

Non-homogeneous Tb Theorem and Random Dyadic Cubes on Metric Measure Spaces

Tuomas Hytönen · Henri Martikainen

Received: 27 November 2009 / Published online: 23 April 2011
© Mathematica Josephina, Inc. 2011

Abstract We prove a Tb theorem on quasimetric spaces equipped with what we call an upper doubling measure. This is a property that encompasses both the doubling measures and those satisfying the upper power bound $\mu(B(x, r)) \leq Cr^d$. Our spaces are only assumed to satisfy the geometric doubling property: every ball of radius r can be covered by at most N balls of radius $r/2$. A key ingredient is the construction of random systems of dyadic cubes in such spaces.

Keywords Calderón–Zygmund operator · Non-doubling measure · Probabilistic constructions in metric spaces

Mathematics Subject Classification (2000) Primary 42B20 · Secondary 30L99 · 60D05

1 Introduction

There is a wide range of Tb theorems. The original formulations are by McIntosh and Meyer [8] and David, Journé, and Semmes [4]. The arguments of [4] carry over to a metric space—however, only if one is working with doubling measures. In the introduction to their celebrated paper on the non-homogeneous Tb theorem

Communicated by Michael Lacey.

The authors are supported by the Academy of Finland through projects 130166 and 218148 (“ L^p methods in harmonic analysis”) and 133264 (“Stochastic and harmonic analysis, interactions and applications”).

T. Hytönen (✉) · H. Martikainen

Department of Mathematics and Statistics, University of Helsinki, P.O.B. 68, 00014 Helsinki, Finland
e-mail: tuomas.hytonen@helsinki.fi

H. Martikainen

e-mail: henri.martikainen@helsinki.fi

[11, p. 153], Nazarov, Treil, and Volberg point out that “a (more or less) complete theory of Calderón–Zygmund operators on non-homogeneous spaces—can be developed in an abstract metric space with measure.” Although a number of results for non-homogeneous singular integral operators (such as the weak-type L^1 inequality under a priori L^2 boundedness [10], estimates on Lipschitz spaces [5], and a certain restricted version of the $T1$ theorem [1]—essentially with $T1, T^*1 \in L^\infty$, and only on bounded spaces) have been established in quite general metric measure spaces, it seems that the quoted remark has not been fully elaborated for the deeper aspects of the theory. The goal of this paper is to close this gap, and in fact obtain a new level of generality even in the context of \mathbb{R}^n .

Let us describe our setting in more detail. We consider quasimetric spaces with the following well-established postulate, which we refer to as geometric doubling: every ball of radius r can be covered by at most N balls of radius $r/2$. This is essentially the original definition of “a space of homogeneous nature” by Coifman and Weiss [3, pp. 66–67], although this name is now commonly used for quasimetric spaces equipped with a doubling measure, the “particularly important case” pointed out by Coifman and Weiss immediately after their general definition. It is known that if a metric space is geometrically doubling and complete, then it also supports some doubling measures [7]; however, we do not assume completeness, and we regard our measure of interest as given by the problem at hand, and not as something that one is free to choose or construct.

We consider so-called upper doubling measures μ , introduced in [6], which constitute a simultaneous generalization of doubling measures and those with the upper power bound property $\mu(B(x, r)) \leq Cr^d$, which are the ones usually considered in the literature on non-homogeneous analysis. But note that power bounded measures are only different from, not more general than, the doubling measures. While the original motivation behind the notion of upper doubling in [6] was to find a natural unified framework for the doubling and power bounded theories, it also encapsulates other examples of interest. Indeed, although it was not our specific goal, this general framework allows to essentially recapture the recent $T1$ theorem of Volberg and Wick [13] for “Bergman-type” operators.

We now discuss the general strategy and some new aspects of the proof. We follow the basic approach from [11] and try to adapt the treatment of their most general cases to our situation. First, the assumption that μ is merely upper doubling causes, for example, the effect that the bounds for $\mu(B(x, r))$ depend not only on r but also on x . We formulate the kernel estimates in a natural way adapted to this, and carry out all the estimates with this extra generality. Second, the fact that we work in abstract quasimetric spaces complicates many things. However, note that parts of the relevant BMO and RBMO aspects of this theory were already dealt with in [6].

A key ingredient behind the proof of Nazarov et al. [11] is the random choice of a system of dyadic cubes, so that certain “bad” situations can be handled by arguing that their occurrence has only a small probability. It is then clear that we need something similar in an abstract quasimetric space. For a system of dyadic cubes as such, there is a well-known construction due to Christ [2], which serves as our starting point. Even then, it is not obvious how to choose a random system, since the randomization procedure of Nazarov et al. heavily relies on the action of the translation group on \mathbb{R}^n .

Our solution to this problem, which is based on randomly choosing new “centers” for the dyadic cubes of generation k among the old centers of the smaller cubes of generation $k + 1$, appears to be new, and it may also be of independent interest, besides the present application to the Tb theorem. In this respect, we note that a family of dyadic systems in a quasi-metric space, rather than just a fixed one, was already exploited by Verbitsky and Wheeden [14] in the context of weighted norm inequalities, but they required the underlying space to have a group structure, so that the new dyadic systems could still be obtained by simply shifting a given one, in analogy to the Euclidean setting.

We shall also employ a closely related construction of random almost-coverings of the space by balls of comparable radius, by which we mean that any given point has a small probability of not being covered. The need for this is related to the fact that, unlike in \mathbb{R}^n , it now seems far more natural to formulate notions like the weak boundedness property and the BMO space in terms of balls rather than cubes, and so we essentially need to cut our dyadic cubes into comparable balls when estimating the “diagonal” part of the operator.

In the following section, we give detailed statements of the results discussed here, which are then proven in the rest of the paper. In the final section, we describe the relation to the above-mentioned results of Volberg and Wick [13].

2 Preliminaries and the Main Result

We now give the detailed definitions, fix some notation and parameters, and then formulate our main theorem, Theorem 2.3.

2.1 Geometrically Doubling Regular Quasimetric Spaces

Recall that a quasimetric is almost like a metric but the triangle inequality is replaced by the requirement that for some $A_0 \geq 1$ it holds that $\rho(x, y) \leq A_0(\rho(x, z) + \rho(z, y))$ for all $x, y, z \in X$. A quasimetric space (X, ρ) is geometrically doubling if every open ball $B(x, r) = \{y \in X : \rho(y, x) < r\}$ can be covered by at most N balls of radius $r/2$. We use this somewhat non-standard name to clearly differentiate this property from other types of doubling properties. We adapt the convention that a ball B is equipped with a fixed center $c_B \in X$ and radius $r_B > 0$ (if no other notation is in place, we use this). Also, we set $n = \log_2 N$, which can be viewed as (an upper bound for) a geometric dimension of the space. Let us state the following well-known lemma.

Lemma 2.1 *In a geometrically doubling quasimetric space, a ball $B(x, r)$ can contain the centers x_i of at most $N\alpha^{-n}$ disjoint balls $B(x_i, \alpha r)$ for $\alpha \in (0, 1]$.*

Note also that there is a uniform constant depending only on N and A_0 so that all subsets of X are geometrically doubling with this constant. Choosing N large enough in the first place, let us use the same constant N everywhere.

For many purposes, quasimetrics are just as good as metrics, only somewhat more annoying to deal with due to the presence of the additional constant in the triangle

inequality. However, for some of the more delicate estimates, it seems to us that general quasimetrics can be a bit too wild, and we always ask that our quasimetric ρ satisfy the following regularity property: for every $\epsilon > 0$ there exists $A(\epsilon) < \infty$ so that

$$\rho(x, y) \leq (1 + \epsilon)\rho(x, z) + A(\epsilon)\rho(z, y).$$

Notice that this property is in particular satisfied by all positive powers of an honest metric, and every quasimetric is equivalent to one of that form by a well-known result of Macías and Segovia [9]. While it is easy to cook up irregular quasimetrics, it seems that practically all reasonable examples of quasimetrics from applications already satisfy the regularity property even without passing to an equivalent version. This is in particular the case for all the examples of quasi-metrics pointed out by Coifman and Weiss [3, p. 68].

Much of our subsequent assumptions will be essentially invariant under the change to an equivalent quasimetric, which we explicitly exploit through the mentioned result of Macías and Segovia, so that a large part of the proof can be carried out in an honest metric space. However, we want to use indicators of balls (of the given quasimetric) as test functions, and it is here that the general quasimetric balls seem to be somewhat too arbitrary for our purposes.

2.2 Upper Doubling Measures

A Borel measure μ in a quasimetric space (X, ρ) is called upper doubling if there exists a dominating function $\lambda: X \times (0, \infty) \rightarrow (0, \infty)$ so that $r \mapsto \lambda(x, r)$ is non-decreasing, $\lambda(x, 2r) \leq C_\lambda \lambda(x, r)$ and $\mu(B(x, r)) \leq \lambda(x, r)$ for all $x \in X$ and $r > 0$. The number $d := \log_2 C_\lambda$ can be thought of as (an upper bound for) a dimension of the measure μ , and it will play a similar role as the quantity denoted by the same symbol in [11].

Lemma 2.2 *We have for every ball $B = B(c_B, r_B)$ and for every $\epsilon > 0$ that*

$$\int_{X \setminus B} \frac{\rho(x, c_B)^{-\epsilon}}{\lambda(c_B, \rho(x, c_B))} d\mu(x) \leq C_\lambda A_\epsilon r_B^{-\epsilon},$$

where $A_\epsilon = 2^\epsilon / (2^\epsilon - 1)$.

Proof We calculate

$$\begin{aligned} \int_{X \setminus B} \frac{\rho(x, c_B)^{-\epsilon}}{\lambda(c_B, \rho(x, c_B))} d\mu(x) &= \sum_{j=0}^{\infty} \int_{2^j r_B \leq \rho(x, c_B) < 2^{j+1} r_B} \frac{\rho(x, c_B)^{-\epsilon}}{\lambda(c_B, \rho(x, c_B))} d\mu(x) \\ &\leq \sum_{j=0}^{\infty} \frac{\mu(B(c_B, 2^{j+1} r_B))}{\lambda(c_B, 2^j r_B)} (2^j r_B)^{-\epsilon} \\ &\leq C_\lambda r_B^{-\epsilon} \sum_{j=0}^{\infty} 2^{-\epsilon j} = C_\lambda A_\epsilon r_B^{-\epsilon}, \end{aligned}$$

where we used that λ is non-decreasing and $\mu(B(c_B, 2^{j+1}r_B)) \leq \lambda(c_B, 2^{j+1}r_B) \leq C_\lambda \lambda(c_B, 2^j r_B)$. \square

In what follows, we work in a geometrically doubling regular quasimetric space (X, ρ) (with the constants N and n as above, and the related function $A(\epsilon)$), which is equipped with an upper doubling measure μ with the related majorant λ .

2.3 Standard Kernels and Calderón–Zygmund Operators

Define $\Delta = \{(x, x) : x \in X\}$. A standard kernel is a mapping $K : X^2 \setminus \Delta \rightarrow \mathbb{C}$ for which we have for some $\alpha > 0$ and $c, C < \infty$ that

$$|K(x, y)| \leq C \min\left(\frac{1}{\lambda(x, \rho(x, y))}, \frac{1}{\lambda(y, \rho(x, y))}\right), \quad x \neq y,$$

$$|K(x, y) - K(x', y)| \leq C \frac{\rho(x, x')^\alpha}{\rho(x, y)^\alpha \lambda(x, \rho(x, y))}, \quad \rho(x, y) \geq c\rho(x, x'),$$

and

$$|K(x, y) - K(x, y')| \leq C \frac{\rho(y, y')^\alpha}{\rho(x, y)^\alpha \lambda(y, \rho(x, y))}, \quad \rho(x, y) \geq c\rho(y, y').$$

The smallest admissible C will be denoted by $\|K\|_{CZ_\alpha}$; it is understood that the parameter c has been fixed, and it will not be indicated explicitly in this notation.

Let $T : f \mapsto Tf$ be a linear operator acting on some functions f (which we shall specify in more detail later). It is called a Calderón–Zygmund operator with kernel K if

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

for x outside the support of f .

2.4 Accretivity

A function $b \in L^\infty(\mu)$ is called accretive if $\operatorname{Re} b \geq a > 0$ almost everywhere. We can also make do with the following weaker form of accretivity: $|\int_A b d\mu| \geq a\mu(A)$ for all Borel sets A which satisfy the condition that $B \subset A \subset CB$ for some ball $B = B(A)$, where C is some large constant which depends on the quasimetric ρ . (One can, e.g., take $C = 500$ if dealing with metrics.)

The point is to have the above estimate whenever A is one of the “cubes” to be constructed below, but there is no easy explicit description of what kind of sets they actually are.

2.5 Weak Boundedness Property

An operator T is said to satisfy the weak boundedness property if $|\langle T\chi_B, \chi_B \rangle| \leq A\mu(\Lambda B)$ for all balls B and for some fixed constants $A > 0$ and $\Lambda > 1$. Here $\langle \cdot, \cdot \rangle$ is

the bilinear duality $\langle f, g \rangle = \int fg \, d\mu$. Let us denote the smallest admissible constant above by $\|T\|_{WBPA}$.

Literature abounds in variants of the weak boundedness property, and it is sometimes of interest to replace the rough test functions χ_B by some more regular ones. It is also possible in the present case. (We had originally not considered this question, which was posed to one of us by E.M. Stein after a presentation of an earlier version of this paper in SNS Pisa, March 2010.) Namely, suppose that for every ball B and every $\epsilon \in (0, 1]$ there is a function $\tilde{\chi}_{B,\epsilon}$ such that $\chi_B \leq \tilde{\chi}_{B,\epsilon} \leq \chi_{(1+\epsilon)B}$, and we have the estimate $|\langle T\tilde{\chi}_{B,\epsilon}, \tilde{\chi}_{B,\epsilon} \rangle| \leq S(\epsilon)A\mu(\Lambda B)$ with some $S(\epsilon) < \infty$ independent of the other quantities. We denote the smallest admissible A by $\|T\|_{WBPA,S}$. Clearly, if T satisfies the weak boundedness property in the original sense, we may take $\tilde{\chi}_{B,\epsilon} := \chi_B$ and $S(\epsilon) := 1$ for all $\epsilon > 0$, and we have $\|T\|_{WBPA,1} \leq \|T\|_{WBPA}$. Note that the generalized weak boundedness property simply asks the above inequality for some set of functions $\tilde{\chi}_{B,\epsilon}$, regular or not. Depending on the structure of the underlying space X , one may find such functions with different degrees of regularity.

In the Tb theorem, the weak boundedness property is demanded from the operator $M_{b_2}TM_{b_1}$, where b_1 and b_2 are accretive functions and $M_b : f \mapsto bf$.

2.6 BMO and RBMO

We say that $f \in L^1_{loc}(\mu)$ belongs to $BMO^p_k(\mu)$, if for any ball $B \subset X$ there exists a constant f_B such that

$$\left(\int_B |f - f_B|^p \, d\mu \right)^{1/p} \leq L\mu(\kappa B)^{1/p},$$

where the constant L does not depend on B .

