ON GENERAL LOCAL Tb THEOREMS

TUOMAS HYTÖNEN AND HENRI MARTIKAINEN

ABSTRACT. In this paper, local Tb theorems are studied both in the doubling and non-doubling situation. We prove a local Tb theorem for the class of upper doubling measures. With such general measures, scale invariant testing conditions are required (L^∞ or BMO). In the case of doubling measures, we also modify the general non-homogeneous method of proof to yield a new proof of the local Tb theorem with L^2 type testing conditions.

1. INTRODUCTION

There are multiple local Tb theorems with a bit different assumptions. In these theorems, one assumes that to every cube Q there exist functions b_Q^1 and b_Q^2, supported on Q, so that we also know something about Tb_Q^1 and $T^*b_Q^2$ (where T is a Calderón–Zygmund operator). One wishes to conclude that $T\colon L^2 \to L^2$ boundedly so that $\|T\|$ has natural dependence on the assumptions.

The first local Tb theorem is by Michael Christ [Chr90], and there it was assumed that $\|b_Q^1\|_\infty \leq C$ and $\|Tb_Q^1\|_\infty \leq C$ (and similarly for b_Q^2 and $T^*b_Q^2$). This was proven for doubling measures (even in metric spaces). Nazarov, Treil and Volberg [NTV02] obtained a version of this theorem for measures satisfying the power bound $\mu(B(x,r)) \lesssim r^m$ for a given number m. So it is a non-homogeneous version of Christ's theorem in $\mathbb{R}^n$ (it also allows BMO control in the operator side if the kernel of T is antisymmetric).

For doubling measures, one can also consider more general L^p type testing conditions introduced by Auscher, Hofmann, Muscalu, Tao and Thiele [AHM$^+$02], and further studied by Hofmann [Hof07], Auscher and Yang [AY09] and Tan and Yan [TY09]. The most general assumption used in these papers is of the form that $\int_Q |b_Q^1|^p \leq |Q|$, $\int_Q |b_Q^2|^q \leq |Q|$, $\int_Q |Tb_Q^1|^{q'} \leq |Q|$ and $\int_Q |T^*b_Q^2|^{p'} \leq |Q|$, where s' denotes the dual exponent of s and $1 < p, q \leq \infty$.

In [AHM$^+$02] a theorem of this type is proved only for very special operators, the so-called perfect dyadic singular integral operators. This was expected to easily generalize for all Calderón–Zygmund operators – but this turned out not to be the case (it being easy, at least). In [Hof07] the theorem is extended for standard

Calderón–Zygmund operators, but only in the case $\int_Q |b_Q^1|^s \leq |Q|$, $\int_Q |b_Q^2|^s \leq |Q|$, $\int_Q |Tb_Q^1|^2 \leq |Q|$ and $\int_Q |T^*b_Q^2|^2 \leq |Q|$ for some $s > 2$. Finally, [AY09] establishes, by reducing the question to the known case of perfect dyadic operators, the theorem for standard Calderón–Zygmund operators in the case $1/p + 1/q \leq 1$. The case $1/p + 1/q > 1$ is still open for general Calderón–Zygmund operators.

Recently we extended global Tb theorems to the general setting of upper doubling measures in metric spaces [HM09]. It is our opinion that the upper doubling theory constitutes a flexible framework yielding, in particular, proofs that work simultaneously for doubling and non-doubling measures. See also [Hyt10a], [Mar10], [HYY10].

The purpose of this paper is twofold. We extend, streamline and modify the general non-homogeneous proof technique of Nazarov, Treil and Volberg [NTV02] to the case of upper doubling measures μ. Also, considering doubling measures ν, we show that this technique, appropriately modified, can also be used to prove local Tb theorems with L^p type testing conditions. As the general case $1/p + 1/q > 1$ is still quite open, a completely new approach may be appreciated. We give such a proof in the case of $\int_Q |b_Q^1|^2 \, d\nu \leq \nu(Q)$, $\int_Q |b_Q^2|^2 \, d\nu \leq \nu(Q)$, $\int_Q |Tb_Q^1|^s \, d\nu \leq \nu(Q)$ and $\int_Q |T^*b_Q^2|^s \, d\nu \leq \nu(Q)$ for any $s > 2$. So this is basically the theorem of Hofmann [Hof07], but with a general doubling measure and with the extra integrability assumptions pushed into the operator side, allowing to demand less integrability from the test functions.

Regarding non-homogeneous analysis in this local situation, there seems to be a previously unnoticed problem with the use of goodness and the implication it may have on the collapse of certain paraproducts. In any case, we add the good cubes into the decomposition in a new way so that no such problem may arise.

While giving finishing touches to this paper, we also learned about a recent related manuscript by Auscher and Routin [AR10]. Using the local $T1$ theorem, the so-called BCR algorithm, and Hardy inequalities, some partial progress on the case $1/p + 1/q > 1$ is achieved there. However, many technical assumptions appear. Furthermore, they also establish a direct proof (in the sense that it is not a reduction to the perfect dyadic case like [AY09]) of the case $1/p + 1/q \leq 1$.

Our proof of the local Tb theorem with L^2 test functions is also direct in many senses. As it uses this general non-homogeneous proof technique, it does not rely on the standard local $T1$ theorem unlike all the other known proofs seem to do. Also, it is not a reduction to the perfect dyadic case. The use of the Hardy type inequalities is completely replaced by the use of non-homogeneous analysis. Such techniques also circumvent many problems with nearby cubes and boundary regions.

2. DEFINITIONS AND THE MAIN RESULT

2.A. **Upper doubling measures and Calderón–Zygmund operators.** Let $\lambda \colon \mathbb{R}^n \times (0, \infty) \to (0, \infty)$ be a function so that $r \mapsto \lambda(x, r)$ is non-decreasing and $\lambda(x, 2r) \le C_\lambda \lambda(x, r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. Let μ be a Borel measure in $\mathbb{R}^n$. We assume that μ is upper doubling with the dominating function λ, that is, $\mu(B(x, r)) \le \lambda(x, r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. In the case of doubling measures one can take $\lambda(x, r) = \mu(B(x, r))$, and in the case of power bounded measures ($\mu(B(x, r)) \le Cr^m$), one can take $\lambda(x, r) = Cr^m$. Let $d = \log_2 C_\lambda$ – this is a convenient number for us, and can be thought of as a dimension of the measure μ.

The kernel estimates are always tied to the particular choice of λ. We say that $K \colon \mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, y) : x = y\} \to \mathbb{C}$ is a standard kernel if there holds for some $C < \infty$ and $\alpha > 0$ that

$$|K(x, y)| \le C \min \left(\frac{1}{\lambda(x, |x - y|)}, \frac{1}{\lambda(y, |x - y|)} \right), \qquad x \ne y,$$

$$|K(x, y) - K(x', y)| \le C \frac{|x - x'|^\alpha}{|x - y|^\alpha \lambda(x, |x - y|)}, \qquad |x - y| \ge 2|x - x'|,$$

and

$$|K(x, y) - K(x, y')| \le C \frac{|y - y'|^\alpha}{|x - y|^\alpha \lambda(y, |x - y|)}, \qquad |x - y| \ge 2|y - y'|.$$

These are the familiar standard estimates in the case of doubling and power bounded measures. Let $\gamma = \alpha/(2\alpha + 2d)$ – another convenient number.

Sometimes the property $\lambda(x, |x - y|) \sim \lambda(y, |x - y|)$ would be convenient. This can be arranged as follows. In [HYY10, Proposition 1.1] it is shown that $\Lambda(x, r) := \inf_{z \in \mathbb{R}^n} \lambda(z, r + |x - z|)$ satisfies that $r \mapsto \Lambda(x, r)$ is non-decreasing, $\Lambda(x, 2r) \le C_\lambda \Lambda(x, r), \mu(B(x, r)) \le \Lambda(x, r), \Lambda(x, r) \le \lambda(x, r)$ and $\Lambda(x, r) \le C_\lambda \Lambda(y, r)$ if $|x - y| \le r$. Even the kernel estimates hold with Λ, since $1/\lambda \le 1/\Lambda$. Thus, we may (and do) assume that λ satisfies the additional symmetry property $\lambda(x, r) \le C\lambda(y, r)$ if $|x - y| \le r$, and then demand the kernel estimates in the form

$$|K(x, y)| \le \frac{C}{\lambda(x, |x - y|)}, \qquad x \ne y,$$

$$|K(x, y) - K(x', y)| \le C \frac{|x - x'|^\alpha}{|x - y|^\alpha \lambda(x, |x - y|)}, \qquad |x - y| \ge 2|x - x'|,$$

and

$$|K(x, y) - K(x, y')| \le C \frac{|y - y'|^\alpha}{|x - y|^\alpha \lambda(x, |x - y|)}, \qquad |x - y| \ge 2|y - y'|.$$

A Calderón–Zygmund operator with a standard kernel K is a bounded linear operator T taking $L^2(\mu)$ into $L^2(\mu)$ so that there holds

$$Tf(x) = \int K(x, y) f(y) \, d\mu(y)$$

for x not in the support of f.

Note that while we assume the boundedness of T a priori, we are interested in quantitative bounds for $\|T\|$, which only depend on some specified information.

2.B. **Systems of accretive functions.** When working with a general upper doubling measure μ, we assume that to every cube $Q \subset \mathbb{R}^n$ there exist two functions b_Q^1 and b_Q^2 so that there holds

(i) $\operatorname{spt} b_Q^1 \subset Q$, $\operatorname{spt} b_Q^2 \subset Q$;
(ii) $\|b_Q^1\|_{L^\infty(\mu)} \leq C$, $\|b_Q^2\|_{L^\infty(\mu)} \leq C$;
(iii) $\|Tb_Q^1\|_{L^\infty(\mu)} \leq C$, $\|T^*b_Q^2\|_{L^\infty(\mu)} \leq C$;
(iv) $\int_Q b_Q^1 \, d\mu = \mu(Q) = \int_Q b_Q^2 \, d\mu$.

We call these *accretive L^∞ systems*. At least in the case of an antisymmetric kernel, one could make do with BMO control in the operator side (see [NTV02]), but we focus only on this case.

When working with a doubling measure ν, we may also use the following set of assumptions: to every cube $Q \subset \mathbb{R}^n$ there exist two functions b_Q^1 and b_Q^2 so that there holds

(i) $\operatorname{spt} b_Q^1 \subset Q$, $\operatorname{spt} b_Q^2 \subset Q$;
(ii) $\int_Q |b_Q^1|^2 \, d\nu \leq C\nu(Q)$, $\int_Q |b_Q^2|^2 \, d\nu \leq C\nu(Q)$;
(iii) $\int_Q |Tb_Q^1|^s \, d\nu \leq C\nu(Q)$, $\int_Q |T^*b_Q^2|^s \, d\nu \leq C\nu(Q)$ for some fixed $s > 2$;
(iv) $\int_Q b_Q^1 = \nu(Q) = \int_Q b_Q^2$.

We call these *accretive L^2 systems* (suppressing from the name the fact that we actually impose the somewhat stronger L^s conditions in (iii)).

We now formulate our main theorem.

2.1. **Theorem.** *Let μ be an upper doubling measure with a dominating function λ and $T \colon L^2(\mu) \to L^2(\mu)$ a Calderón–Zygmund operator with a standard kernel K. Assuming the existence of accretive L^∞ systems (b_Q^1) and (b_Q^2), we have $\|T\| \leq C$, where C depends on the dimension n and on the explicit constants in the definitions of λ, K, (b_Q^1) and (b_Q^2).*

If $\mu = \nu$ for some doubling measure ν, then the same conclusion holds assuming only the existence of accretive L^2 systems (b_Q^1) and (b_Q^2).

The rest of this paper contains a direct proof which simultaneously gives the theorem with either set of assumptions. In particular, the proof is neither a reduction to a local $T1$ theorem, nor to a perfect dyadic case. Notation-wise we work so that we use μ as long as everything works with the use of either set of assumptions, and sometimes write $\mu = \nu$ when we explicitly estimate differently in the doubling L^2 case. We write $X \lesssim Y$ to mean $X \leq CY$ with some constant C like in the theorem. Also, $X \sim Y$ means $Y \lesssim X \lesssim Y$. Sometimes we absorb other parameters, but then it is either explicitly said or written in the notation (e.g. $X \lesssim_\delta Y$ would mean $X \leq C(\delta)Y$).

2.2. *Remark.* While the second part of the theorem concerning the L^2 test function case is not new, the proof is. Certainly some new ideas are still needed to establish the theorem with general p and q. However, the point is not solely in the range of exponents. For example, we point out that the non-homogeneous proof technique completely avoids the use of the so called Hardy type inequalities used and studied in [AR10].

3. PRELIMINARIES

We begin by recording the following basic facts. Let a dyadic system $\mathcal{D}$ be given. The side length of a cube $Q \in \mathcal{D}$ is denoted $\ell(Q)$, and $Q^{(j)}$ denotes the unique cube $S \in \mathcal{D}$ for which $Q \subset S$ and $\ell(S) = 2^j \ell(Q)$. We also set $\langle f \rangle_Q = \mu(Q)^{-1} \int_Q f \, d\mu$.

