, The local Tb theorem with rough test functions, preprint (arXiv.org/abs/1206.0907, 2012).
15. T. P. Hytönen and A. V. Vähäkangas, The local non-homogeneous Tb theorem for vector-valued functions, *Glasgow Math. J.* (2014), doi: 10.1017/S0017089514000123.
16. M. T. Lacey and H. Martikainen, Local Tb theorem with L2 testing conditions and general measures: square functions, preprint (arXiv.org/abs/1308.4571, 2013).
17. M. T. Lacey and A. V. Vähäkangas, Non-homogeneous local $T1$ theorem: dual exponents, preprint (arXiv.org/abs/1301.5858, 2013).
18. M. T. Lacey and A. V. Vähäkangas, The perfect local Tb theorem and twisted martingale transforms, *Proc. Am. Math. Soc.* **142**(5) (2014), 1689–1700.
19. M. T. Lacey, E. T. Sawyer, C.-Y. Shen and I. Uriarte-Tuero, The two weight inequality for Hilbert transform, coronas, and energy conditions, preprint (arXiv.org/abs/1108.2319, 2011).
20. M. T. Lacey, E. T. Sawyer, C.-Y. Shen and I. Uriarte-Tuero, Two weight inequality for the Hilbert transform: a real variable characterization, I, *Duke Math. J.* **163**(15) (2014), 2795–2820.
21. F. Nazarov, S. Treil and A. Volberg, Accretive system Tb-theorems on nonhomogeneous spaces, *Duke Math. J.* **113**(2) (2002), 259–312.
22. F. Nazarov, S. Treil and A. Volberg, The Tb-theorem on non-homogeneous spaces, *Acta Math.* **190**(2) (2003), 151–239.

23. F. Nazarov, S. Treil and A. Volberg, Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures, preprint (arXiv.org/abs/1003.1596, 2004).
24. S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, *C. R. Acad. Sci. Paris I* **330**(6) (2000), 455–460.
25. A. Volberg, *Calderón–Zygmund capacities and operators on nonhomogeneous spaces*, CBMS Regional Conference Series in Mathematics, Volume 100 (American Mathematical Society/CBMS, 2003).