Let $\varrho > 1$. A function $f \in L^1_{loc}(\mu)$ belongs to $RBMO(\mu)$ if there exists a constant L , and for every ball B , a constant f_B , such that one has

$$\int_B |f - f_B| \, d\mu \leq L\mu(\varrho B),$$

and, whenever $B \subset B_1$ are two balls,

$$|f_B - f_{B_1}| \leq L \left(1 + \int_{2B_1 \setminus B} \frac{1}{\lambda(c_B, d(x, c_B))} \, d\mu(x) \right).$$

We do not demand that f_B be the average $\langle f \rangle_B = \frac{1}{\mu(B)} \int_B f \, d\mu$, and this is actually important in the $RBMO(\mu)$ -condition. The useful thing here is that the space $RBMO(\mu)$ is independent of the choice of parameter $\varrho > 1$ and satisfies the John–Nirenberg inequality. For these results in our setting, see [6]. The norms in these spaces are defined in the obvious way as the best constant L .

We do not really need the $RBMO(\mu)$ space here as we formulate our main theorem with respect to $BMO^2_k(\mu)$ rather than $BMO^1_k(\mu)$. However, some reductions are possible here, and we comment on this after the formulation of our main theorem.

2.7 Vinogradov Notation and Implicit Constants

The notation $f \lesssim g$ is used synonymously with $f \leq Gg$ for some constant G . We also use $f \sim g$ if $f \lesssim g \lesssim f$. We now specify the parameters on which the implied constant G is allowed to depend in this notation. We let it depend on N and A_0 , which are related to the space X , and on C_λ , which is related to the measure μ . Next, we let G depend on c and α , the constants from the kernel estimates, and on the constant Λ and the function S related to the weak boundedness property. Then we let G depend on a , the constant related to the accretivity assumption of the test functions involved, and on the L^∞ -norms of the same test functions. We also let G depend on the BMO parameter κ . It is an inconvenient fact that one needs so many constants. Several more auxiliary parameters will be built of the aforementioned ones. Also, there will be quite a few other parameters that are not swallowed by this notation.

We now formulate our main theorem in full detail.

Theorem 2.3 *Let (X, ρ) be a geometrically doubling regular quasimetric space which is equipped with an upper doubling measure μ . Let T be an $L^2(\mu)$ -bounded Calderón–Zygmund operator with a standard kernel K , let b_1 and b_2 be two accretive functions, let $\alpha > 0$ and $\kappa, \Lambda > 1$ be constant, and $S : (0, 1] \rightarrow (0, \infty)$ a function. Then*

$$\|T\| \lesssim \|Tb_1\|_{BMO_\kappa^2(\mu)} + \|T^*b_2\|_{BMO_\kappa^2(\mu)} + \|M_{b_2}TM_{b_1}\|_{WB P_{\Lambda,S}} + \|K\|_{CZ_\alpha},$$

where the first three terms on the right are in turn dominated by $\|T\| + \|K\|_{CZ_\alpha}$.

Here all the estimates depend on the quasimetric space (X, ρ) through the constants N, A_0 and the function $\epsilon \mapsto A(\epsilon)$, the measure μ through the constant C_λ , the test functions b_1 and b_2 through the L^∞ -norms and the accretivity constants, the constants c and α appearing in the standard estimates, as well as the constant Λ and the function $\epsilon \mapsto S(\epsilon)$ in the weak boundedness property.

Remark 2.4

- (i) We have made the assumption that T is a priori known to be a bounded operator on $L^2(\mu)$, but one can reduce to this situation in several well-known ways (and sometimes no reduction is necessary). In what follows, we shall take the two BMO-norms, the weak boundedness constant, and the Calderón–Zygmund constant to be 1, and show that $\|T\| \lesssim 1$ with the conventions agreed upon above. It will be clear from the proof that the dependence on the mentioned quantities is of the asserted form. The converse direction of the theorem is standard.
- (ii) At least in the case that ρ is an honest metric, one can work with the larger space $BMO_\kappa^1(\mu)$ too; see [11] and [6]. Here one passes through the RBMO(μ) space, uses the John–Nirenberg inequality there, and then returns to the $BMO_\kappa^2(\mu)$ setting. In other words, $Tb_1 \in BMO_\kappa^1(\mu)$ implies that actually $Tb_1 \in BMO_\kappa^2(\mu)$. (This utilizes the fact that Tb_1 is in the range of the Calderón–Zygmund operator T , e.g., via the weak boundedness property; nothing of this sort holds for a general $f \in BMO_\kappa^1(\mu)$.)

- (iii) As will be explained in the following section, all assumptions except possibly for the weak boundedness property are stable under the change to an equivalent quasimetric. Hence the only place where the regularity assumption on ρ plays a role is in the balls involved in the weak boundedness property. We do not know whether the theorem remains true if the weak boundedness property is assumed with respect to balls of an irregular quasimetric. On the other hand, since these balls could be quite wild, it is questionable if such a testing condition would even be very useful.

3 Reduction to Metric Spaces

A result of Macías and Segovia (see the proof of [9, Theorem 2]) implies that there is a metric d and a constant $\beta \geq 1$ such that $2^{-\beta}d(x, y)^\beta \leq \rho(x, y) \leq 4^\beta d(x, y)^\beta$ for all $x, y \in X$. Here $3A_0^2 = 2^\beta$. One checks that (X, d) is then geometrically doubling with the constant $N16^{\beta n}$. Choosing N large enough in the first place lets us again use the same constant for both spaces. The measure μ is also upper doubling with respect to d with the function $(x, r) \mapsto \lambda(x, 4^\beta r^\beta)$. As we no longer have any use for the original λ , we replace it with this one.

It also follows that Tb_1 and T^*b_2 belong to the space $BMO_\kappa^2(\mu)$ with respect to d -balls, if we just simply replace the original κ by $8\kappa^{1/\beta}$. Of course, if the accretivity is assumed in the form $\operatorname{Re} b \geq a > 0$, this requires no modifications. If we assume it in the weaker form, one sees that we can, e.g., take $C = 4000^\beta$. Then we have that if $B_d \subset A \subset 500B_d$ for some d -ball B_d , then $B_\rho \subset A \subset CB_\rho$ for some ρ -ball. The proof that follows shows that the constant 500 works, as certain dyadic cubes with respect to d are sets of this form.

There seems to be no easy way to immediately conclude the weak boundedness property also for the d -balls. Thus, we shall not even attempt anything of this sort. Instead, we shall explicitly use the quasimetric ρ in a certain random ball covering, and as the reader will see, this circumvents the problem.

It remains to speak about the kernel estimates with respect to d and the new λ (which works for d -balls). An easy calculation using the facts that λ is non-decreasing and doubling shows that the first kernel estimate holds, just with a larger constant C . The rest of the kernel estimates also hold, just with the original α replaced by $\beta\alpha$ and demanding C to be large enough. This ends our reduction. As stated, from now on for the most part we deal with just the metric space (X, d) , and use the original ρ only in one carefully indicated place.

4 Dyadic Systems of Cubes

We now provide a variant of Christ's [2] construction of dyadic cubes in a metric space. His original result was formulated assuming the presence of a doubling measure, but most of the argument actually employs geometric doubling only. However, Christ only proved the covering property of his cubes in an a.e. sense with respect

to the doubling reference measure. We want to avoid this, which leads us to the construction of a system of “half-open” cubes, which exactly partition the whole space at every length scale, just like in \mathbb{R}^n .

The construction involves a parameter $\delta \leq 1/1000$. For each $k \in \mathbb{Z}$ we are given a collection of points x_α^k for which $d(x_\alpha^k, x_\beta^k) \geq \delta^k/8$ for all $\alpha \neq \beta$ and $\min_\alpha d(x, x_\alpha^k) < 4\delta^k$ for every $x \in X$. The parameters $1/8$ and 4 are used simply because they will do in a certain randomization procedure of points (see Sect. 10).

Let us construct a certain transitive relation \leq among the pairs (k, α) , following [2]. For each (k, α) there exists at least one β for which $d(x_\alpha^k, x_\beta^{k-1}) < 4\delta^{k-1}$. Also, there exists at most one β for which $d(x_\alpha^k, x_\beta^{k-1}) < \delta^{k-1}/16$. The ordering \leq is constructed using the rules we now describe. Consider any pair (k, α) . Check first whether there exists β so that $d(x_\alpha^k, x_\beta^{k-1}) < \delta^{k-1}/16$. If so, set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Otherwise, choose any β for which $d(x_\alpha^k, x_\beta^{k-1}) < 4\delta^{k-1}$, and set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Extend by transitivity to obtain a partial ordering.

The dyadic cubes of Christ are defined by

$$Q_\alpha^k = \bigcup_{(\ell, \beta) \leq (k, \alpha)} B(x_\beta^\ell, \delta^\ell/100).$$

However, we aim to replace them by the “half-open” cubes advertised before.

One can easily check that $\overline{Q}_\alpha^k \subset B(x_\alpha^k, 5\delta^k)$. Also, we have the property that if $\ell \geq k$, then either $Q_\beta^\ell \subset Q_\alpha^k$ or $Q_\beta^\ell \cap \overline{Q}_\alpha^k = \emptyset$ —these still follow more or less as in [2]. We now state and prove a number of lemmas relevant to our modification.

Lemma 4.1 *For all $k \in \mathbb{Z}$, the closures \overline{Q}_α^k cover X .*

Proof Let $x \in X$ be arbitrary. For each $m \geq k$ we find some $x_\beta^m =: x_m$ so that $d(x_m, x) < 4\delta^m$. This especially implies that $x_m \rightarrow x$. We know that $x_m \in Q_{\alpha(m)}^k$ for some $\alpha(m)$. We have $d(x_{\alpha(m)}^k, x) \leq d(x_{\alpha(m)}^k, x_m) + d(x_m, x) < 9\delta^k$. As X is geometrically doubling, this implies that $\alpha(m)$ can take only finitely many values. In particular, one finds an infinite subsequence with $x_m \in Q_\alpha^k$ for some fixed α , and so $x \in \overline{Q}_\alpha^k$. □

We then note that as the collection $(\overline{Q}_\alpha^k)_\alpha$ is locally finite, any union of them is closed. Let us then define the open dyadic cubes

$$\tilde{Q}_\alpha^k = X \setminus \bigcup_{\beta \neq \alpha} \overline{Q}_\beta^k,$$

and note that, by what we have already seen, it holds that $Q_\alpha^k \subset \tilde{Q}_\alpha^k \subset \overline{Q}_\alpha^k$.

Lemma 4.2 *If $\ell \geq k$, we have*

$$\overline{Q}_\alpha^k = \bigcup_{\sigma: (\ell, \sigma) \leq (k, \alpha)} \overline{Q}_\sigma^\ell.$$

Proof As this is obvious for $\ell = k$, we take $\ell > k$. Write

$$Q_\alpha^k = \bigcup_{\sigma: (\ell, \sigma) \leq (k, \alpha)} Q_\sigma^\ell \cup \bigcup_{\substack{(m, \sigma) \leq (k, \alpha) \\ m < \ell}} B(x_\sigma^m, \delta^m/100).$$

As the union on the right-hand side is finite, we have

$$\overline{Q}_\alpha^k = \bigcup_{\sigma: (\ell, \sigma) \leq (k, \alpha)} \overline{Q}_\sigma^\ell \cup \bigcup_{\substack{(m, \sigma) \leq (k, \alpha) \\ m < \ell}} \overline{B}(x_\sigma^m, \delta^m/100).$$

Thus, it suffices to prove that

$$\bigcup_{\substack{(m, \sigma) \leq (k, \alpha) \\ m < \ell}} \overline{B}(x_\sigma^m, \delta^m/100) \subset \bigcup_{\sigma: (\ell, \sigma) \leq (k, \alpha)} \overline{Q}_\sigma^\ell.$$

Fix some $(m, \sigma) \leq (k, \alpha)$, where $m < \ell$, and consider a point $x \in \overline{B}(x_\sigma^m, \delta^m/100)$. Since the sets $(\overline{Q}_\beta^\ell)_\beta$ cover the whole X by Lemma 4.1, there is some β for which $x \in \overline{Q}_\beta^\ell$.

It remains to show that $(\ell, \beta) \leq (k, \alpha)$. Fix the γ for which $(\ell, \beta) \leq (m + 1, \gamma)$. We have

$$\begin{aligned} d(x_\sigma^m, x_\gamma^{m+1}) &\leq d(x_\sigma^m, x) + d(x, x_\beta^\ell) + d(x_\beta^\ell, x_\gamma^{m+1}) \\ &\leq \delta^m/100 + 5\delta^\ell + 5\delta^{m+1} < \delta^m/50 < \delta^m/16 \end{aligned}$$

proving that $(m + 1, \gamma) \leq (m, \sigma)$. Since $(\ell, \beta) \leq (m + 1, \gamma)$ and $(m, \sigma) \leq (k, \alpha)$, it holds that $(\ell, \beta) \leq (k, \alpha)$. □

Lemma 4.3 *If $\ell \geq k$, then either $\tilde{Q}_\beta^\ell \subset \tilde{Q}_\alpha^k$, or else $\tilde{Q}_\beta^\ell \cap \overline{Q}_\alpha^k = \emptyset$.*

Proof Suppose that $\tilde{Q}_\beta^\ell \cap \overline{Q}_\alpha^k \neq \emptyset$. The previous lemma gives that $\tilde{Q}_\beta^\ell \cap \overline{Q}_\sigma^\ell \neq \emptyset$ for some $(\ell, \sigma) \leq (k, \alpha)$. By the definition of open cubes, we have to have $\beta = \sigma$, and thus $(\ell, \beta) \leq (k, \alpha)$. Since there is a unique α with this property, it has to be that $\tilde{Q}_\beta^\ell \cap \overline{Q}_\gamma^k = \emptyset$ for all $\gamma \neq \alpha$. This implies that

$$\tilde{Q}_\beta^\ell \subset X \setminus \bigcup_{\gamma \neq \alpha} \overline{Q}_\gamma^k = \tilde{Q}_\alpha^k. \quad \square$$

We are now ready to construct the exact partition of X using “half-open” cubes.