3.1. **Definition.** We say that a sequence $(a_Q)_{Q \in \mathcal{D}}$ of positive numbers is a Carleson sequence, if there holds

$$\sum_{\substack{S \in \mathcal{D} \\ S \subset Q}} a_S \leq C \mu(Q)$$

for every $Q \in \mathcal{D}$. The condition is called the Carleson (measure) condition.

The following is the famous Carleson embedding theorem.

3.2. **Theorem.** *Given a Carleson sequence (a_Q) there holds for any $f \in L^2(\mu)$ that*

$$\sum_{Q \in \mathcal{D}} a_Q |\langle f \rangle_Q|^2 \leq C \|f\|_{L^2(\mu)}^2.$$

The following is called the unweighted square function estimate.

3.3. **Theorem.** *There holds for any $f \in L^2(\mu)$ that*

$$\sum_{Q \in \mathcal{D}} |\langle f \rangle_Q - \langle f \rangle_{Q^{(1)}}|^2 \mu(Q) \leq C \|f\|_{L^2(\mu)}^2.$$

3.A. **Stopping times and the martingale difference operators** Δ_Q. Let $\mathcal{D}$ be a dyadic system of cubes, and let $Q_0 \in \mathcal{D}$ be a fixed large cube. Let $\mathcal{D}^0 = \{Q_0\}$.

3.A.1. *Stopping time: L^∞ case.* Let $\mathcal{D}^1 = \{Q_1^k\}_k$ consist of the maximal $\mathcal{D}$-cubes $Q \subset Q_0$ for which there holds

$$\left| \int_Q b_{Q_0}^1 \, d\mu \right| < \mu(Q)/2.$$

One easily checks that

$$\mu \Big(\bigcup_k Q_1^k \Big) \leq \tau \mu(Q_0)$$

for some $\tau < 1$.

Next, one fixes a cube Q_1^k and considers all the maximal $\mathcal{D}$-cubes $Q \subset Q_1^k$ for which there holds

$$\left| \int_Q b_{Q_1^k}^1 \, d\mu \right| < \mu(Q)/2.$$

One does this for every $Q_1^k \in \mathcal{D}^1$, and then the resulting collection of cubes is called $\mathcal{D}^2 = \{Q_2^k\}_k$. One proceeds like this to obtain collections $\mathcal{D}^j$ for every j. Of course, we have the property that for every $Q \in \mathcal{D}^j$ there holds

$$\mu\Big(\bigcup_{Q' \in \mathcal{D}^{j+1}, \, Q' \subset Q} Q' \Big) \leq \tau \mu(Q).$$

3.A.2. *Stopping time: L^2 case.* Define

$$M_\nu f(x) = \sup_{r>0} \frac{1}{\nu(B(x,r))} \int_{B(x,r)} |f(y)| \, d\nu(y).$$

Let $\mathcal{D}^1 = \{Q_1^k\}_k$ consist of the maximal $\mathcal{D}$-cubes $Q \subset Q_0$ for which there holds

$$\int_Q |M_\nu b_{Q_0}^1|^2 \, d\nu > \delta^{-1} \nu(Q)$$

or

$$\int_Q |T b_{Q_0}^1|^s \, d\nu > \delta^{-1} \nu(Q)$$

or

$$\left| \int_Q b_{Q_0}^1 \, d\nu \right| < \delta \nu(Q).$$

Fixing δ to be small enough, one easily checks that

$$\nu\Big(\bigcup_k Q_1^k \Big) \leq \tau \nu(Q_0)$$

for some $\tau < 1$. This is then continued just like in the L^∞ case.

3.A.3. *Martingale difference operators.* For every $Q \subset Q_0$ we let Q^a be the smallest cube in the family $\bigcup \mathcal{D}^j$ containing Q. Note that if $Q \subset Q_0$ is such that $Q^a \in \mathcal{D}^t$, there holds for every $j \geq 1$ that

$$\mu\Big(\bigcup_{Q' \in \mathcal{D}^{t+j}, \, Q' \subset Q} Q' \Big) = \sum_{Q' \in \mathcal{D}^{t+j}, \, Q' \subset Q} \mu(Q') \leq \tau^{j-1} \mu(Q).$$

We state a very useful (but immediate) consequence of this as a lemma.

3.4. **Lemma.** *The following is a Carleson sequence: $\alpha_Q = 0$ if Q is not from $\bigcup_j \mathcal{D}^j$, and it equals $\mu(Q)$ otherwise.*

Given a cube Q let $\mathrm{ch}(Q)$ consist of those cubes $Q' \subset Q$ for which $\ell(Q') = \ell(Q)/2$. Let f be a function supported on Q_0. We define

$$\Delta_Q f = \sum_{Q' \in \mathrm{ch}(Q)} \Big[\frac{\langle f \rangle_{Q'}}{\langle b^1_{(Q')^a} \rangle_{Q'}} b^1_{(Q')^a} - \frac{\langle f \rangle_Q}{\langle b^1_{Q^a} \rangle_Q} b^1_{Q^a} \Big] \chi_{Q'}.$$

Note that then we have

$$(\Delta_Q)^* f = \sum_{Q' \in \mathrm{ch}(Q)} \Big[\frac{\langle b^1_{(Q')^a} f \rangle_{Q'}}{\langle b^1_{(Q')^a} \rangle_{Q'}} - \frac{\langle b^1_{Q^a} f \rangle_Q}{\langle b^1_{Q^a} \rangle_Q} \Big] \chi_{Q'}.$$

Also set

$$E_{Q_0} = \frac{\langle f \rangle_{Q_0}}{\langle b^1_{Q_0} \rangle_{Q_0}} b^1_{Q_0}.$$

3.5. **Lemma.** *The identity*

(3.6) $$f = E_{Q_0} f + \sum_{Q \in \mathcal{D}} \Delta_Q f = E_{Q_0} f + \lim_{k \to \infty} \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) > 2^{-k}}} \Delta_Q f$$

holds both pointwise almost everywhere and in $L^2(\mu)$.

Proof. First, few additional notations. Set $b^{a,1}_k = \sum_{Q \in \mathcal{D}_k} \chi_Q b^1_{Q^a}$ and

$$E^{a,1}_k f = \frac{E_k f}{E_k b^{a,1}_k} b^{a,1}_k,$$

where naturally $E_k h = \sum_{Q \in \mathcal{D}_k} \chi_Q \langle h \rangle_Q$. It is immediate to see that the right hand side of (3.6), for a fixed k, is precisely $E^{a,1}_k f$.

It follows from the stopping time construction that almost every $x \in Q_0$ belongs to only finitely many stopping cubes $P \in \bigcup_{t=0}^{\infty} \mathcal{D}^t$. If S is the smallest of them, then $Q^a = S$ for all $Q \ni x$ with $\ell(Q) = 2^{-k} \leq \ell(S)$. Thus $b^{a,1}_k(x) = b^1_S(x)$ and $E_k b^{a,1}_k(x) = \langle b^1_S \rangle_Q = E_k b^1_S(x) \to b^1_S(x)$ as $k \to \infty$ (this happens almost everywhere as the set $\bigcup_{t=0}^{\infty} \mathcal{D}^t$ is countable). Since also $E_k f(x) \to f(x)$ almost everywhere, we have verified the pointwise convergence $E^{a,1}_k f \to f$. In the case of accretive L^∞ systems, the $L^2(\mu)$ convergence is immediate from dominated convergence, since $|E^{a,1}_k f| \lesssim Mf$, where M is the dyadic maximal operator.

It remains to prove that $E^{a,1}_k f \to f$ in $L^2(\nu)$ in the case of accretive L^2 systems. Note that $\|E^{a,1}_k f\|_2 \lesssim \|f\|_2$. Thus, it suffices to prove the convergence for a given bounded function f. As the convergence is in any case fine in the pointwise almost everywhere sense, we just need to find a suitable square integrable majorant. And we have

$$|E^{a,1}_k f(x)| \lesssim_f \sup_{Q = Q^a} |b^1_Q(x)| \leq \Big(\sum_{Q = Q^a} |b^1_Q(x)|^2 \Big)^{1/2},$$

and this is in L^2 since

$$\int \sum_{Q=Q^a} |b_Q^1|^2 \, d\nu = \sum_{Q=Q^a} \int_Q |b_Q^1|^2 \, d\nu \lesssim \sum_{Q=Q^a} \nu(Q) \lesssim \nu(Q_0)$$

by Lemma 3.4.   □

We are usually given two dyadic systems $\mathcal{D}$ and $\mathcal{D}'$. Then we use operators Δ_Q constructed using (b_Q^1) in connection with the family $\mathcal{D}$ and operators Δ_R constructed using (b_R^2) in connection with the family $\mathcal{D}'$ (in the L^2 case the stopping time for the latter also uses T^* instead of T, of course). It would perhaps be better to write Δ_Q^1 and Δ_R^2 to indicate the difference (as we have done above for some operators that we need not use so frequently), but we omit this for brevity. It should nevertheless be clear from the various summing conditions like $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$.

3.B. **Square function estimates.** With accretive L^∞ systems, the estimates

$$\sum_{Q \in \mathcal{D}} \|\Delta_Q f\|_2^2 \lesssim \|f\|_2^2 \qquad \text{and} \qquad \sum_{Q \in \mathcal{D}} \|(\Delta_Q)^* f\|_2^2 \lesssim \|f\|_2^2$$

are quite clear (see [NTV02, chapter 3]). So in the rest of this subsection, we work with a doubling measure ν and L^2 type test functions (the second estimate is actually, perhaps surprisingly, generally false in this setting).

3.7. **Lemma.** *The sequence*

$$\beta_Q = |\langle b_{Q^a}^1 \rangle_Q - \langle b_{Q^a}^1 \rangle_{Q^{(1)}}|^2 \nu(Q), \qquad Q \in \mathcal{D},$$

is Carleson.

Proof. Let $Q \in \mathcal{D}$ be such that $Q^a \in \mathcal{D}^t$. We simply write as follows

$$\sum_{S \subset Q} \beta_S = \sum_{S \subset Q} |\langle b_{S^a}^1 \rangle_S - \langle b_{S^a}^1 \rangle_{S^{(1)}}|^2 \nu(S)$$

$$= \sum_{\substack{S \subset Q \\ S^a = Q^a}} |\langle b_{Q^a}^1 \rangle_S - \langle b_{Q^a}^1 \rangle_{S^{(1)}}|^2 \nu(S)$$

$$+ \sum_{j=1}^\infty \sum_{H \in \mathcal{D}^{t+j}, H \subset Q} \sum_{\substack{S \subset H \\ S^a = H}} |\langle b_H^1 \rangle_S - \langle b_H^1 \rangle_{S^{(1)}}|^2 \nu(S)$$

$$\lesssim \|1_{Q^{(1)}} b_{Q^a}^1\|_2^2 + \sum_{j=1}^\infty \sum_{H \in \mathcal{D}^{t+j}, H \subset Q} \|b_H^1\|_2^2$$

$$\lesssim \nu(Q) + \sum_{j=1}^\infty \sum_{H \in \mathcal{D}^{t+j}, H \subset Q} \nu(H) \lesssim \nu(Q)$$

by the unweighted square function estimate (Theorem 3.3) and Lemma 3.4.   □

3.8. **Proposition.** *There holds*

$$\sum_{Q \in \mathcal{D}} \|\Delta_Q f\|_2^2 \lesssim \|f\|_2^2.$$

Proof. Note that

$$\sum_{Q \in \mathcal{D}} \|\Delta_Q f\|_2^2 = I + II,$$

where

$$I = \sum_{\substack{Q \in \mathcal{D} \\ Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q'}} \int_{Q'} \left| \frac{\langle f \rangle_{Q'}}{\langle b_{Q'}^1 \rangle_{Q'}} b_{Q'}^1 - \frac{\langle f \rangle_Q}{\langle b_{Q^a}^1 \rangle_Q} b_{Q^a}^1 \right|^2 d\nu,$$

$$II = \sum_{\substack{Q \in \mathcal{D} \\ Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q^a}} \int_{Q'} \left| \frac{\langle f \rangle_{Q'}}{\langle b_{Q^a}^1 \rangle_{Q'}} - \frac{\langle f \rangle_Q}{\langle b_{Q^a}^1 \rangle_Q} \right|^2 |b_{Q^a}^1|^2 d\nu.$$

Furthermore, there holds (as ν is doubling) that

$$I \lesssim \sum_{\substack{Q \in \mathcal{D} \\ Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q'}} \left[|\langle f \rangle_{Q'}|^2 \nu(Q') + |\langle f \rangle_{Q'} - \langle f \rangle_Q|^2 \nu(Q') \right] \lesssim \|f\|_2^2.$$

Here we used Lemma 3.4 to bound the first term by $\|f\|_2^2$ (the bound for the second term follows from the unweighted square function estimate, Theorem 3.3).

Next, note that

$$II \lesssim \sum_{\substack{Q \in \mathcal{D} \\ Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q^a}} \left[|\langle f \rangle_{Q'}|^2 |\langle b_{Q^a}^1 \rangle_{Q'} - \langle b_{Q^a}^1 \rangle_Q|^2 \nu(Q') + |\langle f \rangle_{Q'} - \langle f \rangle_Q|^2 \nu(Q') \right].$$

The latter term is yet again bounded by $\|f\|_2^2$ by the unweighted square function estimate (Theorem 3.3), and the first one is, too, bounded by $\|f\|_2^2$ by the previous lemma. $\qquad\square$

The following example is a bit disconcerting. After all, we want to work with accretive L^2 systems of functions, and the failure of such a fundamental estimate seems like a real predicament. A weaker, but sufficient for us, substitute result is offered afterwards.