Theorem 4.4 *There exist sets \hat{Q}_α^k , obtained from closed and open sets by finitely many operations, such that $\tilde{Q}_\alpha^k \subset \hat{Q}_\alpha^k \subset \bar{Q}_\alpha^k$,*

$$X = \bigcup_{\alpha} \hat{Q}_\alpha^k$$

for every $k \in \mathbb{Z}$ and if $\ell \geq k$ then either $\hat{Q}_\alpha^k \cap \hat{Q}_\beta^\ell = \emptyset$ or $\hat{Q}_\beta^\ell \subset \hat{Q}_\alpha^k$. Moreover, for every $\ell \geq k$ we have

$$\hat{Q}_\alpha^k = \bigcup_{\beta: (\ell, \beta) \leq (k, \alpha)} \hat{Q}_\beta^\ell.$$

Proof We may assume that $\alpha \in \mathbb{N}$ (just enumerate them for each k). For $k = 0$, define

$$\hat{Q}_0^0 = \bar{Q}_0^0, \quad \hat{Q}_\alpha^0 = \bar{Q}_\alpha^0 \setminus \bigcup_{\beta=0}^{\alpha-1} \hat{Q}_\beta^0, \quad \alpha \geq 1.$$

For $k < 0$ define

$$\hat{Q}_\alpha^k = \bigcup_{\beta: (0, \beta) \leq (k, \alpha)} \hat{Q}_\beta^0.$$

Finally, for $k > 0$ we proceed by induction as follows. Suppose that the cubes \hat{Q}_α^ℓ , $\ell \leq k-1$, are already defined. For every α , consider the finitely many pairs $(k, \beta) \leq (k-1, \alpha)$, temporarily relabel them $\beta = 0, 1, \dots$ (up to some finite number), and set

$$\hat{Q}_0^k = \hat{Q}_\alpha^{k-1} \cap \bar{Q}_0^k, \quad \hat{Q}_\beta^k = \hat{Q}_\alpha^{k-1} \cap \bar{Q}_\beta^k \setminus \bigcup_{\gamma=0}^{\beta-1} \hat{Q}_\gamma^k, \quad \beta \geq 1.$$

All the properties follow. \square

Note that it trivially holds that $B(x_\alpha^k, \delta^k/100) \subset \tilde{Q}_\alpha^k$. Our final lemma concerning solely these cubes will be of use later in the randomization procedure studied in detail in Sect. 10.

Lemma 4.5 *Let $m \in \mathbb{N}$ and $\epsilon > 0$ be such that $500\epsilon \leq \delta^m$. Suppose that $x \in \bar{Q}_\alpha^k$ is such that $d(x, X \setminus \tilde{Q}_\alpha^k) < \epsilon\delta^k$. Then for any chain*

$$(k+m, \sigma) = (k+m, \sigma_{k+m}) \leq \dots \leq (k+1, \sigma_{k+1}) \leq (k, \sigma_k)$$

such that $x \in \bar{Q}_{\sigma_{k+m}}^{k+m}$, it holds that $d(x_{\sigma_j}^j, x_{\sigma_i}^i) \geq \delta^j/500$ for all $k \leq j < i \leq k+m$.

Proof Let us denote, for brevity, $x_{\sigma_j}^j = x_j$, $k \leq j \leq k+m$. Assume that $d(x_j, x_i) < \delta^j/500$ for some $k \leq j < i \leq k+m$. Let us first consider the case $\sigma_k = \alpha$. We have $B(x_j, \delta^j/100) \subset \tilde{Q}_{\sigma_j}^j \subset \tilde{Q}_\alpha^k$, and thus

$$\begin{aligned}
 \delta^j/100 &\leq d(x^j, X \setminus \tilde{Q}_\alpha^k) \\
 &\leq d(x, X \setminus \tilde{Q}_\alpha^k) + d(x, x_i) + d(x_i, x_j) \\
 &\leq \epsilon\delta^k + 5\delta^i + \delta^j/500 \\
 &\leq \delta^j/500 + \delta^j/200 + \delta^j/500 \\
 &= 9\delta^j/1000 \\
 &< \delta^j/100,
 \end{aligned}$$

which is a contradiction. Suppose that $\sigma_k \neq \alpha$. Then $x \in \overline{Q}_{\sigma_k+m}^{k+m} \subset \overline{Q}_{\sigma_k}^k$. On the other hand, we have $x \in \overline{Q}_\alpha^k \subset X \setminus \tilde{Q}_{\sigma_k}^k$. This implies that $d(x, X \setminus \tilde{Q}_{\sigma_k}^k) = 0 < \epsilon\delta^k$, so we are in the identical situation with α replaced by σ_k , and the same conclusion applies. □

We make the following important remark. In all that follows, all the cubes will be “half-open”, but the hat notation is no longer applied.

5 Carleson’s Embedding Theorem and Martingale Difference Decompositions

We now state the Carleson embedding theorem and a related lemma in our setting. We omit the short proof of the following lemma; it is a straightforward adaptation of a known argument.

Lemma 5.1 *Let $a_\alpha^k \geq 0$ be non-negative numbers such that*

$$\sum_{(\ell,\beta) \leq (k,\alpha)} a_\beta^\ell \mu(Q_\beta^\ell) \leq \mu(Q_\alpha^k)$$

for every (k, α) . Suppose we also have some other collection of non-negative numbers $b_\alpha^k \geq 0$ and $b^(x) = \sup_{Q_\alpha^k \ni x} b_\alpha^k$, when $x \in X$. Then it holds that*

$$\sum_{(k,\alpha)} b_\alpha^k a_\alpha^k \mu(Q_\alpha^k) \leq \int_X b^* d\mu.$$

Theorem 5.2 (Carleson’s Embedding Theorem) *Suppose we are given non-negative numbers $a_\alpha^k \geq 0$ such that*

$$\sum_{(\ell,\beta) \leq (k,\alpha)} a_\beta^\ell \mu(Q_\beta^\ell) \leq \mu(Q_\alpha^k)$$

for every (k, α) . Then we have for every $f \in L^2(\mu)$ that

$$\sum_{(k,\alpha)} |\langle f \rangle_{Q_\alpha^k}|^2 a_\alpha^k \mu(Q_\alpha^k) \lesssim \|f\|_{L^2(\mu)}^2.$$

Proof Follows from the previous lemma and the fact that the dyadic maximal operator M_d related to this set of cubes Q_α^k is of strong type $(2, 2)$. \square

We continue to define the martingale difference decomposition of a function f , a tool fundamental to our study. Set

$$\begin{aligned} E_k f &= \sum_{\alpha} \langle f \rangle_{Q_\alpha^k} \chi_{Q_\alpha^k}, \\ E_{Q_\alpha^k} f &= \chi_{Q_\alpha^k} E_k f, \\ \Delta_k f &= E_{k+1} f - E_k f, \\ \Delta_{Q_\alpha^k} f &= \chi_{Q_\alpha^k} \Delta_k f. \end{aligned}$$

If b is accretive, one has the b -adapted versions $E_k^b, E_{Q_\alpha^k}^b, \Delta_k^b$, and $\Delta_{Q_\alpha^k}^b$:

$$\begin{aligned} E_k^b f &= \sum_{\alpha} \langle f \rangle_{Q_\alpha^k} \langle b \rangle_{Q_\alpha^k}^{-1} \chi_{Q_\alpha^k} b, \\ E_{Q_\alpha^k}^b f &= \chi_{Q_\alpha^k} E_k^b f, \\ \Delta_k^b f &= E_{k+1}^b f - E_k^b f, \\ \Delta_{Q_\alpha^k}^b f &= \chi_{Q_\alpha^k} \Delta_k^b f. \end{aligned}$$

We have for any m the decompositions

$$f = \sum_{(k,\alpha):k \geq m} \Delta_{Q_\alpha^k} f + \sum_{\alpha} E_{Q_\alpha^m} f = \sum_{(k,\alpha):k \geq m} \Delta_{Q_\alpha^k}^b f + \sum_{\alpha} E_{Q_\alpha^m}^b f,$$

and it also holds that

$$\begin{aligned} \|f\|_{L^2(\mu)}^2 &= \sum_{(k,\alpha):k \geq m} \|\Delta_{Q_\alpha^k} f\|_{L^2(\mu)}^2 + \sum_{\alpha} \|E_{Q_\alpha^m} f\|_{L^2(\mu)}^2 \\ &\sim \sum_{(k,\alpha):k \geq m} \|\Delta_{Q_\alpha^k}^b f\|_{L^2(\mu)}^2 + \sum_{\alpha} \|E_{Q_\alpha^m}^b f\|_{L^2(\mu)}^2. \end{aligned}$$

Here the implied constants depend only on b via its accretivity constant and L^∞ -norm. This last fact follows from some algebraic manipulations, Carleson’s embedding theorem, and duality as in [11, Lemma 4.1].

In what follows we are given two dyadic systems \mathcal{D} and \mathcal{D}' of “half-open” cubes as constructed above, and we associate the accretive function b_1 with \mathcal{D} and the accretive function b_2 with \mathcal{D}' . We denote the \mathcal{D} -cubes by Q_α^k and the \mathcal{D}' -cubes by R_γ^m . Since for many purposes this is too heavy a notation, we agree that $\ell(Q_\alpha^k) = \delta^k$ and $\text{gen}(Q_\alpha^k) = k$ (and similarly for the \mathcal{D}' -cubes), and more often than not denote simply $Q = Q_\alpha^k$ and $R = R_\gamma^m$. Also, set $C_0 = 10$, $C_1 = 1/100$, and $C_3 = 4$, so that $d(Q_\alpha^k) < C_0 \delta^k$, $B(x_\alpha^k, C_1 \delta^k) \subset Q_\alpha^k$, and $\min_\alpha d(x, x_\alpha^k) < C_3 \delta^k$ (and similarly for the other grid \mathcal{D}'). Note also that all these cubes are sets as in the definition of accretivity.

We have now disposed of the preliminaries, and will begin the task of estimating the operator T . As the reader probably already knows, the idea is to write the adapted martingale difference decompositions for two functions f and g with respect to the grids \mathcal{D} and \mathcal{D}' , respectively, decompose $\langle Tf, g \rangle$, and study the various pairings $\langle T \Delta_Q^{b_1} f, \Delta_R^{b_2} g \rangle$ thus introduced. Note that the theorems and lemmas formulated below do not cover all the cases per se (we mostly consider $\ell(Q) \leq \ell(R)$ only), but combined with symmetry they do. Everything will be brought together to prove the Tb theorem in the very end. We follow the outline given by the most general aspects of [11], and the main contributions are in the details.

6 Separated Cubes

Here we deal with well-separated cubes $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$.

Lemma 6.1 *Let $Q \in \mathcal{D}$, $R \in \mathcal{D}'$, $\ell(Q) \leq \ell(R)$, and $d(Q, R) \geq CC_0\ell(Q)$. Let φ_Q and ψ_R be $L^2(\mu)$ functions supported by the cubes Q and R , respectively, and assume $\int \varphi_Q = 0$. We have the estimate*

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

Proof This follows from the second kernel estimate via the facts that φ_Q has zero integral and $d(Q, R) \geq CC_0\ell(Q)$. □

A reader familiar with the original proof may recall that a condition of the type $d(Q, R) \geq \ell(Q)^\gamma \ell(R)^{1-\gamma}$ plays a key role. The correct choice for γ in our situation has the same algebraic expression as in [11],

$$\gamma := \frac{\alpha}{2(\alpha + d)},$$

where we recall that $d := \log_2 C_\lambda$ in our setting. We then have the familiar relation $\gamma d + \gamma \alpha = \alpha/2$. Also, set $D(Q, R) = \ell(Q) + \ell(R) + d(Q, R)$.

Lemma 6.2 *Assume that $Q \in \mathcal{D}$, $R \in \mathcal{D}'$, $\ell(Q) \leq \ell(R)$, $d(Q, R) \geq CC_0\ell(Q)$ and, in addition, $d(Q, R) \geq \ell(Q)^\gamma \ell(R)^{1-\gamma}$. Let φ_Q and ψ_R be $L^2(\mu)$ functions supported by the cubes Q and R , respectively, and assume $\int \varphi_Q = 0$. We have the estimate*

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

Proof Consider first the case $d(Q, R) \geq \ell(R)$. Then $d(Q, R) \geq D(Q, R)/3$ and so $\lambda(z, d(Q, R)) \geq \lambda(z, D(Q, R)/3) \gtrsim \lambda(z, D(Q, R))$ as λ is doubling. The estimate then follows from the previous lemma.

Let us now assume that $d(Q, R) \leq \ell(R)$. Note that $C_\lambda = 2^d$ so that

$$C_\lambda^{-\gamma \log_2 \frac{\ell(R)}{\ell(Q)}} = \left(\frac{\ell(R)}{\ell(Q)} \right)^{-\gamma d}.$$

We have

$$\begin{aligned} \lambda(z, \ell(R)) &= \lambda(z, (\ell(R)/\ell(Q))^\gamma \ell(Q)^\gamma \ell(R)^{1-\gamma}) \\ &\leq C_\lambda^{\gamma \log_2 \frac{\ell(R)}{\ell(Q)} + 1} \lambda(z, \ell(Q)^\gamma \ell(R)^{1-\gamma}) \end{aligned}$$

and thus

$$\begin{aligned} \lambda(z, d(Q, R)) &\geq \lambda(z, \ell(Q)^\gamma \ell(R)^{1-\gamma}) \\ &\gtrsim C_\lambda^{-\gamma \log_2 \frac{\ell(R)}{\ell(Q)}} \lambda(z, \ell(R)) \\ &= \left(\frac{\ell(R)}{\ell(Q)} \right)^{-\gamma d} \lambda(z, \ell(R)). \end{aligned}$$

This implies that

$$\begin{aligned} \frac{\ell(Q)^\alpha}{d(Q, R)^\alpha \lambda(z, d(Q, R))} &\lesssim \frac{\ell(Q)^\alpha}{\ell(Q)^\gamma \ell(R)^{(1-\gamma)\alpha} \left(\frac{\ell(R)}{\ell(Q)} \right)^{-\gamma d} \lambda(z, \ell(R))} \\ &= \frac{\ell(Q)^{\alpha - (\gamma d + \gamma \alpha)} \ell(R)^{\gamma d + \gamma \alpha}}{\ell(R)^\alpha \lambda(z, \ell(R))} \\ &= \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{\ell(R)^\alpha \lambda(z, \ell(R))}. \end{aligned}$$

Furthermore, one now has $\ell(R) \geq D(Q, R)/3$ so that the claim follows from the previous lemma. □

We may now forget for the moment under which assumptions these estimates were achieved, and just study the matrix that we got. Namely, let us define the matrix

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

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

$$T_{QR} = 0$$

otherwise.