3.9. **Example.** The estimate

$$\sum_{Q \in \mathcal{D}} \|(\Delta_Q)^* f\|_2^2 \lesssim \|f\|_2^2$$

is not, in general, true for accretive L^2 systems.

Proof. Consider the one-dimensional situation with $Q_0 = [0, 1)$ and $N \in \mathbb{Z}_+$ a fixed but arbitrary parameter. We construct a sequence of examples, where the constant in the dual square function estimate grows without limit as a function of N. Let

$$b_{[0,2^{-j})} := b_j := 2^{(N-j)/2}\chi_{[0,2^{-N})} + \chi_{[2^{-N},2^{-j})}, \quad j = 0, 1, \dots, N,$$

and $b_Q := \chi_Q$ for all other dyadic intervals Q. They satisfy $|Q|^{-1}\int_Q |b_Q|^2 dx \le 2$, and the accretivity of these functions is not an issue; however, the normalized L^2 norm of $b_{[0,2^{-j})}$ on $[0, 2^{-k})$ will increase as k increases. With a suitable choice of the stopping parameters, it follows that the stopping cubes are precisely all the $Q_j := [0, 2^{-j})$, $j = 0, 1, \dots, N$. In particular,

$$\chi_{Q_j}(\Delta_{Q_{j-1}})^* f = \chi_{Q_j}\Big(\frac{\langle b_j f\rangle_{Q_j}}{\langle b_j\rangle_{Q_j}} - \frac{\langle b_{j-1} f\rangle_{Q_{j-1}}}{\langle b_{j-1}\rangle_{Q_{j-1}}}\Big), \quad j = 1, \dots, N.$$

We apply this to the function $f = 2^{N/2}\chi_{Q_N}$ for which

$$\frac{\langle b_j f\rangle_{Q_j}}{\langle b_j\rangle_{Q_j}} = \frac{2^{j/2}}{1 + 2^{(j-N)/2} - 2^{j-N}},$$

yielding, by a simple computation,

$$\chi_{Q_j}(\Delta_{Q_{j-1}})^* f \ge c 2^{j/2}\chi_{Q_j}.$$

Since $\|\chi_{Q_j}\|_2^2 = 2^{-j}$ and $\|f\|_2 = 1$, it follows that

$$\sum_{Q \in \mathcal{D}} \|(\Delta_Q)^* f\|_2^2 \ge \sum_{j=1}^N \|\chi_{Q_j}(\Delta_{Q_{j-1}})^* f\|_2^2 \ge \sum_{j=1}^N c = cN = cN\|f\|_2^2,$$

and this proves the impossibility of the dual square function estimate. $\qquad\square$

The following weaker estimate is, however, true and still useful.

3.10. **Proposition.** *For general accretive L^2 systems, there holds*

$$\sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \|(\Delta_Q)^* f\|_2^2 \lesssim \|\chi_P f\|_2^2.$$

Proof. We write

$$\sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \|(\Delta_Q)^* f\|_2^2 = \sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{Q' \in \mathrm{ch}(Q)} \Big|\frac{\langle b^1_{(Q')^a} f\rangle_{Q'}}{\langle b^1_{(Q')^a}\rangle_{Q'}} - \frac{\langle b^1_{Q^a} f\rangle_Q}{\langle b^1_{Q^a}\rangle_Q}\Big|^2 \nu(Q') \lesssim I + II,$$

where

$$I = \sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{\substack{Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q'}} [|\langle b^1_{Q'} f\rangle_{Q'}|^2 + |\langle b^1_{Q^a} f\rangle_{Q'}|^2 + |\langle b^1_{Q^a} f\rangle_{Q'} - \langle b^1_{Q^a} f\rangle_Q|^2]\nu(Q')$$

and

$$II = \sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{\substack{Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q^a}} [|\langle b^1_{Q^a} f \rangle_{Q'}|^2 |\langle b^1_{Q^a} \rangle_{Q'} - \langle b^1_{Q^a} \rangle_Q|^2 + |\langle b^1_{Q^a} f \rangle_{Q'} - \langle b^1_{Q^a} f \rangle_Q|^2] \nu(Q').$$

Let t be such that $P \in \mathcal{D}^t$. Note that

$$|\langle b^1_{Q'} f \rangle_{Q'}|^2 \nu(Q') \leq \frac{1}{\nu(Q')} \Big(\int_{Q'} |b^1_{Q'}|^2 \, d\nu \Big) \Big(\int_{Q'} |f|^2 \, d\nu \Big) \lesssim \int_{Q'} |f|^2 \, d\nu$$

and (as ν is doubling) that

$$|\langle b^1_P f \rangle_{Q'}|^2 \nu(Q') \leq \frac{1}{\nu(Q')} \Big(\int_Q |b^1_P|^2 \, d\nu \Big) \Big(\int_{Q'} |f|^2 \, d\nu \Big) \lesssim \int_{Q'} |f|^2 \, d\nu$$

showing that

$$\sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{\substack{Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q'}} [|\langle b^1_{Q'} f \rangle_{Q'}|^2 + |\langle b^1_P f \rangle_{Q'}|^2] \nu(Q') \lesssim \sum_{\substack{Q' \in \mathcal{D}^{t+1} \\ Q' \subset P}} \int_{Q'} |f|^2 \, d\nu \leq \|\chi_P f\|_2^2.$$

Here we used the fact that $Q' \in \mathcal{D}^{t+1}$ are disjoint.

For the rest of the terms, we write

$$\langle b^1_P f \rangle_J = \Big\langle \chi_{P \setminus \cup \mathcal{D}^{t+1}} b^1_P f \Big\rangle_J + \Big\langle \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} \chi_S \langle b^1_P f \rangle_S \Big\rangle_J$$

for $J = Q$ or $J = Q'$, where $Q^a = P$. Recalling Lemma 3.7 and the unweighted square function estimate, Theorem 3.3, we have that

$$\sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{\substack{Q' \in \mathrm{ch}(Q) \\ (Q')^a = Q^a}} |\langle b^1_P f \rangle_{Q'}|^2 |\langle b^1_{Q^a} \rangle_{Q'} - \langle b^1_{Q^a} \rangle_Q|^2 \nu(Q')$$

$$+ \sum_{\substack{Q \in \mathcal{D} \\ Q^a = P}} \sum_{Q' \in \mathrm{ch}(Q)} |\langle b^1_P f \rangle_{Q'} - \langle b^1_P f \rangle_Q|^2 \nu(Q')$$

is dominated by

$$\|\chi_{P \setminus \cup \mathcal{D}^{t+1}} b^1_P f\|_2^2 + \Big\| \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} \chi_S \langle b^1_P f \rangle_S \Big\|_2^2 \lesssim \|\chi_P f\|_2^2,$$

where the last estimate follows since on $P \setminus \bigcup \mathcal{D}^{t+1}$ we have L^∞-control of b_P^1 by Lebesgue's differentiation theorem, and

$$
\begin{aligned}
\Big\| \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} \chi_S \langle b_P^1 f \rangle_S \Big\|_2^2 &= \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} |\langle b_P^1 f \rangle_S|^2 \nu(S) \\
&\leq \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} \frac{1}{\nu(S)} \Big(\int_{S^{(1)}} |b_P^1|^2 \, d\nu \Big) \Big(\int_S |f|^2 \, d\nu \Big) \\
&\lesssim \sum_{\substack{S \in \mathcal{D}^{t+1} \\ S \subset P}} \int_S |f|^2 \, d\nu \leq \|\chi_P f\|_2^2.
\end{aligned}
$$

$\square$

3.11. *Remark.* The stronger estimate

$$
\sum_{Q \in \mathcal{D}} \|(\Delta_Q)^* f\|_2^2 \lesssim \|f\|_2^2
$$

is true if our test functions satisfy $\int_Q |b_Q^1|^q \, d\nu \lesssim \nu(Q)$ for some $q > 2$ (and the stopping time argument is modified to use this condition, of course). The point is that then one can cope with summing over the multiple generations of $\mathcal{D}^j$ because of the better estimate $|\langle b_{Q'}^1 f \rangle_{Q'}|^2 + |\langle b_{Q^a}^1 f \rangle_{Q'}|^2 \lesssim |\langle |f|^p \rangle_{Q'}|^{2/p}$ for $p = q' < 2$ (the Hardy–Littlewood maximal function is then bounded on $L^{2/p}$).

4. RANDOM DYADIC CUBES AND THE DECOMPOSITION OF THE PAIRING $\langle Tf, g \rangle$

Start by fixing once and for all two compactly supported functions f and g so that $\|f\|_2 = \|g\|_2 = 1$ and $\|T\|/2 \leq |\langle Tf, g \rangle|$. We choose a big enough integer N so that spt f, spt $g \subset B(0, 2^{N-3})$. Consider two independent random squares $Q_0 = Q_0(w) = w + [-2^N, 2^N)^n$ and $R_0 = R_0(w') = w' + [-2^N, 2^N)^n$, where $w, w' \in [-2^{N-1}, 2^{N-1})^n$. The cubes Q_0 and R_0 are taken to be the starting cubes of the independent grids $\mathcal{D}$ and $\mathcal{D}'$ (only the cubes inside Q_0 and R_0 matter). Of course, the probability measure in question is the normalized Lebesgue measure on the square $[-2^{N-1}, 2^{N-1})^n$. Furthermore, note that always spt f, spt $g \subset \alpha Q_0 \cap \alpha R_0$ with some absolute constant $\alpha < 1$.

A cube $Q \in \mathcal{D}$ is bad (or $\mathcal{D}'$-bad), if there exists a cube R in the dyadic system $\mathcal{D}'$ such that $\ell(Q) \leq 2^{-r}\ell(R)$ and $d(Q, \text{sk } R) \leq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma}$. Here the skeleton of R is the set sk $R = \bigcup \partial R_i$, where R_i are the children of R. Also, recall that $\gamma = \alpha/(2\alpha + 2d)$, where α is the number from the kernel estimates and $d = \log_2 C_\lambda$. The number r is fixed to be large enough (this is quantified later).

We shall use the badness morally in the same line as it is usually used [NTV02, NTV03] – the details are somewhat different, however. There are various reasons for this, and we shall carefully elaborate on those after performing the decomposition, since this seems to us like a genuine source of trouble.

We define $\sum_{Q\in\mathcal{D}}^{k} = \sum_{Q\in\mathcal{D},\,\ell(Q)>2^{-k}}$. Using the facts that $E_k^{a,1}f \to f$ in L^2 and $\|E_k^{a,1}f\|_2 \lesssim \|f\|_2$ combined with dominated convergence (in the probability space) we see that

$$\langle Tf, g\rangle = \lim_{k\to\infty} E\langle T(E_k^{a,1}f), E_k^{a,2}g\rangle,$$

where E is the expectation over the random grids $\mathcal{D}$ and $\mathcal{D}'$; sometimes we will explicitly write it as $E = E_{\mathcal{D}}E_{\mathcal{D}'}$. Since

$$E_k^{a,1}f = E_{Q_0}f + \sum_{Q\in\mathcal{D}}^{k}\Delta_Q f,$$

the pairing on the right hand side can be written in the form

$$\sum_{R\in\mathcal{D}'}^{k}\sum_{Q\in\mathcal{D}}^{k}\langle T(\Delta_Q f), \Delta_R g\rangle + \langle T(E_{Q_0}f), E_k^{a,2}g\rangle + \langle T(E_k^{a,1}f), E_{R_0}g\rangle - \langle T(E_{Q_0}f), E_{R_0}g\rangle.$$

Note that $\operatorname{spt} E_k^{a,2}g \subset Q_0$ for all sufficiently large k. Thus, one can bound $|\langle T(E_{Q_0}f), E_k^{a,2}g\rangle|$ by $\mu(Q_0)^{-1/2}\|\chi_{Q_0}Tb_{Q_0}^1\|_2 \lesssim 1$. For the same reason there holds $|\langle T(E_k^{a,1}f), E_{R_0}g\rangle| \lesssim 1$ (for large k). There seems to be no such equally cheap way to further bound $|\langle T(E_{Q_0}f), E_{R_0}g\rangle| \leq \mu(Q_0)^{-1/2}\mu(R_0)^{-1/2}|\langle Tb_{Q_0}^1, b_{R_0}^2\rangle|$. However, this can be controlled using a much simplified version of the arguments we shall use in Section 7 concerning adjacent cubes of comparable size in the main series $\sum_{R\in\mathcal{D}'}^{k}\sum_{Q\in\mathcal{D}}^{k}\langle T(\Delta_Q f), \Delta_R g\rangle$. We detail on this at the end of that chapter.

Therefore, one is (remembering the above remark) reduced to estimating

$$\left| E\sum_{R\in\mathcal{D}'}^{k}\sum_{Q\in\mathcal{D}}^{k}\langle T(\Delta_Q f), \Delta_R g\rangle \right|$$

with a bound independent of k. The summation after the expectation is finite, and thus all the rearrangements one could want to make are legitimate. In the sequel, the index $k = k_0$ is fixed, and we no longer make any reference to it in the notation. (The symbol k will then be free for other uses.)