Proposition 6.3 *Suppose we are given nonnegative constants x_Q and y_R for each $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$. It holds that*

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

Proof We assume first that $\ell(Q) = \delta^m \ell(R)$ for some $m = 0, 1, 2, \dots$ and then also that $\ell(R) = \delta^k$ for some $k \in \mathbb{Z}$. Define the kernel

$$K_{m,k}(x, y) = \sum_{\ell(Q)=\delta^{k+m}, \ell(R)=\delta^k} T_{QR} \mu(Q)^{-1/2} \mu(R)^{-1/2} \chi_Q(x) \chi_R(y)$$

and set

$$h_1 = \sum_{\ell(Q)=\delta^{k+m}} \mu(Q)^{-1/2} x_Q \chi_Q, \quad h_2 = \sum_{\ell(R)=\delta^k} \mu(R)^{-1/2} y_R \chi_R.$$

Note that

$$\sum_{\ell(Q)=\delta^{k+m}, \ell(R)=\delta^k} T_{QR} x_Q y_R = \int_X \int_X K_{m,k}(x, y) h_1(x) h_2(y) d\mu(x) d\mu(y)$$

and

$$\left(\sum_{\ell(Q)=\delta^{k+m}} x_Q^2 \right)^{1/2} = \|h_1\|_{L^2(\mu)} \quad \text{and} \quad \left(\sum_{\ell(R)=\delta^k} y_R^2 \right)^{1/2} = \|h_2\|_{L^2(\mu)}.$$

Writing out the definitions, one has

$$K_{m,k}(x, y) = \sum_{\ell(Q)=\delta^{k+m}, \ell(R)=\delta^k} \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^\alpha \sup_{z \in Q} \lambda(z, D(Q, R))} \chi_Q(x) \chi_R(y).$$

Consider some pair (x, y) . Then there exists one and only one pair (Q_x, R_y) for which $\ell(Q_x) = \delta^{k+m}$, $\ell(R_y) = \delta^k$, $x \in Q_x$, and $y \in R_y$, and so

$$K_{m,k}(x, y) = \frac{\delta^{\alpha m/2} \delta^{\alpha k}}{D(Q_x, R_y)^\alpha \sup_{z \in Q_x} \lambda(z, D(Q_x, R_y))}.$$

We want to prove that

$$\int_X K_{m,k}(x, y) d\mu(x) \lesssim \delta^{\alpha m/2} \quad \text{and} \quad \int_X K_{m,k}(x, y) d\mu(y) \lesssim \delta^{\alpha m/2}. \quad (6.1)$$

Let us only deal with the first term in detail—it is actually the bit harder of the two. The second integral is estimated basically in the same way as we now deal with the first integral, but one does not have to go through the trouble of fiddling with

the centers (just use directly that $\sup_{z \in Q_x} \lambda(z, D(Q_x, R_y)) \geq \lambda(x, D(Q_x, R_y))$). We have

$$\int_X K_{m,k}(x, y) d\mu(x) = \left(\int_{B(y, \delta^k)} + \int_{X \setminus B(y, \delta^k)} \right) K_{m,k}(x, y) d\mu(x).$$

Note that $D(Q_x, R_y) \geq \ell(R_y) = \delta^k$, and so

$$\int_{B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \leq \delta^{\alpha m/2} \delta^{\alpha k} \int_{B(y, \delta^k)} \frac{d\mu(x)}{\delta^{\alpha k} \sup_{z \in Q_x} \lambda(z, \delta^k)}.$$

We have that $Q_x \subset B(y, 2C_0\delta^k)$ for every $x \in B(y, \delta^k)$, and so $\sup_{z \in Q_x} \lambda(z, \delta^k) \geq \inf_{z \in B(y, 2C_0\delta^k)} \lambda(z, \delta^k)$. This yields

$$\int_{B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \leq \delta^{\alpha m/2} \frac{\mu(B(y, \delta^k))}{\inf_{z \in B(y, 2C_0\delta^k)} \lambda(z, \delta^k)}.$$

We then have that $B(y, \delta^k) \subset B(z, 3C_0\delta^k)$ for every $z \in B(y, 2C_0\delta^k)$, yielding that $\mu(B(y, \delta^k)) \leq \mu(B(z, 3C_0\delta^k)) \leq \lambda(z, 3C_0\delta^k) \lesssim \lambda(z, \delta^k)$. This gives that

$$\int_{B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \lesssim \delta^{\alpha m/2}.$$

We cannot directly employ Lemma 2.2 to deal with the integral over $X \setminus B(y, \delta^k)$. However, we can use its proof together with similar gimmicks as with the previous term. Note that $d(x, y) \lesssim D(Q_x, R_y)$ to get that

$$\begin{aligned} & \int_{X \setminus B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \\ & \lesssim \delta^{\alpha m/2} \delta^{\alpha k} \sum_{j=0}^{\infty} \int_{2^j \delta^k \leq d(x, y) < 2^{j+1} \delta^k} \frac{d(x, y)^{-\alpha}}{\sup_{z \in Q_x} \lambda(z, d(x, y))} d\mu(x). \end{aligned}$$

If $2^j \delta^k \leq d(x, y) < 2^{j+1} \delta^k$, we have $Q_x \subset B(y, C_0 2^{j+2} \delta^k)$, and therefore it holds that $\sup_{z \in Q_x} \lambda(z, d(x, y)) \geq \inf_{z \in B(y, C_0 2^{j+2} \delta^k)} \lambda(z, 2^j \delta^k)$. This establishes that

$$\int_{X \setminus B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \lesssim \delta^{\alpha m/2} \sum_{j=0}^{\infty} 2^{-\alpha j} \frac{\mu(B(y, 2^{j+1} \delta^k))}{\inf_{z \in B(y, C_0 2^{j+2} \delta^k)} \lambda(z, 2^j \delta^k)}.$$

If $z \in B(y, C_0 2^{j+2} \delta^k)$, we have $B(y, 2^{j+1} \delta^k) \subset B(z, C_0 2^{j+3} \delta^k)$, and thus it holds $\mu(B(y, 2^{j+1} \delta^k)) \leq \mu(B(z, C_0 2^{j+3} \delta^k)) \leq \lambda(z, C_0 2^{j+3} \delta^k) \lesssim \lambda(z, 2^j \delta^k)$. This proves that

$$\int_{X \setminus B(y, \delta^k)} K_{m,k}(x, y) d\mu(x) \lesssim \delta^{\alpha m/2}.$$

We have established (6.1). Schur’s lemma then gives that

$$\int_X \int_X K_{m,k}(x, y)h_1(x)h_2(y) d\mu(x) d\mu(y) \lesssim \delta^{\alpha m/2} \|h_1\|_{L^2(\mu)} \|h_2\|_{L^2(\mu)}.$$

As noted above this is the same as

$$\sum_{\ell(Q)=\delta^{k+m}, \ell(R)=\delta^k} T_{QR}x_Qy_R \lesssim \delta^{\alpha m/2} \left(\sum_{\ell(Q)=\delta^{k+m}} x_Q^2 \right)^{1/2} \left(\sum_{\ell(R)=\delta^k} y_R^2 \right)^{1/2}.$$

Sum this over $k \in \mathbb{Z}$, use the Cauchy–Schwarz inequality, and then sum over $m = 0, 1, 2, \dots$, to get that

$$\sum_{Q,R} T_{QR}x_Qy_R = \sum_{\ell(Q) \leq \ell(R)} T_{QR}x_Qy_R \lesssim \left(\sum_Q x_Q^2 \right)^{1/2} \left(\sum_R y_R^2 \right)^{1/2}.$$

□

7 Paraproducts and Cubes Well Inside Another Cube

We begin by proving the following lemma which is needed later in proving that a certain paraproduct is bounded.

Lemma 7.1 *Suppose that $Q \in \mathcal{D}$ is fixed and that b is some pseudoaccretive function. It holds that*

$$\sum_{\substack{R \in \mathcal{D}' : R \subset Q \\ \ell(R) \leq \delta^r \ell(Q) \\ d(R, X \setminus Q) \geq CC_0\kappa \ell(R)}} \|\Delta_R^b(b\varphi)\|_{L^2(\mu)}^2 \lesssim \mu(Q) \|\varphi\|_{BMO_\kappa^2(\mu)}^2$$

if $\varphi \in BMO_\kappa^2(\mu)$ and r is so large that

$$\delta^r \leq \frac{C_1}{CC_0\kappa + C_0 + C_3}.$$

Here the implied constants are exceptionally allowed to depend on the function b in the obvious way.

Proof We start by constructing a Whitney type decomposition using cubes and then we shall associate with each such cube in the covering a certain ball—these balls, as we shall see, will also have finite overlap (even when multiplied with the constant κ) because of the geometry of the construction and also because X is geometrically doubling.

We first prove that there exist cubes $R \in \mathcal{D}'$ for which $R \subset Q$, $\ell(R) = \delta^r \ell(Q)$ and $d(R, X \setminus Q) \geq CC_0\kappa \ell(R)$. Denote the center of Q by z and recall that we

have $B(z, C_1\ell(Q)) \subset Q$. Choose some point w which is the center of a \mathcal{D}' -cube R of generation $\text{gen}(Q) + r$ and satisfies $d(w, z) < C_3\ell(R)$. Suppose that $x \in R \subset B(w, C_0\ell(R))$ and note that $d(x, z) \leq d(x, w) + d(w, z) < (C_0 + C_3)\ell(R)$, that is $R \subset B(z, (C_0 + C_3)\ell(R))$. Suppose that $y \in X \setminus Q$ and $x \in R$. In this case we have $d(y, z) \geq C_1\ell(Q)$ and $d(x, z) < (C_0 + C_3)\ell(R)$, which yields

$$d(x, y) \geq d(y, z) - d(x, z) \geq (C_1\delta^{-r} - C_0 - C_3)\ell(R) \geq CC_0\kappa\ell(R).$$

Thus, $d(R, X \setminus Q) \geq CC_0\kappa\ell(R)$. We choose all such cubes R . Then we choose all those \mathcal{D}' -cubes R of the next generation which are not subcubes of the previously chosen cubes and which still satisfy the condition that $d(R, X \setminus Q) \geq CC_0\kappa\ell(R)$. We continue in this way and obtain a disjoint collection of \mathcal{D}' -cubes R , which have the property that any cube R' in the sum

$$\sum_{\substack{R' \in \mathcal{D}': R' \subset Q \\ \ell(R') \leq \delta^r \ell(Q) \\ d(R', X \setminus Q) \geq CC_0\kappa\ell(R')}} \|\Delta_{R'}^b(b\varphi)\|_{L^2(\mu)}$$

is contained in one of them. To exploit the BMO condition, though, we want a covering consisting of suitable balls. To this end, we associate with each chosen cube R the ball $B_R = B(w_R, C_0\ell(R)) \supset R$ which is centered at the center w_R of the cube R and which has radius $C_0\ell(R)$.

We now wish to demonstrate that every $x \in Q$ belongs to $\lesssim 1$ balls κB_R . We first prove that x can belong to only $\lesssim 1$ balls κB_R associated with a fixed generation $k \geq \text{gen}(Q) + r$ of the chosen cubes. Indeed, suppose that $x \in \kappa B_{R_i^k}$ for some collection i . This implies that $w_{R_i^k} \in B(x, C_0\kappa\delta^k)$ for all i . This means that the ball $B(x, C_0\kappa\delta^k)$ contains the centers $w_{R_i^k}$ of the disjoint balls $B(w_{R_i^k}, C_1\delta^k) = B(w_{R_i^k}, C_1/(C_0\kappa) \cdot C_0\kappa\delta^k)$. As X is geometrically doubling, we have that $\#i \leq N(C_1/(C_0\kappa))^{-n} \lesssim 1$.

We then prove that if $\kappa B_{R_i^k} \cap \kappa B_{R_j^l} \neq \emptyset$ (where R_i^k and R_j^l are chosen cubes, $k, l > r$), then $|k - l| \lesssim 1$. This only utilizes the geometry of the construction. We prove a certain auxiliary estimate from which this follows. Suppose that R^k is a chosen cube of generation $k > \text{gen}(Q) + r$ and that $x \in \kappa B_{R^k}$. As R^k is a chosen cube, its dyadic parent R^{k*} has to satisfy $d(R^{k*}, X \setminus Q) < CC_0\kappa\delta^{k-1}$. We take some point $y \in R^{k*}$ for which $d(y, X \setminus Q) \leq CC_0\kappa\delta^{k-1}$. Let w_{R^k} be the center of R^k and notice that we now have

$$\begin{aligned} d(x, X \setminus Q) &\leq d(x, y) + d(y, X \setminus Q) \\ &\leq d(x, w_{R^k}) + d(y, w_{R^k}) + d(y, X \setminus Q) \\ &\leq C_0\kappa\delta^k + C_0\delta^{k-1} + CC_0\kappa\delta^{k-1} \\ &\leq 3CC_0\kappa\delta^{k-1}. \end{aligned}$$

To the other direction, it holds that (we have $C \geq 2$)

$$d(x, X \setminus Q) \geq d(w_{R^k}, X \setminus Q) - d(x, w_{R^k}) \geq CC_0\kappa\delta^k - C_0\kappa\delta^k \geq C_0\kappa\delta^k.$$

We have established that if R is a chosen cube for which $\text{gen}(R) = k > \text{gen}(Q) + r$, then

$$C_0\kappa\delta^k \leq d(x, X \setminus Q) \leq 3CC_0\kappa\delta^{k-1}$$

holds for all $x \in \kappa B_R$. Taking logarithms one sees that this fixes k to a certain finite range.

We have now done more than enough to show that for every $x \in Q$ one has $\#\{R : x \in \kappa B_R\} \lesssim 1$, which we use to conclude that

$$\sum_R \mu(\kappa B_R) = \int_Q \sum_R \chi_{\kappa B_R} d\mu \lesssim \mu(Q),$$

where we sum over the chosen cubes R . We have

$$\begin{aligned} \sum_{\substack{R' \in \mathcal{D}' : R' \subset Q \\ \ell(R') \leq \delta^r \ell(Q) \\ d(R', X \setminus Q) \geq CC_0\kappa \ell(R')}} \|\Delta_{R'}^b(b\varphi)\|_{L^2(\mu)}^2 &= \sum_R \sum_{R' \subset R} \|\Delta_{R'}^b(\chi_{B_R} b(\varphi - \varphi_{B_R}))\|_{L^2(\mu)}^2 \\ &\lesssim \sum_R \int_{B_R} |\varphi - \varphi_{B_R}|^2 d\mu \\ &\lesssim \sum_R \mu(\kappa B_R) \|\varphi\|_{\text{BMO}_\kappa^2(\mu)}^2 \\ &\lesssim \mu(Q) \|\varphi\|_{\text{BMO}_\kappa^2(\mu)}^2. \end{aligned}$$

□

Let us define the paraproduct

$$\Pi f = \sum_{R' \in \mathcal{D}'} \sum_{\substack{Q' \in \mathcal{D} : Q' \subset R' \\ \ell(Q') = \delta^r \ell(R') \\ d(Q', X \setminus R') \geq CC_0\kappa \ell(Q')}} (E_{R'} b_2)^{-1} \cdot E_{R'} f \cdot (\Delta_{Q'}^{b_1})^*(T^* b_2).$$

We shall always assume that r is at least as large as is required by the previous lemma. However, we make several further assumptions about it later. Of course, basically it could be fixed at the very beginning, and so it is not a problem if we let the implied constants depend also on r .

Theorem 7.2 *The paraproduct Π is bounded on $L^2(\mu)$.*

Proof This follows from the previous lemma via the $L^2(\mu)$ norm estimate related to the adapted martingale difference decomposition, Carleson’s embedding theorem, and the fact that $(\Delta_{Q'}^{b_1})^*(T^* b_2) = b_1^{-1} \Delta_{Q'}^{b_1}(b_1 T^* b_2)$. □

We introduce the concept of good cubes in more detail now. Recall the definition of γ from the previous section. Consider a cube $Q \in \mathcal{D}$. We say that Q is good if

for any cube $R \in \mathcal{D}'$ for which we have $\ell(Q) \leq \delta^r \ell(R)$, we have either $d(Q, R) \geq \ell(Q)^\gamma \ell(R)^{1-\gamma}$ or $d(Q, X \setminus R) \geq \ell(Q)^\gamma \ell(R)^{1-\gamma}$. We denote this set by $\mathcal{D}_{\text{good}}$, and the rest are denoted by \mathcal{D}_{bad} .

Let us now fix $Q \in \mathcal{D}_{\text{good}}$ and $R \in \mathcal{D}'$ so that $Q \subset R$ and $\ell(Q) < \delta^r \ell(R)$. Let us also fix the child $R_1 \subset R$ for which $Q \subset R_1$ and let us denote the other children of R by R_j . We have $\#i \lesssim 1$. We assume that we are given two functions of the form

$$\varphi_Q = \sum_{\substack{Q_j \in \mathcal{D}: Q_j \subset Q \\ \ell(Q_j) = \delta \ell(Q)}} A_{Q_j} \chi_{Q_j} b_1$$

and

$$\psi_R = \sum_{\substack{R_i \in \mathcal{D}': R_i \subset R \\ \ell(R_i) = \delta \ell(R)}} B_{R_i} \chi_{R_i} b_2.$$

We demand that $\int \varphi_Q d\mu = 0$. We aim to prove that

$$|\langle (T - \Pi^*)\varphi_Q, \psi_R \rangle| \lesssim \left(\frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \left(\frac{\mu(Q)}{\mu(R_1)} \right)^{1/2} \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)}. \tag{7.1}$$

In the previous section our matrix looked a little bit different from that in [11]. However, (7.1) is exactly of the same form as in [11, Lemma 7.3]. We note that

$$\langle \varphi_Q, \Pi \psi_R \rangle = \sum_{R' \in \mathcal{D}'} \sum_{\substack{Q' \in \mathcal{D}: Q' \subset R' \\ \ell(Q') = \delta^r \ell(R') \\ d(Q', X \setminus R') \geq CC_0 \kappa \ell(Q')}} \langle \varphi_Q, (E_{R'} b_2)^{-1} \cdot E_{R'} \psi_R \cdot (\Delta_{Q'}^{b_1})^*(T^* b_2) \rangle$$

and that

$$\langle \varphi_Q, (E_{R'} b_2)^{-1} \cdot E_{R'} \psi_R \cdot (\Delta_{Q'}^{b_1})^*(T^* b_2) \rangle = \frac{\langle \psi_R \rangle_{R'}}{\langle b_2 \rangle_{R'}} \langle \Delta_Q^{b_1} \varphi_Q, T^* b_2 \rangle.$$

Since φ_Q is of this particular form and $\int \varphi_Q d\mu = 0$ one sees that $\Delta_Q^{b_1} \varphi_Q = \varphi_Q$ if Q' is the same cube as Q and $\Delta_{Q'}^{b_1} \varphi_Q = 0$ otherwise. This implies in particular that in the non-trivial situation one must also have $R' \subset R_1$ and then $\frac{\langle \psi_R \rangle_{R'}}{\langle b_2 \rangle_{R'}} = B_{R_1}$. If a suitable $R' \supset Q$ exists (and there can only be one) we thus have $\langle \varphi_Q, \Pi \psi_R \rangle = B_{R_1} \langle T \varphi_Q, b_2 \rangle$. Let us now demonstrate that, indeed, such an R' exists provided that we have chosen r to be large enough. As $Q \in \mathcal{D}_{\text{good}}$, $Q \subset R$, and $\ell(Q) < \delta^r \ell(R)$, we must have some $R' \in \mathcal{D}'$ so that $Q \subset R' \subset R_1 \subset R$, $\ell(Q) = \delta^r \ell(R')$ and $d(Q, X \setminus R') \geq \ell(Q)^\gamma \ell(R')^{1-\gamma}$. In particular, demanding that r is at least so large that $\delta^{-(1-\gamma)r} \geq CC_0 \kappa$, we have

$$d(Q, X \setminus R') \geq \ell(Q)^\gamma \ell(R')^{1-\gamma} = \ell(Q) \delta^{-r(1-\gamma)} \geq CC_0 \kappa \ell(Q).$$

We have shown that $\langle \varphi_Q, \Pi \psi_R \rangle = B_{R_1} \langle T \varphi_Q, b_2 \rangle$, and so

$$\begin{aligned} \langle (T - \Pi^*)\varphi_Q, \psi_R \rangle &= \langle T\varphi_Q, \psi_R - B_{R_1}b_2 \rangle \\ &= B_{R_1}\langle T\varphi_Q, (\chi_{R_1} - 1)b_2 \rangle + \sum_{i \neq 1} B_{R_i}\langle T\varphi_Q, \chi_{R_i}b_2 \rangle. \end{aligned}$$

We handle the first term first. Let us calculate (choosing some arbitrary point $z \in Q$)

$$\begin{aligned} |\langle T\varphi_Q, (\chi_{R_1} - 1)b_2 \rangle| &= \left| \int_{X \setminus R_1} \int_Q [K(x, y) - K(x, z)]\varphi_Q(y)b_2(x) d\mu(y) d\mu(x) \right| \\ &\lesssim \ell(Q)^\alpha \|\varphi_Q\|_{L^1(\mu)} \int_{X \setminus R_1} \frac{d(x, z)^{-\alpha}}{\lambda(z, d(x, z))} d\mu(x) \\ &\leq \ell(Q)^\alpha \|\varphi_Q\|_{L^1(\mu)} \int_{X \setminus B(z, d(Q, X \setminus R_1))} \frac{d(x, z)^{-\alpha}}{\lambda(z, d(x, z))} d\mu(x) \\ &\lesssim \ell(Q)^\alpha \|\varphi_Q\|_{L^1(\mu)} d(Q, X \setminus R_1)^{-\alpha}. \end{aligned}$$

We used (for the kernel estimates) that $d(y, z) \leq C_0\ell(Q)$ and $d(x, z) \geq d(Q, X \setminus R_1) \geq \ell(Q)^\gamma \ell(R_1)^{1-\gamma} \geq CC_0\ell(Q)$ as there is a gap of at least r in the generations of Q and R_1 . We then use that $d(Q, X \setminus R_1) \geq \ell(Q)^\gamma \ell(R_1)^{1-\gamma} \geq \ell(Q)^{1/2} \ell(R_1)^{1/2} \gtrsim \ell(Q)^{1/2} \ell(R_1)^{1/2}$ to get that

$$|\langle T\varphi_Q, (\chi_{R_1} - 1)b_2 \rangle| \lesssim \ell(Q)^\alpha \mu(Q)^{1/2} \|\varphi_Q\|_{L^2(\mu)} \ell(Q)^{-\alpha/2} \ell(R)^{-\alpha/2}.$$

Note then that $|B_{R_1}| \lesssim \|\psi_R\|_{L^2(\mu)} \mu(R_1)^{-1/2}$ to infer

$$|B_{R_1}| |\langle T\varphi_Q, (\chi_{R_1} - 1)b_2 \rangle| \lesssim \left(\frac{\ell(Q)}{\ell(R)}\right)^{\alpha/2} \left(\frac{\mu(Q)}{\mu(R_1)}\right)^{1/2} \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)}.$$

We then deal with the other $\#i \lesssim 1$ terms. This time we have, using estimates from Sect. 6 (see the proof of Lemma 6.2), for some fixed $z \in Q$ that

$$\begin{aligned} &|B_{R_i}| |\langle T\varphi_Q, \chi_{R_i}b_2 \rangle| \\ &= |B_{R_i}| \left| \int_{R_i} \int_Q [K(x, y) - K(x, z)]\varphi_Q(y)b_2(x) d\mu(y) d\mu(x) \right| \\ &\lesssim \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)} \mu(R_i)^{1/2} \mu(Q)^{1/2} \frac{\ell(Q)^\alpha}{d(Q, R_i)^\alpha \lambda(z, d(Q, R_i))} \\ &\lesssim \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)} \mu(R_i)^{1/2} \mu(Q)^{1/2} \frac{\ell(Q)^{\alpha/2} \ell(R)^{-\alpha/2}}{\lambda(z, \ell(R_i))} \\ &= \left(\frac{\ell(Q)}{\ell(R)}\right)^{\alpha/2} \mu(Q)^{1/2} \frac{\mu(R_i)^{1/2}}{\lambda(z, \ell(R_i))} \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)}. \end{aligned}$$

We make the following deduction which uses the doubling property of λ :

$$\frac{\mu(R_i)^{1/2}}{\lambda(z, \ell(R_i))} \leq \frac{\mu(R)^{1/2}}{\lambda(z, \ell(R_i))} \leq \frac{\lambda(z, C_0\ell(R))^{1/2}}{\lambda(z, \ell(R_i))} \lesssim \frac{1}{\lambda(z, C_0\ell(R))^{1/2}} \leq \frac{1}{\mu(R_1)^{1/2}}.$$

Insert this into the estimate from above to get that

$$|B_{R_i}| \langle |T\varphi_Q, \chi_{R_i} b_2 \rangle | \lesssim \left(\frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \left(\frac{\mu(Q)}{\mu(R_1)} \right)^{1/2} \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)},$$

and (7.1) follows. Set

$$T_{QR} = \left(\frac{\ell(Q)}{\ell(R)} \right)^{\alpha/2} \left(\frac{\mu(Q)}{\mu(R_1)} \right)^{1/2},$$

if $Q \in \mathcal{D}_{\text{good}}$, $Q \subset R$ and $\ell(Q) < \delta^r \ell(R)$, $T_{QR} = 0$ otherwise.

Proposition 7.3 *The matrix T_{QR} generates a bounded operator on ℓ^2 .*

Proof This follows precisely as in [11, Lemma 7.4], as one just has to deal with the measure μ not using any special assumptions about it. \square

8 Random Almost-Covering by Balls

We construct a probabilistic covering of a large portion of the space with balls, starting from some fixed size and going down in size but only for some controlled amount. This will be used as a substitute for a certain auxiliary third dyadic grid used in [11, Sect. 10.2] in connection with the weak boundedness property. Here we need to explicitly work with our original quasimetric ρ , the reason being that the weak boundedness property does not transfer to the d -balls in any obvious way.

Let $0 < \vartheta < A_0^{-4}/32$. For each $k \in \mathbb{Z}$ fix some maximal collection $z_\alpha^k \in X$ for which $\rho(z_\alpha^k, z_\beta^k) \geq \vartheta^k$ for all $\alpha \neq \beta$. This time we use the following transitive relation \leq . For each (k, α) there exists at least one β for which $\rho(z_\alpha^k, z_\beta^{k-1}) < \vartheta^{k-1}$. Also, there exists at most one β for which $\rho(z_\alpha^k, z_\beta^{k-1}) < (2A_0)^{-1} \vartheta^{k-1}$. The ordering \leq is constructed using the rules we now describe. Consider any pair (k, α) . Check first whether there exists β so that $\rho(z_\alpha^k, z_\beta^{k-1}) < (2A_0)^{-1} \vartheta^{k-1}$. If so, set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Otherwise, choose any β for which $\rho(z_\alpha^k, z_\beta^{k-1}) < \vartheta^{k-1}$, and set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Extend by transitivity.

We introduce another relation \searrow now. Given any (k, α) , pick one β for which $(k+1, \beta) \leq (k, \alpha)$ and set $(k, \alpha) \searrow (k+1, \beta)$ and $(k, \alpha) \not\searrow (k+1, \gamma)$ for $\gamma \neq \beta$. This is the relation which we shall randomize in a natural way, and in this way we shall obtain in a random way a new collection of points for each level k . Indeed, we shall essentially replace z_α^k with z_β^{k+1} if $(k, \alpha) \searrow (k+1, \beta)$, and then remove some points if they end up being too close to each other. Let us now do this in detail. We define a probability \mathbb{P} , on the family of all relations \searrow of the kind described, by setting

$$\mathbb{P}((k, \alpha) \searrow (k+1, \beta)) = \frac{1}{\#\{\gamma : (k+1, \gamma) \leq (k, \alpha)\}}$$

for all $(k + 1, \beta) \leq (k, \alpha)$, and requiring that such events for two different (k, α) and (ℓ, σ) are independent. If $(k, \alpha) \searrow (k + 1, \beta)$, we set $y_\alpha^k = z_\beta^{k+1}$. However, we want to now obtain better separation between the points y_α^k . To this end, we say that y_α^k and y_β^k are in conflict if $\rho(y_\alpha^k, y_\beta^k) < (2A_0)^{-2}\vartheta^k$. If such is the case, we have $\rho(z_\alpha^k, z_\beta^k) < 3A_0^2\vartheta^k$. As X is geometrically doubling, it follows that at most finitely many pairs can conflict with a given pair (k, α) . We enumerate the points y_α^k as y_1^k, y_2^k, \dots . We choose y_1^k and remove all the boundedly many points conflicting with it. Next, choose the point with the smallest index in the remaining sequence, and remove all the boundedly many points conflicting with it. Continue this by induction. The final collection is now denoted by x_α^k . By construction we have that

$$\rho(x_\alpha^k, x_\beta^k) \geq (2A_0)^{-2}\vartheta^k, \quad \text{if } \alpha \neq \beta.$$

Observe also that for an arbitrary $x \in X$ there exists x_α^k so that $\rho(x, x_\alpha^k) < 3A_0^2\vartheta^k$.

Lemma 8.1 *If $(k + 1, \beta) \leq (k, \alpha)$, then $\mathbb{P}(z_\beta^{k+1} = x_\alpha^k) \geq \pi_0 > 0$.*

Proof The event $z_\beta^{k+1} = x_\alpha^k$ requires that $(k, \alpha) \searrow (k + 1, \beta)$ and then that the point $y_\alpha^k = z_\beta^{k+1}$ was not removed in the above described removal process. As X is geometrically doubling, we have that $\#\{\gamma : (k + 1, \gamma) \leq (k, \sigma)\} \lesssim 1$ and so always $\mathbb{P}((k, \sigma) \searrow (k + 1, \gamma)) \geq \pi'_0 > 0$ if $(k + 1, \gamma) \leq (k, \sigma)$. So all we need to prove is that for any of the $\lesssim 1$ pairs (k, σ) with the potential of conflicting with (k, α) , we have some γ so that $(k + 1, \gamma) \leq (k, \sigma)$ and $\rho(z_\beta^{k+1}, z_\gamma^{k+1}) \geq (2A_0)^{-2}\vartheta^k$, for then there is a positive probability that $(k, \sigma) \searrow (k + 1, \gamma)$ and no conflict with (k, α) will arise.

Consider any such (k, σ) . There is z_γ^{k+1} so that $\rho(z_\sigma^k, z_\gamma^{k+1}) < \vartheta^{k+1}$. In particular, $\rho(z_\sigma^k, z_\gamma^{k+1}) < (2A_0)^{-1}\vartheta^k$ and so certainly $(k + 1, \gamma) \leq (k, \sigma)$. Note also that $(k + 1, \beta) \leq (k, \alpha) \neq (k, \sigma)$, and so $\rho(z_\beta^{k+1}, z_\sigma^k) \geq (2A_0)^{-1}\vartheta^k$. This yields that

$$\rho(z_\beta^{k+1}, z_\gamma^{k+1}) \geq A_0^{-1} \rho(z_\beta^{k+1}, z_\sigma^k) - \rho(z_\gamma^{k+1}, z_\sigma^k) \geq [2^{-1}A_0^{-2} - \vartheta]\vartheta^k > (2A_0)^{-2}\vartheta^k.$$

This proves the assertion. □

Take a new random variable τ , uniformly distributed on $[1, 2]$ and independent of all the previous random quantities. Then define the random ρ -balls

$$B_\alpha^k = B(x_\alpha^k, \tau A_0^{-4}\vartheta^k/32).$$

We now note that $\mathbb{P}(x \in \bigcup_\alpha B_\alpha^k) \geq \pi_0 > 0$ for all $x \in X$ and $k \in \mathbb{Z}$. Indeed, for a given $x \in X$ there exists z_β^{k+1} so that $\rho(x, z_\beta^{k+1}) < \vartheta^{k+1} < A_0^{-4}\vartheta^k/32$, and $\mathbb{P}(z_\beta^{k+1} = x_\alpha^k) \geq \pi_0$ as proved above. We also have

$$\rho(B_\alpha^k, B_\gamma^k) \geq A_0^{-2}(2A_0)^{-2}\vartheta^k - 2^{-3}A_0^{-4}\vartheta^k = 2^{-3}A_0^{-4}\vartheta^k.$$

So we have separation for balls of the same generation.