We continue to write the summation

$$\sum_{R\in\mathcal{D}'}\sum_{\substack{Q\in\mathcal{D}\\\ell(Q)\leq\ell(R)}}$$

in the form

$$\sum_{R\in\mathcal{D}'}\left(\sum_{\substack{Q\in\mathcal{D}\\\ell(Q)\leq\ell(R)\\d(Q,R)>2n^{1/2}\ell(Q)^{\gamma}\ell(R)^{1-\gamma}}} + \sum_{\substack{Q\in\mathcal{D}\\\ell(Q)\leq2^{-r}\ell(R)\\d(Q,R)\leq2n^{1/2}\ell(Q)^{\gamma}\ell(R)^{1-\gamma}}} + \sum_{\substack{Q\in\mathcal{D}\\2^{-r}\ell(R)<\ell(Q)\leq\ell(R)\\d(Q,R)\leq2n^{1/2}\ell(Q)^{\gamma}\ell(R)^{1-\gamma}}} \right).$$

We denote the corresponding parts of the sum by Σ_i, $i = 1, 2, 3$. Goodness will be separately inserted only in the middle sum Σ_2. We shall now study these

sums one by one in the following sections (using both set of assumptions). Note that the sum $\ell(R) < \ell(Q)$ will then also be in check by the symmetry of our assumptions.

4.1. *Remark.* We now give a few technical comments to compare our strategy with previous works based on the use of random dyadic grids. One can safely ignore these, especially if one is not too familiar with non-homogeneous analysis.

It is natural (if one follows the beautiful strategy pioneered by Nazarov, Treil and Volberg in their deep papers [NTV97], [NTV03], [NTV02] and some others) to define

$$f_{\text{good}} = \sum_{Q \in \mathcal{D}_{\text{good}}} \Delta_Q f \qquad \text{and} \qquad f_{\text{bad}} = \sum_{Q \in \mathcal{D}_{\text{bad}}} \Delta_Q f,$$

and then write $f = f_{\text{good}} + f_{\text{bad}}$. One does the similar thing also for g but using the grid $\mathcal{D}'$ and operators Δ_R. Then one decomposes

$$\langle Tf, g \rangle = \langle Tf_{\text{good}}, g_{\text{good}} \rangle + \langle Tf_{\text{good}}, g_{\text{bad}} \rangle + \langle Tf_{\text{bad}}, g \rangle.$$

One usually wants to reduce the considerations to the pairing $\langle Tf_{\text{good}}, g_{\text{good}} \rangle$ by arguing that the bad parts are small. However, getting a hold of this smallness would typically exploit the dual square function estimate, the failure of which we already saw in our general context of accretive L^2 systems (see Example 3.9). However, with a moderate amount of work and a certain trick we managed to show (also in the L^2 case) that, after all, $E\|f_{\text{bad}}\|_2 \lesssim c(r)\|f\|_2$, where $c(r) \to 0$ when $r \to \infty$. So this reduction could, nevertheless, always be made.

Here comes another unfortunate snag: in our local situation even the good part, as defined above, seems not so good after all. Let us explain. In the global Tb theorems there holds $\Delta_Q f_{\text{good}} = \Delta_Q f$, if $Q \in \mathcal{D}_{\text{good}}$, and $\Delta_Q f_{\text{good}} = 0$, if $Q \in \mathcal{D}_{\text{bad}}$. However, there is no reason for this to be true in this local situation with the more complicated operators Δ_Q, which in general fail the pairwise orthogonality $\Delta_Q \Delta_R = 0$ for $Q \neq R$. This means that in the pairing

$$\langle Tf_{\text{good}}, g_{\text{good}} \rangle = \sum_{R \in \mathcal{D}'_{\text{good}}} \sum_{Q \in \mathcal{D}_{\text{good}}} \langle T(\Delta_Q f), \Delta_R g \rangle$$

one cannot remove any goodness from the summation – which one can in the global situation, if one replaces $\Delta_Q f = \Delta_Q f_{\text{good}}$ (and similarly for g), and then notes that adding some bad cubes to the sum just amounts to adding zeroes.

One works hard to add the restriction to good cubes only, so why would one need to remove some of it? The answer is that in the paraproduct part of the argument there is a subtle phenomenon, where it is essential that the bigger cube has no restrictions for a certain telescoping sum to collapse. If the bigger cubes are restricted to be good, the sum does not collapse, and the resulting object seems to be way too complicated to handle.

This is the reason why we choose to modify this earlier strategy, and insert the goodness in a different way. However, the paraproduct still does not become quite as simple as usually, and it is basically for this reason that in the L^2 test

function case we need the stronger integrability exponent $s > 2$ on the operator side.

There are subtle tricks which depend on independence to add and remove goodness, see [Hyt09], [Hyt10b] and [Mar10]. These cannot be used here either, and this is basically because Δ_Q depends not only on the cube Q and its children (like in the global Tb theorems), but also, through the stopping time argument, on the whole grid $\mathcal{D}$ (and this stops one from using certain independence properties).

5. SEPARATED CUBES

The following is the long interaction lemma. For a proof in this general upper doubling situation, see [HM09, Lemma 6.1 and Lemma 6.2].

5.1. **Lemma.** *Suppose that $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$ are such that $\ell(Q) \leq \ell(R)$ and $d(Q, R) \geq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma}$, and that φ_Q and ψ_R are $L^2(\mu)$ functions supported on Q and R respectively. Assume also that $\int \varphi_Q \, d\mu = 0$. Then there holds*

$$|\langle T\varphi_Q, \psi_R \rangle| \lesssim \frac{\ell(Q)^\alpha}{d(Q, R)^\alpha \sup_{z \in Q} \lambda(z, d(Q, R))} \mu(Q)^{1/2}\mu(R)^{1/2}\|\varphi_Q\|_{L^2(\mu)}\|\psi_R\|_{L^2(\mu)}$$

$$\lesssim \frac{\ell(Q)^{\alpha/2}\ell(R)^{\alpha/2}}{D(Q, R)^\alpha \sup_{z \in Q} \lambda(z, D(Q, R))} \mu(Q)^{1/2}\mu(R)^{1/2}\|\varphi_Q\|_{L^2(\mu)}\|\psi_R\|_{L^2(\mu)},$$

where $D(Q, R) = \ell(Q) + \ell(R) + d(Q, R)$.

The fact that the corresponding matrix generates a bounded operator in ℓ^2 is the content of the next proposition (this is [HM09, Proposition 6.3]).

5.2. **Proposition.** *Let*

$$T_{QR} = \frac{\ell(Q)^{\alpha/2}\ell(R)^{\alpha/2}}{D(Q, R)^\alpha \sup_{z \in Q} \lambda(z, D(Q, R))} \mu(Q)^{1/2}\mu(R)^{1/2},$$

if $Q \in \mathcal{D}$, $R \in \mathcal{D}'$ and $\ell(Q) \leq \ell(R)$, and

$$T_{QR} = 0$$

otherwise. Then there holds with any $x_Q, y_R \geq 0$ that

$$\sum_{Q, R} T_{QR} x_Q y_R \lesssim \Big(\sum_Q x_Q^2 \Big)^{1/2} \Big(\sum_R y_R^2 \Big)^{1/2}.$$

The above combined with the square function estimates $\big(\sum_Q \|\Delta_Q f\|_2^2 \big)^{1/2} \lesssim \|f\|_2 = 1$ and $\big(\sum_R \|\Delta_R g\|_2^2 \big)^{1/2} \lesssim \|g\|_2 = 1$ yield the following.

5.3. **Proposition.** *There holds $|\Sigma_1| \leq C$.*

The long range interaction lemma will still have further use to us when dealing with the sum Σ_2 in the next section.

6. Cubes well inside another cube and the related bad part

We shall now deal with Σ_2. We define

$$\Sigma_{2,\,\mathrm{bad}} = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq 2^{-r}\ell(R) \\ d(Q,R) \leq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma} \\ Q \text{ is bad w.r.t. a cube of the size of } R \text{ or larger}}} \langle T(\Delta_Q f), \Delta_R g \rangle.$$

The last summing condition just means that there is a cube $S \in \mathcal{D}'$ such that $\ell(S) \geq \ell(R)$ and $d(Q, \mathrm{sk}\, S) \leq 2n^{1/2}\ell(Q)^\gamma \ell(S)^{1-\gamma}$. Then $\Sigma_2 = \Sigma_{2,\,\mathrm{good}} + \Sigma_{2,\,\mathrm{bad}}$, where

$$\Sigma_{2,\,\mathrm{good}} = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq 2^{-r}\ell(R) \\ d(Q,R) \leq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma} \\ Q \text{ is good w.r.t. all the cubes of the size of } R \text{ or larger}}} \langle T(\Delta_Q f), \Delta_R g \rangle.$$

6.A. **The disposal of the bad bart $\Sigma_{2,\,\mathrm{bad}}$.** Define $\mathcal{D}_{\mathrm{bad},\,A}$ to be the collection of those cubes $Q \in \mathcal{D}$ which are bad with respect to some $\mathcal{D}'$-cube of side length A or larger. We do not always explicitly write the summing conditions $\ell(Q) \leq 2^{-r}\ell(R)$ and $d(Q, R) \leq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma}$, but these are in force, nevertheless. We then estimate as follows

$$|\Sigma_{2,\,\mathrm{bad}}| \leq \sum_{R \in \mathcal{D}'} \Big| \Big\langle T\Big(\sum_{\substack{Q \in \mathcal{D} \\ Q \in \mathcal{D}_{\mathrm{bad},\,\ell(R)}}} \Delta_Q f \Big), \Delta_R g \Big\rangle \Big|$$

$$\leq \|T\| \sum_{R \in \mathcal{D}'} \Big\| \sum_{\substack{Q \in \mathcal{D} \\ Q \in \mathcal{D}_{\mathrm{bad},\,\ell(R)}}} \Delta_Q f \Big\|_2 \|\Delta_R g\|_2$$

$$= \|T\| \sum_{R \in \mathcal{D}'} \Big\| \sum_{k \geq r} \sum_{\substack{Q \in \mathcal{D}, \ell(Q) = 2^{-k}\ell(R) \\ Q \in \mathcal{D}_{\mathrm{bad},\,\ell(R)}}} \Delta_Q f \Big\|_2 \|\Delta_R g\|_2$$

$$\leq \|T\| \sum_{k \geq r} \sum_{R \in \mathcal{D}'} \Big(\sum_{\substack{Q \in \mathcal{D}, \ell(Q) = 2^{-k}\ell(R) \\ Q \in \mathcal{D}_{\mathrm{bad},\,2^k\ell(Q)}}} \|\Delta_Q f\|_2^2 \Big)^{1/2} \|\Delta_R g\|_2$$

$$\leq \|T\| \sum_{k \geq r} \Big(\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{D}, \ell(Q) = 2^{-k}\ell(R) \\ Q \in \mathcal{D}_{\mathrm{bad},\,2^k\ell(Q)}}} \|\Delta_Q f\|_2^2 \Big)^{1/2} \Big(\sum_{R \in \mathcal{D}'} \|\Delta_R g\|_2^2 \Big)^{1/2}$$

$$\lesssim \|T\| \|g\|_2 \sum_{k \geq r} \Big(\sum_{\substack{Q \in \mathcal{D} \\ Q \in \mathcal{D}_{\mathrm{bad},\,2^k\ell(Q)}}} \|\Delta_Q f\|_2^2 \Big)^{1/2},$$

where the last estimate used Proposition 3.8 and the fact that given Q, there are $\lesssim 1$ cubes R so that $\ell(R) = 2^k \ell(Q)$ and $d(Q, R) \leq 2n^{1/2} \ell(Q)^\gamma \ell(R)^{1-\gamma}$. Thus, we have (here the expectation $E = E_{\mathcal{D}} E_{\mathcal{D}'} = E_w E_{w'}$)

$$E|\Sigma_{2,\text{bad}}| \lesssim \|T\|\|g\|_2 E_{\mathcal{D}} \sum_{k \geq r} \Big(\sum_{Q \in \mathcal{D}} \mathbb{P}(Q \in \mathcal{D}_{\text{bad}, 2^k \ell(Q)}) \|\Delta_Q f\|_2^2 \Big)^{1/2}$$

$$\lesssim \|T\|\|g\|_2 E_{\mathcal{D}} \sum_{k \geq r} \Big(2^{-\gamma k} \sum_{Q \in \mathcal{D}} \|\Delta_Q f\|_2^2 \Big)^{1/2}$$

$$\lesssim \|T\|\|f\|_2\|g\|_2 \sum_{k \geq r} (2^{-\gamma/2})^k = c(r)\|T\|,$$

where $c(r) \to 0$, when $r \to \infty$ (recall $\|f\|_2 = \|g\|_2 = 1$). We now fix a large r so that $E|\Sigma_{2,\text{bad}}| \leq \|T\|/16$. We are done with the bad part.

6.B. **Reduction of the good part $\Sigma_{2,\text{good}}$ to a paraproduct.** Note that if $Q \in \mathcal{D}$ is good with respect to R and $d(Q, R) \leq 2n^{1/2} \ell(Q)^\gamma \ell(R)^{1-\gamma}$, then there actually is a child R_1 of R so that $Q \subset R_1$ and $d(Q, \partial R_1) > 2n^{1/2} \ell(Q)^\gamma \ell(R)^{1-\gamma}$.

6.B.1. *Case $(R_1)^a = R^a$.* We begin by assuming that $(R_1)^a = R^a$. In this case $\Delta_R g = B_{R_1} \chi_{R_1} b_{R^a}^2 + \chi_{R \setminus R_1} \Delta_R g$. One may then perform the usual decomposition

$$(6.1) \quad \langle T(\Delta_Q f), \Delta_R g \rangle = \langle T(\Delta_Q f), B_{R_1} b_{R^a}^2 \rangle - \langle T(\Delta_Q f), B_{R_1}(1 - \chi_{R_1}) b_{R^a}^2 \rangle + \langle T(\Delta_Q f), \chi_{R \setminus R_1} \Delta_R g \rangle.$$

The last term, where $\chi_{R \setminus R_1} = \sum_{i=2}^{2^n} \chi_{R_i}$, can be readily estimated using the long range interaction lemma:

$$|\langle T(\Delta_Q f), \chi_{R_i} \Delta_R g \rangle| \lesssim \Big(\frac{\ell(Q)}{\ell(R)} \Big)^{\alpha/2} \mu(Q)^{1/2} \frac{\mu(R_i)^{1/2}}{\sup_{z \in Q} \lambda(z, \ell(R_i))} \|\Delta_Q f\|_2 \|\Delta_R g\|_2$$

$$\lesssim \Big(\frac{\ell(Q)}{\ell(R)} \Big)^{\alpha/2} \Big(\frac{\mu(Q)}{\mu(R_1)} \Big)^{1/2} \|\Delta_Q f\|_2 \|\Delta_R g\|_2, \qquad i \neq 1.$$

The corresponding matrix is a bounded operator in ℓ^2 by [NTV02, Lemma 6.1] (this is a lemma which uses no special properties of the measure). The first term will be part of the soon to be formed paraproduct.

Let us now bound the term $\langle T(\Delta_Q f), B_{R_1}(1 - \chi_{R_1}) b_{R^a}^2 \rangle$ in the middle. We have with any fixed $z \in Q$ that

$$|\langle T(\Delta_Q f), (1 - \chi_{R_1}) b_{R^a}^2 \rangle|$$

$$= \Big| \int_{R^a \setminus R_1} \int_Q [K(x, y) - K(x, z)] \Delta_Q f(y) b_{R^a}^2(x) \, d\mu(y) \, d\mu(x) \Big|$$

$$\lesssim \|\Delta_Q f\|_1 \int_{R^a \setminus R_1} \frac{\ell(Q)^\alpha}{|x - z|^\alpha \lambda(z, |x - z|)} |b_{R^a}^2(x)| d\mu(x).$$

Let us first bound this in the easier case of the L^∞ test functions. We have that

$$\int_{R^a \setminus R_1} \frac{\ell(Q)^\alpha}{|x-z|^\alpha \lambda(z, |x-z|)} |b_{R^a}^2(x)| d\mu(x) \lesssim \ell(Q)^\alpha \int_{\mathbb{R}^n \setminus B(z, d(Q, \partial R_1))} \frac{|x-z|^{-\alpha}}{\lambda(z, |x-z|)} d\mu(x)$$

$$\lesssim \ell(Q)^\alpha d(Q, \partial R_1)^{-\alpha} \lesssim \Big(\frac{\ell(Q)}{\ell(R)}\Big)^{\alpha/2},$$

where we used [HM09, Lemma 2.4] and the fact that $d(Q, \partial R_1) \gtrsim \ell(Q)^{1/2} \ell(R)^{1/2}$.

Let us now establish the same bound in the case of L^2 test functions (we do not even need a doubling measure for this – so this gives another proof of the above estimate too). Here we need to use the fact that Q is good with respect to R and all the bigger cubes. Let M be such that $(R_1)^{(M+1)} = R^a$. We have

$$\int_{R^a \setminus R_1} \frac{\ell(Q)^\alpha}{|x-z|^\alpha \lambda(z, |x-z|)} |b_{R^a}^2(x)| d\mu(x)$$

$$= \ell(Q)^\alpha \sum_{j=0}^{M} \int_{(R_1)^{(j+1)} \setminus (R_1)^{(j)}} \frac{|b_{R^a}^2(x)|}{|x-z|^\alpha \lambda(z, |x-z|)} d\mu(x).$$

There holds (since $\gamma(\alpha + d) = \alpha/2$) that

$$|x-z|^\alpha \lambda(z, |x-z|) \gtrsim d(Q, \partial R_1^{(j)})^\alpha \lambda(z, d(Q, \partial R_1^{(j)}))$$

$$\gtrsim \ell(Q)^{\gamma \alpha} \ell((R_1)^{(j)})^{\alpha - \gamma \alpha} \Big(\frac{\ell((R_1)^{(j)})}{\ell(Q)}\Big)^{-\gamma d} \mu((R_1)^{(j+1)})$$

$$= \ell(Q)^{\alpha/2} \ell((R_1)^{(j)})^{\alpha/2} \mu((R_1)^{(j+1)}),$$

and so using the fact that $\int_{(R_1)^{(j+1)}} |b_{R^a}^2| \, d\mu \lesssim \mu((R_1)^{(j+1)})$, we have

$$\int_{R^a \setminus R_1} \frac{\ell(Q)^\alpha}{|x-z|^\alpha \lambda(z, |x-z|)} |b_{R^a}^2(x)| d\mu(x) \lesssim \Big(\frac{\ell(Q)}{\ell(R)}\Big)^{\alpha/2} \sum_{j=0}^{\infty} 2^{-\alpha j/2} \lesssim \Big(\frac{\ell(Q)}{\ell(R)}\Big)^{\alpha/2}.$$

As $|B_{R_1}| \lesssim \mu(R_1)^{-1/2} \|\Delta_R g\|_2$, we have shown that

$$|\langle T(\Delta_Q f), B_{R_1}(1 - \chi_{R_1}) b_{R^a}^2 \rangle| \lesssim \Big(\frac{\mu(Q)}{\mu(R_1)}\Big)^{1/2} \Big(\frac{\ell(Q)}{\ell(R)}\Big)^{\alpha/2} \|\Delta_Q f\|_2 \|\Delta_R g\|_2,$$

and this is known to be acceptable (see again [NTV02, Lemma 6.1]).

6.B.2. *Case* $(R_1)^a = R_1$. We then assume that $(R_1)^a = R_1$. In this case we write $B_{R_1} = \langle g \rangle_{R_1} / \langle b_{R_1}^2 \rangle_{R_1}$ and $C_R = \langle g \rangle_R / \langle b_{R^a}^2 \rangle_R$, and then decompose as follows:

$$(6.2) \quad \begin{aligned} \langle T(\Delta_Q f), \Delta_R g \rangle &= \langle T(\Delta_Q f), B_{R_1} b_{R_1}^2 \rangle - \langle T(\Delta_Q f), C_R b_{R^a}^2 \rangle \\ &\quad + C_R \langle T(\Delta_Q f), (1 - \chi_{R_1}) b_{R^a}^2 \rangle + \langle T(\Delta_Q f), \chi_{R \setminus R_1} \Delta_R g \rangle. \end{aligned}$$

The last term, being identical to the last term in (6.1), is again handled using the long range interaction lemma. The next to last term is also estimated as above,

except that this time we have $|C_R| \lesssim |\langle g \rangle_R|$, so we get

$$|C_R \langle T(\Delta_Q f), (1 - \chi_{R_1}) b_{R^a}^2 \rangle| \lesssim \left(\frac{\mu(Q)}{\mu(R_1)} \right)^{1/2} \left(\frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \|\Delta_Q f\|_2 \mu(R_1)^{1/2} |\langle g \rangle_R|.$$

This is again fine by [NTV02, Lemma 6.1], since

$$\left(\sum_{\substack{R \in \mathcal{D}' \\ (R_1)^a = R_1}} |\langle g \rangle_R|^2 \mu(R_1) \right)^{1/2} \lesssim \|g\|_2$$

by Carleson's embedding theorem. The first two terms in (6.2) will be part of the paraproduct.

6.C. **The paraproduct and its boundedness.** Let $\mathcal{D}_{\text{good}, k}$ be the collection of those $Q \in \mathcal{D}$ which are good with respect to all $\mathcal{D}'$-cubes of side length $2^k \ell(Q)$ and larger. If $Q \in \bigcup_{k \geq r} \mathcal{D}_{\text{good}, k}$, let $\alpha(Q)$ be the smallest index k so that $Q \in \mathcal{D}_{\text{good}, k}$. So collecting the terms that we did not yet estimate in (6.1) and (6.2), we see that we need to bound

$$\sum_{\substack{Q \in \cup_{k \geq r} \mathcal{D}_{\text{good}, k} \\ Q \subset R_0}} \sum_{\substack{R \in \mathcal{D}' \\ \ell(R) \geq 2^{\alpha(Q)} \ell(Q) \\ d(Q,R) \leq 2n^{1/2} \ell(Q)^\gamma \ell(R)^{1-\gamma}}} \left\langle T(\Delta_Q f), \frac{\langle g \rangle_{R_1}}{\langle b_{(R_1)^a}^2 \rangle_{R_1}} b_{(R_1)^a}^2 - \frac{\langle g \rangle_R}{\langle b_{R^a}^2 \rangle_R} b_{R^a}^2 \right\rangle.$$

Note that there is a unique R of each side length in the inner sum, the one with $R \supset Q$. In the above summation, let $S(Q) \in \mathcal{D}'$ be R_1, when $\ell(R) = 2^{\alpha(Q)} \ell(Q)$. Then bringing the R summation inside the pairing, we see that the sum collapses to

$$\sum_{\substack{Q \in \cup_{k \geq r} \mathcal{D}_{\text{good}, k} \\ Q \subset R_0}} \left\langle T(\Delta_Q f), \frac{\langle g \rangle_{S(Q)}}{\langle b_{S(Q)^a}^2 \rangle_{S(Q)}} b_{S(Q)^a}^2 - \frac{\langle g \rangle_{R_0}}{\langle b_{R_0}^2 \rangle_{R_0}} b_{R_0}^2 \right\rangle.$$

We write this in the form

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \cup_{k \geq r} \mathcal{D}_{\text{good}, k} \\ S(Q) = R}} \left\langle T(\Delta_Q f), \frac{\langle g \rangle_R}{\langle b_{R^a}^2 \rangle_R} b_{R^a}^2 \right\rangle - \sum_{\substack{Q \in \cup_{k \geq r} \mathcal{D}_{\text{good}, k} \\ Q \subset R_0}} \left\langle T(\Delta_Q f), \frac{\langle g \rangle_{R_0}}{\langle b_{R_0}^2 \rangle_{R_0}} b_{R_0}^2 \right\rangle.$$

So we were able to collapse the sum because we introduced the goodness in a more restricted way than is usually done (see Remark 4.1). But the result is somewhat different from the usual paraproducts, since $S(Q)$ can be arbitrarily larger than Q.

At this stage we bring the absolute values inside the summations. We may then consider the following, somewhat more general, situation. Let us be given a collection $\mathcal{F} \subset \mathcal{D}$ so that to every cube $Q \in \mathcal{F}$ there is associated a unique cube

$F(Q) \in \mathcal{D}'$ for which there holds $Q \subset F(Q)$. The rest of this section is concerned with proving that

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} |\langle T(\Delta_Q f), \langle g \rangle_R b_{R^a}^2 \rangle| \lesssim 1.$$

We begin by recalling from [NTV02, p. 271] that

$$\Delta_Q f = (\Delta_Q)^2 f + \sum_{\substack{P \in \operatorname{ch}(Q) \\ P^a = P}} \varphi_P,$$

where

$$\varphi_P = \frac{\langle f \rangle_Q}{\langle b_{Q^a}^1 \rangle_Q} \Big[\frac{\langle b_{Q^a}^1 \rangle_P}{\langle b_P^1 \rangle_P} b_P^1 - b_{Q^a}^1 \Big] \chi_P.$$

It follows that always

$$\|\varphi_P\|_2 \lesssim |\langle f \rangle_Q| (|\langle b_{Q^a}^1 \rangle_P| \|b_P^1\|_2 + \|\chi_P b_{Q^a}^1\|_2).$$

This can be further bounded by $|\langle f \rangle_Q| \mu(P)^{1/2}$ (in the L^2 case, the doubling property is needed here).