We now make the final construction of the balls. We are given some small $\nu \in (0, 1)$ and a fixed starting size k . We construct the level k balls B_α^k as above. We take some small parameter $\omega \in (0, 1)$ which we shall fix momentarily. We introduced the random variable τ to make the proof of the following fact easy: it is unlikely for a point to belong to the set $\bigcup_\alpha (1 + \omega)B_\alpha^k \setminus B_\alpha^k$.

Let us spell this out. One notes that $x \in (1 + \omega)B_\alpha^k$ can only happen for certain boundedly many different α , where the bound depends on the geometric doubling property. We then estimate the probability that $x \in (1 + \omega)B_\alpha^k \setminus B_\alpha^k$ for one of these balls B_α^k . The mentioned inclusion happens if and only if

$$\frac{\tau \vartheta^k}{32A_0^4} \leq \rho(x, x_\alpha^k) < (1 + \omega) \frac{\tau \vartheta^k}{32A_0^4}.$$

This means that τ must belong to a certain interval of length

$$32A_0^4 \rho(x, x_\alpha^k) \vartheta^{-k} \frac{\omega}{1 + \omega} \leq 4\omega,$$

since necessarily $\rho(x, x_\alpha^k) \leq (8A_0^4)^{-1} \vartheta^k$. Given the uniform distribution of τ on $[1, 2]$, this implies that

$$\mathbb{P}\left(x \in \bigcup_\alpha (1 + \omega)B_\alpha^k \setminus B_\alpha^k\right) \leq \eta_0 = \eta_0(\omega) \lesssim \omega.$$

We now choose ω so small, and then $M \in \mathbb{N}$ so large, depending on ν , that

$$\frac{\pi_0}{\pi_0 + \eta_0} \geq 1 - \nu/2, \quad 1 - (1 - \pi_0 - \eta_0)^M > \frac{1 - \nu}{1 - \nu/2}.$$

Let further $\epsilon > 0$ be so small, and then $s \in \mathbb{N}$ so large, that

$$(1 + \epsilon)^2 < 1 + \omega/3, \quad 2A(\epsilon) \vartheta^s < \omega/3.$$

Then, ultimately, ω , M , and ϵ may be chosen as functions of ν , and s as a function of ν and ϑ .

We now continue to make the above random ball covering with k replaced by $k + s$ and X replaced by $X \setminus \bigcup_\alpha (1 + \omega)B_\alpha^k$. We repeat this procedure M times. We denote the collection of balls we obtain by \mathcal{B} .

We are in the following situation. At stage one a point belongs to some ball with probability $\pi_1 \geq \pi_0$ and to the ω -buffer of some ball with probability $\eta_1 \leq \eta_0$. Thus, a point belongs to none of these sets with probability $1 - \pi_1 - \eta_1$. Note that subsets of X are geometrically doubling with the same constant N , and thus a point belongs to some ball at stage two with probability $\pi_2 \geq \pi_0$ and to the ω -buffer of some ball with probability $\eta_2 \leq \eta_0$. We have this situation at every stage. Therefore, it holds that

$$\begin{aligned} \mathbb{P}\left(x \in \bigcup_{B \in \mathcal{B}} B\right) &= \pi_1 + (1 - \pi_1 - \eta_1)\pi_2 + \dots + \left(\prod_{i=1}^{M-1} (1 - \pi_i - \eta_i)\right)\pi_M \\ &\geq \pi_0 \sum_{i=0}^{M-1} (1 - \pi_0 - \eta_0)^i = \frac{\pi_0}{\pi_0 + \eta_0} [1 - (1 - \pi_0 - \eta_0)^M] \\ &> 1 - \nu. \end{aligned}$$

We got ρ -balls of generation $k, k + s, \dots, k + (M - 1)s$ so that it is very likely for a point to belong to one of them. Also, balls of the same generation $k + ms$ are $2^{-3}A_0^{-4}\vartheta^{k+ms}$ -separated. Now we need to utilize the regularity of the quasimetric ρ . Indeed, this is to guarantee that we can keep the buffer small but still separate balls of different generations. Let us study two balls B^{k+ms} and B^{k+ns} of centers x^{k+ms} and x^{k+ns} , where $m < n$. (We suppress the lower indices for this argument.) Choose $x \in B^{k+ms}$ and $y \in B^{k+ns}$. Let R and r be the radii of B^{k+ms} and B^{k+ns} , so that $r \leq 2\vartheta^s R$. We have

$$\begin{aligned} (1 + \omega)R &\leq \rho(x^{k+ms}, x^{k+ns}) \\ &\leq (1 + \epsilon)\rho(x^{k+ms}, y) + A(\epsilon)\rho(y, x^{k+ns}) \\ &\leq (1 + \epsilon)^2\rho(x^{k+ms}, x) + (1 + \epsilon)A(\epsilon)\rho(x, y) + A(\epsilon)\rho(y, x^{k+ns}) \\ &\leq (1 + \epsilon)^2R + (1 + \epsilon)A(\epsilon)\rho(x, y) + 2A(\epsilon)\vartheta^s R \\ &\leq R + (2\omega/3)R + (1 + \epsilon)A(\epsilon)\rho(x, y) \end{aligned}$$

implying that $\rho(x, y) \geq cR$, where c depends on ω and ϵ , thus ultimately on ν , whereas

$$C\vartheta^k = \frac{1}{16A_0^4}\vartheta^k \geq R > r \geq \frac{1}{32A_0^4}\vartheta^{k+(M-1)s} = c(\vartheta, \nu)\vartheta^k,$$

since also M and s depend only on ϑ and ν .

Let us now formulate the above given construction of the random almost-covering by balls as a proposition.

Proposition 8.2 *Let $0 < \vartheta < A_0^{-4}/32$, $k \in \mathbb{Z}$, and $\nu \in (0, 1)$ be given. Then we may randomly construct a disjoint family \mathcal{B} of ρ -balls as follows: if $B, B' \in \mathcal{B}$ are two different balls, then*

$$c(\vartheta, \nu)\vartheta^k \leq r_B \leq C\vartheta^k, \quad \rho(B, B') \geq c(\vartheta, \nu)\vartheta^k,$$

and for every $x \in X$,

$$\mathbb{P}\left(x \in \bigcup_{B \in \mathcal{B}} B\right) > 1 - \nu.$$

9 Estimates for Adjacent Cubes of Comparable Size

We are given adjacent ($d(Q, R) < CC_0 \min(\ell(Q), \ell(R))$ say) $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$ of comparable size, that is $\ell(Q) \sim \ell(R)$ (meaning $|\text{gen}(Q) - \text{gen}(R)| \leq r$). We are also given some fixed small $\epsilon > 0$. We define $\Delta = Q \cap R$, $\delta_Q = \{x : d(x, Q) \leq \epsilon \ell(Q) \text{ and } d(x, X \setminus Q) \leq \epsilon \ell(Q)\}$ and $\delta_R = \{x : d(x, R) \leq \epsilon \ell(R) \text{ and } d(x, X \setminus R) \leq \epsilon \ell(R)\}$. Also, set

$$Q_b = Q \cap \bigcup_{R' \in \mathcal{D}' : \ell(R') \sim \ell(Q)} \delta_{R'} \quad (9.1)$$

and

$$R_b = R \cap \bigcup_{Q' \in \mathcal{D} : \ell(Q') \sim \ell(R)} \delta_{Q'}. \quad (9.2)$$

We define $Q_s = Q \setminus \Delta \setminus \delta_R$, $Q_\partial = Q \setminus \Delta \setminus Q_s$, $R_s = R \setminus \Delta \setminus \delta_Q$ and $R_\partial = R \setminus \Delta \setminus R_s$. Furthermore, set $\tilde{\Delta} = \Delta \setminus \delta_Q \setminus \delta_R$. We now finally fix $\delta = A_0^{-4}/1000$, and then fix the smallest k for which $\delta^k \leq \Lambda^{-1}(8^{-1}\epsilon \min(\ell(Q), \ell(R)))^\beta$. Consider some small enough $\nu \in (0, 1)$, and set $\vartheta = \delta$. Recall Proposition 8.2, that is, the random way to construct a collection of ρ -balls \mathcal{B} starting from the fixed level k with parameter ν (and the related parameters ω , s , and M that all depend on ν but not on k). As we have $\mathbb{P}(x \in \bigcup_{B \in \mathcal{B}} B) > 1 - \nu$ for all $x \in X$, we have $\mathbb{E}(\mu(\tilde{\Delta} \setminus \bigcup_{B \in \mathcal{B}} B)) < \nu \mu(\tilde{\Delta})$. So we may now fix some such ball covering \mathcal{B} for which $\mu(\tilde{\Delta} \setminus \bigcup_{B \in \mathcal{B}} B) \leq \nu \mu(\tilde{\Delta})$ as we have positive probability to obtain one. We now remove from the collection \mathcal{B} those balls that do not touch $\tilde{\Delta}$.

First we want to estimate $\#\mathcal{B}$. Observe that $\text{diam}(\Delta) \leq C_0 \min(\ell(Q), \ell(R))$ and fix some $x_0 \in \Delta$. If $B \in \mathcal{B}$, fix some $x \in B \cap \Delta$. Denote the center of B by z_B . Note that

$$\begin{aligned} d(z_B, x_0) &\leq d(z_B, x) + d(x, x_0) \\ &\leq 2\rho(z_B, x)^{1/\beta} + d(x, x_0) \\ &\leq 2\delta^{k/\beta} + C_0 \min(\ell(Q), \ell(R)) \\ &< [2 + C_0 \Lambda^{1/\beta} \delta^{-1/\beta} 8\epsilon^{-1}] \delta^{k/\beta} \\ &\leq 16C_0 \Lambda^{1/\beta} \delta^{-1/\beta} \epsilon^{-1} \delta^{k/\beta}. \end{aligned}$$

This means that the d -ball $B(x_0, 16C_0 \Lambda^{1/\beta} \delta^{-1/\beta} \epsilon^{-1} \delta^{k/\beta})$ contains the centers z_B of the disjoint d -balls $B(z_B, 4^{-1}(A_0^{-4} \delta^{(M-1)s} / 32)^{1/\beta} \delta^{k/\beta})$. This implies that

$$\#\mathcal{B} \leq N \left(\frac{\epsilon}{64C_0} \left(\frac{\delta^{(M-1)s+1}}{32\Lambda A_0^4} \right)^{1/\beta} \right)^{-n}.$$

This is a dependence we can live with, as all the quantities in the upper bound will be eventually fixed.

Next, let us check that $\Lambda B \subset \Delta$ for every $B \in \mathcal{B}$. There exists $x \in B$ so that $d(x, X \setminus \Delta) \geq \epsilon \min(\ell(Q), \ell(R))$ as $B \cap \tilde{\Delta} \neq \emptyset$. If $w \in \Lambda B$, we have

$$\begin{aligned} d(x, X \setminus \Delta) - d(w, X \setminus \Delta) &\leq d(w, x) \\ &\leq d(w, z_B) + d(z_B, x) \\ &\leq 2\rho(w, z_B)^{1/\beta} + 2\rho(z_B, x)^{1/\beta} \\ &\leq 4\Lambda^{1/\beta} \delta^{k/\beta} \leq \frac{\epsilon}{2} \min(\ell(Q), \ell(R)), \end{aligned}$$

and so $d(\Lambda B, X \setminus \Delta) \geq (1/2)\epsilon \min(\ell(Q), \ell(R)) > 0$.

In a forthcoming decomposition we shall have plenty of separated terms. For these the following lemma comes in handy, and we use it without further mention in what follows.

Lemma 9.1 *Let S_1 and S_2 be two sets so that we have $d(S_1) \sim d(S_2)$ and $d(S_1, S_2) \gtrsim \epsilon \min(d(S_1), d(S_2))$. Suppose we are also given functions φ and ψ so that $\|\varphi\|_{L^\infty(\mu)} + \|\psi\|_{L^\infty(\mu)} \lesssim 1$, $\text{spt } \varphi \subset S_1$ and $\text{spt } \psi \subset S_2$. Then it holds that*

$$|\langle T\varphi, \psi \rangle| \lesssim \epsilon^{-d} \mu(S_1)^{1/2} \mu(S_2)^{1/2}.$$

Proof Using the first kernel estimate we have

$$|\langle T\varphi, \psi \rangle| \lesssim \int_{S_2} \int_{S_1} \min\left(\frac{1}{\lambda(x, d(x, y))}, \frac{1}{\lambda(y, d(x, y))}\right) d\mu(y) d\mu(x).$$

It holds that $\lambda(y, d(x, y)) \geq \lambda(y, d(S_1, S_2)) \gtrsim \lambda(y, \epsilon d(S_1)) \gtrsim \epsilon^{-d} \lambda(y, d(S_1)) \geq \epsilon^{-d} \mu(S_1)$. Similarly $\lambda(x, d(x, y)) \gtrsim \epsilon^{-d} \mu(S_2)$. Thus, it follows that

$$|\langle T\varphi, \psi \rangle| \lesssim \epsilon^{-d} \min(\mu(S_1), \mu(S_2)) \leq \epsilon^{-d} \mu(S_1)^{1/2} \mu(S_2)^{1/2}. \quad \square$$