We estimate

$$\sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} |\langle T(\Delta_Q f), \langle g \rangle_R b_{R^a}^2 \rangle| \leq I + II,$$

where

$$I = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} |\langle \Delta_Q f, \langle g \rangle_R (\Delta_Q)^* T^* b_{R^a}^2 \rangle|,$$

$$II = \sum_{R \in \mathcal{D}'} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} \sum_{\substack{P \in \operatorname{ch}(Q) \\ P^a = P}} |\langle \varphi_P, \langle g \rangle_R \chi_P T^* b_{R^a}^2 \rangle|.$$

There holds that

$$I \leq \Big(\sum_{Q \in \mathcal{D}} \|\Delta_Q f\|_2^2 \Big)^{1/2} \Big(\sum_{R \in \mathcal{D}'} |\langle g \rangle_R|^2 \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} \|(\Delta_Q)^* T^* b_{R^a}^2\|_2^2 \Big)^{1/2}$$

$$\lesssim \Big(\sum_{R \in \mathcal{D}'} |\langle g \rangle_R|^2 a_R \Big)^{1/2},$$

where $a_R = \sum_{Q \in \mathcal{F}, \, F(Q)=R} \|(\Delta_Q)^* T^* b_{R^a}^2\|_2^2$. Also, there holds

$$II \le \Big(\sum_{Q \in \mathcal{D}} |\langle f \rangle_Q|^2 \sum_{\substack{P \in \text{ch}(Q) \\ P^a = P}} \mu(P) \Big)^{1/2} \Big(\sum_{R \in \mathcal{D}'} |\langle g \rangle_R|^2 \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} \sum_{\substack{P \in \text{ch}(Q) \\ P^a = P}} \|\chi_P T^* b_{R^a}^2\|_2^2 \Big)^{1/2}$$

$$\lesssim \Big(\sum_{R \in \mathcal{D}'} |\langle g \rangle_R|^2 b_R \Big)^{1/2},$$

where $b_R = \sum_{Q \in \mathcal{F}, \, F(Q)=R} \sum_{P \in \text{ch}(Q), \, P^a = P} \|\chi_P T^* b_{R^a}^2\|_2^2$.

We are reduced to showing that (a_R) and (b_R) form Carleson sequences (both in the L^∞ test function and in the L^2 test function case).

6.3. Lemma. *The sequence*

$$a_R = \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} \|(\Delta_Q)^* T^* b_{R^a}^2\|_2^2$$

is Carleson.

Proof. Consider an arbitrary $R \in \mathcal{D}'$. We write

$$\sum_{\substack{S \in \mathcal{D}' \\ S \subset R}} a_S = \sum_{\substack{S \in \mathcal{D}' \\ S \subset R}} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=S}} \|(\Delta_Q)^* T^* b_{S^a}^2\|_2^2$$

$$= \Big(\sum_{\substack{S \in \mathcal{D}' \\ S \subset R \\ S^a = R^a}} + \sum_{\substack{H \subset R \\ H^a = H}} \sum_{\substack{S \in \mathcal{D}' \\ S^a = H}} \Big) \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=S}} \|(\Delta_Q)^* T^* b_{S^a}^2\|_2^2.$$

We are reduced to showing that for an arbitrary $H \in \mathcal{D}'$ there holds

$$I_H := \sum_{\substack{S \in \mathcal{D}' \\ S \subset H \\ S^a = H^a}} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=S}} \|(\Delta_Q)^* T^* b_{H^a}^2\|_2^2 \lesssim \mu(H).$$

We estimate as follows

$$I_H \le \sum_{\substack{Q \in \mathcal{D} \\ Q \subset H}} \|(\Delta_Q)^* T^* b_{H^a}^2\|_2^2 = \sum_{Q \in M(H)} \sum_{\substack{U \in \mathcal{D} \\ U \subset Q}} \|(\Delta_U)^* T^* b_{H^a}^2\|_2^2,$$

where $M(H)$ consists of maximal $Q \in \mathcal{D}$ for which $Q \subset H$. Now the claim is very easy in the L^∞ case. Just use the dual square function estimate and the fact that $\|\chi_Q T^* b_{H^a}^2\|_2^2 \lesssim \mu(Q)$.

We are thus reduced to the case $\mu = \nu$ with L^2 test functions. For a given $Q \in M(H)$ we estimate using Proposition 3.10 that

$$\sum_{\substack{U \in \mathcal{D} \\ U \subset Q}} \|(\Delta_U)^* T^* b_{H^a}^2\|_2^2 = \Big(\sum_{\substack{U \in \mathcal{D} \\ U \subset Q \\ U^a = Q^a}} + \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \sum_{\substack{U \in \mathcal{D} \\ U^a = K}} \Big) \|(\Delta_U)^* T^* b_{H^a}^2\|_2^2$$

$$\lesssim \|\chi_Q \chi_H T^* b_{H^a}^2\|_2^2 + \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \|\chi_K \chi_H T^* b_{H^a}^2\|_2^2.$$

Thus, there holds

$$I_H \lesssim \int \Big(\sum_{Q \in M(H)} \chi_Q \Big) \chi_H |T^* b_{H^a}^2|^2 \, d\nu + \int \Big(\sum_{Q \in M(H)} \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big) \chi_H |T^* b_{H^a}^2|^2 \, d\nu$$

$$\leq \Big(\Big\| \sum_{Q \in M(H)} \chi_Q \Big\|_{p'} + \Big\| \sum_{Q \in M(H)} \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'} \Big) \Big(\int_H |T^* b_{H^a}^2|^s \, d\nu \Big)^{1/p}$$

$$\lesssim \Big(\Big\| \sum_{Q \in M(H)} \chi_Q \Big\|_{p'} + \Big\| \sum_{Q \in M(H)} \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'} \Big) \nu(H)^{1/p},$$

where $p = s/2 > 1$.

Note that

$$\Big\| \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'}^{p'} = \Big\| \sum_j \sum_{\substack{K \in \mathcal{D}^j \\ K \subset Q}} \chi_K \Big\|_{p'}^{p'}$$

$$\leq \Big(\sum_j \Big\| \sum_{\substack{K \in \mathcal{D}^j \\ K \subset Q}} \chi_K \Big\|_{p'} \Big)^{p'} \lesssim \Big(\sum_j \tau^{j/p'} \Big)^{p'} \nu(Q) \lesssim \nu(Q),$$

and so

$$\Big\| \sum_{Q \in M(H)} \chi_Q \Big\|_{p'}^{p'} + \Big\| \sum_{Q \in M(H)} \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'}^{p'} = \sum_{Q \in M(H)} \Big(\nu(Q) + \Big\| \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'}^{p'} \Big)$$

$$\lesssim \sum_{Q \in M(H)} \nu(Q) \leq \nu(H).$$

This establishes that $I_H \lesssim \nu(H)^{1/p'} \nu(H)^{1/p} = \nu(H)$, as was the goal. $\square$

6.4. **Lemma.** *The sequence*

$$b_R = \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=R}} \sum_{\substack{P \in \mathrm{ch}(Q) \\ P^a = P}} \|\chi_P T^* b_{R^a}^2\|_2^2$$

is Carleson.

Proof. As in the proof of the previous lemma, this reduces to showing that for an arbitrary $H \in \mathcal{D}'$ there holds

$$I_H := \sum_{\substack{S \in \mathcal{D}' \\ S \subset H \\ S^a = H^a}} \sum_{\substack{Q \in \mathcal{F} \\ F(Q)=S}} \sum_{\substack{P \in \mathrm{ch}(Q) \\ P^a = P}} \|\chi_P T^* b_{H^a}^2\|_2^2 \lesssim \mu(H).$$

Letting $M(H)$ consist of the maximal $Q \in \mathcal{D}$ for which $Q \subset H$, we have

$$I_H \leq \sum_{\substack{Q \in \mathcal{D} \\ Q \subset H}} \sum_{\substack{P \in \mathrm{ch}(Q) \\ P^a = P}} \|\chi_P T^* b_{H^a}^2\|_2^2 = \sum_{Q \in M(H)} \sum_{\substack{U \in \mathcal{D} \\ U \subset Q}} \sum_{\substack{P \in \mathrm{ch}(U) \\ P^a = P}} \|\chi_P T^* b_{H^a}^2\|_2^2.$$

The L^∞ case is again clear from this (recalling Lemma 3.4). Otherwise, we have as in the proof of the previous lemma that

$$I_H \leq \int \Big(\sum_{Q \in M(H)} \sum_{\substack{U \in \mathcal{D} \\ U \subset Q}} \sum_{\substack{P \in \mathrm{ch}(U) \\ P^a = P}} \chi_P \Big) \chi_H |T^* b_{H^a}^2|^2 \, d\nu$$

$$\lesssim \Big(\sum_{Q \in M(H)} \Big\| \sum_{\substack{K \in \mathcal{D} \\ K \subset Q \\ K^a = K}} \chi_K \Big\|_{p'}^{p'} \Big)^{1/p'} \nu(H)^{1/p} \lesssim \nu(H),$$

where $p = s/2 > 1$. $\square$

6.5. *Remark.* The proofs of the previous two lemmata are the only places of the paper where we use, in the case of accretive L^2 systems, the stronger integrability exponent $s > 2$ on the operator side. The lemmata are true with $s = 2$, if one always has $\nu(F(Q)) \lesssim \nu(Q)$. Unfortunately, if $F(Q) = S(Q)$, as in the proof of the main theorem, then this does not have to be the case. It does not seem to be easy to arrange the collapse of the paraproduct in such a way that $S(Q)$ would be, say, always precisely r generations larger than Q (and still know how to estimate the bad part to be small).

The above two lemmata end our proof of the boundedness of the paraproduct. Recalling that $\Sigma_2 = \Sigma_{2,\mathrm{good}} + \Sigma_{2,\mathrm{bad}}$, $E|\Sigma_{2,\mathrm{bad}}| \leq \|T\|/16$, and that $\Sigma_{2,\mathrm{good}}$ decomposes into the paraproduct and some other terms, all of which we have shown to be bounded, we have established the following proposition.

6.6. **Proposition.** *There holds, after fixing the parameter r to be large enough, that*

$$|E\Sigma_2| \leq C + \|T\|/16.$$

7. ADJACENT CUBES OF COMPARABLE SIZE

We shall sum over those $Q \in \mathcal{D}$, $R \in \mathcal{D}'$ for which $2^{-r}\ell(R) < \ell(Q) \leq \ell(R)$ and $d(Q,R) \leq 2n^{1/2}\ell(Q)^\gamma \ell(R)^{1-\gamma}$. For a given R, there are only boundedly many such Q. Thus, this reduces to considering a finite number of subseries

$$\sum_{R \in \mathcal{D}'} \langle T(\Delta_Q f), \Delta_R g \rangle,$$

where $Q = Q(R)$. Moreover, one may assume that $R \mapsto Q(R)$ is invertible.

There holds

$$\langle T(\Delta_Q f), \Delta_R g \rangle = \sum_{i,j=1}^{2^n} \langle T(\chi_{Q_i} \Delta_Q f), \chi_{R_j} \Delta_R g \rangle.$$

If $Q \in \mathcal{D}_k$, one can write

$$\chi_{Q_i} \Delta_Q f = \chi_{Q_i} b_{Q_i^a}^1 \langle s_k \rangle_{Q_i} + \chi_{Q_i} b_{Q_i^a}^1 \langle h_k \rangle_{Q_i} + \chi_{Q_i} b_{Q^a}^1 \langle u_k \rangle_{Q_i},$$

where

$$s_k = \chi_{\{b_{k+1}^{a,1} = b_k^{a,1}\}} \left(\frac{E_{k+1}f}{E_{k+1}b_{k+1}^{a,1}} - \frac{E_k f}{E_k b_k^{a,1}} \right),$$

$$h_k = \chi_{\{b_{k+1}^{a,1} \neq b_k^{a,1}\}} \frac{E_{k+1}f}{E_{k+1}b_{k+1}^{a,1}}, \qquad u_k = -\chi_{\{b_{k+1}^{a,1} \neq b_k^{a,1}\}} \frac{E_k f}{E_k b_k^{a,1}}.$$

Here we interpret

$$\{b_{k+1}^{a,1} = b_k^{a,1}\} = \bigcup_{\substack{Q \in \mathcal{D}_{k+1} \\ Q^a = (Q^{(1)})^a}} Q, \qquad \{b_{k+1}^{a,1} \neq b_k^{a,1}\} = \bigcup_{\substack{Q \in \mathcal{D}_{k+1} \\ Q^a = Q}} Q.$$

Hence, we can dominate our series with nine summands of the form

$$\sum_{i,j=1}^{2^n} \left| \sum_{R \in \mathcal{D}'} \langle f_Q \rangle_{Q_i} \langle T(\chi_{Q_i} \varphi_{Q,i}), \chi_{R_j} \psi_{R,j} \rangle \langle g_R \rangle_{R_j} \right|,$$

where $\langle f_Q \rangle_{Q_i} = \langle \chi_Q f_k \rangle_{Q_i} = \langle f_k \rangle_{Q_i}$ (if $Q \in \mathcal{D}_k$), and the summands are determined by the choices

$$(f_k, \varphi_{Q,i}) \in \{(s_k, b_{Q_i^a}^1), (h_k, b_{Q_i^a}^1), (u_k, b_{Q^a}^1)\}$$

and analogous choices for g. Observe that in each case we have

$$\|\chi_{Q_i} \varphi_{Q,i}\|_2 + \|\chi_{Q_i} T\varphi_{Q,i}\|_2 \lesssim \mu(Q_i)^{1/2}$$

by the construction of the stopping time, using doubling in the case of accretive L^2 systems.

We fix the parameters i, j now.

7.1. Lemma. *There holds*

$$\left(\sum_{Q \in \mathcal{D}} \mu(Q_i) |\langle f_Q \rangle_{Q_i}|^2 \right)^{1/2} \lesssim \|f\|_2.$$

Proof. If $Q_i^a = Q^a$, we have $\mu(Q_i)|\langle f_Q\rangle_{Q_i}|^2 \leq \|\tilde\Delta_Q f\|_2^2$, where $\tilde\Delta_Q$ is the operator Δ_Q without the multiplying b functions:

$$\tilde\Delta_Q f = \sum_{Q'\in\mathrm{ch}(Q)} \Big[\frac{\langle f\rangle_{Q'}}{\langle b_{(Q')^a}^1\rangle_{Q'}} - \frac{\langle f\rangle_{Q}}{\langle b_{Q^a}^1\rangle_{Q}}\Big]\chi_{Q'}.$$

Otherwise, we have the bound $\mu(Q_i)\,|\langle f_Q\rangle_{Q_i}|^2 \lesssim \mu(Q_i)[|\langle f\rangle_{Q_i}|^2 + |\langle f\rangle_Q|^2]$. Stopping cubes (those cubes H for which $H^a = H$) form a Carleson sequence, and the estimate

$$\sum_{Q\in\mathcal{D}} \|\tilde\Delta_Q f\|_2^2 \lesssim \|f\|_2^2$$

is shown in the same way as Proposition 3.8, so we are done. $\qquad\square$

7.A. **Surgery.** We then begin the delicate surgery part of the argument – this is done a bit differently than in [NTV02] (e.g. the concept of badly intersected cubes is not needed). Also, the L^2 test function case needs several modifications.

To handle the various separated terms that we shall encounter in a unified manner, estimates in the spirit of the following lemma are useful (this is a small modification of [HM09, Lemma 9.3]).

7.2. **Lemma.** *Let S_1 and S_2 be two sets so that we have $d(S_1) \sim d(S_2)$ and $d(S_1, S_2) \gtrsim \delta \min(d(S_1), d(S_2))$. Suppose we are also given functions φ and ψ supported on S_1 and S_2 respectively. Then there holds that*

$$|\langle T\varphi, \psi\rangle| \lesssim \delta^{-d}\|\varphi\|_2\|\psi\|_2.$$

Let $\eta > 0$. Define $\delta_Q^\eta = (1+\eta)Q \setminus (1-\eta)Q$. If $R \in \mathcal{D}'$ and $Q = Q(R) \in \mathcal{D}$, we set

$$Q_{i,\partial} = Q_i \cap \delta_{R_j}^\eta, \quad Q_{i,s} = Q_i \setminus Q_{i,\partial} \setminus (Q_i \cap R_j), \quad \Delta_{Q_i} = (Q_i \cap R_j) \setminus Q_{i,\partial},$$

and define the analogous sets also for R. (The subscript s refers to separation from R_j.) Of course, e.g. $Q_{i,\partial}$ depends also on j, but the dependence is suppressed, as j is considered fixed here, in any case. We may then decompose

$$\begin{aligned}
\langle T(\chi_{Q_i}\varphi_{Q,i}), \chi_{R_j}\psi_{R,j}\rangle &= \langle T(\chi_{Q_i}\varphi_{Q,i}), \chi_{R_{j,s}}\psi_{R,j}\rangle \\
&+ \langle T(\chi_{Q_i}\varphi_{Q,i}), \chi_{R_{j,\partial}}\psi_{R,j}\rangle \\
&+ \langle T(\chi_{Q_{i,\partial}}\varphi_{Q,i}), \chi_{\Delta_{R_j}}\psi_{R,j}\rangle \\
&+ \langle T(\chi_{Q_{i,s}}\varphi_{Q,i}), \chi_{\Delta_{R_j}}\psi_{R,j}\rangle \\
&+ \langle T(\chi_{\Delta_{Q_i}}\varphi_{Q,i}), \chi_{\Delta_{R_j}}\psi_{R,j}\rangle.
\end{aligned}$$

7.A.1. *Arguments involving η-boundary regions.* We always have that $\|\chi_{Q_i}\varphi_{Q,i}\|_2 \lesssim \mu(Q_i)^{1/2}$. Thus, by separation,

$$|\langle T(\chi_{Q_i}\varphi_{Q,i}), \chi_{R_{j,s}}\psi_{R,j}\rangle| + |\langle T(\chi_{Q_{i,s}}\varphi_{Q,i}), \chi_{\Delta_{R_j}}\psi_{R,j}\rangle| \lesssim_\eta \mu(Q_i)^{1/2}\mu(R_j)^{1/2}.$$

The relevant series with these matrix elements are then bounded by

$$C(\eta) \sum_{R \in \mathcal{D}'} \mu(Q_i)^{1/2} |\langle f_Q \rangle_{Q_i}| \cdot \mu(R_j)^{1/2} |\langle g_R \rangle_{R_j}|$$

$$\leq C(\eta) \Big(\sum_{Q \in \mathcal{D}} \mu(Q_i) |\langle f_Q \rangle_{Q_i}|^2 \Big)^{1/2} \Big(\sum_{R \in \mathcal{D}'} \mu(R_j) |\langle g_R \rangle_{R_j}|^2 \Big)^{1/2}$$

$$\lesssim C(\eta) \|f\|_2 \|g\|_2 = C(\eta).$$

Next, we have

$$|\langle T(\chi_{Q_i} \varphi_{Q,i}), \chi_{R_{j,\partial}} \psi_{R,j} \rangle| \lesssim \|T\| \mu(Q_i)^{1/2} \|\chi_{\delta_{Q_i}^\eta} \chi_{R_j} \psi_{R,j}\|_2.$$

Thus, there holds

$$E_{\mathcal{D}} \Big| \sum_{R \in \mathcal{D}'} \langle f_Q \rangle_{Q_i} \langle T(\chi_{Q_i} \varphi_{Q,i}), \chi_{R_{j,\partial}} \psi_{R,j} \rangle \langle g_R \rangle_{R_j} \Big|$$

$$\leq \|T\| E_{\mathcal{D}} \Big(\sum_{Q \in \mathcal{D}} \mu(Q_i) |\langle f_Q \rangle_{Q_i}|^2 \Big)^{1/2} \Big(\sum_{R \in \mathcal{D}'} \|\chi_{\delta_{Q_i}^\eta} \chi_{R_j} \psi_{R,j}\|_2^2 |\langle g_R \rangle_{R_j}|^2 \Big)^{1/2}$$

$$\lesssim \|T\| \|f\|_2 E_{\mathcal{D}} \Big(\sum_{R \in \mathcal{D}'} \|\chi_{\delta_{Q_i}^\eta} \chi_{R_j} \psi_{R,j}\|_2^2 |\langle g_R \rangle_{R_j}|^2 \Big)^{1/2}$$

$$\lesssim \|T\| \|f\|_2 \Big(\sum_{R \in \mathcal{D}'} \|E_{\mathcal{D}}(\chi_{\delta_{Q_i}^\eta}) \chi_{R_j} \psi_{R,j}\|_2^2 |\langle g_R \rangle_{R_j}|^2 \Big)^{1/2}$$

$$\leq c(\eta) \|T\| \|f\|_2 \Big(\sum_{R \in \mathcal{D}'} \mu(R_j) |\langle g_R \rangle_{R_j}|^2 \Big)^{1/2}$$

$$\leq c(\eta) \|T\| \|f\|_2 \|g\|_2 = c(\eta) \|T\|,$$

where $c(\eta) \to 0$ if $\eta \to 0$. A similar estimate holds also with the matrix element $\langle T(\chi_{Q_{i,\partial}} \varphi_{Q,i}), \chi_{\Delta_{R_j}} \psi_{R,j} \rangle$.

We are left to deal with $\langle T(\chi_{\Delta_{Q_i}} \varphi_{Q,i}), \chi_{\Delta_{R_j}} \psi_{R,j} \rangle$. Choose $j(\eta) \in \mathbb{Z}$ so that $\eta/64 \leq 2^{j(\eta)} < \eta/32$. Let $\mathcal{D}^*$ be another independent grid (e.g. choose a large cube U_0 at random so that $Q_0 \cup R_0 \subset U_0$ always, and use that as the starting cube of the grid $\mathcal{D}^*$). Let $s = 2^{j(\eta)} \ell(Q_i)$ and $G = G(R) = \mathcal{D}^*_{-\log_2 s}$.

We enlarge the sets Δ_{Q_i} and Δ_{R_j} to obtain new sets $\Delta_{Q_i}^G$ and $\Delta_{R_j}^G$ so that $\Delta_{Q_i}^G \cap \Delta_{R_j}^G = \bigcup \{g : g \in G, g \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G\}$. This is done so that $\Delta_{Q_i} = \Delta_{Q_i}^G \setminus \Delta_{Q_i}^\partial$, $\Delta_{R_j} = \Delta_{R_j}^G \setminus \Delta_{R_j}^\partial$, where $\Delta_{Q_i}^\partial \subset Q_{i,\partial}$ and $\Delta_{R_j}^\partial \subset R_{j,\partial}$. Furthermore, we may perform this so that $5g \subset Q_i \cap R_j$ if $g \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G$.

Let us now write

$$\langle T(\chi_{\Delta_{Q_i}} \varphi_{Q,i}), \chi_{\Delta_{R_j}} \psi_{R,j} \rangle = \langle T(\chi_{\Delta_{Q_i}^G} \varphi_{Q,i}), \chi_{\Delta_{R_j}^G} \psi_{R,j} \rangle$$

$$- \langle T(\chi_{\Delta_{Q_i}^G} \varphi_{Q,i}), \chi_{\Delta_{R_j}^\partial} \psi_{R,j} \rangle - \langle T(\chi_{\Delta_{Q_i}^\partial} \varphi_{Q,i}), \chi_{\Delta_{R_j}} \psi_{R,j} \rangle.$$

The series which has the sum of the last two terms as its matrix element is, after averaging, dominated by $c(\eta)\|T\|$ by the very same argument used above. We fix at this point η to be so small that the above four η-boundary region terms contribute no more than $c(\eta)C\|T\| < \|T\|/32$.

7.A.2. *Arguments involving ϵ-boundary regions.* We are reduced to consider the pairing $\langle T(\chi_{\Delta_{Q_i}^G}\varphi_{Q,i}), \chi_{\Delta_{R_j}^G}\psi_{R,j}\rangle$. Let $\epsilon > 0$, and set $G_\epsilon = G_\epsilon(R) = \bigcup_{g \in G} \delta_g^\epsilon$, where $\delta_g^\epsilon = (1+\epsilon)g \setminus (1-\epsilon)g$.

We define $\Delta_{Q_i}' = \Delta_{Q_i}^G \cap G_\epsilon$ and $\tilde{\Delta}_{Q_i} = \Delta_{Q_i}^G \setminus G_\epsilon$ (and similarly for R). We then write

$$\langle T(\chi_{\Delta_{Q_i}^G}\varphi_{Q,i}), \chi_{\Delta_{R_j}^G}\psi_{R,j}\rangle = \langle T(\chi_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}), \chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle$$
$$+ \langle T(\chi_{\Delta_{Q_i}^G}\varphi_{Q,i}), \chi_{\Delta_{R_j}'}\psi_{R,j}\rangle + \langle T(\chi_{\Delta_{Q_i}'}\varphi_{Q,i}), \chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle.$$

There holds $|\langle T(\chi_{\Delta_{Q_i}^G}\varphi_{Q,i}), \chi_{\Delta_{R_j}'}\psi_{R,j}\rangle| \lesssim \|T\|\mu(Q_i)^{1/2}\|\chi_{G_\epsilon}\chi_{R_j}\psi_{R,j}\|_2$. Again, we have $E_{\mathcal{D}^*}\chi_{G_\epsilon}(x) \le c(\epsilon)$, where $c(\epsilon) \to 0$ when $\epsilon \to 0$. Therefore, the series which has the sum of the last two terms as its matrix element is, after averaging, dominated by $c(\epsilon)\|T\|$.

We are now left with $\langle T(\chi_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}), \chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle$. It suffices to consider pairings $\langle T(\chi_{g_1}\chi_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}), \chi_{g_2}\chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle$ for $g_1, g_2 \in G$, as there are only boundedly many (depending on η – but this is fixed) cubes in G which matter. Suppose first that $g_1 \neq g_2$. Then, because of separation, $|\langle T(\chi_{g_1}\chi_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}), \chi_{g_2}\chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle| \lesssim_\epsilon \mu(Q_i)^{1/2}\mu(R_j)^{1/2}$. This implies, like above, that the relevant series with this matrix element is dominated by $C(\epsilon)\|f\|_2\|g\|_2 = C(\epsilon)$.