Recall that any two $B, B' \in \mathcal{B}$ are separated by a distance comparable to their radii. Then the same is still true for some expanded balls $(1 + \epsilon')B$, which contain the supports of the functions $\tilde{\chi}_B := \tilde{\chi}_{B, \epsilon'}$ from the weak boundedness property. We write

$$\chi_\Delta = \sum_{B \in \mathcal{B}} \tilde{\chi}_B + \tilde{\chi}_{\Delta \setminus \cup B}, \quad \tilde{\chi}_{\Delta \setminus \cup B} := \chi_\Delta - \sum_{B \in \mathcal{B}} \tilde{\chi}_B \leq \chi_{\Delta \setminus \cup B};$$

note also that $\tilde{\chi}_{\Delta \setminus \cup B} \geq 0$ since $(1 + \epsilon')B \subset \Lambda B \subset \Delta$. Let further

$$\tilde{\chi}_{\Delta \setminus \cup B} = \tilde{\chi}_{\Delta \setminus \cup B} (\chi_{\Omega_i} + \chi_{\Omega_Q} + \chi_{\Omega_R}) =: \tilde{\chi}_{\Omega_i} + \tilde{\chi}_{\Omega_Q} + \tilde{\chi}_{\Omega_R},$$

where $\Delta \setminus \cup B$ is a disjoint union of $\Omega_i = \tilde{\Delta} \setminus \cup B$ and some sets $\Omega_Q \subset Q_b$ and $\Omega_R \subset R_b$. We now decompose

$$\begin{aligned} \langle T(\chi_Q b_1), \chi_R b_2 \rangle &= \langle T(\chi_Q b_1), \chi_{R_\partial} b_2 \rangle + \langle T(\chi_Q b_1), \chi_{R_s} b_2 \rangle \\ &\quad + \langle T(\chi_{Q_\partial} b_1), \chi_\Delta b_2 \rangle + \langle T(\chi_{Q_s} b_1), \chi_\Delta b_2 \rangle \\ &\quad + \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Delta \setminus \cup_B} b_2 \rangle + \left\langle T(\tilde{\chi}_{\Delta \setminus \cup_B} b_1), \sum \tilde{\chi}_B b_2 \right\rangle \\ &\quad + \left\langle T\left(\sum \tilde{\chi}_B b_1\right), \sum \tilde{\chi}_B b_2 \right\rangle \\ &= A + B + C + D + E + F + G. \end{aligned}$$

Furthermore, we decompose

$$\begin{aligned} E &= \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Delta \setminus \cup_B} b_2 \rangle \\ &= \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_Q} b_2 \rangle + \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_R} b_2 \rangle + \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_i} b_2 \rangle \\ &= E_1 + E_2 + E_3 \end{aligned}$$

and

$$\begin{aligned} F &= \left\langle T(\tilde{\chi}_{\Delta \setminus \cup_B} b_1), \sum \tilde{\chi}_B b_2 \right\rangle \\ &= \left\langle T(\tilde{\chi}_{\Omega_Q} b_1), \sum \tilde{\chi}_B b_2 \right\rangle + \left\langle T(\tilde{\chi}_{\Omega_R} b_1), \sum \tilde{\chi}_B b_2 \right\rangle + \left\langle T(\tilde{\chi}_{\Omega_i} b_1), \sum \tilde{\chi}_B b_2 \right\rangle \\ &= F_1 + F_2 + F_3. \end{aligned}$$

We still write

$$\begin{aligned} G &= \left\langle T\left(\sum \tilde{\chi}_B b_1\right), \sum \tilde{\chi}_B b_2 \right\rangle \\ &= \sum_B \langle T(\tilde{\chi}_B b_1), \tilde{\chi}_B b_2 \rangle + \sum_{B \neq B'} \langle T(\tilde{\chi}_B b_1), \tilde{\chi}_{B'} b_2 \rangle \\ &= G_1 + G_2. \end{aligned}$$

Let us deal with these terms now. We have for the terms

$$A = \langle T(\chi_Q b_1), \chi_{R_\partial} b_2 \rangle, \quad C = \langle T(\chi_{Q_\partial} b_1), \chi_\Delta b_2 \rangle, \quad E_1 = \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_Q} b_2 \rangle$$

and

$$\begin{aligned} E_2 &= \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_R} b_2 \rangle, \quad F_1 = \left\langle T(\tilde{\chi}_{\Omega_Q} b_1), \sum \tilde{\chi}_B b_2 \right\rangle, \\ F_2 &= \left\langle T(\tilde{\chi}_{\Omega_R} b_1), \sum \tilde{\chi}_B b_2 \right\rangle \end{aligned}$$

that

$$|C| + |E_1| + |F_1| \lesssim \|T\| \|\chi_{Q_\partial} b_1\|_{L^2(\mu)} \mu(R)^{1/2}$$

and

$$|A| + |E_2| + |F_2| \lesssim \|T\| \mu(Q)^{1/2} \|\chi_{R_\partial} b_2\|_{L^2(\mu)},$$

where we have used the facts that $|b_1| \sim 1$ and $|b_2| \sim 1$. Next, we observe that for the terms

$$E_3 = \langle T(\chi_\Delta b_1), \tilde{\chi}_{\Omega_i} b_2 \rangle \quad \text{and} \quad F_3 = \left\langle T(\tilde{\chi}_{\Omega_i} b_1), \sum \tilde{\chi}_B b_2 \right\rangle$$

we have

$$|E_3| + |F_3| \lesssim \nu^{1/2} \|T\| \mu(Q)^{1/2} \mu(R)^{1/2}.$$

For the separated terms

$$B = \langle T(\chi_Q b_1), \chi_{R_s} b_2 \rangle \quad \text{and} \quad D = \langle T(\chi_Q, b_1), \chi_\Delta b_2 \rangle$$

we have the estimate

$$|B| + |D| \lesssim \epsilon^{-d} \mu(Q)^{1/2} \mu(R)^{1/2}.$$

It remains to deal with the term G . We invoke the weak boundedness property, the fact that $\Lambda B \subset \Delta = Q \cap R$ for all the boundedly many $B \in \mathcal{B}$, and recall that the parameter ϵ' implicit in the functions $\tilde{\chi}_B = \tilde{\chi}_{B, \epsilon'}$ only depends on ν to get that

$$|G_1| = \left| \sum_B \langle T(\tilde{\chi}_B b_1), \tilde{\chi}_B b_2 \rangle \right| \lesssim C(\epsilon, \nu) \mu(Q)^{1/2} \mu(R)^{1/2}.$$

Using the separation of different balls $(1 + \epsilon')B$ supporting the $\tilde{\chi}_B$, we also obtain that

$$|G_2| = \left| \sum_{B \neq B'} \langle T(\tilde{\chi}_B b_1), \tilde{\chi}_{B'} b_2 \rangle \right| \lesssim C(\epsilon, \nu) \mu(Q)^{1/2} \mu(R)^{1/2}.$$

We now recapitulate what we have done in the form of a proposition.

Proposition 9.2 *Let $Q \in \mathcal{D}$ and $R \in \mathcal{D}'$ be two adjacent cubes of comparable size, that is, $d(Q, R) < CC_0 \min(\ell(Q), \ell(R))$ and $|\text{gen}(Q) - \text{gen}(R)| \leq r$. Let $\epsilon > 0$ and $\nu \in (0, 1)$. It holds that*

$$\begin{aligned} |\langle T(\chi_Q b_1), \chi_R b_2 \rangle| &\lesssim \|T\| \|\chi_{Q_b} b_1\|_{L^2(\mu)} \mu(R)^{1/2} + \|T\| \mu(Q)^{1/2} \|\chi_{R_b} b_2\|_{L^2(\mu)} \\ &\quad + \nu^{1/2} \|T\| \mu(Q)^{1/2} \mu(R)^{1/2} + C(\epsilon, \nu) \mu(Q)^{1/2} \mu(R)^{1/2}, \end{aligned}$$

where Q_b and R_b are as in (9.1) and (9.2), respectively.

10 Random Dyadic Systems

We now randomize our dyadic grids. We first fix a reference system of dyadic points (z_α^k) and the relation \leq essentially as in the case of the random ball covering (but working with the metric d instead). Indeed, for each $k \in \mathbb{Z}$ fix some maximal collection $z_\alpha^k \in X$ for which $d(z_\alpha^k, z_\beta^k) \geq \delta^k$ for all $\alpha \neq \beta$. For each (k, α) there exists at least one β for which $d(z_\alpha^k, z_\beta^{k-1}) < \delta^{k-1}$. Also, there exists at most one β

for which $d(z_\alpha^k, z_\beta^{k-1}) < \delta^{k-1}/2$. The ordering \leq is constructed using the rules we now describe. Consider any pair (k, α) . Check first whether there exists β so that $d(z_\alpha^k, z_\beta^{k-1}) < \delta^{k-1}/2$. If so, set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Otherwise, choose any β for which $d(z_\alpha^k, z_\beta^{k-1}) < \delta^{k-1}$, and set $(k, \alpha) \leq (k-1, \beta)$ and $(k, \alpha) \not\leq (k-1, \gamma)$ for $\gamma \neq \beta$. Extend by transitivity.

Next, we introduce the transitive relation \searrow exactly as before, and equipped with the same probabilistic notions. The new dyadic points y_α^k are built as before. The points are said to conflict if $d(y_\alpha^k, y_\beta^k) < \delta^k/4$. Then we do the familiar removal procedure and get the final dyadic points x_α^k , which satisfy $d(x_\alpha^k, x_\beta^k) \geq \delta^k/4$ if $\alpha \neq \beta$. Observe also that for an arbitrary $x \in X$ there exists x_α^k so that $d(x, x_\alpha^k) < 3\delta^k$. Using these one may then build a new relation \leq' , similar to \leq but related to these new points, and the corresponding “half-open” cubes Q_α^k .

Consider a given point $x \in X$. There exists β so that $d(x, z_\beta^{k+1}) < \delta^{k+1} < \delta^k/500$. Let $(k+1, \beta) \leq (k, \alpha)$. We have $\mathbb{P}(z_\beta^{k+1} = x_\alpha^k) \geq \pi_0 > 0$ by an analogue of Lemma 8.1. In particular, final dyadic points of consecutive generations k and $k+1$ may well end up close to each other in this sense. Recalling Lemma 4.5 this is relevant for the proof of the next lemma.

Lemma 10.1 *For some fixed $x \in X$ and $k \in \mathbb{Z}$, it holds*

$$\mathbb{P}(x \in \delta_{Q_\alpha^k} \text{ for some } \alpha) \lesssim \epsilon^\eta$$

for some $\eta > 0$.

Proof Recall the open and closed cubes \tilde{Q}_α^k and \overline{Q}_α^k and how they are related to the “half-open” cubes (we no longer use the hat notation so it may be a bit confusing). One advantage of these is that they are determined by the centers a little bit differently from the “half-open” ones. Namely, to know these cubes for some generation M , it suffices to know the centers x_β^ℓ for generations $\ell \geq M$.

Fix the largest m so that $500\epsilon \leq \delta^m$. Now the point is to simply combine Lemma 4.5 with the last observation preceding this lemma. Indeed, let the relation \searrow be fixed from the level $k+m$ up. Choose some σ_{k+m} so that $x \in \overline{Q}_{\sigma_{k+m}}^{k+m}$. We then randomly choose the relation \searrow between the levels $k+m$ and $k+m-1$. We have $\mathbb{P}(d(x_{\sigma_{k+m}}^{k+m}, x_\beta^{k+m-1}) < \delta^{k+m-1}/500 \text{ for some } \beta) \geq \pi_0$ so that $\mathbb{P}(d(x_{\sigma_{k+m}}^{k+m}, x_\beta^{k+m-1}) \geq \delta^{k+m-1}/500 \text{ for all } \beta) \leq 1 - \pi_0 =: \pi_1 < 1$. Let $(k+m, \sigma_{k+m}) \leq' (k+m-1, \sigma_{k+m-1})$. We again have $\mathbb{P}(d(x_{\sigma_{k+m-1}}^{k+m-1}, x_\beta^{k+m-2}) \geq \delta^{k+m-2}/500 \text{ for all } \beta) \leq \pi_1 < 1$. We continue this way. Let $x \in \overline{Q}_\alpha^k$. Lemma 4.5 implies together with independence that

$$\begin{aligned} \mathbb{P}(d(x, X \setminus \tilde{Q}_\alpha^k) < \epsilon^k) &\leq \mathbb{P}(d(x_{\sigma_{k+j}}^{k+j}, x_{\sigma_{k+j-1}}^{k+j-1}) \geq \delta^{k+j-1}/500 \text{ for all } j = 1, \dots, m) \\ &\leq \pi_1^m = (\delta^m)^{\log(\pi_1)/\log \delta} \lesssim \epsilon^\eta, \end{aligned}$$

where $\eta = \log(\pi_1)/\log \delta > 0$. This was actually a conditional probability with the condition that the relation \searrow was fixed in some way from the level $k+m$ up, but as

this was arbitrary, the same estimate holds without any conditionality. It remains to note that what we have done actually proves the whole lemma. \square

Theorem 10.2 *For a fixed Q_α^k we have under the random choice of the other dyadic system that*

$$\mathbb{P}(Q_\alpha^k \in \mathcal{D}_{\text{bad}}) \lesssim \delta^{r\gamma\eta}.$$

Proof We make yet another assumption about the largeness of r . Namely, we assume that r is so large that $\delta^{r(1-\gamma)} < 1$, say. Let R^{k-s} be the unique \mathcal{D}' -cube of generation $k - s$ containing the center x_α^k of Q_α^k . We have that

$$d(Q_\alpha^k, X \setminus R^{k-s}) \geq d(x_\alpha^k, X \setminus R^{k-s}) - C_0\delta^k.$$

If $d(x_\alpha^k, X \setminus R^{k-s}) \geq 2\delta^{k\gamma}\delta^{(k-s)(1-\gamma)}$, then $d(Q_\alpha^k, X \setminus R^{k-s}) \geq \delta^{k\gamma}\delta^{(k-s)(1-\gamma)}$ for $s \geq r$ by the above inequality and the assumption that $\delta^{r(1-\gamma)} < 1$. Using a variant of the previous lemma we thus get that

$$\mathbb{P}(Q_\alpha^k \in \mathcal{D}_{\text{bad}}) \lesssim \sum_{s=r}^{\infty} (\delta^{\gamma\eta})^s \lesssim \delta^{r\gamma\eta}. \quad \square$$

11 Synthesis

We now combine all these estimates to prove the non-trivial side of our main theorem, Theorem 2.3. Indeed, we now prove that $\|T\| \lesssim 1$. To this end, choose f and g so that $|\langle Tf, g \rangle| \geq (1/2)\|T\|$, $\|f\|_{L^2(\mu)} = \|g\|_{L^2(\mu)} = 1$ and $\text{spt } f \subset B(x_0, \delta^m)$, $\text{spt } g \subset B(x_1, \delta^m)$ for some $x_0, x_1 \in X$ and $m \in \mathbb{Z}$. Write

$$f = \sum_{\substack{Q \in \mathcal{D} \\ \text{gen}(Q) \geq m}} \Delta_Q^{b_1} f + \sum_{\substack{Q \in \mathcal{D} \\ \text{gen}(Q) = m}} E_Q^{b_1} f = f_{\text{good}} + f_{\text{bad}},$$

where

$$f_{\text{good}} = \sum_{\substack{Q \in \mathcal{D}_{\text{good}} \\ \text{gen}(Q) \geq m}} \Delta_Q^{b_1} f + \sum_{\substack{Q \in \mathcal{D} \\ \text{gen}(Q) = m}} E_Q^{b_1} f.$$

We write the similar decomposition also for g , and then estimate

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

Furthermore, we have

$$\begin{aligned} |\langle Tf_{\text{good}}, g_{\text{good}} \rangle| &\leq \sum |\langle T(\Delta_Q^{b_1} f), \Delta_R^{b_2} g \rangle| + \sum |\langle T(\Delta_Q^{b_1} f), E_R^{b_2} g \rangle| \\ &\quad + \sum |\langle T(E_Q^{b_1} f), \Delta_R^{b_2} g \rangle| + \sum |\langle T(E_Q^{b_1} f), E_R^{b_2} g \rangle|, \end{aligned}$$

where we sum over the obvious sets.