This time we are left with $T(\chi_g\chi_{\tilde{\Delta}_{Q_i}}\varphi_{Q,i}), \chi_g\chi_{\tilde{\Delta}_{R_j}}\psi_{R,j}\rangle$ for some $g \in G$. We may write this in the form $\langle T(\chi_H\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle$, where $H = g \setminus G_\epsilon \subset \Delta_{Q_i}^G \cap \Delta_{R_j}^G$ is a cube (otherwise the pairing vanishes by construction). We continue to decompose

$$\langle T(\chi_H\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle = \langle T(\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle$$
$$- \langle T(\chi_{\mathbb{R}^n \setminus 5H}\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle$$
$$- \langle T(\chi_{5H \setminus (1+\epsilon)H}\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle$$
$$- \langle T(\chi_{(1+\epsilon)H \setminus H}\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle.$$

We have $|\langle T(\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle| \le \|\chi_{Q_i}T(\varphi_{Q,i})\|_2\|\chi_{R_j}\psi_{R,j}\|_2 \lesssim \mu(Q_i)^{1/2}\mu(R_j)^{1/2}$. We have by separation (recall that $\ell(H) \sim \ell(Q_i) \sim \ell(R_j)$) and the fact that $5H \subset Q_i \cap R_j$ that $|T(\chi_{5H \setminus (1+\epsilon)H}\varphi_{Q,i}), \chi_H\psi_{R,j})| \lesssim_\epsilon \mu(Q_i)^{1/2}\mu(R_j)^{1/2}$. Also, there yet again holds that $|\langle T(\chi_{(1+\epsilon)H \setminus H}\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle| \lesssim \|T\|\|\chi_{G_\epsilon}\chi_{Q_i}\varphi_{Q,i}\|_2\mu(R_j)^{1/2}$, which is fine after averaging as before.

Having disposed of the terms above, we are to handle $\langle T(\chi_{\mathbb{R}^n \setminus 5H}\varphi_{Q,i}), \chi_H\psi_{R,j}\rangle$. We let $\tau = T(\chi_{\mathbb{R}^n \setminus 5H}\varphi_{Q,i})$ and $\beta_H = \langle b_H^2/\mu(H), \tau\rangle$, and then decompose

$$\langle \tau, \chi_H\psi_{R,j}\rangle = \langle \tau - \beta_H, \chi_H\psi_{R,j}\rangle + \beta_H \int_H \psi_{R,j}\, d\mu.$$

7.3. **Lemma.** *We have*

$$|\langle \tau - \beta_H, \chi_H \psi_{R,j} \rangle| \lesssim \mu(Q_i)^{1/2} \mu(R_j)^{1/2}.$$

Proof. There holds

$$|\langle \tau - \beta_H, \chi_H \psi_{R,j} \rangle| \leq \int |\chi_H(x) \psi_{R,j}(x)||\tau(x) - \beta_H| \, d\mu(x).$$

We have (using $\frac{1}{\mu(H)} \int_H b_H^2 \, d\mu = 1$) that

$$\begin{aligned}
|\tau(x) - \beta_H| &= \left| \tau(x) - \frac{1}{\mu(H)} \int_H b_H^2(y) \tau(y) \, d\mu(y) \right| \\
&= \left| \frac{1}{\mu(H)} \int_H b_H^2(y)[\tau(x) - \tau(y)] \, d\mu(y) \right| \\
&\leq \frac{1}{\mu(H)} \int_H |b_H^2(y)||\tau(x) - \tau(y)| \, d\mu(y).
\end{aligned}$$

Furthermore, we have for $x, y \in H$ that

$$\begin{aligned}
|\tau(x) - \tau(y)| &\leq \int_{\mathbb{R}^n \setminus 5H} |K(x,z) - K(y,z)||\varphi_{Q,i}(z)| \, d\mu(z) \\
&\lesssim \ell(H)^\alpha \int_{|x-z|>c\ell(H)} \frac{|\varphi_{Q,i}(z)|}{|x-z|^\alpha \lambda(x, |x-z|)} \, d\mu(z) \lesssim M_\mu \varphi_{Q,i}(x).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
|\tau(x) - \beta_H| &\lesssim \left(\frac{1}{\mu(H)} \int_H |b_H^2(y)| \, d\mu(y) \right) M_\mu \varphi_{Q,i}(x) \\
&\leq \left(\frac{1}{\mu(H)} \int_H |b_H^2(y)|^2 \, d\mu(y) \right)^{1/2} M_\mu \varphi_{Q,i}(x) \lesssim M_\mu \varphi_{Q,i}(x).
\end{aligned}$$

We conclude that

$$\begin{aligned}
|\langle \tau - \beta_H, \chi_H \psi_{R,j} \rangle| &\lesssim \int |\chi_H(x) \psi_{R,j}(x)||\chi_H(x) M_\mu \varphi_{Q,i}(x)| \, d\mu(x) \\
&\leq \left(\int_{Q_i} |M_\mu \varphi_{Q,i}|^2 \, d\mu \right)^{1/2} \left(\int_{R_j} |\psi_{R,j}|^2 \, d\mu \right)^{1/2} \lesssim \mu(Q_i)^{1/2} \mu(R_j)^{1/2},
\end{aligned}$$

where the last estimate follows by noting that in the L^∞ case $|M_\mu \varphi_{Q,i}| \leq \|\varphi_{Q,i}\|_\infty \lesssim 1$, and that in the L^2 case this also works out by the stopping time and the doubling property of the measure. $\square$

Finally, we are to deal with $\beta_H \int_H \psi_{R,j} \, d\mu$. The absolute value of this is dominated by

$$|\beta_H| \mu(H) \left(\frac{1}{\mu(H)} \int_H |\psi_{R_j}|^2 \, d\mu \right)^{1/2} \lesssim |\beta_H| \mu(H),$$

where we used that $H \subset Q_i \cap R_j$ and that in the doubling case $\mu = \nu$ we have $\nu(R) \lesssim \nu(R_j) \lesssim \nu(H)$. We then write

$$\beta_H \mu(H) = \langle b_H^2, T(\varphi_{Q,i}) \rangle - \langle b_H^2, T(\chi_{5H \setminus (1+\epsilon)H} \varphi_{Q,i}) \rangle$$
$$- \langle b_H^2, T(\chi_{(1+\epsilon)H \setminus H} \varphi_{Q,i}) \rangle - \langle b_H^2, T(\chi_H \varphi_{Q,i}) \rangle.$$

We can now deal with all of these remaining terms (after which we are finally left with nothing more to estimate). Let us do this now. Recalling that $5H \subset Q_i$, we have

$$|\langle b_H^2, T(\varphi_{Q,i}) \rangle| \le \|b_H^2\|_2 \|\chi_{Q_i} T(\varphi_{Q,i})\|_2 \lesssim \mu(H)^{1/2} \mu(Q_i)^{1/2},$$

$$|\langle b_H^2, T(\chi_{5H \setminus (1+\epsilon)H} \varphi_{Q,i}) \rangle|$$
$$\lesssim_\epsilon \Big(\int_H |b_H^2|^2 \, d\mu \Big)^{1/2} \Big(\int_{Q_i} |\varphi_{Q,i}|^2 \, d\mu \Big)^{1/2} \lesssim \mu(H)^{1/2} \mu(Q_i)^{1/2},$$

$$|\langle b_H^2, T(\chi_{(1+\epsilon)H \setminus H} \varphi_{Q,i}) \rangle| \lesssim \|T\| \mu(H)^{1/2} \|\chi_{G_\epsilon} \chi_{Q_i} \varphi_{Q,i}\|_2$$

and

$$|\langle b_H^2, T(\chi_H \varphi_{Q,i}) \rangle| = |\langle T^* b_H^2, \chi_H \varphi_{Q,i} \rangle|$$
$$\le \|\chi_H T^* b_H^2\|_2 \|\chi_{Q_i} \varphi_{Q,i}\|_2 \lesssim \mu(H)^{1/2} \mu(Q_i)^{1/2}.$$

Thus, using $H \subset R_j$, we have shown that

$$\Big| \beta_H \int_H \psi_{R,j} \, d\mu \Big| \le C(\epsilon) \mu(H)^{1/2} \mu(Q_i)^{1/2} + C\|T\| \mu(H)^{1/2} \|\chi_{G_\epsilon} \chi_{Q_i} \varphi_{Q,i}\|_2$$
$$\le C(\epsilon) \mu(R_j)^{1/2} \mu(Q_i)^{1/2} + C\|T\| \mu(R_j)^{1/2} \|\chi_{G_\epsilon} \chi_{Q_i} \varphi_{Q,i}\|_2.$$

Furthermore, there holds

$$E_{\mathcal{D}^*} \|\chi_{G_\epsilon} \chi_{Q_i} \varphi_{Q,i}\|_2 \le c(\epsilon) \|\chi_{Q_i} \varphi_{Q_i}\|_2 \lesssim c(\epsilon) \mu(Q_i)^{1/2}.$$

We have proved that the series with the coefficient $\langle T(\chi_{\mathbb{R}^n \setminus 5H} \varphi_{Q,i}), \chi_H \psi_{R,j} \rangle$ can be dominated by $C(\epsilon) + c(\epsilon)\|T\|$. We now fix ϵ to be so small (this depends on the already fixed parameter η) that all of the ϵ-boundary region terms contribute no more than $c(\epsilon) C\|T\| < \|T\|/32$.

We have proven the following proposition in this section.

7.4. **Proposition.** *There holds*

$$|E\Sigma_3| \le C + \|T\|/16.$$

7.5. *Remark.* We still detail on the term $\mu(Q_0)^{-1/2} \mu(R_0)^{-1/2} |\langle T b_{Q_0}^1, b_{R_0}^2 \rangle|$, which we did not yet estimate in Section 4. On the right hand side of the pairing write

$$\chi_{R_0} = \chi_{R_0 \cap Q_0} + \chi_{(R_0 \setminus Q_0) \cap \delta_{Q_0}^\theta} + \chi_{R_0 \setminus Q_0 \setminus \delta_{Q_0}^\theta}$$

with some small parameter $\theta > 0$. Notice that to estimate the first pairing thus formed is this time trivial, since it can be bounded by $\|\chi_{Q_0} T b_{Q_0}^1\| \|b_{R_0}^2\|_2$, and this

is dominated by $\mu(Q_0)^{1/2}\mu(R_0)^{1/2}$. The third pairing so formed can be bounded using separation, and this yields $C(\theta)\mu(Q_0)^{1/2}\mu(R_0)^{1/2}$. Finally, we have

$$E_{\mathcal{D}}[\mu(Q_0)^{-1/2}\mu(R_0)^{-1/2}|\langle Tb_{Q_0}^1, \chi_{(R_0\setminus Q_0)\cap\delta_{Q_0}^\theta}b_{R_0}^2\rangle|] \leq C\|T\|E_{\mathcal{D}}[\mu(R_0)^{-1/2}\|\chi_{\delta_{Q_0}^\theta}b_{R_0}^2\|_2]$$
$$\leq c(\theta)C\|T\| < \|T\|/12$$

fixing θ to be small enough.

8. COMPLETION OF THE PROOF

Collecting the above estimates for Σ_i, $i = 1, 2, 3$, (and for their symmetric counterparts and for the term $\mu(Q_0)^{-1/2}\mu(R_0)^{-1/2}|\langle Tb_{Q_0}^1, b_{R_0}^2\rangle|$), we have established that

$$\|T\|/2 \leq |\langle Tf, g\rangle| \leq C + \|T\|/3,$$

and from this we may conclude that $\|T\| \leq 6C$.

REFERENCES

[AHM+02] P. Auscher, S. Hofmann, C. Muscalu, T. Tao, and C. Thiele, *Carleson measures, trees, extrapolation, and T(b) theorems*, Publicacions Matemátiques **46** (2002), no. 2, 257–325.

[AR10] P. Auscher and E. Routin, *Local Tb theorems and Hardy inequalities*, manuscript (2010).

[AY09] Pascal Auscher and Qi Xiang Yang, *BCR algorithm and the T(b) theorem*, Publ. Mat. **53** (2009), no. 1, 179–196.

[Chr90] Michael Christ, *A T(b) theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601–628.

[HM09] Tuomas Hytönen and Henri Martikainen, *Non-homogeneous Tb theorem and random dyadic cubes on metric measure spaces*, preprint (2009), arXiv:0911.4387.

[Hof07] Steve Hofmann, *A proof of the local Tb theorem for standard Calderón-Zygmund operators*, unpublished manuscript (2007), arXiv:0705.0840.

[Hyt09] Tuomas Hytönen, *The vector-valued non-homogeneous Tb theorem*, preprint (2009), arXiv:0809.3097.

[Hyt10a] Tuomas Hytönen, *A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa*, Publ. Mat. **54** (2010), no. 2, 485–504.

[Hyt10b] Tuomas Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, preprint (2010), arXiv:1007.4330.

[HYY10] Tuomas Hytönen, Dachun Yang, and Dongyong Yang, *The Hardy space H^1 on non-homogeneous metric spaces*, preprint (2010), arXiv:1008.3831.

[Mar10] Henri Martikainen, *Vector-valued non-homogeneous Tb theorem on metric measure spaces*, preprint (2010), arXiv:1004.3176.

[NTV97] F. Nazarov, S. Treil, and A. Volberg, *Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces*, Internat. Math. Res. Notices (1997), no. 15, 703–726.

[NTV02] ———, *Accretive system Tb-theorems on nonhomogeneous spaces*, Duke Math. J. **113** (2002), no. 2, 259–312.

[NTV03] ———, *The Tb-theorem on non-homogeneous spaces*, Acta Math. **190** (2003), no. 2, 151–239.

[TY09] C. Tan and L. Yan, *Local Tb theorem on spaces of homogeneous type*, Z. Anal. Anwend. **28** (2009), no. 3, 333–347.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 HELSINKI, FINLAND

E-mail address: tuomas.hytonen@helsinki.fi, henri.martikainen@helsinki.fi