The first three series are similar so we only deal with the first one (by the above theory, it suffices that the term with the smaller support has zero integral). To this end, let us estimate the first series by (we agree that naturally all the time $Q \in \mathcal{D}_{\text{good}}$ and $R \in \mathcal{D}'_{\text{good}}$)

$$\left(\sum_{\ell(Q) \sim \ell(R)} + \sum_{\ell(Q) \not\sim \ell(R): Q \cap R = \emptyset} + \sum_{\ell(Q) \not\sim \ell(R): Q \cap R \neq \emptyset} \right) |\langle T(\Delta_Q^{b_1} f), \Delta_R^{b_2} g \rangle|.$$

The second series in the above decomposition is $\lesssim \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} = 1$ by Sect. 6 (see Lemma 6.2 and Proposition 6.3), while the third series is $\lesssim \|f\|_{L^2(\mu)} \times \|g\|_{L^2(\mu)} = 1$ by Sect. 7 (see (7.1), Theorem 7.2, and Proposition 7.3). We then write the first series in the above decomposition in the form

$$\begin{aligned} & \sum_{\substack{\ell(Q) \sim \ell(R) \\ d(Q, R) \geq CC_0 \min(\ell(Q), \ell(R))}} |\langle T(\Delta_Q^{b_1} f), \Delta_R^{b_2} g \rangle| \\ & + \sum_{\substack{\ell(Q) \sim \ell(R) \\ d(Q, R) < CC_0 \min(\ell(Q), \ell(R))}} |\langle T(\Delta_Q^{b_1} f), \Delta_R^{b_2} g \rangle| \end{aligned}$$

noting that the first series is $\lesssim \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} = 1$ by the techniques used in Sect. 6 (see the beginning of the proof of Lemma 6.2 and Proposition 6.3). For a given cube Q there exists only $\lesssim 1$ cubes R such that $\ell(Q) \sim \ell(R)$ and $d(Q, R) \lesssim \min(\ell(Q), \ell(R))$. Thus, we have by Sect. 9 that (see Proposition 9.2)

$$\begin{aligned} & \sum_{\substack{\ell(Q) \sim \ell(R) \\ d(Q, R) < CC_0 \min(\ell(Q), \ell(R))}} |\langle T(\Delta_Q^{b_1} f), \Delta_R^{b_2} g \rangle| \\ & \lesssim \|T\| \left(\sum_{\text{gen}(Q) > m} |A_Q|^2 \| \chi_{Q_b} b_1 \|_{L^2(\mu)}^2 \right)^{1/2} \\ & + \|T\| \left(\sum_{\text{gen}(R) > m} |B_R|^2 \| \chi_{R_b} b_2 \|_{L^2(\mu)}^2 \right)^{1/2} \\ & + \nu^{1/2} \|T\| + C(\epsilon, \nu), \end{aligned}$$

where the constants A_Q are related to the decomposition $\Delta_Q^{b_1} f = \sum A_{Q'} \chi_{Q'} b_1$, where we sum over the subcubes Q' of Q . The constants B_R are similarly related to g .

The fourth series involving the factors $E_Q^{b_1} f$ and $E_R^{b_2} g$ has no terms with zero integral, but the point is that there are only $\lesssim 1$ nonzero terms as the functions f and g are supported on balls of radius δ^m and $\text{gen}(Q) = \text{gen}(R) = m$ in that sum. One can deal with the well-separated terms using the first kernel estimate and use the estimates of Sect. 9 for the rest (see Proposition 9.2). The net result is that

$$\begin{aligned}
 (1/2)\|T\| &\leq |\langle Tf, g \rangle| \\
 &\lesssim \|T\| \left(\sum_{\text{gen}(Q) \geq m} |A_Q|^2 \|\chi_{Q_b} b_1\|_{L^2(\mu)}^2 \right)^{1/2} \\
 &\quad + \|T\| \left(\sum_{\text{gen}(R) \geq m} |B_R|^2 \|\chi_{R_b} b_2\|_{L^2(\mu)}^2 \right)^{1/2} \\
 &\quad + \nu^{1/2} \|T\| + \|T\| \|g_{\text{bad}}\|_{L^2(\mu)} + \|T\| \|f_{\text{bad}}\|_{L^2(\mu)} + C(\epsilon, \nu),
 \end{aligned}$$

where the constants for the level m come from writing $E_Q^{b_1} f = A_Q \chi_Q b_1$ and $E_R^{b_2} g = B_R \chi_R b_2$. Choosing r large enough, ϵ and ν small enough, and choosing the dyadic grids \mathcal{D} and \mathcal{D}' so that the first five terms (together with the implicit constants in front) contribute less than $(1/4)\|T\|$ yields that $\|T\| \lesssim 1$ as desired. These details follow pretty much as in [11] now that the lemmas in the previous section have been proven.

Let us quickly sketch the details for completeness. We can estimate $\mathbb{E}\|f_{\text{bad}}\|_{L^2(\mu)}^2$ (over the grids \mathcal{D} and \mathcal{D}') by introducing the square function

$$Sh(x) = \sum_{\substack{Q \in \mathcal{D} \\ \text{gen}(Q) \geq m}} \|\Delta_Q^{b_1} h\|_{L^2(\mu)}^2 \mu(Q)^{-1} \chi_Q + \sum_{\substack{Q \in \mathcal{D}' \\ \text{gen}(Q) = m}} \|E_Q^{b_1} h\|_{L^2(\mu)}^2 \mu(Q)^{-1} \chi_Q.$$

The point is that $\int_X Sh \, d\mu \sim \|h\|_{L^2(\mu)}^2$ for all h . By Theorem 10.2 we have for any fixed grid \mathcal{D} taking the expectation over the grids \mathcal{D}' that

$$\mathbb{E}\|f_{\text{bad}}\|_{L^2(\mu)}^2 \lesssim \int_X \mathbb{E}Sf_{\text{bad}} \, d\mu \lesssim \delta^{r\gamma\eta} \int_X Sf \, d\mu \lesssim \delta^{r\gamma\eta} \|f\|_{L^2(\mu)}^2 = \delta^{r\gamma\eta},$$

and then the same holds if we take the expectation over all the grids \mathcal{D} and \mathcal{D}' too. We now once and for all fix r to be so large that everything we have done above works and that we have $\mathbb{E}\|f_{\text{bad}}\|_{L^2(\mu)}^2 \leq 2^{-100}$. The same argument shows that $\mathbb{E}\|g_{\text{bad}}\|_{L^2(\mu)}^2 \leq 2^{-100}$.

To deal with the remaining terms, write

$$f^k = \sum_{\text{gen}(Q)=k} A_Q \chi_Q b_1 \quad \text{and} \quad f_b^k = \sum_{\text{gen}(Q)=k} A_Q \chi_{Q_b} b_1.$$

We have by Lemma 10.1 that $\mathbb{E}\chi_{Q_b}(x) \lesssim \epsilon^\eta \chi_Q(x)$ for all $x \in X$ and for all Q in a fixed grid \mathcal{D} (taking the expectation over the grids \mathcal{D}'), and thus

$$\mathbb{E} \sum_{k \geq m} \|f_b^k\|_{L^2(\mu)}^2 = \sum_{k \geq m} \int_X \mathbb{E}|f_b^k|^2 \, d\mu \lesssim \epsilon^\eta \sum_{k \geq m} \|f^k\|_{L^2(\mu)}^2 \lesssim \epsilon^\eta \|f\|_{L^2(\mu)}^2 = \epsilon^\eta.$$

The same then holds if we take the expectation over all the grids \mathcal{D} and \mathcal{D}' too. The same argument shows that also (with the obvious notation)

$$\mathbb{E} \sum_{k \geq m} \|\delta_b^k\|_{L^2(\mu)}^2 \lesssim \epsilon^\eta.$$

This proves that we may choose the grids \mathcal{D} and \mathcal{D}' so that

$$(1/2 - 2^{-50} - C_4 v^{1/2} - C_5 \epsilon^{\eta/2}) \|T\| \lesssim C(\epsilon, v),$$

from which the claim follows by choosing the constants ϵ and v properly.

12 Application to Bergman-Type Operators

Volberg and Wick [13] recently obtained a characterization of measures μ in the unit ball \mathbb{B}_{2n} of \mathbb{C}^n for which the analytic Besov–Sobolev space $B_2^\sigma(\mathbb{B}_{2n})$ embeds continuously into $L^2(\mu)$. (More precisely, they completed the picture by settling the remaining difficult case concerning $\sigma \in (1/2, n/2)$.) Their proof goes through a new $T1$ theorem for what they call “Bergman-type” operators. Let us describe the situation to see that this application (although, unfortunately, not their abstract $T1$ theorem behind it) could also be obtained as a consequence of our theory.

The measures μ in [13] satisfy the upper power bound $\mu(B(x, r)) \leq r^m$, except possibly when $B(x, r) \subseteq H$, where H is a fixed open set. However, in the exceptional case it holds that $r \leq \delta(x) := d(x, H^c)$, and hence

$$\mu(B(x, r)) \leq \lim_{\epsilon \rightarrow 0^+} \mu(B(x, \delta(x) + \epsilon)) \leq \lim_{\epsilon \rightarrow 0^+} (\delta(x) + \epsilon)^m = \delta(x)^m.$$

Thus we find that their measures are actually upper doubling with

$$\mu(B(x, r)) \leq \max(\delta(x)^m, r^m) =: \lambda(x, r).$$

The Bergman-type kernels $K(x, y)$ of [13] are required to have the pointwise estimate

$$|K(x, y)| \leq C \min\left(\frac{1}{d(x, y)^m}, \frac{1}{\max(\delta(x)^m, \delta(y)^m)}\right),$$

where the upper bound is seen to be precisely the same

$$C \min\left(\frac{1}{\lambda(x, d(x, y))}, \frac{1}{\lambda(y, d(x, y))}\right)$$

as required by our theory. However, the Hölder-continuity estimate is only assumed in the form

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

for $d(x, y) > 2d(x, x')$, which is weaker than our condition when $\delta(x) \gg d(x, y)$. Hence the abstract main result of Volberg and Wick [13, Theorem 1], is not as such included in our theory, but consists of a different extension of the non-homogeneous analysis of [11].

However, when it comes to the main application concerning the Besov–Sobolev spaces [13, Theorem 2], the relevant kernel has the specific form

$$K(x, y) = (1 - \bar{x} \cdot y)^{-m}, \quad x, y \in \bar{\mathbb{B}}_{2n} \subset \mathbb{C}^n.$$

Here \bar{x} stands for the componentwise complex conjugation, and dot designates the usual dot product of n -vectors. Moreover, one equips $\bar{\mathbb{B}}_{2n} \setminus \frac{1}{2}\mathbb{B}_{2n}$ with the regular quasi-distance (see [12, Lemma 2.6])

$$d(x, y) := ||x| - |y|| + \left| 1 - \frac{\bar{x} \cdot y}{|x||y|} \right|.$$

Finally, the set H related to the exceptional balls is now the open unit ball \mathbb{B}_{2n} . It is noteworthy that $\delta(x) = d(x, H^c) = 1 - |x|$ is the same as the distance of x and H^c in the Euclidean metric [12, Lemma 2.8].

In [13] it is checked that this kernel K , the quasi-metric d , and the set H indeed satisfy the Bergman-type kernel estimates. We now observe that even the standard estimates of our theory are verified. It is shown in [12, Eq. (6)] that $|1 - \bar{x} \cdot y| \geq 3^{-1}d(x, y)$, and obviously $|1 - \bar{x} \cdot y| \geq 1 - |x| = \delta(x)$; similarly with y in place of x . This confirms the first standard estimate, which we already knew. As for the Hölder-continuity, the proof of [12, Proposition 2.13] contains the bound

$$|K(x, y) - K(x', y)| \lesssim \left(\frac{d(x, x')}{|1 - \bar{x} \cdot y|} \right)^{1/2} \frac{1}{|1 - \bar{x} \cdot y|^m},$$

for $d(x, y) > Cd(x, x')$, and it suffices to use $|1 - \bar{x} \cdot y| \gtrsim d(x, y)$ in the first factor and $|1 - \bar{x} \cdot y| \gtrsim \max(d(x, y), \delta(x))$ in the second. The Hölder estimate with respect to the second variable is of course completely analogous.

In the formulation of their weak boundedness property and the BMO conditions, Volberg and Wick use certain “cubes” [13, Sect. 7] which, just like in \mathbb{R}^n with the usual distance, can actually be viewed as balls with respect to an equivalent regular quasimetric of ℓ^∞ -type. So even this is compatible with our theory. Volberg and Wick conclude their paper [13] with essentially the same remark, with which Nazarov, Treil, and Volberg started theirs [11], that “these considerations can be extended to the case of metric spaces.” And indeed they can!

References

1. Bramanti, M.: Singular integrals in nonhomogeneous spaces: L^2 and L^p continuity from Hölder estimates. *Rev. Mat. Iberoam.* **26**(1), 347–366 (2010)
2. Christ, M.: A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.* **60/61**(2), 601–628 (1990)

3. Coifman, R.R., Weiss, G.: Analyse harmonique non-commutative sur certains espaces homogènes. In: Étude de certaines intégrales singulières. Lecture Notes in Mathematics, vol. 242. Springer, Berlin (1971)
4. David, G., Journé, J.-L., Semmes, S.: Opérateurs de Calderón-Zygmund, fonctions para-accretives et interpolation. Rev. Mat. Iberoam. **1**(4), 1–56 (1985)
5. García-Cuerva, J., Gatto, A.E.: Lipschitz spaces and Calderón-Zygmund operators associated to non-doubling measures. Publ. Mat. **49**(2), 285–296 (2005)
6. Hytönen, T.: A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa. Publ. Mat. **54**(2), 485–504 (2010)
7. Luukkainen, J., Saksman, E.: Every complete doubling metric space carries a doubling measure. Proc. Am. Math. Soc. **126**(2), 531–534 (1998)
8. McIntosh, A., Meyer, Y.: Algèbres d'opérateurs définis par des intégrales singulières. C. R. Acad. Sci. Paris Sér. I Math. **301**(8), 395–397 (1985)
9. Macías, R.A., Segovia, C.: Lipschitz functions on spaces of homogeneous type. Adv. Math. **33**(3), 257–270 (1979)
10. Nazarov, F., Treil, S., Volberg, A.: Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces. Int. Math. Res. Not. **9**, 463–487 (1998)
11. Nazarov, F., Treil, S., Volberg, A.: The Tb -theorem on non-homogeneous spaces. Acta Math. **190**(2), 151–239 (2003)
12. Tchoundja, E.: Carleson measures for the generalized Bergman spaces via a $T(1)$ -type theorem. Ark. Mat. **46**(2), 377–406 (2008)
13. Volberg, A., Wick, B.D.: Bergman-type singular operators and the characterization of Carleson measures for Besov–Sobolev spaces on the complex ball. Am. J. Math. (in press). Preprint. [arXiv:0910.1142](https://arxiv.org/abs/0910.1142)
14. Verbitsky, I.E., Wheeden, R.L.: Weighted norm inequalities for integral operators. Trans. Am. Math. Soc. **350**(8), 3371–3391 (1998)